# THE DISTRIBUTION OF RANGE IN SAMPLES FROM A NORMAL POPULATION, EXPRESSED IN TERMS OF AN INDEPENDENT ESTIMATE OF STANDARD DEVIATION

By D. NEWMAN

*Department of Statistics, University College, London*

## 1. Introduction

Starting from the contribution of Tippett (1925), a considerable amount of computational work has been carried out in recent years with the object of making possible the use of range, i.e. the distance between the highest and lowest observation, when dealing with samples from a normal population. Thus Tippett's tables of the mean range expressed in terms of the population standard deviation, $\sigma$, for sample sizes $n = 2$ to 1000 have been republished in *Tables for Statisticians and Biometricians*, Part II, Table XXII (K. Pearson, 1931). Later E. S. Pearson (1932) gave a table containing the standard deviation of range, and also the approximate upper and lower 10, 5, 1 and 0·5 % probability levels for sample sizes $n = 2$ to 100, again in terms of $\sigma$ as unit. In doing this he used empirical Pearson-type curves with correct moments, and checked his results against some experimental sampling distributions.

If a number of small samples are available, it has been shown that a rapid estimate of $\sigma$ may be obtained from the mean value of the range, which is only slightly less accurate than the estimate obtained from the sums of squares. Again, owing to the high correlation between range and standard deviation in a sample of size 10 or less, it was pointed out by Pearson & Haines (1935) that range may be usefully substituted for standard deviation in control charts used to study changes in the variation of quality in industry. In all these cases, however, the basic sampling distribution used has been that of the ratio of range to $\sigma$.

Not very long ago "Student" (the late Mr W. S. Gosset) suggested to Prof. E. S. Pearson that it might be useful to know more about the sampling distribution of the ratio

$$q = w/s,$$

where $w$ is the range in a sample of $n$ observations from a normal population with standard deviation $\sigma$, and $s^2$ is an *independent* and unbiassed estimate of $\sigma^2$ based on $f$ degrees of freedom, obtained from a sum of squares in the usual manner. The type of problem which "Student" had in mind was one in which, as a result of an experiment, a number of 'treatment' means, say $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$, are available, and also an independent estimate, $s^2$, of their sampling variance. Then a rapid method of judging whether any treatment differences exist would consist in comparing the difference between the highest and lowest treatment means, say

$w = \bar{x}_n - \bar{x}_1$, with $s$. Should this difference be clearly significant, having regard to the values of $n$ and $f$, the more divergent of the extremes, say $\bar{x}_1$, could be set aside, and the difference $\bar{x}_n - \bar{x}_2$ compared with $s$, using $n-1$ and $f$. This procedure would in fact be similar to that suggested by "Student" himself in his paper on "Errors in routine analysis" (1927, pp. 161–2), except that the ratio used would now be $w/s$ rather than $w/\sigma$. Of course the probability levels for the former will tend to those of the latter ratio as $f \to \infty$.

In the following sections I shall first make use of the results obtained by Prof. Pearson in computing probability levels for $w/\sigma$, to determine appropriate levels for $w/s$, and then illustrate "Student's" suggestion on three practical examples. It should be noted that in a recent paper H. O. Hartley (1938) has suggested a systematic method of obtaining probability levels for "studentized" functions. It is hoped that before long fuller and more accurate tables may be available to supplement the present tables which rest to some extent on an empirical basis.

## 2. The expectation of $q = w/s$

Before describing the method of quadrature by which the probability levels were obtained, it may be useful to give a table from which the expectation of $q$ for various values of $n$ and $f$ can be calculated. Since $s^2 = \sum_{i=1}^{f+1} (x_i - \bar{x})^2/f$, we have for the probability distribution of $s$,

$$p(s) = \frac{f^{\frac{1}{2}f}}{2^{\frac{1}{2}(f-2)}\Gamma(\frac{1}{2}f)\,\sigma^f} s^{f-1} e^{-fs^2/2\sigma^2} \tag{1}$$

If we write $p(w)$ for the probability distribution of range, a function whose precise value is only known for the cases $n = 2, 3$, and denote an expectation by the symbol $E$, we have

$$\begin{aligned}
E(q) &= \int_0^\infty p(q)\,q\,dq \\
&= \int_0^\infty \int_0^\infty w s^{-1} p(w)\,p(s)\,dw\,ds, \text{ since } w \text{ and } s \text{ are independent,} \\
&= \int_0^\infty wp(w)\,dw \times \int_0^\infty s^{-1} p(s)\,ds \\
&= E(w) \times \int_0^\infty s^{-1} p(s)\,ds.
\end{aligned}$$

