# A STATISTICAL NOTE ON THE METHOD OF COMPARING MEAN VALUES BASED ON SMALL SAMPLES

BY

P. C. MAHALANOBIS, M. A. (Cantab.), B. Sc. (Cal.),

*Professor of Physics, Presidency College, Calcutta.*

(Received for publication on the 9th October 1931.)

1. Mr. K. V. Joshi, Cotton Physiologist, of the Cotton Research Laboratories, Surat, Gujrat, recently sent me certain data (given in Table I) for the "percentage success of bolls from flowers opening on a plant" for 20 plants each of 5 types of cotton*: (1) A. (2) B, (3) C. (4) D, (5) E. and enquired how the data could be used for comparative purposes.

Table I.

*Percentage success of bolls from flowers (Surat Cotton) (Mr. K. V. Joshi's Data).*

| No | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | 43·4 | 51·8 | 49·3 | 39·7 | 41·5 |
| 2 | 60·0 | 44·6 | 63·2 | 42·0 | 37·4 |
| 3 | 34·6 | 43·6 | 39·2 | 43·7 | 42·9 |
| 4 | 33·3 | 51·3 | 43·6 | 52·6 | 39·8 |
| 5 | 37·5 | 58·1 | 47·6 | 40·9 | 50·9 |
| 6 | 29·4 | 58·9 | 42·9 | 40·3 | 49·1 |
| 7 | 35·0 | 47·2 | 41·7 | 48·3 | 39·3 |
| 8 | 30·6 | 48·5 | 55·8 | 37·5 | 44·7 |
| 9 | 38·8 | 43·7 | 41·9 | 35·3 | 37·7 |
| 10 | 38·1 | 44·6 | 25·0 | 35·4 | 40·3 |
| 11 | 33·7 | 43·2 | 55·3 | 48·0 | 46·0 |
| 12 | 35·4 | 47·3 | 44·6 | 36·0 | 52·2 |
| 13 | 33·3 | 55·6 | 41·3 | 37·8 | 51·0 |

* As Mr. Joshi's experiments have not yet been concluded I refrain from giving the actual names of the various types of cotton used by him, and have referred to them by the fictitious names A, B, C, D, E. The actual names will become available when Mr. Joshi publishes full details of his experiments.

TABLE I—*contd.*

*Percentage success of bolls from flowers (Surat Cotton) (Mr. K. V. Joshi's Data)*—contd.

| No. | A | B | C | D | E |
|---|---|---|---|---|---|
| 14 | 33·8 | 44·4 | 52·9 | 34·9 | 43·0 |
| 15 | 36·6 | 41·9 | 45·5 | 32·3 | 36·4 |
| 16 | 33·0 | 45·5 | 38·2 | 35·9 | 44·4 |
| 17 | 29·8 | 50·0 | 51·8 | 38·7 | 50·0 |
| 18 | 30·0 | 44·4 | 53·0 | 36·5 | 50·0 |
| 19 | 35·9 | 47·9 | 35·0 | 40·0 | 36·1 |
| 20 | 27·3 | 44·2 | 51·1 | 29·2 | 49·2 |

2. For the comparison of mean values based on moderately large samples (of size greater than say 25 or 30) it is usually convenient to apply the classical theory of errors *. Let $s_1$ and $s_2$ be the standard deviations of the two samples of size $N_1$ and $N_2$ respectively. The standard deviations of the respective mean values $m_1$ and $m_2$ are then given by $s_1/\sqrt{N_1}$ and $s_2/\sqrt{N_2}$ respectively. The standard deviation of the differences between the two means. that is of $(m_1 - m_2)$ will be given by

$$S_d = \sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

in order to test the significance of the differences between the two means we divide $(m_1 - m_2)$ by $S_d$ and write

$$x = (m_1 - m_2)/s_d \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Using a table of the probability integral (for example, Table II of the Tables for Statisticians and Biometricians edited by Karl Pearson). it is possible to calculate the probability of occurrences of a difference as large as or greater than the observed difference. (A numerical illustration is given in paragraph 9.)

