CONTRIBUTIONS TO STATISTICAL DISTRIBUTION THEORY

by

Arthur William Davis

B.A. (Hons), M.A. (Melb), Ph.D. (A.N.U.)

Thesis submitted to the Department of Statistics in the Faculty of Mathematical Sciences, University of Adelaide, in fulfilment of the requirements for the degree of Doctor of Science.

June, 1979.

Awarded July 1980

Contents.

13. "On the ratios of the individual latent roots to the trace of a Wishart matrix." J. Multiv. Anal., 2, (1972), 440-3.

14. "On the differential equation for Meijer's $G^{p,o}_{p,p}$ function, and further tabulation of Wilks's likelihood ratio criterion." To appear in Biometrika.

15. "A differential equation approach to linear combinations of independent chi-squares." J. Amer. Statist. Assoc. 72, (1977), 212-4.

Section 3. Approximations to statistical distributions.

16. "Generalized asymptotic expansions of Cornish-Fisher type." (with G.W. Hill). Ann. Math. Statist., 39, (1968), 1264-73.

17. "Percentile approximations for ordered F ratios". Biometrika, 57, (1970), 457-9.

18. "Percentile approximations for a class of likelihood ratio criteria". Biometrika, 58, (1971), 349-56.

19. "Tables of some multivariate test criteria" (with J.B.F. Field). CSIRO Divn. Math. Statistics Tech. Paper 32, (1971).

20. "On the k-sample Behrens-Fisher distribution" (with A.J. Scott). CSIRO Divn. Math. Statistics Tech. Paper 33, (1971).

21. "An approximation to the k-sample Behrens-Fisher distribution " (with A.J. Scott). Sankhya, B, 35, (1973), 45-50.

22. "On the asymptotic distribution of Gower's m^2 goodness-of-fit criterion in a particular case." Ann. Inst. Statist. Math. 30, (1978), 71-9.

23. "On certain ratio statistics in weather modification experiments." To appear in Technometrics.

<u>Section 4. A generalization of the zonal polynomials, with applications to noncentral multivariate distributions.</u>

24. "Invariant polynomials with two matrix arguments extending the zonal polynomials." To appear in Multivariate Analysis - V, P.R. Krishnaiah (ed.)

25. "Invariant polynomials with two matrix arguments extending the zonal polynomials: applications to multivariate distribution theory." Submitted.

<u>Section 5. Distribution theory under nonnormality.</u>

26. "Statistical distributions in univariate and multivariate Edgeworth populations." Biometrika <u>63</u>, (1976), 661-70.

27. "On the effects of moderate nonnormality on Wilks's likelihood ratio criterion." Submitted.

28. "Asymptotic theory for principal component analysis: non-normal case." Austral. J. Statist., <u>19</u>, (1977), 206-12.

<u>Section 6. Miscellaneous.</u>

29. "A note on a problem posed by Fisher." Divn. Math. Statistics Tech. Paper <u>26</u>, (1969).

30. "Cyclic change-over designs " (with W.B. Hall). Biometrika, <u>56</u>, (1969), 283-93.

INTRODUCTION

With one exception (paper 30), the papers in this thesis are contributions to the theory of statistical distributions, the principal area of interest being that of multivariate distribution theory. The papers are grouped into six sections, the arrangement within each section being in accordance with the subject matter rather than the date of publication.

Section 1. Distributions arising from stochastic processes. The papers in this section are based on my Ph.D. thesis, and concern distributions of age and location of individuals in certain stochastic populations. The fundamental technique used in these papers is the characteristic functional.

Section 2. Exact null distributions of univariate and multivariate test statistics. The unifying concept throughout most of this section is the characterization of certain statistical distributions as solutions of linear homogeneous differential equations with the object of tabulating the distributions by analytic continuation of an initial series solution. Related systems of differential equations (see paper 12) are derived for the distributions of three of the principal test criteria in multivariate analysis of variance (Hotelling's generalized T_0^2, Pillai's trace, and Roy's largest root). The other main criterion, Wilks's likelihood ratio criterion, is treated as a special case of the differential equation for Meijer's $G_{p,p}^{p,o}$ function (paper 14).

Section 3. Approximations to statistical distributions. Most of the papers in this section are based on a general formulation of the Cornish-Fisher asymptotic expansion which allows for the case of non-normal limiting distributions. This formulation was originally derived by me, but published in a joint paper (16) with Dr G.W. Hill. An important case is that of a limiting chi-square distribution, as required for likelihood ratio criteria and the Hotelling and Pillai traces.

Section 4. A generalization of the zonal polynomials, with applications to noncentral multivariate distributions. Application of a technique described in Section 5 to multivariate distributions has resulted in a generalization of Professor A.T. James's zonal polynomials to a new class of invariant polynomials in two symmetric matrices. These have applications also to multinormal distributions under certain departures from null hypothesis.

Section 5. Distribution theory under nonnormality. A general technique is described for the formal construction of statistical distributions in nonnormal populations, by taking certain expectations of appropriate normal-theory noncentral distributions. This approach unified a number of previous ad hoc investigations, and preliminary results are presented for the effects of moderate nonnormality on the multivariate analysis of variance test criteria.

STATEMENT

The papers in this thesis have not been previously submitted for any other degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person, except where due reference is made in the text of the papers. With regard to papers of joint authorship:

Paper 16: In his Ph. D. thesis, Dr Hill applied Lagrange's inversion formula to Edgeworth's expansion to derive the results presented in Sections 5 and 6 for the original Cornish-Fisher expansion, in which $\Phi(x)$ is the unit normal distribution function. I contributed the extension to "arbitrary" $\Phi(x)$, also the reasonably compact mathematical presentation given in the paper.

Paper 19: Mr Field programmed and computed the tables.

Paper 20: Prof. Scott suggested the problem, supplied the Introduction, and computed the tables. I contributed the application of the Cornish-Fisher expansion and the algebra.

Paper 21: This paper was almost entirely the work of Prof. Scott, based on my derivations in Paper 20.

Paper 30: Mr Hall suggested the problem of constructing and analyzing cyclic change-over designs, and also carried out the calculations in Section 7; we collaborated in drafting the paper. The mathematical formulation is mainly mine.

A.W. Davis

Section 1.

Distributions arising from stochastic processes.

DAVIS A. W.
1964
THE JOURNAL OF THE AUSTRALIAN
MATHEMATICAL SOCIETY
Vol. IV, Part. 2, p.p. 233-243

ON THE CHARACTERISTIC FUNCTIONAL FOR A REPLACEMENT MODEL

BY

A. W. DAVIS

Reprinted from
THE JOURNAL OF THE AUSTRALIAN
MATHEMATICAL SOCIETY
Volume IV, Part 2, p.p. 233-243
1964

ON THE CHARACTERISTIC FUNCTIONAL FOR A REPLACEMENT MODEL

A. W. DAVIS

(received 2 July 1963)

1. Introduction

J. Gani and G. F. Yeo ([2] and [3]) have recently investigated certain age-distributions associated with a replacement model which serves, in particular, as a model for phage reproduction. In this paper, the characteristic functional for this model will be obtained explicitly.

We suppose that at time $t = 0$ there is a system of n ancestors with ranked ages $0 \leqq x_1 < x_2 < \cdots < x_n$. At successive instants (regeneration points) one individual in the system is replaced by another of age zero, so that the system remains of fixed size n. The probability that the ith ranked individual is replaced at a regeneration point is $p_i > 0$ $(i = 1, \cdots, n)$, where $p_1 + \cdots + p_n = 1$; this approximates to an ideal model in which the probability of an individual being replaced would depend on its age. We shall further suppose that the regeneration times are identically and independently distributed with the arbitrary distribution function $G(t)$. In the case $n = 1$, $p_1 = 1$, we have the ordinary renewal process.

The characteristic functional for the age-distribution of the system at any time $t > 0$ is defined to be

$$C[\theta(u); t] = E\left\{\exp\left[i\int_0^\infty \theta(u)dN(u, t)\right]\right\}, \tag{1.1}$$

where $N(u, t)$ is the number of individuals with ages $\leqq u$ at time t, and $\theta(u)$ is any bounded function which is R-integrable over each finite interval. This may be written

$$C[\theta(u); t] = E\left\{\exp\left[i\sum_{k=1}^{n} \theta(u_k(t))\right]\right\} \tag{1.2}$$

where $u_k(t)$ is the age of the kth ranked individual at time t.

If we take

$$\theta(u) = \begin{cases} \phi, & (0 \leqq u \leqq x) \\ 0, & \text{elsewhere,} \end{cases} \tag{1.3}$$

the characteristic functional reduces to

$$C(\phi, t) = E \exp [i\phi N(x, t)], \tag{1.4}$$

the characteristic function of $N(x, t)$.

2. Evaluation of the characteristic functional

Following Bartlett and Kendall [1], we write the characteristic functional in the form

$$C[w_1, \cdots, w_n; \theta(u); t] = E\{w_1^{l_1} \cdots w_n^{l_n} \exp [i \sum_k \theta(u_k(t))]\} \tag{2.1}$$

where

$$w_k = \exp [i\theta(t+x_k)], \qquad (k = 1, \cdots, n), \tag{2.2}$$

and the random variable l_k is equal to 1 while the kth ranked ancestor survives, and is equal to 0 after it is replaced. The summation is extended over all descendants alive at time t.

During the arbitrary time interval $(0, t)$, there will either be no regeneration point, or else a first regeneration point occurs at time τ, $0 \leqq \tau \leqq t$. In the latter event, the new individual becomes the first ranking ancestor for the interval (τ, t). Considering the possibilities, we find that the characteristic functional (2.1) satisfies the following integral equation:

$$\begin{aligned} C[w_1, \cdots, w_n; \theta(u); t] &= [1-G(t)]w_1 \cdots w_n \\ &+ \sum_{j=1}^{n} pj\, C[e^{i\theta(t)}, w_1, \cdots, w_{j-1}, w_{j+1}, \cdots, w_n; \theta(u); t] * G(t) \end{aligned} \tag{2.3}$$

where

$$F * G(t) = \int_{0-}^{t} F(t-\tau)dG(\tau).$$

Since the characteristic functional is clearly linear in the 2^n products $w_1^{l_1} \cdots w_n^{l_n}$ $(l_k = 0, 1)$, we may conveniently write

$$C[w_1, \cdots, w_n; \theta(u); t] = w_1 \cdots w_n[V_0+\sum_{(k)} V_{k_1\cdots k_s}(w_{k_1} \cdots w_{k_s})^{-1}], \tag{2.4}$$

where the summation is extended over all selections of up to n integers $k_1 < k_2 < \cdots < k_s$ from the set $1, 2, \cdots, n$. The V's do not involve the w's, while $V_0 = 1$, $V_{k_1\cdots k_s} = 0$ at $t = 0$ for all $\theta(u)$.

Substituting (2.4) in (2.3),

$$\begin{aligned} V_0+\sum_{(k)} V_{k_1\cdots k_s}(w_{k_1} \cdots w_{k_s})^{-1} &= [1-G(t)] \\ &+ \sum_{j=1}^{n} p_j \{\sum_{(k)} V_{1k_2\cdots k_s}(w_{k_2-1} \cdots w_{k_J-1} w_j w_{k_{J+1}} \cdots w_{k_s})^{-1} \\ &+ e^{i\theta(t)}[V_0 w_j^{-1}+\sum_{k_1>1} V_{k_1\cdots k_s}(w_{k_1-1} \cdots w_{k_J-1} w_j w_{k_{J+1}} \cdots w_{k_s})^{-1}]\} * G(t), \end{aligned} \tag{2.4}$$

k_J being the largest member of $(k_1, \cdots, k_s)$ which is $\leqq j$. The following equations for the V's may now be obtained by equating coefficients:

(2.5) $$V_0 = 1-G(t),$$

(2.6) $$V_{k_1} = p_{k_1}[V_1+e^{i\theta(t)}V_0] * G(t),$$

and for $s > 1$,

(2.7) $$\begin{aligned} V_{k_1\cdots k_s} = \sum_{r=1}^{s} p_{k_r}[&V_{1k_1+1\cdots k_{r-1}+1k_{r+1}\cdots k_s} \\ &+ e^{i\theta(t)}V_{k_1+1\cdots k_{r-1}+1k_{r+1}\cdots k_s}] * G(t). \end{aligned}$$

We will now show that if we define

(2.8) $$q_r = p_{r+1}+\cdots+p_n, \qquad (r = 0, \cdots, n-1),$$

(2.9) $$Q_s = \prod_{r=0}^{s-1}(q_r-q_s), \qquad (s = 1, \cdots, n),$$

(2.10) $$\gamma_s(t) = \sum_{m=0}^{\infty}(1-q_s)^m G^{(m+1)*}(t), \qquad (s = 1, \cdots, n),$$

taking $q_n = 0$, $Q_0 = 1$, and $\gamma_0(t) = G(t)$, then the solution of the system of equations (2.6) and (2.7) is given by the recurrence relations:

(2.11) $$V_{12\cdots s} = \frac{Q_s}{Q_{s-1}}\{e^{i\theta(t)}V_{12\cdots s-1}\} * \gamma_s(t), \qquad (s = 1, \cdots, n),$$

(2.12) $$V_{k_1\cdots k_s} = Q_s^{-1}\prod_{r=1}^{s}(q_{k_r-1}-q_{k_r+s-r})V_{12\cdots s}.$$

Clearly, the solutions for V_0 and the $V_{k_1\cdots k_s}$ satisfy the initial conditions at $t = 0$.

To prove (2.11), we observe that on taking $k_1 = 1, \cdots, k_s = s$, equation (2.7) reduces to

$$\begin{aligned} V_{12\cdots s} &= \sum_{r=1}^{s} p_r[V_{12\cdots s}+e^{i\theta(t)}V_{2\cdots s}] * G(t), \\ &= (1-q_s)[e^{i\theta(t)}V_{2\cdots s} * G(t)+V_{12\cdots s} * G(t)]. \end{aligned}$$

The solution to this equation is

$$V_{12\cdots s} = \{e^{i\theta(t)}V_{2\cdots s}\} * \{\sum_{m=1}^{\infty}(1-q_s)^m G^{m*}(t)\},$$

which will yield (2.11) once we have established (2.12), since the latter gives

$$V_{2\cdots s} = (1-q_s)^{-1}\frac{Q_s}{Q_{s-1}}V_{12\cdots s-1}.$$

We shall now indicate how (2.12) may be built up by successive inductions on the suffixes. Considering first the V_{k_1} we must prove that

$$\frac{V_{k_1}}{p_{k_1}} = \frac{V_1}{p_1}. \tag{2.13}$$

But this follows immediately from (2.6). Now assuming that the relations have been proved for the $V_{k_1 \cdots k_{s-1}}$, we proceed to establish them for the $V_{k_1 \cdots k_s}$. From (2.7),

$$V_{12\cdots s-1k_s} = (1-q_{s-1})[V_{12\cdots s-1k_s}+e^{i\theta(t)}V_{2\cdots s-1k_s}] * G(t) + p_{k_s}[V_{12\cdots s}+e^{i\theta(t)}V_{2\cdots s}] * G(t),$$

whence, using the induction hypothesis,

$$V_{12\cdots s-1k_s} = (1-q_{s-1})V_{12\cdots s-1k_s} * G(t) + p_{k_s}\left[V_{12\cdots s}+p_s^{-1}\frac{Q_s}{Q_{s-1}}e^{i\theta(t)}V_{12\cdots s-1}\right] * G(t).$$

Since $V_{12\cdots s}$ satisfies the same equation with $k_s = s$,

$$\frac{V_{12\cdots s-1k_s}}{p_{k_s}} - \frac{V_{12\cdots s}}{p_s} = (1-q_{s-1})\left[\frac{V_{12\cdots s-1k_s}}{p_{k_s}} - \frac{V_{12\cdots s}}{p_s}\right] * G(t),$$

so that

$$\frac{V_{12\cdots s-1k_s}}{p_{k_s}} = \frac{V_{12\cdots s}}{p_s},$$

which establishes (2.12) for the $V_{12\cdots s-1k_s}$. Similarly one can deduce the required relations for the $V_{12\cdots s-2k_{s-1}k_s}$, and so on for all the $V_{k_1\cdots k_s}$. We shall omit the details.

Equations (2.2), (2.4), (2.11) and (2.12) provide the required evaluation of the characteristic functional.

3. The distribution of $N(x, t)$

Calculations of product-densities for the process based on the characteristic functional obtained above would clearly involve extensive algebra, and will not be attempted here[1]. However, to give some idea of the type of algebra that would be encountered, we shall outline the evaluation of the characteristic function (1.3) of $N(x, t)$, and deduce the age-distribution of the ith ranked individual at any time t. The explicit formulae for the latter were in fact first obtained from the characteristic functional, although more direct approaches are possible in two particular cases (Gani [3]).

[1] See the author's note in *J. Appl. Prob.* 1, No. 1 (1964).

To find the V's for the particular $\theta(u)$ defined by (1.2), it is convenient to introduce the functions

$$A_s(t) = \begin{cases} V_{12\cdots s}(t), & (0 \leqq t \leqq x), \\ 0, & \text{elsewhere}; \end{cases} \tag{3.1}$$

$$B_s(t) = \begin{cases} V_{12\cdots s}(t+x), & (t > 0), \\ 0, & (t \leqq 0). \end{cases} \tag{3.2}$$

From (2.11), we find that these functions satisfy the recurrence relations

$$A_s(t) = e^{i\phi} \frac{Q_s}{Q_{s-1}} A_{s-1} * \gamma_s(t), \qquad (0 \leqq t \leqq x), \tag{3.3}$$

$$B_s(t) = \frac{Q_s}{Q_{s-1}} \left[e^{i\phi} \int_t^{t+x} A_{s-1}(t+x-\tau)\, d\gamma_s(\tau) + B_{s-1} * \gamma_s(t) \right], \qquad (t > 0). \tag{3.4}$$

Solving, and using the readily-proved formula

$$\gamma_{r_1} * \cdots * \gamma_{r_s}(t) = \sum_{k=1}^{s} \frac{\gamma_{r_k}(t)}{\prod\limits_{\substack{j=1 \\ j \neq k}}^{s} (q_{r_j} - q_{r_k})}, \qquad (0 \leqq r_1 < r_2 < \cdots < r_s \leqq n), \tag{3.5}$$

for convolutions of the γ's, it eventually follows that, for $s = 1, \cdots, n$,

$$A_s(t) = -e^{is\phi} Q_s \sum_{k=0}^{s} \frac{q_k \gamma_k(t)}{\prod\limits_{\substack{j=0 \\ j \neq k}}^{s} (q_j - q_k)}, \tag{3.6}$$

and

$$\begin{aligned} B_s(t) = e^{i\phi} Q_s \Bigg\{ & \int_t^{t+x} \frac{A_{s-1}(t+x-\tau)}{Q_{s-1}}\, d\gamma_s(\tau) \\ & + \sum_{r=2}^{s} \sum_{k=r}^{s} \frac{1}{\prod\limits_{\substack{j=r \\ j \neq k}}^{s} (q_j - q_k)} \int_0^t d\gamma_k(\tau) \int_{t-\tau}^{t+x-\tau} \frac{A_{r-2}(t+x-\tau-\sigma)}{Q_{r-2}}\, d\gamma_{r-1}(\sigma) \Bigg\} \\ & + Q_s \sum_{k=1}^{s} \frac{1}{\prod\limits_{\substack{j=1 \\ j \neq k}}^{s} (q_j - q_k)} \int_0^t V_0(t+x-\tau)\, d\gamma_k(\tau). \end{aligned} \tag{3.7}$$

Hence, for $0 \leqq x \leqq t$,

$$V_{12\cdots s}(t) = -e^{is\phi} Q_s \sum_{k=0}^{s} \frac{q_k \gamma_k(t)}{\prod\limits_{\substack{j=0 \\ j \neq k}}^{s} (q_j - q_k)}, \qquad (s = 1, \cdots, n), \tag{3.8}$$

while for $t > x$,

$$V_{12\cdots s}(t) = e^{i\phi} Q_s \Bigg\{ \int_{t-x}^{t} \frac{A_{s-1}(t-\tau)}{Q_{s-1}} d\gamma_s(\tau)$$

(3.9)
$$+ \sum_{r=2}^{s} \sum_{k=r}^{s} \frac{1}{\prod_{\substack{j=r \\ j\neq k}}^{s} (q_j - q_k)} \int_0^{t-x} d\gamma_k(\tau) \int_{t-x-\tau}^{t-\tau} \frac{A_{r-2}(t-\tau-\sigma)}{Q_{r-2}} d\gamma_{r-1}(\sigma) \Bigg\}$$

$$+ Q_s \sum_{k=1}^{s} \frac{1}{\prod_{\substack{j=1 \\ j\neq k}}^{s} (q_j - q_k)} \int_0^{t-x} V_0(t-\tau) d\gamma_k(\tau), \qquad (s = 1, \cdots, n),$$

remembering that V_0 is given by (2.5).

We may now calculate the characteristic function for $N(x, t)$. From (2.2) and (2.4) we have, for $0 < x < t$,

$$C(\phi, t) = V_0(t) + \sum_{(k)} V_{k_1 \cdots k_s}(t).$$

After some algebra, it follows from (2.12) that

(3.10)
$$C(\phi, t) = \sum_{s=0}^{r} \left(\prod_{h=0}^{s-1} q_h \right) Q_s^{-1} V_{12\cdots s}(t), \qquad (0 < x < t).$$

Taking $x_0 = 0$, $x_{n+1} = +\infty$, we find similarly that for $t+x_m \leqq x < t+x_{m+1}$ $(m = 0, \cdots, n)$

$$C(\phi, t) = e^{im\phi} \Bigg[V_0(t) + \sum_{s=1}^{n} \sum_{k_1 > m} V_{k_1 \cdots k_s}(t)$$
$$+ \sum_{l=1}^{m} e^{-il\phi} \sum_{s=l}^{n+l-m} \sum_{k_1 \cdots k_l = 1}^{m} \sum_{k_{l+1} > m} V_{k_1 \cdots k_s}(t) \Bigg],$$

which reduces to

(3.11)
$$C(\phi, t) = \sum_{l=0}^{m} e^{il\phi} \sum_{s=m-l}^{n-l} \left(\prod_{h=m}^{l+s-1} q_h \right)$$
$$\cdot \left\{ Q_s^{-1} V_{12\cdots s}(t) \sum_{k_1, \cdots, k_{m-l}=1}^{m} \left[\prod_{r=1}^{m-l} (q_{k_r - 1} - q_{k_r + s - r}) \right] \right\}.$$

Substituting (3.9) in (3.10), it may be shown that if $0 < x < t$, then

$$C(\phi, t) = V_0(t) + \int_0^{t-x} V_0(t-\tau) d\gamma_n(\tau)$$
$$+ e^{i\phi} \left[\int_{t-x}^{t} V_0(t-\tau) d\gamma_1(\tau) + q_1 \int_0^{t-x} d\gamma_n(\tau) \int_{t-x-\tau}^{t-\tau} V_0(t-\tau-\sigma) d\gamma_1(\sigma) \right]$$
$$+ \sum_{s=2}^{n} \left(\prod_{h=0}^{s-1} q_h \right) \left[\int_{t-x}^{t} \frac{A_{s-1}(t-\tau)}{Q_{s-1}} d\gamma_s(\tau) \right.$$
$$\left. + q_s \int_0^{t-x} d\gamma_n(\tau) \int_{t-x-\tau}^{t-\tau} \frac{A_{s-1}(t-\tau-\sigma)}{Q_{s-1}} d\gamma_s(\sigma) \right],$$

which from (3.5) and (3.6) reduces to

$$
(3.12)\qquad \begin{aligned} C(\phi, t) &= V_0(t) + \int_0^{t-x} V_0(t-\tau)d\gamma_n(\tau) \\ &\quad - \sum_{s=1}^{n} e^{is\phi}\left(\prod_{h=0}^{s-1} q_h\right)\sum_{\kappa=0}^{s} \frac{q_k}{\prod\limits_{\substack{j=0\\ j\neq k}}^{s}(q_j-q_k)}\int_{t-x}^{t}\{1-q_k\gamma_k(t-\tau)\}d\gamma_n(\tau). \end{aligned}
$$

On the other hand, if $t+x_m \leqq x < t+x_{m+1}$ $(m = 0, \cdots, n)$,

$$
(3.13)\qquad C(\phi, t) = e^{im\phi} - \sum_{s=m}^{n} e^{is\phi}\left(\prod_{h=m}^{s-1} q_h\right)\sum_{k=m}^{s}\frac{q_k\gamma_k(t)}{\prod\limits_{\substack{j=m\\ j\neq k}}^{s}(q_j-q_k)}.
$$

Hence, writing

$$
(3.14)\qquad p_s(x, t) = Pr\{N(x, t) = s\},
$$

the probability that at time t there are s individuals with ages $\leqq x$, it follows from (3.12) and (3.13) that

$$
(3.15)\qquad p_0(x, t) = \begin{cases} 1 - \int_{t-x}^{t}\{1-q_0\gamma_0(t-\tau)\}d\gamma_n(\tau), & (0 < x < t), \\ 1 - q_0\gamma_0(t), & (t \leqq x < t+x_1), \\ 0, & (x \geqq t+x_1), \end{cases}
$$

while, for $s = 1, \cdots, n$,

$$
(3.16)\qquad p_s(x, t) = \begin{cases} -\left(\prod_{h=0}^{s-1} q_h\right)\sum_{k=0}^{s}\frac{q_k}{\prod\limits_{\substack{j=0\\ j\neq k}}^{s}(q_j-q_k)}\int_{t-x}^{t}\{1-q_k\gamma_k(t-\tau)\}d\gamma_n(\tau), & \\ & (0 < x < t), \\ -\left(\prod_{h=m}^{s-1} q_h\right)\sum_{k=m}^{s}\frac{q_k\gamma_k(t)}{\prod\limits_{\substack{j=m\\ j\neq k}}^{s}(q_j-q_k)}, & (t+x_m \leqq x < t+x_{m+1}; \\ & m = 0, \cdots, s-1), \\ 1-q_s\gamma_s(t), & (t+x_s \leqq x < t+x_{s+1}), \\ 0, & (x \geqq t+x_{s+1}), \end{cases}
$$

noting that $p_n(x, t) = 1$ for $x \geqq t+x_n$.

The age-distribution of the ith ranked individual at time t is given by

$$
(3.17)\qquad \begin{aligned} f_i(x, t) &= Pr\,\{i\text{th ranked individual has age} \leqq x \text{ at time } t\} \\ &= \sum_{s=i}^{n} p_s(x, t). \end{aligned}
$$

This distribution is found to be

$$(3.18)\qquad f_i(x,t)=\begin{cases}\displaystyle\sum_{k=0}^{i-1}\frac{q_k}{\prod_{\substack{j=0\\ j\neq k}}^{i-1}(1-q_k/q_j)}\int_{t-x}^{t}\{1-q_k\gamma_k(t-\tau)\}d\gamma_n(\tau), & (0<x<t),\\ \displaystyle\sum_{k=m}^{i-1}\frac{q_k\gamma_k(t)}{\prod_{\substack{j=m\\ j\neq k}}^{i-1}(1-q_k/q_j)}, & (t+x_m\leqq x<t+x_{m+1};\ m=0,\cdots,i-1),\\ 1, & (x\geqq t+x_i),\end{cases}$$

for $i=1,\cdots,n$.

In particular, if $n=1$ and $p_1=1$ we have a renewal process in which the initial component has age x at $t=0$, and its residual useful life has distribution $G(t)$. Also, $\gamma_0(t)=G(t)$, while $\gamma_1(t)=\sum_{m=1}^{\infty}G^{m*}(t)=H(t)$ is the renewal function. From (3.18), the age-distribution of the article in use at time t is

$$(3.19)\qquad f_1(x,t)=\begin{cases}\int_{t-x}^{t}\{1-G(t-\tau)\}dH(t), & (0<x<t),\\ G(t), & (t\leqq x<t+x_1),\\ 1, & (x\geqq t+x_1),\end{cases}$$

which agrees with Smith [4] equation (5.2).

4. The Poisson case

If the replacement process is Poisson, with $G(t)=1-e^{-\lambda t}$ $(t\geqq 0)$, we readily find that if $k=0,\cdots,n-1$,

$$(4.1)\qquad \begin{aligned}\gamma_k(t)&=q_k^{-1}(1-e^{-\lambda q_k t})\\ 1-q_k\gamma_k(t)&=e^{-\lambda q_k t},\end{aligned}$$

for $t\geqq 0$, these functions vanishing for $t<0$.

On the other hand, $\gamma_n(t)=\sum_{m=1}^{\infty}G^{m*}(t)$, the renewal function for the replacements, is given by

$$(4.2)\qquad \gamma_n(t)=\begin{cases}\lambda t, & t\geqq 0,\\ 0, & t<0.\end{cases}$$

From (4.1) and (4.2), the system of probabilities (3.15) and (3.16) reduce to

$$p_s(x, t) = \begin{cases} \left(\prod_{h=0}^{s-1} q_h\right) \sum_{k=0}^{s} \frac{e^{-\lambda q_k x}}{\prod_{\substack{j=0 \\ j \neq k}}^{s} (q_j - q_k)}, & (0 < x < t), \\ \left(\prod_{h=m}^{s-1} q_h\right) \sum_{k=m}^{s} \frac{e^{-\lambda q_k t}}{\prod_{\substack{j=m \\ j \neq k}}^{s} (q_j - q_k)}, & (t+x_m \leqq x < t+x_{m+1};\ m = 0, \cdots, s), \\ 0, & (x \geqq t+x_{s+1}), \end{cases} \tag{4.3}$$

for $s = 0, \cdots, n$, with $p_n(x, t) = 1$ for $x \geqq t+x_n$.

The age distribution of the ith ranked individual at time t is found to be:

$$f_i(x, t) = \begin{cases} 1 - \sum_{k=0}^{i-1} \frac{e^{-\lambda q_k x}}{\prod_{\substack{j=0 \\ j \neq k}}^{i-1} (1 - q_k/q_j)}, & (0 < x < t), \\ 1 - \sum_{k=m}^{i-1} \frac{e^{-\lambda q_k t}}{\prod_{\substack{j=m \\ j \neq k}}^{i-1} (1 - q_k/q_j)}, & (t+x_m \leqq x < t+x_{m+1};\ m = 0, \cdots, i-1), \\ 1, & (x \geqq t+x_i). \end{cases} \tag{4.4}$$

This result has also been obtained by Gani ([3], equation 1.10), using a more direct argument.

5. The limiting distributions

It is easily seen from (2.10) that the $\gamma_k(t)$ are non-decreasing functions of t. Furthermore, as $t \to \infty$,

$$\gamma_k(t) \to q_k^{-1}, \qquad (k = 0, \cdots, n-1), \tag{5.1}$$

while, by the elementary renewal theorem,

$$\gamma_n(t) \sim t/\mu, \tag{5.2}$$

where μ is the mean time between replacements.

Hence if we define

$$\Gamma_k(u) = \begin{cases} 1 - q_k \gamma_k(u), & (0 \leqq u < x) \\ 0, & \text{elsewhere}, \end{cases} \tag{5.3}$$

$\Gamma_k(u)$ is certainly of bounded variation in the finite interval $0 \leqq u < x$, and so, by Corollary (1.1) of Smith's paper [4],

$$\lim_{t\to\infty}\int_{t-x}^{t}\{1-q_k\gamma_k(t-\tau)\}d\gamma_n(\tau)$$

(5.4)
$$= \lim_{t\to\infty} \Gamma_k * \gamma_n(t)$$
$$= \mu^{-1}\int_0^\infty \Gamma_k(u)du$$
$$= \mu^{-1}\int_0^x \{1-q_k\gamma_k(u)\}du.$$

For convenience, we have assumed that $G(t)$ is a nonlattice distribution.

Hence, the limiting distribution of $N(x, t)$ is, from (3.15) and (3.16),

(5.5)
$$p_0(x, \infty) = 1-\mu^{-1}\int_0^x \{1-q_0\gamma_0(u)\}du,$$
$$p_s(x, \infty) = -\mu^{-1}\left(\prod_{h=0}^{s-1} q_h\right)\sum_{k=0}^{s}\frac{q_k}{\prod_{\substack{j=0\\ j\neq k}}^{s}(q_j-q_k)}\int_0^x \{1-q_k\gamma_k(u)\}du, \qquad (s = 1, \cdots, n).$$

The limiting age-distribution of the ith ranked individual follows from (3.18):

(5.6)
$$f_i(x, \infty) = \mu^{-1}\sum_{k=0}^{i-1}\frac{q_k}{\prod_{\substack{j=0\\ j\neq k}}^{i-1}(1-q_k/q_j)}\int_0^x \{1-q_k\gamma_k(u)\}du.$$

In the Poisson case, this becomes

(5.7)
$$f_i(x, \infty) = 1-\sum_{k=0}^{i-1}\frac{e^{-\lambda q_k x}}{\prod_{\substack{j=0\\ j\neq k}}^{i-1}(1-q_k/q_j)}.$$

Finally, we shall relate the limiting distribution (5.7) to the result obtained by Gani and Yeo ([2], p. 59) for the stationary age-distribution $F_i(x)$ immediately after a regeneration point. Using equations (13) of [2] and (1.2) of [3], Mr. Yeo has shown that

(5.8)
$$f_i(x, \infty) = F_i(x) * (1-e^{-\lambda x}),$$

a relation which is intuitively obvious. Taking $F_i(x)$ in the form [1]

[1] The gamma-type terms appearing in Gani and Yeo's expression for $F_i(x)$ are redundant, as may be seen by substituting $\varphi(\theta) = \mu/(\mu+\theta)$ in their equation (14) for the Laplace transform of $F_i(x)$.

$$F_i(x) = 1 - \sum_{r=1}^{i-1} \frac{e^{-\lambda q_r x}}{\prod_{\substack{j=1 \\ j \neq r}}^{i-1} (1 - q_r/q_j)}, \tag{5.9}$$

(noting that q_i has been redefined), we may readily deduce (5.7) from (5.8).

The author is indebted to Dr. J. Gani for having suggested the problems treated in this paper, and to Mr. G. F. Yeo for some useful discussions.

References

[1] Bartlett, M. S. and Kendall, D. G., On the use of the characteristic functional in the analysis of some stochastic processes occurring in physics and biology, *Proc. Camb. Phil. Soc.* 47 (1951), 65—76.

[2] Gani, J. and Yeo, G. F., On the age-distribution of n ranked elements after several replacements, *Aust. J. Statist.* 4 (1962), 55—60.

[3] Gani, J., On the age distribution of replaceable ranked elements, *to appear*.

[4] Smith, W. L., Asymptotic renewal theorems, *Proc. Roy. Soc. Edinburgh*, Sec. A, 64 (1954), 9—48.

Australian National University,
Canberra.

2

Reprinted from JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS Vol. 18, No 2 May 1967

Branching-Diffusion Processes with No Absorbing Boundaries. I.

A. W. DAVIS

*C.S.I.R.O., Division of Mathematical Statistics, Adelaide, South Australia**

Submitted by Richard Bellman

1. INTRODUCTION

Several papers have appeared recently on the asymptotic properties of stochastic populations which multiply and "diffuse" randomly throughout a given region. Interest has centred principally on cases in which the region has an absorbing boundary, and results have been obtained for the asymptotic behavior of the population size (Sevast'yanov [1], [2], Conner [3], [4]), and more recently for the number of individuals in a subset of the region (Watanabe [5]). Roughly speaking, the mainspring of these investigations is the existence of eigenvalues of maximum modulus, and corresponding positive bounded eigenfunctions, for positive operators derived from the mean distribution of the population.

The present author follows Adke and Moyal ([6], [7], [8]) in studying simpler branching-diffusion processes in which there are no absorbing boundaries, and the branching mechanism is independent of location. This provides some scope for a more concrete discussion of the "spatial" properties of the population. The motivating problem has been to investigate certain natural measures of average position and dispersion. We are therefore led to consider possibly unbounded functions of position whose asymptotic behavior may depend on higher eigenvalues of the mean distribution, and we are unable to rely upon positivity and boundedness properties. Certain rather "rough and ready" restrictions are placed on these functions in order to develop an asymptotic theory which applies to some simple cases of interest.

The present paper summarizes and slightly extends the results of the author's earlier paper [9]. In Section 2 the model is introduced, and in the following section the defining integral equation for the characteristic functional of the process is stated. In Section 4, recurrence relations are obtained

* The work reported here was carried out during the tenure of a research scholarship at the Australian National University, Canberra.

for the moment distributions, and sufficient conditions are set up for the existence of moment functionals. After a summary of the Bellman-Harris asymptotic theory, further restrictions are placed on the functions considered, and a mean-square convergence theory is presented for the branching-diffusion process in Section 5. The treatment of [9] is extended in that allowance is made for possible time-dependent transformations of the position variables. Applications of the theory are given in the final section.

In the accompanying paper [10], we shall consider the quotient random variable $N(\xi_t \mid x_0; t)/N(t)$ which leads to results for the average position and dispersion of the population. It is assumed in both papers that $m > 1$, m being the expected number of offspring at a birth. Some asymptotic results for $m \leqslant 1$ will, it is hoped, be presented elsewhere.

2. The Model

We shall assume that the population diffuses through a region $\mathscr{X}$ with no absorbing boundaries, and multiplies according to the Bellman-Harris age-dependent branching process ([11], Chapter VI); i.e., at time $t = 0$ the population consists of a single ancestor aged zero, situated at $x_0 \in \mathscr{X}$, having a random life-length ℓ whose distribution function

$$G(t) = \text{Prob}\,\{\ell \leqslant t\} \tag{2.1}$$

is continuous on the right, with $G(0+) = 0$. Suppose that at the end of its life the ancestor is at the point x. Then it is replaced by n similar and indistinguishable offspring with probability p_n ($\sum_0^\infty p_n = 1$), all of age zero, and situated at the same point x initially. Each of the offspring then proceeds to diffuse and to generate a subpopulation independently of the others.

We shall let

$$f(z) = \sum_0^\infty p_n z^n \qquad (|z| \leqslant 1) \tag{2.2}$$

denote the probability generating function of the p_n, which are assumed to be constant. The factorial moments of $\{p_n\}$ will be denoted by

$$m = f'(1-), \qquad m_{(n)} = f^{(n)}(1-), \qquad (n = 2, 3, \ldots). \tag{2.3}$$

In order to ensure that $N(t)$, the population size at time t, is finite with probability one, we shall take $m < \infty$ throughout.

The region $\mathscr{X}$ is assumed to be a locally compact Hausdorff space, with a σ-field $\mathscr{B}_x$ generated by the Borel subsets X of $\mathscr{X}$. The time axis $[0, \infty)$ will be denoted by $\mathscr{T}$, and $\mathscr{B}_t$ will denote the ordinary Borel σ-field of subsets of $\mathscr{T}$.

The "diffusion" of each individual through $\mathscr{X}$ will be determined by a time-homogeneous Markovian transition probability $\chi(X \mid x; t)$, which represents the probability that the individual will move from $x \in \mathscr{X}$ to a point in $X \in \mathscr{B}_x$ during a time interval of length t, conditional upon its survival during this interval.

Let $\mathscr{S} = \mathscr{X} \times \mathscr{T}$, and let $\mathscr{B}_s = \mathscr{B}_x \times \mathscr{B}_t$ denote the minimal σ-field of subsets of $\mathscr{S}$ containing all rectangle sets $X \times T$ $(X \in \mathscr{B}_x, T \in \mathscr{B}_t)$. Then χ has the following properties:

(a) $\chi(X \mid x; t)$ is a $\mathscr{B}_s$-measurable function on $\mathscr{S}$ for fixed $X \in \mathscr{B}_x$, and a probability distribution on $\mathscr{B}_x$ for fixed $(x, t) \in \mathscr{S}$: $\chi(\mathscr{X} \mid x; t) = 1$.

(b) χ satisfies the Chapman-Kolmogorov equation:

$$\int \chi(X \mid x; u)\, \chi(dx \mid x_0; v) = \chi(X \mid x_0; u + v),$$
$$(u > 0, v > 0, X \in \mathscr{B}_x, x_0 \in \mathscr{X}). \tag{2.4}$$

(c) $$\chi(X \mid x; 0) = \delta(X \mid x), \qquad (x \in \mathscr{X}, X \in \mathscr{B}_x), \tag{2.5}$$

where $\delta(X \mid \cdot)$ is the indicator function of the set X.

Although χ is Markovian, it is seen that the process will be non-Markovian except in the case of negative-exponential G.

3. The Characteristic Functionals

The process defined in the preceding section is a particular case of the stochastic population processes with general individual state space $\mathscr{X}$ whose theory has been given by Moyal [12]. If at any given instant the population consists of n indistinguishable individuals situated at points $x_1, \ldots, x_n \in \mathscr{X}$, then it is convenient to identify the state of the system with the set of all distinct ordered n-tuples $x^{(n)} = (x_1, \ldots, x_n)$ obtained by permuting $x_1, \ldots, x_n$. We may then define the population state space to be $\Omega = \sum_0^\infty \mathscr{X}^n$, $\mathscr{X}^0$ denoting the empty set. On each $\mathscr{X}^n$ we define the minimal σ-field $\mathscr{B}_x{}^n$ containing all product sets $X_1 \times \cdots \times X_n$ $(X_i \in \mathscr{B}_x)$, and on Ω the minimal σ-field $\mathscr{U}$ containing all the sets of the $\mathscr{B}_x{}^n$ $(n = 1, 2, \ldots)$. The population is specified at any instant t by a symmetric conditional probability measure $P(\cdot \mid x_0; t)$ on $\mathscr{U}$; by symmetric, we mean that P is invariant under all coordinate permutations in Ω. $P^{(n)}$ will denote the restriction of P to $\mathscr{X}^n$.

Associated with a population process is the *counting process* $N(X \mid x_0; t)$ which denotes the number of individuals with states in $X \in \mathscr{B}_x$ at time t, conditional upon an ancestor in state x_0 at time $t = 0$.

Now let $\mathcal{M}$ be the class of all real finite-valued $\mathcal{B}_x$-measurable functions $\xi(x)$ on $\mathcal{X}$. If at time t there are $N(t) = N(\mathcal{X} \mid x_0; t)$ individuals in the population, states $x_1, \ldots, x_{N(t)}$, then

$$N(\xi \mid x_0; t) = \sum_{1}^{N(t)} \xi(x_i) = \int_{\mathcal{X}} \xi(x)\, N(dx \mid x_0; t), \qquad (\xi \in \mathcal{M}), \tag{3.1}$$

is the *counting functional* at time t. This may be used to introduce the *characteristic functional* for the process at time t:

$$\begin{aligned} \Gamma[\xi \mid x_0; t] &= \mathcal{E} \exp \{iN(\xi \mid x_0; t)\} \\ &= \sum_{n=0}^{\infty} \int_{\mathcal{X}^n} \exp \left\{ i \sum_{j=1}^{n} \xi(x_j) \right\} P^{(n)}(dx^{(n)} \mid x_0; t). \end{aligned} \tag{3.2}$$

On the basis of the description given in Section 2, we *define* the characteristic functional for our branching-diffusion process to be a solution of the integral equation:

$$\begin{aligned} \Gamma[\xi \mid x_0; t] &= [1 - G(t)] \int_{\mathcal{X}} e^{i\xi(x)}\, \chi(dx \mid x_0; t) \\ &+ \iint_{\mathcal{X}\times[0,t]} f\{\Gamma[\xi \mid x; t-u]\}\, \chi(dx \mid x_0; u)\, G(du), \qquad (\xi \in \mathcal{M}), \end{aligned} \tag{3.3}$$

such that $|\Gamma| \leqslant 1$. It was shown in [9], Theorem 3.1, that there is exactly one such solution, and that the latter defines a probability measure $P(\cdot \mid x_0; t)$ on $\mathcal{U}$ to which it is related by Eq. (3.2).

Taking $\xi(x) = \theta$, (real), we see that the characteristic functional reduces to the characteristic function of the population size $N(t)$ for the Bellman-Harris process. This will be denoted by $\Gamma[\theta; t]$, and is the unique solution of the integral equation

$$\Gamma[\theta; t] = [1 - G(t)]\, e^{i\theta} + \int_{[0,t]} f\{\Gamma[\theta; t-u]\}\, G(du) \tag{3.4}$$

such that $|\Gamma| \leqslant 1$.

In order to completely specify the evolution of the process, it is necessary to define a conditional probability distribution on $\mathcal{U}^{\mathcal{T}}$, the product σ-field on $\Omega^{\mathcal{T}}$, which is the space of all realizations of the process. This is achieved by defining joint characteristic functionals at all finite subsets of $\mathcal{T}$ (see [9], Section 3.)

P. E. Ney ([13], [14]) has considered a model which has some points of similarity with the above. The branching mechanism is also taken to be the Bellman-Harris model, while $\mathcal{X}$ is the non-negative real axis. It is assumed that the initial states $x_1, \ldots, x_n$ of the offspring of a parent in state z have the conditional joint distribution $P(x_1, \ldots, x_n \mid z)$, the state of each individual then remaining fixed throughout its lifetime. By imposing certain restrictions on P, Ney obtains asymptotic results on $\sum_1^{N(t)} x_j$, i.e., $N(\xi \mid x_0; t)$ with $\xi(x) \equiv x$.

4. The Moment Functionals

The population mean size for the Bellman-Harris process,

$$M(t) = \mathscr{E}N(t) \tag{4.1}$$

is the unique solution of the renewal equation

$$M(t) = [1 - G(t)] + m \int_{[0,t]} M(t-u)\, G(du), \tag{4.2}$$

which is bounded on each finite t-interval. Equation (4.2) is obtained by differentiation of (3.4), and its solution may be written in the form

$$M(t) = 1 + (m-1)\, H(t), \tag{4.3}$$

where

$$H(t) = \sum_1^{\infty} m^{n-1} G_n(t), \tag{4.4}$$

$G_n(t)$ being the nth convolution of $G(t)$.

If $m_{(n)}$ is finite (see (2.3)), then all the moments

$$M_k(t) = \mathscr{E}N^k(t) \tag{4.5}$$

are finite for $k = 1, \ldots, n$, and recurrence relations may be obtained for them from (3.4).

The *mean distribution* of a stochastic population process is a measure on $\mathscr{B}_x$ defined by

$$\begin{aligned} M(X \mid x_0; t) &= \mathscr{E}N(X \mid x_0; t) \\ &= \left\{ \frac{\partial}{i\partial\theta}\, \Gamma[\theta\delta(X \mid \cdot) \mid x_0; t] \right\}_{\theta=0}. \end{aligned} \tag{4.6}$$

For the present process, it follows from (3.3) that

$$\begin{aligned} M(X \mid x_0; t) = {} & [1 - G(t)]\, \chi(dx \mid x_0; t) \\ & + m \iint_{\mathscr{X}\times[0,t]} M(X \mid x; t-u)\, \chi(dx \mid x_0; u)\, G(du). \end{aligned} \tag{4.7}$$

This equation has the form of the generalized renewal equation

$$R(x_0, t) = F(x_0, t) + m \iint_{\mathscr{X}\times[0,t]} R(x; t-u)\, \chi(dx \mid x_0; u)\, G(du), \tag{4.8}$$

where R and F are members of the class $\mathscr{F}$ consisting of all measurable functions on $\mathscr{S} = \mathscr{X} \times \mathscr{T}$ which are bounded on each cylinder set $\mathscr{X} \times T$, where T is a finite t-interval. The following result was proved in [9] Lemma 4.1 :

THEOREM 1. *If $F \in \mathscr{F}$, then Eq. (4.8) has exactly one solution R in $\mathscr{F}$, namely,*

$$R(x_0, t) = F(x_0, t) + m \iint_{\mathscr{X}\times[0,t]} F(x, t-u)\, \chi(dx \mid x_0; u)\, H(du), \tag{4.9}$$

where H was defined in (4.4).

The Chapman-Kolmogorov relation (2.4) for χ is central in the proof of Theorem 1.

Solving (4.7), we obtain

$$M(X \mid x_0; t) = M(t)\, \chi(X \mid x_0; t). \tag{4.10}$$

As is characteristic of branching-diffusion processes, the mean distribution is proportional to the transition probability χ at any given instant.

More generally, the nth *moment distribution* $M_n(\cdot \mid x_0; t)$ at time t, $(n = 1, 2, \ldots)$, is defined at each product set

$$X^{(n)} = X_1 \times \cdots \times X_n \in \mathscr{B}_x{}^n (X_j \in \mathscr{B}_x)$$

by

$$M_n(X^{(n)} \mid x_0; t) = \mathscr{E} \prod_{j=1}^{n} N(X_j \mid x_0; t)$$

$$= \left\{ \frac{\partial^n}{i^n \partial\theta_1 \cdots \partial\theta_n} \Gamma\left[\sum_{j=1}^{n} \theta_j \delta(X_j \mid \cdot) \mid x_0; t \right] \right\}_{\theta_1 = \ldots = \theta_n = 0}, \tag{4.11}$$

and may be extended to a symmetric non-negative measure on $\mathscr{B}_x{}^n$. If $m_{(n)} < \infty$, then $M_k(\cdot \mid x_0; t)$ is necessarily a finite measure for $k = 1, \ldots, n$. Integral equations defining each M_k in terms of earlier moment distributions may be obtained by differentiation of (3.3). These have the general form (4.8), and may be solved by Theorem 1. We first note that

$$\frac{\partial^k}{\partial\theta_1 \cdots \partial\theta_k} f[g(\theta_1, \ldots, \theta_k)] = \sum_{j=1}^{k} f^{(j)}[g] \sum_{\Pi_j(k)} \left\{ \prod_{h=1}^{j} \frac{\partial^{r_h} g}{\partial\theta^{(\pi_h)}} \right\}, \tag{4.12}$$

where the inner summation is extended over all (unordered) partitions $\prod_j(k)$ of the set $\{1, 2,..., k\}$ into j nonempty classes $\pi_1,..., \pi_j$ of sizes $r_1,..., r_j$, respectively. The symbol $\partial^{r_h}g/\partial\theta^{(\pi_h)}$ denotes the r_h-fold partial derivative of g with respect to those θ_i with suffixes in π_h. Now letting $X^{(\pi_h)}$ be the cartesian product of the X_i whose suffixes are members of π_h, it may be shown that:

THEOREM 2. *If $m_{(n)} < \infty$, then for $k = 2,..., n$,*

$$M_k(X^{(k)} \mid x_0 ; t) = M\left(\bigcap_{j=1}^{k} X_j \mid x_0 ; t\right)$$

$$+ \sum_{j=2}^{k} m_{(j)} \sum_{\Pi_j(k)} \iint_{\mathscr{X}\times[0,t]} \prod_{h=1}^{j} M_{r_h}(X^{(\pi_h)} \mid x; t-u)\, \chi(dx \mid x_0; u)\, H(du) \quad (4.13)$$

defines a finite measure on $\mathscr{B}_x{}^k$.

Let $\xi^{(n)} = (\xi_1,..., \xi_n) \in \mathscr{M}^n$. We define the nth *moment functional* at $\xi^{(n)}$ by:

$$M_n(\xi^{(n)} \mid x_0; t) = \mathscr{E} \prod_{j=1}^{n} N(\xi_j \mid x_0; t) = \int_{\mathscr{X}^n} \prod_{j=1}^{n} \xi_j(x_j)\, M_n(dx^{(n)} \mid x_0; t), \quad (4.14)$$

provided that the integral with respect to the nth moment distribution is absolutely convergent. If $\xi_1 = \xi_2 = \cdots = \xi_n, = \xi$ say, then we shall simply write

$$M_n(\xi \mid x_0; t) = \mathscr{E}N^n(\xi \mid x_0; t). \quad (4.15)$$

As a first step toward setting up sufficient conditions for the existence of the nth moment functional, let us define for each $p \geqslant 1$:

$$\mathscr{L}_p = \bigcap_{(x,t)\in\mathscr{S}} L_p\{\chi(\cdot \mid x; t)\}, \quad (4.16)$$

where $L_p\{\chi(\cdot \mid x; t)\}$ is the Banach space of real $\mathscr{B}_x$-measurable functions $\xi(x)$ on $\mathscr{X}$ for which

$$\| \xi \|_{(x,t)}^{(p)} = \left\{\int_{\mathscr{X}} | \xi(y) |^p \chi(dy \mid x; t)\right\}^{1/p} < \infty. \quad (4.17)$$

From (2.5), each member of $\mathscr{L}_p$ is finite-valued, and so is a member of $\mathscr{M}$. Since χ is a probability distribution, it follows that if $p \geqslant r \geqslant 1$, then

$$\| \xi \|_{(x,t)}^{(r)} \leqslant \| \xi \|_{(x,t)}^{(p)}, \qquad ((x, t) \in \mathscr{S}), \quad (4.18)$$

and so

$$\mathscr{L}_p \subseteq \mathscr{L}_r, \qquad (p \geqslant r \geqslant 1). \tag{4.19}$$

Next let $\mathscr{K}_p$ be the linear space of all $\xi \in \mathscr{L}_p$ such that

$$\| \xi \|_{(x,t)}^{(p)} \leqslant \gamma_\xi^{(p)}(x)\, \delta_\xi^{(p)}(t), \qquad ((x, t) \in \mathscr{S}), \tag{4.20}$$

where $\gamma_\xi^{(p)}(x)$ is a finite-valued function on $\mathscr{X}$, and $\delta_\xi^{(p)}(t)$ is a $\mathscr{B}_t$-measurable function bounded on each finite t-interval. Clearly, these functions are not unique for a given ξ. From (4.18),

$$\mathscr{K}_p \subseteq \mathscr{K}_r, \qquad (p \geqslant r \geqslant 1). \tag{4.21}$$

Writing $\mathscr{K}_p(\gamma; 1) \equiv \mathscr{K}_p$, we may successively define $\mathscr{K}_p(\gamma; n)$ $(n = 2, 3,...)$ to be the linear subspace of $\mathscr{K}_p(\gamma; n-1)$ consisting of those $\xi(x)$ having a $\gamma_\xi^{(p)}$ also in $\mathscr{K}_p(\gamma; n-1)$. $\mathscr{K}_p(\gamma; 0)$ will denote the class of all real finite-valued functions on $\mathscr{X}$. From (4.21)

$$\mathscr{K}_p(\gamma; m) \subseteq \mathscr{K}_r(\gamma; n), \qquad (p \geqslant r \geqslant 1, m \geqslant n). \tag{4.22}$$

If $\xi(x)$ is a bounded $\mathscr{B}_x$-measurable function on $\mathscr{X}$, then trivially $\xi \in \mathscr{K}_p(\gamma; n)$ for all $p \geqslant 1$, $n = 0, 1, 2,...$.

In order to obtain satisfactory asymptotic results for $N(\xi \mid x_0; t)$, it may be necessary to apply a time-dependent transformation to the location variable x. We therefore introduce $\mathscr{B}_s$-measurable functions $\xi_t(x)$ on $\mathscr{X} \times \mathscr{T}$. Let $\Lambda_p(\gamma; n)$, $(p \geqslant 1; n = 1, 2,...)$ be the class of $\xi_t(x)$ such that, for fixed t, $\xi_t(\cdot) \in \mathscr{M}$, and

$$\| \xi_{t+\tau} \|_{(x,t)}^{(p)} \leqslant g_\xi^{(p)}(x)\, d_\xi^{(p)}(t+\tau), \qquad (t, \tau \geqslant 0, x \in \mathscr{X}), \tag{4.23}$$

where $g_\xi^{(p)} \in \mathscr{K}_p(\gamma; n-1)$ and $d_\xi^{(p)}(t)$ is measurable and bounded on each finite interval in $\mathscr{T}$. It is easily seen that

$$\mathscr{K}_p(\gamma; n) \subseteq \Lambda_p(\gamma; n), \tag{4.24}$$

in the sense that we may identify $\xi(x)$ with $\xi_t(x)$ if $\xi_t \equiv \xi$ for all t. Furthermore,

$$\Lambda_p(\gamma; m) \subseteq \Lambda_r(\gamma; n), \qquad (p \geqslant r \geqslant 1, m \geqslant n). \tag{4.25}$$

We shall write

$$\xi_t^{(k)} = (\xi_{1,t}, ..., \xi_{k,t}). \tag{4.26}$$

THEOREM 3. *Let* $m_{(n)} < \infty$, *and* $\xi_{j,t} \in \Lambda_n(\gamma; n)$, $(j = 1,..., n)$. *Then the* $M_k(\xi_{t+\tau}^{(k)} \mid x_0; t)$ *exist for* $k = 1,..., n$ *and* $(x_0, t) \in \mathscr{S}$. *For* $k = 1$,

$$M(\xi_{t+\tau} \mid x_0; t) = M(t) \int_{\mathscr{X}} \xi_{t+\tau}(x)\, \chi(dx \mid x_0; t), \tag{4.27}$$

while for $k = 2,..., n$ *they are given recursively by (4.13) with* $\bigcap_{j=1}^{k} X_j$ *replaced by* $\prod_{j=1}^{k} \xi_{j,t+\tau}$, *and* $X^{(\pi_h)}$ *by* $\xi_{t+\tau}^{(\pi_h)}$.

We shall merely indicate the proof for $n = 2$. If $\xi_t \in \Lambda_2(\gamma; 2)$, then taking $k = 1$:

$$\begin{aligned} M(\mid \xi_{t+\tau} \mid \mid x_0; t) &= M(t) \parallel \xi_{t+\tau} \parallel_{(x_0,t)}^{(1)} \\ &\leqslant M(t) \parallel \xi_{t+\tau} \parallel_{(x_0,t)}^{(2)} \\ &\leqslant g_\xi^{(2)}(x_0) \cdot M(t)\, d_\xi^{(2)}(t+\tau), \end{aligned} \tag{4.28}$$

which is finite. Hence, using Hölder's inequality:

$$\begin{aligned} M_2(\mid \xi_{t+\tau}^{(2)} \mid \mid x_0; t) &= M(t) \left\| \prod_{j=1}^{2} \xi_{j,t+\tau} \right\|_{(x_0,t)}^{(1)} \\ &\quad + m_{(2)} \iint_{\mathscr{X}\times[0,t]} \prod_{j=1}^{2} M(\mid \xi_{j,t+\tau} \mid \mid x; t-u)\, \chi(dx \mid x_0; u)\, H(du) \\ &\leqslant M(t) \prod_{j=1}^{2} \parallel \xi_{j,t+\tau} \parallel_{(x_0,t)}^{(\ )} \\ &\quad + m_{(2)} \int_{[0,t]} M^2(t-u) \left\{ \prod_{j=1}^{2} \parallel g_{\xi_j}^{(2)} \parallel_{(x_0,u)}^{(2)} d_{\xi_j}^{(2)}(t+\tau) \right\} H(du), \end{aligned} \tag{4.29}$$

which is seen to be finite since $g_{\xi_j}^{(2)} \in \mathscr{K}_2$, $(j = 1, 2)$, and $M(t)$ is bounded on each finite t-interval.

5. MEAN-SQUARE CONVERGENCE WHEN $m > 1$

The asymptotic properties of the population size $N(t)$ and its moments have been investigated by Bellman and Harris ([11], Chapter VI) when $m > 1$

and $G(t)$ is a nonlattice distribution. If a is the unique positive root of the equation

$$m \int_{\mathscr{T}} e^{-at}\, G(dt) = 1, \tag{5.1}$$

and we write

$$N_0 = \frac{m-1}{am^2 \int_{\mathscr{T}} te^{-at}\, G(dt)}, \tag{5.2}$$

then

$$M(t) \sim N_0 e^{at} \qquad \text{as} \qquad t \to \infty. \tag{5.3}$$

Furthermore, if $m_{(2)} < \infty$, the random variable

$$W(t) = \frac{N(t)}{N_0 e^{at}} \tag{5.4}$$

converges in mean square to a random variable W, whose characteristic function $L(\theta)$ is the unique solution of the integral equation

$$L(\theta) = \int_{\mathscr{T}} f\{L(\theta e^{-au})\}\, G(du) \tag{5.5}$$

such that $L(0) = 1$, $(dL/i\, d\theta)_{\theta=0} = 1$, $|L(\theta)| \leqslant 1$.

The first two moments of W are

$$\mathscr{E}W = 1, \qquad \mathscr{E}W^2 = m_{(2)} \int_{\mathscr{T}} e^{-2au}\, H(du). \tag{5.6}$$

In order to develop an analogous theory for $N(\xi_t \mid x_0; t)$, we shall have to introduce some further function classes. Let $\mathscr{K}_p(\delta; 0)$ be the class of real finite-valued functions on $\mathscr{X}$. For $n \geqslant 1$, let us recursively define $\mathscr{K}_p(\delta; n)$ to be the linear subspace of $\mathscr{K}_p(\delta; n-1)$ consisting of those ξ for which (4.20) is satisfied by a $\gamma_\xi^{(p)} \in \mathscr{K}_p(\delta; n-1)$, and a $\delta_\xi^{(p)}(t)$ which is $\mathscr{B}_t$-measurable and bounded on each finite t-interval and $O(t^{i_\xi})$ for large $t (i_\xi \geqslant 0)$. Clearly

$$\begin{aligned} &\mathscr{K}_p(\delta; n) \subseteq \mathscr{K}_p(\gamma; n), \\ &\mathscr{K}_p(\delta; m) \subseteq \mathscr{K}(\delta; n), \qquad (p \geqslant r \geqslant 1, m \geqslant n). \end{aligned} \tag{5.7}$$

We may correspondingly define $\Lambda_p(\delta; n)$ to be the subclass of $\Lambda_p(\gamma; n)$ consisting of those $\xi_t(x)$ for which there exists a $g_\xi^{(p)} \in \mathscr{K}_p(\delta; n-1)$, and a corresponding $d_\xi^{(p)}(t)$ which is $O(t^{i_\xi})$ for large $t > 0$, $(i_\xi \geqslant 0)$. Then

$$\begin{aligned} &\mathscr{K}_p(\delta; n) \subseteq \Lambda_p(\delta; n), \\ &\Lambda_p(\delta; m) \subseteq \Lambda_r(\delta; n), \qquad (p \geqslant r \geqslant 1, m \geqslant n). \end{aligned} \tag{5.8}$$

Now let $\boldsymbol{\Phi}$ be the class of real finite-valued $\mathscr{B}_t$-measurable functions on $\mathscr{T}$. Then we define $\mathscr{J}_n$ $(n = 1, 2,...)$ to be the subclass of $\Lambda_n(\delta; n) \times \boldsymbol{\Phi}$ consisting of those pairs (ξ_t, Φ_ξ) such that the following limits exist for all $x \in \mathscr{X}$ and $\tau \geqslant 0$:

$$\text{(i)} \quad \lim_{t\to\infty} \Phi_\xi(t) \int_{\mathscr{X}} \xi_{t+\tau}(y)\, \chi(dy \mid x; t) = J_\xi(x; \tau),$$

$$\text{(ii)} \quad \lim_{t\to\infty} \frac{\Phi_\xi(t+\tau)}{\Phi_\xi(t)} = \Psi_\xi(\tau),$$

and the following conditions are also fulfilled for all $t \geqslant 0$:

$$\text{(iii)} \quad \left| \Phi_\xi(t) \int_{\mathscr{X}} \xi_{t+\tau}(y)\, \chi(dy \mid x; t) \right| \leqslant K_\xi(x)\, D_\xi(\tau),$$

where $K_\xi \in \mathscr{K}_n(\delta; n-1)$, and $D_\xi(t)$ is a $\mathscr{B}_t$-measurable function bounded on each finite t-interval, being $O(t^{I_\xi})$ for large t, $(I_\xi \geqslant 0)$;

$$\text{(iv)} \quad \left| \frac{\Phi_\xi(t+\tau)}{\Phi_\xi(t)} \right| \leqslant C_\xi e^{\kappa_\xi \tau},$$

where C_ξ and κ_ξ are non-negative constants.

It can be seen that

$$\mathscr{J}_p \subseteq \mathscr{J}_r, \qquad (p \geqslant r \geqslant 1). \tag{5.9}$$

Generalizing $W(t)$, let us write

$$W_\tau(\xi \mid x_0; t) = \frac{\Phi_\xi(t)\, N(\xi_{t+\tau} \mid x_0; t)}{N_0 e^{at}}, \qquad (\tau \geqslant 0). \tag{5.10}$$

THEOREM 4. *Suppose that $m > 1$, $m_{(2)} < \infty$, and G is a nonlattice distribution. If $(\xi_t, \Phi_\xi) \in \mathscr{J}_2$ and $2\kappa_\xi < a$, then for each $\tau \geqslant 0$ $W_\tau(\xi \mid x_0; t)$ converges in mean square to a random variable $W_\tau(\xi \mid x_0)$.*

The characteristic function $L_\tau[\theta; \xi \mid x_0]$ of $W_\tau(\xi \mid x_0)$ is a solution of the integral equation

$$L_\tau[\theta; \xi \quad x_0] = \int_{\mathscr{S}} \int f\{L_{\tau+u}[\theta e^{-au} \Psi_\xi(u); \xi \mid x]\}\, \chi(dx \mid x_0; u)\, G(du) \tag{5.11}$$

which satisfies the conditions

$$L_\tau[0; \xi \mid x_0] \equiv 1, \qquad | L_\tau[\theta; \xi \mid x_0] | \leqslant 1,$$

$$\left| \frac{L_\tau[\theta; \xi \mid x_0] - 1}{i\theta} - J_\xi(x_0; u) \right| \leqslant | \theta |\, \Delta_\xi(\tau)\, \Theta_\xi(x_0) \tag{5.12}$$

for all real $\theta \neq 0$, *where* $\Theta_\xi \in \mathscr{K}_1(\delta; 1)$, *and* Δ_ξ *is a* $\mathscr{B}_t$*-measurable function of* t, *bounded on each finite* t*-interval and* $O(t^s)$ *for large* t, *where* $s \geqslant 0$.

If $|\Psi_\xi(\tau)| \leqslant e^{\kappa_\xi \tau}$, *there is exactly one such solution.*

We shall omit the proof of this theorem, which follows closely the lines of [9], Theorem 6.3.

Corollary 1. *If* $J_\xi(x; \tau)$ *is independent of* τ, *then* $W_\tau(\xi \mid x_0) \equiv W(\xi \mid x_0)$ *is also independent of* τ.

2. *If* (ξ_t, Φ_ξ), $(\eta_t, \Phi_\eta) \in \mathscr{J}_2$ *and* $\Psi_\xi(t) \equiv \Psi_\eta(t)$, $J_\xi(x; \tau) \equiv cJ_\eta(x; \tau)$, *(c constant), then*

$$W_\tau(\xi \mid x_0) \equiv cW_\tau(\eta \mid x_0). \tag{5.13}$$

3. *In particular, if* $\Psi_\xi(t) \equiv 1$, *and* $J_\xi(x; \tau) \equiv J_\xi$ *is independent of* x *and* τ, *then*

$$W_\tau(\xi \mid x_0) \equiv J_\xi W, \tag{5.14}$$

where W *is the limiting random variable of Bellman and Harris.*

Finally in this section we consider the moments of $W_\tau(\xi \mid x_0)$. Let us write

$$Q_{k,\tau}(\xi \mid x_0; t) = \mathscr{E} W_\tau^{\,k}(\xi \mid x_0; t) = \frac{\Phi_\xi^{\,k}(t)\, M_k(\xi_{t+\tau} \mid x_0; t)}{N_0^{\,k} e^{k\alpha t}},$$

$$Q_{k,\tau}(\xi \mid x_0) = \mathscr{E} W_\tau^{\,k}(\xi \mid x_0), \tag{5.15}$$

whenever these exist.

In the following theorem, $\mathscr{P}_j(k) = [s_1^{\rho_1}, \ldots, s_\ell^{\rho_\ell}]$ denotes a partition of the positive integer k into a sum of j positive integers:

$$\sum_{h=1}^{\ell} \rho_h s_h = k, \qquad \sum_{h=1}^{\ell} \rho_h = j, \qquad (s_h > 0, \rho_h > 0), \tag{5.16}$$

while $A(\cdot)$ denotes the elementary partitional function:

$$A(\mathscr{P}_j(k)) = \frac{k!}{\left\{\prod_{h=1}^{\ell} (s_h!)^{\rho_h} \rho_h!\right\}}. \tag{5.17}$$

Theorem 5. *If* $m > 1$, $m_{(n)} < \infty$ *(n positive even),* $(\xi_t, \Phi_\xi) \in \mathscr{J}_n$, *and* $2\kappa_\xi < a$, *then all moments of* $W_\tau(\xi \mid x_0)$ *up to the* $(n-1)$th *exist, and are given recursively by:*

$$Q_{1,\tau}(\xi \mid x_0) = J_\xi(x_0; \tau), \tag{5.18}$$

$$Q_{k,\tau}(\xi \mid x_0) = \sum_{j=2}^{k} m_{(j)} \sum_{\mathscr{P}_j(k)} A(\mathscr{P}_j(k)) \int_{\mathscr{S}}\!\!\int e^{-kau}\, \Psi_\xi^{\,k}(u)$$

$$\times \prod_{h=1}^{\ell} Q^{\rho_h}_{s_h,\tau+u}(\xi \mid x)\, \chi(dx \mid x_0; u)\, H(du), \qquad (k = 2,\dots, n-1). \tag{5.19}$$

In order to indicate the proof of this result, we first note that in virtue of the mean-square convergence of $W_\tau(\xi \mid x_0; t)$ to $W_\tau(\xi \mid x_0)$:

$$Q_{j,\tau}(\xi \mid x_0) = \lim_{t\to\infty} Q_{j,\tau}(\xi \mid x_0; \tau), \qquad (j = 1, 2). \tag{5.20}$$

But, from (5.3) and (i):

$$Q_{1,\tau}(\xi \mid x_0; t) = \frac{M(t)}{N_0 e^{at}} \cdot \Phi_\xi(t) \int_{\mathscr{X}} \xi_{t+\tau}(x)\, \chi(dx \mid x_0; t) \to J_\xi(x_0; \tau), \tag{5.21}$$

which proves (5.17). Also, letting b denote $\sup_{t\geqslant 0} M(t)/N_0 e^{at}$,

$$| Q_{1,\tau}(\xi \mid x_0; t) | \leqslant b K_\xi(x_0)\, D_\xi(\tau), \qquad (K_\xi \in \mathscr{K}_2). \tag{5.22}$$

From (4.31),

$$Q_{2,\tau}(\xi \mid x_0; t) = N_0^{-2}\, e^{-2at}\, \Phi_\xi^{\,2}(t)\, M(t) \int_{\mathscr{X}} \xi^2_{t+\tau}(x)\, \chi(dx \mid x_0; t)$$

$$+ m_{(2)} \iint_{\mathscr{X}\times[0,t]} e^{-2au} \left\{\frac{\Phi_\xi(t)}{\Phi_\xi(t-u)} Q_{1,\tau+u}(\xi \mid x; t-u)\right\}^2 \chi(dx \mid x_0; u)\, H(du). \tag{5.23}$$

The integrand in the second term is dominated for all t and fixed τ by $b^2 C_\xi^{\,2} e^{-2(a-\kappa_\xi)u}\, D_\xi^{\,2}(\tau + u)\, K_\xi^{\,2}(x)$; this has a finite integral over $\mathscr{S} = \mathscr{X} \times \mathscr{T}$ with respect to $\chi(dx \mid x_0; u)\, H(du)$ in virtue of [9] Lemma 6.1. The first term on the right-hand side of (5.23) is dominated by

$$\text{const. } e^{-(a-2\kappa_\xi)t} \{g_\xi^{(2)}(x_0)\, d_\xi^{(2)}(t + \tau)\}^2.$$

Hence, letting $t \to \infty$, we obtain

$$Q_{2,\tau}(\xi \mid x_0) = m_{(2)} \int_{\mathscr{S}}\!\!\int e^{-2au}\, \Psi_\xi^{\,2}(u)\, J_\xi^{\,2}(x; \tau + u)\, \chi(dx \mid x_0; u)\, H(du). \tag{5.24}$$

Applying this method inductively to Eqs. (4.31), we find that under the conditions of the theorem, $\lim_{t\to\infty} Q_{k,\tau}(\xi \mid x_0; t)$ exists for $k = 1,\ldots, n$, the limits being given by (5.18) and (5.19). That these limits are in fact the moments of $W_\tau(\xi \mid x_0)$ for $k = 1,\ldots, n-1$ follows by a theorem due to Kendall and Rao ([15], Theorem (5.3.3)).

Equation (5.24) shows that if $J_\xi(x_0; \tau)$ is null, then so also is $W_\tau(\xi \mid x_0)$. Interest therefore centers on finding weighting functions $\Phi_\xi(t)$ which yield non-null $J_\xi(x_0; \tau)$.

It should also be noted that all the function classes introduced have been defined in terms of the diffusion process only. The "interlock" with the branching process is expressed by the requirement $2\kappa_\xi < a$.

6. Applications

Example (a). *Gaussian diffusion on the Real Line.* In this example, $\mathscr{X}$ denotes the real line, and χ has the density

$$\psi(y \mid x; t) = (2\pi\sigma^2 t)^{-1/2} \exp\left[-\frac{(y-x)^2}{2\sigma^2 t}\right], \tag{6.1}$$

where $\sigma^2/2$ is the constant coefficient of diffusion.

If $\xi(x) = |x|^s$, $(s \geqslant 0)$, then

$$\| \xi \|_{(x,t)}^{(p)} \leqslant \text{const.} \{|x|^s + \nu_{sp}^{1/p}(\sigma\sqrt{t})^s\}, \tag{6.2}$$

where

$$\nu_s = \frac{1}{\sqrt{2\pi}} \int_{\mathscr{X}} |y|^s e^{-y^2/2}\, dy. \tag{6.3}$$

Hence, if $\mathbf{P}$ denotes the class of Borel functions on the real line which are bounded on each finite interval and are of polynomial order for large $|x|$, then

$$\mathbf{P} \subseteq \bigcap_{n,p\geqslant 1} \mathscr{K}_p(\delta; n).$$

Since

$$\lim_{t\to\infty} \sigma\sqrt{t}\, \psi(\sigma\sqrt{t}\, x \mid x_0; t) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \tag{6.4}$$

we are led to consider the functions

$$\xi_t(x) = \xi\left(\frac{x}{\sigma \max(1, \sqrt{t})}\right). \tag{6.5}$$

If $\xi(x) \in \mathbf{P}$, then it is found that $\xi_t(x) \in \bigcap_{n,p\geqslant 1} \Lambda_p(\delta; n)$.

In order to discuss the classes $\mathscr{J}_n$, we first note that on writing $\theta = \sqrt{\tau/(t+\tau)}$ we obtain the following expansion using Mehler's formula:

$$\sigma\sqrt{t+\tau}\,\psi(y\sigma\sqrt{t+\tau} \mid x_0; t)$$

$$= \frac{\exp(-y^2/2)}{\sqrt{2\pi(1-\theta^2)}} \exp\left\{-\frac{\theta^2 y^2 - 2\theta y(x_0/\sigma\sqrt{\tau}) + \theta^2(x_0/\sigma\sqrt{\tau})^2}{2(1-\theta^2)}\right\}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\tau}{t+\tau}\right)^{n/2} H_n(y)\, H_n\left(\frac{x_0}{\sigma\sqrt{\tau}}\right), \tag{6.6}$$

where the $H_n(y)$ are Hermite's polynomials with respect to the weight function $e^{-y^2/2}$. Let us now set

$$(\xi, H_n) = \frac{1}{n!\sqrt{2\pi}} \int_{\mathscr{X}} \xi(y)\, H_n(y)\, e^{-y^2/2}\, dy. \tag{6.7}$$

It may be shown without difficulty that if $\xi(x)$ is a polynomial in x, and j is the smallest integer such that $(\xi, H_j) \neq 0$, then

$$(\xi_t, \Phi_\xi) \in \bigcap_{n \geqslant 1} \mathscr{J}_n,$$

with

$$\Phi_\xi(t) = \sigma^j \max(1, t^{j/2}), \qquad \Psi_\xi(t) \equiv 1,$$

$$J_\xi(x_0; \tau) = (\xi, H_j)\,(\sigma\sqrt{\tau})^j\, H_j\left(\frac{x_0}{\sigma\sqrt{\tau}}\right). \tag{6.8}$$

Applying Theorem 4 and Corollary 2, it follows that

$$W_0(\xi \mid x_0; t) \xrightarrow[\text{m.s.}]{} (\xi, H_j)\, W_0(H_j \mid x_0). \tag{6.9}$$

In particular, from Corollary 3, $W_0(H_0 \mid x_0) \equiv W$. Thus, if $\xi(x) = \delta(X \mid x)$, where X is a Borel set, then

$$\frac{N(\sigma\sqrt{t}\,X \mid x_0; t)}{N_0 e^{at}} \xrightarrow[\text{m.s.}]{} \frac{W}{\sqrt{2\pi}} \int_X e^{-y^2/2}\, dy. \tag{6.10}$$

Next, considering $\xi(x) = x^n$, we see that if n is *even*, then

$$\frac{\sum_{i=1}^{N(t)} x_i^n}{(\sigma\sqrt{t})^n N_0 e^{at}} \xrightarrow[\text{m.s.}]{} \nu_n W, \tag{6.11}$$

while if n is *odd*

$$W_0(x^n \mid x_0) = \nu_{n+1} W_0(x \mid x_0). \tag{6.12}$$

Further information on the asymptotic behavior of $\sum_{i=}^{N(t)} x_i^2$ may be obtained by taking $\xi(x) = H_2(x)$:

$$N_0^{-1} e^{-at} \left\{ \sum_{i=1}^{N(t)} x_i^2 - \sigma^2 t \, N(t) \right\} \xrightarrow[\text{m.s.}]{} W_0(H_2 \mid x_0). \tag{6.13}$$

In the case of *binary splitting* $f(z) = z^2$, it is known that the characteristic function of W is analytic. However, we shall now show that the characteristic function of $W_\tau(\xi \mid x_0)$ is not in general analytic even when the branching mechanism is of *simple birth and death* type:

$$f(z) = \frac{\mu + \lambda z^2}{\mu + \lambda}, \qquad G(t) = 1 - e^{-(\lambda+\mu)t}, \qquad (t \geqslant 0), \tag{6.14}$$

where λ, μ are the constant birth and death rates, respectively, $\lambda > \mu$. Taking $\xi(x) = H_1(x) = x$, all the moments of $W_0(\xi \mid x_0)$ exist by Theorem 5, and are given recursively by:

$$Q_1(\xi \mid x_0) = x_0 ,$$

$$Q_n(\xi \mid x_0) = \lambda \sum_{r=1}^{n-1} \binom{n}{r} \int_{\mathscr{S}} \int e^{-(n-1)(\lambda-\mu)u} Q_r(\xi \mid x) \, Q_{n-r}(\xi \mid x) \, \psi(x \mid x_0; u) \, dx \, du, \tag{6.15}$$

where $\binom{n}{r}$ is the binomial coefficient. It follows by induction that

$$Q_n(\xi \mid x_0) = n! \left(\frac{\lambda}{\lambda - \mu}\right)^{n-1} \sum_{s=0}^{[n/2]} \frac{a_s^{(n)}}{(\lambda - \mu)^s} x_0^{n-2s}, \tag{6.16}$$

the coefficients $a_s^{(n)}$ being defined by the recurrence relations:

$$a_0^{(n)} = 1,$$

$$(n-1) \, a_s^{(n)} = \frac{1}{2} \sigma^2 (n - 2s + 2)(n - 2s + 1) \, a_{s-1}^{(n)} + 2 \sum_{r=2s}^{n-1} a_s^{(r}$$

$$+ \sum_{\ell=0}^{s-1} \sum_{r=2s-2\ell}^{n-2\ell} a_{s-\ell}^{(r)} a_\ell^{(n-r)}, \qquad \left(1 \leqslant s \leqslant \left[\frac{n}{2}\right]\right). \tag{6.17}$$

Clearly all $a_s^{(n)} > 0$, and so we obtain the rough inequality

$$(n-1) \, a_s^{(n)} > \tfrac{1}{2} \sigma^2 (n - 2s + 2)(n - 2s + 1) \, a_{s-1}^{(n} . \tag{6.18}$$

Taking $x_0 = 0$, we see from (6.16) that all odd moments of $W_0(\xi \mid 0)$ vanish, while from (6.18):

$$Q_{2n}(\xi \mid 0) = \frac{(2n)!\,\lambda^{2n-1}}{(\lambda - \mu)^{3n-1}}\,a_n^{(2n)} > \{(2n)!\}^2 \frac{\lambda^{2n-1}}{(\lambda - \mu)^{3n-1}} \left(\frac{\frac{1}{2}\sigma^2}{2n-1}\right)^n. \tag{6.19}$$

Hence $\sum_{n=1}^{\infty} (Q_{2n}(\xi \mid 0)/(2n)!)\,z^{2n}$ is convergent for no $z \neq 0$.

EXAMPLE (b). *Random Walk on the Integers.* Let $\mathscr{X}$ denote the set of all positive and negative integers, and let $\chi(x \mid x_0; t)$ be the transition probability for an unrestricted random walk on $\mathscr{X}$ in continuous time: an individual at x at time t has the probabilities $\sigma\delta t$, ρdt of moving to $x + 1$, $x - 1$, respectively, during $(t, t + \delta t)$. Clearly,

$$\chi(x \mid x_0; t) = \chi(x - x_0 \mid 0; t), \tag{6.20}$$

where $\chi(\cdot \mid 0; t)$ has the characteristic function

$$\phi(\theta; t) = \exp\{[-(\sigma + \rho) + (\sigma e^{i\theta} + \rho e^{-i\theta})]\,t\}. \tag{6.21}$$

If $\mathbf{P}^*$ denotes the class of functions on $\mathscr{X}$ of polynomial order for large $|x|$, then it is found that

$$\mathbf{P}^* \subset \bigcap_{n,p \geqslant 1} \mathscr{K}_p(\delta; n).$$

Let $x(t)$ denote the position at time t of an individual initially at x_0. It follows easily from the characteristic function (6.21) that

$$y(t) = \frac{x(t) - (\sigma - \rho)\,t}{\sqrt{(\sigma + \rho)\,t}} \tag{6.22}$$

is asymptotically unit normal. This suggests that we consider the functions

$$\xi_t(x) = \xi\left(\frac{x - (\sigma - \rho)\,t}{\sqrt{(\sigma + \rho)\,t}}\right), \qquad (x \in \mathscr{X}), \tag{6.23}$$

where $\xi(\cdot)$ is defined on the real line $\mathscr{R}$. If $\xi(x)$ is of polynomial order for large $|x|$, then $\xi_t(x) \in \bigcap_{n,p\geqslant 1} \Lambda_p(\delta; n)$.

It is found that

$$\sum_{\mathscr{X}} \xi_{t+\tau}(x)\,\chi(x \mid x_0; t) \sim \frac{1}{\sqrt{2\pi}} \int_{\mathscr{R}} \xi(y)\,e^{-y^2/2} \times \left\{1 + \frac{1}{(t+\tau)^{1/2}}\left[\frac{H_1(y)\,(x_0 - (\sigma - \rho)\,\tau)}{(\sigma + \rho)^{1/2}} + \frac{H_3(y)\,(\sigma - \rho)}{6(\sigma + \rho)^{3/2}}\right] + \frac{\zeta(x_0, \tau, y)}{(t + \tau)} + \cdots\right\} dy, \tag{6.24}$$

whence, if $\xi(x)$ is a polynomial in x, then

$$(\xi_t, \Phi_\xi) \in \bigcap_{n \geqslant 1} \mathscr{I}_n,$$

where $\Phi_\xi(t)$ and $J_\xi(x_0; \tau)$ may be derived from the following table (or its extension):

Φ_ξ and J_ξ for the random walk

$\Phi_\xi(t) = (\sigma + \rho)^{r/2} \max(1, t^{r/2})$	$\xi(x)$	$J_\xi(x_0; \tau)$
$r = 0$	$H_0(x)$	1
$r = 1$	$H_1(x)$	$[x_0 - (\sigma - \rho)\tau]$
	$H_3(x)$	$(\sigma - \rho)/(\sigma + \rho) = \omega$, say.
$r = 2$	$H_2(x)$	$[x_0 - (\sigma - \rho)\tau]^2 - (\sigma + \rho)\tau$
	$H_4(x)$	$4\omega[x_0 - (\sigma - \rho)\tau] + 1$
	$H_6(x)$	$10\omega^2$

In particular, we have

$$N_0^{-1} e^{-at} \left\{ \sum_{i=1}^{N(t)} x_i - (\sigma - \rho)\, t\, N(t) \right\} \xrightarrow[\text{m.s.}]{} W_0(H_1 \mid x_0). \tag{6.25}$$

EXAMPLE (c). *Diffusion in a compact region.* Suppose that $\mathscr{X}$ is a compact region in Euclidean space, and let χ have a density ψ with respect to a finite measure $\mu(\cdot)$ on $\mathscr{B}_x$ such that ψ may be expanded in terms of its eigenfunctions:

$$\psi(x \mid x_0; t) = \sum_{n=0}^{\infty} e^{-\kappa_n t} \alpha_n(x)\, \alpha_n(x_0). \tag{6.26}$$

Hence $\{\alpha_n\}$ denotes a complete orthonormal set in the space of real functions square-integrable with respect to μ:

$$\int_{\mathscr{X}} \alpha_i(x)\, \alpha_j(x)\, \mu(dx) = \delta_{ij} \qquad \text{(Kronecker's delta)}, \tag{6.27}$$

and

$$0 = \kappa_0 < \kappa_1 < \kappa_2 < \cdots. \tag{6.28}$$

We shall assume that the α_n are bounded:

$$|\alpha_n(x)| \leqslant k_n, \qquad (x \in \mathscr{X}, n = 0, 1, \ldots), \tag{6.29}$$

and that for some $t_0 > 0$:

$$\sum_{n=0}^{\infty} e^{-\kappa_n t_0} k_n^2 < \infty. \tag{6.30}$$

If $\xi(x)$ is any bounded Borel function on $\mathscr{X}$, let us write

$$(\xi, \alpha_n) = \int_{\mathscr{X}} \xi(x)\, \alpha_n(x)\, \mu(dx) \tag{6.31}$$

and let j denote the smallest integer for which $(\xi, \alpha_j) \neq 0$. Then $(\xi, \Phi_\xi) \in \bigcap_{n \geqslant 1} \mathscr{I}_n$, with

$$\Phi_\xi(t) = e^{\kappa_j t} = \Psi_\xi(t), \qquad J_\xi(x; \tau) \equiv J_\xi(x) = (\xi, \alpha_j)\, \alpha_j(x_0). \tag{6.32}$$

For example, if $\mathscr{X}$ denotes the closed interval $[-\frac{1}{2}L, \frac{1}{2}L]$ on the real line, where $\pm \frac{1}{2}L$ are reflecting boundaries, we have for simple Gaussian diffusion:

$$\mu(dx) = dx, \qquad \kappa_n = \frac{1}{2}\left(\frac{n\pi\sigma}{L}\right)^2, \qquad \alpha_0 = L^{-1/2},$$

$$\alpha_n(x) = \begin{cases} \sqrt{\dfrac{2}{L}} \sin\left(\dfrac{\pi n x}{L}\right), & (n \text{ odd}) \\ \sqrt{\dfrac{2}{L}} \cos\left(\dfrac{\pi n x}{L}\right), & (n \text{ positive even}). \end{cases} \tag{6.33}$$

If $\xi(x) = x^n$ (n even), then $j = 0$ and

$$N_0^{-1} e^{-at} \sum_{i=1}^{N(t)} x_i^n \xrightarrow[\text{m.s.}]{} \frac{(L/2)^n}{(n+1)} W. \tag{6.34}$$

If n is odd, then $j = 1$. Hence if $2\kappa_1 < a$,

$$N_0^{-1} e^{-(a-\kappa_1)t} \sum_{i=1}^{N(t)} x_i^n \xrightarrow[\text{m.s.}]{} \left\{ 2\sqrt{\frac{2}{L}} \int_0^{L/2} x^n \sin\frac{\pi x}{L}\, dx \right\} W(\alpha_1 \mid x_0). \tag{6.35}$$

To obtain more information on $\sum_{i=1}^{N(t)} x^n$, for n even, we may take $\xi(x) = x^n - (x^n, \alpha_0)\, \alpha_0$.

Then if $2\kappa_2 < a$:

$$N_0^{-1} e^{-(a-\kappa_2)t} \left\{ \sum_{i=1}^{N(t)} x_i^n - \frac{(L/2)^n}{(n+1)} N(t) \right\} \xrightarrow[\text{m.s.}]{} (x^n, \alpha_2)\, W(\alpha_2 \mid x_0). \tag{6.36}$$

Example (d). Eigenfunction expansions of the type (6.26) may exist even when $\mathscr{X}$ is not compact. Let $\mathscr{X}$ be the real line, and let χ be the transition probability for simple diffusion under a force proportional to the displacement x and directed toward the origin. Then χ has the density

$$\psi(x \mid x_0; t) = \frac{1}{\sqrt{1 - e^{-2ct}}} \exp\left\{ - \frac{c[x^2 e^{-2ct} - 2xx_0 e^{-ct} + x_0^2 e^{-2ct}]}{\sigma^2 (1 - e^{-2ct})} \right\}$$

$$(\sigma^2,\ c \text{ constants}) \tag{6.37}$$

with respect to the finite measure

$$\mu(dx) = \sqrt{\frac{c}{\pi\sigma^2}} \exp\left(- \frac{cx^2}{\sigma^2} \right) dx. \tag{6.38}$$

By Mehler's formula, ψ may be expressed in the form (6.26) with

$$\kappa_n = nc,$$

$$\alpha_n(x) = H_n\left(\frac{x\sqrt{2c}}{\sigma} \right). \tag{6.39}$$

For this diffusion process, it may be shown that $\xi \in \mathscr{K}_p(\gamma; n)$ provided that $\xi(x)$ is dominated by $e^{x^2/s}$, where $s > np\sigma^2/c$.

All polynomials $\xi(x)$ are members of $\bigcap_{n \geq 1} \mathscr{I}_n$, and Eqs. (6.31), (6.32) carry over directly in this case.

References

1. B. A. Sevast'yanov. Branching stochastic processes for particles diffusing in a restricted domain with absorbing boundaries. *Th. Prob. Appl.* **3** (1958), 111-126.
2. B. A. Sevast'yanov. Extinction conditions for branching stochastic processes with diffusion. *Th. Prob. Appl.* **6** (1961), 253-263.
3. Howard E. Conner. Extinction probabilities for age- and position-dependent branching processes. *J. Soc. Indust. Appl. Math.* **12** (1964), 899-909.
4. Howard E. Conner. Limiting behavior for age- and position-dependent branching processes. *J. Math. Anal. Appl.* **13** (1966), 265-295.
5. S. Watanabe. On the branching process for Brownian particles with an absorbing boundary. *J. Math. Kyoto Univ.* **4** (1965), 385-398.

6. S. R. Adke. A stochastic population diffusing on a finite interval. *J. Indian Statist. Assoc.* **2** (1964), 32-40.
7. S. R. Adke. The generalized birth and death process and Gaussian diffusion. *J. Math. Anal. Appl.* **9** (1964), 336-340.
8. S. R. Adke and J. E. Moyal. A birth, death and diffusion process. *J. Math. Anal. Appl.* **7** (1963), 209-224.
9. A. W. Davis. On the theory of birth, death and diffusion processes. *J. Appl. Prob.* **2** (1965), 293-322.
10. A. W. Davis. Branching-diffusion processes with no absorbing boundaries II. To appear in *J. Math. Anal. Appl.* **18** (1967).
11. T. E. Harris. "The Theory of Branching Processes." Springer Verlag, Berlin, 1963.
12. J. E. Moyal. The general theory of stochastic population processes. *Acta Math.* **108** (1962), 1-31.
13. P. E. Ney. Generalized branching processes I: Existence and uniqueness theorems. *Illinois J. Math.* **8** (1964), 316-331.
14. P. E. Ney. Generalized branching processes II: Asymptotic theory. *Illinois J. Math.* **8** (1964), 332-350.
15. S. S. Wilks. "Mathematical Statistics." Wiley, New York, 1962.

PRINTED IN BRUGES, BELGIUM, BY THE ST CATHERINE PRESS, LTD.

Reprinted from JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS Vol. 19, No 1, July 1967

Branching-Diffusion Processes with No Absorbing Boundaries. II

323

A. W. DAVIS

C.S.I.R.O., Division of Mathematical Statistics, Adelaide, South Australia

Submitted by Richard Bellman

1. INTRODUCTION

The present paper is a continuation of the author's previous paper [1], and the notation and results of [1] will be used without further explanation. It is assumed throughout that $m > 1$.

Our aim in this paper is to consider the ratio

$$\frac{N(\xi_t \mid x_0; t)}{N(t)} = \frac{\sum_{j=1}^{N(t)} \xi_t(x_j)}{N(t)}, \tag{1.1}$$

which is of interest since it represents the *average value* of $\xi_t(x)$ over the members of the population at time t. This leads to results for certain natural measures of the spatial location of the population, such as the average position:

$$\bar{x}(t \mid x_0) = \frac{\sum^{N(t)} x_j}{N(t)}. \tag{1.2}$$

In Section 2, a general statement on convergence in probability is given. It does not seem possible to obtain the limiting distribution of the ratio (1.1) in nontrivial cases when $m > 1$. As a first step toward the limiting moments, recurrence relations are obtained in Section 3 for conditional moment distributions and functionals. These are applied to discuss the asymptotic behavior of the moments in the case of $(k + 1)$-fold splitting (Section 5), and also in the case of the simple birth and death process (Section 6). The formulas for these moments involve the Lauricella F_D-type hypergeometric functions, some of whose properties are summarized in Appendix A. In Section 7, we present results for some simple "diffusion" processes.

2. Convergence in Probability

It is convenient to consider the restriction of the random function (1.1) to the subspace of $\Omega^{\mathcal{T}}$ (the space of all realizations of the process, [1] Section 3) on which $N(t) \nrightarrow 0$. This subspace has total probability $1 - P_0$, where P_0 is the probability of the population becoming extinct. It is well-known that P_0 is the smallest non-negative root of the equation

$$P = f(P). \tag{2.1}$$

When $m > 1$, P_0 is strictly less than one, so that $1 - P_0$ is positive. Furthermore, the subspaces $\{N(t) \nrightarrow 0\}$ and $\{W > 0\}$ are known to be equivalent for the Bellman-Harris process ([2] p. 147, Remark 1). We therefore consider

$$\begin{aligned} Y(\xi \mid x_0; t) &= \Phi_\xi(t) \frac{N(\xi_t \mid x_0; t)}{N(t)} \bigg|_{N(t) \nrightarrow 0} \\ &= \frac{W_0(\xi \mid x_0; t)}{W(t)} \bigg|_{W>0}. \end{aligned} \tag{2.2}$$

Theorem 1. *Under the conditions of* [1] *Theorem 4,* $Y(\xi \mid x_0; t)$ *converges in probability as* $t \to \infty$ *to the random variable*

$$Y(\xi \mid x_0) = \frac{W(\xi \mid x_0)}{W} \bigg|_{W>0}. \tag{2.3}$$

The restrictions of $W_0(\xi \mid x_0; t)$ and $W(t)$ to the subspace $\{W > 0\}$ converge in mean square to the restrictions of $W_0(\xi \mid x)$ and W, respectively. The theorem follows by [3] Lemma 6.4. From [1], Theorem 4, Corollary 3,

Corollary. *If* $\Psi_\xi(t) \equiv 1$, *and* $J_\xi(x; \tau) \equiv J_\xi$ *is independent of* x *and* τ, *then*

$$Y(\xi \mid x_0; t) \xrightarrow[\text{i.p.}]{} J_\xi, \tag{2.4}$$

a nonrandom limit.

3. Conditional Moment Generating Functions

As a first step towards investigating the moments of the limiting random variable $Y(\xi \mid x_0)$, we shall consider in this section the moments of $N(\xi_{t+\tau} \mid x_0; t)$ *conditional on there being* r *survivors at time* t.

Let us write

$$\Gamma[\xi, \phi \mid x_0; t] = \Gamma[\xi + \phi \mid x_0; t], \qquad (\phi \text{ real}). \tag{3.1}$$

It was remarked in [1] Section 3 that $\Gamma[\xi \mid x_0; t]$ specifies a probability distribution on the population state space $\Omega = \sum_{r=0}^{\infty} \mathscr{X}^r$; we have accordingly

$$\Gamma[\xi, \phi \mid x_0; t] = \sum_{r=0}^{\infty} e^{ir\phi}\, \Gamma^{(r)}[\xi \mid x_0; t], \tag{3.2}$$

where

$$\Gamma^{(r)}[\xi \mid x_0; t] = \int_{\mathscr{X}^r} \exp\left\{ i \sum_{j=1}^{r} \xi(x_j) \right\} P^{(r)}(dx^{(r)} \mid x_0; t); \tag{3.3}$$

i.e., if $p_r(t) = \text{Prob}\,\{N(t) = r\} \neq 0$, then $\Gamma^{(r)}[\xi \mid x_0; t]/p_r(t)$ is the characteristic functional at time t conditional on r survivors. If $p_r(t) = 0$ then $\Gamma^{(r)} \equiv 0$.

We shall obtain recurrence relations for the conditional moment distributions. Let us suppose that $m_{(n)} < \infty$. Then the moment distributions $M_k(\cdot \mid x_0; t)$ are finite measures on $\mathscr{B}_x{}^k$ for $k = 1,\ldots, n$, ([1] Theorem 2), and if $X_j \in \mathscr{B}_x$ ($j = 1,\ldots, n$) we may write

$$M_k(X^{(k)} \mid x_0; t) = \sum_{r=0}^{\infty} M_k^{(r)}(X^{(k)} \mid x_0; t), \tag{3.4}$$

where

$$M_k^{(r)}(X^{(k)} \mid x_0; t) = \int_{\mathscr{X}^r} \prod_{j=1}^{k} \left[\sum_{h=1}^{r} \delta(X_j \mid x_h) \right] P^{(r)}(dx^{(r)} \mid x_0; t)$$

$$= \left\{ \frac{\partial^k}{i^k\, \partial\theta_1 \cdots \partial\theta_k} \Gamma^{(r)} \left[\sum_{j=1}^{k} \theta_j \delta(X_j \mid \cdot) \mid x_0; t \right] \right\} \quad \theta_1 = \cdots = \theta_k = 0. \tag{3.5}$$

$M_k^{(r)}$ may be extended to a finite measure on $\mathscr{B}_x{}^k$ for all $k = 1,\ldots, n$ and all r. If $p_r(t) = 0$, this measure is identically zero, while if $p_r(t) \neq 0$, $M_k^{(r)}(\cdot \mid x_0; t)/p_r(t)$ is the kth moment distribution for the point process whose characteristic functional is $\Gamma^{(r)}[\xi \mid x_0; t]/p_r(t)$, and whose size is r with probability one.

The generating function of the $M_k^{(r)}$ is

$$M_k(X^{(k)}, \phi \mid x_0; t) = \sum_{r=0}^{\infty} e^{ir\phi}\, M_k^{(r)}(X^{(k)} \mid x_0; t)$$

$$= \left\{ \frac{\partial^k}{i^k\, \partial\theta_1 \cdots \partial\theta_k} \Gamma \left[\sum_{j=1}^{k} \theta_j \delta(X_j \mid \cdot), \phi \mid x_0; t \right] \right\} \quad \theta_1 = \cdots = \theta_k = 0. \tag{3.6}$$

Differentiating the integral equation for $\Gamma[\xi, \phi \mid x_0; t]$ derived from [1] Eq. (3.3), we obtain by virtue of [1] Eq. (4.12)

$$M_k(X^{(k)}, \phi \mid x_0; t) = M_k^{(0)}(X^{(k)}, \phi \mid x_0; t)$$

$$+ \iint_{\mathscr{X}\times[0,t]} f'\{\Gamma[\phi; t-u]\} M_k(X^{(k)}, \phi \mid x; t-u)\, \chi(dx \mid x_0; u)\, G(du), \quad (3.7)$$

where

$$M_k^{(0)}(X^{(k)}, \phi \mid x_0; t) = [1 - G(t)]\, e^{i\phi} \chi\left(\bigcap_{j=1}^{k} X_j \mid x_0; t\right)$$

$$+ \iint_{\mathscr{X}\times[0,t]} \sum_{j=2}^{k} f^{(j)}\{\Gamma[\phi; t-u)\} \sum_{\Pi_j(k)} \prod_{h=1}^{j} M_{r_h}(X^{(\pi_h)}, \phi \mid x; t-u)$$

$$\times\, \chi(dx \mid x_0; u)\, G(du). \quad (3.8)$$

Here $\Gamma[\phi; t]$ is the characteristic function of $N(t)$ ([1] Eq. (3.4)).

The author has not succeeded in solving (3.7) in the case of a general life-span distribution $G(t)$. However, progress becomes possible if we take G to be negative exponential:

$$G(t) = 1 - e^{-qt}, \qquad (t \geqslant 0,\, q > 0). \quad (3.9)$$

Equation (3.7) then takes the form

$$R(x_0, t) = F(x_0, t) + q \int_{\mathscr{X}} \int_0^t f'\{\Gamma[\phi; t-u]\} R(x, t-u)\, \chi(dx \mid x_0; u)\, e^{-qu}\, du. \quad (3.11)$$

Now

$$| M_k(X^{(k)}, \phi \mid x_0; t) | \leqslant M_k(X^{(k)} \mid x_0; t) \leqslant M_k(t), \quad (3.11)$$

and so $M_k(X^{(k)}, \phi \mid x_0; t)$ is a member of the class $\mathscr{F}$, defined in [1] Section 4.

LEMMA 1. *Equation (3.11) has at most one solution R in $\mathscr{F}$. A formal solution is*

$$R(x_0; t) = F(x_0, t) + \frac{\partial \Gamma[\phi; t]}{i\partial\phi} \int_{\mathscr{X}} \int_0^t F(x, t-u) \frac{\partial}{\partial u} \left\{ e^{qu} \frac{\delta \Gamma[\phi; t-u]}{i\partial\phi} \right\}$$

$$\times\, \chi(dx \mid x_0; u)\, e^{-qu}\, du. \quad (3.12)$$

PROOF. Suppose that there exist two solutions in $\mathscr{F}$. Then their difference, $D(x_0, t)$ say, is a member of $\mathscr{F}$ and satisfies the equation

$$D(x_0; t) = \int_{\mathscr{X}} \int_0^t f'\{\Gamma[\phi; t-u]\} D(x, t-u)\, \chi(dx \mid x_0; u)\, G(du) \tag{3.13}$$

where G is given by Eq. (3.9). Hence

$$\begin{aligned} | D(x_0; t)| &\leqslant m \int_{\mathscr{X}} \int_0^t | D(x, t-u) |\, \chi(dx \mid x_0; u)\, G(du), \\ &\leqslant mG(t)\, D^*(t), \end{aligned} \tag{3.14}$$

where

$$D^*(t) = \sup_{\mathscr{X}\times[0,t]} | D(x, t) | . \tag{3.15}$$

Iterating the first inequality in (3.4), we find that for each positive integer n

$$| D(x_0; t) | \leqslant m^n G_n(t) \cdot D^*(t) \to 0, \qquad \text{as} \qquad n \to \infty. \tag{3.16}$$

Thus $D(x_0, t) \equiv 0$ and the solutions are identical.

It is natural to seek a solution of (3.10) having the form

$$R(x_0, t) = F(x_0, t) + \int_{\mathscr{X}} \int_0^t r(t-u, t) F(x, t-u)\, \chi(dx \mid x_0; u)\, e^{-qu}\, du. \tag{3.17}$$

By substituting this expression, it is sufficient that r satisfy the integral equation

$$r(u, t) = qf'\{\Gamma[\phi; u]\} + q \int_u^t r(u, v) f'\{\Gamma[\phi; v]\}\, dv. \tag{3.18}$$

Differentiating both sides of (3.18) with respect to t, and solving the resulting differential equation with the side condition

$$r(u, u) = qf'\{\Gamma[\phi; u]\}, \tag{3.19}$$

we obtain

$$r(t-u, t) = \frac{\partial}{\partial u} \exp \int_{t-u}^t qf'\{\Gamma[\phi; v]\}\, dv. \tag{3.20}$$

A more convenient expression can be found by first noting that, for negative exponential G,

$$\frac{\partial \Gamma[\phi; t]}{\partial t} = q[f\{\Gamma[\phi; t]\} - \Gamma[\phi; t]], \tag{3.21}$$

(see [2] Chapt. V, Eq. (9.1)).

Differentiating both side of (3.21) with respect to $i\phi$, we obtain

$$\frac{\partial}{\partial t}\left\{\frac{\partial}{i\partial\phi}\Gamma[\phi;t]\right\} = q[f'\{\Gamma[\phi;t]\} - 1]\frac{\partial}{i\partial\phi}\Gamma[\phi;t], \tag{3.22}$$

whence

$$\frac{\dfrac{\partial}{i\partial\phi}\Gamma[\phi;t]}{\dfrac{\partial}{i\partial\phi}\Gamma[\phi;t-u]} = \exp\int_{t-u}^{t} q[f'\{\Gamma[\phi;t]\} - 1]\,dv. \tag{3.23}$$

Substituting in (3.20) and (3.17), we obtain the formal solution (3.12).

Applying the lemma to Eq. (3.7), we are led to propose the following solution in the case of negative exponential G:

$$M_1(X,\phi \mid x_0;t) = \frac{\partial\Gamma[\phi;t]}{i\partial\phi}\chi(X \mid x_0;t). \tag{3.24}$$

$$M_k(X^{(k)},\phi \mid x_0;t) = \frac{\partial\Gamma[\phi;t]}{i\partial\phi}\chi\left(\bigcap_{j=1}^{k} X_j \mid x_0;t\right)$$

$$+ q\frac{\partial\Gamma[\phi;t]}{i\partial\phi}\int_{\mathscr{X}}\int_0^t\left\{\frac{\partial\Gamma[\phi;t-u]}{i\partial\phi}\right\}^{-1}\sum_{j=2}^{k} f^{\,j}\{\Gamma[\phi;t-u]\}$$

$$\times\sum_{\Pi_j(k)}\prod_{h=1}^{j} M_{r_h}(X^{(\pi_h)},\phi \mid x;t-u)\,\chi(dx \mid x_0;u)\,du, \qquad (k = 1,\ldots,n). \tag{3.25}$$

This may be shown to yield the required sequence of moment generating functions. For, using the following inequality, derived from (3.23),

$$\left|\frac{\dfrac{\partial}{i\partial\phi}\Gamma[\phi;t]}{\dfrac{\partial}{i\partial\phi}\Gamma[\phi;t-u]}\right| \leqslant e^{q(m-1)u}, \tag{3.26}$$

it may be shown inductively that (3.24-3.25) defines a sequence of functions of (x_0, t) which are members of $\mathscr{F}$. Also, it may be verified by direct substitution that (3.24-3.25) is a solution of (3.7), and the result follows by the uniqueness statement of the lemma.

Now suppose that $\xi_{j,t} \in \Lambda_n(\gamma; n)$, $(j = 1,..., n)$. Then

$$M_k^{(r)}(\xi_{t+\tau}^{(k)} \mid x_0; t) = p_r(t)\, \mathscr{E}\left\{\prod_{j=1}^{k} N(\xi_{j,t+\tau} \mid x_0; t) \mid N(t) = r\right\}$$

$$= \int_{\mathscr{X}^k} \prod_{j=1}^{k} \xi_{j,t+\tau}(x_j)\, M_k^{(r)}(dx^{(k)} \mid x_0; t) \tag{3.27}$$

certainly exists for $k = 1,..., n$ and $r = 1, 2,...,$ since by [1] Theorem 3:

$$\sum_{r=0}^{\infty} M_k^{(r)}(|\, \xi_{t+\tau}^{(k)} \,|\mid x_0; t) = M_k(|\, \xi_{t+\tau}^{(k)} \,|\mid x_0; t) < \infty. \tag{3.28}$$

Writing

$$M_k(\xi_{t+\tau}^{(k)}, \phi \mid x_0; t) = \sum_{r=0}^{\infty} e^{ir\phi} M_k^{(r)}(\xi_{t+\tau}^{(k)} \mid x_0; t) \tag{3.29}$$

it follows that these moment generating functions exist and are given recursively by Eqs. (3.24-3.25), with $X^{(\pi_h)}$ replaced by $\xi_{t+\tau}^{(\pi_h)}$, and $\bigcap_{j=1}^{k} X_j$ by $\prod_{j=1}^{k} \xi_{j,t+\tau}$.

4. Moments of $Y(\xi \mid x_0; t)$

The random variable $Y(\xi \mid x_0; t)$ was defined by Eq. (2.2). We shall write

$$M_n^{(Y)}(\xi^{(n)} \mid x_0; t) = \mathscr{E}\left\{\prod_{j=1}^{n} Y(\xi_j \mid x_0; t)\right\}$$

$$= \mathscr{E}\left\{\prod_{j=1}^{n} \left[\frac{\Phi_{\xi_j}(t)\, N(\xi_{j,t} \mid x_0; t)}{N(t)}\right] \mid N(t') \nrightarrow 0\right\}$$

$$= \sum_{r=1}^{\infty} \frac{\operatorname{Prob}\{N(t') \nrightarrow 0 \mid N(t) = r\}}{\operatorname{Prob}\{N(t') \nrightarrow 0\}}$$

$$\times\, p_r(t)\, \mathscr{E}\left\{\prod_{j=1}^{n} [\cdots] \mid N(t) = r\right\}$$

$$= (1 - P_0)^{-1} \left\{\prod_{j=1}^{n} \Phi_{\xi_j}(t)\right\} \sum_{r=1}^{n} \frac{(1 - P_0{}^r)}{r^n} M_n^{(r)}(\xi_t^{(n)} \mid x_0; t), \tag{4.1}$$

where P_0 is the extinction probability. This moment certainly exists for $m_{(n)} < \infty$ and $\xi_{j,t} \in \Lambda_n(\gamma; n)$, $(j = 1,..., n)$, since the series is dominated by (3.28) with $\tau = 0$.

Lemma 2. *If* $(\xi_t, \Phi_\xi) \in \mathscr{J}_1$, *then*

$$\lim_{t\to\infty} M^{(Y)}(\xi \mid x_0; t) = J_\xi(x_0; 0). \tag{4.2}$$

Proof. From (3.24),

$$M_1^{(r)}(\xi_t \mid x_0; t) = r p_r(t) \int_{\mathscr{X}} \xi_t(x)\, \chi(dx \mid x_0; t), \qquad (r = 1, 2, \ldots). \tag{4.3}$$

Substituting in (4.1), and using the Markovian property of the population size $N(t)$ for negative exponential G:

$$\begin{aligned} M^{(Y)}(\xi \mid x_0; t) &= (1 - P_0)^{-1} (1 - P_0)\, \Phi_\xi(t) \int_{\mathscr{X}} \xi_t(x)\, \chi(dx \mid x_0; t) \\ &\to J_\xi(x; 0) \qquad \text{as} \qquad t \to \infty. \end{aligned} \tag{4.4}$$

We next consider $M_2^{(Y)}(\xi \mid x_0; t)$. Writing

$$j_\xi(x_0; t) = \int_{\mathscr{X}} \xi(x)\, \chi(dx \mid x_0; t), \tag{4.5}$$

we obtain from (3.25)

$$\begin{aligned} M(\xi_t, \phi \mid x_0; t) = \frac{\partial \Gamma[\phi; t]}{i\partial\phi} \Big[j_{\xi_t}^2(x_0; t) + q \int_0^t \frac{\partial \Gamma[\phi; t-u]}{i\partial\phi} f^{(2)}\{\Gamma[\phi; t-u]\} \\ \times du \int_{\mathscr{X}} j_{\xi_t}^2(x; t-u)\, \chi(dx \mid x_0; u) \Big]. \end{aligned} \tag{4.6}$$

An important feature of this equation is the separation of the contributions from the diffusion process and the branching mechanism in the repeated integral. Let us introduce the following functions, which depend entirely on the branching process:

$$\rho_{(1)}(t) = (1 - P_0)^{-1} \sum_{r=1}^{\infty} \frac{(1 - P_0^r)}{r^2} p_r(t),$$

$$\begin{aligned} \rho_{(2)}(u; t) &= q(1 - P_0)^{-1} \sum_{r=1}^{\infty} \frac{(1 - P_0^r)}{r^2} \\ &\times \left[\text{coefficient of } e^{ir\phi} \text{ in } \frac{\partial \Gamma[\phi; t]}{i\partial\phi} \frac{\partial \Gamma[\phi; t-u]}{i\partial\phi} f^{(2)}\{\Gamma[\phi; t-u]\} \right]. \end{aligned} \tag{4.7}$$

Then from (4.1)

$$\begin{aligned} M_2^{(Y)}(\xi^{(2)} \mid x_0; t) = \Phi_{\xi_1}(t)\, \Phi_{\xi_2}(t) \Big\{ \rho_{(1)}(t)\, j_{\xi_{1,t}\xi_{2,t}}(x_0; t) \\ + \int_{\mathscr{X}} \int_0^t \rho_{(2)}(u; t)\, j_{\xi_{1,t}}(x; t-u)\, j_{\xi_{2,t}}(x; t-u)\, \chi(dx \mid x_0; u)\, du \Big\}. \end{aligned} \tag{4.8}$$

We are therefore led to consider the asymptotic behavior of $\rho_{(1)}(t)$ and $\rho_{(2)}(u; t)$ as $t \to \infty$. It appears difficult to make progress in this direction unless the probabilities $p_r(t)$ are known explicitly. In the following sections we shall deal with the cases of $(k + 1)$-fold splitting and the simple birth and death process.

5. Moments of $Y(\xi \mid x_0)$ in the Case of $(k + 1)$-Fold Splitting

For $(k + 1)$-fold splitting, (k a positive integer),

$$f(z) = z^{k+1}, \tag{5.1}$$

and taking $q = 1$, we may write

$$\Gamma[\phi; t] = e^{i\phi}e^{-t}\,\gamma^{k^{-1}}(\phi; t), \tag{5.2}$$

where

$$\gamma(\phi; t) = [1 - (1 - e^{-kt})\,e^{ik\phi}]^{-1}. \tag{5.3}$$

(See [2] Chapt. V, Section 13.2).
Let us also write

$$M_n(\xi_{t+\tau}^{(n)}, \phi \mid x_0; t) = \frac{\partial \Gamma[\phi; t]}{i\partial\phi}\,\mathscr{M}_n(\xi_{t+\tau}^{(n)}, \phi \mid x_0; t). \tag{5.4}$$

Then substituting in (3.25) and noting that

$$\frac{\partial \Gamma[\phi; t]}{i\partial\phi} = e^{i\phi}e^{-t}\,\gamma^{k^{-1}+1}(\phi; t), \tag{5.5}$$

$$f^{(s)}(z) = (k - s + 2)_s\, z^{k-s+1}, \tag{5.6}$$

where

$$(\alpha)_s = \alpha(\alpha + 1) \cdots (\alpha + s - 1), \qquad (\alpha_0) = 1, \tag{5.7}$$

we obtain

$$\begin{aligned}\mathscr{M}_h(\xi_{t+\tau}^{(h)}, \phi \mid x_0; t) &= j_{\Pi_1^h \xi_{j,t+\tau}}(x_0; t)\\ &+ \sum_{s=2}^{\min(k+1,h)} (k - s + 2)_s \int_0^t e^{ik\phi}e^{-k(t-u)}\,\gamma^s(\phi; t - u)\,du\\ &\times \sum_{\Pi_s(h)} \int_{\mathscr{X}} \prod_{\ell=1}^{s} \mathscr{M}_{\pi_\ell}(\xi_{t+\tau}^{(\pi_\ell)}, \phi \mid x; t - u)\,\chi(dx \mid x_0; u).\end{aligned} \tag{5.8}$$

This is valid for $h = 1,\ldots, n$, $\xi_{j,t} \in \Lambda_n(\gamma; n)$, $(j = 1,\ldots, n)$.

It is necessary to consider the coefficients of the $e^{ir\phi}$ in $\mathscr{M}_n$. It may be shown by induction that $\mathscr{M}_n$ is expressible as a linear combination of integrals of the form

$$\int_{u_1=0}^{t}\int_{u_2=u_1}^{t}\cdots\int_{u_p=u_{p-1}}^{t}\prod_{h=1}^{p}[e^{ik\phi}e^{-k(t-u_h)}\,\gamma^{s_h}(\phi;t-u_h)]\,du^{(p)}\int_{\mathscr{X}}\cdots\int_{\mathscr{X}}\cdots,$$

where

(a) the expressions to the right of the x-integrals are independent of ϕ,
(b) $0\leqslant p\leqslant n-1$ ($p=0$ corresponding to the case of no t-integral),
(c) $s_h\geqslant 2$, ($h=1,\ldots,p$),
(d) $\sum_{h=1}^{p}s_h\leqslant n+p-1$.

Hence $M_n^{(r)}(\xi_{t+\tau}^{(n)}\mid x_0;t)$ depends on r only through the coefficients of $e^{ir\phi}$ in the terms

$$e^{i\phi}e^{-t}\,\gamma^{k^{-1}+1}(\phi;t)\prod_{h=1}^{p}[e^{ik\phi}e^{-k(t-u_h)}\,\gamma^{s_h}(\phi;t-u_h)], \tag{5.10}$$

and so, in order to discuss the $M_n^{(Y)}(\xi^{(n)}\mid x_0)$ in the case of $(k+1)$-fold splitting, it is necessary from (4.1) to determine the asymptotic behavior as $t\to\infty$ of the functions

$$\rho_{n,k}\begin{pmatrix}u_1,\ldots,u_h\\ s_1,\ldots,s_p\end{pmatrix};t\Big)=\sum_{r=1}^{\infty}\frac{1}{r^n}\{\text{coefficient of } e^{ir\phi}\text{ in (5.10)}\}. \tag{5.11}$$

These are subject to the conditions (a)-(d), and also $0<u_1<\cdots<u_\rho<t$. We note that $P_0=0$ since $k\geqslant 1$. When $p=0$, we shall write simply $\rho_{n,k}(t)$.

For $n=2$ and 3, $(\xi_{j,t},\Phi_\xi)\in\Lambda_n(\gamma;n)$, $(j=1,2,3)$, the $M_n^{(Y)}(\xi^{(n)}\mid x_0;t)$ are as follows:

$$M_2^{(Y)}(\xi^{(2)}\mid x_0;t)=\prod_{h=1}^{2}\Phi_{\xi_h}(t)\Big\{\rho_{2,k}(t)j_{\Pi_1^2\xi_{\ell,t}}(x_0;t)$$

$$+k(k+1)\int_0^t\rho_{2,k}\binom{u}{2};t\Big)\,du\int_{\mathscr{X}}\prod_{\ell=1}^{2}j_{\xi_{\ell,t}}(x;t-u)\,\chi(dx\mid x_0;u)\Big\}, \tag{5.12}$$

$$M_3^{(Y)}(\xi^{(3)}\mid x_0;t)=\prod_{h=1}^{3}\Phi_{\xi_h}(t)\Big\{\rho_{3,k}(t)j_{\Pi_1^3\xi_{\ell,t}}(x_0;t)$$

$$+k(k+1)\int_0^t\rho_{3,k}\binom{u}{2};t\Big)\,du\int_{\mathscr{X}}[j_{\xi_{1,t}}(x;t-u)\,j_{\Pi_2^3\xi_{\ell,t}}(x;t-u)$$

$$+ \text{ terms obtained by cyclically interchanging } \xi_1, \xi_2, \xi_3] \, \chi(dx \mid x_0; u)$$

$$+ k^2(k+1)^2 \int_0^t \int_u^t \rho_{3,k} \begin{pmatrix} u, v \\ 2, 2 \end{pmatrix}; t \Big) \, du \, dv \iint_{\mathscr{X}^2} \Big[j_{\xi_1,t}(x; t-u) \prod_{\ell=2}^{3} j_{\xi_\ell,t}(y; t-v)$$

$$+ \text{ terms by cyclic interchange} \Big] \, \chi(dy \mid x; v-u) \, \chi(dx \mid x_0; u)$$

$$+ (k-1)_3 \int_0^t \rho_{3,k} \left(\begin{matrix} u \\ 3 \end{matrix}; t \right) du \int_{\mathscr{X}} \prod_{\ell=1}^{3} j_{\xi_\ell,t}(x; t-u) \, \chi(dx \mid x_0; u) \Big\} . \tag{5.13}$$

The asymptotic behavior of the ρ's is discussed in the following theorem. It is shown that if $\sum_{h=1}^{p} s_h < n + p - 1$, then $\rho_{n,k}$ converges to zero exponentially fast as $t \to \infty$. However, if $\sum_{h=1}^{p} s_h = n + p - 1$, then $\rho_{n,k}$ converges to a limiting function which may be expressed in terms of Lauricella's F_D-type generalized hypergeometric function:

$$F_D(\alpha; \beta_1, \dots, \beta_p; \gamma; x_1, \dots, x_p)$$

$$= \sum_{\ell_1, \dots, \ell_p = 0}^{\infty} \frac{(\alpha)_{\ell_1 + \cdots + \ell_p} (\beta_1)_{\ell_1} \cdots (\beta_p)_{\ell_p}}{(\gamma)_{\ell_1 + \cdots + \ell_p} \ell_1! \cdots \ell_p!} x_1^{\ell_1} \cdots x_p^{\ell_p}. \tag{5.14}$$

(See [4], premiere partie, Chapt. VII, Section xxxvii). When $p = 2$, F_D reduces to Appell's first type of hypergeometric function of two variables,

$$F_1(\alpha; \beta_1, \beta_2; \gamma; x_1, x_2) = \sum_{\ell_1, \ell_2 = 0}^{\infty} \frac{(\alpha)_{\ell_1 + \ell_2} (\beta_1)_{\ell_1} (\beta_2)_{\ell_2}}{(\gamma)_{\ell_1 + \ell_2} \ell_1! \, \ell_2!} x_1^{\ell_1} x_2^{\ell_2}, \tag{5.15}$$

(see [5], Chapt. IX, Section 9.1), while when $p = 1$ it reduces to $F(\alpha, \beta; \gamma; x)$, the familiar hypergeometric function of Gauss.

THEOREM 2. *Let* $n \geqslant 2$. (1) *If* $\sum_{h=1}^{p} s_h < n + p - 1$, *then for large* t

$$\rho_{n,k} \begin{pmatrix} u_1, \dots, u_p \\ s_1, \dots, s_p \end{pmatrix}; t \Big) = \begin{cases} O(te^{-t}) & \text{for} \quad p = 0, \\ O(e^{-t}) & \text{for} \quad p \geqslant 1, \end{cases} \tag{5.16}$$

uniformly for $0 < u_i < t$, $(i = 1, \dots, p)$.

(2) *Suppose that* $p \geqslant 1$, $\sum_{h=1}^{p} s_h = n + p - 1$, $0 < u_1 < \cdots < u_p < \infty$. *If we write*

$$\sigma_0 = 1 - e^{-ku_p}, \qquad \sigma_j = 1 - e^{-k(u_p - u_j)}, \qquad (j = 1, \dots, p-1), \tag{5.17}$$

then

$$\lim_{t\to\infty} \rho_{n,k}\binom{u_1,\ldots,u_p}{s_1,\ldots,s_p}; t) = \rho_{n,k}\binom{u_1,\ldots,u_p}{s_1,\ldots,s_p}$$

$$= \frac{(1-\sigma_0)^{k-1}\prod_{j=1}^{p-1}(1-\sigma_j)}{k^n(k^{-1}+p)_n}$$

$$\times F_D(k^{-1}+p; k^{-1}+1, s_1,\ldots, s_{p-1}; k^{-1}+n+p; \sigma_0,\ldots,\sigma_{p-1}). \quad (5.18)$$

$\rho_{n,k}\binom{u_1,\ldots,u_p}{s_1,\ldots,s_p}; t)$ *is dominated for* $0 < u_1 < \cdots < u_p < t$ *by* $e^{-u_p}(cu_p + d)$, *where c and d are constants.*

(3) *The transformation defined by (5.17) mapping* $0 < u_1 < \cdots < u_p < \infty$ *onto* $0 < \sigma_{p-1} < \cdots < \sigma_1 < \sigma_0 < 1$ *has the jacobian*

$$\frac{\partial(u_1,\ldots,u_p)}{\partial(\sigma_0,\ldots,\sigma_{p-1})} = \left\{k^p \prod_{j=0}^{p-1}(1-\sigma_j)\right\}^{-1}. \quad (5.19)$$

PROOF. We first prove (1) in the case $p = 0$. From (5.3) and (5.11)

$$\rho_{n,k}(t) = e^{-t}\sum_{h=0}^{\infty}\binom{k^{-1}+h}{h}\frac{(1-e^{-kt})^h}{(kh+1)^n}$$

$$\leqslant e^{-t}\sum_{h=0}^{\infty}\binom{1+h}{h}\frac{(1-e^{-kt})^h}{(1+h)^2}$$

$$= \frac{kte^{-t}}{(1-e^{-kt})}, \quad (5.20)$$

and the result follows.

Let us next derive an expression for $\rho_{n,k}\binom{u_1,\ldots,u_p}{s_1,\ldots,s_p}; t)$ when $p \geqslant 1$. It is convenient to write

$$a_0 = 1 - e^{-kt}, \qquad a_h = 1 - e^{-k(t-u_h)}, \qquad (h = 1,\ldots, p), \quad (5.21)$$

$$s_0 = k^{-1} + 1. \quad (5.22)$$

Then (5.10) may be written as

$$e^{i(kp+1)\phi}(1-a_0)^{k-1}\prod_{j=1}^{p}(1-a_j)\cdot\prod_{h=0}^{p}\{(1-a_pe^{ik\phi}) - (a_h - a_p)e^{ik\phi}\}^{-s_h}. \quad (5.23)$$

Since $0 < u_1 < \cdots < u_p < t$, it follows that $0 < a_p < \cdots < a_0 < 1$, so that for real ϕ

$$| a_h - a_p) e^{ik\phi} | < | 1 - a_p e^{ik\phi} | , \qquad (h = 0,\dots, p-1); \tag{5.24}$$

hence, writing

$$S = \sum_{h=1}^{p} s_h , \qquad L = \sum_{h=0}^{p-1} \ell_h , \tag{5.25}$$

we obtain

$$\rho_{n,k}\binom{u_1,\dots,u_p}{s_1,\dots,s_p}; t\Big) = k^{-n}(1-a_0)^{k^{-1}} \prod_{j=1}^{p} (1-a_j)$$
$$\times \sum_{\ell_0,\dots,\ell_{p-1}=0}^{\infty} \left\{ \prod_{r=0}^{p-1} \binom{s_r + \ell_r - 1}{\ell_r} (a_r - a_p)^{\ell_r} \right\}$$
$$\times \sum_{h=0}^{\infty} \binom{L + S + k^{-1} + h}{h} \frac{a_p{}^h}{(L + h + p + k^{-1})^n} . \tag{5.26}$$

The remainder of the proof will be illustrated by a derivation of (5.18) when $S = n + p - 1$. Let

$$c_n = \sup_{S \geqslant r} \frac{(r)_n - r^n}{r^{n-1}}; \tag{5.27}$$

then

$$\left| \binom{r + h + n - 1}{h} \frac{(r)_n}{(r+h)^n} - \binom{r + h - 1}{h} \right| \leqslant \frac{c_n}{(r-1)} \binom{r + h - 2}{h} . \tag{5.28}$$

Hence, noting that

$$\frac{a_h - a_p}{1 - a_p} = \sigma_h , \qquad (h = 0,\dots, p-1), \tag{5.29}$$

it is found that

$$\left| \rho_{n,k}\binom{u_1,\dots,u_p}{s_1,\dots,s_p}; t\Big) - \rho_{n,k}\binom{u_1,\dots,u_p}{s_1,\dots,s_p} \right|$$
$$\leqslant c_n k^{-n}(1-a_0)^{k^{-1}} \prod_{j=1}^{p} (1-\sigma_j)$$
$$\times \sum_{\ell_0,\dots,\ell_{p-1}=0}^{\infty} \frac{(\ell_0 + 1)\, \sigma_0^{\ell_0}}{(L + p + k^{-1} - 1)_{n+1}} \prod_{r=1}^{p-1} \left\{ \binom{s_r + \ell_r - 1}{s_r - 1} \sigma_r^{\ell_r} \right\}$$
$$\leqslant c_n k^{-n}(1-a_0)^{k^{-1}} \prod_{j=1}^{p} (1-\sigma_j) \prod_{r=1}^{p-1} (1-\sigma_r)^{-1} \sum_{\ell_0=0}^{\infty} (\ell_0 + 1)^{-2}$$
$$= \frac{\pi^2}{6} c_n k^{-n} e^{-t}, \tag{5.30}$$

which converges to zero as $t \to \infty$. (Note that $(l_0 + 1) \prod_{r=1}^{p-1} (l_r + 1)_{s_r - 1}$ is a product of $S - s_p - p + 2 \leqslant n - 1$ terms.)

We may now in principle deduce formulas for the $\lim_{t\to\infty} M_n^{(Y)}(\xi^{(n)} \mid x_0; t)$ for suitably restricted $\xi_t^{(n)}$. There are apparently no recurrence relations for these limiting moments, and because of their complicated structure we shall give explicit results only for those of fourth and lower orders.