Since the values of $E(w/\sigma)$ for changing $n$ have been tabled by Tippett, it is only necessary to consider the integral

$$\begin{aligned}
\int_0^\infty \sigma s^{-1} p(s)\,ds &= \frac{f^{\frac{1}{2}f}}{2^{\frac{1}{2}(f-2)}\Gamma(\frac{1}{2}f)\,\sigma^{f-1}} \int_0^\infty s^{f-2} e^{-fs^2/2\sigma^2}\,ds \\
&= \sqrt{(\tfrac{1}{2}f)}\,\Gamma\{\tfrac{1}{2}(f-1)\}/\Gamma(\tfrac{1}{2}f).
\end{aligned}$$

Hence it follows that

$$E(q) = E(w/\sigma) \times \sqrt{(\tfrac{1}{2}f)}\, \Gamma\{\tfrac{1}{2}(f-1)\}/\Gamma(\tfrac{1}{2}f). \tag{2}$$

A brief table of the second function is given below; the values of $E(w/\sigma)$ may be obtained from Tippett (1925), pp. 386–7, or from *Tables for Statisticians and Biometricians*, Part II, Table XXII.

TABLE I

*Factors by which to multiply $E(w/\sigma)$ to obtain $E(q)$*

| Degrees of freedom | 3 | 5 | 7 | 10 | 20 | 30 | $\infty$ |
|---|---|---|---|---|---|---|---|
| Factor | 1·382 | 1·189 | 1·126 | 1·084 | 1·040 | 1·026 | 1·000 |

### 3. Computation of table of 5 and 1% significance levels for $q = w/s$

The problem is to determine, for different values of $n$ and $f$, values $q_\alpha$ such that

$$\begin{aligned}\alpha &= \int_{q_\alpha}^{\infty} p(q)\, dq \\ &= \int_0^{\infty} \{p(w) \int_0^{w/q_\alpha} p(s)\, ds\}\, dw,\end{aligned} \tag{3}$$

where $\alpha = 0{\cdot}05$ and $0{\cdot}01$. Since the distribution of $q$ will clearly be independent of the population standard deviation $\sigma$, we may take $\sigma$ as unity in the probability functions used in (3). Except in the cases $n = 2, 3$, which will be referred to again below, the procedure adopted was as follows:

(*a*) The ordinates of the empirical curves, say $y(w)$, obtained by E. S. Pearson (1932) for the cases $n = 4, 6, 10$, and 20 were used in place of the unknown $p(w)$. These ordinates had been calculated at equal intervals of 0·1 for $w$, the population standard deviation being the unit.

(*b*) Taking a trial value of $q_\alpha$, the integrals $J(w, q_\alpha) = \int_0^{w/q_\alpha} p(s)\, ds$ were calculated, with the help of the *Tables of the Incomplete Gamma Function* (K. Pearson, 1922), for each value of $w$ used in (*a*). Thus $J(w, q_\alpha) = I(u, p)$ in the notation of the tables, where

$$u = (\sqrt{\tfrac{1}{2}f})\left(\frac{w}{q_\alpha}\right)^2, \quad p = \tfrac{1}{2}f - 1. \tag{4}$$

(*c*) It was then necessary to apply quadrature to the products $y(w) \times J(w, q_\alpha)$, calculated at intervals 0·5 for $w$, through as much of the range $w = 0$ to $w = \infty$ as was necessary to obtain the required degree of accuracy.

(*d*) The resulting expression would of course not correspond to $\alpha$ exactly, as $q_\alpha$ had been guessed. Other trial values were taken for $q_\alpha$, and the final value corresponding to $\alpha = 0{\cdot}05$ or $0{\cdot}01$ obtained by backward interpolation. Starting with the case $n = 2$, where the exact value of $q_\alpha$ could be obtained for all $f$, and knowing for $n > 2$ the limits to which $q_\alpha$ tended as $f \to \infty$, this process of trial and error was not found too laborious.