The observed difference is sometimes taken to be significant when it is twice its standard deviation (*i.e.*. $x=2$). This corresponds approximately to a probability of 1 in 20, and represents a lower limit below which it will almost never be safe to assert significance. On the other hand observed differences of 2·5 or 3 times the standard deviation may be safely considered significant.

*See discussion in paragraph 12.

3. The classical theory of errors involves three assumptions, namely (*a*) that the errors or deviations from mean values for both the variates are 'normally' distributed (*i.e.*, conform to the Gauss-Laplacian distribution), (*b*) that the two variates under comparison are statistically independent, and (*c*) the observed mean values and observed standard deviations are based on large samples.

It is clear, therefore, that the classical theory of errors cannot be legitimately used for the comparison of mean values based on small samples. During the last 15 years a good deal of theoretical work has been done on small samples by a large number of statisticians. To R. A. Fisher belongs the credit of giving certain general formulæ and also calculating necessary tables for facilitating the comparison of mean values based on small samples. An excellent description of the method is given in Chapter V of Fisher's book: "Statistical Methods for Research Workers "†. This method, however does not appear to be familiar to all workers in India. A detailed explanation of the numerical calculations may, therefore, prove useful.

4. Let us consider the data given in Table I. We must calculate (*a*) the mean values, and (*b*) the standard deviations for the different strains. For small samples it is usually more convenient to proceed directly (*i.e.*, without grouping). The individual figures in Table I are squared (with the help of a table of squares like Barlow's Tables of Squares, Cubes, etc.). and added. The work is shown in detail for (A) in Table II (*a*). The gross sum of the squares is 26069·51. We require, however, the sum of squares of deviations from the mean. In order to find this quantity we calculate a 'correction' which must be subtracted from the gross sum. The sum of (x) is + 709·5. Squaring 709·5 we have 503390·25. Dividing this quantity by 20 (the size of the sample). we get the correction —25169·51. Applying this correction (*i.e.*, subtracting 25169·51 from the gross

† Fisher's theory is free from the assumption of large samples, but still involves the two other assumptions of (*a*) normal distributions, and (*b*) statistical independence. During the last 5 years the first assumption (of normal populations) has been examined to some extent, and it has been shown that small samples drawn from skew populations (within wide limits) conform to Fisher's theory with closer approximation than to the classical theory. Hence the application of Fisher's theory to small samples from even fairly skew populations would not usually lead to seriously erroneous results. Alternative and more satisfactory tests to suit special purposes have also been devised by Egon Pearson, Neyman and others. But Fisher's test continues to be the most convenient one for every day use. The question of correlation between the variates has just begun to engage the attention of statisticians, and it is not unlikely that further modifications will become necessary when this point is more fully investigated. In fact what is urgently required is a fuller development of the theory of small samples drawn from skew correlated populations.

sum of squares) we obtain finally 900·00 as the sum of squares of deviations from the mean. Also, the mean value is obtained by dividing 709·5 by 20, or 35·475.

