Mr. R. A. Fisher
Chief Statistician
Rothamsted Experiment Station
Rothamsted, England

Dear Sir:-

I wish to ask your opinion upon a matter which has been bothering me for sometime. The problem concerns the comparison of the means of two samples or the significance of a mean difference. For an example we may use the data in Table 27, page 106, "Statistical methods for Research Workers", second edition. You have treated this data in two ways in Exs. 19 and 20. In Ex. 20, it was assumed that the single observations of one series did not correspond to any particular observations of the other series. The thing I wish to propose is that in such examples the separate observations always correspond in certain respects and that the method of (Ex. 19) taking single differences may always be used.

The basis of correspondence, I conceive as follows. The differences between corresponding observations in Ex. 19, were taken not primarily because they were observations upon the same patient. The primary reason is because it is judged that a single patient will tend to vary less in other factors or causes of variance from one test to another than will two different patients. If the patients in the two tests were different persons as was assumed in Ex. 20, and it were possible to measure the causes of variance in each patient separately this would furnish a basis for judging which patients of the one series most truly corresponded with which patients of the other series.

In field tests differences are taken between adjacent plats. The primary reason however is not because the plats are adjacent, but because it is believed that adjacent plats are usually more alike with respect to some of the causes of variance, viz. soil fertility, insect damage, etc. If it were possible to determine which plats were most alike with respect to the totality of all causes of variance, these plats would be judged to most truly correspond and differences would be taken between them, rather than between adjacent plats.

The yield of plats are the results of many factors, some constant, some variable. Those which are constant do not affect the argument. Those which are variable may be positive or negative in effect upon yield. They are causes of variance. And the algebraic sum of their effects is expressed by the yield of a given plat. It is judged then that the yields of plats are rather
good measures of which plats, in two series, most truly correspond. If this can be granted it is only necessary to rank the plat yields in the two series separately according to magnitude. Those of corresponding rank are corresponding members of the two sets. Proceed as in Ex. 19. The argument given above, if valid, may be generally applied.

It may be added that in taking differences between adjacent plats or the effects of two drugs on the same patient, the basis of correspondence is only a part of the causes of variance. The only innovations then are the use of the total effect of all causes of variance rather than just part of the total as a basis of correspondence between single observations in the two series, and the measuring of the effect of all causes of variance by the effect itself. Presumably those observations which are most alike in size are also most alike in the totality of effect of all causes of variance.

Applying this method to the data of Table 27, (It is unimportant whether the same set of patients was used in both series or not so far as the method of analysis is concerned.)

<table>
<thead>
<tr>
<th>Rank</th>
<th>1(Dextro -)</th>
<th>2(Levo-)</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient number</td>
<td>Hours sleep gained</td>
<td>Patient number</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>3.7</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>3.4</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>2.0</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>0.6</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.7</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>0.0</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>-0.1</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>-0.2</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>-1.2</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>-1.6</td>
<td>5</td>
</tr>
<tr>
<td>Mean</td>
<td>+0.75</td>
<td></td>
<td>+2.33</td>
</tr>
</tbody>
</table>

\[
\bar{x} = -1.58
\]

\[
\frac{s^2}{10} = 0.0102977
\]

\[
s/\sqrt{10} = 0.1725
\]

\[
t = 9.1594
\]
The maximum correlation and hence minimum variance of difference is thus secured. If the method is valid, the (s) which it gives must be a rather efficient statistic. Whether, in your terminology, it is a sufficient statistic, or not, I am unable to determine.

I would be very glad to get your opinion of this method. Perhaps you can point out some fault which I overlook.

If you could send me a reprint or separate of the paper or papers in which your use of "degrees of freedom" is developed or elaborated, also the method of maximum likelihood, I would appreciate it very much. We do not have the journals in which your papers appear.

There is one other matter upon which I would value your opinion. A certain worker has conceived the following scheme. Suppose that in the example above the effect of the drug in each series was correlated with age of the patient. The standard error of estimate where effect of the drug is estimated from age, would represent the standard error which would be expected if age were constant. Or the square of the standard error of estimate is the variance of, hours of sleep gained, when all of the patients are of the same age. Substitute then the standard error of estimate in each series for the actual standard error, and calculate the standard error of the difference equal to \( \sqrt{s_1^2 + s_2^2} \). The sum under this radical indicates the respective standard errors of estimate. The result is the standard error of the difference in the conditions of constant age. To make the illustration good it should be assumed that the two series of patients are different persons and that the average age of the two sets is not necessarily the same. The observed difference is not corrected for the part which may be due to different mean ages of the two series. My point is that this correction should be made. Otherwise if it is possible to thus remove all of the variance due to age, it should also be possible to remove the variance due to any other factor in which case all of the variance may be removed and the observed difference will have a variance of zero. My colleague replies that each time a factor is held constant or the variance due to it is eliminated that one degree of freedom is lost and hence my argument against his method does not hold. For if, \( \lambda = 10 \), after nine factors have been held constant no more degrees of freedom remain and it is then impossible to hold any more factors constant or remove the variance due to them. This sounds very far-fetched to me. But I must confess that I do not understand the theory of degrees of freedom.

This method is essentially the same as that presented by Mr. F. D. Richey, "The moving average --", Jour. Agr. Research 33: 1161-, June 15, 1926. Except that Mr. Richey does correct the differences for that part which may be due to the factor which is being held constant, while my colleague deems this correction unnecessary. My position is that the correction must be made in order to present any appearance of respectability. And further, if the correction to the difference is made, that in small samples such corrections to the variance of difference or to the difference are of such slight precision as to be useless. I would value very highly a brief statement of your opinion of these points.

Thanking you, I am

Very truly yours,

[Signature]

FRED H. HULL
Assistant Agronomist

(Over)
P.S.

I trust that you will be able to grasp my argument for ranking observations according to magnitude and taking differences between two sets of members of corresponding rank. A great deal more might have been said, but the argument for is rather obvious. The argument against if any exists must be rather elusive. At least it escapes me entirely.

F.F.W.