*Special case when $n = 2$*

In this case, taking $\sigma = 1$, the distribution of $w$ assumes the simple form

$$p(w) = \frac{1}{\sqrt{\pi}} e^{-\frac{1}{4}w^2} \quad \text{for } 0 \leqslant w < \infty. \tag{5}$$

Hence the joint probability distribution of $w$ and $s$ is

$$p(w, s) = \frac{f^{\frac{1}{2}f}}{2^{\frac{1}{2}(f-2)}\pi^{\frac{1}{2}}\Gamma(\frac{1}{2}f)} s^{f-1} e^{-\frac{1}{4}w^2 - \frac{1}{2}fs^2} \tag{6}$$

Transforming to variables $q = w/s$ and $s$, since

$$\left|\frac{\partial(w, s)}{\partial(q, s)}\right| = s, \tag{7}$$

we obtain
$$p(q, s) = \text{constant} \times s^f e^{-\frac{1}{4}s^2(q^2+2f)}. \tag{8}$$

Now integrating for $s$ between the limits 0 and $\infty$, we obtain for the probability distribution of $q$

$$p(q) = \sqrt{\left(\frac{2}{f\pi}\right)} \frac{\Gamma\{\frac{1}{2}(f+1)\}}{\Gamma(\frac{1}{2}f)} \left(1 + \frac{q^2}{2f}\right)^{-\frac{1}{2}(f+1)} \quad \text{for } q \geqslant 0. \tag{9}$$

This corresponds to the positive half of a "Student" distribution having $f$ degrees of freedom. Values of $q_\alpha$ satisfying the relation $\alpha = \int_{q_\alpha}^{\infty} p(q)\,dq$ may therefore be obtained from R. A. Fisher's (1938) tables of the percentage points for $t$. Thus

$$q_\alpha = t_\alpha \sqrt{2}, \tag{10}$$

where $t_\alpha$ will be respectively the 5 and 1 % levels for $t$.*

*Special case when $n = 3$*

For this case McKay & Pearson (1933) have given an expression for the true distribution of $w$. The quadrature method employed when $n > 3$ was again used, but the true values of $p(w)$ were taken, and not the ordinates of the empirical curve upon which E. S. Pearson (1932) based his original tables of percentage limits for $w$.

The following table shows the framework of values for $q_{0\cdot05}$ and $q_{0\cdot01}$ obtained as has been described. Values for $f = \infty$ were of course already known, and values for $n = 2$ and 3 are exact.

From Table II, the more complete working Tables III and IV were obtained by interpolation. It was found that the changes in percentage levels, both with increasing $n$ and $f$, ran most smoothly if the arguments $60/n$ and $60/f$ were used in place of $n$ and $f$. On this basis, interpolation from the framework values was effected, using five and six-point Lagrangian formulae. Various checks were

* These levels correspond to deviations at which the ordinates cut off 2·5 and 0·5 % from each end of the $t$-distribution, but they are termed by Fisher the 5 and 1 % levels.

TABLE II

*Framework values for $q_{0\cdot05}$ and $q_{0\cdot01}$*

| $f$ \ $n$ | 5% points | | | | | | 1% points | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 6 | 10 | 20 | 2 | 3 | 4 | 6 | 10 | 20 |
| 5 | 3·64 | 4·60 | 5·22 | 6·03 | 7·00 | 8·21 | 5·70 | 6·99 | 7·83 | 8·94 | 10·26 | 11·95 |
| 10 | 3·15 | 3·88 | 4·34 | 4·92 | 5·60 | 6·47 | 4·48 | 5·27 | 5·77 | 6·42 | 7·21 | 8·22 |
| 20 | 2·95 | 3·58 | 3·97 | 4·45 | 5·01 | 5·71 | 4·02 | 4·63 | 5·01 | 5·50 | 6·08 | 6·82 |
| 30 | 2·89 | 3·49 | 3·86 | 4·31 | 4·82 | 5·47 | 3·89 | 4·45 | 4·78 | 5·23 | 5·75 | 6·40 |
| $\infty$ | 2·77 | 3·31 | 3·65 | 4·04 | 4·48 | 5·01 | 3·64 | 4·12 | 4·38 | 4·74 | 5·15 | 5·64 |

carried out, as for example that of comparing the values obtained by this method with the known true values in the case of $n = 2$. Finally, the figures were reduced to one place of decimals, and these are given in Tables III and IV.