TABLE II.
*Calculations for Type (A).*

| No. | (a) | | (b) | | | |
|---|---|---|---|---|---|---|
| | x | $x^2$ | No. | x(+) | x(−) | $x^2$ |
| 1 | 43·4 | 1883·56 | 1 | +8·4 | .. | 70·56 |
| 2 | 60·0 | 3600·00 | 2 | +25·0 | .. | 625·00 |
| 3 | 34·6 | 1197·16 | 3 | .. | −0·4 | 0·16 |
| 4 | 33·3 | 1108·89 | 4 | .. | −1·7 | 2·89 |
| 5 | 37·5 | 1406·25 | 5 | +2·5 | .. | 6·25 |
| 6 | 29·4 | 864·36 | 6 | .. | −5·6 | 31·36 |
| 7 | 35·0 | 1225·00 | 7 | 0·0 | .. | 0·0 |
| 8 | 30·6 | 936·36 | 8 | .. | −4·4 | 19·36 |
| 9 | 38·8 | 1505·44 | 9 | +3·8 | .. | 14·44 |
| 10 | 38·1 | 1451·61 | 10 | +3·1 | .. | 9·61 |
| 11 | 33·7 | 1135·69 | 11 | .. | −1·3 | 1·69 |
| 12 | 35·4 | 1253·16 | 12 | +0·4 | .. | 0·16 |
| 13 | 33·3 | 1108·89 | 13 | .. | −1·7 | 2·89 |
| 14 | 33·8 | 1142·44 | 14 | .. | −1·2 | 1·44 |
| 15 | 36·6 | 1339·56 | 15 | +1·6 | .. | 2·56 |
| 16 | 33·0 | 1089·00 | 16 | .. | −2·0 | 4·00 |
| 17 | 29·8 | 884·04 | 17 | .. | −5·2 | 27·04 |
| 18 | 30·0 | 900·00 | 18 | .. | −5·0 | 25·00 |
| 19 | 35·9 | 1288·81 | 19 | +0·9 | .. | 0·81 |
| 20 | 27·3 | 745·29 | 20 | .. | −7·7 | 59·29 |
| Total | 709·5 | 26069·51 | Sum | +45·7 | −36·2 | 904·51 |
| | | | Total | +9·5 | .. | |
| | | −25169·51 | | | | −4·51 |
| | | 900·00 | | | | 900·00 |

A great deal of unnecessary arithmetical labour can, however, be usually saved by a slight modification of the procedure.* We first of all subtract a suitable constant quantity, say 35, from all individual figures. The resulting figures (some *plus*, and some *minus*) are given in columns 2 and 3 of Table II (b). The squares of these deviations are entered in column 4, and added, yielding 904·51. (The average value is easily seen to be given by 35+9·5/20=35·475 agreeing with the previous value). The correction is calculated in the same way. Squaring + 9·5 (the sum of deviations) we have 90·25. Dividing this quantity by 20, we get 4·51; subtracting 4·51 from 904·51 we again have 900·00 as the corrected sum of squares, which necessarily agrees with the value found by the more laborious process.

5. The mean value and sum of squares of deviations from the mean values for the other strains are calculated in the same way and entered as shown in Table III. The sum of squares are next divided by 19 to give the estimated † variances which are entered in column 6 of Table III. We shall also require the quantities $S^2/20$, and these are given in column 7 of the same table.

TABLE III.

*Mean values and variances.*

| (1) | (2) Variety | (3) | (4) Mean value | (5) Sum of squares | (6) $S^2$ | (7) $S^2/20$ |
|---|---|---|---|---|---|---|
| 1 | A . . . . . | 20 | 35·47 | 900·00 | 47·3684 | 2·36842 |
| 2 | B . . . . . | 20 | 47·68 | 418·75 | 22·0395 | 1·10197 |
| 3 | C . . . . . | 20 | 45·94 | 1392·07 | 73·2205 | 3·66102 |
| 4 | D . . . . . | 20 | 39·25 | 602·11 | 31·6895 | 1·58447 |
| 5 | E . . . . . | 20 | 44·09 | 558·67 | 29·4037 | 1·47018 |

* A recalculation of the standard deviation by the shorter method is in any case desirable as a check on the arithmetic.

† The observed variances (or squares of observed standard deviations) would be given by dividing the sum of squares of deviations by 20, the size of the sample. In Fisher's method it is, however, necessary to use the estimated variances which are obtained by dividing the sum of squares of deviations by the number of degrees of freedom. This number represents the number of independent comparisons possible within the sample, and is given by (N – 1) where N is the size of the sample. The estimated variance is the best estimate of the variance of the population from which the sample is drawn.

