

# STATISTICAL NOTES FOR AGRICULTURAL WORKERS.\*

## No. 9.—CERTAIN VARIETAL STUDIES ON THE COTTON PLANT IN SURAT.

BY

P. O. MAHALANOBIS AND SUBIENDU SEKHAR BOSE.

*Statistical Laboratory, Presidency College, Calcutta.*

(Received for publication on 3rd October 1932)

1. Mr. K. V. Joshi, of the Cotton Research Laboratory, Surat, has sent us the data given in Tables I and II showing the percentage success of bolls from flowers for 5 different strains of the cotton plant sown in 1929-30 and 1930-31, respectively.

TABLE I.  
(1929-30)

	I	II	III	IV	V
Control . . . . .	38.2	37.7	38.9	37.9	38.2
Selection A . . . . .	43.2	41.0	42.3	41.2	40.2
Selection B . . . . .	46.5	45.3	45.0	45.6	44.7
Selection C . . . . .	46.8	47.4	49.3	47.1	46.5
Selection D . . . . .	49.5	46.6	48.7	49.6	47.6

TABLE II.  
(1930-31)

	I	II	III	IV	V
Control . . . . .	38.4	35.9	36.0	35.0	34.4
Selection A . . . . .	42.9	42.8	37.7	38.8	40.1
Selection B . . . . .	42.5	41.1	38.4	39.8	40.7
Selection C . . . . .	48.8	46.3	45.1	44.5	45.8
Selection D . . . . .	50.4	47.3	47.0	45.5	45.9

\* A large number of enquiries of a statistical nature are being received from agricultural workers in different parts of India. Many of these enquiries are of considerable general interest, and it is proposed to publish notes on selected topics from time to time. These notes will deal mainly with statistical methods and procedure, and it is not intended that they should always contain new matter.

2. Fisher's method of analysis of variance has been used for studying the above results. The value of observed  $z$  is given in each case by :

$$z = \frac{1}{2} \log_e v/v_0$$

where  $v_0$  = residual variance (error). The 5 per cent. (or one per cent.) values of  $z$  have been taken from Fisher's Table VI.

The analysis for 1929-30 is given in Table III.

TABLE III.  
(1929-30)

	D. F.	Sum of squares	Mean variance	Value of $z$	
				Observed	5 per cent.
Control w. Selection . . . . .	1	226.80	226.80	2.8903	0.7514
Selections . . . . .	3	135.69	45.23	2.0793	0.5876
Varieties . . . . .	4	362.49	90.62	..	..
Soil differences . . . . .	4	8.95	2.24	..	..
Error . . . . .	16	11.13	0.696	..	..
	21	382.57			

The varietal differences in yield are very pronounced and statistically significant.

The mean percentage of success and the respective differences are shown in Table IV.

TABLE IV.  
(1929-30)

	Mean percentage of success	Difference from				
		Control	Selec. 13	Selec. 65	Selec. 32	Selec. 9
Control . . . . .	38.18	..	-3.42	-7.24	-9.24	-10.22
Selection A . . . . .	41.60	3.42	..	-3.82	-5.82	-6.80
Selection B . . . . .	45.42	7.24	3.82	..	-2.00	-2.98
Selection C . . . . .	47.42	9.24	5.82	2.00	..	-0.98
Selection D . . . . .	48.40	10.22	6.80	2.98	0.98	..

Standard error of difference in means 0.53.

We find :-

- (i) Compared to control, all the selected varieties exhibit significantly higher percentage success of boll-formation from flowers.
- (ii) Compared to Selection A, the three other varieties Selections B, C and D are definitely superior.
- (iii) Selections D and C appear to be better than Selection B.
- (iv) The difference in mean percentage of success of boll-formation from flowers between Selections D and C is inappreciable.

3. The analysis for the data for 1930-31 is shown in Tables V and VI.

TABLE V.  
(1930-31)

	D. F.	Sum of squares	Mean variance	Value of z	
				Observed	5 per cent.
Control vs. Selections . . . . .	1	242.12	242.12	2.6578	0.7514
Selections . . . . .	3	177.05	59.02	1.9519	0.5876
Varieties . . . . .	4	419.17	104.79	..	..
Soil differences . . . . .	4	58.16	14.54	..	..
Error . . . . .	16	10.01	1.10	..	..
	24	496.34			

The mean percentages and differences are given in Table VI.

TABLE VI.  
(1930-31).

	Mean percentage of success	Difference from				
		Control	Selec. 43	Selec. 65	Selec. 32	Selec. 9
Control . . . . .	35.0	..	-4.6	-5.2	-10.2	-11.3
Selection A . . . . .	40.5	+4.6	..	-0.6	-5.0	-6.7
Selection B . . . . .	41.1	+5.2	+0.6	..	-5.0	-0.1
Selection C . . . . .	46.1	+10.2	+5.6	+5.0	..	-1.1
Selection D . . . . .	47.2	+11.3	+0.7	+6.1	+1.1	..