## 4. Illustrative examples

In the following examples the range test is used as an alternative to the $z$-test; the latter is, on theoretical grounds, the more efficient of the two in the sense that it is the more likely to detect the presence of real differences if they exist. Both "Student" and L. H. C. Tippett have, however, held that situations may be met, particularly when dealing with industrial problems, where the gain in speed following the use of range justifies the relatively small loss in efficiency. No doubt other types of examples besides those illustrated below will occur to the reader.

### *Example A*

Fisher (1937, p. 93) has given the results of a $6 \times 6$ Latin square experiment in which six different fertilizer treatments were applied to a crop of potatoes. Denoting these treatments by the letters $A, B, \ldots, F$, the mean yields per plot in lb. were as follows:

| $A$ | $B$ | $C$ | $D$ | $E$ | $F$ |
|---|---|---|---|---|---|
| 345·0 | 426·5 | 477·8 | 405·2 | 520·2 | 601·8 |

The analysis of variance table is as follows:

| Sources of variation | Degrees of freedom | Mean squares |
|---|---|---|
| Rows | 5 | 10,839·72 |
| Columns | 5 | 4,893·45 |
| Treatments | 5 | 49,635·98 |
| Error | 20 | 1,527·05 |

## TABLE III

*5 % points for $q = w/s$*

| $f$ \ $n$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 3·64 | 4·6 | 5·2 | 5·7 | 6·0 | 6·3 | 6·6 | 6·8 | 7·0 | 7·2 | 7·3 | 8·2 |
| 6 | 3·46 | 4·4 | 4·9 | 5·3 | 5·6 | 5·9 | 6·1 | 6·3 | 6·5 | 6·7 | 6·8 | 7·6 |
| 7 | 3·34 | 4·2 | 4·7 | 5·1 | 5·4 | 5·6 | 5·8 | 6·0 | 6·2 | 6·3 | 6·5 | 7·2 |
| 8 | 3·26 | 4·1 | 4·5 | 4·9 | 5·2 | 5·4 | 5·6 | 5·8 | 5·9 | 6·0 | 6·2 | 6·9 |
| 9 | 3·20 | 4·0 | 4·4 | 4·7 | 5·0 | 5·2 | 5·4 | 5·6 | 5·7 | 5·8 | 5·9 | 6·6 |
| 10 | 3·15 | 3·9 | 4·3 | 4·7 | 4·9 | 5·1 | 5·3 | 5·5 | 5·6 | 5·7 | 5·8 | 6·5 |
| 11 | 3·11 | 3·8 | 4·3 | 4·6 | 4·8 | 5·0 | 5·2 | 5·4 | 5·5 | 5·6 | 5·7 | 6·3 |
| 12 | 3·08 | 3·8 | 4·2 | 4·6 | 4·8 | 5·0 | 5·1 | 5·3 | 5·4 | 5·5 | 5·6 | 6·2 |
| 13 | 3·05 | 3·7 | 4·2 | 4·5 | 4·7 | 4·9 | 5·0 | 5·2 | 5·3 | 5·4 | 5·5 | 6·1 |
| 14 | 3·03 | 3·7 | 4·1 | 4·4 | 4·6 | 4·8 | 5·0 | 5·2 | 5·3 | 5·4 | 5·5 | 6·0 |
| 15 | 3·01 | 3·7 | 4·1 | 4·4 | 4·6 | 4·8 | 4·9 | 5·1 | 5·2 | 5·3 | 5·4 | 6·0 |
| 16 | 3·00 | 3·6 | 4·1 | 4·4 | 4·6 | 4·7 | 4·9 | 5·0 | 5·1 | 5·2 | 5·3 | 5·9 |
| 17 | 2·98 | 3·6 | 4·0 | 4·3 | 4·5 | 4·7 | 4·8 | 5·0 | 5·1 | 5·2 | 5·3 | 5·8 |
| 18 | 2·97 | 3·6 | 4·0 | 4·3 | 4·5 | 4·7 | 4·8 | 5·0 | 5·1 | 5·2 | 5·3 | 5·8 |
| 19 | 2·96 | 3·6 | 4·0 | 4·3 | 4·5 | 4·6 | 4·8 | 4·9 | 5·0 | 5·1 | 5·2 | 5·8 |
| 20 | 2·95 | 3·6 | 4·0 | 4·2 | 4·5 | 4·6 | 4·7 | 4·9 | 5·0 | 5·1 | 5·2 | 5·7 |
| 24 | 2·92 | 3·5 | 3·9 | 4·2 | 4·4 | 4·6 | 4·7 | 4·8 | 4·9 | 5·0 | 5·1 | 5·6 |
| 30 | 2·89 | 3·5 | 3·9 | 4·1 | 4·3 | 4·5 | 4·6 | 4·7 | 4·8 | 4·9 | 5·0 | 5·5 |
| 40 | 2·86 | 3·4 | 3·8 | 4·0 | 4·2 | 4·3 | 4·5 | 4·6 | 4·7 | 4·8 | 4·9 | 5·4 |
| 60 | 2·83 | 3·4 | 3·8 | 4·0 | 4·2 | 4·3 | 4·4 | 4·5 | 4·6 | 4·7 | 4·8 | 5·2 |
| $\infty$ | 2·77 | 3·31 | 3·65 | 3·87 | 4·04 | 4·18 | 4·29 | 4·39 | 4·48 | 4·55 | 4·62 | 5·01 |