6. The fundamental formulæ in Fisher's method may be now given. Let $m_1$ and $m_2$ be two mean values based on samples of size $N_1$ and $N_2$, and let $S_1^2$, $S_2^2$ be the two corresponding estimated variances. We then calculate the quantity

$$t = \frac{(m_1 - m_2)}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

When $N_1 = N_2 = N$, we have

$$t = \frac{(m_1 - m_2)\sqrt{N}}{\sqrt{S_1^2 + S_2^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

In the present case $N_1 = N_2 = 20$, and formula (4) would be slightly more convenient to use. But formula (3) is more general, and as we have already tabulated the values of $S^2/20$ in Table III, column 7, we shall use it in the manner shown below.

Let us compare B and A.

We have—

$$m_1 = 47{\cdot}68\,, \qquad \text{also} \qquad S_1^2/20 = 1{\cdot}10197$$

$$m_2 = 35{\cdot}47\,, \qquad\qquad\qquad S_2^2/20 = 2{\cdot}36842$$

$$\text{Thus } (m_1 - m_2) = 12{\cdot}21\,, \quad \text{and} \quad (S_1^2/20) + (S_2^2/20) = 3{\cdot}47039$$

From Barlow's Tables we find the square root of 3·47039 to be 1·863. Hence "t" =12·21/1·863=6·554.

In the same way we can calculate the value of "t" for any pair of plants. It is convenient to proceed systematically. We first tabulate the differences between the mean values in Table IV.

## Table IV.

*Differences between mean values.*

| | A | B | C | D | E |
|---|---|---|---|---|---|
| A . . . . . . | .. | +12·21 | +10·47 | +3·78 | +8·62 |
| B . . . . . . | −12·21 | .. | −1·74 | −8·43 | −3·59 |
| C . . . . . . | −10·47 | +1·74 | .. | −6·69 | −1·85 |
| D . . . . . . | −3·78 | +8·43 | +6·69 | .. | +4·84 |
| E . . . . . . | −8·62 | +3·59 | +1·85 | −4·84 | .. |

The corresponding values of $(S_1^2/20)$ are then added (for each pair of plants) and entered in Table V.

TABLE V.

*Sums of* $(S_1^2/n_1+S_2^2/n_2)$.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| A . . . . . . | .. | 3·47039 | 6·02944 | 3·95289 | 3·83860 |
| B . . . . . . | 3·47039 | .. | 4·76299 | 2·68644 | 2·57215 |
| C . . . . . . | 6·02944 | 4·76299 | .. | 5·24549 | 5·13120 |
| D . . . . . . | 3·95289 | 2·68644 | 5·24549 | .. | 3·05465 |
| E . . . . . . | 3·83860 | 2·57215 | 5·13120 | 3·05465 | .. |

The corresponding square roots are next found from Barlow's Tables, and entered in Table VI.

TABLE VI.

*Standard deviations of differences.*

| | A | B | C | D | E |
|---|---|---|---|---|---|
| A . . . . . . | .. | 1·863 | 2·455 | 1·988 | 1·959 |
| B . . . . . . | 1·863 | .. | 2·182 | 1·639 | 1·604 |
| C . . . . . . | 2·455 | 2·182 | .. | 2·290 | 2·265 |
| D . . . . . . | 1·988 | 1·639 | 2·290 | .. | 1·748 |
| E . . . . . . | 1·959 | 1·604 | 2·265 | 1·748 | .. |

Finally the differences $m_1 - m_2$ (given in Table IV) are divided by the corresponding standard deviations (given in Table VI), and entered in Table VII, which thus furnishes the values of "t" corresponding to the comparison of any pair of plants.

TABLE VII.