THEOREM 3. (i) *If* $(\xi_{h,t}, \Phi_\xi) \in \mathscr{J}_2$, $(h = 1, 2)$ *and* $\sum_1^2 \kappa_{\xi_h} < 1$, *then*

$$\lim_{t\to\infty} M_2^{(Y)}(\xi^{(2)} \mid x_0; t)$$

$$= k(k+1) \int_0^\infty \rho_{2,k} \binom{u}{2} \prod_{i=1}^{2} \Psi_{\xi_i}(u)\, du \int_{\mathscr{X}} \prod_{j=1}^{2} J_{\xi_j}(x; u)\, \chi(dx \mid x_0; u). \tag{5.31}$$

(ii) *If* $(\xi_{h,t}, \Phi_\xi) \in \mathscr{J}_3$, $(h = 1, 2, 3)$, *then we require* $\sum_1^3 \kappa_{\xi_h} < 1$ *for convergence. If* $\Psi_{\xi_h} \equiv 1$ *for each h, then*

$$\lim_{t\to\infty} M_3^{(Y)}(\xi^{(3)} \mid x_0; t) = k^2(k+1)^2 \int_0^\infty \int_u^\infty \rho_{3,k} \binom{u,\ v}{2,\ 2}\, du\, dv$$

$$\times \left\{ \iint_{\mathscr{X}^2} J_{\xi_1}(x; u) \prod_{j=2}^{3} J_{\xi_j}(y; v)\, \chi(dy \mid x; v-u)\, \chi(dx \mid x_0; u) \right.$$

$$\left. + \text{ terms by cyclic interchange} \right\}$$

$$+ (k-1)_3 \int_0^\infty \rho_{3,k} \binom{u}{3}\, du \int_{\mathscr{X}} \prod_{j=1}^{3} J_{\xi_j}(x; u)\, \chi(dx \mid x_0; u). \tag{5.32}$$

(iii) *If* $(\xi_t, \Phi_\xi) \in \mathscr{J}_4$, $4\kappa_\xi < 1$ *and* $\Psi_\xi \equiv 1$, *then*

$$\lim_{t\to\infty} M_4^{(Y)}(\xi \mid x_0; t)$$

$$= 6k^2(k+1)^2 \int_0^\infty \int_u^\infty \int_v^\infty \rho_{4,k} \binom{u,\ v,\ w}{2,\ 2,\ 2}\, du\, dv\, dw$$

$$\times \left\{ 2 \iiint_{\mathscr{X}^3} J_\xi(x; u)\, J_\xi(y; v)\, J_\xi^2(z; w)\, \chi(dz \mid y; w-v)\, \chi(dy \mid x; v-u) \right.$$

$$\times\, \chi(dx \mid x_0; u) + \iiint_{\mathscr{X}^3} J_\xi^2(y; v)\, J_\xi^2(z; w)\, \chi(dy \mid x; v-u)$$

$$\left. \times\, \chi(dz \mid x; w-u)\, \chi(dx \mid x_0; u) \right\}$$

$$+ 4(k-1)\, k^2(k+1)^2 \int_0^\infty \int_u^\infty \rho_{4.k} \binom{u,\ v}{2,\ 3}\, du\, dv$$

$$\times \iint_{\mathfrak{X}^2} J_\xi(x; u)\, J_\xi^3(y; v)\, \chi(dy \mid x; v-u)\, \chi(dx \mid x_0; u)$$

$$+ 6(k-1)\, k^2(k+1)^2 \int_0^\infty \int_u^\infty \rho_{4.k} \binom{u,\ v}{3,\ 2}\, du\, dv$$

$$\times \iint_{\mathfrak{X}^2} J_\xi^2(x; u)\, J_\xi^2(y; v)\, \chi(dy \mid x; v-u)\, \chi(dx \mid x_0; u)$$

$$+ (k-2)_4 \int_0^\infty \rho_{4.k} \binom{u}{4}\, du \int_{\mathfrak{X}} J_\xi^4(x; u)\, \chi(dx \mid x_0; u). \tag{5.33}$$

By virtue of Theorem 2 (1) and (2), these formulas may be proved by the method of [1] Theorem 5.

It is plausible that $\lim_{t\to\infty} M_n^{(Y)}(\xi \mid x_0; t)$ will exist if $(\xi_t, \Phi_\xi) \in \mathscr{J}_n$ and $n\kappa_\xi < 1$. From Theorem 1 and the result due to Kendall and Rao quoted in [1] Theorem 5, we infer that if n is even, then under the conditions just stated,

$$M_r^{(Y)}(\xi \mid x_0) = \mathscr{E}\{Y^r(\xi \mid x_0)\}$$

$$= \lim_{t\to\infty} M_n^{(Y)}(\xi \mid x_0; t) \tag{5.34}$$

for $r = 1, \ldots, n-1$. For all moments of $Y(\xi \mid x_0)$ to exist, it appears to be sufficient that $(\xi_t, \Phi_\xi) \in \bigcap_{n \geq 2} \mathscr{J}_n$, and that κ_ξ may be taken arbitrarily small positive. This is in contrast with the results of [1] Theorem 5, so that $N(t)$ is seen to be a weaker "normalizing" factor than e^{at}.

6. Moments of $Y(\xi \mid x_0)$ in the Case of the Simple Birth and Death Process ($\lambda > \mu$)

This process is defined by [1] Eq. (6.14). It is wellknown that

$$\Gamma[\phi; t] = p_0(t) + e^{i\phi} p_1(t) \left[1 - e^{i\phi} \frac{\lambda}{\mu} p_0(t)\right]^{-1}, \tag{6.1}$$

where

$$p_0(t) = \mu[1 - e^{-(\lambda-\mu)t}]\, [\lambda - \mu e^{-(\lambda-\mu)t}]^{-1},$$
$$p_1(t) = (\lambda-\mu)^2\, e^{-(\lambda-\mu)t}\, [\lambda - \mu e^{-(\lambda-\mu)t}]^{-2}. \tag{6.2}$$

When $\lambda > \mu$, the extinction probability is clearly

$$P_0 = \frac{\mu}{\lambda}. \tag{6.3}$$

As we would expect by a comparison with Eqs. (5.2) and (5.5), the analysis is essentially a special case of that encountered in the preceding section, with $k = 1$.

For $n = 2, 3$ the $M_n^{(Y)}(\xi^{(n)} \mid x_0; t)$ are given by (5.12) and (5.13) with $k = 1$, and $\rho_{n,k}\binom{u_1,\ldots,u_p}{2,\ldots,2}; t)$ replaced by $\rho_n(u_1, \ldots, u_p; t)$ say.

Corresponding to Theorem 2, we obtain:

THEOREM 4. *Let $\lambda > \mu$, $n \geqslant 2$. (1) If $p \leqslant n - 2$ then*

$$\rho_n(u_1, \ldots, u_p; t) = \begin{cases} O(te^{-(\lambda-\mu)t}) & \text{for} \quad p = 0, \\ O(e^{-(\lambda-\mu)t}) & \text{for} \quad p \geqslant 1, \end{cases} \tag{6.4}$$

uniformly for $0 < u_i < t$ ($i = 1, \ldots, p$).

(2) *If $0 < u_1 < \cdots < u_{n-1} < \infty$, then*

$$\lim_{t\to\infty} \rho_n(u_1, \ldots, u_{n-1}; t) = \rho_n(u_1, \ldots, u_{n-1})$$
$$= \frac{(\lambda - \mu)^{n-1} (n-1)!}{(2n-1)!} \prod_{j=0}^{n-2} (1 - \sigma_j)$$
$$\times F_D(n; 2, \ldots, 2; 2n; \sigma_0, \ldots, \sigma_{n-2}), \tag{6.5}$$

where

$$\sigma_0 = 1 - e^{-(\lambda-\mu)u_{n-1}}, \qquad \sigma_j = 1 - e^{-(\lambda-\mu)(u_{n-1}-u_j)}, \qquad (j = 1, \ldots, n-2). \tag{6.6}$$

For all t, $\rho_n(u_1, \ldots, u_{n-1}; t)$ is dominated for $0 < u_1 < \cdots < u_{n-1} < t$ by $e^{-(\lambda-\mu)u_{n-1}}(cu_{n-1} + d)$.

$$\frac{\partial(u_1, \ldots, u_{n-1})}{\partial(\sigma_0, \ldots, \sigma_{n-2})} = \left\{(\lambda - \mu)^{n-1} \prod_{j=0}^{n-2} (1 - \sigma_j)\right\}^{-1}. \tag{6.7}$$

The limit function ρ_n may be expressed in closed form, although this appears to be less useful in the calculation of moments than the form (6.5); for $n \geqslant 2$

$$\rho_n(u_1, \ldots, u_{n-1}) = \frac{1}{(n-1)!} \frac{\partial^{n-1}}{\partial u_1 \cdots \partial u_{n-1}} \left\{ \frac{1}{\prod_{j=1}^{n-1} [e^{(\lambda-\mu)u_j} - 1]} - \sum_{j=1}^{n-1} \frac{(\lambda - \mu)\, u_j\, e^{(n-1)(\lambda-\mu)u_j}}{[e^{(\lambda-\mu)u_j} - 1]^2 \prod_{\substack{k=1 \\ k\neq j}}^{n-1} [e^{(\lambda-\mu)u_k} - e^{(\lambda-\mu)u_j}]} \right\}. \tag{6.8}$$

In particular,

$$\rho_2(u) = \frac{d^2}{du^2}\left[\frac{u}{e^{(\lambda-\mu)u} - 1}\right]. \tag{6.9}$$

Theorem 4 may be used to prove an analogue of Theorem 3. The statement of the theorem may be obtained by taking $k = 1$ in Theorem 3, replacing $\rho_{n,k}\binom{u_1,\ldots,u_{n-1}}{2,\ldots,2}$ by $\rho_n(u_1, \ldots, u_{n-1})$, and replacing the conditions $\Sigma\kappa_\xi < 1$ by $\Sigma\kappa_\xi < \lambda - \mu$.

7. Applications

The examples in this section correspond to those of [1] Section 6. The suffix d will be used to distinguish parameters of the "diffusion" processes.

Example (a). Gaussian Diffusion on the Real Line. Taking $\xi(x) = \delta(X \mid x)$, where X is a Borel set, it follows from [1] (6.10) that for a general life-span distribution

$$\frac{N(\sigma_d \sqrt{t}\, X \mid x_0; t)}{N(t)}\Big|_{N(t) \nrightarrow 0} \xrightarrow[\text{i.p.}]{} \frac{1}{\sqrt{2\pi}} \int_X e^{-y^2/2}\, dy, \tag{7.1}$$

a nonrandom measure. Let us write

$$\overline{x^n}\,(t \mid x_0) = \frac{1}{N(t)} \sum_{i=1}^{N(t)} x_i \Big|_{N(t')\nrightarrow 0}. \tag{7.2}$$

Then if n is a positive *even* integer,

$$\frac{\overline{x^n}\,(t \mid x_0)}{(\sigma_d \sqrt{t}\,)^n} \xrightarrow[\text{i.p.}]{} \nu_n\,, \tag{7.3}$$

while if n is *odd*,

$$\frac{\overline{x^n}(t \mid x_0)}{(\sigma_d \sqrt{t}\,)^{n-1}} \xrightarrow[\text{i.p.}]{} \nu_{n+1} Y(x \mid x_0), \tag{7.4}$$

where

$$Y(x \mid x_0) = p\text{-}\lim_{t\to\infty} \bar{x}(t \mid x_0). \tag{7.5}$$

The random function $\bar{x}(t \mid x_0)$ measures the *average position* of the members of the population at time t. Now taking $\xi(x) = H_2(x)$, we have from [1] (6.13)

$$\overline{x^2}(t \mid x_0) - \sigma_d{}^2 t \xrightarrow[\text{i.p.}]{} Y(H_2 \mid x_0). \tag{7.6}$$

The random function

$$\begin{aligned} s^2(t \mid x_0) &= \frac{1}{N(t)} \sum_{i=1}^{N(t)} [x_i - \bar{x}(t \mid x_0)]^2 \Big|_{N(t') \not\to 0} \\ &= \overline{x^2}(t \mid x_0) - \{\bar{x}(t \mid x_0)\}^2 \end{aligned} \tag{7.7}$$

measures the *dispersion* of the members of the population about their average position at time t. From (7.5-7.6)

$$s^2(t \mid x_0) - \sigma_d{}^2 t \xrightarrow[\text{i.p.}]{} Y(H_2 \mid x_0) - Y^2(x \mid x_0) = \Sigma\,(x_0), \text{ say.} \tag{7.8}$$

Adke and Moyal ([5] Section IV) have considered the asymptotic behaviour of $s^2(t \mid x_0)$ conditional upon a fixed finite number of survivors, when the branching process is of simple birth and death type. They found that under this condition the average position gets progressively more diffuse, while the dispersion does not grow without limit. When we stipulate only that $N(t) \not\to 0$, the conclusion drawn from (7.5) and (7.8) is essentially the reverse.

The evaluation of even the first few moments of $Y(x \mid x_0)$ and $\Sigma(x_0)$ in the case of $(k + 1)$-fold splitting is extremely lengthy, and for an outline we refer to Appendices A and B.

The distribution of $Y(x \mid 0)$ is symmetrical about the origin, as we would expect. It may be shown that

$$\text{Var}\{Y(x \mid 0)\} = (k^{-1} + 1)\,\sigma_d{}^2, \tag{7.9}$$

$$\mathscr{E}\{Y^4(x \mid 0)\} = \sigma_d{}^4\{2k^{-3}(k + 1)\,\psi'(k^{-1}) + \tfrac{1}{4}\,k^{-4}(2k + 1)\,(11k^3 - 8k^2 - 9k + 18)\}, \tag{7.10}$$

where $\psi(z)$ is the logarithmic derivative of the gamma function, i.e.,

$$\psi'(z) = \sum_{\ell=0}^{\infty} (\ell + z)^{-2}. \tag{7.11}$$

It might also be expected that $Y(x \mid 0)$, the limiting average position, is normally distributed. That this is not the case will be clear from (7.9-7.10); in fact, the *kurtosis* is given by

$$\gamma(k) = \frac{2}{k(k + 1)}\,\psi'(k^{-1}) + \frac{(10k^4 - 29k^3 - 38k^2 + 27k + 18)}{4k^2(k + 1)^2}. \tag{7.12}$$

Some values of $\gamma(k)$ are given in the following table:

TABLE 1

k	$\gamma(k)$	k	$\gamma(k)$
1	0.895	7	2.788
2	0.589	8	2.970
3	1.308	9	3.117
4	1.859	10	3.239
5	2.259	∞	$4\frac{1}{2}$
6	2.558		

For $k \geqslant 10$, $\gamma(k)$ is given approximately by

$$\gamma(k) \sim 4\tfrac{1}{2} - 14.25k^{-1} + 17.8k^{-2} - 16k^{-3}. \tag{7.13}$$

In the case of the simple birth and death process ($\lambda > \mu$):

$$\operatorname{Var}\{Y(x \mid 0)\} = \frac{2\sigma_d^2}{\lambda - \mu}, \tag{7.14}$$

$$\mathscr{E}\{Y^4(x \mid 0)\} = \frac{(\frac{2}{3}\pi^2 + 9)\,\sigma_d^4}{(\lambda - \mu)^2}. \tag{7.15}$$

The first two moments of $Y(H_2 \mid 0)$ for $(k + 1)$-fold splitting are:

$$\mathscr{E}Y(H_2 \mid 0) = 0, \qquad \operatorname{Var}\{Y(H_2 \mid 0)\} = 4k^{-2}(k^{-1} + 1)\,\psi'(k^{-1})\,\sigma_d^4, \tag{7.16}$$

while in the simple birth and death case we also have:

$$\mathscr{E}\{Y^3(H_2 \mid 0)\} = \frac{96\sigma_d^6}{(\lambda - \mu)^3}\left[\sum_{n=1}^{\infty}\frac{1}{n^2} - \sum_{n=1}^{\infty}\frac{1}{n^3}\right], \tag{7.17}$$

skewness $Y(H_2|0) = 0.890.$

Finally, we consider $\Sigma(0)$ defined by Eq. (7.8). From (7.9) and (7.14), we have for $(k + 1)$-fold splitting:

$$\mathscr{E}\Sigma(0) = -(k^{-1} + 1)\,\sigma_d^2. \tag{7.18}$$

Also,

$$\mathscr{E}\Sigma^2(0) = \operatorname{Var} Y(H_2 \mid 0) - 2\mathscr{E}\{Y(H_2 \mid 0)\,Y^2(x \mid 0)\} + \mathscr{E}\{Y^4(x \mid 0)\}. \tag{7.19}$$

The mixed moment on the right-hand side is evaluated by taking $\xi_1(x) = H_2(x)$, $\xi_2(x) = \xi_3(x) = x$ in (5.32).

It is found that

$$\operatorname{Var} \Sigma(0) = \sigma_d^4 \left\{ 6\tfrac{1}{2} + k^{-4} \left[2(k+1)(4k+3) \sum_{n=1}^{\infty} \frac{1}{(n+k^{-1})^2} - \tfrac{1}{4}(13k^3 + 62k^2 - 3k - 18) \right] \right\}. \tag{7.20}$$

TABLE 2

k	$\sigma_d^{-4} \operatorname{Var} \Sigma(0)$	k	$\sigma_d^{-4} \operatorname{Var} \Sigma(0)$
1	11.058	7	6.004
2	5.231	8	6.068
3	5.401	9	6.117
4	5.637	10	6.157
5	5.803	∞	$6\frac{1}{2}$
6	5.918		

For $k \geqslant 10$, Var $\Sigma(0)$ is given approximately by

$$\operatorname{Var} \Sigma(0) \sim \sigma_d^4[6\tfrac{1}{2} - 3.25k^{-1} - 2.3k^{-2} + 5k^{-3}]. \tag{7.21}$$

~~Table 2 leads to the interesting conclusion that the population is most strongly concentrated about the average position in the case of three-fold-splitting~~.

EXAMPLE (b). RANDOM WALK ON THE INTEGERS. Taking $\xi(x) = \delta(X \mid x)$, where X is a Borel set on the real line,

$$\left.\frac{N(\sqrt{(\sigma_d + \rho_d)\, t}\, X + (\sigma_d - \rho_d)\, t \mid x_0;\, t)}{N(t)}\right|_{N(t') \to \infty} \xrightarrow[\text{i.p.}]{} \frac{1}{\sqrt{2\pi}} \int_X e^{-y^2/2}\, dy. \tag{7.22}$$

Taking $\xi(x) = x$:

$$\bar{x}(t \mid x_0) - (\sigma_d - \rho_d)\, t \xrightarrow[\text{i.p.}]{} Y(x \mid x_0). \tag{7.23}$$

whose first four moments in the case of $(k + 1)$-fold splitting are:

$$\mathscr{E} Y(x \mid 0) = 0;$$

$$\operatorname{Var}\{Y(x \mid 0)\} = (k^{-1} + 1)(\sigma_d + \rho_d);$$

$$\mathscr{E}\{Y^3(x \mid 0)\} = \tfrac{1}{2}(k^{-1} + 2)(\sigma_d - \rho_d);$$

$$\mathscr{E}\{Y^4(x \mid 0)\} = (\sigma_d + \rho_d)^2 \{2k^{-3}(k+1)\, \psi'(k^{-1}) + \tfrac{1}{4} k^{-4}(2k+1)(11k^3 - 8k^2 - 9k + 18\} + \tfrac{1}{3}(\sigma_d + \rho_d)(k^{-1} + 3). \tag{7.24}$$

EXAMPLE (c). DIFFUSION IN A COMPACT REGION. In the case of Gaussian diffusion on the finite interval ([1] (6.33)), and general life-span distribution:

$$\frac{N(X \mid x_0;\, t)}{N(t)}\bigg|_{N(t')\to 0} \xrightarrow[\text{i.p.}]{} \alpha_0 \mu(X), \tag{7.25}$$

where $\mu(X)$ is the ordinary Lebesgue measure of the Borel set X.

If n is even,

$$\overline{x^n}(t \mid x_0) \xrightarrow[\text{i.p.}]{} \frac{(L/2)^n}{(n+1)}, \tag{7.26}$$

whereas if n is odd, and $2\kappa_1 < a$

$$e^{\kappa_1 t}\, \overline{x^n}(t \mid x_0) \xrightarrow[\text{i.p.}]{} \left\{ 2\sqrt{\frac{2}{L}} \int_0^{L/2} x^n \sin \frac{\pi x}{L}\, dx \right\} Y(\alpha_1 \mid x_0), \tag{7.27}$$

where

$$\mathscr{E} Y(\alpha_1 \mid x_0) = \alpha_1(x_0). \tag{7.28}$$

From [1] (6.36), we have for n even and $2\kappa_2 < a$

$$e^{\kappa_2 t} \left\{ \overline{x^n}(t \mid x_0) - \frac{(L/2)^n}{(n+1)} \right\} \xrightarrow[\text{i.p.}]{} \left\{ 2\sqrt{\frac{2}{L}} \int_0^{L/2} x^n \cos \frac{2\pi x}{L}\, dx \right\} Y(\alpha_2 \mid x_0), \tag{7.29}$$

where

$$\mathscr{E} Y(\alpha_2 \mid x_0) = \alpha_2(x_0). \tag{7.30}$$

In the case of $(k+1)$-fold splitting

$$\mathscr{E} Y^2(\alpha_1 \mid x_0) = \frac{k(k+1)}{L} \left\{ B_k(2\kappa_1) - B_k(2\kappa_1 - \kappa_2) \cos \frac{2\pi x_0}{L} \right\}, \quad (4\kappa_1 < 1),$$

$$\mathscr{E} Y^2(\alpha_2 \mid x_0) = \frac{k(k+1)}{L} \left\{ B_k(2\kappa_2) - B_k(\kappa_2) \cos \frac{4\pi x_0}{L} \right\}, \quad (4\kappa_2 < 1), \tag{7.31}$$

where

$$B_k(\beta) = \int_0^\infty e^{\beta u} \rho_{2,k} \binom{u}{2} du$$

$$= (k+1)^{-1} + \beta k^{-1} + \beta^2 k^{-1} (1-\beta)^{-1} \times {}_3F_2 \left[\begin{matrix} 1,\ k^{-1}, & k^{-1}; & 1 \\ k^{-1}+1,\ (1-\beta) & k^{-1}+1 & \end{matrix} \right]. \tag{7.32}$$

Here ${}_3F_2$ denotes the generalized hypergeometric function of one variable ([5], p. 8). In particular, taking $k = 1$ (binary splitting)

$$B_1(\beta) = \frac{1}{2} + \beta + \beta^2 \sum_{n=1}^{\infty} \frac{1}{(n-\beta)^2}\,. \tag{7.33}$$

No worthwhile simplification appears to take place for higher moments, even in the case $k = 1$.

Appendix A: Elementary Properties of the Lauricella F_D-Type Hypergeometric Functions

Some easily derived contraction formulas are:

$$ {}_D(\alpha; \beta_1, \dots, \beta_p; \gamma; x_1, \dots, x_{p-1}, 0) = F_D(\alpha; \beta_1, \dots, \beta_{p-1}; \gamma; x_1, \dots, x_{p-1}); \tag{A.1}$$

$$F_D(\alpha; \beta_1, \dots, \beta_p; \gamma; x_1, \dots, x_{p-1}, x_{p-1}) = F_D(\alpha; \beta_1, \dots, \beta_{p-1} + \beta_p; \gamma; x_1, \dots, x_{p-1}); \tag{A.2}$$

$$F_D(\alpha; \beta_1, \dots, \beta_p; \gamma; x_1, \dots, x_{p-1}, 1) = \frac{\Gamma(\gamma)\,\Gamma(\gamma-\alpha-\beta_p)}{\Gamma(\gamma-\alpha)\,\Gamma(\gamma-\beta_p)} \times F_D(\alpha; \beta_1, \dots, \beta_{p-1}; \gamma-\beta_p; x_1, \dots, x_{p-1}), \tag{A.3}$$

where $\Gamma(\cdot)$ here denotes the gamma function. The following differentiation formulas are generalizations of [6] Vol. 1, Section 2.8, formulas (20) and (25), respectively:

$$F_D(\alpha; \beta_1, \dots, \beta_p; \gamma; x_1, \dots, x_p) = \frac{(\gamma-1)}{(\alpha-1)(\beta_p-1)} \frac{\partial}{\partial x^p} F_D(\alpha-1; \beta_1, \dots, \beta_{p-1}, \beta_p-1; \gamma-1; x_1, \dots, x_p);$$

$$(1-x_p)^{\beta_p-n-1} F_D(\alpha; \beta_1, \dots, \beta_p; \gamma; x_1, \dots, x_p) \tag{A.4}$$

$$= \frac{(\gamma-n)_n}{(\beta_p-n)_n\,(\gamma-n-\alpha)_n} \frac{\partial^n}{\partial {x_p}^n} \times (-1)^n (1-x_p)^{\beta_p-1} F_D(\alpha; \beta_1, \dots, \beta_{p-1}, \beta_p-n; \gamma-n; x_1, \dots, x_p)]. \tag{A.5}$$

Appendix B: Evaluation of Integrals Involving Hypergeometric Functions

In order to prove (7.9), we note that when $\xi(x) = x$, we have

$$\Psi_\xi \equiv 1, \qquad J_\xi(x; \tau) \equiv x. \tag{B.1}$$

Hence, from (5.32)

$$\mathscr{E}\{Y^2(x \mid 0)\} = \sigma_d^2 k(k+1) \int_0^\infty \rho_{2,k} \binom{u}{2} u \cdot du. \tag{B.2}$$

Making the substitution

$$\sigma_0 = 1 - e^{-ku}, \tag{B.3}$$

it follows from Theorem 2 (2-3) and Eq. (A.5) that

$$\begin{aligned}
\mathscr{E}\{Y^2(x \mid 0)\} &= \frac{\sigma_d^2}{k^2(k^{-1}+2)} \int_0^1 \log \frac{1}{(1-\sigma_0)} (1-\sigma_0)^{k^{-1}-1} \\
&\qquad \times F(k^{-1}+1, k^{-1}+1; k^{-1}+3; \sigma_0)\, d\sigma \\
&= k^{-1}\sigma_d^2 \int_0^1 \log \frac{1}{(1-\sigma_0)} \frac{d}{d\sigma_0} \\
&\qquad \times [-(1-\sigma_0)^{k^{-1}} F(k^{-1}, k^{-1}+1; k^{-1}+2; \sigma_0)]\, d\sigma_0 \\
&= (k^{-1}+1)\,\sigma_d^2 \int_0^1 \frac{d}{d\sigma_0} \\
&\qquad \times [-(1-\sigma_0)^{k^{-1}} F(k^{-1}, k^{-1}; k^{-1}+1; \sigma_0)]\, d\sigma_0 \\
&= (k^{-1}+1)\,\sigma_d^2.
\end{aligned} \tag{B.4}$$

As an example of a more involved case, we shall indicate the derivation of (7.10). Substituting from (B.1) in (5.33), and making the transformations

$$\sigma_0 = 1 - e^{-kw}, \qquad \sigma_1 = 1 - e^{-k(w-u)}, \qquad \sigma_2 = 1 - e^{-k(w-v)} \tag{B.5}$$

in the (u, v, w) integral,

$$\sigma_0 = 1 - e^{-kv}, \qquad \sigma_1 = 1 - e^{-k(v-u)} \tag{B.6}$$

in the two (u, v) integrals, and (B.3) in the last term, it is found that

$$\begin{aligned}
\mathscr{E}\{Y^4(x \mid 0)\} = \sigma_d^4 \Bigg\{ &\frac{6(k^{-1}+1)^3}{k^3(k^{-1}+3)_4} [9I_1 + 2I_2 - 10I_3 + 4I_4 - 5I_5] \\
&+ \text{integrals of lower order} \Bigg\},
\end{aligned} \tag{B.7}$$

where

$$\begin{aligned}
I_1 = \int_0^1 \int_0^{\sigma_0} \int_0^{\sigma_1} &\log^2 \frac{1}{(1-\sigma_0)} (1-\sigma_0)^{k^{-1}-1} \\
&\times F_D(k^{-1}+3; k^{-1}+1, 2, 2; k^{-1}+7; \sigma_0, \sigma_1, \sigma_2)\, d\sigma_0\, d\sigma_1\, d\sigma_2 .
\end{aligned} \tag{B.8}$$

I_2, I_3, I_4, and I_5 are obtained by replacing $\log^2 1/(1-\sigma_0)$ by $\log^2 1/(1-\sigma_1)$, $\log 1/(1-\sigma_0) \log 1/(1-\sigma_1)$, $\log 1/(1-\sigma_1) \log 1/(1-\sigma_2)$ and $\log 1/(1-\sigma_0)$ $\log 1/(1-\sigma_2)$, respectively. Using the formulas of Appendix A, these integrals may be reduced to linear combinations of a few basic types; for example,

$$\begin{aligned} I_3 = &-\frac{(k^{-1}+4)_3}{3(k^{-1}+1)} \mathscr{I}_1 + \frac{(k^{-1}+3)_4}{6(k^{-1}+2)^2} \mathscr{I}_2(k^{-1}+2) \\ &\qquad + \frac{1}{6} k(k^{-1}+4)_3 [\mathscr{I}_1 + \mathscr{I}_3] \\ &- \frac{1}{12}(k^{-1}+3)_4 \left\{ \frac{1}{(k^{-1}+1)^2} + \frac{k}{(k^{-1}+1)} \mathscr{I}_2(k^{-1}+1) \right. \\ &\qquad \left. + \frac{k}{(k^{-1}+2)} [\mathscr{I}_4 - \mathscr{I}_5] \right\}, \end{aligned} \tag{B.9}$$

where

$$\begin{aligned} \mathscr{I}_1 &= \int_0^1 \log \frac{1}{(1-\sigma)} (1-\sigma)^{k^{-1}-1} F(k^{-1}+1, k^{-1}+3; k^{-1}+4; \sigma)\, d\sigma \\ &= \tfrac{1}{2}(k^{-1}+3)\{\psi'(k^{-1}) - k^2 + (k^{-1}+1)^{-1}\}; \\ \mathscr{I}_2(\alpha) &= \int_0^1 (1-\sigma)^{\alpha-1} F(\alpha, \alpha; \alpha+1; \sigma)\, d\sigma = \alpha\psi'(\alpha); \\ \mathscr{I}_3 &= \int_0^1\int_0^\sigma (1-\sigma)^{k^{-1}-1} F_1(k^{-1}+3; k^{-1}, 1; k^{-1}+4; \sigma, \sigma')\, d\sigma\, d\sigma' \\ &= \tfrac{1}{4}(k^{-1}+3)[3k + (k^{-1}+1)^{-1}], \\ \mathscr{I}_4 &= \int_0^1\int_0^\sigma (1-\sigma)^{k^{-1}-1} F_1(k^{-1}+2; k^{-1}, 1; k^{-1}+3; \sigma, \sigma')\, d\sigma\, d\sigma' \\ &= k(k^{-1}+2), \\ \mathscr{I}_5 &= \int_0^1 (1-\sigma)^{k^{-1}-1} F(k^{-1}, k^{-1}+2; k^{-1}+3; \sigma)\, d\sigma \\ &= \tfrac{1}{2}(k^{-1}+2)[k + (k^{-1}+1)^{-1}]. \end{aligned} \tag{B.10}$$

Hence

$$I_3 = \frac{(k^{-1}+3)_4}{6(k^{-1}+1)_2} \left\{ -\psi'(k^{-1}) + \frac{(4k^2+11k+2)}{2(k^{-1}+1)} \right\}. \tag{B.11}$$

References

1. A. W. Davis. Branching-diffusion processes with no absorbing boundaries I. *Math. Anal. Appl.*, in press.
2. T. E. Harris. "The Theory of Branching Processes." Springer-Verlag, Berlin, 1963.
3. A. W. Davis. On the theory of birth, death and diffusion processes. *J. Appl. Prob.* **2** (1965), 293-322.

4. P. Appell and J. Kampe de Feriet. "Fonctions hypergéométriques et hypersphériques." Gauthier Villars, Paris, 1926.
5. S. R. Adke and J. E. Moyal. A birth, death and diffusion process. *J. Math. Anal. Appl.* **7** (1963), 209-224.
6. A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi. "Higher Transcendental Functions." McGraw-Hill, New York, 1953.
7. W. N. Bailey. "Generalized Hypergeometric Series," Cambridge Tracts in Mathematics and Mathematical Physics, No. 32.

PRINTED IN BRUGES, BELGIUM, BY THE ST CATHERINE PRESS, LTD.

Adv. Appl. Prob. **2**, 83–109 (1970)
Printed in Israel

SOME GENERALIZATIONS OF BAILEY'S BIRTH DEATH AND MIGRATION MODEL

A. W. DAVIS, *C.S.I.R.O., Adelaide*

Summary

Some results for a general Markov branching-diffusion process are presented, and applied to a model recently considered by Bailey. Moments of the limiting distributions of certain natural measures of the spatial location and dispersion of the population are shown to be expressible in terms of the Lauricella F_D-type hypergeometric functions, when the population multiplies according to the simple birth and death process with $\lambda > \mu$.

1. Introduction

In a recent article, Bailey (1968) has derived some results for a simple birth, death and migration process as a preliminary to studies of the spatial distribution of individuals in more complex epidemic processes. Bailey assumes in his model that the population is distributed over the points of a discrete set: the integer points of the real line, or the nodes of a two- or three-dimensional lattice. The colony at each location multiplies according to the simple birth and death process, and migration takes place to its nearest neighbours at a constant rate. Initially, the population consists of an arbitrary number of individuals (possibly infinitely many) distributed in a prescribed manner throughout the lattice. As possible fields of application of the model, Bailey suggests the growth of cells in biology, and the spread of population in the theory of urban development. He obtains generating functions for the expected sizes of the colonies and, in the one-dimensional case, for variances and covariances of colony size. In the case of a single ancestor, this model is equivalent to a special case of a general "branching-diffusion" process considered by Davis ((1965), (1967a), (1967b)), in which the population multiplies according to the Bellman-Harris age-dependent branching process, and the state of each individual varies according to an arbitrary homogeneous Markov process, independently of the remainder of the population, throughout a region with no absorbing boundaries. A model equivalent to Bailey's is obtained by taking this Markov process to be the simple symmetrical random walk on the lattice points. Another case of the general model has been con-

Received in revised form 28 April 1969.

sidered by Adke and Moyal (1963), who assumed that the population size was determined by the simple birth and death process, and that the individuals were subject to ordinary Brownian diffusion along the real line. These authors discussed the asymptotic distribution of mean position and spatial dispersion of the population, conditional upon a fixed finite number of survivors. In later sections of this paper we shall present some results for these quantities in the case of Bailey's process with $\lambda > \mu$, conditional only upon the population not becoming extinct. Reference should also be made to investigations of more complex branching-diffusion processes in which the state space is a compact region with an absorbing boundary. Sevast'yanov (1958) and Conner (1966) have considered asymptotic properties of the population size, and Watanabe (1965) has investigated the numbers of individuals in subsets of the region.

For the general theory of stochastic population processes, in which the state (location, age, etc.) of each individual can range over an arbitrary (not necessarily finite set), we refer the reader to Moyal ((1962), (1964)). In particular, Moyal ((1962), Section 6) gives the foundations of a rigorous treatment of populations which have infinitely many members with positive probability. In the present paper, however, we shall assume only a finite number of ancestors; the population is therefore finite at any time with probability one, although some results would remain formally valid for the infinite case. Moreover, in order to avoid complications arising from the ages of the ancestors, we shall consider only Markov branching-diffusion processes, in which the life-time of each individual has a negative exponential distribution. Some preliminary results for this model were given by Moyal ((1964), Section 7).

The relevant theory and definitions are summarised in Section 2, and the basic integral equation for the characteristic functional (c.fl.) of the process is introduced. An iterative form of solution of the latter in terms of conditional c.fl.'s is discussed in Section 3. Recurrence relations for the moment distributions are then presented in the following section, Bailey's results being indicated as corollaries. In Section 5 the author's mean-square convergence theory for branching-diffusion processes is outlined, and applications to Bailey's model with $\lambda > \mu$ are made in Section 6. The m.s. theory implies the convergence in probability of certain natural measures of spatial location and dispersion, and as a first step towards investigating the moments of the limiting distributions, recurrence relations for conditional moment distributions are presented in Section 7.

When the population develops from n ancestors according to the simple birth and death process, it is shown in Section 8 that the limiting moments are expressible in terms of the Lauricella F_D-type hypergeometric functions, some of whose elementary properties are given in the Appendix. These results are finally applied in Section 9 to obtain the first four central moments of the limiting mean position in Bailey's model, and also the first two moments of a random variable related to the measure of dispersion. These moments are found to be polynomials in the central moments of the initial spatial distribution of the ancestors.

2. The model

Let $\mathscr{X}$ denote the arbitrary state space of the members of the population, and suppose that at time $t = 0$ there are n individuals located at the (not necessarily distinct) points $y_1, \cdots, y_n$ of $\mathscr{X}$. The unordered set of points will be denoted by $y^{(n)}$. Each individual is assumed to have an exponentially distributed life-time l,

$$G(t) = \Pr\{l \leqq t\} = 1 - e^{-\alpha t}, \qquad (\alpha > 0, t \geqq 0), \tag{2.1}$$

and at the end of its life it is replaced by r similar offspring with probability $p_r(\sum_{r=0}^{\infty} p_r = 1)$, all located initially at the final position of their parent. The generating function of the distribution (p_r) will be denoted by

$$f(z) = \sum_{r=0}^{\infty} p_r z^r, \tag{2.2}$$

and its factorial moments by

$$m = f'(1-), \; m_{(s)} = f^{(s)}(1-), \quad (s = 2, 3, \cdots). \tag{2.3}$$

These will be assumed finite to any required order.

The "motion" of each individual is determined by a temporally homogeneous Markov transition probability distribution $\chi(\cdot \mid x; t)$, defined on the Borel subsets X of $\mathscr{X}$. Here $\chi(X \mid x; t)$ denotes the probability that an individual initially at x in $\mathscr{X}$ will be located at some point in the set X after time t, conditional upon its survival during this interval. The Markov property of χ is expressed by the Chapman-Kolmogorov equation,

$$\int_{x \in \mathscr{X}} \chi(X \mid x; u) \chi(dx \mid y; v) = \chi(X \mid y; u + v), \tag{2.4}$$

and we also have

$$\chi(X \mid x; 0) = \delta(X \mid x) = \begin{cases} 1 & \text{if } x \in X, \\ 0 & \text{otherwise.} \end{cases} \tag{2.5}$$

Since $\mathscr{X}$ has no absorbing boundaries, the population size $N(t)$ at time t constitutes a temporally homogeneous Markov branching process of the type discussed in Harris ((1963), Section V.9). If $\Gamma(\theta; t)$ denotes the characteristic function (c.f.) of $N(t)$ conditional upon a single ancestor,

$$\Gamma(\theta;t) = \sum_{s=0}^{\infty} p_s(t)e^{is\theta}, \tag{2.6}$$

then $\Gamma(\theta;t)$ satisfies the differential equation

$$\frac{\partial}{\partial t}\Gamma(\theta;t) = \alpha[f(\Gamma(\theta;t)) - 1], \qquad \Gamma(\theta;0) = e^{i\theta}, \tag{2.7}$$

or, equivalently, the integral equation

$$\Gamma(\theta;t) = e^{-\alpha t + i\theta} + \alpha\int_0^t f(\Gamma(\theta;t-u))e^{-\alpha u}du. \tag{2.8}$$

The c.f. of $N(t)$ given n ancestors is of course $(\Gamma(\theta;t))^n$. Following Moyal (1962), the spatial distribution of the population conditional on n ancestors at $y^{(n)}$ may be specified by the counting process $N(\cdot \mid y^{(n)};t)$, defined on the Borel subsets X of $\mathscr{X}$: $N(X \mid y^{(n)};t)$ denotes the number of individuals in the set X at time t. The complete joint distribution of these random variables is in turn specified by the c.fl. $\Gamma(\xi \mid y^{(n)};t)$, defined on the class of all real, finite-valued Borel functions $\xi(\cdot)$ on $\mathscr{X}$:

$$\begin{aligned}\Gamma(\xi \mid y^{(n)};t) &= \mathscr{E}\exp\left\{i\int_{\mathscr{X}}\xi(x)N(dx \mid y^{(n)};t)\right\} \\ &= \mathscr{E}\exp\left\{i\sum_{s=1}^{N(t)}\xi(x_s)\right\},\end{aligned} \tag{2.9}$$

where $\mathscr{E}$ denotes expectation, and the x_s ($s = 1,\cdots,N(t)$) are the locations of the individuals existing at time t. It is also convenient to introduce the random functional

$$N(\xi \mid y^{(n)};t) = \sum_{s=1}^{N(t)}\xi(x_s) = \int_{\mathscr{X}}\xi(x)N(dx \mid y^{(n)};t), \tag{2.10}$$

whose c.f. for given ξ is obviously $\Gamma(\theta\xi \mid y^{(n)};t)$, ($\theta$ real).

Since each ancestor generates a subpopulation independently of the others, we have the multiplicative property

$$\Gamma(\xi \mid y^{(n)};t) = \prod_{j=1}^{n}\Gamma(\xi \mid y_j;t), \tag{2.11}$$

which is equivalent to

$$N(\xi \mid y^{(n)};t) = \sum_{j=1}^{n}N(\xi \mid y_j;t). \tag{2.12}$$

Also, taking $\xi(x) \equiv \theta$, the c.fl. reduces to the c.f. of the population size. By analogy with (2.8), the c.fl. conditional upon a single ancestor initially at y is readily seen to be a solution of the integral equation

(2.13)
$$\Gamma(\xi \mid y;t) = e^{-\alpha t}\int_{\mathscr{X}} e^{i\xi(x)}\chi(dx \mid y;t) + \alpha\int_0^t \int_{\mathscr{X}} f(\Gamma(\xi \mid x; t-u))\chi(dx \mid y;u)e^{-\alpha u}du,$$

(cf. Moyal (1964), Equation (7.7)). In the case of a general life-time distribution, it was shown in Davis ((1965), Section 3) that there is a unique solution of this equation such that $|\Gamma| \leqq 1$ which is in fact the c.fl. of a population process.

Bailey's (1968) model is equivalent to the particular case of the above scheme in which the branching mechanism is of simple birth and death type, with constant birth rate λ and death rate μ,

(2.14)
$$\alpha = \lambda + \mu, \quad f(z) = \frac{\mu + \lambda z^2}{\mu + \lambda},$$
$$\Gamma(\theta;t) = p_0(t) + e^{i\theta}p_1(t)\left[1 - e^{i\theta}\frac{\lambda}{\mu}p_0(t)\right]^{-1},$$

where

(2.15)
$$p_0(t) = \mu[1 - e^{-(\lambda-\mu)t}][\lambda - \mu e^{-(\lambda-\mu)t}]^{-1}$$
$$p_1(t) = (\lambda-\mu)^2 e^{-(\lambda-\mu)t}[\lambda - \mu e^{-(\lambda-\mu)t}]^{-2}.$$

In the one-dimensional case, each individual executes a simple symmetric random walk on the integers; i.e., an individual at x at time t has equal probabilities $\frac{1}{2}v.\delta t$ of moving to either $x+1$ or $x-1$ during the interval $(t, t+\delta t)$. Thus the transition probability χ may be expressed in terms of the modified Bessel function of the first kind,

(2.16)
$$\chi(x \mid y;t) = e^{-vt}I_{x-y}(vt),$$

with generating function

(2.17)
$$\phi(z \mid y;t) = \sum_{x=-\infty}^{\infty} z^x\chi(x \mid y;t) = z^y\exp\{-vt[1 - \tfrac{1}{2}(x+z)^{-1})]\}.$$

Equation (2.13) for the c.fl. with integration over $\mathscr{X}$ interpreted as summation over the positive and negative integers is the analogue of Bailey's ((1968), Equation (3)) infinite system of differential equations.

Some asymptotic results for χ so defined and a general branching process have been given in Davis ((1967a), page 292, (1967b), page 20).

For the two-dimensional case, an individual at (x_1, x_2) has equal probabilities $\frac{1}{4}v.\delta t$ of migrating to any one of its four nearest neighbours during $(t, t+\delta t)$, whence

(2.18) $$\chi((x_1,x_2)\,|\,(y_1,y_2);t) \;=\; \chi(x_1\,|\,y_1;t)\,\chi(x_2\,|\,y_2;t),$$

where $\chi(x_i\,|\,y_i;t)$ is defined by (2.16). The three-dimensional case is similarly specified.

3. Conditional probabilities

The basic integral equation (2.13) cannot of course be solved directly. However, an iterative form of solution can be given in terms of the c.fl.'s for the process conditional upon the population size at time t being $N(t)=s$, along the lines of Adke and Moyal ((1963), Sections II and IV). Substituting

(3.1) $$\Gamma(\xi+\theta\,|\,y;t) = \sum_{s=0}^{\infty} e^{is\theta}\,\Gamma_s(\xi\,|\,y;t) \qquad (\theta \text{ real})$$

in (2.13), we obtain

(3.2) $$\Gamma_0(\xi\,|\,y;t) \;\equiv\; p_0(t),$$

the probability of zero survivors at time t, while for $s \geqq 1$:

(3.3) $$\Gamma_s(\xi\,|\,y;t) \;=\; F_s(y;t) + \alpha\int_0^t\int_{\mathscr{X}} f'(p_0(u))\Gamma_s(\xi|x;u)\chi(dx|y;t-u)e^{-\alpha(t-u)}\,du,$$

where

(3.4) $$F_1(\xi\,|\,y;t) \;=\; e^{-\alpha t}\int_{\mathscr{X}} e^{i\xi(x)}\chi(dx\,|\,y;t),$$

$$F_s(\xi\,|\,y;t)$$
$$= \alpha\sum_{k=2}^{s}\frac{1}{k!}\int_0^t\int_{\mathscr{X}} f^{(k)}(p_0(u))\sum_{\substack{s_1+\cdots+s_k=s\\(s_i>0)}}\left(\prod_{j=1}^{k}\Gamma_{s_j}(\xi\,|\,x;u)\right)\chi(dx\,|\,y;t-u)e^{-\alpha(t-u)}du,$$
$$(s \geqq 2).$$

Taking $\xi \equiv 0$, we obtain recurrence relations for the $p_r(t)$. In particular,

(3.5) $$e^{\alpha t}p_1(t) = 1+\alpha\int_0^t f'(p_0(u))p_1(u)e^{\alpha u}du,$$

and it follows by direct substitutions in (3.3), using the Chapman-Kolmogorov relation (2.4), that

(3.6) $$\Gamma_1(\xi\,|\,y;t) \;=\; p_1(t)\int_{\mathscr{X}} e^{i\xi(x)}\chi(dx\,|\,y;t).$$

When F_s has the form

(3.7) $$F_s(\xi\,|\,y;t) \;=\; \alpha\int_0^t\int_{\mathscr{X}} H_s(\xi\,|\,x;u)\chi(dx\,|\,y;t-u)e^{-\alpha(t-u)}du,$$

as is the case for $s \geqq 2$, the Equation (3.3) has the solution

$$\Gamma_s(\xi \mid y;t) = \alpha \int_0^t \int_{\mathscr{X}} \frac{p_1(t)}{p_1(u)} H_s(\xi \mid x;u)\chi(dx \mid y;t-u)du. \tag{3.8}$$

To see this directly, substitution in the right-hand side of (3.3) yields

$$\alpha \int_0^t \int_{\mathscr{X}} \frac{1}{p_1(u)} H_s(\xi \mid x;u)\chi(dx \mid y;t-u)\Big\{ e^{-\alpha t}[e^{\alpha u}p_1(u) + \alpha \int_u^t f'(p_0(v))p_1(v)e^{\alpha v}dv] \Big\} du, \tag{3.9}$$

which leads to (3.8), since from (3.5) the second summand in the bracket is equal to $e^{\alpha t}p_1(t) - e^{\alpha u}p_1(u)$.

In principle, Equations (3.8), (3.7) and (3.4) yield a set of recurrence relations for the Γ_s. Some reduction of these will now be attempted when the branching mechanism is of simple birth and death type and the "diffusion" is one-dimensional, with $\chi(dx \mid y;t)$ depending on x and y only through the difference $x - y$; i.e., the c.f. of the transition probability is of the form

$$\int_{\mathscr{X}} e^{i\theta x}\chi(dx \mid y;t) = e^{i\theta y}\phi(\theta \mid t), \text{ say.} \tag{3.10}$$

For $s \geqq 2$, we then have

$$\Gamma_s(\xi \mid y;t) = \lambda \int_0^t \int_{\mathscr{X}} \frac{p_1(t)}{p_1(u)} \sum_{r=1}^{[\frac{1}{2}s]} \varepsilon_r \Gamma_r(\xi \mid x;u)\Gamma_{s-r}(\xi \mid x;u)\chi(dx \mid y;t-y)du, \tag{3.11}$$

where $[\frac{1}{2}s]$ denotes the largest integer not exceeding $\frac{1}{2}s$, and $\varepsilon_r = 1$ if $r = \frac{1}{2}s$ (s even), $\varepsilon_r = 2$ otherwise. It may be shown by induction on s that Γ_s can be written as

$$\Gamma_s(\xi \mid y;t) = \int_{\mathscr{X}^s} \exp\Big\{ i \sum_{j=1}^{s} \xi(x_j) \Big\} \Lambda_s(dx_1, \cdots, dx_s \mid y;t), \tag{3.12}$$

where Λ_s is a non-negative measure on $\mathscr{X}^s$ (the s-fold Cartesian product of $\mathscr{X}$ with itself), having total mass $p_s(t)$ and c.f. of the form

$$\begin{aligned} &\int_{\mathscr{X}^s} \exp\Big\{ i \sum_{j=1}^{s} \theta_j x_j \Big\} \Lambda_s(dx_1, \cdots, dx_s \mid y;t) \\ &= \lambda^{s-1} p_1(t) \exp\Big\{ iy \sum_{j=1}^{s} \theta_j \Big\} \int_{u_1=0}^{t} \cdots \int_{u_{s-1}=0}^{u_{s-2}} \Big(\prod_{j=1}^{s-1} p_1(u_j) \Big) \\ &\quad \beta_s(\theta_1, \cdots, \theta_s; u_1, \cdots, u_{s-1} \mid t) du_1 \cdots du_{s-1}. \end{aligned} \tag{3.13}$$

The β_s depend only on the "diffusion" process, and are given recursively by

$$\beta_1(\theta_1 \mid t) = \phi(\theta_1 \mid t),$$

$$\beta_s(\theta_1, \cdots, \theta_s; u_1, \cdots, u_{s-1} \mid t) = \phi\left(\sum_{j=1}^{s} \theta_j \mid t - u_1\right) \tag{3.14}$$
$$\times \sum_{r=1}^{[\frac{1}{2}s]} \varepsilon_r \sum_{2 \leqq k_1 < \cdots < k_{r-1} \leqq s-1} \beta_r(\theta_1, \cdots, \theta_r; u_{k_1}, \cdots, u_{k_{r-1}} \mid u_1)$$
$$\times \beta_{s-r}(\theta_{r+1}, \cdots, \theta_s; u_{l_1}, \cdots, u_{l_{s-r-1}} \mid u_1), \qquad (s \geqq 2),$$

where $l_1 < \cdots < l_{s-r-1}$ is obtained by deleting $k_1, \cdots, k_{r-1}$ from the set of integers $2, \cdots, s-1$. We have in particular

$$\beta_2(\theta_1, \theta_2; u_1 \mid t) = \phi(\theta_1 \mid u_1)\phi(\theta_2 \mid u_1)\phi(\theta_1 + \theta_2 \mid t - u_1);$$
$$\beta_3(\theta_1, \theta_2, \theta_3; u_1, u_2 \mid t) = 2\phi(\theta_1 \mid u_1)\phi(\theta_2 \mid u_2)\phi(\theta_3 \mid u_2)$$
$$\times \phi(\theta_2 + \theta_3 \mid u_1 - u_2)\phi(\theta_1 + \theta_2 + \theta_3 \mid t - u_1); \tag{3.15}$$
$$\beta_4(\theta_1, \theta_2, \theta_3, \theta_4; u_1, u_2, u_3 \mid t) = \phi(\theta_2 \mid u_2)\phi(\theta_3 \mid u_3)\phi(\theta_4 \mid u_4)$$
$$\times \{4\phi(\theta_1 \mid u_1)\phi(\theta_2 + \theta_3 + \theta_4 \mid u_1 - u_2)\phi(\theta_3 + \theta_4 \mid u_2 - u_3)$$
$$+ 2\phi(\theta_1 \mid u_2)\phi(\theta_1 + \theta_2 \mid u_1 - u_2)\phi(\theta_3 + \theta_4 \mid u_1 - u_3)\}\phi\left(\sum_{j=1}^{4} \theta_j \mid t - u_1\right).$$

Taking

$$\xi(x) = \sum_{j=1}^{h} \gamma_j \delta(X_j \mid x), \tag{3.16}$$

where $\{X_j\}$ is a collection of non-overlapping Borel subsets of $\mathscr{X}$ with union $\mathscr{X}$, $\Gamma_s(\xi \mid y; t)$ becomes the joint c.f. of the $N(X_j \mid y; t) (j = 1, \cdots, h)$, the numbers of individuals in $X_1, \cdots, X_h$ respectively, conditional upon the total population size being s. From (3.12) it follows that

$$\Pr\{N(X_j \mid y; t) = n_j; j = 1, \cdots, h \mid N(t) = s\}$$
$$= \sum_{(j_1, \cdots, j_s)} \Lambda_s(X_{j_1}, \cdots, X_{j_s} \mid y; t), \tag{3.17}$$

where the summation is extended over the $\binom{s}{n_1, \cdots, n_k}$ distinct permutations $(j_1, \cdots, j_s)$ of the set $[1^{n_1}, 2^{n_2}, \cdots, h^{n_h}]$.

In particular, for Bailey's one-dimensional model with a single ancestor initially at y, suppose that a and b are two distinct integer points, and let $N(a, t)$, $N(b, t)$ respectively denote the numbers of individuals at these locations at time t. Then, conditional upon $N(t) = 2$,

$$\Pr\{N(a,t) = 1,\ N(b,t) = 1 \mid N(t) = 2\} = 2\Lambda_2(a, b \mid y; t)$$
$$\Pr\{N(a,t) = 2 \mid N(t) = 2\} = \Lambda_2(a, a \mid y; t), \tag{3.18}$$

where $\Lambda_2(a,b\,|\,y;t)$ is the coefficient of $\theta_1^a\theta_2^b$ in the expansion of

$$\begin{aligned}(3.19)\quad &\lambda e^{-\nu t}p_1(t)(\theta_1\theta_2)^y \exp\left\{\tfrac{1}{2}\nu t\left(\theta_1\theta_2 + \frac{1}{\theta_1\theta_2}\right)\right\}\\ &\times \int_u^t e^{-\nu u}p_1(u)\exp\left\{\tfrac{1}{2}\nu u\left(\theta_1 + \frac{1}{\theta_1} + \theta_2 + \frac{1}{\theta_2} - \theta_1\theta_2 - \frac{1}{\theta_1\theta_2}\right)\right\} du.\end{aligned}$$

Even in this simple instance, the integral can be obtained in closed form only when the death rate $\mu = 0$, and $p_1(u) = e^{-\lambda u}$. The individual probabilities unfortunately seem to be so complicated as to be of doubtful practical use.

4. Moment distributions

The moments of the counting process $N(\cdot\,|\,y^{(n)};t)$ define a family of measures on the product spaces $\mathscr{X}^s$ called the *moment distributions* (Moyal (1962), Section 3). The mean distribution

$$(4.1)\qquad M(X\,|\,y^{(n)};t) = \mathscr{E}N(X\,|\,y^{(n)};t)$$

is a measure on the Borel sets X of $\mathscr{X}$ with total mass

$$(4.2)\qquad M(t) = \mathscr{E}N(t) = ne^{\alpha(m-1)t},$$

where m is the expected number of offspring at a birth (Harris (1963), Section V.6). More generally, the rth moment distribution is defined for each product set $X_1 \times \cdots \times X_r$ by

$$(4.3)\qquad M_r(X_1\times\cdots\times X_r\,|\,y^{(n)};t) = \mathscr{E}\prod_{j=1}^{r} N(X_j\,|\,y^{(n)};t),$$

and may be extended to a non-negative symmetric measure on $\mathscr{X}^r$.

Recurrence relations for these measures may be obtained from (2.13) since

$$(4.4)\quad M_r(X_1\times\cdots\times X_r\,|\,y^{(n)};t) = \left[\frac{\partial}{i_r\partial\theta_1\cdots\partial\theta_r}\Gamma\left(\sum_{j=1}^{r}\theta_j\delta(X_j\,|\,\cdot)\,|\,y^{(n)};t\right)\right]_{\theta_1=\cdots=\theta_r=0}$$

(Davis (1965), Section 4 and (1967a), Section 4).

In the case of a single ancestor, the mean distribution is a solution of the integral equation

$$(4.5)\quad M(X\,|\,y;t) = e^{-\alpha t}\chi(X\,|\,y;t) + \alpha m\int_0^t\int_{\mathscr{X}} M(X\,|\,x;t-u)\chi(dx\,|\,y;u)e^{-\alpha u}du.$$

It may be verified by direct substitution that

$$(4.6)\qquad M(X\,|\,y;t) = M(t)\chi(X\,|\,y;t),$$

(cf. Moyal (1964), Equation (7.9)), noting that

$$(4.7) \qquad M(t) = e^{-\alpha t} + \alpha m \int_0^t M(t-u)e^{-\alpha u} du\,.$$

The higher order moment distributions satisfy similar integral equations to (4.5), and may be solved to yield, in particular:

$$(4.8) \quad M_2(X_1 \times X_2 \,|\, y;t) =$$

$$M(X_1 \cap X_2 \,|\, y;t) + \alpha m_{(2)} \int_0^t \int_{\mathscr{X}} M(X_1|x;t-u)M(X_2|x;t-u)\chi(dx|y;u)e^{\alpha(m-1)u}]du$$

(cf. Moyal (1964), Equation (7.10)),

$$(4.9) \quad \begin{aligned} M_3(X_1 \times X_2 \times X_3 \,|\, y;t) &= M(X_1 \cap X_2 \cap X_3 \,|\, y;t) \\ &+ \alpha \int_0^t \int_{\mathscr{X}} [m_{(2)} M(X_1 \,|\, x;t-u) M_2(X_2 \times X_3 \,|\, x;t-u) \\ &+ \text{terms obtained by cyclic interchange of } X_1, X_2, X_3 \\ &+ m_{(3)} \prod_{j=1}^{3} M(X_j \,|\, x;t-u)]\chi(dx \,|\, y;u)e^{\alpha(m-1)u} du\,. \end{aligned}$$

From (2.11), we have for n ancestors

$$(4.10) \qquad M(X \,|\, y^{(n)};t) = \sum_{j=1}^{n} M(X \,|\, y_j;t),$$

$$(4.11) \quad \begin{aligned} M_2(X_1 \times X_2 \,|\, y^{(n)};t) &= \sum_{j=1}^{n} M_2(X_1 \times X_2 \,|\, y_j;t) \\ &+ \sum_{j_1,j_2=1(j_1 \neq j_2)}^{n} M(X_1 \,|\, y_{j_1};t)M(X_2 \,|\, y_{j_2};t). \end{aligned}$$

In the case of the simple birth and death process,

$$(4.12) \qquad M(t) = e^{(\lambda-\mu)t}, \quad \alpha(m-1) = \lambda - \mu,$$

and the formulas given by Bailey ((1968), Equations (18), (48), and (55)) for the expected sizes of colonies follow from (4.10), (4.6), (2.16), (2.18) and the corresponding three-dimensional formula. The variances and covariances of colony sizes are given by

$$(4.13) \quad \begin{aligned} &\mathrm{cov}[N(x_1 \,|\, y^{(n)};t)\, N(x_2 \,|\, y^{(n)};t)] \\ &= \sum_{j=1}^{n} [M_2(x_1, x_2 \,|\, y_j;t) - M(x_1 \,|\, y_j;t)M(x_2 \,|\, y_j;t)], \end{aligned}$$

which in the one-dimensional case reduces to

$$e^{(\lambda-\mu-\nu)t} = \sum_{j=1}^{n} [\delta(x_1 | x_2) I_{x_1-y_j}(\nu t) - e^{(\lambda-\mu-\nu)t} I_{x_1-y_j}(\nu t) I_{x_2-y_j}(\nu t)$$
(4.14)
$$+ 2\lambda \int_0^t e^{(\lambda-\mu-\nu)u} \sum_{x=-\infty}^{\infty} I_{x_1-x}(\nu u) I_{x_2-x}(\nu u) I_{x-y_j}(\nu(t-u)) du].$$

This is the required coefficient in Equation (31) of Bailey's paper.

5. Mean square convergence when $m > 1$

If $m > 1$, $m_{(2)} < \infty$, and there is a single ancestor, the random variable

(5.1) $$W(t) = e^{-\alpha(m-1)t} N(t)$$

converges in mean square as $t \to \infty$ (in fact, with probability one) to a random variable W, whose c.f. $L(\theta)$ is the unique solution of the integral equation

(5.2) $$L(\theta) = \alpha \int_0^\infty f(L(\theta e^{-\alpha(m-1)u})) e^{-\alpha u} du,$$

(Harris (1963) Sections V.11 and VI.20). The distribution of W has a mass at the origin equal to p, the probability of extinction, which is the smallest non-negative root of the equation

(5.3) $$p = f(p).$$

For $W > 0$, the distribution is continuous. In the case of the simple birth and death process with $\lambda > \mu$,

(5.4) $$p = \frac{\mu}{\lambda},$$

and for $W > 0$ the distribution has the density

(5.5) $$q^2 e^{-qW}, \qquad (q = 1 - p).$$

An attempt has been made in Davis ((1965), (1967a)) to adapt this asymptotic theory for branching-diffusion processes with no absorbing boundaries. Sufficient conditions are given for the m.s. convergence of the random functional $N(\xi | y; t)$. To obtain satisfactory results, it may be necessary to apply a time-dependent transformation ω_t to the location variable x. Writing

(5.6) $$\xi_t(x) \equiv \xi(\omega_t x)$$

it is found in general that the limiting behaviour of

(5.7) $$N(\xi_t | y; t) = \sum_{j=1}^{N(t)} \xi_t(x_j)$$

depends upon that of its expectation

$$(5.8)\qquad \begin{aligned} \mathscr{E}N(\xi_t \mid y;t) &= \int_{\mathscr{X}} \xi_t(x) M(dx \mid y;t) \\ &= M(t) \int_{\mathscr{X}} \xi_t(x)\chi(dx \mid y;t). \end{aligned}$$

Suppose that for a given ξ and ω_t a function $\Phi(\xi \mid t)$ can be found such that the following limits exist.

$$(5.9)\qquad \lim_{t\to\infty} \Phi(\xi \mid t) \int_{\mathscr{X}} \xi_{t+u}(x)\chi(dx \mid y;t) = J(\xi \mid y;u), \quad (u \geqq 0),$$

$$(5.10)\qquad \lim_{t\to\infty} \Phi(\xi \mid t+u)/\Phi(\xi \mid t) = \Psi(\xi \mid u) = O(e^{\kappa(\xi)u}).$$

Then for a wide range of ξ with these properties, such that

$$(5.11)\qquad 2\kappa(\xi) < \alpha(m-1),$$

the random variable

$$(5.12)\qquad W_t(\xi \mid y;u) = e^{-\alpha(m-1)t}\Phi(\xi \mid t)N(\xi_{t+u} \mid y;t)$$

converges in m.s. as $t \to \infty$ to a limiting variable $W(\xi \mid y;u)$ for each $u \geqq 0$. The c.f. $L(\theta;\xi \mid y;u)$ of $W(\xi \mid y;u)$ is a solution of the integral equation

$$(5.13)\ L(\theta;\xi|y;u) = \alpha\int_0^\infty \int_{\mathscr{X}} f(L(\theta e^{-\alpha(m-1)v}\Psi(\xi \mid v);\xi \mid x;u+v))\chi(dx|y;v)e^{-\alpha v}\,dv.$$

If $J(\xi \mid y;u) \equiv J(\xi)$ is independent of y and u, then

$$(5.14)\qquad W(\xi \mid y;u) \equiv J(\xi)\cdot W,$$

where W is the m.s. limit of $W(t)$ defined in (5.1). If $J(\xi \mid y;u)$ is independent of u, then so also is $W(\xi \mid y;u)$.

Recurrence relations for the moments of $W(\xi \mid y;u)$ may be derived from those of the moment distributions (Section 4). Writing

$$(5.15)\qquad M_r^{(W)}(\xi \mid y;u) = \mathscr{E}[W(\xi \mid y;u)]^r,$$

we have in particular

$$(5.16)\qquad M_1^{(W)}(\xi \mid y;u) = J(\xi \mid y;u)$$

$$(5.17)\quad M_2^{(W)}(\xi \mid y;u) = \alpha m_{(2)}\int_0^\infty \int_{\mathscr{X}} e^{-\alpha(m-1)v}\Psi^2(\xi|v)J^2(\xi|x;u+v)\chi(dx|y;v)dv.$$

$$(5.18)\qquad \begin{aligned} M_3^{(W)}(\xi \mid y;u) = \alpha\int_0^\infty \int_{\mathscr{X}} & e^{-2\alpha(m-1)v}\Psi^3(\xi \mid v)[3m_{(2)}J(\xi \mid x;u+v) \\ & \times M_2^{(W)}(\xi \mid x;u+v) + m_{(3)}J^3(\xi|x;u+v)]\chi(dx|y;v)dv. \end{aligned}$$

$$M_4^{(W)}(\xi \mid y; u) = \alpha \int_0^\infty \int_{\mathscr{X}} e^{-3\alpha(m-1)v} \Psi^4(\xi \mid v) [4m_{(2)} J(\xi \mid x; u+v)$$

(5.19)
$$\times M_3^{(W)}(\xi \mid x; u+v) + 3m_{(2)}(M_2^{(W)}(\xi \mid x; u+v))^2$$
$$+ 6m_{(3)} J^2(\xi \mid x; u+v) M_2^{(W)}(\xi \mid x; u+v)$$
$$+ m_{(4)} J^4(\xi \mid x; u+v)] \chi(dx \mid y; v) dv.$$

6. Applications to Bailey's model when $\lambda > \mu$.

Taking

(6.1)
$$\xi(x) = \begin{cases} 1 & \text{if } x = x_1 \\ 0 & \text{otherwise} \end{cases}$$

we may consider the m.s. convergence of the colony size at $x = x_1$ in Bailey's one-dimensional model. Since

(6.2)
$$\sum_{x=-\infty}^{\infty} \xi(x)\chi(x \mid y; t) = e^{-vt} I_{x_1 - y}(vt) \sim (2\pi vt)^{-\frac{1}{2}} \quad \text{as} \quad t \to \infty$$

(cf. Bailey (1968), Equation (21)) we may take ω_t to be the identity operator, and

(6.3)
$$\Phi(\xi \mid t) = (2\pi vt)^{\frac{1}{2}}, \; J(\xi \mid y; u) \equiv 1, \; \Psi(\xi \mid u) \equiv 1.$$

Since J is independent of y and u, it follows that

(6.4)
$$e^{-(\lambda-\mu)t}(2\pi vt)^{\frac{1}{2}} N(x_1 \mid y_1; t) \underset{\text{m.s.}}{\to} W_1,$$

the limiting random variable for the simple birth and death process. For n ancestors, the limit would be $W^{(n)} = \sum_{i=1}^{n} W_i$, the W_i being independent.

More information on the asymptotic spatial distribution may be obtained by applying the transformation

(6.5)
$$\omega_t x = (vt)^{-\frac{1}{2}} x$$

to the locations of the individuals, considered as points on the complete real line, R. If we define the probability distribution function on R as

(6.6)
$$F(\zeta \mid y; t) = \sum_{x \leq \zeta} \chi(x \mid y; t), \qquad (\zeta \in R),$$

then the cf. of $F(\zeta(v(t+u))^{\frac{1}{2}} \mid y; t)$ is

(6.7)
$$\exp\{i\theta y/(v(t+u))^{\frac{1}{2}} - vt[1 - \cos(\theta/(v(t+u))^{\frac{1}{2}})]\}$$
$$\sim e^{-\frac{1}{2}\theta^2}\{1 + (vt)^{-\frac{1}{2}} i\theta y + (vt)^{-1}[\tfrac{1}{2}(i\theta)^2(y^2 - vu) + \tfrac{1}{24}\theta^4] + \cdots\}.$$

Hence $F(\zeta(v(t+u))^{\frac{1}{2}} \mid y; t)$ tends to the unit normal distribution as $t \to \infty$, and for suitable $\xi(\zeta)$ on R,

$$\sum_{x=-\infty}^{\infty} \xi_{t+u}(x)\chi(x\,|\,y;t) \sim \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\zeta^2}\xi(\zeta) \tag{6.8}$$
$$\times \{1 + (\nu t)^{-\frac{1}{2}}H_1(\zeta)y + (\nu t)^{-1}[\tfrac{1}{2}H_2(\zeta)(y^2 - \nu u) + \tfrac{1}{24}H_4(\zeta)] + \cdots\}d\zeta,$$

where $H_r(\zeta)$ is Hermite's polynomial of order r. This expansion determines the asymptotic behaviour of the L.H.S., and hence of $W_t(\xi\,|\,y;u)$. Since

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\zeta^2} H_r(\zeta)H_s(\zeta)d\zeta = r!\delta_{rs}, \tag{6.9}$$

we obtain the following table.

TABLE 1

$\xi(x)$	$\Phi(\xi\,|\,t)$	$J(\xi\,|\,y;u)$
1	1	1
$H_1(x)$	$(\nu t)^{\frac{1}{2}}$	y
$H_2(x)$	νt	$y^2 - \nu u$
$H_3(x)$	$(\nu t)^{\frac{3}{2}}$	$y^3 - 3\nu u y$
$H_4(x)$	νt	1
$H_5(x)$	$(\nu t)^{\frac{3}{2}}$	$5y$
$H_6(x)$	$(\nu t)^2$	$15(y^2 - \nu u) + 1$
$H_8(x)$	$(\nu t)^2$	35

Examples. (i) $\xi(x) = \delta(X\,|\,x)$, where X is any subset of R, the real line:

$$e^{-(\lambda-\mu)t}N(\omega_t^{-1}X\,|\,y;t) \underset{\text{m.s.}}{\rightarrow} \left(\frac{1}{\sqrt{2\tau}} \int_X e^{-\frac{1}{2}\eta^2}d\eta\right)W, \tag{6.10}$$

where $\omega_t^{-1}X$ is the set obtained by multiplying the coordinate of each point of X by $(\nu t)^{\frac{1}{2}}$. In the case of n ancestors, W is to be replaced by $W^{(n)}$.

(ii) $\xi(x) = x = H_1(x)$. Then from Table 1,

$$\Phi(H_1\,|\,t) = (\nu t)^{\frac{1}{2}},\ J(H_1\,|\,y;u) = y. \tag{6.11}$$

Hence

$$W_t(H_1\,|\,y;0) = e^{-(\lambda-\mu)t}\sum_{j=1}^{N(t)} x_j \underset{\text{m.s.}}{\rightarrow} W(H_1\,|\,y), \tag{6.12}$$

say. From (5.16) and (5.17), this limiting random variable has expected value y and variance

$$\frac{2\lambda\nu}{(\lambda-\mu)^2} + \frac{(\lambda+\mu)y^2}{(\lambda-\mu)}. \tag{6.13}$$

For n ancestors, the corresponding limit is

$$W(H_1 \mid y^{(n)}) = \sum_{j=1}^{n} W(H_1 \mid y_j), \tag{6.14}$$

a sum of mutually independent random variables.

(iii) $\xi(x) = H_2(x) = x^2 - 1$. Then with Φ and J obtained from Table 1,

$$\begin{aligned} W_t(H_2 \mid y;0) &= e^{-(\lambda-\mu)t}\left\{\sum_{j=1}^{N(t)} x_j^2 - \nu t N(t)\right\} \\ &\underset{\text{m.s.}}{\rightarrow} W(H_2 \mid y;0), \end{aligned} \tag{6.15}$$

where

$$\begin{aligned} M_1^{(W)}(H_2 \mid y;0) &= y^2 \\ \operatorname{Var} W(H_2 \mid y;0) &= \frac{8\lambda\nu^2}{(\lambda-\mu)^3} + \frac{2\lambda\nu(4y^2+1)}{(\lambda-\mu)^2} + \frac{(\lambda+\mu)y^4}{(\lambda-\mu)}. \end{aligned} \tag{6.16}$$

The integral equations of form (5.13) for the c.f.'s of $W(H_1 \mid y)$ and $W(H_2 \mid y;u)$ are intractable, but presumably their distributions will have mass μ/λ at the origin and will otherwise be continuous. In the case of $W(H_1 \mid y)$ at least, it would follow as in Davis ((1967a), page 291) that the c.f. is non-analytic in the neighbourhood of the origin.

7. Conditional moment distributions

The results of the previous section suggest an investigation of such natural measures of the location and dispersion of the population as the *mean position*,

$$(N(t))^{-1} \sum_{j=1}^{N(t)} x_j = (N(t))^{-1} N(H_1 \mid y^{(n)};t), \tag{7.1}$$

and the variance of the spatial distribution about this mean, conditional upon the population not becoming extinct. If ξ is such that $W_t(\xi \mid y;u)$ converges in mean square, then (Davis (1967b), Theorem 1)

$$Y(\xi \mid y^{(n)};t) = \Phi(\xi \mid t) \left.\frac{N(\xi_t \mid y^{(n)};t)}{N(t)}\right|_{N(t)\not\to 0} \tag{7.2}$$

converges in probability (i.p.) to

$$Y(\xi \mid y^{(n)}) = \left.\frac{W(\xi \mid y^{(n)};0)}{W^{(n)}}\right|_{W^{(n)}>0}. \tag{7.3}$$

If $J(\xi \mid y; u) = J(\xi)$ is independent of y and u, then

$$Y(\xi \mid y^{(n)}) = J(\xi), \tag{7.4}$$

a non-random limit.

Considering Example (i) of the previous section, with $\lambda > \mu$, we have for any Borel subset X of the real line

$$\frac{N(\omega_t^{-1}X \mid y^{(n)}; t)}{N(t)} \bigg|_{N(t) \not\to 0} \underset{\text{i.p.}}{\to} \frac{1}{\sqrt{2\pi}} \int_X e^{-\frac{1}{2}\eta^2} d\eta, \tag{7.5}$$

where the numerator on the left-hand side denotes the number of individuals in the set obtained by multiplying the coordinate of each point of X by $(vt)^{\frac{1}{2}}$. Similarly, from Equation (6.4), we have for the colony size at x_1:

$$(2\pi vt)^{\frac{1}{2}} \frac{N(x_1 \mid y^{(n)}; t)}{N(t)} \bigg|_{N(t) \not\to 0} \underset{\text{i.p.}}{\to} 1. \tag{7.6}$$

In the case of the *mean position*,

$$\begin{aligned} \bar{x}(t) &= \frac{1}{N(t)} \sum_{j=1}^{N(t)} x_j \bigg|_{N(t) \not\to 0} \\ &\underset{\text{i.p.}}{\to} Y(H_1 \mid y^{(n)}) = \bar{x}(\infty), \text{ say.} \end{aligned} \tag{7.7}$$

From Example (iii), the *spatial dispersion*

$$s^2(t) = \frac{1}{N(t)} \sum_{j=1}^{N(t)} (x_j - \bar{x}(t))^2 \bigg|_{N(t) \not\to 0} \tag{7.8}$$

has the property

$$s^2(t) - vt \underset{\text{i.p.}}{\to} Y(H_2 \mid y^{(n)}) - (\bar{x}(\infty))^2. \tag{7.9}$$

Thus in Bailey's one-dimensional model, the mean position tends to a limiting distribution, whereas the spatial dispersion grows more and more diffuse. These results are in contrast to those of Adke and Moyal (1963), who considered the asymptotic distributions of these quantities conditional upon a fixed number of survivors in the case of Brownian motion along the real line. This would be due in part to the fact that, when $\lambda > \mu$, $N(t)$ converges to zero with probability $p = \mu/\lambda$, or else grows without limit.

The distributions of the $Y(\xi \mid y^{(n)})$ are again intractable. However, methods of evaluating their moments may be sought through the moment distributions conditional upon given population sizes. These are obtained by differentiation of the Γ_r, defined in (3.1).

Thus

$$M_k^{(r)}(X_1 \times \cdots \times X_k \mid y;t) = \mathscr{E}\left\{\prod_{j=1}^{k} N(X_j \mid y;t) \mid N(t) = r\right\}$$

(7.10)

$$= \left\{\frac{\partial^k}{i^k \partial\theta_1 \cdots \partial\theta_k}\Gamma_r\left(\sum_{j=1}^{k}\theta_j\delta(X_j \mid \cdot) \mid y;t\right)\right\}_{\theta_1=\cdots=\theta_k=0}.$$

Let

(7.11) $$M_k(\cdot,\theta \mid y;t) = \sum_{r=0}^{\infty} e^{ir\theta} M_k^{(r)}(\cdot \mid y;t), \qquad (k = 1,2,\cdots).$$

The recurrence relations for these generating functions may be obtained from (2.13), taking

(7.12) $$\xi(x) = \theta + \sum_{j=1}^{k}\theta_j\delta(X_j \mid x).$$

For the general Markov branching-diffusion process, we have for $k = 1$,

$$M(X,\theta \mid y;t) = e^{-\alpha t + i\theta}\chi(X \mid y;t)$$

(7.13)

$$+ \alpha\int_0^t\int_{\mathscr{X}} f'(\Gamma(\theta;t-u))M(X,\theta \mid y;t-u)\chi(dx \mid y;u)e^{-\alpha u}du.$$

One may verify by direct substitution that the required solution is

(7.14) $$M(X,\theta \mid y;t) = \frac{\partial\Gamma(\theta;t)}{i\partial\theta}\chi(X \mid y;t),$$

so that

(7.15) $$\mathscr{E}\{N(X \mid y;t) \mid N(t) = r\} = rp_r(t)\chi(X \mid y;t), \qquad (r = 0,1,2,\cdots).$$

Similarly

$$M_2(X_1 \times X_2,\theta \mid y;t) = \frac{\partial\Gamma(\theta;t)}{i\partial\theta}\{\chi(X_1 \cap X_2 \mid y;t) + \alpha\int_0^t\int_{\mathscr{X}} f''(\Gamma(\theta;t-u))$$

(7.16)

$$\times \frac{\partial\Gamma(\theta;t-u)}{i\partial\theta}\chi(X_1 \mid x;t-u)\,\chi(X_2 \mid x;t-u)\chi(dx \mid y;u)du\}.$$

A general recurrence formula for higher M_k is given in Davis ((1967b), Equation (3.25)). When $\theta = 0$ these results reduce to the formulae for the moment distributions $M_k(\cdot \mid y;t)$ of Section 4. For n ancestors, the conditional moment distributions are obtained in terms of those for a single ancestor using the multiplicative property (2.11) of the c.fl. In particular:

$$(7.17)\quad M(X,\theta\,|\,y^{(n)};t) = \frac{\partial}{i\partial\theta}(\Gamma(\theta;t))^n \cdot \frac{1}{n}\sum_{j=1}^{n}\chi(X\,|\,y_j;t)$$

$$M_2(X_1 \times X_2,\theta\,|\,y^{(n)};t) = (\Gamma(\theta;t))^{n-1}\sum_{j=1}^{n} M_2(X_1 \times X_2,\theta\,|\,y_j;t)$$

$$(7.18)\qquad + (\Gamma(\theta;t))^{n-2}\sum_{j_1\neq j_2} M(X_1,\theta\,|\,y_{j_1};t)M(X_2,\theta\,|\,y_{j_2};t).$$

Thus

$$(7.19)\quad E\{N(X\,|\,y^{(n)};t)\,|\,N(t)=r\} = rp_r^{(n)}(t).\frac{1}{n}\sum_{j=1}^{n}\chi(X\,|\,y_j;t),$$

where $p_r^{(n)}(t)$ denotes the probability of r survivors at time t conditional upon n ancestors.

8. Moments of $Y(\xi\,|\,y^{(n)})$ for the simple birth and death process

Using the Markov property of the process, it may be shown that (Davis (1967b), page 7)

$$M_r^{(Y)}(\xi_1,\cdots,\xi_r\,|\,y^{(n)};t) = \mathscr{E}\prod_{j=1}^{r} Y(\xi_j\,|\,y^{(n)};t)$$

$$(8.1)\qquad = (1-p^n)^{-1}\left(\prod_{j=1}^{r}\Phi(\xi_j\,|\,t)\right)$$

$$\times \sum_{s=1}^{\infty} s^{-r}(1-p^s)\int_{\mathscr{X}^r}\prod_{j=1}^{r}\xi_{j,t}(x_j)M_r^{(s)}(dx_1\times\cdots\times dx_r\,|\,y^{(n)};t),$$

where p is the extinction probability for a single ancestor. In particular, taking $r=1$,

$$(8.2)\quad M^{(Y)}(\xi\,|\,y^{(n)};t) = \frac{1}{n}\sum_{j=1}^{n}\Phi(\xi\,|\,t)\int_{\mathscr{X}}\xi_t(x)\chi(dx\,|\,y_j;t),$$

whence, letting $t\to\infty$, the expected value of $Y(\xi\,|\,y^{(n)})$ is seen to be

$$(8.3)\quad M^{(Y)}(\xi\,|\,y^{(n)}) = \frac{1}{n}\sum_{j=1}^{n} J(\xi\,|\,y_j;0).$$

Let

$$(8.4)\quad M_r^{(Y)}(\xi_1,\cdots,\xi_r\,|\,y^{(n)}) = \mathscr{E}\prod_{j=1}^{r} Y(\xi_j\,|\,y^{(n)}).$$

Considering the case $r=2$,

(8.5)
$$M_2^{(Y)}(\xi_1,\xi_2\,|\,y^{(n)}) = \Phi(\xi_1\,|\,t)\,\Phi(\xi_2\,|\,t)\sum_{j=1}^{n}\Big\{\rho_1(t)\int_{\mathscr{X}}\xi_{1,t}(x)\xi_{2,t}(x)\chi(dx\,|\,y_j;t)$$
$$+\int_0^t\int_{\mathscr{X}}\rho(u;t)\Big(\prod_{h=1}^{2}\int\xi_{h,t}(x_1)\chi(dx_1|x_2;t-u)\Big)\chi(dx_2\,|\,y_j;u)du\Big\}$$
$$+\ \rho_2(t)\sum_{j_1\neq j_2}\prod_{h=1}^{2}\Big(\Phi(\xi_h\,|\,t)\int_{\mathscr{X}}\xi_{h,t}(x)\chi(dx\,|\,y_{j_h};t)\Big),$$

where

$$\rho_1(t) = (1-p^n)^{-1}\sum_{r=1}^{\infty} r^{-1}(1-p^r)p_r^{(n)}(t),$$

$$\rho(u;t) = (1-p^n)^{-1}\sum_{r=1}^{\infty} r^{-2}(1-p^r)\times \text{coefficient of } e^{ir\theta}$$

(8.6)
$$\text{in } (\Gamma(\theta;t))^{n-1}\,\frac{\partial\Gamma(\theta;t)}{i\partial\theta}\,\frac{\partial\Gamma(\theta;t-u)}{i\partial\theta}f''(\Gamma(\theta;t-u))$$

$$\rho_2(t) = (1-p^n)^{-1}\sum_{r=1}^{\infty} r^{-2}(1-p^r)\times \text{coefficient of } e^{ir\theta}$$
$$\text{in } (\Gamma(\theta;t))^{n-2}\left(\frac{\partial\Gamma(\theta;t)}{i\partial\theta}\right)^2.$$

The asymptotic behaviour of (8.5) as $t\to\infty$ depends upon the limiting properties of these functions. In general, it appears difficult to make further progress unless the individual probabilities $p_r(t)$ are known explicitly, and are of fairly simple functional form. The case of $(k+1)$-fold splitting for a single ancestor has been considered in Davis ((1967b), Section 5). In the case of the simple birth and death process and a single ancestor, the moments of $Y(\xi\,|\,y^{(n)};t)$ are found from (8.1) and the conditional moment distributions of Section 7 are expressible in terms of the functions

(8.7)
$$\rho_{r,k}(u_1,\cdots,u_h;t) = (1-p^n)^{-1}\sum_{j=1}^{\infty} j^{-r}(1-p^j)\times\text{coefficient of } e^{ij\theta} \text{ in}$$
$$\left\{(\Gamma(\theta;t))^{n-k}\left(\frac{\partial\Gamma(\theta;t)}{i\partial\theta}\right)^k\prod_{l=1}^{h}\frac{\partial\Gamma(\theta;t-u_l)}{i\partial\theta}\right\},$$
$$(0<u_1<\cdots<u_h<t).$$

The dependence of ρ upon n will be omitted from the notation for convenience in writing. Let

(8.8)
$$a_0 = \frac{\lambda}{\mu}p_0(t),\quad a_j = \frac{\lambda}{\mu}p_0(t-u_j),\qquad (j=1,\cdots,h).$$

Then

$$0 < a_h < a_{h-1} < \cdots < a_0 < 1, \tag{8.9}$$

and from the Formula (2.14) for the c.f. $\Gamma(\theta; t)$ for the simple birth and death process, $\rho_{r,k}$ may be expanded in a multiple power series as

$$\begin{aligned}\rho_{r,k}(u_1, \cdots, u_h; t) = &\left(\prod_{j=1}^{h} p_1(t-u_j)\right) \cdot \sum_{s=k}^{n} \binom{n-k}{s-k} \left(\frac{\mu}{\lambda} a_0\right)^{n-s} (p_1(t))^s \\ &\times \sum_{l_0, \cdots, l_{h-1}=0}^{\infty} \binom{s+k+l_0-1}{l_0} (l_1+1) \cdots (l_{h-1}+1)(a_0-a_h)^{l_0} \cdots (a_{h-1}-a_h)^{l_{h-1}} \\ &\times \sum_{j=0}^{\infty} \binom{L+s+2h+k+j-1}{j} \frac{a_h^j(1-p^{L+s+h+j})}{(L+s+h+j)^r},\end{aligned} \tag{8.10}$$

where L denotes $l_0 + \cdots + l_{h-1}$. The limiting behaviour of this series may be discussed by the methods of the author's earlier paper ((1967b), Theorem 2). It is possible to show that, if $h < r - k$, then $\rho_{r,k}(u_1, \cdots, u_h; t)$ converges to zero exponentially fast. On the other hand, if $h = r-k$,

$$\lim_{t \to \infty} \rho_{r,k}(u_1, \cdots, u_h; t) = \rho_{r,k}(u_1, \cdots, u_h) \tag{8.11}$$

exists, and may be expressed in terms of the Lauricella F_D-type hypergeometric functions

$$F_D(a; b_1, \cdots, b_h; c; z_1, \cdots, z_h) = \sum_{j_1, \dots, j_h=0}^{\infty} \frac{(a)_{j_1+\cdots+j_h}(b_1)_{j_1} \dots (b_h)_{j_h}}{(c)_{j_1+\cdots+j_h} j_1! \cdots j_h!} z_1^{j_1} \cdots z_h^{j_h}, \tag{8.12}$$

where

$$(b)_j = b(b+1) \cdots (b+j-1). \tag{8.13}$$

(Appell and Kampé de Feriet (1926), première partie, Chapter VII, Section XXXVII). When $h = 2$, F_D reduces to Appell's first type of hypergeometric function of two variables; for $h = 1$ it is the Gaussian hypergeometric function ${}_2F_1$.

Writing

$$\sigma_j = \lim_{t \to \infty} \frac{a_j - a_h}{1 - a_h} = 1 - \exp\{-(\lambda-\mu)(u_h - u_j)\}, \; (j = 0, \cdots, h-1), \tag{8.14}$$

$$q = 1 - p = 1 - \frac{\mu}{\lambda}, \quad Q = 1 - p^n, \tag{8.15}$$

(taking $u_0 \equiv 0$), the limit function in (8.11) is given by

(8.16) $$\rho_{r,k}(u_1,\cdots,u_h) = Q^{-1}q^h \prod_{j=1}^{h-1}(1-\sigma_j) \times$$

$$\sum_{s=k}^{n}\binom{n-k}{s-k}\frac{p^{n-s}q^s(1-\sigma_0)^s}{(s+h)_r}F_D(s+h;s+k,2,\cdots 2;s+h+r;\sigma_0,\sigma_1,\cdots\sigma_{h-1}).$$

When $h = 0$

(8.17) $$\rho_{r,r} = Q^{-1}\sum_{s=r}^{n}\binom{n-r}{s-r}\frac{p^{n-s}q^s}{(s)_r}.$$

The transformation of $0<u_1<\cdots<u_h<\infty$ onto $0<\sigma_{h-1}<\cdots<\sigma_1<\sigma_0<1$ defined by (8.14) has the jacobian

(8.18) $$\frac{\partial(u_1,\cdots,u_h)}{\partial(\sigma_0,\cdots,\sigma_{h-1})} = \left\{(\lambda-\mu)^h\prod_{j=0}^{h-1}(1-\sigma_j)\right\}^{-1}.$$

For arbitrary χ, the simple birth and death process with $\lambda>\mu$, and functions ξ_i such that $\Psi(\xi_i|t)\equiv 1$, the early moments of $Y(\xi|y^{(n)})$ are

(8.19) $$M_2^{(Y)}(\xi_1,\xi_2|y^{(n)}) = 2\lambda\sum_{j=1}^{n}\int_0^\infty\int_{\mathscr{X}}\rho_{2,1}(u)J(\xi_1|x;u)J(\xi_2|x;u)\chi(dx|y_j;u)du + \rho_{2,2}\sum_{j_1\neq j_2}J(\xi_1|y_{j_1};0)J(\xi_2|y_{j_2};0);$$

(8.20) $$M_3^{(Y)}(\xi_1,\xi_2,\xi_3|y^{(n)}) = 4\lambda^2\sum_{j=1}^{n}\int_0^\infty\int_u^\infty\rho_{3,1}(u,v)\int_{\mathscr{X}^2}[J(\xi_1|x;u)J(\xi_2|y;v)$$
$$J(\xi_3|y;v) + (\text{cyclic interchange of }\xi_1,\xi_2,\xi_3)]\chi(dy|x;v-u)\chi(dx|y_j;u)dudv$$
$$+ 2\lambda\sum_{j_1\neq j_2}J(\xi_1|y_{j_1};0)\int_0^\infty\int_{\mathscr{X}}\rho_{3,2}(u)J(\xi_2|x;u)J(\xi_3|x;u)\chi(dx|y_{j_2};u)du$$
$$+ (\text{cyclic interchange of }\xi_1,\xi_2,\xi_3) + \rho_{3,3}\sum_{j_1\neq j_2\neq j_3}\prod_{l=3}^{3}J(\xi_{j_l}|y_{j_l};0);$$

(8.21) $$M_4^{(Y)}(\xi|y^{(n)}) = 48\lambda^3\sum_{j=1}^{n}\int_0^\infty\int_u^\infty\int_v^\infty\rho_{4,1}(u,v,w)$$
$$\int_{\mathscr{X}^3}[2J(\xi|x;u)J(\xi|y;v)J^2(\xi|z;w)\chi(dz|y;w-v)\chi(dy|x;v-u)$$
$$+ J^2(\xi|y;v)J^2(\xi|z;w)\chi(dy|x;v-u)\chi(dz|x;w-u)]\chi(dx|y_j;u)\ du\,dv\,dw$$

$$
\begin{aligned}
&+ 24\lambda^2 \sum_{j_1 \neq j_2} \int_0^\infty \int_u^\infty \rho_{4,2}(u,v) \int_{\mathscr{X}^2} [2J(\xi \mid y_{j_1};0) \\
&\times J(\xi \mid x;u) J^2(\xi \mid y;v) \chi(dy \mid x; v-u) \chi(dx \mid y_{j_2};u) \\
&+ J^2(\xi \mid x;u) J^2(\xi \mid y;v) \chi(dx \mid y_{j_1};u) \chi(dy \mid y_{j_2};v)] du dv \\
&+ 12\lambda \sum_{j_1 \neq j_2 \neq j_3} J(\xi \mid y_{j_1};0) J(\xi \mid y_{j_2};0) \int_0^\infty \rho_{4,3}(u) \\
&\times \int_{\mathscr{X}} J^2(\xi \mid x;u) \chi(dx \mid y_{j_3};u) du + \rho_{4,4} \sum_{j_1 \neq j_2 \neq j_3 \neq j_4} \prod_{l=1}^{4} J(\xi \mid y_{j_l};0),
\end{aligned}
\tag{8.21}
$$

where if the ξ_i are equal, we write simply $M_r^{(Y)}(\xi \mid y^{(n)})$.

9. Moments of the mean position and spatial dispersion in Bailey's model

(a) *Mean position.* We have seen in (7.7) that the mean position $\bar{x}(t)$for Bailey's model converges in probability to a limiting variable $\bar{x}(\infty)$, say. From (8.3),

$$
M_1 = \mathscr{E}\bar{x}(\infty) = \frac{1}{n} \sum_{j=1}^{n} y_j = \mu_1, \text{ say,} \tag{9.1}
$$

the average initial position of the n ancestors. The higher central moments

$$
M_r = \mathscr{E}(\bar{x}(\infty) - \mu_1)^r, \qquad (r = 2,3,\cdots), \tag{9.2}
$$

may also be expressed in terms of the central moments μ_r of the initial distribution

$$
\mu_r = \frac{1}{n} \sum_{j=1}^{n} (y_j - \mu_1)^r, \qquad (r = 2,3,\cdots). \tag{9.3}
$$

Considering the variance of $\bar{x}(\infty)$, we have from (8.17)

$$
\mathscr{E}(\bar{x}(\infty))^2 = 2\lambda[vnI_1^{(2,1)} + I_0^{(2,1)} \sum_{j=1}^{n} y_j^2] + \rho_{2,2} \sum_{j_1 \neq j_2} y_{j_1} y_{j_2}, \tag{9.4}
$$

where

$$
I_0^{(2,1)} = \int_0^\infty \rho_{2,1}(u) du, \quad I_1^{(2,1)} = \int_0^\infty \rho_{2,1}(u) u du. \tag{9.5}
$$

The integrals encountered in the evaluation of these moments are similar to those encountered in the case of $(k+1)$-fold splitting and a single ancestor (Davis (1967b).) Some elementary properties of the F_D hypergeometric functions are listed in the Appendix.

The integrals (9.5) may be evaluated by means of the substitutions

$$
\sigma = 1 - e^{-(\lambda-\mu)u}. \tag{9.6}
$$

From (8.16)

$$I_0^{(2,1)} = \lambda^{-1}Q^{-1} \sum_{s=1}^{n} \binom{n-1}{s-1} \frac{p^{n-s}q^s}{(s+1)(s+2)} \times \int_0^1 (1-\sigma)^{s-1}{}_2F_1(s+1, s+1; s+3; \sigma)\, d\sigma. \tag{9.7}$$

Using Formula (A.4) of the Appendix, the integral on the right-hand side is equal to

$$s^{-1}(s+2)\int_0^1 \frac{d}{d\sigma}[-(1-\sigma)^s {}_2F_1(s+1; s; s+2; \sigma)]\, d\sigma = s^{-1}(s+2), \tag{9.8}$$

whence

$$I_0^{(2,1)} = [n(n+1)\lambda q Q]^{-1}[1-(n+1)p^n + np^{n+1}]. \tag{9.9}$$

From (8.17),

$$\rho_{2,2} = [(n-1)_3 qQ]^{-1}[(n-1)(1-p^{n+1}) - (n+1(p-p^n)]. \tag{9.10}$$

Noting that

$$n\mu_2 = (n-1)\sum_{j=1}^{n} y_j^2 - \sum_{j_1 \neq j_2} y_{j_1}y_{j_2}, \tag{9.11}$$

we obtain for the variance

$$M_2 = (qQ)^{-1}\Big\{2\left(\frac{v}{\lambda}\right)p^n \sum_{s=1}^{n} s^{-1}(p^{-s}-1) + \frac{\mu_2}{(n^2-1)}[(n-1)+(n+1)p-(n+1)(2n-1)p^n+(n-1)(2n+1)p^{n+1}\Big\} \tag{9.12}$$

For a single ancestor, the variance is independent of its initial position

$$M_2 = \frac{2v}{\lambda-\mu}. \tag{9.13}$$

For two ancestors initially at y_1 and y_2,

$$M_2 = (\lambda^2-\mu^2)^{-1}[v(\lambda+3\mu) + \tfrac{1}{12}(y_1-y_2)^2(\lambda-\mu)(\lambda+5\mu)]. \tag{9.14}$$

There would appear to be some interest in computing the third and fourth central moments of $\bar{x}(\infty)$, thus enabling a Pearson curve to be fitted to its distribution.

After some reduction, the third moment of $\bar{x}(\infty)$ can be expressed in terms of μ_3, the third central moment of the initial distribution, as follows:

$$M_3 = \mathscr{E}(\bar{x}(\infty))^3 - 3M_1M_2 - M_1^3$$

$$= \mu_3[(n^2-1)(n^2-4)q^2Q]^{-1}[2(n-1)(n-2) + 2(n^2-4)p$$

(9.15)

$$+ 2(n+1)(n+2)p^2 - (n+1)(n+2)(3n^2-6n+2)p^n$$

$$+ 2(n^2-4)(3n^2-1)p^{n+1} - (n-1)(n-2)(3n^2+6n+2)p^{n+2}], (n \geqq 3).$$

For $n = 1$ or 2, this moment vanishes identically.

The fourth central moment of $\bar{x}(\infty)$ is expressible in terms of μ_2 and μ_4 as

$$M_4 = A_1\nu^2 + \nu(B_1 + C_1\mu_2) + D_1\mu_2^2 + E_1\mu_4, \tag{9.16}$$

where

$$A_1 = 2(\lambda^2q^2Q)^{-1}\Big\{\tfrac{1}{3}\pi^2Q - 2\sum_{s=2}^{n}\binom{n}{s}p^{n-s}q^s\sum_{j=1}^{s-1}j^{-2}$$

(9.17)

$$+ 8\sum_{s=1}^{n}s^{-2}\binom{n}{s}p^{n-s}q^s - 5p^n\sum_{s=1}^{n}s^{-1}(p^{-s}-1)$$

$$+ 3(n+1)^{-1}q^{-1}[1-(n+1)p^n + np^{n+1}]\Big\};$$

$$B_1 = 4(\lambda qQ)^{-1}\Big\{p^n\sum_{s=1}^{n}s^{-1}(p^{-s}-1) - [(n+1)(n+2)q^2]^{-1}$$

(9.18)

$$\times [(n+1)-(n+2)p-\tfrac{3}{2}(n+1)(n+2)p^n+(n+2)(3n+1)p^{n+1}-\tfrac{1}{2}n(3n+5)p^{n+2}]\Big\}.$$

$$C_1 = 4[(n^2-1)\lambda q^2Q]^{-1}\Big\{[4n+(n-1)(5n+3)q]p^n\sum_{s=1}^{n}s^{-1}(p^{-s}-1)$$

$$+ n[2(n+1)(n+2)q]^{-1}[-6(n+1)-2(n+2)(5n+1)q+5(n+1)^2(n+2)p^n$$

(9.19)

$$- 2(n+2)(5n^2+2n+1)p^{n+1} + n(n-1)(5n+9)p^{n+2}]\Big\}, \quad (n \geqq 2);$$

$$D_1 = 3n^2[(n-3)_7q^2Q]^{-1}\{(n-3)_3 + (n-2)(n^2-9)p - (n+2)(n^2-9)p^2$$

$$-(n+1)_3p^3+\tfrac{1}{3}(n+1)_3(2n^2-6n+3)p^n-(n+2)(n^2-9)(2n^2-2n-1)p^{n+1}$$

(9.20)

$$+(n-2)(n^2-9)(2n^2+2n-1)p^{n+2}-\tfrac{1}{3}(n-3)_3(2n^2+6n+3)p^{n+3}\}, (n \geqq 4);$$

$$E_1 = 6n[(n-3)_7q^3Q]^{-1}\{(n-3)_3+(n-2)(n^2-9)p+(n+2)(n^2-9)p^2$$

$$+(n+1)_3p^3-\tfrac{1}{3}(n+1)_3(2n-3)(n^2-3n+1)p^n+(n+2)(n^2-9)(2n-1)$$

(9.21)

$$\times (n^2-n-1)p^{n+1} - (n-2)(n^2-9)(2n+1)(n^2+n-1)p^{n+2}$$

$$+ \tfrac{1}{3}(n-3)_3(2n^3+9n^2+11n+3)p^{n+3}\}, \quad (n \geqq 4).$$

For $n = 1, 2, 3$, the required values may be obtained by taking formal limits. Thus, for $n = 1$, (9.22) $$M_4 = \frac{\nu^2}{(\lambda-\mu)^2}\left(\tfrac{2}{3}\pi^2+9\right) + \tfrac{4}{3}\,\frac{\nu}{\lambda-\mu},$$

while for two ancestors initially at y_1 and y_2:

(9.23) $$\begin{aligned} M_4 = (\lambda^2-\mu^2)^{-1}\Big\{&\frac{\nu^2}{(\lambda-\mu)}\left[\tfrac{2}{3}\pi^2(\lambda+\mu)-3(\lambda-7\mu)\right] \\ &+\tfrac{1}{3}\nu(\lambda+7\mu)+\tfrac{1}{9}\nu(5\lambda+49\mu)(y_1-y_2)^2\Big\}+\tfrac{1}{80}\,\frac{(\lambda+9\mu)}{(\lambda+\mu)}(y_1-y_2)^4. \end{aligned}$$

(b) *Spatial dispersion.* From (7.9), $s^2(t)-\nu t$ converges in probability to

(9.24) $$Y(H_2 \mid y^{(n)}) - (\bar{x}(\infty))^2 = S(\infty), \text{ say.}$$

From Table 1, (8.3) and (9.15) we obtain

(9.25) $$\begin{aligned} \mathscr{E}S(\infty) = (qQ)^{-1}\Big\{&-2\binom{\nu}{\lambda}\, p^n \sum_{s=1}^{n} s^{-1}(p^{-s}-1) \\ &+\mu_2 n(n^2-1)^{-1}[(n-1)(1-p^{n+1})-(n+1)(p-p^n)]\Big\}. \end{aligned}$$

Also,

(9.26) $$\mathscr{E}(S(\infty))^2 = M_2^{(Y)}(H_2 \mid y^{(n)}) - 2M_3^{(Y)}(H_1,H_1,H_2 \mid y^{(n)}) + \mathscr{E}(\bar{x}(\infty))^4.$$

It may be finally shown that

(9.27) $$\mathscr{E}(S(\infty))^2 = A_2\nu^2 + \nu(B_2+C_2\mu_3) + D_2\mu_2^2 + E_2\mu_4,$$

where

(9.28) $$\begin{aligned} A_2 = 2(\lambda^2 q^2 Q)^{-1}\Big\{&\tfrac{1}{3}\pi^2 Q - 2\sum_{s=2}^{n}\binom{n}{s}p^{n-s}q^s\sum_{j=1}^{s-1} j^{-2} \\ &- p^n\sum_{s=1}^{n} s^{-1}(p^{-s}-1) + 3(n+1)^{-1}q^{-1}[1-(n+1)p^n+np^{n+1}]\Big\}; \end{aligned}$$

(9.29) $$B_2 = 2[(n+1)(n+2)\lambda q^3 Q]^{-1}[n(1-p^{n+2})-(n+2)pQ];$$

(9.30) $$\begin{aligned} C_2 = 4n^2[(n-1)_4\lambda q^3 Q]^{-1}[&(n-1)-(n+2)p+\tfrac{1}{2}(n+1)(n+2)p^n \\ &-(n-1)(n+2)p^{n+1}+\tfrac{1}{2}n(n-1)p^{n+2}]; \end{aligned}$$

(9.31) $$\begin{aligned} D_2 = n^2[(n-3)_7 q^3 Q]^{-1}[&(n-3)_3(n^2+3n+3)(1-p^{n+3}) \\ &-3(n-2)(n^2-9)(n^2+n-1)(p-p^{n+2}) \\ &+3(n+2)(n^2-9)(n^2-n-1)(p^2-p^{n+1}) \\ &-(n+1)_3(n^2-3n+3)(p^3-p^n)]; \end{aligned}$$

$$(9.32)\quad \begin{aligned} E_2 &= n^2[(n-3)_7 q^3 Q]^{-1}\{(n-3)_3[(n+1)(1-p^{n+3})-(n+3)(p-p^{n+2})] \\ &- (n+1)_3[(n-3)(p^2-p^{n+1})-(n-1)(p^3-p^n)]\} \end{aligned}$$

For a single ancestor,

$$(9.33)\quad \mathscr{E}S(\infty) = -\frac{2\nu}{(\lambda-\mu)},\quad \operatorname{Var} S(\infty) = \frac{\nu^2}{(\lambda-\mu)}\left(\tfrac{2}{3}\pi^2-3\right)+\tfrac{1}{3}\frac{\nu}{(\lambda-\mu)}.$$

For two ancestors initially at y_1 and y_2

$$(9.34)\quad \mathscr{E}S(\infty) = -\frac{\nu(\lambda+3\mu)}{(\lambda^2-\mu^2)}+\tfrac{1}{6}\frac{(\lambda-\mu)}{(\lambda+\mu)}(y_1-y_2)^2;$$

$$(9.35)\quad \begin{aligned} \operatorname{Var} S(\infty) &= \frac{\nu^2}{(\lambda-\mu)^2}\left[\tfrac{2}{3}\pi^2-4+\frac{4\lambda\mu}{(\lambda+\mu)^2}\right]+\tfrac{1}{3}\frac{\nu}{(\lambda-\mu)} \\ &+ \tfrac{1}{6}\frac{\nu(3\lambda+7\mu)}{(\lambda+\mu)^2}(y_1-y_2)^2+\tfrac{1}{180}\frac{(\lambda-\mu)(\lambda+11\mu)}{(\lambda+\mu)^2}(y_1-y_2{}^4). \end{aligned}$$

Appendix: Elementary properties of the F_D hypergeometric functions

Contraction formulas

$$(A.1)\quad F_D(a;b_1,\cdots,b_n;c;z_1,\cdots,z_{h-1},0) = F_D(a;b_1,\cdots,b_{h-1};c;z_1,\cdots,z_{h-1});$$

$$(A.2)\quad F_D(a;b_1,\cdots,b_n;c;z_1,\cdots,z_{h-1},z_{h-1}) = F_D(a;b_1,\cdots,b_{h-1}+b_h;c;z_1,\cdots,z_{h-1});$$

Differentiation formulas:

$$(A.3)\quad \begin{aligned} &F_D(a;b_1,\cdots,b_h;c;z_1,\cdots,z_h) \\ &= \frac{(c-1)}{(a-1)(b-1)}\frac{\partial}{\partial z_h}F_D(a-1;b_1,\cdots,b_{h-1},b_h-1;c-1;z_1,\cdots,z_h); \end{aligned}$$

$$(A.4)\quad \begin{aligned} &(1-z_h)^{b_h-s-1}F_D(a;b_1,\cdots,b_h;c;z_1,\cdots,z_h) \\ &= \frac{(c-s)_s}{(b_h-s)_s(c-s-a)_s}\frac{\partial^s}{\partial z_h^s}(-1)^s(1-z_h)^{b_h-1}F_D(a;b_1,\cdots,b_{h-1}, \\ &\qquad b_h-s;\ c-s;z_1,\cdots,z_h). \end{aligned}$$

References

Adke, S. R. and Moyal, J. E. (1963) A birth, death and diffusion process. *J. Math. Anal. Appl.* **7**, 209–224.

Appell, P. and Kampe de Feriet, J. (1926) *Fonctions Hypergéometriques et Hypersphériques.* Gauthier Villar.

Bailey, N. T. J. (1968) Stochastic birth. death and migration processes for spatially distributed populations. *Biometrika* **55**, 189–198.

CONNER, H. E. (1966) Limiting behavior for age- and position-dependent branching processes. *J. Math. Anal. Appl.* **13,** 265–295.

DAVIS, A. W. (1965) On the theory of birth, death and diffusion processes. *J. Appl. Prob.* **2,** 293–322.

DAVIS, A. W. (1967a) Branching-diffusion processes with no absorbing boundaries. I. *J. Math. Anal. Appl.* **18,** 276–296.

DAVIS, A. W. (1967b) Branching-diffusion processes with no absorbing boundaries. II. *J. Math. Anal. Appl.* **19,** 1–25.

HARRIS, T. E. (1963) *The Theory of Branching Processes.* Springer-Verlag, Berlin.

MOYAL, J. E. (1962) The general theory of stochastic population processes. *Acta Math.* **108,** 1–31.

MOYAL, J. E. (1964) Multiplicative population processes. *J. Appl. Prob.* **1,** 267–283.

SEVAST'YANOV, B. A. (1958) Branching stochas ic processes for particles diffusing in a restricted domain with absorbing boundaries. *Theor. Probability Appl.* **3,** 111–126.

WATANABE, S. (1965) On the branching process for Brownian particles with an absorbing boundary. *J. Math. Kyoto Univ.* **4,** 385–398.

Section 2.

Exact null distributions of univariate and multivariate test statistics.

Reprinted from THE ANNALS OF MATHEMATICAL STATISTICS
Vol. 38, No. 2, April, 1967
Printed in U.S.A.

5

A COUNTER-EXAMPLE RELATING TO CERTAIN MULTIVARIATE GENERALIZATIONS OF t AND F

BY A. W. DAVIS

C.S.I.R.O., Adelaide, South Australia.

Summary. It is shown that the distributions of certain multivariate analogues of t and F are dependent on the population covariance matrix.

Suppose that $\mathbf{z} = (x_1, \cdots, x_p)'$ has the multivariate normal distribution with mean vector $\boldsymbol{\xi}$ and covariance matrix $\boldsymbol{\Sigma}$. Let $\mathbf{S}$, $\mathbf{S}^*$ be independent Wishart matrices also with covariance matrix $\boldsymbol{\Sigma}$, based on n, n^* degrees of freedom respectively. Then if $\mathbf{T}$ is a $p \times p$ matrix such that

$$\mathbf{T}\mathbf{T}' = \mathbf{S} \tag{1}$$

natural candidates for the multivariate analogues of t and F are

$$\mathbf{t} = \mathbf{T}^{-1}(\mathbf{z} - \boldsymbol{\xi}), \tag{2}$$

$$\mathbf{W} = \mathbf{T}^{-1}\mathbf{S}^*\mathbf{T}'^{-1}. \tag{3}$$

Olkin and Rubin (1964), Theorems 3.2 and 4.2, have shown that if $\mathbf{T}$ is taken to be upper or lower triangular, then $\mathbf{t}$ and $\mathbf{W}$ do in fact have distributions which independent of $\boldsymbol{\Sigma}$. However, if $\mathbf{T}$ is taken to be symmetrical and positive definite, $\mathbf{T} = \mathbf{S}^{\frac{1}{2}}$, then they remark (Section 3) that the distribution of $\mathbf{W}$ is unknown for general $\boldsymbol{\Sigma}$. It seems worthwhile to present the following example, which shows that for $\mathbf{T} = \mathbf{S}^{\frac{1}{2}}$ the distributions of $\mathbf{t}$ and $\mathbf{W}$ depend on $\boldsymbol{\Sigma}$; the contrary is occasionally asserted (see Bennett and Cornish (1964), p. 907).

Let $p = 2$, and assume that

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}. \tag{4}$$

Since $\mathbf{S}^{\frac{1}{2}}$ is positive definite, it may be written in the form

$$\mathbf{S}^{\frac{1}{2}} = \begin{bmatrix} x, & (xy)^{\frac{1}{2}}q \\ (xy)^{\frac{1}{2}}q, & y \end{bmatrix} \quad (0 < x, y < \infty, q^2 < 1). \tag{5}$$

Noting the Jacobian

$$\partial(\mathbf{S})/\partial(x, y, q) = 4(1 - q^2)(xy)^{\frac{3}{2}}(x + y), \tag{6}$$

it is readily found by transforming the Wishart distribution that $\mathbf{S}^{\frac{1}{2}}$ has the distribution

$$\begin{aligned} f(\mathbf{S}^{\frac{1}{2}})\, d\mathbf{S}^{\frac{1}{2}} &= \{(nab)^n/\pi\Gamma(n-1)\}(1 - q^2)^{n-2}(xy)^{n-\frac{3}{2}}(x + y) \\ &\quad \cdot \exp\{-\tfrac{1}{2}n[a^2x^2 + (a^2 + b^2)xyq^2 + b^2y^2]\}\, dx\, dy\, dq, \\ &\qquad (0 < x, y < \infty, q^2 < 1). \end{aligned} \tag{7}$$

Received 26 September 1966.

(See Olkin and Rubin, p. 266.)

Now

$$\mathcal{E}(\mathbf{tt}') = \mathcal{E}\mathbf{W}$$

(8) $$= \mathcal{E}(\mathbf{S}^{-\frac{1}{2}}\mathbf{\Sigma}\,\mathbf{S}^{-\frac{1}{2}})$$

$$= \begin{bmatrix} \nu_1/a^2 + \nu_3/b^2, & -\nu_4/a^2 - \nu_5/b^2 \\ -\nu_4/a^2 - \nu_5/b^2, & \nu_3/a^2 + \nu_2/b^2 \end{bmatrix},$$

where

(9) $$\nu_1 = \mathcal{E}[x^{-2}(1-q^2)^{-2}], \qquad \nu_2 = \mathcal{E}[y^{-2}(1-q^2)^{-2}], \qquad \nu_3 = \mathcal{E}[q^2/xy(1-q^2)^2],$$

$$\nu_4 = \mathcal{E}[q/x^{\frac{3}{2}}y^{\frac{1}{2}}(1-q^2)^2], \qquad \nu_5 = \mathcal{E}[q/x^{\frac{1}{2}}y^{\frac{3}{2}}(1-q^2)^2].$$

It will be sufficient to show that the matrices in (8) are not independent of a and b. In order to simplify the calculations still further, *we shall take the case* $n = 4$, and consider only the elements $\mathcal{E}(t_1^2)$, $\mathcal{E}(t_2^2)$ on the main diagonal in (8). The density (7) now takes the form:

(10) $$f(\mathbf{S}^{\frac{1}{2}})\, d\mathbf{S}^{\frac{1}{2}} = 2^7\pi^{-1}(ab)^4(1-q^2)^2(xy)^{\frac{5}{2}}(x+y)$$
$$\cdot \exp\{-2[a^2x^2 + (a^2+b^2)xyq^2 + b^2y^2]\}\, dx\, dy\, dq,$$

and we wish to evaluate ν_1, ν_2, ν_3.

For ν_1, let us first consider

$$I(q) = \int_0^\infty\int_0^\infty x^{-2}(1-q^2)^{-2}f(\mathbf{S}^{\frac{1}{2}})\, dx\, dy$$

(11) $$= 2^7\pi^{-1}(ab)^4\int_0^\infty\int_0^\infty (x^{\frac{3}{2}}y^{\frac{5}{2}} + x^{\frac{1}{2}}y^{7/2})e^{-2a^2x^2-2b^2y^2}$$
$$\cdot \sum_{k=0}^\infty k!^{-1}[-2(a^2+b^2)xyq^2]^k\, dx\, dy.$$

Making use of the formulas

$$\Gamma(\omega)\Gamma(\omega+\tfrac{1}{2}) = \pi^{\frac{1}{2}}\Gamma(2\omega)/2^{2\omega-1},$$

(12) $$\int_0^\infty \omega^\lambda e^{-a^2\omega^2}\, d\omega = \Gamma((\lambda+1)/2)/2a^{\lambda+1},$$

$$\sum_{k=0}^\infty x^k\Gamma(k+\lambda)/k! = \Gamma(\lambda)(1-x)^{-\lambda}, \qquad (|x|<1),$$

it is found that

(13) $$I(q) = (2a^3/\pi b)^{\frac{1}{2}}\sum_{k=0}^\infty k!^{-1}[-(a^2+b^2)q^2/2ab]^k$$
$$\cdot\{a\Gamma(k+\tfrac{3}{2}) + (a+b)\Gamma(k+\tfrac{5}{2})\}$$

(14) $$= (a^3/2b)^{\frac{1}{2}}\{a\phi^{-\frac{3}{2}}(q) + \tfrac{3}{2}(a+b)\phi^{-\frac{5}{2}}(q)\},$$

where

(15) $$\phi(q) = 1 + (a^2+b^2)q^2/2ab.$$

The series in (13) is convergent only for

(16) $$|q|^2 < 2ab/(a^2+b^2).$$

However, since $I(q)$ and the expression (14) are both analytic functions of the complex variable $q = u + iv$ over the region $u^2 - v^2 > -2ab/(a^2 + b^2)$ in the q plane, it follows by analytic continuation that $I(q)$ is certainly equal to (14) for all real q.

Noting that

$$\int_0^1 \phi^{-\frac{3}{2}}(q)\,dq = (2ab)^{\frac{1}{2}}/(a + b), \tag{17}$$
$$\int_0^1 \phi^{-\frac{5}{2}}(q)\,dq = \tfrac{2}{3}(2ab)^{\frac{1}{2}}(a^2 + 3ab + b^2)/(a + b)^3,$$

we have

$$\begin{aligned} \nu_1 &= 2 \int_0^1 I(q)\,dq \\ &= 2a^2[2 - (b/(a + b))^2], \end{aligned} \tag{18}$$

ν_2 being obtained by interchanging a and b.

Similarly, it is found that

$$\nu_3 = 2(ab/(a + b))^2. \tag{19}$$

Hence

$$\begin{aligned} \mathcal{E}(t_1^2) &= \nu_1/a^2 + \nu_3/b^2 = 4 + 2(a - b)/(a + b), \\ \mathcal{E}(t_2^2) &= \nu_3/a^2 + \nu_2/b^2 = 4 - 2(a - b)/(a + b), \end{aligned} \tag{20}$$

and the matrices in (8) are not independent of a and b.

As a check on the working, we have:

$$\mathcal{E}(t_1^2 + t_2^2) = \mathcal{E}(\mathbf{z} - \boldsymbol{\xi})'\mathbf{S}^{-1}(\mathbf{z} - \boldsymbol{\xi}) = 8, \tag{21}$$

which is independent of $\boldsymbol{\Sigma}$ and in accordance with the known distribution of Hotelling's T^2.

REFERENCES

Bennett, G. W. and Cornish, E. A. (1964). A comparison of the simultaneous fiducial distributions derived from the multivariate normal distribution. *Bull. Inst. Internat. Statist.* **40.2** 902–919.

Olkin, Ingram and Rubin, Herman (1964). Multivariate beta distributions and independence properties of the Wishart distribution. *Ann. Math. Statist.* **35** 261–269.

Commonwealth of Australia, Commonwealth Scientific and Industrial Research Organisation

The Annals of Mathematical Statistics
1968, Vol. 39, No. 3, 815–832

A SYSTEM OF LINEAR DIFFERENTIAL EQUATIONS FOR THE DISTRIBUTION OF HOTELLING'S GENERALIZED T_0^2

By A. W. Davis

C.S.I.R.O., Adelaide

1. Introduction and summary. Let $\mathcal{S}_1$, $\mathcal{S}_2$ be independent $m \times m$ matrices on n_1, n_2 degrees of freedom respectively, $\mathcal{S}_2$ having a Wishart distribution and $\mathcal{S}_1$ having a possibly non-central Wishart distribution with the same covariance matrix. Hotelling's generalized T_0^2 statistic is then defined [7] by

$$T = n_2^{-1}T_0^2 = \operatorname{tr} \mathcal{S}_1\mathcal{S}_2^{-1}. \tag{1.1}$$

The complete distribution of this statistic is known only in particular cases. If $m = 1$, then $(n_2/n_1)T$ is simply non-central F. In the case $n_1 = 1$, T reduces to Hotelling's generalization of "Student's" t, which also has a non-central F distribution. When $m = 2$, Hotelling [7] has shown that in the null case the density function of T is

$$f(T) = [\Gamma(n_1 + n_2 - 1)/\Gamma(n_1)\Gamma(n_2 - 1)](\tfrac{1}{2}T)^{n_1-1}(1 + \tfrac{1}{2}T)^{-(n_1+n_2)} \cdot {}_2F_1(1, \tfrac{1}{2}(n_1 + n_2); \tfrac{1}{2}(n_2 + 1); v), \tag{1.2}$$

where $v = T^2/(T + 2)^2$, and ${}_2F_1$ is the Gaussian hypergeometric function.

When n_2 becomes large, the distribution of T_0^2 approaches that of χ^2 based on mn_1 degrees of freedom. Ito [9] has derived asymptotic expansions both for the cumulative distribution function (cdf) of T_0^2, and for the percentiles of T_0^2 in terms of the corresponding $\chi^2_{mn_1}$ percentiles.

Other approximations to the distribution requiring large n_2 for validity have been obtained by Pillai and Samson [12]. These authors have used the method of fitting a Pearson curve by means of moment quotients to tabulate upper 5% and 1% points for $m = 2, 3, 4$.

The exact distribution of T over the range $0 \leqq T < 1$ has been obtained in the general non-central case by Constantine [3], using the methods of zonal polynomials and hypergeometric functions of matrix argument developed by James and Constantine ([2] and [10], for example). Constantine's solution has the form

$$f(T) = [\Gamma_m(\tfrac{1}{2}(n_1 + n_2))/\Gamma(\tfrac{1}{2}mn_1)\Gamma_m(\tfrac{1}{2}n_2)]T^{\frac{1}{2}mn_1-1}\mathcal{P}(T), \tag{1.3}$$

where $\mathcal{P}(T)$ is a power series in T convergent in the unit circle, and

$$\Gamma_m(z) = \pi^{\frac{1}{4}m(m-1)} \prod_{i=0}^{m-1} \Gamma(z - \tfrac{1}{2}i). \tag{1.4}$$

In Section 2 of the present paper, it is shown that in the null case the density function $f(T)$ (or rather, its analytic continuation into the complex T-plane) satisfies an ordinary linear differential equation of degree m of Fuchsian type, having regular singularities at $T = 0, -1, \cdots, -m$ and infinity. More specifi-

Received 27 April 1967; revised 22 September 1967.

cally, an equivalent first-order system is obtained, and the problem is most conveniently treated in this form. Constantine's series (1.3) in the null case is shown in Section 3 to be the relevant solution for $f(T)$ in the neighbourhood of the regular singularity at $T = 0$. The differential equations lead to convenient recurrence relations for the coefficients in $\mathcal{P}(T)$. In Section 4 an alternative derivation of Ito's asymptotic formula is presented. Preliminary results are then given (Section 5) for the regular singularity at $T = \infty$, and a heuristic treatment of the limiting distribution as $n_1 \to \infty$ is presented in Section 6. Finally, it is shown in Section 7 that the moments of T may be obtained from the differential equations for the Laplace transform of $f(T)$ given in Section 1.

One objective in deriving the differential equations for $f(T)$ has been to obtain a convenient exact method for computing the distribution and its percentiles. This work is in progress, and it is hoped that results will be available shortly.

2. The system of linear differential equations. Let $w_1, \cdots, w_m$ denote the latent roots of $\mathcal{S}_1\mathcal{S}_2^{-1}$; then from (1.1)

$$T = \sum_{i=1}^m w_i. \tag{2.1}$$

Assuming that $\mathcal{S}_1$ has the central Wishart distribution, the joint density function of the w_i when $n_1, n_2 \geqq m$ is

$$\begin{aligned}\phi_{m;n_1,n_2}(\mathbf{w}) &= [\pi^{\frac{1}{2}m^2}\Gamma_m(\tfrac{1}{2}(n_1+n_2))/\Gamma_m(\tfrac{1}{2}n_1)\Gamma_m(\tfrac{1}{2}n_2)\Gamma_m(\tfrac{1}{2}m)] \\ &\quad\cdot(\textstyle\prod_{i=1}^m w_i)^{\frac{1}{2}(n_1-m-1)}\prod_{i=1}^m(1+w_i)^{-\frac{1}{2}(n_1+n_2)}\prod_{i<j}^m(w_i-w_j), \\ &\quad (0 < w_m < \cdots < w_1 < \infty).\end{aligned} \tag{2.2}$$

(See [5], [8], [13]). The case of singular $\mathcal{S}_1$, $n_1 < m$, does not require separate treatment, since the distribution of T then has a simple relation to the case $n_1 \geqq m$ ([3] Section 4). The following proof holds strictly for $m \geqq 2$.