## TABLE IV

*1 % points for $q = w/s$*

| $f$ \ $n$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 5·70 | 7·0 | 7·8 | 8·5 | 8·9 | 9·3 | 9·6 | 10·0 | 10·3 | 10·5 | 10·7 | 12·0 |
| 6 | 5·24 | 6·4 | 7·0 | 7·5 | 7·9 | 8·3 | 8·5 | 8·9 | 9·1 | 9·3 | 9·5 | 10·5 |
| 7 | 4·95 | 5·9 | 6·5 | 7·0 | 7·4 | 7·7 | 7·9 | 8·2 | 8·4 | 8·6 | 8·7 | 9·6 |
| 8 | 4·74 | 5·7 | 6·2 | 6·6 | 6·9 | 7·2 | 7·4 | 7·7 | 7·9 | 8·1 | 8·2 | 9·0 |
| 9 | 4·60 | 5·4 | 6·0 | 6·3 | 6·6 | 6·9 | 7·1 | 7·3 | 7·5 | 7·7 | 7·8 | 8·6 |
| 10 | 4·48 | 5·3 | 5·8 | 6·1 | 6·4 | 6·7 | 6·8 | 7·0 | 7·2 | 7·4 | 7·5 | 8·2 |
| 11 | 4·39 | 5·1 | 5·6 | 5·9 | 6·2 | 6·4 | 6·6 | 6·8 | 7·0 | 7·2 | 7·3 | 8·0 |
| 12 | 4·32 | 5·0 | 5·5 | 5·8 | 6·1 | 6·3 | 6·5 | 6·7 | 6·8 | 6·9 | 7·0 | 7·7 |
| 13 | 4·26 | 5·0 | 5·4 | 5·7 | 6·0 | 6·2 | 6·4 | 6·6 | 6·7 | 6·8 | 6·9 | 7·6 |
| 14 | 4·21 | 4·9 | 5·3 | 5·7 | 5·9 | 6·1 | 6·2 | 6·4 | 6·5 | 6·6 | 6·7 | 7·4 |
| 15 | 4·17 | 4·8 | 5·2 | 5·6 | 5·8 | 6·0 | 6·1 | 6·3 | 6·4 | 6·5 | 6·6 | 7·3 |
| 16 | 4·13 | 4·8 | 5·2 | 5·5 | 5·7 | 5·9 | 6·0 | 6·2 | 6·3 | 6·4 | 6·5 | 7·2 |
| 17 | 4·10 | 4·7 | 5·1 | 5·5 | 5·7 | 5·9 | 6·0 | 6·2 | 6·3 | 6·4 | 6·5 | 7·1 |
| 18 | 4·07 | 4·7 | 5·1 | 5·4 | 5·6 | 5·8 | 5·9 | 6·1 | 6·2 | 6·3 | 6·4 | 7·0 |
| 19 | 4·05 | 4·7 | 5·0 | 5·3 | 5·5 | 5·7 | 5·8 | 6·0 | 6·1 | 6·2 | 6·3 | 6·9 |
| 20 | 4·02 | 4·6 | 5·0 | 5·3 | 5·5 | 5·7 | 5·8 | 6·0 | 6·1 | 6·2 | 6·3 | 6·8 |
| 24 | 3·96 | 4·5 | 4·9 | 5·2 | 5·4 | 5·6 | 5·7 | 5·8 | 5·9 | 6·0 | 6·1 | 6·6 |
| 30 | 3·89 | 4·5 | 4·8 | 5·0 | 5·2 | 5·4 | 5·5 | 5·6 | 5·8 | 5·8 | 5·9 | 6·4 |
| 40 | 3·82 | 4·4 | 4·7 | 4·9 | 5·1 | 5·3 | 5·4 | 5·5 | 5·6 | 5·7 | 5·8 | 6·2 |
| 60 | 3·76 | 4·3 | 4·6 | 4·8 | 5·0 | 5·1 | 5·2 | 5·3 | 5·4 | 5·5 | 5·6 | 6·0 |
| $\infty$ | 3·64 | 4·12 | 4·38 | 4·59 | 4·74 | 4·87 | 4·98 | 5·07 | 5·15 | 5·22 | 5·28 | 5·64 |