*Values of "t"* $(n=38)$.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| A . . . . . . | .. | 6·554 | 4·265 | 1·901 | 4·400 |
| B . . . . . . | 6·554 | .. | 0·797 | 5·143 | 2·238 |
| C . . . . . . | 4·265 | 0·797 | .. | 2·921 | 0·817 |
| D . . . . . . | 1·901 | 5·143 | 2·921 | .. | 2·769 |
| E . . . . . . | 4·400 | 2·238 | 0·817 | 2·769 | .. |

7. We may now use Fisher's Table IV (p. 139) to test the significance of the observed values of "t". The appropriate value for 'n' in Table IV will be given by $(N_1+N_2-2)$ or $2(N-1)$ when the size of each sample is N. In the present case n=(20+20—2)=38. Unfortunately Fisher's Table does not extend beyond n=30. This, however, will not seriously hamper its use as the following examples will clearly show.*

Let us consider the comparison between A and B. Here t=6·554. From Fisher's Table IV we notice that for even n=30, the probability of occurrence of a value of "t" as large as 2·75 is only ·01, that is 1 in 100. The probability of occurrence of a value as great as 6·554 (for n=30, and hence also for n=38) must therefore be less than ·01. Hence we conclude that the observed difference is definitely significant, that is, it is practically certain that B gives a greater "percentage success" than A.

Let us now take B and C. From Table VII, we find t=0·797. Looking up Fisher's Table, we find that for n=∞, the probability of occurrence of a value 't' equal to or greater than 0·797 must be greater than 0·4, *i.e.*, greater than 40 per cent., but less than 50 per cent. That is, even when there is no real difference between the two strains, such a value of 't' as 0·797 will occur no less than between 40 and 50 times in 100 random trials. We conclude, therefore, that on the given data, we cannot assert that B would give better results than C.

8. Proceeding in the same way, we have entered in Table VIII the limiting probability for each comparison.†

TABLE VIII.

*Significance of differences: Limits of probability.*

| | A | B | C | D | E |
|---|---|---|---|---|---|
| A . . . . . | .. | <0·01 | <0·01 | ·05 & ·10 | < |
| B . . . . | (<0·01) | .. | (0·4 & 0·5) | (<0·01) | (·02 & ·05) |
| C . . . . . | (<0·01) | 0·4 & 0·5 | .. | (<0·01) | (0·4 & 0·5) |
| D . . . . . | (·05 & ·10) | <0·01 | <0·01 | .. | About ·01 |
| E . . . . . | (<0·01) | ·02 & ·05 | 0·4 & 0·5 | (About ·01) | .. |

* We can, of course, if we so desire extend Fisher's Table IV by interpolation. Let us take the present case for n=38. For P=·01 we notice that for n=30 the value of 't'=2·750 is higher than the value for n=∞ (*i.e.*, for infinitely large samples) t=2·57582, by (2·750—2·57582)=0·17418. We observe that [illegible]=0·8895 approximately. The interpolated difference then is 0·8895×0·17418. Hence the interpolated value of 't' for n=38 will be given by 2·5758+0·1549=2·7307. It is usually sufficient, however, to find the limiting probabilities as shown in the text.

† The brackets in Table VIII correspond to the *minus* sign in the differences given in Table IV. For example, in column 5, the brackets indicate that E gives a lower yield than B and C, while the absence of the bracket indicates that it gives a better yield than A or D.

B, C and E yield significantly greater "percentage success" than A; while B and C are definitely superior to D. Adopting a slightly lower level of significance, C is seen to be better than D. At a still lower level of significance, we find that B is better than E, but much reliance should not be placed on this result. The difference between D and A or between B and C or between C and E are statistically insignificant.

9. For large samples we can find a numerical value of the probability with the help of a table of the 'Probability Integral'. I have already mentioned that with samples of 20, the probability integral will not give absolutely correct results. The present figures may, however, be used for purposes of illustration.