Throughout this section, the suffixes on ϕ will be omitted for convenience. The Laplace transform (Lt) $L_0(s)$ of $f(T)$ may then be written in the form

$$L_0(s) = \int_{\mathfrak{D}_m} e^{-s\Sigma w_i}\,\phi(\mathbf{w})\,d\mathbf{w}, \qquad (s \geqq 0), \tag{2.3}$$

where $\mathfrak{D}_m$ denotes the region $\{0 < \omega_m < \cdots < w_1 < \infty\}$. In general, it may be seen that the functional

$$\mathcal{L}(\psi) = \int_{\mathfrak{D}_m} e^{-s\Sigma w_i}\,\psi(\mathbf{w})\,d\mathbf{w} \tag{2.4}$$

is the ordinary Lt of the following function of a single variable:

$$\Psi(T) = \int_{\mathfrak{D}_{m-1}(T)} \psi(T - w_2 - \cdots - w_m, w_2, \cdots, w_m)\,dw_2 \cdots dw_m, \tag{2.5}$$

where

$$\begin{aligned}\mathfrak{D}_{m-1}(T) &= \mathfrak{D}_{m-1} \cap \{2w_2 + w_3 + \cdots + w_m < T\} \\ &= \{0 < w_m < m^{-1}T;\, w_m < w_{m-1} < (m-1)^{-1}(T - w_m);\, \cdots; \\ &\quad w_3 < w_2 < \tfrac{1}{2}(T - w_3 - \cdots - w_m)\}.\end{aligned} \tag{2.6}$$

Taking $\psi = \phi$ in (2.5), we obtain an integral form of $f(T)$.
Thus $L_0(s) = \mathfrak{L}(\phi)$, and the following Lt's will also be introduced:

$$(2.7)\quad L_r(s) = \mathfrak{L}\{\phi(\mathbf{w}) \textstyle\sum_{k_1<k_2<\cdots<k_r}^{m} [(1 + w_{k_1}) \cdots (1 + w_{k_r})]^{-1}\}, \quad (r = 1, 2, \cdots, m).$$

The summation in curly brackets is extended over the $\binom{m}{r}$ selections of r distinct roots, and is a symmetric function of the w_i. Clearly, the $L_r(s)$ exist for all $s \geqq 0$. Differentiating under the sign of integration:

$$(2.8)\quad -L_r'(s) = \mathfrak{L}\{\phi(\mathbf{w}) \textstyle\sum_{k_1<\cdots<k_r}^{m} [(1 + w_{k_1}) \cdots (1 + \omega_{k_r})]^{-1} \cdot[(1 + w_{k_1}) + \cdots + (1 + w_{k_r}) - r + (w_{l_1} + \cdots + w_{l_{m-r}})]\},$$

where $(l_1, \cdots, l_{m-r})$ is the set of suffixes complementary to $(k_1, \cdots, k_r)$. Hence, writing

$$(2.9)\quad \Phi_r(s) = \mathfrak{L}\{\phi(\mathbf{w}) \textstyle\sum_{k_1<\cdots<k_r}^{m} (w_{l_1} + \cdots + w_{l_{m-r}}) \cdot[(1 + w_{k_1}) \cdots (1 + w_{k_r})]^{-1}\},$$

it is seen that

$$(2.10)\quad -L_r'(s) = (m - r + 1)L_{r-1}(s) - rL_r(s) + \Phi_r(s), \quad (r = 1, 2, \cdots, m).$$

Now let l denote any suffix distinct from each of $k_1, \cdots, k_r$. It follows by integration by parts that

$$s\mathfrak{L}\{\phi(\mathbf{w})w_l[(1 + w_{k_1}) \cdots (1 + w_{k_r})]^{-1}\}$$
$$= -\textstyle\int_{\mathfrak{D}_m} ((\partial/\partial w_l)e^{-s\Sigma w_i})w_l\phi(\mathbf{w})/(1 + w_{k_1}) \cdots (1 + w_{k_r})\, d\mathbf{w}$$
$$(2.11)\quad = \mathfrak{L}\{\phi(\mathbf{w})[(1 + w_{k_1}) \cdots (1 + w_{k_r})]^{-1}[-\tfrac{1}{2}(n_2 + m - 1) + \tfrac{1}{2}(n_1 + n_2)(1 + w_l)^{-1} + (\textstyle\prod_{i<j}^{m} (w_i - w_j))^{-1}w_l(\partial/\partial w_l) \cdot\textstyle\prod_{i<j}^{m} (w_i - w_j)]\}.$$

Since

$$(2.12)\quad (\textstyle\prod_{i<j}^{m} (w_i - w_j))^{-1}(\partial/\partial w_l) \textstyle\prod_{i<j}^{m} (w_i - w_j) = \textstyle\sum_{i=1, i\neq l}^{m} (w_l - w_i)^{-1},$$

it is necessary to consider the following summation in connection with (2.9):

$$(2.13)\quad \textstyle\sum_{k_1<\cdots<k_r}^{m} [(1 + w_{k_1}) \cdots (1 + w_{k_r})]^{-1} \cdot\{w_{l_1} \textstyle\sum_{i=1, i\neq l_1}^{m} (w_{l_1} - w_i)^{-1} + \cdots + w_{l_{m-r}} \textstyle\sum_{i=1, i\neq l_{m-r}}^{m} (w_{l_{m-r}} - w_i)^{-1}\}.$$

The coefficient of $(w_i - w_j)^{-1}$, $(i < j)$, is seen to be

$$w_i \textstyle\sum_{k_1<\cdots<k_r (\text{all } k_n \neq i)}^{m} [(1 + w_{k_1}) \cdots (1 + w_{k_r})]^{-1} - w_j \textstyle\sum_{k_1<\cdots<k_r (\text{all } k_n \neq j)}^{m} [(1 + w_{k_1}) \cdots (1 + w_{k_r})]^{-1}$$

$$(2.14)\quad \begin{aligned} &= w_i\{\textstyle\sum_{k_1<\cdots<k_r(\text{all } k_n\neq i,j)}^m [(1+w_{k_1})\cdots(1+w_{k_r})]^{-1} \\ &\quad + (1+w_j)^{-1}\textstyle\sum_{k_1<\cdots<k_{r-1}(\text{all } k_n\neq i,j)}^m ([1+w_{k_1})\cdots(1+w_{k_{r-1}})]^{-1}\} \\ &\quad - (\text{similar term with } i\,, j \text{ interchanged}) \\ &= (w_i - w_j)\textstyle\sum_{k_1<\cdots<k_r(\text{all } k_n\neq i,j)}^m [(1+w_{k_1})\cdots(1+w_{k_r})]^{-1} \\ &\quad + (w_i - w_j)(1+w_i+w_j)[(1+w_i)(1+w_j)]^{-1} \\ &\quad \cdot \textstyle\sum_{k_1<\cdots<k_{r-1}(\text{all } k_n\neq i,j)}^m [(1+w_{k_1})\cdots(1+w_{k_{r-1}})]^{-1}. \end{aligned}$$

The summation (2.13) may therefore be written as

$$(2.15)\quad \begin{aligned} &\textstyle\sum_{i<j}^m \sum_{k_1<\cdots<k_n(\text{all } k_n\neq i,j)}^m [(1+w_{k_1})\cdots(1+w_{k_r})]^{-1} \\ &\quad + \textstyle\sum_{i<j}^m \{(1+w_i)^{-1} + (1+w_j)^{-1}\} \\ &\qquad \cdot \textstyle\sum_{k_1<\cdots<k_{r-1}(\text{all } k_n\neq i,j)}^m [(1+w_{k_1})\cdots(1+w_{k_{r-1}})]^{-1} \\ &\quad - \textstyle\sum_{i<j}^m [(1+w_i)(1+w_j)]^{-1} \\ &\qquad \cdot \textstyle\sum_{k_1<\cdots<k_{r-1}(\text{all } k_n\neq i,j)}^m [(1+w_{k_1})\cdots(1+w_{k_{r-1}})]^{-1}. \end{aligned}$$

Since the second term in (2.15) is

$$(2.16)\quad \begin{aligned} &\textstyle\sum_{i\neq j}^m (1+w_i)^{-1}\sum_{k_1<\cdots<k_{r-1}(\text{all } k_n\neq i,j)}^m [(1+w_{k_1})\cdots(1+w_{k_{r-1}})]^{-1} \\ &\qquad = r(m-r)\textstyle\sum_{k_1<\cdots<k_r}^m [(1+w_{k_1})\cdots(1+w_{k_r})]^{-1}, \end{aligned}$$

it follows that (2.15) reduces to

$$(2.17)\quad \begin{aligned} &[\tbinom{m-r}{2} + r(m-r)]\textstyle\sum_{k_1<\cdots<k_r}^m [(1+w_{k_1})\cdots(1+w_{k_r})]^{-1} \\ &\qquad - \tbinom{r+1}{2}\textstyle\sum_{k_1<\cdots<k_{r+1}}^m [(1+w_{k_1})\cdots(1+w_{k_{r+1}})]^{-1}. \end{aligned}$$

Hence, from (2.11) and (2.17):

$$(2.18)\quad s\Phi_r(s) = -\tfrac{1}{2}(m-r)(n_2-r)L_r(s) + \tfrac{1}{2}(r+1)(n_1+n_2-r)L_{r+1}(s).$$

Substituting in (2.10) we obtain:

$$(2.19)\quad \begin{aligned} &(m-r+1)sL_{r-1} + [s((d/ds)-r) - \tfrac{1}{2}(m-r)(n_2-r)]L_r \\ &\qquad + \tfrac{1}{2}(r+1)(n_1+n_2-r)L_{r+1} = 0, \quad (r = 1, 2, \cdots, m-1). \end{aligned}$$

The same result holds for $r = 0$ and m if L_{-1} and L_{m+1} are defined to be identically zero. If $r = m$, a factor s may be cancelled, yielding

$$(2.20)\quad ((d/ds) - m)L_m + L_{m-1} = 0.$$

It now remains to invert the Laplace transforms. As seen earlier, $L_r(s)$ is the Lt of a certain function of a single variable which will be denoted by $H_r(T)$, $(H_{-1} \equiv 0, H_0 = f)$. In virtue of (2.5) and (2.7), it is seen that $H_r(T)$ is dominated for all $T \geqq 0$ by a constant multiple of $f(T)$. Constantine's result (1.3) shows that $f(T) = O(T^{\frac{1}{2}mn_1-1})$ as $T \to 0+$. (This may also be obtained by taking $\psi = \phi$ in (2.5)). Hence $sL_r(s)$ is the Lt of $H_r'(T)$, and since also $L_r'(s)$ is the Lt of

$-TH_r(T)$, the required system of first order differential equations is obtained:

$$(2.21)\quad -(m-r+1)\,dH_{r-1}/dT + \{(T+r)\,d/dT + a_r\}H_r - b_rH_{r+1} = 0, \qquad (r = 0, 1, \cdots, m-1),$$

$$(2.22)\quad -H_{m-1} + (T+m)H_m = 0,$$

where

$$(2.23)\quad a_r = \tfrac{1}{2}(m-r)(n_2-r)+1, \qquad b_r = \tfrac{1}{2}(r+1)(n_1+n_2-r), \qquad (r = 0, 1, \cdots, m).$$

Although these equations have been derived for $m \geqq 2$, they also hold for $m = 1$.

Elimination of $H_1, \cdots, H_m$ from equations (2.21–22) will clearly yield a linear homogeneous differential equation of order m for $f = H_0$. The coefficient of $f^{(r)}$ is a polynomial in T of order $r+1$, that of the highest derivative $f^{(m)}$ being $T(T+1)\cdots(T+m)$. The differential equation is therefore of Fuchsian type with regular singularities at $0, -1, \cdots, -m$, and infinity. In particular, when $m = 2$:

$$(2.24)\quad T(T+1)(T+2)f'' + [\tfrac{1}{2}(3n_2+5)T^2 + 2(n_2-n_1+4)T - 2(n_1-2)]f' + \tfrac{1}{2}(n_2+1)[(n_2+1)T - 2(n_1-2)]f = 0.$$

If the transformations

$$(2.25)\quad f(T) = T^{n_1-1}(1+\tfrac{1}{2}T)^{-(n_1+n_2)}g(T), \qquad v = T^2/(T+2)^2$$

are made in (2.24), g may be shown to satisfy a hypergeometric equation in conformity with Hotelling's result (1.2).

In the general case, however, it is preferable to work with the linear system (2.21–22) itself. An extensive literature exists dealing with such systems (see [1]). To express the result in matrix form, the following notation will be introduced for $(m+1)\times(m+1)$ matrices, all of whose elements are zero except those on their leading, upper and lower diagonals:

$$(2.26)\quad \{(\lambda_0, \cdots, \lambda_{m-1}), (\mu_0, \cdots, \mu_m), (\nu_1, \cdots, \nu_m)\} = \begin{bmatrix} \mu_0, \lambda_0, 0, \cdots, 0 \\ \nu_1, \mu_1, \lambda_1, \cdots, 0 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ 0 \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot\cdot \\ \cdot \quad \cdot \quad \cdot \quad 0 \\ \nu_{m-1}, \mu_{m-1}, \lambda_{m-1} \\ 0, \cdots, 0, \nu_m, \mu_m \end{bmatrix}.$$

Columns and rows will be numbered $0, 1, \cdots, m$.

Differentiating (2.22) with respect to T to achieve symmetry, and introducing the column vector

$$\mathbf{H} = (H_0, H_1, \cdots, H_m)', \tag{2.27}$$

the system may be written in the form

$$(T\mathbf{E}_{m+1} + \mathbf{A})\, d\mathbf{H}/dT = \mathbf{BH}, \tag{2.28}$$

where $\mathbf{E}_{m+1}$ is the $(m+1) \times (m+1)$ unit matrix, and

$$\mathbf{A} = \{(0, 0, \cdots, 0), (0, 1, 2, \cdots, m), (-m, -(m-1), \cdots, -1)\}, \tag{2.29}$$
$$\mathbf{B} = \{(b_0, \cdots, b_{m-1}), (-a_0, -a_1, \cdots, -a_m), (0, \cdots, 0)\}.$$

3. The regular singularity at the origin. Equation (2.28) may also be written as

$$d\mathbf{H}/dT = (T^{-1}\mathbf{R} + \textstyle\sum_{r=0}^{\infty} \mathbf{S}_r T^r)\mathbf{H}, \tag{3.1}$$

where $\mathbf{R}$, $\mathbf{S}_r$ are constant $(m+1) \times (m+1)$ matrices. The standard procedure for discussing the solution of (3.1) in the vicinity of the origin is to reduce $\mathbf{R}$ to its canonical Jordan form by means of a suitable linear transformation of $\mathbf{H}$ ([1], Chapter 4). For present purposes it is sufficient to find a matrix $\mathbf{P}$ reducing $\mathbf{A}$ to its canonical form:

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{\Lambda} = \text{diag}\,\{0, 1, \cdots, m\}. \tag{3.2}$$

The right-hand side denotes an $(m+1) \times (m+1)$ diagonal matrix. A suitable $\mathbf{P}$, together with its inverse, is given by

$$\mathbf{P} = \{p_{ij}\}, \qquad p_{ij} = \binom{m-j}{m-i}, \tag{3.3}$$
$$\mathbf{P}^{-1} = \{p_{ij}^*\}, \qquad p_{ij}^* = (-1)^{i+j}\binom{m-j}{m-i}, \qquad (i, j = 0, 1, \cdots, m).$$

Clearly, both $\mathbf{P}$ and $\mathbf{P}^{-1}$ are lower triangular. It may be shown without difficulty that

$$\mathbf{P}^{-1}\mathbf{BP} = \mathbf{C} = \{(\beta_0, \cdots, \beta_{m-1}), (\alpha_0, \cdots, \alpha_m), (\gamma_1, \cdots, \gamma_m)\}, \tag{3.4}$$

where

$$\alpha_i = \tfrac{1}{2}[(m - 2i)n_1 - in_2 + (2i^2 - mi - i - 2)],$$
$$\beta_i = \tfrac{1}{2}(i+1)(n_1 + n_2 - i), \tag{3.5}$$
$$\gamma_i = -\tfrac{1}{2}(m - i + 1)(n_1 - i + 1).$$

Thus, if we write

$$\mathbf{H} = \mathbf{PM}, \tag{3.6}$$

equation (2.28) becomes

$$(T\mathbf{E}_{m+1} + \mathbf{\Lambda})\, d\mathbf{M}/dT = \mathbf{CM}, \tag{3.7}$$

or, alternatively,

$$d\mathbf{M}/dT = \{T^{-1}\mathbf{V}_0 + \textstyle\sum_{r=1}^{m} (T + r)^{-1}\mathbf{V}_r\}\mathbf{M}, \tag{3.8}$$

where

$$\mathbf{V}_0 = \{(\tfrac{1}{2}(n_1 + n_2), 0, \cdots, 0), (\tfrac{1}{2}mn_1 - 1, 0, \cdots, 0), (0, 0, \cdots, 0)\},$$

$$\mathbf{V}_r = \{(0, \cdots, 0, \beta_r, 0, \cdots, 0), (0, \cdots, 0, \alpha_r, 0, \cdots, 0), (0, \cdots, 0, \gamma_r, 0, \cdots, 0)\} \quad (r = 1, \cdots, m). \tag{3.9}$$

The characteristic roots of $\mathbf{V}_0$ are zero (with multiplicity m) and $\frac{1}{2}mn_1 - 1$. The presence of equal roots makes discussion of the complete solution difficult. However, in view of Constantine's result (1.3), the relevant solution in the vicinity of $T = 0$ is of the form

$$\mathbf{M} = k(m; n_1, n_2)T^{\frac{1}{2}mn_1 - 1}\textstyle\sum_{r=0}^{\infty} \mathbf{W}_r T^r, \qquad (|T| < 1), \tag{3.10}$$

where $k(m; n_1, n_2)$ is the constant in square brackets in (1.3). The vectors $\mathbf{W}_r = (W_{0r}, \cdots, W_{mr})'$ may be determined without difficulty since $\frac{1}{2}mn_1 - 1$ is the largest root ([1] Chapter 4, Problem 13) except when $m = n_1 = 1$. Substitution in (3.7) yields the following recurrence relations:

$$\mathbf{W}_0 = (1, 0, \cdots, 0)';$$

$$\begin{aligned} i(r + \tfrac{1}{2}mn_1 - 1)W_{ir} &= \gamma_i W_{i-1,r-1} + [\alpha_i - (r + \tfrac{1}{2}mn_1 - 2)]W_{i,r-1} + \beta_i W_{i+1,r-1} \\ &\qquad (i = 1, \cdots, m; r = 1, 2, \cdots); \end{aligned} \tag{3.11}$$

$$rW_{0r} = \tfrac{1}{2}(n_1 + n_2)W_{1r}, \qquad (r = 1, 2, \cdots).$$

When $m = n_1 = 1$, the roots are $-\frac{1}{2}$, 0, and do not differ by an integer.

4. Ito's asymptotic expansion for large n_2. Let us write

$$t = n_2 T = T_0^2. \tag{4.1}$$

Then the cdf $F(t)$ of t has a power series representation, convergent for $|t| < n_2$, which may be obtained by term-by-term integration of Constantine's series. Essentially, Ito's expansion of $F(t)$ is obtained by rearranging this series as a convergent power series in n_2^{-1}, and multiplying by the Stirling-type asymptotic expansion of $k(m; n_1, n_2)$.

It is readily seen by induction from (3.11) that each W_{ir} is a polynomial in n_2 of order $(r - i)$ at most. Hence, making the substitution (4.1) in (3.10),

$$\mathbf{M} = k(m; n_1, n_2)(t/n_2)^{\frac{1}{2}mn_1 - 1}\textstyle\sum_{r=0}^{\infty} \mathbf{W}_r^* t^r, \qquad (\mathbf{W}_r^* = n_2^{-r}\mathbf{W}_r), \tag{4.2}$$

where the components W_{ir}^* of W_r^* are polynomials in n_2^{-1}, lower powers $\geq i$. To verify that rearrangement of (4.2) as a power series in n_2^{-1} is valid, we note that this series is dominated by the corresponding solution of

$$d\mathbf{M}/dt = \{t^{-1}\mathbf{V}_0 + \textstyle\sum_{r=1}^{m} (r - n_2^{-1}t)^{-1} n_2^{-1}\mathbf{V}_r^{+}\}\mathbf{M}, \tag{4.3}$$

where the $\mathbf{V}_r^+$ are obtained from the $\mathbf{V}_r$ $(r = 1, \cdots, m)$ by replacing all negative signs by positive signs in the expressions for α_i, β_i, γ_i. This solution is a double power-series in t and n_2^{-1} with positive coefficients, convergent for $|t| < n_2$.

In order to obtain the rearranged series for the component $M_0 = f$, it is convenient to first remove the factor n_2^{-i} from each W^*_{ir}, i.e. we make the shearing transformation

$$\mathbf{M} = \operatorname{diag}\{1, n_2^{-1}, \cdots, n_2^{-m}\}\mathbf{N}. \tag{4.4}$$

Then

$$\mathbf{N} = n_2^{-(\frac{1}{2}mn_1-1)}k(m; n_1, n_2)\sum_{r=0}^{\infty} n_2^{-r}\mathbf{Y}_r(t), \tag{4.5}$$

where

$$\begin{aligned} \mathbf{Y}_0(t) &= t^{\frac{1}{2}mn_1-1}\mathbf{y}_0(t), & y_{00}(0) &= 1, \\ \mathbf{Y}_r(t) &= t^{\frac{1}{2}mn_1}\mathbf{y}_r(t), & (r &= 1, 2, \cdots), \end{aligned} \tag{4.6}$$

and the components $y_{ir}(t)$ of the $\mathbf{y}_r(t)$ are power series in t. The second requirement in (4.6) arises because $\mathbf{W}_0^*$ and its transform under (4.4) are both $(1, 0, \cdots, 0)'$, which is independent of n_2^{-1}.

From (3.7), $\mathbf{N}$ satisfies the equation:

$$\begin{aligned} &(n_2^{-1}t\mathbf{E}_{m+1} + \mathbf{\Lambda})\, d\mathbf{N}/dt \\ &= \{(n_2^{-2}\beta_0, \cdots, n_2^{-2}\beta_{m-1}), (n_2^{-1}\alpha_0, \cdots, n_2^{-1}\alpha_m), (\gamma_1, \cdots, \gamma_m)\}\mathbf{N} \\ &= \{\mathbf{\Delta}_0 + n_2^{-1}\mathbf{\Delta}_1 + n_2^{-2}\mathbf{\Delta}_2\}\mathbf{N}. \end{aligned} \tag{4.7}$$

The matrices $\mathbf{\Delta}_i$ are given by:

$$\begin{aligned} \mathbf{\Delta}_0 &= \{(0, \cdots, 0)(0, -\tfrac{1}{2}, -1, \cdots, -\tfrac{1}{2}m), (\gamma_1, \cdots, \gamma_m)\}, \\ \mathbf{\Delta}_1 &= \{(\tfrac{1}{2}, 1, \cdots, \tfrac{1}{2}m), (\bar{\alpha}_0, \bar{\alpha}_1, \cdots, \bar{\alpha}_m), (0, \cdots, 0)\}, \\ \mathbf{\Delta}_2 &= \{(\bar{\beta}_0, \cdots, \bar{\beta}_{m-1}), (0, \cdots, 0), (0, \cdots, 0)\}, \end{aligned} \tag{4.8}$$

where

$$\begin{aligned} \bar{\alpha}_i &= \tfrac{1}{2}[(m - 2i)n_1 + 2i^2 - mi - i - 2], \\ \bar{\beta}_i &= \tfrac{1}{2}(i + 1)(n_1 - i). \end{aligned} \tag{4.9}$$

Substituting (4.5) in (4.7), it is found that:

$$[t(d/dt) - (\tfrac{1}{2}mn_1 - 1)]Y_{0,r} - \tfrac{1}{2}Y_{1,r} = \tfrac{1}{2}n_1 Y_{1,r-1}, \quad (r = 0, 1, \cdots), \tag{4.10}$$

while for $i = 1, \cdots, m$:

$$\begin{aligned} &i((d/dt) + \tfrac{1}{2})Y_{i,r} - \gamma_i Y_{i-1,r} \\ &\qquad = (-t(d/dt) + \bar{\alpha}_i)Y_{i,r-1} + \tfrac{1}{2}(i + 1)Y_{i+1,r-1} + \bar{\beta}_i Y_{i+1,r-2}. \end{aligned} \tag{4.11}$$

Y_{-1} and Y_{-2} are taken to be identically zero.

These equations determine the $Y_{i,r}$ uniquely in virtue of (4.6). It is seen that for $r = 0, 1, \cdots$, $Y_{0,r}$ and $Y_{1,r}$ are given by a pair of simultaneous linear differential equations, inhomogeneous except when $r = 0$. The corresponding homogeneous equations are of the form:

$$\begin{bmatrix} t, & 0 \\ 0, & 1 \end{bmatrix} d\mathbf{Z}/dt = \begin{bmatrix} \frac{1}{2}\, mn_1 - 1, & \frac{1}{2} \\ -\frac{1}{2}\, mn_1, & -\frac{1}{2} \end{bmatrix} \mathbf{Z}, \qquad (\mathbf{Z} = (Z_0, Z_1)'), \tag{4.12}$$

or, equivalently:

$$d\mathbf{Z}/dt = \left\{ t^{-1} \begin{bmatrix} \frac{1}{2}mn_1 - 1, & \frac{1}{2} \\ 0, & 0 \end{bmatrix} + \begin{bmatrix} 0, & 0 \\ -\frac{1}{2}mn_1, & -\frac{1}{2} \end{bmatrix} \right\} \mathbf{Z}. \tag{4.13}$$

The solution corresponding to the root $\frac{1}{2}mn_1 - 1$ of the leading matrix is easily verified to be

$$\mathbf{Z} = \text{const}\ \gamma(t)(1, -t)', \tag{4.14}$$

where

$$\gamma(t) = e^{-\frac{1}{2}t} t^{\frac{1}{2}mn_1 - 1}. \tag{4.15}$$

For $(Y_{0,0}, Y_{1,0})$ it follows from (4.6) that the zero root does not apply, and that the constant in (4.14) is unity in this case. Hence it follows from (4.11) that for $r = 0$:

$$\begin{aligned} Y_{i,0} = {} & (-t)^i \gamma(t)(m-1)(m-2) \cdots (m-i+1)(n_1-1)(n_1-2) \\ & \cdots (n_1 - i + 1)[i!\,(mn_1+2)(mn_1+4) \cdots (mn_1 + 2(i-1))]^{-1}, \\ & (i = 2, \cdots, m). \end{aligned} \tag{4.16}$$

Similarly, in solving for $Y_{0,r}, Y_{1,r}$, $(r \geqq 1)$, it is found that only the solution (4.14) of the homogeneous part applies. The constant is determined by the requirement that the lowest power of t occurring is at least $\frac{1}{2}mn_1$. The same requirement eliminates the general term const $e^{-\frac{1}{2}t}$ in the solution for $Y_{i,r}$ $(i = 2, \cdots, m)$.

In order to derive the expansion ([9], equation (4.3)) of Ito's paper, it is sufficient to calculate the following $Y_{i,r}$:

$$\begin{aligned} Y_{0,1} = {} & t\gamma(t)\{-\tfrac{1}{2}n_1 + \tfrac{1}{4}t(m + n_1 + 1)(mn_1 + 1)^{-1}\}, \\ Y_{0,1} = {} & t^2\gamma(t)(mn_1+2)^{-1}\{\tfrac{1}{2}[m(n_1^2+2) + (4n_1+2)] - \tfrac{1}{4}t(m + n_1 + 1)\}, \\ Y_{2,1} = {} & (m-1)(n_1-1)t^3\gamma(t)[(mn_1+2)(mn_1+4)]^{-1} \\ & \cdot\{-\tfrac{1}{4}[m(n_1^2+2) + (4n_1+2)] \\ & + \tfrac{1}{8}t(mn_1+6)^{-1}[mn_1 + m(n_1^2 + n_1 + 8) + (8n_1+4)]\}, \\ Y_{0,2} = {} & t^2\gamma(t)(mn_1+2)^{-1}\{\tfrac{1}{8}n_1\,[m(n_1^2+2) + (4n_1+2)] \\ & - \tfrac{1}{24}t(mn_1+4)^{-1}[m^2(3n_1^2+4) + 3m(n_1^3 + n_1^2 + 8n_1 + 4) \\ & + (16n_1^2 + 24n_1 + 16)] \\ & + \tfrac{1}{32}t^2[(mn_1+4)(mn_1+6)]^{-1}[m^3n_1 + m^2(2n_1^2 + 2n_1 + 8) \\ & + m(n_1^3 + 2n_1^2 + 21n_1 + 20) + (8n_1^2 + 20n_1 + 20)]\}. \end{aligned} \tag{4.17}$$

The density function of t is then

$$f(t) = n_2^{-\frac{1}{2}mn_1}k(m; n_1, n_2)\{\gamma(t) + n_2^{-1}Y_{0,1} + n_2^{-2}Y_{0,2} + \cdots\}, \quad (|t| < n_2). \tag{4.18}$$

Integrating with respect to t to obtain $F(t)$, and using the expansion

$$\begin{aligned} n_2^{-\frac{1}{2}mn_1}k(m; n_1, n_2) \sim [2^{\frac{1}{2}mn_1}\Gamma(\tfrac{1}{2}mn_1)]^{-1}\{1 &+ \tfrac{1}{4}mn_1n_2^{-1}(n_1 - m - 1) \\ &+ mn_1(96n_2^2)^{-1}[3m^3n_1 - 2m^2(3n_1^2 - 3n_1 + 4) \\ &+ 3m(n_1^3 - 2n_1^2 + 5n_1 - 4) \\ &+ (-8n_1^2 + 12n_1 + 4)] + \cdots\}, \end{aligned} \tag{4.19}$$

Ito's result is obtained as an asymptotic expansion uniformly valid for t in any bounded interval. The corresponding expansion of the percentiles of t for large n_2 ([9] equation (3.33)) may be derived formally from that of the cdf by means of an algorithm found by G. W. Hill and the present author ([6]).

5. The regular singularity at infinity. Letting

$$z = T^{-1}, \tag{5.1}$$

equation (3.8) takes the form

$$d\mathbf{M}/dz = \{-z^{-1}\mathbf{C} + \textstyle\sum_{r=1}^{m} (z + r^{-1})^{-1}\mathbf{V}_r\}\mathbf{M}. \tag{5.2}$$

Thus (3.8) has a regular singularity at $T = \infty$, with linearly independent solutions corresponding to the $(m + 1)$ latent roots of $-\mathbf{C}$, convergent for $|T| > m$. Since $\mathbf{C}$ is similar to $\mathbf{B}$, these roots are

$$a_r = \tfrac{1}{2}(m - r)(n_2 - r) + 1, \qquad (r = 0, 1, \cdots, m). \tag{5.3}$$

The a_r form a decreasing sequence for increasing r. We now seek to relate these solutions to $f(T)$. Let $l(m; n_1, n_2)$ denote the constant in (2.2). From (2.5) we have, as $T \to \infty$,

$$\begin{aligned} &f(T)T^{\frac{1}{2}(n_2-m+3)} \\ &= l(m; n_1, n_2) \int_{\mathfrak{D}_{m-1}(T)} (1 - T^{-1} \textstyle\sum_{i=2}^m w_i)^{\frac{1}{2}(n_1-m-1)} \\ &\quad \cdot (\textstyle\prod_{i=2}^m w_i)^{\frac{1}{2}(n_1-m-1)}[1 + T^{-1}(1 - \textstyle\sum_{i=2}^m w_i)]^{-\frac{1}{2}(n_1+n_2)} \\ &\quad \cdot \textstyle\prod_{i=2}^m (1 + w_i)^{-\frac{1}{2}(n_1+n_2)} \prod_{j=2}^m [1 - T^{-1}(\sum_{i=2}^m w_i + w_j)] \\ &\quad \cdot \textstyle\prod_{2 \le i < j \le m} (w_i - w_j)\, dw_2 \cdots dw_m \\ &\to [l(m; n_1, n_2)/l(m - 1; n_1 - 1, n_2 + 1)] \int_{\mathfrak{D}_{m-1}} \phi_{m-1; n_1-1, n_2+1}(\mathbf{w})\, d\mathbf{w} \\ &= \pi^{\frac{1}{2}}\Gamma(\tfrac{1}{2}(n_1 + n_2 - m + 1))\Gamma(\tfrac{1}{2}(n_2 + 1))/[\Gamma(\tfrac{1}{2}m) \\ &\quad \cdot \Gamma(\tfrac{1}{2}n_1)\Gamma(\tfrac{1}{2}(n_2 - m + 1))\, \Gamma(\tfrac{1}{2}(n_2 - m + 2))] \\ &= \kappa(m; n_1, n_2), \quad \text{say}. \end{aligned} \tag{5.4}$$

In order to justify the limit, we note that since

$$w_2 + \cdots + w_m < T - w_2 < \tfrac{1}{2}T, \tag{5.5}$$

the integrand is dominated on $\mathfrak{D}_{m-1}$ by a multiple of $\phi_{m-1;n_1-1,n_2+1}$. Since

$$a_{m-1} = \tfrac{1}{2}(n_2 - m + 3), \tag{5.6}$$

equation (5.4) shows that in the neighbourhood of $T = \infty$, the relevant **M** may be a linear combination of the solutions corresponding to all roots of $-\mathbf{C}$ except $a_m = 1$. Although (5.4) yields the coefficient of the a_{m-1} solution, the problem of determining the other coefficients in this linear combination remains unsolved. A further complication consists in the fact that the a_r differ by integers if $n_2 - m + 1$ is an even integer, while if $n_2 - m + 1$ is odd then $a_r - a_{r+2}$ is integral. The solutions therefore involve logarithmic terms. This situation is unfortunate because the tabulation of higher percentile points of $f(T)$ is of considerable interest.

If $n_2 > m + 1$, the first four terms in the expansion of $F(T)$ for large T are:

$$\begin{aligned} F(T) &= 1 - \kappa(m; n_1, n_2)T^{-\frac{1}{2}(n_2-m+1)}\{2(n_2 - m + 1)^{-1} \\ &+ [T(n_2 - m + 3)]^{-1}[m(n_1 - 1) - (2n_1 + n_2 - 1)] \\ &+ [4T^2(n_2 - m + 5)]^{-1}[m^2(n_1 - 1)^2 - 2m(n_1 - 1)(3n_1 + n_2 - 2) \\ &+ n_2^2 - 2n_2(n_1^2 - 5n_1 + 2) + 3(2n_1^2 - 2n_1 + 1) \\ &+ 2(n_1 - 1)(n_1 - 2)n_2(n_2 + 1)(n_2 - m + 2)^{-1}] + \cdots\}. \end{aligned} \tag{5.7}$$

This result may be derived by applying to **H** the transformation given later in Section 7. However, the details will be omitted.

6. The limiting distribution for increasing n_1. As n_1 becomes large, it is clear that the random variable

$$\tau = n_1^{-1}T = \operatorname{tr}\{(n_1^{-1}\mathcal{S}_1)\mathcal{S}_2^{-1}\} \tag{6.1}$$

will converge in distribution. The limiting distribution is perhaps of mainly mathematical interest, but it has the merit of giving completeness to tables of the Hotelling statistic. Our discussion in this section will be rather heuristic.

We first obtain an integral form of the limiting density function, which will be denoted by $\theta(\tau)$. Take $\psi = \phi_{m;n_1,n_2}$ in (2.5). Substituting

$$w_i = n_1 u_i, \qquad (i = 2, \cdots, m), \tag{6.2}$$

and letting $n_1 \to \infty$ it is found that

$$\begin{aligned} \theta(\tau) &= \lim_{n_1\to\infty} n_1 f(n_1 \tau) \\ &= \alpha(m; n_2) \int_{\mathfrak{D}_{m-1}(\tau)} \exp\{-\tfrac{1}{2}[(\tau - \textstyle\sum_{i=2}^m u_i)^{-1} + \sum_{i=2}^m u_i^{-1}]\} \\ &\cdot[(\tau - \textstyle\sum_{i=2}^m u_i) \prod_{i=2}^m u_i]^{-\frac{1}{2}(n_2+m+1)} \prod_{j=2}^m (\tau - \sum_{i=2}^m u_i - u_j) \\ &\cdot \textstyle\prod_{2\leq j<i\leq m} (u_i - u_j)\, du_2 \cdots du_m, \end{aligned} \tag{6.3}$$

where

$$\alpha(m; n_2) = \lim_{n_1\to\infty} n_1^{-\frac{1}{2}mn_2} l(m; n_1, n_2) = \pi^{\frac{1}{2}m^2}/2^{\frac{1}{2}mn_2}\Gamma_m(\tfrac{1}{2}n_2)\Gamma_m(\tfrac{1}{2}m). \tag{6.4}$$

For large τ, one easily shows that

$$\theta(\tau) = O(\tau^{-\frac{1}{2}(n_2-m+3)}). \tag{6.5}$$

The behaviour for small τ is more complicated. Setting

$$v_i = \tau u_i, \qquad (i = 2, \cdots, m), \tag{6.6}$$

in equation (6.3), we obtain:

$$\begin{aligned} \theta(\tau) = {} & \alpha(m; n_2)\tau^{-\frac{1}{2}mn_2-1} \\ & \cdot \int_{\mathfrak{R}_{m-1}} \exp\{-(2\tau)^{-1}[(1 - \textstyle\sum_{i=2}^m v_i)^{-1} + \sum_{i=2}^m v_i^{-1}]\} \\ & \cdot [(1 - \textstyle\sum_{i=2}^m v_i) \prod_{i=2}^m v_i]^{-\frac{1}{2}(n_2+m+1)} \prod_{j=2}^m (1 - \sum_{i=2}^m v_i - v_j) \\ & \cdot \textstyle\prod_{2\leq i<j\leq m} (v_i - v_j)\, dv_2 \cdots dv_m \end{aligned} \tag{6.7}$$

where

$$\mathfrak{R}_{m-1} = \{0 < v_m < m^{-1}; v_m < v_{m-1} < (m-1)^{-1}(1 - v_m); \cdots; v_3 < v_2 < \tfrac{1}{2}(1 - v_2 - \cdots - v_m)\}. \tag{6.8}$$

An asymptotic estimate of $\theta(\tau)$ as $\tau \to 0+$ may be obtained by means of a method due to Laplace ([4] Chapter I, Section 3). Consider an integral of the form

$$\mathfrak{F}(\tau) = \int_\eta^\xi e^{\tau^{-1}\omega(v)}\rho(v)\, dv. \tag{6.9}$$

Suppose that $\omega(v)$ is real on (η, ξ) and has its greatest value in the interval at $v = \xi$, with $\omega'(\xi) = 0$, $\omega''(\xi) < 0$. Under wide conditions on ρ, ω, we may expect that, as $\tau \to 0+$,

$$\mathfrak{F}(\tau) \sim \tfrac{1}{2}e^{\tau^{-1}\omega(\xi)} \textstyle\sum_{r=0}^\infty \Gamma(\tfrac{1}{2}(r+1))c_r\tau^{\frac{1}{2}(r+1)}, \tag{6.10}$$

where

$$\textstyle\sum_{r=0}^\infty c_r u^r = \rho(\zeta(u))\zeta'(u) \tag{6.11}$$

and $v = \zeta(u)$ is the inverse of

$$u = +[\omega(\xi) - \omega(v)]^{\frac{1}{2}}. \tag{6.12}$$

The integral involving v_2 in (6.7) is:

$$\begin{aligned} \mathfrak{g}_2 = {} & \int_{v_2=v_3}^{\frac{1}{2}(1-\Sigma_3)} \exp\{-(2\tau)^{-1}[(1 - \textstyle\sum_3 - v_2)^{-1} + v_2^{-1}]\} \\ & \cdot [(1 - \textstyle\sum_3 - v_2)v_2]^{-\frac{1}{2}(n_2+m+1)} \\ & \cdot (1 - \textstyle\sum_3 - 2v_2) \prod_{j\geq 3} [(1 - \sum_3 - v_j - v_2)(v_2 - v_j)]\, dv_2, \end{aligned} \tag{6.13}$$

where $\sum_3$ denotes $\sum_3^m v_i$.

Clearly

$$\omega(v_2) = -\tfrac{1}{2}[(1 - \textstyle\sum_3 - v_2)^{-1} + v_2^{-1}] \tag{6.14}$$

has its greatest value at $\xi = \frac{1}{2}(1 - \sum_3)$:

$$\omega(\xi) = -2(1 - \textstyle\sum_3)^{-1}. \tag{6.15}$$

Also,

$$\begin{aligned} u &= +[\omega(\xi) - \omega(v_2)]^{\frac{1}{2}} \\ &= [\tfrac{1}{2}(1 - \textstyle\sum_3)]^{-\frac{3}{2}}(\xi - v_2) + \cdots , \end{aligned} \tag{6.16}$$

whence

$$v_2 = \zeta(u) = \tfrac{1}{2}(1 - \textstyle\sum_3) - [\tfrac{1}{2}(1 - \textstyle\sum_3)]^{\frac{3}{2}}u + \cdots . \tag{6.17}$$

Hence, obtaining ρ from (6.13):

$$\begin{aligned} &\rho(\zeta(u))\zeta'(u) \\ &\quad = u \cdot 2^{n_2 - m + 3}(1 - \textstyle\sum_3)^{-(n_2 + m - 2)} \prod_{j \geqq 3} [(1 - \textstyle\sum_3) - 2v_j]^2 + \cdots . \end{aligned} \tag{6.18}$$

It follows from (6.10) that

$$\begin{aligned} \mathcal{J}_2 \sim d_2 \tau \exp \{-2\tau^{-1}(1 - \textstyle\sum_3)^{-1}\} \\ \cdot (1 - \textstyle\sum_3)^{-(n_2+m-2)} \prod_{j \geqq 3} [(1 - \textstyle\sum_3) - 2v_j]^2, \end{aligned} \tag{6.19}$$

where

$$d_2 = 2^{n_2 - m + 2}. \tag{6.20}$$

The same method may be applied successively to the $v_3, \cdots, v_m$ integrals. The integral with respect to v_r is found by induction to be:

$$\begin{aligned} \mathcal{J}_r \sim d_r \tau^{\frac{1}{2}r} \exp \{-r^2(2\tau)^{-1}(1 - \textstyle\sum_{r+1})^{-1}\}(1 - \textstyle\sum_{r+1})^{-\mu_r} \\ \cdot \prod_{j \geqq r+1} [(1 - \textstyle\sum_{r+1}) - rv_j]^r, \end{aligned} \tag{6.21}$$

where

$$\begin{aligned} d_r &= 2^{\frac{1}{2}r - 1}\Gamma(\tfrac{1}{2}r) r^{\frac{1}{2}r(n_2 - m) + \frac{1}{4}(r^2 + r + 2)}(r - 1)^{-\frac{1}{2}(r-1)(n_2 - m) - \frac{1}{4}(r^2 - r + 4)} \\ \mu_r &= \tfrac{1}{2}r(n_2 + m - 1) - \tfrac{3}{4}(r - 1)(r + 2), \ \textstyle\sum_{r+1} \equiv \sum_{r+1}^m v_i . \end{aligned} \tag{6.22}$$

Hence, as $\tau \to 0+$,

$$\begin{aligned} \theta(\tau) &\sim \alpha(m; n_2)(\textstyle\prod_{r=2}^m d_r) e^{-m^2/2\tau} \tau^{-\frac{1}{2}mn_2 - 1 + \frac{1}{2}\sum_2^m r} \\ &= [\pi^{\frac{1}{4}(m-1)(m+2)} m^{\frac{1}{4}m(2n_2 - m + 1) + 1} / 2^{\frac{1}{2}mn_2 - \frac{1}{4}(m-1)(m-2)} (m!)^{\frac{1}{2}} \Gamma_m(\tfrac{1}{2}n_2)] \\ &\quad \cdot e^{-m^2/2\tau} \tau^{-\frac{1}{2}mn_2 + \frac{1}{4}(m-2)(m+3)}. \end{aligned} \tag{6.23}$$

Although the above derivation strictly requires that $m \geqq 2$, it is seen that when $m = 1$ (6.23) reduces to

(6.24) $$\theta(\tau) \sim [2^{\frac{1}{2}n_2}\Gamma(\tfrac{1}{2}n_2)]^{-1}e^{-1/2\tau}\tau^{-(\frac{1}{2}n_2+1)},$$

which is the density function of $\tau = 1/\chi^2$, where χ^2 is based on n_2 degrees of freedom.

A system of linear differential equations for $\theta(\tau)$ will now be derived. Let the following substitutions be made in (2.28):

(6.25) $$\tau = n_1^{-1}T, \qquad \mathbf{H} = \operatorname{diag}\{1, n_1^{-1}, \cdots, n_1^{-m}\}\,\mathbf{J}.$$

Then

(6.26) $$\begin{aligned}[\tau\mathbf{E}_{m+1} + \{(0, \cdots, 0), (0, n_1^{-1}, 2n_1^{-1}, \cdots, mn_1^{-1}),\\ (-m, -(m-1), \cdots, -1)\}]\,d\mathbf{J}/d\tau\\ = \{(n_1^{-1}b_0, \cdots, n_1^{-1}b_{m-1}), (-a_0, \cdots, -a_m), (0, \cdots, 0)\}\mathbf{J}.\end{aligned}$$

Formally letting $n_1 \to \infty$, we obtain

(6.27) $$(\tau\mathbf{E}_{m+1} + \boldsymbol{\Omega})\,d\mathbf{J}/d\tau = -\boldsymbol{\Gamma}\mathbf{J},$$

where

(6.28) $$\begin{aligned}\boldsymbol{\Omega} &= \{(0, \cdots, 0), (0, \cdots, 0), (-m, -(m-1), \cdots, -1)\},\\ \boldsymbol{\Gamma} &= \{(\tfrac{1}{2}, 1, \cdots, \tfrac{1}{2}m), (a_0, a_1, \cdots, a_m), (0, 0, \cdots, 0)\}.\end{aligned}$$

In order to show that $\mathbf{J}$ has a regular singularity at $\tau = \infty$, set $x = \tau^{-1}$. Then

(6.29) $$\begin{aligned}d\mathbf{J}/dx &= x^{-1}(\mathbf{E}_{m+1} + x\boldsymbol{\Omega})^{-1}\boldsymbol{\Gamma}\mathbf{J}\\ &= \{x^{-1}\boldsymbol{\Gamma} + \textstyle\sum_{r=0}^{m}(-\boldsymbol{\Omega})^{r+1}\boldsymbol{\Gamma}x^r\}\mathbf{J}.\end{aligned}$$

The $(m + 1)$ linearly independent solutions in the neighbourhood of ∞ correspond to the latent roots of $\boldsymbol{\Gamma}$, viz a_r $(r = 0, 1, \cdots, m)$. In view of (6.5), the required $\mathbf{J}$ may be a linear combination of the solutions obtained from $a_0, \cdots, a_{m-1}$, and the problem of determining the coefficients arises as in Section 5.

Turning next to consider the solution near $\tau = 0$, let

(6.30) $$\mathbf{J} = \operatorname{diag}\{1, \tau^{-1}, \cdots, \tau^{-m}\}\mathbf{K}.$$

Then

(6.31) $$d\mathbf{K}/d\tau = \tau^{-2}(\mathbf{D}_0 + \tau\mathbf{D}_1)\mathbf{K},$$

where

(6.32) $$\begin{aligned}\mathbf{D}_0 &= \tfrac{1}{2}(\mathbf{E}_{m+1} + \boldsymbol{\Omega})^{-1}\boldsymbol{\Xi},\\ \boldsymbol{\Xi} &= \{(1, 2, \cdots, m), (0, \cdots, 0), (0, \cdots, 0)\},\\ \mathbf{D}_1 &= +(\mathbf{E}_{m+1} + \boldsymbol{\Omega})^{-1}\{(0, \cdots, 0, \cdots, 0), (a_0, \cdots, a_j - j, \cdots, a_m - m),\\ &\qquad (0, \cdots, (j-1)(m-j+1), \cdots, 0)\}.\end{aligned}$$

The system (6.31) is seen to have a singular point of the second kind (in fact an irregular singularity) at $\tau = 0$ ([1], Chapter 5). In general, such systems have formal solutions which provide asymptotic expansions of the actual solutions for small τ. As in the case of the regular singular points, we first seek the latent roots of the leading matrix $\mathbf{D}_0$. These may be obtained from the determinantal equation

$$\det \{\lambda(\mathbf{E}_{m+1} + \mathbf{\Omega}) - \mathbf{\Xi}\} = 0. \tag{6.33}$$

The left-hand side is a continuant ([11], Chapter XIII), and may be written in centrosymmetric form (loc. cit. Section 549):

$$\det \{(-\lambda^{\frac{1}{2}}, -2\lambda^{\frac{1}{2}}, \cdots, -m\lambda^{\frac{1}{2}}), (\lambda, \lambda, \cdots, \lambda), (-m\lambda^{\frac{1}{2}}, \cdots, -2\lambda^{\frac{1}{2}}, -\lambda^{\frac{1}{2}})\} = 0. \tag{6.34}$$

Hence, a theorem on continuants (loc. cit. Section 576) may be used to evaluate the determinant:

$$\prod_{i=0}^{m} [\lambda - (m - 2i)\lambda^{\frac{1}{2}}] = \lambda^{[\frac{1}{2}m]+1} \prod_{i=0}^{[\frac{1}{2}(m-1)]} [\lambda - (m - 2i)^2] = 0, \tag{6.35}$$

where [ ] denotes the greatest integer part. The matrix $\mathbf{D}_0$ therefore has $[\frac{1}{2}(m + 1)]$ positive latent roots $\frac{1}{2}m^2, \frac{1}{2}(m - 2)^2, \cdots$, and $[\frac{1}{2}m] + 1$ zero roots.

Again, the presence of equal roots makes any discussion of the complete solution of (6.31) extremely difficult (see [14]). However, (6.23) implies that the relevant solution is that corresponding to the largest latent root $\frac{1}{2}m^2$, since this solution approaches zero more rapidly as $\tau \to 0+$ than any other.

Noting that $\mathbf{D}_0$ has rank m, there exists a non-singular matrix $\mathbf{Q} = \{q_{ij}\}$ reducing $\mathbf{D}_0$ to its canonical Jordon form:

$$\begin{aligned} \mathbf{Q}^{-1}\mathbf{D}_0\mathbf{Q} = \{&(0, 0, \cdots, 0, \cdots, 0), \\ &(\tfrac{1}{2}m^2, \tfrac{1}{2}(m - 2)^2, \cdots, \tfrac{1}{2}(m - 2\nu)^2, 0, \cdots, 0), \\ &(0, 0, \cdots, 0, 1, \cdots, 1)\} = \mathbf{\Upsilon}_0' \end{aligned} \tag{6.36}$$

where $\nu = [\frac{1}{2}(m - 1)]$, and there are $[\frac{1}{2}m]$ ones in the lower diagonal.

A suitable set of q_{ij} may be obtained from the following recurrence relations:

$$\begin{aligned} &q_{0j} = 1, \quad (j = 0, \cdots, m); \quad q_{im} = 0, \ (i = 1, \cdots, m); \\ &q_{ij} = i^{-1}(m - 2j)^2[q_{i-1,j} - (m - i + 2)q_{i-2,j}], \\ &\qquad\qquad (i = 1, \cdots, m; j = 0, \cdots, \nu); \\ &q_{ij} = 2i^{-1}[q_{i-1,j+1} - (m - i + 2)q_{i-2,j+1}], \\ &\qquad\qquad (i = 1, \cdots, m; j = \nu + 1, \cdots, m - 1). \end{aligned} \tag{6.37}$$

Writing

$$\mathbf{K} = \mathbf{Q}\mathbf{G}, \tag{6.38}$$

equation (6.31) takes the form

$$d\mathbf{G}/d\tau = \tau^{-2}(\mathbf{\Upsilon}_0 + \tau\mathbf{\Upsilon}_1)\mathbf{G}, \tag{6.39}$$

where

(6.40) $$\Upsilon_1 = Q^{-1}D_1Q = (v_{ij}).$$

We now seek a formal solution

(6.41) $$G = \beta(m; n_2)e^{-m^2/2\tau}\tau^{\delta} \sum_{r=0}^{\infty} X_r\tau^r, \qquad (X_{00} = 1),$$

where the constant $\beta(m; n_2)$ is given by (6.23). The $X_r = (X_{0r}, \cdots, X_{mr})'$ and δ are to be determined. Substitution in (6.39) yields:

(6.42) $$X_0 = (1, 0, \cdots, 0)'; \qquad (v_{00} + \delta)\, X_{00} = 0;$$

while for $r = 1, 2, \cdots$:

(6.43)
$$X_{ir} = \tfrac{1}{4}A_{i,r-1}/i(\tfrac{1}{2}m - i), \qquad (i = 1, \cdots, \nu);$$
$$X_{\nu+1,r} = m^{-2}A_{\nu+1,r-1};$$
$$X_{ir} = m^{-2}(A_{i,r-1} + 2X_{i-1,r}), \qquad (i = \nu + 2, \cdots, m);$$
$$(v_{00} + r + \lambda)X_{0r} = -\sum_{j=1}^{m} v_{0j}X_{jr},$$

where

(6.44) $$A_{i,r} = -2[\sum_{j=0}^{m} v_{ij}X_{jr} + (r + \delta)X_{ir}], \quad (i = 1, \cdots, m; r = 0, 1, \cdots).$$

From (6.42),

(6.45) $$\delta = -v_{00}$$

so that the last relation in (6.43) becomes

(6.46) $$X_{0r} = -r^{-1} \sum_{j=1}^{m} v_{0j}X_{jr}.$$

The author has not succeeded in proving in general that

(6.47) $$v_{00} = \tfrac{1}{2}mn_2 - \tfrac{1}{4}(m - 2)(m + 3)$$

as required by (6.23) and (6.45). However, this has been verified for some early values of m.

It has thus been shown that a formal solution (6.41) of (6.39) exists, and we conjecture that this provides an asymptotic expansion of the required solution as $\tau \to 0+$. The cdf, $\Theta(\tau)$ say, is given by an expansion

(6.48) $$\Theta(\tau) \sim \beta(m; n_2)e^{-m^2/2\tau}\tau^{\delta+2} \sum_{r=0}^{\infty} U_r\tau^r,$$

where the U_r may be obtained from the relation

(6.49) $$d\Theta/d\tau = \theta(\tau) = \sum_{j=0}^{m} G_j, \quad (G = (G_0, \cdots, G_m)').$$

An asymptotic development of $f(T)$ for large n_1 based on $\theta(\tau)$ would clearly present a difficult problem, and will not be attempted here.

7. The moments of T. Constantine has shown ([3], Section 5) that the moments of T exist up to the jth, where j is the largest integer such that $j < \frac{1}{2}(n_2 - m + 1)$, and he has obtained expressions for these moments in terms of zonal polynomials. In the present section it is shown that recurrence relations for the moments may be derived from the differential equations (2.19–20) for the Laplace transforms $L_r(s)$. These equations may be written:

$$dL/ds = [s^{-1}B^* + A]L, \tag{7.1}$$

where **A** was defined in (2.29) and

$$B^* = \{(-b_0, \cdots, -b_{m-1}), (a_0^*, \cdots, a_m^*), (0, \cdots, 0)\}, \tag{7.2}$$
$$a_r^* = a_r - 1 = \tfrac{1}{2}(m - r)(n_2 - r).$$

Since $L(0) = (1, \cdots)'$ and $a_m^* = 0$ is the smallest root of B^*, it follows that **L** may be a linear combination of the independent solutions corresponding to the latent roots a_r^* $(r = 0, \cdots, m)$ (which differ by integers in the same manner as the a_r (Section 5)). In any case, however,

$$L(s) = \textstyle\sum_{r=0}^{j^*} l_r s^r + o(s^{j^*}), \quad (l_r = (l_{0r}, \cdots, l_{mr})') \tag{7.3}$$

where j^* is the largest integer such that

$$j^* < a_{m-1}^* = \tfrac{1}{2}(n_2 - m + 1). \tag{7.4}$$

Thus $j^* = j$, and all moments of T up to the jth exist by a standard result on Laplace transforms, in agreement with Constantine's result.

The matrix B^* may be reduced to diagonal form by the transformation

$$L = \Pi L^*, \tag{7.5}$$

where

$$\Pi = \{\pi_{ik}\}, \qquad \pi_{ik} = \binom{k}{i}(n_1 + n_2 - i)![(m + n_2 - i - k)!]^{-1} \tag{7.6}$$

It may also be shown that

$$\Pi^{-1} = \{\pi_{ik}^*\}, \qquad \pi_{ik}^* = (-1)^{i+k}\binom{k}{i} \cdot (m + n_2 - i - k - 1)![(n_1 + n_2 - k)!]^{-1}(m + n_2 - 2i), \tag{7.7}$$

and that

$$\Pi^{-1}A\Pi = \Sigma = \{(\lambda_0^*, \cdots, \lambda_{m-1}^*), (\mu_0^*, \cdots, \mu_m^*), (\nu_1^*, \cdots, \nu_m^*)\}, \tag{7.8}$$

where

$$\lambda_r^* = (r + 1)(n_2 - r - 1)(m + n_2 - r)(n_1 - m + 1 + r) \cdot [(m + n_2 - 2r - 2)(m + n_2 - 2r - 1)]^{-1},$$
$$\mu_r^* = [-r(m + n_2 - r)(m + 2n_1 + n_2 + 1) + m(n_1 + n_2)(m + n_2 + 1)] \cdot [(m + n_2 - 2r - 1)(m + n_2 - 2r + 1)]^{-1}, \tag{7.9}$$
$$\nu_r^* = -(m - r + 1)(n_1 + n_2 - r + 1)[(m + n_2 - 2r + 1) \cdot (m + n_2 - 2r + 2)]^{-1}.$$

The differential equation (7.1) reduces to

$$dL^*/ds = [s^{-1} \operatorname{diag}\{a_0^*, \cdots, a_m^*\} + \Sigma]L^*. \tag{7.10}$$

Taking

$$\mathbf{L}^* = \sum_{r=0}^{j} \mathbf{l}_r^* s^r + \mathbf{o}(s^j), \tag{7.11}$$

the following recurrence relations are obtained for the $\mathbf{l}_r^* = (l_{0r}^*, \cdots, l_{mr}^*)'$

$$\begin{aligned} \underset{\sim}{l}_0^* &= n_2![(n_1+n_2)!]^{-1}(0, \cdots, 0, 1)', \\ \underset{\sim}{l}_r^* &= \operatorname{diag}\{(r-a_0^*)^{-1}, \cdots, (r-a_m^*)^{-1}\}\Sigma \underset{\sim}{l}_{r-1}^*, \qquad (r = 1, \cdots, j). \end{aligned} \tag{7.12}$$

The moments of T are then given by

$$\mathcal{E}(T^r) = (-1)^r r!\, l_{0r} = (-1)^r r!(n_1+n_2)! \sum_{k=0}^{m} l_{kr}^*/(m+n_2-k)!. \tag{7.13}$$

In particular, taking $r = 1$,

$$\mathbf{l}_1^* = n_2![(n_1+n_2)!]^{-1}(0, \cdots, 0, -2\lambda_{m-1}^*/(n_2-m-1), \mu_m^*)', \tag{7.14}$$

whence it is easily found that

$$\mathcal{E}(T) = mn_1/(n_2-m-1). \tag{7.15}$$

(Constantine, loc. cit).

8. Acknowledgment. The author wishes to thank Professor A. T. James for his helpful suggestions during the preparation of this paper.

REFERENCES

[1] Coddington, E. A. and Levinson, N. (1955). *Theory of Ordinary Differential Equations*. McGraw-Hill, New York.

[2] Constantine, A. G. (1963). Some non-central distribution problems in multivariate analysis. *Ann. Math. Statist.* **34** 1270–1285.

[3] Constantine, A. G. (1966). The distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **37** 215–225.

[4] Evgrafov, M. A. (1961). *Asymptotic Estimates and Entire Functions*. Gordon and Breach, New York.

[5] Fisher, R. A. (1939). The sampling distribution of some statistics obtained from non-linear equations. *Ann. Eugenics* **9** 238–249.

[6] Hill, G. W. and Davis, A. W. (1968). Generalized asymptotic expansions of Cornish-Fisher type. To appear in *Ann. Math. Statist.* **39,** No. 4.

[7] Hotelling, H. (1951). A generalized T-test and measure of multivariate dispersion. *Proc. Second Berkeley Symp. Math. Statist. Prob.* 23–42. Univ. of California Press.

[8] Hsu, P. L. (1939). On the distribution of the roots of certain determinantal equations. *Ann. Eugenics* **9** 250–258.

[9] Ito, K. (1956). Asymptotic formulae for the distribution of Hotelling's generalized T_0^2 statistic. *Ann. Math. Statist.* **27** 1091–1105.

[10] James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475–501.

[11] Muir, T. and Metzler, W. H. (1930). *A Treatise on the Theory of Determuinants*. Albany, New York.

[12] Pillai, K. C. S. and Pablo Samson, Jr. (1959). On Hotelling's generalization of T^2. *Biometrika* **46** 160–168.

[13] Roy, S. N. (1939). p-statistics, or some generalizations in analysis of variance appropriate to multivariate problems. *Sankhyā* **4** 381–396.

[14] Turrittin, H. L. (1955). Convergent solutions of ordinary linear homogeneous equations in the neighborhood of an irregular singular point. *Acta Math.* **93** 27–66.

Biometrika (1970), **57**, 1, *p.* 187
With 1 *text-figure*
Printed in Great Britain

Exact distributions of Hotelling's generalized T_0^2

By A. W. DAVIS
C.S.I.R.O., Adelaide

Summary

Percentile tables are given for the null distribution of Hotelling's generalized T_0^2 statistic. These are obtained by analytic continuation of Constantine's series using a system of linear differential equations. The accuracy of certain approximations given by Pillai and Ito is discussed.

1. Introduction

Hotelling's (1951) T_0^2 statistic is defined by

$$T_0^2 = n_2 \operatorname{tr}(\mathbf{S}_1 \mathbf{S}_2^{-1}) = n_2 T,$$

say, where $\mathbf{S}_1$ and $\mathbf{S}_2$ are independent $m \times m$ Wishart matrices on n_1 and n_2 degrees of freedom, respectively, estimating the same covariance matrix, with $n_2 \geqslant m$. The exact null distribution when $m = 2$ was given by Hotelling in terms of the Gaussian hypergeometric function, and tabulated by F. E. Grubbs in 1954, in an unpublished report of Aberdeen Proving Ground, Maryland. In the general null case, the density function of T has been shown by the present author (1968) to satisfy a linear homogeneous differential equation of order m, with regular singularities at $T = 0, -1, \ldots, -m$ and infinity. Constantine's (1966) series reduces in this case to the relevant solution of the differential equation in the unit circle about $T = 0$.

Table 1 presents accurate percentiles of $T_0^2/n_1 = (n_2/n_1)T$, for $m = 3$ and 4 obtained by using an equivalent system of first order differential equations to carry out an analytic continuation of Constantine's series along the positive real axis. The accuracy of the procedure was checked by mapping the differential equation on to the unit interval using the change of variable $Y = T/(T+1)$. Agreement was generally to at least five significant figures. A similar tabulation has been carried out for $m = 5$.

2. Discussion of approximations

Figure 1 presents the results of comparing the accurate percentiles with Ito's (1956) percentile approximation to $O(n_2^{-2})$ and Pillai & Samson's (1959) Pearson-curve approximation. The author has extended Ito's formula to $O(n_2^{-3})$, and reports the following reduced form of the Pearson parameter β_2:

$$\beta_2 = \frac{3(c-2)(c+1)A}{mn_1(n_2-1)(c-6)(c-4)(c-1)(c+2)(c+3)(c+n_1)},$$

where

$$c = n_2 - m - 1,$$

$$A = n_1(c+n_1)\{m(n_2-1)(c^3-5c^2+78c-72)+4c^2(5c-6)\} + 4c^2\{m(n_2-1)(5c-6)+c(c^2-c+2)\}.$$

Table 1. *Upper percentage points for Hotelling's generalized* T_0^2/n_1

For $n_1 < m$, make the transformations $n_1 \to m$, $n_2 \to n_1 + n_2 - m$, $m \to n_1$ before calculating T_0^2/n_1.

$m = 3$	$n_2 \backslash n_1$	3	4	5	6	8	10	12	15	20	25	40	60
5%	3	25·930*	26·996*	27·665*	28·125*	28·712*	29·073*	29·316*	29·561*	29·809*	29·959*	30·19*	30·31*
	4	1·1880*	1·1929*	1·1959*	1·1978*	1·2003*	1·2018*	1·2028*	1·2038*	1·2048*	1·2054*	1·2063*	1·2068*
	5	42·474	41·764	41·305	40·983	40·562	40·300	40·120	39·937	39·750	39·635	39·462	39·366
	6	25·456	24·715	24·235	23·899	23·458	23·182	22·992	22·799	22·600	22·479	22·294	22·190
	7	18·752	18·056	17·605	17·288	16·870	16·608	16·427	16·241	16·051	15·934	15·755	15·653
	8	15·308	14·657	14·233	13·934	13·540	13·290	13·118	12·941	12·758	12·646	12·473	12·375
	10	11·893	11·306	10·921	10·649	10·287	10·057	9·8974	9·7320	9·5603	9·4541	9·2897	9·1955
	12	10·229	9·6825	9·3234	9·0680	8·7271	8·5088	8·3566	8·1982	8·0330	7·9301	7·7700	7·6777
	14	9·2550	8·7356	8·3935	8·1495	7·8225	7·6122	7·4649	7·3110	7·1497	7·0488	6·8908	6·7991
	16	8·6180	8·1183	7·7884	7·5526	7·2355	7·0307	6·8868	6·7360	6·5772	6·4774	6·3204	6·2287
	18	8·1701	7·6851	7·3644	7·1347	6·8251	6·6244	6·4830	6·3343	6·1771	6·0780	5·9212	5·8292
	20	7·8384	7·3649	7·0513	6·8263	6·5224	6·3249	6·1853	6·0383	5·8822	5·7834	5·6266	5·5341
	25	7·2943	6·8407	6·5394	6·3227	6·0287	5·8365	5·7001	5·5555	5·4010	5·3025	5·1446	5·0503
	30	6·9654	6·5245	6·2311	6·0196	5·7319	5·5431	5·4085	5·2654	5·1116	5·0129	4·8535	4·7575
	35	6·7453	6·3132	6·0253	5·8175	5·5341	5·3476	5·2143	5·0720	4·9185	4·8195	4·6586	4·5608
	40	6·5877	6·1621	5·8783	5·6732	5·3929	5·2081	5·0757	4·9340	4·7806	4·6813	4·5189	4·4195
	50	6·3773	5·9606	5·6823	5·4809	5·2050	5·0224	4·8911	4·7502	4·5967	4·4968	4·3319	4·2297
	60	6·2433	5·8324	5·5577	5·3587	5·0856	4·9044	4·7739	4·6334	4·4798	4·3793	4·2123	4·1078
	70	6·1504	5·7436	5·4715	5·2742	5·0031	4·8229	4·6929	4·5526	4·3988	4·2979	4·1292	4·0227
	80	6·0823	5·6786	5·4084	5·2122	4·9426	4·7632	4·6336	4·4935	4·3395	4·2381	4·0680	3·9600
	100	5·9891	5·5896	5·3220	5·1276	4·8601	4·6817	4·5525	4·4126	4·2583	4·1563	3·9840	3·8734
	200	5·8099	5·4186	5·1562	4·9653	4·7017	4·5252	4·3970	4·2574	4·1023	3·9988	3·8212	3·7042
	∞	5·6397	5·2565	4·9992	4·8116	4·5519	4·3773	4·2499	4·1104	3·9541	3·8487	3·6642	3·5384
1%	3	6·4845†	6·7500†	6·9169†	7·0313†	7·1778†	7·2675†	7·3281†	7·3891†	7·4511†	7·4883†	—	—
	4	5·9896*	5·9946*	5·9976*	5·9996*	6·0021*	6·0035*	6·0046*	6·0056*	6·0067*	6·0071*	6·008*	6·008*
	5	1·2738*	1·2420*	1·2219*	1·2080*	1·1901*	1·1790*	1·1715*	1·1638*	1·1561*	1·1514*	1·144*	1·141*
	6	59·507	57·032	55·462	54·377	52·973	52·102	51·509	50·906	50·292	49·918	49·349	49·04
	7	37·994	35·993	34·721	33·840	32·695	31·984	31·498	31·002	30·496	30·188	29·718	29·452
	8	28·308	26·599	25·511	24·755	23·771	23·157	22·737	22·308	21·868	21·599	21·188	20·955
	10	19·737	18·355	17·471	16·855	16·050	15·544	15·197	14·840	14·472	14·246	13·899	13·702
	12	15·973	14·765	13·990	13·448	12·737	12·288	11·978	11·659	11·328	11·124	10·809	10·628
	14	13·905	12·803	12·096	11·599	10·945	10·530	10·243	9·9462	9·6377	9·4463	9·1490	8·9780
	16	12·610	11·581	10·918	10·452	9·8359	9·4444	9·1724	8·8900	8·5955	8·4121	8·1260	7·9605
	18	11·729	10·751	10·120	9·6756	9·0870	8·7117	8·4503	8·1782	7·8934	7·7154	7·4365	7·2743
	20	11·091	10·152	9·5452	9·1173	8·5492	8·1861	7·9325	7·6679	7·3901	7·2159	6·9419	6·7818
	25	10·075	9·2005	8·6339	8·2333	7·6992	7·3560	7·1152	6·8627	6·5958	6·4273	6·1598	6·0019
	30	9·4785	8·6441	8·1022	7·7183	7·2050	6·8739	6·6407	6·3953	6·1346	5·9690	5·7042	5·5464
	35	9·0874	8·2798	7·7548	7·3822	6·8829	6·5598	6·3317	6·0909	5·8339	5·6700	5·4063	5·2478
	40	8·8113	8·0233	7·5108	7·1460	6·6564	6·3392	6·1147	5·8771	5·6227	5·4598	5·1962	5·0367
	50	8·4479	7·6861	7·1894	6·8358	6·3599	6·0503	5·8305	5·5970	5·3457	5·1838	4·9196	4·7578
	60	8·2195	7·4745	6·9882	6·6416	6·1744	5·8696	5·6528	5·4218	5·1722	5·0108	4·7455	4·5815
	70	8·0627	7·3295	6·8504	6·5087	6·0474	5·7460	5·5312	5·3019	5·0535	4·8922	4·6258	4·4598
	80	7·9485	7·2239	6·7502	6·4120	5·9551	5·6562	5·4428	5·2147	4·9670	4·8058	4·5383	4·3706
	100	7·7932	7·0805	6·6141	6·2809	5·8300	5·5344	5·3230	5·0965	4·8497	4·6883	4·4190	4·2484
	200	7·4980	6·8083	6·3561	6·0323	5·5930	5·3037	5·0961	4·8725	4·6270	4·4650	4·1906	4·0124
	∞	7·2220	6·5542	6·1156	5·8009	5·3725	5·0892	4·8849	4·6638	4·4190	4·2557	3·9738	3·7843

TABLE 1 (cont.)

$m = 4$	$n_2 \backslash n_1$	4	5	6	8	10	12	15	20	25	40	60
5 %	4	49·964*	51·204*	52·054*	53·142*	53·808*	54·258*	54·71*	55·17*	55·46*	—	—
	5	1·9964*	2·0013*	2·0046*	2·0087*	2·0112*	2·0128*	2·0145*	2·0161*	2·0171*	2·019*	—
	6	65·715	64·999	64·497	63·841	63·432	63·151	62·866	62·573	62·396	62·13	—
	7	37·343	36·629	36·129	35·474	35·064	34·782	34·495	34·200	34·019	33·75	—
	8	26·516	25·868	25·413	24·814	24·437	24·178	23·912	23·639	23·471	23·214	23·072
	10	17·875	17·326	16·938	16·424	16·098	15·872	15·640	15·399	15·250	15·021	14·891
	12	14·338	13·848	13·500	13·037	12·741	12·535	12·321	12·099	11·961	11·747	11·624
	14	12·455	12·002	11·680	11·248	10·972	10·778	10·577	10·366	10·234	10·029	9·9103
	16	11·295	10·868	10·563	10·154	9·8904	9·7054	9·5119	9·3085	9·1810	8·9808	8·8644
	18	10·512	10·104	9·8121	9·4190	9·1647	8·9857	8·7978	8·5996	8·4748	8·2778	8·1626
	20	9·9500	9·5560	9·2736	8·8926	8·6453	8·4708	8·2871	8·0926	7·9696	7·7748	7·6601
	25	9·0585	8·6884	8·4223	8·0616	7·8261	7·6590	7·4821	7·2933	7·1730	6·9805	6·8659
	30	8·5377	8·1825	7·9265	7·5784	7·3502	7·1876	7·0147	6·8291	6·7101	6·5181	6·4026
	35	8·1968	7·8517	7·6026	7·2631	7·0397	6·8801	6·7099	6·5262	6·4079	6·2156	6·0989
	40	7·9566	7·6188	7·3746	7·0413	6·8214	6·6640	6·4955	6·3131	6·1952	6·0023	5·8844
	50	7·6404	7·3125	7·0751	6·7501	6·5350	6·3804	6·2143	6·0334	5·9157	5·7214	5·6011
	60	7·4417	7·1202	6·8872	6·5676	6·3555	6·2027	6·0381	5·8581	5·7403	5·5446	5·4222
	70	7·3054	6·9884	6·7584	6·4426	6·2325	6·0809	5·9173	5·7378	5·6200	5·4230	5·2987
	80	7·2061	6·8924	6·6646	6·3515	6·1430	5·9924	5·8294	5·6503	5·5323	5·3343	5·2084
	100	7·0711	6·7619	6·5372	6·2279	6·0215	5·8721	5·7101	5·5313	5·4131	5·2133	5·0849
	200	6·8143	6·5139	6·2952	5·9933	5·7910	5·6439	5·4836	5·3053	5·1863	4·9819	4·8471
	∞	6·5741	6·2821	6·0692	5·7743	5·5758	5·4309	5·2721	5·0940	4·9737	4·7629	4·6190
1 %	4	12·491†	12·800†	13·012†	13·283†	13·449†	13·561†	13·67†	13·79†	13·87†	—	—
	5	9·9992*	10·004*	10·008*	10·012*	10·014*	10·016*	10·018*	10·02*	10·02*	—	—
	6	1·9377*	1·9064*	1·8848*	1·8570*	1·8398*	1·8281*	1·8162*	1·8041*	1·7969*	—	—
	7	85·053	82·731	81·125	79·047	77·759	76·882	75·989	75·082	74·522	—	—
	8	51·991	50·178	48·921	47·290	46·276	45·583	44·877	44·156	43·715	43·04	—
	10	29·789	28·478	27·566	26·376	25·632	25·121	24·597	24·060	23·731	23·224	22·95
	12	21·965	20·889	20·138	19·154	18·534	18·108	17·668	17·215	16·936	16·505	16·261
	14	18·142	17·199	16·539	15·670	15·121	14·742	14·349	13·943	13·691	13·301	13·077
	16	15·916	15·059	14·457	13·662	13·157	12·807	12·444	12·066	11·831	11·466	11·255
	18	14·473	13·674	13·112	12·368	11·894	11·564	11·221	10·863	10·639	10·289	10·086
	20	13·466	12·710	12·177	11·470	11·018	10·703	10·374	10·030	9·8138	9·4748	9·2771
	25	11·924	11·237	10·751	10·103	9·6871	9·3951	9·0890	8·7658	8·5618	8·2383	8·0476
	30	11·055	10·409	9·9509	9·3382	8·9430	8·6646	8·3715	8·0602	7·8626	7·5468	7·3586
	35	10·499	9·8801	9·4405	8·8511	8·4695	8·2000	7·9153	7·6115	7·4177	7·1059	6·9186
	40	10·114	9·5138	9·0872	8·5142	8·1424	7·8791	7·6002	7·3015	7·1102	6·8006	6·6132
	50	9·6141	9·0400	8·6308	8·0795	7·7204	7·4652	7·1938	6·9015	6·7131	6·4054	6·2168
	60	9·3053	8·7472	8·3490	7·8114	7·4603	7·2101	6·9434	6·6548	6·4679	6·1606	5·9704
	70	9·0954	8·5485	8·1578	7·6297	7·2840	7·0373	6·7736	6·4875	6·3015	5·9940	5·8022
	80	8·9437	8·4048	8·0197	7·4984	7·1567	6·9124	6·6510	6·3666	6·1812	5·8733	5·6799
	100	8·7388	8·2111	7·8334	7·3214	6·9851	6·7443	6·4858	6·2036	6·0189	5·7100	5·5139
	200	8·3542	7·8476	7·4842	6·9900	6·6639	6·4293	6·1763	5·8979	5·7138	5·4012	5·1976
	∞	8·0000	7·5132	7·1633	6·6857	6·3691	6·1402	5·8920	5·6164	5·4323	5·1133	4·8981

* Multiply entry by 100.   † Multiply entry by 10^4.

In the regions below the curves, the corresponding approximations achieve three-decimal accuracy. The Pearson curve method is seen to be applicable over a remarkably wide range of parameter values, both for increasing n_1 and n_2. In the shaded regions of Fig. 1, the error

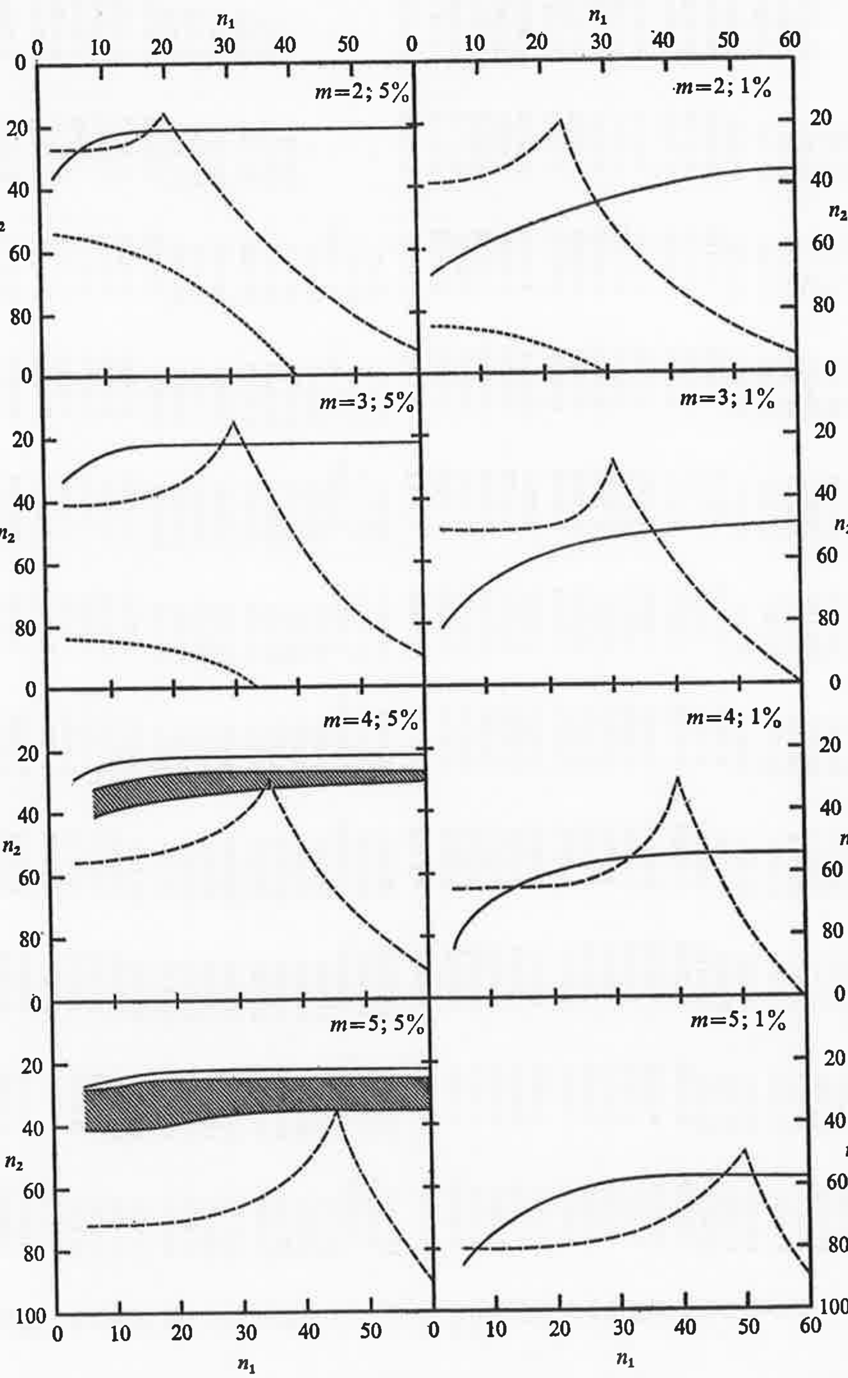

Fig. 1. Regions of three-decimal accuracy for approximations to T_0^2/n_1. ——, Pillai; – – –, Ito (n_2^{-3}); ……, Ito (n_2^{-2}). In the shaded regions Pillai's approximation is not accurate to three decimal places.

of the Pillai approximation increases slightly after three-decimal accuracy has been attained for small n_2, and then decreases again as n_2 increases. This effect becomes more pronounced for increasing m.

The accuracy of the Ito-type expansions falls off sharply for large n_1, but nevertheless these provide useful direct formulae. The cusps occur because for fixed n_2 the error in the $O(n_2^{-3})$ approximation changes sign from negative to positive as n_1 increases, briefly producing a sharp rise in accuracy.

The author's grateful thanks are due to Miss J. Hawkes for preparing the tables, and to Mr Murray Childs for drawing the diagrams.

References

Constantine, A. G. (1966). The distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **37**, 215–25.

Davis, A. W. (1968). A system of linear differential equations for the distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **39**, 815–32.

Hotelling, H. (1951). A generalized T-test and measure of multivariate dispersion. *Proc. 2nd Berkeley Symp.*, pp. 23–42.

Ito, K. (1956). Asymptotic formulae for the distribution of Hotelling's generalized T_0^2 statistic. *Ann. Math. Statist.* **27**, 1091–105.

Pillai, K. C. S. & Samson, P. (1959). On Hotelling's generalization of T^2. *Biometrika* **46**, 160–8.

[*Received May* 1969. *Revised July* 1969]

[Reprinted from ANNALS OF THE INSTITUTE OF STATISTICAL MATHEMATICS, Vol. 22, No. 1, 1970]

FURTHER APPLICATIONS OF A DIFFERENTIAL EQUATION FOR HOTELLING'S GENERALIZED T_0^2

A. W. DAVIS

FURTHER APPLICATIONS OF A DIFFERENTIAL EQUATION FOR HOTELLING'S GENERALIZED T_0^2 *

A. W. DAVIS

(Received Oct. 22, 1969)

1. Summary

The null distribution of the Hotelling-Lawley generalized T_0^2 statistic has been shown [4] to satisfy a homogeneous linear differential equation (d.e.). The latter has been used to tabulate some exact percentage points of T_0^2 by analytic continuation of Constantine's [3] series, and a table for the 5-variate case is presented in this paper. The Ito-Siotani [9], [16] asymptotic expansions for the distribution function and percentage points of T_0^2 are also extended.

2. The differential equation

The T_0^2 statistic of Hotelling [8] and Lawley [11] is defined by

$$T_0^2 = n_2 \operatorname{tr} S_1 S_2^{-1} = n_2 T\,, \tag{2.1}$$

say, where S_1 and S_2 are independent $m\times m$ Wishart matrices based on n_1 and n_2 degrees of freedom, respectively, estimating the same population covariance matrix. In the general case when S_1 has a non-central Wishart distribution, Constantine [3] has given a power-series representation of the density function $f(T)$ of T, which is valid for $0<T<1$. The exact null distribution when $m=2$ was found by Hotelling [8] in terms of the Gaussian hypergeometric function. This result was extended by the present author [4], who showed that $f(T)$ satisfies a d.e. of fuchsian type of order m, with regular singularities at $T=0, -1, -2, \cdots, -m$ and ∞. Constantine's series reduces in the null case to the relevant solution of this d.e. in the unit circle about $T=0$.

When $n_1, n_2 \geqq m$, the d.e. for $f(T)$ is equivalent to a first order matrix d.e. ([4], Section 3):

* This work was partly supported by The National Institutes of Health Grant GM-12868-05.

$$\text{diag}\{T, T+1, \cdots, T+m\}dM/dT = CM , \tag{2.2}$$

where

$$M(T) = (M_0(T), \cdots, M_m(T))' \tag{2.3}$$

is an $(m+1)$-vector with

$$M_0(T) \equiv f(T) , \tag{2.4}$$

$$\sum_{j=0}^{m} (T+j)M_j(T) \equiv 0 . \tag{2.5}$$

The constant $(m+1)\times(m+1)$ matrix C has the form

$$C = \begin{bmatrix} \alpha_0, \beta_0, 0, 0, \cdots\cdots, 0 \\ \gamma_1, \alpha_1, \beta_1, 0, \cdots\cdots, 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \qquad\qquad \gamma_{m-1}, \alpha_{m-1}, \beta_{m-1} \\ 0 \cdots\cdots 0, \quad \gamma_m, \quad \alpha_m \end{bmatrix} \tag{2.6}$$

$$= \{(\beta_0, \beta_1, \cdots, \beta_{m-1}), (\alpha_0, \cdots, \alpha_m), (\gamma_1, \cdots, \gamma_m)\} ,$$

say, where

$$\begin{aligned} \alpha_i &= [(m-2i)n_1 - in_2 + (2i^2 - mi - i - 2)]/2 , \\ \beta_i &= (i+1)(n_1+n_2-i)/2 , \\ \gamma_i &= -(m-i+1)(n_1-i+1)/2 . \end{aligned} \tag{2.7}$$

Since (2.2) may be rewritten as

$$dM/dT = \left\{ \sum_{r=0}^{m} (T+r)^{-1} V_r \right\} M , \tag{2.8}$$

where V_r is obtained by replacing all elements of C by zeros except those in the rth row $(r=0, 1, \cdots, m)$, it follows from the general theory of such systems ([2], Chapter 4) that the d.e. has regular singularities at $T=0, -1, \cdots, -m$ and ∞. The $(m+1)$ linearly independent solutions of (2.2) in the unit circle about $T=0$ correspond to the latent roots of V_0, i.e. zero (with multiplicity m) and

$$a_0 = mn_1/2 - 1 . \tag{2.9}$$

In virtue of Constantine's result [3], the relevant solution is given by the latter root:

$$M(T)=k(m\,;n_1,n_2)T^{mn_1/2-1}\sum_{j=0}^{\infty}W_jT^j\,,$$
(2.10)
$$(W_j=(W_{0j},\cdots,W_{mj})'\,;\ W_{00}=1,\ |T|<1)\,.$$

Here the $M_0(T)$ component is the null case of Constantine's series, and ([3], equation (2)),

(2.11) $$k(m\,;n_1,n_2)=\Gamma_m((n_1+n_2)/2)/\Gamma(mn_1/2)\Gamma_m(n_2/2)\,,$$

where

(2.12) $$\Gamma_m(z)=\pi^{m(m-1)/4}\prod_{i=0}^{m-1}\Gamma(z-i/2)\,.$$

The following recurrence relations for the W_{ij} are obtained from (2.2):

$$W_{i0}=\delta_{i0}\,,\qquad \text{(Kronecker's delta)}\,;$$
(2.13)
$$i(j+mn_1/2-1)W_{ij}=\gamma_iW_{i-1,j}+[\alpha_i-(j+mn_1/2-2)]W_{i,j-1}+\beta_iW_{i+1,j-1}\,,\qquad (i=1,\cdots,m\,;\ j=1,2,\cdots)\,;$$
$$jW_{0j}=(n_1+n_2)W_{1j}/2\,.$$

The d.e. (2.2) may be used to carry out an analytic continuation of the solution (2.10) along the positive T axis, noting that at any "ordinary" point $T^*>0$ it is sufficient to know $M(T^*)=(M_0(T^*),\cdots,M_m(T^*))'$ in order to construct the solution in the neighbourhood of T^*. A computer program has been written which effects this analytic continuation, and calculates percentage points of T_0^2/n_1 by the Newton-Raphson method. Table 1 presents results for $m=5$. Brief tables for $m=3$ and 4 will be reported elsewhere [5]. For $n_1<m$, the distribution of T is obtained by the transformations [3]

(2.14) $$n_1\to m\,,\qquad m\to n_1\,,\qquad n_2\to n_1+n_2-m\,,$$

so that Table 1 may also be used for $n_1=5$, and values of m between 6 and 60.

As a check on the accuracy of the program, the range of T was mapped onto the unit interval (0, 1) by $Y=T/(T+1)$, and the percentage points recalculated using the transformed d.e.. For $m=2, 3, 4$ and 5, $n_1\leqq 60$ and $n_2\geqq m$, the results agreed to five significant figures, except for small n_2 and large n_1. However, even for such relatively extreme values of the parameters as $m=5$, $n_1=40$, $n_2=20$, the values arrived at (a) by analytic continuation of Constantine's series and (b) after transformation onto (0, 1), were

Table 1. Upper percentage points for

	$n_2 \backslash n_1$	5	6	8	10
5%	5	81.991 +	83.352 +	85.093 +	86.160 +
	6	3.0093+	3.0142+	3.0204+	3.0241+
	7	93.762	93.042	92.102	91.515
	8	51.339	50.646	49.739	49.170
	10	27.667	27.115	26.387	25.927
	12	20.169	19.701	19.079	18.683
	14	16.643	16.224	15.666	15.309
	16	14.624	14.239	13.722	13.389
	18	13.326	12.963	12.476	12.161
	20	12.424	12.078	11.612	11.310
	25	11.046	10.728	10.297	10.016
	30	10.270	9.9689	9.5592	9.2907
	35	9.7739	9.4836	9.0879	8.8277
	40	9.4292	9.1469	8.7613	8.5070
	50	8.9825	8.7107	8.3385	8.0921
	60	8.7057	8.4406	8.0769	7.8355
	70	8.5174	8.2570	7.8991	7.6612
	80	8.3811	8.1241	7.7705	7.5351
	100	8.1969	7.9446	7.5969	7.3649
	200	7.8505	7.6070	7.2706	7.0451
	∞	7.5305	7.2955	6.9698	6.7505
1%	5	20.495 *	20.834 *	21.267 *	21.53 *
	6	15.014 +	15.019 +	15.025 +	15.029 +
	7	2.7354+	2.7045+	2.6646+	2.6400+
	8	1.1498+	1.1276+	1.0989+	1.0811+
	10	48.048	46.670	44.877	43.758
	12	31.108	30.065	28.701	27.846
	14	24.016	23.145	22.001	21.279
	16	20.240	19.472	18.459	17.817
	18	17.929	17.228	16.302	15.713
	20	16.380	15.727	14.862	14.310
	25	14.107	13.529	12.759	12.265
	30	12.880	12.345	11.629	11.167
	35	12.115	11.607	10.926	10.486
	40	11.593	11.105	10.448	10.022
	50	10.928	10.465	9.8408	9.4336
	60	10.523	10.076	9.4712	9.0758
	70	10.251	9.8142	9.2229	8.8354
	80	10.055	9.6261	9.0446	8.6629
	100	9.7929	9.3742	8.8058	8.4319
	200	9.3055	8.9065	8.3629	8.0036
	∞	8.8628	8.4820	7.9613	7.6154

+ Multiply entry by 100. * Multiply entry by 10^4.

	(a)	(b)
Upper 5% point	10.25171	10.25169
Upper 1% point	12.43142	12.43134.

The entries given in Table 1 are 10.252 and 12.431, respectively. It may be noted that the null distribution of Pillai's trace [13]

$$V = \operatorname{tr} S_1(S_1+S_2)^{-1} \tag{2.15}$$

satisfies a d.e. obtained from (2.2) by the transformations

Hotelling's generalized T_0^2/n_1, $(m=5)$

12	15	20	25	40	60
86.88 +	—	—	—	—	—
3.0266+	3.0291+	3.032 +	—	—	—
91.113	90.705	90.29	90.04	—	—
48.780	48.382	47.973	47.723	47.35	—
25.610	25.284	24.947	24.740	24.422	—
18.409	18.124	17.830	17.647	17.365	17.20
15.059	14.800	14.530	14.361	14.100	13.95
13.157	12.914	12.659	12.499	12.250	12.105
11.939	11.708	11.463	11.310	11.068	10.928
11.097	10.874	10.637	10.488	10.252	10.113
9.8168	9.6061	9.3814	9.2386	9.0102	8.8745
9.0995	8.8964	8.6785	8.5389	8.3141	8.1790
8.6419	8.4437	8.2301	8.0926	7.8693	7.7339
8.3250	8.1303	7.9195	7.7833	7.5607	7.4247
7.9150	7.7248	7.5177	7.3829	7.1605	7.0229
7.6615	7.4741	7.2692	7.1351	6.9124	6.7730
7.4894	7.3039	7.1004	6.9667	6.7434	6.6024
7.3648	7.1807	6.9782	6.8448	6.6208	6.4785
7.1968	7.0145	6.8133	6.6801	6.4550	6.3103
6.8811	6.7023	6.5032	6.3702	6.1416	5.9908
6.5902	6.4144	6.2171	6.0838	5.8499	5.6899
—	—	—	—	—	—
15.033 +	15.03 +	15.06 +	—	—	—
2.6232+	2.6064+	2.590 +	2.579 +	—	—
1.0689+	1.0567+	1.0440+	1.0364+	—	—
42.992	42.210	41.408	40.921	—	—
27.257	26.653	26.031	25.648	25.06	24.71
20.781	20.268	19.736	19.408	18.90	18.61
17.373	16.913	16.435	16.138	15.678	15.412
15.304	14.878	14.435	14.159	13.727	13.478
13.925	13.525	13.105	12.843	12.431	12.192
11.918	11.555	11.172	10.930	10.547	10.322
10.842	10.500	10.136	9.9059	9.5378	9.3188
10.174	9.8453	9.4944	9.2706	8.9106	8.6946
9.7204	9.4006	9.0581	8.8387	8.4838	8.2691
9.1441	8.8361	8.5041	8.2901	7.9404	7.7261
8.7938	8.4930	8.1674	7.9563	7.6090	7.3940
8.5586	8.2626	7.9411	7.7319	7.3858	7.1697
8.3899	8.0973	7.7787	7.5708	7.2251	7.0078
8.1638	7.8758	7.5611	7.3547	7.0093	6.7897
7.7448	7.4652	7.1572	6.9532	6.6062	6.3798
7.3650	7.0929	6.7903	6.5878	6.2361	5.9984

(2.16)
$$T \to -V\,,$$
$$n_2 \to m - n_1 - n_2 + 1\,,$$

m and n_1 remaining unchanged. This d.e. therefore has regular singularities at $0, 1, 2, \cdots, m$ and ∞, i.e. within the range $(0, m)$ of V. However, the program written for the present paper may with trivial modifications be used to calculate accurate percentage points of V in the ranges $(0, 1)$ and $(m-1, m)$, and some investigation has been made of approximations to its distribution [6].

3. Extension of the Ito-Siotani asymptotic expansions

As $n_2 \to \infty$, the null distribution of T_0^2 approaches the chi-squared distribution on mn_1 degrees of freedom, and we may show that its cumulant generating function

$$K(\theta)=\log E \exp(i\theta T_0^2) \qquad (\theta \text{ real}) \tag{3.1}$$

may be developed in an asymptotic expansion for large n_2 of the type considered by Box [1]:

$$K(\theta)-(mn_1/2)\log(1-2i\theta)+\sum_{r=1}^{\infty}\omega_r[(1-2i\theta)^{-r}-1] . \tag{3.2}$$

The ω_r are functions of m, n_1 and n_2, and will be given below to order n_2^{-4}. Writing $t=T_0^2$, introduce the vector of functions

$$\begin{aligned} N(t)&=(N_0(t), \cdots, N_m(t))' \\ &=n_2 \operatorname{diag}\{1, n_2, \cdots, n_2^m\} M(t/n_2) , \end{aligned} \tag{3.3}$$

$N_0(t)$ being the density function of t. Then N may be shown [4] to satisfy the d.e.:

$$\operatorname{diag}(n_2^{-1}, n_2^{-1}t+1, \cdots, n_2^{-1}t+m)dN/dt=[\Delta_0+n_2^{-1}\Delta_1+n_2^{-2}\Delta_2]N , \tag{3.4}$$

where

$$\begin{aligned} \Delta_0&=\{(0, \cdots, 0), (0, -1/2, -1, \cdots, -m/2), (\gamma_1, \cdots, \gamma_m)\} , \\ \Delta_1&=\{(1/2, 1, \cdots, m/2), (\bar{\alpha}_0, \cdots, \bar{\alpha}_m), (0, \cdots, 0)\} , \\ \Delta_2&=\{(\bar{\beta}_0, \cdots, \bar{\beta}_{m-1}), (0, \cdots, 0), (0, \cdots, 0)\} , \end{aligned} \tag{3.5}$$

and

$$\begin{aligned} \bar{\alpha}_i&=[(m-2i)n_1+2i^2-mi-i-2]/2 , \\ \bar{\beta}_i&=(i+1)(n_1-i)/2 . \end{aligned} \tag{3.6}$$

If we take the Fourier transform of N,

$$C(\theta)=\int_0^{\infty} e^{it\theta}N(t)dt=(C_0(\theta), \cdots, C_m(\theta))' , \qquad (\theta \text{ real}) , \tag{3.7}$$

so that $C_0(\theta)$ is the characteristic function of t, then C clearly satisfies a first-order matrix d.e. with respect to $i\theta$ which we omit.

Let

$$\begin{aligned} Q(\theta)&=\log[(1-2i\theta)^{mn_1/2}C_0(\theta)]=(mn_1/2)\log(1-2i\theta)+K(\theta) , \\ R_j(\theta)&=C_j(\theta)/C_0(\theta) , \qquad (j=1, \cdots, m) , \end{aligned} \tag{3.8}$$

and assume expansions of the form

$$\text{(3.9)}\quad \begin{aligned} Q(\theta) &\sim \sum_{r=1}^{\infty} \omega_r[(1-2i\theta)^{-r}-1] \\ R_j(\theta) &\sim \sum_{r=0}^{\infty} \xi_{j,r}[(1-2i\theta)^{-r}-1]\,, \qquad (j=1,\cdots,m)\,. \end{aligned}$$

The following recurrence relations may be obtained:

$$\text{(3.10)}\qquad 2r\omega_r=2(r-1)\omega_{r-1}+mn_1\delta_{1,r}+(1+n_2^{-1}n_1)\xi_{1,r}\,,$$

where the $\xi_{j,r}$ are given recursively by:

$$\xi_{0,r}=\xi_{r,0}=\delta_{0,r}\,,$$

$$\text{(3.11)}\quad \begin{aligned} j\xi_{j,r} &=2\gamma_j\xi_{j-1,r-1}-n_2^{-1}[\lambda_j+2(r-1)]\xi_{j,r-1} \\ &\quad+[n_2^{-1}(j+1)2n_2^{-2}\bar{\beta}_j]\xi_{j+1,r-1}+n_2^{-1}[mn_1+2(r-2)]\xi_{j,r-2} \\ &\quad+2n_2^{-1}\sum_{s=1}^{r-2} s\omega_s(\xi_{j,r-s-2}-\xi_{j,r-s-1})\,, \\ &\qquad\qquad (j=1,\cdots,m;\ r=1,2,\cdots)\,, \end{aligned}$$

$$\lambda_j=j(m+2n_1-2j+1)\,, \qquad (j=1,\cdots,m)\,.$$

The first eight ω_r are necessary to give the expansion (3.2) to order n_2^{-4}:

$$\text{(3.12)}\quad \left\{ \begin{aligned} \omega_1 &=-mn_1^2/2n_2\,, \\ \omega_2 &=(1/4)mn_1[n_2^{-1}(m+n_1+1)+n_2^{-2}n_1(m+2n_1+1)]\,, \\ \omega_3 &=-(1/6)mn_1\{(n_2^{-2}+n_2^{-3}n_1)[m^2+3m(2n_1+1) \\ &\qquad +(4n_1^2+6n_1+4)]+n_2^{-3}n_1^3\}\,, \\ \omega_4 &=(1/8)mn_1\{n_2^{-2}[2m^2+5m(n_1+1)+(2n_1^2+5n_1+5)] \\ &\quad +n_2^{-3}[m^3+2m^2(7n_1+3)+m(34n_1^2+39n_1+21) \\ &\qquad +(15n_1^3+34n_1^2+47n_1+20)] \\ &\quad +n_2^{-4}n_1[m^3+6m^2(2n_1+1)+m(29n_1^2+34n_1+21) \\ &\qquad +(14n_1^3+29n_1^2+42n_1+20)]\}\,, \\ \omega_5 &=-(1/10)mn_1\{5n_2^{-3}[m^3+m^2(7n_1+5)+2m(5n_1^2+9n_1+7) \\ &\quad +(3n_1^3+10n_1^2+19n_1+12)]+n_2^{-4}[m^4+5m^3(5n_1+2) \\ &\qquad +5m^2(26n_1^2+27n_1+13)+10m(18n_1^3+35n_1^2+42n_1+16) \\ &\qquad +4(14n_1^4+45n_1^3+100n_1^2+95n_1+37)]+0(n_2^{-5})\}\,, \\ \omega_6 &=(1/12)mn_1\{n_2^{-3}[5m^3+22m^2(n_1+1)+2m(11n_1^2+27n_1+26) \\ &\quad +(5n_1^3+22n_1^2+52n_1+41)]+n_2^{-4}[9m^4+15m^3(8n_1+5) \\ &\qquad +m^2(388n_1^2+585n_1+397)+3m(122n_1^3+332n_1^2+531n_1+289) \end{aligned} \right.$$

$$+(84n_1^4+366n_1^3+1048n_1^2+1350n_1+732)]+0(n_2^{-5})\} ,$$

$$\omega_7=-(1/2)mn_1[3m^4+m^3(27n_1+22)+m^2(62n_1^2+122n_1+100) +m(44n_1^3+154n_1^2+299n_1+199) +4(2n_1^4+11n_1^3+38n_1^2+60n_1+39)]/n_2^4+0(n_2^{-5}) ,$$

$$\omega_8=(1/16)mn_1[14m^4+93m^3(n_1+1)+m^2(164n_1^2+398n_1+374) +m(93n_1^3+398n_1^2+899n_1+690) +(14n_1^4+93n_1^3+374n_1^2+690n_1+509)]/n_2^4+0(n_2^{-5}) .$$

For $r\geqq 9$, the ω_r do not exceed $0(n_2^{-5})$; apparently, $\omega_r=0(n_2^{-[r/2]})$ where [ ] denotes the integer part. The first six ω_r have been given to $0(n_2^{-3})$ by Muirhead [12], using partial d.e.'s for Constantine's hypergeometric functions of matrix argument. Exponentiation of (3.2), followed by inversion of the resulting linear combination of chi-squared characteristic functions, in principle yields an extension of Ito's and Siotani's expansions of the cumulative distribution function of T_0^2 in the null case. Their formula for T_0^2 percentiles in terms of percentage points, u say, of $\chi^2_{mn_1}$ may be extended using a general inversion of Box-type expansions [7] and the above ω_r. To order n_2^{-3}:

$$(3.13)\quad T_0^2\sim u+(1/2n_2)[u(m-n_1+1)+u^2(m+n_1+1)/(mn_1+2)] +(1/24n_2^2)\{u[7m^2-12m(n_1-1)+(7n_1^2-12n_1+1)] +u^2[13m^2+24m-11n_1^2+7]/(mn_1+2) +u^3[4m^3n_1+2m^2(3n_1^2+3n_1+10)+2m(2n_1^3+3n_1^2+17n_1+18) +4(5n_1^2+9n_1+2)]/(mn_1+2)^2(mn_1+4) +6u^4(m-1)(m+2)(n_1-1)(n_1+2)/(mn_1+2)^2(mn_1+4)(mn_1+6)\} +(1/48n_2^3)\{3u[3m^3-7m^2(n_1-1)+m(7n_1^2-12n_1+1) -(3n_1^3-7n_1^2+n_1+3)]+u^2[23m^3-m^2(19n_1-59) -m(13n_1^2+36n_1-29)+(17n_1^3-13n_1^2-13n_1-7)]/(mn_1+2) +2u^3[7m^4n_1+2m^3(2n_1^2+8n_1+17)-m^2(2n_1^3-9n_1^2-29n_1-88) -m(5n_1^4+2n_1^3+13n_1^2-46n_1-46)-(26n_1^3+20n_1^2-22n_1+8]/ (mn_1+2)^2(mn_1+4)+2u^4[m^5n_1^2+2m^4n_1(7n_1^2+7n_1-6) -m^3(4n_1^4-21n_1^3-83n_1^2+4)+m^2(n_1^5-4n_1^4-7n_1^3+70n_1^2+196n_1+16) +4m(6n_1^4+4n_1^2+57n_1+25)+4(17n_1^3+22n_1^2-11n_1+20)]/ (mn_1+2)^3(mn_1+4)(mn_1+6)-4u^5(m-1)(m+2)(n_1-1)(n_1+2) \times[m^2n_1+m(n_1^2+7n_1-28)-4(7n_1+4)]/(mn_1+2)^3(mn_1+4) \times(mn_1+6)(mn_1+8)+8u^6(m-1)(m+2)(n_1-1)(n_1+2) \times[m^2n_1+m(n_1^2+4n_1-10)-(10n_1+4)]/(mn_1+2)^3(mn_1+4) \times(mn_1+6)(mn_1+8)(mn_1+10)\}+0(n_2^{-4}) .$$

We note also that, corresponding to (2.16), the cumulant generating function of Pillai's n_2V may be expanded in the form (3.2), with coefficients $\omega_{r,v}$ related to those for T_0^2 by

$$(3.14)\qquad (-n_2)^r\omega_{r,v}=mn_1(n_1-m-1)^r/2r + \sum_{s=1}^{r}\binom{r-1}{s-1}(m-n_1-n_2+1)^s(n_1-m-1)^{r-s}\omega_s^*\,,$$

where ω_s^* is obtained by replacing n_2 by $m-n_1-n_2+1$ in ω_s. The $\omega_{r,v}$ thus specified agree with those given by Muirhead [12] to order n_2^{-3}. A percentile expansion for n_2V corresponding to (3.13) may also be derived [6].

4. Discussion of approximations

Another approach to approximating the distribution of T_0^2 has been taken by Pillai and his associates ([14], [15], [17]). They have given formulas for the moment-ratios β_1 and β_2 required for fitting a Pearson curve. The rth central moment μ_r of T_0^2 exists if $r<(n_2-m+1)/2$, ([3], Section 5), so that β_1 and β_2 are defined if $n_2>m+7$, independently of the value of n_1. In the present notation,

$$(4.1)\qquad \beta_1=\mu_3^2/\mu_2^3=\frac{8(n_2-m-3)(n_2-m)(n_2+m-1)^2(2n_1+n_2-m-1)^2}{mn_1(n_2-1)(n_2-m-5)^2(n_2-m+1)^2(n_1+n_2-m-1)}\,.$$

Also, writing

$$(4.2)\qquad c=n_2-m-1\,,$$

the following reduced form of β_2 has been derived by the present author from the recurrence relations for moments given in [4], Section 7.

$$(4.3)\qquad \beta_2=\mu_4/\mu_2^2=\frac{3(c-2)(c+1)A}{mn_1(n_2-1)(c-6)(c-4)(c-1)(c+2)(c+3)(c+n_1)}\,,$$

where

$$(4.4)\qquad A=n(c+n_1)[m(n_2-1)(c^3-5c^2+78c-72)+4c^2(5c-6)] + 4c^2[m(n_2-1)(5c-6)+c(c^2-c+2)]\,.$$

The mean and variance of T_0^2 are also required in fitting a Pearson curve:

$$(4.5)\qquad \begin{aligned}\mu_1&=mn_1n_2/(n_2-m-1)\,,\\ \mu_2&=2mn_1n_2^2(n_2-1)(n_1+n_2-m-1)/(n_2-m-3)(n_2-m-1)^2(n_2-m)\,.\end{aligned}$$

It may be noted that the transformation (2.16) converts the above quantities into those required for fitting a Pearson curve to V.

The extended tables of the Pearson curve given in [10] have been used to compute the Pillai approximation to T_0^2, and its accuracy has been compared with that of the Ito-Siotani approximation. The results will be presented in diagram form in [5]. Briefly, the Pillai method gives remarkably accurate percentage points for $m \leqq 5$ provided n_1 and n_2 are not both small. On the other hand, the accuracy of the Ito-Siotani approximations (3.13) falls away as n_1 increases. However, the latter has the advantage of being a direct formula, and its accuracy is considerably improved by the addition of the n_2^{-3} term. For $m \leqq 4$, $n_1 \leqq 30$, formula (3.13) achieves 3-decimal place accuracy for the 1% points of T_0^2/n_1 for smaller values of n_2 than Pillai's approximation.

Acknowledgement

The author's thanks are due to Miss J. Hawkes of CSIRO for writing the computer program and preparing the tables.

C.S.I.R.O., ADELAIDE AND UNIVERSITY OF NORTH CAROLINA, CHAPEL HILL

REFERENCES

[1] Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria, *Biometrika*, **36**, 317–346.

[2] Coddington, E. A. and Levinson, N. (1955). *Theory of Ordinary Differential Equations*, McGraw-Hill, New York.

[3] Constantine, A. G. (1966). The distribution of Hotelling's generalized T_0^2, *Ann. Math. Statist.*, **37**, 215–225.

[4] Davis, A. W. (1968). A system of linear differential equations for the distribution of Hotelling's generalized T_0^2, *Ann. Math. Statist.*, **39**, 815–832.

[5] Davis, A. W. (1969). Exact distributions of Hotelling's generalized T_0^2, To appear in *Biometrika*.

[6] Davis, A. W. (1970). On the null distribution of the sum of the roots of a multivariate beta distribution, (to appear in *Ann. Math. Statist.*).

[7] Davis, A. W. (1969). Cornish-Fisher-type inversions of Box's asymptotic series, Submitted for publication.

[8] Hotelling, H. (1951). A generalized T-test and measure of multivariate dispersion, *Proc. Second Berkeley Symp. Math. Statist. Prob.*, University of California Press, 23–42.

[9] Ito, K. (1956). Asymptotic formulae for the distribution of Hotelling's generalized T_0^2 statistic, *Ann. Math. Statist.*, **27**, 1091–1105.

[10] Johnson, N. L., Nixon, E., Amos D. E. and Pearson, E. S. (1963). Table of percentage points of Pearson curves, for given $\sqrt{\beta_1}$ and β_2, expressed in standard measure, *Biometrika*, **50**, 459–498.

[11] Lawley, D. N. (1938). A generalization of Fisher's Z-test, *Biometrika*, **30**, 180–187.

[12] Muirhead, R. J. (1969). Asymptotic distributions of some multivariate tests, Submitted for publication.

[13] Pillai, K. C. S. (1955). Some new test criteria in multivariate analysis, *Ann. Math. Statist.*, **26**, 117–121.

[14] Pillai, K. C. S. (1960). *Statistical Tables for Tests of Multivariate Hypotheses*, The Statistical Center, University of the Philippines.

[15] Pillai K. C. S. and Pablo, Samson Jr. (1959). On Hotelling's generalization of T^2, *Biometrika*, **46**, 160–168.

[16] Siotani, M. (1956). On the distributions of the Hotelling's T^2-statistics, *Ann. Inst. Statist. Math.*, 8, 1–14.

[17] Ting, A. (1959). On distributions of traces of certain matrices in multivariate analysis, Unpublished thesis, The Statistical Center, University of the Philippines.

Submitted

9

FURTHER TABULATION OF HOTELLING'S GENERALIZED T_0^2

A.W. Davis

C.S.I.R.O. Division of Mathematics and Statistics,
Glen Osmond, South Australia.

ABSTRACT

A linear homogeneous differential equation of the author's is applied to tabulate accurate upper 5% and 1% points of the null distribution of Hotelling's generalized T_0^2 for m = 6 through 10 variates, extending previous tabulations of this statistic.

1. INTRODUCTION

Hotelling's (1951) generalized T_0^2 is defined by

$$T = T_0^2/n_2 = \text{tr}\ S_1 S_2^{-1} \tag{1.1}$$

where S_1 and S_2 are $m \times m$ matrices with independent Wishart distributions on n_1 and n_2 degrees of freedom respectively. This statistic may be used, for example, as a multivariate analogue of the F statistic in testing linear hypotheses based on the multivariate linear model; in this case, S_1 and S_2 are the hypothesis and error sum of products matrices, respectively.

As yet, no explicit expression for the exact null distribution of T has been given in the general case. Constantine (1966) gave a series expansion which is valid for $|T| < 1$. Hotelling (1951) derived the distribution for $m = 2$, and explicit results for $m \leq 4$ and small n_1 were obtained by Pillai and Young (1971) and Pillai and Sudjana (1974). The latter paper also discusses a rearrangement of Constantine's series in terms of beta densities, suggested by Pillai. As indicated in Section 2, this series is convergent for $0 < T < m/(m - 2)$; thus, while it may yield some useful results for $m \leq 6$ and small n_1 (see Tables VI, VII and VIII of Pillai and Sudjana, 1974), for larger m and n_1 it will only be useful for increasingly large n_2.

The present author (Davis, 1968) showed that the null density of T satisfies a linear homogeneous differential equation (d.e.) of order m. Constantine's series reduces in the null case to a single fundamental solution of this d.e. at $T = 0$, convergent in the unit circle. The d.e. was used to tabulate upper 5% and 1% points of T for $m = 3$, 4 and 5 (Davis 1970a, 1970b) by analytic continuation of Constantine's series outside the unit circle.

Apart from some percentage points for $m = 6$, $n_1 = 6$, 8 and large n_2, computed by Pillai and Sudjana (1974) using the series mentioned above, no extensions of the author's tabulations have been made subsequently. It thus seems worthwhile to show that the d.e. approach continues to be useful for larger m by presenting 5% and 1% points for $m = 6$ through 10, and n_1 ranging up to 35 (see Table). The method did however encounter difficulties for $n_2 \lesssim m + 4$.

2. THE DIFFERENTIAL EQUATION

It was shown in Davis (1968) that an (m + 1)-vector

$$\underline{M}(T) = (M_0(T), \ldots, M_m(T))' \tag{2.1}$$

may be defined, in which $M_0(T) \equiv f(T)$, the null density function of T, and the remaining $M_i(T)$ are certain auxiliary functions, such that $\underline{M}$ satisfies a differential equation

$$(TI_{m+1} + \Lambda)d\underline{M}/dT = C\underline{M}. \tag{2.2}$$

Here I_{m+1} denotes the (m + 1) × (m + 1) unit matrix, Λ is a diagonal matrix

$$\Lambda = \text{diag}\{0, 1, \ldots, m\}, \tag{2.3}$$

and

$$C = \begin{bmatrix} \alpha_0, & \beta_0, & 0, & 0, & \cdots\cdots & , 0 \\ \gamma_1, & \alpha_1, & \beta_1, & 0, & \cdots\cdots & , 0 \\ 0 & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & \gamma_{m-1}, & \alpha_{m-1}, & \beta_{m-1} \\ 0 & \cdots\cdots & & 0 & , \gamma_m & , \alpha_m \end{bmatrix} \tag{2.4}$$

where

$$\begin{aligned} \alpha_i &= \{(m - 2i)n_1 - in_2 + (2i^2 - mi - i - 2)\}/2, \\ \beta_i &= (i + 1)(n_1 + n_2 - i)/2, \\ \gamma_i &= -(m - i + 1)(n_1 - i + 1)/2. \end{aligned} \tag{2.5}$$

Since (2.2) may be rewritten as

$$d\underline{M}/dT = \{\sum_{r=0}^{m} (T + r)^{-1} V_r\}\underline{M}, \qquad (2.6)$$

where V_r is obtained by replacing all elements of C by zeros except those in the r th row (r = 0, 1, ..., m), it follows from the general theory of such systems (Hille, 1969, Ch. 5, Appendix B) that the d.e. has regular singularities at 0, -1, ..., -m and infinity. The (m + 1) linearly independent solutions of (2.2) in the unit circle about T = 0 correspond to the latent roots of V_0, i.e. zero (with multiplicity m) and

$$\alpha_0 = mn_1/2 - 1. \qquad (2.7)$$

In virtue of Constantine's (1963) result, the relevant solution is given by the latter root,

$$\underline{M}(T) = k(m; n_1, n_2)T^{mn_1/2-1} \sum_{j=0}^{\infty} \underline{W}_j T^j, \qquad (2.8)$$

$$\underline{W}_j = (W_{0j}, \ldots, W_{mj})', \quad W_{00} = 1, \quad |T| < 1,$$

where

$$k(m; n_1, n_2) = \Gamma_m((n_1 + n_2)/2)/\Gamma(mn_1/2)\Gamma_m(n_2/2), \qquad (2.9)$$

$$\Gamma_m(z) = \pi^{m(m-1)/4} \Pi_{i=0}^{m-1} \Gamma(z - i/2).$$

The following recurrence relations for the W_{ij} are obtained from (2.2),

$$W_{i0} = \delta_{i0}, \text{ (Kronecker's delta)},$$

$$i(j + mn_1/2 - 1)W_{ij} = \gamma_i W_{i-1,j} + \{\alpha_i - (j + mn_1/2 - 2)\}W_{i,j-1} + \beta_i W_{i+1,j-1}, \quad (i = 1, .., m;\ j = 1,2,..), \qquad (2.10)$$

$$jW_{0j} = (n_1 + n_2)W_{1j}/2.$$

In the neighbourhood of any point a > 0, M(T) may be expanded in a series

$$\underline{M}(T) = \sum_{j=0}^{\infty} \underline{W}_j^* (T - a)^j. \tag{2.11}$$

Recurrence relations for the $\underline{W}_j^*$ are derivable from (2.2), requiring only the value of $\underline{W}_0^* = \underline{M}(a)$ as a starting point. This is the basis of the analytic continuation procedure. The flexibility of the d.e. approach as a computational tool is further indicated by the ease with which T may be transformed into another variable, for example

$$V_\theta = \frac{T/\theta}{1 + T/\theta} \tag{2.12}$$

so that the results obtained from (2.2) may be checked by recalculation using the transformed d.e. In particular, the solution at the origin in the case V_m is Pillai's transformed series (Pillai and Sudjana, 1974). Applying this transformation to the singularities of (2.2), we see that the latter map into the following singularities of the transformed d.e.,

$-s/(m - s)$ $(s = 0, 1, \ldots, m - 1)$, 1, and infinity.

The radius of convergence of the series expansion at the origin is the distance to the nearest singularity, $(m - 1)^{-1}$ in this case, yielding a range of convergence $0 < T < m/(m - 2)$ for Pillai's series, as stated in the Introduction.

3. COMPUTATION

In solving (2.2), it was found convenient to make a change of scale, $T^* = T/\theta$, where θ was generally assigned the value n_1. The series (2.8) was expanded to forty terms using (2.10), and a value $T^* = a > 0$ was constructed for which the last six terms in each component series did not exceed 10^{-10} in absolute value. The cumulative distribution function of T^* was specified by the integrated series for $M_0(T)$, and if its value at a did not exceed the required probability level, $\underline{M}(a)$ was calculated, the series was re-expanded about a, and the process was repeated as often as required. The percentage point was then obtained by the Newton-Raphson method. All results presented in the Table were confirmed to the number of figures quoted by transforming (2.2) in accordance with (2.12) with $\theta = n_1$, and recalculating. The percentage points are those of T_0^2/n_1, and the ∞ line represents the limiting $\chi^2_{mn_1}/n_1$ values.

Computations were carried out on the CSIRO Control Data Cyber 70 model 76 in Canberra.

BIBLIOGRAPHY

Constantine, A.G. (1966). The distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* *37*, 215-25.

Davis, A.W. (1968). A system of linear differential equations for the distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* *39*, 815-32.

Davis, A.W. (1970a). Exact distributions of Hotelling's generalized T_0^2. *Biometrika* *57*, 187-91.

Davis, A.W. (1970b). Further applications of a differential equation for Hotelling's generalized T_0^2. *Ann. Inst. Statist. Math.* *22*, 77-87.

Hille, E. (1969). *Lectures on Ordinary Differential Equations.* Reading, Mass: Addison-Wesley Publishing Co.

Hotelling, H. (1951). A generalized T-test and measure of multivariate dispersion. *Proc. Second Berkeley Symp.*,23-42.

Pillai, K.C.S. and Sudjana (1974). On the distribution of Hotelling's trace and power comparisons. *Commun. Statist. 3*, 433-54,

Pillai, K.C.S. and Young, D.L. (1971). On the exact distribution of Hotelling's generalized T_0^2. *J. Multivariate Anal. 1*, 90-107.

TABLE

Upper percentage points for the null distribution of Hotelling's generalized T_0^2/n_1

m = 6. Upper 5% points

n_2 \ n_1	6	8	10	12	15	20	25	30	35
10	45.722	44.677	44.019	43.567	43.103	42.626	42.334	42.136	41.993
12	28.959	28.121	27.590	27.223	26.843	26.451	26.209	26.044	25.925
14	22.321	21.600	21.141	20.821	20.489	20.144	19.929	19.783	19.677
16	18.858	18.210	17.795	17.505	17.202	16.886	16.688	16.553	16.455
18	16.755	16.157	15.772	15.501	15.218	14.921	14.735	14.607	14.513
20	15.351	14.788	14.424	14.168	13.899	13.615	13.436	13.313	13.223
25	13.293	12.786	12.456	12.222	11.975	11.711	11.544	11.428	11.343
30	12.180	11.705	11.395	11.173	10.939	10.687	10.526	10.414	10.331
35	11.484	11.031	10.733	10.520	10.293	10.049	9.8921	9.7820	9.7003
40	11.009	10.571	10.282	10.075	9.8535	9.6142	9.4596	9.3508	9.2699
50	10.402	9.9832	9.7060	9.5067	9.2927	9.0598	8.9082	8.8009	8.7207
60	10.031	9.6246	9.3547	9.1602	8.9507	8.7215	8.5717	8.4651	8.3851
70	9.7813	9.3830	9.1182	8.9269	8.7204	8.4938	8.3450	8.2388	8.1589
80	9.6014	9.2093	8.9480	8.7591	8.5548	8.3300	8.1819	8.0759	7.9959
100	9.3598	8.9760	8.7197	8.5340	8.3326	8.1102	7.9629	7.8572	7.7771
200	8.9099	8.5419	8.2950	8.1153	7.9193	7.7011	7.5552	7.4494	7.3685
500	8.6594	8.3002	8.0587	7.8823	7.6894	7.4734	7.3280	7.2219	7.1403
1000	8.5788	8.2226	7.9827	7.8075	7.6155	7.4002	7.2550	7.1487	7.0668
∞	8.4997	8.1463	7.9082	7.7340	7.5430	7.3284	7.1832	7.0768	6.9945

TABLE (continued)

$m = 6$. Upper 1% points

n_2 \ n_1	6	8	10	12	15	20	25	30	35
10	86.397	83.565	81.804	80.602	79.376	78.124	77.360	76.845	76.474
12	46.027	44.103	42.899	42.073	41.227	40.359	39.826	39.466	39.206
14	32.433	30.918	29.966	29.309	28.634	27.936	27.507	27.215	27.004
16	25.977	24.689	23.875	23.311	22.729	22.126	21.753	21.498	21.314
18	22.292	21.146	20.418	19.913	19.389	18.844	18.505	18.273	18.105
20	19.935	18.886	18.217	17.752	17.267	16.761	16.445	16.229	16.071
25	16.642	15.737	15.156	14.749	14.324	13.875	13.592	13.397	13.254
30	14.944	14.118	13.586	13.211	12.816	12.398	12.133	11.949	11.814
35	13.913	13.138	12.635	12.281	11.906	11.506	11.252	11.074	10.943
40	13.223	12.482	12.000	11.659	11.298	10.911	10.663	10.490	10.361
50	12.358	11.661	11.206	10.882	10.538	10.167	9.9271	9.7587	9.6333
60	11.839	11.169	10.730	10.417	10.083	9.7206	9.4860	9.3202	9.1963
70	11.493	10.841	10.413	10.107	9.7795	9.4238	9.1922	9.0281	8.9050
80	11.246	10.607	10.187	9.8859	9.5634	9.2121	8.9826	8.8195	8.6968
100	10.917	10.295	9.8857	9.5918	9.2758	8.9301	8.7033	8.5414	8.4193
200	10.312	9.7231	9.3330	9.0517	8.7476	8.4121	8.1897	8.0295	7.9075
500	9.9799	9.4089	9.0296	8.7553	8.4576	8.1275	7.9072	7.7473	7.6249
1000	9.8738	9.3085	8.9328	8.6607	8.3651	8.0366	7.8168	7.6570	7.5344
∞	9.7699	9.2103	8.8379	8.5680	8.2744	7.9475	7.7283	7.5685	7.4456

TABLE (continued)

m = 7. Upper 5% points

n_2 \ n_1	8	10	12	15	20	25	30	35
10	85.040	84.082	83.426	82.755	82.068	81.648	81.364	81.159
12	42.850	42.126	41.627	41.113	40.583	40.257	40.037	39.877
14	29.968	29.373	28.961	28.534	28.091	27.817	27.631	27.495
16	24.038	23.519	23.158	22.781	22.389	22.145	21.978	21.857
18	20.692	20.222	19.893	19.549	19.189	18.964	18.809	18.696
20	18.561	18.125	17.819	17.498	17.159	16.947	16.800	16.694
25	15.587	15.202	14.930	14.642	14.337	14.143	14.009	13.911
30	14.049	13.693	13.440	13.172	12.884	12.701	12.573	12.478
35	13.113	12.776	12.535	12.278	12.002	11.825	11.700	11.608
40	12.485	12.160	11.927	11.679	11.411	11.237	11.115	11.025
50	11.695	11.386	11.165	10.927	10.668	10.500	10.381	10.292
60	11.219	10.921	10.706	10.475	10.221	10.056	9.9383	9.8500
70	10.901	10.610	10.400	10.173	9.9233	9.7596	9.6429	9.5550
80	10.674	10.388	10.181	9.9567	9.7102	9.5478	9.4317	9.3440
100	10.371	10.091	9.8886	9.6688	9.4259	9.2652	9.1498	9.0622
200	9.8118	9.5448	9.3504	9.1384	8.9021	8.7441	8.6295	8.5419
500	9.5037	9.2438	9.0539	8.8462	8.6134	8.4568	8.3424	8.2543
1000	9.4051	9.1475	8.9591	8.7527	8.5211	8.3648	8.2504	8.1621
∞	9.3085	9.0531	8.8662	8.6612	8.4306	8.2747	8.1603	8.0718

TABLE (continued)

m = 7. Upper 1% points

n_2 \ n_1	8	10	12	15	20	25	30	35
10	185.93	182.94	180.90	178.83	176.73	175.44	174.57	173.92
12	71.731	69.978	68.779	67.552	66.296	65.528	65.010	64.636
14	44.255	42.978	42.099	41.197	40.269	39.698	39.311	39.032
16	33.097	32.057	31.339	30.599	29.834	29.361	29.039	28.806
18	27.273	26.374	25.750	25.105	24.435	24.019	23.735	23.529
20	23.757	22.949	22.388	21.804	21.195	20.816	20.556	20.367
25	19.117	18.440	17.965	17.469	16.947	16.619	16.392	16.227
30	16.848	16.239	15.810	15.360	14.882	14.580	14.370	14.216
35	15.512	14.945	14.544	14.121	13.670	13.383	13.183	13.036
40	14.634	14.095	13.713	13.309	12.876	12.599	12.405	12.262
50	13.553	13.049	12.691	12.310	11.899	11.634	11.448	11.309
60	12.914	12.432	12.088	11.720	11.323	11.065	10.882	10.746
70	12.492	12.024	11.690	11.332	10.942	10.689	10.509	10.374
80	12.193	11.736	11.408	11.056	10.673	10.422	10.244	10.110
100	11.797	11.353	11.034	10.691	10.316	10.070	9.8935	9.7607
200	11.077	10.658	10.356	10.028	9.6670	9.4273	9.2545	9.1228
500	10.685	10.280	9.9866	9.6679	9.3140	9.0776	8.9060	8.7744
1000	10.561	10.160	9.8693	9.5533	9.2017	8.9663	8.7950	8.6634
∞	10.439	10.043	9.7547	9.4413	9.0920	8.8575	8.6865	8.5548

TABLE (continued)

m = 8. Upper 5% points

n_2 \ n_1	8	10	12	15	20	25	30	35
14	42.516	41.737	41.198	40.641	40.066	39.711	39.470	39.296
16	31.894	31.242	30.788	30.318	29.829	29.525	29.318	29.167
18	26.421	25.847	25.446	25.028	24.591	24.319	24.132	23.996
20	23.127	22.605	22.239	21.856	21.454	21.201	21.028	20.902
25	18.770	18.324	18.009	17.677	17.325	17.102	16.947	16.834
30	16.626	16.221	15.934	15.629	15.303	15.095	14.950	14.843
35	15.356	14.977	14.707	14.418	14.109	13.910	13.771	13.668
40	14.518	14.156	13.898	13.621	13.322	13.129	12.994	12.893
50	13.482	13.142	12.898	12.636	12.351	12.165	12.034	11.936
60	12.866	12.540	12.305	12.051	11.774	11.593	11.465	11.368
70	12.459	12.142	11.912	11.665	11.393	11.215	11.088	10.992
80	12.169	11.858	11.634	11.390	11.122	10.946	10.820	10.725
100	11.785	11.483	11.264	11.026	10.763	10.590	10.465	10.370
200	11.084	10.798	10.589	10.362	10.108	9.9389	9.8159	9.7218
500	10.701	10.423	10.221	9.9993	9.7509	9.5836	9.4614	9.3673
1000	10.579	10.304	10.104	9.8840	9.6371	9.4704	9.3484	9.2543
∞	10.459	10.188	9.9892	9.7712	9.5258	9.3598	9.2379	9.1437

TABLE (continued)

$m = 8$. Upper 1% points

n_2 \ n_1	8	10	12	15	20	25	30	35
14	65.793	64.035	62.828	61.592	60.323	59.545	59.019	58.639
16	44.977	43.633	42.707	41.754	40.771	40.164	39.753	39.456
18	35.265	34.146	33.373	32.573	31.745	31.232	30.882	30.629
20	29.786	28.808	28.129	27.425	26.691	26.235	25.924	25.697
25	23.001	22.212	21.661	21.085	20.480	20.100	19.838	19.647
30	19.867	19.173	18.686	18.173	17.631	17.288	17.051	16.876
35	18.077	17.440	16.991	16.516	16.011	15.690	15.466	15.301
40	16.924	16.324	15.900	15.451	14.970	14.662	14.447	14.288
50	15.528	14.975	14.582	14.163	13.711	13.420	13.216	13.063
60	14.715	14.190	13.815	13.414	12.980	12.698	12.499	12.351
70	14.184	13.677	13.313	12.925	12.502	12.226	12.031	11.885
80	13.810	13.315	12.960	12.580	12.165	11.894	11.701	11.556
100	13.317	12.839	12.496	12.127	11.722	11.457	11.267	11.124
200	12.429	11.983	11.660	11.311	10.925	10.669	10.484	10.343
500	11.951	11.521	11.210	10.871	10.495	10.244	10.061	9.9208
1000	11.800	11.375	11.067	10.732	10.359	10.109	9.9270	9.7870
∞	11.652	11.233	10.928	10.597	10.227	9.9778	9.7963	9.6564

TABLE (continued)

m = 9. Upper 5% points

n_2 \ n_1	10	12	15	20	25	30	35
14	61.915	61.196	60.456	59.694	59.224	58.907	58.68
16	42.157	41.583	40.988	40.372	39.990	39.730	39.542
18	33.171	32.680	32.169	31.637	31.305	31.079	30.914
20	28.140	27.702	27.245	26.766	26.466	26.260	26.110
25	21.911	21.548	21.165	20.759	20.503	20.326	20.196
30	19.020	18.695	18.350	17.982	17.747	17.584	17.464
35	17.361	17.059	16.737	16.391	16.170	16.015	15.900
40	16.287	16.000	15.694	15.363	15.150	15.000	14.889
50	14.982	14.714	14.427	14.115	13.912	13.768	13.661
60	14.218	13.962	13.687	13.385	13.188	13.049	12.944
70	13.717	13.469	13.201	12.907	12.714	12.577	12.473
80	13.364	13.121	12.859	12.570	12.380	12.243	12.141
100	12.898	12.663	12.407	12.125	11.938	11.804	11.703
200	12.055	11.833	11.591	11.321	11.141	11.010	10.910
500	11.600	11.385	11.150	10.887	10.710	10.580	10.480
1000	11.455	11.243	11.011	10.749	10.573	10.444	10.344
∞	11.315	11.105	10.874	10.615	10.440	10.311	10.211

TABLE (continued)

m = 9. Upper 1% points

n_2 \ n_1	10	12	15	20	25	30	35
14	101.87	100.15	98.387	96.583	95.478	94.74	94.2
16	60.990	59.770	58.518	57.229	56.437	55.90	55.51
18	44.668	43.697	42.697	41.662	41.022	40.587	40.272
20	36.244	35.419	34.564	33.676	33.126	32.750	32.477
25	26.601	25.960	25.292	24.591	24.152	23.850	23.629
30	22.443	21.890	21.310	20.696	20.308	20.040	19.843
35	20.153	19.650	19.120	18.556	18.198	17.949	17.765
40	18.708	18.238	17.741	17.210	16.870	16.632	16.457
50	16.993	16.563	16.105	15.612	15.294	15.071	14.905
60	16.011	15.604	15.169	14.698	14.393	14.177	14.016
70	15.375	14.983	14.563	14.107	13.810	13.599	13.441
80	14.930	14.548	14.139	13.693	13.401	13.194	13.038
100	14.348	13.981	13.585	13.152	12.868	12.664	12.511
200	13.310	12.968	12.597	12.188	11.915	11.719	11.569
500	12.756	12.427	12.070	11.673	11.407	11.214	11.066
1000	12.581	12.257	11.904	11.511	11.247	11.054	10.907
∞	12.412	12.092	11.743	11.353	11.091	10.899	10.752

TABLE (continued)

m = 10. Upper 5% points

n_2 \ n_1	10	12	15	20	25	30	35
14	98.999	98.013	97.002	95.963	95.326	94.9	94.6
16	58.554	57.814	57.050	56.260	55.772	55.44	55.20
18	43.061	42.454	41.824	41.169	40.762	40.485	40.284
20	35.146	34.620	34.071	33.497	33.140	32.895	32.716
25	26.080	25.660	25.219	24.753	24.458	24.255	24.107
30	22.140	21.773	21.384	20.970	20.706	20.523	20.388
35	19.955	19.618	19.260	18.876	18.630	18.458	18.331
40	18.569	18.252	17.914	17.550	17.316	17.151	17.029
50	16.913	16.622	16.309	15.969	15.748	15.592	15.476
60	15.960	15.684	15.385	15.059	14.847	14.695	14.582
70	15.341	15.074	14.786	14.469	14.261	14.113	14.002
80	14.907	14.647	14.365	14.055	13.851	13.705	13.595
100	14.338	14.087	13.814	13.513	13.313	13.170	13.061
200	13.319	13.085	12.828	12.542	12.351	12.212	12.106
500	12.774	12.548	12.301	12.023	11.836	11.699	11.594
1000	12.602	12.379	12.134	11.859	11.674	11.538	11.432
∞	12.434	12.214	11.972	11.700	11.515	11.380	11.275

TABLE (continued)

m = 10. Upper 1% points

n_2 \ n_1	10	12	15	20	25	30	35
14	180.90	178.28	175.62	172.91	171.24	170	--
16	89.068	87.414	85.270	83.980	82.91	82.2	81.7
18	59.564	58.328	57.055	55.742	54.933	54.384	53.990
20	45.963	44.951	43.905	42.821	42.150	41.693	41.362
25	31.774	31.029	30.253	29.440	28.932	28.583	28.328
30	26.115	25.489	24.832	24.139	23.701	23.399	23.177
35	23.116	22.556	21.966	21.338	20.939	20.663	20.459
40	21.267	20.749	20.201	19.615	19.241	18.980	18.787
50	19.114	18.646	18.148	17.611	17.266	17.023	16.842
60	17.901	17.462	16.992	16.484	16.154	15.922	15.748
70	17.124	16.703	16.252	15.762	15.443	15.216	15.046
80	16.583	16.175	15.738	15.260	14.948	14.726	14.559
100	15.881	15.490	15.069	14.608	14.305	14.088	13.925
200	14.641	14.280	13.889	13.457	13.169	12.962	12.803
500	13.986	13.641	13.266	12.848	12.569	12.366	12.210
1000	13.780	13.441	13.070	12.658	12.381	12.179	12.023
∞	13.581	13.246	12.881	12.472	12.198	11.997	11.842

Key words and phrases: linear homogeneous differential equation; multivariate distribution; analytic continuation; percentage points.