To test the significance of treatment differences as a whole we find $z = 1{\cdot}7407$, while for degrees of freedom $f_1 = 5, f_2 = 20$, R. A. Fisher's tables give $z_{0{\cdot}01} = 0{\cdot}7058$. There are clearly, therefore, significant treatment differences present. The individual treatment means have been plotted in the accompanying figure, and the tabled significance levels for $q = w/s$ may be used, with discretion, as a foot-rule in investigating the situation.

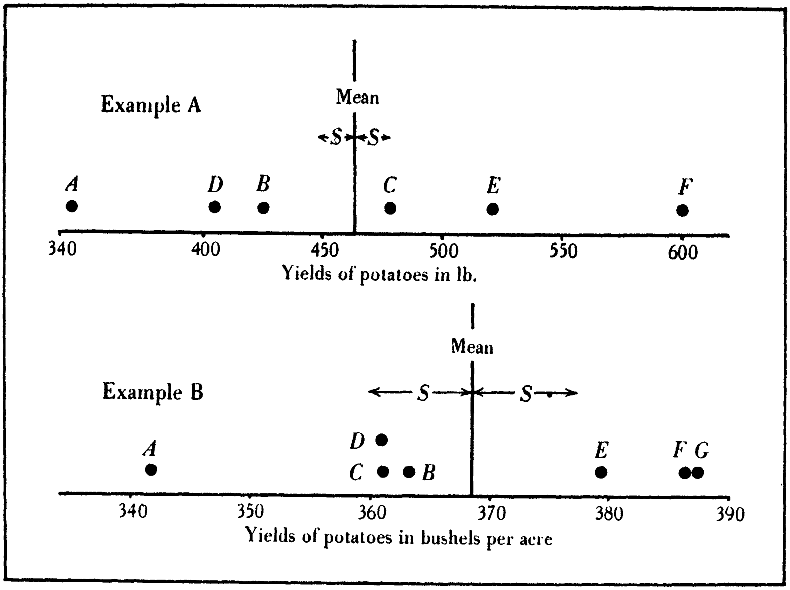


Fig. 1.

In the first place the appropriate pooled estimate, based on 20 degrees of freedom, of the standard error of a treatment mean is

$$s = \frac{39{\cdot}08}{\sqrt{6}} = 15{\cdot}95.$$

The range of the six treatment means is

$$w = 601{\cdot}8 - 345{\cdot}0 = 256{\cdot}8,$$

so that $$q = \frac{w}{s} = 16{\cdot}1.$$

Table IV above gives for $n = 6, f = 20$ a 1 % level for $q$ of $5{\cdot}5$, confirming the very significant scatter of treatment means brought out by the $z$-test.

We may now ask whether, if we were to exclude the most divergent treatment $F$, there is evidence of a significant difference among the remaining five treatments. We find that $w = 520{\cdot}2 - 345{\cdot}0 = 175{\cdot}2$, and, using the same estimate of standard error, $s = 15{\cdot}95$, find that $q = 11{\cdot}0$. This value is still well beyond $q_{0\cdot01} = 5{\cdot}3$, the figure obtained from Table IV with $n = 5, f = 20$.

Making successive trials we find that:

(i) Omitting $A$ and $F$, the range of the four treatments $B$, $C$, $D$, and $E$ is still significant, since $q = 115{\cdot}0/15{\cdot}95 = 7{\cdot}2$, while for $n = 4$, $f = 20$ we have $q_{0\cdot01} = 5{\cdot}0$.

(ii) Omitting $A$, $E$, and $F$, the range of the three treatments $B$, $C$, and $D$ is significant, at the 5 % but just not significant at the 1 % level. For in this case $q = 72{\cdot}6/15{\cdot}95 = 4{\cdot}5$, while for $n = 3, f = 20$ we find $q_{0\cdot05} = 3{\cdot}6$ and $q_{0\cdot01} = 4{\cdot}6$.