Let us consider the comparison between B and C. Here t=0·797, this is the same quantity as "x" in Table II of the Tables for Statisticians and Biometricians (edited by Karl Pearson, Cambridge University Press). From this Table II (p. 2), we find:—

For $x=0{\cdot}79$, $\frac{1}{2}(1+\alpha)=0{\cdot}7852361$. The tabulated difference '$\Delta$' in $\frac{1}{2}(1+\alpha)$ for a change of $+{\cdot}01$ in '$x$' is $+{\cdot}0029085$. For a change of $(0{\cdot}797-0{\cdot}79)={\cdot}007$ in '$x$', the difference in $\frac{1}{2}(1+\alpha)$ must be $\frac{{\cdot}007}{{\cdot}01}\times(+{\cdot}0029085)=+{\cdot}0020359$. We thus have

| | | |
|---|---|---|
| For $x=0{\cdot}79$ | $\frac{1}{2}(1+\alpha)=$ | 0·7852361 |
| | Adding | ·0020359 |
| For $x=0{\cdot}797$, | $\frac{1}{2}(1+\alpha)=$ | 0·7872720 |
| Thus, | $\frac{1}{2}(1-\alpha)=$ | 0·2127280 |
| And | $2\times\frac{1}{2}(1-\alpha)=$ | 0·4254560 |
| Also | $1-(1-\alpha)=$ | 0·5745440 |

Hence if we assume that the observed deviation is as likely to be in excess as in defect, then the probability that we should reach or exceed the observed deviation is given by 0·425456. The probability that we shall not reach or exceed the observed deviation is, therefore, given by 1—0·425456=0·574544. The odds against the observed result are, therefore, roughly 58 to 42. We cannot assert that the observed difference is significant.

It will be remembered that from Fisher's t-table we had concluded by inspection that the probability of occurrence of a value of t as great as 0·797 (for n=38) lies between 0·4 and 0·5. From the probability integral we find that the calculated probability is 0·43. The two results are therefore consistent. (In fact the probability integral will yield fairly reliable results for samples of 20, although such samples cannot be strictly called large samples).

10. It is sometimes necessary to form a rough idea of the size of samples required to yield a significant difference in mean values.* For any given observed

* For example, Mr. Joshi has enquired what is the minimum number of plants for which data should be collected in order to get a significant difference in mean values.

difference in mean values the size of the sample will naturally depend on the variability as measured by the standard deviation.

Let $m_1$ and $m_2$ be the two mean values, $s_1$, $s_2$ the two estimated standard deviations based on N plants each from the two strains.

Writing $s = \sqrt{(s_1^2 + s_2^2)/2}$ for the average standard deviation, it is easy to see that

$$t = x = \frac{m_1 - m_2}{s} \sqrt{\frac{N}{2}} \quad \ldots \ldots \quad (5)$$

Let us write $$f = (m_1 - m_2)/s \quad \ldots \ldots \quad (6)$$

Then $$t = f \sqrt{\frac{N}{2}} \quad \ldots \ldots \quad (7)$$

That is, $$f = t \sqrt{\frac{2}{N}} \quad \ldots \ldots \quad (8)$$

Also the value of 'n' for entering Fisher's Table IV will be given by n = 2 (N−1) . . (9)

With the help of the above two equations (7) and (8) it is now easy to calculate the value of 'f' from the corresponding value of 't' given in Fisher's Table I (p. 139).

### TABLE IX.

*Values of "f". (Theory of small samples.)*

| N† | P‡ | | | |
|---|---|---|---|---|
| | 0·10 | 0·05 | 0·02 | 0·01 |
| 2 | 2·920 | 4·303 | 6·965 | 9·925 |
| 3 | 1·761 | 2·267 | 3·059 | 3·759 |
| 4 | 1·374 | 1·730 | 2·222 | 2·621 |
| 5 | 1·176 | 1·458 | 1·831 | 2·122 |
| 6 | 1·046 | 1·286 | 1·596 | 1·829 |
| 7 | 0·952 | 1·165 | 1·433 | 1·633 |
| 8 | 0·881 | 1·073 | 1·312 | 1·488 |
| 9 | 0·823 | 0·999 | 1·218 | 1·377 |
| 10 | 0·775 | 0·940 | 1·141 | 1·287 |
| 11 | 0·736 | 0·889 | 1·078 | 1·213 |
| 12 | 0·701 | 0·847 | 1·024 | 1·151 |

*That is, the variance is obtained by dividing the sum of squares of deviations by (N—1) the appropriate degrees of freedom.