10

Commonwealth of Australia
COMMONWEALTH SCIENTIFIC INDUSTRIAL RESEARCH ORGANIZATION
The Annals of Mathematical Statistics
1970, Vol. 41, No. 5, 1557–1562

ON THE NULL DISTRIBUTION OF THE SUM OF THE ROOTS OF A MULTIVARIATE BETA DISTRIBUTION[1]

BY A. W. DAVIS

C.S.I.R.O., Adelaide and University of North Carolina at Chapel Hill

1. Introduction. The distribution of Pillai's V statistic [8] is shown to satisfy a homogeneous linear differential equation (d.e.) of Fuchsian type, which is related by a simple transformation to the author's d.e. for Hotelling's generalized T_0^2 [3]. This transformation implies certain relationships between the moments and asymptotic expansions of the two distributions. The adequacy of some approximations to V is checked by using the d.e. to tabulate some accurate percentage points.

2. Systems of differential equations. Let $\mathbf{S}_1$, $\mathbf{S}_2$ denote $m \times m$ matrices with independent null Wishart distributions on n_1, n_2 degrees of freedom respectively ($n_1, n_2 \geqq m$), estimating the same covariance matrix. The joint distribution of the latent roots $\theta_1, \cdots, \theta_m$ of $\mathbf{S}_1(\mathbf{S}_1+\mathbf{S}_2)^{-1}$ is well known to be

$$(2.1)\quad \phi_{n_1,n_2}(\theta_1, \cdots, \theta_m) = C(m; n_1, n_2)(\prod_{i=1}^m \theta_i)^{\frac{1}{2}(n_1-m-1)}(\prod_{i=1}^m (1-\theta_i))^{\frac{1}{2}(n_2-m-1)} \cdot \prod_{i<j}(\theta_i-\theta_j), \qquad (0 < \theta_m < \cdots < \theta_1 < 1),$$

where

$$(2.2)\qquad C(m; n_1, n_2) = \pi^{\frac{1}{2}m^2}\Gamma_m(\tfrac{1}{2}(n_1+n_2))/\Gamma_m(\tfrac{1}{2}m)\Gamma_m(\tfrac{1}{2}n_1)\Gamma_m(\tfrac{1}{2}n_2).$$

Pillai's V statistic is defined by

$$(2.3)\qquad V = \sum_{i=1}^m \theta_i$$

and Hotelling's generalized T_0^2 statistic by

$$(2.4)\qquad T = \sum_{i=1}^m \theta_i/(1-\theta_i) = T_0^2/n_2.$$

Following the method of [3], Section 2, we introduce the Laplace transforms (Lt's)

$$(2.5)\quad L_r(s) = \int_{R_m} \exp(-s\sum\theta_i)\, \phi_{n_1,n_2}(\theta_1, \cdots, \theta_m) \sum_{k_1<\cdots<k_r}[(1-\theta_{k_1})\cdots(1-\theta_{k_r})]^{-1} \cdot d\theta_1 \cdots d\theta_m, \qquad (r = 0, 1, \cdots, m),$$

where R_m is the region defined in (2.1), and the summation is extended over the $\binom{m}{r}$ selections of r distinct integers $k_1, \cdots, k_r$ from the set $1, 2, \cdots, m$. Thus, $L_0(s)$ is the Lt of $f_{n_1,n_2}(V)$, the density function of V. For $r \geqq 1$, the integrands in (2.5) are dominated by ϕ_{n_1,n_2-2}, and so the $L_r(s)$ exist only for $n_2 \geqq m+2$. This restriction will be preserved for the present. In general, we see that

$$(2.6)\qquad \int_{R_m} \exp(-s\sum\theta_i)\psi(\theta_1, \cdots, \theta_m)\, d\theta_1 \cdots d\theta_m$$

Received September 3, 1969; revised January 5, 1970.

[1] This research was partially supported by the National Institutes of Health Grant No. GM-12868.

is the ordinary Lt of

$$\Psi(V) = \int_{R_{m-1}(V)} \Psi(V-\theta_2-\cdots-\theta_m, \theta_2, \cdots, \theta_m)\, d\theta_2 \cdots d\theta_m, \tag{2.7}$$

where

$$\begin{aligned} R_{m-1}(V) &= R_{m-1} \cap \{\theta_2+\cdots+\theta_m > V-1\} \cap \{2\theta_2+\theta_3+\cdots+\theta_m < V\} \\ &= \{\max[\theta_{s+1}, V-(s-1)-\theta_{s+1}-\cdots-\theta_m] < \theta_s < s^{-1} \\ &\quad \cdot (V-\theta_{s+1}-\cdots-\theta_m);\ s = 2, \cdots, m\}, \qquad (\theta_{m+1} \equiv 0). \end{aligned} \tag{2.8}$$

Hence $L_r(s)$ is the Lt of $H_r(V)$, say, $(r = 0, 1, \cdots, m)$, which may be obtained in integral form from (2.5) and (2.7). Clearly, if $V = j$, $(j = 1, 2, \cdots, m-1)$, the left-hand sides of the inequalities in (2.8) reduce to θ_{s+1} for $s = j+1, \cdots, m$. The boundary of $R_{m-1}(V)$ therefore alters its character as V passes through the integer values $1, 2, \cdots, m-1$, corresponding to the passage of the hyperplane $\sum \theta_i = V$ through the vertices $(1, 0, \cdots, 0)$, $(1, 1, 0, \cdots, 0)$, $\cdots$, $(1, 1, \cdots, 1, 0)$ of R_m. This results in $f_{n_1, n_2}(V)$ having a piecewise analytic nature which is reflected in the d.e.'s derived below.

A first-order system of d.e.'s relating the $L_r(s)$ may be obtained along the lines of [3], Section 2; in fact, the integrands at any stage of the argument may be derived formally from those given in this reference by making the transformation $w_i \to -\theta_i$, $n_2 \to m-n_1-n_2+1$, $s \to -s$.

This leads to the following system of d.e.'s:

$$\begin{aligned} -(m-r+1)sL_{r-1} + [s((d/ds)+r)+a_r+1]L_r - b_r L_{r+1} &= 0, \\ (r = 0, 1, \cdots, m-1;\ L_{-1} &\equiv 0), \\ ((d/ds)+m)L_m - L_{m-1} &= 0, \end{aligned} \tag{2.9}$$

where

$$a_r = \tfrac{1}{2}(m-r)(n_1+n_2-m+r-1)-1, \qquad b_r = \tfrac{1}{2}(r+1)(n_2-m+r-1). \tag{2.10}$$

Equation (2.9) may be obtained from [3] equations (2.19) and (2.20) by the transformations

$$s \to -s, \qquad n_2 \to m-n_1-n_2+1. \tag{2.11}$$

Inverting the Lt's, the following system of first order d.e.'s is found for the $H_r(V)$, $(n_2 \geqq m+2)$:

$$\begin{aligned} (m-r+1)\, dH_{r-1}/dV + [(V-r)\, d/dV - a_r]H_r + b_r H_{r+1} &= 0, \\ (r = 0, 1, \cdots, m-1;\ H_{-1} &\equiv 0), \\ H_{m-1} + (V-m)H_m &= 0. \end{aligned} \tag{2.12}$$

This system is related to [3] equations (2.21) and (2.22) by the transformations

$$T \to -V, \qquad n_2 \to m-n_1-n_2+1. \tag{2.13}$$

Since $b_r > 0$ for $n_2 \geqq m+2$, elimination of $H_1, \cdots, H_m$ from equation (2.12) will result in a linear homogeneous d.e. of order m for $H_0 = f$, having regular singularities at $V = 0, 1, \cdots, m$ and infinity.

3. Nature of the solution. The solution of (2.12) in the unit circle about $V = 0$ follows from [3] Section 3, using (2.13). Again the characteristic roots of the d.e. are $\frac{1}{2}mn_1 - 1$ and zero (with multiplicity m), the relevant solution following from the non-zero root. Recurrence relations for the coefficients in the power series for $f_{n_1, n_2}(V)$, $0 < V < 1$, are obtainable from [3] equation (3.11), and the multiplicative constant is the same as that for T, namely,

$$k(m; n_1, n_2) = \Gamma_m(\tfrac{1}{2}(n_1 + n_2))/\Gamma(\tfrac{1}{2}mn_1)\Gamma_m(\tfrac{1}{2}n_2). \tag{3.1}$$

(Constantine [2]). This solution also serves to define the distribution in the interval $m-1 < V < m$, since from the definition of V,

$$f_{n_1, n_2}(V) = f_{n_2, n_1}(m - V), \qquad (0 < V < m). \tag{3.2}$$

Unfortunately, however, in the intervals between the singularities $1, 2, \cdots, m-1$, $f_{n_1, n_2}(V)$ will be specified by certain linear combinations of the full set of m linearly independent solutions. The calculation of the numerical coefficients in these linear combinations presents a formidable unsolved problem.

In the bivariate case $m = 2$, the differential equation for f_{n_1, n_2} is found to be

$$\begin{aligned} V(1-V)(2-V)f'' - [\tfrac{1}{2}(3n_1 + 3n_2 - 14)V^2 - 2(2n_1 + n_2 - 7)V + 2(n_1 - 2)]f' \\ + \tfrac{1}{2}(n_1 + n_2 - 4)[(n_1 + n_2 - 4)V - 2(n_1 - 2)]f = 0, \end{aligned} \tag{3.3}$$

and the density function may be expressed in terms of the Gaussian hypergeometric function:

$$\begin{aligned} f_{n_1, n_2}(V) &= [2B(n_1, n_2 - 1)]^{-1}(\tfrac{1}{2}V)^{n_1 - 1}(1 - \tfrac{1}{2}V)^{n_2 - 3} \cdot \\ &\quad {}_2F_1(1, \tfrac{1}{2}(3 - n_2); \tfrac{1}{2}(n_1 + 1); r^2), \qquad (0 < V < 1), \\ f_{n_1, n_2}(V) &= [2B(n_2, n_1 - 1)]^{-1}(\tfrac{1}{2}V)^{n_1 - 3}(1 - \tfrac{1}{2}V)^{n_2 - 1} \cdot \\ &\quad {}_2F_1(1, \tfrac{1}{2}(3 - n_1); \tfrac{1}{2}(n_2 + 1); r^{-2}), \qquad (1 < V < 2), \end{aligned} \tag{3.4}$$

where $r = V/(2-V)$. These functions reduce to polynomials in V for odd $n_2 \geqq 3$ and odd $n_1 \geqq 3$, respectively.

So far, it has been assumed that $n_2 \geqq m+2$. In the cases $n_2 = m, m+1$ we note that f_{n_1, n_2} is a numerical multiple of the H_m function corresponding to f_{n_1, n_2+2}. Elimination of $H_0, \cdots, H_{m-1}$ from (2.12) with n_2 replaced by $n_2 + 2$ would show that f_{n_1, n_2} satisfies the general mth order d.e. in these cases. However, when $n_2 = m$, $m+1$, we have $b_1 = 0$, $b_0 = 0$ respectively, and the system (2.12), regarded as a d.e. for $H_0 = f_{n_1, n_2}$, degenerates into a second or first order d.e.:

$$\begin{aligned} V(1-V)H_0'' + [V(mn_1 - \tfrac{1}{2}m - \tfrac{1}{2}n_1 - 3) - (\tfrac{1}{2}mn_1 - 2)]H_0' \\ - (\tfrac{1}{2}mn_1 - \tfrac{1}{2}m - 1)(\tfrac{1}{2}mn_1 - \tfrac{1}{2}n_1 - 1)H_0 = 0, \qquad (n_2 = m), \end{aligned} \tag{3.5}$$

$$VH_0' - (\tfrac{1}{2}mn_1 - 1)H_0 = 0, \qquad (n_2 = m+1). \tag{3.6}$$

It may be shown that these d.e.'s validly specify the distribution in (0, 1), the solutions being

$$(3.7)\qquad f_{n_1,m}(V) = k(m; n_1, m)V^{\frac{1}{2}mn_1-1}\,{}_2F_1(\tfrac{1}{2}m, \tfrac{1}{2}n_1; \tfrac{1}{2}mn_1; V), \qquad (0 < V < 1),$$
$$f_{n_1,m+1}(V) = k(m; n_1, m+1)V^{\frac{1}{2}mn_1-1}, \qquad (0 < V < 1).$$

In virtue of (3.2), these results also define $f_{m,n}$ and $f_{m+1,n}$ in the interval $(m-1, m)$, V being replaced by $(m-V)$. It must be emphasized, however, that the degenerate d.e.'s (3.5)(3.6) do not hold throughout the entire range of V (with the exception of (3.5) when $m = 2$), although the general mth order d.e. does. The situation may be illustrated in the case $m = 3$, when

$$(3.8)\qquad f_{4,3}(V) = (6/7)(3-V)^{7/2}, \qquad (2 < V < 3),$$
$$f_{4,4}(V) = (3/8)(3-V)^5, \qquad (2 < V < 3).$$

These functions are not solutions of the d.e.'s (3.5), (3.6) respectively, but by taking each in turn as H_3 in (2.12) with $n_2 = 5, 6$, they may be shown to satisfy the general 3rd order d.e. for $m = 3$. That f_{n_1,n_2} may be cusped, with discontinuous first derivative, may be seen by taking $n_1 = n_2 = 3$ in (3.4).

4. Moments of V. From [3], Section 7, the system of d.e.'s (2.7) for the Lt. $L_0(s)$ of $f_{n_1,n_2}(V)$ has characteristic roots $-(a_r+1)$ at the regular singularity $s=0$. These are all negative with the exception of $-(a_m+1) = 0$, and the system has an analytic solution at the origin as we would expect, since V has a finite range, and all its moments exist.

By virtue of (2.11), a recurrence relation for $\mathscr{E}V^r$ may be obtained from equation (7.13) of [3] for $\mathscr{E}T^r$ by replacing n_2 by $m-n_1-n_2+1$ and multiplying by $(-1)^r$, $(r = 1, 2, \cdots)$. Pillai [9] has used the first four moments of V to fit a Pearson curve to the distribution. The following reduced form of Pearson's coefficient β_2 has been derived using the above recurrence relation:

$$(4.1)\quad \beta_2 = 3(N-1)(N+2)A/mn_1n_2(N-m)(N-3)(N-2)(N+1)(N+4)(N+6),$$

where

$$N = n_1+n_2,$$
$$A = n_1n_2[(Nm-m^2)(N^3+5N^2+78N+72)-4N^2(5N+6)]$$
$$+4N^2[(m^2-Nm)(5N+6)+N(N^2+N+2)].$$

5. Itô-type expansions for large n_2. For completeness, we note that an Itô-type expansion [6] for the distribution of n_2V for large n_2 may be obtained from [3] Section 4. Noting that n_2V is asymptotically distributed as χ^2 on mn_1 degrees of freedom, a convenient approach is to expand the cumulant generating function of the statistic in a series of the type considered by Box [1]:

$$(5.1)\qquad \log L_0(s/n_2) \sim -\tfrac{1}{2}mn_1 \log(1+2s) + \textstyle\sum_{r=1}^{\infty} \omega_{r,V}[(1+2s)^{-r}-1].$$

Using the differential equations, the following set of recurrence relations may be obtained for the $\omega_{r,V}$:

(5.2) $$2r\omega_{r,V} = 2(r-1)\omega_{r-1,V} + mn_1\delta_{1,r} - (1-(m+1)/n_2)\xi_{1,r}, \qquad (r = 1, 2, \cdots),$$

where the $\xi_{j,r}$ are defined by

$$\xi_{0,r} = \xi_{r,0} = \delta_{0,r},$$

(5.3) $$\begin{aligned} j\xi_{j,r} = {} & \alpha_j\xi_{j-1,r-1} + (\beta_j + 2(r-1))\xi_{j,r-1}/n_2 \\ & + [(j+1)/n_2 - \gamma_j/n_2^2]\xi_{j+1,r-1} - [mn_1 + 2(r-2)]\xi_{j,r-2}/n_2 \\ & + 2n_2^{-1}\textstyle\sum_{s=1}^{r-2} s\omega_{s,V}(\xi_{j,r-s-1} - \xi_{j,r-s-2}), \\ & \qquad (j = 1, \cdots, m;\ r = 1, 2, \cdots), \end{aligned}$$

$$\alpha_j = (m-j+1)(n_1-j+1), \qquad \beta_j = j(2m+n_1-2j+2), \qquad \gamma_j = (j+1)(m-j+1),$$

ξ's with negative subscripts being zero. Thus, in particular,

(5.4) $$\begin{aligned} \omega_{1,V} &= mn_1(m+1)/2n_2, \\ \omega_{2,V} &= -\tfrac{1}{4}mn_1[(m+n_1+1)/n_2 - (m+1)(2m+n_1+2)/n_2^2]. \end{aligned}$$

The first six ω's to order n_2^{-3} have been derived by Muirhead [7] using an independent approach, and the first eight to order n_2^{-4} by the present author. An analogue of Itô's expansion of T_0^2 percentiles in terms of $\chi^2_{mn_1}$ percentiles may be derived from a general Cornish–Fisher inversion of Box-type series given by the author [4]. To order n_2^{-2},

(5.5) $$\begin{aligned} n_2V \sim {} & \chi^2 + 1/2n_2[\chi^2(m-n_1+1) - \chi^4(m+n_1+1)/(mn_1+2)] \\ & + 1/24n_2^2\{\chi^2[7m^2 - 12m(n_1-1) + (7n_1^2 - 12n_1 + 1)] \\ & - \chi^4[11m^2 + 24m - 13n_1^2 + 17]/(mn_1+2) \\ & + 2\chi^6[2m^3n_1 + m^2(2n_1 + 3n_1 + 10) + m(2n_1^3 + 3n_1^2 + 17n_1 + 18) \\ & + 2(5n_1^2 + 9n_1 + 2)]/(mn_1+2)^2(mn_1+4) \\ & + 6\chi^8(m-1)(m+2)(n_1-1)(n_1+2)/(mn_1+2)^2(mn_1+4)(mn_1+6)\} \\ & + O(n_2^{-3}), \end{aligned}$$

and the n_2^{-3} term has also been obtained.

An expansion of the type (5.1) also exists for T_0^2. In view of (2.11), the following relationship exists between the coefficients $\omega_{r,T}$ in this series and the $\omega_{r,V}$:

(5.6) $$\begin{aligned} (-n_2)^r\omega_{r,T} = {} & mn_1(n_1-m-1)^r/2r \\ & + \textstyle\sum_{s=1}^{r}\binom{r-1}{s-1}(m-n_1-n_2+1)^s(n_1-m-1)^{r-s}\omega_{s,V}, \end{aligned}$$

where, in the $\omega_{s,V}$, n_2 is to be replaced by $m-n_1-n_2+1$. T and V may also be interchanged in this formula. The $\omega_{r,T}$ obtained from (5.6) and (5.2) check with those obtained to order n_2^{-3} by Muirhead (*loc. cit.*).

6. Examination of the approximations. In principle, the solution of (2.12) at the regular singularity $V=0$ specifies the distribution of V in (0, 1) (or in $(m-1, m)$ when n_1 and n_2 are interchanged). For sufficiently large n_2 (or n_1, respectively), the upper 5% and 1% points of V lie in these intervals, and some investigation may be made of the accuracy of the available approximations. A corresponding study has been made for T in [5], where the d.e. was used to compute accurate percentage points by analytic continuation of the solution at $T=0$. The same computer program, with the trivial modification (2.13), has been used to tabulate some percentiles of V in the range $m \leqq 5$, n_1 and $n_2 \leqq 200$. Except when n_1 and n_2 are both small integers, Pillai's Pearson curve approximation is accurate to four decimal places. The Itô-type approximation (5.5) is considerably improved by adding the n_2^{-3} term, and is a useful direct formula for large n_2 and small n_1, but its accuracy falls off rapidly as n_1 increases. In virtue of (3.2), a similar statement holds with n_1 and n_2 interchanged.

REFERENCES

[1] Box, G. E. P. (1958). A general distribution theory for a class of likelihood criteria. *Biometrika* **36** 317–46.

[2] Constantine, A. G. (1966). The distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **37** 215–225.

[3] Davis, A. W. (1968). A system of linear differential equations for the distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **39** 815–832.

[4] Davis, A. W. (1970). Cornish–Fisher-type inversions of Box's asymptotic series. To be submitted.

[5] Davis, A. W. (1970). Exact distributions of Hotelling's generalized T_0^2. *Biometrika* **57** 187–191.

[6] Itô, K. (1956). Asymptotic formulae for the distribution of Hotelling's generalized T_0^2 statistic. *Ann. Math. Statist.* **27** 1091–1105.

[7] Muirhead, R. J. (1970). Asymptotic distributions of some multivariate tests. *Ann. Math. Statist.* **41**, 1002–1010.

[8] Pillai, K. C. S. (1955). Some new test criteria in multivariate analysis. *Ann. Math. Statist.* **27** 117–121.

[9] Pillai, K. C. S. (1960). *Statistical Tables for Tests of Multivariate Hypotheses.* The Statistical Center Univ. of the Philippines

11

The Annals of Mathematical Statistics
1972, Vol. 43, No. 5, 1664-1670

ON THE MARGINAL DISTRIBUTIONS OF THE LATENT ROOTS OF THE MULTIVARIATE BETA MATRIX[1]

BY A. W. DAVIS

C.S.I.R.O., Adelaide, and University of North Carolina

The marginal distributions of the latent roots of the multivariate beta matrix are shown to constitute a complete system of solutions of an ordinary differential equation (d.e.), which is related to the author's d.e.'s for Hotelling's generalized T_0^2 and Pillai's $V^{(m)}$ statistics. Results may be derived for the latent roots of the multivariate F and Wishart matrices ($\Sigma = I$). Pillai's approximations to the distributions of the largest and smallest roots are interpreted as exact solutions, the contributions of higher order solutions being neglected.

1. Introduction. Let $S(m \times m)$ and $T(m \times m)$ have independent Wishart distributions $W(q, \Sigma)$ and $W(n, \Sigma)$, respectively, where Σ is the population covariance matrix and $q, n \geqq m$. The latent roots $l_1 > \cdots > l_m > 0$ of the multivariate beta matrix $B = S(S+T)^{-1}$ are well known to have the joint density function

$$(1) \qquad \phi_{m;q,n}(\mathbf{l}) = K(m; q, n) \prod_{i=1}^m l_i^{\frac{1}{2}(q-m-1)} \prod_{i=1}^m (1-l_i)^{\frac{1}{2}(n-m-1)} \prod_{i<j} (l_i - l_j)\,,$$

where $\mathbf{l} = (l_1, \cdots, l_m)'$ and

$$(2) \qquad K(m; q, n) = \pi^{\frac{1}{2}m} \prod_{i=0}^{m-1} [\Gamma(\tfrac{1}{2}(q+n-i))/\Gamma(\tfrac{1}{2}(m-i))\Gamma(\tfrac{1}{2}(q-i))\Gamma(\tfrac{1}{2}(n-i))]\,.$$

The marginal distributions of the individual l_i have been investigated by Roy [14], [15], who showed that the largest root l_1 is of basic importance in testing hypotheses and constructing confidence regions in multivariate analysis of variance; also by Pillai [10], Khatri [8], Sugiyama and Fukutomi [17], Sugiyama [16], and Al-Ani [1]. Pillai [11] gave very accurate approximations to the upper and lower tails of the distributions of l_1 and l_m, respectively, and l_1 has been extensively tabulated by Heck [7] for $m \leqq 5$, using Pillai's approximation, and Pillai ([11], [12], etc.) for $m \leqq 20$. Studies of the noncentral distributions have been made by Khatri [9] and Pillai and Dotson [13].

As $n \to \infty$, $nB \to W_I$, say, having the distribution $W(q, I)$ where $I(m \times m)$ is the unit matrix. Hanumara and Thompson [6] have tabulated the largest and smallest roots of W_I using limiting forms of Pillai's approximations, and discussed their application.

The present author [3], [5] has shown that the null distributions for tr B (Pillai's $V^{(m)}$) and tr F (Hotelling's generalized T_0^2), where $F = ST^{-1}$, satisfy certain ordinary linear differential equations (d.e.'s) of order m which are related by a simple transformation. In Section 2 it is shown that the marginal

Received July 10, 1970; revised May 17, 1971.

[1] This research was supported by the National Institutes of Health Grant No. GM-12868.

distributions of the l_i form a complete system of solutions of a similar d.e. Thus, the power-series of Sugiyama and Fukutomi are solutions at the regular singularities 0 and 1. Pillai's approximations are also shown in Section 4 to be exact solutions of the d.e., but approximations to the distributions insofar as contributions from higher-order solutions are neglected. Corresponding results for the latent roots of F and W_I are readily deduced.

2. The differential equation. Let $D^r(s, l) = \{0 < x_r < \cdots < x_s < l < x_{s-1} < \cdots < x_1 < 1\} \subset R^r$, where R is the real line. The marginal density function $f_s(l)$ of l_s is given by

$$f_s(l) = \int_{D^{m-1}(s,l)} \phi_{m;q,n}(x_1, \cdots, x_{s-1}, l, x_s, \cdots x_{m-1})\, d\mathbf{x} \tag{3}$$

where $d\mathbf{x} = \prod_{i=1}^{m-1} dx_i$; it is proportional to

$$l^{\frac{1}{2}(q-m-1)}(1-l)^{\frac{1}{2}(n-m-1)} \int_{D^{m-1}(s,l)} \Phi(\mathbf{x}) \prod_{i=1}^{m-1} (l - x_i)\, d\mathbf{x}\,, \tag{4}$$

in which Φ denotes $\phi_{m-1;q-1,n-1}$. Define

$$\Psi_r(l; x) = \Phi(\mathbf{x}) \sum_\alpha (l - x_{\alpha(1)}) \cdots (l - x_{\alpha(m-1-r)})\,, \quad (r = 0, 1, \cdots, m-1)\,, \tag{5}$$

the summation being extended over the $\binom{m-1}{r}$ selections of integers $\alpha(1) < \cdots < \alpha(m-1-r)$ from the set $1, 2, \cdots, m-1$. When $r = m-1$, the sum is taken to be unity. We now introduce the m functions

$$L_{s,r}(l) = \int_{D^{m-1}(s,l)} \Psi_r(l; \mathbf{x})\, d\mathbf{x}\,, \qquad (r = 0, 1, \cdots, m-1)\,, \tag{6}$$

noting that $f_s(l)$ is proportional to $l^{\frac{1}{2}(q-m-1)}(1-l)^{\frac{1}{2}(n-m-1)}L_{s,0}$. Our object is to show that, for each s, the $L_{s,r}$ are related by a system of first-order differential equations which are independent of s. Differentiating (6),

$$L'_{s,r}(l) = -Z_{s,r}^{(1)} + Z_{s,r}^{(2)} + (r+1)L_{s,r+1}\,, \tag{7}$$

where

$$\begin{aligned} Z_{s,r}^{(1)} &= \int_{D^{m-2}(s-1,l)} \Psi_r(l; x_1, \cdots, x_{s-2}, l, x_{s-1}, \cdots, x_{m-2})\, d\mathbf{x}\,, \\ Z_{s,r}^{(2)} &= \int_{D^{m-2}(s,l)} \Psi_r(l; x_1, \cdots, x_{s-1}, l, x_s, \cdots, x_{m-2})\, d\mathbf{x}\,. \end{aligned} \tag{8}$$

Now let $\beta(1), \cdots, \beta(r)$ denote the set of subscripts complementary to $\alpha(1), \cdots, \alpha(m-1-r)$. We have

$$\begin{aligned} rlL_{s,r} &= \int_{D^{m-1}(s,l)} \Phi(\mathbf{x}) \sum_\alpha (l - x_{\alpha(1)}) \cdots (l - x_{\alpha(m-1-r)}) \\ &\qquad \times [(l - x_{\beta(1)}) + \cdots + (l - x_{\beta(r)}) + (x_{\beta(1)} + \cdots + x_{\beta(r)})]\, d\mathbf{x} \\ &= (m-r)L_{s,r-1} + \Theta_{s,r}\,, \end{aligned} \tag{9}$$

say. Integration by parts with respect to the $x_{\beta(i)}$ yields

$$\tfrac{1}{2}(q + n - 2m + 2)\Theta_{s,r} = l(1-l)[Z_{s,r}^{(1)} - Z_{s,r}^{(2)}] + \tfrac{1}{2}r(q - m + r)L_{s,r} + \Psi_{s,r}\,, \tag{10}$$

where

$$\begin{aligned} \Psi_{s,r} = \int_{D^{m-1}(s,l)} &\Phi(\mathbf{x}) \sum_\alpha (l - x_{\alpha(1)}) \cdots \\ &(l - x_{\alpha(m-1-r)}) \sum_{j=1}^{r} x_{\beta(j)} (1 - x_{\beta(j)}) \sum_{k \neq \beta(j)} (x_{\beta(j)} - x_k)^{-1}\, d\mathbf{x}\,. \end{aligned} \tag{11}$$

A term similar to (11) occurred in the derivation of a d.e. for Hotelling's generalized T_0^2 ([3] (2.13)), and the same approach yields

$$\Psi_{s,r} = \tfrac{1}{2}(m-r)(m+r-3)L_{s,r-1} + \tfrac{1}{2}r(r-1)(1-2l)L_{s,r} + \tfrac{1}{2}r(r+1)l(1-l)L_{s,r+1}\,. \tag{12}$$

Finally, eliminating the Z's, Θ's and Ψ's from (7), (9), (10) and (12), we find that

$$\begin{aligned} l(1-l)L'_{s,r} &= \tfrac{1}{2}(m-r)(q+n-m+r-1)L_{s,r-1} \\ &\quad + \tfrac{1}{2}r[(1-l)(q-m+r) - l(n-m+r)]L_{s,r} \\ &\quad + \tfrac{1}{2}(r+1)(r+2)l(1-l)L_{s,r+1}\,, \\ &\qquad (r = 0, 1, \cdots, m-1)\,, \end{aligned} \tag{13}$$

where $L_{s,-1} \equiv L_{s,m} \equiv 0$.

We observe that the system (13) is independent of s, and in principle one could successively eliminate $L_{s,1}, \cdots, L_{s,m-1}$, arriving at a homogeneous linear d.e. of order m having each $L_{s,0}$ as a solution. Clearly the f_s will be solutions of a similar d.e.; furthermore, they will constitute a linearly independent and hence complete system of solutions, since as $l \to 0^+$

$$f_s(l)/l^{\frac{1}{2}(q-m-1)} \sim k_s(m; q, n) l^{\frac{1}{2}(m-s)(q-s+2)}\,, \qquad (s = 1, \cdots, m)\,, \tag{14}$$

where

$$k_s(m; q, n) = K(m; q, n)/[K(s-1; q+m-s+1, n-m+s-1) \times K(m-s; q-s, m-s+3)]\,. \tag{15}$$

This is easily proved by writing $x_j = lw_j$ $(j = s, \cdots, m-1)$ in (4) and letting $l \to 0$. We note in addition that (13) is invariant under

$$q \to n\,, \qquad n \to q\,, \qquad l \to 1-l\,, \tag{16}$$

provided that we also replace $L_{s,r}$ by $(-1)^r L_{s,r}$; this reflects the obvious result that (16) transforms $f_s(l)$ into $f_{m+1-s}(l)$.

3. Solutions of the d.e. It is convenient to introduce $H_r = (1-l)^r L_{s,r}$ $(r = 0, 1, \cdots, m-1)$ and to express (13) as a matrix d.e. for $H = (H_0, \cdots, H_{m-1})'$:

$$dH/dl = [l^{-1}A + (1-l)^{-1}C]H\,, \tag{17}$$

where

$$A = \begin{bmatrix} a_0 & & & 0 \\ b_1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & b_{m-1} & a_{m-1} \end{bmatrix}, \qquad C = \begin{bmatrix} c_0 & d_0 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & d_{m-2} \\ 0 & & & c_{m-1} \end{bmatrix} \tag{18}$$

$$a_r = \tfrac{1}{2}r(q-m+r)\,, \qquad b_r = \tfrac{1}{2}(m-r)(q+n-m+r-1)$$
$$c_r = -\tfrac{1}{2}r(n-m+r+2)\,, \qquad d_r = \tfrac{1}{2}(r+1)(r+2)\,.$$

The d.e. (17) is of Fuchsian type, with regular singularities at $l = 0$, 1 and infinity, and we refer to [2], Chapter 4, for the general theory of such d.e.'s.

Assuming a series solution $H = \sum_{r=0}^{\infty} h_r l^{\rho+r}$ in $|l| < 1$, we obtain $Ah_0 = \rho h_0$, so that ρ must be one of the latent roots $a_0, \cdots, a_{m-1}$ of A, and h_0 the corresponding latent vector. To relate this fundamental set of solutions to the $f_s(l)$, we first obtain a non-singular transformation $H = PM$, where $P(m \times m)$ is independent of l, such that $P^{-1}AP = \text{diag}(a_r)$. A suitable choice is

$$P = \{p_{ij}\}, \quad p_{ij} = (-1)^{i-j}\binom{m-j-1}{i-j} \prod_{r=j}^{i-1} (q+n-m+r)/(q-m+j+r+1), \tag{19}$$

with inverse

$$P^{-1} = \{p_{ij}^*\}, \qquad p_{ij}^* = \binom{m-j-1}{i-j} \prod_{r=j}^{i-1} (q+n-m+r)/(q-m+i+r). \tag{20}$$

Both P and P^{-1} are lower triangular, and $M_0 = H_0$. It may be shown that

$$P^{-1}CP = G = \begin{bmatrix} \mu_0, & \nu_0 & & & 0 \\ \lambda_1, & \mu_1 & \nu_1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \nu_{m-2} \\ 0 & & & \lambda_{m-1} & \mu_{m-1} \end{bmatrix} \tag{21}$$

where

$$\begin{aligned} \lambda_i &= (m-i)(n-i)(q+i-1)(q-m+i-1)(q-m+i-2) \\ &\quad \times (q+n-m+i-1) \\ &\quad \div [2(q-m+2i-2)(q-m+2i-1)^2(q-m+2i)], \\ \mu_i &= i^2 - \tfrac{1}{2}i(m+n-3) - m + 1 \\ &\quad + \tfrac{1}{2}i(i+1)(m-i)(n-i)/(q-m+2i-1) \\ &\quad - \tfrac{1}{2}(i+1)(i+2)(m-i-1)(n-i-1)/(q-m+2i+1), \\ \nu_i &= \tfrac{1}{2}(i+1)(i+2). \end{aligned} \tag{22}$$

The d.e. (17) now takes the form

$$dM/dl = [l^{-1}\,\text{diag}(a_r) + (1-l)^{-1}G]M, \tag{23}$$

and assuming a solution $M = \sum_{r=0}^{\infty} \eta_r l^{a_p+r}$ corresponding to the latent root a_p of A, we obtain the following recurrence relations for the components $(\eta_{0,r}, \cdots, \eta_{m-1,r})$ of the η_r:

$$\begin{aligned} &\eta_{p,0} = 1, \qquad \eta_{i,0} = 0 \qquad (i \neq p), \\ &(r - a_i + a_p)\eta_{i,r} \\ &\quad = \lambda_i \eta_{i-1,r-1} + [\mu_i + (r-1) - a_i + a_p]\eta_{i,r-1} + \nu_i \eta_{i+1,r-1}, \\ &\qquad (i = 0, \cdots, m-1; \quad r = 1, 2, \cdots). \end{aligned} \tag{24}$$

This form of solution unfortunately breaks down if $a_i - a_p$ is a positive integer for some i. In fact, $a_{p+1} - a_p = \frac{1}{2}(q-m+1) + p$ $(p \leqq m-2)$, which is an integer if $q - m$ is odd, while $a_{p+2} - a_p = q - m + 2(p+1)$ $(p \leqq m-3)$ is always an integer. Generally in such situations the solution must be obtained by limiting procedures which may produce logarithmic terms. However, it may be seen from (6) that the $L_{s,r}$ are in fact representable by power series, and it appears that if $a_i - a_p$ is a positive integer for $i > p$, then the right-hand side of the ith equation in (24) vanishes identically when $r = a_i - a_p$. Thus $\eta_{i,r}$ is an undetermined constant introducing the a_i-solution at this stage, and the power series form is preserved.

We also see from (24) that the $\eta_{0,r}$ are zero for $r < p$ in the a_p-solution, while

(25) $$\eta_{0,p} = (p+1)!/\prod_{i=1}^{p}(q-m+p+i+1) = \xi_p, \quad \text{say}.$$

Hence $M_0(l) = O(l^{a_p+p})$ as $l \to 0^+$, and since $a_{m-s} + (m-s) = \frac{1}{2}(m-s)(q-s+2)$, it follows from (14) that $L_{0,s}$ must be some linear combination of the $a_{m-s}, a_{m-s+1}, \cdots$, and a_{m-1}-solutions. The coefficient of the a_{m-s}-solution is clearly $k_s(m; q, n)/\xi_{m-s}$, but the remaining coefficients have not been determined for general s. However, the density function $f_1(l)$ of the largest root corresponds to the largest root a_{m-1} of A, and is thus completely specified by (24). The resulting power series coincides with the result of Sugiyama and Fukutomi [17].

4. Pillai's approximations. A particular solution corresponding to the smallest root $a_0 = 0$ of A may be given explicitly as an $(m-1)$th degree polynomial. Writing $(z)_i = z(z+1)\cdots(z+i-1)$, $(z)_{-i} = z(z-1)\cdots(z-i+1)$, it may be shown that

(26) $$\eta_{i,r} = (-1)^r \binom{m-1}{r}\binom{r}{i}(q)_i(n-1)_{-i}(q+n-m)_r \div [(q-m+1)_{i+r}(q-m+i)_i] .$$

Thus we obtain the following approximation to the lower tail of $f_m(l)$ for large q by neglecting the $a_1, \cdots, a_{m-1}$ solutions (i.e., terms of order l^{q-m+1} at least):

(27) $$f_m(l) \approx k(m; q, n) l^{\frac{1}{2}(q-m-1)}(1-l)^{\frac{1}{2}(n-m-1)} \times \sum_{r=0}^{m-1}\binom{m-1}{r}(q+n-m)_r(q-1)_{-(m-1-r)}(-l)^r ,$$

where

(28) $$k(m; q, n) = \pi^{\frac{1}{2}}\Gamma(\tfrac{1}{2}(q+n-m+1))/[2^{m-1}\Gamma(\tfrac{1}{2}m)\Gamma(\tfrac{1}{2}q)\Gamma(\tfrac{1}{2}n)] .$$

Using (16), a corresponding approximation to the upper tail of $f_1(l)$ for large n (near the regular singularity $l = 1$) is obtained. The result is found to be simply the right-hand side of (27) multiplied by $(-1)^{m-1}$.

The integrated form of the approximation was arrived at by Pillai [11] using a different approach, and used in a series of tabulations of the upper 5% and 1% points of l_1. Its accuracy to essentially five places of decimals when $n_2 \geqq m + 11$ was demonstrated at least for $m \leqq 10$ by substituting in explicit expressions for the distribution function [10]. In order to investigate the usefulness of the d.e. (23), some percentage points were calculated by following the a_{m-1}-solution out from the origin, using the same computation procedure as in [4]. The method appeared to be effective at least up to $m = 7$, since on comparing the 1% points, i.e., the less accurate results of the d.e. and the more accurate results of the approximation, these were generally found to differ by no more than a unit in the fifth decimal place. On the other hand, the 5% points obtained from the d.e. tended to exceed Pillai's by about three units in the fifth decimal place. The d.e. approach should be more accurate at lower significance levels, and a tabulation of upper 10% points has been made.

5. Some remarks. The success of the Pillai approximation suggests a similar approach to the other roots, approximating the lower tail of f_s by the a_{m-s}-solution for large q, and deducing a corresponding result for the upper tail of f_{m-s+1} when n is large using (16). No general results corresponding to (27) have been obtained, but it has been found, for instance that when $m = 3$ the distribution of l_2, the median root, is closely approximated by a beta density with parameters $q - 1$ and $n - 1$. Upper 5% and 1% points based on this approximation are identical to five decimal places with those published by Pillai and Dotson [13], except where the latter have employed interpolation.

Differential equations for the latent roots of F and W_I (defined in Section 1) are readily deduced from (13). The approximation used by Hanumara and Thompson [6] corresponds of course to an exact solution of the d.e. in the Wishart case. The Wishart d.e. is in fact closely related to the author's d.e. [5] for the moment generating function of Pillai's $V^{(m)} = \operatorname{tr} B$. If we write $\lambda_{m;q,n} = E\exp(-sV^{(m)})$, then it is easily seen that the density of the largest root of W_I is proportional to $e^{-\frac{1}{2}u}u^{\frac{1}{2}mq-1}\lambda_{m-1;q-1,m+2}(\frac{1}{2}u)$.

Acknowledgment. The author's thanks are due to Mrs. Darien Wei and Dr. R. W. Helms of the Department of Biostatistics, School of Public Health, University of North Carolina, for their computing assistance.

REFERENCES

[1] Al-Ani, S. (1968). On the distribution of the ith latent root under null hypotheses concerning complex multivariate normal populations. Mimeo Series 145, Department of Statistics, Purdue University.

[2] Coddington, E. A. and Levinson, N. (1955). *Theory of Ordinary Differential Equations.* McGraw-Hill, New York.

[3] Davis, A. W. (1968). A system of linear differential equations for the distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **39** 815–832.

[4] Davis, A. W. (1970a). Exact distributions of Hotelling's generalized T_0^2. *Biometrika* **57** 187–191.

[5] Davis, A. W. (1970b). On the null distribution of the sum of the roots of a multivariate beta distribution. *Ann. Math. Statist.* **41** 1557–1562.

[6] Hanumara, R. C. and Thompson, W. A. (1968). Percentage points of the extreme roots of a Wishart matrix. *Biometrika* **55** 505–512.

[7] Heck, D. L. (1960). Charts of some upper percentage points of the distribution of the largest characteristic root. *Ann. Math. Statist.* **31** 625–642.

[8] Khatri, C. G. (1964). Distribution of the largest or the smallest characteristic root under null hypothesis concerning complex multivariate normal populations. *Ann. Math. Statist.* **35** 1807–1810.

[9] Khatri, C. G. (1969). Non-central distributions of ith largest characteristic roots of three matrices concerning complex multivariate normal populations. *Ann. Inst. Statist. Math.* **21** 23–32.

[10] Pillai, K. C. S. (1954). On some distribution problems in multivariate analysis. Mimeo Series 88, Institute of Statistics, University of North Carolina.

[11] Pillai, K. C. S. (1965). On the distribution of the largest characteristic root of a matrix in multivariate analysis. *Biometrika* **52** 405–414.

[12] Pillai, K. C. S. (1967). Upper percentage points of the largest root of a matrix in multivariate analysis. *Biometrika* **54** 189–194.

[13] Pillai, K. C. S. and Dotson, C. O. (1969). Power comparisons of tests of two multivariate hypotheses based on individual characteristic roots. *Ann. Inst. Statist. Math.* **21** 49–66.

[14] Roy, S. N. (1945). The individual sampling distribution of the maximum, the minimum and any intermediate of the p-statistics on the null hypothesis. *Sankhyā* **7** 133–158.

[15] Roy, S. N. (1957). *Some Aspects of Multivariate Analysis.* Wiley, New York.

[16] Sugiyama, T. (1967). Distribution of the largest latent root and the smallest latent root of the generalized B statistic and F statistic in multivariate analysis. *Ann. Math. Statist.* **38** 1152–1159.

[17] Sugiyama, T. and Fukutomi, K. (1967). On the distribution of the extreme characteristic roots of the matrices in multivariate analysis. *Rep. Statist. Appl. Res. Un. Japan. Sci. Eng.* **14** 156–160.

C.S.I.R.O.
Division of Mathematical Statistics
Glen Osmond
South Australia 5064

Reprinted from JOURNAL OF MULTIVARIATE ANALYSIS
All Rights Reserved by Academic Press, New York and London

Vol. 2, No. 2, June 1972
Printed in Belgium

12

On the Distributions of the Latent Roots and Traces of Certain Random Matrices

A. W. DAVIS

Division of Mathematical Statistics, C.S.I.R.O., Adelaide, Australia

It is shown that differential equations given by the author may be used recursively to construct certain multivariate null distributions in reduced form. These include the distributions of individual latent roots of $\mathbf{B} = \mathbf{S}_1(\mathbf{S}_1 + \mathbf{S}_2)^{-1}$, and distributions of Tr $\mathbf{B}$ and Tr $\mathbf{S}_1\mathbf{S}_2^{-1}$, for small numbers of variates.

1. INTRODUCTION

Let $\mathbf{S}_1(m \times m)$ and $\mathbf{S}_2(m \times m)$ have independent central Wishart distributions with n_1, n_2 degrees of freedom, respectively ($n_1, n_2 \geqslant m$), and the same population covariance matrix $\mathbf{\Sigma}$. Denote the multivariate F and beta matrices by

$$\mathbf{F} = \mathbf{S}_1\mathbf{S}_2^{-1}, \qquad \mathbf{B} = \mathbf{S}_1(\mathbf{S}_1 + \mathbf{S}_2)^{-1}, \tag{1.1}$$

respectively. General results are now available for certain statistics which are expressible in terms of the latent roots of these matrices. We shall be concerned in this paper with the null distributions of the statistics $T = \text{Tr}\,\mathbf{F}$, $V = \text{Tr}\,\mathbf{B}$, and the individual latent roots $l_1 > \cdots > l_m > 0$ of $\mathbf{B}$. Reduction procedures for the exact distributions of the roots l_i were first considered by Roy [18], and later by Nanda [12, 13]. Pillai [15] gave expressions in the important case l_1 for $m \leqslant 10$. Lawley [10] introduced the statistic T and Hotelling [6] derived its exact distribution for $m = 2$. Nanda [14] and Pillai [15] considered V and its distribution for small m. The present author has shown [1, 4, 5] that the density functions of these statistics are solutions of ordinary differential equations (DE's). The power series solutions of these DE's have been applied to the derivation of approximations, and to the accurate tabulation of T[2, 3]. Some unpublished

Received July 22, 1971; revised February 14, 1972.

AMS 1970 subject classification: Primary 62H10; secondary 62E15.

Key words and phrases: exact distributions; random matrices; individual latent roots; trace; differential equations; Laplace transforms.

tabulations have also been made of V and l_1, and compared with certain approximations [4, 5].

Recently, Krishnaiah and his associates [7, 8], using a very elegant method based on [11], have expressed the distributions of the l_k (and subsets of them) as finite linear combinations of products of double integrals, for arbitrary m. These results are not only of theoretical importance, but are well adapted for computing purposes, and this work is in progress. Pillai and his associates [16, 17] and Krishnaiah and Chang [9] have also applied their methods to the traces T and V.

Any of these approaches seems likely to encounter heavy algebra when we seek to reduce these distributions to their simplest and most compact analytic forms. In the present paper, we present expressions for the marginal distributions of the individual l_k for $m \leqslant 5$ which appear to be simpler than those published previously (Appendix A). A DE is applied recursively (Section 2) to construct the distributions for m from those obtained for $m - 1$. Results are derived for the roots of the Wishart matrix $\mathbf{W}_I = \mathbf{S}_1\mathbf{\Sigma}^{-1}$ by letting n_2 approach infinity (Appendix B). In Section 3, we show that the distribution of the largest root of $\mathbf{W}_I$ may be used in conjunction with another DE to construct Laplace transforms of the density of V (Appendix C). Results may then be deduced for T (Section 4). This approach also involved extensive calculations, but it was felt that the method provided some guidance as to the structure of the distributions, as well as a scheme of relationships among them. It is to be emphasized that the method is not being presented here as a computing algorithm. Our intention in this paper is to present exact distributions which have been reduced as far as possible algebraically, in cases where this has been found feasible. Such results would appear to have some intrinsic interest, and may possibly serve as a reference for future work.

2. Marginal Distributions of the Roots of B and $\mathbf{W}_I$

The l_k have the well-known joint density

$$\phi_{m;n_1,n_2}(l^{(m)}) = K(m; n_1, n_2) \prod_{i=1}^{m} l_i^{\frac{1}{2}(n_1-m-1)}(1 - l_i)^{\frac{1}{2}(n_2-m-1)} \prod_{i<j} (l_i - l_j), \tag{2.1}$$

where $0 < l_m < \cdots < l_1 < 1$, $l^{(m)} = (l_1, \ldots, l_m)'$, and

$$K(m; n_1, n_2) = \pi^{\frac{1}{2}m} \prod_{i=0}^{m-1} [\Gamma(\tfrac{1}{2}(n_1 + n_2 - i))/\Gamma(\tfrac{1}{2}(m - i))\, \Gamma(\tfrac{1}{2}(n_1 - i))\, \Gamma(\tfrac{1}{2}(n_2 - i))]. \tag{2.2}$$

Letting $D^m(k, l)$ denote the region $\{0 < l_m < \cdots < l_k < l < l_{k-1} < \cdots < l_1 < 1\}$ the marginal density of l_k is given by

$$f^{(k)}_{m;n_1,n_2}(l) = \int_{D^{m-1}(k,l)} \phi_{m;n_1,n_2}(l_1, \ldots, l_{k-1}, l, l_k, \ldots, l_{m-1})\, d\boldsymbol{l}^{(m-1)}, \quad (2.3)$$

where $d\boldsymbol{l}^{(m-1)} = \prod_{i=1}^{m-1} dl_i$. Now define

$$\Psi_r(l, \boldsymbol{l}^{(m-1)}) = \phi_{m-1;n_1-1,n_2-1}(\boldsymbol{l}^{(m-1)}) \sum_{\alpha} (l - l_{\alpha(1)}) \cdots (l - l_{\alpha(m-1-r)}), \quad (2.4)$$

the summation being extended over the $\binom{m-1}{r}$ selections of integers

$$\alpha(1) < \cdots < \alpha(m - 1 - r)$$

from the set $1, 2, \ldots, m - 1$. When $r = m - 1$, the sum is taken to be unity. If we introduce the auxiliary functions

$$H_{k,r}(l) = \int_{D^{m-1}(k,l)} \Psi_r(l, \boldsymbol{l}^{(m-1)})\, d\boldsymbol{l}^{(m-1)}, \qquad (r = 0, \ldots, m - 1), \quad (2.5)$$

then

$$f^{(k)}_{m;n_1,n_2}(l) = C(m; n_1, n_2)\, l^{\frac{1}{2}(n_1-m-1)}(1 - l)^{\frac{1}{2}(n_2-m-1)} H_{k,0}(l), \quad (2.6)$$

where

$$C(m; n_1, n_2) = (-1)^{k-1} K(m; n_1, n_2)/K(m - 1; n_1 - 1, n_2 - 1). \quad (2.7)$$

It may be shown [5] that the $H_{k,r}$ satisfy the system of DE's:

$$\begin{aligned} l(1 - l)\, H'_{k,r} &= \tfrac{1}{2}(m - r)(n_1 + n_2 - m + r - 1)\, H_{k,r-1} + \tfrac{1}{2}r[(1 - l)(n_1 - m + r) \\ &\quad - l(n_2 - m + r)]\, H_{k,r} + \tfrac{1}{2}(r + 1)(r + 2)\, l(1 - l)\, H_{k,r+1}, \end{aligned} \quad (2.8)$$

($r = 0, \ldots, m - 1$; $H_{k,-1} \equiv H_{k,m} \equiv 0$). This system is independent of the particular root l_k.

The basis of the present method is that the $(m - 1)$-th auxiliary function

$$H_{k,m-1}(l) = \int_{D^{m-1}(k,l)} \phi_{m-1;n_1-1,n_2-1}(\boldsymbol{l}^{(m-1)})\, d\boldsymbol{l}^{(m-1)} \quad (2.9)$$

is the probability that $l_k < l < l_{k-1}$ when the parameters are $m - 1$, $n_1 - 1$,

$n_2 - 1$. Hence if we let $F^{(k)}_{m;n_1,n_2}(l)$ denote the cumulative distribution function (CDF) of l_k, ($F^{(0)} \equiv 0, F^{(m+1)} \equiv 1$), we have

$$H_{k,m-1}(l) = F^{(k)}_{m-1;n_1-1,n_2-1}(l) - F^{(k-1)}_{m-1;n_1-1,n_2-1}(l), \tag{2.10}$$

so that if the CDF's of the roots are known for $m - 1$, they may be used to compute the density functions for m, via (2.8) and (2.6). $H_{k,m-2}$ is first derived, using the $r = m - 1$ equation, and so on until we obtain $H_{k,0}$ (hence $f^{(k)}_{m;n_1,n_2}$) from the $r = 1$ DE. The calculations may then be checked by substituting in the $r = 0$ DE. We note that the distribution of the largest root for m is obtained from the corresponding distribution for $m - 1$. Since

$$F^{(k)}_{m;n_1,n_2}(l) = 1 - F^{(m-k+1)}_{m;n_2,n_1}(1 - l), \tag{2.11}$$

it is sufficient to consider $k \leqslant \frac{1}{2}(m + 1)$.

It should also be pointed out that, although substitution in the DE to obtain the density function is a mechanical, if tedious, process, the integration and algebraic reduction of this result to derive the CDF is in general far from straightforward. A basic tool in the reduction is the identity

$$(x + y + 1) \int_0^l u^x(1 - u)^y \, du = -l^x(1 - l)^{y+1} + x \int_0^l u^{x-1}(1 - u)^y \, dy. \tag{2.12}$$

Results for $m = 2$ through 5 are listed in Appendix A. It turns out that these distributions may be concisely expressed in terms of the unstandardized beta density

$$\beta_0(an + b, l) = l^{an_1+b}(1 - l)^{an_2+b}, \qquad (0 < l < 1), \tag{2.13}$$

($n = (n_1, n_2)$), and its r-fold integrals

$$\begin{aligned} \beta_r(an + b, l) &= \Gamma(r)^{-1} \int_0^l (l - u)^{r-1} \beta_0(an + b; u) \, du, \\ \bar{\beta}_r(an + b, l) &= \Gamma(r)^{-1} \int_l^1 (u - l)^{r-1} \beta_0(an + b; u) \, du. \end{aligned} \tag{2.14}$$

For positive integral r, these may obviously be expanded in terms of the incomplete beta functions $\beta \equiv \beta_1$ and $\bar{\beta} \equiv \bar{\beta}_1$, respectively. The ordinary beta function $\beta(an + b, 1)$ will be denoted $B(an + b + 1)$.

EXAMPLE. An interesting result is obtained for the median root l_2 when $m = 3$. From (A.1) and (A.2),

$$\begin{aligned} H_{2,2} &= F^{(2)}_{2;n_1-1,n_2-1} - F^{(1)}_{2;n_1-1,n_2-1} \\ &= \tfrac{1}{2}B(n-2)^{-1}\, B(\tfrac{1}{2}n - 1)\, \beta_0(\tfrac{1}{2}n - 1, l), \end{aligned} \tag{2.15}$$

and we obtain successively from the DE (2.8) (omitting the multiplicative constant in (2.15)):

$$\begin{aligned} (n_1 + n_2 - 2)\, H_{2,1} &= [-n_1(1-l) + n_2 l]\, \beta_0(\tfrac{1}{2}n - 1, l), \\ (n_1 + n_2 - 2)\, H_{2,0} &= -2\beta_0(\tfrac{1}{2}n, l), \\ f^{(2)}_{3;n_1,n_2} &= B(n-1)^{-1}\, \beta_0(n-2, l). \end{aligned} \tag{2.16}$$

Thus l_2 has a simple beta distribution with parameters $n_1 - 1$, $n_2 - 1$, and the result mentioned in [5, Section 5] as an approximation is seen to be exact. This was first proved by S. Eckert (unpublished work) in the Wishart case.

The form of the results in Appendix A makes it a simple matter to deduce expressions for the individual roots of the Wishart matrix $\mathbf{W}_I$. Appendix B lists formulas for the CDF of the largest root,

$$G_{m;q}(l) = \lim_{n_2 \to \infty} F^{(1)}_{m;q,n_2}(l/n_2). \tag{2.17}$$

These are given partly for their own interest and partly because they will be required in Section 3. The results are expressed in terms of the gamma density and its integrals:

$$\gamma_0(aq + b, l) = e^{-al} l^{aq+b}, \quad \gamma_r(aq + b, l) = \Gamma(r)^{-1} \int_0^l (l-u)^{r-1}\, \gamma_0(aq + b, u)\, du \tag{2.18}$$

($\gamma_1 \equiv \gamma$). Clearly γ_r may be expanded in terms of incomplete gamma functions for positive integral r.

3. THE STATISTIC $V = \text{TR}\,\mathbf{B}$

The Laplace transform (Lt) of the density function of V will be denoted by

$$\Lambda_{m;n_1,n_2}(s) = \mathscr{E} \exp\left(-s \sum l_i\right). \tag{3.1}$$

It is easy to show (Nanda [14]) that in the case $n_2 = m + 1$,

$$\Lambda_{m;n_1,m+1}(s) = \kappa(m; n_1)\, s^{-\frac{1}{2}mn_1} G_{m;n_1}(2s), \tag{3.2}$$

where $G_{m;n_1}(l)$ is the CDF of the largest root of the Wishart matrix $\mathbf{W}_l$ on n_1 degrees of freedom, and

$$\kappa(m; n_1) = \prod_{i=2}^{m+1} [\Gamma(\tfrac{1}{2}(n_1 + i))/\Gamma(\tfrac{1}{2}i)]. \tag{3.3}$$

Thus, for $m = 2$ through 5, $\Lambda_{m;n_1,m+1}(s)$ may be obtained directly from Appendix B. We note that

$$\gamma_r(an_1 + b, 2s) = \Gamma(r)^{-1}(s/a)^{an_1+b+r}\lambda(an_1 + b, r - 1; s), \tag{3.4}$$

where

$$\lambda(an_1 + b, r; s) = \int_0^{2a} e^{-su}u^{an_1+b}(2a - u)^r \, du \tag{3.5}$$

is the Lt of the function which is equal to $u^{an_1+b}(2a - u)^r$ on $(0, 2a)$ and zero elsewhere.

Another DE of the author [4] may now be applied recursively to construct $\Lambda_{m;n_1,n_2}$ from $\Lambda_{m;n_1,m+1}$ when $n_2 - (m + 1)$ is a positive even integer. Introduce the Lt's

$$L_r(s) = \int_{D^m} \exp\left(-s \sum l_i\right) \phi_{m;n_1,n_2}(l^{(m)}) \sum_\alpha [(1 - l_{\alpha(1)}) \cdots (1 - l_{\alpha(r)})]^{-1} \, dl^{(m)}, \quad (r = 0, 1,..., m), \tag{3.6}$$

where D^m is the region $\{0 < l_m < \cdots < l_1 < 1\}$ and the summation is as defined in (2.4). Then writing $(z)_r = z(z + 1) \cdots (z + r - 1)$,

$$L_0(s) = \Lambda_{m;n_1,n_2}(s), \tag{3.7}$$

$$L_m(s) = [(n_1 + n_2 - m - 1)_m/(n_2 - m - 1)_m] \Lambda_{m;n_1,n_2-2}(s), \tag{3.8}$$

and for $n_2 \geqslant m + 2$ the L_r satisfy the DE's

$$-(m - r + 1) sL_{r-1} + [s(d/ds + r) + \tfrac{1}{2}(m - r)(n_1 + n_2 - m + r - 1)] L_r - \tfrac{1}{2}(r + 1)(n_2 - m + r - 1) L_{r+1} = 0, \quad (r = 0, 1,..., m - 1; L_{-1} \equiv 0), \tag{3.9}$$

$$(d/ds + m) L_m - L_{m-1} = 0. \tag{3.10}$$

It is thus possible to derive $\Lambda_{m;n_1,n_2}$ from $\Lambda_{m;n_1,n_2-2}$ by substituting (3.8) in (3.10), and successively calculating $L_{m-1},..., L_0$. The $r = 0$ DE provides a check on the working.

In applications, one is more interested in the distribution of V for large n_2 and moderate n_1. It is readily shown from the definition of V that

$$\Lambda_{m;n_1,n_2}(s) = e^{-ms}\Lambda_{m;n_2,n_1}(-s). \tag{3.11}$$

This transformation is facilitated by the relation

$$\begin{aligned}\lambda(an_1 + b, r; -s) &= e^{2as} \int_0^{2a} e^{-su} u^r (2a - u)^{an_1+b}\, du \\ &= e^{2as}\lambda(r, an_1 + b; s), \quad \text{say.}\end{aligned} \tag{3.12}$$

The distribution of V is known for $m = 2$, being expressible in terms of the Gaussian hypergeometric function (see, for example, [4, Section 3]). Further results are given in Appendix C of the present paper.

Example. When $m = 3$, we obtain from (B.2), (3.2) and (3.3)

$$\begin{aligned}\Lambda_{3;n_1,4}(s) = 2^{-n_1-2}(n_1 - 1)_4\, n_1 s^{-1}[\lambda(\tfrac{1}{2}n_1 - 1, 0; s)\, \lambda(n_1 - 2, 0; s) \\ - e^{-s}\lambda(n_1 - 2, 1; s)].\end{aligned} \tag{3.13}$$

Substituting in (3.10) and (3.9), it may be shown that

$$\begin{aligned}\Lambda_{3;n_1,6}(s) = (4!2^{n_1+3})^{-1}\, (\tfrac{1}{2}n_1)_2\, (n_1 - 1)_6\, s^{-2}\{\lambda(\tfrac{1}{2}n_1 - 1, 0; s)[s\lambda(n_1 - 2, 3; s) \\ + 3\lambda(n_1 - 2, 2; s) - 12\lambda(n_1 - 2, 1; s)] + 6e^{-s}\lambda(n_1 - 2, 2; s)\}.\end{aligned} \tag{3.14}$$

The powers of s in (3.13) and (3.14) may be eliminated by integration by parts, resulting in forms more suitable for inversion. For example,

$$\Lambda_{3;n_1,4}(s) = k(n_1)[2\lambda(\tfrac{1}{2}n_1, 0; s)\, \lambda(n_1 - 2, 0; s) - e^{s}\lambda(n_1 - 1, 1; s)], \tag{3.15}$$

where

$$k(n_1) = 2^{-n_1-2}(n_1 - 1)_4\,. \tag{3.16}$$

Using (3.11) and (3.12) we obtain (C.1); (C.2) is derived similarly from (3.14).

Denoting the density function of V by $\theta_{m;n_1,n_2}(V)$ we obtain on inverting (C.1)

$$\theta_{3;4,n_2}(V) = k(n_2)\left[2\int_0^V (1-V+u)^{\frac{1}{2}n_2}(2-u)^{n_2-2}\,du - V(2-V)^{n_2-1}\right], \quad (0 < V < 1),$$

$$\theta_{3;4,n_2}(V) = k(n_2)\left[2\int_{V-1}^V (1-V+u)^{\frac{1}{2}n_2}(2-u)^{n_2-2}\,du - V(2-V)^{n_2-1}\right], \quad (1 < V < 2),$$

$$\theta_{3;4,n_2}(V) = 2k(n_2)\int_{V-1}^2 (1-V+u)^{\frac{1}{2}n_2}(2-u)^{n_2-2}\,du, \quad (2 < V < 3). \tag{3.17}$$

The piecewise smooth nature of $\theta_{m;n_1,n_2}(G)$ arises from taking convolutions of functions which vanish outside (0, 1) or (0, 2). It is reflected also in the DE for this function [4], which has regular singularities at $V = 0, 1, \ldots, m$.

4. The Statistic $T = \mathrm{TR}\,\mathbf{F}$

Let $\psi_{m;n_1,n_2}(T)$ denote the density function of T. The DE's for θ and ψ are related by the transformation [4]:

$$V \to -T, \qquad n_2 \to \bar{n}_2 = m - n_1 - n_2 + 1, \tag{4.1}$$

and this relation also holds between their power series expansions in the unit circle about the origin. The analytic continuation of ψ into the complex plane thus has singularities at $T = 0, -1, \ldots, -m$.

It follows from (4.1) and (3.17), for example, that for $0 < T < 1$

$$\psi_{3;4,n_2}(T) = k(n_2)\left[2\int_0^T (1+T-u)^{-\frac{1}{2}n_2}(u+2)^{-n_2-2}\,du - T(T+2)^{-n_2-1}\right], \tag{4.2}$$

and by analytic continuation this expression will hold for $0 < T < \infty$. When n_2 is even, the convolution may be evaluated by partial fractions [17].

Introducing the Lt $M_{m;n_1,n_2}$ of $\psi_{m;n_1,n_2}$, and writing

$$\mu(r, an_2 + b; s) = \int_0^\infty e^{-su} u^r (2a + u)^{-an_2-b}\,du, \tag{4.3}$$

we obtain from (4.2)

$$M_{3;4,n_2}(s) = k(n_2)[2\mu(0, \tfrac{1}{2}n_2\,; s)\,\mu(0, n_2 + 2; s) - \mu(1, n_2 + 1; s)]. \tag{4.4}$$

Comparison with (3.16) suggests a simple formal relationship between M

and Λ. Corresponding to (4.1), the DE for M [1] is related to (3.9), (3.10) by the transformation

$$s \to -s, \qquad n_2 \to \bar{n}_2 . \tag{4.5}$$

M and Λ are not themselves related by (4.5), but the following rule may be given for obtaining M from the corresponding Λ: (1) Apply (4.5) throughout, then (2) replace $\lambda(r, a\bar{n}_2 + b; -s)$ by $(-1)^{r+1}\mu(r, -a\bar{n}_2 - b; s)$. From (C.4), we obtain, for example,

$$\begin{aligned} M_{4;5,n_2}(s) = {} & [(n_2 - 3)_6\, 2^{n_2-1}/4!][2^{n_2+3}\mu(0, n_2 - 1; s)\, \mu(0, n_2 + 3; s) \\ & + 2n_2\mu(0, \tfrac{1}{2}(n_2 - 1); s)\, \mu(2, n_2 + 3; s) \\ & - \mu(1, n_2 ; s) - 4\mu(1, n_2 + 2; s)]. \end{aligned} \tag{4.6}$$

APPENDIX A: Marginal Distributions of the Roots of B

$m = 2$

$$F^{(1)}_{2;n_1,n_2}(l) = B(n - 1)^{-1}\, [\beta(n - 2, l) - \tfrac{1}{2}\beta_0(\tfrac{1}{2}(n - 1), l)\, \beta(\tfrac{1}{2}(n - 3), l)], \tag{A.1}$$

$$F^{(2)}_{2;n_1,n_2}(l) = B(n - 1)^{-1}\, [\beta(n - 2, l) + \tfrac{1}{2}\beta_0(\tfrac{1}{2}(n - 1), l)\, \bar{\beta}(\tfrac{1}{2}(n - 3), l)]. \tag{A.2}$$

$m = 3$

$$\begin{aligned} F^{(1)}_{3;n_1\,n_2}(l) = {} & B(\tfrac{1}{2}n)^{-1}\, B(n - 1)^{-1} \\ & \times [\beta(\tfrac{1}{2}n - 1, l)\, \beta(n - 2, l) - 2\beta_0(\tfrac{1}{2}n - 1, l)\, \beta_2(n - 2, l)], \end{aligned} \tag{A.3}$$

$$F^{(2)}_{3;n_1,n_2}(l) = B(n - 1)^{-1}\, \beta(n - 2, l). \tag{A.4}$$

$m = 4$

$$\begin{aligned} F^{(1)}_{4;n_1,n_2}(l) = {} & B(n - 1)^{-1}\, B(n - 2)^{-1} \\ & \times [\beta(n - 2, l)\, \beta(n - 3, l) - \beta_0(n - 2, l)\, \beta_2(n - 3, l) \\ & - \tfrac{1}{2}(n_1 + n_2 - 4)\, \beta_0(\tfrac{1}{2}(n - 3), l)\, \beta(\tfrac{1}{2}(n - 3), l)\, \beta_3(n - 3, l)], \end{aligned} \tag{A.5}$$

$$\begin{aligned} F^{(2)}_{4;n_1,n_2}(l) = {} & B(n - 1)^{-1}\, B(n - 2)^{-1} \\ & \times [\beta(n - 2, l)\, \beta(n - 3, l) - \beta_0(n - 2, l)\, \beta_2(n - 3, l) \\ & + \tfrac{1}{2}(n_1 + n_2 - 4)\, \beta_0(\tfrac{1}{2}(n - 3), l)\, \bar{\beta}(\tfrac{1}{2}(n - 3), l)\, \beta_3(n - 3, l)]. \end{aligned} \tag{A.6}$$

$m = 5$

$$F^{(1)}_{5;n_1,n_2}(l) = [B(\tfrac{1}{2}n-1)\,B(n-1)\,B(n-3)]^{-1}\{\beta(\tfrac{1}{2}n-2,l)\,\beta(n-2,l)$$
$$\times\,\beta(n-4,l) + 2(n_1+n_2-3)^{-1}\,(n_1+n_2-5)$$
$$\times\,[\beta_0(\tfrac{1}{2}n-1,l)\,\beta(n-4,l)\,\beta_2(n-3,l) - \beta_0(\tfrac{3}{2}n-4,l)\,\beta_3(n-4,l)$$
$$-\,2(n_1+n_2-6)\,\beta_0(\tfrac{1}{2}n-2,l)\,\beta(n-3,l)\,\beta_4(n-4,l)]$$
$$-\,(n_1+n_2-6)^{-1}\,\beta_0(n-3,l)\,\beta(\tfrac{1}{2}n-2,l)$$
$$\times\,[(n_1+n_2-4)\,\beta_2(n-3,l) + (n_2-n_1)\,\beta_3(n-4,l)$$
$$+\,(n_1+n_2-6)(n_1+n_2-8)\,\beta_4(n-4,l)]\}, \tag{A.7}$$

$$F^{(2)}_{5;n_1,n_2}(l) = B(n-1)^{-1}\,B(n-3)^{-1}\{\beta(n-2,l)\,\beta(n-4,l) - (n_1+n_2-6)^{-1}$$
$$\times\beta_0(n-3,l)[(n_1+n_2-4)\,\beta_2(n-3,l) + (n_2-n_1)\,\beta_3(n-4,l)$$
$$+\,(n_1+n_2-6)(n_1+n_2-8)\,\beta_4(n-4,l)]\}, \tag{A.8}$$

and $F^{(3)}_{5;n_1,n_2}(l)$ has also been obtained.

APPENDIX B: Marginal Distributions of the Largest Root of $\mathbf{W}_l$

$$G_{2;q}(l) = \Gamma(q-1)^{-1}[\gamma(q-2,l) - \tfrac{1}{2}\gamma_0(\tfrac{1}{2}(q-1),l)\,\gamma(\tfrac{1}{2}(q-3),l)], \tag{B.1}$$

$$G_{3,q}(l) = [2^{\frac{1}{2}q}\Gamma(\tfrac{1}{2}q)\,\Gamma(q-1)]^{-1}[\gamma(\tfrac{1}{2}q-1,l)\,\gamma(q-2,l)$$
$$-\,2\gamma_0(\tfrac{1}{2}q-1,l)\,\gamma_2(q-2,l)], \tag{B.2}$$

$$G_{4;q}(l) = \Gamma(q-1)^{-1}\Gamma(q-2)^{-1}[\gamma(q-2,l)\,\gamma(q-3,l) - \gamma_0(q-2,l)\,\gamma_2(q-3,l)$$
$$-\,\tfrac{1}{2}\gamma_0(\tfrac{1}{2}(q-3),l)\,\gamma(\tfrac{1}{2}(q-3),l)\,\gamma_3(q-3,l)], \tag{B.3}$$

$$G_{5;q}(l) = [2^{\frac{1}{2}q-1}\Gamma(\tfrac{1}{2}q-1)\Gamma(q-1)\Gamma(q-3)]^{-1}\{\gamma(\tfrac{1}{2}q-2,l)\gamma(q-2,l)\gamma(q-4,l)$$
$$+\,2[\gamma_0(\tfrac{1}{2}q-1,l)\gamma(q-4,l)\gamma_2(q-3,l) - \gamma_0(\tfrac{3}{2}q-4,l)\gamma_3(q-4,l)$$
$$-\,2\gamma_0(\tfrac{1}{2}q-2,l)\,\gamma(q-3,l)\,\gamma_4(q-4,l)] - \gamma_0(q-3,l)\,\gamma(\tfrac{1}{2}q-2,l)$$
$$\times\,[\gamma_2(q-3,l) + \gamma_3(q-4,l) + \gamma_4(q-4,l)]\}. \tag{B.4}$$

APPENDIX C: Laplace Transforms of the Density of V

$m = 3$:

$$\Lambda_{3;4,n_2}(s) = 2^{-n_2-2}(n_2-1)_4\,[2\lambda(0,\tfrac{1}{2}n_2\,;s)\,\lambda(0,n_2-2;s) - \lambda(1,n_2-1;s)], \tag{C.1}$$

$$\Lambda_{3;6,n_2}(s) = (4!2^{n_2+3})^{-1}(n_2-1)_6\{\lambda(0, \tfrac{1}{2}n_2+1; s)[(n_2-2)\lambda(3, n_2-3; s) - 12\lambda(1, n_2-2; s)] + \lambda(2, n_2; s) + 2\lambda(2, n_2-2; s)\}, \tag{C.2}$$

$$\begin{aligned}\Lambda_{3;8,n_2}(s) = (6!2^{n_2+7})^{-1}(n_2-1)_8\{&\lambda(0, \tfrac{1}{2}n_2+2; s)[(n_2-2)(n_2-3)\\ &\times\lambda(6, n_2-4; s) + 3(n_2-2)(\lambda(4, n_2-2; s) - 12\lambda(4, n_2-3; s))\\ &+ 240\lambda(2, n_2-2; s)] + 2[(n_2-2)\lambda(5, n_2-3; s)\\ &- 2\lambda(3, n_2+1; s) - 6\lambda(3, n_2-1; s) - 12\lambda(3, n_2-2; s)]\}.\end{aligned} \tag{C.3}$$

$m = 4$:

$$\begin{aligned}\Lambda_{4;5,n_2}(s) = (4!2^{n_2+1})^{-1}(n_2-6)_6[&2^{-n_2+3}\lambda(0, n_2+1; s)\lambda(0, n_2-3; s)\\ &- 2n_2\lambda(0, \tfrac{1}{2}(n_2+1); s)\lambda(2, n_2-3; s) - \lambda(1, n_2; s)\\ &- 4\lambda(1, n_2-2; s)],\end{aligned} \tag{C.4}$$

$$\begin{aligned}\Lambda_{4;7,n_2}(s) = (6!2^{n_2+4})^{-1}(n_2-2)_8\{&2^{-n_2}\lambda(0, n_2+3; s)[(n_2-3)(n_2-4)\\ &\times\lambda(4, n_2-5; s) - 4(n_2-3)(\lambda(3, n_2-4; s) + 6\lambda(2, n_2-4; s))\\ &+ 48(\lambda(1, n_2-3; s) + \lambda(0, n_2-3; s))]\\ &+ 2n_2(n_2+2)\lambda(0, \tfrac{1}{2}(n_2+3); s)\lambda(4, n_2-3; s)\\ &+ 4(n_2-3)[\lambda(3, n_2-2; s) + 4\lambda(3, n_2-4; s)]\\ &- 3[\lambda(1, n_2+2; s) + 16\lambda(3, n_2-3; s) + 48\lambda(1, n_2-2; s)]\}.\end{aligned} \tag{C.5}$$

$\Lambda_{3;10,n_2}$ and $\Lambda_{4;9,n_2}$ have been obtained in similar form.

References

[1] Davis, A. W. (1968). A system of linear differential equations for the distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **39** 815–832.

[2] Davis, A. W. (1970). Exact distributions of Hotelling's generalized T_0^2. *Biometrika* **57** 187–191.

[3] Davis, A. W. (1970). Further applications of a differential equation for Hotelling's generalized T_0^2. *Ann. Inst. Statist. Math.* **22** 77–87.

[4] Davis, A. W. (1970). On the null distribution of the sum of the roots of a multivariate beta distribution. *Ann. Math. Statist.* **41** 1557–1562.

[5] Davis, A. W. (1972). On the marginal distributions of the latent roots of the multivariate beta matrix. *Ann. Math. Statist.*, to appear.

[6] Hotelling, H. (1951). A generalized T-test and measure of multivariate dispersion. In *Proc. Second Berkeley Symp. Math. Statist. Prob.*, pp. 23–42. University of California Press, Berkeley, California.
[7] Krishnaiah, P. R. and Chang, T. C. (1971). On the exact distribution of the extreme roots of the Wishart and MANOVA matrices. *J. Multivariate Anal.* **1** 108–117.
[8] Krishnaiah, P. R. and Waikar, V. B. (1971). Exact joint distributions of any few ordered roots of a class of random matrices. *J. Multivariate Anal.* **1** 308–315.
[9] Krishnaiah, P. R. and Chang, T. C. (1972). On the exact distributions of the traces of $S_1(S_1 + S_2)^{-1}$ and $S_1S_2^{-1}$. *Sankhya* **34**.
[10] Lawley, D. N. (1938). A generalization of Fisher's Z-test. *Biometrika* **30** 180–187.
[11] Mehta, M. L. (1967). *Random Matrices and the Statistical Theory of Energy Levels*. Academic Press, New York.
[12] Nanda, D. N. (1947). Distribution of a root of a determinantal equation. *Ann. Math. Statist.* **19** 47–57.
[13] Nanda, D. N. (1948). Limiting distribution of a root of a determinantal equation. *Ann. Math. Statist.* **19** 340–350.
[14] Nanda, D. N. (1950). Distribution of the sum of the roots of a determinantal equation under a certain condition. *Ann. Math. Statist.* **21** 432–439.
[15] Pillai, K. C. S. (1954). On some distribution problems in multivariate analysis. Mimeo Series No. 88, Institute of Statistics, University of North Carolina, Chapel Hill.
[16] Pillai, K. C. S. and Jayachandran, K. (1970). On the exact distribution of Pillai's $V^{(s)}$ criterion. *J. Amer. Statist. Assoc.* **65** 447–454.
[17] Pillai, K. C. S. and Young, D. L. (1971). On the exact distribution of Hotelling's generalized T_0^2. *J. Multivariate Anal.* **1** 90–107.
[18] Roy, S. N. (1945). The individual sampling distribution of the maximum, the minimum and any intermediate of the p-statistics on the null hypothesis. *Sankhya* **7** 133–158.

Printed by the St Catherine Press Ltd., Tempelhof 37, Bruges, Belgium.

Reprinted from JOURNAL OF MULTIVARIATE ANALYSIS Vol. 2, No. 4, December 19
All Rights Reserved by Academic Press, New York and London Printed in Belgi

On the Ratios of the Individual Latent Roots to the Trace of a Wishart Matrix

A. W. DAVIS

C.S.I.R.O., Adelaide, Australia

A simple relationship is given between the exact null distribution $g_{m,n}^{(J)}$ of the J-th largest latent root of an $m \times m$ Wishart matrix on n degrees of freedom, and the distribution $f_{m,n}^{(J)}$ of the ratio of this root to the trace of the matrix. Explicit expressions for certain $f_{m,n}^{(J)}$ may thus be obtained from previous results for the corresponding $g_{m,n}^{(J)}$.

1. INTRODUCTION

In a recent paper, Krishnaiah and Waikar [5] have discussed tests of equality of the latent roots of certain matrices against various classes of alternatives. These tests are based on ratios of the latent roots of certain random matrices; in particular, on the ratios of the individual roots to the trace. In this note, we present a simple relationship between the marginal distribution of a latent root of a central Wishart matrix $\mathbf{S}$ and the distribution of the ratio of this root to the trace. Unfortunately, no corresponding results have yet been found for the matrices $\mathbf{S}_1\mathbf{S}_2^{-1}$ and $\mathbf{S}_1(\mathbf{S}_1 + \mathbf{S}_2)^{-1}$.

2. RATIOS OF THE INDIVIDUAL ROOTS TO THE TRACE

Let $\mathbf{S}$ be an $m \times m$ central Wishart matrix on n degrees of freedom, having latent roots $0 < l_m < \cdots < l_1 < \infty$, and let $u_i = l_i/\sum l_j$ $(i = 1,..., m; \sum_{i=1}^m u_i = 1)$. If $f_{m,n}^{(J)}, g_{m,n}^{(J)}$ denote the marginal densities of u_J, l_J respectively, and

$$\mathscr{L}(h(w)) = \int_0^\infty e^{-sw} h(w)\, dw$$

Received March 22, 1972; revised August 21, 1972.

AMS 1970 subject classification: Primary 62H10; Secondary 62E15.

Key words and phrases: Exact distributions; Wishart matrix, ratio of root to trace; Laplace transforms.

denotes the Laplace transform, then we prove in the next section that

$$\mathscr{L}\left((1+w)^{\frac{1}{2}mn-2} f^{(J)}_{m,n}\left(\frac{1}{1+w}\right)\right) = 2\Gamma(\tfrac{1}{2}mn)\, e^{s} s^{-\frac{1}{2}mn+1} g^{(J)}_{m,n}(2s). \quad (2.1)$$

In principle, this result allows us to deduce the $f^{(J)}_{m,n}$ from expressions for the $g^{(J)}_{m,n}$ [6, 3]. In particular, if $p_{m;n_1,n_2}$ is the null density of trace $\mathbf{S}_1(\mathbf{S}_1 + \mathbf{S}_2)^{-1}$, where $\mathbf{S}_1$, $\mathbf{S}_2$ are independent $m \times m$ Wishart matrices on n_1, n_2 degrees of freedom, respectively, then we obtain for the largest ratio u_1

$$f^{(1)}_{m,n}(u) = k_1(m,n)\, u^{\frac{1}{2}mn-2} p_{m-1;n-1,m+2}\left(\frac{1-u}{u}\right), \quad (2.2)$$

where

$$k_1(m,n) = [\pi^{1/2}\Gamma(\tfrac{1}{2}mn)/\Gamma(\tfrac{1}{2}m)\Gamma(\tfrac{1}{2}n)] \prod_{i=0}^{m-2} [\Gamma(\tfrac{1}{2}(m+2-i))/\Gamma(\tfrac{1}{2}(n+m+1-i))]. \quad (2.3)$$

It follows from known results for p [2] that u_1 has range $(m^{-1}, 1)$, and that $f^{(1)}_{m,n}(u)$ is piecewise analytic in the intervals between the points $u = m^{-1}$, $(m-1)^{-1}, \ldots, 2^{-1}, 1$. As $n \to \infty$, $n^{1/2}(mu_1 - 1)$ tends to a limiting distribution, so that for very large n the distribution of u_1 is mainly concentrated in $(m^{-1}, (m-1)^{-1})$.

When $m = 2$,

$$\text{Prob}(u_1 < u) = 1 - 2^{n-1}(u - u^2)^{\frac{1}{2}(n-1)}, \qquad (\tfrac{1}{2} < u < 1), \quad (2.4)$$

and $\frac{1}{2}n(2u_1 - 1)^2$ is asymptotically negative exponential with mean 1.

When $m = 3$, we obtain from [2, Eq. (3.4)] and (2.2)

$$f^{(1)}_{3,n}(u) = k(n)\{2^n u^{n-2}(1-2u)^{\frac{1}{2}n} + u^{\frac{1}{2}n-2}[\tfrac{1}{2}n(1-u)^{n-2}(3u-1)^2 - (1-u)^n]\}, \quad (\tfrac{1}{3} < u < \tfrac{1}{2}),$$

$$f^{(1)}_{3,n}(u) = k(n)\, u^{\frac{1}{2}n-2}[\tfrac{1}{2}n(1-u)^{n-2}(3u-1)^2 - (1-u)^n], \qquad (\tfrac{1}{2} < u < 1), \quad (2.5)$$

where $k(n) = 1/2\,B(\frac{1}{2}n + 1, n - 1)$. The limiting density of $w = n^{1/2}(3u_1 - 1)$ in this case is

$$h(w) = \left(\frac{3}{2\pi}\right)^{1/2} [e^{-3w^2/2} + (\tfrac{9}{8}w^2 - 1)\, e^{-3w^2/8}], \qquad (0 < w < \infty), \quad (2.6)$$

with mean $3(3/2\pi)^{1/2} = 2.0730$, variance $(29\pi - 81)/6\pi = 0.5361$, and Pearson parameters $\beta_1^{1/2} = 0.4983$, $\beta_2 = 3.2071$. Approximate upper 5% and 1% points are 3.381 and 4.021, respectively.

Since the median root l_2 in the case $m = 3$ is a gamma variate with parameter $n - 1$, it follows from (2.1) that $2u_2$ is a beta variate with parameters $n - 1$, $\frac{1}{2}n + 1$.

For $m > 3$, the densities become increasingly complicated, with convolution-type terms. No further approximations have been found for the limiting distributions.

3. Proof of (2.1)

The joint nonnull density of $u_1, \ldots, u_{m-1}$ has been given in [4] and [5], and reduces in the null case (see [1]) to

$$\theta(u_1, \ldots, u_{m-1}) = k_2(m, n) \prod_{i=1}^{m} u_i^{\frac{1}{2}(n-m-1)} \prod_{i<j} (u_i - u_j),$$

where

$$k_2(m, n) = \pi^{\frac{1}{2}m} \Gamma(\tfrac{1}{2}mn) \Big/ \Big[\prod_{i=0}^{m-1} \Gamma(\tfrac{1}{2}(m-i))\, \Gamma(\tfrac{1}{2}(n-i)) \Big].$$

The u_i lie in the $(m - 1)$ dimensional region $\mathscr{E}_{m-1} = \{0 < u_m < \cdots < u_1 < 1, \sum_{i=1}^{m} u_i = 1\}$. Writing the left-hand side of (2.1) in the form

$$\int_{\mathscr{E}_{m-1}} \exp[-s(1 - u_J)/u_J]\, u_J^{-\frac{1}{2}mn} \theta(u_1, \ldots, u_{m-1}) \prod_{i=1}^{m-1} du_i, \tag{3.1}$$

we make the transformation

$$v_i = 2su_i/u_J \qquad (i = 1, \ldots, m;\ s > 0) \tag{3.2}$$

to new variables $v_1, \ldots, v_{J-1}, v_{J+1}, \ldots, v_m$. This maps $\mathscr{E}_{m-1}$ onto the region

$$\mathscr{F}_{m-1}(J, s) = \{0 < v_m < \cdots < v_{J+1} < v_J \equiv 2s < v_{J-1} < \cdots < v_1 < \infty\}. \tag{3.3}$$

Since

$$s(1 - u_J)/u_J = \tfrac{1}{2} \sum_{\substack{i=1 \\ i \neq J}}^{m} v_i, \tag{3.4}$$

and the transformation (3.2) has Jacobian

$$\partial(v_1, \ldots, v_{J-1}, v_{J+1}, \ldots, v_m)/\partial(u_1, \ldots, u_{m-1}) = (2s)^{m-1}/u_J^{\,m}, \tag{3.5}$$

the integral (3.1) becomes

$$2\Gamma(\tfrac{1}{2}mn)\, e^{s} s^{-\frac{1}{2}mn+1} \int_{\mathscr{F}_{m-1}(J,s)} \psi(v_1, \ldots, v_m) \prod_{\substack{i=1 \\ i \neq J}}^{m} dv_i\,, \tag{3.6}$$

where $\psi(v_1, \ldots, v_m)$ is the joint null density of $l_1, \ldots, l_m$. This is equivalent to the right-hand side of (2.1).

References

[1] Bartlett, M. S. (1951). The effect of standardization on a χ^2 approximation in factor analysis. *Biometrika* **38** 337–344.

[2] Davis, A. W. (1970). On the null distribution of the sum of the roots of a multivariate beta distribution. *Ann. Math. Statist.* **41** 1557–1562.

[3] Davis, A. W. (1972). On the distributions of the latent roots and traces of certain random matrices. *J. Multivariate Anal.* **2** 189–200.

[4] James, A. T. (1966). Inference on latent roots by calculation of hypergeometric functions of matrix argument. In *Multivariate Analysis* (P. R. Krishnaiah, Ed.), pp. 209–235. Academic Press, New York.

[5] Krishnaiah, P. R. and Waikar, V. B. (1971). Simultaneous tests for equality of latent roots against certain alternatives—I. *Ann. Inst. Statist. Math.* **23** 451–468.

[6] Krishnaiah, P. R. and Waikar, V. B. (1972). Exact joint distributions of any few ordered roots of a class of random matrices. *J. Multivariate Anal.* **1** 308–315.

Printed by the St Catherine Press Ltd., Tempelhof 37, Bruges, Belgium.

To appear in Biometrika

14

On the differential equation for Meijer's $G^{p,o}_{p,p}$ function,

and further tabulation of Wilks's likelihood ratio criterion

A.W. Davis

CSIRO Division of Mathematics and Statistics, Glen Osmond, South Australia.

Summary

The null distributions of a number of likelihood ratio criteria are well known to be expressible in terms of Meijer's $G^{p,o}_{p,p}$ function. It is shown that such distributions correspond to a single fundamental solution of the differential equation for this function at the regular singularity 1. Computation of percentage points of Wilks's likelihood ratio criterion by "analytic continuation" of this solution fills some gaps in the previously published tabulations.

Some key words: Meijer's G function, linear differential equation, Wilks's likelihood ratio criterion, percentage points, analytic continuation.

Introduction

It was pointed out by Nair (1938) that the exact distributions of many likelihood ratio criteria L ($0 \le L \le 1$) may be derived by applying Mellin's inversion theorem to the s th moments of the criteria. For the null distributions, the latter are frequently of the form

$$E(L^s) = K\cdot\Pi_{i=1}^{p}\{\Gamma(b_i + s)/\Gamma(a_i + s)\}, \quad (a_i, b_i > 0, i = 1, \ldots, p), \tag{1}$$

where

$$K = \Pi_{i=1}^{p}\{\Gamma(a_i)\ \Gamma(b_i)\}.$$

Applying the inversion theorem, it follows that the density of L is given by

$$f(L) = K(2\pi i)^{-1} \int_{c-i\infty}^{c+i\infty} L^{-s-1}\ \Pi_{i=1}^{p}\ \{\Gamma(b_i + s)/\Gamma(a_i + s)\}\ ds \quad (c \ge 0)$$

$$= KL^{-1}G_{p,p}^{p,o}\left(L\middle|\begin{matrix}a_1, \ldots\ldots, a_p\\ b_1, \ldots\ldots, b_p\end{matrix}\right),$$

a special case of Meijer's $G_{p,q}^{m,n}$ function (Erdelyi (1953) Section 5.3). This result provides powerful methods for deriving explicit forms of these distributions in a number of cases (see references in Consul (1969), Mathai (1973), for example). However, computational problems may arise in tabulating the distributions when the latter are obtained as infinite series. One possibility for overcoming the problem of a slowly convergent series is to expand the series at intermediate points, and so approach the required percentage points by a process of "analytic continuation". If the distribution can be shown to satisfy a differential equation, the latter may provide a convenient tool for this process.

As is well known (Erdelyi (1953), Section 5.4), Meijer's G-function satisfies a certain homogeneous linear differential equation which has a regular singularity at 0, an irregular singularity at ∞, and in the case of $G^{p,o}_{p,p}$ functions, a regular singularity at 1. Differential equations of this type were obtained by Nair (1938), who approached their solution by way of the regular singularity at 0. Unfortunately, the required solution is then a linear combination of the complete fundamental system of solutions at 0, with constants which Nair evaluated using the calculus of residues. Furthermore, the characteristic equation for the solution at 0 has roots which will generally differ by integers in statistical applications, so that Frobenius's method is required to construct the fundamental solutions. The shortcomings of this approach in connection with multivariate distributions were justifiably pointed out by Mathai (1973).

However, the situation is different at the regular singularity 1. In Section 2, the differential equation for $G^{p,o}_{p,p}$ will be formulated in a way which clearly displays this singularity. The uniqueness of the required solution, and the evaluation of the multiplicative constant, will be shown to follow quite simply from the well known representation of L satisfying (1) as a product of independent beta variates.

In Section 3, the approach is applied to Wilks's (1932) likelihood ratio criterion

$$\Lambda = |W| \; / \; |B + W|, \tag{2}$$

where B and W are the p × p hypothesis and error sum-of-products matrices, on q and n degrees of freedom respectively, used in the hypothesis testing for the general linear model. The complete explicit form of the distribution has been given by Mathai (1971). Tabulations have been made by Schatzoff (1966), Pillai and Gupta (1969), Mathai (1971), and Lee (1972). Infinite series are encountered when p and q are both odd and such values are omitted from the above tabulations except in the work of Lee, who reduced the problem to the evaluation of certain univariate integrals and tabulated the cases $p, q \leq 9$. Also, it was generally found necessary to replace the explicit forms by asymptotic expansions for larger n. In the present paper, the computational utility of the differential equation approach is indicated, at least in the case of Λ, by filling most of the gaps in the published tabulations for $p \leq 10$, $q \leq 22$, $pq \leq 140$ (Table 1). No modifications were required for particular p, q or large n, although for larger p and q the method tended to break down for very small n. The usefulness of the technique for other likelihood ratio criteria remains to be investigated.

The differential equation approach has been used by the present author in connection with the null distributions of other multivariate test criteria whose moments are not of form (1) (Davis 1968, 1970a, 1970b, 1972), and also for the linear combination of chi-squares (Davis 1977).

2. The differential equation for $G^{p,o}_{p,p}$

For r = 0, 1, ..., p define

$$H_r(L) = (2\pi i)^{-1} \int_{c-i\infty}^{c+i\infty} L^{-s} \, \Pi_{i=1}^{p}\{\Gamma(b_i + s)/\Gamma(a_i + s)\} \, ds/\Pi_{j=1}^{r}(a_j + s),$$

noting that $H_0(L) = G^{p,o}_{p,p}\left(L\middle|{}^{a_i}_{b_i}\right)$. It follows easily that

$$(Ld/dL - a_r)H_r = -H_{r-1} \; (r = 1, \ldots, p) \tag{3}$$

$$(1 - L)H_0 = \textstyle\sum_{r=1}^{p} \theta_r^{(p)} H_r$$

where the $\theta_r^{(p)}$ are constants such that

$$\Pi_{i=1}^{p}\{(b_i + s)/(a_i + s)\} = 1 - \sum_{r=1}^{p} \theta_r^{(p)}/\Pi_{j=1}^{r} (a_j + s). \qquad (4)$$

If we now define the vector of functions $\bar{H} = (H_1, \ldots, H_p)'$, then from (3)

$$d\bar{H}/dL = \{L^{-1}(A - T) + (L - 1)^{-1}T\}\bar{H} \qquad (5)$$

where

$$A = \begin{pmatrix} a_1 & 0 & \ldots & 0 \\ -1 & a_2 & & \cdot \\ 0 & & & \cdot \\ 0 & & -1 & a_p \end{pmatrix}, \quad T = \begin{pmatrix} \theta_1^{(p)}, & \ldots, & \theta_p^{(p)} \\ 0, & \ldots, & 0 \\ \cdot & & \\ \cdot & & \\ 0, & \ldots, & 0 \end{pmatrix}.$$

Equation (5) displays $\bar{H}$ as the solution of a homogeneous linear differential equation of order p, with regular singularities a + 0 and 1. The theory of such differential equations is presented, for example, in Hille (1969), Chapter 5, Appendix B. The fundamental systems of solutions at L = a correspond to the latent roots of the matrix coefficient of $(L - a)^{-1}$. Expanding by the first row, the latent roots of A + T satisfy the polynomial equation

$$\Pi_{i=1}^{p}(a_i - \lambda) - \sum_{r=1}^{p}\theta_r^{(p)} \Pi_{j=r+1}^{p}(a_j - \lambda) = 0,$$

so that in virtue of (4) the roots at L = 0 are simply $b_1, \ldots, b_p$. As stated in the Introduction these may differ by integers in statistical applications. On the other hand, T has a zero root of multiplicity p - 1 and a single nonzero root $\theta_1^{(p)}$.

The solution corresponding to the latter has the form

$$\bar{H} = C\sum_{r=0}^{\infty} h_r(1 - L)^{r+\theta_1^{(p)}} \qquad (6)$$

where h_0, the right eigenvector of T corresponding to the root $\theta_1^{(p)}$, is the p-vector $(1, 0, \ldots, 0)'$. It remains to show that (6) is the required solution, and also to evaluate C and the $\theta_r^{(p)}$; the h_r may then be obtained recursively from (5).

Considering partial fractions in (4), the $\theta_r^{(p)}$ are seen to satisfy the linear equations

$$\sum_{r=1}^{p} \theta_r^{(p)} \prod_{u=r+1}^{p} (a_u - a_i) = -\prod_{j=1}^{p} (b_j - a_i), \quad (i = 1, \ldots, p),$$

which may be shown to have the unique solution

$$\theta_r^{(p)} = -\sum_{j=r}^{p} \{\prod_{u=1}^{p} (b_u - a_j)/\prod_{v=r, v\neq j}^{p} (a_v - a_j)\}, \quad (r = 1, .., p).$$

Using this result and (4), the θ's may also be shown to satisfy the recursive relations

$$\theta_r^{(p+1)} = \theta_r^{(p)} + (b_{p+1} - a_{p+1})\sum_{j=0}^{r-1} \theta_j^{(p)} \prod_{u=j+1}^{r-1} (a_u - a_{p+1}), \qquad (7)$$

$(\theta_0^{(p)} \equiv -1)$. In particular, we obtain

$$\theta_1^{(p)} = \sum_{i=1}^{p} (a_i - b_i).$$

Now, a statistic L with moments (1) can clearly be represented as the product of independent beta variates with parameters b_i, $a_i - b_i (i = 1, \ldots, p)$, or equivalently $-\log L = \sum_{i=1}^{p} w_i$ where the w_i are independent variates with density functions

$$\{\Gamma(a_i)/\Gamma(b_i)\Gamma(a_i - b_i)\}e^{-b_i w}(1 - e^{-w})^{a_i - b_i - 1} \quad (i = 1, .., p). \qquad (8)$$

Hence as $L \to 1$, the behaviour of $f(L)$ may be derived from that of the convolution of the densities (8) as $w \to 0$; it is simple to show that

$$f(L) \sim \{K/\Gamma(\theta_1^{(p)})\}(1 - L)^{\theta_1^{(p)} - 1} \quad \text{as } L \to 1.$$

I.e. $f(L) = L^{-1}H_0(L)$, where $H_0(L)$ is given by (3) and (6) with multiplicative constant

$$C = K/\Gamma(\theta_1^{(p)} + 1).$$

Writing $H_r = L^{-1}H_r (r = 0, 1, \ldots, p)$, we obtain the system (5) with a_r replaced by $a_r - 1$.

3. Wilks's likelihood ratio criterion

For the statistic Λ given by (2), we have

$$a_i = \tfrac{1}{2}(q + n - i + 1),\; b_i = \tfrac{1}{2}(n - i + 1), \quad (i = 1, \ldots, p).$$

From (7), by induction,

$$\theta_r^{(p)} = (-1)^{r-1}2^{-r} \binom{p}{r} q(q - 1) \ldots (q - r + 1), \; (r = 1, \ldots, p),$$

so that $\theta_1^{(p)} = \tfrac{1}{2}pq$. In solving (5), it is convenient to write $H_r^* = \theta_r^{(p)} H_r$; the equations then become

$$(Ld/dL - a_r)H_r^* = \phi_r H_{r-1}^*, \qquad (r = 1, \ldots, p),$$

$$(1 - L)H_0^* = \textstyle\sum_{r=1}^{p} H_r^*,$$

where

$$\phi_r = -\tfrac{1}{2}r^{-1}(p - r + 1)(q - r + 1), \qquad (r = 1, \ldots, p).$$

These equations may also be derived by a method similar to that of Davis (1968).

The numerical solution was carried out essentially on the lines described in Davis (1977), except that the differential equation was not transformed to check the accuracy. Agreement of the $(p, q \pm 1, n)$ percentiles with the published values was considered to be a sufficient confirmation of the (p, q, n) results, since regardless of whether p and q are odd or even, we are attempting to sum infinite series. The results are presented in Table 1 in a similar format to that used in previous tabulations, i.e. we give correction ratios

$$C = - \{n - \tfrac{1}{2}(p - q + 1)\} \log_e \Lambda(\alpha) \,/\, \chi^2_{pq}(\alpha),$$

where $\Lambda(\alpha)$ and $\chi^2_{pq}(\alpha)$ denote the upper and lower α-points of Λ and chi-square on pq degrees of freedom respectively, ($\alpha = 0.1, 0.05, 0.025, 0.01, 0.005$).

References

Consul, P.C. (1969). The exact distributions of likelihood criteria for different hypotheses. In Multivariate Analysis - II (P.R. Krishnaiah, Ed.) New York: Academic Press.

Davis, A.W. (1968). A system of differential equations for the distribution of Hotelling's generalized T_0^2. Ann. Math. Statist., 39, 815-32.

Davis, A.W. (1970a). Exact distributions of Hotelling's generalized T_0^2. Biometrika 57, 187-91.

Davis, A.W. (1970b). On the null distribution of the sum of the roots of a multivariate beta distribution. Ann. Math. Statist. 41, 1557-62.

Davis, A.W. (1972). On the marginal distributions of the latent roots of the multivariate beta distribution. Ann. Math. Statist. 43, 1664-70.

Davis, A.W. (1977). A differential equation approach to linear combinations of independent chi-squares. J. Am. Statist. Ass. 72, 212-4.

Erdélyi, A. (Ed.) (1953). Higher Transcendental Functions, vol. 1. New York: McGraw-Hill.

Hille, E. (1969). Lectures on Ordinary Differential Equations. Reading, Mass.: Addison-Wesley Publishing Co.

Lee, Y. - S. (1972). Some results on the distribution of Wilks's likelihood ratio criterion. Biometrika 59, 649-64.

Mathai, A.M. (1971). On the distribution of the likelihood-ratio criterion for testing linear hypotheses on regression coefficients. Ann. Inst. Statist. Math., 23, 181-97.

Mathai, A.M. (1973). A review of the various methods of obtaining the exact distributions of multivariate test criteria. Sankhya, Ser. A., 35, 39-60.

Nair, U.S. (1938). The application of the moment function in the study of distribution laws in statistics. Biometrika 30, 274-94.

Pillai, K.C.S. and Gupta, A.K. (1969). On the exact distribution of Wilks's criterion. Biometrika 56, 109-18.

Schatzoff, M. (1966). Exact distributions of Wilks's likelihood ratio criterion. Biometrika 53, 347-58.

Wilks, S.S. (1932). Certain generalizations in the analysis of variance. Biometrika 24, 471-94.

Table 1. Chi-squared adjustments to Wilks's criterion Λ. Factor C for lower percentiles of Λ(upper percentiles of χ^2).

	p = 3, q = 11					p = 3, q = 13				
M \ α	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
1	1.685	1.754	1.821	1.907	1.969	1.750	1.824	1.896	1.988	2.055
2	1.358	1.385	1.410	1.442	1.466	1.405	1.434	1.462	1.497	1.522
3	1.237	1.252	1.266	1.284	1.297	1.274	1.291	1.306	1.326	1.340
4	1.173	1.183	1.192	1.204	1.213	1.203	1.214	1.225	1.238	1.247
5	1.133	1.140	1.147	1.156	1.162	1.158	1.167	1.174	1.184	1.191
6	1.106	1.112	1.117	1.124	1.128	1.128	1.134	1.140	1.148	1.153
7	1.087	1.092	1.096	1.101	1.105	1.106	1.111	1.116	1.122	1.126
8	1.073	1.077	1.080	1.084	1.087	1.089	1.094	1.098	1.102	1.106
9	1.062	1.065	1.068	1.072	1.074	1.076	1.080	1.083	1.088	1.090
10	1.054	1.056	1.059	1.062	1.064	1.066	1.069	1.072	1.076	1.078
12	1.041	1.043	1.045	1.047	1.049	1.052	1.054	1.056	1.059	1.061
14	1.033	1.034	1.036	1.037	1.039	1.041	1.043	1.045	1.047	1.048
16	1.027	1.028	1.029	1.030	1.031	1.034	1.035	1.037	1.038	1.040
18	1.022	1.023	1.024	1.025	1.026	1.028	1.029	1.031	1.032	1.033
20	1.019	1.020	1.020	1.021	1.022	1.024	1.025	1.026	1.027	1.028
24	1.014	1.014	1.015	1.016	1.016	1.018	1.019	1.019	1.020	1.021
30	1.009	1.010	1.010	1.011	1.011	1.012	1.013	1.013	1.014	1.014
40	1.006	1.006	1.006	1.007	1.007	1.008	1.008	1.008	1.009	1.009
60	1.003	1.003	1.003	1.003	1.003	1.004	1.004	1.004	1.004	1.004
120	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	43.745	47.400	50.725	54.776	57.648	50.660	54.572	58.120	62.428	65.476

Table 1 (cont.)

p = 3, q = 15

M \ α	0.100	0.050	0.025	0.010	0.005
1	1.808	1.887	1.964	2.061	2.133
2	1.449	1.480	1.510	1.547	1.575
3	1.309	1.327	1.344	1.365	1.381
4	1.232	1.244	1.256	1.270	1.280
5	1.183	1.192	1.200	1.211	1.218
6	1.149	1.156	1.163	1.171	1.177
7	1.124	1.130	1.135	1.142	1.147
8	1.105	1.110	1.115	1.120	1.124
9	1.091	1.095	1.099	1.103	1.107
10	1.079	1.083	1.086	1.090	1.093
12	1.062	1.065	1.067	1.070	1.072
14	1.050	1.052	1.054	1.056	1.058
16	1.041	1.043	1.045	1.047	1.048
18	1.035	1.036	1.037	1.039	1.040
20	1.030	1.031	1.032	1.033	1.034
24	1.022	1.023	1.024	1.025	1.026
30	1.016	1.016	1.017	1.017	1.018
40	1.010	1.010	1.010	1.011	1.011
60	1.005	1.005	1.005	1.005	1.005
120	1.001	1.001	1.001	1.001	1.002
∞	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	57.505	61.656	65.410	69.957	73.166

p = 3, q = 17

M \ α	0.100	0.050	0.025	0.010	0.005
1	1.861	1.944	2.025	2.127	2.203
2	1.489	1.522	1.554	1.594	1.623
3	1.341	1.361	1.379	1.402	1.419
4	1.259	1.273	1.285	1.300	1.312
5	1.206	1.216	1.225	1.237	1.245
6	1.169	1.177	1.184	1.193	1.200
7	1.142	1.149	1.154	1.162	1.167
8	1.122	1.127	1.132	1.138	1.142
9	1.105	1.110	1.114	1.119	1.123
10	1.092	1.096	1.100	1.104	1.107
12	1.073	1.076	1.079	1.082	1.084
14	1.059	1.061	1.064	1.066	1.068
16	1.049	1.051	1.053	1.055	1.056
18	1.041	1.043	1.044	1.046	1.047
20	1.035	1.037	1.038	1.040	1.041
24	1.027	1.028	1.029	1.030	1.031
30	1.019	1.020	1.020	1.021	1.022
40	1.012	1.012	1.013	1.013	1.013
60	1.006	1.006	1.006	1.006	1.007
120	1.002	1.002	1.002	1.002	1.002
∞	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	64.295	68.669	72.616	77.386	80.747

Table 1 (cont.)

	p = 3, q = 19					p = 3, q = 21				
M \ α	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
1	1.909	1.996	2.080	2.188	2.267	1.954	2.044	2.131	2.243	2.325
2	1.526	1.561	1.595	1.637	1.668	1.561	1.598	1.633	1.677	1.709
3	1.372	1.393	1.412	1.437	1.454	1.401	1.423	1.444	1.470	1.488
4	1.285	1.300	1.313	1.330	1.341	1.310	1.325	1.340	1.357	1.370
5	1.229	1.240	1.250	1.262	1.271	1.250	1.262	1.273	1.286	1.295
6	1.189	1.198	1.205	1.215	1.222	1.208	1.217	1.226	1.236	1.243
7	1.160	1.167	1.173	1.181	1.186	1.177	1.184	1.191	1.200	1.205
8	1.137	1.143	1.148	1.155	1.159	1.153	1.159	1.165	1.172	1.176
9	1.119	1.124	1.129	1.134	1.138	1.133	1.139	1.144	1.150	1.154
10	1.105	1.109	1.113	1.118	1.121	1.118	1.122	1.127	1.132	1.135
12	1.084	1.087	1.090	1.093	1.096	1.094	1.098	1.101	1.105	1.108
14	1.068	1.071	1.073	1.076	1.078	1.077	1.080	1.083	1.086	1.088
16	1.057	1.059	1.061	1.063	1.065	1.065	1.067	1.059	1.072	1.074
18	1.048	1.050	1.052	1.054	1.055	1.055	1.057	1.059	1.061	1.063
20	1.041	1.043	1.044	1.046	1.047	1.048	1.049	1.051	1.053	1.054
24	1.032	1.033	1.034	1.035	1.036	1.036	1.038	1.039	1.040	1.041
30	1.022	1.023	1.024	1.025	1.025	1.026	1.027	1.028	1.029	1.029
40	1.014	1.015	1.015	1.016	1.016	1.016	1.017	1.018	1.018	1.019
60	1.007	1.007	1.008	1.008	1.008	1.008	1.009	1.009	1.009	1.009
120	1.002	1.002	1.002	1.002	1.002	1.002	1.002	1.003	1.003	1.003
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	71.040	75.624	79.752	84.733	88.236	77.745	82.529	86.830	92.010	95.649

Table 1 (cont.)

	p = 4, q = 11					p = 4, q = 13				
M \ α	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.330	1.352	1.374	1.402	1.422	1.369	1.393	1.417	1.446	1.468
3	1.222	1.235	1.247	1.262	1.274	1.254	1.268	1.281	1.298	1.310
4	1.164	1.173	1.181	1.191	1.198	1.190	1.200	1.209	1.220	1.228
5	1.127	1.134	1.140	1.147	1.152	1.150	1.157	1.163	1.171	1.177
6	1.103	1.107	1.112	1.118	1.122	1.122	1.127	1.132	1.139	1.143
7	1.085	1.089	1.092	1.097	1.100	1.102	1.106	1.110	1.115	1.118
8	1.071	1.075	1.077	1.081	1.084	1.086	1.090	1.093	1.097	1.100
9	1.061	1.064	1.066	1.069	1.071	1.074	1.077	1.080	1.083	1.086
10	1.053	1.055	1.057	1.060	1.062	1.065	1.067	1.070	1.073	1.075
12	1.041	1.043	1.044	1.046	1.047	1.050	1.052	1.054	1.056	1.058
14	1.033	1.034	1.035	1.037	1.038	1.041	1.042	1.044	1.045	1.047
16	1.027	1.028	1.029	1.030	1.031	1.033	1.035	1.036	1.037	1.038
18	1.022	1.023	1.024	1.025	1.026	1.028	1.029	1.030	1.031	1.032
20	1.019	1.020	1.020	1.021	1.022	1.024	1.025	1.026	1.027	1.027
24	1.014	1.015	1.015	1.016	1.016	1.018	1.019	1.019	1.020	1.020
30	1.010	1.010	1.010	1.011	1.011	1.012	1.013	1.013	1.014	1.014
40	1.006	1.006	1.006	1.007	1.007	1.008	1.008	1.008	1.008	1.009
60	1.003	1.003	1.003	1.003	1.003	1.004	1.004	1.004	1.004	1.004
120	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	56.369	60.481	64.201	68.710	71.893	65.422	69.832	73.810	78.616	82.001

Table 1 (cont.)

M \ α	p = 4, q = 15					p = 4, q = 17				
	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.406	1.432	1.456	1.488	1.511	1.440	1.468	1.494	1.527	1.551
3	1.284	1.299	1.313	1.331	1.344	1.313	1.329	1.344	1.363	1.377
4	1.216	1.226	1.236	1.248	1.256	1.240	1.252	1.262	1.275	1.284
5	1.172	1.179	1.187	1.195	1.202	1.193	1.201	1.209	1.218	1.225
6	1.141	1.147	1.153	1.159	1.164	1.160	1.166	1.172	1.180	1.185
7	1.118	1.123	1.128	1.133	1.137	1.135	1.140	1.145	1.151	1.155
8	1.101	1.105	1.109	1.113	1.116	1.116	1.120	1.124	1.129	1.133
9	1.087	1.091	1.094	1.098	1.101	1.101	1.105	1.108	1.112	1.115
10	1.077	1.080	1.082	1.085	1.088	1.089	1.092	1.095	1.098	1.101
12	1.060	1.063	1.065	1.067	1.069	1.070	1.073	1.075	1.078	1.080
14	1.049	1.051	1.052	1.054	1.056	1.057	1.059	1.061	1.063	1.065
16	1.040	1.042	1.043	1.045	1.046	1.048	1.049	1.051	1.053	1.054
18	1.034	1.035	1.036	1.038	1.039	1.040	1.042	1.043	1.045	1.046
20	1.029	1.030	1.031	1.032	1.033	1.035	1.036	1.037	1.038	1.039
24	1.022	1.023	1.023	1.024	1.025	1.026	1.027	1.028	1.029	1.030
30	1.015	1.016	1.016	1.017	1.017	1.019	1.019	1.020	1.020	1.021
40	1.010	1.010	1.010	1.011	1.011	1.012	1.012	1.012	1.013	1.013
60	1.005	1.005	1.005	1.005	1.005	1.006	1.006	1.006	1.006	1.007
120	1.001	1.001	1.001	1.001	1.001	1.002	1.002	1.002	1.002	1.002
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	74.397	79.082	83.298	88.379	91.952	83.308	88.250	92.689	98.028	101.776

Table 1 (cont.)

	p = 4, q = 19					p = 4, q = 21				
M \ α	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.473	1.502	1.529	1.563	1.588	1.504	1.533	1.562	1.598	1.624
3	1.340	1.357	1.373	1.393	1.408	1.367	1.384	1.401	1.422	1.437
4	1.264	1.276	1.287	1.300	1.310	1.287	1.299	1.311	1.325	1.335
5	1.214	1.223	1.231	1.241	1.248	1.234	1.243	1.252	1.262	1.270
6	1.178	1.185	1.191	1.199	1.205	1.196	1.203	1.210	1.218	1.224
7	1.151	1.157	1.162	1.169	1.173	1.167	1.173	1.179	1.186	1.190
8	1.130	1.135	1.140	1.145	1.149	1.145	1.150	1.155	1.160	1.164
9	1.114	1.118	1.122	1.126	1.130	1.127	1.132	1.136	1.140	1.144
10	1.101	1.104	1.107	1.111	1.114	1.113	1.116	1.120	1.124	1.127
12	1.080	1.083	1.086	1.089	1.091	1.091	1.094	1.096	1.099	1.102
14	1.066	1.068	1.070	1.073	1.074	1.075	1.077	1.079	1.082	1.084
16	1.055	1.057	1.059	1.061	1.062	1.063	1.065	1.066	1.069	1.070
18	1.047	1.048	1.050	1.051	1.053	1.054	1.055	1.057	1.059	1.060
20	1.040	1.042	1.043	1.044	1.045	1.046	1.048	1.049	1.051	1.052
24	1.031	1.032	1.033	1.034	1.035	1.036	1.037	1.038	1.039	1.040
30	1.022	1.023	1.023	1.024	1.025	1.025	1.026	1.027	1.028	1.028
40	1.014	1.014	1.015	1.015	1.015	1.016	1.017	1.017	1.018	1.018
60	1.007	1.007	1.007	1.008	1.008	1.008	1.008	1.009	1.009	1.009
120	1.002	1.002	1.002	1.002	1.002	1.002	1.002	1.003	1.003	1.003
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	92.166	97.351	101.999	107.583	111.495	100.980	106.395	111.242	117.057	121.126

Table 1 (cont.)

M \ α	p = 5, q = 11					p = 5, q = 13				
	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.312	1.333	1.352	1.378	1.396	1.345	1.367	1.387	1.414	1.433
3	1.213	1.225	1.236	1.250	1.260	1.241	1.253	1.265	1.280	1.290
4	1.159	1.167	1.174	1.183	1.190	1.182	1.191	1.199	1.208	1.215
5	1.125	1.130	1.136	1.142	1.147	1.145	1.151	1.157	1.164	1.169
6	1.101	1.105	1.110	1.115	1.118	1.118	1.123	1.128	1.133	1.137
7	1.084	1.087	1.091	1.095	1.098	1.099	1.103	1.107	1.111	1.114
8	1.071	1.074	1.077	1.080	1.082	1.084	1.088	1.091	1.094	1.097
9	1.061	1.063	1.066	1.068	1.070	1.073	1.076	1.078	1.081	1.083
10	1.053	1.055	1.057	1.059	1.061	1.064	1.066	1.068	1.071	1.073
12	1.041	1.043	1.044	1.046	1.047	1.050	1.052	1.054	1.055	1.057
14	1.033	1.034	1.035	1.037	1.038	1.040	1.042	1.043	1.045	1.046
16	1.027	1.028	1.029	1.030	1.031	1.033	1.035	1.036	1.037	1.038
18	1.023	1.023	1.024	1.025	1.026	1.028	1.029	1.030	1.031	1.032
20	1.019	1.020	1.021	1.021	1.022	1.024	1.025	1.026	1.026	1.027
24	1.014	1.015	1.015	1.016	1.016	1.018	1.019	1.019	1.020	1.020
30	1.010	1.010	1.011	1.011	1.011	1.013	1.013	1.013	1.014	1.014
40	1.006	1.006	1.006	1.007	1.007	1.008	1.008	1.008	1.009	1.009
60	1.003	1.003	1.003	1.003	1.003	1.004	1.004	1.004	1.004	1.004
120	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	68.796	73.311	77.380	82.292	85.749	79.973	84.821	89.177	94.422	98.105

le 1 (t.)

M \ α	p = 5, q = 15					p = 5, q = 17				
	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.377	1.399	1.421	1.449	1.469	1.407	1.431	1.453	1.482	1.503
3	1.267	1.281	1.293	1.309	1.320	1.293	1.307	1.320	1.336	1.348
4	1.205	1.214	1.223	1.233	1.240	1.227	1.237	1.246	1.257	1.265
5	1.164	1.171	1.177	1.185	1.190	1.184	1.191	1.198	1.206	1.212
6	1.136	1.141	1.146	1.152	1.156	1.153	1.159	1.164	1.170	1.175
7	1.115	1.119	1.123	1.127	1.131	1.130	1.134	1.139	1.144	1.147
8	1.098	1.102	1.105	1.109	1.112	1.112	1.116	1.119	1.124	1.127
9	1.085	1.088	1.091	1.094	1.097	1.098	1.101	1.104	1.108	1.110
10	1.075	1.078	1.080	1.083	1.085	1.086	1.089	1.092	1.095	1.097
12	1.059	1.061	1.063	1.065	1.067	1.069	1.071	1.073	1.075	1.077
14	1.048	1.050	1.051	1.053	1.054	1.056	1.058	1.060	1.062	1.063
16	1.040	1.041	1.043	1.044	1.045	1.047	1.048	1.050	1.051	1.052
18	1.034	1.035	1.036	1.037	1.038	1.040	1.041	1.042	1.044	1.044
20	1.029	1.030	1.031	1.032	1.033	1.034	1.035	1.036	1.037	1.038
24	1.022	1.023	1.023	1.024	1.025	1.026	1.027	1.028	1.029	1.029
30	1.015	1.016	1.016	1.017	1.017	1.019	1.019	1.020	1.020	1.021
40	1.010	1.010	1.010	1.011	1.011	1.012	1.012	1.012	1.013	1.013
60	1.005	1.005	1.005	1.005	1.005	1.006	1.006	1.006	1.006	1.006
120	1.001	1.001	1.001	1.001	1.002	1.002	1.002	1.002	1.002	1.002
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	91.061	96.217	100.839	106.393	110.286	102.079	107.522	112.393	118.236	122.325

Table 1 (cont.)

M \ α	p = 5, q = 19					p = 6, q = 15				
	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.436	1.460	1.483	1.513	1.535	1.357	1.377	1.397	1.422	1.440
3	1.318	1.332	1.346	1.363	1.375	1.256	1.268	1.279	1.293	1.303
4	1.249	1.259	1.268	1.280	1.288	1.198	1.206	1.214	1.223	1.230
5	1.203	1.210	1.217	1.226	1.232	1.160	1.166	1.171	1.178	1.183
6	1.170	1.176	1.181	1.188	1.193	1.132	1.137	1.142	1.147	1.151
7	1.145	1.150	1.154	1.160	1.164	1.112	1.116	1.120	1.124	1.127
8	1.126	1.130	1.134	1.138	1.141	1.097	1.100	1.103	1.106	1.109
9	1.110	1.114	1.117	1.121	1.124	1.084	1.087	1.089	1.092	1.095
10	1.097	1.101	1.103	1.107	1.109	1.074	1.076	1.079	1.081	1.083
12	1.078	1.081	1.083	1.085	1.087	1.059	1.061	1.062	1.064	1.066
14	1.064	1.066	1.068	1.070	1.072	1.048	1.050	1.051	1.053	1.054
16	1.054	1.056	1.057	1.059	1.060	1.040	1.041	1.042	1.044	1.045
18	1.046	1.047	1.049	1.050	1.051	1.034	1.035	1.036	1.037	1.038
20	1.040	1.041	1.042	1.043	1.044	1.029	1.030	1.031	1.032	1.032
24	1.031	1.031	1.032	1.033	1.034	1.022	1.023	1.023	1.024	1.025
30	1.022	1.022	1.023	1.024	1.024	1.016	1.016	1.017	1.017	1.017
40	1.014	1.014	1.015	1.015	1.015	1.010	1.010	1.010	1.011	1.011
60	1.007	1.007	1.007	1.008	1.008	1.005	1.005	1.005	1.005	1.005
120	1.002	1.002	1.002	1.002	1.002	1.001	1.001	1.001	1.002	1.002
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	113.038	118.752	123.858	129.973	134.247	107.565	113.145	118.136	124.116	128.299

Table 1 (cont.)

M \ α	p = 6, q = 17					p = 6, q = 19				
	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.383	1.404	1.424	1.450	1.469	1.408	1.430	1.451	1.477	1.497
3	1.279	1.291	1.303	1.317	1.328	1.301	1.314	1.326	1.341	1.352
4	1.218	1.226	1.234	1.244	1.251	1.237	1.246	1.255	1.265	1.273
5	1.177	1.184	1.190	1.197	1.202	1.195	1.201	1.208	1.215	1.221
6	1.148	1.153	1.158	1.164	1.168	1.164	1.169	1.174	1.180	1.184
7	1.126	1.130	1.134	1.139	1.142	1.140	1.145	1.149	1.154	1.157
8	1.109	1.113	1.116	1.120	1.123	1.122	1.126	1.129	1.133	1.136
9	1.096	1.099	1.101	1.105	1.107	1.107	1.110	1.113	1.117	1.119
10	1.085	1.087	1.090	1.092	1.094	1.095	1.098	1.101	1.104	1.106
12	1.068	1.070	1.072	1.074	1.075	1.077	1.079	1.081	1.083	1.085
14	1.056	1.057	1.059	1.061	1.062	1.063	1.065	1.067	1.069	1.070
16	1.047	1.048	1.049	1.051	1.052	1.053	1.055	1.056	1.058	1.059
18	1.040	1.041	1.042	1.043	1.044	1.046	1.047	1.048	1.049	1.050
20	1.034	1.035	1.036	1.037	1.038	1.039	1.041	1.041	1.043	1.043
24	1.026	1.027	1.028	1.028	1.029	1.030	1.031	1.032	1.033	1.033
30	1.019	1.019	1.020	1.020	1.021	1.022	1.022	1.023	1.023	1.024
40	1.012	1.012	1.012	1.013	1.013	1.014	1.014	1.015	1.015	1.015
60	1.006	1.006	1.006	1.006	1.007	1.007	1.007	1.007	1.008	1.008
120	1.002	1.002	1.022	1.002	1.002	1.002	1.002	1.002	1.002	1.002
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	120.679	126.574	131.838	138.134	142.532	133.729	139.921	145.441	152.037	156.637

Table 1 (cont.)

M \ α	p = 7, q = 11					p = 7, q = 13				
	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.297	1.315	1.332	1.354	1.371	1.320	1.338	1.356	1.378	1.395
3	1.207	1.218	1.227	1.239	1.248	1.228	1.238	1.248	1.261	1.270
4	1.157	1.164	1.171	1.179	1.184	1.175	1.182	1.189	1.197	1.203
5	1.125	1.130	1.135	1.140	1.145	1.141	1.146	1.151	1.157	1.161
6	1.102	1.106	1.110	1.114	1.118	1.116	1.120	1.124	1.129	1.132
7	1.086	1.089	1.092	1.095	1.098	1.098	1.102	1.105	1.108	1.111
8	1.073	1.076	1.078	1.081	1.083	1.084	1.087	1.090	1.093	1.095
9	1.063	1.065	1.067	1.070	1.072	1.073	1.076	1.078	1.080	1.082
10	1.055	1.057	1.059	1.061	1.062	1.064	1.066	1.068	1.071	1.072
12	1.043	1.045	1.046	1.048	1.049	1.051	1.053	1.054	1.056	1.057
14	1.035	1.036	1.037	1.038	1.039	1.042	1.043	1.044	1.045	1.046
16	1.029	1.030	1.031	1.032	1.032	1.035	1.036	1.036	1.038	1.038
18	1.024	1.025	1.026	1.027	1.027	1.029	1.030	1.031	1.032	1.032
20	1.021	1.021	1.022	1.023	1.023	1.025	1.026	1.026	1.027	1.028
24	1.016	1.016	1.017	1.017	1.017	1.019	1.020	1.020	1.021	1.021
30	1.011	1.011	1.012	1.012	1.012	1.013	1.014	1.014	1.014	1.015
40	1.007	1.007	1.007	1.007	1.008	1.008	1.009	1.009	1.009	1.009
60	1.003	1.003	1.004	1.004	1.004	1.004	1.004	1.004	1.004	1.005
120	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001	1.001
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	93.270	98.484	103.158	108.771	112.704	108.661	114.268	119.282	125.289	129.491

Table 1 (cont.)