(iii) On the other hand, if we divided the six treatments into two groups, one consisting of $A$, $D$, $B$, and the other of $C$, $E$, $F$, we find that the value of $q$ in both groups falls beyond $q_{0\cdot01} = 4{\cdot}6$.

We are therefore led to conclude that the high value of $z$ obtained from the comprehensive test cannot be explained by one, or even two, treatments differing from the others. It is doubtful, even, whether any three treatments out of the six could be regarded as forming a homogeneous group.

Two final points should be noted. In the first place, after omitting successive treatments regarded as divergent, analysis of variance procedure could be applied to test for significant differences among the remaining treatments. The calculation would, however, not be as quick as that involved in the successive trials (i), (ii), and (iii) above, using $q$. Finally, as mentioned above, the method followed, whether $z$ or $q$ is used, must be employed with discretion, as is always the case when observations are rejected and a hypothesis tested using the selected data that remain.

*Example B*

A similar example has been taken from Snedecor (1937, p. 214). The following are seven treatment means expressed in terms of bushels per acre, obtained from a $7 \times 7$ Latin square experiment, again comparing the effect of different fertilizers on potatoes.

| *A* | *B* | *C* | *D* | *E* | *F* | *G* |
|---|---|---|---|---|---|---|
| 341·86 | 363·14 | 360·57 | 360·43 | 379·86 | 386·29 | 387·14 |

The appropriate estimate of the standard deviation of a mean of seven plots calculated from the error sum of squares of the analysis of variance table is $s = 9{\cdot}52$. Testing the significance of treatment differences as a whole, it is found that $z = 0{\cdot}5574$, while for degrees of freedom $f_1 = 6$, $f_2 = 30$, $z_{0\cdot05} = 0{\cdot}4420$ and $z_{0\cdot01} = 0{\cdot}6226$. The ratio $q = w/s$ for all seven treatments is $45{\cdot}3/9{\cdot}52 = 4{\cdot}7$ a value lying between $q_{0\cdot05} = 4{\cdot}5$ and $q_{0\cdot01} = 5{\cdot}4$ (entering Tables III and IV with $n = 7, f = 30$). Thus using either test we should conclude that there were probably significant treatment differences.

The seven treatment means have been plotted in the lower half of the figure. There is a suggestion that either if (i) treatments $F$ and $G$ or (ii) treatment $A$

were regarded as exceptional, the remaining treatments would form a homogeneous group.

This is confirmed on investigation:

(i) Removing $F$ and $G$, we find $q = 38\cdot 0/9\cdot 52 = 4\cdot 0$, which is just below the 5 % level ($q_{0\cdot 05} = 4\cdot 1$) for $n = 5, f = 3$.

(ii) Removing $A$, we find $q = 26\cdot 7/9\cdot 52 = 2\cdot 8$, which is well below the 5 % level ($q_{0\cdot 05} = 4\cdot 3$) for $n = 6$, $f = 3$. The more homogeneous group appears to be left in the second case, i.e. on removing $A$ alone. Beyond this indication we cannot go, as it would need a knowledge of the character of the treatments to draw more definite conclusions.

*Example C*

In cases where a number of duplicate observations are available, an estimate of variability may be obtained rapidly from summing the squares of the differences between pairs. Thus

$$s^2 = \sum_{t=1}^{k} (x_{t1} - x_{t2})^2/2k$$

and has $k$ degrees of freedom. We may now compare the range in the means of pairs with $s$, in order to determine whether there is too much variation between pairs having regard to the variation within pairs. It may be noted that an estimate of $\sigma$ may be obtained even more rapidly by calculating the mean range in pairs and multiplying by 0·8862,* so that

$$s' = 0\cdot 8862 \times \sum_{t=1}^{k} | x_{t1} - x_{t2} | /k,$$

but since the sampling distribution of $(s')^2$ is not that of $\chi^2$, we cannot justifiably take $q = w/s'$ and refer to the tables of percentage limits I have given.