† N=size of samples. ‡ P=probability of occurrence of 'f'.

TABLE IX—*contd.*

*Values of "f". (Theory of small samples*—contd.)

| N* | P† | | | |
|---|---|---|---|---|
| | 0·10 | 0·05 | 0·02 | 0·01 |
| 13 | 0·671 | 0·810 | 0·978 | 1·097 |
| 14 | 0·645 | 0·777 | 0·937 | 1·050 |
| 15 | 0·621 | 0·748 | 0·901 | 1·009 |
| 16 | 0·600 | 0·722 | 0·869 | 0·972 |
| 17 | 0·581 | 0·6986 | 0·8399 | 0·9395 |
| 18 | 0·564 | 0·6774 | 0·8139 | 0·9098 |
| 19 | 0·548 | 0·6581 | 0·7901 | 0·8828 |
| 20 | 0·5332 | 0·6403 | 0·7683 | 0·8580 |
| 21 | 0·5197 | 0·6239 | 0·7482 | 0·8352 |
| 22 | 0·5072 | 0·6086 | 0·7296 | 0·8141 |
| 23 | 0·4955 | 0·5944 | 0·7123 | 0·7946 |
| 24 | 0·4844 | 0·5809 | 0·6958 | 0·7760 |
| 25 | 0·4745 | 0·5689 | 0·6811 | 0·7593 |
| 30 | 0·4317 | 0·5170 | 0·6181 | 0·6883 |
| 35 | 0·3987 | 0·4772 | 0·5699 | 0·6341 |
| 40 | 0·3723 | 0·4453 | 0·5314 | 0·5909 |
| 50 | 0·3321 | 0·3970 | 0·4733 | 0·5258 |
| 75 | 0·2703 | 0·3228 | 0·3842 | 0·4264 |
| 100 | 0·2337 | 0·2789 | 0·3318 | 0·3680 |

Such a table of f-values is given in Table IX for 4 different levels of significance; P=0·10, 0·05, 0·02 and 0·01. In this table 'N' represents the size of each of the two samples on which the two mean values to be compared are based, and 'f' gives the largest value of $\frac{(m_1-m_2)}{s}$ which is significant within the assigned degree of probability (10 per cent., 5 per cent., 2 per cent. or 1 per cent.).

*N=size of samples †P=probability of occurrence of 'f'.

From N=2 to N=16, the 'f'=values are calculated directly from the 't'= values given in Fisher's Table IV. From N=17 to N=100, they are obtained from interpolated values of 't' in the same table.

11. If "s" is known even approximately, it is possible with the help of the above Table IX to find roughly the value of N, the size of the samples required to make any observed difference in mean values statistically significant.

For example, for Mr. Joshi's data, we can find an average value of ($s^2$) by adding the different values of sums of squares given in Table III, for the 5 different strains and dividing by 100. The average variance comes out to be 40·74 leading to an average value of s=6·4 approximately.

We can now use this value of s=6·4, in conjunction with Table IX, to give a rough idea of the size of the samples required for specific purposes.

*Example* (*i*). Let us compare B and A. The observed difference is 12·21 in favour of B. Dividing 12·21 by 6·4, we get f=1·91 approximately. Adopting P=·01, (that is odds of 100 to 1), we notice from Table IX that for N=6, 'f' is 1·829. We conclude that for this particular comparison, the observed difference would have been significant even if the mean values were based on small samples of size 6 or 7.

*Ex.* (*ii*). For C and D the observed difference is 6·69. That is f=1·05. In Table IX, with P=·01, and N=14, 'f' is 1·05. In this case samples of size 14 would be sufficient.

*Ex.* (*iii*). For D and A the observed difference is 3·78, giving a value f=0·59. For P=·01, N=40, f is 0·59 in Table IX. With odds of 100 to 1, we shall therefore require samples of size 40. Working with a 5 per cent. probability, we notice that N=23 gives a value of f=·5944. Hence with odds of 20 to 1, samples of size 23 would be just sufficient.