M \ α	p = 7, q = 15					p = 7, q = 17				
	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.343	1.362	1.380	1.403	1.420	1.365	1.385	1.403	1.427	1.445
3	1.248	1.259	1.270	1.283	1.292	1.268	1.280	1.291	1.304	1.314
4	1.193	1.201	1.208	1.216	1.223	1.211	1.219	1.226	1.235	1.242
5	1.157	1.162	1.168	1.174	1.179	1.173	1.179	1.184	1.191	1.196
6	1.131	1.135	1.139	1.144	1.148	1.145	1.150	1.154	1.159	1.163
7	1.111	1.115	1.118	1.122	1.125	1.124	1.128	1.131	1.136	1.139
8	1.096	1.099	1.102	1.105	1.107	1.108	1.111	1.114	1.117	1.120
9	1.084	1.086	1.089	1.092	1.094	1.095	1.097	1.100	1.103	1.105
10	1.074	1.076	1.078	1.081	1.082	1.084	1.086	1.088	1.091	1.093
12	1.059	1.061	1.062	1.064	1.066	1.067	1.069	1.071	1.073	1.074
14	1.048	1.050	1.051	1.053	1.054	1.056	1.057	1.058	1.060	1.061
16	1.040	1.042	1.043	1.044	1.045	1.047	1.048	1.049	1.050	1.051
18	1.034	1.035	1.036	1.037	1.038	1.040	1.041	1.042	1.043	1.044
20	1.030	1.030	1.031	1.032	1.033	1.034	1.035	1.036	1.037	1.038
24	1.023	1.023	1.024	1.024	1.025	1.026	1.027	1.028	1.028	1.029
30	1.016	1.016	1.017	1.017	1.018	1.019	1.019	1.029	1.020	1.021
40	1.010	1.010	1.011	1.011	1.011	1.012	1.012	1.013	1.013	1.013
60	1.005	1.005	1.005	1.005	1.006	1.006	1.006	1.006	1.007	1.007
120	1.001	1.001	1.002	1.002	1.002	1.002	1.002	1.002	1.002	1.002
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	123.947	129.918	135.247	141.620	146.070	139.149	145.461	151.084	157.800	162.481

Table 1 (cont.)

M \ α	p = 7, q = 19					p = 8, q = 11				
	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.388	1.408	1.427	1.451	1.469	1.294	1.312	1.328	1.350	1.366
3	1.288	1.300	1.311	1.325	1.335	1.208	1.218	1.227	1.239	1.247
4	1.229	1.237	1.245	1.254	1.261	1.159	1.165	1.172	1.179	1.185
5	1.189	1.195	1.201	1.208	1.213	1.127	1.132	1.136	1.142	1.146
6	1.159	1.164	1.169	1.174	1.178	1.104	1.108	1.112	1.116	1.119
7	1.137	1.141	1.145	1.149	1.152	1.088	1.091	1.094	1.097	1.100
8	1.119	1.123	1.126	1.130	1.132	1.075	1.078	1.080	1.083	1.085
9	1.105	1.108	1.111	1.114	1.116	1.065	1.067	1.069	1.072	1.073
10	1.094	1.096	1.099	1.101	1.103	1.057	1.059	1.061	1.063	1.064
12	1.076	1.078	1.080	1.082	1.083	1.045	1.046	1.048	1.049	1.050
14	1.063	1.065	1.066	1.068	1.069	1.037	1.038	1.039	1.040	1.041
16	1.053	1.054	1.056	1.057	1.058	1.030	1.031	1.032	1.033	1.034
18	1.045	1.047	1.048	1.049	1.050	1.026	1.026	1.027	1.028	1.028
20	1.039	1.040	1.041	1.042	1.043	1.022	1.022	1.023	1.024	1.024
24	1.030	1.031	1.032	1.033	1.033	1.017	1.017	1.017	1.018	1.018
30	1.022	1.022	1.023	1.023	1.024	1.012	1.012	1.012	1.013	1.013
40	1.014	1.014	1.015	1.015	1.015	1.007	1.007	1.008	1.008	1.008
60	1.007	1.007	1.007	1.008	1.008	1.004	1.004	1.004	1.004	1.004
120	1.002	1.002	1.002	1.002	1.002	1.001	1.001	1.001	1.001	1.001
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	154.283	160.915	166.816	173.854	178.755	105.372	110.898	115.841	121.767	125.913

Table 1 (cont.)

	p = 8, q = 13					p = 8, q = 15				
M \ α	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.313	1.331	1.347	1.369	1.385	1.333	1.351	1.368	1.389	1.406
3	1.225	1.235	1.245	1.257	1.266	1.243	1.253	1.263	1.275	1.284
4	1.174	1.181	1.188	1.196	1.201	1.190	1.198	1.204	1.212	1.218
5	1.141	1.146	1.151	1.156	1.161	1.155	1.160	1.165	1.171	1.176
6	1.117	1.121	1.125	1.129	1.132	1.130	1.134	1.138	1.143	1.146
7	1.099	1.102	1.105	1.109	1.111	1.111	1.114	1.117	1.121	1.124
8	1.085	1.088	1.090	1.093	1.096	1.096	1.099	1.101	1.105	1.107
9	1.074	1.077	1.079	1.081	1.083	1.084	1.087	1.089	1.091	1.093
10	1.066	1.067	1.069	1.071	1.073	1.074	1.076	1.078	1.081	1.082
12	1.052	1.054	1.055	1.057	1.058	1.060	1.061	1.063	1.065	1.066
14	1.043	1.044	1.045	1.046	1.047	1.049	1.050	1.052	1.053	1.054
16	1.035	1.036	1.037	1.038	1.039	1.041	1.042	1.043	1.044	1.045
18	1.030	1.031	1.032	1.033	1.033	1.035	1.036	1.037	1.038	1.038
20	1.026	1.027	1.027	1.028	1.028	1.030	1.031	1.032	1.032	1.033
24	1.020	1.020	1.021	1.021	1.022	1.023	1.024	1.024	1.025	1.025
30	1.014	1.014	1.015	1.015	1.015	1.016	1.017	1.017	1.018	1.018
40	1.009	1.009	1.009	1.009	1.010	1.010	1.011	1.011	1.011	1.011
60	1.004	1.004	1.005	1.005	1.005	1.005	1.005	1.005	1.006	1.006
120	1.001	1.001	1.001	1.001	1.001	1.002	1.002	1.002	1.002	1.002
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	122.858	128.804	134.111	140.459	144.891	140.233	146.567	152.211	158.950	163.648

Table 1 (cont.)

M \ α	p = 8, q = 17 0.100	0.050	0.025	0.010	0.005	p = 9, q = 11 0.100	0.050	0.025	0.010	0.005
2	1.353	1.371	1.388	1.410	——	1.294	1.311	1.327	1.348	1.364
3	1.261	1.272	1.282	1.294	1.303	1.210	1.219	1.229	1.240	1.248
4	1.207	1.214	1.221	1.229	1.235	1.161	1.168	1.174	1.182	1.187
5	1.170	1.175	1.180	1.187	1.191	1.130	1.134	1.139	1.144	1.148
6	1.143	1.147	1.151	1.156	1.160	1.107	1.111	1.114	1.119	1.122
7	1.123	1.126	1.130	1.134	1.136	1.091	1.094	1.096	1.100	1.102
8	1.107	1.110	1.113	1.116	1.118	1.078	1.080	1.083	1.085	1.087
9	1.094	1.097	1.099	1.102	1.104	1.068	1.070	1.072	1.074	1.076
10	1.084	1.086	1.088	1.090	1.092	1.059	1.061	1.063	1.065	1.066
12	1.067	1.069	1.071	1.073	1.074	1.047	1.048	1.050	1.051	1.052
14	1.056	1.057	1.058	1.060	1.061	1.038	1.039	1.040	1.042	1.043
16	1.047	1.048	1.049	1.050	1.051	1.032	1.033	1.034	1.035	1.035
18	1.040	1.041	1.042	1.043	1.044	1.027	1.028	1.028	1.029	1.030
20	1.035	1.036	1.036	1.037	1.038	1.023	1.024	1.024	1.025	1.025
24	1.027	1.027	1.028	1.029	1.029	1.018	1.018	1.018	1.019	1.019
30	1.019	1.020	1.020	1.021	1.021	1.012	1.013	1.013	1.013	1.014
40	1.012	1.013	1.013	1.013	1.013	1.008	1.008	1.008	1.008	1.008
60	1.006	1.006	1.007	1.007	1.007	1.004	1.004	1.004	1.004	1.004
120	1.002	1.002	1.002	1.002	1.002	1.001	1.001	1.001	1.001	1.001
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	157.518	164.216	170.175	177.280	182.226	117.407	123.225	128.422	134.642	138.987

Table 1 (cont.)

M \ α	p = 9, q = 13					p = 9, q = 15				
	0.100	0.050	0.025	0.010	0.005	0.100	0.050	0.025	0.010	0.005
2	1.310	1.326	1.343	1.363	1.379	1.326	1.343	1.359	—	—
3	1.224	1.234	1.243	1.255	1.263	1.240	1.250	1.259	1.271	1.279
4	1.175	1.181	1.188	1.195	1.201	1.189	1.196	1.202	1.210	1.216
5	1.142	1.147	1.151	1.157	1.161	1.155	1.160	1.165	1.170	1.174
6	1.118	1.122	1.126	1.130	1.133	1.130	1.134	1.138	1.142	1.145
7	1.101	1.104	1.107	1.110	1.113	1.111	1.115	1.118	1.121	1.124
8	1.087	1.089	1.092	1.095	1.097	1.097	1.099	1.102	1.105	1.107
9	1.076	1.078	1.080	1.083	1.084	1.085	1.087	1.089	1.092	1.094
10	1.067	1.069	1.071	1.073	1.074	1.075	1.077	1.079	1.081	1.083
12	1.054	1.055	1.056	1.058	1.059	1.061	1.062	1.064	1.065	1.066
14	1.044	1.045	1.046	1.047	1.048	1.050	1.051	1.052	1.054	1.055
16	1.037	1.038	1.039	1.040	1.040	1.042	1.043	1.044	1.045	1.046
18	1.031	1.032	1.033	1.034	1.034	1.036	1.037	1.037	1.038	1.039
20	1.027	1.028	1.028	1.029	1.029	1.031	1.032	1.032	1.033	1.034
24	1.020	1.021	1.021	1.022	1.022	1.024	1.024	1.025	1.025	1.026
30	1.015	1.015	1.015	1.016	1.016	1.017	1.017	1.018	1.018	1.018
40	1.009	1.009	1.010	1.010	1.010	1.011	1.011	1.011	1.012	1.012
60	1.005	1.005	1.005	1.005	1.005	1.005	1.006	1.006	1.006	1.006
120	1.001	1.001	1.001	1.001	1.001	1.002	1.002	1.002	1.002	1.002
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	136.982	143.246	148.829	155.496	160.146	156.440	163.116	169.056	176.138	181.070

Table 1 (cont.)

p = 10, q = 11

M \ α	0.100	0.050	0.025	0.010	0.005
2	1.296	1.313	1.329	1.349	——
3	1.213	1.222	1.231	1.243	1.251
4	1.165	1.171	1.177	1.185	1.190
5	1.133	1.138	1.142	1.148	1.152
6	1.111	1.114	1.118	1.122	1.125
7	1.094	1.097	1.099	1.103	1.105
8	1.081	1.083	1.085	1.088	1.090
9	1.070	1.072	1.074	1.077	1.078
10	1.062	1.064	1.065	1.067	1.069
12	1.049	1.051	1.052	1.054	1.055
14	1.040	1.041	1.042	1.044	1.044
16	1.034	1.034	1.035	1.036	1.037
18	1.028	1.029	1.030	1.031	1.031
20	1.024	1.025	1.026	1.026	1.027
24	1.019	1.019	1.020	1.020	1.020
30	1.013	1.013	1.014	1.014	1.014
40	1.008	1.008	1.009	1.009	1.009
60	1.004	1.004	1.004	1.004	1.005
120	1.001	1.001	1.001	1.001	1.001
∞	1.000	1.000	1.000	1.000	1.000
χ^2_{pq}	129.385	135.480	140.917	147.414	151.948

Reprinted from: © Journal of the American Statistical Association
March 1977, Volume 72, Number 357
Theory and Methods Section
Pages 212–214

A Differential Equation Approach to Linear Combinations of Independent Chi-Squares

A. W. DAVIS

A Differential Equation Approach to Linear Combinations of Independent Chi-Squares

A. W. DAVIS*

The distribution of a linear combination of m independent central chi-square variables with positive coefficients is shown to be derivable from an mth order linear differential equation. The computation of percentage points by analytic continuation of the series solution at the origin is shown to be feasible for a useful range of parameter values.

KEY WORDS: Tabulation of distributions; Linear combinations of independent chi-squares; Differential equations; Power series; Analytic continuation.

1. INTRODUCTION

Many papers have been published on the distribution of linear combinations

$$y = \sum_{i=1}^{m} \alpha_i \chi_{\nu_i}^2 , \qquad (1.1)$$

where the α_i are known constants and the $\chi_{\nu_i}^2$ are independent central chi-square variables with ν_i degrees of freedom. Alternatively, such linear combinations may be regarded as quadratic forms in normal variables. For a comprehensive survey of the literature, including applications, distribution theory, and tabulation, see Johnson and Kotz [5, Ch. 29]. Various representations of the distribution have been given when all α_i are positive (positive definite case), including certain mixtures of chi-square distributions, Laguerre series expansions, and power series expansions about $y = 0$.

The most successful approach so far to tabulating y appears to have been that of Johnson and Kotz [4], based on Laguerre series; they give percentage points for selected linear combinations of up to five independent chi-squares on single degrees of freedom. They remark [5, p. 163] that the power series about $y = 0$ "appears to be useful for computation only for small values of y." However, we seek to show in this note that the problem of slow convergence of the power series for larger y may be overcome, at least for a useful range of parameter values, by deriving an mth order linear differential equation (d.e.) for the distribution. This provides a convenient means of expanding the distribution about points $y = a \neq 0$, so that the solution may be followed out from the origin by analytic continuation. A check on the accuracy of the procedure described next is obtained by transforming the d.e. and recalculating; results appear to be accurate to about seven significant figures for $m \leq 10$, all $\nu_i = 1$, at least for the particular parameter values considered. A similar approach has been used by the author in connection with other complex distributions [1, 2].

2. THE DIFFERENTIAL EQUATION

Denoting the Laplace transform of the distribution of y by $L_0(s)$, we have

$$L_0(s) = \prod_{i=1}^{m} (1 + 2s\alpha_i)^{-\frac{1}{2}\nu_i} , \quad s > 0 , \qquad (2.1)$$

so that

$$\frac{dL_0(s)}{ds} = - \sum_{i=1}^{m} \nu_i L_i(s) , \qquad (2.2)$$

where

$$L_i(s) = \alpha_i L_0(s)/(1 + 2s\alpha_i) , \quad i = 1, \ldots, m . \qquad (2.3)$$

These functions satisfy

$$sL_i(s) = \tfrac{1}{2}[L_0(s) - \alpha_i^{-1} L_i(s)] , \quad i = 1, \ldots, m . \qquad (2.4)$$

Now let $H_i(y)$ be the inverse Laplace transform of $L_i(s)$ $(i = 0, \ldots, m)$; i.e., $H_0(y)$ is the density function of y and the auxiliary function $H_i(y)$ is proportional to the density of the linear combination obtained by increasing ν_i to $\nu_i + 2$, $(i = 1, \ldots, m)$. Write

$$\mathbf{H} = (H_1, \ldots, H_m)' , \quad \boldsymbol{\nu} = (\nu_1, \ldots, \nu_m)' ,$$
$$\mathbf{A} = \text{diag}\,(\alpha_1^{-1}, \ldots, \alpha_m^{-1}) ,$$

and let $\mathbf{1}$ denote the m vector $(1, 1, \ldots, 1)'$. Then from (2.2) and (2.4), respectively, we obtain

$$yH_0 = \boldsymbol{\nu}'\mathbf{H} , \qquad (2.5)$$

and

$$d\mathbf{H}/dy = \tfrac{1}{2}[\mathbf{1}H_0 - \mathbf{A}\mathbf{H}]$$
$$= \tfrac{1}{2}[y^{-1}\mathbf{1}\boldsymbol{\nu}' - \mathbf{A}]\mathbf{H} . \qquad (2.6)$$

Equation (2.6) is seen to be a homogeneous linear d.e. of order m for $\mathbf{H}$, with a regular singularity at $y = 0$. The theory of such d.e.'s is presented, for example, in [3, Ch. 5, Appendix B]; in general, a d.e. of this type has m linearly independent solutions, any particular solution being some linear combination of these. A fundamental system of solutions in the neighborhood of

* A. W. Davis is Senior Principal Research Scientist, CSIRO Division of Mathematics and Statistics, Glen Osmond, South Australia 5064. The author is grateful to Miss J. A. Hawkes for her assistance with the computing.

March 1977, Volume 72, Number 357
Theory and Methods Section

$y = 0$ corresponds to the latent roots of the matrix coefficient of y^{-1}, in this case $\frac{1}{2}\mathbf{1v}'$. This particular matrix is of rank 1, having a zero root of multiplicity $m - 1$ and a single positive root

$$\lambda = \tfrac{1}{2}\mathbf{v}'\mathbf{1} = \tfrac{1}{2} \sum_{i=1}^{m} \nu_i \ . \tag{2.7}$$

Since (2.6) has no other singularities for finite y, the series solutions at $y = 0$ will be convergent for $|y| < \infty$.

In the positive definite case, with all α_i positive, it may be shown, by expressing the distribution in convolution form, that the density function

$$H_0(y) \sim Ky^{\lambda-1} \quad \text{as} \quad y \to 0+ \ , \tag{2.8}$$

where

$$K = [2^\lambda \Gamma(\lambda) \prod_{i=1}^{m} \alpha_i^{\frac{1}{2}\nu_i}]^{-1} \ .$$

Alternatively, (2.8) follows from the power series expansion of Robbins [8]. In view of (2.5), (2.8) is sufficient to identify the required solution of (2.6) as the one corresponding to the root λ, since any linear combination of solutions corresponding to the zero root must involve terms of lower order as $y \to 0$. Substituting

$$H_i(y) = \sum_{j=0}^{\infty} C_{ij} y^{\lambda+j-1} \ , \quad i = 0, \ldots, m \ , \tag{2.9}$$

in (2.5) and (2.6), we obtain

$$C_{00} = K \ , \quad C_{i0} = 0 \ , \quad i = 1, \ldots, m \ , \tag{2.10}$$

and for $j = 1, 2, \ldots,$

$$\begin{aligned} C_{ij} &= (C_{0,j-1} - \alpha_i^{-1} C_{i,j-1})/2(j + \lambda - 1) \\ C_{0j} &= -(2j)^{-1} \sum_{i=1}^{m} \alpha_i^{-1} \nu_i C_{ij} \ . \end{aligned} \tag{2.11}$$

The resulting power series for $H_0(y)$ is identical with that given by Robbins [6], who gave a different recurrence relation for the coefficients, however.

In the case where some α_i are negative, the distribution still corresponds to a solution of the d.e., but (2.8) no longer holds, and there are difficulties in identifying and constructing the solution. The d.e. approach is thus probably not feasible in the indefinite case.

As indicated in the introduction, the advantage claimed for the d.e. is that it enables (2.9) to be re-expanded about a point $a \neq 0$. Substituting

$$H_i(y) = \sum_{j=0}^{\infty} D_{ij}(y - a)^j \ , \quad i = 0, \ldots, m \ , \tag{2.12}$$

in (2.5) and (2.6) yields

$$\begin{aligned} D_{i0} &= H_i(a) \ , \quad i = 1, \ldots, m \ , \\ D_{00} &= a^{-1} \sum_{i=1}^{m} \nu_i D_{i0} = H_0(a) \end{aligned} \tag{2.13}$$

and for $j = 1, 2, \ldots,$

$$\begin{aligned} D_{0j} &= a^{-1}[\sum_{i=1}^{m} \nu_i D_{ij} - D_{0,j-1}] \ , \\ D_{ij} &= (2j)^{-1}[D_{0,j-1} - \alpha_i^{-1} D_{i,j-1}] \ , \\ & \qquad i = 1, \ldots, m \ . \end{aligned} \tag{2.14}$$

Hence, only the values of the H_i at $y = a$ are required to generate the expansion (2.12).

The cumulative distribution function

$$F(y) = \int_0^y H_0(u) du$$

may be obtained at any stage by integrating H_0 term by term and summing the resulting series.

3. COMPUTATION AND RESULTS

The d.e. (2.5)–(2.6) may be applied either to calculate the probability $F(y_1)$ corresponding to an assigned value $y = y_1$, or to calculate the percentile y_ϵ for an assigned probability ϵ. The computing procedure used may be summarized as follows:

1. The series (2.9) was expanded to forty terms using (2.10) and (2.11), and a value $y = a > 0$ was constructed for which the last six terms ine ach expansion did not exceed 10^{-10} in absolute value,
$$a = 1/\max \{(10^{10}|C_{ij}|)^{1/j}, \quad i = 0, 1, \ldots, m; \ j = 35, \ldots, 40\} \ . \tag{3.1}$$
2. If a fell short of the required value y_1 or y_ϵ, the series was expanded about a using (2.13) and (2.14), and the process was repeated as often as required. Percentiles y_ϵ were obtained by iteration.
3. To check on accuracy, the d.e. was transformed onto a finite interval $(0, c)$ by substituting
$$z = cy/(y + 1) \tag{3.2}$$
in (2.5)–(2.6); the procedure was then carried out on the resulting d.e. The value $c = 10$ was found convenient for the parameter range considered. For smaller c, the gradient of the transformed distribution became excessive as $y \to c-$, making accurate tabulation of the upper tail increasingly difficult.

Calculations were carried out in double precision on the CSIRO Control Data Cyber 70 model 76 in Canberra. Execution time for the results in Tables 1 through 3 was 1.102 seconds. It is hoped that the algorithm will be published elsewhere.

Table 1 presents percentage points of $\chi_{10}^2(\alpha_i = \nu_i = 1, i = 1, \ldots, 10)$ obtained using (a) the d.e. (2.5)–(2.6) and (b) the transformed d.e. For purposes of comparison, the a_i and y_ϵ values in (b) have been obtained by inversion of (3.2). It is seen that in the case of the original d.e., the series expansion about the origin yielded $a = a_1 = 13.00$, with $F(a_1) = 0.7763$. Expanding in a new series about a_1, the value $a_2 = 24.29$ was obtained, with $F(a_2) = 0.9931$. The a_1 series was thus used to compute the y_ϵ for $\epsilon = .900, .950$, and $.990$. In the case of the transformed d.e., five changes of origin were re-

1. Upper Percentage Points of χ_{10}^2

i	a_i	$F(a_i)$	ϵ	y_ϵ
		a. Differential equation		
1	13.00	.7763		
			.900	15.98718
			.950	18.30704
			.990	23.20925
2	24.29	.9931		
			.995	25.18818
3	37.22	.9999		
		b. Transform[a]		
1	.6224	$.1878 \times 10^{-4}$		
2	1.409	$.8070 \times 10^{-3}$		
3	2.796	.01418		
4	5.242	.1256		
5	9.456	.5106		
			.900	15.98718
6	16.91	.9236		
			.950	18.30704
			.990	23.20925
			.995	25.18818
7	30.39	.9993		

[a] The a_i and y_ϵ values in b have been obtained by inversion of (3.2) with $c = 10$.

quired before the .900 point could be calculated, and an additional step was necessary to obtain the remaining points. The results are accurate to five decimal places.

2. Upper Percentage Points of $y(\nu_1 = \ldots = \nu_5 = 1)$[a]

α_1	α_2	α_3	α_4	α_5	ϵ	Johnson and Kotz	d.e. and transform	Steps[b] d.e.	Steps[b] Trans.
3.0	.5	.5	.5	.5	.900	10.407	10.40708	1	3
					.950	13.778	13.77754 (+1)	2	4
					.990	22.259	22.12712	3	5
					.995	25.856	25.85486		
2.5	1.2	.5	.4	.4	.900	10.101	10.10127	2	3
					.950	12.932	12.92783	3	4
					.990	19.826	19.82597		
					.995	22.907	22.90679	4	5
1.8	1.8	.6	.4	.4	.900	9.921	9.92111	2	3
					.950	12.419	12.41859		4
					.990	18.214	18.21316	3	
					.995	20.709	20.70851 (−1)		
1.1	1.1	1.0	.9	.9	.900	9.248	9.24823	0	4
					.950	11.096	11.09629		
					.990	15.156	15.15568	1	
					.995	16.843	16.84227		

[a] (±1) indicates that the result given by the transformed d.e. is obtained by adding ±1 to the digit in the fifth decimal place.

[b] The accumulated number of changes of origin required.

Some comparisons with the Johnson and Kotz [4] tabulation are presented in Table 2. Taking $c = 10$, the results obtained from the d.e. and its transform generally agreed to five decimal places, confirming the values given by the Laguerre expansion to the number of figures presented (except in one or two instances). The total number of changes of origin required to obtain each percentage point is also shown for the two methods.

Some percentage points for higher values of m are given in Table 3 to illustrate the scope of the method.

3. Upper Percentage Points of y(all $\nu_i = 1$)

α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	ϵ	d.e. and transform	Steps d.e.	Steps Trans.
3	1	1	1	.5	.3	.2				.900	13.23966	3	4
										.950	16.57786 (−1)	4	
										.990	24.80656	5	5
										.995	28.50349	6	
3	2	1	1	.4	.3	.2	.1			.900	15.13968	5	4
										.950	18.73166	6	5
										.990	27.20669	9	
										.995	30.92252	10	
4	2	1	1	.2	.2	.2	.2	.2		.900	17.41478	4	5
										.950	21.95221	5	
										.990	32.98203	6	
										.995	37.90350	7	6
5	2	1	1	.5	.1	.1	.1	.1	.1	.900	19.82018	7	5
										.950	25.41102	9	
										.990	39.20686 (+1)	13	7
										.995	45.38733 (−1)	15	8

[*Received May 1975. Revised December 1975.*]

REFERENCES

[1] Davis, A.W., "Exact Distributions of Hotelling's Generalized T_0^2," *Biometrika*, 57 (April 1970), 187–91.

[2] ———, "On the Distributions of the Latent Roots and Traces of Certain Random Matrices," *Journal of Multivariate Analysis*, 2 (June 1972), 189–200.

[3] Hille, E., *Lectures on Ordinary Differential Equations*, Reading, Mass.: Addison-Wesley Publishing Co., 1969.

[4] Johnson, N.L., and Kotz, S., "Tables of Distributions of Positive Definite Quadratic Forms in Central Normal Variables," *Sankhyā, Ser. B*, 30 (December 1968), 303–14.

[5] ———, and Kotz, S., *Distributions in Statistics: Continuous Univariate Distributions Vol. 2*, Boston: Houghton Mifflin Co., 1970.

[6] Robbins, H.E., "The Distribution of a Definite Quadratic Form," *Annals of Mathematical Statistics*, 19 (1948), 266–70.

Section 3.

Approximations to statistical distributions.

16

Commonwealth of Australia
COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANIZATION
Reprinted from "Annals of Mathematical Statistics".
Vol. 39, No. 4, Pages 1264–1273, Aug. 1968

GENERALIZED ASYMPTOTIC EXPANSIONS OF CORNISH-FISHER TYPE

G. W. HILL AND A. W. DAVIS

C.S.I.R.O., Adelaide

1. Introduction and summary. Let $\{F_n(x)\}$ be a sequence of distribution functions depending on a parameter n, and converging to a limiting distribution $\Phi(x)$ as n increases. Then a generalized expansion of Cornish-Fisher type is an asymptotic relation between the quantiles of F_n and Φ. The original Cornish-Fisher formulae [3], [5] provided leading terms of these expansions in the case of normal Φ, expressing a normal deviate in terms of the corresponding quantile of F_n and its cumulants (the "normalizing" expansion) and, conversely, the quantiles of F_n in terms of its cumulants and the corresponding quantiles of Φ (the "inverse" expansion). The value of both these asymptotic formulae has been well illustrated by their use in approximating the quantiles of complicated distributions (Johnson and Welch [9], Fisher [4], Goldberg and Levine [6]), and for obtaining random quantiles for distribution sampling applications (Teichroew [13], Bol'shev [2]). For a survey of the literature on Cornish-Fisher expansions, and some discussion of their validity, see Wallace ([14], Section 4).

In Sections 2, 3 of the present paper, formal expansions are obtained which generalize the Cornish-Fisher relations to arbitrary analytic Φ. Essentially, these expansions provide algorithms for transforming an asymptotic expansion of F_n in terms of the "standard" distribution Φ into asymptotic relations between the quantiles of these distributions. The "standardizing" expansion of the quantile u of Φ in terms of the corresponding quantile x of F_n is expressed (Section 2) in terms of a sequence of functions defined by a differential recurrence operator. A similar differential operator appears in the generalized "inverse" expansion for x in terms of u (Section 3), which arises from the application of Lagrange's inversion formula to the equation of quantiles. An asymptotic expansion for quantiles of the Wilks likelihood ratio criterion is given as an example.

Formal expansions in terms of the cumulants of F_n and Φ are obtained in Section 4 by developing F_n about Φ as a Charlier differential series and collecting terms of like degree in the resulting exponential series. For known cumulants and for normal Φ these formal expressions reduce, as shown in Section 5, to a general form of the Cornish-Fisher expansions, in which the polynomial terms are represented as sums of products of Hermite polynomials. This representation is shown in Section 6 to account for some properties of the Cornish-Fisher polynomials.

2. The general standardizing expansion. If x and u are corresponding quantiles of F_n and Φ respectively, then

$$F_n(x) = \Phi(u) \tag{1}$$

and it is required to solve this equation for u in terms of x.

Received 1 May 1967.

The density function $\phi(u)$ of the distribution Φ will be assumed to be arbitrarily differentiable. Then writing

$$Z_n(x) = F_n(x) - \Phi(x), \tag{2}$$

it follows from (1) that

$$Z_n(x) = \int_x^u \phi(t)\,dt. \tag{3}$$

If the equation

$$\zeta = \int_x^u \phi(t)\,dt \tag{4}$$

is regarded as defining a function $u(\zeta)$ with $u(0) = x$, then $u(\zeta)$ may be developed in a formal Taylor series about $\zeta = 0$.

Differentiation of (4) yields

$$du/d\zeta = [\phi(u)]^{-1}. \tag{5}$$

Now writing

$$D_u \equiv d/du, \qquad \psi(u) = -\phi'(u)/\phi(u) = D_u \log (1/\phi(u)); \tag{6}$$

it is found by induction that

$$d^r u/d\zeta^r = c_r(u)[\phi(u)]^{-r}, \tag{7}$$

where the c_r are defined recursively by

$$c_1(u) \equiv 1, \qquad c_{r+1}(u) = (r\psi(u) + D_u)c_r(u), \qquad (r = 1, 2, \cdots). \tag{8}$$

Since $u = x$ when $\zeta = 0$, the Taylor series is seen to be

$$u(\zeta) = x + \sum_{r=1}^{\infty} c_r(x)(\zeta/\phi(x))^r/r!; \tag{9}$$

and applying this result to (3), the general standardizing expansion is obtained in the form

$$u = x + \sum_{r=1}^{\infty} c_r(x)(Z_n(x)/\phi(x))^r/r!. \tag{10}$$

In many applications $F_n(x)$ is known to have an asymptotic expansion of the form

$$\begin{aligned} F_n(x) &= \Phi(x) + \phi(x)[n^{-1}p_1(x) + n^{-2}p_2(x) + \cdots] \\ &= \Phi(x) + \phi(x)z_n(x), \end{aligned} \tag{11}$$

say, where the $p_r(x)$ may be polynomials in x. In terms of $z_n(x)$, (10) becomes

$$u = x + \sum_{r=1}^{\infty} c_r(x)(z_n(x))^r/r!, \tag{12}$$

which expresses the quantile u directly as a series in terms of x whose rth term is $O(n^{-r})$.

When the limiting distribution, $\Phi(x)$, is the unit normal distribution,

$$\psi(x) = D_x \log ((2\pi)^{\frac{1}{2}}e^{x^2/2}) = x, \tag{13}$$

and $c_r(x)$ is an $(r - 1)$th degree polynomial in x:

$$c_1(x) \equiv 1, \quad c_2(x) = x, \quad c_3(x) = 2x^2 + 1, \quad c_4(x) = 6x^3 + 7x, \cdots. \tag{14}$$

In this case, (12) is essentially the Cornish-Fisher normalizing expansion, which will be considered in more detail in Sections 5, 6.

For other applications, however, the appropriate limiting function Φ may be the distribution of χ^2 with ν degrees of freedom. Then

$$\psi(x) = \tfrac{1}{2} - (\tfrac{1}{2}\nu - 1)x^{-1}, \tag{15}$$

and $c_r(x)$ is an $(r-1)$th degree polynomial in x^{-1}.

3. The general inverse expansion. The solution of (1) for x in terms of u could be obtained from (10) or (12) by inverting the series, which suggests the application of Lagrange's inversion formula. This formula provides under certain conditions (see [15], p. 133) that, if γ and θ are analytic functions and

$$w = v + \gamma(w), \tag{16}$$

then

$$\theta(w) = \theta(v) + \textstyle\sum_{r=1}^{\infty} D_v^{\,r-1}[\theta'(v)(\gamma(v))^r]/r!\,. \tag{17}$$

Cornish and Fisher ([3], p. 316) in effect used the early terms of the Lagrange formula when inverting their normalizing expansion. Riordan [11] applied (17) with $\theta(w) \equiv w$ to derive general relations between the polynomials occurring in the two expansions.

Although the solution can be established by inverting the series (10) or (12), it appears more instructive to apply the Lagrange formula directly to equation (1), rewritten in the form

$$\Phi(x) = \Phi(u) - Z_n(x). \tag{18}$$

If new variables v and w are defined by

$$v = \Phi(u), \qquad w = \Phi(x), \tag{19}$$

then equivalently

$$w = v - Z_n(\Phi^{-1}(w)) \tag{20}$$

where Φ^{-1} denotes the inverse function of Φ. Since (20) is of the form (16), this functional equation can be solved for $\Phi^{-1}(w) = x$ by taking $\theta \equiv \Phi^{-1}$ in (17):

$$\Phi^{-1}(w) = \Phi^{-1}(v) + \textstyle\sum_{r=1}^{\infty} (-1)^r(r!)^{-1}D_v^{\,r-1}\{[Z_n(\Phi^{-1}(v))]^r/\phi(\Phi^{-1}(v))\}, \tag{21}$$

or, on reverting to the original variables:

$$x = u - \textstyle\sum_{r=1}^{\infty} (r!)^{-1}(-[\phi(u)]^{-1}D_u)^{r-1}[(Z_n(u))^r/\phi(u)]. \tag{22}$$

In cases where Z_n is a multiple of ϕ as in (11), (22) takes the form

$$x = u - \textstyle\sum_{r=1}^{\infty} D_{(r)}(z_n(u))^r/r!\,, \tag{23}$$

where $D_{(1)}$ denotes the identity operator and

$$D_{(r)} = (\psi(u) - D_u)(2\psi(u) - D_u) \cdots ((r-1)\psi(u) - D_u), \quad (r = 2, 3, \cdots). \tag{24}$$

As in the case of the standardizing expansion the rth term in the general inverse expansion (23) is seen to be $O(n^{-r})$.

EXAMPLE 1. *For normal* Φ, $\psi(u) = u$, and if z_n is expressed in terms of cumulants, (23) becomes the Cornish-Fisher inverse expansion as shown in the following sections.

EXAMPLE 2. *The Wilks likelihood ratio criterion.* Let $\mathbf{X}$, $\mathbf{Y}$ be $p \times m$ $(m > p)$ and $p \times q$ matrices respectively with the joint probability density function

$$(25) \quad (2\pi)^{-\frac{1}{2}p(m+q)} |\mathbf{\Sigma}|^{-\frac{1}{2}(m+q)} \exp \{ -\tfrac{1}{2} \operatorname{tr} \mathbf{\Sigma}^{-1}[\mathbf{XX}' + (\mathbf{Y} - \boldsymbol{\mu})(\mathbf{Y} - \boldsymbol{\mu})'] \}$$

where $\mathbf{\Sigma}$ is a $p \times p$ positive definite matrix. Then the likelihood ratio criterion for testing the hypothesis $\boldsymbol{\mu} = 0$ is

$$(26) \qquad \Lambda = \det (\mathbf{XX}')/\det (\mathbf{XX}' + \mathbf{YY}').$$

Let

$$(27) \qquad n = m - \tfrac{1}{2}(p - q + 1).$$

Then $-n \log \Lambda$ is asymptotically distributed as χ^2 with $\nu = pq$ degrees of freedom, and Rao ([10], see also [1] Theorem 8.6.2) has developed the distribution function in an asymptotic expansion of the form (11) to order n^{-4}. Let x and u be corresponding quantiles for $-n \log \Lambda$ and χ_ν^2. Schatzoff [12] has tabulated exact values of the correction factor x/u for $q = 4, 6, 8, 10$ and p such that $pq \leqq 70$. Applying the first three terms of (23) to Rao's expansion with ψ given by (15), it is found that

$$(28) \quad \begin{aligned} x/u \sim{}& 1 + n^{-2} 2\gamma_2(u + (\nu + 2))[\nu(\nu + 2)]^{-1} \\ &+ n^{-4}\{2\gamma_4[u^3 + (\nu + 6)u^2 + (\nu + 4)(\nu + 6)u \\ &+ (\nu + 2)(\nu + 4)(\nu + 6)][\nu(\nu + 2)(\nu + 4)(\nu + 6)]^{-1} \\ &- \gamma_2^2[u^3 + (\nu - 2)u^2 + (\nu + 2)(\nu - 6)u + (\nu - 2)(\nu + 2)^2 \\ &\quad \cdot [\nu^2(\nu + 2)^2]^{-1}\} \\ &+ O(n^{-6}), \end{aligned}$$

where

$$(29) \quad \begin{aligned} \gamma_2 &= pq(48)^{-1}(p^2 + q^2 - 5), \\ \gamma_4 &= \tfrac{1}{2}\gamma_2^2 + pq(1920)^{-1}[3p^4 + 3q^4 + 10p^2q^2 - 50(p^2 + q^2) + 159]. \end{aligned}$$

At the upper 5% and 1% levels, formula (28) gives results agreeing with those of Schatzoff to within 0.1% for n ranging from $n \doteq 4$ for $p = 3$, $q = 4$ to $n \doteq 10$ for $p = 7$, $q = 10$. Presumably, similar accuracy would apply for q odd.

EXAMPLE 3. *Hotelling's generalized* T_0^2. Similarly, Ito's expansion of Hotelling's generalized T_0^2 statistic in terms of χ^2 quantiles ([8] equation (3.33)) may be obtained directly from his expansion of the cumulative distribution function ([8] equation (4.3)), by applying the first three terms of (23) with ψ defined by

(15). This expansion includes the case of the F-ratio, which is asymptotic to χ^2 when the number of degrees of freedom in the denominator is large.

4. Expansion in terms of cumulants. If the cumulants of F_n and Φ are $\{\kappa_r\}$ and $\{\gamma_r\}$, respectively, then ([14], Section 3) F_n may be formally expanded about Φ in the Charlier differential series

$$F_n(x) = \exp [\textstyle\sum_{r=1}^{\infty} \lambda_r(-D_x)^r/r!]\Phi(x), \tag{30}$$

where

$$\lambda_r = \kappa_r - \gamma_r. \tag{31}$$

In developing the exponential series of (30), it is convenient to denote by π a partition of the positive integer m into l positive integers:

$$\pi = [s_1^{\rho_1}, \cdots, s_k^{\rho_k}], \qquad m = \textstyle\sum_{i=1}^k \rho_i s_i, \qquad l = \textstyle\sum_{i=1}^k \rho_i, \tag{32}$$

and let $a(\pi)$ denote the elementary partition function:

$$a(\pi) = m! [(s_1 !)^{\rho_1} \cdots (s_k !)^{\rho_k} \rho_1 ! \cdots \rho_k !)]^{-1}. \tag{33}$$

Then defining

$$\lambda_\pi = a(\pi)\lambda_{s_1}^{\rho_1} \cdots \lambda_{s_k}^{\rho_k}/m!, \tag{34}$$

terms of like degree in the exponential series may be collected:

$$F_n(x) = \Phi(x) - \phi(x) \textstyle\sum_\pi \lambda_\pi\{[\phi(x)]^{-1}(-D_x)^{m-1}\phi(x)\}, \tag{35}$$

where the summation is extended over all partitions π of all positive integers.

The functions

$$\psi_m(x) = [\phi(x)]^{-1}(-D_x)^{m-1}\phi(x) \tag{36}$$

satisfy the recurrence relation

$$\psi_m(x) = (\psi(x) - D_x)\psi_{m-1}(x), \tag{37}$$

where, (see also (6)),

$$\psi_1(x) \equiv 1, \qquad \psi_2(x) = \psi(x) = -\phi'(x)/\phi(x) = D_x \log (1/\phi(x)). \tag{38}$$

From (11) and (35) a series in terms of cumulants is obtained:

$$z_n(x) = - \textstyle\sum_\pi \lambda_\pi \psi_m(x) \tag{39}$$

and a similar expression may be sought for $(z_n(x))^r/r!$

For (non-empty) partitions $\pi_1, \cdots, \pi_r$ of positive integers $m_1, \cdots, m_r$ having as their union the partition π of the integer m,

$$\pi = [s_1^{\rho_1}, \cdots, s_k^{\rho_k}], \qquad \pi_j = [s_i^{\rho_{ij}}, \cdots, s_k^{\rho_{kj}}],$$
$$\rho_i = \textstyle\sum_{j=1}^r \rho_{ij}, \qquad (j = 1, \cdots, r); \tag{40}$$

we define the partition function

$$p(\pi_1, \cdots, \pi_r) = \textstyle\prod_{i=1}^k \binom{\rho_i}{\rho_{i1}, \cdots, \rho_{ir}}, \tag{41}$$

where $\binom{\rho_i}{\rho_{i1}, \cdots, \rho_{ir}}$ denotes the multinomial coefficient. Then using the definition (33) of $a(\pi)$, it is found that

$$(42) \qquad (z_n(x))^r/r! = (-1)^r \textstyle\sum_\pi \lambda_\pi \psi_\pi^{(r)}(x),$$

where

$$(43) \qquad \psi_\pi^{(r)} = (r!)^{-1} \textstyle\sum_{\pi_1+\cdots+\pi_r=\pi} p(\pi_1, \cdots, \pi_r)\psi_{m_1} \cdots \psi_{m_r}, \quad (\psi_\pi^{(1)} = \psi_m).$$

The summation is extended over all distinct *arrangements* of π_i's having union π; i.e. if the π_i's are identical in groups of sizes $\sigma_1, \cdots, \sigma_R$ then $p(\pi_1, \cdots, \pi_r)\psi_{m_1} \cdots \psi_{m_r}$ occurs $\binom{r}{\sigma_1, \cdots, \sigma_R}$ times. It follows also that

$$(44) \qquad \psi_\pi^{(r)} = \textstyle\sum_{\pi_1 \cup \cdots \cup \pi_r = \pi} q(\pi_1, \cdots, \pi_r)\psi_{m_1} \cdots \psi_{m_r},$$

where

$$(45) \qquad q(\pi_1, \cdots, \pi_r) = p(\pi_1, \cdots, \pi_r)/\sigma_1! \cdots \sigma_R!,$$

and the summation in (44) is extended over all distinct *combinations* of π_i's with union π.

Substituting (42) in (12), the *standardizing expansion* takes the form

$$(46) \qquad u = x + \textstyle\sum_\pi \lambda_\pi \sum_{r=1}^{l} (-1)^r c_r(x)\psi_\pi^{(r)}(x),$$

and similarly the *inverse expansion* (23) becomes

$$(47) \qquad x = u - \textstyle\sum_\pi \lambda_\pi \sum_{r=1}^{l} (-1)^r D_{(r)}\psi_\pi^{(r)}(x).$$

These expansions relate the quantiles u and x in terms of cumulant differences and functions derived by application of differential operators of the type $r\psi \pm D$, where ψ is determined by the frequency function of the limiting distribution.

5. The Cornish-Fisher expansions. In the case of Cornish-Fisher expansions, $\Phi(x)$ is the unit normal distribution and

$$(48) \qquad \lambda_2 = \kappa_2 - 1, \quad \lambda_i = \kappa_i \ (i \neq 2), \quad \psi(x) = x,$$
$$\psi_r(x) = h_r(x) = e^{x^2/2}(-D_x)^{r-1}e^{-x^2/2}, \quad (r = 1, 2, \cdots),$$

where h_r is Hermite's $(r - 1)$th polynomial. The Cornish-Fisher *normalizing expansion* may be obtained from (46):

$$(49) \qquad u = x + \textstyle\sum_\pi \lambda_\pi N_\pi(x),$$

where the polynomials N_π are defined by:

$$(50) \qquad N_\pi = \textstyle\sum_{r=1}^{l} (-1)^r c_r h_\pi^{(r)},$$

and the polynomials c_r are to be derived from equations (8) and (13). Similarly (47) becomes the Cornish-Fisher *inverse expansion*:

$$(51) \qquad x = u + \textstyle\sum_\pi \lambda_\pi P_\pi(u),$$

where the polynomials P_π are given by:

$$(52) \qquad P_\pi = \textstyle\sum_{r=1}^{l} (-1)^{r-1} D_{(r)} h_\pi^{(r)},$$
$$D_{(r)} \equiv (u - D)(2u - D) \cdots ((r-1)u - D).$$

The Cornish-Fisher assumption that

$$(53)\quad \lambda_1 = O(n^{-\frac{1}{2}}), \qquad \lambda_2 = O(n^{-1}), \qquad \lambda_r = O(n^{-r/2+1}), \quad (r = 3, 4, \cdots),$$

leads to a classification of the λ_π into successive "adjustments": if $\lambda_\pi = O(n^{-M/2})$ then λ_π (and hence N_π and P_π) belongs to the Mth adjustment. The Cornish-Fisher polynomials, whose numeric coefficients were determined in the case of leading terms in the formulae [3], [5], are seen to involve sums of products of Hermite polynomials. Indeed, all components, including the Hermite polynomials, in the general forms of the Cornish-Fisher polynomials can be derived from $c_1 = h_1 = 1$ by application of differential operators of the type $nx \pm D$.

6. Properties of Cornish-Fisher polynomials. In the case $\pi = [s^m]$ it is easily shown from (44) that

$$(54)\qquad h_{[s^m]}^{(r)} = \Sigma a(\pi)(h_{t_1 s})^{\rho_1} \cdots (h_{t_k s})^{\rho_k},$$

where the summation is extended over all partitions $\pi = [t_1^{\rho_1}, \cdots, t_k^{\rho_k}]$ of m into r parts. In particular,

$$(55)\qquad h_{[s^m]}^{(m)} = (h_s)^m, \qquad h_{[s^m]}^{(m-1)} = \binom{m}{2}(h_s)^{m-2}h_{2s}\,.$$

Next consider $h_{\pi,s}^{(r)}$, where π, s denotes the partition obtained by adjoining the singleton $[s]$ to the arbitrary partition π. From (43) it is seen that a given arrangement $(\pi_1, \cdots, \pi_r)$ is effectively given the multiplicity $p(\pi_1, \cdots, \pi_r)$ in the sum defining $h_\pi^{(r)}$. But (41) shows that $p(\pi_1, \cdots, \pi_r)$ is the number of ways of constructing this arrangement if each set of s_i's is considered as being composed of ρ_i distinct individuals. Hence,

$$(56)\quad h_{\pi,s}^{(r)} = h_s h_\pi^{(r-1)} + (r!)^{-1} \sum\nolimits_{\pi_1+\cdots+\pi_r=\pi} p(\pi_1, \cdots, \pi_r) \cdot [h_{m_1+s}h_{m_2} \cdots h_{m_r} + \cdots + h_{m_1} \cdots h_{m_{r-1}}h_{m_r+s}],$$

where the term $h_s h_\pi^{(r-1)}$ corresponds to the sum over all partitions of π, s containing s as a singleton and the second term corresponds to all other partitions.

The well known relations for Hermite polynomials:

$$(57)\qquad (x - D_x)h_s = h_{s+1}\,,$$

$$(58)\qquad h_{s+2} = xh_{s+1} - sh_s\,,$$

may now be generalized for the $h_\pi^{(r)}$ by means of (56):

$$(59)\qquad (rx - D_x)h_\pi^{(r)} = h_{\pi,1}^{(r)} - h_\pi^{(r-1)}, \qquad (h_\pi^{(0)} \equiv 0),$$

$$(60)\qquad h_{\pi,2}^{(r)} = xh_{\pi,1}^{(r)} - mh_\pi^{(r)}$$

for all partitions π of m.

By inspection of the expressions obtained by Cornish and Fisher for the leading adjustments in their expansions the polynomials may be observed to satisfy the following identities:

$$(61)\qquad -N_{[s]} = P_{[s]} = h_s\,,$$

for $[s]$ a singleton, while for all partitions π of m:

$$N_{\pi,1} = -D_x N_\pi ; \tag{62}$$

$$N_{\pi,2} = xN_{\pi,1} - mN_\pi = -x^{1-m}D_x(x^m N_\pi); \tag{63}$$

$$P_{\pi,1} \equiv 0; \tag{64}$$

$$P_{\pi,2} = -(m-1)P_\pi . \tag{65}$$

These identities follow from the expressions (50) and (52) of N and P in terms of symmetric sums of products of Hermite polynomials.

To prove (62), equation (59) and the defining relation (8) of the c_r may be used; clearly

$$\begin{aligned} D_x N_\pi &= \Sigma_r(-1)^r[-c_r(rx - D_x)h_\pi^{(r)} + h_\pi^{(r)}(rx + D_x)c_r] \\ &= \Sigma_r(-1)^r[-c_r h_{\pi,1}^{(r)} + c_r h_\pi^{(r-1)} + c_{r+1}h_\pi^{(r)}] \\ &= -N_{\pi,1} . \end{aligned} \tag{66}$$

Equation (63) is a trivial consequence of (60) and (62), while (64) follows immediately from (59) and (52). Lastly, in virtue of (59) and (60):

$$P_{\pi,2} = \Sigma_r(-1)^r[-D_{(r+1)}(xh_\pi^{(r)}) + D_{(r)}(xh_\pi^{(r-1)}) + (1-m)D_{(r)}h_\pi^{(r)}], \tag{67}$$

which implies (65).

The identities (62) and (64) reflect the fact that changes in the location parameter of x affect λ_1 only. Equivalent identities apply to symmetric sums associated with $\lambda_{\pi,1}$ in the general expansions (46) and (47), but the identities (63) and (65) for elements involving π, 2 arise from properties of Hermite polynomials and hold only in asymptotic expansions about normal Φ. In practice, terms involving λ_1 and λ_2 can be excluded by relating quantiles of u to quantiles of $x = (x' - \mu)/\sigma$, for which $\lambda_1 = \lambda_2 = 0$, and treating x as an intermediate variate, whose quantiles are linearly related to the quantiles of x'.

Since the Hermite polynomial h_s is an odd or even function according as $(s - 1)$ is odd or even, the parities of the polynomials $h_\pi^{(r)}$, P_π and N_π, where π is a partition of m, are those of the integers $(m - r)$, $(m - 1)$ and $(m - 1)$, respectively. The order of the adjustment to which λ_π belongs is clearly of the same parity as m. Hence, the polynomials P_π, N_π in odd order adjustments are even, whereas those in even order adjustments are odd.

These results are illustrated in the Table, which presents polynomials for the first three adjustments of the normalizing and inverse expansions. The first four adjustments listed by Cornish and Fisher ([3] pp. 316–317) and the first six adjustments of the inverse expansion listed by Fisher and Cornish [5] have been checked against the expressions presented in Section 5. Using these formulations, the first twelve normalizing and inverse adjustments have been tabulated [7] by means of a computer program, which used algorithms for generating the partitions (40) and partition functions (45) and for polynomial operations, including the application of the operator $nx \pm D$, arising in (8), (52) and (57).

TABLE 1

Cornish-Fisher Polynomials

M	π	λ_π	$h_\pi^{(1)}$	$h_\pi^{(2)}$	$h_\pi^{(3)}$	N_π	P_π
1	[1]	λ_1	1	0	0	-1	1
	[3]	$\lambda_3/6$	h_3	0	0	$-(x^2 - 1)$	$x^2 - 1$
2	[2]	$\lambda_2/2$	h_2	0	0	$-x$	x
	[4]	$\lambda_4/24$	h_4	0	0	$-(x^3 - 3x)$	$x^3 - 3x$
	[1^2]	$\lambda_1^2/2$	h_2	1	0	0	0
	[1, 3]	$\lambda_1\lambda_3/6$	h_4	$x^2 - 1$	0	$2x$	0
	[3^2]	$\lambda_3^2/72$	h_6	$(x^2 - 1)^2$	0	$2(4x^3 - 7x)$	$-2(2x^3 - 5x)$
3	[5]	$\lambda_5/120$	h_5	0	0	$-(x^4 - 6x^2 + 3)$	$x^4 - 6x^2 + 3$
	[1, 2]	$\lambda_1\lambda_2/2$	h_3	x	0	1	0
	[1, 4]	$\lambda_1\lambda_4/24$	h_5	$x^3 - 3x$	0	$3(x^2 - 1)$	0
	[2, 3]	$\lambda_2\lambda_3/12$	h_5	$x^3 - x$	0	$5x^2 - 3$	$-2(x^2 - 1)$
	[3, 4]	$\lambda_3\lambda_4/144$	h_6	$x^5 - 4x^3 + 3x$	0	$11x^4 - 42x^2 + 15$	$-2(x^4 - 5x^2 + 2)$
	[1^2, 3]	$\lambda_1^2\lambda_3/12$	h_5	$3x^3 - 7x$	$x^2 - 1$	-2	0
	[1, 3^2]	$\lambda_1\lambda_3^2/72$	h_7	$3x^5 - 18x^3 + 21x$	$(x^2 - 1)^2$	$-2(12x^2 - 7)$	0
	[3^3]	$\lambda_3^3/1296$	h_9	$3(x^7 - 11x^5 + 25x^3 - 15x)$	$(x^2 - 1)^3$	$-2(69x^4 - 187x^2 + 52)$	$4(12x^4 - 53x^2 + 17)$

Acknowledgment. The authors wish to express their gratitude to Professor A. T. James for his comments on an early draft of this paper.

REFERENCES

[1] ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.

[2] BOL'SHEV, L. N. (1959). O prebrazovaniyakh sluchainykh velichin (Transformation of random variables). C.S.I.R.O. translation No. 4695. *Teor. Verayatnost. i primen* **4** 137–149.

[3] CORNISH, E. A. and FISHER, R. A. (1937). Moments and cumulants in the specification of distributions. *Revue de l'Institut Internat. de Statist.* **4** 307–320.

[4] FISHER, R. A. (1941). The asymptotic approach to Behren's integral with further tables for the d test of significance. *Ann. Eugen.* **11** 141–172.

[5] FISHER, R. A. and CORNISH, E. A. (1960). The percentile points of distributions having known cumulants. *Technometrics* **2** 209–226.

[6] GOLDBERG, H. and LEVINE, HARRIET. (1946). Approximate formulas for the percentage points and normalization of t and χ^2. *Ann. Math. Statist.* **17** 216–225.

[7] HILL, G. W. (1964). Cornish-Fisher polynomials for asymptotic expansions to order n^{-6} of distributions specified by cumulants. Working paper No. 25, Graduate School of Business, Stanford Univ. 72pp.

[8] ITO, K. (1956). Asymptotic formulae for the distribution of Hotelling's generalized T_0^2 statistic. *Ann. Math. Statist.* **27** 1091–1105.

[9] JOHNSON, N. L. and WELCH, B. L. (1940). Applications of the non-central t-distribution. *Biometrika* **31** 362–389.

[10] RAO, C. R. (1948). Tests of significance in multivariate analysis. *Biometrika* **35** 58–79.

[11] RIORDAN, J. (1949). Inversion formulas in normal variable mapping. *Ann. Math. Statist.* **20** 417–425.

[12] SCHATZOFF, MARTIN. (1966). Exact distributions of Wilks's likelihood ratio criterion. *Biometrika* **53** 347–358.

[13] TEICHROEW, D. (1953). Distribution sampling with high speed computers. Ph.D. Thesis. Univ. of North Carolina.

[14] WALLACE, D. L. (1958). Asymptotic approximations to distributions. *Ann. Math. Statist.* **29** 635–654.

[15] WHITTAKER, E. T. and WATSON, G. N. *A Course of Modern Analysis* (3rd edition). Cambridge Univ. Press.

Biometrika (1970), **57**, 2, *p.* 457
Printed in Great Britain

Percentile approximations for ordered *F* ratios

By A. W. DAVIS
C.S.I.R.O., Adelaide and University of North Carolina, Chapel Hill

Summary

A direct percentile approximation is given for the rth largest in a set of k F ratios.

1. Introduction

Table 19 of *Biometrika Tables for Statisticians* gives percentage points for the largest of k F ratios in an analysis of variance, each based on $m = 1$ and ν degrees of freedom. This table is due to Nair (1948), and was extended by Chambers (1967). These authors used Hartley's (1944) expansion of the studentized integral. In the present context, let ${}_\nu P_r(Q)$ denote the distribution function of the square root of the rth largest member F_r of a set of k variance ratios, each based on m and ν degrees of freedom, and let $P_r(Q)$ be the limiting distribution as $\nu \to \infty$. Then

$$P'_r(Q) = \frac{k!}{(r-1)!\,(k-r)!}\{\Phi_m(Q)\}^{k-r}\{1-\Phi_m(Q)\}^{r-1}\,\Phi'_m(Q), \tag{1}$$

where

$$\Phi_m(Q) = 2\Gamma(\tfrac{1}{2}m)^{-1}(\tfrac{1}{2}m)^{\frac{1}{2}m}\int_0^Q x^{m-1}e^{-\frac{1}{2}mx^2}\,dx, \tag{2}$$

$${}_\nu P_r(Q) \sim P_r(Q) + a_1/\nu + a_2/\nu^2 + a_3/\nu^3 + \dots. \tag{3}$$

Nair (1948, equation (7)) has given a_1–a_4 in terms of the derivatives of $P_r(Q)$.

2. The approximation

The expansion (3) may be formally inverted by the method of Hill & Davis (1968, Section 3) to yield a direct expansion of 100δ % points $\sqrt{F_{r,\delta}}$ in terms of the 100δ % points Q_δ of $P_r(Q)$. Thus, writing

$${}_\nu P_r(Q) \sim P_r(Q) + P'_r(Q)\,{}_\nu z_r(Q),$$

$$\psi_r(Q) = P''_r(Q)/P'_r(Q), \tag{4}$$

and introducing the sequence of differential operators

$$D_{(n+1)} \equiv D_{(n)}\left\{\frac{d}{dQ} + n\psi_r(Q)\right\} \quad (n = 1, 2, \dots), \tag{5}$$

where $D_{(1)}$ denotes the identity operator, we have

$$\sqrt{F_{r,\delta}} \sim Q_\delta + \left[\sum_{n-1}^{\infty} D_{(n)}\{-{}_\nu z_r(Q)\}^n/n!\right]_{Q=Q_\delta}. \tag{6}$$

In reducing (6), we note that the ratios $P_r^{(s)}/P'_r$ in the asymptotic expansion of ${}_\nu z_r$ may be expressed in terms of the polynomials

$$h_n(Q) = \Phi_m^{(n)}(Q)/\Phi'_m(Q) \quad (n = 2, 3, \dots), \tag{7}$$

and the functions

$$\theta_j(Q) = \sum_{i=0}^{j}\binom{j}{i}(-1)^i(r-1)_i(k-r)_{j-i}\{\phi_1(Q)\}^i\{\phi(Q)\}^{j-i}, \tag{8}$$

where

$$\phi(Q) = \Phi'_m(Q)/\Phi_m(Q), \quad \phi_1(Q) = \Phi'_m(Q)/\{1-\Phi_m(Q)\}, \quad (k)_i = k(k-1)\dots(k-i+1). \tag{9}$$

Thus

$$P''_r/P'_r = \psi_r = h_2 + \theta_1, \quad P'''_r/P'_r = h_3 + 3\theta_1 h_2 + \theta_2, \quad \text{etc.} \tag{10}$$

The h_r may be conveniently written in terms of

$$u = mQ^2 - m, \tag{11}$$

and in particular,

$$\left.\begin{aligned} h_2 &= -Q^{-1}(u+1),\\ h_3 &= Q^{-2}\{u^2+u-2(m-1)\},\\ h_4 &= Q^{-3}\{-u^3+(6m-3)\,u+(8m-6)\}.\end{aligned}\right\} \tag{12}$$

Using the differentiation formula

$$\frac{d\theta_j}{dQ} = j\theta_j h_2 + \theta_{j+1} - \theta_1\theta_j, \tag{13}$$

and introducing also $\lambda_j = Q^j\theta_j$, we obtain to order ν^{-3}

$$\sqrt{F_{r,\delta}}/Q_\delta \sim 1+\alpha_1/\nu+\alpha_2/\nu^2+\alpha_3/\nu^3+O(\nu^{-4}), \tag{14}$$

where

$$\begin{aligned}\alpha_1 &= \tfrac{1}{4}(u+2-\lambda_1),\\ \alpha_2 &= \frac{1}{6(4)^2}[5u^2-(6m-32)\,u-4(m-7)-6\lambda_1\{u^2+3u-2(m-3)\}+4\lambda_2(3u+2)-3\lambda_1^2(3u-1)\\ &\quad -3(\lambda_3-2\lambda_1\lambda_2+\lambda_1^3)],\\ \alpha_3 &= \frac{1}{6(4)^3}[3u^3-(10m-38)\,u^2+(12m^2-64m+116)\,u+8(m^2-3m+5)-\lambda_1\{4u^4+2u^3-(52m-77)\,u^2\\ &\quad -(78m-152)\,u+(52m^2-148m+140)\}+4\lambda_2\{8u^3+9u^2-u(30m-36)-20(m-1)\}\\ &\quad -3\lambda_1^2\{6u^3-u^2-(26m-21)\,u-10\}-\lambda_3(38u^2+41u-34m+46)+2\lambda_1\lambda_2(30u^2+7u-30m+26)\\ &\quad -\lambda_1^3(26u^2-10u-28m+31)+4(\lambda_4-\lambda_2^2)(3u+2)-\lambda_1\lambda_3(27u+5)+2\lambda_1^2\lambda_2(21u+1)-3\lambda_1^4(5u-1)\\ &\quad +(-\lambda_5+3\lambda_1\lambda_4+3\lambda_2\lambda_3-6\lambda_1^2\lambda_3+6\lambda_1\lambda_2^2+10\lambda_1^3\lambda_2-3\lambda_1^5)].\end{aligned} \tag{15}$$

The term α_4 has also been obtained. The first step in using the formula is to solve the equation $P_r(Q) = \delta$, regarded as a polynomial in Φ_m. Obtaining $\Phi_m(Q_\delta) = \epsilon$, say, mQ_δ^2 is seen to be the 100ϵ % point of χ_m^2, whence u and the λ_j may be calculated. Given the α_j for a set of k, m and r, the $F_{r,\delta}$ may be approximated directly without further interpolation.

3. Applications

Tables 1 and 2 compare results obtained using the above approximation to terms of order ν^{-3} and ν^{-4} with upper 5 % points tabulated by Chambers in the case $m = r = 1$, and with Finney's (1941) exact values in the case $m = 2, r = 1$. There is good agreement with Finney's values when ν is sufficiently large for the $O(\nu^{-3})$ and $O(\nu^{-4})$ approximations to agree in the second or third decimal place. Inspection of Table 1 of the present note and Table 3 of Chambers (1967) indicates that the direct percentile expansion converges more rapidly than the interpolation procedure based on (3); Chambers's method also apparently tends to underestimate the correct values. Table 3 of this paper presents in addition some results for $m = 1, r = 2$, the $O(\nu^{-3})$ and $O(\nu^{-4})$ approximations agreeing to the number of decimals quoted.

The author's thanks are due to Miss J. Hawkes for her programming assistance.

Table 1. *Upper* 5 % *points of the largest* F: $m = 1$

	$k = 8$			$k = 20$		
ν	Chambers	$O(\nu^{-3})$	$O(\nu^{-4})$	Chambers	$O(\nu^{-3})$	$O(\nu^{-4})$
10	10·85	11·20	11·23	14·65	14·47	14·63
15	9·72	9·77	9·77	12·32	12·47	12·50
20	9·11	9·119	9·121	11·49	11·545	11·554
30	8·51	8·516	8·517	10·66	10·672	10·674
60	7·96	7·956	7·956	9·85	9·855	9·855
∞	7·44	7·437	7·437	9·10	9·096	9·096

Table 2. *Upper 5 % points of the largest F: m = 2*

ν	$k = 2$ Finney	$O(\nu^{-3})$	$O(\nu^{-4})$	$k = 3$ Finney	$O(\nu^{-3})$	$O(\nu^{-4})$
6	6·90	6·874	6·906	8·03	7·970	8·074
8	5·86	5·852	5·861	6·75	6·733	6·764
10	5·32	5·319	5·323	6·09	6·085	6·097
20	4·41	4·410	4·410	4·98	4·975	4·975
∞	3·68	3·676	3·676	4·08	4·077	4·077

Table 3. *Upper 5 % points of the second largest F: m = 1*

ν \ k	2	3	4	5	10	20
10	1·886	3·007	3·842	4·517	—	—
15	1·735	2·713	3·423	3·990	5·86	—
20	1·665	2·579	3·235	3·754	5·444	—
30	1·600	2·455	3·060	3·534	5·057	6·67
60	1·539	2·339	2·897	3·330	4·697	6·112
∞	1·481	2·230	2·744	3·139	4·361	5·590

References

Chambers, C. (1967). Extension of tables of percentage points of the largest variance ratio, s^2_{max}/s^2_0. *Biometrika* **54**, 225–7.

Finney, D. J. (1941). The joint distribution of variance ratios based on a common error mean square. *Ann. Eugen.* **11**, 136–40.

Hartley, H. O. (1944). Studentization or the elimination of the standard deviation of the parent population from the random sample—distribution of statistics. *Biometrika* **33**, 173–80.

Hill, G. W. & Davis, A. W. (1968). Generalized asymptotic expansions of Cornish–Fisher type. *Ann. Math. Statist.* **39**, 1264–73.

Nair, K. R. (1948). The studentized form of the extreme mean square test in the analysis of variance. *Biometrika* **35**, 16–31.

[*Received September* 1969. *Revised December* 1969]

Biometrika (1971), 58, 2, p. 349
Printed in Great Britain

Percentile approximations for a class of likelihood ratio criteria

By A. W. DAVIS
C.S.I.R.O., Adelaide, and University of North Carolina, Chapel Hill

Summary

The general asymptotic series developed by Box (1949) for the distributions of a large class of likelihood ratio criteria has been widely used to obtain good approximations to these distributions. A direct percentile approximation based upon this expansion, bearing a relationship to it similar to that of the Cornish–Fisher expansion to Edgeworth's series is given.

1. Introduction

The asymptotic expansions of Cornish & Fisher (1937) and Fisher & Cornish (1960) have become well known as a tool for obtaining accurate percentage points of distributions which approach normality in large samples, and are adequately specified by their early cumulants. Essentially, these expansions are obtained by inverting the Edgeworth series for the distributions in question. A reasonably concise formulation of the Cornish–Fisher normalizing and inverse expansions has been given by Hill & Davis (1968), together with procedures for calculating the terms of these series using Hermite polynomials. Nothing appears to have been published concerning the validity of the Cornish–Fisher expansions as asymptotic series in the strict sense, although according to Wallace (1958) this may be established for absolutely continuous distributions whose Edgeworth series are valid.

The likelihood ratio criteria of Neyman & Pearson (1928*a*, *b*) form another important class of statistics whose distributions may be specified by their moments. Exact null distributions have recently been given for a number of these tests, Consul (1969) presenting a general formulation in terms of Meijer's G function. Schatzoff (1966) and Pillai & Gupta (1969) have tabulated the Wilks likelihood ratio criterion using its exact distribution. However, accurate percentage points are generally difficult to obtain. In large samples, the distributions of these statistics are asymptotically chi-squared, and Box (1949) has given a method of asymptotic expansion in cases where the moments are known and of a certain structure. His approach has been applied extensively, and it seems worthwhile to present a Cornish–Fisher inversion of Box's series, expressing the percentage points of a given distribution in terms of chi-squared percentiles. The method has the advantage of providing a direct polynomial approximation, not requiring the calculation of chi-squared probabilities and inverse interpolation as in applications of Box's expansion. Actually, it is convenient to take as a starting point a general expansion of the cumulant generating function which holds for a larger class of statistics than that explicitly considered by Box. This class includes the Lawley–Hotelling T_0^2 and Pillai's V, proposed as tests of linear hypotheses in multivariate analysis of variance. The inversion is effected using an algorithm given by Hill & Davis (1968). Special cases were considered in the latter paper, while Sugiura & Nagao (1969) and Sugiura in an unpublished report have applied the algorithm in calculations of the powers of certain criteria. No attempt is made in the present paper to establish the asymptotic validity of the percentile expansion. However, numerical applications in §4 show that it can be used to obtain accurate approximations, even for comparatively small sample sizes.

2. The percentile expansion

We consider an absolutely continuous distribution function $F(x)$ whose cumulant generating function

$$K(\theta) = \log \int_{-\infty}^{\infty} e^{i\theta x} F'(x)\,dx \tag{1}$$

may be validly represented by an asymptotic series

$$K(\theta) \sim -\tfrac{1}{2} f \log(1-2i\theta) + \sum_{r=1}^{\infty} \omega_r \{(1-2i\theta)^{-r} - 1\}, \tag{2}$$

corresponding to a limiting chi-squared distribution with f degrees of freedom. The characteristic function $C(\theta) = \exp\{K(\theta)\}$ then has the formal expansion

$$C(\theta) \sim (1-2i\theta)^{-\frac{1}{2}f} \left[1 + \sum_{\pi} \omega_\pi \prod_{j=1}^{k} \{(1-2i\theta)^{-s_j} - 1\}^{\nu_j}\right], \tag{3}$$

where the summation is extended over the partitions $\pi = \{s_1^{\nu_1}, \ldots, s_k^{\nu_k}\}$ of all positive integers $n(s_1 > \ldots > s_k;\ \nu_1 s_1 + \ldots + \nu_k s_k = n)$ and

$$\omega_\pi = \omega_{s_1}^{\nu_1} \ldots \omega_{s_k}^{\nu_k} / (\nu_1! \ldots \nu_k!). \tag{4}$$

On inverting $C(\theta)$, it follows that

$$F(x) \sim \Phi_f(x) - 2\phi_f(x) \sum_{\pi} \omega_\pi h_\pi(x), \tag{5}$$

where the $h_\pi(x)$ are polynomials and Φ_f is the cumulative distribution function of χ_f^2 with density function ϕ_f,

$$\phi_f(x) = \{2\Gamma(\tfrac{1}{2}f)\}^{-1} e^{-\frac{1}{2}x} (\tfrac{1}{2}x)^{\frac{1}{2}f-1}, \quad \Phi_f(x) = \int_0^x \phi_f(y)\,dy \quad (x > 0). \tag{6}$$

The h_π can be readily constructed from the basic set $h_{[n]}(x)$ $(n = 1, 2, \ldots)$. Denoting the latter simply by $h_n(x)$, we have

$$h_n(x) = \{\Phi_f(x) - \Phi_{f+2n}(x)\}/\{2\phi_f(x)\} = \sum_{r=1}^{n} x^r / f_r, \tag{7}$$

where $f_r = f(f+2)\ldots(f+2r-2)$. In particular,

$$h_{[p,q]}(x) = h_{p+q} - (h_p + h_q), \quad h_{[p,q,r]}(x) = h_{p+q+r} - (h_{p+q} + h_{q+r} + h_{r+p}) + (h_p + h_q + h_r), \tag{8}$$

and the extension to higher order π is obvious.

In the examples considered by Box, $\omega_r = O(m^{-r})$ $(r = 1, 2, \ldots)$, where m depends on the sample size. It follows that if π is a partition of n then ω_π is $O(m^{-n})$, and Box showed that valid expansions of the $F(x)$ in his examples are obtained by grouping the terms of the same order in m.

The expansion (5) may now be formally inverted to express an arbitrary $100(1-\alpha)\%$ point x_α of F in terms of the corresponding chi-squared percentile $u = \chi^2_{f,\alpha}$. Let

$$\psi(u) = -2\phi_f'(u)/\phi_f(u) = 1 - (f-2)/u, \tag{9}$$

and define the sequence of differential operators $D_{(r)}$ by

$$D_1 = 1, \quad D_r = 2\frac{d}{du} - (r-1)\psi(u) \quad (r = 2, 3, \ldots), \quad D_{(r)} = D_1 D_2 \ldots D_r \quad (r = 1, 2, \ldots). \tag{10}$$

Then (Hill & Davis, 1968, § 3),

$$x_\alpha \sim u + 2\sum_{r=1}^{\infty} D_{(r)}\{z(u)\}^r/r!, \tag{11}$$

where $z(u) = \Sigma\omega_\pi h_\pi(u)$. On formally expanding $\{z(u)\}^r$, the products of ω_π's can again be expressed as ω_π's, and we may write

$$\{z(u)\}^r/r! = \sum_\pi \omega_\pi h_\pi^{(r)}(u) \quad (r = 1, 2, \ldots;\ h_\pi^{(1)} \equiv h_\pi), \tag{12}$$

noting that $h_\pi^{(r)} = 0$ if $r > l_\pi = \nu_1 + \ldots + \nu_k$, the total number of parts in π. Thus (11) takes the form

$$x_\alpha \sim u + 2\sum_\pi \omega_\pi P_\pi(u), \tag{13}$$

where the polynomials P_π are given by

$$P_\pi(u) = \sum_{r=1}^{l_\pi} D_{(r)} h_\pi^{(r)}(u). \tag{14}$$

In particular, $P_{[n]} = h_n$ $(n = 1, 2, \ldots)$. The construction of the $h_\pi^{(r)}$ and P_π is illustrated in the following section. It should be noted that the P_π in the percentile expansion (13) depend only on f, the degrees of freedom of the limiting chi-squared distribution, and are otherwise independent of the particular distribution $F(x)$.

3. Construction of the polynomials

The derivation of (13) from (2) is analogous to the construction of the Cornish–Fisher expansion from Edgeworth's series as presented by Hill & Davis (1968). There the Hermite polynomials played the role of the h_π in the present context. It follows that the $h_\pi^{(r)}$ may be computed using the method of the earlier paper, and this is perhaps best conveyed by an example. The first step is to find all (unordered) sets of r partitions having union π. Thus, if $\pi = [2^2, 1^2] = \pi_0$, say, and $r = 2$, we have the possibilities $[2^2]$, $[1^2]$, and $[2, 1]$, $[2,1]$.

In this case

$$h_{\pi_0}^{(2)} = \binom{2}{2, 0}\binom{2}{0, 2} h_{[2^2]}h_{[1^2]} + \binom{2}{1, 1}\binom{2}{1, 1}(h_{[2,1]})^2/2!,$$

where the multinomial coefficients in each term are constructed from the multiplicities of 2 and 1 in π_0 and the corresponding subpartitions; the divisor 2! in the second term arises because [2,1] occurs twice in the second subpartitioning. By (7) and (8),

$$h_{\pi_0}^{(2)}(u) = u^6(3f+16)/(f_3f_4) + 4u^5/(f_2f_4) - u^4(6f+14)/(f_2f_3) - 2u^3/(ff_2) + 3u^2/f^2.$$

Similarly, if $r = 4$ the only subpartitioning is [2], [2], [1], [1], and

$$h_{\pi_0}^{(4)}(u) = \binom{2}{1, 1, 0, 0}\binom{2}{0, 0, 1, 1} h_2^2h_1^2/2!\,2! = \{u^3/(ff_2) + u_2/f^2\}^2.$$

If the polynomials for $r = 1, 3$ are also calculated, we obtain

$$\begin{aligned} P_{\pi_0} &= h_{\pi_0} + D_2[h_{\pi_0}^{(2)} + D_3\{h_{\pi_0}^{(3)} + D_4 h_{\pi_0}^{(4)}\}] \\ &= -128u^6(14f^3 + 165f^2 + 622f + 720)/(ff_2f_3f_6) \\ &\quad + 64u^5(32f^3 + 387f^2 + 1480f + 1728)/(ff_2f_3f_5) \\ &\quad + 16u^4(13f^2 - 40f - 56)/(ff_2^2f_4) - 16u^3(29f + 66)/(f^2f_2f_3) - 8u^2/(f^3f_2) + 8u/f^4. \end{aligned}$$

It was shown by Box that in many applications ω_1 may be reduced to zero, thus eliminating all partitions containing 1, and greatly increasing the rate of convergence of the series. Accordingly, Table 1 presents the P_π for those partitions of integers 2 through 7 which do not contain unity. Results for partitions of 8 are also available. In certain applications the partitions involving 1 are required, and the complete set has been tabulated to order 6.

Table 1. *Polynomials in the percentile expansion* (13): $x_\alpha \sim u + 2\sum_\pi \omega_\pi P_\pi(u)$ $(u = \chi^2_{f,\alpha})$

π	ω_π	$P_\pi(u)$
n	ω_n	$u/f_1 + \ldots + u^n/f_n$
2^2	$\frac{1}{2}\omega_2^2$	$-8u^4(f+3)/(f_2 f_4) + 8u^3/(f_2 f_3) + 6u^2/(ff_2) + 2u/f^2$
$3, 2$	$\omega_3\omega_2$	$-12u^5(f+4)/(f_2 f_5) - 2u^4(f-6)/(f_2 f_4) + 2u^3(3f+10)/(f_2 f_3)$ $+ 6u^2/(ff_2) + 2u/f^2$
$4, 2$	$\omega_4\omega_2$	$-16u^6(f+5)/(f_2 f_6) - 4u^5(f-4)/(f_2 f_5) + 2u^4(3f+14)/(f_2 f_4)$ $+ 2u^3(3f+10)/(f_2 f_3) + 6u^2/(ff_2) + 2u/f^2$
3^2	$\frac{1}{2}\omega_3^2$	$-6u^6(3f^2+30f+80)/(f_3 f_6) - 6u^5(f^2+2f-16)/(f_3 f_5) + 4u^4(f+12)/(f_2 f_4)$ $+ 4u^3(3f+8)/(f_2 f_3) + 6u^2/(ff_2) + 2u/f^2$
2^3	$\frac{1}{6}\omega_2^3$	$32u^6(7f^2+62f+120)/(f_2^2 f_6) - 32u^5(2f^2+37f+96)/(f_2^2 f_5) - 8u^4$ $\times (23f^2+124f+132)/(f_2^2 f_4) - 8u^3(f-10)/(ff_2 f_3) + 28u^2/(f^2 f_2) + 4u/f^3$
$5, 2$	$\omega_5\omega_2$	$-20u^7(f+6)/(f_2 f_7) - 2u^6(3f-10)/f_2 f_6 + 2\sum_{r=2}^{5} u^r(3f+4r-2)/(f_2 f_r) + 2u/f^2$
$4, 3$	$\omega_4\omega_3$	$-24u^7(f^2+12f+40)/(f_3 f_7) - 2u^6(5f^2+18f-80)/(f_3 f_6)$ $+ 2u^5(f^2+42f+176)/(f_3 f_5) + 4u^4(3f+16)/(f_2 f_4) + 4u^3(3f+8)/(f_2 f_3)$ $+ 6u^2/(ff_2) + 2u/f^2$
$3, 2^2$	$\frac{1}{2}\omega_3\omega_2^2$	$192u^7(2f^3+31f^2+154f+240)/(f_2 f_3 f_7)$ $-16u^6(4f^3+153f^2+1106f+2160)/(f_2 f_3 f_6) - 8u^5(35f^3$ $+420f^2+1540f+1632)/(f_2 f_3 f_5) - 4u^4(25f^2+80f+12)/(f_2^2 f_4)$ $+ 4u^3(7f+38)/(ff_2 f_3) + 28u^2/(f^2 f_3) + 4u/f^3$

4. APPLICATIONS

Once the polynomials P_π have been programmed for the computer, it remains to specify the ω's for the distribution considered. Box (1949) gives methods which are sufficient to obtain the ω_r in a number of instances; see also Anderson (1958, Chapters 8–10) for a presentation of the basic likelihood ratio criteria in multivariate analysis.

Wilks's test of independence (Anderson, 1958, Chapter 9; Consul, 1969, § 3). Given N observations from a p variate normal population, suppose that the variates are partitioned into k groups of sizes p_i $(i = 1, \ldots, k; \Sigma p_i = p)$, and it is required to test the independence of the groups. Box gives detailed results for ω_1 to ω_6 for the likelihood ratio criterion, thus carrying the expansion to terms of order N^{-6}. In the important case $k = 2$, the distribution of the statistic is identical with that of Wilks's test $U_{p_1, p_2, N-p_2-1}$ for linear hypotheses about regression coefficients. Exact percentage points for the latter have been tabulated by Schatzoff (1966) and Pillai & Gupta (1969). Rao (1948) and Box have derived the first six ω's explicitly in this case, and we also have

$$\omega_7 = 0,$$

$$\omega_8 = (184,320m^8)^{-1} p_1 p_2 [5(p_1^8+p_2^8) - 300(p_1^6+p_2^6) + 6,678(p_1^4+p_2^4) - 59,900(p_1^2+p_2^2)$$
$$+ \{60(p_1^4+p_2^4) + 126p_1^2 p_2^2 - 2,100(p_1^2+p_2^2) + 22,260\} p_1^2 p_2^2 + 151,053], \qquad (15)$$

where $m = N - \frac{1}{2}(p_1+p_2+3)$. The $O(m^{-8})$ percentile expansion gives results differing from the exact values by at most a unit in the third decimal place for $M = N - p_1 - p_2$ ranging from as low as $M = 3$ for $p_1 = 3$, $p_2 = 4$ to $M = 7$ or 8 for the largest p_1 and p_2 tabulated.

Also, the present method does not require that p_1 or p_2 should be even. Some tabulation for $k > 2$ has been made in an unpublished report of A. W. Davis and J. B. F. Field.

The generalized test of homoscedasticity (Anderson, 1958, Chapter 9). Let k independent samples be drawn from k p variate normal populations, $\mathbf{S}_i$ being the ith sample covariance matrix on N_i degrees of freedom ($i = 1, \ldots, k$), and let $\mathbf{S} = \Sigma N_i \mathbf{S}_i / N$, with $N = \Sigma N_i$, be the pooled covariance matrix. Box (1949, § 2) considers the multivariate generalization

$$M^* = N \log |\mathbf{S}| - \sum_{i=1}^{k} N_i \log |\mathbf{S}_i|$$

of Bartlett's test to investigate the equality of the k population covariance matrices, and shows how to construct the ω's. Korin (1969) has used Box's series to compute percentage points of M^* for equal sample sizes $N_i = N_0$ ($i = 1, \ldots, k$). Table 2 compares some of these results with values obtained from series (13), the $O(N_0^{-7})$ and $O(N_0^{-8})$ expansions agreeing to the number of decimal places quoted. The direct approximation appears to have some advantages, both in accuracy and the range of parameter values for which useful results can be obtained.

Table 2. *Upper 5 % points of the test M^* for homoscedasticity, $k = 10$ samples*

p	2		4		7
N_0	Korin	Series (13)	Korin	Series (13)	Series (13)
2	—	71	—	—	—
3	55·87	55·863	—	—	—
4	50·64	50·6378	—	—	—
5	48·02	48·0194	—	183	—
6	46·45	46·4454	—	162·33	—
8	44·65	44·6431	144·6	144·484	—
10	43·64	43·6398	136·3	136·2369	425
12	—	—	131·5	131·4506	389·2
15	—	—	—	127·1216	361·34
$\infty(\chi^2_f)$	—	40·1133	—	113·1453	290·0285

Mauchly's sphericity test (Anderson, 1958, Chapter 10; Consul, 1969). In testing the hypothesis that a sample of size N is drawn from a p variate normal population whose covariance matrix is proportional to a given matrix $\mathbf{\Sigma}_0$, the likelihood ratio criterion is a power of

$$W = |\mathbf{S}\mathbf{\Sigma}_0^{-1}|^{\frac{1}{2}n} \{\operatorname{tr}(\mathbf{S}\mathbf{\Sigma}_0^{-1})/p\}^{-\frac{1}{2}pn}, \tag{16}$$

where $n = N-1$ and $\mathbf{S}$ is the sample covariance matrix. Defining

$$f = \tfrac{1}{2}(p-1)(p+2), \quad \rho = 1-(2p^2+p+2)/(6pn), \tag{17}$$

we find that $-2\rho \log W$ is asymptotically distributed as χ^2_f, and its cumulant generating function has an expansion of the form (2). In this case $\omega_1 = 0$, and for $r \geqslant 2$,

$$\omega_r = \frac{2(-1)^r}{r(r+1)(r+2)\rho^r} \sum_{s=1}^{r+1} \binom{r+2}{s+1} (1-\rho)^{r+1-s} \{\delta_s + \tfrac{1}{2}(s+1) B_s/p^{s-1}\}/(\tfrac{1}{2}n)^{s-1}, \tag{18}$$

where the B_s are Bernoulli's numbers, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_3 = 0$, ..., and the δ_s are certain polynomials in p defined by Box (1949). Table 3 was constructed by comparing the $O(n^{-7})$ and $O(n^{-8})$ expansions, and is merely intended to indicate their range of applicability. The entries are correction factors x_α/u.

Table 3. *Correction factors for Mauchly's sphericity test* $-2\rho \log W$

$n \backslash p$	3		6		10	
	5 %	1 %	5 %	1 %	5 %	1 %
4	1·074	1·091	—	—	—	—
5	1·0376	1·0463	—	—	—	—
6	1·0228	1·0280	—	—	—	—
7	1·0153	1·0188	—	—	—	—
8	1·0110	1·0135	1·105	1·12	—	—
10	1·0064	1·0079	1·0508	1·058	—	—
12	1·0042	1·0052	1·0303	1·0343	1·15	—
15	1·0026	1·0031	1·0169	1·0191	1·067	1·073
20	1·0014	1·0017	1·0084	1·0095	1·0287	1·0311
χ_f^2	11·0705	15·0863	31·4104	37·5662	72·1532	81·0687

Test of a given covariance matrix. The hypothesis that a sample of N observations is drawn from a p variate population with a given covariance matrix $\mathbf{\Sigma}_0$ may be tested using the following modification of the likelihood ratio statistic (Korin, 1968):

$$\lambda_1 = e^{\frac{1}{2}pn}\,|\mathbf{S}\mathbf{\Sigma}_0^{-1}|^{\frac{1}{2}n} \exp\{-\tfrac{1}{2}n \operatorname{tr}(\mathbf{S}\mathbf{\Sigma}_0^{-1})\} \quad (n = N-1). \tag{19}$$

The moments of λ_1 are not of the form considered by Box, but the cumulant generating function of $-2\rho \log \lambda_1$ has an expansion (2), with

$$f = \tfrac{1}{2}p(p+1), \quad \rho = 1-(2p^2+3p-1)/\{6n(p+1)\}, \tag{20}$$

$$\omega_r = \frac{2(-1)^r}{r(r+1)(r+2)\rho^r} \sum_{s=1}^{r+1} \binom{r+2}{s+1} (1-\rho)^{r+1-s}\, \delta_s/(\tfrac{1}{2}n)^{s-1} \quad (\omega_1 = 0). \tag{21}$$

Table 4 compares values obtained from the percentile expansion with Korin's (1968) results. The 5 % points given by Korin for $p = 4$ should apparently be given with their degrees of freedom reduced by one.

Table 4. *Upper percentage points of* $-2 \log \lambda_1$ *for testing* $\mathbf{\Sigma} = \mathbf{\Sigma}_0$

$n \backslash p$	6				10			
	5 %		1 %		5 %		1 %	
	Korin	Series (13)	Korin	Series (13)	Korin	Series (13)	Korin	Series (13)
10	—	43·6	—	52·3	—	—	—	—
12	—	40·919	—	48·96	—	—	—	—
15	—	38·7138	—	46·234	—	101·8	—	114·9
20	36·87	36·8638	43·99	43·9754	—	91·28	—	102·70
25	35·89	35·8847	42·80	42·7890	—	86·516	—	97·234
30	35·28	35·2774	42·07	42·0559	—	83·7679	—	94·1027
40	—	—	—	—	80·7	80·7067	90·7	90·6287
50	—	—	—	—	79·13	79·0361	88·83	88·7388
χ_f^2	—	32·6706	—	38·9321	—	73·3115	—	82·2921

Test of a given mean and covariance matrix. If it is required to test not only that $\mathbf{\Sigma} = \mathbf{\Sigma}_0$, but also that the population mean vector $\boldsymbol{\mu} = \boldsymbol{\mu}_0$, then a suitable criterion is

$$\lambda_2 = (ne/N)^{\frac{1}{2}pN}\,|\mathbf{S}\mathbf{\Sigma}_0^{-1}|^{\frac{1}{2}N} \exp\{-\tfrac{1}{2}n \operatorname{tr}(\mathbf{S}\mathbf{\Sigma}_0^{-1}) + N(\bar{\mathbf{x}}-\boldsymbol{\mu})'\,\mathbf{\Sigma}_0^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)\}, \tag{22}$$

where $\bar{\mathbf{x}}$ is the sample mean vector (Anderson, 1958, §10.9). In this case $-2\rho\log\lambda_2$ is asymptotically χ_f^2 with

$$f = \tfrac{1}{2}p(p+3), \quad \rho = 1-(2p^2+9p+11)/\{6N(p+3)\}, \tag{23}$$

while an expansion (2) exists, with $\omega_1 = 0$ and

$$\omega_r = \frac{2(-1)^r}{r(r+1)(r+2)\rho^r}\sum_{s=1}^{r+1}\binom{r+2}{s+1}(1-\rho-N^{-1})^{r+1-s}\{\delta_s+p(s+1)/2^{s+1}\}/(\tfrac{1}{2}N)^{s-1}. \tag{24}$$

Some percentage points are given in the unpublished report of Davis and Field mentioned earlier.

The Lawley–Hotelling and Pillai traces. The Lawley–Hotelling generalized T_0^2 and Pillai's (1955) V statistic, defined respectively by

$$T_0^2 = n\,\mathrm{tr}\,(\mathbf{AB}^{-1}), \quad V = n\,\mathrm{tr}\,\{\mathbf{A}(\mathbf{A}+\mathbf{B})^{-1}\},$$

have been suggested as alternatives to Wilks's criterion for testing multivariate linear hypotheses. Here $\mathbf{A}$ and $\mathbf{B}$ are independent $p\times p$ Wishart matrices on q and n degrees of freedom respectively. For discussions of the exact distribution of T_0^2 and V, see Constantine (1966) and Davis (1968, 1970*a*, *b*, *c*). As $n\to\infty$, both criteria are asymptotically distributed as χ_{pq}^2, and their cumulant generating functions have expansions (2). Muirhead (1970) has calculated the ω_r to order n^{-3}, while the author (1970*b*, *c*) has given recurrence relations for them, together with a formula relating the ω's for T_0^2 and V. We also note that Ito's (1956) percentile expansion of T_0^2 to order n^{-2} may be derived from (13), the polynomials P_π for partitions containing 1 being required in this connection. The accuracy of Ito's result and its extension to order n^{-3} has been investigated in the author's paper (1970*a*), and similar results are given for V in Davis (1970*b*).

The author wishes to thank Mr J. B. F. Field, of the Division of Mathematical Statistics, C.S.I.R.O., for his assistance with the computing. The work was partly supported by the National Institutes of Health.

References

Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis.* New York: Wiley.

Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika* **36**, 317–46.

Constantine, A. G. (1966). The distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **37**, 215–25.

Consul, P. C. (1969). The exact distributions of likelihood criteria for different hypotheses. In *Multivariate Analysis*, ed. P. R. Krishnaiah, pp. 171–81. New York: Academic Press.

Cornish, E. A. & Fisher, R. A. (1937). Moments and cumulants in the specification of distributions. *Rev. Inst. Int. Statist.* **4**, 307–20.

Davis, A. W. (1968). A system of linear differential equations for the distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **39**, 815–32.

Davis, A. W. (1970*a*). Exact distributions of Hotelling's generalized T_0^2. *Biometrika* **57**, 187–91.

Davis, A. W. (1970*b*). On the null distribution of the sum of the roots of a multivariate beta distribution. *Ann. Math. Statist.* **41**, 1557–62.

Davis, A. W. (1970*c*). Further applications of a differential equation for Hotelling's generalized T_0^2. *Ann. Inst. Statist. Math.* **22**, 77–87.

Fisher, R. A. & Cornish, E. A. (1960). The percentile points of distributions having known cumulants. *Technometrics* **2**, 209–26.

Hill, G. W. & Davis, A. W. (1968). Generalized asymptotic expansions of Cornish–Fisher type. *Ann. Math. Statist.* **39**, 1264–73.

Ito, K. (1956). Asymptotic formulae for the distribution of Hotelling's generalized T_0^2 statistic. *Ann. Math. Statist.* **27**, 1091–105.
Korin, B. P. (1968). On the distribution of a statistic used for testing a covariance matrix. *Biometrika* **55**, 171–8.
Korin, B. P. (1969). On testing the equality of k covariance matrices. *Biometrika* **56**, 216–8.
Muirhead, R. J. (1970). Asymptotic distributions of some multivariate tests. *Ann. Math. Statist.* **41**, 1002–10.
Neyman, J. & Pearson, E. S. (1928*a*). On the use and interpretation of certain test criteria for purposes of statistical inference, Part I. *Biometrika* **20** A, 175–240.
Neyman, J. & Pearson, E. S. (1928*b*). On the use and interpretation of certain test criteria for purposes of statistical inference, Part II. *Biometrika* **20** A, 263–94.
Pillai, K. C. S. (1955). Some new test criteria in multivariate analysis. *Ann. Math. Statist.* **26**, 117–21.
Pillai, K. C. S. & Gupta, A. K. (1969). On the exact distribution of Wilks's criterion. *Biometrika* **56**, 109–18.
Rao, C. R. (1948). Tests of significance in multivariate analysis. *Biometrika* **35**, 58–79.
Schatzoff, M. (1966). Exact distributions of Wilks's likelihood ratio criterion. *Biometrika* **53**, 347–58.
Sugiura, N. & Nagao, H. (1969). On Bartlett's test and Lehmann's test for homogeneity of variances. *Ann. Math. Statist.* **40**, 2018–32.
Wallace, D. L. (1958). Asymptotic approximations to distributions. *Ann. Math. Statist.* **29**, 635–54.

[*Received April* 1970. *Revised December* 1970]

Some key words: Asymptotic expansions; Box's expansion; Chi-squared distribution; Cornish–Fisher expansion; Cumulant generating function; Likelihood ratio test; Multivariate statistics; Percentage point approximation; Wilks's criterion.

19

Tables of Some Multivariate Test Criteria

A. W. DAVIS and J. B. F. FIELD

DIVISION OF MATHEMATICAL STATISTICS TECHNICAL PAPER NO. 32
COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANIZATION, AUSTRALIA 1971

Tables of Some Multivariate Test Criteria

By A. W. Davis and J. B. F. Field

Division of Mathematical Statistics Technical Paper No. 32

Commonwealth Scientific and Industrial
Research Organization, Australia
1971

Printed by CSIRO, Melbourne

TABLES OF SOME MULTIVARIATE TEST CRITERIA

By A. W. Davis* and J. B. F. Field†

Summary

Tables of upper 5% and 1% points are presented for a number of likelihood ratio criteria encountered in multivariate analysis. The tables are constructed using a Cornish–Fisher-type inversion of Box's asymptotic series. The latter is a direct percentile expansion in terms of χ^2 percentiles and the polynomials in this expansion are listed to a sufficient order.

I. A Cornish–Fisher Inversion of Box's Asymptotic Series

The asymptotic expansions of Cornish and Fisher (1937) and Fisher and Cornish (1960) have become well known as a tool for obtaining accurate percentile approximations for distributions which approach normality and which are adequately specified by their early cumulants. Essentially, these expansions are obtained by formal inversion of the Edgeworth series for the distribution in question. Nothing has been published concerning the validity of the Cornish–Fisher series as asymptotic expansions although, according to Wallace (1958), this can be established for absolutely continuous distributions whose Edgeworth series are valid.

The likelihood ratio criteria of Neyman and Pearson (1928) form another important class of statistics whose distributions are specified by their moments and whose percentage points are difficult to obtain by exact methods. In large samples, their distributions are approximately of chi-squared type. Box (1949) has developed a general asymptotic expansion of the cumulative distribution functions (c.d.f.'s) for those cases in which the moments can be expressed as products of gamma functions. As indicated in Davis (1971), it is possible to obtain inversions of this series which bear the same relationship to it as the Cornish–Fisher expansions have to the Edgeworth series. These are derived by applying the general algorithms given in Hill and Davis (1968).

It is convenient to take as starting point the following expansion of the cumulant generating function $k(\theta)$, which holds for a larger class of statistics than that explicitly considered by Box, including the Hotelling generalized T_0^2 and Pillai's V (Davis 1970),

$$k(\theta) \sim -\tfrac{1}{2}f\log(1-2i\theta)+\sum_{r=1}^{\infty}\omega_r\{(1-2i\theta)^{-r}-1\}, \qquad (1)$$

where f denotes the degrees of freedom of the limiting χ^2 distribution. Formulae for the coefficients ω_r are listed in Section II for a number of statistics. Then if

* Division of Mathematical Statistics, CSIRO, Private Bag No. 1, Glen Osmond, S.A. 5064.
† Division of Mathematical Statistics, CSIRO, Private Bag, Townsville, Qld. 4810.

x denotes a percentile of the given distribution, and u is the corresponding χ_f^2 percentile, we have

$$x \sim u + 2 \sum_{\pi} \omega_\pi P_\pi(u) , \tag{2}$$

where the summation is extended over all partitions $\pi = [s_1^{g_1}, \ldots, s_k^{g_k}]$ of all positive numbers, and

$$\omega_\pi = \omega_{s_1}^{g_1} \ldots \omega_{s_k}^{g_k} / g_1! \ldots g_k! .$$

The polynomials $P_\pi(u)$ are listed in Table 1 for partitions of integers up to 8. In many cases it is possible to choose $\omega_1 = 0$. This greatly increases the rate of convergence of (2) and in these cases it is sufficient to consider partitions not containing the integer 1. However, the P_π have been computed for all partitions up to those of 6 and these have been used to obtain Ito-type percentile approximations for T_0^2 and V (Davis 1970).

II. Tabulated Criteria

(i) *Wilks's Generalized Test of Independence.*—(See Box 1949, Section 6; Anderson 1958, Ch. 9). Given N observations from a p-variate normal population, suppose that the variates are partitioned into k groups with sizes

$$p_1, \ldots, p_k \qquad \left(\sum_{i=1}^{k} p_i = p \right) .$$

The Wilks likelihood ratio criterion for testing the null hypothesis that the k groups are independent is

$$\lambda = \left(|\mathbf{R}| \Big/ \prod_{i=1}^{k} |\mathbf{R}_{ii}| \right)^{\frac{1}{2}N} ,$$

where $|\mathbf{R}|$ and $|\mathbf{R}_{ii}|$ are the determinants of the sample correlation coefficients for the complete sample and the ith group respectively ($i = 1, \ldots, k$).

Then $-2\rho \log \lambda$ has a limiting chi-squared distribution with

$$f = \tfrac{1}{2}\left(p^2 - \sum_{i=1}^{k} p_i^2 \right) \tag{3}$$

degrees of freedom. Box has shown that

$$\omega_r = \alpha_r / m^r , \qquad m = \rho N, \quad (r = 1, 2, \ldots) ,$$

where for arbitrary ρ the α_r are expressible in terms of their values α_r' at $\rho = 1$. The α_r' are in turn linear combinations of the quantities

$$\Sigma_s = p^s - \sum_{i=1}^{k} p_i^s , \qquad (s = 2, 3, \ldots) ,$$

and supplementing the list given by Box (1949, Section 6.1) in order to make full use of Table 1, we have

$$\alpha_1' = (1/24)(2\Sigma_3 + 3\Sigma_2),$$

$$\alpha_2' = (1/48)(\Sigma_4 + 2\Sigma_3 - \Sigma_2),$$

$$\alpha_3' = (1/720)(6\Sigma_5 + 15\Sigma_4 - 10\Sigma_3 - 30\Sigma_2),$$

$$\alpha_4' = (1/480)(2\Sigma_6 + 6\Sigma_5 - 5\Sigma_4 - 20\Sigma_3 + 3\Sigma_2),$$

$$\alpha_5' = (1/840)(2\Sigma_7 + 7\Sigma_6 - 7\Sigma_5 - 35\Sigma_4 + 7\Sigma_3 + 49\Sigma_2),$$

$$\alpha_6' = (1/2016)(3\Sigma_8 + 12\Sigma_7 - 14\Sigma_6 - 84\Sigma_5 + 21\Sigma_4 + 196\Sigma_3 - 10\Sigma_2),$$

$$\alpha_7' = (1/10080)(10\Sigma_9 + 45\Sigma_8 - 60\Sigma_7 - 420\Sigma_6 + 126\Sigma_5 + 1470\Sigma_4 - 100\Sigma_3 - 1860\Sigma_2),$$

$$\alpha_8' = (1/2880)(2\Sigma_{10} + 10\Sigma_9 - 15\Sigma_8 - 120\Sigma_7 + 42\Sigma_6 + 588\Sigma_5 - 50\Sigma_4 - 1240\Sigma_3 + 21\Sigma_2).$$

Then

$$\alpha_r = \sum_{i=0}^{r-1} \binom{r-1}{i}(-\beta)^i \alpha_{r-i}' + (f/2r)(-\beta)^r,$$

where

$$\beta = N(1-\rho) - 1.$$

Thus α_1, and hence ω_1, will be reduced to zero by taking

$$\beta = 2\alpha_1'/f,$$

whence

$$\rho = 1 - (2\Sigma_3 + 9\Sigma_2)/6N\Sigma_2 \qquad \text{and} \qquad m = N - \tfrac{3}{2} - \Sigma_3/3\Sigma_2. \tag{4}$$

In the important case of two groups ($k = 2$), the distribution of $\lambda^{2/N}$ is identical with that of Wilks's likelihood ratio statistic $U_{p,q,n}$, if we take

$$p = p_1, \qquad q = p_2, \qquad n = N - q - 1.$$

From (3) and (4) we have in this case

$$f = pq, \qquad m = N - \tfrac{1}{2}(p+q+3),$$

and the expressions for the α_i reduce considerably (Rao 1948; Box 1949, Section 6.21):

$$\alpha_1 = 0, \qquad \alpha_2 = (1/48)pq(p^2+q^2-5),$$

$$\alpha_3 = 0, \qquad \alpha_4 = (1/1920)pq\{3(p^4+q^4)+10p^2q^2-50(p^2+q^2)+159\},$$

$$\alpha_5 = 0, \qquad \alpha_6 = (1/16128)pq[3(p^6+q^6)-105(p^4+q^4)+1113(p^2+q^2) + \{21(p^2+q^2)-350\}p^2q^2-2995].$$

TABLE 1

The Polynomials $P_\pi(u)$

$f_r = f(f+2)\ \dots\ (f+2(r-1))$,   ($r = 1, 2, \dots$)

π	$P_\pi(u)$
2	$u^2/f_2+u/f$
3	$\Sigma_{i=1}^{3} u^i/f_i$
4	$\Sigma_{i=1}^{4} u^i/f_i$
2^2	$-8u^4(f+3)/f_2f_4+8u^3/f_2f_3+6u^2/ff_2+2u/f^2$
5	$\Sigma_{i=1}^{5} u^i/f_i$
2,3	$-12u^5(f+4)/f_2f_5-2u^4(f-6)/f_2f_4+2u^3(3f+10)/f_2f_3+6u^2/ff_2+2u/f^2$
6	$\Sigma_{i=1}^{6} u^i/f_i$
2,4	$-16u^6(f+5)/f_2f_6-4u^5(f-4)/f_2f_5+2u^4(3f+14)/f_2f_4+2u^3(3f+10)/f_2f_3+6u^2/ff_2+2u/f^2$
3^2	$-6u^6(3f^2+30f+80)/f_3f_6-6u^5(f^2+2f-16)/f_3f_5+4u^4(f+12)/f_2f_4+4u^3(3f+8)/f_2f_3$ $+6u^2/ff_2+2u/f^2$
2^3	$32u^6(7f^2+62f+120)/f_2^2f_6-32u^5(2f^2+37f+96)/f_2^2f_5-8u^4(23f^2+124f+132)/f_2^2f_4$ $-8u^3(f-10)/ff_2f_3+28u^2/f^2f_2+4u/f^3$
7	$\Sigma_{i=1}^{7} u^i/f_i$
2,5	$-20u^7(f+6)/f_2f_7-2u^6(3f-10)/f_2f_6+2\ \Sigma_{i=2}^{5} u^i(3f+4i-2)/f_2f_i+2u/f^2$
3,4	$-24u^7(f^2+12f+40)/f_3f_7-2u^6(5f^2+18f-80)/f_3f_6+2u^5(f^2+42f+176)/f_3f_5$ $+4u^4(3f+16)/f_2f_4+4u^3(3f+8)/f_2f_3+6u^2/ff_2+2u/f^2$
$2^2,3$	$192u^7(2f^3+31f^2+154f+240)/f_2f_3f_7-16u^6(4f^3+153f^2+1106f+2160)/f_2f_3f_6$ $-8u^5(35f^3+420f^2+1540f+1632)/f_2f_3f_5-4u^4(25f^2+80f+12)/f_2^2f_4$ $+4u^3(7f+38)/ff_2f_3+28u^2/f^2f_3+4u/f^3$

TABLE 1 (*Continued*)

π	$P_\pi(u)$
8	$\Sigma_{i=1}^{8}\ u^i/f_i$
2,6	$-24u^8(f+7)/f_2f_8-8u^7(f-3)/f_2f_7+2\ \Sigma_{i=2}^{6}\ u^i(3f+4i-2)/f_2f_i+2u/f^2$
3,5	$-30u^8(f^2+14f+56)/f_3f_8-2u^7(7f^2+34f-120)/f_3f_7+8u^6(13f+70)/f_3f_6$ $+12u^5(f+6)^2/f_3f_5+4u^4(3f+16)/f_2f_4+4u^3(3f+8)/f_2f_3+6u^2/ff_2+2u/f^2$
4^2	$-32u^8(f^3+21f^2+158f+420)/f_4f_8-16u^7(f^3+12f^2+20f-120)/f_4f_7$ $-2u^6(f^2-54f-400)/f_3f_6+2u^5(5f^2+90f+304)/f_3f_5+20u^4(f+4)/f_2f_4$ $+4u^3(3f+8)/f_2f_3+6u^2/ff_2+2u/f^2$
$2^2,4$	$64u^8(9f^4+216f^3+1914f^2+7356f+10{,}080)/f_2f_4f_8$ $-64u^7(f^4+66f^3+890f^2+4404f+7200)/f_2f_4f_7$ $-4u^6(100f^3+1388f^2+5960f+7360)/f_2f_3f_6-8u^5(17f^3+168f^2+532f+480)/f_2f_3f_5$ $+4u^4(-9f^2-16f+52)/f_2^2f_4+u^3(28f+152)/ff_2f_3+28u^2/f^2f_2+4u/f^3$
$2,3^2$	$24u^8(27f^3+522f^2+3336f+6720)/f_2f_3f_8-24u^7(f^3+158f^2+1752f+4800)/f_2f_3f_7$ $-4u^6(97f^3+1692f^2+8612f+12{,}000)/f_2f_3f_6-4u^5(79f^3+780f^2+2108f+1056)/f_2f_3f_5$ $-16u^4(f^2-22f-60)/f_2^2f_4+32u^3(2f+7)/ff_2f_3+28u^2/f^2f_2+4u/f^3$
2^4	$-2304u^8(5f^5+143f^4+1580f^3+8376f^2+21{,}096f+20{,}160)/f_2^2f_4f_8$ $+256u^7(23f^5+873f^4+11{,}536f^3+69{,}360f^2+192{,}312f+198{,}720)/f_2^2f_4f_7$ $+128u^6(81f^4+1202f^3+5733f^2+9178f+1680)/f_2^2f_3f_6$ $-384u^5(5f^4+129f^3+913f^2+245f+2160)/f_2^2f_3f_5$ $-16u^4(175f^2+998f+1056)/ff_2^2f_4-16u^3(9f-32)/f^2f_2f_3+120u^2/f^3f_2+8u/f^4$

TABLE 2

Wilks's Generalized Test of Independence : Correction Factors for $-2\rho \log \lambda$

Partitions:	2,1,1		3,1,1		2,2,1		4,1,1		3,2,1		
N	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	N
6	1.07	1.08	1.32	1.40	1.29						6
7	1.034	1.042	1.12	1.152	1.109	1.13					7
8	1.020	1.025	1.067	1.0813	1.058	1.071	1.18	1.21	1.15	1.18	8
9	1.0135	1.0168	1.042	1.0509	1.036	1.044	1.100	1.12	1.083	1.10	9
10	1.0097	1.0119	1.0291	1.0350	1.0250	1.0300	1.0646	1.077	1.0536	1.063	10
11	1.0072	1.0089	1.0213	1.0256	1.0182	1.0218	1.0454	1.054	1.0376	1.044	11
12	1.0056	1.0069	1.0162	1.0195	1.0139	1.0166	1.0338	1.0399	1.0279	1.033	12
13	1.0045	1.0055	1.0128	1.0154	1.0110	1.0131	1.0261	1.0308	1.0216	1.0252	13
14	1.0037	1.0045	1.0104	1.0124	1.0088	1.0105	1.0209	1.0245	1.0172	1.0201	14
15	1.0031	1.0038	1.0086	1.0103	1.0073	1.0087	1.0170	1.0200	1.0140	1.0163	15
16	1.0026	1.0032	1.0072	1.0086	1.0061	1.0073	1.0142	1.0166	1.0117	1.0136	16
17	1.0022	1.0027	1.0061	1.0073	1.0052	1.0062	1.0120	1.0141	1.0099	1.0115	17
18	1.0019	1.0024	1.0053	1.0063	1.0045	1.0053	1.0103	1.0120	1.0084	1.0098	18
19	1.0017	1.0021	1.0046	1.0055	1.0039	1.0046	1.0089	1.0104	1.0073	1.0085	19
20	1.0015	1.0018	1.0040	1.0048	1.0034	1.0041	1.0078	1.0091	1.0064	1.0074	20
21	1.0013	1.0016	1.0036	1.0043	1.0030	1.0036	1.0069	1.0080	1.0056	1.0065	21
22	1.0012	1.0014	1.0032	1.0038	1.0027	1.0032	1.0061	1.0072	1.0050	1.0058	22
23	1.0011	1.0013	1.0029	1.0034	1.0024	1.0029	1.0055	1.0064	1.0045	1.0052	23
24	1.0010	1.0012	1.0026	1.0031	1.0022	1.0026	1.0049	1.0058	1.0040	1.0047	24
25	1.0009	1.0011	1.0023	1.0028	1.0020	1.0024	1.0044	1.0052	1.0036	1.0042	25
30	1.0006	1.0007	1.0015	1.0018	1.0013	1.0015	1.0029	1.0034	1.0024	1.0027	30
35	1.0004	1.0005	1.0011	1.0013	1.0009	1.0011	1.0020	1.0024	1.0016	1.0019	35
40	1.0003	1.0004	1.0008	1.0010	1.0007	1.0008	1.0015	1.0017	1.0012	1.0014	40
45	1.0002	1.0003	1.0006	1.0007	1.0005	1.0006	1.0011	1.0013	1.0009	1.0011	45
50	1.0002	1.0002	1.0005	1.0006	1.0004	1.0005	1.0009	1.0011	1.0007	1.0009	50
55	1.0002	1.0002	1.0004	1.0005	1.0003	1.0004	1.0007	1.0009	1.0006	1.0007	55
60	1.0001	1.0002	1.0003	1.0004	1.0003	1.0003	1.0006	1.0007	1.0005	1.0006	60
120	1.0000	1.0000	1.0001	1.0001	1.0001	1.0001	1.0001	1.0002	1.0001	1.0001	120
∞	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	∞
χ_f^2	11.0705	15.0863	14.0671	18.4753	15.5073	20.0902	16.9190	21.6660	19.6751	24.7250	χ_f^2

TABLE 2 (*Continued*)

Partitions:	1^4		1^5		1^6		1^7		1^8		
N	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	*N*
7	1.034	1.043	1.115	1.14							7
8	1.021	1.0255	1.062	1.08							8
9	1.0136	1.0169	1.039	1.047							9
10	1.0097	1.0120	1.0267	1.032	1.055	1.06	1.10	1.12			10
11	1.0073	1.0089	1.0195	1.0233	1.0386	1.045	1.07	1.08	1.12	1.13	11
12	1.0056	1.0069	1.0149	1.0177	1.0288	1.033	1.049	1.057	1.08	1.09	12
13	1.0045	1.0055	1.0118	1.0139	1.0223	1.026	1.037	1.043	1.059	1.07	13
14	1.0037	1.0045	1.0095	1.0112	1.0178	1.0205	1.0293	1.033	1.045	1.051	14
15	1.0031	1.0037	1.0079	1.0093	1.0145	1.0167	1.0236	1.027	1.036	1.040	15
16	1.0026	1.0032	1.0066	1.0078	1.0121	1.0139	1.0195	1.0220	1.0293	1.033	16
17	1.0022	1.0027	1.0056	1.0066	1.0102	1.0117	1.0163	1.0184	1.0243	1.027	17
18	1.0019	1.0023	1.0048	1.0057	1.0088	1.0101	1.0139	1.0157	1.0205	1.0229	18
19	1.0017	1.0020	1.0042	1.0049	1.0076	1.0087	1.0120	1.0135	1.0176	1.0195	19
20	1.0015	1.0018	1.0037	1.0043	1.0066	1.0076	1.0104	1.0117	1.0152	1.0169	20
21	1.0013	1.0016	1.0033	1.0038	1.0059	1.0067	1.0092	1.0103	1.0133	1.0148	21
22	1.0012	1.0014	1.0029	1.0034	1.0052	1.0060	1.0081	1.0091	1.0117	1.0130	22
23	1.0011	1.0013	1.0026	1.0031	1.0047	1.0053	1.0072	1.0081	1.0104	1.0116	23
24	1.0010	1.0012	1.0024	1.0028	1.0042	1.0048	1.0065	1.0073	1.0093	1.0103	24
25	1.0009	1.0010	1.0021	1.0025	1.0038	1.0043	1.0059	1.0066	1.0084	1.0093	25
30	1.0006	1.0007	1.0014	1.0016	1.0025	1.0028	1.0038	1.0042	1.0053	1.0059	30
35	1.0004	1.0005	1.0010	1.0012	1.0017	1.0020	1.0026	1.0029	1.0037	1.0041	35
40	1.0003	1.0004	1.0007	1.0009	1.0013	1.0014	1.0019	1.0022	1.0027	1.0030	40
50	1.0002	1.0002	1.0004	1.0005	1.0008	1.0009	1.0012	1.0013	1.0016	1.0018	50
60	1.0001	1.0001	1.0003	1.0004	1.0005	1.0006	1.0008	1.0009	1.0011	1.0012	60
90	1.0001	1.0001	1.0001	1.0001	1.0002	1.0002	1.0003	1.0004	1.0004	1.0005	90
120	1.0000	1.0000	1.0001	1.0001	1.0001	1.0001	1.0002	1.0002	1.0002	1.0003	120
∞	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	∞
χ_f^2	12.5916	16.8119	18.3070	23.2093	24.9958	30.5779	32.6705	38.9321	41.3372	48.2782	χ_f^2

It may also be shown that

$$\alpha_7 = 0, \qquad \alpha_8 = (1/184320)pq[5(p^8+q^8)-300(p^6+q^6)+6678(p^4+q^4) -59900(p^2+q^2)+\{60(p^4+q^4)+126p^2q^2-2100(p^2+q^2) +22260\}p^2q^2+151053].$$

Exact correction factors x/u in the case $k = 2$ have been tabulated by Schatzoff (1966) and Pillai and Gupta (1969). The $O(m^{-8})$ expansion gives values differing from their exact results by no more than one digit in the third decimal place for $M = n-p+1$ ranging from as low as $M = 3$ for $p = 3$, $q = 4$ to $M \geqslant 7$ or 8 for larger p, q.

Table 2 gives correction factors for selected partitions with $k \geqslant 3$ obtained from the Cornish–Fisher expansion.

(ii) *Generalized Test of Homoscedasticity*.—(See Box 1949, Section 2; Anderson 1958, Ch. 9). Consider k samples of sizes N_i ($i = 1, \ldots, q$; $\Sigma N_i = N$) from q p-variate normal populations. The hypothesis that these populations have equal covariance matrices may be tested by the following generalization of Bartlett's univariate test.

$$W_1 = \prod_{i=1}^{q} |\mathbf{S}_i|^{\frac{1}{2}n_i} / |\mathbf{S}|^{\frac{1}{2}n},$$

where

$$n_i = N_i - 1, \qquad \text{and} \qquad n = \sum_{i=1}^{q} n_i = N - q.$$

$\mathbf{S}_i$ is the ith sample covariance matrix and $\mathbf{S} = \sum_{i=1}^{q} n_i \mathbf{S}_i/n$. The moments of W_1 have Box's form and it is found that

$$f = \tfrac{1}{2}(q-1)p(p+1),$$

while for $\omega_1 = 0$ we require

$$\rho = 1-\left(\sum_{i=1}^{q}\frac{1}{n_i}-\frac{1}{n}\right)\frac{(2p^2+3p-1)}{6(p+1)(q-1)}.$$

The higher ω_r are given by

$$\omega_r = \frac{(-1)^r}{r(r+1)(r+2)\rho^r}\sum_{s=1}^{r+1}\binom{r+2}{s+1}2^s\theta_s\,\delta_s(1-\rho)^{r+1-s},$$

where

$$\theta_s = \sum_{i=1}^{q}\frac{1}{n_i^{s-1}}-\frac{1}{n^{s-1}},$$

and the first eight ω's may be obtained using

$$\delta_1 = \tfrac{1}{4}p(p+1),$$

$$\delta_2 = -(1/16)p(2p^2+3p-1),$$

$\delta_3 = (1/16)p(p-1)(p+1)(p+2)$,

$\delta_4 = -(1/192)p(6p^4+15p^3-10p^2-30p+3)$,

$\delta_5 = (1/128)p(p-1)(p+1)(p+2)(2p^2+2p-7)$,

$\delta_6 = -(1/768)p(6p^6+21p^5-21p^4-105p^3+21p^2+147p-5)$,

$\delta_7 = (1/768)p(p-1)(p+1)(p+2)(3p^4+6p^3-23p^2-26p+62)$,

$\delta_8 = -(1/5120)p(10p^8+45p^7-60p^6-420p^5+126p^4+1470p^3-100p^2-1860p+21)$,

$\delta_9 = (1/2048)p(p-1)(p+1)(p+2)(2p^6+6p^5-25p^4-60p^3+149p^2+180p-381)$.

Table 3 gives 5% points for $p = 2(1)6$, $q = 2(1)5$, with equal sample sizes $n_i = n$. These may be compared with Korin's (1969) tabulation.

(iii) *Sphericity Test.*—(See Anderson 1958, Section 10.7). In testing the hypothesis that a sample of size N is drawn from a p-variate normal population whose covariance matrix $\mathbf{\Sigma}$ is proportional to a given matrix $\mathbf{\Sigma}_0$,

$$\mathbf{\Sigma} = \sigma^2\mathbf{\Sigma}_0\,,$$

the likelihood ratio criterion is a power of

$$W = |\,\mathbf{S\Sigma}_0^{-1}\,|^{\frac{1}{2}n}(\text{trace}\,\mathbf{S\Sigma}_0^{-1}/p)^{-\frac{1}{2}pn}\,, \qquad (n = N-1)\,,$$

where $\mathbf{S}$ is the sample covariance matrix. Then

$$f = \tfrac{1}{2}(p-1)(p+2)\,,$$

and ω_1 vanishes for

$$\rho = 1-2\beta_1/pn\,, \qquad \beta_1 = (1/12)(2p^2+p+2)\,.$$

In this case

$$\omega_r = \alpha_r/(np\rho)^r\,,$$

where

$$\alpha_r = \frac{(-1)^r 2^{r+1}}{r(r+1)(r+2)}\sum_{s=1}^{r+1}\binom{r+2}{s+1}\beta_1^{r+1-s}\{p^{s-1}\delta_s + \tfrac{1}{2}(s+1)B_s\}\,.$$

Here the B_s are Bernoulli's numbers, and the δ_s have been given above. In particular,

$\alpha_2 = (1/288)(p-2)(p-1)(p+2)(2p^3+6p^2+3p+2)$,

$\alpha_3 = (1/3240)(p-2)^3(p-1)(p+1)^2(p+2)(p+4)(2p-1)$,

$\alpha_4 = (1/34560)(p-2)(p-1)(p+2)$

$\quad \times(16p^7+64p^6-60p^5-388p^4-277p^3-258p^2-188p-152)$,

$\alpha_5 = (1/136080)(p-2)^2(p-1)(p+1)^2(p+2)(p+4)(2p-1)(8p^4+8p^3-81p^2-16p-52)$.