*Determinations of percentage fibre*

| Analyst | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1st determination ($x_{t1}$) | 12·66 | 12·51 | 12·32 | 13·15 | 12·73 | 12·48 | 12·30 | 12·14 | 12·48 | 13·30 |
| 2nd determination ($x_{t2}$) | 12·47 | 12·62 | 12·55 | 12·93 | 12·43 | 12·46 | 12·73 | 12·03 | 12·44 | 12·68 |
| Difference ($d_t$) | 0·19 | −0·11 | −0·23 | 0·22 | 0·30 | 0·02 | −0·43 | 0·11 | 0·04 | 0·62 |
| Sum | 25·13 | 25·13 | 24·87 | 26·08 | 25·16 | 24·94 | 25·03 | 24·17 | 24·92 | 25·98 |

* The reciprocal of 1·12838, since the expectation of range in a sample of two individuals is 1·12838$\sigma$.

The data shown on p. 28 above consist of ten duplicate determinations of the percentage fibre in carefully mixed samples taken from the same supply of Soya Cotton Cake, each pair of values being obtained by one analyst.* Ten different analysts were concerned. The problem is to determine whether these few observations provide any evidence of systematic differences in technique between the analysts.

The full analysis of variance is as follows:

| | Sum of squares | D.F. | Mean square |
|---|---|---|---|
| Within pairs | 0·411450 | 10 | 0·041145 |
| Between pairs | 1·352045 | 9 | 0·150227 |
| Total | 1·763495 | 19 | |

Testing for differences between analysts, we obtain $z = 0{\cdot}6475$, a value falling between the 1 % (0·7990) and 5 % (0·5527) levels.

Using the range method, we note that the estimate of the variance of the *sum* of two determinations is

$$\frac{1}{k}\sum_{t=1}^{k}(x_{t1}-x_{t2})^2 = \frac{1}{10}\sum_{t=1}^{10}(d_t^2) = 0{\cdot}08229,$$

giving an estimate of the standard error of a sum of 0·2868.† The range in the ten sums in the table is $w = 26{\cdot}08 - 24{\cdot}17 = 1{\cdot}91$, so that $q = 1{\cdot}91/0{\cdot}2868 = 6{\cdot}66$. For $n = 10$, $f = 10$, Tables III and IV show $q_{0{\cdot}05} = 5{\cdot}6$, $q_{0{\cdot}01} = 7{\cdot}2$. Thus the difference is significant at the 5 % level, a result similar to that found using the $z$-test.

If now we omit determinations of analyst D (who gave the highest readings) we find $q = (25{\cdot}98 - 24{\cdot}17)/0{\cdot}2868 = 6{\cdot}31$, and is still significant at the 5 % level, since for $n = 9$, $f = 10$, $q_{0{\cdot}05} = 5{\cdot}5$. If, however, we omit the determinations of analyst H (who gave the lowest readings) we find $q = (26{\cdot}08 - 24{\cdot}87)/0{\cdot}2868 = 4{\cdot}22$, and is no longer significant. We should conclude, therefore, that except possibly in the case of analyst H, there is no evidence on the data available of systematic differences in technique.

I should like to take this opportunity of thanking Prof. E. S. Pearson not only for suggesting the problem to me, but also for his kind help in putting this paper together.

* The figures have been taken from more extensive data made available through the kindness of Dr J. F. Tocher.

† The corresponding estimate obtained, as described on p. 28 above, from the mean difference between pairs is 0·291.

## REFERENCES

FISHER, R. A. (1938). *Statistical Methods for Research Workers.* 7th ed. Edinburgh: Oliver and Boyd.

—— (1937). *The Design of Experiments,* 2nd ed. Edinburgh: Oliver and Boyd.

HARTLEY, H. O. (1938). *J.R. Statist. Soc. Suppl.* **5**, 80.

MCKAY, A. T. & PEARSON, E. S. (1933). *Biometrika,* **25**, 415.

PEARSON, E. S. (1926). *Biometrika,* **18**, 173.

—— (1932). *Biometrika,* **24**, 404.

PEARSON, E. S. & HAINES, JOAN (1935). *J.R. Statist. Soc. Suppl.* **2**, 83.

PEARSON, KARL (1922). *Tables of the Incomplete Gamma Function.* London: Biometrika Office.

—— (1931). *Tables for Statisticians and Biometricians,* Part II. London: Biometrika Office.

SNEDECOR, G. W. (1937). *Statistical Methods.* Iowa Collegiate Press, Inc.

"STUDENT" (1927). *Biometrika,* **19**, 151.

TIPPETT, L. H. C. (1925). *Biometrika,* **17**, 364.