Standard error of difference in mean = 0.69.

We again observe that:—

- (i) All the selected strains are significantly better than the control.
- (ii) Selections D and C are better than Selections A and B.
- (iii) The difference between Selections A and B, as also the difference between Selections D and C are not significant.
- (iv) Finally, the two years' data may be combined when the analysis takes a slightly different form.

TABLE VII.  
(1929-30 and 1930-31)

—	D. F.	Sum of squares	Mean variance	Value of z	
				Observed	5 per cent.
Control vs. Selections . . . . .	1	468·80	468·86	1·9214	0·7514
Selections . . . . .	3	291·57	98·19	1·1397	0·5876
Varieties . . . . .	4	763·43	190·86	..	..
Soil differences . . . . .	8	67·11	8·29	..	..
Seasons . . . . .	1	52·02	52·02	1·8221	0·7072
Error . . . . .	30	48·43	1·36	..	..
	49	930·99			

TABLE VIII.  
(1929-30 and 1930-31)

—	Mean percentage of success	Difference from				
		Control	Selec. 43	Selec. 65	Selec. 32	Selec. 9
Control . . . . .	37·08	..	-3·97	-6·20	-9·70	-10·75
Selection A . . . . .	41·03	+3·97	..	-2·23	-5·73	-6·78
Selection B . . . . .	43·26	+6·20	+2·23	..	-3·50	-4·55
Selection C . . . . .	46·76	+9·70	+5·73	+3·50	..	-1·05
Selection D . . . . .	47·81	+10·75	+6·78	+4·55	+1·05	..

Standard error of difference in mean = 0·51 or 1·2 per cent. of mean value.

We find that—

(a) The superiority of—

- (i) All the selections over the Control,
  - (ii) Selections B, C and D over Selection A,
  - (iii) Selections C and D over Selection B
- is clearly established.

(b) The difference between Selections C and D is on the verge of significance.

(c) The seasonal difference is clearly significant.

5. The performance of Selection A was the subject of specific inquiry. The data for several seasons for control and Selection A are given in Table IX.

TABLE IX.

Season	Variety	No. of flowers	No. of bolls	Percentage success of bolls
1924-25	Control . . . .	84.3	34.9	41.4
	Selection A . . . .	68.4	25.5	37.3
1925-26	Control . . . .	87.4	30.0	34.3
	Selection A . . . .	92.9	32.3	35.3
1926-27	Control . . . .	66.5	24.1	36.3
	Selection A . . . .	70.6	29.6	41.9
1927-28	Control . . . .	87.0	35.5	40.8
	Selection A . . . .	85.2	30.7	43.1
1928-29	Control . . . .	90.8	34.3	37.7
	Selection A . . . .	89.0	35.0	40.3
1929-30	Control . . . .	55.6	21.2	38.1
	Selection A . . . .	61.9	26.7	43.2
1930-31	Control . . . .	69.3	24.3	35.0
	Selection A . . . .	65.2	25.4	38.9

A direct difference method for comparison may be adopted with advantage. The differences between Selection A and the control are shown separately for flowers, bolls and percentage success for each season in Table X.

TABLE X.

Season	Flowers	Bolls	Percentage success
1924-25 . . . . .	-15.0	-9.4	-4.1
1925-26 . . . . .	+5.5	+2.8	+1.0
1926-27 . . . . .	+4.1	+5.5	+5.6
1927-28 . . . . .	-1.8	+1.2	+2.3
1928-29 . . . . .	-1.8	+1.6	+2.6
1929-30 . . . . .	+6.3	+5.5	+5.1
1930-31 . . . . .	-4.1	+1.1	+3.9

We can now find the value of the standard deviation of each set of differences and obtain the values of  $z$  given in Table XI. The probability of occurrence of the observed value of  $z$  in each case is next found from Table XXV of the Tables for Statisticians and Biometricians, Part I.

TABLE XI.

—	Mean diff.	S. D.	$z$	$P$	Odds (approximate)
Flowers . . . . .	-1.1	7.12	-15	0.6276	2 : 3
Bolls . . . . .	+1.18	4.66	+25	0.7174	3 : 1
Percentage success of flowers to bolls	2.63	3.02	+87	0.9610	24 : 1

Selection A is thus significantly superior to the control so far as the percentage success of flowers to bolls is concerned; but the difference in the number of flowers or the number of bolls per plant is insignificant.

It is worth noting that the precision of the comparison is 1.2 per cent, which compares favourably with the highest degree of precision reached in varietal trials anywhere else.

We would like to add that this note is written purely from a statistical standpoint. Details of the experiments with full discussion of their agricultural bearing will be published in due course by Mr. Joshi himself.