TABLE 3

Generalized Test of Homoscedasticity : Upper 5% Points of $-2 \log W_1$

n	2 Samples ($q = 2$)					3 Samples ($q = 3$)					n
	$p =$ 2	3	4	5	6	$p =$ 2	3	4	5	6	
3	12.186					18.70					3
4	10.7017					16.6527					4
5	9.9689					15.6335					5
6	9.5325	17.57	30.07	51.14		15.0233	28.235	48.63			6
7	9.2430	16.5910	27.310	43.40		14.6171	26.8404	44.690			7
8	9.0370	15.9342	25.6100	39.29	59.24	14.3271	25.9002	42.2406	65.15		8
9	8.8829	15.4632	24.4548	36.704	53.571	14.1098	25.2228	40.5629	61.39		9
10	8.7633	15.1089	23.6174	34.924	49.952	13.9409	24.7113	39.3394	58.781	84.42	10
11	8.6678	14.8325	22.9820	33.6193	47.425	13.8058	24.3112	38.4066	56.853	80.69	11
12	8.5898	14.6109	22.4830	32.6207	45.554	13.6953	23.9897	37.6714	55.3698	77.90	12
13	8.5249	14.4293	22.0807	31.8310	44.110	13.6033	23.7257	37.0768	54.1915	75.735	13
14	8.4699	14.2776	21.7493	31.1906	42.9604	13.5254	23.5049	36.5857	53.2324	74.004	14
15	8.4229	14.1492	21.4716	30.6605	42.0229	13.4586	23.3175	36.1733	52.4361	72.586	15
16	8.3822	14.0389	21.2354	30.2145	41.2433	13.4008	23.1565	35.8220	51.7643	71.404	16
17	8.3466	13.9432	21.0321	29.8338	40.5846	13.3502	23.0167	35.5191	51.1897	70.401	17
18	8.3152	13.8594	20.8553	29.5051	40.0204	13.3056	22.8941	35.2552	50.6925	69.5406	18
19	8.2873	13.7854	20.7001	29.2183	39.5318	13.2659	22.7858	35.0232	50.2580	68.7935	19
20	8.2623	13.7196	20.5627	28.9660	39.1044	13.2304	22.6893	34.8177	49.8751	68.1387	20
21	8.2399	13.6606	20.4402	28.7422	38.7273	13.1984	22.6029	34.6343	49.5350	67.5600	21
22	8.2196	13.6075	20.3304	28.5423	38.3922	13.1695	22.5250	34.4697	49.2310	67.0448	22
23	8.2011	13.5595	20.2314	28.3627	38.0923	13.1432	22.4545	34.3211	48.9575	66.5832	23
24	8.1843	13.5158	20.1416	28.2005	37.8224	13.1193	22.3903	34.1863	48.7101	66.1672	24
25	8.1689	13.4758	20.0598	28.0532	37.5781	13.0973	22.3316	34.0635	48.4853	65.7903	25
30	8.1077	13.3187	19.7404	27.4821	36.6383	13.0101	22.1006	33.5829	47.6120	64.3362	30
35	8.0646	13.2091	19.5195	27.0909	36.0010	12.9486	21.9390	33.2498	47.0120	63.3465	35
40	8.0325	13.1282	19.3576	26.8062	35.5403	12.9028	21.8197	33.0053	46.5743	62.6291	40
45	8.0077	13.0660	19.2339	26.5896	35.1916	12.8675	21.7280	32.8182	46.2409	62.0852	45
50	7.9880	13.0167	19.1363	26.4193	34.9185	12.8393	21.6553	32.6704	45.9785	61.6586	50
55	7.9720	12.9768	19.0572	26.2819	34.6989	12.8164	21.5962	32.5507	45.7665	61.3150	55
60	7.9587	12.9437	18.9920	26.1688	34.5183	12.7974	21.5473	32.4518	45.5917	61.0323	60
120	7.8862	12.7650	18.6421	25.5663	33.5642	12.6937	21.2829	31.9207	44.6593	59.5339	120
χ^2_f	7.8147	12.5916	18.3070	24.9958	32.6705	12.5916	21.0261	31.4104	43.7729	58.1240	χ^2_f

TABLE 3 (*Continued*)

n	4 Samples ($q = 4$)					5 Samples ($q = 5$)					n
	$p =$ 2	3	4	5	6	$p =$ 2	3	4	5	6	
3	24.55					30.09					3
4	22.003					27.0732					4
5	20.7318	40.95				25.5641	50.950				5
6	19.9695	38.0589				24.6583	47.487				6
7	19.4614	36.2929	60.90			24.0541	45.367	76.56			7
8	19.0984	35.0997	57.767	89.45		23.6224	43.9322	72.774			8
9	18.8262	34.2384	55.617	84.62		23.2984	42.8960	70.171			9
10	18.6144	33.5872	54.0454	81.25		23.0464	42.1120	68.266	103.06		10
11	18.4450	33.0773	52.8450	78.754	112.17	22.8447	41.4977	66.8091	100.025	142.84	11
12	18.3063	32.6672	51.8975	76.831	108.55	22.6796	41.0034	65.6587	97.679	138.40	12
13	18.1908	32.3301	51.1302	75.300	105.73	22.5420	40.5969	64.7265	95.810	134.95	13
14	18.0930	32.0480	50.4960	74.0529	103.466	22.4256	40.2567	63.9555	94.2865	132.18	14
15	18.0092	31.8085	49.9628	73.0161	101.612	22.3257	39.9678	63.3071	93.0191	129.906	15
16	17.9366	31.6027	49.5083	72.1404	100.0625	22.2392	39.1793	62.7542	91.9480	128.005	16
17	17.8730	31.4238	49.1162	71.3907	98.7481	22.1635	39.5034	62.2769	91.0308	126.391	17
18	17.8169	31.2669	48.7744	70.7416	97.6185	22.0966	39.3140	61.8608	90.2362	125.003	18
19	17.7671	31.1282	48.4738	70.1741	96.6371	22.0372	39.1465	61.4948	89.5412	123.796	19
20	17.7225	31.0047	48.2074	69.6735	95.7763	21.9841	38.9974	61.1703	88.9281	122.7377	20
21	17.6823	30.8940	47.9696	69.2287	95.0150	21.9362	38.8637	60.8806	88.3831	121.8012	21
22	17.6460	30.7942	47.7560	68.8308	94.3369	21.8929	38.7431	60.6204	87.8954	120.9668	22
23	17.6130	30.7038	47.5632	68.4728	93.7289	21.8536	38.6339	60.3854	87.4565	120.2185	23
24	17.5828	30.6215	47.3882	68.1488	93.1808	21.8176	38.5345	60.1721	87.0594	119.5436	24
25	17.5552	30.5463	47.2287	67.8544	92.6840	21.7847	38.4436	59.9776	86.6983	118.9318	25
30	17.4455	30.2500	46.6042	66.7092	90.7650	21.6540	38.0855	59.2162	85.2933	116.5674	30
35	17.3681	30.0427	46.1710	65.9216	89.4571	21.5617	37.8349	58.6877	84.3265	114.9546	35
40	17.3106	29.8896	45.8528	65.3466	88.5081	21.4931	37.6497	58.2994	83.6203	113.7839	40
45	17.2661	29.7718	45.6092	64.9083	87.7880	21.4400	37.5072	58.0021	83.0819	112.8951	45
50	17.2307	29.6784	45.4167	64.5631	87.2228	21.3978	37.3942	57.7670	82.6577	112.1974	50
55	17.2018	29.6025	45.2608	64.2842	86.7674	21.3634	37.3025	57.5766	82.3149	111.6350	55
60	17.1779	29.5396	45.1318	64.0541	86.3926	21.3348	37.2264	57.4191	82.0321	111.1721	60
120	17.0474	29.1997	44.4392	62.8258	84.4037	21.1792	36.8150	56.5729	80.5215	108.7142	120
χ_f^2	16.9190	28.8693	43.7729	61.6562	82.5287	21.0261	36.4151	55.7585	79.0819	106.3948	χ_f^2

TABLE 4

Sphericity Test $\Sigma = \sigma^2 \Sigma_0$: Correction Factors for $-2\rho \log W$

p	3		4		5		6		*p*
n	5%	1%	5%	1%	5%	1%	5%	1%	*n*
6	1.0228	1.0280	1.066	1.079	1.15	1.18			6
7	1.0153	1.0188	1.0414	1.0494	1.09	1.10			7
8	1.0110	1.0135	1.0286	1.0340	1.057	1.067	1.11	1.12	8
9	1.0083	1.0101	1.0210	1.0248	1.040	1.047	1.070	1.08	9
10	1.0064	1.0079	1.0160	1.0190	1.0301	1.0348	1.051	1.058	10
11	1.0052	1.0063	1.0127	1.0150	1.0233	1.0270	1.039	1.044	11
12	1.0042	1.0052	1.0103	1.0121	1.0187	1.0215	1.0303	1.0343	12
13	1.0035	1.0043	1.0085	1.0100	1.0153	1.0176	1.0244	1.0277	13
14	1.0030	1.0037	1.0071	1.0084	1.0127	1.0146	1.0202	1.0228	14
15	1.0026	1.0031	1.0061	1.0072	1.0108	1.0124	1.0169	1.0191	15
16	1.0022	1.0027	1.0053	1.0062	1.0092	1.0106	1.0144	1.0162	16
17	1.0020	1.0024	1.0046	1.0054	1.0080	1.0092	1.0124	1.0140	17
18	1.0017	1.0021	1.0040	1.0047	1.0070	1.0080	1.0108	1.0122	18
19	1.0015	1.0019	1.0036	1.0042	1.0062	1.0071	1.0095	1.0107	19
20	1.0014	1.0017	1.0032	1.0037	1.0055	1.0063	1.0084	1.0095	20
21	1.0012	1.0015	1.0029	1.0034	1.0049	1.0056	1.0075	1.0085	21
22	1.0011	1.0014	1.0026	1.0030	1.0044	1.0051	1.0068	1.0076	22
23	1.0010	1.0012	1.0023	1.0028	1.0040	1.0046	1.0061	1.0069	23
24	1.0009	1.0011	1.0021	1.0025	1.0037	1.0042	1.0055	1.0062	24
25	1.0009	1.0010	1.0020	1.0023	1.0033	1.0038	1.0051	1.0057	25
30	1.0006	1.0007	1.0013	1.0016	1.0022	1.0026	1.0034	1.0038	30
35	1.0004	1.0005	1.0010	1.0011	1.0016	1.0018	1.0024	1.0027	35
40	1.0003	1.0004	1.0007	1.0008	1.0012	1.0014	1.0018	1.0020	40
45	1.0002	1.0003	1.0006	1.0007	1.0009	1.0011	1.0014	1.0016	45
50	1.0002	1.0002	1.0005	1.0005	1.0008	1.0009	1.0011	1.0013	50
60	1.0001	1.0002	1.0003	1.0004	1.0005	1.0006	1.0008	1.0009	60
120	1.0000	1.0000	1.0001	1.0001	1.0001	1.0001	1.0002	1.0002	120
∞	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	∞
χ^2_f	11.0705	15.0863	16.9190	21.6660	23.6848	29.1413	31.4104	37.5662	χ^2_f

TABLE 4 (*Continued*)

p	7		8		9		10		p
n	5%	1%	5%	1%	5%	1%	5%	1%	n
11	1.060	1.068	1.09	1.10	1.14				11
12	1.046	1.052	1.07	1.08	1.10	1.11			12
13	1.037	1.041	1.054	1.060	1.08	1.08	1.11	1.12	13
14	1.0300	1.033	1.043	1.048	1.060	1.07	1.08	1.09	14
15	1.0249	1.0278	1.035	1.039	1.049	1.054	1.07	1.07	15
16	1.0211	1.0234	1.0296	1.033	1.040	1.044	1.055	1.059	16
17	1.0181	1.0201	1.0252	1.0277	1.034	1.037	1.045	1.049	17
18	1.0156	1.0174	1.0217	1.0238	1.0292	1.032	1.039	1.042	18
19	1.0137	1.0152	1.0189	1.0207	1.0252	1.0275	1.033	1.036	19
20	1.0121	1.0134	1.0166	1.0182	1.0220	1.0240	1.0287	1.031	20
21	1.0107	1.0119	1.0147	1.0161	1.0194	1.0212	1.0252	1.0273	21
22	1.0096	1.0107	1.0131	1.0143	1.0173	1.0188	1.0223	1.0241	22
23	1.0087	1.0096	1.0118	1.0129	1.0155	1.0168	1.0199	1.0215	23
24	1.0078	1.0087	1.0106	1.0116	1.0139	1.0151	1.0179	1.0193	24
25	1.0071	1.0079	1.0096	1.0105	1.0126	1.0137	1.0161	1.0174	25
30	1.0047	1.0052	1.0063	1.0069	1.0082	1.0088	1.0103	1.0111	30
35	1.0033	1.0037	1.0044	1.0049	1.0057	1.0062	1.0072	1.0077	35
40	1.0025	1.0028	1.0033	1.0036	1.0042	1.0046	1.0053	1.0057	40
45	1.0019	1.0021	1.0026	1.0028	1.0033	1.0035	1.0041	1.0044	45
50	1.0015	1.0017	1.0020	1.0022	1.0026	1.0028	1.0032	1.0035	50
60	1.0010	1.0012	1.0014	1.0015	1.0017	1.0019	1.0022	1.0023	60
120	1.0002	1.0003	1.0003	1.0004	1.0004	1.0004	1.0005	1.0005	120
∞	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	∞
χ^2_f	40.1133	46.9630	49.8019	57.3421	60.4809	68.7095	72.1532	81.0688	χ^2_f

TABLE 5

Test of Hypothesis $\Sigma = \Sigma_0$: Percentiles of $-2 \log \lambda_1$

p	2		3		4		5		6		*p*
n	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	*n*
6	8.9415	13.0019	15.805	21.229							6
7	8.7539	12.7231	15.1854	20.358	24.06	30.75					7
8	8.6198	12.5246	14.7676	19.7750	23.002	29.32					8
9	8.5193	12.3761	14.4663	19.3577	22.278	28.357	32.47	39.97			9
10	8.4411	12.2608	14.2387	19.0439	21.749	27.657	31.36	38.55			10
11	8.3786	12.1687	14.0605	18.7992	21.3456	27.1268	30.549	37.51	42.08		11
12	8.3274	12.0935	13.9173	18.6029	21.0276	26.7102	29.922	36.710	40.92	48.96	12
13	8.2847	12.0308	12.7995	18.4421	20.7702	26.3743	29.424	36.079	40.02	47.84	13
14	8.2486	11.9778	13.7010	18.3078	20.5576	26.0975	29.0182	35.567	39.303	46.96	14
15	8.2177	11.9325	13.6174	18.1940	20.3789	25.8655	28.6812	35.1435	38.714	46.234	15
16	8.1909	11.8932	13.5456	18.0964	20.2266	25.6681	28.3967	34.7866	38.222	45.632	16
17	8.1674	11.8588	13.4832	18.0116	20.0953	25.4982	28.1532	34.4818	37.806	45.122	17
18	8.1467	11.8285	13.4284	17.9373	19.9808	25.3503	27.9425	34.2185	37.4475	44.686	18
19	8.1283	11.8015	13.3800	17.8718	19.8801	25.2204	27.7582	33.9886	37.1365	44.3069	19
20	8.1118	11.7774	13.3369	17.8134	19.7909	25.1054	27.5958	33.7862	36.8638	43.9754	20
21	8.0970	11.7557	13.2983	17.7611	19.7113	25.0029	27.4514	33.6066	36.6227	43.6827	21
22	8.0835	11.7361	13.2634	17.7141	19.6398	24.9109	27.3224	33.4461	36.4080	43.4224	22
23	8.0713	11.7183	13.2319	17.6715	19.5753	24.8279	27.2062	33.3018	36.2155	43.1892	23
24	8.0602	11.7020	13.2032	17.6327	19.5167	24.7527	27.1011	33.1713	36.0420	42.9792	24
25	8.0500	11.6871	13.1769	17.5973	19.4633	24.6841	27.0056	33.0529	35.8847	42.7890	25
26	8.0406	11.6734	13.1529	17.5648	19.4144	24.6214	26.9184	32.9448	35.7415	42.6160	26
27	8.0319	11.6608	13.1307	17.5349	19.3695	24.5638	26.8385	32.8458	35.6106	42.4579	27
28	8.0239	11.6490	13.1102	17.5073	19.3281	24.5107	26.7650	32.7547	35.4904	42.3128	28
29	8.0164	11.6382	13.0912	17.4817	19.2899	24.4617	26.6971	32.6707	35.3797	42.1793	29
30	8.0095	11.6280	13.0735	17.4579	19.2543	24.4162	26.6342	32.5930	35.2774	42.0559	30
35	7.9809	11.5864	13.0012	17.3606	19.1094	24.2307	26.3788	32.2774	34.8635	41.5575	35
40	7.9597	11.5554	12.9478	17.2887	19.0029	24.0946	26.1924	32.0474	34.5632	41.1965	40
45	7.9432	11.5314	12.9067	17.2335	18.9214	23.9906	26.0503	31.8723	34.3353	40.9229	45
50	7.9301	11.5124	12.8741	17.1898	18.8570	23.9085	25.9384	31.7345	34.1565	40.7083	50
60	7.9106	11.4840	12.8257	17.1249	18.7617	23.7870	25.7734	31.5315	33.8937	40.3935	60
120	7.8624	11.4139	12.7071	16.9660	18.5300	23.4922	25.3751	31.0424	33.2642	39.6405	120
∞	7.8147	11.3449	12.5916	16.8119	18.3070	23.2093	24.9958	30.5779	32.6705	38.9321	∞

TABLE 5 (*Continued*)

p	7		8		9		10		*p*
n	5%	1%	5%	1%	5%	1%	5%	1%	*n*
15	50.70	59.38	64.94						15
16	49.90	58.41	63.66	73.39					16
17	49.221	57.60	62.60	72.14					17
18	48.645	56.911	61.71	71.08	76.86				18
19	48.149	56.318	60.95	70.19	75.72	86.11			19
20	47.716	55.801	60.290	69.41	74.75	84.97	91.28		20
21	47.336	55.348	59.714	68.733	73.90	83.99	90.06	101.30	21
22	46.9982	54.947	59.206	68.138	73.16	83.13	89.01	100.08	22
23	46.6971	54.5889	58.754	67.609	72.502	82.37	88.08	99.02	23
24	46.4266	54.2678	58.350	67.137	71.918	81.697	87.25	98.08	24
25	46.1823	53.9780	57.986	66.712	71.394	81.092	86.52	97.23	25
26	45.9605	53.7151	57.6573	66.328	70.922	80.548	85.855	96.48	26
27	45.7582	53.4756	57.3580	65.9787	70.494	80.054	85.258	95.799	27
28	45.5730	53.2563	57.0847	65.6601	70.104	79.605	84.716	95.181	28
29	45.4027	53.0549	56.8340	65.3681	69.747	79.195	84.221	94.618	29
30	45.2456	52.8693	56.6032	65.0996	69.4190	78.818	83.768	94.103	30
35	44.6133	52.1228	55.6790	64.0253	68.1134	77.3197	81.9731	92.064	35
40	44.1575	51.5856	55.0172	63.2576	67.1852	76.2565	80.7067	90.6287	40
45	43.8132	51.1804	54.5197	62.6812	66.4910	75.4625	79.7646	89.5624	45
50	43.5440	50.8637	54.1321	62.2325	65.9521	74.8466	79.0361	88.7388	50
60	43.1499	50.4008	53.5669	61.5790	65.1694	73.9533	77.9823	87.5489	60
120	42.2125	49.3019	52.2325	60.0394	63.3356	71.8646	75.5328	84.7885	120
∞	41.3371	48.2782	50.9985	58.6192	61.6562	69.9568	73.3115	82.2921	∞

TABLE 6

Test of Hypothesis $\Sigma = \Sigma_0$, $\mu = \mu_0$: Correction Factors for $-2\rho \log \lambda_2$

p	2		3		4		5		6		p
N	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	N
4	1.08	1.11									4
5	1.044	1.055	1.13	1.16							5
6	1.0271	1.0336	1.069	1.085	1.16						6
7	1.0185	1.0228	1.044	1.053	1.09	1.11					7
8	1.0134	1.0165	1.0305	1.0367	1.059	1.070	1.11	1.13			8
9	1.0102	1.0125	1.0224	1.0269	1.0416	1.049	1.072	1.08	1.12	1.14	9
10	1.0080	1.0098	1.0172	1.0205	1.0311	1.036	1.052	1.060	1.08	1.10	10
11	1.0064	1.0079	1.0137	1.0162	1.0242	1.028	1.039	1.045	1.061	1.07	11
12	1.0053	1.0065	1.0111	1.0132	1.0194	1.0225	1.0309	1.035	1.047	1.053	12
13	1.0044	1.0054	1.0092	1.0109	1.0159	1.0183	1.0250	1.028	1.037	1.042	13
14	1.0038	1.0046	1.0077	1.0092	1.0132	1.0153	1.0206	1.0234	1.0304	1.034	14
15	1.0032	1.0040	1.0066	1.0078	1.0112	1.0129	1.0173	1.0196	1.0253	1.028	15
16	1.0028	1.0034	1.0057	1.0067	1.0096	1.0111	1.0147	1.0167	1.0214	1.0239	16
17	1.0025	1.0030	1.0050	1.0059	1.0083	1.0096	1.0127	1.0144	1.0183	1.0204	17
18	1.0022	1.0027	1.0044	1.0052	1.0073	1.0084	1.0111	1.0125	1.0159	1.0177	18
19	1.0019	1.0024	1.0039	1.0046	1.0065	1.0074	1.0098	1.0110	1.0139	1.0154	19
20	1.0017	1.0021	1.0035	1.0041	1.0058	1.0066	1.0086	1.0097	1.0123	1.0136	20
21	1.0016	1.0019	1.0031	1.0037	1.0052	1.0059	1.0077	1.0087	1.0109	1.0121	21
22	1.0014	1.0017	1.0028	1.0033	1.0046	1.0053	1.0069	1.0078	1.0098	1.0108	22
23	1.0013	1.0016	1.0026	1.0030	1.0042	1.0048	1.0063	1.0070	1.0088	1.0098	23
24	1.0012	1.0015	1.0023	1.0027	1.0038	1.0044	1.0057	1.0064	1.0080	1.0088	24
25	1.0011	1.0013	1.0021	1.0025	1.0035	1.0040	1.0052	1.0058	1.0072	1.0080	25
30	1.0007	1.0009	1.0014	1.0017	1.0023	1.0027	1.0035	1.0039	1.0048	1.0053	30
35	1.0005	1.0007	1.0010	1.0012	1.0017	1.0019	1.0025	1.0028	1.0034	1.0038	35
40	1.0004	1.0005	1.0008	1.0009	1.0013	1.0014	1.0018	1.0021	1.0025	1.0028	40
45	1.0003	1.0004	1.0006	1.0007	1.0010	1.0011	1.0014	1.0016	1.0020	1.0022	45
50	1.0003	1.0003	1.0005	1.0006	1.0008	1.0009	1.0011	1.0013	1.0016	1.0017	50
60	1.0002	1.0002	1.0003	1.0004	1.0005	1.0006	1.0008	1.0009	1.0011	1.0012	60
80	1.0001	1.0001	1.0002	1.0002	1.0003	1.0003	1.0004	1.0005	1.0006	1.0006	80
100	1.0001	1.0001	1.0001	1.0001	1.0002	1.0002	1.0003	1.0003	1.0004	1.0004	100
120	1.0000	1.0001	1.0001	1.0001	1.0001	1.0001	1.0002	1.0002	1.0003	1.0003	120
∞	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	∞
χ^2_f	11.0705	15.0863	16.9190	21.6660	23.6848	29.1413	31.4104	37.5662	40.1133	46.9630	χ^2_f

TABLE 6 (*Continued*)

p	7		8		9		10		*p*
N	5%	1%	5%	1%	5%	1%	5%	1%	**N**
10	1.13								10
11	1.09	1.10	1.14						11
12	1.07	1.08	1.10	1.11					12
13	1.054	1.060	1.08	1.09	1.11	1.12			13
14	1.043	1.048	1.061	1.07	1.08	1.09	1.12	1.13	14
15	1.036	1.040	1.049	1.054	1.07	1.07	1.09	1.10	15
16	1.0299	1.033	1.041	1.045	1.055	1.060	1.07	1.08	16
17	1.0254	1.028	1.034	1.038	1.046	1.050	1.060	1.07	17
18	1.0219	1.0241	1.0294	1.032	1.039	1.042	1.050	1.054	18
19	1.0190	1.0210	1.0254	1.028	1.033	1.036	1.043	1.046	19
20	1.0167	1.0184	1.0222	1.0242	1.0289	1.031	1.037	1.040	20
21	1.0148	1.0163	1.0196	1.0213	1.0254	1.0274	1.032	1.035	21
22	1.0132	1.0145	1.0174	1.0189	1.0224	1.0243	1.0285	1.031	22
23	1.0119	1.0130	1.0156	1.0169	1.0200	1.0216	1.0253	1.0272	23
24	1.0107	1.0118	1.0140	1.0152	1.0180	1.0194	1.0226	1.0243	24
25	1.0097	1.0107	1.0127	1.0138	1.0162	1.0175	1.0203	1.0218	25
26	1.0089	1.0097	1.0115	1.0125	1.0147	1.0158	1.0184	1.0197	26
27	1.0081	1.0089	1.0105	1.0114	1.0134	1.0144	1.0167	1.0179	27
28	1.0075	1.0082	1.0097	1.0105	1.0123	1.0132	1.0153	1.0164	28
29	1.0069	1.0075	1.0089	1.0097	1.0113	1.0121	1.0140	1.0150	29
30	1.0064	1.0070	1.0082	1.0089	1.0104	1.0112	1.0129	1.0138	30
35	1.0045	1.0049	1.0058	1.0062	1.0072	1.0078	1.0089	1.0095	35
40	1.0033	1.0036	1.0043	1.0046	1.0053	1.0057	1.0065	1.0070	40
45	1.0026	1.0028	1.0033	1.0036	1.0041	1.0044	1.0050	1.0053	45
50	1.0021	1.0022	1.0026	1.0028	1.0032	1.0035	1.0040	1.0042	50
60	1.0014	1.0015	1.0018	1.0019	1.0022	1.0023	1.0026	1.0028	60
120	1.0003	1.0004	1.0004	1.0004	1.0005	1.0005	1.0006	1.0006	120
∞	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	∞
χ_f^2	49.8019	57.3421	60.4809	68.7095	72.1532	81.0688	84.8207	94.4221	χ_f^2

When $p = 2$ all α_i vanish and $-2(1-n^{-1})\log W$ is distributed as χ^2_2. Table 4 gives correction factors for $-2\rho\log W$ for p in the range 3(1)10.

The hypothesis that a sample of N observations is drawn from a p-variate normal population with a given covariance matrix $\mathbf{\Sigma}_0$ may be tested using the following modification of the likelihood ratio statistic (Korin 1968):

$$\lambda_1 = e^{\frac{1}{2}pn} \,|\, \mathbf{S}\mathbf{\Sigma}_0^{-1} \,|^{\frac{1}{2}n} \exp(-\tfrac{1}{2}n \operatorname{trace} \mathbf{S}\mathbf{\Sigma}_0^{-1}), \qquad (n = N-1),$$

where $\mathbf{S}$ is the sample covariance matrix on n degrees of freedom. The moments of λ_1 are not of Box's form, but the c.g.f. of $-2\rho\log\lambda_1$ may be expanded in a series (1), with

$$f = \tfrac{1}{2}p(p+1), \qquad \omega_r = \alpha_r/(n\rho)^r,$$

$$\alpha_r = \frac{(-1)^r 2^{r+1}}{r(r+1)(r+2)} \sum_{s=1}^{r+1} \binom{r+2}{s+1} \beta_1^{r+1-s} \delta_s,$$

$$\beta_1 = \tfrac{1}{2}n(1-\rho).$$

Thus $\omega_1 = 0$ when

$$\beta_1 = \frac{(2p^2+3p-1)}{12(p+1)}, \qquad \rho = 1 - \frac{(2p^2+3p-1)}{6n(p+1)},$$

and in this case

$$\alpha_2 = p(2p^4+6p^3+p^2-12p-13)/288(p+1),$$

$$\alpha_3 = p(2p^6+9p^5-3p^4-45p^3-39p^2-18p-34)/3240(p+1)^2,$$

$$\alpha_4 = p(16p^8+96p^7+100p^6-444p^5-1197p^4-228p^3+2518p^2$$
$$+3048p+955)/34560(p+1)^3.$$

The correction factor ρ has been suggested by Bartlett and Kullback (see Korin 1968). Comparisons indicated that the 5% percentiles given by Korin for $p = 4$ should appear with their degrees of freedom reduced by one. Table 5 gives percentiles for $-2\log\lambda_1$ with $p = 2(1)10$.

If it is required to test not only that $\mathbf{\Sigma} = \mathbf{\Sigma}_0$ but also that the population mean vector $\boldsymbol{\mu} = \boldsymbol{\mu}_0$, then a suitable statistic is

$$\lambda_2 = (e/N)^{\frac{1}{2}pN} \,|\, \mathbf{A}\mathbf{\Sigma}_0^{-1} \,|^{\frac{1}{2}N} \exp\{-\tfrac{1}{2}\operatorname{trace} \mathbf{A}\mathbf{\Sigma}_0^{-1} + N(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\mathbf{\Sigma}_0^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)\},$$

where $\bar{\mathbf{x}}$ is the sample mean (Anderson 1958, Section 10.9), $\mathbf{A}=n\mathbf{S}$. In this case,

$$f = \tfrac{1}{2}p(p+3), \qquad \omega_r = \alpha_r/(N\rho)^r,$$

$$\alpha_r = \frac{(-1)^r 2^{r+1}}{r(r+1)(r+2)} \sum_{s=1}^{r+1} \binom{r+2}{s+1} \beta_2^{r+1-s} \{\delta_s + 2^{-s-1}p(s+1)\},$$

$$\beta_2 = \tfrac{1}{2}[N(1-\rho)-1].$$

Taking

$$\beta_2 = \frac{(2p^2+3p-7)}{12(p+3)}, \qquad \rho = 1-\frac{(2p^2+9p+11)}{6N(p+3)},$$

we obtain $\omega_1 = 0$ and

$$\alpha_2 = p(2p^4+18p^3+49p^2+36p-13)/288(p+3),$$

$$\alpha_3 = p(2p^6+27p^5+123p^4+135p^3-489p^2-1404p-1006)/3240(p+3)^2,$$

$$\alpha_4 = p(16p^8+288p^7+2020p^6+6732p^5+9363p^4-1404p^3-12458p^2$$
$$+3816p+15355)/34560(p+3)^3.$$

Table 6 gives correction factors for $-2\rho \log \lambda_2$ for $p = 2(1)10$.

III. References

Anderson, T. W. (1958).—"An Introduction to Multivariate Statistical Analysis." (Wiley: New York.)

Box, G. E. P. (1949).—A general distribution theory for a class of likelihood criteria. *Biometrika* **36**, 317–46.

Cornish, E. A., and Fisher, R. A. (1937).—Moments and cumulants in the specification of distributions. *Revue Inst. int. Statist.* **4**, 307–20.

Davis, A. W. (1970).—On the null distribution of the sum of the roots of a multivariate beta distribution. *Ann. math. Statist.* **41**, 1557–62.

Davis, A. W. (1971).—Percentile approximations for a class of likelihood ratio criteria. *Biometrika* (in press).

Fisher, R. A., and Cornish, E. A. (1960).—The percentile points of distributions having known cumulants. *Technometrics* **2**, 209–26.

Hill, G. W., and Davis, A. W. (1968).—Generalized asymptotic expansions of Cornish–Fisher type. *Ann. math. Statist.* **39**, 1264–73.

Korin, B. P. (1968).—On the distribution of a statistic used for testing a covariance matrix. *Biometrika* **55**, 171–8.

Korin, B. P. (1969).—On testing the equality of k covariance matrices. *Biometrika* **56**, 216–18.

Neyman, J., and Pearson, E. S. (1928).—On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika* **20A**, Pt. I 175–240, Pt. II 263–94.

Pillai, K. C. S., and Gupta, A. K. (1969).—On the exact distribution of Wilks's criterion. *Biometrika* **56**, 109–18.

Rao, C. R. (1948).—Tests of significance in multivariate analysis. *Biometrika* **35**, 58–79.

Schatzoff, M. (1966).—Exact distributions of Wilks's likelihood ratio criterion. *Biometrika* **53**, 347–58.

Wallace, D. L. (1958).—Asymptotic approximations to distributions. *Ann. math. Statist.* **29**, 635–54.

On the k-Sample Behrens–Fisher Distribution

A. W. Davis and A. J. Scott

DIVISION OF MATHEMATICAL STATISTICS TECHNICAL PAPER NO. 33
COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANIZATION, AUSTRALIA 1971

On the k-Sample Behrens–Fisher Distribution

By A. W. Davis and A. J. Scott

Division of Mathematical Statistics Technical Paper No. 33

Commonwealth Scientific and Industrial
Research Organization, Australia
1971

Printed by CSIRO, Melbourne

ON THE k-SAMPLE BEHRENS–FISHER DISTRIBUTION*

By A. W. DAVIS† and A. J. SCOTT‡

Abstract

Fisher's expansion for the percentage points of the two-sample Behrens–Fisher distribution is extended to the k-sample case. The expansion is used to construct tables of a few common percentage points including upper 5% points in the two-sample case, and a simple approximation is suggested. An expansion is also given for the probability content, in the frequency sense, of intervals based on the Behrens-Fisher distribution.

I. INTRODUCTION

The construction of interval estimates for a linear combination of means, say $\eta = \sum_1^k \lambda_i \mu_i$, is important in many applications. Suppose that m_i independent observations are taken from the ith group and the resulting sample mean and variance are $\bar{x}_i$ and s_i^2 ($i = 1, ..., k$). The standard assumption is that the observations are normally distributed about mean μ_i with variance σ_i^2. If it is known that $\sigma_i^2 = \sigma^2$ ($i = 1, ..., k$) then all common systems of inference (frequentist, Bayesian, fiducial) lead to essentially the same intervals, but if the variances cannot be assumed equal the Bayesian and fiducial solution differs from the frequentist one (Kendall and Stuart 1961, p. 146, give a comprehensive account for $k = 2$). In the usual Bayesian (or fiducial) approach, with little *a priori* knowledge about the value of σ_i^2, the posterior distribution of

$$d = \sum_1^k \lambda_i(\mu_i - \bar{x}_i)/(\sum_1^k \lambda_i^2 s_i^2/m_i)^{1/2} \tag{1}$$

is equivalent to the distribution of

$$\sum_1^k \xi_i t_{n_i}, \quad (\sum_1^k \xi_i^2 = 1) ,$$

where $\xi_1, ..., \xi_k$ are known constants $[\xi_i = (\lambda_i s_i/m_i^{1/2})/(\sum_1^k \lambda_i^2 s_i^2/m_i)^{1/2}]$, t_{n_i} has a Student t-distribution with $n_i = m_i - 1$ degrees of freedom, and $t_{n_1}, ..., t_{n_k}$ are mutually independent (Lindley 1965). Interval estimates for $\eta = \sum_1^k \lambda_i \mu_i$ follow directly from the percentage points of $\sum_1^k \xi_i t_{n_i}$.

For $k = 2$, tables of the percentage points have been constructed by Sukhatme (1938), and Sukhatme *et al.* (1951) for a few values of n_1, n_2, and $\sin^{-1}\xi_1$, and by Fisher and Healy (1956) for small odd degrees of freedom. Both sets of tables are

* This research was carried out under National Institutes of Health Grant GM–12868, while both authors were on leave at the Department of Biostatistics, University of North Carolina.

† Division of Mathematical Statistics, CSIRO, Private Bag No. 1, Glen Osmond, S.A. 5064.

‡ London School of Economics and Political Science, University of London, Aldwych, London, W.C.2.

given in Fisher and Yates (1963). Fisher (1941) has given an asymptotic expansion for the percentage points of d and this can be used for larger n_1, n_2 when the parameters lie outside the range covered by the tables. There are no tables for larger values of k. Even if tables could be constructed, they would be complicated and difficult to use since there are $(2k-1)$ parameters involved. In these circumstances it is natural to search for a reasonably simple approximation, and several have been proposed in the literature (Quenouille 1958, p. 139; Patil 1965, 1969; Smith and Scott, unpublished data).

The difficulty is that at present there is no way of judging the accuracy of the approximations for $k>2$. To overcome this difficulty, we have extended Fisher's expansion for the percentage points of d to include arbitrary values of k. Fisher's approach, involving convolutions of t-variables, is difficult to extend directly, and we have adopted a different approach based on the cumulant generating function. This is considerably simpler even for $k = 2$, and has the additional advantage that direct use can be made of the tabulated Cornish–Fisher polynomials (Cornish and Fisher 1937; Fisher and Cornish 1960; Hill and Davis 1968). The details are given in Section II. The expansion has been used to construct tables of the upper 5% point of the distribution of d for $k = 2$, and of the upper 2.5% and upper 0.5% points of the distribution of d for $k = 3, 4, 5, 10$ in the important special case of equal degrees of freedom. To get the accuracy required for the tables, it was found necessary to take the expansion to terms of $O(n_i^{-5})$, one term further than Fisher. A simple modification of the expansion, given in Section III, suggests a rather simple approximation closely related to the approximation given by Welch (1947, 1949) in his attack on a frequentist solution of the problem. The tables given in Section II have been used to compare this approximation with existing approximations in Davis and Scott (unpublished data).

It has often been noted that the Behrens–Fisher intervals for $k = 2$ seem to be wider than Welch's approximate confidence intervals. An expansion for the actual confidence level, in the conventional long-run frequency sense, is given in Section IV. This suggests that the Behrens–Fisher intervals will be conservative in this sense for all values of k, certainly when the degrees of freedom are large.

II. The Cornish–Fisher Expansion of d_θ

The cumulant generating function (c.g.f.) $K_n(\phi) = \log \mathcal{E} \exp(i\phi t_n)$ of Student's t on n degrees of freedom may be expanded in an asymptotic series in n^{-1}, the terms to fifth order being

$$\begin{aligned} K_n(\phi) \sim & -\tfrac{1}{2}\phi^2 + (1/4n)(\phi^4 - 2\phi^2) - (1/3n^2)(\phi^6 - 6\phi^4 + 6\phi^2) \\ & + (1/24n^3)(15\phi^8 - 128\phi^6 + 264\phi^4 - 96\phi^2) \\ & + (1/15n^4)(-21\phi^{10} + 240\phi^8 - 800\phi^6 + 780\phi^4 - 120\phi^2) \\ & + (1/10n^5)(35\phi^{12} - 512\phi^{10} + 2450\phi^8 - 4320\phi^6 + 2280\phi^4 - 160\phi^2). \end{aligned} \quad (2)$$

Hence, if we define

$$\alpha_i = \Sigma_{j=1}^{k} \xi_j^{2i}/n_j, \quad \beta_i = \Sigma_{j=1}^{k} \xi_j^{2i}/n_j^2, \quad \gamma_i = \Sigma_{j=1}^{k} \xi_j^{2i}/n_j^3, \quad (3)$$

$$\delta_i = \Sigma_{j=1}^{k} \xi_j^{2i}/n_j^4 , \quad \epsilon_i = \Sigma_{j=1}^{k} \xi_j^{2i}/n_j^5 , \quad (i = 1, 2, ...) , \tag{3}$$

the c.g.f. $\Lambda(\phi)$ of $d = \Sigma_{i=1}^{k} \xi_i t_{n_i}$ may be formally expressed as

$$\Lambda(\phi) = -\tfrac{1}{2}\phi^2 + \Sigma_{j=1}^{\infty} \rho_{2j}(i\phi)^{2j} , \tag{4}$$

where, to the fifth order,

$$\begin{aligned}
\rho_2 &= \alpha_1 + 2\beta_1 + 4\gamma_1 + 8\delta_1 + 16\epsilon_1 &&= O(n_i^{-1}) , \\
\rho_4 &= \tfrac{1}{4}\alpha_2 + 2\beta_2 + 11\gamma_2 + 52\delta_2 + 228\epsilon_2 &&= O(n_i^{-1}) , \\
\rho_6 &= \tfrac{1}{3}(\beta_3 + 16\gamma_3 + 160\delta_3 + 1296\epsilon_3) &&= O(n_i^{-2}) , \\
\rho_8 &= \tfrac{5}{8}\gamma_4 + 16\delta_4 + 245\epsilon_4 &&= O(n_i^{-3}) , \\
\rho_{10} &= \tfrac{1}{5}(7\delta_5 + 256\epsilon_5) &&= O(n_i^{-4}) , \\
\rho_{12} &= \tfrac{7}{2}\epsilon_6 &&= O(n_i^{-5}) .
\end{aligned} \tag{5}$$

The series (4) is thus of the general form considered by Fisher and Cornish (1937) and Cornish and Fisher (1960). Hence, the expansion of the $100(1-\theta)\%$ point d_θ of d in terms of the corresponding unit normal deviate u_θ may be represented as

$$d_\theta \sim u_\theta + \Sigma_\pi \rho_\pi P_\pi(u_\theta) \tag{6}$$

(Hill and Davis 1968), where the summation is extended over all partitions

$$\pi = [s_1^{t_1}, ..., s_l^{t_l}]$$

of all positive integers, the $P_\pi(u)$ are polynomials in u, and

$$\rho_\pi = \rho_{s_1}^{t_1} \; .... \; \rho_{s_l}^{t_l} / t_1! \; .... \; t_l! \tag{7}$$

Fisher and Cornish tabulated sufficient P_π to give the expansion (6) to order n_i^{-3}. The additional polynomials required for the fifth-order expansion may be constructed by the methods given in Hill and Davis (1968, Sections 4 and 5), and a listing is given in Table 1. We note that, in virtue of (4), only partitions into even integers occur in (6), and a further reduction is effected by the relationship

$$P_{\pi,2}(u) = -(m-1)P_\pi(u) , \tag{8}$$

where π is any partition of the integer m. Thus it is necessary to consider only partitions into integers not less than 4. In particular, $P_m(u)$ is Hermite's polynomial of order $(m-1)$.

A valid asymptotic expansion is obtained when terms of like order are grouped. We finally obtain

$$d_\theta \sim u_\theta + \Sigma_{r=1}^{\infty} Q_r(u_\theta) , \tag{9}$$

where

$$Q_r(u) = O(n_i^{-r})$$

and

$$
\begin{aligned}
Q_1 &= \alpha_1 P_2 + \tfrac{1}{4}\alpha_2 P_4 \,, \\
Q_2 &= \tfrac{1}{2}(4\beta_1 - \alpha_1^2)P_2 + \tfrac{1}{4}(8\beta_2 - 3\alpha_1\alpha_2)P_4 + \tfrac{1}{3}\beta_3 P_6 + \tfrac{1}{32}\alpha_2^2 P_{4^2} \,, \\
Q_3 &= \tfrac{1}{2}(8\gamma_1 - 4\alpha_1\beta_1 + \alpha_1^3)P_2 + \tfrac{1}{8}(88\gamma_2 - 12\alpha_2\beta_1 - 48\alpha_1\beta_2 + 15\alpha_1^2\alpha_2)P_4 \\
&\quad + \tfrac{1}{3}(16\gamma_3 - 5\alpha_1\beta_3)P_6 + \tfrac{5}{8}\gamma_4 P_8 + \tfrac{1}{32}(16\alpha_2\beta_2 - 7\alpha_1\alpha_2^2)P_{4^2} + \tfrac{1}{12}\alpha_2\beta_3 P_{6,4} \\
&\quad + \tfrac{1}{384}\alpha_2^3 P_{4^3} \,, \\
Q_4 &= \tfrac{1}{8}(64\delta_1 - 32\alpha_1\gamma_1 - 16\beta_1^2 + 24\alpha_1^2\beta_1 - 5\alpha_1^4)P_2 \\
&\quad + \tfrac{1}{8}(416\delta_2 - 24\alpha_2\gamma_1 - 96\beta_1\beta_2 - 264\alpha_1\gamma_2 + 60\alpha_1\alpha_2\beta_1 + 120\alpha_1^2\beta_2 - 35\alpha_1^3\alpha_2)P_4 \\
&\quad + \tfrac{1}{6}(320\delta_3 - 20\beta_1\beta_3 - 160\alpha_1\gamma_3 + 35\alpha_1^2\beta_3)P_6 + \tfrac{1}{8}(128\delta_4 - 35\alpha_1\gamma_4)P_8 \\
&\quad + \tfrac{7}{5}\delta_5 P_{10} + \tfrac{1}{64}(176\alpha_2\gamma_2 + 128\beta_2^2 - 28\alpha_2^2\beta_1 - 224\alpha_1\alpha_2\beta_2 + 63\alpha_1^2\alpha_2^2)P_{4^2} \\
&\quad + \tfrac{1}{12}(16\alpha_2\gamma_3 + 8\beta_2\beta_3 - 9\alpha_1\alpha_2\beta_3)P_{6,4} + \tfrac{1}{18}\beta_3^2 P_{6^2} + \tfrac{5}{32}\alpha_2\gamma_4 P_{8,4} \\
&\quad + \tfrac{1}{96}\alpha_2^2\beta_3 P_{6,4^2} + \tfrac{1}{384}(24\alpha_2^2\beta_2 - 11\alpha_1\alpha_2^3)P_{4^3} + \tfrac{1}{6144}\alpha_2^4 P_{4^4} \,, \\
Q_5 &= [16\epsilon_1 - 8(\alpha_1\delta_1 + \beta_1\gamma_1) + 6(\alpha_1^2\gamma_1 + \alpha_1\beta_1^2) - 5\alpha_1^3\beta_1 + \tfrac{7}{8}\alpha_1^5]P_2 \\
&\quad + [228\epsilon_2 - 6(\alpha_2\delta_1 + 4\beta_2\gamma_1 + 11\beta_1\gamma_2 + 26\alpha_1\delta_2) \qquad\qquad (10) \\
&\quad + \tfrac{15}{2}(11\alpha_1^2\gamma_2 + 8\alpha_1\beta_1\beta_2 + 2\alpha_1\alpha_2\gamma_1 + \alpha_2\beta_1^2) - \tfrac{35}{4}(4\alpha_1^3\beta_2 + 3\alpha_1^2\alpha_2\beta_1) \\
&\quad + \tfrac{315}{12}\alpha_1^4\alpha_2]P_4 + [432\epsilon_3 - \tfrac{20}{3}(40\alpha_1\delta_3 + 8\beta_1\gamma_3 + \beta_3\gamma_1) \\
&\quad + \tfrac{70}{3}(4\alpha_1^2\gamma_3 + \alpha_1\beta_1\beta_3) - \tfrac{35}{2}\alpha_1^3\beta_3]P_6 + [245\epsilon_4 - \tfrac{7}{4}(64\alpha_1\delta_4 + 5\beta_1\gamma_4) \\
&\quad + \tfrac{315}{16}\alpha_1^2\gamma_4]P_8 + \tfrac{1}{5}(256\epsilon_5 - 63\alpha_1\delta_5)P_{10} + \tfrac{7}{2}\epsilon_6 P_{12} + [13\alpha_2\delta_2 + 22\beta_2\gamma_2 \\
&\quad - \tfrac{7}{8}(16\alpha_1\beta_2^2 + 22\alpha_1\alpha_2\gamma_2 + 8\alpha_2\beta_1\beta_2 + \alpha_2^2\gamma_1) + \tfrac{63}{16}(4\alpha_1^2\alpha_2\beta_2 + \alpha_1\alpha_2^2\beta_1) \\
&\quad - \tfrac{231}{64}\alpha_1^3\alpha_2^2]P_{4^2} + [\tfrac{40}{3}\alpha_2\delta_3 + \tfrac{32}{3}\beta_2\gamma_3 + \tfrac{11}{3}\beta_3\gamma_2 \\
&\quad - 3(4\alpha_1\alpha_2\gamma_3 + \tfrac{1}{2}\alpha_2\beta_1\beta_3 + 2\alpha_1\beta_2\beta_3) + \tfrac{33}{8}\alpha_1^2\alpha_2\beta_3]P_{6,4} \\
&\quad + \tfrac{1}{32}(40\beta_2\gamma_4 + 128\alpha_2\delta_4 - 55\alpha_1\alpha_2\gamma_4)P_{8,4} + \tfrac{1}{18}(32\beta_3\gamma_3 - 11\alpha_1\beta_3^2)P_{6^2} \\
&\quad + [\tfrac{1}{2}\alpha_2\beta_2^2 + \tfrac{11}{32}\alpha_2^2\gamma_2 - \tfrac{11}{192}(\alpha_2^3\beta_1 + 12\alpha_1\alpha_2^2\beta_2) + \tfrac{143}{768}\alpha_1^2\alpha_2^3]P_{4^3} \\
&\quad + \tfrac{7}{20}\alpha_2\delta_5 P_{10,4} + \tfrac{5}{24}\beta_3\gamma_4 P_{8,6} + \tfrac{1}{96}[16(\alpha_2^2\gamma_3 + \alpha_2\beta_2\beta_3) - 13\alpha_1\alpha_2^2\beta_3]P_{6,4^2} \\
&\quad + \tfrac{5}{256}\alpha_2^2\gamma_4 P_{8,4^2} + \tfrac{1}{72}\alpha_2\beta_3^2 P_{6^2,4} + (\tfrac{1}{192}\alpha_2^3\beta_2 - \tfrac{5}{2048}\alpha_1\alpha_2^4)P_{4^4} \\
&\quad + \tfrac{1}{1152}\alpha_2^3\beta_3 P_{6,4^3} + \tfrac{1}{122{,}880}\alpha_2^5 P_{4^5} \,.
\end{aligned}
$$

To the fourth order, this reduces to Fisher's result (1941, Table 2) in the two-sample case, when $\xi_1 = \sin\psi$, $\xi_2 = \cos\psi$. Sukhatme *et al.* (1951) have taken the two-sample expansion to fifth order but have not published their result. However, our numerical results are in essential agreement with theirs.

TABLE 1

CORNISH–FISHER POLYNOMIALS

π	$P_\pi(u)$
2	u
4	$u^3 - 3u$
6	$u^5 - 10u^3 + 15u$
8	$u^7 - 21u^5 + 105u^3 - 105u$
10	$u^9 - 36u^7 + 378u^5 - 1260u^3 + 945u$
12	$u^{11} - 55u^9 + 990u^7 - 6930u^5 + 17{,}325u^3 - 10{,}395u$
4^2	$-3(3u^5 - 24u^3 + 29u)$
6, 4	$-15(u^7 - 17u^5 + 69u^3 - 57u)$
8, 4	$-21(u^9 - 30u^7 + 262u^5 - 730u^3 + 465u)$
6^2	$-5(5u^9 - 140u^7 + 1170u^5 - 3180u^3 + 1989u)$
4^3	$27(9u^7 - 131u^5 + 451u^3 - 321u)$
10, 4	$-9(3u^{11} - 141u^9 + 2162u^7 - 12{,}894u^5 + 27{,}615u^3 - 14{,}385u)$
8, 6	$-35(u^{11} - 43u^9 + 618u^7 - 3534u^5 + 7377u^3 - 3771u)$
$6, 4^2$	$15(33u^9 - 834u^7 + 6158u^5 - 14{,}654u^3 + 8097u)$
$8, 4^2$	$3(273u^{11} - 10{,}941u^9 + 143{,}444u^7 - 736{,}176u^5 + 1{,}370{,}355u^3 - 628{,}635u)$
$6^2, 4$	$15(65u^{11} - 2455u^9 + 30{,}830u^7 - 153{,}930u^5 + 281{,}589u^3 - 127{,}467u)$
4^4	$-27(429u^9 - 9624u^7 + 62{,}478u^5 - 131{,}088u^3 + 64{,}573u)$
$6, 4^3$	$-135(195u^{11} - 6763u^9 + 76{,}912u^7 - 345{,}036u^5 + 566{,}949u^3 - 232{,}209u)$
4^5	$81(9945u^{11} - 313{,}107u^9 + 3{,}205{,}782u^7 - 12{,}938{,}186u^5 + 19{,}203{,}825u^3 - 7{,}160{,}835u)$

TABLE 2

95% POINTS OF THE DISTRIBUTION OF d FOR $k = 2$

$\sin^{-1}\xi_1$		0°	15°	30°	45°	60°	75°	90°
n_1	n_2							
6	6	1.943	1.949	1.965	1.974	1.965	1.949	1.943
8		1.943	1.941	1.937	1.926	1.900	1.871	1.859^5
12		1.943	1.934^5	1.913	1.882	1.840	1.799	1.782
24		1.943	1.929	1.893	1.843	1.785	1.733	1.711
∞		1.943	1.925	1.875^5	1.808	1.734	1.671	1.645
8	8	1.860	1.863	1.873	1.878	1.873	1.863	1.860
12		1.860	1.856	1.849	1.835	1.813	1.791	1.782
24		1.860	1.851	1.828	1.796	1.758	1.724	1.711
∞		1.860	1.846	1.810^5	1.761	1.707	1.663	1.645
12	12	1.782	1.784	1.789	1.792	1.789	1.784	1.782
24		1.782	1.778^5	1.768	1.753	1.734	1.717	1.711
∞		1.782	1.774	1.751	1.718	1.684	1.656	1.645
24	24	1.711	1.712	1.713^5	1.714^5	1.713^5	1.712	1.711
∞		1.711	1.707	1.695^5	1.680	1.663	1.650	1.645

The expansion has been used to obtain the values given in Tables 2 and 3, Table 2 gives 95% points of d for $k = 2$ over the range of parameter values used by

TABLE 3

CRITICAL VALUES OF THE BEHRENS–FISHER STATISTIC FOR $k = 3, 4, 5, 10$ AND $n_i = n$ $(i = 1, \ldots, k)$

				Probability = 0.95				Probability = 0.99			
				$n = 6$	8	12	24	$n = 6$	8	12	24
$k = 3$											
ξ_1^2	ξ_2^2										
0.8	0.1			2.431	2.292	2.168	2.058	3.559	3.246	2.984	2.763
0.7	0.2			2.431	2.289	2.166	2.057	3.511	3.209	2.959	2.750
0.6	0.3			2.430	2.288	2.164	2.056	3.480	3.185	2.943	2.742
0.6	0.2			2.430	2.287	2.163	2.055	3.468	3.176	2.937	2.740
0.5	0.4			2.430	2.288	2.164	2.055	3.464	3.173	2.935	2.738
0.5	0.3			2.429	2.286	2.162	2.054	3.445	3.158	2.925	2.734
0.4	0.4			2.429	2.286	2.162	2.054	3.437	3.152	2.921	2.732
0.4	0.3			2.429	2.286	2.162	2.054	3.430	3.146	2.916	2.730
$k = 4$											
ξ_1^2	ξ_2^2	ξ_3^2									
0.7	0.1	0.1		2.427	2.287	2.164	2.056	3.492	3.198	2.953	2.448
0.6	0.2	0.1		2.426	2.285	2.162	2.055	3.450	3.165	2.932	2.738
0.5	0.3	0.1		2.426	2.284	2.161	2.054	3.427	3.147	2.919	2.732
0.5	0.2	0.2		2.426	2.284	2.160	2.053	3.415	3.139	2.914	2.729
0.4	0.4	0.1		2.426	2.284	2.161	2.054	3.419	3.142	2.915	2.730
0.4	0.3	0.2		2.425	2.283	2.160	2.053	3.400	3.127	2.906	2.725
0.4	0.2	0.2		2.425	2.282	2.159	2.053	3.389	3.118	2.901	2.723
0.3	0.3	0.3		2.425	2.283	2.159	2.053	3.393	3.120	2.901	2.723
0.3	0.3	0.2		2.425	2.282	2.159	2.052	3.382	3.112	2.896	2.721
$k = 10$											
ξ_1^2	ξ_2^2	ξ_3^2	ξ_4^2								
0.6	0.1	0.1	0.1	2.423	2.283	2.161	2.054	3.431	3.155	2.926	2.735
0.5	0.2	0.1	0.1	2.423	2.282	2.159	2.053	3.397	3.128	2.908	2.727
0.4	0.3	0.1	0.1	2.423	2.281	2.159	2.053	3.382	3.116	2.900	2.723
0.4	0.2	0.2	0.1	2.423	2.280	2.158	2.052	3.372	3.108	2.895	2.721
0.3	0.3	0.2	0.1	2.423	2.280	2.158	2.052	3.365	3.102	2.890	2.719
0.3	0.2	0.2	0.2	2.423	2.280	2.157	2.052	3.354	3.094	2.886	2.716
0.2	0.2	0.2	0.2	2.423	2.279	2.156	2.051	3.344	3.086	2.881	2.714
$k = 10$											
$\xi_i^2 = 0.1$, $i = 1, \ldots, 10$				2.415	2.272	2.152	2.049	3.261	3.034	2.852	2.702
$k = \infty$				2.401	2.264	2.147	2.047	3.155	2.974	2.822	2.690

Sukhatme in his tabulation of other percentage points, and Table 3 gives 97.5% and 99.5% points for $k = 3, 4, 5, 10$ when the sample sizes are equal.

III. An Approximation to the Posterior Distribution of d

If (9) is now expressed in terms of the standardized deviate d_θ/s, where s^2 is the variance of d,

$$s^2 = \Sigma_{i=1}^{k} \xi_i^2 n_i/(n_i - 2) \sim 1 + 2\alpha_1 + 4\beta_1 + 8\gamma_1 + \ldots , \tag{11}$$

it is found that all products formed from α_1, β_1, γ_1, and δ_1 are removed:

$$d_\theta/su_\theta \sim 1 + \tfrac{1}{4}\alpha_2(u_\theta^2 - 3) + [(2\beta_2 - \alpha_1\alpha_2)(u_\theta^2 - 3) + \tfrac{1}{3}\beta_3(u_\theta^4 - 10u_\theta^2 + 15) - \tfrac{3}{32}\alpha_2^2(3u_\theta^4 - 24u_\theta^2 + 29)] \ldots . \tag{12}$$

Comparing this expansion with the special case of t_n ($k = 1$, $\xi_1 = 1$), for which $\alpha_2 = n^{-1}$, we are led to the following simple approximation to the posterior distribution of d:

$$d \approx st_\nu[\nu/(\nu - 2)]^{1/2} , \tag{13}$$

where

$$\nu = \alpha_2^{-1} = (\Sigma_{j=1}^{k} \xi_j^4/n_j)^{-1} . \tag{14}$$

Welch (1947, 1949) gave a similar approximation for his confidence intervals. The approximation has been found to work particularly well at the 5% level (Davis and Scott, unpublished data).

IV. Asymptotic Expansion of the Confidence Level

It is well known that the posterior probability intervals for $\eta = \Sigma_{i=1}^{k} \lambda_i\mu_i$ based on the Behrens–Fisher distribution appear, at least in the two-sample case, to be wider than the approximate confidence intervals given by Welch. There may therefore be some interest in the actual confidence level of the Behrens–Fisher interval when interpreted in the frequency sense. The required probability is

$$p = P(d < d_\theta) = \mathcal{E}\Phi[H(s_1^2, \ldots, s_k^2, \theta)] , \tag{15}$$

where $\Phi(u)$ is the unit normal cumulative distribution function, and

$$H(s_1^2, \ldots, s_k^2, \theta) = d_\theta[(\Sigma\lambda_i^2 s_i^2/m_i)/(\Sigma\lambda_i^2 \sigma_i^2/m_i)]^{1/2} .$$

The expectation in (15) is taken with respect to the sample variances s_i^2, independently distributed as $\sigma_i^2\chi_{n_i}^2/n_i$ ($i = 1, \ldots, k$). Welch's (1947) method may be applied to develop an asymptotic series for p:

$$p = \Theta\Phi(H) , \tag{16}$$

where Θ is the differential operator

$$\Theta = \mathcal{E}\exp[\Sigma_{i=1}^{k}(s_i^2 - \sigma_i^2)\partial_i] , \quad \partial_i = (\partial/\partial s_i^2)_{s_i^2 = \sigma_i^2} , \quad (i = 1, \ldots, k) . \tag{17}$$

To the second order

$$\Theta = 1 + \Sigma_{i=1}^{k}\sigma_i^4\partial_i^2/n_i + [\tfrac{4}{3}\Sigma\sigma_i^6\partial_i^3/n_i^2 + \tfrac{1}{2}(\Sigma\sigma_i^4\partial_i^2/n_i)^2] + \ldots . \tag{18}$$

We shall denote the values of the α_i, β_i, γ_i, and δ_i at the population values $s_j^2 = \sigma_j^2$ ($j = 1, ..., k$) by the corresponding upper-case Greek letters. Thus

$$p \sim \Phi(u_\theta + [\mathrm{A}_1 u_\theta + \tfrac{1}{4}\mathrm{A}_2(u_\theta^3 - 3u_\theta)] + ...) + \Sigma_{i=1}^{k}\sigma_i^4(\partial_i^2\Phi)/n_i + O(n_i^{-2}) \tag{19}$$

where

$$\begin{aligned}\partial_i^2\Phi = {} & \Phi'(u_\theta)[-\lambda_i^4 u_\theta/4m_i^2(\textstyle\sum_1^k \lambda_i\sigma_i^2/m_i)^2 + O(n_i^{-1})] \\ & + \Phi''(u_\theta)[\lambda_i^2 u_\theta/2m_i \textstyle\sum_1^k \lambda_i\sigma_i^2/m_i)^2 + O(n_i^{-1})] \,.\end{aligned} \tag{20}$$

Hence

$$p \sim (1-\theta) + (2\pi)^{-\frac{1}{2}} u_\theta \exp(-\tfrac{1}{2}u_\theta^2)[(\mathrm{A}_1 - \mathrm{A}_2) + \Sigma_{r=2}^{\infty} S_r(u_\theta)] \,, \tag{21}$$

where $S_r = O(n_i^{-r})$. Since $\mathrm{A}_1 \geqslant \mathrm{A}_2$, the first-order correction term is positive, as expected. The remaining terms to third order are

$$\begin{aligned}S_2(u) = {} & \tfrac{1}{2}u^2(-\mathrm{A}_1^2 + \mathrm{A}_1\mathrm{A}_2 + 3\mathrm{A}_2^2 + 3\mathrm{B}_2 - 6\mathrm{B}_3) \\ & + \tfrac{1}{2}(-\mathrm{A}_1^2 + 6\mathrm{A}_1\mathrm{A}_2 - 12\mathrm{A}_2^2 + 4\mathrm{B}_1 - 17\mathrm{B}_2 + 20\mathrm{B}_3) \,, \\ S_3(u) = {} & \tfrac{1}{6}u^4(\mathrm{A}_1^3 - 9\mathrm{A}_1\mathrm{A}_2^2 - 15\mathrm{A}_2^3 - 9\mathrm{A}_1\mathrm{B}_2 + 16\mathrm{A}_1\mathrm{B}_3 + 44\mathrm{A}_2\mathrm{B}_3 \\ & + 14\Gamma_3 - 42\Gamma_4) + \tfrac{1}{24}u^2(8\mathrm{A}_1^3 - 36\mathrm{A}_1^2\mathrm{A}_2 - 189\mathrm{A}_1\mathrm{A}_2^2 \\ & + 1185\mathrm{A}_2^3 - 48\mathrm{A}_1\mathrm{B}_1 + 24\mathrm{A}_2\mathrm{B}_1 + 48\mathrm{A}_1\mathrm{B}_2 + 828\mathrm{A}_2\mathrm{B}_2 \\ & + 208\mathrm{A}_1\mathrm{B}_3 - 3316\mathrm{A}_2\mathrm{B}_3 + 288\Gamma_2 - 1712\Gamma_3 + 2712\Gamma_4) \\ & + \tfrac{1}{2}(\mathrm{A}_1^3 - 15\mathrm{A}_1^2\mathrm{A}_2 + 84\mathrm{A}_1\mathrm{A}_2^2 - 180\mathrm{A}_2^3 - 4\mathrm{A}_1\mathrm{B}_1 + 12\mathrm{A}_2\mathrm{B}_1 \\ & + 51\mathrm{A}_1\mathrm{B}_2 - 195\mathrm{A}_2\mathrm{B}_2 - 100\mathrm{A}_1\mathrm{B}_3 + 496\mathrm{A}_2\mathrm{B}_3 + 8\Gamma_1 \\ & - 96\Gamma_2 + 334\Gamma_3 - 396\Gamma_4) \,.\end{aligned} \tag{22}$$

As a check on the working, the expansion reduces to $1 - \theta$ when $k = 1$.

TABLE 4

MAXIMUM PROBABILITY CONTENT OF BEHRENS–FISHER INTERVALS IN THE SPECIAL CASE $n_i = n$

k	Probability = 0.95				Probability = 0.99			
	n = 6	8	12	24	n = 6	8	12	24
2	0.968	0.964	0.960	0.955	0.996	0.995	0.993	0.992
3	0.974^5	0.969	0.963	0.956	0.997^5	0.996	0.994	0.992
4	0.977	0.971	0.964	0.957	0.998	0.996	0.994	0.992
5	0.979	0.972	0.965	0.958	0.998	0.997	0.995	0.992
10	0.982	0.974^5	0.967	0.958^5	0.998	0.997	0.995	0.993

If $n_i = n$ ($i = 1, ..., k$) the first-order correction takes on its maximum value,

$$c_1 = (2\pi)^{-1} u_\theta \exp(-\tfrac{1}{2}u_\theta^2)[1 - (1/k)]/n\,,$$

when $\lambda_i^2 \sigma_i^2 / m_i$ = constant. Obviously c_1 increases with k and we would expect intervals based on the Behrens–Fisher distribution to become more conservative as k increases, at least for larger n. This is confirmed by the values in Table 4, which shows the maximum probability content of such intervals as given by the expansion to third order when $n_i = n$. The situation is slightly more complicated if the n_i's are not all equal except for $k = 2$ when the first-order term reduces to a multiple of $(1/n_1 + 1/n_2)$. In general, the maximum probability always lies between the values given in Table 4 for $n = \max(n_i)$ and $n = \min(n_i)$.

V. References

Cornish, E. A., and Fisher, R. A. (1937).–Moments and cumulants in the specification of distributions. *Revue Inst. int. Statist.* **5**, 307–22.

Fisher, R. A. (1941).–The asymptotic approach to Behrens integral, with further tables for the *d* test of significance. *Ann. Eugen.* **11**, 141–72.

Fisher, R. A., and Cornish, E. A. (1960).–The percentile points of distributions having known cumulants. *Technometrics* **2**, 209–25.

Fisher, R. A., and Healy, M. J. R. (1956).–New tables of Behrens' test of significance. *Jl R. statist. Soc.* B **18**, 212–16.

Fisher, R. A., and Yates, F. (1963).–"Statistical Tables." 6th Ed. (Oliver and Boyd: London.)

Hill, G. W., and Davis, A. W. (1968).–Generalized asymptotic expansions of Cornish–Fisher type. *Ann. math. Statist.* **39**, 1264–73.

Kendall, M. G., and Stuart, A. (1961).–"The Advanced Theory of Statistics." Vol. 2. (Griffin: London.)

Lindley, D. V. (1965).–"Introduction to Probability and Statistics from a Bayesian Viewpoint." Pt. 2. Inference. (Cambridge Univ. Press.)

Patil, V. H. (1965).–Approximations to the Behrens–Fisher distributions. *Biometrika* **52**, 267–71.

Patil, V. H. (1969).–Approximation to the generalized Behrens–Fisher distribution involving three variates. *Biometrika* **56**, 687–9.

Quenouille, M. H. (1958).–"The Fundamentals of Statistical Reasoning." p. 139. (Griffin: London.)

Sukhatme, P. V. (1938).–On Fisher and Behrens' test of significance for the difference in means of two normal samples. *Sankhya* **4**, 39–48.

Sukhatme, P. V., Thawani, V. D., Pendharkar, V. G., and Natu, N. P. (1951).–Revised tables for the *d*-test of significance. *J. Indian Soc. agric. Statist.* **3**, 9–19.

Welch, B. L. (1947).–The generalization of 'Student's' problem when several population variances are involved. *Biometrika* **34**, 28–35.

Welch, B. L. (1949).–Further note on Mrs. Aspin's tables and on certain approximations to the tabled function. *Biometrika* **36**, 293–96.

21

Reprinted from Sankhyā : The Indian Journal of Statistics, Series B, Vol. 35, Part 1, 1973

AN APPROXIMATION TO THE k-SAMPLE BEHRENS-FISHER DISTRIBUTION

By A. W. DAVIS and A. J. SCOTT*

C.S.I.R.O., Adelaide and the London School of Economics

SUMMARY. An approximation to the k-sample Behrens-Fisher distribution is suggested. It is based on an asymptotic expansion of the percentage points and is closely related to the approximation suggested by Welch in his attack on a frequency solution to the Behrens-Fisher problem. The approximation is shown to compare favourably with others that have been proposed.

1. Introduction

The construction of interval estimates for linear combinations of means is important in many applications. If the observations are independent and normally distributed about the means with unknown, and not necessarily equal, variances it is well known (see Lindley, 1965, for example) that in a Bayesian approach with conventional conjugate prior distributions, or in a fiducial approach, the posterior distribution of any linear combination of the means is equivalent to the distribution of a linear combination of Student-t variables, say

$$d = \sum_1^k \xi_i t_{n_i} \left(\sum_1^k \xi_i^2 = 1 \right), \qquad \dots \quad (1)$$

where $\xi_1, \dots, \xi_k$ are known constants, t_{n_i} has a Student-t distribution with n_i d.f. and $t_{n_1}, \dots, t_{n_k}$ are mutually independent. The distribution (1) is also of interest in other contexts (see Chapman, 1950, for example).

When $k = 2$, percentage points for some values of ξ_1, ξ_2, n_1, n_2 have been tabulated by Sukhatme (1938). The tables are reproduced in Fisher and Yates (1963), who also give tables constructed by Fisher and Healy (1956) for small odd degrees of freedom. Further tables are given by Weir (1969a). For larger n_1, n_2 the expansion given by Fisher (1941) can be used for values of $\theta = \sin^{-1}\xi_1$ outside the range of the tables. The relative accuracies of the values given by the expansion and a direct evaluation are discussed in Sukhatme *et al* (1951) and Weir (1969b). Davis and Scott (1971) have given tables for some special cases when $k > 2$, but the number of parameters becomes so large as k increases that the tables are cumbersome and difficult to use. The natural approach in these circumstances is to look for a good approximation. Approximations based on the t-distribution have been suggested by Quenouille (1958) and Patil (1965, 1969).

* This research was supported by the National Institutes of Health Grant GM-12868 while both authors were on leave at the University of North Carolina at Chapel Hill.

In this paper we derive a simple approximation based on the extension of Fisher's expansion to the k-sample case. This approximation is closely related to that suggested by Welch in his attack on a frequentist solution to the Behrens-Fisher problem. The accuracy of this approximation and the two suggested earlier are investigated in the final section.

2. The approximation

We define

$$V_{ij} = \sum_{l=1}^{k} \xi_l^{2i}/n_l^j, \quad \ldots \quad (2)$$

following the notation of Aspin (1948) in her development of Welch's method for finding an approximate frequentist solution to the Behrens-Fisher problem (Welch, 1947). Then the Cornish-Fisher expansion of the $100\left(1-\frac{\alpha}{2}\right)\%$ point, d_α, of d in terms of the corresponding standard normal deviate u_α may be represented as

$$d_\alpha \sim u_\alpha + \sum_{r=1}^{\infty} Q_r(u_\alpha) \quad \ldots \quad (3)$$

where $Q_r(u_\alpha) = O(n_i^{-r})$. (See Fisher and Cornish, 1937; 1960 and Hill and Davis, 1968). Expressions for $Q_r(u)$ have been given in Davis and Scott (1970) for $r = 1, \ldots, 5$. In particular

$$Q_1(u) = u\left[V_{11} + \frac{1}{4} V_{21}(u^2-3)\right],$$

$$Q_2(u) = u\Big[2V_{12} - \frac{1}{2}V_{11}^2 + \left(2V_{22} - \frac{3}{4}V_{11}V_{21}\right)(u^2-3)$$

$$+\frac{1}{3}V_{32}(u^2-10u^2+15) - \frac{3}{32}V_{21}^2(3u^4-24u^2+29)\Big]. \quad \ldots \quad (4)$$

In his approach to finding a frequentist solution to the Behrens-Fisher problem, Welch (1947) equates the first-order term in his series expansion with that of the t-distribution and suggests approximating the distribution of d by the t-distribution with degrees of freedom

$$\nu = V_{21}^{-1} = \left(\sum_1^k \xi_j^4/n_j\right)^{-1}.$$

The form of expansion (3) as it stands does not lead directly to a similar approximation for d_α, for equating the 1st order term

$Q_1(u_\alpha) = u_\alpha[V_{11}+V_{21}(u_\alpha^2-3)/4]$ to the 1st order term of t_ν, $u_\alpha(u_\alpha^2+1)/4\nu$,

gives an expression involving the particular percentage point u_α. However if we express the expansion in terms of the standardized deviate d_α/s where s^2 is the variance of d,

$$s^2 = \sum_1^k \xi_j^2/(n_j-2) \sim 1+2V_{11}+4V_{12}+8V_{13}+\ldots, \quad \ldots \quad (5)$$

it is found that all products formed from V_{11}, V_{12}, V_{13} etc. are removed to give

$$\frac{d_\alpha}{s} \sim u_\alpha\Big[1+\frac{1}{4}V_{21}(u_\alpha^2-3)+(2V_{22}-V_{11}V_{21})(u_\alpha^2-3)$$

$$+\frac{1}{3}V_{32}(u_\alpha^4-10u^2+15)-\frac{3}{32}V_{21}^2(3u_\alpha^4-24u_\alpha^2+29)+\ldots\Big].$$

Comparison with the special case of t_n ($k=1$) and equating the 1st order terms now lead directly to the simple approximation to the distribution of d :

$$d \approx t_\nu s\sqrt{\frac{\nu-2}{\nu}} \qquad \ldots \quad (6)$$

where

$$\nu = V_{21}^{-1} = \Big(\sum_1^k \xi_j^4/n_j\Big)^{-1}. \qquad \ldots \quad (7)$$

The degrees of freedom are the same as in Welch's approximation, but the approximation is for the standardized d in terms of the standardized t.

3. A comparison with other approximations

Quenouille (1958) has suggested approximating the distribution of d by a t-distribution with the same variance. Patil (1965, 1969) has proposed using a multiple of a t-variate with the multiplying constant and degrees of freedom chosen by equating the first four moments. All three approximations can be expressed in the form (6), $d \approx t_\nu s\sqrt{\frac{\nu-2}{\nu}}$ for different values of ν. The values of ν are

$$\nu_{DS} = \Big[\sum_1^k \xi_j^4/n_j\Big]^{-1} \qquad \ldots \quad (8)$$

for the approximation of Section 3 ($D-S$);

$$\nu_Q = \frac{\Big[\sum_1^k \xi_j^2\, n_j/(n_j-2)\Big]}{\Big[\sum_1^k \xi_j^2/(n_j-2)\Big]} \qquad \ldots \quad (9)$$

for Quenouille's approximation (Q);

$$\nu_P = 4+\frac{\Big[\sum_1^k \xi_j^2\, n_j/(n_j-2)^2\Big]}{\Big[\sum_1^k \xi_j^4\, n_j^2/(n_j-2)^2(n_j-4)\Big]} \qquad \ldots \quad (10)$$

for Patil's approximation (P).

The approximations all become more accurate for fixed $\xi_1, \ldots, \xi_k$ as $\min_{1\leqslant j\leqslant k}\{n_j\}$ increases. As k increases with $\min_{1\leqslant j\leqslant k}\{n_j\} \geqslant 3$, d/s converges in distribution to a standard normal random variable provided $\max_{1\leqslant j\leqslant k} \xi_j \to 0$ (see, for example, Hájek and Sidák, 1969, p. 153). Since $\nu_{DS} \geqslant 3/\max\{\xi_i\}$ the $D-S$ approximation also approaches the appropriate normal percentage point under these conditions and the error converges to zero as k increases. The same is true of the P approximation if $\min\{n_j\} \geqslant 5$. The value of ν_Q, however, remains finite so that the error in the Q approximation does not converge to zero as k increases.

The results of a numerical investigation into the accuracy of the three approximations are summarized in Tables 1 and 2. Table 1 gives the maximum absolute error of the approximations over the range of values of $\theta = \sin^{-1}\xi_1$ used in Sukhatme's Tabulation (1938, 1951) for $k = 2$ and $n_1, n_2 = 6, 8, 12, 24, \infty$. Table 2 covers the special case $n_1 = n_2 = \ldots = n_k$ for $k = 3, 4, 5, 10$ and gives the maximum absolute error of the approximations over the grid of all $\{\xi_1, \xi_2, \ldots, \xi_k\}$ with $10\xi_i^2$ a positive integer. All the approximations perform exceptionally well at the 5% level, not so well at the 1% level. As expected, the accuracy of the $D-S$ and P approximations improves ultimately as k increases, but the Q approximation seems to get worse. The performances of the $D-S$ and P approximations are very similar and it is difficult to give a clear-cut recommendation for one or the other. The $D-S$ approximation is simpler and more accurate at the 5% level, while the P approximation becomes better at the 1% level. It was found that the true value at the 1% level always lay between the values given by the P and $D-S$ approximations for the range of parameters considered. The Q approximation is not as accurate as the other two. However, it is simpler to use since it is based directly on the t-distribution, and it gives reasonable values at the 5% level.

TABLE 1. MAXIMUM ABSOLUTE ERROR OF THE APPROXIMATIONS FOR $k = 2$

α		.05			.01		
n_1	n_2	$D-S$	P	Q	$D-S$	P	Q
6	6	.008	.011	.012	.10	.05	.19
8		.009	.015	.015	.08	.06	.17
12		.012	.019	.020	.07	.07	.15
24		.013	.022	.024	.09	.08	.11
∞		.013	.028	.027	.10	.09	.08
8	8	.004	.006	.014	.05	.02	.15
12		.004	.009	.014	.04	.03	.12
24		.005	.011	.016	.04	.03	.10
∞		.005	.014	.017	.05	.03	.06
12	12	.001	.002	.012	.02	.01	.10
24		.001	.004	.011	.02	.01	.07
∞		.001	.005	.010	.02	.01	.04
24	24	.000	.001	.007	.00	.00	.05
∞		.000	.001	.004	.01	.00	.02

TABLE 2. MAXIMUM ABSOLUTE ERROR OF THE k APPROXIMATIONS FOR 3, $n_i = n$

α		.05			.01		
k	n	$D-S$	P	Q	$D-S$	P	Q
3	6	.012	.017	.018	.10	.07	.28
	8	.004	.009	.020	.05	.03	.21
	12	.003	.004	.017	.02	.01	.14
	24	.001	.001	.010	.00	.00	.07
4	6	.010	.019	.022	.10	.08	.33
	8	.004	.010	.024	.05	.03	.24
	12	.002	.004	.020	.02	.01	.16
	24	.001	.001	.012	.00	.00	.08
5	6	.007	.020	.024	.09	.09	.36
	8	.004	.010	.027	.04	.03	.27
	12	.002	.004	.023	.02	.01	.17
	24	.001	.001	.013	.00	.00	.08
10	6	.006	.005	.032	.06	.02	.45
	8	.003	.002	.034	.03	.00	.32
	12	.001	.000	.027	.01	.00	.20
	24	.000	.000	.015	.00	.00	.10

References

Aspin, A. A. (1948): An examination and further development of a formula arising in the problem of comparing two mean values. *Biometrika*, **35**, 88-96.

Chapman, D. G. (1950): Some two sample tests. *Ann. Math. Statist.*, **21**, 601-06.

Cornish, E. A. and Fisher, R. A. (1937): Moments and cumulants in the specification of distributions. *Rev. de l' Inst. Int. de Stat.*, **5**, 307-22.

Davis, A. W. and Scott, A. J. (1970): On the k-sample Behrens-Fisher distribution. C.S.I.R.O. Technical Papers, **33**, 1-11.

Fisher, R. A. (1941): The asymptotic approach to Behrens integral, with further tables for the d test of significance. *Ann. Eugen.*, **11**, 141-72.

Fisher, R. A. and Cornish, E. A. (1960): The percentile points of distributions having known cumulants *Technometrics*, **2**, 209-25.

Fisher, R. A. and Healy, M. J. R. (1956): New tables of Behrens' test of significance. *J. R. Statist. Soc.*, B, **18**, 212-16.

Fisher, R. A. and Yates, F. (1963): *Statistical Tables*, 6th Ed. Oliver and Boyd, London.

Hájek, J. and Šidák, Z. (1967): *Theory of Rank Tests*, 153, Academic Press, New York.

Hill, G. W. and Davis, A. W. (1968): Generalized asymptotic expansions of Cornish-Fisher type. *Ann. Math. Statist.*, **39**, 1264-73.

LINDLEY, D. V. (1965) : *Introduction to Probability and Statistics from a Bayesian Viewpoint*, Part 2, *Inference*. The University Press. Cambridge.

PATIL, V. M. (1965) : Approximations to the Behrens-Fisher distributions. *Biometrika*, **52**, 267-71.

——— (1969) : Approximation to the generalized Behrens-Fisher distribution involving three variates. *Biometrika*, **56**, 687-89.

QUENOUILLE, M. H. (1958) : *The Fundamentals of Statistical Reasoning*, 139, Griffin, London.

SUKHATME, P. V. (1938) : On Fisher and Behrens' test of significance for the difference in means of two noarmal samples. *Sankhyā*, **4**, 39-48.

SUKHATME, P. V., THAWANI, V. D., PENDHARKAR, V. G. and NATU, N. P. (1951) : Revised tables for the d-test of significance. *Indian Soc. Agricultural Statist. Journal*, **3**, 9-19.

WEIR, J. B. DE V. (1969) : Table of 0.2 percentage points of Behrens d. *Sankhyā*, Series B, **31**, 103-04.

——— (1969) : Revised end-figures of Behrens d. *Sankhyā*, Series B, **31**, 105-06.

WELCH, B. L. (1957) : The generalization of 'Student's' problem when several population variances are involved. *Biometrika*, **34**, 28-35.

——— (1949) : Further note on Mrs. Aspin's tables and on certain approximations to the tabled function. *Biometrika*, **36**, 293-96.

Paper received : March, 1971.
Revised : February, 1972.

22

[Reprinted from ANNALS OF THE INSTITUTE OF STATISTICAL MATHEMATICS, Vol. 30, No. 1, A, 1978]

ON THE ASYMPTOTIC DISTRIBUTION OF GOWER'S m^2 GOODNESS-OF-FIT CRITERION IN A PARTICULAR CASE

A. W. DAVIS

Ann. Inst. Statist. Math.
30 (1978), Part A, 71–79

ON THE ASYMPTOTIC DISTRIBUTION OF GOWER'S m^2 GOODNESS-OF-FIT CRITERION IN A PARTICULAR CASE

A. W. DAVIS

(Received Nov. 25, 1976)

1. Summary

The asymptotic behaviour of Gower's m^2 goodness-of-fit criterion is considered in the case of two independent sets of Mahalanobis's D^2 distances, obtained from the same p-variate normal population (equal numbers of samples, equal sample sizes). The asymptotic distribution is approximated by a multiple of chi-square, and some Monte Carlo results are presented to illustrate the approach to this distribution.

2. Introduction

The m^2 (originally R^2) statistic was introduced by Gower ([2], [3]) to compare sets of distances constructed between multivariate samples or populations. For example, we may wish to compare (i) distances derived by different approaches (e.g. Mahalanobis's D^2 and Pearson's coefficient of racial likeness) from the same observations on the same samples, (ii) distances derived by the same or different methods from different subsets of observations on the same samples, or (iii) distances derived from different samples from the same populations.

Given k populations and two sets of distances between them, (d_{ij}) and (d_{ij}^*) $(i, j=1,\cdots,k)$ say, Gower suggested (a) applying principal coordinate analysis for example, (Gower [1]), to map the distances onto two sets of geometric points (P_i), (P_i^*) $(i=1,\cdots,k)$ in Euclidean p-space, in such a way that $d_{ij}(d_{ij}^*)$ is the Euclidean distance between P_i and P_j (P_i^* and P_j^*); then (b) moving the P_i^* relative to the P_i in p-space by means of translations, rotations, reflections and scalings until the "residual" sum of squares

$$m^2=\sum_{i=1}^{k} \delta^2(P_i, P_i^*) \tag{1}$$

Key words: Multivariate analysis; Gower's goodness-of-fit criterion; Asymptotic distribution; Non-central Wishart distribution; Canonical analysis.

is minimum, where δ denotes Euclidean distance.

Gower [2] showed that the required translation consists of shifting the two sets of points to a common centroid, which we shall take to be the origin. If $\boldsymbol{X}$ and $\boldsymbol{Y}$ denote the $k\times p$ matrices whose ith rows are the resulting coordinates of P_i and P_i^* respectively, then the minimum value of (1) following suitable rotation and reflection is

$$m^2=\text{trace}\,[\boldsymbol{X}'\boldsymbol{X}+\boldsymbol{Y}'\boldsymbol{Y}-2(\boldsymbol{X}'\boldsymbol{Y}\boldsymbol{Y}'\boldsymbol{X})^{1/2}]\ . \tag{2}$$

The reader is referred to Gower [3] for the generalization to more than two sets of distances, with allowance for scaling (Generalized Procrustes Analysis).

It was pointed out by Gower [2] that the distributional properties of m^2 in "null" situations are fundamental for any statistical inference based on the above approach. As a starting point for the investigation of these he suggested a particular case of (iii), namely, that in which Mahalanobis's D^2 distances are constructed for two independent sets of samples from the same p-variate normal populations $N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$ $(i=1, \cdots, k)$, with mean vectors $\boldsymbol{\mu}_i$ and common covariance matrix $\boldsymbol{\Sigma}$. The rows of $\boldsymbol{X}$ and $\boldsymbol{Y}$ are then the canonical mean vectors (centroid at the origin) arising from separate canonical variate analyses of the two sets of samples. In the present note we shall further assume that the individual samples have equal size n, and show that, as $n\to\infty$, nm^2 is asymptotically distributed as a central positive definite quadratic form in normal variables. The complexity of this result suggests that the exact distribution of m^2 may be exceedingly difficult to obtain. The asymptotic mean and variance are also derived, and used to obtain an asymptotic approximation $nm^2\sim c\chi^2_\alpha$.

3. Asymptotic distribution

We shall refer to the two sets of samples as the x-set and y-set respectively, and introduce the following notation for the x-set:

$\bar{\boldsymbol{x}}_i$=mean vector of the x-sample (size n) from $N(\mu_i, \boldsymbol{\Sigma})$, $(i=1,\cdots,k)$

$\bar{\boldsymbol{x}}=k^{-1}\sum_{i=1}^{k}\bar{\boldsymbol{x}}_i$=grand mean vector,

$\mathcal{X}(k\times p)=(\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}},\cdots,\bar{\boldsymbol{x}}_k-\bar{\boldsymbol{x}})'$

$\boldsymbol{S}_x$=pooled within-samples covariance matrix on $\nu=k(n-1)$ degrees of freedom,

with a corresponding notation for the y-set. We assume $k\geqq p$. Now let

$$\boldsymbol{z}_i'=\sqrt{n}\,(\bar{\boldsymbol{x}}_i', \bar{\boldsymbol{y}}_i')\ , \qquad (i=1,\cdots,k)\ ,$$

$$\bar{\boldsymbol{z}}=k^{-1}\sum_{i=1}^{k}\boldsymbol{z}_i\ ,$$

$$\mathcal{Z}(k\times 2p)=(z_1-\bar{z},\cdots,z_k-\bar{z})'$$

$$B(2p\times 2p)=\mathcal{Z}'\mathcal{Z}=n\begin{bmatrix}\mathcal{X}'\mathcal{X} & \mathcal{X}'\mathcal{Y}\\ \mathcal{Y}'\mathcal{X} & \mathcal{Y}'\mathcal{Y}\end{bmatrix}=\begin{bmatrix}B_{xx} & B_{xy}\\ B'_{xy} & B_{yy}\end{bmatrix},$$

where B_{xx}, B_{yy} are the between-samples sums of squares and products matrices for the x- and y-sets, respectively.

In carrying out a canonical variate analysis for the x-set, say ([8], Chapter 7), an orthogonal matrix H_x is found such that

$$S_x^{-1/2}B_{xx}S_x^{-1/2}=H_x\Lambda_xH_x' ,$$

where Λ_x is the diagonal matrix of eigenvalues of the left-hand side matrix. Defining

$$X(k\times p)=\mathcal{X}S_x^{-1/2}H_x ,$$

the rows of X are seen to be the canonical mean vectors of the x-samples, with centroid at the origin as required. Similarly, we construct

$$Y(k\times p)=\mathcal{Y}S_y^{-1/2}H_y$$

for the y-set, and Gower's m^2 statistic (equation (2)) takes the form

$$(3)\qquad m^2=n^{-1}\,\text{trace}\,[B_{xx}S_x^{-1}+B_{yy}S_y^{-1}-2\{S_y^{-1/2}B'_{xy}S_y^{-1}B_{xy}S_y^{-1/2}\}^{1/2}] .$$

(Note that if H is orthogonal and A is positive definite symmetric, then trace $(HAH')^{1/2}$=trace $HA^{1/2}H'$=trace $A^{1/2}$.)

To discuss the asymptotic properties of this quantity for large n, we first note that since the z_i are independent $2p$-variate normal vectors with means $\sqrt{n}(\mu_i', \mu_i')'$ and covariance matrix

$$(4)\qquad \begin{bmatrix}\Sigma & O_p\\ O_p & \Sigma\end{bmatrix},$$

where O_p is the $p\times p$ zero matrix, the matrix B has the non-central Wishart distribution [6] with $q=k-1$ degrees of freedom, population covariance matrix (4), and matrix of non-centrality parameters $n\Omega/2$, where

$$\Omega=\begin{bmatrix}\Theta & \Theta\\ \Theta & \Theta\end{bmatrix},$$

$$(5)\qquad \Theta=\Sigma^{-1/2}\left\{\sum_{i=1}^{k}(\mu_i-\bar{\mu})(\mu_i-\bar{\mu})'\right\}\Sigma^{-1/2},$$

$$\bar{\mu}=k^{-1}\sum_{i=1}^{k}\mu_i .$$

Now let $\boldsymbol{K}$ be an orthogonal matrix reducing $\boldsymbol{\Theta}$ to diagonal form,

(6) $$\boldsymbol{\Theta}=\boldsymbol{K\Delta K}' , \qquad \boldsymbol{\Delta}=\mathrm{diag}\,(\theta_1,\cdots,\theta_p) .$$

It will be assumed that $\boldsymbol{\Theta}$ is positive definite, that is, all $\theta_i>0$. If we transform to new variables

$$\boldsymbol{x}^*=\boldsymbol{K\Sigma}^{-1/2}\boldsymbol{x} , \qquad \boldsymbol{y}^*=\boldsymbol{K\Sigma}^{-1/2}\boldsymbol{y} ,$$

then m^2 is given by (3) in terms of the corresponding "starred" quantities; νS_x^* and νS_y^* have central Wishart distributions with ν degrees of freedom and unit population covariance matrices $\boldsymbol{I}_p$, and $\boldsymbol{B}^*$ has a non-central Wishart distribution with q degrees of freedom, population covariance matrix $\boldsymbol{I}_{2p}$, and non-centrality matrix $n\boldsymbol{\Omega}^*/2$, where

(7) $$\boldsymbol{\Omega}^*=\begin{bmatrix}\boldsymbol{\Delta} & \boldsymbol{\Delta}\\ \boldsymbol{\Delta} & \boldsymbol{\Delta}\end{bmatrix}=(\omega_{ij}) .$$

Now

(8) $$\mathrm{E}\,(\boldsymbol{B}^*)=n\boldsymbol{\Omega}^*+q\boldsymbol{I}_{2p}$$

(see for example [4]), so that we may write

(9) $$n^{-1}\boldsymbol{B}^*=\boldsymbol{\Omega}^*+(q/n)\boldsymbol{I}_{2p}+\boldsymbol{U} ,$$

where

(10) $$\boldsymbol{U}=\begin{bmatrix}\boldsymbol{U}_{xx} & \boldsymbol{U}_{xy}\\ \boldsymbol{U}_{xy}' & \boldsymbol{U}_{yy}\end{bmatrix}=(U_{ij})$$

has zero expectation. Also let

(11) $$\boldsymbol{S}_x^*=\boldsymbol{I}_p+\boldsymbol{V}_x , \qquad (\boldsymbol{V}_x=(V_{ij}^x))$$

with a similar notation for $\boldsymbol{S}_y^*$. Then $\boldsymbol{U}$, $\boldsymbol{V}_x$ and $\boldsymbol{V}_y$ are $O(n^{-1/2})$ (equations (A7), (A8) in the Appendix), and to order n^{-1} (Appendix (a))

(12) $$\begin{aligned} m^2\sim 2pq/n&+\sum_{i=1}^{p}(U_{ii}+U_{i+p,i+p}-2U_{i,i+p})\\ &-\sum_{i=1}^{p}\sum_{j=1}^{p}(U_{ij}V_{ij}^x+U_{i+p,j+p}V_{ij}^y)\\ &+\frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}(\theta_i+\theta_j)^{-1}\{-(U_{i,j+p}U_{j,i+p})^2\\ &+4U_{i,j+p}(\theta_jV_{ij}^x+\theta_iV_{ij}^y)+\theta_i\theta_j(V_{ij}^x-V_{ij}^y)^2\} .\end{aligned}$$

It follows from (9) that the first two terms sum to trace $(\mathcal{X}^*-\mathcal{Y}^*)'(\mathcal{X}^*-\mathcal{Y}^*)$, which is the trace of a central Wishart matrix with q degrees of freedom and population covariance matrix $(2/n)\boldsymbol{I}_p$, and is thus distributed as $(2/n)\chi^2_{pq}$. Considering the remaining terms, we note that the elements of the non-centrality matrix $n\Omega^*/2\to\infty$ as $n\to\infty$, so that the variates $\sqrt{n}\,U_{ij}$ are asymptotically jointly normal [4] with zero means, and covariances given by equation (A7). The $\sqrt{n}\,V^x_{ij}$ and $\sqrt{n}\,V^y_{ij}$ also approach joint normality, with zero means and covariances given by (A8), and hence nm^2 is asymptotically distributed as a central quadratic form in normal variables ([5], Chapter 29). Since nm^2 is non-negative, this form is necessarily positive-definite, although it seems difficult to show this explicitly. Generally such distributions are well approximated by multiples of chi-square, $c\chi^2_\alpha$, on non-integer degrees of freedom, and in order to evaluate c and α it is sufficient to have the asymptotic mean μ'_1 and variance μ_2 of nm^2; these are derived in Appendix (b):

$$\mu'_1=2\left[p\left(q-\frac{1}{2}(p-1)\right)+k^{-1}\sum_{1\le i\le j\le p}\sum\frac{\theta_i\theta_j}{\theta_i+\theta_j}\right], \tag{13}$$

$$\mu_2=8\left[p\left(q-\frac{1}{2}(p-1)\right)+k^{-1}(p+1)\sum_{i=1}^{p}\theta_i \right.$$
$$\left.-k^{-1}\sum_{1\le i\le j\le p}\sum\frac{\theta_i^2+\theta_j^2}{\theta_i+\theta_j}+k^{-2}\sum_{1\le i\le j\le p}\sum\frac{\theta_i^2\theta_j^2}{(\theta_i+\theta_j)^2}\right]. \tag{14}$$

Then

$$c=\mu_2/2\mu'_1\,,\qquad \alpha=2(\mu'_1)^2/\mu_2\,. \tag{15}$$

4. Some Monte Carlo results

Ten 5-variate population mean vectors were constructed by sampling uniformly over a hypersphere of radius 6 (Table 2), and 200 x-sets and y-sets of samples (sizes 50, 100, 200) were generated with population covariance matrices $\boldsymbol{\Sigma}=\boldsymbol{I}_5$. The matrix $\boldsymbol{\Theta}$ thus reduces to the sum of squares and products matrix of the mean vectors,

$$\boldsymbol{\Theta}=\begin{bmatrix} 72.96 & 13.30 & 6.29 & -3.55 & 6.56\\ * & 46.40 & 11.71 & -23.88 & -8.91\\ * & * & 35.82 & -11.09 & -23.88\\ * & * & * & 27.65 & 7.74\\ * & * & * & * & 43.31\end{bmatrix} \tag{16}$$

with latent roots

$$\theta_1=88.89\,,\quad \theta_2=71.72\,,\quad \theta_3=40.73\,,\quad \theta_4=14.66\,,\quad \theta_5=10.14\,. \tag{17}$$

Canonical analyses were carried out, and 200 values of nm^2 were calculated for each sample size using the ROTATE directive in the program GENSTAT [7].

Table 1 shows the means and variances of the sampled values, together with the asymptotic values obtained from (13), (14) and (17). The corresponding approximate asymptotic distribution (equation (15)) was

$$5.273\chi^2_\alpha\,, \quad \text{with} \quad \alpha=23.680\,. \tag{18}$$

As a rough indication of the approach of the simulated distributions to (18), Pearson's goodness-of-fit criterion was calculated for each set of 200 values, based on the deciles of (18). By sample size 50, (18) appears to be giving a reasonable approximation to the overall shape of the curve. It would be desirable to investigate the approximation in higher dimensions.

Table 1 Approach of nm^2 to the approximate asymptotic distribution (10 5-variate populations, 200 trials)

n	Mean (nm^2)	Var (nm^2)	Pearson χ^2_9	P
50	129.33 (2.86)	1639 (333)	12.8	0.17
100	125.01 (2.51)	1264 (145)	4.0	0.91
200	129.30 (2.65)	1407 (144)	4.2	0.90
∞	124.86	1317		

Figures in brackets denote standard errors.

Table 2 Mean vectors used in Table 1

Population	Mean Vector				
1	2.1368	4.3781	−0.3241	−2.2281	0.9611
2	−2.6930	−1.3939	−1.7595	−0.1107	3.4107
3	0.1313	3.2676	1.9097	−4.1451	−1.1622
4	3.4693	2.7769	1.4386	−0.9864	−2.1049
5	−2.8560	−1.0431	−3.5809	−0.0487	−0.4427
6	−3.6216	0.4996	2.0489	0.8789	−3.0896
7	0.1676	1.8274	−2.1315	1.5259	3.1300
8	1.8050	−1.7579	1.4265	0.3312	1.4627
9	4.1422	−1.9152	−0.9020	1.5625	0.1427
10	3.3105	1.3766	−1.9964	−0.4910	2.3302

Appendix

(a) *Derivation of* (12): Write

(A1) $$S_x^{*-1}=I_p+\mathcal{V}_x\ , \qquad S_x^{*-1/2}=I_p+\bar{\mathcal{V}}_x(\mathcal{V}_x=(\mathcal{V}^x_{ij}),\ \bar{\mathcal{V}}_x=(\bar{\mathcal{V}}^x_{ij}))\ ,$$

with a similar notation for S_y^*. Equations (9) and (A1) are to be substituted in the starred version of (3), retaining terms of order 1, $n^{-1/2}$ and n^{-1}; thus

(A2) $$n^{-1}\,\mathrm{trace}\,[B^*_{xx}S_x^{*-1}]=\sum_{i=1}^{p}\theta_i+pq/n+\sum_{i=1}^{p}[U_{ii}+(\theta_i+q/n)\mathcal{V}^x_{ii}] +\sum_{i=1}^{p}\sum_{j=1}^{p}U_{ij}\mathcal{V}^*_{ij}\ ,$$

and a similar expression is obtained for $n^{-1}\,\mathrm{trace}\,[B^*_{yy}S_y^{*-1}]$.

To derive a corresponding result for the final term in (3), let

(A3) $$\begin{aligned}C^2&=(I_p+\bar{\mathcal{V}}_x)(\Delta+U_{xy})(I_p+\mathcal{V}_y)(\Delta+U'_{xy})(I_p+\bar{\mathcal{V}}_x)\\&=\Delta^2+A\ , \qquad (A=(A_{ij}))\ ,\end{aligned}$$

say, where

$$C=\sum_{l=0}^{\infty}C_l\ , \qquad C_l=(C^{(l)}_{ij})=O(n^{-l/2})\ ,$$

($l=0, 1, 2, \cdots$) and each C_l is symmetric. Equating terms of like order,

$$\begin{aligned}&C_0=\Delta\ ,\\&C_1C_0+C_0C_1=A\ ,\\&C_2C_0+C_1^2+C_0C_2=O_p\ , \qquad \text{etc.}\end{aligned}$$

Hence

$$\begin{aligned}&C^{(1)}_{ij}=(\theta_i+\theta_j)^{-1}A_{ij}\ ,\\&C^{(2)}_{ij}=-(\theta_i+\theta_j)^{-1}\sum_{l=1}^{p}[A_{il}A_{jl}/(\theta_i+\theta_l)(\theta_j+\theta_l)]\ ,\end{aligned}$$

etc., so that

(A4) $$2\,\mathrm{trace}\,C=2\sum_{i=1}^{p}\theta_i+\sum_{i=1}^{p}A_{ii}/\theta_i-\sum_{i=1}^{p}\sum_{j=1}^{p}A^2_{ij}/\theta_i(\theta_i+\theta_j)^2+\cdots\ .$$

To terms of order $n^{-1/2}$ and n^{-1},

(A5) $$\begin{aligned}A_{ij}=&[\theta_iU_{i,j+p}+\theta_jU_{j,i+p}+(\theta_i^2+\theta_j^2)\mathcal{V}^x_{ij}+\theta_i\theta_j\mathcal{V}^y_{ij}]\\&+\sum_{l=1}^{p}[U_{i,l+p}U_{j,l+p}+\theta_iU_{l,i+p}\bar{\mathcal{V}}^x_{jl}+\theta_jU_{l,j+p}\bar{\mathcal{V}}^x_{il}\end{aligned}$$

$$+\theta_i U_{j,l+p}\mathcal{V}^y_{il}+\theta_j U_{i,l+p}\mathcal{V}^y_{jl}+\theta_l\{U_{i,l+p}\bar{\mathcal{V}}^x_{jl}$$
$$+U_{j,l+p}\bar{\mathcal{V}}^x_{il}+\theta_i\bar{\mathcal{V}}^x_{jl}\mathcal{V}^y_{il}+\theta_j\bar{\mathcal{V}}^x_{il}\mathcal{V}^y_{jl}+\theta_l\bar{\mathcal{V}}^x_{il}\bar{\mathcal{V}}^x_{jl}\}] ,$$

and substituting in (A4), the resulting expression takes its most convenient form if we replace $\bar{\mathcal{V}}_x$ by $\mathcal{V}_x/2-\mathcal{V}_x^2/8$:

$$\text{(A6)} \qquad 2 \operatorname{trace} \boldsymbol{C}=2\sum_{i=1}^{p}\theta_i+\sum_{i=1}^{p}[2U_{i,i+p}+\theta_i(\mathcal{V}^x_{ii}+\mathcal{V}^y_{ii})]$$
$$+\frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}(\theta_i+\theta_j)^{-1}[(U_{i,j+p}-U_{j,i+p})^2$$
$$+4U_{i,j+p}(\theta_j\mathcal{V}^x_{ij}+\theta_i\mathcal{V}^y_{ij})-\theta_i\theta_j(\mathcal{V}^x_{ij}-\mathcal{V}^y_{ij})^2] .$$

Finally, (12) is obtained from (A2) and (A6). Quantities of order 1 and $n^{-1/2}$ cancel, and to order n^{-1} it has been possible to replace $\mathcal{V}_x$, $\mathcal{V}_y$ by $-\boldsymbol{V}_x$, $-\boldsymbol{V}_y$ respectively.

(b) *Derivation of mean and variance*: For (13), we require only

$$\text{(A7)} \qquad \mathrm{E}[U_{ij}U_{kl}]=\kappa(ij, kl)/n-q\lambda(ij, kl)/n^2$$

$$\text{(A8)} \qquad \mathrm{E}[V^x_{ij}V^x_{kl}]=\lambda(ij, kl)/\nu ,$$

(Jensen [4]), where

$$\text{(A9)} \qquad \begin{aligned}\kappa(ij, kl)&=\delta_{ik}\omega_{jl}+\delta_{il}\omega_{jk}+\delta_{jk}\omega_{il}+\delta_{jl}\omega_{ik} ,\\ \lambda(ij, kl)&=\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jl} ,\end{aligned}$$

and δ_{ij} is Kronecker's delta. Note that $\boldsymbol{U}$, $\boldsymbol{V}_x$ and $\boldsymbol{V}_y$ are independently distributed.

In deriving the variance of (12), the $2pq/n$ term may be ignored, and the remainder consists of five terms, $1, 2, \cdots, 5$ say, defined by the bracketing. Let (i, j) denote the covariance of terms i and j. Then from the independence, these vanish apart from the five variances, (1, 3), and (2, 4), and to evaluate these we require the following formulae in addition to (A7) and (A8):

$$\text{(A10)} \qquad \mathrm{E}[U_{ij}U_{kl}U_{mn}]=n^{-2}\sum\omega_{ik}\lambda(jl, mn)+O(n^{-3}) ,$$

where the summation extends over the twelve possible selections of the subscripts of ω from distinct pairs of U's; and

$$\text{(A11)} \qquad \mathrm{E}[U_{ij}U_{kl}U_{mn}U_{pq}]=n^{-2}\sum\kappa(ij, kl)\kappa(mn, pq)+O(n^{-3}) ,$$

$$\text{(A12)} \qquad \mathrm{E}[V_{ij}V_{kl}V_{mn}V_{pq}]=\nu^{-2}\sum\lambda(ij, kl)\lambda(mn, pq)+O(n^{-3}) ,$$

where the summations extend over the three distinct arrangements of (ij), (kl), (mn) and (pq) into pairs. These results may be obtained from the moment generating function of the non-central Wishart distribution ([6], p. 175). The required covariances are found to be

$$(1,1)=8pq/n^2\,,\qquad (2,2)=(8/n\nu)(p+1)\sum_{i=1}^{p}\theta_i$$

$$(3,3)=4p(p-1)/n^2=-(1,3)/2\,,$$

(A13)

$$(4,4)=(8/n\nu)\sum\sum_{1\le i\le j\le p}(\theta_i^2+\theta_j^2)/(\theta_i+\theta_j)=-(2,4)/2$$

$$(5,5)=(8/\nu^2)\sum\sum_{1\le i\le j\le p}\theta_i^2\theta_j^2/(\theta_i+\theta_j)^2\,.$$

On summing we obtain (14).

Acknowledgments

The author's thanks are due to Mr. J. C. Gower for suggesting this investigation and to Mr. L. G. Veitch for assistance with the programming.

CSIRO Division of Mathematics and Statistics, Adelaide

References

[1] Gower, J. C. (1966). Some distance properties of latent and vector methods used in multivariate analysis, *Biometrika*, **53**, 325–338.

[2] Gower, J. C. (1971). Statistical methods of comparing different multivariate analyses of the same data, *Mathematics in the Archaeological and Historical Sciences*, Hodson, F. R., Kendall, D. G., and Tautu, P. (Eds.), Edinburgh University Press, 138–149.

[3] Gower, J. C. (1975). Generalized Procrustes Analysis, *Psychometrika*, **40**, (in the press).

[4] Jensen, D. R. (1972). The limiting form of the non-central Wishart distribution, *Aust. J. Statist.*, **14**, 10–16.

[5] Johnson, N. L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions*-2, Hougton Mifflin, Boston.

[6] Johnson, N. L. and Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions*, Wiley, New York.

[7] Nelder, J. A. and Members of the Rothamsted Statistics Department (1975). *Genstat Reference Manual*, Inter-University/Research Council Series, Report No. 3, Third Edition, Edinburgh Regional Computing Centre.

[8] Seal, H. (1974). *Multivariate Statistical Analysis for Biologists*, Methuen, London.

To appear in Technometrics

On certain ratio statistics in weather modification experiments

A.W. Davis

Division of Mathematics and Statistics, CSIRO, Adelaide, Australia.

ABSTRACT

Permutation tests based on certain ratio statistics have been used (a) in the analysis of rainfall stimulation experiments and (b) in simulation studies aimed at providing guidance in the planning of such experiments. The paper presents asymptotic approximations to the permutation distributions of these statistics, and also to their power functions under a simple multiplicative model for the seeding effect. Simulation results are presented for comparison with the theoretical approximations.

Key words

Permutation distribution

Weather modification experiments

Double-ratio statistic

Power of permutation test

Asymptotic normality

1. Introduction

In the planning of rainfall experiments, Australian workers (e.g. Smith [4]) have favored the target-control type of design whenever the meteorological conditions have made this appropriate. They have also made extensive use of permutation tests in their analyses of such experiments, based in particular on the double-ratio and root-double-ratio statistics. No useful approximations to the permutation distributions of these statistics have been given in the literature, however, and for any given experiment the distribution has been approximated by random sampling.

The power of such tests is also an important consideration, particularly in the planning of experiments. Simulation investigations have been carried out in this connection (Section 5). The amount of computing required in such simulations is quite considerable, particularly if it is required to assess the effects of varying the period length and duration of the experiment.

It thus seems worthwhile to develop an approximate analytic approach to the distributional properties of the double-ratio and root-double-ratio statistics, and the present paper is an attempt in this direction. The approximate permutation distributions under the null hypothesis of no seeding effect are considered in Section 2 for arbitrary block size, and are shown to be asymptotically normal in Section 3. In Section 4 the power of the tests against the alternative hypothesis of a variable multiplicative seeding effect is also approximated. Section 5 presents simulation results which indicate that the theoretical approximations yield useful results, when applied to actual rainfall records on which artificial seeding effects have been superimposed in accordance with the above model. Finally, the method is applied to investigate the effects of varying period length and block size for a set of rainfall data.

2. The double-ratio and root-double-ratio statistics

The experimental design considered in connection with the double-ratio and root-double-ratio statistics is as follows. Two geographical areas are selected to be as climatologically similar as possible, and are designated "Target Area" and "Control Area" respectively. The duration of the experiment is divided into 2kn periods, arranged in n blocks each of 2k consecutive periods. Cloud seeding is then carried out according to a randomized block design with k replications per block, i.e. the target area is seeded during k randomly selected periods within each block. If T_s and C_s denote the total rainfalls in the target and control areas during the seeded periods, and T_u and C_u are the unseeded periods, then the double-ratio statistic is defined by

$$D = T_s C_u / T_u C_s .$$

D is thus the ratio of T_s to $(C_s/C_u)T_u$, the latter being a crude estimate of the total target rainfall during the seeded periods, assuming that seeding had not taken place. Hence D-1 estimates the average proportional increase in rainfall due to seeding.

In experiments of cross-over type, the "control" area is seeded during periods in which the target area is left unseeded. An estimate of the average proportional increase in rainfall may thus be obtained from the root-double-ratio statistic $S = \sqrt{D}$.

Under the null hypothesis of no seeding effect, the $\binom{2k}{k}^n$ equally probable values of D or S corresponding to the allocation of all possible seeding sequences to the observed rainfalls define the permutation distributions of the statistics. We first consider the approximate moments of these distributions.

Let $T = T_s + T_u$, $C = C_s + C_u$ denote the total target and control rainfalls for the duration of the experiment, and define

$$A = (T_s - T_u)/T, \quad B = (C_s - C_u)/C.$$

Then D may be written as

$$D = (1 + A)(1 - B)/(1 - A)(1 + B) = 1 + d, \qquad (1)$$

say, where

$$d = 2(A - B)/(1 - A)(1 + B).$$

Clearly $|A| \leq 1$, $|B| \leq 1$, and the denominator of d^r may be validly expanded as a product series in which terms of like degree in A and B are grouped together, provided that neither $|A|$ nor $|B|$ is unity for any seeding sequence; i.e. at least one block must have non-zero target rainfalls in more than k periods, and similarly for the control rainfalls. If we define

$$H_{\ell,m} = E_P\{(A - B)^{\ell}(AB)^m\}, \qquad (\ell, m = 0, 1, 2, \ldots),$$

where E_P denotes expectation with respect to the permutation distribution, then grouping terms of like degree we obtain

$$\begin{aligned}\mu'_r &= E_P(d^r) \\ &= 2^r[H_{r,0} + rH_{r+1,0} + \{rH_{r,1} + \tfrac{1}{2}r(r + 1)H_{r+2,0}\} + \ldots.] , \qquad (2) \\ & \qquad (r = 1, 2, \ldots).\end{aligned}$$

This form can be seen by writing the denominator of d as $1 - \{(A - B) + (AB)\}$; the resulting expansion is not valid for all $|A| \leq 1$, $|B| \leq 1$, until rearranged as in (2). The moments may thus be expressed in terms of the doubly-subscripted array $H_{\ell,m}$, with the further simplification that $H_{\ell,m} = 0$ when ℓ is odd. This follows because to each seeding sequence corresponds a complementary sequence which merely interchanges T_s, C_s with T_u, C_u, thus reversing the signs of A and B. Hence in particular

$$\begin{aligned}\mu'_1 &= 2[H_{2,0} + (2H_{2,1} + H_{4,0}) + \ldots], \\ \mu'_2 &= 4[H_{2,0} + (2H_{2,1} + 3H_{4,0}) + \ldots], \qquad (3) \\ \mu'_3 &= 24H_{4,0} + \ldots, \quad \mu'_4 = 16H_{4,0} + \ldots,\end{aligned}$$

and D and d have the same central moments. The $H_{\ell,m}$ are expressible in terms of the rainfalls t_{ij} and c_{ij} in the target and control areas, respectively, during the j th period of the i th block, and also the block means $t_{i.}$, $c_{i.}$ ($i = 1, \ldots, n$; $j = 1, \ldots, k$) (see Appendix).

Under the assumptions of Section 3, A, B $= O(n^{-\frac{1}{2}})$, $H_{\ell,m} = O(n^{-\frac{1}{2}\ell-m})$, and to order n^{-1} the variance and mean of the permutation distribution of D are given by

$$\sigma_P^2(D) = 4H_{2,0}$$

$$= \{8k/(2k-1)\} \sum_{i=1}^{n}\sum_{j=1}^{2k} \{(t_{ij} - t_{i.})/T - (c_{ij} - c_{i.})/C\}^2$$

and

$$\mu_P(D) = 1 + 2H_{2,0} = 1 + \tfrac{1}{2}\sigma_P^2(D),$$

respectively. The variance is seen to be proportional to the pooled within -block sum of squares of the weighted differences $t_{ij}/T - c_{ij}/C$ of the target and control rainfalls.

A similar approach shows that the root-double-ratio statistic S has mean and variance

$$\mu_P(S) = 1 + \tfrac{1}{8}\sigma_P^2(D), \quad \sigma_P^2(S) = \tfrac{1}{4}\sigma_P^2(D)$$

to order n^{-1}. The variance of S is thus approximately one quarter that of D in long experiments. Each statistic is also seen to have a positive bias equal to about one half of its variance. Flueck and Holland [1] have discussed the bias in these statistics, assuming gamma probability models for the rainfall distributions. A referee has pointed out that the bias may be eliminated by considering $\ln D$ ($\ln S = \frac{1}{2}\ln D$), since it follows from the remarks preceding equation (3) that the permutation distribution of this statistic is symmetrical about zero. The large-sample variance is approximately $\sigma_P^2(D)$.

3. Asymptotic normality of the permutation distributions

Let $t_i^{(s)}$ and $c_i^{(s)}$ denote the target and control rainfall totals for the seeded periods, $1 \lesssim j_1 < \ldots < j_k \lesssim 2k$ say, in the i th block. Then

$$(T_s, C_s) = \sum_{i=1}^{n} (t_i^{(s)}, c_i^{(s)})$$

is a sum of n independent and non-identically distributed random vectors under the permutation distribution, where

$$t_i^{(s)} - kt_{i.} = \sum_{\ell=1}^{k} (t_{ij_\ell} - t_{i.}),\ c_i^{(s)} - kc_{i.} = \sum_{\ell=1}^{k} (c_{ij_\ell} - c_{i.}).$$

Using the method indicated in the Appendix, we find that $(t_i^{(s)}, c_i^{(s)})$ has mean vector $(kt_{i.}, kc_{i.})$ and covariance matrix $\frac{1}{2}kS_i$, where S_i is the sample covariance matrix of the (t_{ij}, c_{ij}) $(j = 1, \ldots, 2k)$ in the i th block. Thus (T_s, C_s) has mean vector $\frac{1}{2}(T, C)$ and covariance matrix $\frac{1}{2}kn\ \Sigma_n$, where

$$\Sigma_n = \begin{pmatrix} \sigma_T^2 & \rho\sigma_T\sigma_C \\ \rho\sigma_T\sigma_C & \sigma_C^2 \end{pmatrix}$$

is the pooled within-blocks covariance matrix of the (t_{ij}, c_{ij}); that is,

$$\sigma_T^2 = \{n(2k - 1)\}^{-1} \sum_{i=1}^{n} \sum_{j=1}^{2k} (t_{ij} - t_{i.})^2,$$

etc.

We now make the following assumptions concerning the particular rainfall record in the absence of seeding effect:

I. The rainfalls t_{ij}, c_{ij} remain bounded as $n \to \infty$.

II. $\lim_{n\to\infty} T/2kn = \mu_t$, $\lim_{n\to\infty} C/2kn = \mu_c$,

the overall period means for target and control areas, respectively.

III. $\lim_{n\to\infty} \Sigma_n = \Sigma$, the overal covariance matrix of period totals.

Assumption I is clearly plausible on physical grounds. Assumptions II and III may be expected to hold for almost all observed sets of rainfalls, provided that the underlying joint distribution is sufficiently "regular". Sufficient conditions for such regularity could no doubt be formulated, but at the risk of imposing artificial restrictions which may be unacceptable to meteorologists.

From I and III, $n^{-\frac{1}{2}}(T_s - \frac{1}{2}T, C_s - \frac{1}{2}C)$ is asymptotically bivariate normal with zero mean vector and covariance matrix $\frac{1}{2}k\Sigma$, in virtue of a multivariate analogue of the Lindeberg-Feller conditions for the Central Limit Theorem (Rao [3] p. 118). Thus, in the notation of Section 2, (A, B) is asymptotically bivariate normal with zero mean vector and covariance matrix

$$2kn \begin{pmatrix} \sigma_T^2/T^2 & \rho\sigma_T\sigma_C/TC \\ \rho\sigma_T\sigma_C/TC & \sigma_C^2/C^2 \end{pmatrix},$$

which is $O(n^{-1})$ in virtue of II and III. Hence from (1)

$$D \lessapprox 1 + 2(A - B)$$

is asymptotically $N(1, \sigma_P^2(D))$. As a consequence, S is asymptotically $N(1, \frac{1}{4}\sigma_P^2(D))$. Under the assumptions made, $\sigma_P^2(D)$ and hence the biases in D and S are $O(n^{-1})$; the biases are thus neglected in the asymptotic formulae.

4. Power of the tests

In this section, the asymptotic power of the D and S tests will be discussed under a simple multiplicative model for the seeding effect. No doubt more realistic models can and will be found, but it is hoped that the methods of this paper may be adaptable to deal with some of these.

Suppose that the target area contains g gages and let $t_{ij}^{(q)}$ denote the reading that would have been obtained from the q th gage during the j th period of the i th block in the absence of seeding. In connection with the D test, the following model for a variable seeding effect will be considered:

(a) If the j th period is seeded, the observed rainfall for the q th gage is $t_{ij}^{(q)*} = z_{ij}^{(q)} t_{ij}^{(q)}$ $(i = 1, \ldots, n; j = 1, \ldots, 2k; q = 1, \ldots, g)$.

Conditional upon an assigned seeding sequence, δ, say.

(b) The $z_{ij}^{(q)}$ have mean $(1 + \alpha)$ and variance ϕ^2, where α is the average proportional increase in rainfall due to seeding;

(c) Between periods, the $z_{ij}^{(q)}$ are independently distributed; within periods, the multipliers $z_{ij}^{(q)}$ and $z_{ij}^{(r)}$ for the q th and r th gages have correlation $\rho_{q,r}$ $(q, r = 1, \ldots, g)$.

Let T_s^* denote the total target rainfall during the seeded periods; then the double ratio statistic for the assigned δ and multipliers $(z_{ij}^{(q)})$ is

$$D^* = T_s^* C_u / T_u C_s .$$

We first consider the distribution of D^* conditional upon a particular sequence of underlying rainfalls $(t_{ij}^{(q)}, c_{ij})$, i.e. as δ and the $z_{ij}^{(q)}$ are permitted to vary. Conditional upon δ it follows from the proposed seeding model that D^* and D^{*2} have expected values $(1 + \alpha)D$ and

$$(1 + \alpha)^2 D^2 + \phi^2 \{C_u^2 / T_u^2 C_s^2\} \sum_{i=1}^{n} \sum_{\ell=1}^{k} \sum_{q=1}^{g} \sum_{r=1}^{g} \rho_{q,r} t_{ij}^{(q)} t_{ij_\ell}^{(r)}$$

respectively, where D is the double ratio statistic for the underlying rainfalls. Averaging over δ, and replacing C_u, T_u and C_s by their expected values under the permutation distribution, D^* has mean and approximate variance

$$\mu(D^*) = (1 + \alpha)\mu_P(D)$$

$$\sigma^2(D^*) \approx (1 + \alpha)^2 \sigma_P^2(D) + (2\phi^2/T^2) \sum_{i=1}^{n} \sum_{j=1}^{k} \sum_{q=1}^{g} \sum_{r=1}^{g} \rho_{q,r} t_{ij}^{(q)} t_{ij}^{(r)} \qquad (4)$$

respectively. A similar argument to that used in Section 3 shows that for almost all rainfall sequences D^* will be as asymptotically $N(1 + \alpha, \sigma^2(D^*))$ as $n \to \infty$ under fairly wide conditions. Furthermore $n\sigma^2(D^*)$ will be asymptotically equivalent to its expected value, and so this result may be expected to hold asymptotically for the unconditional distribution of D^*.

The power function of the D^* test is the probability that D^* exceeds the upper ε-point of the permutation distribution constructed from the observed rainfalls. From Section 3, the permutation distribution is asymptotically $N(1, \sigma_P^2(D^*))$, where for large n, $\sigma_P^2(D^*)$ may be approximated by its expected value as δ and the $z_{ij}^{(q)}$ vary. The latter may be shown to be asymptotically equivalent to

$$\sigma^2_{P,E}(D^*) = \sigma_P^2(D) + \frac{2}{(1 + \tfrac{1}{2}\alpha)^2 T^2} \left[\frac{2k\alpha^2}{(2k - 1)^2} \{k \textstyle\sum_{i=1}^{n} t_{i.}^2 + (k - 1)\sum_{i=1}^{n}\sum_{j=1}^{2k} t_{ij}^2\} + \phi^2 \sum_{i=1}^{n}\sum_{j=1}^{2k}\sum_{q=1}^{g}\sum_{r=1}^{g} \rho_{q,r} t_{ij}^{(q)} t_{ij}^{(r)} \right]. \tag{5}$$

The approximate power of the D test under the above seeding model may thus be calculated using the normal approximations based on (4) and (5).

Similar results may be derived for the S test. Assuming that the control area contains h gages, with correlations $\bar{\rho}_{q,r}$ between the period rainfalls $c_{ij}^{(q)}$ and $c_{ij}^{(r)}$ for the q the and r th gages, we obtain the required variances from (4) and (5) by adding $c_{i.}/C$, c_{ij}/C and $\sum_{q=1}^{h}\sum_{r=1}^{h} \bar{\rho}_{q,r} c_{ij}^{(q)} c_{ij}^{(r)}/C^2$ to $t_{i.}/T$, t_{ij}/T and $\sum_{q=1}^{g}\sum_{r=1}^{g} \rho_{q,r} t_{ij}^{(q)} t_{ij}^{(r)}/T^2$ respectively, where the latter occur explicitly, and dividing the resulting expressions by 4.

5. Examples and Discussion

The present investigation was suggested by a paper [5] of Twomey and Robertson (TR), who carried out a computer simulation aimed at estimating probabilities of "success" in experiments of various lengths. Actual rainfall records for selected Australian regions were taken, and artificial variable seeding effects were applied to these in accordance with the model of Section 4, with k = 1 (2 periods per block). The $z_{ij}^{(q)}$ were assumed to be independent ($\rho_{q,r} = \bar{\rho}_{q,r} = 0$, $q \neq r$) and $N(1 + \alpha, \phi^2)$, for selected values of α and ϕ. A "successful" experiment was one in which the D(or S) test yielded a significant result at the 5% level, and the estimated seeding effect was in error by not more than 50%. The probability, P say, of "success" may be approximated using the results of Section 4.

The TR approach has received some criticism, and D.E.Shaw (unpublished report) has suggested an alternative approach in which the variable seeding effect is applied directly to the period totals (all $\rho_{q,r} = \bar{\rho}_{q,r} = 1$), and a significant result at the 5% level is regarded as a success. This corresponds to the power function with $\varepsilon = 0.05$.

Table 1 presents the results of simulations designed to test the approximations made in Section 4. Two data sets were used, namely

I. Western Australia Coast (Perth to Albany). 19 stations. Monthly rainfall for four years (1950 - 1953). Block size 2 (k = 1), n = 24 blocks.

II. Tasmania. 54 stations. Weekly rainfall for selected periods during 1965, 1967, 1969, 1972. Block size 10 (k = 5), n = 18 blocks.

The simulation results were based on 400 runs, with 100 re-randomizations per run to estimate the permutation distributions. In order to test the approximations under the least favorable conditions in the

TR set-up, α and ϕ were assigned the largest values used by them ($\alpha = 0.2$, $\phi = 0.25$), and a four year record was taken for data set I. It should be noted that normality of the seeding effects was not assumed in Section 4; clearly it would be desirable to compare the asymptotic results with simulations based on, e.g. a positively-skewed distribution for the seeding effects.

Points illustrated by Table 1 are

(a) The small positive bias in the means μ_p of the permutation distributions, equal to about one-half of the corresponding variances (Section 2).

(b) A fair degree of stability in the standard deviations σ_p of the permutation distributions, as indicated by their standard deviation over the 400 runs, shown in brackets. Also quite good agreement between the theoretical $\sigma_{P,E}$ and the average of the simulation σ_p's.

(c) Rough normality of the permutation, D^* and S^* distributions, indicated by the skewness and kurtosis coefficients β_1, β_2.

(d) Good agreement between the mean and standard deviation of the 400 D^* values and the theoretical approximations to $\mu(D^*)$ and $\sigma(D^*)$. Similarly for S^*.

(e) Useful agreement between the theoretical approximations to the probabilities of success (P) and the estimates yielded by the simulations.

These results could almost certainly have been improved by using ln D and ln S, as suggested by the referee.

Table 2 compares results from Tables 1 and 2 of [5] with approximations based on Section 4. Agreement is generally satisfactory, suggesting that to the degree of accuracy warranted by the approach, the theoretical results may be usefully applied, with considerable saving in computing time.

(Tables near here)

In Table 3, the theoretical results have been applied to the Tasmanian data set II to investigate the effects of varying period length and block size (k). The choice of block size 2 (k = 1) for experiments is seen to be justified, at least from the viewpoint of Shaw's seeding model, and clearly the chances of "success" generally decrease with increasing period length. Due to the great variability of the rainfall and the finite length of the rainfall record, the results may be quite markedly affected by small changes in the starting point of the experiment. This is illustrated by the column headed " + 2 days", in which the starting point has been advanced by two days. A particularly striking instance is the anomalous result for 7 day periods and block size 2; this appears to be due to a succession of very high daily rainfalls which occur in a single period under the original starting point, but are divided among two periods when the origin is altered. However, in connection with the planning of future experiments, the overall trends in the table seem sufficiently clear.

6. Conclusions

Asymptotic approximations have been presented for the permutation distributions of the double-ratio and root-double-ratio statistics used in the analysis of target-control-type experiments in rainfall modification. The asymptotic powers of these tests in the case of a variable multiplicative seeding effect have also been obtained. Comparison of the approximations with simulation results using actual rainfall records with artificial seeding effects superimposed indicates that the results may be useful in the planning of future experiments, with considerable saving in computing time.

7. Appendix

Let $1 \leq j_1 < \dots < j_k \leq 2k$ denote any one of the $\binom{2k}{k}$ possible sets of k periods randomly selected for seeding in a given block. Then the calculation of the various moments required reduces to finding moments and product moments of quantities $\sum_{\ell=1}^{k}\alpha_{j_\ell}$, $\sum_{\ell=1}^{k}\beta_{j_\ell}$, ..., where $\sum_{r=1}^{2k}\alpha_r = \sum_{r=1}^{2k}\beta_r = \dots 0$. For example

$$E_P(\sum_{\ell=1}^{k}\alpha_{j_\ell}) = \binom{2k}{k}^{-1}\binom{2k-1}{k-1}\sum_{r=1}^{2k}\alpha_r = 0,$$

and similarly

$$E_P\{(\sum_{\ell=1}^{k}\alpha_{j_\ell})(\sum_{m=1}^{k}\beta_{j_m})\} = \{k/2(2k-1)\}\sum_{r=1}^{2k}\alpha_r\beta_r.$$

Such formulae are special cases of results given by Pitman [2], p. 125. Let

$$\alpha_{ij} = (t_{ij} - t_{i\cdot})/T, \ \beta_{ij} = (c_{ij} - c_{i\cdot})/C, \ \gamma_{ij} = \alpha_{ij} - \beta_{ij}.$$

Then using the independence of the blocks under the permutation distribution, we obtain in particular

$$H_{2,0} = \{2k/(2k-1)\}\sum_{i=1}^{n}\sum_{j=1}^{2k}\gamma_{ij}^2,$$

$$H_{4,0} = \{4k/(2k-1)^2(2k-3)\}\ \{-2k(2k-1)\Sigma\Sigma\gamma_{ij}^4 + 3k(2k-3)(\Sigma\Sigma\gamma_{ij}^2)^2 + 3\Sigma_i(\Sigma_j\gamma_{ij}^2)^2\}.$$

REFERENCES

[1] Flueck, J.A. and Holland, B.S. (1976). Ratio estimators and some inherent problems in their utilization. J. Appl. Meteorology, 15, 535-543.

[2] Pitman, E.J.G. (1937). Significance tests which may be applied to samples from any populations. I. Suppl. J.R. Statist. Soc., 4, 119-130.

[3] Rao, C.R. (1965). Linear Statistical Inference and Its Applications. New York: John Wiley and Sons.

[4] Smith, E.J. (1967). Cloud seeding experiments in Australia. Proc. Fifth Berkeley Symp., 5, 161-176.

[5] Twomey, S. and Robertson, I. (1973). Numerical simulation of cloud seeding experiments in selected Australian areas. J. Appl. Meteorology, 12, 473-478.

TABLE 1. Comparison of simulation results (Sim.) with theoretical applications (Th.) for TR- and Shaw models. $\alpha = 0.2$, $\phi = 0.25$. 400 runs. Standard deviations shown in brackets.

(a.) Double-Ratio Statistic

Data	2k	Blocks n		Permutation Distribution $\mu_P(D)$	$\sigma_P(D)$	β_1	β_2	D* Distribution $\mu(D^*)$	$\sigma(D^*)$	β_1	β_2	Probability 100P
I(TR)	2	24	Sim.	1.005	0.0950	0.223	2.663	1.205	0.0975	0.120	2.797	68.75
				(0.002)	(0.0151)	(0.045)	(0.177)					(2.38)
			Th.	1.000	0.0953	0	3	1.200	0.0973	0	3	65.47
I(Shaw)	2	24	Sim.	1.006	0.109	0.249	2.652	1.119	0.120	0.350	3.235	48.75
				(0.003	(0.023)	(0.056)	(0.236)					(2.50)
			Th.	1.000	0.112	0	3	1.200	0.117	0	3	52.97
II(Shaw)	10	18	Sim.	1.004	0.0859	0.250	2.989	1.201	0.1059	0.208	2.687	66.25
				(0.005)	(0.0064)	(0.115)	(0.259)					(2.28)
			Th.	1.000	0.0862	0	3	1.200	0.0997	0	3	70.74

(b) Root-double-ratio Statistic

Data	2k	Blocks n		Permutation Distribution $\mu_P(S)$	$\sigma_P(S)$	β_1	β_2	D* Distribution $\mu(S^*)$	$\sigma(S^*)$	β_1	β_2	Probability 100P
I(TR)	2	24	Sim.	1.002	0.0626	0.152	2.653	1.201	0.0491	0.113	2.945	100.00
				(0.001)	(0.0101)	(0.040)	(0.140)					
			Th.	1.000	0.0632	0	3	1.200	0.0496	0	3	97.11
I(Shaw)	2	24	Sim.	1.003	0.0735	0.174	2.631	1.201	0.0692	0.260	3.279	92.75
				(0.002)	(0.0140)	(0.044)	(0.070)					(1.72)
			Th.	1.000	0.0754	0	3	1.200	0.0671	0	3	86.22
II(Shaw)	10	18	Sim.	1.001	0.0486	0.142	2.906	1.199	0.0563	0.097	2.824	98.5
				(0.002)	(0.0044)	(0.106)	(0.221)					(0.67)
			Th.	1.000	0.0490	0	3	1.200	0.0531	0	3	98.70

TABLE 2. Comparison of original Twomey-Robertson results with theoretical approximations (given underneath).

Values of 20P for α = 0.1, 0.2; ϕ = (a) 0.01, (b) 0.10, (c) 0.25.

Location	α	Double-Ratio Statistic									Root-Double-Ratio Statistic								
		4 years			8 years			16 years			4 years			8 years			16 years		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
Western Australia (coast,	0.1	3	1	1	2	3	1	11	9	10	11	9	8	16	16	13	20	20	19
Perth to Albany, 19 stations).		0	0	0	2	2	2	11	10	10	11	11	10	14	14	13	19	19	19
Mean Rainfall 41 inches.	0.2	10	9	10	14	17	12	19	20	18	19	19	19	20	20	20	20	20	20
		10	10	10	13	13	13	19	19	19	19	19	19	20	20	20	20	20	20
Western Queensland and	0.1	0	0	0	0	0	0	0	0	0	0	0	0	3	2	2	10	8	9
N.S.W. (21 stations).		0	0	0	0	0	0	0	0	0	0	0	0	2	2	2	7	7	7
Mean Rainfall 10 inches.	0.2	0	0	0	1	0	1	4	7	4	4	3	0	9	14	11	17	17	18
		0	0	0	2	2	1	6	6	6	4	4	4	13	13	12	17	17	17
Eastern Riverina	0.1	0	1	0	3	6	4	5	10	10	7	10	7	14	13	10	19	20	18
(40 stations).		0	0	0	3	3	2	9	9	9	8	8	7	15	14	14	19	19	19
Mean Rainfall 17 inches.	0.2	9	7	6	9	14	17	20	18	18	18	20	18	19	20	19	20	20	20
		7	7	7	14	14	13	18	18	18	18	18	18	20	20	20	20	20	20

TABLE 3. Theoretical approximation: percentages 100P of success experiments (Shaw model) for various period lengths and block sizes. Data set II.

Period length	Block size (2k)	Double-Ratio Statistic: α = 0.1, ϕ = 0.01	(+ 2 days)	Double-Ratio Statistic: α = 0.2, ϕ = 0.25	(+ 2 days)	Root-Double-Ratio Statistic: α = 0.1, ϕ = 0.01	(+ 2 days)	Root-Double-Ratio Statistic: α = 0.2, ϕ = 0.25	(+ 2 days)
1 day	2	59.80	59.77	95.47	95.46	97.76	97.75	100.00	100.00
	4	57.54	57.84	94.43	94.60	97.11	97.24	100.00	100.00
	6	56.18	55.99	93.79	93.69	96.64	96.57	100.00	100.00
	8	56.03	56.65	93.72	94.02	96.59	96.81	100.00	100.00
	10	54.53	55.39	92.92	93.38	95.98	96.34	100.00	100.00
2 days	2	53.19	54.23	91.15	91.76	95.24	95.72	99.99	99.99
	4	50.74	51.92	89.62	90.42	94.00	94.66	99.99	99.99
	6	49.05	49.77	88.42	88.95	92.96	93.42	99.98	99.98
	8	48.18	48.44	87.72	87.92	92,37	92.55	99.97	99.98
	10	47.12	47.25	86.87	86.97	91.60	91.70	99.97	99.99
4 days	2	45.17	46.50	83.45	84.71	89.68	90.81	99.87	99.90
	4	42.62	42.28	81.10	80.80	87.43	87.08	99.80	99.79
	6	41.99	41.98	80.54	80.58	86.81	86.81	99.79	99.79
	8	40.69	40.44	79.04	78.82	85.39	85.10	99.72	99.71
	10	39.95	39.30	78.23	77.53	84.53	83.74	99.67	99.63
7 days	2	51.76	44.38	86.32	80.60	94.07	88.65	99.88	99.66
	4	38.19	36.37	74.09	71.94	81.99	79.49	99.15	98.89
	6	37.92	36.28	73.68	71.77	81.67	79.40	99.10	98.86
	8	35.22	34.04	70.14	68.66	77.80	75.96	98.58	98.34
	10	35.57	34.01	70.74	68.72	78.36	75.93	98.70	98.37

Section 4.

A generalization of the zonal polynomials, with applications to noncentral multivariate distributions.

P. R. Krishnaiah, ed., *Multivariate Analysis–V*

INVARIANT POLYNOMIALS WITH TWO MATRIX ARGUMENTS, EXTENDING THE ZONAL POLYNOMIALS

A.W. Davis
C.S.I.R.O. Division of Mathematics and Statistics, Adelaide, South Australia.

A class of invarant polynomials in two matrices is defined, generalizing the zonal polynomials. Basic properties of the polynomials are derived, and their applications to distribution theory indicated.

1. Introduction

The theory of group representations and zonal spherical functions was applied by James [7] to the evaluation of certain integrals arising in multivariate distribution theory. The integrals were expanded as power series in the zonal polynomials of the real positive definite symmetric matrices, which had been previously studied by Hua [6]. Subsequently, Constantine [2] showed that Herz's [5] hypergeometric functions of matrix argument could be expanded in terms of zonal polynomials, and expressed a number of noncentral multivariate distributions in terms of these functions. A survey of the area was given by James [8]. Since then, many authors have utilized and developed these techniques, so that an extensive literature now exists.