*Ex.* (*iv*). For B and C the observed difference is 1·74, with f=·272. With 5 per cent. probability we find that f=·2789 for N=100. We thus conclude that samples of size greater than 100 must be collected in order to make this particulars observed difference statistically significant.

I need scarcely note that unless we have some idea (even if rough and tentative) regarding the magnitude of the standard deviation, it is not possible to say anything about the required size of the sample.

Although in the above discussion the size of the sample has been assumed to be equal for both varieties, it is not necessary that this should be so in actual practice. Fisher's formula is quite general, and would apply for the comparison of mean values for two samples when the size of the sample is different in the two cases. It should be remembered that in using Fisher's Table IV, n should be taken equal to $(N_1+N_2-1)$, and we must write the t=formula as :—

$$t = f\sqrt{\frac{N_1 . N_2}{N_1+N_2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7,1)$$

$$\text{and } f = t\sqrt{\frac{N_1+N_2}{N_1 . N_2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8,1)$$

TABLE X.

*Values of "f". (Theory of large samples.)*

| N* | P† | | N* | P† | |
|---|---|---|---|---|---|
| .. | 0·05 | 0·01 | .. | 0·05 | 0·01 |
| 2 | 1·9600 | 2·5758 | 17 | 0·6723 | 0·8835 |
| 3 | 1·6003 | 2·1032 | 18 | 0·6533 | 0·8586 |
| 4 | 1·3859 | 1·8214 | 19 | 0·6359 | 0·8357 |
| 5 | 1·2396 | 1·6291 | 20 | 0·6198 | 0·8146 |
| 6 | 1·1316 | 1·4872 | 21 | 0·6049 | 0·7949 |
| 7 | 1·0476 | 1·3768 | 22 | 0·5909 | 0·7766 |
| 8 | 0·9800 | 1·2879 | 23 | 0·5779 | 0·7596 |
| 9 | 0·9240 | 1·2142 | 24 | 0·5655 | 0·7432 |
| 10 | 0·8765 | 1·1519 | 25 | 0·5514 | 0·7286 |
| 11 | 0·8357 | 1·0983 | 30 | 0·5060 | 0·6651 |
| 12 | 0·8001 | 1·0514 | 35 | 0·4685 | 0·6157 |
| 13 | 0·7689 | 1·0105 | 40 | 0·4382 | 0·5760 |
| 14 | 0·7409 | 0·9737 | 50 | 0·3920 | 0·5162 |
| 15 | 0·7157 | 0·9406 | 75 | 0·3201 | 0·4206 |
| 16 | 0·6929 | 0·9107 | 100 | 0·2772 | 0·3643 |

12. In this connexion it will not be out of place to say a few words regarding the magnitude of the inaccuracy involved in the use of the theory of large samples. In Table X will be found the values of 'f' calculated on the basis of the theory of large samples. For N=2, an observed value of f=2·58 would be considered significant (with P=·01) on the classical theory, whereas using the theory of small

* N=size of samples.

† P=probability of occurrence of 'f'.

samples we notice that '$f$' must reach the value of 9·93 for samples of 2 before the observed difference can be asserted to be significant with the same degree of probability.

It will be noticed that the difference between '$f$' values in the two Tables IX and X at first decreases rapidly with the increase of N, and then decreases more gradually. For N=10, the difference in the value of '$f$' is about 11 per cent. (P=·01). For N=25, the difference is just over 4 per cent., while for N=100, the difference is just about 1 per cent. The theory of large samples may, therefore, be used with safety for N greater than 100. It may be used without introducing appreciable errors practically for samples of size greater than 25.

The numerical calculations were completed by my assistants Messrs. Prabhat Ranjan Ray and Rabindra Nath Sen. My acknowledgments are also due to the Imperial Council of Agricultural Research for a research grant for work in agricultural statistics.