A STATISTICAL STUDY OF THE DISTRIBUTION OF ALCOHOL CONSUMPTION AND CONSEQUENT INFERENTIAL PROBLEMS

by

John B.F. Field
B.Sc. (Hons.) (Adel.)

Thesis submitted for the Degree of
Doctor of Philosophy
in the University of Adelaide
Department of Statistics

Adelaide

June 1985

CONTENTS

SUMMARY

The thesis consists of two parts. Part I examines the distribution of alcohol consumption (that is, the distribution of individual consumers of alcohol according to their consumption averaged over a suitable time period), in relation to Australian data, while Part II considers some more general inferential problems raised in Part I.

After a review of the literature concerning the distribution of alcohol consumption, Part I presents a detailed review of the controversial Ledermann model, providing a new interpretation of some of Ledermann's work. A substantial body of quantitative Australian data is collected together, and then other models, notably various lognormal distributions, are examined in the light of this data. It is found that the most commonly used model of the distribution of alcohol consumption, the two parameter lognormal, spuriously uses information about the light drinkers to make inferences about the heavy drinkers, because of the symmetry of the distribution on the logarithmic scale.

Part II examines this apparent paradox, and suggests some possible solutions. This is done using linear functionals of the class probabilities ("contrasts"). These linear functionals have considerable utility in precisely quantifying important inferential questions, and the mathematics necessary to use them is established. The approach is then to decompose a linear functional to show that a nonparametric estimator of a contrast is partitioned into the parametric estimator plus a second component whose expected value is zero if we can assume the validity of the specification. If we have some

doubt as to the validity of a particular aspect of the parametric specification, we may modify it and so transfer a further component to the parametric estimator, and be confident that the new reduced second component has zero expectation.

We show that, in the case of inferences concerned with the upper tail of the distribution of alcohol consumption, modifying the two parameter log-normal by the addition of a third parameter, or altering the fitting procedure by censoring the lower class frequencies, may ensure valid inferences.

Finally we present a method for fitting a probability distribution subject to a constraint on a linear function of the fitted class probabilities.