However, there remain related distributional problems which cannot be solved in terms of zonal polynomials, including

(a) the cumulative distribution functions of the basic noncentral distributions (Constantine [2]),

(b) the latent roots of the noncentral Wishart distribution (Pillai [12, p.30]), and

(c) the doubly noncentral multivariate F distribution.

The latter arises in particular from a formal approach to the MANOVA distributions in multivariate Edgeworth populations (Davis [3]), and was the initial stimulus for this investigation. In the above and certain other situations, integrals arise having the form (see Section 6)

$$\left[\int_{O(m)} C_\kappa(AH'XH)C_\lambda(BH'YH)\mathrm{d}H\right] \qquad (1.1)$$

where A, B, X, Y are $m\times m$ symmetric matrices, dH is the invariant Haar measure over the group $O(m)$ of $m\times m$ orthogonal matrices, and C_κ, C_λ are the zonal polynomials indexed by the ordered partitions κ, λ of the integrals k, l respectively into $\leqslant m$ parts.

In seeking to evaluate (1.1), we observe that since dH is invariant under left translation $H\to\mathcal{H}H(\mathcal{H}\in O(m))$, (1.1) is a homogeneous polynomial of degree k, l in the elements of X, Y respectively, invarant under the *simultaneous* transformations

$$X\to\mathcal{H}'X\mathcal{H},\quad Y\to\mathcal{H}'Y\mathcal{H},\quad \mathcal{H}\in O(m). \tag{1.2}$$

Invariance of dH under right translation implies that the same holds for A, B. We thus require a basis for polynomials having the property (1.2) which will facilitate the evaluation of (1) the integrals (1.1) and (2) expectations with respect to the Wishart distribution. Similar requirements in the one matrix case are fulfilled by the zonal polynomials.

Section 2 defines a class of invariant polynomials with two matrix arguments as a direct extension of the zonal polynomials. Elementary properties of the polynomials are derived in Section 3, and in Section 4 an expansion of (1.1) is given which constitutes the cornerstone of the subsequent theory (equation (4.13)). Derivation of this result leads to the definition of a class of 'orthogonal' invariant polynomials whose properties are discussed in Section 5. Owing to restrictions on space, their use in multivariate distribution theory can only be briefly indicated (Section 6) but it is hoped that these applications, together with details about the construction of the polynomials, will be presented in subsequent papers. In particular, the application to MANOVA distributions in nonnormal cases requires the low-degree invariant polynomials. This probably constitutes the main practical justification for studying these polynomials at the present stage, since the problems associated with construction, and the convergence of infinite series in the invariant polynomials, are considerably more serious than in the case of the zonal polynomials.

We finally note that the invariant polynomials of this paper are not zonal polynomials, and lack certain properties of the latter. In particular, there is no guarantee that they will possess an analogue of the Laplace-Beltrami operator for the zonal polynomials (James [9]). Hence, although one could define hypergeometric functions of two matrix arguments generalizing e.g. those of Appell in the scalar case, they may not satisfy useful systems of differential equations as do the hypergeometric functions of Herz and Constantine (Muirhead [11]).

2. Invariant polynomials in two matrix arguments

Let $X=(x_{ij})$, $Y=(y_{ij})$ be $m\times m$ complex symmetric matrices; $P_k[X]$ the class of homogeneous polynomials of degree k in the x_{ij}; $P_{k,l}[X,Y]$ the class of homogeneous polynomials of degree k, l in the x_{ij}, y_{ij} respectively; ξ_k the vector of all monomials $\Pi_{i\le j}x_{ij}^{k_{ij}}$ $(\Sigma_{i\le j}k_{ij}=k,\ k_{ij}\ge 0)$; $Gl(m,R)$ the group of $m\times m$ real nonsingular matrices L. Then ξ_k is a basis for $P_k[X]$, and defining a similar basis η_l for $P_l[Y]$,

$$P_{k,l}[X,Y]=P_k[X]\otimes P_l[Y] \tag{2.1}$$

in the sense that $\xi_k\otimes\eta_l$ is a basis for $P_{k,l}[X,Y]$. The simultaneous congruence transformations

$$X\to L'XL,\quad Y\to L'YL,\quad L\in Gl(m,R), \tag{2.2}$$

produce linear transformations in $P_k[X]$, $P_l[Y]$ and $P_{k,l}[X,Y]$ with matrices $T_{2k}(L)$, $T_{2l}(L)$ and $T_{2k,2l}(L)$, respectively, such that

$$\xi_k\to T_{2k}(L)'\xi_k,\quad \eta_l\to T_{2l}(L)'\eta_l,$$
$$\xi_k\otimes\eta_l\to T_{2k,2l}(L)'(\xi_k\otimes\eta_l). \tag{2.3}$$

These constitute representations of $Gl(m,R)$ in the respective vector spaces,

$$T_{2k}(L_1L_2)=T_{2k}(L_1)T_{2k}(L_2),\quad \text{etc.,} \tag{2.4}$$

of polynomial degrees $2k,2l,2f(f=k+l)$ respectively in the elements of L.

It is a classical result (e.g. Boerner [1, Ch. 5] that a vector space in which a polynomial representation of $Gl(m,R)$ is defined may be decomposed into a direct sum of irreducible invariant subspaces, each carrying an irreducible representation of $Gl(m,R)$. The inequivalent representations of degree $2k$, say, may be indexed by the ordered partitions $\mathfrak{K}$ of $2k$ into $\le m$ parts; we shall denote these by $\tilde{T}_{\mathfrak{K}}(L)$. Then A.T. James (unpublished lecture notes) has shown that

(a) $\tilde{T}_{\mathfrak{K}}(L)$ occurs in the decomposition of $\tilde{T}_{2k}(L)$ if it contains the identity representation when L is restricted to the subgroup $O(m)$;

(b) The only such $\tilde{T}_{\mathfrak{K}}(L)$ are those for which $\mathfrak{K}=2\kappa$, where κ is any ordered partition of k into $\le m$ parts;

(c) $\tilde{T}_{2\kappa}(L)$ contains the identity representation exactly once when restricted to $O(m)$, and

(d) $\tilde{T}_{2\kappa}(L)$ occurs with multiplicity one in the decomposition of $T_{2k}(L)$, so that

$$T_{2k}(L) = \bigoplus_{\kappa} \tilde{T}_{2\kappa}(L), \quad T_{2l}(L) = \bigoplus_{\lambda} \tilde{T}_{2\lambda}(L) \tag{2.5}$$

where κ, λ run through the ordered partitions of k, l respectively into $\leq m$ parts. The decompositions (2.5) were originally obtained by Littlewood and Ffoulkes (see references in [7]). Correspondingly, we have the decompositions into irreducible invariant subspaces

$$P_k[X] = \bigoplus_{\kappa} \mathcal{V}_\kappa[X], \quad P_l[Y] = \bigoplus_{\lambda} \mathcal{V}_\lambda[Y], \tag{2.6}$$

where in virtue of (c) $\mathcal{V}_\kappa[X]$ for example contains exactly one one-dimensional subspace generated by a polynomial in X invariant under $X \to H'XH$ $(H \in O(m))$, which when suitably normalized is the zonal polynomial $C_\kappa(X)$. From (2.3) and (2.5)

$$\begin{aligned} T_{2k,2l}(L) &= T_{2k}(L) \otimes T_{2l}(L) \\ &= \bigoplus_{\kappa} \bigoplus_{\lambda} \{\tilde{T}_{2\kappa}(L) \otimes \tilde{T}_{2\lambda}(L)\}, \end{aligned} \tag{2.7}$$

and correspondingly

$$P_{k,l}[X, Y] = \bigoplus_{\kappa} \bigoplus_{\lambda} \{\mathcal{V}_\kappa[X] \otimes \mathcal{V}_\lambda[Y]\}. \tag{2.8}$$

The Kronecker products $\tilde{T}_{2\kappa}(L) \otimes \tilde{T}_{2\lambda}(L)$ are also representations of $Gl(m, R)$ having polynomial degree $2f$, but are not in general irreducible; they may therefore be decomposed into direct sums of representations

Table 1
Decomposition of Kronecker products of irreducible representations

$2f$	2κ	2λ	Φ					
4	2	2	4	31	2^2			
6	4	2	6	51	42			
	2^2	2	42	321	2^3			
8	6	2	8	71	62			
	42	2	62	53	521	4^2	431	42^2
	2^3	2	42^2	32^21	2^4			

$\tilde{T}_\Phi(L)$, Φ being an ordered partition of $2f$ into $\leqslant m$ parts. Rules for determining the particular $\tilde{T}_\Phi(L)$ occurring in such a decomposition, together with their multiplicities, are given for example by Robinson [13, Section 3.3]. Table 1 presents the decompositions for some low order κ, λ.

We thus obtain the decomposition

$$P_{k,l}[X,Y] = \bigoplus_{\kappa} \bigoplus_{\lambda} \bigoplus_{\Phi} \mathcal{V}_\Phi^{\kappa,\lambda}[X,Y] \tag{2.9}$$

into irreducible subspaces, where Φ runs through the irreducible representations in the decomposition of $\tilde{T}_{2\kappa} \otimes \tilde{T}_{2\lambda}$. If $\Phi = 2\phi$, where ϕ is a partition of $f = k+l$, then by (c) $\mathcal{V}_{2\phi}^{\kappa,\lambda}[X,Y]$ contains a one-dimensional subspace generated by a polynomial $\Gamma_\phi^{\kappa,\lambda}(X,Y)$, say, which is invariant under (1.2). The question of normalization of this polynomial will be discussed in Section 4. We note that

(i) Subspaces $\mathcal{V}_\Phi^{\kappa,\lambda}$ occur with Φ not of the form 2ϕ; these contain no invariant polynomials;

(ii) A representation 2ϕ may occur in (2.9) with multiplicity greater than 1 for a given κ, λ, so that strictly an additional subscript should be used; however, for notational simplicity we shall omit this, and write for example $\phi' \equiv \phi$ to indicate equivalent representations. A more important consequence is that the $\mathcal{V}_{2\phi}^{\kappa,\lambda}$, and hence the corresponding polynomials $\Gamma_\phi^{\kappa,\lambda}$, are not uniquely defined when 2ϕ occurs with multiplicity > 1 for a given κ, λ. However, the direct sum of equivalent subspaces

$$\mathcal{U}_{2\phi}^{\kappa,\lambda} = \bigoplus_{\phi' \equiv \phi} \mathcal{V}_{2\phi'}^{\kappa,\lambda} \tag{2.10}$$

is uniquely defined. A sufficient resolution of the non-uniqueness problem will be described in Section 4.

Multiplicity first occurs in the case of the polynomials of degree 6, with $k = l = 3$, $\kappa = \lambda = [2,1]$, when $2\phi = 2[3,2,1]$ occurs with multiplicity 3.

3. Elementary properties of the $\Gamma_\phi^{\kappa,\lambda}$.

(a) $$\Gamma_\phi^{\kappa,\lambda}(X,X) = \{\Gamma_\phi^{\kappa,\lambda}(I,I)/C_\phi(I)\}\, C_\phi(X), \tag{3.1}$$

(b) $$\Gamma_\phi^{\kappa,\lambda}(X,I) = \{\Gamma_\phi^{\kappa,\lambda}(I,I)/C_\kappa(I)\}\, C_\kappa(X), \tag{3.2}$$

with a corresponding result for $\Gamma_{\phi}^{\kappa,\lambda}(I, Y)$;

(c) $\Gamma_{\kappa}^{\kappa,O}(X,Y) \stackrel{\text{def}}{=} C_{\kappa}(X), \Gamma_{\lambda}^{O,\lambda}(X,Y) \stackrel{\text{def}}{=} C_{\lambda}(Y),$ (3.3)

(d) $$\int_{O(m)} \Gamma_{\phi}^{\kappa,\lambda}(A'H'XHA, A'H'YHA)\mathrm{d}H$$

$$= \Gamma_{\phi}^{\kappa,\lambda}(X,Y)C_{\phi}(AA')/C_{\phi}(I) \quad (3.4)$$

(e) If $VV' \sim W_m(n, \Sigma, O)$ (the central Wishart distribution with n degrees of freedom and population covariance Σ), and $RR' = \Sigma$, then

$$E_V \Gamma_{\phi}^{\kappa,\lambda}(V'XV, V'YV) = 2^f \left(\tfrac{1}{2}n\right)_{\phi} \Gamma_{\phi}^{\kappa,\lambda}(R'XR, R'YR) \quad (3.5)$$

where $(a)_{\phi}$ is the generalized hypergeometric coefficient (Constantine [2]).

For brevity we shall merely indicate the salient points in the proofs.

(a) Setting $Y = X$ defines a group homomorphism of $P_{k,l}[X,Y]$ onto $P_k[X]$; since $\mathcal{V}_{\phi}^{\kappa,\lambda}[X,Y]$ may be mapped into the null space, $\Gamma_{\phi}^{\kappa,\lambda}(X,X)$ may be identically zero.

(b) Setting $Y = I$ maps $\mathcal{V}_{\kappa}[X] \otimes \mathcal{V}_{\lambda}[Y]$ onto $\mathcal{V}_{\kappa}[X]$.

(c) $\bar{T}_0(L)$ is the identity representation.

(d) Choose $\Gamma_{\phi}^{\kappa,\lambda}(X,Y)$ as the first basis vector in $\mathcal{V}_{2\phi}^{\kappa,\lambda}[X,Y]$, and the remainder in subspaces invariant under the non-identity representations of $O(m)$; then James has shown that the (1,1) element of $\bar{T}_{2\phi}(A)$ is $C_{\phi}(AA')/C_{\phi}(I)$ and that

$$\int_{O(m)} \bar{T}_{2\phi}(H)\mathrm{d}H = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \vdots & 0 \end{bmatrix}$$

No result corresponding to (3.4) holds if A is replaced by B in one argument of $\Gamma_{\phi}^{\kappa,\lambda}$.

(e) $W = R^{-1}VV'R'^{-1} \sim W_m(n, I, O)$, and $V = RW^{\frac{1}{2}}H_v$, where $H_v \in O(m)$; the result follows using (d).

4. Evaluation of the integral (1.1)

The evaluation of (1.1) constitutes the most important single result for the invariant polynomials introduced in this paper, and the derivation also yields a sufficient resolution of the non-uniqueness mentioned in Section 4.

This is effected by the construction of a new set of invariant polynomials $C_\phi^{\kappa,\lambda}(X,Y)$ which are linear combinations of $\Gamma_\phi^{\kappa,\lambda}$ for $\phi'\equiv\phi$, and are orthogonal in a sense to be defined below (extending the orthogonality of the zonal polynomials, James [7]).

Since $C_\kappa(AX)C_\lambda(BY)$ is invariant under (2.2) together with the contragredient transformation

$$A\to L^{-1}AL^{-1'}, B\to L^{-1}BL^{-1'}, \quad L\in Gl(m,R), \tag{4.1}$$

an argument of James [7] Section 4 (see also Hannan [4] Section 2.2) may be adapted to show that ? 10.2

$$\int_{O(m)} C_\kappa(AH'XH)C_\lambda(BH'YH)\mathrm{d}H=\sum_{\phi\in\kappa\cdot\lambda}\sum_{\phi'\equiv\phi} q_{\phi,\phi'}^{\kappa,\lambda}\Gamma_\phi^{\kappa,\lambda}(A,B)\Gamma_{\phi'}^{\kappa,\lambda}(X,Y) \tag{4.2}$$

where $\phi\in\kappa\cdot\lambda$ is an abbreviation for $2\phi\in 2\kappa\otimes 2\lambda$. Since $C_\kappa(AX)C_\lambda(BY)$ may be shown to span $\mathcal{V}_\kappa[X]\otimes\mathcal{V}[Y]$ as A, B vary over the symmetric matrices, the integral (1.1) must span the polynomials in this space invariant under (1.2), and hence the bilinear form in (4.2) must have a non-singular matrix $Q_{\kappa,\lambda}=(q_{\phi,\phi'}^{\kappa,\lambda})$ ($q_{\phi,\phi'}^{\kappa,\lambda}=0$ unless $\phi\equiv\phi'$). At this stage, however, $Q_{\kappa,\lambda}$ is unknown.

We now follow an approach based on Saw [14]. Let $U=(u_{ij})$ be an $m\times m$ matrix of independent unit normal variables, and consider

$$g=E_u\operatorname{etr}\{\alpha AU'XU+\beta BU'YU\}, \qquad \alpha,\beta \text{ real},$$

$$=\sum_{k,l=0}^{\infty}\alpha^k\beta^l E_{k,l}/k!l! \tag{4.3}$$

where, since $U=W^{\frac{1}{2}}H$, $W\sim W_m(m,I,O)$,

$$E_{k,l}=\sum_{\kappa,\lambda}E_w\int_{O(m)}C_\kappa(AH'W^{\frac{1}{2}}XW^{\frac{1}{2}}H)C_\lambda(BH'W^{\frac{1}{2}}YW^{\frac{1}{2}}H)\mathrm{d}H$$

$$=\Gamma_{k,l}(A,B)'\tilde{Q}_{k,l}\,\Gamma_{k,l}(X,Y). \tag{4.4}$$

Here $\bar{\Gamma}_{k,l}(X,Y)$ denotes the vector of all $\Gamma_\phi^{\kappa,\lambda}(X,Y)$ for fixed k, l and

$$\tilde{Q}_{k,l}=\operatorname{diag}\{\tilde{Q}_{\kappa,\lambda}\},\quad \tilde{Q}_{\kappa,\lambda}=\left(2^f\left(\tfrac{1}{2}m\right)_\phi q_{\phi,\phi'}^{\kappa,\lambda}\right). \tag{4.5}$$

On the other hand, writing $G=\alpha A\otimes X+\alpha B\otimes Y$,

$$g=|I_{m^2}-G|^{-\frac{1}{2}}=\sum_{f=0}^{\infty}1.3\cdots(2f-1)C_f(G)/f! \tag{4.6}$$

whence

$$E_{k,l}=\pi_{k,l}(A,B)'\Delta_{k,l}\pi_{k,l}(X,Y) \tag{4.7}$$

where $\Delta_{k,l}$ is a diagonal matrix with positive diagonal elements, and $\pi_{k,l}(X,Y)$ denotes the vector of all *distinct* products of traces

$$(\operatorname{tr}X^{a_1}Y^{b_1}X^{c_1}\cdots)^{r_1}(\operatorname{tr}X^{a_2}Y^{b_2}X^{c_2}\cdots)^{r_2}\cdots \tag{4.8}$$

of total degree k,l in the elements of the symmetric matrices X, Y respectively (i.e. account must be taken of the symmetry of X and Y, and the properties $\operatorname{tr}XY=\operatorname{tr}YX$, $\operatorname{tr}Z'=\operatorname{tr}Z$).

Example. When $k=2$, $l=1$ it is found from $C_3(G)$ or otherwise that

$$\pi_{2,1}(X,Y)'=\left((\operatorname{tr}X)^2\operatorname{tr}Y,\operatorname{tr}XY\operatorname{tr}X,\operatorname{tr}X^2\operatorname{tr}Y,\operatorname{tr}X^2Y\right),$$

$$\Delta_{2,1}=5\operatorname{diag}(1,4,2,8). \tag{4.9}$$

It seems clear that the distinct products (4.8) for a given k, and l (e.g. the components in (4.9) for $k=2$, $l=1$) are functionally independent, although this may be difficult to prove rigorously. If we assume independence then it follows from (4.4) and (4.7) that

(a) the products (4.8) constitute an elementary basis for the invariant polynomials, generalizing the familiar basis of the zonal polynomials constructed from the $s_i=\operatorname{tr}(X^i)$. The number of such terms should thus equal the sum of the multiplicities of the irreducible representations $\bar{T}_{2\phi}$ occurring in $T_{2k}\otimes T_{2l}$.

(b) $E_{k,l}=\sigma_{k,l}(A,B)'\sigma_{k,l}(X,Y)$ iff the components of $\sigma_{k,l}(X,Y)$ are '$\Delta_{k,l}$-orthonormal' linear combinations of the (4.8) in the sense that

$$\sigma_{k,l}(X,Y)=S\pi_{k,l}(X,Y),\quad S\Delta_{k,l}^{-1}S'=I. \tag{4.10}$$

(c) from (4.4) and (4.5) the $\Gamma_\phi^{\kappa,\lambda}$ are $\Delta_{k,l}$-orthogonal for inequivalent ϕ.

(d) for ϕ' equivalent to ϕ, choose *any set* of $\Delta_{k,l}$-orthonormal polynomials $\bar{\Gamma}_{\phi'}^{\kappa,\lambda}(X,Y)$ say in the unique invariant subspace $\mathfrak{U}_{2\phi}^{\kappa,\lambda}$ (equation (2.10)). Then from (4.2) and (4.5) the integral (1.1) is given by

$$\sum_{\phi\in\kappa.\lambda} \bar{\Gamma}_{\phi'}^{\kappa,\lambda}(A,B)\bar{\Gamma}_{\phi}^{\kappa,\lambda}(X,Y)/2^{f}\left(\tfrac{1}{2}m\right)_{\phi'}, \tag{4.11}$$

noting that the denominators are constant for equivalent ϕ. Renormalizing to

$$C_{\phi}^{\kappa,\lambda}(X,Y)=\sqrt{z_{\phi}}\,\bar{\Gamma}_{\phi}^{\kappa,\lambda}(X,Y) \tag{4.12}$$

where $z_{\phi}=C_{\phi}(I_m)/2^{f}(\frac{1}{2}m)_{\phi}$ is independent of m (Constantine [2]), and the sign of the square root is chosen to make the coefficient of $(\operatorname{tr}X)^k(\operatorname{tr}Y)^l$ in $C_{\phi}^{\kappa,\lambda}$ positive if it is non-zero, we finally obtain

$$\int_{O(m)} C_{\kappa}(AH'XH)C_{\lambda}(BH'YH)\,\mathrm{d}H= \sum_{\phi\in\kappa.\lambda} C_{\phi}^{\kappa,\lambda}(A,B)C_{\phi}^{\kappa,\lambda}(X,Y)/C_{\phi}(I_m). \tag{4.13}$$

The $\Delta_{k,l}$-orthogonality property provides the required resolution of the non-uniqueness, and considerably facilitates the construction of the polynomials. The $C_{\phi}^{\kappa,\lambda}$ are of course not unique if 2ϕ has multiplicity >1 in $2\kappa\otimes 2\lambda$, corresponding to the fact that (4.13) is invariant under 'orthogonal' transformations within the $\mathfrak{U}_{2\phi}^{\kappa,\lambda}$. Details of construction must be deferred to a subsequent paper. Table 2 presents some low degree $C_{\phi}^{\kappa,\lambda}$; the

Table 2
Polynomials $C_{\phi}^{\kappa,\lambda}(X,Y)$

f	k	l	$C_{\phi}^{\kappa,\lambda}$	multiplier	$\operatorname{tr}X\operatorname{tr}Y$	$\operatorname{tr}XY$		
2	1	1	$C_{2}^{\kappa,\lambda}$	1/3	1	2		
			$C_{1^2}^{1,1}$	2/3	1	-1		
					$(\operatorname{tr}X)^2\operatorname{tr}Y$	$\operatorname{tr}XY\operatorname{tr}X$	$\operatorname{tr}X^2\operatorname{tr}Y$	$\operatorname{tr}X^2Y$
3	2	1	$C_{3}^{2,1}$	1/15	1	4	2	8
			$C_{21}^{2,1}$	2/5	1	-1	2	-2
			$C_{21}^{1^2,1}$	$1/\sqrt{5}$	1	2	-1	-2
			$C_{1^3}^{1^2,1}$	1/3	1	-2	-1	2

complete set to $f=k+l=6$ has been tabulated and is available from the author. It is convenient for tabulation purposes to define $C_\phi^{\kappa,\lambda}(X,Y)=C_\phi^{\lambda,\kappa}(Y,X)$ when $\kappa\neq\lambda$, but this symmetry may neither be convenient nor possible to maintain in cases of multiplicity >1 when $\kappa=\lambda$. We do not therefore assume symmetry to hold in general.

5. Properties of the $C_\phi^{\kappa,\lambda}$.

Since the $C_\phi^{\kappa,\lambda}$ are linear combinations of the $\Gamma_{\phi'}^{\kappa,\lambda}$ for $\phi'\equiv\phi$, we have from Section 3:

$$C_\phi^{\kappa,\lambda}(X,X)=\theta_\phi^{\kappa,\lambda}C_\phi(X), \text{ where } \theta_\phi^{\kappa,\lambda}=C_\phi^{\kappa,\lambda}(I,I)/C_\phi(I) \qquad (5.1)$$

may be zero.

$$C_\phi^{\kappa,\lambda}(X,I)=\{\theta_\phi^{\kappa,\lambda}\,C_\phi(I)/C_\kappa(I)\}\,C_\kappa(X) \qquad (5.2)$$

and similarly for $C_\phi^{\kappa,\lambda}(I,Y)$.

$$C_\kappa^{\kappa,0}(X,Y)\overset{\text{def}}{=}C_\kappa(X),\, C_\lambda^{0,\lambda}(X,Y)\overset{r\text{def}}{=}C_\lambda(Y). \qquad (5.3)$$

The basis (4.8) allows us to write without ambiguity

$$\int_{O(m)}C_\phi^{\kappa,\lambda}(AH'XH,AH'YH)\mathrm{d}H=C_\phi^{\kappa,\lambda}(X,Y)C_\phi(A)/C_\phi(I), \qquad (5.4)$$

$$E_w C_\phi^{\kappa,\lambda}(XW,YW)=2^f\left(\tfrac{1}{2}n\right)_\phi C_\phi^{\kappa,\lambda}(X\Sigma,Y\Sigma) \qquad (5.5)$$

where $W\sim W_m(n,\Sigma,O)$. Hence using (4.13) and (5.1):

$$E_w\{C_\kappa(XW)C_\lambda(YW)\}=\sum_{\phi\in\kappa\cdot\lambda}2^f\left(\tfrac{1}{2}n\right)_\phi\theta_\phi^{\kappa,\lambda}C_\phi^{\kappa,\lambda}(X\Sigma,Y\Sigma). \qquad (5.6)$$

$$C_\phi^{\kappa,\lambda}(\alpha X,\beta Y)=\alpha^k\beta^l C_\phi^{\kappa,\lambda}(X,Y). \qquad (5.7)$$

From (4.13)

$$C_\kappa(X)C_\lambda(Y)=\sum_{\phi\in\kappa\cdot\lambda}\theta_\phi^{\kappa,\lambda}C_\phi^{\kappa,\lambda}(X,Y), \tag{5.8}$$

$$(\operatorname{tr}X)^k(\operatorname{tr}Y)^l=\sum_{\kappa,\lambda;\phi\in\kappa\cdot\lambda}\theta_\phi^{\kappa,\lambda}C_\phi^{\kappa,\lambda}(X,Y), \tag{5.9}$$

$$C_\kappa(X)C_\lambda(X)=\sum_{\phi\in\kappa\cdot\lambda}\left(\theta_\phi^{\kappa,\lambda}\right)^2C_\phi(X) \tag{5.10}$$

whence in the usual notation (e.g. Khatri and Pillai [10])

$$g_{\kappa,\lambda}^\phi=\sum_{\phi'\equiv\phi}\left(\theta_{\phi'}^{\kappa,\lambda}\right)^2. \tag{5.11}$$

Thus $g_{\kappa,\lambda}^\phi\geq 0$. If $g_{\kappa,\lambda}^\phi>0$ so that not all $\theta_{\phi'}^{\kappa,\lambda}=0(\phi'\equiv\phi)$, we may choose the first $C_\phi^{\kappa,\lambda}$ in $\mathcal{U}_{2\phi}^{\kappa,\lambda}$ to be proportional to the component of $(\operatorname{tr}X)^k(\operatorname{tr}Y)^l$ in this space with $\theta_\phi^{\kappa,\lambda}=+\sqrt{g_{\kappa,\lambda}^\phi}$, the remaining $C_{\phi'}^{\kappa,\lambda}$ having $\theta_{\phi'}^{\kappa,\lambda}=0\quad(\phi'\equiv\phi)$.

From (4.13) in an obvious notation

$$\int_{O(m)}\operatorname{etr}(AH'XH+BH'YH)\mathrm{d}H=\sum_{\kappa,\lambda;\phi}^{\infty}C_\phi^{\kappa,\lambda}(A,B)C_\phi^{\kappa,\lambda}(X,Y)/k!l!C_\phi(I). \tag{5.12}$$

This expansion may be used to derive a number of useful results, including the following.

$$\int_{O(m)}C_\phi^{\kappa,\lambda}(A'H'XHA,B)\mathrm{d}H=C_\phi^{\kappa,\lambda}(A'A,B)C_\kappa(X)/C_\kappa(I), \tag{5.13}$$

with a corresponding result for $C_\phi^{\kappa,\lambda}(A,B'H'YHB)$.

Laplace transform

$$\int_{R>0}\operatorname{etr}(-RW)|R|^{t-p}C_\phi^{\kappa,\lambda}(ARA',B)\mathrm{d}R=\Gamma_m(t,\kappa)|W|^{-t}C_\phi^{\kappa,\lambda}(AW^{-1}A',B), \tag{5.14}$$

where $p=\frac{1}{2}(m+1)$, and $\Gamma_m(t,\kappa)$ is defined in [2]. A similar result holds for $C_\phi^{\kappa,\lambda}(A, BRB')$. From (5.5), $C_\phi^{\kappa,\lambda}(XR, YR)$ has Laplace transform $\Gamma_m(t,\phi)|W|^{-t}C_\phi^{\kappa,\lambda}(XW^{-1}, YW^{-1})$.

Binomial expansion

$$C_\phi(X+Y)=\sum_{\kappa,\lambda(\phi\in\lambda)} \binom{f}{k}\theta_\phi^{\kappa,\lambda}C_\phi^{\kappa,\lambda}(X,Y), \tag{5.15}$$

and in particular

$$C_f(X+Y)=\sum_{k+l=f}\binom{f}{k}C_f^{k,l}(X,Y) \tag{5.16}$$

so that $\binom{f}{k}$ $C_f^{k,l}(X,Y)$ is given by the terms of degree k, l in X, Y respectively in $C_f(X+Y)$. It follows from (4.6) that the diagonal elements of $\Delta_{k,l}$ in Section 4 are proportional to the coefficients of the terms (4.8) in $C_f^{k,l}(X,Y)$ so that the orthogonality of the $C_\phi^{\kappa,\lambda}$ generalizes that of the C_ϕ (James [7])

$C_{1^f}^{1^k,1^l}$ may similarly be obtained from $C_{1^f}(X+Y)$.

6. Latent roots of the noncentral Wishart distribution

By way of illustration, we shall apply the invariant polynomials to this distributional problem, which was mentioned in the Introduction. Let $S\sim W_m(n,\Sigma,\Omega)$; then for real q we may write

$$f(S)=f_q(S)|q\Sigma|^{-\frac{1}{2}n}\operatorname{etr}\left\{-\tfrac{1}{2}(\Sigma^{-1}-qI)S\right\}{}_0F_1\left(\tfrac{1}{2}n;\tfrac{1}{2}\Sigma^{-\frac{1}{2}}\Omega\Sigma^{-\frac{1}{2}}S\right), \tag{6.1}$$

where $f_q(S)$ is the $W_m(n,q^{-1}I,0)$ density function, and ${}_0F_1$ is the Bessel function of matrix argument [2]. Hence the joint density of the latent roots $\Lambda=\operatorname{diag}(l_1,\ldots,l_m)$ of S is given by

$$f(\Lambda)=\left\{\pi^{\frac{1}{2}m^2}/\Gamma_m\left(\tfrac{1}{2}m\right)|q\Sigma|^{\frac{1}{2}n}\right\}f_q(\Lambda)\Pi_{i<j}(l_i-l_j)$$

$$\int_{O(m)}\operatorname{etr}\left\{-\tfrac{1}{2}(\Sigma^{-1}-qI)H'\Lambda H\right\}{}_0F_1\left(\tfrac{1}{2}n;\tfrac{1}{2}\Sigma^{-\frac{1}{2}}\Omega\Sigma^{-\frac{1}{2}}H'\Lambda H\right)dH$$

$$=\left\{\pi^{\frac{1}{2}m^2}/\Gamma_m\left(\tfrac{1}{2}m\right)|q\Sigma|^{\frac{1}{2}n}\right\}f_q(\Lambda)\Pi_{i<j}(l_i-l_j)$$

$$\sum_{\kappa,\lambda;\phi}^{\infty} (-1)^k \theta_\phi^{\kappa,\lambda} C_\phi\left(\tfrac{1}{2}\Lambda\right) C_\phi^{\kappa,\lambda}(\Sigma^{-1} - qI, \Omega\Sigma^{-1})/k!l!\left(\tfrac{1}{2}n\right)_\lambda C_\phi(I). \tag{6.2}$$

This expansion may be numerically useful if qI can be chosen close to Σ^{-1}.

References

[1] Boerner, H. (1963). *Representations of Groups.* North-Holland, Amsterdam.

[2] Constantine, A. G. (1963). Some non-central distribution problems in multivariate analysis. *Ann. Math. Statist.* 34, 1270–1285.

[3] Davis, A. W. (1976). Statistical distributions in univariate and multivariate Edgeworth populations. *Biometrika* 63, 661–670.

[4] Hannan, E. J. (1965). Group representations and applied probability. *J. Appl. Prob.* 2, 1–68.

[5] Herz, C. S. (1955). Bessel functions of matrix argument. *Ann. Math.* 61, 474–523.

[6] Hua, L. K. (1955). On the theory of functions of several complex variables. III. A complete orthonormal system in the hyperbolic space of symmetric and skew symmetric matrices. *Acta. Math. Sinica.* 5, 205–242 (In Chinese). *Amer. Math. Soc. Transl.* (2), 32, 205–242.

[7] James, A. T. (1960). The distribution of the latent roots of the covariance matrix. *Ann. Math. Statist.* 31, 151–158.

[8] James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* 35, 475–501.

[9] James, A. T. (1968). Calculations of zonal polynomial coefficients by use of the Laplace-Beltrami operator. *Ann. Math. Statist.* 39, 1711–1718.

[10] Khatri, C. G. and Pillai, K. C. S. (1968). On the noncentral distribution of two test criteria in multivariate analysis of variance. *Ann. Math. Statist.* 39, 215–226.

[11] Muirhead, R. J. (1970). Systems of partial differential equations for hypergeometric functions of matrix argument. *Ann. Math. Statist.* 41, 991–1001.

[12] Pillai, K. C. S. (1976). Distributions of characteristic roots. Mimeo Series No. 462, Department of Statistics, Purdue University.

[13] Robinson, G. De B. (1961). *Representation theory of the symmetric group.* Edinburgh University Press, Edinburgh.

[14] Saw, J. G. (1977). Zonal polynomials: an alternative approach. *J. Multivariate Analysis* 7, 461–467.

25

Submitted

Invariant polynomials with two matrix arguments extending the zonal polynomials: applications to multivariate distribution theory.

A.W.Davis

CSIRO, Division of Mathematics and Statistics, Adelaide, South Australia

1. Introduction

In a recent article [6] the author has defined a class of polynomials $C_\phi^{\kappa,\lambda}(X, Y)$ in the elements of the $m \times m$ symmetric complex matrices X and Y, having the property of invariance under the simultaneous transformations

(1.1) $$X \to H'XH, \quad Y \to H'YH, \quad H \in O(m),$$

where $O(m)$ is the group of $m \times m$ orthogonal matrices. These satisfy the basic relationship

(1.2) $$\int_{O(m)} C_\kappa(AH'XH)C_\lambda(BH'YH)dH = \sum_{\phi\in\kappa.\lambda} C_\phi^{\kappa,\lambda}(A, B)C_\phi^{\kappa,\lambda}(X, Y)/C_\phi(I),$$

where C_κ, C_λ, C_ϕ are zonal polynomials (James [12]), indexed by the ordered partitions κ, λ, ϕ of the nonnegative integers k, ℓ, $f = k + \ell$ respectively into $\leq m$ parts. Letting $G\ell(m, R)$ denote the group of $m \times m$ real nonsingular matrices, "$\phi\in\kappa.\lambda$" signifies that the irreducible representation of $G\ell(m, R)$ indexed by 2ϕ occurs in the decomposition of the Kronecker product $2\kappa \otimes 2\lambda$ of the irreducible representations indexed by 2κ and 2λ. Irreducible representations of $G\ell(m, R)$ indexed by ordered partitions of the form 2κ are fundamental to the theory of zonal polynomials. Further properties of the $C_\phi^{\kappa,\lambda}$ are summarized in Section 2.

The present paper applies the polynomials to some problems in multivariate normal distribution theory. The joint distribution of the latent roots of the noncentral Wishart matrix $S \sim W_m(n, \Sigma, \Omega)$ was presented in [6]. In Section 3 expansions are given of multivariate incomplete gamma and beta functions which are relevant to the cumulative distribution functions (c.d.f.'s) of the noncentral Wishart and MANOVA matrices (cf. Constantine [2] p. 1270). A further application is to the noncentral quadratic form (Section 7), since it is shown in Section 4 that certain polynomials in two matrices defined by Hayakawa [8] and Khatri [13] are expressible in terms of the $C_\phi^{\kappa,\lambda}$. A complete orthogonal set of Laguerre polynomials with two matrix arguments is constructed in Section 5, following the approach of Herz [11] and Constantine [3]. Section 6 presents a number of expansions generally based on the following corollary to (1.2)

$$(1.3) \qquad \int_{O(m)} \operatorname{etr}(AH'XH + BH'YH)dH = \sum\nolimits_{\kappa,\lambda;\phi}^{\infty} C_\phi^{\kappa,\lambda}(A, B)C_\phi^{\kappa,\lambda}(X, Y)/k!\ell!C_\phi(I),$$

where the summation on the right denotes $\sum_{k,\ell=0}^{\infty}\sum_{\kappa,\lambda;\phi\in\kappa.\lambda}$. Results for noncentral F in the case of unequal covariance matrices (Pillai [14], Pillai and Sudjana [15]) are obtained in Section 8, and finally the distribution of doubly noncentral multivariate F with equal covariance matrices is derived in Section 9. The latter will be applied in a subsequent paper to consider the effects of moderate nonnormality on the MANOVA tests, following the approach of Davis [5]; this objective constituted the original stimulus for the present investigation.

2. Basic properties of the invariant polynomials $C_\phi^{\kappa,\lambda}(X, Y)$

Proofs of the following results are indicated in [6]. It should be noted that a representation 2ϕ may occur in (1.2) with multiplicity greater than one, so that strictly an additional subscript is required, but we shall omit this for notational convenience, and assume that each ϕ is given the required multiplicity. Furthermore, in such cases the polynomials $C_\phi^{\kappa,\lambda}$ are not uniquely defined, but it is sufficient that they be "orthogonal" in a sense defined in [6] for (1.2) and other basic properties to hold. This non-uniqueness first occurs for polynomials of degree 6, when if $\kappa,\lambda = [2,1]$, $\phi = [3, 2, 1]$ occurs with multiplicity 3.

(2.1) $C_\phi^{\kappa,\lambda}(X, X) = \theta_\phi^{\kappa,\lambda} C_\phi(X)$, where $\theta_\phi^{\kappa,\lambda} = C_\phi^{\kappa,\lambda}(I, I)/C_\phi(I)$

may be zero.

(2.2) $C_\phi^{\kappa,\lambda}(X, I) = \{\theta_\phi^{\kappa,\lambda} C_\phi(I)/C_\kappa(I)\}C_\kappa(X)$,

with a corresponding result for $C_\phi^{\kappa,\lambda}(I, Y)$.

(2.3) $C_\kappa^{\kappa,0}(X, Y) \overset{\text{def}}{=} C_\kappa(X)$, $C_\lambda^{0,\lambda}(X, Y) \overset{\text{def}}{=} C_\lambda(Y)$.

(2.4) $\int_{O(m)} C_\phi^{\kappa,\lambda}(AH'XH, AH'YH)dH = C_\phi^{\kappa,\lambda}(X, Y)C_\phi(A)/C_\phi(I)$.

If $W \sim W_m(n, \Sigma, 0)$ then

(2.5) $E_W C_\phi^{\kappa,\lambda}(XW, YW) = 2^f(\tfrac{1}{2}n)_\phi C_\phi^{\kappa,\lambda}(X\Sigma, Y\Sigma)$ $(f = k + \ell)$,

(2.6) $E_W\{C_\kappa(XW)C_\lambda(YW)\} = \sum_{\phi\in\kappa.\lambda} 2^f(\tfrac{1}{2}n)_\phi \theta_\phi^{\kappa,\lambda} C_\phi^{\kappa,\lambda}(X\Sigma, Y\Sigma)$,

where we note that $(\tfrac{1}{2}n)_\phi$ is constant for equivalent representations 2ϕ.

(2.7) $C_\phi^{\kappa,\lambda}(\alpha X, \beta Y) = \alpha^k\beta^\ell C_\phi^{\kappa,\lambda}(X, Y)$, ($\alpha$, β complex constants).

The following are consequences of (1.2):

(2.8) $C_\kappa(X)C_\lambda(Y) = \sum_{\phi\in\kappa.\lambda} \theta_\phi^{\kappa,\lambda} C_\phi^{\kappa,\lambda}(X, Y)$,

(2.9) $(\mathrm{tr}X)^k(\mathrm{tr}Y)^\ell = \sum_{\kappa,\lambda;\phi\in\kappa.\lambda} \theta_\phi^{\kappa,\lambda} C_\phi^{\kappa,\lambda}(X, Y)$,

(2.10) $C_\kappa(X)C_\lambda(X) = \sum_{\phi^*\in\kappa.\lambda} g_{\kappa,\lambda}^{\phi^*} C_{\phi^*}(X)$, $g_{\kappa,\lambda}^{\phi^*} = \sum_{\phi\equiv\phi^*} (\theta_\phi^{\kappa,\lambda})^2$,

where $\sum_{\phi^*\in\kappa.\lambda}$ implies that we sum over the inequivalent representations $2\phi^*$ occurring in $2\kappa \otimes 2\lambda$, and $\sum_{\phi\equiv\phi^*}$ denotes summation over the representations equivalent to $2\phi^*$ in $2\kappa \otimes 2\lambda$.

(2.11) $$\int_{O(m)} C_\phi^{\kappa,\lambda}(A'H'XHA,\ B)dH = C_\phi^{\kappa,\lambda}(A'A,\ B)C_\kappa(X)/C_\kappa(I),$$

with a corresponding result for $C_\phi^{\kappa,\lambda}(A,\ B'H'YHB)$.

Laplace transform:

(2.12) $$\int_{R>0} \operatorname{etr}(-RW)|R|^{t-p}C_\phi^{\kappa,\lambda}(ARA',\ B)dR = \Gamma_m(t,\kappa)|W|^{-t}C_\phi^{\kappa,\lambda}(AW^{-1}A',B)$$

where $p = \tfrac{1}{2}(m + 1)$, and $\Gamma_m(t,\ \kappa)$ is defined in [2]. Similarly for $C_\phi^{\kappa,\lambda}(A,\ BRB')$. From (2.5), $|R|^{t-p}C_\phi^{\kappa,\lambda}(XR,\ YR)$ has Laplace transform $\Gamma_m(t,\ \phi)|W|^{-t}C_\phi^{\kappa,\lambda}(XW^{-1},\ YW^{-1})$. Setting $R = W^{-\frac{1}{2}}HSH'W^{-\frac{1}{2}}$ and integrating over $O(m)$, it follows directly from [3] equation (10) that $|R|^{t-p}\ C_\phi^{\kappa,\lambda}(XR^{-1},\ YR^{-1})$ has Laplace transform $[(-1)^f\Gamma_m(t)/(-t + p)_\phi]$ $|W|^{-t}C_\phi^{\kappa,\lambda}(XW,\ YW)$.

Binomial expansion:

(2.13) $$C_\phi(X + Y) = \sum_{\kappa,\lambda(\phi\in\kappa.\lambda)}\sum_{\phi'\equiv\phi} \binom{f}{k}\theta_{\phi'}^{\kappa,\lambda}C_{\phi'}^{\kappa,\lambda}(X,\ Y),$$

where in particular

(2.14) $$C_f(X + Y) = \sum_{k+\ell=f}\binom{f}{k}C_f^{k,\ell}(X,\ Y).$$

Thus $\binom{f}{k}C_f^{k,\ell}(X,\ Y)$ is given by the terms of degree k,ℓ in X, Y respectively in the expansion of $C_f(X + Y)$. The polynomials for $\kappa = [1^k]$, $\lambda = [1^\ell]$, $\phi = [1^f]$ may similarly be obtained from $C_{1^f}(X + Y)$.

If R, S are $r \times r$, $s \times s$ symmetric matrices respectively, $r + s = m$, then it may be shown using (1.3) and comparing coefficients of $(\operatorname{tr}R)^k(\operatorname{tr}S)^\ell$ that

(2.15) $$C_\phi^{\kappa,\lambda}\left(\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 \\ 0 & S \end{pmatrix}\right) = \theta_\phi^{\kappa,\lambda}\ z_\phi C_\kappa(R)C_\lambda(S)/z_\kappa z_\lambda$$

where $z_\phi = C_\phi(I_m)/2^m(\tfrac{1}{2}m)_\phi$ is the coefficient of $(\operatorname{tr}X)^f$ in $C_\phi(X)$. Hence from (2.13)

$$(2.16)\quad C_\phi\left(\begin{pmatrix} R & 0 \\ 0 & S \end{pmatrix}\right) = z_\phi \sum_{\kappa,\lambda(\phi\in\kappa.\lambda)} \binom{f}{k} g^{\phi}_{\kappa,\lambda} C_\kappa(R) C_\lambda(S)/z_\kappa z_\lambda,$$

so that Hayakawa's [7] coefficient $a^{\phi}_{\kappa,\lambda}$ is related to $g^{\phi}_{\kappa,\lambda}$ by

$$(2.17)\quad a^{\phi}_{\kappa,\lambda} = \binom{f}{k} g^{\phi}_{\kappa,\lambda} z_\phi / z_\kappa z_\lambda .$$

For a given k, ℓ, the $C^{\kappa,\lambda}_{\phi}(X, Y)$ are linear combinations of the distinct products of traces

$$(\mathrm{tr}\, X^{a_1} Y^{b_1} X^{c_1} \ldots)^{r_1} (\mathrm{tr}\, X^{a_2} Y^{b_2} X^{c_2} \ldots)^{r_2} \ldots$$

of total degree k, ℓ in the elements of X, Y respectively. In constructing these, account must be taken of the symmetry of X and Y, and the trace properties tr XY = tr YX, tr Z' = tr Z. The number of distinct terms of this type should thus equal the sum of multiplicities of the irreducible representations $2\phi \in 2\kappa \otimes 2\lambda$, for all ordered partitions κ, λ of k, ℓ respectively into ≤ m parts. An algorithm for determining these multiplicities is given in Robinson [16] Section 3.3.

The polynomials have been tabulated up to degree f = k + ℓ = 6, and are available from the author. Polynomials up to f = 5 are listed in the Appendix. When κ ≠ λ it is convenient for purposes of construction to define $C^{\kappa,\lambda}_{\phi}(X, Y) = C^{\lambda,\kappa}_{\phi}(Y, X)$. However, when κ = λ it may neither be convenient nor possible to insist on symmetry in cases of multiplicity > 1; thus we shall not assume symmetry to hold in general.

3. Incomplete gamma and beta functions

Incomplete gamma function: Generalizing Constantine [2] equation (60) we have

$$(3.1)\quad \int_0^X \mathrm{etr}(-AS)|S|^{t-p} C_\lambda(BS)dS = \{\Gamma_m(t)\Gamma_m(p)/\Gamma_m(t+p)\}|X|^t$$

$$\sum_{k=0}^{\infty} \sum_{\kappa;\phi\in\kappa.\lambda} (t)_\phi \, \theta_\phi^{\kappa,\lambda} C_\phi^{\kappa,\lambda}(-AX,\ BX)/k!(t+p)_\phi .$$

Proof: Let $S = X^{\frac{1}{2}}H'THX^{\frac{1}{2}}$, $H\in O(m)$. Expanding the exponential, we average over $O(m)$ and use

$$(3.2) \qquad \int_0^I |T|^{t-p} C_\phi(T)/C_\phi(I)dT = \Gamma_m(t,\ \phi)\Gamma_m(p)/\Gamma_m(t+p,\ \phi).$$

Equation (3.1) implies an expansion of the c.d.f. of the noncentral Wishart distribution $W_m(n,\ \Sigma,\ \Omega)$ (equation (4.9)), and in particular of the largest root.

Incomplete beta function

Equation (61) of [2] may similarly be generalized as follows:

$$(3.3) \quad \int_0^X |R|^{t-p}|I-R|^{u-p}\ C_\lambda(AR)dR = \{\Gamma_m(t)\Gamma_m(p)/\Gamma_m(t+p)\}|X|^t$$

$$\sum_{k=0}^{\infty} \sum_{\kappa;\phi\in\kappa.\lambda} (t)_\phi(-u+p)_\kappa\ \theta_\phi^{\kappa,\lambda}\ C_\phi^{\kappa,\lambda}(X,\ AX)/k!(t+p)_\phi .$$

This yields an expansion for the c.d.f. of the multivariate noncentral beta matrix, and in particular of the largest root.

A further result is

$$(3.4) \qquad \int_0^I |R|^{t-p}|I-R|^{u-p} C_\phi^{\kappa,\lambda}(R,\ I-R)dR$$

$$= [\Gamma_m(t,\ \kappa)\Gamma_m(u,\ \lambda)/\Gamma_m(t+u,\ \phi)]\theta_\phi^{\kappa,\lambda} C_\phi(I).$$

4. Polynomials of Hayakawa and Khatri

Explicit representations of these polynomials in terms of the $C_\phi^{\kappa,\lambda}$ will be presented in this section. In connection with the multivariate noncentral quadratic form, Hayakawa has defined a polynomial $P_\phi(T,\ A)$, which may be obtained as the coefficient of $C_\phi(UU')/f!(\frac{1}{2}m)_\phi C_\phi(I_n)$ in the generating function ([8] Theorem 7)

(4.1) $$\int_{O(m)}\int_{O(n)} \operatorname{etr}(-UH_2AH_2'U' + 2H_1UH_2A^{\frac{1}{2}}T')dH_1dH_2$$

$$= \int_{O(n)} \operatorname{etr}(-U'UH_2AH_2')\,{}_0F_1(\tfrac{1}{2}m;\ U'UH_2A^{\frac{1}{2}}T'TA^{\frac{1}{2}}H_2')dH_2$$

$$= \sum_{\kappa,\lambda;\phi}^{\infty} \frac{C_\phi(U'U)\ \theta_\phi^{\kappa,\lambda}\ C_\phi^{\kappa,\lambda}(-A,\ T'TA)}{k!\ell!(\frac{1}{2}m)_\lambda C_\phi(I_n)}$$

where U, T are m × n, A is n × n positive definite symmetric. Hence

(4.2) $$P_\phi(T,\ A) = (\tfrac{1}{2}m)_\phi \sum_{\kappa,\lambda(\phi\in\kappa.\lambda)} \sum_{\phi'\equiv\phi} \binom{f}{k}\theta_{\phi'}^{\kappa,\lambda}C_{\phi'}^{\kappa,\lambda}(-A,\ T'TA)/(\tfrac{1}{2}m)_\lambda .$$

If we set $A = I_n$, then (2.2), (2.10) and Bingham's [1] identity for the generalized binomial coefficient [3]

(4.3) $$a_{\phi,\lambda} = \binom{f}{\ell}\sum_\kappa g_{\kappa,\lambda}^{\phi}$$

implies that

(4.4) $$P_\phi(T,\ I_n) = (-1)^f\ L_\phi^{\frac{1}{2}n-p}\ (TT')$$

([7] equation (35)), where L_ϕ^t is Constantine's [3] generalized Laguerre polynomial. Similarly, using (2.11), we obtain

(4.5) $$\int_{O(m)}P(TH,\ A)dH = \int_{O(m)}P(T,\ HAH')dH = (-1)^f\ \frac{C_\phi(A)}{C_\phi(I_n)}\ L_\phi^{\frac{1}{2}n-p}\ (TT')$$

([7] equation (38)). Hayakawa [9] has tabulated the P_ϕ up to f = 4; further tabulation could be based on (4.2).

Also in connection with the noncentral quadratic form, Khatri [13] has defined a two-matrix generalization $L_\phi^t(X,\ A)$ of $L_\phi^t\ (X)$ (X, A are m × m symmetric), such that $|X|^t L_\phi^t(X,\ A)$ has Laplace transform with respect to X

(4.6) $$\Gamma_m(t + p,\ \phi)|Z|^{-t-p}C_\phi((I - Z^{-1})A)$$

$$= \Gamma_m(t + p,\ \phi)|Z|^{-t-p}\sum_{\kappa,\lambda(\phi\in\kappa.\lambda)}\sum_{\phi'\equiv\phi}\binom{f}{k}\theta_{\phi'}^{\kappa,\lambda}C_{\phi'}^{\kappa,\lambda}(a,\ -Z^{-1}A)$$

by (2.13), which may be inverted by (2.12) to yield

$$(4.7)\quad L_\phi^t(X, A) = (t+p)_\phi \sum_{\kappa,\lambda(\phi\in\kappa.\lambda)} \sum_{\phi'\equiv\phi} \binom{f}{k} \theta_{\phi'}^{\kappa,\lambda} C_{\phi'}^{\kappa,\lambda}(A, -XA)/(t+p)_\lambda .$$

From (4.2), expressions for L_ϕ^t may be obtained by replacing $\frac{1}{2}m$ by $t + p$ in $(-1)^f P_\phi$. The three-matrix polynomials defined by Crowther [4] and Khatri [13] are not obtainable using the methods of this paper.

Khatri's expansion of the Wishart density $W_m(n, \Sigma, \Omega)$ ([13] equation (4.10)) may be obtained directly using (2.8) and (4.7),

$$(4.8)\quad W_m(n, \Sigma, 0)\, \mathrm{etr}(-\Omega)\, {}_0F_1(\tfrac{1}{2}n; \tfrac{1}{2}\Sigma^{-\frac{1}{2}}S\Sigma^{-\frac{1}{2}}\Omega)$$

$$= W_m(n, \Sigma, 0) \sum_{\kappa,\lambda;\phi}^{\infty} \theta_\phi^{\kappa,\lambda} C_\phi^{\kappa,\lambda}(-\Omega, \tfrac{1}{2}\Sigma^{-\frac{1}{2}}S\Sigma^{-\frac{1}{2}}\Omega)/k!\ell!(\tfrac{1}{2}n)_\lambda$$

$$= W_m(n, \Sigma, 0) \sum_{f=0}^{\infty} \sum_\phi L_\phi^{\frac{1}{2}n-p} (\tfrac{1}{2}\Sigma^{-\frac{1}{2}}S\Sigma^{-\frac{1}{2}}, -\Omega)/f!(\tfrac{1}{2}n)_\phi .$$

The c.d.f. of this distribution may be obtained from (3.1) in the form

$$(4.9)\quad P\{S < X\} = \{\Gamma_m(p)/2^{\frac{1}{2}mn}\Gamma_m(\tfrac{1}{2}n+p)|\Sigma|^{\frac{1}{2}n}\}\, \mathrm{etr}(-\Omega)|X|^{\frac{1}{2}n}$$

$$\sum_{f=0}^{\infty} \sum_\phi (-1)^f L_\phi^{\frac{1}{2}n-p}(\Omega, \tfrac{1}{2}\Sigma^{-\frac{1}{2}}X\Sigma^{-\frac{1}{2}})/f!(\tfrac{1}{2}n+p)_\phi .$$

Taking $X = \lambda_1 I$ we obtain the c.d.f. of the largest root λ_1.

5. A complete system of generalized Laguerre polynomials with two matrix arguments

From (4.7) or Khatri's original definition, the $L_\phi^t(X, A)$ are to be regarded as a multivariate generalization of $a^f L_f^t(x)$, rather than of $L_k^t(x)\, L_\ell^u(y)$ as would be required for functions invariant under (1.1). In the present section we present such a generalization along the lines of Herz [11] and Constantine [3] . Define

$$(5.1)\quad L^{t,u}_{\kappa,\lambda;\phi}(X, Y) = \mathrm{etr}(X + Y) \int_{R>0}\int_{S>0} \mathrm{etr}(-R - S)|R|^t|S|^u C^{\kappa,\lambda}_{\phi}(R, S)$$

$$A_t(RX)A_u(SY)dRdS, \quad (t, u > -1)$$

where A_t is Herz's Bessel function (see above references), noting that the $L^{t,u}_{\kappa,\lambda;\phi}$ will depend on the particular basis $\{C^{\kappa,\lambda}_{\phi}\}$, in the spirit of Herz's definition.

Laplace transform. Noting that ([11] equation (15))

$$(5.2)\quad \int_{R>0} \mathrm{etr}(-RW)|R|^t A_t(RM)dR = |W|^{-t-p}\, \mathrm{etr}(-MW^{-1}),$$

we obtain using (2.12)

$$(5.3)\quad \int_{R>0}\int_{S>0} \mathrm{etr}(-RW - SZ)|R|^t|S|^u L^{t,u}_{\kappa,\lambda;\phi}(R, S)dRdS$$

$$= \Gamma_m(t + p,\kappa)\Gamma_m(u + p, \lambda)|W|^{-t-p}|Z|^{-u-p}C^{\kappa,\lambda}_{\phi}(I - W^{-1}, I - Z^{-1}).$$

This yields an explicit representation of the polynomials, analogous to [3] equation (20). Let

$$(5.4)\quad C^{\kappa,\lambda}_{\phi}(I + X, I + Y)/C_{\phi}(I) = \sum_{r=0}^{k}\sum_{s=0}^{\ell}\sum_{\rho,\sigma;\tau} a^{\kappa,\lambda;\phi}_{\rho,\sigma;\tau}C^{\rho,\sigma}_{\tau}(X, Y)/C_{\tau}(I)$$

where ρ, σ, τ are partitions of r, s, $r + s$ respectively. Again using (2.12),

$$(5.5)\quad L^{t,u}_{\kappa,\lambda;\phi}(X, Y) = (t + p)_{\kappa}(u + p)_{\lambda}C_{\phi}(I)\sum_{r=0}^{k}\sum_{s=0}^{\ell}$$

$$\sum_{\rho,\sigma;\tau} \frac{(-1)^{r+s}\, a^{\kappa,\lambda;\phi}_{\rho,\sigma;\theta}}{(t + p)_{\rho}(u + p)_{\sigma}} \frac{C^{\rho,\sigma}_{\tau}(X, Y)}{C_{\tau}(I)}.$$

Generating function $L^{t,u}_{\kappa,\lambda;\phi}(X, Y)$ is the coefficient of $C^{\kappa,\lambda}_{\phi}(W, Z)/k!\ell!C_{\phi}(I)$ in the expansion of

$$(5.6)\quad |I - W|^{-t-p}|I - Z|^{u-p} \int_{O(m)} \mathrm{etr}\{-XH'W(I - W)^{-1}H - YH'Z(I - Z)^{-1}H\}dH.$$

A relation with Khatri's polynomial proved by taking Laplace transforms is

(5.7) $\int_{O(m)} L^t_\kappa(HAH', X)L^u_\lambda(HBH', Y)dH = \sum_{\phi\in\kappa.\lambda} C^{\kappa,\lambda}_\phi(X, Y)L^{t,u}_{\kappa,\lambda;\phi}(A, B)/C_\phi(I)$.

<u>Orthogonality.</u> Multiplying (5.6) by $\mathrm{etr}(-X - Y)|X|^t|Y|^u C^{\rho,\sigma}_\tau(X, Y)$ and integrating over $X > 0$, $Y > 0$, we find that

(5.8) $\int_{X>0}\int_{Y>0} \mathrm{etr}(- X - Y)|X|^t|Y|^u C^{\rho,\sigma}_\tau(X, Y)\, L^{t,u}_{\kappa,\lambda;\phi}(X, Y)dXdY$

is the coefficient of $C^{\kappa,\lambda}_\phi(W, Z)/k!\ell!C_\phi(I)$ in the expansion of

(5.9) $\Gamma_m(t + p, \kappa)\Gamma_m(u + p, \lambda)C^{\rho,\sigma}_\tau(I - W, I - Z)$.

A similar argument to [3] Theorem 2 shows that (5.8) is zero for $k \geq r$, $\ell \geq s$ unless $(\kappa, \lambda;\phi) = (\rho, \sigma; \tau)$ and hence that $L^{t,u}_{\kappa,\lambda;\phi}$ is orthogonal to all $L^{t,u}_{\rho,\sigma;\tau}$ with respect to $\mathrm{etr}(-X - Y)|X|^t|Y|^u$ for $k \geq r$, $\ell \geq s$ and $(\rho, \sigma; \tau) \neq (\kappa, \lambda; \phi)$, with

(5.10) $\int_{X>0}\int_{Y>0} \mathrm{etr}(-X - Y)\, |X|^t|Y|^u\{L^{t,u}_{\kappa,\lambda;\phi}(X, Y)\}^2 dXdY$

$= k!\,\ell!\,\Gamma_m(t + p, \kappa)\Gamma_m(u + p, \lambda)C_\phi(I)$.

<u>Completeness.</u> On the basis of (5.10), we could follow Herz [11] Sections 3 and 4 in considering the Hilbert space $\mathcal{L}^2_{t,u}$ of functions $f(X, Y)$ defined for $X > 0$, $Y > 0$, such that

(5.11) $||f||^2_{t,u} = \int_{R>0}\int_{S>0} |f(R, S)|^2|R|^t|S|^u\, dRdS < \infty$,

and show that the "(t, u)-Hankel transform" of f, $g = \mathcal{U}_{t,u}f$ say, where

(5.12) $g(X, Y) \sim \int_{R>0}\int_{S>0} A_t(RX)A_u(SY)|R|^t|S|^u f(R, S)dRdS$,

has the properties of a Watson transform; i.e. $\mathcal{U}_{t,u}$ is unitary, self-adjoint and self-inversive on $\mathcal{L}^2_{t,u}$. In (5.12), $\sim$ signifies equality whenever the integral is absolutely convergent, otherwise a limit in the above norm.

The Laplace transforms F, G of f, g respectively are related by

(5.13) $G(W, Z) = |W|^{-t-p}|Z|^{-u-p} F(W^{-1}, Z^{-1})$.

Hence if we set

(5.14) $\ell^{t,u}_{\kappa,\lambda;\sigma}(X, Y) = \text{etr}(-X-Y)\, L^{t,u}_{\kappa,\lambda;\phi}(2X, 2Y)$

it follows from (5.1) that $\ell^{t,u}_{\kappa,\lambda;\phi} \in L^2_{t,u}$, while from (5.3) and (5.13) $\ell^{t,u}_{\kappa,\lambda;\phi}$ and $(-1)^f U_{t,u}\ell^{t,u}_{\kappa,\lambda;\phi}$ each has double Laplace transform

(5.15) $\Gamma_m(t + p, \kappa)\Gamma_m(u + p, \lambda)|I + W|^{-t-p}|I + Z|^{-u-p}C^{\kappa,\lambda}_{\phi}\left(\frac{W - I}{W + I}, \frac{Z - I}{Z + I}\right)$.

Thus the $\ell^{t,u}_{\kappa,\lambda;\phi}$ are orthogonal eigenfunctions of $U_{t,u}$ with eigenvalues $(-1)^f$. We may now show that they are complete in the closed subspace $\tilde{L}^2_{t,u}$ consisting of the functions in $L^2_{t,u}$ invariant under (1.1) ([11] p.501). Suppose that $f \in \tilde{L}^2_{t,u}$ and that

(5.16) $\int_{R>0}\int_{S>0} \ell^{t,u}_{\kappa,\lambda;\phi}(R,S)f(R,S)\,|R|^t|S|^u\, dRdS = 0$ for all $\kappa, \lambda; \phi$.

Then since each $C^{\kappa,\lambda}_{\phi}(R, S)$ can be expressed as a linear combination of $L^{t,u}_{\kappa,\lambda;\phi}$ and lower degree Laguerre polynomials,

(5.17) $\omega_{\kappa,\lambda;\phi} = \int_{R>0}\text{etr}(-R - S)C^{\kappa,\lambda}_{\phi}(R, S)f(R, S)\,|R|^t|S|^u\, dRdS$

$= 0$ for all $\kappa, \lambda; \phi$.

Hence, since f satisfies (1.1) its double Laplace transform can be written

(5.18) $F(W, Z) = \int_{R>0}\int_{S>0} \text{etr}(-R - S)f(R, S)|R|^t|S|^u \int_{O(m)}\text{etr}\{(I - W)H'RH + (I - Z)H'SH\}dH$

$= \sum^{\infty}_{\kappa,\lambda;\phi} \omega_{\kappa,\lambda;\phi}C^{\kappa,\lambda}_{\phi}(I - W, I - Z)/k!\ell!C_{\phi}(I)$

$= 0$

for $0 \leq \text{Re } W, \text{Re } Z < I$. But F is complex analytic for Re W > 0, Re Z > 0; hence F = 0, and f is a null function.

6. Some useful expansions

(a) Multiplying both sides of (1.3) by etr(-X),

$$(6.1)\qquad \mathrm{etr}(-X)\sum_{\kappa,\lambda;\phi}^{\infty} \frac{C_\phi^{\kappa,\lambda}(A,\ B)C_\phi^{\kappa,\lambda}(X,\ Y)}{k!\ell!C_\phi(I)} = \sum_{\kappa,\lambda;\phi}^{\infty} \frac{C_\phi^{\kappa,\lambda}(A-I,\ B)(C_\phi^{\kappa,\lambda}(X,\ Y)}{k!\ell!C_\phi(I)} .$$

A number of expansions may be derived from (6.1) by taking and inverting Laplace transforms.

Substitute $X \to Z^{\frac{1}{2}}XZ^{\frac{1}{2}}$, $Y \to Z^{\frac{1}{2}}YZ^{\frac{1}{2}}$, multiply by $\mathrm{etr}(-Z)|Z|^{a-p}$ and integrate over $Z > 0$; then

$$(6.2)\quad |I+X|^{-a}\sum_{\kappa,\lambda;\phi}^{\infty} (a)_\phi C_\phi^{\kappa,\lambda}(A,\ B)C_\phi^{\kappa,\lambda}(X(I+X)^{-1},\ Y(I+X)^{-1})/k!\ell!C_\phi(I)$$

$$= \sum_{\kappa,\lambda;\phi}^{\infty} (a)_\phi C_\phi^{\kappa,\lambda}(A-I,\ B)C_\phi^{\kappa,\lambda}(X,\ Y)/k!\ell!C_\phi(I).$$

On the other hand, replacing B by B^{-1}, multiplying by $|B|^{-u}$ and inverting the Laplace transform yields

$$(6.3)\quad \mathrm{etr}(-X)\sum_{\kappa,\lambda;\phi}^{\infty} C_\phi^{\kappa,\lambda}(A,\ B)C_\phi^{\kappa,\lambda}(X,\ Y)/k!\ell!(u)_\lambda C_\phi(I)$$

$$= \sum_{\kappa,\lambda;\phi}^{\infty} C_\phi^{\kappa,\lambda}(A-I,\ B)C_\phi^{\kappa,\lambda}\ (X,\ Y)/k!\ell!(u)_\lambda C_\phi(I).$$

To obtain a result used later, set $Y = X$, $B \to -AB$ in the R.H.S., multiply by $\mathrm{etr}\{-B(I+Z)\}|B|^{u-p}/\Gamma_m(u)$ and integrate over $B > 0$; using the binomial expansion (2.13), the transform is

$$(6.4)\qquad |I+Z|^{-u}\,{}_0F_0^{(m)}(I - AZ(I+Z)^{-1},\ -X).$$

This may be expanded in terms of Constantine's Laguerre polynomials; applying the binomial expansion again, and inverting the Laplace transform, we obtain

$$(6.5)\quad \mathrm{etr}(-B)\ \sum_{\kappa,\lambda;\phi}^{\infty} C_\phi(X)\theta_\phi^{\kappa,\lambda}C_\phi^{\kappa,\lambda}(A-I,\ -AB)/k!\ell!(u)_\lambda C_\phi(I)$$

$$= |A|^{-u}\ \sum_{\kappa,\lambda;\phi}^{\infty} L_\phi^{u-p}(-X)\theta_\phi^{\kappa,\lambda}\ C_\phi^{\kappa,\lambda}(I-A^{-1},\ -A^{-1}B)/k!\ell!(u)_\lambda C_\phi(I).$$

(b) We now show that $C_\phi^{\kappa,\lambda}(I + A, B)$ has an expansion of the form

(6.6) $C_\phi^{\kappa,\lambda}(I + A, B)/C_\phi(I) = \sum_{r=0}^{k} \sum_{\rho;\tau\in\rho.\lambda} b_{\rho,\lambda;\tau}^{\kappa,\lambda;\phi}\, C_\tau^{\rho,\lambda}(A, B)/C_\tau(I).$

Proof: There certainly exists an expansion

(6.7) $C_\phi^{\kappa,\lambda}(I + A, B)/C_\phi(I) = \sum_{r=0}^{k} \sum_{s=0}^{\ell} \sum_{\rho,\sigma;\tau} b_{\rho,\sigma;\tau}^{\kappa,\lambda;\phi}\, C_\tau^{\rho,\sigma}(A, B)/C_\tau(I),$

so that from (2.11)

$$\int_{O(m)} C_\phi^{\kappa,\lambda}(I + A, B'H'XHB)dH/C_\phi(I)$$
$$= \sum_{r=0}^{k}\sum_{s=0}^{\ell} \sum_{\rho,\sigma;\tau} b_{\rho,\sigma;\tau}^{\kappa,\lambda;\phi}\, C_\tau^{\rho,\sigma}(A, B'B)C_\sigma(X)/C_\sigma(I)C_\tau(I).$$

But by (2.11) the L.H.S. is also

(6.8) $C_\phi^{\kappa,\lambda}(I + A, B'B)C_\lambda(X)/C_\lambda(I)C_\phi(I)$

$$= \{C_\lambda(X)/C_\lambda(I)\}\sum_{r=0}^{k}\sum_{s=0}^{\ell}\sum_{\rho,\sigma;\tau} b_{\rho,\sigma;\tau}^{\kappa,\lambda;\phi}\, C_\tau^{\rho,\sigma}(A, B'B)/C_\tau(I),$$

whence if $b_{\rho,\sigma;\tau}^{\kappa,\lambda;\phi} \neq 0$ we must have $\sigma = \lambda$. Q.E.D.

Thus, we may define another generalized Laguerre polynomial

(6.9) $L_{\kappa,\lambda;\phi}^{t}(X, Y) = \operatorname{etr}(X)\int_{R>0}\operatorname{etr}(-R)|R|^t C_\phi^{\kappa,\lambda}(R, Y)A_t(RX)dR$

with Laplace transform

(6.10) $\int_{X>0} \operatorname{etr}(- XW)|X|^t\, L_{\kappa,\lambda;\phi}^{t}(X, Y)dX = \Gamma_m(t + p, \kappa)\, |W|^{-t-p}C_\phi^{\kappa,\lambda}(I - W^{-1}, Y)$

so that

(6.11) $L_{\kappa,\lambda;\phi}^{t}(X, Y) =$

$$(t + p)_\kappa C_\phi(I)\sum_{r=0}^{k} \sum_{\rho;\tau\in\rho.\lambda}(-1)^r\, b_{\rho,\lambda;\tau}^{\kappa,\lambda;\phi}\, C_\tau^{\rho,\lambda}(X, Y)/(t + p)_\rho C_\tau(I).$$

Replace A, B in (6.2) by A^{-1}, B^{-1} and multiply by $|A|^{-t}|B|^{-u}$; inverting the Laplace transform

(6.12) $|I + X|^{-a}\sum_{\kappa,\lambda;\phi}^{\infty}(a)_\phi C_\phi^{\kappa,\lambda}(A, B)C_\phi^{\kappa,\lambda}(X(I + X)^{-1}, Y(I + X)^{-1})/k!\ell!(t)_\kappa(u)_\lambda C_\phi(I$

$$= \sum_{\kappa,\lambda;\phi}^{\infty} (a)_\phi C_\phi^{\kappa,\lambda}(-X, Y)L_{\kappa,\lambda;\phi}^{t-p}(A, B)/k!\ell!(t)_\kappa(u)_\lambda C_\phi(I).$$

(c) Multiplying (1.3) by etr(-X - Y),

$$(6.13)\qquad \mathrm{etr}(-X - Y)\sum_{\kappa,\lambda;\phi}^{\infty} C_{\phi}^{\kappa,\lambda}(A,\ B)C_{\phi}^{\kappa,\lambda}(X,\ Y)/k!\ell!C_{\phi}(I)$$

$$= \sum_{\kappa,\lambda;\phi}^{\infty} C_{\phi}^{\kappa,\lambda}(A - I,\ B - I)C_{\phi}^{\kappa,\lambda}(X,\ Y)/k!\ell!C_{\phi}(I).$$

Replace A, B by A^{-1}, B^{-1}, multiply by $|A|^{-t}|B|^{-u}$ and invert the Laplace transform:

$$(6.14)\qquad \mathrm{etr}(-X - Y)\int_{O(m)} {}_0F_1(t;\ AHXH')\,{}_0F_1(u;\ BHYH')dH$$

$$= \sum_{\kappa,\lambda;\phi}^{\infty}(-1)^{f}C_{\phi}^{\kappa,\lambda}(X,\ Y)\ L_{\kappa,\lambda;\phi}^{t-p,u-p}(A,\ B)/k!\ell!(t)_{\kappa}(u)_{\lambda}C_{\phi}(I).$$

(d) Expanding each side of $|I - (X + Y)|^{-a} =$ $|I + Z|^{-a}|I - (X + Z + Y)(I + Z)^{-1}|^{-a}$ and using (2.13),

$$(6.15)\qquad \sum_{\kappa,\lambda;\phi}^{\infty}(a)_{\phi}\ \theta_{\phi}^{\kappa,\lambda}\ C_{\phi}^{\kappa,\lambda}(X,\ Y)/k!\ell!$$

$$= |I + Z|^{-a}\sum_{\kappa,\lambda;\phi}^{\infty}(a)_{\phi}\ \theta_{\phi}^{\kappa,\lambda}C_{\phi}^{\kappa,\lambda}\left(\frac{X + Z}{I + Z},\ \frac{Y}{I + Z}\right)/k!\ell!\ .$$

Replace Y by Y^{-1}, multiply by $|Y|^{-u}$ and invert the Laplace transform:

$$(6.16)\qquad \sum_{\kappa,\lambda;\phi}^{\infty}(a)_{\phi}\ \theta_{\phi}^{\kappa,\lambda}\ C_{\phi}^{\kappa,\lambda}(X,\ Y)/k!\ell!(u)_{\lambda}$$

$$= |I + Z|^{-a}\sum_{\kappa,\lambda;\phi}^{\infty}(a)_{\phi}\ \theta_{\phi}^{\kappa,\lambda}C_{\phi}^{\kappa,\lambda}\left(\frac{X + Z}{I + Z},\ \frac{Y}{I + Z}\right)/k!\ell!(u)_{\lambda}.$$

Taking X = -Z,

$$(6.17)\qquad |I + Z|^{-a}\,{}_1F_1(a;u;Y(I + Z)^{-1}) = \sum_{\kappa,\lambda;\phi}^{\infty}(a)_{\phi}\ \theta_{\phi}^{\kappa,\lambda}C_{\phi}^{\kappa,\lambda}(-Z,Y)/k!\ell!(u)_{\lambda}$$

$$= \sum_{f=0}^{\infty}\sum_{\phi}(-1)^{f}(a)_{\phi}L_{\phi}^{u-p}(YZ^{-1},\ Z)/f!(u)_{\phi}.$$

In particular, if a = u,

$$(6.18)\qquad |I + Z|^{-a}\ \mathrm{etr}\,[YZ(I + Z)^{-1}] = \sum_{f=0}^{\infty}\sum_{\phi}(-1)^{f}L_{\phi}^{a-p}(Y,\ Z)/f!.$$

Taking $Z = -A^{\frac{1}{2}}H'XHA^{\frac{1}{2}}$ and integrating over O(m) yields Khatri's generating function ([13] equation (2.26)). Equation (6.18) also provides an expansion of the noncentral Wishart distribution ([2] equation (34)).

Setting $Y = A^{\frac{1}{2}}H'XHA^{\frac{1}{2}}$ in (6.18) and integrating over O(m), we obtain

$$(6.19)\quad |I - Z|^{-a}C_{\lambda}(A(I - Z)^{-1}) = (a)_{\lambda}^{-1} \sum_{k=0}^{\infty} \sum_{\kappa,\phi\in\kappa.\lambda} (a)_{\phi}\, \theta_{\phi}^{\kappa,\lambda} C_{\phi}^{\kappa,\lambda}(Z, A)/k!$$

7. Noncentral quadratic form

Suppose that X is $N_{m,n}(M, I_n \otimes \Sigma)$ and A is n × n positive definite symmetric. The quadratic form $\Sigma^{-\frac{1}{2}}XAX'\Sigma^{-\frac{1}{2}} = YY'$, where $Y = \Sigma^{-\frac{1}{2}}XA^{\frac{1}{2}}$, has the same latent roots $\lambda_1 \geq \dots \geq \lambda_m \geq 0$ as $\tilde{Y}\tilde{Y}'$, where $\tilde{Y} = HYK$ $(H \in O(m),\ K \in O(n))$. Integrating the density of $\tilde{Y}$ over O(m), O(n) we obtain

$$(7.1)\quad f(\tilde{Y}) = (2\pi)^{-\frac{1}{2}mn}|A|^{-\frac{1}{2}m}\operatorname{etr}(-\Omega)\sum_{\kappa,\lambda;\phi}^{\infty} C_{\phi}(\tfrac{1}{2}\tilde{Y}\tilde{Y}')\theta_{\phi}^{\kappa,\lambda}C_{\phi}^{\kappa,\lambda}(-A^{-1}, \tfrac{1}{2}M'\Sigma^{-1}MA^{-1}) / k!\ell!(\tfrac{1}{2}m)_{\lambda}C_{\phi}(I_n)$$

$$= (2\pi)^{-\frac{1}{2}mn}|A|^{-\frac{1}{2}m}\operatorname{etr}(-\Omega)\sum_{f=0}^{\infty}\sum_{\phi} C_{\phi}(\tfrac{1}{2}\tilde{Y}\tilde{Y}')P_{\phi}(\frac{1}{\sqrt{2}}\Sigma^{-\frac{1}{2}}M, A) / f!(\tfrac{1}{2}n)_{\phi}C_{\phi}(I_m)$$

by (4.2) (Hayakawa [8], Theorem 8). The density of $\Lambda = \operatorname{diag}(\lambda_i)$ follows directly by the usual method. From (6.3), the joint density of the roots may be written in the following form for real $q > 0$,

$$(7.2)\quad f(\Lambda) = [\pi^{\frac{1}{2}m^2}/\Gamma_m(\tfrac{1}{2}m)]\ |qA|^{-\frac{1}{2}m}\ \operatorname{etr}(-\Omega)W_m(n, q^{-1}I, 0)\Pi_{i<j}(\lambda_i - \lambda_j)$$
$$\sum_{\kappa,\lambda;\phi}^{\infty} C_{\phi}(\tfrac{1}{2}q\Lambda)C_{\phi}^{\kappa,\lambda}(I - q^{-1}A^{-1}, \tfrac{1}{2}q^{-1}M'\Sigma^{-1}MA^{-1})/k!\ell!(\tfrac{1}{2}m)_{\lambda}C_{\phi}(I_n),$$

(Hayakawa [9] Theorem 2). Setting

$$(7.3)\quad m \to n,\ A \to q^{-1}A^{-1},\ B = \tfrac{1}{2}M'\Sigma^{-1}M,\ X = -\tfrac{1}{2}q\Lambda,\ u = \tfrac{1}{2}m$$

in (6.5), we obtain Khatri's [13] equation (5.2) in the form

(7.4) $[\pi^{\frac{1}{2}m^2}/\Gamma_m(\frac{1}{2}m)]\ W_m(n,\ q^{-1}I,\ 0)\ \Pi_{i<j}(\lambda_i - \lambda_j)$

$$\sum_{\kappa,\lambda;\phi}^{\infty} L_\phi^{\frac{1}{2}n-p}(\tfrac{1}{2}q\Lambda)\theta_\phi^{\kappa,\lambda} C_\phi^{\kappa,\lambda}(I - q\Lambda,\ -\tfrac{1}{2}q\Lambda M'\Sigma^{-1}M)/k!\ell!(\tfrac{1}{2}m)_\lambda C_\phi(I_n),$$

expressible also in terms of the P_ϕ. It may be noted that this expression facilitates the derivation of Hayakawa's [10] asymptotic expansion of the distribution of $\sqrt{n/2m}\ \log|q\Lambda/n|$, since the characteristic function may be obtained from

(7.5) $$E_\Lambda\{|q\Lambda/n|^\theta\} = [\Gamma_n(\tfrac{1}{2}n + \theta)/(\tfrac{1}{2}n)^{m\theta}\Gamma_n(\tfrac{1}{2}n)]$$

$$\sum_{\kappa,\lambda;\phi}^{\infty} e_\phi(\tfrac{1}{2}m)_\phi\ \theta_\phi^{\kappa,\lambda}\ C_\phi^{\kappa,\lambda}(I - B,\ -B\tilde{\Omega})/k!\ell!(\tfrac{1}{2}m)_\lambda$$

where

(7.6) $$B = q\Lambda,\ \tilde{\Omega} = \tfrac{1}{2}M'\Sigma^{-1}M,$$

$$e_\phi = E_W\{L_\phi^{\frac{1}{2}n-p}(\tfrac{1}{2}W)/(\tfrac{1}{2}n)_\phi C_\phi(I_m)\}\ ,\ W \sim W_m(n + 2\theta,\ I,\ 0),$$

$$= \sum_{r=0}^{f}\sum_\rho (-1)^r a_{\phi,\rho}(\tfrac{1}{2}n + \theta)_\rho/(\tfrac{1}{2}n)_\rho .$$

With $\theta = it\sqrt{n/2m}$, we have $e_\phi = O(n^{-\frac{1}{2}f})$, at least for small f.

8. Noncentral F with unequal covariance matrices

In the present section we consider the joint distribution of the latent roots $\lambda_1 \geq \dots \geq \lambda_m \geq 0$ of $F = S_2^{-\frac{1}{2}}S_1S_2^{-\frac{1}{2}}$, where $S_1 \sim W_m(n_1, \Sigma_1, \Omega)$ and $S_2 \sim W_m(n_1, \Sigma_2, 0)$, $\Sigma_1 \neq \Sigma_2$. This distribution has been investigated by Pillai [14], and Pillai and Sudjana [15], under certain "randomness" assumptions on the parameter matrices which facilitate the application of zonal polynomials, and also provide scope for an exact study of the robustness of some standard test criteria against nonnormality and inequality of covariance matrices.

Pillai [14] showed that the roots of F have the same distribution as those of $\tilde{F}$ with density function

$$(8.1)\quad f(\tilde{F}) = C_1(m;\, n_1,\, n_2)\operatorname{etr}(-\Omega)|\Psi|^{-\frac{1}{2}n_1}\ |\tilde{F}|^{\frac{1}{2}n_1-p}|I + \Psi^{-1}\tilde{F}|^{-\frac{1}{2}(n_1+n_2)}$$

$$ {}_1F_1(\tfrac{1}{2}(n_1 + n_2),\, \tfrac{1}{2}n_1;\ \Phi\tilde{F}(I + \Psi^{-1}\tilde{F})^{-1})$$

where

$$(8.2)\quad \Psi = \Sigma_1^{\frac{1}{2}}\Sigma_2^{-1}\Sigma_1^{\frac{1}{2}},\qquad \Phi = \Psi^{-\frac{1}{2}}\Omega\Psi^{-\frac{1}{2}},$$

$$C_1(m;\, n_1,\, n_2) = \Gamma_m(\tfrac{1}{2}(n_1 + n_2))/\Gamma_m(\tfrac{1}{2}n_1)\Gamma_m(\tfrac{1}{2}n_2).$$

From (6.17), (8.1) may be expressed in terms of Khatri's Laguerre polynomials of two matrix arguments,

$$(8.3)\quad f(\tilde{F}) = C_1(m;\, n_1,\, n_2)\ \operatorname{etr}(-\Omega)|\Psi|^{-\frac{1}{2}n_1}|\tilde{F}|^{\frac{1}{2}n_1-p}$$

$$\sum_{\kappa,\lambda;\phi}^{\infty} (\tfrac{1}{2}(n_1 + n_2))_\phi\ \theta_\phi^{\kappa,\lambda}C_\phi^{\kappa,\lambda}(-\Psi^{-1}\tilde{F},\ \Phi\tilde{F})/k!\ell!(\tfrac{1}{2}n_1)_\lambda$$

$$(8.4)\qquad = C_1(m;\, n_1,\, n_2)\operatorname{etr}(-\Omega)\ |\Psi|^{-\frac{1}{2}n_1}\ |\tilde{F}|^{\frac{1}{2}n_1-p}$$

$$\sum_{f=0}^{\infty} \sum_\phi (-1)^f(\tfrac{1}{2}(n_1 + n_2))_\phi\ L_\phi^{\frac{1}{2}n_1-p}(\Omega,\ \Psi^{-\frac{1}{2}}\ \tilde{F}\Psi^{-\frac{1}{2}})/f!(\tfrac{1}{2}n_1)_\phi.$$

Applying (6.16), with $Z = q\tilde{F}$ for suitable real q, we obtain

$$(8.5)\quad f(\tilde{F}) = C_1(m;\, n_1,\, n_2)\operatorname{etr}(-\Omega)|\Psi|^{-\frac{1}{2}n_1}|\tilde{F}|^{\frac{1}{2}n_1-p}|I + q\tilde{F}|^{-\frac{1}{2}(n_1+n_2)}$$

$$\sum_{\kappa,\lambda;\phi}^{\infty} (\tfrac{1}{2}(n_1+ n_2))_\phi\ \theta_\phi^{\kappa,\lambda}C_\phi^{\kappa,\lambda}((I - q^{-1}\ \Psi^{-1})\tilde{B},\ q^{-1}\Phi\tilde{B})/k!\ell!(\tfrac{1}{2}n_1)_\lambda$$

where $\tilde{B} = q\tilde{F}(I + q\tilde{F})^{-1}$. Expansions for the distribution of $\Lambda = \operatorname{diag}(\lambda_i)$ may be obtained from (8.3), (8.4) and (8.5).

It may be shown by taking a Laplace transform with respect to B, or from (6.3) with X = Y using Bingham's identity (4.3), that

$$(8.6)\quad \sum_{\kappa,\lambda(\phi\in\kappa.\lambda)} \binom{f}{k}\phi_\phi^{\kappa,\lambda}\ C_\phi^{\kappa,\lambda}(I - A,\ B)/(t)_\lambda$$

$$= C_\phi(I)\sum_{\rho,\sigma;\nu} a_{\phi,\nu}\binom{n}{r}\theta_\nu^{\rho,\sigma}C_\nu^{\rho,\sigma}(-A,\ B)/(t)_\sigma C_\nu(I).$$

Hence from (8.5)

$$(8.7)\quad f(\Lambda) = C_2(m;\, n_1,\, n_2)\operatorname{etr}(-\Omega)|\Psi|^{-\frac{1}{2}n_1}|\Lambda|^{\frac{1}{2}n_1-p}|I + q\Lambda|^{-\frac{1}{2}(n_1+n_2)}\Pi_{i<j}(\lambda_i - \lambda_j)$$

$$\sum_{f=0}^{\infty}\sum_{\phi}\{(\tfrac{1}{2}(n_1 + n_2))_{\phi}C_{\phi}(q\Lambda/I + q\Lambda)/f!\}$$

$$\sum_{n=0}^{f}\sum_{\nu} a_{\phi,\nu}\ (-1)^n L_{\nu}^{\frac{1}{2}n_1-p}(\Omega,\, q^{-1}\Psi^{-1})/(\tfrac{1}{2}n_1)_{\nu}C_{\nu}(I),$$

where $C_2(m;\, n_1,\, n_2) = \pi^{\frac{1}{2}m^2}C_1(m;\, n_1,\, n_2)/\Gamma_m(\frac{1}{2}m)$. Taking Ω and Ψ "random" in Pillai's sense, we obtain his [14] equations (3.2) and (3.7) from (8.3) and (8.7), respectively, using (2.11).

<u>Hotelling trace</u> $T = \operatorname{tr} F$. For $S > 0$ let $s = \operatorname{tr} S$, $S = sS_1$; then

$$(8.7)\qquad \int_{S_1} L_{\phi}^{t}(X,\, A^{\frac{1}{2}}SA^{\frac{1}{2}})|S_1|^b dS_1 = s^f\Gamma_m(b,\, \phi)L_{\phi}^{t}(X,\, A)/\Gamma(bm + f).$$

This may be obtained using an unpublished result of A.T.James

$$(8.8)\qquad \int_{S_1} C_{\phi}(AS_1)|S_1|^b dS_1 = \Gamma_m(b,\, \phi)C_{\phi}(A)/\Gamma(bm + f).$$

James has also shown that $|S|^{-p}dS = dS_1 ds/s$; hence from (8.3),

$$(8.9)\qquad f(T) = C_3(m;\, n_1,\, n_2)\operatorname{etr}(-\Omega)|\Psi|^{-\frac{1}{2}n_1}$$

$$T^{\frac{1}{2}mn_1-1}\sum_{f=0}^{\infty}[(-T)^f/f!(\tfrac{1}{2}mn_1)_f]\sum_{\phi}(\tfrac{1}{2}(n_1 + n_2))_{\phi}L_{\phi}^{\frac{1}{2}n_1-p}(\Omega,\, \Psi^{-1}),$$

where $C_3(m;\, n_1,\, n_2) = \Gamma_m(\frac{1}{2}(n_1+ n_2))/\Gamma(\frac{1}{2}mn_1)\Gamma_m(\frac{1}{2}n_2)$; this extends Constantine [3] equation (1) to the case of unequal covariance matrices.

Using (3.3), the c.d.f. of the largest root f_1 of F may be obtained from (3.1) in the form

$$(8.10)\ P\{f_1 < x\} = C_4(m;\, n_1,\, n_2)\operatorname{etr}(-\Omega)|\Psi|^{-\frac{1}{2}n_1}(q^{-1}\beta)^{\frac{1}{2}mn_1}\sum_{f=0}^{\infty}\sum_{\phi}[(\tfrac{1}{2}n_1)_{\phi}$$

$$C_{\phi}(\beta I)/f!(\tfrac{1}{2}n_1 + p)_{\phi}]\sum_{\kappa,\lambda(\phi\in\kappa.\lambda)}\binom{f}{k}g_{\kappa,\lambda}^{\phi}$$

$$(-\tfrac{1}{2}n_2 + p)_{\kappa}(\tfrac{1}{2}(n_1 + n_2))_{\lambda}A_{\lambda}$$

where $\beta = qx/(1 + qx)$, $q > 0$ real, $C_4(m;\ n_1,\ n_2) = \Gamma_m(\frac{1}{2}(n_1 + n_2))\Gamma_m(p)$
$/\Gamma_m(\frac{1}{2}n_2)\Gamma_m(\frac{1}{2}n_1 + p)$,

and

$$(8.11)\quad A_\lambda = \sum_{\rho,\sigma(\lambda\in\rho.\sigma)} \binom{\ell}{r}\theta_\lambda^{\rho,\sigma} C_\lambda^{\rho,\sigma}(I - q^{-1}\Psi^{-1},\ q^{-1}\Phi)/(\tfrac{1}{2}n_1)_\sigma.$$

9. Doubly noncentral F

Finally, we consider the latent roots of multivariate F when $S_1 \sim W_m(n_1,\ I,\ \Omega_1)$, $S_2 \sim W_m(n_2,\ I,\ \Omega_2)$.

(a) Case $n_1 \geq m$. F has the same roots as $\tilde{F} = \tilde{S}_2^{-\frac{1}{2}}\tilde{S}_1\tilde{S}_2^{-\frac{1}{2}}$, where $\tilde{S}_1 = H'S_1H$, $\tilde{S}_2 = H'S_2H$; averaging over $O(m)$,

$$(9.1)\quad f(\tilde{S}_1,\ \tilde{S}_2)$$
$$= K(m;\ n_1,\ n_2)\mathrm{etr}(-\Omega_1 - \Omega_2)\mathrm{etr}[-\tfrac{1}{2}(\tilde{S}_1 + \tilde{S}_2)]|\tilde{S}_1|^{\frac{1}{2}n_1-p}|\tilde{S}_2|^{\frac{1}{2}n_2-p}$$
$$\int_{O(m)} {}_0F_1(\tfrac{1}{2}n_1;\ \tfrac{1}{2}\Omega_1 H'\tilde{S}_1H)\,{}_0F_1(\tfrac{1}{2}n_2;\ \tfrac{1}{2}\Omega_2 H'\tilde{S}_2H)dH$$

$$(9.2)\quad = K(m;\ n_1,\ n_2)\mathrm{etr}(-\Omega_1 - \Omega_2)\mathrm{etr}[-\tfrac{1}{2}(\tilde{S}_1 + \tilde{S}_2)]|\tilde{S}_1|^{\frac{1}{2}n_1-p}|\tilde{S}_2|^{\frac{1}{2}n_2-p}$$
$$\sum_{\kappa,\lambda;\phi}^{\infty} C_\phi^{\kappa,\lambda}(\Omega_1,\ \Omega_2)C_\phi^{\kappa,\lambda}(\tfrac{1}{2}\tilde{S}_1,\ \tfrac{1}{2}\tilde{S}_2)/k!\ell!(\tfrac{1}{2}n_1)_\kappa(\tfrac{1}{2}n_2)_\lambda C_\phi(I)$$

$$(9.3)\quad = K(m;\ n_1,\ n_2)\mathrm{etr}[-\tfrac{1}{2}(\tilde{S}_1 + \tilde{S}_2)]|S_1|^{\frac{1}{2}n_1-p}|\tilde{S}_2|^{\frac{1}{2}n_2-p}$$
$$\sum_{\kappa,\lambda;\phi}^{\infty}(-1)^f C_\phi^{\kappa,\lambda}(\Omega_1,\ \Omega_2)L_{\kappa,\lambda;\phi}^{\frac{1}{2}n_1-p,\frac{1}{2}n_2-p}(\tfrac{1}{2}S_1,\ \tfrac{1}{2}S_2)/k!\ell!(\tfrac{1}{2}n_1)_\kappa(\tfrac{1}{2}n_2)_\lambda C_\phi(I)$$

by (6.14), where $K(m;\ n_1,\ n_2) = [2^{\frac{1}{2}m(n_1+n_2)}\Gamma_m(\frac{1}{2}n_1)\Gamma_m(\frac{1}{2}n_2)]^{-1}$. Hence we obtain

$$(9.4)\quad f(\tilde{F}) = C_1(m;\ n_1,\ n_2)\mathrm{etr}(-\Omega_1 - \Omega_2)|\tilde{F}|^{\frac{1}{2}n_1-p}|I + \tilde{F}|^{-\frac{1}{2}(n_1+n_2)}$$
$$\sum_{\kappa,\lambda;\phi}^{\infty}(\tfrac{1}{2}(n_1+n_2))_\phi C_\phi^{\kappa,\lambda}(\Omega_1,\ \Omega_2)C_\phi^{\kappa,\lambda}(\tilde{F}(I + \tilde{F})^{-1},\ (I + \tilde{F})^{-1})$$
$$/k!\ell!(\tfrac{1}{2}n_1)_\kappa(\tfrac{1}{2}n_2)_\lambda C_\phi(I)$$

(9.5) $$= C_1(m;\ n_1,\ n_2)|\tilde{F}|^{\frac{1}{2}n_1-p}|I + \tilde{F}|^{-\frac{1}{2}(n_1+n_2)}$$

$$\sum^{\infty}_{\kappa,\lambda;\phi} (-1)^f C^{\kappa,\lambda}_{\phi}(\Omega_1,\ \Omega_2)E_S\{L^{\frac{1}{2}n_1-p,\frac{1}{2}n_2-p}_{\kappa,\lambda;\phi}(\tfrac{1}{2}FS,\ \tfrac{1}{2}S)\}$$

$$/k!\ell!(\tfrac{1}{2}n_1)_{\kappa}(\tfrac{1}{2}n_2)_{\lambda}C_{\phi}(I)$$

where $S \sim W_m(n_1 + n_2,\ (I + \tilde{F})^{-1},\ 0)$. From (5.5),

(9.6) $$E_S\{L^{\frac{1}{2}n_1-p,\frac{1}{2}n_2-p}_{\kappa,\lambda;\phi}(\tfrac{1}{2}FS,\ \tfrac{1}{2}S)\} = (\tfrac{1}{2}n_1)_{\kappa}(\tfrac{1}{2}n_2)_{\lambda}C_{\phi}(I)\sum_{\rho,\sigma;\tau}(-1)^{r+s}a^{\kappa,\lambda;\phi}_{\rho,\sigma;\tau}(\tfrac{1}{2}(n_1 + n_2$$

$$C^{\rho,\sigma}_{\tau}(\tilde{F}(I + \tilde{F})^{-1},\ (I + \tilde{F})^{-1})/(\tfrac{1}{2}n_1)_{\rho}(\tfrac{1}{2}n_2)_{\sigma}C_{\tau}(I).$$

Expressions for the joint distributions of the roots of F follow from (9.4) and (9.5). From (9.4) and (6.12)

(9.7) $$f(\tilde{F}) = C_1(m;\ n_1,\ n_2)\mathrm{etr}(-\Omega_1 - \Omega_2)|\tilde{F}|^{\frac{1}{2}n_1-p}$$

$$\sum^{\infty}_{\kappa,\lambda;\phi} (\tfrac{1}{2}(n_1 + n_2))_{\phi}C^{\kappa,\lambda}_{\phi}(-\tilde{F},\ I)L^{\frac{1}{2}n_1-p}_{\kappa,\lambda;\phi}(\Omega_1,\ \Omega_2)/k!\ell!(\tfrac{1}{2}n_1)_{\kappa}(\tfrac{1}{2}n_2)_{\lambda}C_{\phi}(I).$$

Hence using (8.8) we obtain the distribution of Hotelling's $T = \mathrm{tr}\,F$ in the doubly noncentral case,

(9.8) $$f(T) = C_3(m;\ n_1,\ n_2)\mathrm{etr}(-\Omega_1 - \Omega_2)T^{\frac{1}{2}mn_1-1}\sum^{\infty}_{k=0}\{(-T)^k/k!(\tfrac{1}{2}mn_1)_k\}$$

$$\sum^{\infty}_{\ell=0}\sum_{\kappa,\lambda;\phi}(\tfrac{1}{2}(n_1 + n_2))_{\phi}\theta^{\kappa,\lambda}_{\phi}L^{\frac{1}{2}n_1-p}_{\kappa,\lambda;\phi}(\Omega_1,\ \Omega_2)/\ell!(\tfrac{1}{2}n_2)_{\lambda},$$

which also generalizes [3] equation (1).

(b) Case $n_1 \leq m$. Consider $F = X_1'S_2^{-1}X_1$ where $X_1 \sim N_{m,\ n_1}(M_1,\ I_{mn_1})$, $\frac{1}{2}M_1M_1' = \Omega_1$ and $S_2 \sim W_m(n_2,\ I,\ \Omega_2)$. The roots of F are invariant under $X_1 \to HX_1K$, $S_2 \to HS_2H'$ ($H \in O(m)$, $K \in O(n_1)$), so averaging over H, K

(9.9) $$f(X_1,\ S_2) = C_5(m;\ n_1,\ n_2)\mathrm{etr}(-\Omega_1 - \Omega_2)\mathrm{etr}(-\tfrac{1}{2}X_1X_1' - \tfrac{1}{2}S_2)|S_2|^{\frac{1}{2}n_2-p}$$

$$\int_{O(m)} {}_0F_1(\tfrac{1}{2}n_1;\ \tfrac{1}{2}\Omega_1HX_1X_1'H')\,{}_0F_1(\tfrac{1}{2}n_2;\ \tfrac{1}{2}\Omega_2HS_2H')dH,$$

where $C_5(m;\ n_1,\ n_2) = [2^{\frac{1}{2}m(n_1+n_2)}\pi^{\frac{1}{2}mn_1}\Gamma_m(\tfrac{1}{2}n_2)]^{-1}$. Let $Y = S_2^{-\frac{1}{2}}X_1$, so that

$F = Y'Y$ and

(9.10) $f(Y, S_2) = C_5(m; n_1, n_2)\text{etr}(-\Omega_1 - \Omega_2)\text{etr}\{-\tfrac{1}{2}S_2(I + YY')\}$

$$|S_2|^{\frac{1}{2}(n_1+n_2)-p}\int_{O(m)} {}_0F_1(\tfrac{1}{2}n_1; \tfrac{1}{2}\Omega_1 HS_2^{\frac{1}{2}}YY'S_2^{\frac{1}{2}}H')\, {}_0F_1(\tfrac{1}{2}n_2; \tfrac{1}{2}\Omega_2 HS_2H')dH.$$

Now

(9.11) $YY' = J_Y F^* J_Y'$, where $J_Y = [YF^{-\frac{1}{2}}|J_2|\in O(m)$, $F^* = \begin{pmatrix} F & 0 \\ 0 & 0 \end{pmatrix}$,

and F^* is $m \times m$. Making the transformation

(9.12) $S_2^* = J_Y' S_2 J_Y$, $H^* = HJ_Y$.

and integrating over $S_2^* > 0$, we finally obtain

(9.13) $f(F) = [\Gamma_m(\tfrac{1}{2}(n_1+n_2)/\Gamma_{n_1}(\tfrac{1}{2}m)\Gamma_m(\tfrac{1}{2}n_2)]\ \text{etr}(-\Omega_1 - \Omega_2)$

$$|F|^{\frac{1}{2}(m-n_1-1)}|I + F^*|^{-\frac{1}{2}(n_1+n_2)} E_{S_2^*}\int_{O(m)} {}_0F_1(\tfrac{1}{2}n_1;\tfrac{1}{2}\Omega_1 HS_2^{*\frac{1}{2}}F^*S_2^{*\frac{1}{2}}H')$$

$${}_0F_1(\tfrac{1}{2}n_2;\tfrac{1}{2}\Omega_2 HS_2^*H')dH$$

where $S_2^* \sim W_m(n_1 + n_2, (I + F^*)^{-1}, 0)$. Forms corresponding to (9.4), (9.5) and (9.7) may now be obtained .

CSIRO, Division of Mathematics and Statistics, Adelaide, South Australia.

REFERENCES

[1] Bingham, C. (1974). An identity involving partitional generalized binomial coefficients, J. Multivariate Analysis, 4, 210-223.

[2] Constantine, A.G. (1963). Some non-central distribution problems in multivariate analysis, Ann. Math. Statist., 34, 1270-1285.

[3] Constantine, A.G. (1966). The distribution of Hotelling's generalized T_0^2, Ann. Math. Statist., 37, 215-225.

[4] Crowther, N.A.S. (1975). The exact non-central distribution of a quadratic form in normal vectors, South African Statist. J., 9, 27-36

[5] Davis, A.W. (1976). Statistical distributions in univariate and multivariate Edgeworth populations, Biometrika, 63, 661-670.

[6] Davis, A.W. (1978). Invariant polynomials with two matrix arguments, extending the zonal polynomials, to appear in Multivariate Analysis - V (P.R.Krishnaiah, Ed.).

[7] Hayakawa, T. (1967). On the distribution of the maximum latent root of a positive definite symmetric random matrix, Ann. Inst. Statist. Math., 19, 1-17.

[8] Hayakawa, T. (1969). On the distribution of the latent roots of a positi definite random symmetric matrix I, Ann. Inst. Statist. Math., 21, 1-21.

[9] Hayakawa, T. (1972). On the distribution of the multivariate quadratic form in multivariate normal samples, Ann. Inst. Statist. Math., 24, 205-230.

[10] Hayakawa, T. (1973). An asymptotic expansion for the distribution of the determinant of a multivariate quadratic form in a normal sample, Ann. Inst. Statist. Math., 25, 395-406.

[11] Herz, C.S. (1955). Bessel functions of matrix argument, Ann. Math., 61, 474-523.

[12] James, A.T. (1964). Distributions of matrix variates and latent roots derived from normal samples, Ann. Math. Statist., 35, 475-501.

[13] Khatri, C.G. (1977). Distribution of a quadratic form in noncentral normal vectors using generalised Laguerre polynomials, South African Statist. J., 11, 167-179.

[14] Pillai, K.C.S. (1975). The distribution of the characteristic roots of $S_1S_2^{-1}$ under violations, Ann. Statist., 3, 773-779.

[15] Pillai, K.C.S. and Sudjana (1975). Exact robustness studies of tests of two multivariate hypotheses based on four criteria and their distribution problems under violations, Ann. Statist., 3, 617-636.

[16] Robinson, G.de B., Representation theory of the symmetric group, Edinburgh University Press, Edinburgh.

APPENDIX

Orthonormal polynomials $z_{\phi}^{-\frac{1}{2}}C_{\phi}^{\kappa,\lambda}(X, Y)$, where $z_{\phi} = \chi_{[2\kappa]}(1)/1.2\ ...(2f-1)$.

C = multiplier. (X) = tr X .

F	K	L								
2	1	1	κ,λ	1,1						
			ϕ	2	1^2					
			C^2	1/3	2/3					
			(XY)	2	-1					
			(X)(Y)	1	1					
3	2	1	κ,λ	2,1		$1^2,1$				
			ϕ	3	21	21	1^3			
			C^2	1/15	4/15	1/3	1/3			
			(X^2Y)	8	-2	-2	2			
			$(X^2)(Y)$	2	2	-1	-1			
			(XY)(X)	4	-1	2	-2			
			$(X)^2(Y)$	1	1	1	1			
4	3	1	κ,λ	3,1		21,1			$1^3,1$	
			ϕ	4	31	31	2^2	21^2	21^2	1^4
			C^2	1/105	2/35	2/15	2/15	1/3	1/5	2/15
			(X^3Y)	48	-8	-8	-2	4	4	-6
			$(X^3)(Y)$	8	8	-2	-2	-2	2	2
			$(X^2Y)(X)$	24	-4	6	-6	0	-4	6
			$(X^2)(XY)$	12	-2	-2	7	-2	-2	3
			$(X^2)(X)(Y)$	6	6	1	1	1	-3	-3
			$(XY)(X)^2$	6	-1	4	1	-2	2	-3
			$(X)^3(Y)$	1	1	1	1	1	1	1

F	K	L									
4	2	2	κ,λ	2,2			$2,1^2$		$1^2,1^2$		
			ϕ	4	31	2^2	31	21^2	2^2	21^2	1^4
			C^2	1/105	8/189	8/135	2/27	4/27	2/27	32/135	2/15
			(X^2Y^2)	32	4	-8	-8	4	4	1	-4
			$(XYXY)$	16	-12	6	0	0	-6	3	-2
			$(X^2Y)(Y)$	16	2	-4	8	-4	-4	-1	4
			$(XY^2)(X)$	16	2	-4	-4	2	-4	-1	4
			$(X^2)(Y^2)$	4	4	4	-2	-2	1	1	1
			$(XY)^2$	8	-6	3	0	0	6	-3	2
			$(X^2)(Y)^2$	2	2	2	2	2	-1	-1	-1
			$(XY)(X)(Y)$	8	1	-2	4	-2	4	1	-4
			$(Y^2)(X)^2$	2	2	2	-1	-1	-1	-1	-1
			$(X)^2(Y)^2$	1	1	1	1	1	1	1	1

F	K	L													
5	4	1	κ,λ	4,1		31,1			$2^2,1$		$21^2,1$			1^4	
			ϕ	5	41	41	32	31^2	32	2^21	31^2	2^21	21^3	21^3	1^5
			C^2	1/945	8/945	1/35	16/315	1/9	2/45	4/45	8/63	8/45	8/35	4/45	2/45
			(X^4Y)	384	-48	-48	-8	16	-8	4	16	4	-12	-12	24
			$(X^4)(Y)$	48	48	-8	-8	-8	-2	-2	4	4	4	-6	-6
			$(X^3Y)(X)$	192	-24	32	-18	0	-24	12	-12	6	2	12	-24
			$(X^3)(XY)$	64	-8	-8	22	-8	-8	4	4	-8	4	4	-8
			$(X^2Y)(X^2)$	96	-12	-12	-2	4	28	-14	-8	-2	6	6	-12
			$(X^3)(X)(Y)$	32	32	4	4	4	-8	-8	-2	-2	-2	8	8
			$(X^2Y)(X)^2$	48	-6	22	-8	-2	4	-2	4	-8	4	-6	12
			$(X^2)^2(Y)$	12	12	-2	-2	-2	7	7	-2	-2	-2	3	3
			$(X^2)(XY)(X)$	48	-6	8	13	-8	4	-2	-8	7	-1	-6	12
			$(X^2)(X)^2(Y)$	12	12	5	5	5	2	2	-1	-1	-1	-6	-6
			$(XY)(X)^3$	8	-1	6	1	-2	4	-2	4	1	-3	2	-4
			$(X)^4(Y)$	1	1	1	1	1	1	1	1	1	1	1	1

5 3 2 κ,λ	3,2			21,2				1^3,2		3,1^2		21,1^2				1^3,1^2		
ϕ	5	41	32	41	32	31^2	2^21	31^2	21^3	41	31^2	32	31^2	2^21	21^3	2^21	21^3	1^5
C^2	1/945	4/675	8/525	1/50	8/225	1/18	4/45	1/21	4/63	1/90	1/30	2/45	32/315	1/9	1/7	1/15	1/9	2/45
(X^3Y^2)	192	48	-32	-32	-12	0	12	16	-12	-48	16	12	6	0	-8	-8	0	12
(X^2YXY)	192	-96	24	-16	4	16	-8	0	0	0	0	-20	10	4	-4	12	-12	12
$(X^3Y)(Y)$	96	24	-16	-16	-6	0	6	8	-6	48	-16	-12	-6	0	8	8	0	-12
$(X^2Y^2)(X)$	96	24	-16	24	-16	8	-4	-16	12	-24	8	-4	-7	8	0	8	0	-12
$(XYXY)(X)$	48	-24	6	16	-4	-16	8	0	0	0	0	-10	5	2	-2	-6	6	-6
$(X^3)(Y^2)$	16	16	16	-4	-4	-4	-4	4	4	-8	-8	2	2	2	2	-2	-2	-2
$(X^2Y)(XY)$	96	-48	12	-8	2	8	-4	0	0	0	0	20	-10	-4	4	-12	12	-12
$(XY^2)(X^2)$	48	12	-8	-8	22	-8	-2	-8	6	-12	4	-2	4	-8	4	4	0	-6
$(X^3)(Y)^2$	8	8	8	-2	-2	-2	-2	2	2	8	8	-2	-2	-2	-2	2	2	2
$(X^2Y)(X)(Y)$	48	12	-8	12	-8	4	-2	-8	6	24	-8	4	7	8	0	-8	0	12
$(XY^2)(X)^2$	24	6	-4	16	6	0	-6	8	-6	-6	2	-6	-3	0	4	-4	0	6
$(X^2)(XY)(Y)$	24	6	-4	-4	11	-4	-1	-4	3	12	-4	2	-4	8	-4	-4	0	6
$(XY)^2(X)$	24	-12	3	8	-2	-8	4	0	0	0	0	10	-5	-2	2	6	-6	6
$(X^2)(Y^2)(X)$	12	12	12	2	2	2	2	-6	-6	-6	-6	-1	-1	-1	-1	3	3	3
$(X^2)(X)(Y)^2$	6	6	6	1	1	1	1	-3	-3	6	6	1	1	1	1	-3	-3	-3
$(XY)(X)^2(Y)$	12	3	-2	8	3	0	-3	4	-3	6	-2	6	3	-	-4	4	0	-6
$(Y^2)(X)^3$	2	2	2	2	2	2	2	2	2	-1	-1	-1	-1	-1	-1	-1	-1	-1
$(X)^3(Y)^2$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Section 5.

Distribution theory under nonnormality.

Biometrika (1976), **63**, 3, *pp.* 661–70
Printed in Great Britain

Statistical distributions in univariate and multivariate Edgeworth populations

BY A. W. DAVIS
C.S.I.R.O. Division of Mathematics and Statistics, Adelaide

SUMMARY

A general method is presented for the formal construction of statistical distributions in univariate and multivariate populations, when the underlying distributions are specifiable by Edgeworth-type expansions. It is shown that the distribution of a statistic may be derived in principle from the appropriate normal-theory noncentral distribution, by means of a symbolic expectation operator. This provides a unified approach to the results of numerous investigations into the effects of nonnormality on various well-known statistics.

Some key words: Gram–Charlier type A series; Multivariate statistics; Noncentral distribution; Nonnormal distribution.

1. INTRODUCTION

Historically, the many studies of the behaviour of familiar statistics in populations specifiable by Edgeworth expansions have played an important role in stimulating the development of nonparametric and robust procedures. Except for the more recent work of Tiku (1964, etc.) on F-type statistics, explicit representations of their distributions have been cumbersome, obscuring the underlying structure. In § 2, we show that these distributions may be formally constructed by (*a*) deriving an appropriate normal-theory noncentral distribution, then (*b*) applying an expectation-type operator. Chambers (1967) presented an iterative approach.

No attempt is made in this paper to discuss the asymptotic convergence of the expansions obtained. We shall take the Gram–Charlier type A expansion as our starting-point, assuming that the terms can subsequently be grouped according to their orders of magnitude.

Sections 4 and 5 deal with t and F statistics respectively. The remaining sections extend the approach to the multivariate case.

2. CONSTRUCTION OF UNIVARIATE DISTRIBUTIONS

Consider a random variable x with density function $f(x)$, raw moments μ'_r $(r = 1, 2, \ldots)$ and cumulants κ_r $(r = 1, 2, \ldots)$, $\kappa_1 = \mu'_1 = \mu$, $\kappa_2 = \sigma^2$. If $\phi(x|\mu, \sigma^2)$ denotes the normal density with mean μ and variance σ^2, then the Gram–Charlier type A series for $f(x)$ (Johnson & Kotz, 1970, Vol. 1, p. 16) may be written as

$$f(x) = \exp\left\{\sum_{r=3}^{\infty} \frac{\kappa_r}{r!}\left(-\frac{d}{dx}\right)^r\right\}\phi(x|\mu, \sigma^2) = \sum_{r=0}^{\infty}\frac{c_r}{r!}\left(-\frac{d}{dx}\right)^r \phi(x|\mu, \sigma^2). \qquad (2\cdot1)$$

In particular $c_0 = 1$, $c_1 = c_2 = 0$, $c_3 = \kappa_3$, $c_4 = \kappa_4$, $c_5 = \kappa_5$ and $c_6 = \kappa_6 + 10\kappa_3^2, \ldots$.

The method of this paper is based on the simple observation that c_r is obtained by setting $\kappa_1 = \kappa_2 = 0$ in the expression for μ'_r in terms of the cumulants. That is, if we introduce a dummy variate z and an associated linear operator E_z defined by $E_z(z^r) = c_r$ $(r = 0, 1, \ldots)$, then z may be regarded as a pseudovariate with zero mean and variance and the same higher-order cumulants $\kappa_3, \kappa_4, \ldots$, as x; E_z functions as the expectation operator with respect to z. By applying Taylor's theorem, (2·1) takes the form

$$f(x) = E_z \phi(x-z|\mu, \sigma^2) = E_z \phi(x|\mu+z, \sigma^2). \tag{2·2}$$

That is, conditional upon z, x is normal with mean $\mu+z$ and variance σ^2; in a sense, therefore, z carries the information concerning the departure of x from normality.

Next consider a random sample $X = (x_1, \ldots, x_N)'$, where x_i has density $f_i(x)$ and cumulants $\mu_i, \sigma_i^2, \kappa_{i3}, \kappa_{i4}, \ldots$ $(i = 1, \ldots, N)$, not necessarily the same for all i. We may associate with X a vector $Z = (z_1, \ldots, z_N)'$ of independent pseudovariates z_i, and express the likelihood of the sample X as

$$\prod_{i=1}^{N} f_i(x_i) = E_Z \left\{ \prod_{i=1}^{N} \phi(x_i|\mu_i + z_i, \sigma_i^2) \right\}. \tag{2·3}$$

From (2·3), construction of the distribution of a statistic $S(X)$ in a sample from possibly distinct nonnormal distributions specifiable by their Gram–Charlier or Edgeworth expansions, may thus be reduced to two steps.

Step 1. Obtain the density $g(S|Z)$ of S conditional upon Z, that is in a normal sample with means $\mu_i + z_i$ and variances σ_i^2 $(i = 1, \ldots, N)$. This is the analytic problem of constructing a normal-theory noncentral distribution of S.

Step 2. Compute the 'unconditional' density $g(S)$,

$$g(S) = E_Z\{g(S|Z)\}, \tag{2·4}$$

where the $\{z_i\}$ function as independent variates with zero means, and the same higher-order cumulants as the corresponding x_i. This is usually a combinatorial problem.

3. Sample mean and sum of squares

Consider the sample sum of squares $S = \Sigma(x_i - \bar{x})^2$, when the x_i are drawn from a single density $f(x)$, with cumulants $\mu, \sigma^2, \kappa_3, \kappa_4, \ldots$.

Step 1. In a sample from normal distributions with means $\mu + z_1, \ldots, \mu + z_N$ and variances σ^2, S/σ^2 has the noncentral χ^2 distribution, with $n = N-1$ degrees of freedom and noncentrality parameter S_Z/σ^2, where $S_Z = \Sigma(z_i - \bar{z})^2$. Thus

$$g_1(S|Z) = \psi_n(S) \sum_{r=0}^{\infty} \frac{(-\tfrac{1}{2}S_Z/\sigma^2)^r}{(\tfrac{1}{2}n)_r} L_r^{(\frac{1}{2}n-1)}(\tfrac{1}{2}S/\sigma^2), \tag{3·1}$$

where ψ_n is the density function of central $\sigma^2\chi_n^2$, $(a)_r = a(a+1)\ldots(a+r-1)$, and $L_r^{(\alpha)}$ is the generalized Laguerre polynomial (Johnson & Kotz, 1970, Vol. 2, p. 133).

Step 2. The unconditional density is

$$g_1(S) = \psi_n(S) \sum_{r=0}^{\infty} \frac{(-\tfrac{1}{2})^r \theta_r}{(\tfrac{1}{2}n)_r} L_r^{(\frac{1}{2}n-1)}(\tfrac{1}{2}S/\sigma^2), \tag{3·2}$$

where

$$\theta_r = E_Z(S_Z^r) \quad (r = 0, 1, \ldots). \tag{3·3}$$

To find the distribution of S, we thus express the moments of the sample sum of squares of the $\{z_i\}$ in terms of the population cumulants, noting the simplification $\kappa_1 = \kappa_2 = 0$. Early terms may be calculated, for example, using results of Fisher (1929)

$$\theta_1 = 0, \quad \theta_2 = \frac{n^2}{n+1}\kappa_4, \quad \theta_3 = \frac{n^3}{(n+1)^2}\kappa_6 + \frac{4n(n-1)}{(n+1)}\kappa_3^2. \tag{3·4}$$

The formulae up to θ_6 were derived by Quensel (1938); however, the general relationship (3·3) was not recognized.

The procedure obviously extends to joint distributions of several statistics. In deriving Quensel's (1938) joint distribution of $\bar{x}$ and S, we first note that in step 1 these statistics are independent, with joint density function

$$g_2(S, \bar{x}|Z) = g_1(S|Z)\,\phi(\bar{x}|\mu + \bar{z}, N^{-1}\sigma^2). \tag{3·5}$$

Expanding the right-hand side as a double power series in $\bar{z}$ and S_Z and applying E_Z, we find that the product moments

$$\theta_{rs} = E_Z(S_Z^r \bar{z}^s) \quad (r, s = 0, 1, \ldots) \tag{3·6}$$

are required. Again the early terms may be derived using results of Fisher (1929), and we obtain,

$$\theta_{r0} = \theta_r \quad (r = 0, 1, \ldots), \quad \theta_{01} = \theta_{02} = 0, \quad \theta_{11} = \frac{n}{n+1}\kappa_3,$$

$$\theta_{21} = \frac{n^2}{(n+1)^2}\kappa_5, \quad \theta_{12} = \frac{n}{(n+1)^2}\kappa_4, \quad \theta_{03} = \frac{\kappa_3}{(n+1)^2}. \tag{3·7}$$

Some further terms were given by Quensel.

4. Student's t statistics

4·1. *General considerations*

The normal theory noncentral distribution required in this section is that of doubly noncentral t,

$$t = n^{\frac{1}{2}}(u + \delta + \zeta)/\chi_n(\omega), \tag{4·1}$$

where δ and ζ are constants, u is standard normal, and $\chi_n^2(\omega)$ is an independent noncentral chi-squared variate with n degrees of freedom and noncentrality parameter ω. The constant ζ has been included to deal with the contribution of the z's. Various forms of the distribution have been given (Johnson & Kotz, 1970, Vol. 2, p. 213), but none is convenient here. However, a suitable form may be derived by integrating the joint distribution of u and $\chi_n^2(\omega)$ over the region $t \leqslant t_0$. If we define $a^2 = 1 + t_0^2/n$ and $b = (t_0\delta)/(n^{\frac{1}{2}}a)$, the cumulative distribution function of t is

$$F_n(t_0|\delta, \zeta, \omega) = F_n(t_0|\delta, 0, 0) - \frac{e^{-\frac{1}{2}\delta^2/a^2}}{2^{\frac{1}{2}n-\frac{1}{2}}\pi^{\frac{1}{2}}\Gamma(\frac{1}{2}n)\,a^n}\left\{\frac{t_0}{n^{\frac{3}{2}}a}\sum_{r=1}^{\infty}\frac{(-\frac{1}{2}\omega)^r}{r(\frac{1}{2}n+1)_{r-1}}\right.$$

$$\times \int_{y=0}^{\infty} e^{-\frac{1}{2}(y-b)^2}\, y^n\, L_{r-1}^{(\frac{1}{2}n)}\left(\frac{y^2}{2a^2}\right) dy$$

$$\left. + \sum_{r=0}^{\infty}\sum_{s=1}^{\infty}\frac{(-\frac{1}{2}\omega)^r\zeta^s}{(\frac{1}{2}n)_r\, s!}\int_{y=0}^{\infty} e^{-\frac{1}{2}(y-b)^2}\, y^{n-1}\, L_r^{(\frac{1}{2}n-1)}\left(\frac{y^2}{2a^2}\right) H_{s-1}\left(\frac{t_0 y}{a n^{\frac{1}{2}}} - \delta\right) dy\right\}, \tag{4·2}$$

where $F_n(t_0|\delta, 0, 0)$ is the cumulative distribution function of ordinary noncentral t with parameter δ, and $H_s(x)$ is the Hermite polynomial of order s with weight function $e^{-\frac{1}{2}x^2}$. The integrals may be expressed as finite linear combinations of the functions (Fisher, 1931)

$$\mathrm{Hh}_r(-b) = \frac{1}{r!}\int_{y=0}^{\infty} e^{-\frac{1}{2}(y-b)^2}\, y^r\, dy.$$

4·2. *One-sample t*

In a sample from normal distributions with means $\mu+z_1, \ldots, \mu+z_N$ and variances σ^2,

$$t = N^{\frac{1}{2}}(x-\mu_0)/(S/n)^{\frac{1}{2}}$$

has the doubly noncentral t distribution (4·2) with parameters $\delta = N^{\frac{1}{2}}(\mu-\mu_0)/\sigma$, $\zeta = N^{\frac{1}{2}}\bar{z}/\sigma$ and $\omega = S_Z/\sigma^2$. The power function of Student's t in a sample from a single distribution $f(x)$ with cumulants μ, σ^2, κ_3, $\kappa_4, \ldots$ is thus derived from (4·2) by replacing the products $\omega^r\zeta^s$ by $E_Z(\omega^r\zeta^s) = N^{\frac{1}{2}s}\theta^*_{rs}$, where the θ^*_{rs} are obtained by substituting the standardized cumulants $\lambda_i = \kappa_i/\sigma^i$ in (3·6). In particular, the $r = 0$, $s = 3$ and $r = s = 1$ terms in (4·2) yield the coefficient of the skewness measure λ_3 in the form

$$\frac{-e^{-\frac{1}{2}\delta^2/a^2}\, n!}{6\times 2^{\frac{1}{2}(n-1)}\,\Gamma(\frac{1}{2}n)\,\{\pi(n+1)\}^{\frac{1}{2}}\,a^{n+2}}\{(n+1)\,(a^2+2)\,\mathrm{Hh}_{n+1}(-b)-2a^2b\,\mathrm{Hh}_n(-b)$$
$$+(a^2/n)\,(\delta^2-3n-1)\,\mathrm{Hh}_{n-1}(-b)\},$$

as given by Ghurye (1949). The λ_4 and λ_3^2 terms were obtained by Srivastava (1958). Gayen (1949) gave the terms in λ_3, λ_4, λ_3^2, λ_5, $\lambda_3\lambda_4$ and λ_3^3 for the central case $\mu_0 = \mu$ ($\delta = 0$). Tiku (1971*a*) has investigated the power function of t for two-sided alternatives; see § 5·4.

4·3. *Two-sample t*

Assume that samples $x_{i1}, \ldots, x_{iN_i}$ are drawn from populations with cumulants

$$\mu_i,\ \sigma^2,\ \kappa_{3i},\ \kappa_{4i}, \ldots \quad (i = 1, 2).$$

If $\bar{x}_i$ and S_i denote the sample means and sums of squares, then the two-sample t statistic is defined by

$$t = \frac{(\nu')^{\frac{1}{2}}(\bar{x}_1-\bar{x}_2)}{\{(S_1+S_2)/\nu''\}^{\frac{1}{2}}}, \quad \frac{1}{\nu'} = \frac{1}{N_1}+\frac{1}{N_2}, \quad \nu'' = N_1+N_2-2.$$

Step 1. In two normal samples with means $\mu_i+z_{i1}, \ldots, \mu_i+z_{iN_i}$ ($i = 1, 2$) and variances σ^2, t has the doubly noncentral t distribution (4·2) with

$$\delta = (\nu')^{\frac{1}{2}}(\mu_1-\mu_2)/\sigma, \quad \zeta = (\nu')^{\frac{1}{2}}(\bar{z}_1-\bar{z}_2)/\sigma, \quad \omega = (S_{1Z}+S_{2Z})/\sigma^2,$$

using an obvious notation.

Step 2. The power function of t is obtained by replacing $\omega^r\zeta^s$ by

$$\eta_{rs} = E_Z(\omega^r\zeta^s).$$

Since the z's are independent in the two samples, the η_{rs} may be evaluated using the θ^*_{rs}. The resulting power function has not been studied numerically; however, early terms of the null distribution were obtained by Gayen (1950*b*).

5. *F* statistics

5·1. *General considerations*

The basic normal-theory distribution here is that of doubly noncentral F,

$$F = \frac{\chi^2_{\nu_1}(\omega_1)/\nu_1}{\chi^2_{\nu_2}(\omega_2)/\nu_2}, \tag{5·1}$$

with ν_1 and ν_2 degrees of freedom and noncentrality parameters ω_1 and ω_2. Tiku (1972, equation (1)) has given the density in the form

$$p(F) = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} \frac{(-\frac{1}{2}\omega_1)^r(-\frac{1}{2}\omega_2)^s}{r!\,s!} q_{rs}(F), \tag{5·2}$$

where

$$q_{rs}(F) = \sum_{i=0}^{r}\sum_{j=0}^{s}(-1)^{i+j}\binom{r}{i}\binom{s}{j}\frac{(\nu_1/\nu_2)^{\frac{1}{2}\nu_1+i}}{B(\frac{1}{2}\nu_1+i, \frac{1}{2}\nu_2+j)}\frac{F^{\frac{1}{2}\nu_1+i-1}}{\{1+(\nu_1/\nu_2)F\}^{\frac{1}{2}\nu_1+\frac{1}{2}\nu_2+i+j}}. \tag{5·3}$$

5·2. *Two-sample F*

We assume two independent samples as in §4·3. In testing equality of variances it is common to use $F = (S_1/\nu_1)(S_2/\nu_2)^{-1}$, where $\nu_1 = N_1-1$ and $\nu_2 = N_2-1$.

The step 1 distribution of F is thus given by (5·2), with $\omega_1 = S_{1Z}/\sigma^2$ and $\omega_2 = S_{2Z}/\sigma^2$. Hence, since the z's in the two samples are independent, the null distribution in the non-normal case is derived by replacing the products $\omega_1^r\omega_2^s$ by $\pi_{rs} = E_Z(\omega_1^r\omega_2^s) = \theta_r^{(1)*}\theta_s^{(2)*}$, where $\theta_r^{(i)*}$ is obtained by substituting the standardized cumulants of the ith sample in θ_r, equation (3·3). The results are consistent with those of Tiku (1964). Gayen (1950*a*) also studied this distribution.

5·3. *One-way analysis of variance*

For k samples $x_{i1}, \ldots, x_{iN_i}$ from distributions with cumulants $\mu+\epsilon_i$, σ^2, κ_{3i}, κ_{4i}, … ($i = 1, \ldots, k$; $\Sigma N_i\epsilon_i = 0$), the standard test for equality of the means is $F = (S_B/\nu_1)(S_W/\nu_2)^{-1}$, where S_B and S_W are the between- and within-groups sums of squares respectively, $\nu_1 = k-1$, $\nu_2 = N-k$ ($N = \Sigma N_i$). In step 1, x_{ij} is normal with mean $\mu+\epsilon_1+z_{ij}$ and variance σ^2; thus S_B and S_W have independent noncentral chi-squared distributions, with non-centrality parameters $\omega_1 = S_{BZ}/\sigma^2$ and $\omega_2 = S_{WZ}/\sigma^2$ under the null hypothesis, for all $\epsilon_i = 0$, where S_{BZ} and S_{WZ} have the obvious meaning. To obtain the null distribution of F in the nonnormal case, we require the quantities $\rho_{rs} = E_Z(\omega_1^r\omega_2^s)$. Early terms are thus available from David & Johnson (1951), noting the simplification that the means and variances are to be taken as zero; in particular

$$\rho_{00} = 1, \quad \rho_{10} = \rho_{01} = 0, \quad \rho_{20} = \sum_{i=1}^{k}\frac{\lambda_{4i}}{N_i}\left(1-\frac{N_i}{N}\right)^2,$$

$$\rho_{11} = \sum_{i=1}^{k}\left(1-\frac{1}{N_i}\right)\left(1-\frac{N_i}{N}\right)\lambda_{4i}, \quad \rho_{02} = \sum_{i=1}^{k}\frac{\lambda_{4i}}{N_i}(N_i-1)^2,$$

where the λ's are the standardized cumulants. Tiku (1964) considered the case of equal group sizes, and his coefficients are given by

$$\beta_{rs} = (-\tfrac{1}{2})^{r+s}\rho_{rs}/\{(\tfrac{1}{2}\nu_1)_r(\tfrac{1}{2}\nu_2)_s\}.$$

The Λ_3^2 term in his β_{30} apparently omits a divisor ν_1. Gayen (1950*a*) allowed for unequal N_i but assumed that the error distribution was the same for all groups.

The power function has been studied by Tiku (1971*a, b*) and Srivastava (1959). The non-centrality parameter of S_B in step 1 is now

$$\omega_1 = \sum_{i=1}^{k}N_i(\bar{z}_i-\bar{z}+\epsilon_i)^2/\sigma^2,$$

and by defining

$$\xi = \sum_{i=1}^{k}N_i(\epsilon_i/\sigma)^2$$

we may write the density (5·2) in the form

$$p(F) = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} \frac{\{-\frac{1}{2}(\omega_1-\xi)\}^r(-\frac{1}{2}\omega_2)^s}{r!\,s!}\tilde{q}_{rs}(F),$$

where

$$\tilde{q}_{rs}(F) = \sum_{t=0}^{\infty}\frac{(-\frac{1}{2}\xi)^t}{t!}q_{r+t,\,s}(F).$$

It can be shown that the latter is expressible in terms of Srivastava's functions $p_r(.\,;r_1,r_2)$:

$$\tilde{q}_{rs}(F) = (-1)^r\sum_{i=0}^{s}\binom{s}{i}(-1)^i p_r(F;\nu_1,\nu_2+2i).$$

Calculation of the power function thus requires the quantities

$$\tilde{\rho}_{rs} = E_Z\{(\omega_1-\xi)^r\omega^s\} = \mu(\omega_1^r\omega_2^s)$$

in David & Johnson's (1951) notation, since ω_1 and ω_2 are the between- and within-groups sums of squares for the variables $(z_{ij}+\epsilon_1)/\sigma$, and $E_Z(\omega_1)=\xi$ and $E_Z(\omega_2)=0$. Noting that the variances are to be taken as zero, we have

$$\tilde{\rho}_{00}=1,\quad \tilde{\rho}_{10}=\tilde{\rho}_{01}=0,\quad \tilde{\rho}_{20}=\rho_{20}+4\sum_{i=1}^{k}\lambda_{3i}(\epsilon_i/\sigma)(1-N_i/N)$$

$$\tilde{\rho}_{11}=\rho_{11}+2\sum_{i=1}^{k}(N_i-1)(\epsilon_i/\sigma)\lambda_{3i},\quad \tilde{\rho}_{02}=\rho_{02}.$$

Srivastava assumed equal group sizes, and identical error distributions.

5·4. *Tiku's power functions*

Tiku's approach to the power functions associated with F is based on Patnaik's (1949) approximation to noncentral chi-squared $\chi^2_{\nu_1}(\xi) \simeq q\chi^2_p$, where $p=(\nu_1+\xi)^2/(\nu_1+2\xi)$, $q=(\nu_1+2\xi)/(\nu_1+\xi)$. In applying step 1 of the present paper, we note that the density of $v=\chi^2_{\nu_1}(\omega_1)/q$ may be expanded as a Laguerre series based on the density ψ^0_p of χ^2_p (Tiku, 1965),

$$h(v) = \psi^0_p(v)\sum_{r=0}^{\infty}\alpha_r L_r^{(\frac{1}{2}p-1)}(\tfrac{1}{2}v).$$

By using the orthogonality properties of the polynomials,

$$\alpha_r = \binom{\frac{1}{2}p+r-1}{r}^{-1}E_v\{L_r^{(\frac{1}{2}p-1)}(\tfrac{1}{2}v)\} = \sum_{i=0}^{r}\binom{r}{i}\frac{(-\frac{1}{2})^i}{(\frac{1}{2}p)_i}E_v(v^i).$$

Thus

$$\alpha_0=1,\quad \alpha_1=-(\omega_1-\xi)/(\nu_1+\xi),$$

$$\alpha_2=\frac{1}{(\nu_1+2\xi)(p+2)}\left\{(\omega_1-\xi)^2-\frac{4\xi}{\nu_1+\xi}(\omega_1-\xi)\right\},$$

and so on. The density of $w=(\nu_1F)/(pq)$, with F defined by (5·1), may now be obtained in the form

$$\tilde{p}(w) = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{(\frac{1}{2}p)_r\alpha_r(-\frac{1}{2}\omega_2)^s}{r!\,s!}q_{rs}(w), \tag{5·4}$$

where ν_1 is replaced by p in (5·3) which defines q_{rs}.

Results for the power functions of one-sample t for two-sided alternatives (Tiku, 1971 *a*) and one-way analysis of variance F (Tiku, 1971 *b*) may be derived from (5·4) by making appropriate substitutions for ν_1, ν_2, ξ, ω_1 and ω_2, and applying step 2.

6. Multivariate extensions

The method extends in principle to the multivariate case. Given a multivariate sample $X = (x_1, \ldots, x_N)'$, where x_i has an m-variate density function $f_i(x)$ with mean vector μ_i and covariance matrix Σ_i, we may define an associated sample $Z = (z_1, \ldots, z_N)'$ of independent pseudorandom vectors, z_i having zero mean vector and covariance matrix, and the same higher-order multivariate cumulants as x_i $(i = 1, \ldots, N)$. The relations (2·3) and (2·4) carry straight over, and formal expansions of sampling distributions in multivariate Edgeworth populations may be obtained by applying steps 1 and 2 as before. It will be convenient to use Kaplan's (1952) 'tensor' notation; i.e. if $x = (x_1, \ldots, x_m)$, then product moments of order r are written

$$\tilde{\mu}'_{j_1 \ldots j_r} = E_x(x_{j_1} \ldots x_{j_r}),$$

with a corresponding notation $\tilde{\kappa}_{j_1 \ldots j_r}$ for rth-order cumulants; the cumulants are uniquely defined by restricting them to be symmetric functions of the subscripts. The tilde is a reminder that the notation is to be distinguished from that used earlier in this paper for the univariate case.

Certain multivariate noncentral distributions required for step 1 may be formulated in terms of the theory of zonal polynomials and hypergeometric functions of matrix argument developed by James (1964) and Constantine (1963). An introduction to this theory is also given by Johnson & Kotz (1972, Chapters 38, 39). However, the application of the method to the usual multivariate analysis of variance test statistics would require the doubly noncentral distributions, which have not yet been investigated. Some applications based on the noncentral Wishart distribution are indicated in the remaining sections.

7. Sample sum of squares and products matrix

Let W denote the sum of squares and products matrix of a sample X drawn from a single nonnormal p-variate distribution with mean vector μ and covariance matrix Σ. In step 1, W has the noncentral Wishart distribution with $n = N-1$ degrees of freedom and noncentrality matrix $\Sigma^{-1}W_Z$, where $W_Z = ((u_{ij}))$ is the sums of squares and products matrix of the Z sample. James (1964) has given the distributions in terms of a Bessel-type hypergeometric function of matrix argument.

In the bivariate case $(m = 2)$ with

$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

ρ being the correlation coefficient, the distribution may be obtained from Anderson & Girshick (1944), following a minor correction,

$$g(W|Z) = \{4\pi(1-\rho^2)(n-2)!\}^{-1} |W|^{\frac{1}{2}(n-3)} \exp\{-\tfrac{1}{2}\operatorname{tr}(\Sigma^{-1}W)\} \times \exp\{-\tfrac{1}{2}\operatorname{tr}(\Sigma^{-1}W_Z)\} \sum_{s=0}^{\infty} \frac{1}{4^s(\frac{1}{2}n)_s} \sum_{2k+l=s} \frac{|\Phi|^k\{\operatorname{tr}(\Phi)\}^l}{k!\,l!\{\frac{1}{2}(n-1)\}_k}, \quad (7{\cdot}1)$$

where $\Phi = \Sigma^{-1}W_Z\Sigma^{-1}W$. Step 2 thus requires expectations of the form $E_Z[|W_Z|^k\{\operatorname{tr}(\Phi)\}^l]$.

The results of Quensel (1938) $(\rho = 0)$ and Gayen (1951) $(\rho \neq 0)$ may be reproduced, but we shall omit the details.

8. THE CORRELATION COEFFICIENT

Gayen (1951) studied the effect of nonnormality on the distribution of the ordinary product-moment correlation coefficient r. The step 1 distribution may be derived from (7·1) by means of Fisher's transformation

$$W = y\begin{bmatrix} e^t & r \\ r & e^{-t}\end{bmatrix} = yT. \tag{8·1}$$

Integrating over $(0 < y < \infty, \; -\infty < t < \infty)$, we obtain the distribution of r in a bivariate normal sample with unequal mean vectors,

$$g(r|Z) = \frac{n-1}{\pi}(1-r^2)^{\frac{1}{2}(n-3)}(1-\rho^2)^{\frac{1}{2}n}\exp\{-\tfrac{1}{2}\operatorname{tr}(\Sigma^{-1}W_Z)\}\sum_{s=0}^{\infty}\frac{(n)_s}{4^s(\frac{1}{2}n)_s}\sum_{2k+l=s}\frac{|W_Z|^k(1-r^2)^k}{k!\,l!\{\frac{1}{2}(n-1)\}_k}$$
$$\times\int_{-\infty}^{\infty}\frac{\{\operatorname{tr}(VT)/(1-\rho^2)\}^l}{2(\cosh t-\rho r)^{n+s}}\,dt, \tag{8·2}$$

where

$$\begin{aligned} V = (1-\rho^2)^2\Sigma^{-1}W_Z\Sigma^{-1} = ((v_{ij})), \quad v_{11} &= u_{11}-2\rho u_{12}+\rho^2u_{22},\\ v_{12} = -\rho u_{11}+(1+\rho^2)\,u_{12}-\rho u_{22}, \quad v_{22} &= \rho^2u_{11}-2\rho u_{12}+u_{22}.\end{aligned} \tag{8·3}$$

The term of zero order in the z's is Fisher's distribution of r,

$$\psi(r,\rho) = \frac{n-1}{\pi}(1-r^2)^{\frac{1}{2}(n-3)}(1-\rho^2)^{\frac{1}{2}n}I_n,$$

where

$$I_n = \int_0^{\infty}\frac{dt}{(\cosh t-\rho r)^n}. \tag{8·4}$$

The terms in (8·2) may be expressed as linear combinations of I_j's; Gayen's results indicate that these may in turn be reduced to linear combinations of the partial derivatives of ψ with respect to ρ, with coefficients independent of r. To the fourth order in the z's,

$$g(r|Z) = \psi-\frac{1}{2n}\{\rho(u_{11}+u_{22})-2u_{12}\}\frac{\partial\psi}{\partial\rho}+\frac{1}{8n(n+2)}\Bigg[\Bigg\{3\rho(u_{11}^2+u_{22}^2)-4u_{12}(u_{11}+u_{12})+2\rho u_{11}u_{22}$$
$$+\frac{4\rho}{n-1}|W_Z|\Bigg\}\frac{\partial\psi}{\partial\rho}+\Bigg\{(\rho u_{11}-2u_{12}+\rho u_{22})^2-\frac{4(1-\rho^2)}{n-1}|W_Z|\Bigg\}\frac{\partial^2\psi}{\partial\rho^2}\Bigg]+\dots. \tag{8·5}$$

Operating with E_Z, we find that the second-order term vanishes since the z's have zero correlation matrix; from equation (3) of Kaplan (1952),

$$E_Z(u_{ij}u_{kl}) = \frac{n^2}{n+1}\tilde{\kappa}_{ijkl}, \tag{8·6}$$

the fourth-order term passes over into Gayen's term involving fourth order cumulants. Note that in the bivariate case $E_Z|W_Z| = 0$.

9. LATENT ROOTS OF W

We return to the m-variate situation, the latent roots $l_1, \dots, l_m$ of W in the case $\Sigma = I_m$ may be shown to have the step 1 joint distribution

$$g(\Lambda|Z) \propto e^{-\frac{1}{2}\operatorname{tr}(\Lambda)}|\Lambda|^{\frac{1}{2}(n-m-1)}\prod_{i<j}(l_i-l_j)\sum_{k=0}^{\infty}\sum_{\pi}\frac{(-\frac{1}{2})^k C_\pi(W_Z)}{k!(\frac{1}{2}n)_\pi C_\pi(I_m)}L_\pi^{\frac{1}{2}(n-m-1)}(\tfrac{1}{2}\Lambda), \tag{9·1}$$

where $\Lambda = \text{diag}(l_1, \ldots, l_m)$, Σ_π is extended over all partitions π of the integer k into not more than m parts, C_π is the zonal polynomial corresponding to the partition π, $(\frac{1}{2}n)_\pi$ is a generalized hypergeometric coefficient, and L_π^γ is Constantine's (1966) generalized Laguerre polynomial of order γ. Equation (9·1) is thus a multivariate generalization of (3·1). The formal construction of the joint distribution of the roots in the nonnormal case thus involves multivariate generalizations of (3·3), $\theta_\pi = E_Z\{C_\pi(W_Z)\}$.

Since

$$C_{[1]}(W_Z) = \Sigma u_{ii}, \quad C_{[2]}(W_Z) = \tfrac{1}{3}\Sigma_i \Sigma_j (u_{ii} u_{jj} + 2u_{ij}^2)$$

we have $\theta_{[1]} = 0$, and from (8·6),

$$\theta_{[2]} = N^{-1}(N-1)^2 \Sigma_i \Sigma_j \tilde{\kappa}_{iijj}.$$

Further results based on Kaplan's (1952) formulae are

$$\begin{aligned} \theta_{[1^2]} = \theta_{[1^3]} = 0, \quad \theta_{[2,1]} &= \{12(N-1)(N-2)/(5N)\} \Sigma_{i,j,k} (\tilde{\kappa}_{ijk}^2 - \tilde{\kappa}_{ijj} \tilde{\kappa}_{ikk}), \\ \theta_{[3]} = \{(N-1)^4/N^3\} \Sigma_{i,j,k} \tilde{\kappa}_{iijjkk} &+ \{4(N-1)(N-2)/(5N)\} \Sigma_{i,j,k} (2\tilde{\kappa}_{ijk}^2 + 3\tilde{\kappa}_{ijj} \tilde{\kappa}_{ikk}). \end{aligned} \quad (9{\cdot}6)$$

It is hoped that these results may prove useful, for example, in investigating the effect of nonnormality on the largest root of the Wishart matrix.

References

Anderson, T. W. & Girshick, M. A. (1944). Some extensions of the Wishart distribution. *Ann. Math. Statist.* **15**, 345–57.

Chambers, J. M. (1967). On methods of asymptotic approximation for multivariate distributions. *Biometrika* **54**, 367–83.

Constantine, A. G. (1963). Some non-central distribution problems in multivariate analysis. *Ann. Math. Statist.* **34**, 1270–84.

Constantine, A. G. (1966). The distribution of Hotelling's generalized T_0^2. *Ann. Math. Statist.* **37**, 215–25.

David, F. N. & Johnson, N. L. (1951). The effect of non-normality on the power function of the *F*-test in the analysis of variance. *Biometrika* **38**, 43–57.

Fisher, R. A. (1929). Moments and product moments of sampling distributions. *Proc. Lond. Math. Soc.* (2), **30**, 199–238.

Fisher, R. A. (1931). The sampling error of estimated deviates, together with other illustrations of the properties and applications of the integrals and derivatives of the normal error function. *Brit. Assoc. Advance. Sci. Math. Tables* **1**, 26–35.

Gayen, A. K. (1949). The distribution of 'Student's' *t* in random samples of any size drawn from non-normal universes. *Biometrika* **36**, 353–69.

Gayen, A. K. (1950*a*). The distribution of the variance ratio in random samples of any size drawn from non-normal universes. *Biometrika* **37**, 236–55.

Gayen, A. K. (1950*b*). Significance of difference between the means of two non-normal samples. *Biometrika* **37**, 399–408.

Gayen, A. K. (1951). The frequency distribution of the product-moment correlation coefficient in random samples of any size drawn from non-normal universes. *Biometrika* **38**, 219–47.

Ghurye, S. G. (1949). On the use of Student's *t*-test in an asymmetrical population. *Biometrika* **36**, 426–30.

James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35**, 475–501.

Johnson, N. L. & Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions* Vols 1, 2. Boston: Houghton Mifflin.

Johnson, N. L. & Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions.* New York: Wiley.

Kaplan, E. L. (1952). Tensor notation and the sampling cumulants of *k*-statistics. *Biometrika* **39**, 319–23.

PATNAIK, P. B. (1949). The non-central χ^2 and F-distributions and their applications. *Biometrika* **36**, 202–32.

QUENSEL, C. E. (1938). The distributions of the second moment and of the correlation coefficient in samples from populations of type A. *Lunds. Univ. Arsskr.* (2) **34** (4), 1–111.

SRIVASTAVA, A. B. L. (1958). Effect of non-normality on the power function of t-test. *Biometrika* **45**, 421–9.

SRIVASTAVA, A. B. L. (1959). Effect of non-normality on the power of the analysis of variance test. *Biometrika* **46**, 114–22.

TIKU, M. L. (1964). Approximating the general non-normal variance-ratio sampling distributions. *Biometrika* **51**, 83–95.

TIKU, M. L. (1965). Laguerre series forms of non-central χ^2 and F distributions. *Biometrika* **52**, 415–27.

TIKU, M. L. (1971*a*). Student's t distribution under non-normal situations. *Aust. J. Statist.* **13**, 142–8.

TIKU, M. L. (1971*b*). Power function of the F-test under non-normal situations. *J. Am. Statist. Assoc.* **66**, 913–6.

TIKU, M. L. (1972). A note on the distribution of the doubly non-central F distribution. *Aust. J. Statist.* **14**, 37–40.

[*Received August* 1975. *Revised February* 1976]

Submitted

On the effects of moderate multivariate nonnormality on Wilks's likelihood ratio criterion

A.W. Davis

CSIRO Division of Mathematics and Statistics, Glen Osmond, South Australia.

Summary

A technique for the formal construction of statistical distributions in multivariate Edgeworth populations is applied in an investigation of the effects of moderate multivariate nonnormality on the test statistics used in connection with the general linear model, in particular, Wilks's likelihood ratio criterion. The first-order approximation to the sampling distributions under nonnormality are shown to involve Mardia's measures of multivariate skewness and kurtosis, together with a supplementary skewness measure. Correction terms for tests at the 5% level are constructed for the one-way analysis of variance case, and the resulting approximations are compared with simulation results for a range of bivariate distributions.

Some key words: Multivariate linear model; Nonnormal distribution; Doubly noncentral beta distribution; Wilks's likelihood ratio statistic; Mardia's measures of nonnormality; Invariant polynomials of two matrix arguments; Bivariate distribution.

1. INTRODUCTION

Essentially, this paper presents a multivariate generalization of the investigations by Gayen (1950) and Tiku (1964) into the effects of nonnormality on the F-test. These authors obtained expansions of the distribution of F for the one-way analysis of variance in terms of the early cumulants of the underlying error distribution, and tabulated the required correction terms. In the present paper we derive an analogous expansion for the joint distribution of the latent roots of the multivariate beta matrix, which underlies the usual tests of linear hypotheses associated with the general multivariate linear model. The technique used is based on Davis (1976), and requires the normal-theory doubly noncentral multivariate beta distribution, which in turn involves a generalization of the zonal polynomials of A.T. James (1964) to invariant polynomials with two matrix arguments (Davis, 1979 a). The basic theory is outlined in Sections 2 and 3, and the expansions obtained constitute multivariate generalizations of Tiku's (1964) results.

In Section 4 it is shown that to the first order, the effects of nonnormality on a class of tests based on the multivariate beta matrix may be specified in terms of just three parameters, Mardia's (1970) measures of multivariate skewness and kurtosis, together with a supplementary skewness measure whose effect appears to be fairly negligible. In Section 5 the problem is raised of computing the marginal integrals required for calculating correction terms for the various test statistics, and discussed in connection with Wilks's likelihood ratio statistic, W. Under multinormality, the null distribution of this statistic is a Meijer's G-function, and tabulation of W based on the numerical solution of the differential equation for this function has been investigated by Davis (1979 b). In Section 5 we present a related system of differential-recursive equations, from which the required correction terms can be constructed, and a tabulation of these quantities is given in Table 4 for the one-way multivariate analysis of variance.

Finally, results based the first order approximation to the distribution of W under various nonnormal bivariate distributions are compared with simulation results in Section 6 and Table 5 for the one-way analysis of variance. Although the derivations are based on the Edgeworth expansion, central limit effects arising from the construction of the statistics may be expected to entail that the expansion will be asymptotically valid for a range of parent populations differing quite widely from the normal. This seems to be borne out by the results given in Table 5, at least for distributions whose skewness and kurtosis are not too large.

Several studies on the robustness of Hotelling's generalized T^2 are mentioned in Section 6. An investigation into the robustness of the multivariate analysis of variance test criteria has been made by Pillai and Sudjana (1975). In their work, the nonnormality has been generated by assigning prior distributions to the normal-theory parameters and averaging over these; unfortunately it appears rather difficult to specify the type of nonnormality produced.

2. The multivariate linear model and test criteria

In the case of the multivariate linear model

$$Y = A\xi + \varepsilon$$

where Y, ε are $N \times m$, A is $N \times q$ (rank q) and ξ is $q \times m$, tests of the null hypotheses

$$C\xi = 0 \qquad (2.1)$$

(C : t × q, of rank t) may be based on the latent roots $0 \leq \beta_m \leq .. \beta_1 \leq 1$ of the multivariate beta matrix $B = S_1(S_1 + S_2)^{-1}$, where

$$S_1 = Y'HY, \quad S_2 = Y'EY$$

are the m × m "hypothesis" and "error" sum of products matrices. $H = (h_{ij})$ and $E = (e_{ij})$ are N × N matrices expressible in terms of A and C (see, for example, Smith, Gnanadesikan and Hughes (1962)). The most frequently used criteria for testing (2.1) are

Wilks's likelihood ratio criterion $W = \det(I_m - B)$,

Hotelling's trace $\mathrm{tr}\{B(I - B)^{-1}\}$,

Pillai's trace tr B, and

Roy's largest root β_1,

where I_m is the m × m unit matrix. In seeking to approximate the joint distribution of the roots β_i when the rows of the error matrix ε are independently distributed with nonnormal multivariate distributions $f_\alpha(\varepsilon_\alpha)(\alpha = 1, .., N)$ having zero mean vectors and covariance matrices $\Sigma_\alpha(\alpha = 1, ..., N)$ respectively, we may apply the method of Davis (1976) as follows.

Step 1. Obtain the joint distribution $g(\mathcal{B}|Z)$ of $\mathcal{B} = \mathrm{diag}\{\beta_1, ..., \beta_m\}$ under the model

$$Y = A\xi + Z + \varepsilon, \quad C\xi = 0, \tag{2.2}$$

where Z is an arbitrary N × m matrix and the rows of ε are multinormal with zero mean vectors and covariance matrices $\Sigma_\alpha(\alpha = 1, ..., N)$.

Step 2. Compute the required distribution

$$g(\mathcal{B}) = E_Z\{g(\mathcal{B}|Z)\},$$

where the "expectation" is to be calculated as if the rows z_α of Z were independent random vectors with zero mean vectors <u>and</u> covariance matrices, but with the same third and higher order multivariate cumulants as $f_\alpha(\epsilon_\alpha)$ $(\alpha = 1, \ldots, N)$.

Throughout the rest of this paper, the f_α will be assumed to be equal, in which case we may take all $\Sigma_\alpha = I_m$ without loss of generality.

In carrying out Step 1, under (2.2) and the assumption of multinormality S_1 and S_2 have independent noncentral Wishart distributions with $n_1 = t$, $n_2 = N - q$ degrees of freedom, and noncentrality matrices

$$\Omega_1 = \tfrac{1}{2}Z'HZ, \quad \Omega_2 = \tfrac{1}{2}Z'EZ \tag{2.3}$$

respectively. Thus $g(B|Z)$ is the distribution of latent roots for the doubly noncentral multivariate beta distribution, which is discussed in the next section.

3. The latent roots of doubly noncentral multivariate B.

The noncentral distributions required for studying the normal-theory power functions of multivariate analysis of variance test statistics have been expanded in series form using the zonal polynomials and hypergeometric functions of matrix argument defined by A.T. James (1964) and A.G. Constantine (1963). To obtain similar expansions in the doubly noncentral case we require an extension of the zonal polynomials to certain polynomials $C_\phi^{\kappa,\lambda}(R, S)$ of two symmetric $m \times m$ matrices R, S defined by Davis (1979); these are invariant under the simultaneous transformations $R \to KRK'$, $S \to KSK'$, where K is an arbitrary $m \times m$ orthogonal matrix. Here κ, λ are partitions of the integers k, ℓ into $\leq m$ parts $(k, \ell = 0, 1, 2, \ldots)$ and ϕ is a partition of $f = k + \ell$ such that the irreducible representation of the real linear group $G\ell(m)$ of non-singular $m \times m$ matrices indexed by 2ϕ occurs in the decomposition of the Kronecker product of the irreducible representations indexed by 2κ and 2ℓ. The $C_\phi^{\kappa,\lambda}$ have been tabulated up to degree $f = 6$ by the author; they are expressible in terms of the distinct products of traces of products of R and S of total degree k, ℓ in the elements of R, S respectively. In particular,

$$C_\kappa^{\kappa,0}(R, S) = C_\kappa(R),\ C_\lambda^{0,\lambda}(R, S) = C_\lambda(S),$$

where C_κ, C_λ are zonal polynomials. The remaining low order $C_\phi^{\kappa,\lambda}$ required in this paper are presented in Table 1.

(Table 1 near here)

When $n_1 \geq m$, the density function of $\mathcal{B}$ defined in Section 2 has been expressed by Davis in unpublished work as

$$g(\mathcal{B}/Z) = B(m;\, n_1,\, n_2) \det \mathcal{B}^{\frac{1}{2}(n_1-m-1)} \det(I_m - \mathcal{B})^{\frac{1}{2}(n_2-m-1)} \Pi_{i<j}(\beta_i - \beta_j)$$

$$\sum_{k,\ell=0}^{\infty} \sum_{\kappa,\lambda;\phi}(-1)^{k+\ell} C_\phi^{\kappa,\lambda}(\Omega_1,\, \Omega_2)p(\kappa,\, \lambda;\, \phi;\, n_1,\, n_2,\, \mathcal{B})/$$

$$(\tfrac{1}{2}n_1)_\kappa(\tfrac{1}{2}n_2)_\lambda k!\ell!, \qquad (3.1)$$

where

$$B(m;\, n_1,\, n_2) = \Gamma_m(\tfrac{1}{2}(n_1 + n_2))/\Gamma_m(\tfrac{1}{2}n_1)\Gamma_m(\tfrac{1}{2}n_2),$$

$\Gamma_m(z)$ being the multivariate gamma function and $(q)_\kappa$ the multivariate hypergeometric coefficient (Constantine, 1963). Here, p is the polynomial

$$p(\kappa,\, \lambda;\, \phi;\, n_1,\, n_2;\, \mathcal{B}) = (\tfrac{1}{2}n_1)_\kappa(\tfrac{1}{2}n_2)_\lambda \sum_{r=0}^{k}\sum_{s=0}^{\ell}\sum_{\rho,\sigma;\tau}(-1)^{r+s}\, a_{\rho,\sigma;\tau}^{\kappa,\lambda;\phi}$$

$$(\tfrac{1}{2}(n_1 + n_2))_\tau\, C_\tau^{\rho,\sigma}(\mathcal{B},\, I_m - \mathcal{B})/(\tfrac{1}{2}n_1)_\rho(\tfrac{1}{2}n_2)_\sigma C_\tau(I_m), \qquad (3.2)$$

where ρ, σ, τ are partitions or r, s and $t = r + s$ into $\leq m$ parts, and $a_{\rho,\sigma;\tau}^{\kappa,\lambda;\phi}$ is the coefficient of $C_\tau^{\rho,\sigma}(X, Y)/C_\tau(I_m)$ in the expansion of $C_\phi^{\kappa,\lambda}(I_m + X,\, I_m + Y)/C_\phi(I_m)$. A similar expansion may be obtained in the case $n_1 < m$. Multiplied by the null density of the roots preceding the summation in (3.1), the p's are seen to be multivariate generalizations of the functions $p_{k,\ell}$ defined by Tiku (1964).

4. Expansion of g(B) in the nonnormal case

From (3.1), Stage 2 in the construction of g(B) involves taking the formal expectations of the $C_\phi^{\kappa,\lambda}(\Omega_1, \Omega_2)$ with respect to Z, Ω_1 and Ω_2 being given by (2.3).

Example: Let $Z = (z_{\alpha i};\ \alpha = 1, \ldots, N;\ i = 1, \ldots, m)$. Then

$$\mathrm{trace}(\Omega_1\Omega_2) = \tfrac{1}{4} \sum_{i,j=1}^{M} \sum_{\alpha,\beta,\gamma,\delta=1}^{N} z_{\alpha i} z_{\beta j} z_{\gamma i} z_{\delta j} h_{\alpha\beta} e_{\gamma\delta}.$$

Since expectations are to be calculated as though the $z_{\alpha i}$ are independent for different α and their first and second order moments are zero, we need only consider situations in which the Greek subscripts are equal in groups of size 3 at least. In this example, we need only consider the case where the four Greek subscripts are equal, so that

$$E_Z\{\mathrm{trace}(\Omega_1\Omega_2)\} = \tfrac{1}{4}\left(\sum_{\alpha=1}^{N} h_{\alpha\alpha} e_{\alpha\alpha}\right) . \sum_{i,j=1}^{m} \kappa_{iijj},$$

using Kaplan's (1952) symmetric tensor notation for multivariate cumulants. The same value is obtained for E_Z {trace Ω_1.trace Ω_2}, whence we obtain the results in the first two lines of Table 2.

In a number of situations, including the one-way analysis of variance, the elements h_{ij} of the hypothesis matrix H will be $O(N^{-1})$ and the e_{ij} will be O(1). It is not difficult to show by applying the kind of reasoning indicated in the above example to $E_Z C_\phi^{\kappa,\lambda}(\Omega_1, \Omega_2)$ that this expectation will be a sum of $O(N^{[2f/3]})$ terms, each of order N^{-k}, where [x] denotes the greatest integer not exceeding x. Hence

$$E_Z C_\phi^{\kappa,\lambda}(\Omega_1, \Omega_2)/(\tfrac{1}{2}n_2)_\lambda = O(N^{-[(f+2)/3]})$$

for large N. On the other hand, consideration of the limiting distribution of $T = n_2 B$ in the normal case leads to the result that $p(\kappa, \lambda; \phi; n_1, n_2; n_2^{-1}T)$ is $O(1)$ for fixed T and large N. Hence in the construction of correction terms for fixed nominal significance levels under normal theory, the terms of (3.1) for $k + \ell = f$ are $O(N^{-[(f+2)/3]})$ for large N. To obtain the $O(N^{-1})$ correction terms we thus consider f = 0 to 3. For f = 1 the expectations are identically zero, and certain f = 2, 3 expectations are given in Table 2. The other expectations required may be obtained by equating or interchanging H and E.

(Table 2 near here)

As mentioned in the Introduction, these expectations are expressible in terms of Mardia's (1970) measures of skewness

$$\beta_{1,m} = \sum_{i,j,k=1}^{m} \kappa_{ijk}^2$$

and excess of kurtosis

$$\gamma_{2,m} = \sum_{i,j=1}^{m} \kappa_{iijj} = \beta_{2,m} - m(m + 2)$$

(noting that we have taken $\Sigma = I_m$), together with a supplementary skewness measure

$$\psi_m = \sum_{i,j,k=1}^{m} (\kappa_{ijj}\kappa_{ikk} - \kappa_{ijk}^2).$$

Further reductions would now be required for any specific design matrix A and hypothesis $C\xi = 0$. Two important situations are:

(a) Hotelling's one-sample T^2 tests for zero mean vector, $\xi = 0$. In this case $d_{\alpha\beta} = N^{-1}$, $a_{\alpha\beta} = \delta_{\alpha\beta} - N^{-1}(\alpha, \beta = 1, \ldots, N)$, where $\delta_{\alpha\beta}$ is Kronecker's delta; the expectations are given in Table 3.

(b) One-way analysis of variance, with k samples of sizes $N_1, \ldots, N_k$, $\sum_{i=1}^{k} N_i = N$, assuming all N_i are $O(N^{-1})$. In this case, $n_1 = k - 1$, $n_2 = N - k$. As in the univariate case (Gayen, 1950), the results to the required order are expressible in terms of

$$\Delta = N \sum_{i=1}^{k} N_i^{-1} - k^2,$$

and are presented in Table 3. Δ is positive unless the sample sizes are equal, in which case it is zero.

(Table 3 near here)

5. Calculation of correction terms

To obtain multivariate extensions of Gayen's (1950) approximation, we finally require marginal integrals of $g(\beta)$ corresponding to the various test statistics. In view of the elementary basis for the $C_\phi^{\kappa,\lambda}(R, S)$ described in Section 3, it is easily seen that the polynomials p defined by (3.2) are expressible in terms of products of the power sums $s_r = \sum_{i=1}^{m} \beta_i^r$ of the roots. Thus, in the case of Wilks's criterion W, we require integrals of the form

$$H(s_{r_1} \ldots s_{r_t}) = \int_{\{W \leq W_\theta\}} s_{r_1} \ldots s_{r_t} \cdot B(m; n_1, n_2) \, \Pi\beta_i^{\frac{1}{2}(n_1-m-1)}$$

$$\Pi(1-\beta_i)^{\frac{1}{2}(n_2-m-1)} \, \Pi_{i<j} (\beta_i - \beta_j) \, \Pi d\beta_i,$$

where W_θ is the normal theory θ-point of W. Marginal integrals of this type have been evaluated by Pillai and his co-workers for the bivariate case in the course of work on power and robustness of the test statistics (for example, Pillai and Sudjana, 1975). For arbitrary m, replacing W_θ by the continuous variable w, it may be shown that

$$w\frac{d}{dw}H'(s_{r_1}\dots s_{r_t}) = \tfrac{1}{2}(n_1+n_2+r_1-3)H'(s_{r_1}\dots s_{r_t})$$

$$-\tfrac{1}{2}(n_2+r-1)H'(s_{r_1-1}\dots s_{r_t}) + \sum_{q=2}^{t} r_q\{H'(s_{r_2}\dots s_{r_q+r_1}\dots s_{r_t})$$

$$-H'(s_{r_2}\dots s_{r_q+r_1-1}\dots s_{r_t})\} + \tfrac{1}{2}\sum_{q=1}^{r_1-1}\{H'(s_{r_1-q}\,s_q s_{r_2}\dots s_{r_t})$$

$$-H'(s_{r_1-q-1}\,s_q s_{r_2}\dots s_{r_t})\},$$

where the prime denotes differentiation with respect to w. These differential-reduction equations may be solved in parallel with a differential equation of the author's (Davis,1979b) for the density function $(m^{-1}H(s_0))$ of W, and the correction terms calculated. For the one-way analysis of variance we obtain an expansion of the form

$$\text{Prob}\{W \leq W_\theta\} = \theta + \gamma_{2,m}[Q_1 + \Delta\cdot Q_2] + \beta_{1,m}[Q_3 + \Delta\cdot Q_4]$$
$$+ \psi_m[Q_5 + \Delta\cdot Q_6], \qquad (5.1)$$

where the Q_i depend only on m, n_1, n_2 and θ, and it may be shown that

$$Q_2 = -\{(n_1 + n_2 + 2)/2n_1n_2\}Q_1.$$

Table 4 presents some tabulations of Q_1 and Q_3, .., Q_6 for $\theta = 0.05$, m = 2, 4, 6, and n_1 = 1, 3, 5, 7; a more comprehensive tabulation up to m = 7 has been made. Similar correction terms have also been calculated for the one-sample Hotelling's T^2.

(Table 4 near here)

Table 3 reflects the essential robustness of W in that the correction terms are small and approach zero as n_2 becomes large. However, for sufficiently large skewness and kurtosis the effects on the real θ-level may be quite serious for moderate n_2; increases in kurtosis tend to lower the θ-level, whereas increases in skewness tend to raise it, the effect of skewness being more pronounced for lower n_2, that of kurtosis for larger n_2. Increasing inequality of sample sizes generally appears to influence the actual significance level in the opposite directions to the basic skewness and kurtosis effects. This continues the trends in Gayen's (1950) Table 4 (noting the change in sign of Δ), to which our correction terms reduce in the univariate case.

6. Comparisons with simulation results in the bivariate case.

Mardia (1975) has collected the results of various simulation studies on the effects of bivariate nonnormality on Hotelling's generalized T^2 in the one and two sample cases, and related them in broad terms to his measures of multivariate skewness and kurtosis. In Table 5 we compare approximations derived from (5.1) with simulation results for several bivariate distributions, most of them appearing in Chase and Bulgren's (1971) simulation study on Hotelling's generalized T^2. Only the k-sample situation with equal sample sizes has been considered. The simulations are based on 15,000 replications of the 8 samples of size 10 case, the generated variates being subdivided for the other parameter combinations. The distributions and their associated measures are as follows.

(a) Mixture of bivariate normal distributions (Hopkins and Clay, 1963),

$$f(x, y) = 0.8\nu(x, y|0, I_2) + 0.2\nu(x, y|0, \sigma^2 I_2)$$

where $\nu(x, y|\mu, \Sigma)$ denotes the bivariate normal density function with mean vector μ and covariance matrix Σ. For this distribution

$$\gamma_{2,2} = 32\ \{(\sigma^2 - 1)/(\sigma^2 + 4)\}^2,\ \beta_{1,2} = \psi_2 = 0.$$

(b) Bivariate gamma, $f(x, y) = (3!)^{-2}(xy)^3\exp(-x - y)$, $(x, y > 0)$.

$$\gamma_{2,2} = 3,\ \beta_{1,2} = 2,\ \psi_2 = 0.$$

(c) Bivariate double exponential, $f(x, y) = \frac{1}{4}\exp(-|x| - |y|)$, $(-\infty < x, y < \infty)$.

$$\gamma_{2,2} = 6,\ \beta_{1,2} = \psi_2 = 0.$$

(d) Bivariate exponential (Marshall and Olkin, 1967). Joint distribution of variates X, Y such that

$$\text{Prob}(X \geq x, Y \geq y) = \exp\{-x - y - \max(x, y)\},\ (0 < x, y < \infty).$$

The correlation between X and Y is $\rho = \lambda/(2 + \lambda)$. The skewness and kurtosis coefficients are given in Mardia (1975); we also have

$$\psi_2 = 2\rho^2(2 + \rho)/(1 + \rho)^2.$$

(e) Morgenstern bivariate uniform, $f(x, y) = 1 + 3\rho(1 - 2x)(1 - 2y)$, $(0 < x, y < 1)$, where ρ is the correlation. Skewness and kurtosis are also given in Mardia (1975), and $\psi_2 = 0$.

As mentioned in the Introduction, the approximation is seen to give a reasonable indication of the magnitude and direction of the effect of nonnormality on the actual significance level, even for distributions differing quite widely from the multinormal, provided their skewness and kurtosis parameters are not excessive. The supplementary skewness measure ψ_m appears to have a negligible effect in general.

REFERENCES

Chase, G.R. and Bulgren, W.G. (1971). A Monte Carlo investigation of the robustness of T^2. J. Amer. Statist. Ass. 66, 499-502.

Constantine, A.G. (1963). Some non-central distribution problems in multivariate analysis. Ann. Math. Statist. 34, 1270-84.

Davis, A.W. (1976). Statistical distributions in univariate and multivariate Edgeworth populations. Biometrika 63, 661-70.

Davis, A.W. (1979a). Invariant polynomials with two matrix arguments, extending the zonal polynomials. Multivariate Analysis V, Ed. P.R. Krishnaiah. To appear.

Davis, A.W. (1979b). On the differential equation for Meijer's $G^{p,o}_{p,p}$ function, and further tabulation of Wilks's likelihood ratio criterion. Biometrika. To appear.

Gayen, A.K. (1950). The distribution of the variance ratio in random samples of any size drawn from non-normal universes. Biometrika 37, 236-55.

Hopkins, J.W. and Clay, P.P.F. (1963). Some empirical distributions of bivariate T^2 and homoscedasticity criterion M under unequal variance and leptokursis. J. Amer. Statist. Ass. 58, 1048-53.

James, A.T. (1964). Distributions of matrix variates and latent roots derived from normal samples. Ann. Math. Statist. 35, 475-501.

Kaplan, E.L. (1952). Tensor notation and the sampling cumulants of k-statistics. Biometrika 39, 319-23.

Mardia, K.V. (1970). Measures of multivariate skewness and kurtosis with applications. Biometrika 57, 519-30.

Mardia, K.V. (1975). Assessment of multinormality and the robustness of Hotelling's T^2 test. Appl. Statist. 24, 163-71.

Marshall, A.W. and Olkin, I. A multivariate exponential distribution. J. Amer. Statist. Ass. 62, 30-44.

Pillai, K.C.S. and Sudjana (1975). Exact robustness studies of tests of multivariate hypotheses based on four criteria and their distribution problems under violations. Ann. Statist. 3, 617-36.

Smith, H., Gnanadesikan, R. and Hughes, J.B. (1962). Multivariate analysis of variance (MANOVA). Biometrics 18, 22-41.

Tiku, M.L. (1964). Approximating the general non-normal variance-ratio sampling distributions. Biometrika 51, 83-95.

Table 1. Low order invariant polynomials ((R) = trace R).

κ	λ	ϕ	$C_{\phi}^{\kappa,\lambda}(R,S)$
1	1	2	$\frac{1}{3}\{(R)(S) + 2(RS)\}$
		1^2	$\frac{2}{3}\{(R)(S) - (RS)\}$
2	1	3	$\frac{1}{15}\{(R)^2(S) + 4(RS)(R) + 2(R^2)(S) + 8(R^2S)\}$
		21	$\frac{2}{5}\{(R)^2(S) - (RS)(R) + 2(R^2)(S) - 2(R^2S)\}$
1^2	1	21	$5^{-\frac{1}{2}}\{(R)^2(S) + 2(RS)(R) - (R^2)(S) - 2(R^2S)\}$
		1^3	$\frac{1}{3}\{(R)^2(S) - 2(RS)(R) - (R^2)(S) + 2(R^2S)\}$

Table 2. Expectations of the $C_{\phi}^{\kappa,\lambda}(\Omega_1, \Omega_2)$.

κ	λ	ϕ	$E_Z\{C_{\phi}^{\kappa,\lambda}(\Omega_1, \Omega_2)\}$
1	1	2	$\frac{1}{4}\gamma_{2,m}\sum_{\alpha=1}^{N} d_{\alpha\alpha}e_{\alpha\alpha}$
		1^2	0
2	1	3	$\frac{1}{4}(\beta_{1,m} + \frac{3}{5}\psi_m)\sum_{\alpha,\beta}^{N}(2h_{\alpha\alpha}h_{\alpha\beta}e_{\beta\beta} + h_{\alpha\alpha}e_{\alpha\beta}h_{\beta\beta} + 2h_{\alpha\beta}^2e_{\alpha\beta})$
		21	$\frac{1}{15}\gamma_m\sum_{\alpha,\beta}^{N}(3h_{\alpha\alpha}h_{\alpha\beta}e_{\beta\beta} - h_{\alpha\alpha}e_{\alpha\beta}h_{\beta\beta} - 2h_{\alpha\beta}^2e_{\alpha\beta})$
1^2	1	21	$\frac{1}{6}\gamma_m\sum_{\alpha,\beta}^{N}(h_{\alpha\alpha}h_{\beta\beta} - h_{\alpha\beta}^2)e_{\alpha\beta}$
		1^3	0

Table 3. Expectations (×N) for (a) one-sample T^2 and (b) k-sample tests ($\zeta_m = \beta_{1,m} + \frac{3}{5}\psi_m$).

κ	λ	φ	(a)	(b)
2	0	2	$\frac{1}{4}\gamma_{2,m}$	$\frac{1}{4}(n_1^2 + \Delta)\gamma_{2,m}$
1	1	2	$(N-1)\gamma_{2,m}$	$\frac{1}{4}(n_1 n_2 - \Delta)\gamma_{2,m}$
0	2	2	$\frac{1}{4}(N-1)^2\gamma_{2,m}$	$\frac{1}{4}(n^2 + \Delta)\gamma_{2,m}$
3	0	3	$\frac{5}{4}\zeta_m$	$\frac{1}{4}\{2n_1(n_1 - 1) + 5\Delta\}\zeta_m$
21	0	21	0	$-\frac{3}{10}n_1(n_1 - 1)\psi_m$
2	1	3	$\frac{1}{2}(N-1)\zeta_m$	$-\frac{1}{2}\Delta\cdot\zeta_m$
2	1	21	$\frac{1}{5}(N-1)\psi_m$	$-\frac{1}{5}\Delta\cdot\gamma_m$
1	2	3	$\frac{1}{4}(N^2-1)\zeta_m$	$\frac{1}{4}(2n_1 n_2 - \Delta)\zeta_m$
1	2	21	$-\frac{1}{15}(N^2-1)\psi_m$	$\frac{1}{15}(-2n_1 n_2 + \Delta)\psi_m$
1	1^2	21	$\frac{1}{6}(N-1)(N-2)\psi_m$	$+\frac{1}{6}(-n_1 n_2 + 2\Delta)\psi_m$
0	3	3	$\frac{1}{2}(N-1)(N-2)\zeta_m$	$\frac{1}{2}\{n_2(n_2 - n_1 - 1) + 2\Delta\}\zeta_m$
0	21	21	$\frac{12}{5}(N-1)(N-2)\psi_m$	$-\frac{3}{10}\{n_2(n_2 - n_1 - 1) + 2\Delta\}\psi_m$

Table 4. Correction terms ($\times 10^5$) for Wilks's likelihood ratio criterion in the k-sample case. $\theta = 0.05$.

$m = 2,\ n_1 = 1$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
2	-87	150	-87	159	-139
3	-171	274	-102	233	-130
4	-213	302	-93	238	-105
5	-224	276	-83	211	-85
6	-220	236	-76	177	-72
8	-197	163	-69	121	-57
12	-150	81	-61	60	-43
20	-96	28	-48	21	-31
24	-81	18	-43	14	-28
30	-65	11	-38	8	-24
40	-49	5	-31	4	-19
60	-32	2	-22	2	-13
120	-16	0	-12	0	-7

$m = 2,\ n_1 = 3$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
2	-92	199	-47	157	-47
3	-160	331	-50	204	-38
4	-195	376	-42	225	-31
5	-211	374	-33	225	-26
6	-216	350	-26	214	-21
8	-209	284	-17	179	-14
12	-180	177	-9	118	-8
20	-131	76	-6	56	-5
24	-114	53	-5	42	-4
30	-95	32	-4	28	-3
40	-75	16	-4	17	-3
60	-52	4	-3	8	-2
120	-27	-2	-2	2	-1

$m = 2,\ n_1 = 5$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
2	-81	215	-33	155	-29
3	-139	358	-36	187	-21
4	-171	413	-30	205	-17
5	-186	419	-24	208	-14
6	-193	402	-19	202	-11
8	-191	346	-12	178	-7
12	-171	238	-5	128	-4
20	-131	119	-2	68	-2
24	-116	89	-1	52	-1
30	-99	60	-1	37	-1
40	-79	34	-1	23	-1
60	-56	14	-1	11	0
120	-29	2	0	3	0

$m = 2,\ n_1 = 7$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
2	-70	222	-26	154	-21
3	-121	373	-28	179	-15
4	-149	435	-24	193	-12
5	-164	448	-19	197	-9
6	-171	437	-15	193	-8
8	-173	387	-10	174	-5
12	-159	282	-4	130	-2
20	-126	155	-1	74	-1
24	-113	120	-1	58	0
30	-97	85	0	42	0
40	-79	53	0	26	0
60	-56	25	0	13	0
120	-30	6	0	4	0

Table 4 (continued)

$m = 4,\ n_1 = 1$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
4	-22	29	-11	23	-11
5	-46	59	-15	45	-16
6	-64	76	-15	56	-16
8	-79	80	-12	57	-14
12	-78	56	-11	39	-10
20	-59	26	-11	18	-8
24	-51	19	-11	13	-7
30	-43	12	-10	8	-7
40	-33	7	-9	5	-6
60	-23	3	-7	2	-4
120	-12	1	-4	1	-3

$m = 4,\ n_1 = 3$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
4	-32	53	-7	38	-7
5	-56	91	-10	58	-8
6	-70	108	-9	66	-7
8	-82	111	-7	66	-5
12	-81	87	-3	51	-3
20	-65	48	-1	28	-1
24	-58	36	-1	22	-1
30	-50	25	0	15	0
40	-40	15	0	10	0
60	-28	7	0	5	0
120	-15	2	0	1	0

$m = 4,\ n_1 = 5$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
4	-33	66	-6	45	-5
5	-56	110	-7	66	-5
6	-68	129	-7	72	-5
8	-78	132	-5	71	-3
12	-78	107	-2	55	-2
20	-64	63	-1	32	0
24	-58	50	0	25	0
30	-50	37	0	18	0
40	-40	24	0	12	0
60	-29	12	0	6	0
120	-16	4	0	2	0

$m = 4,\ n_1 = 7$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
4	-31	74	-5	50	-4
5	-52	123	-6	70	-4
6	-64	144	-6	77	-4
8	-73	148	-4	74	-3
12	-73	123	-2	58	-1
20	-61	76	-1	35	0
24	-56	61	0	27	0
30	-49	46	0	20	0
40	-40	31	0	13	0
60	-29	17	0	7	0
120	-16	6	0	2	0

Table 4 (continued)

$m = 6, n_1 = 1$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
6	-9	9	-3	7	-3
8	-27	28	-4	20	-5
12	-41	32	-3	22	-3
20	-38	19	-2	13	-2
24	-34	14	-3	10	-2
30	-29	10	-3	7	-2
40	-24	6	-3	4	-2
60	-17	3	-3	2	-2
120	-9	1	-2	0	-1

$m = 6, n_1 = 3$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
6	-15	20	-2	14	-2
8	-34	44	-3	27	-2
12	-44	47	-2	27	-2
20	-40	30	0	17	0
24	-37	24	0	14	0
30	-32	18	0	10	0
40	-26	11	0	6	0
60	-19	6	0	3	0
120	-10	2	0	1	0

$m = 6, n_1 = 5$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
6	-16	27	-2	18	-1
8	-36	56	-3	32	-2
12	-44	58	-1	31	-1
20	-40	39	0	19	0
24	-36	32	0	16	0
30	-32	24	0	11	0
40	-26	16	0	7	0
60	-19	9	0	4	0
120	-11	3	0	1	0

$m = 6, n_1 = 7$

n_2	Q_1	Q_3	Q_4	Q_5	Q_6
6	-17	32	-2	21	-1
8	-35	65	-2	36	-1
12	-42	66	-1	34	-1
20	-38	46	0	21	0
24	-35	38	0	17	0
30	-31	29	0	13	0
40	-26	20	0	8	0
60	-19	11	0	4	0
120	-11	4	0	1	0

Table 5. Actual 100 θ % significance level of Wilks's likelihood ratio criterion.

Comparison of approximate (A) and simulation levels (S), standard errors (SE), based on ≳ 15,000 replications. Nominal 5% level.

		Mixture of normals			Bivariate Gamma	Double Exponential	Bivariate Exponential			Bivariate Uniform	
		$\sigma = 1.5$	$\sigma = 2.0$	$\sigma = 2.5$			$\lambda = 0$	$\lambda = 2/3$	$\lambda = 2$	$\rho = 0$	$\rho = 0.25$
$\gamma_{2,2}$		1.28	4.5	8.395	3	6	12	14	18.667	-2.4	-2.368
$\beta_{1,2}$		0	0	0	2	0	8	8.284	8.926	0	0
ψ_2		0	0	0	0	0	0	0.18	0.556	0	0
Sample sizes		2 samples									
5	A	4.75	4.11	3.35	4.74	3.82	3.94	3.61	2.84	5.47	5.47
	S	4.70	4.31	3.93	4.75	3.93	3.96	3.54	2.95	5.70	5.72
	SE	0.06	0.06	0.06	0.06	0.06	0.06	0.05	0.05	0.07	0.07
6	A	4.77	4.19	3.50	4.69	3.93	3.76	3.45	2.71	5.43	5.42
	S	4.79	4.40	3.97	4.74	4.12	3.98	3.71	3.05	5.61	5.60
	SE	0.07	0.06	0.06	0.07	0.06	0.06	0.06	0.05	0.07	0.07
8	A	4.83	4.40	3.88	4.72	4.20	3.88	3.64	3.08	5.32	5.31
	S	4.79	4.47	4.07	4.72	4.21	4.01	3.80	3.35	5.39	5.22
	SE	0.08	0.08	0.07	0.08	0.07	0.07	0.07	0.07	0.08	0.08
10	A	4.86	4.52	4.11	4.75	4.36	4.01	3.81	3.35	5.25	5.25
	S	4.95	4.72	4.38	4.79	4.42	4.21	3.93	3.52	5.24	5.22
	SE	0.09	0.09	0.08	0.08	0.08	0.08	0.08	0.08	0.09	0.09

Table 5 (continued)

Sample sizes		4 samples									
5	A	4.80	4.31	3.72	4.77	4.08	4.07	3.81	3.20	5.37	5.26
	S	4.75	4.40	4.02	4.72	4.24	4.25	4.09	3.59	5.44	5.38
	SE	0.09	0.08	0.08	0.09	0.08	0.08	0.08	0.08	0.09	0.09
6	A	4.83	4.41	3.90	4.76	4.21	4.04	3.81	3.26	5.31	5.31
	S	4.94	4.53	4.21	4.78	4.32	4.26	4.11	3.70	5.33	5.33
	SE	0.10	0.09	0.09	0.10	0.09	0.09	0.09	0.08	0.10	0.10
8	A	4.87	4.55	4.15	4.77	4.39	4.09	3.91	3.45	5.24	5.24
	S	4.81	4.61	4.31	4.85	4.37	4.25	4.23	3.90	5.25	5.17
	SE	0.11	0.11	0.10	0.11	0.11	0.10	0.10	0.10	0.12	0.11
10	A	4.90	4.63	4.31	4.80	4.51	4.18	4.03	3.67	5.20	5.19
	S	5.07	4.89	4.55	4.95	4.68	4.31	4.14	4.05	5.06	5.23
	SE	0.13	0.12	0.12	0.13	0.12	0.12	0.11	0.11	0.13	0.13

Table 5 (continued). Actual 100 θ % significance level of Wilks's likelihood ratio criterion. Comparison of approximate (A) and simulation levels (S), standard errors (SE), based on ≥ 15,000 replications. Nominal 5% level.

		Mixture of normals			Bivariate Gamma	Double Exponential	Bivariate Exponential			Bivariate Uniform	
		$\sigma = 1.5$	$\sigma = 2.0$	$\sigma = 2.5$			$\lambda = 0$	$\lambda = 2/3$	$\lambda = 2$	$\rho = 0$	$\rho = 0.25$
$\gamma_{2,2}$		1.28	4.5	8.395	3	6	12	14	18.667	-2.4	-2.368
$\beta_{1,2}$		0	0	0	2	0	8	8.284	8.926	0	0
ψ_2		0	0	0	0	0	0	0.18	0.556	0	0
Sample sizes					8 samples						
5	A	4.88	4.58	4.22	4.87	4.44	4.49	4.33	3.96	5.22	5.22
	S	5.09	4.77	4.52	4.67	4.52	4.61	4.57	4.27	4.94	5.34
	SE	0.13	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.13	0.13
6	A	4.90	4.64	4.34	4.87	4.53	4.48	4.34	4.01	5.19	5.19
	S	4.94	4.72	4.46	4.96	4.54	4.62	4.53	4.52	5.04	5.24
	SE	0.14	0.13	0.13	0.14	0.13	0.13	0.13	0.13	0.14	0.14
8	A	4.92	4.73	4.50	4.88	4.64	4.51	4.40	4.15	5.14	5.14
	S	5.06	4.68	4.43	4.86	4.58	4.69	4.66	4.36	4.91	4.97
	SE	0.16	0.15	0.15	0.16	0.15	0.15	0.15	0.15	0.16	0.16
10	A	4.94	4.78	4.60	4.89	4.71	4.57	4.48	4.27	5.12	5.11
	S	5.08	4.89	4.71	4.98	4.63	4.59	4.67	4.82	4.66	5.33
	SE	0.18	0.18	0.17	0.18	0.17	0.17	0.17	0.17	0.17	0.18

Austral. J. Statist., **19** (**3**), 1977, 206–212.

ASYMPTOTIC THEORY FOR PRINCIPAL COMPONENT ANALYSIS: NON-NORMAL CASE[1]

A. W. DAVIS
CSIRO Division of Mathematics and Statistics, Adelaide

Summary

This paper considers the asymptotic distributions of latent roots and vectors in principal components analysis when the parent population is non-normal. It is shown that sufficient of T. W. Anderson's asymptotic theory in the multivariate normal case carries over for some results to be obtained.

1. Introduction

Consider a p-variate distribution $\Phi(x)$ whose fourth order moments exist, and whose mean vector and covariance matrix are μ and Σ, respectively. Let $\delta_1 \geqslant \delta_2 \geqslant \ldots \geqslant \delta_p > 0$ be the latent roots of Σ, and $\gamma_1, \ldots, \gamma_p$ the corresponding normalized latent vectors, so that

$$\Gamma'\Sigma\Gamma = \Delta = \text{diag}(\delta_i), \tag{1.1}$$

where $\Gamma = (\gamma_1, \ldots, \gamma_p)$, $\Gamma'\Gamma = I_p$, the $p \times p$ unit matrix. Then the linear combination $\gamma_i'x$ is referred to as the ith principal component of x, and has variance $\delta_i (i = 1, \ldots, p)$.

Suppose that $x_1, \ldots, x_N$ is a sample of $N = n+1$ independent observations drawn from $\Phi(x)$. Then Σ has the unbiased estimate

$$S = n^{-1} \sum_{\alpha=1}^{N} (x_\alpha - \bar{x})(x_\alpha - \bar{x})', \tag{1.2}$$

where $\bar{x}$ is the sample mean vector, and the δ_i and γ_i are estimated by the latent roots $d_1 \geqslant \ldots \geqslant d_p$ and associated latent vectors $c_1, \ldots, c_p$ of S, respectively. Corresponding to (1.1) we have

$$C'SC = D = \text{diag}(d_i) \tag{1.3}$$

where $C = (c_i, \ldots, c_p)$, $C'C = I_p$.

In the case of p-variate normal $\Phi(x)$, Anderson (1963) (referred to hereafter as TWA) has discussed the asymptotic distribution theory of the d's and c's as $n \to \infty$, and applied this to certain problems of

[1] Manuscript received March 29, 1977.

inference relating to the δ's and γ's. Our purpose in this paper is to show that his approach carries over to yield some asymptotic results for non-normal $\Phi(x)$. The main difficulty encountered is that the distributions derived involve the fourth order moments of $\Phi(x)$, as one would expect in the asymptotic theory of second-order statistics; and since only sample estimates of these moments will generally be available, the results must be regarded as approximations for inferential purposes. On the other hand, they may provide some indication of the errors involved in applying the normal-theory results to non-normal data. The importance of investigating the effects of non-normality on traditional multivariate methods has been recently emphasized, in particular, by Moran (1975).

Asymptotic distribution theory of the d's and c's in the non-normal case is presented in Section 2, and applied in Section 3 to prove asymptotic normality of d_i and c_i for a root δ_i of multiplicity 1. Sections 4 and 5 consider the implications of non-normality for certain normal-theory inference techniques associated with the γ's and δ's, respectively. The notation in TWA has been adhered to for convenience.

This research was stimulated by the work of Dudziński *et al.* (1975).

2. Asymptotic Distribution of the Sample Roots and Vectors

Following TWA, we work in terms of the transformed variate $y=\Gamma'(x-\mu)$, which has mean vector 0 and covariance matrix Δ. The sample now becomes $y_1,\ldots,y_N$ and the corresponding sample covariance matrix is $T=\Gamma'S\Gamma$. T has the same latent roots d_i as S, and is reduced to diagonal form by the orthogonal matrix $E=(e_{ij})=\Gamma C$,

$$E'TE=D. \tag{2.1}$$

For uniqueness, the diagonal elements of E are taken to be positive. The argument in TWA is based on the application of the Lindeberg–Levy multivariate central limit theorem (Rao (1965) p. 108) to the matrix T in the normal case; asymptotic normality of T continues to hold in the non-normal case under the assumption that the fourth moments of $\Phi(x)$ exist. If

$$n^{1/2}(T-\Delta)=U=(u_{ij}), \tag{2.2}$$

then the u_{ij} are asymptotically jointly normal with zero means, and covariances

$$\operatorname{cov}(u_{ij},u_{kl})=\kappa_{ijkl}+\delta_i\delta_j(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}). \tag{2.3}$$

Hence δ_{ij} denotes Kronecker's delta, and the fourth order cumulants of $y=(y^{(1)},\ldots,y^{(p)})'$ are defined by

$$E(y^{(i)}y^{(j)}y^{(k)}y^{(l)})=\kappa_{ijkl}+\delta_{ij}(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})+\delta_i\delta_k\delta_{ij}\delta_{kl}, \tag{2.4}$$

$(i, j, k, l=1,\ldots,p)$. The κ_{ijkl} vanish of course in the normal case.

To investigate the asymptotic behaviour of the c_i and d_i, suppose that the roots of Σ have multiplicities $q_1,\ldots,q_r$ ($\sum_{i=1}^r q_i=p$),

$$\begin{aligned} \delta_1&=\ldots=\delta_{q_1}=\lambda_1,\\ &\vdots\\ \delta_{q_1+\ldots+q_{n-1}+1}&=\ldots=\delta_p=\lambda_r \end{aligned} \tag{2.5}$$

where $\lambda_1>\ldots>\lambda_r>0$. Corresponding to these multiplicities, we partition the matrices Δ, U, D, E into submatrices (Δ_{ij}) etc,., as in equations (2.15)–(2.18) of TWA. We also define the diagonal matrix

$$H=n^{1/2}(D-\Delta)=\operatorname{diag}(h_i). \tag{2.6}$$

Writing $n^{1/2}E_{kl}=F_{kl}$ $(k\neq l)$, the derivation of the following equations ((2.21)–(2.24) of TWA) carries through without change and in the same notation,

$$\begin{aligned} E_{kk}E'_{kk}&=I_{q_k}-n^{-1}W_{kk},\\ U_{kk}&=E_{kk}H_kE'_{kk}+n^{-1/2}(M_{kk}-\lambda_kW_{kk}),\\ U_{kl}&=\lambda_kE_{kk}F'_{lk}+\lambda_lF_{kl}E'_{ll}+n^{-1/2}M_{kl},\quad k\neq l,\\ 0&=E_{kk}F'_{lk}+F_{kl}E'_{ll}+n^{-1/2}W_{kl},\quad k\neq l. \end{aligned} \tag{2.7}$$

In TWA Section 7 it is shown that the H_k, E_{kk}, and F_{kl} have limiting distributions uniquely defined in terms of the distribution of U by the equations

$$E_{kk}E'_{kk}=I_{q_k}, \tag{2.8}$$

$$U_{kk}=E_{kk}H_kE'_{kk}, \tag{2.9}$$

$$U_{kl}=\lambda_kE_{kk}F'_{lk}+\lambda_lF_{kl}E'_{ll}, \tag{2.10}$$

$$0=E_{kk}F'_{lk}+F_{kl}E'_{ll}. \tag{2.11}$$

This proof depends only upon the existence of a continuous limiting distribution of U, and in no way involves the normality of the parent distribution; it thus applies to the non-normal case as well.

From (2.8) and (2.9), the limiting joint distribution of E_{kk} and H_k is uniquely determined by the requirement that E_{kk} is orthogonal with positive diagonal elements, H_k is diagonal, and $E_{kk}H_kE'_{kk}$ is the symmetric matrix U_{kk}. In fact, if $f_k(U_{kk})$ denotes the normal joint density of the functionally independent elements of U_{kk}, defined by

(2.3), then the joint density of E_{kk} and H_k has the form

$$C_k f_k(E_{kk}H_kE'_{kk})\prod_{i<j}(h_i-h_j)\prod_l dh_i(dE_{kk}), \tag{2.12}$$

where i, j, $l \in L_K$, the set of subscripts in the kth line of (2.5), (dE_{kk}) denotes the conditional Haar invariant distribution over the $q_k \times q_k$ orthogonal matrices with positive diagonal elements, and C_k is a normalizing constant. (See, for example, James (1966), Section 4). Unfortunately the marginal distributions of E_{kk} and H_k will not in general take the simple forms obtained in the normal case (TWA Theorem 1), because of the correlation structure of the u_{ij} (equation (2.3)).

From (2.10) and (2.11), however, $E_{kk}F'_{lk}$ and $-F_{kl}E'_{ll}$ are asymptotically equivalent to each other and to $U_{kl}/(\lambda_k-\lambda_l)$. The elements of the latter are jointly normal with zero means, and covariances obtainable from (2.3). From (2.7), $n^{1/2}(E_{kk}E'_{kk}-I_{q_k})$ converges in probability to 0.

We note that in terms of the original coordinates, the d_i are the latent roots of S, and the latent vectors are given by $C=\Gamma E$. I.e. if C is partitioned as $(C_1,\ldots,C_r)$ and Γ as $(\Gamma_1,\ldots,\Gamma_r)$, then

$$C_k=\Gamma_kE_{kk}+n^{-1/2}\sum_{l\neq k}\Gamma_lF_{lk},\quad (k=1,\ldots,r). \tag{2.13}$$

3. Root of Multiplicity 1

Suppose that $q_k=1$, so that C_k is a vector, say c_i, the ith column of C. Then $E_{kk}=e_{ii}$, and $n^{1/2}(e_{ii}-1)$ converges in probability to 0. Hence, from (2.9), $n^{1/2}(d_i-\delta_i)=h_i$ is asymptotically equivalent to u_{ii}, i.e. is asymptotically normal with mean 0 and variance $2\delta_i^2+\kappa_{iiii}$, (regardless of the multiplicity of the other roots). Also, from (2.13),

$$c_i=e_{ii}\gamma_i+n^{-1/2}\sum_{\substack{l=1\\l\neq k}}^{r}\Gamma_lF_{lk}, \tag{3.1}$$

so that

$$n^{1/2}(c_i-\gamma_i)\sim\sum_{\substack{l=1\\l\neq k}}^{r}\Gamma_lF_{lk}\sim\sum_{\substack{l=1\\l\neq k}}^{r}\Gamma_lU_{lk}/(\lambda_k-\lambda_l). \tag{3.2}$$

Let $\Gamma^*_{(i)}$ denote Γ with its ith column deleted, $U_{(i)}$ the ith column of U with u_{ii} deleted, and

$$P_{(i)}=\operatorname{diag}((\delta_i-\delta_j)^{-1},\quad j\neq i). \tag{3.3}$$

Then (3.2) may be written in the form

$$n^{1/2}(c_i-\gamma_i)\sim\Gamma^*_{(i)}P_{(i)}U_{(i)}, \tag{3.4}$$

which is asymptotically jointly normal with zero means and covariance matrix $\Gamma^*_{(i)}P_{(i)}\Psi_{(i)}P_{(i)}\Gamma^{*\prime}_{(i)}$, where $\Psi_{(i)}$ is the covariance matrix of $U_{(i)}$. Thus $A_{(i)} = P_{(i)}\Psi_{(i)}P_{(i)}$ has (g, h) element

$$\frac{\kappa_{iigh} + \delta_i\delta_g\delta_{gh}}{(\delta_i - \delta_g)(\delta_i - \delta_h)}, \quad (g, h = 1, \ldots, p;\ g, h \neq i). \tag{3.5}$$

If θ_i denotes the angle between c_i and γ_i, then

$$n(1 - \cos \theta_i) = n(c_i - \gamma_i)'(c_i - \gamma_i) \sim U'_{(i)}P^2_{(i)}U_{(i)},$$

a quadratic form in the elements of $U_{(i)}$. Thus if u is a vector of $p-1$ independent standard normal variates, then $n(1 - \cos \theta_i)$ has the same distribution as $u'A_{(i)}u$. This result is relevant to the investigations in Dudziński *et al.* (1975). A good approximation to the distribution of the above quadratic form is $\alpha\chi^2_\beta$, where

$$\begin{aligned} \alpha &= \text{trace}\,(A^2_{(i)})/\text{trace}\,(A_{(i)}), \\ \beta &= (\text{trace}\, A_{(i)})^2/\text{trace}\,(A^2_{(i)}). \end{aligned} \tag{3.6}$$

It should be noted that the κ_{ijkl} are fourth order cumulants of the transformed variate y, not of the original variate x.

4. Asymptotic Test for a Given Principal Component

In TWA Appendix B, a test is presented for the null hypothesis that γ_1, the vector of coefficients of the first principal component, is a specified unit vector γ_1^0, under the assumption that the corresponding root δ_1 has multiplicity 1. Under normal theory, $n^{1/2}(c_1 - \gamma_1^0)$ has covariance matrix $\Gamma^*_{(1)}A^0_{(1)}\Gamma^{*\prime}_{(1)}$, where $A_1^{(0)}$ is the matrix $A_{(1)}$ with the fourth order cumulants replaced by zeros ($A^0_{(1)} = \Lambda^2$ in Anderson's notation). Anderson forms the statistic

$$Q = n(c_1 - \gamma_1^0)'\Gamma^*_{(1)}A^{0-1}_{(1)}\Gamma^{*\prime}_{(1)}(c_1 - \gamma_1^0), \tag{4.1}$$

which has a limiting χ^2_{p-1} distribution in the normal case.

In the non-normal case, Q has the same limiting distribution as the quadratic form $u'Bu$, where u is a vector of $(p-1)$ independent standard normal variates and $B = A^{0-1/2}_{(1)}A_{(1)}A^{0-1/2}_{(1)}$, with (g, h) element

$$\delta_{gh} + \kappa_{11gh}/\delta_1\sqrt{\delta_g\delta_h}, \quad (g, h = 2, \ldots, p). \tag{4.2}$$

5. Statistical Inference on the Latent Roots

Let L_k denote the set of subscripts in the kth line of (2.5). Anderson shows in TWA Section 3 that the normal theory likelihood ratio criterion for testing equality of the roots δ_i, $i \in L_k$, is

$$l_k = \left[\prod_{i \in L_k} d_i \Big/ \left(q_k^{-1} \sum_{j \in L_k} d_j\right)^{q_k}\right]^{N/2}. \tag{5.1}$$

Hence, in virtue of (2.6) and (2.9), $-2 \log l_k$ is asymptotically equivalent to

$$(1/2\lambda_k^2)\left\{2\sum_{i<j} u_{ij}^2 + \sum_i u_{ii}^2 - q_k^{-1}\left(\sum_i u_{ii}\right)^2\right\}, \tag{5.2}$$

as in TWA, where the summations extend over L_k. If we define the $\frac{1}{2}q_k(q_k+1)$ vector

$$U_{(k)} = [\ldots, u_{ii}, \ldots \mid \ldots, u_{jl}, \ldots] \tag{5.3}$$

where $i, j, l \in L_k$, $j<l$, then from (2.3)

$$\begin{aligned} &\operatorname{cov}(u_{ii}, u_{jj}) = 2\delta_{ij}\lambda_k^2 + \kappa_{iijj}, \\ &\operatorname{cov}(u_{ii}, u_{jl}) = \kappa_{iijl}, \quad (j<l), \\ &\operatorname{cov}(u_{ij}, u_{lm}) = \delta_{il}\delta_{jm}\lambda_k^2 + \kappa_{ijlm}, \quad (i<j, l<m). \end{aligned} \tag{5.4}$$

Let

$$\kappa_{klm} = q_k^{-1} \sum_{i \in L_k} \kappa_{iilm}, \quad \kappa_{kk} = q_k^{-1} \sum_{l \in L_k} \kappa_{kll} \tag{5.5}$$

and define κ_{ijk} correspondingly. Then the quadratic form (5.2) has the same distribution as $v'Av$, where v is a $\frac{1}{2}q_k(q_k+1)$ vector of independent standard normal variates, and $A=(a_{ij;lm})$, partitioned in accordance with (5.3), where

$$\begin{aligned} &a_{ii;ll} = (\delta_{il} - q_k^{-1}) + (\kappa_{iill} - \kappa_{iik} - \kappa_{kll} + \kappa_{kk})/2\lambda_k^2 \\ &a_{ii,lm} = (\kappa_{iilm} - \kappa_{klm})/\sqrt{2}\,\lambda_k^2, \quad (l<m), \\ &a_{ij,lm} = (\delta_{il}\delta_{jm} + \kappa_{ijlm}/\lambda_k^2), \quad (i<j, l<m). \end{aligned} \tag{5.6}$$

If it is assumed that q_k roots are equal, say the last q_r, then in the normal case a confidence interval for λ_r may be obtained based on $\bar{d} = q_r^{-1}\sum_{p-q_r+1}^{p} d_i$, the average of the q_r smallest roots of S (TWA Section 3). In the non-normal case,

$$n^{1/2}(\bar{d} - \lambda_r) = q_r^{-1} \operatorname{trace} U_{rr} = q_r^{-1} \sum_{p-q_r+1}^{p} u_{ii}, \tag{5.7}$$

is asymptotically normal with mean 0 and variance $2\lambda_r^2/q_r + \kappa_{rr}$. A confidence interval based on this result would thus depend on κ_{rr}; clearly the normal theory interval will tend to be conservative if $\kappa_{rr}<0$, and too short if $\kappa_{rr}>0$. If the last q_r roots are in fact not all equal, and have variance $\operatorname{var}(\delta)$, the result will depend on the sign of $\kappa_{rr} + 2\operatorname{var}(\delta)/q_r$.

Finally, in testing whether the ratio of "unexplained variance" to the total does not exceed the fraction f, i.e. that

$$\sum_{i=q+1}^{p} \delta_i \leq f \sum_{i=1}^{p} \delta_i \tag{5.8}$$

for some $q(\delta_q > \delta_{q+1})$, the normal theory test based on $-f\sum_i^q d_i + (1-f)\sum_{q+1}^p d_i$ will be conservative or not according as

$$f^2 \sum_{i,j=1}^{q} \kappa_{iijj} - 2f(1-f) \sum_{i=1}^{q} \sum_{j=q+1}^{p} \kappa_{iijj} + (1-f)^2 \sum_{i,j=q+1}^{p} \kappa_{iijj} \tag{5.9}$$

is negative or positive.

Bibliography

Anderson, T. W. (1963). "Asymptotic theory for principal component analysis." *Ann. Math. Statist.* 34, 122–48.

Dudziński, M. L., Norris, J. M., Chmura, J. T. & Edwards, C. B. H. (1975). "Repeatability of principal components in samples: normal and non-normal data sets compared." *Multivariate Behavioral Research* 10, 109–18.

James, A. T. (1966). "Inference on latent roots by calculation of hypergeometric functions of matrix argument." *Multivariate Analysis* (P. R. Krishnaiah, Ed.). Academic Press, New York, 209–35.

Moran, P. A. P. (1975). "What should a professor of statistics do?" *Austral. J. Statist.* 17, 121–33.

Rao, C. R. (1965). *Linear Statistical Inference and its Applications.* John Wiley and Sons, New York.

Section 6.

Miscellaneous.

A Note on a Problem Posed by Fisher

A. W. Davis

DIVISION OF MATHEMATICAL STATISTICS TECHNICAL PAPER NO. 26
COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANIZATION, AUSTRALIA 1969

A Note on a Problem Posed by Fisher

By A. W. Davis

Division of Mathematical Statistics Technical Paper No. 26

Commonwealth Scientific and Industrial
Research Organization, Australia

Melbourne 1969

Printed by CSIRO, Melbourne

A NOTE ON A PROBLEM POSED BY FISHER

By A. W. Davis*

Summary

Fisher's asymptotic expansion of the mean of the fiducial distribution of the binomial parameter is shown to be termwise identical with the expansion of a Bayesian mean, as apparently conjectured by Fisher.

I. Introduction

If the binomial parameter p is postulated to have the probability element *a priori*

$$\frac{1}{\pi\sqrt{(pq)}}\mathrm{d}p\,, \qquad (q = 1-p,\ 0 < p < 1)\,, \tag{1}$$

then for a observed successes in $n = (a+b)$ trials, the mean $\bar{p}_{\mathrm{B}}$ of the corresponding Bayesian distribution *a posteriori* is

$$\bar{p}_{\mathrm{B}} = \frac{a+\frac{1}{2}}{n+1} = \frac{a}{n} + \frac{(b-a)}{2}\sum_{k=2}^{\infty}\left(-\frac{1}{n}\right)^{k}. \tag{2}$$

Since the data are discontinuous, Fisher admitted (1959, p. 60) that his fiducial argument is not strictly applicable. However, he claimed that for moderately large frequencies (of order 1000 or 10 000) the discontinuity may be ignored for practical purposes, and in several places (Fisher 1959, p. 62; Fisher and Cornish 1960, §IV(b)) he sought to obtain a formal asymptotic expansion for $\bar{p}_{\mathrm{F}}$, the mean of the fiducial distribution of p. This was found to agree with (2) as far as the term in n^{-4}. To clear up the purely computational issue, it is shown below that Fisher's method does yield a series which is termwise identical with (2). However, a rigorous justification of his method lies beyond the scope of the present note.

II. Fisher's Expansion of $\bar{p}_{\mathrm{F}}$

As a first step, Fisher assumes that for large n, the departure from normality of the binomial variate a may be corrected using the Cornish–Fisher expansion. The latter expresses percentiles of asymptotically normal variates in terms of their cumulants and the corresponding standard normal percentiles (Fisher and Cornish 1960), and is known to be a valid asymptotic expansion for a large class of continuous

* Division of Mathematical Statistics, CSIRO, Private Bag No. 1, Glen Osmond, S.A. 5064.

distributions (Wallace 1958). Following his claim that for large n the discontinuity may be ignored, Fisher obtains (Fisher and Cornish 1960, p. 220)

$$\left.\begin{aligned} a/n &\sim p + x\sqrt{(pq/n)} + \tfrac{1}{6}(q-p)(x^2-1)/n + \ldots \\ &\sim p - \phi_{x,n}(p) \quad \text{say}, \end{aligned}\right\} \tag{3}$$

(x standard normal), in the sense that this relationship holds asymptotically between the corresponding percentiles of a and x. He then applies the fiducial argument, solving (3) for an asymptotic expansion of p in terms of a (the observed variate) and x (the pivotal variate). The general term in this expansion may be obtained by a formal application of Lagrange's inversion formula:

$$\begin{aligned} p - a/n &\sim \phi_{x,n}(a/n) + \phi'_{x,n}(a/n)(p-a/n) + \ldots \\ &\sim \phi_{x,n}(a/n) + \phi'_{x,n}(a/n)[\phi_{x,n}(a/n) + \phi'_{x,n}(a/n)(p-a/n) + \ldots] + \ldots \\ &\sim \sum_{r=1}^{\infty} \frac{1}{r!}\left\{\left(\frac{\mathrm{d}}{\mathrm{d}p^*}\right)^{r-1} [\phi_{x,n}(p^*)]^r\right\}_{p^*=a/n}. \end{aligned} \tag{4}$$

The early terms are

$$p - a/n \sim -x\sqrt{(ab/n^3)} + \tfrac{1}{6}(b-a)(2x^2+1)/n^2 + \ldots, \tag{5}$$

in accordance with Fisher and Cornish (1960), equation (6). Fisher thus defines the fiducial distribution of p by equation (4), and obtains $\bar{p}_{\mathrm{F}}$ by taking the expectation of the R.H.S. with respect to x (a fixed):

$$\bar{p}_{\mathrm{F}} \sim \frac{a}{n} + \sum_{r=1}^{\infty} \frac{1}{r!}\left\{\left(\frac{\mathrm{d}}{\mathrm{d}p^*}\right)^{r-1} E_x[\phi_{x,n}(p^*)]^r\right\}_{p^*=a/n}. \tag{6}$$

But $\phi_{x,n}(p^*)$ specifies the Cornish–Fisher expansion of a binomial variate a^* with parameter p^*. It follows from the formal definition of this expansion that we have, identically in the powers of n,

$$E_x[\phi_{x,n}(p^*)]^r = E_{a^*}(p^* - a^*/n)^r = M_r/(-n)^r, \tag{7}$$

where M_r is the rth central moment of a^*. Hence the required expansion is

$$\bar{p}_{\mathrm{F}} \sim \frac{a}{n} + \sum_{r=1}^{\infty} \frac{(-1)^r}{r!n^r}\left\{\left(\frac{\mathrm{d}}{\mathrm{d}p^*}\right)^{r-1} M_r\right\}_{p^*=a/n}. \tag{8}$$

III. Reduction of the Expansion

The reduction of (8) to the form (2) may be effected by introducing Stirling numbers of the first and second kinds, defined respectively by

$$\left.\begin{aligned} \prod_{k=0}^{r} (1+kx) &= \sum_{k=0}^{r} S_1(r,k)x^k, \\ \prod_{k=0}^{r} (1-kx)^{-1} &= \sum_{k=0}^{\infty} S_2(r,k)x^k. \end{aligned}\right\} \tag{9}$$

In terms of the S_2, the rth non-central moment of the binomial distribution may be written as follows (deleting the asterisks on a and p for convenience):

$$M'_r = E(a^r) = \sum_{k=0}^{r} S_2(k, r-k)(n)_k p^k, \tag{10}$$

where

$$(n)_0 = 1, \qquad (n)_k = n(n-1)\dots(n-k+1), \qquad (k \geqslant 1), \tag{11}$$

(see Riordan 1958, p. 41). It follows that M_r is a polynomial in p of order r:

$$M_r = \sum_{k=0}^{r} p^k \sum_{s=0}^{k} \binom{r}{k-s}(-n)^{k-s}(n)_s S_2(s, r-k) \tag{12}$$

$$= \alpha_r p^r + \beta_r p^{r-1} +, \dots, \quad \text{say},$$

where $\alpha_1 = \beta_1 = 0$. Hence

$$\left(\frac{\mathrm{d}}{\mathrm{d}p}\right)^{r-1} M_r = r!\alpha_r p + (r-1)!\beta_r, \tag{13}$$

and equation (8) becomes

$$\tilde{p}_\mathrm{F} \sim \frac{a}{n} + \frac{a}{n}\sum_{r=2}^{\infty} \frac{(-1)^r \alpha_r}{n^r} + \sum_{r=2}^{\infty} \frac{(-1)^r \beta_r}{rn^r}. \tag{14}$$

We next show that

$$\alpha_r = -2\beta_r/r, \qquad (r = 2, 3, \dots), \tag{15}$$

and that α_r is a polynomial in n of order $r/2$ (i.e. the largest integer not exceeding $r/2$). The author is indebted to a referee for the following proof of these properties of α_r, which is simpler than that originally given. Introduce the generating functions

$$f(s) = 1 + \sum_{r=1}^{\infty} \frac{\alpha_r s^r}{r!}, \qquad g(s) = \sum_{r=1}^{\infty} \frac{\beta_r s^r}{r!}, \tag{16}$$

and the moment generating function of the binomial distribution

$$G(s) = \sum_{r=0}^{\infty} \frac{M_r s^r}{r!} = (q + p\,\mathrm{e}^s)^n\,\mathrm{e}^{-nps}. \tag{17}$$

It follows immediately from the definition of α_r, β_r that

$$G(s/p) = f(s) + \frac{1}{p} g(s) + 0\left(\frac{1}{p^2}\right), \tag{18}$$

whence,

$$f(s) = [(1+s)\,\mathrm{e}^{-s}]^n, \qquad g(s) = \tfrac{1}{2} n s^2 (1+s)^{n-1}\,\mathrm{e}^{-ns}. \tag{19}$$

Equation (15) follows since

$$\frac{\mathrm{d}f}{\mathrm{d}s} = -\frac{2}{s} g(s), \tag{20}$$

and α_r is easily seen from (19) to be a polynomial in n of the required order. Hence

$$\bar{p}_{\mathrm{F}} \sim \frac{a}{n}+\frac{(b-a)}{2}\sum_{r=2}^{\infty}\alpha_r\left(-\frac{1}{n}\right)^{r+1}, \tag{21}$$

while from (12)

$$\begin{aligned}\alpha_r &= \sum_{s=0}^{r}\tbinom{r}{s}(-n)^{r-s}(n)_s \\ &= (-n)^r \sum_{[k=\frac{1}{2}(r+1)]}^{r}\left(-\frac{1}{n}\right)^k \sum_{s=k+1}^{r}\tbinom{r}{s}(-1)^s S_1(s-1,k)\,.\end{aligned} \tag{22}$$

The formal expansion of $\bar{p}_{\mathrm{F}}$ is therefore

$$\bar{p}_{\mathrm{F}} \sim \frac{a}{n}+\frac{(b-a)}{2}\sum_{k=1}^{\infty}A_k\left(-\frac{1}{n}\right)^{k+1}, \tag{23}$$

where

$$A_k = \sum_{s=k+1}^{2k}\tbinom{2k+1}{s+1}(-1)^s S_1(s-1,k)\,. \tag{24}$$

It is therefore sufficient to prove that $A_k \equiv 1$. Using the following relation between the S_1 and S_2 due to Schläfli (e.g. Gould 1960, p. 447):

$$S_1(s-1,k) = \sum_{j=0}^{k}\tbinom{k+s}{k-j}\tbinom{k-s}{k+j}S_2(j,k) \tag{25}$$

it is found that

$$A_k = \sum_{j=0}^{k}(-1)^{j+1}S_2(j,k)\sum_{s=k+1}^{2k}\tbinom{2k+1}{s+1}\tbinom{k+s}{k-j}\tbinom{s+j-1}{k+j}(-1)^{s-k-1}\,. \tag{26}$$

However, the inner sum is equal to $\binom{2k+1}{1-j}$; or, equivalently,

$$\sum_{s=k+1}^{2k}\frac{(k+s)!}{(s+1)!(s-k-1)!(2k-s)!}\,\frac{(-1)^{s-k-1}}{(s+j)} = \frac{(k-j)!(k+j)!}{(1-j)!(2k+j)!}, \tag{27}$$

which holds since the LHS is the partial-fraction expansion of the RHS considered as a rational function of j. Hence, for $k \geqslant 1$,

$$A_k = -(2k+1)S_2(0,k)+S_2(1,k) = 1\,, \tag{28}$$

as required. A corollary is the identity relating binomial coefficients and Stirling numbers implied by (24) and (28). The author has been unable to find this identity in the literature, and would be interested to know if it is new.

IV. Acknowledgment

My thanks are due to Dr. E. A. Cornish, Division of Mathematical Statistics, CSIRO, for having directed my attention to this problem.

V. References

Fisher, R. A. (1959).—"Statistical Methods and Scientific Inference." 2nd Edn. (Oliver and Boyd: Edinburgh.)

Fisher, R. A., and Cornish, E. A. (1960).—The percentile points of distributions having known cumulants. *Technometrics* **2**, 209–25.

Gould, H. W. (1960).—Stirling number representation problems. *Proc. Am. math. Soc.* **11**, 447–51.

Riordan, J. (1958).—"An Introduction to Combinatorial Analysis." (Wiley: New York.)

Wallace, D. L. (1958).—Asymptotic approximations to distributions. *Ann. math. Statist.* **29**, 635–54.

Biometrika (1969), **56**, 2, *p.* 283
Printed in Great Britain

Cyclic change-over designs

BY A. W. DAVIS AND W. B. HALL
Division of Mathematical Statistics, C.S.I.R.O., Adelaide

SUMMARY

A class of cyclic change-over designs, existing for any number of treatments and periods, is defined as a simple extension of the cyclic incomplete block designs. The analysis is presented for direct and first-order residual effects of treatments but the method generalizes for residual effects of higher order. Cyclic change-over designs may be analysed after any number of periods and extra periods may be added. A table of selected designs is given for 6 to 20 treatments in 3, 4 or 5 periods. The efficiencies of these designs compare favourably with existing designs and in general the cyclic designs require fewer units.

1. INTRODUCTION

Cyclic incomplete block (CIB) designs have recently been shown by several authors (Kempthorne, 1953; Das, 1960; David & Wolock, 1965; John, 1966) to have a number of desirable features. These include existence for any number of treatments and block size, conciseness of representation by initial blocks, ease of analysis and two-way elimination of variation. By suitable choice of the initial blocks, balanced and nearly balanced designs may be obtained which have high average efficiencies.

The purpose of this paper is to show that these features carry over when CIB designs are interpreted as change-over designs, the blocks now being considered as treatment sequences applied to the experimental units and the rows constituting periods. We thus obtain the class of cyclic change-over (CCO) designs. These are constructed by the cyclic development of one or more generating sequences of treatments, corresponding to the initial blocks of the CIB designs, and a general method of analysis is developed for this class of designs. An attractive feature of CCO designs is that they may be analysed after any number of periods, and further periods may be added if required, thus enabling change-over trials to be conducted sequentially. Only first-order residual effects will be considered in this paper, although the methods used generalize directly to higher order effects. Balanced (BCO) and partially balanced change-over (PBCO) designs have been developed in detail by Patterson & Lucas (1962). In Table 4 of the present paper selected CCO designs are presented requiring fewer units but nevertheless possessing average efficiencies which compare favourably with those of Patterson & Lucas.

2. PRELIMINARIES

CIB designs in t treatments (labelled $0, 1, \ldots, t-1$) and block size p are obtained by developing some number, b, of initial blocks mod t. Each such design is a partially balanced incomplete block design with up to $[\frac{1}{2}t]$ associate classes, but the analysis is straightforward since the reduced normal matrix for the estimation of treatment effects is circulant. Hence there may exist up to $[\frac{1}{2}t]$ distinct efficiencies for the differences between treatment effects, although John (1966) has shown that the range of efficiencies may be quite small. Fractional

CIB designs may be constructed when p has a common factor with t (David & Wolock, 1965), but the latter are excluded from the present discussion. Introducing the residual effects model, we thus obtain CCO designs in t treatments, p periods and $n = bt$ units arranged in b blocks, obtained by cyclic development of the b generating sequences.

Example. For $t = 6$, $p = 3$, $b = 2$ and generating sequences (034), (051), we obtain the following CCO design, based on a three-associate class CIB design:

	Block 1						Block 2					
Period 1	0	1	2	3	4	5	0	1	2	3	4	5
Period 2	3_0	4_1	5_2	0_3	1_4	2_5	5_0	0_1	1_2	2_3	3_4	4_5
Period 3	4_3	5_4	0_5	1_0	2_1	3_2	1_5	2_0	3_1	4_2	5_3	0_4

(1)

The columns of array (1) represent the treatment sequences applied to the 12 units. For example, in period 2, the first unit in block 1 receives treatment 3, the subscript 0 denoting that there is a residual effect from the treatment 0 applied in period 1. The subscripts denoting residual effects are arranged in the cyclic design obtained by deleting the final period of the treatments design. Furthermore, certain cyclic associations exist between the symbols for treatment and residual effects. Thus, i occurs in the same column but one period later than $1+i$, $3+i$, $4+i$ and $5+i$ (mod t; $i = 0, 1, \ldots, t-1$), while in the columns containing i the treatment symbols occurring in the first two periods are i, $i+1$, $i+2$, $i+3$, $i+4$ and $i+5$ (mod t; $i = 0, 1, \ldots, t-1$) with multiplicities 4, 1, 1, 2, 1 and 3 respectively. In virtue of these relationships, the matrix of the reduced normal equations for direct and residual effects may in general be partitioned into circulant matrices, this partitioning being the basis of the analysis. The development of a computer program for CCO designs is further facilitated by the fact that the elements of these circulants and the various totals required in the analysis may be specified directly in terms of the generating sequences; see §4.

The following average efficiencies for the CCO design (1) are given in Table 4 (design number 1):

$E_t = 79\%$ for treatments ignoring residual effects,

$E_d = 58\%$ for direct effects of treatments,

$E_r = 34\%$ for residual effects of treatments,

$E_p = 15\%$ for permanent effects of treatments.

These efficiencies are defined in §5, and are consistent with those of Patterson & Lucas (1962, §2·2). The above values compare with those of the PBCO design requiring twice the number of units given by Patterson & Lucas (1962, Table 5·0, design 100) for which $E_t = 77\%$, $E_d = 62\%$, $E_r = 34\%$ and $E_p = 15\%$. The distinct efficiencies for differences between the various effects in the two designs are (see Table 1 and §5):

	CCO design	PBCO design
Treatments (ignoring residual effects)	(77·0, 77·4, 83)	(67, 80)
Direct effects	(56, 59, 63)	(53, 64)
Residual effects	(31, 33, 45)	(30, 36)

3. Model and reduced normal equations

The model postulated for CCO designs is

$$\mathbf{y} = \mu\mathbf{E}_{np,1} + \mathbf{D}\boldsymbol{\delta} + \mathbf{R}\boldsymbol{\rho} + \mathbf{U}\boldsymbol{\upsilon} + \mathbf{P}\boldsymbol{\pi} + \boldsymbol{\epsilon}, \tag{2}$$

where $\mathbf{y}$ is the column vector of np observations; μ is the mean value; $\mathbf{E}_{k,q}$ denotes the

$k \times q$ matrix with all elements unity; $\boldsymbol{\delta}$, $\boldsymbol{\rho}$, $\boldsymbol{\upsilon}$ and $\boldsymbol{\pi}$ are the vectors of direct, residual, unit and period effects, respectively; $\mathbf{D}, \mathbf{R}, \mathbf{U}$ and $\mathbf{P}$ are the corresponding incidence matrices (i.e. $\mathbf{D} = \{D_{hi}\}$, where $D_{hi} = 1$ if treatment i is applied at the hth observation and is zero otherwise, etc.); $\boldsymbol{\epsilon}$ is a vector of uncorrelated normal variables each having zero mean and variance σ^2. It is assumed that

$$\mathbf{E}_{1,t}\boldsymbol{\delta} = \mathbf{E}_{1,t}\boldsymbol{\rho} = \mathbf{E}_{1,n}\boldsymbol{\upsilon} = \mathbf{E}_{1,p}\boldsymbol{\pi} = 0. \tag{3}$$

Now let $\mathbf{I}_q$ denote the $q \times q$ identity matrix and $\mathbf{0}_q$ the q-dimensional zero vector. Write

$$\mathbf{L} = \mathbf{D}'\mathbf{R}, \quad \mathbf{M} = \mathbf{R}'\mathbf{U}, \quad \mathbf{N} = \mathbf{D}'\mathbf{U}; \tag{4}$$

i.e. $\mathbf{N} = \{N_{il}\}$, where N_{il} is the number of times the ith treatment is applied to the lth unit; $\mathbf{M} = \{M_{il}\}$, where M_{il} is the number of times the ith treatment is applied to the lth unit in the first $(p-1)$ periods; $\mathbf{L} = \{L_{ij}\}$, where L_{ij} is the number of times the ith treatment is immediately preceded by the jth treatment. Then from the identities

$$\left.\begin{array}{lll} \mathbf{D}'\mathbf{D} = bp\mathbf{I}_t, & \mathbf{R}'\mathbf{R} = b(p-1)\,\mathbf{I}_t, & \mathbf{U}'\mathbf{U} = p\mathbf{I}_n, \\ \mathbf{P}'\mathbf{P} = n\mathbf{I}_p, & \mathbf{D}'\mathbf{P} = b\mathbf{E}_{t,p}, & \mathbf{U}'\mathbf{P} = \mathbf{E}_{n,p}, \end{array}\right\} \tag{5}$$

it may be shown that μ is estimated by the grand mean and that the reduced normal equations are

$$\mathbf{C}\begin{bmatrix}\hat{\boldsymbol{\delta}}\\ \hat{\boldsymbol{\rho}}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Theta} & \boldsymbol{\Phi}\\ \boldsymbol{\Phi}' & \boldsymbol{\Psi}\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{\delta}}\\ \hat{\boldsymbol{\rho}}\end{bmatrix} = \begin{bmatrix}\mathbf{Q}\\ \mathbf{S}\end{bmatrix}, \tag{6}$$

where

$$\left.\begin{array}{l} \boldsymbol{\Theta} = bp\mathbf{I}_t - p^{-1}\mathbf{N}\mathbf{N}', \quad \boldsymbol{\Phi} = \mathbf{L} - p^{-1}\mathbf{N}\mathbf{M}', \\ \boldsymbol{\Psi} = b(p-1)\,\mathbf{I}_t - p^{-1}\mathbf{M}\mathbf{M}', \quad \mathbf{Q} = (\mathbf{D}' - p^{-1}\mathbf{N}\mathbf{U}')\,\mathbf{y}, \\ \mathbf{S} = [\mathbf{R}' - p^{-1}\mathbf{M}\mathbf{U}' - t^{-1}\{b^{-1}\mathbf{R}'\mathbf{P}\mathbf{P}' - (1-p^{-1})\,\mathbf{E}_{t,np}\}]\,\mathbf{y}, \\ \mathbf{E}_{1,t}\mathbf{Q} = \mathbf{E}_{1,t}\mathbf{S} = 0. \end{array}\right\} \tag{7}$$

Since also

$$\mathbf{R}'\mathbf{P} = b[\mathbf{0}_t, \mathbf{E}_{t,p-1}], \tag{8}$$

the adjusted treatment and residual effect totals Q_i and S_i reduce to special cases of the quantities given by Patterson & Lucas (1962, § 1·3):

$$Q_i = T_i - p^{-1}U_i, \quad S_i = R_i - p^{-1}U_i^{(1)} + t^{-1}(P_1 - p^{-1}G), \tag{9}$$

where $i = 0, 1, \ldots, t-1$ and T_i = total for treatment i, R_i = total for observations involving the residual effect of treatment i, $U_i^{(1)}$ = total for all units receiving treatment i in the first $(p-1)$ periods, U_i = total for all units receiving treatment i, P_1 = total for first period and G = grand total.

4. Construction of C

It is seen that $\boldsymbol{\Theta}$ is the reduced normal matrix for the CIB design obtained by ignoring residual effects. Clearly $\boldsymbol{\Theta}$, $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are circulant matrices (§ 2) and may be specified by their initial rows:

$$\boldsymbol{\Theta} = [\theta_0, \ldots, \theta_{t-1}], \quad \boldsymbol{\Phi} = [\phi_0, \ldots, \phi_{t-1}], \quad \boldsymbol{\Psi} = [\psi_0, \ldots, \psi_{t-1}]. \tag{10}$$

Thus, in $\boldsymbol{\Theta}$, the element in the ith row and jth column $(i, j = 0, 1, \ldots, t-1)$ is θ_{j-i}, where $j-i$ is to be reduced mod t.

The elements θ_j, ϕ_j and ψ_j are readily obtained from the generating sequences of the design.

If $(0, a_1, \ldots, a_{p-1})$ is a generating sequence, with all a_i distinct, then the p sequences obtained by development mod t and containing treatment 0 are

$$\left.\begin{array}{llll} 0 & -a_1 & \ldots & -a_{p-1} \\ a_1 & 0 & \ldots & a_1-a_{p-1} \\ a_2 & a_2-a_1 & \ldots & a_2-a_{p-1} \\ \vdots & \vdots & & \vdots \\ a_{p-2} & a_{p-2}-a_1 & \ldots & a_{p-2}-a_{p-1} \\ a_{p-1} & a_{p-1}-a_1 & \ldots & 0 \end{array}\right\}. \tag{11}$$

This is the array of p^2 differences (mod t) between the elements of the generating sequence. The element in row 0 and column j of the matrix $\mathbf{NN}'$ is therefore

$(\mathbf{NN}')_{0j}$ = number of units receiving both treatments 0 and j;

= number of times j occurs in array (11), $(j = 0, 1, \ldots, t-1)$.

Similarly

$(\mathbf{NM}')_{0j}$ = number of times j occurs in array (11) with the final row omitted;

$(\mathbf{MM}')_{0j}$ = number of times j occurs in array (11) with the final row and column omitted;

L_{0j} = number of units in which treatment 0 is immediately preceded by treatment j

= number of times j occurs among the superdiagonal elements $a_{m-1}-a_m$ $(m = 1, \ldots, p-1;\ a_0 = 0)$ of array (11).

When $b > 1$, the elements of $\mathbf{NN}'$, etc., are obtained by summing the respective quantities over the different generating sequences. For the example given in §2, the first rows are:

$$\mathbf{NN}' = [6, 3, 2, 2, 2, 3], \quad \mathbf{NM}' = [4, 1, 1, 2, 1, 3],$$
$$\mathbf{MM}' = [4, 1, 0, 2, 0, 1], \quad \mathbf{L} = [0, 1, 0, 1, 1, 1],$$
$$3\mathbf{\Theta} = [12, -3, -2, -2, -2, -3], \quad 3\mathbf{\Phi} = [-4, 2, -1, 1, 2, 0],$$
$$3\mathbf{\Psi} = [8, -1, 0, -2, 0, -1].$$

Because in the array (11) each difference j and its complement $t-j$ (mod t) occur equally many times, it is easily seen that

$$(\mathbf{NN}')_{0j} = (\mathbf{NN}')_{0,t-j}, \quad (\mathbf{MM}')_{0j} = (\mathbf{MM}')_{0,t-j} \quad (j = 1, \ldots, t-1), \tag{12}$$

whence
$$\theta_j = \theta_{t-j}, \quad \psi_j = \psi_{t-j} \quad (j = 1, \ldots, t-1). \tag{13}$$

The quantities appearing in the equations (9) for Q_i and S_i may also be given explicitly in terms of the generating sequences. Let $y(k, l)$ denote the observation in the kth period for the lth unit and $y(.,l)$ the total for the lth unit $(k = 0, \ldots, p-1;\ l = 0, \ldots, t-1 \pmod t)$. Then for the generating sequence in (11),

$$\left.\begin{array}{ll} T_i = \sum_{l=0}^{p-1} y(l, i-a_l), & R_i = \sum_{l=0}^{p-2} y(l+1, i-a_l), \\ U_i = \sum_{l=0}^{p-1} y(.,i-a_l), & U_i^{(1)} = \sum_{l=0}^{p-2} y(.,i-a_l). \end{array}\right\} \tag{14}$$

In general, these quantities are added over all generating sequences. The methods of this section remain valid when treatments are applied more than once to a unit.

5. Solution of the normal equations

Let $\mathbf{\Gamma}_j$ denote the circulant matrix $[0, 0, \ldots, 0, 1, 0, \ldots, 0]$, where 1 occurs in the jth position ($j = 0, 1, \ldots, t-1$). Then the $\mathbf{C}$ matrix may be expressed as a sum of Kronecker products,

$$\mathbf{C} = \sum_{j=0}^{t-1} \begin{bmatrix} \theta_j & \phi_j \\ \phi_{t-j} & \psi_j \end{bmatrix} \otimes \mathbf{\Gamma}_j. \tag{15}$$

In order to invert the normal equations, it is convenient to refer the cyclic group of matrices $\mathbf{\Gamma}_j$ to a basis of mutually orthogonal idempotents $\mathbf{\Gamma}^*_u$ defined by

$$\mathbf{\Gamma}^*_u = t^{-1} \sum_{j=0}^{t-1} \omega_u^{-j} \mathbf{\Gamma}_j, \tag{16}$$

where the ω_u are the tth roots of unity, that is

$$\omega_u = \exp(2\pi i u/t) \quad (u = 0, 1, \ldots, t-1). \tag{17}$$

For this, see, e.g. Hannan (1960, § 1·2). Then

$$\left.\begin{aligned} \mathbf{\Gamma}^*_u \mathbf{\Gamma}^*_u = \mathbf{\Gamma}^*_u, \quad \mathbf{\Gamma}^*_u \mathbf{\Gamma}^*_v = 0 \quad (u \neq v), \\ \overline{\mathbf{\Gamma}^*_u} = (\mathbf{\Gamma}^*_u)' = \mathbf{\Gamma}^*_{t-u}, \quad \mathbf{\Gamma}^*_0 = t^{-1}\mathbf{E}_{tt}, \end{aligned}\right\} \tag{18}$$

and

$$\sum_{u=0}^{t-1} \mathbf{\Gamma}^*_u = \mathbf{I}_t, \tag{19}$$

$$\mathbf{\Gamma}_j = \sum_{u=0}^{t-1} \omega_u^j \mathbf{\Gamma}^*_u. \tag{20}$$

With (20), $\mathbf{C}$ may now be expressed in terms of the $\mathbf{\Gamma}^*_u$. In virtue of (3), the term in $\mathbf{\Gamma}^*_0 = t^{-1}\mathbf{E}_{tt}$ may be omitted, yielding

$$\mathbf{C}^* = \sum_{u=1}^{t-1} \begin{bmatrix} \theta^*_u & \phi^*_u \\ \phi^*_{t-u} & \psi^*_u \end{bmatrix} \otimes \mathbf{\Gamma}^*_u, \tag{21}$$

where

$$\theta^*_u = \sum_{j=0}^{t-1} \omega_u^j \theta_j \quad (u = 1, \ldots, t-1), \tag{22}$$

and ϕ^*_u and ψ^*_u are defined similarly. From (13), θ^*_u and ψ^*_u are real numbers, and

$$\theta^*_u = \theta^*_{t-u}, \quad \psi^*_u = \psi^*_{t-u}, \quad \phi^*_u = \overline{\phi^*_{t-u}} \quad (u = 1, \ldots, t-1). \tag{23}$$

The solution of the normal equations has thus been reduced to the inversion of the 2×2 matrices in (21). Provided that the latter are non-singular (see § 6) we may add $\mathbf{I}_2 \otimes \mathbf{\Gamma}^*_0$ to $\mathbf{C}^*$ and, using (18) and (19), invert to obtain

$$\begin{bmatrix} \hat{\boldsymbol{\delta}} \\ \hat{\boldsymbol{\rho}} \end{bmatrix} = \left\{ \sum_{u=1}^{t-1} \begin{bmatrix} \theta^*_u & \phi^*_u \\ \phi^*_{t-u} & \psi^*_u \end{bmatrix}^{-1} \otimes \mathbf{\Gamma}^*_u \right\} \begin{bmatrix} \mathbf{Q} \\ \mathbf{S} \end{bmatrix} = \mathbf{V} \begin{bmatrix} \mathbf{Q} \\ \mathbf{S} \end{bmatrix}, \quad \text{say}, \tag{24}$$

again deleting the term in $\mathbf{\Gamma}^*_0$ since $\mathbf{E}_{1t}\mathbf{Q} = \mathbf{E}_{1t}\mathbf{S} = 0$. It is readily shown that

$$\operatorname{cov} \begin{bmatrix} \mathbf{Q} \\ \mathbf{S} \end{bmatrix} = \sigma^2 \mathbf{C}^*, \tag{25}$$

whence the covariance matrix of the estimated direct and residual effects is also given in terms of the matrix $\mathbf{V}$, that is

$$\operatorname{cov} \begin{bmatrix} \hat{\boldsymbol{\delta}} \\ \hat{\boldsymbol{\rho}} \end{bmatrix} = \sigma^2 \mathbf{V}. \tag{26}$$

In terms of the original Γ_j, $$\mathbf{V} = \sum_{j=0}^{t-1} \begin{bmatrix} \xi_j & \eta_j \\ \eta_{t-j} & \zeta_j \end{bmatrix} \otimes \Gamma_j, \tag{27}$$

where
$$\left.\begin{aligned} \xi_j &= \operatorname{cov}(\hat{\delta}_0, \hat{\delta}_j)/\sigma^2 = t^{-1} \sum_{u=1}^{t-1} \omega_u^{-j} \psi_u^*/\kappa_u, \\ \eta_j &= \operatorname{cov}(\hat{\delta}_0, \hat{\rho}_j)/\sigma^2 = -t^{-1} \sum_{u=1}^{t-1} \omega_u^{-j} \phi_u^*/\kappa_u, \\ \zeta_j &= \operatorname{cov}(\hat{\rho}_0, \hat{\rho}_j)/\sigma^2 = t^{-1} \sum_{u=1}^{t-1} \omega_u^{-j} \theta_u^*/\kappa_u, \\ \kappa_u &= \theta_u^* \psi_u^* - |\phi_u^*|^2. \end{aligned}\right\} \tag{28}$$

In obtaining the estimated effects $\hat{\delta}_i$ and $\hat{\rho}_i$ from (24), it is sufficient to use

$$S_i^* = R_i - p^{-1} U_i^{(1)}, \quad \sum_{i=0}^{t-1} S_i^* = p^{-1} G - P_1, \tag{29}$$

deleting the constant term $t^{-1}(P_1 - p^{-1}G)$ from S_i since the η_j and ζ_j sum to zero. Hence

$$\left.\begin{aligned} \hat{\delta}_i &= \sum_{l=0}^{t-1} (\xi_l Q_{i+l} + \eta_l S_{i+l}^*), \\ \hat{\rho}_i &= \sum_{l=0}^{t-1} (\eta_{l-l} Q_{i+l} + \zeta_l S_{i+l}^*) \quad (i = 0, 1, \ldots, t-1), \end{aligned}\right\} \tag{30}$$

where $i + l$ is to be reduced modulo t. An example of the analysis of a CCO design is given in § 7.

The efficiency E_d for direct treatment effects is defined to be $2\sigma^2/(bp)$ divided by the average of the variances

$$\operatorname{var}(\hat{\delta}_j - \hat{\delta}_{j'}) = 2(\xi_0 - \xi_{j-j'})\sigma^2, \tag{31}$$

where $j - j'$ is to be reduced modulo t. It follows that

$$E_d = (t-1)/(np\xi_0). \tag{32}$$

Similarly $$\operatorname{var}(\hat{\rho}_j - \hat{\rho}_{j'}) = 2(\zeta_0 - \zeta_{j-j'})\sigma^2, \tag{33}$$

and E_r, the average efficiency for residual effects, is given by

$$E_r = (t-1)/(np\zeta_0). \tag{34}$$

For permanent effects of treatments $(\hat{\delta}_j + \hat{\rho}_j;\, j = 0, 1, \ldots, t-1)$, we have

$$\left.\begin{aligned} \operatorname{var}\{(\hat{\delta}_j + \hat{\rho}_j) - (\hat{\delta}_{j'} + \hat{\rho}_{j'})\} &= 2\{(\xi_0 - \xi_{j-j'}) + (\zeta_0 - \zeta_{j-j'}) + (2\eta_0 - \eta_{j-j'} - \eta_{j'-j})\}, \\ E_p &= (t-1)/\{np(\xi_0 + 2\eta_0 + \zeta_0)\}. \end{aligned}\right\} \tag{35}$$

The techniques of §§ 4 and 5 may be generalized for p'th order residuals $(p' \leqslant p-1)$, the solution of the normal equations requiring the inversion of $(p'+1)$th order matrices.

6. Discussion

In the construction of Table 4, CIB designs having high efficiencies E_t for $t \leqslant 15$ treatments were selected from Table 1 of John (1966). For $t > 15$ the choice of designs was facilitated by requiring that $|\theta_i - \theta_j| \leqslant 1$ (or in certain cases, $\leqslant 2$) for the off-diagonal elements of $\mathbf{\Theta}$ (John, 1966, p. 349). The treatments in the generating sequences were then permuted to obtain CCO designs with good efficiencies E_d for direct effects. Here the range of possibilities was narrowed by restricting the differences between the off-diagonal elements of $\mathbf{\Phi}$ to be as small as possible. It was found that designs with high E_d also had high efficiencies E_r, and that there were many good designs differing only marginally in their efficiency factors. By restricting the designs considered as above, no singular matrices were encountered in the inversion of (21); however, such cases can arise (e.g. $t = 12$, generating sequences (0 1 *11*) (0 5 7)). Table 4 gives selected CCO designs, together with their average efficiencies and ranges of efficiencies for individual comparisons for $t = 6(1)\,20$: $p = 3$, $b = 2$; $p = 4$, $b = 2$; $p = 5$, $b = 1$. The BCO design number 17 was given by Patterson (1951); see also Patterson & Lucas (1962, Table 5·1, design number 19). A program (Fortran for CDC 3200 computer) has been written which analyses any CCO design and enables suitable designs to be chosen for any given set of parameter values. Copies may be obtained from the authors on request.

CCO designs may be analysed after any number of periods, and further periods may be added in such a way as to maximize E_d or E_r, these features adding great flexibility to the experimental situation. As in the designs of Patterson & Lucas, repetition of the final period increases E_r at the expense of E_d. However, orthogonality of direct and residual effects is not achieved by this means as in their designs.

In the example discussed in §2, repetition of the final period leads to efficiencies $E_t = 74\,\%$, $E_d = 72\,\%$, $E_r = 61\,\%$ and $E_p = 33\,\%$. If the generating sequences are increased to (0 3 4 2), (0 5 1 2), we obtain $E_t = 90\,\%$, $E_d = 80\,\%$, $E_r = 57\,\%$ and $E_p = 26\,\%$.

In using the designs of Table 4, the randomization to be followed is: (*a*) allocate blocks of the design to groups of experimental units at random; (*b*) allocate units to treatment sequences of the design at random independently for each block; (*c*) allocate treatment numbers to treatments at random. It should be noted that randomization of rows is not permissible.

In order to carry out overall F tests for direct and residual effects separately, sums of squares are required (i) for direct effects ignoring residual effects and (ii) for residual effects ignoring direct effects (Patterson & Lucas, 1962, §1·3). This involves the analysis of (i) the corresponding CIB design and (ii) the CIB design obtained by deleting the final period, respectively. The covariance matrix of estimated treatment effects for the CIB design (i) is circulant with elements

$$\alpha_j = t^{-1} \sum_{u=1}^{t-1} \omega_u^{-j} (\theta_u^*)^{-1} \quad (j = 0, 1, \ldots, t-1), \tag{36}$$

while if we write

$$b(p-1)\,\mathbf{I}_t - (p-1)^{-1}\mathbf{MM}' = [\lambda_0, \lambda_1, \ldots, \lambda_{t-1}], \quad \lambda_u^* = \sum_{j=0}^{t-1} \omega_u^j \lambda_j, \tag{37}$$

then the elements of the covariance matrix for design (ii) are

$$\gamma_j = t^{-1} \sum_{u=1}^{t-1} \omega_u^{-j} (\lambda_u^*)^{-1} \quad (j = 0, 1, \ldots, t-1). \tag{38}$$

In terms of these quantities, the direct effects $\hat{\tau}_i$ of treatments ignoring residual effects, and the residual effects $\hat{\nu}_i$ of treatments ignoring direct effects are given by

$$\hat{\tau}_i = \sum_{l=0}^{t-1} \alpha_l Q_{l+i}, \quad \hat{\nu}_i = \sum_{l=0}^{t-1} \gamma_l S_{l+i}^*. \tag{39}$$

Then

$$\operatorname{var}(\hat{\tau}_j - \hat{\tau}_{j'}) = 2(\alpha_0 - \alpha_{j-j'}), \tag{40}$$

$$E_t = (t-1)/(np\alpha_0). \tag{41}$$

7. Example

Table 1 presents the elements of the variance–covariance matrices required for the analysis of the example given in §2.

Table 1. *Variances and covariances for* CCO *design number* 1

i	ξ_i	η_i	ζ_i	α_i	γ_i
0	0·23827	0·13429	0·41397	0·17484	0·39722
1	−0·04462	−0·07889	−0·09721	−0·02612	−0·08611
2	−0·06144	0·02633	−0·13006	−0·04054	−0·13611
3	−0·02615	0·00208	0·04057	−0·04151	0·04722
4	−0·06144	−0·10293	−0·13006	−0·04054	−0·13611
5	−0·04462	0·01912	−0·09721	−0·02612	−0·08611

Table 2. *Example of the analysis*

(*a*) *Design and observations*

	Block 1							Block 2							Period
Units	1	2	2	4	5	6	Total	1	2	3	4	5	6	Total	total
Treatment	0	1	2	3	4	5		0	1	2	3	4	5		
Period 1	38·7	48·9	35·2	34·6	32·9	30·4	220·7	25·7	30·8	25·4	21·8	21·4	22·8	147·9	368·6
Treatment	3	4	5	0	1	2		5	0	1	2	3	4		
Period 2	37·4	46·9	33·5	32·3	33·1	29·5	212·7	26·1	29·3	26·0	23·9	22·0	21·0	148·3	361·0
Treatment	4	5	0	1	2	3		1	2	3	4	5	0		
Period 3	34·3	42·0	28·4	28·5	27·5	26·7	187·4	23·4	26·4	23·9	21·7	19·4	18·6	133·4	320·8
Total	110·4	137·8	97·1	95·4	93·5	86·6	620·8	75·2	86·5	75·3	67·4	62·8	62·4	429·6	1050·4

(*b*) *Derived totals and estimates*

Treatment	T_i	R_i	U_i	$U_i^{(1)}$	$3Q_i$	$3S_i^*$	$\hat{\delta}_i$	$\hat{\rho}_i$	$\hat{\tau}_i$	$\hat{\nu}_i$
0	173·0	118·4	527·0	367·5	−8·0	−12·3	−0·810	−0·684	−0·531	−0·470
1	190·7	127·6	563·7	393·1	8·4	−10·3	0·746	0·366	0·552	−0·142
2	167·9	107·9	506·4	326·4	−2·7	−2·7	0·102	0·498	−0·147	0·877
3	166·4	109·9	497·9	336·0	1·3	−6·3	0·189	0·602	0·084	0·440
4	178·2	115·7	534·3	356·5	0·3	−9·4	0·035	0·125	0·028	−0·132
5	174·2	102·3	521·9	321·3	0·7	−14·4	−0·262	−0·907	0·014	−0·573
Total	1050·4	681·8	3151·2	2100·8	0·0	−55·4	0·000	0·000	0·000	0·000

To illustrate the analysis, Table 2(*a*) gives uniformity data on the milk production of dairy cows (average daily productions of fat-corrected milk over periods of five weeks) taken from Patterson & Lucas (1962, Table 3·1) with CCO design number 1 superimposed. Table 2(*b*) contains the totals and derived quantities defined in equations (9) and (29) and also the estimated effects defined in equations (30) and (39). For example,

$$\begin{aligned}3\hat{\rho}_1 = {} & \{8{\cdot}4(0{\cdot}13429) - 2{\cdot}7(0{\cdot}01912) + 1{\cdot}3(-0{\cdot}10293) \\ & + 0{\cdot}3(0{\cdot}00208) + 0{\cdot}7(0{\cdot}02633) - 8{\cdot}0(-0{\cdot}07889)\} \\ & + \{-10{\cdot}3(0{\cdot}41397) - 2{\cdot}7(-0{\cdot}09721) - 6{\cdot}3(-0{\cdot}13006) \\ & - 9{\cdot}4(0{\cdot}04057) - 14{\cdot}4(-0{\cdot}13006) - 12{\cdot}3(-0{\cdot}09721)\} \\ = {} & 1{\cdot}098.\end{aligned}$$

Table 3. *Analysis of variance of cyclic change-over design*

Source of variation	Degrees of freedom		Sums of squares	
Blocks	$b-1$	1	†	1015·484
Periods	$p-1$	2	†	714·884
Blocks × periods	$(b-1)(p-1)$	2	†	109·962
Units within blocks	$b(t-1)$	10	†	14·782
Direct and residual effects of treatments	$2(t-1)$	10	$\sum_{i=0}^{t-1}(\hat{\delta}_i Q_i + \hat{\rho}_i S_i^*)$	7·981
Error	$(t-1)\{b(p-1)-2\}$	10	By difference	6·629
Total	$np-1$	35	†	1869·722

† Calculated in the usual way.

Partition of direct and residual effects of treatments

(1)	Degrees of freedom		Sums of squares	
Direct effects (eliminating residual effects)	$t-1$	5	By difference	4·116
Residual effects (ignoring direct effects)	$t-1$	5	$\sum_{i=0}^{t-1}\hat{\nu}_i S_i^*$	3·865
Total	$2(t-1)$	10	$\sum_{i=0}^{t-1}(\hat{\delta}_i Q_i + \hat{\rho}_i S_i^*)$	7·981
(2)				
Direct effects (ignoring residual effects)	$t-1$	5	$\sum_{i=0}^{t-1}\hat{\tau}_i Q_i$	3·136
Residual effects (eliminating direct effects)	$t-1$	5	By difference	4·845
Total	$2(t-1)$	10	$\sum_{i=0}^{t-1}(\hat{\delta}_i Q_i + \hat{\rho}_i S_i^*)$	7·981

The analysis of variance is set out in Table 3 and follows that given by Patterson & Lucas (1962, Table 1·1).

The error variance is estimated by $\hat{\sigma}^2 = 0{\cdot}663$, whence estimated standard errors of comparisons between the various effects may be obtained from (31), (33), (35), (40) and Table 1. For example,

$$\text{SE}\,(\hat{\hat{\delta}}_1 - \hat{\hat{\delta}}_2) = \{2(0{\cdot}23827 + 0{\cdot}04462)\,(0{\cdot}663)\}^{\frac{1}{2}} = 0{\cdot}612;$$

$$\text{SE}\,(\hat{\rho}_2 - \hat{\rho}_5) = \{2(0{\cdot}41397 - 0{\cdot}04057)\,(0{\cdot}663)\}^{\frac{1}{2}} = 0{\cdot}704.$$

We wish to thank Miss J. A. Hawkes, of the Division of Mathematical Statistics, C.S.I.R.O., for valuable assistance with programming, and the referee for his comments.

Table 4. *Cyclic change-over designs*

Design number	t	Generating sequences	E_t	E_d	E_r	E_p
		$p = 3, \quad b = 2$				
1	6	(0 3 4) (0 5 1)	79 (77, 83)	58 (56, 63)	34 (31, 45)	15
2	7	(0 3 1) (0 4 5)	78 (78, 78)	62 (62, 63)	35 (31, 43)	16
3	8	(0 4 1) (0 6 5)	76 (72, 78)	54 (52, 60)	31 (27, 38)	14
4	9	(0 3 8) (0 6 7)	74 (71, 77)	55 (51, 58)	31 (28, 40)	14
5	10	(0 1 3) (0 5 4)	73 (71, 77)	51 (45, 55)	30 (26, 39)	13
6	11	(0 1 7) (0 *10* 2)	73 (71, 77)	52 (48, 56)	30 (25, 38)	13
7	12	(0 1 7) (0 *11* 3)	71 (65, 76)	52 (49, 55)	30 (25, 37)	13
8	13	(0 1 4) (0 *11* 6)	72 (72, 72)	45 (40, 49)	28 (26, 32)	12
9	14	(0 1 9) (0 *13* 3)	70 (65, 75)	50 (46, 55)	30 (27, 37)	13
10	15	(0 8 2) (0 *14* 4)	71 (66, 72)	44 (40, 49)	28 (25, 31)	12
11	16	(0 6 9) (0 4 *15*)	70 (66, 72)	45 (52, 50)	28 (25, 30)	12
12	17	(0 1 5) (0 *15* 9)	70 (66, 71)	45 (42, 48)	28 (25, 30)	12
13	18	(0 1 5) (0 *12* 2)	69 (65, 71)	44 (38, 45)	27 (24, 31)	12
14	19	(0 4 *18*) (0 *17* 7)	69 (65, 71)	45 (42, 48)	28 (25, 31)	12
15	20	(0 2 *13*) (0 *19* 4)	68 (64, 71)	44 (40, 47)	27 (25, 31)	12
		$p = 4, \quad b = 2$				
16	6	(0 1 3 2) (0 3 1 4)	90 (87, 94)	81 (79, 83)	57 (56, 60)	26
17	7	(0 1 3 6) (0 6 4 1)	88 (88, 88)	80 (80, 80)	57 (57, 57)	26
18	8	(0 2 1 4) (0 1 5 3)	85 (84, 88)	77 (76, 79)	55 (52, 58)	25
19	9	(0 1 4 2) (0 5 2 6)	84 (84, 84)	74 (73, 76)	54 (53, 54)	24
20	10	(0 4 2 1) (0 5 7 4)	83 (81, 84)	73 (71, 75)	51 (47, 57)	23
21	11	(0 5 1 2) (0 6 5 2)	82 (81, 84)	73 (70, 75)	50 (45, 57)	23
22	12	(0 *11* 1 4) (0 *10* 3 4)	82 (81, 84)	71 (68, 73)	51 (49, 53)	23
23	13	(0 1 3 9) (0 8 6 9)	81 (81, 81)	69 (67, 71)	50 (49, 52)	23
24	14	(0 7 1 5) (0 6 4 3)	81 (78, 81)	69 (67, 71)	50 (48, 52)	23
25	15	(0 7 1 5) (0 5 *11* *12*)	80 (78, 81)	69 (66, 71)	51 (50, 52)	23
26	16	(0 5 7 1) (0 *14* 6 *13*)	79 (77, 81)	68 (66, 69)	49 (47, 51)	22
27	17	(0 1 7 5) (0 3 2 9)	79 (78, 81)	67 (65, 69)	49 (48, 51)	22
28	18	(0 1 8 5) (0 *11* 2 *12*)	79 (77, 81)	66 (62, 70)	48 (44, 51)	21
29	19	(0 7 6 9) (0 *12* *16* *11*)	79 (78, 81)	65 (63, 67)	47 (44, 50)	21
30	20	(0 1 *18* 6) (0 *19* 4 7)	78 (74, 80)	65 (61, 70)	47 (44, 53)	21
		$p = 5, \quad b = 1$				
31	6	(0 1 3 2 5)	96 (96, 96)	86 (82, 89)	66 (62, 67)	30
32	7	(0 2 3 1 5)	93 (92, 96)	80 (78, 82)	61 (58, 66)	28
33	8	(0 1 3 2 5)	91 (88, 92)	79 (76, 81)	59 (53, 66)	28
34	9	(0 1 3 2 5)	90 (87, 92)	76 (72, 81)	56 (49, 65)	27
35	10	(0 3 1 8 7)	89 (88, 92)	74 (71, 80)	57 (54, 65)	26
36	11	(0 4 7 1 2)	88 (88, 88)	67 (61, 72)	52 (47, 55)	24
37	12	(0 1 5 4 7)	87 (84, 88)	71 (70, 74)	55 (50, 57)	26
38	13	(0 2 3 7 4)	86 (83, 88)	71 (71, 72)	56 (55, 57)	26
39	14	(0 3 4 *11* 5)	86 (83, 88)	71 (69, 73)	55 (52, 60)	26
40	15	(0 1 5 4 7)	85 (83, 87)	68 (62, 71)	51 (45, 56)	25
41	16	(0 8 5 6 1)	85 (83, 87)	68 (64, 70)	53 (51, 57)	24
42	17	(0 2 9 3 4)	84 (83, 88)	66 (63, 67)	52 (50, 54)	24
43	18	(0 9 *12* 8 7)	84 (83, 88)	64 (61, 70)	49 (46, 56)	23
44	19	(0 7 3 8 9)	84 (84, 88)	62 (57, 66)	49 (46, 54)	22
45	20	(0 1 4 2 9)	83 (79, 87)	63 (57, 67)	47 (41, 54)	22

References

DAS, M. N. (1960). Circular designs. *J. Indian Soc. Agric. Statist.* **12**, 46–56.

DAVID, H. A. & WOLOCK, F. N. (1965). Cyclic designs. *Ann. Math. Statist.* **36**, 1526–34.

HANNAN, E. J. (1960). *Time Series Analysis*. London: Methuen.

JOHN, J. A. (1966). Cyclic incomplete block designs. *J. R. Statist. Soc.* B **28**, 345–60.

KEMPTHORNE, O. (1953). A class of experimental designs using blocks of two plots. *Ann. Math. Statist.* **24**, 76–84.

PATTERSON, H. D. (1951). Change-over trials. *J. R. Statist. Soc.* B **13**, 256–71.

PATTERSON, H. D. & LUCAS, H. L. (1962). *Change-over designs. N. Carolina Agric. Exp. Station Tech. Bull.*, no. 147.

[*Received July* 1968. *Revised November* 1968]