FUNDAMENTALS OF

GENETICS

(Garden Pea to Gene Synthesis)

A. S. ISLAM

FUNDAMENTALS OF
GENETICS

(Garden Pea to Gene Synthesis)

A. S. ISLAM
Prof. of Botany
Dacca University, Bangladesh

VIKAS PUBLISHING HOUSE PVT LTD

VIKAS PUBLISHING HOUSE PVT LTD
Regd. Office: 5 Ansari Road, New Delhi 110002
H.O. Vikas House, 20/4 Industrial Area, Sahibabad,
Distt. Ghaziabad, U.P. (India)

1V2I1218

ISBN 0-7069-1798-7

Printed by Typographers (India) at Rashtravani Printers
A-49/1 Phase-I, Mayapuri, New Delhi 110064.

PREFACE TO THE SECOND EDITION

The first edition of the book was well received by the student community. They say that they find the book easy to follow compared to nicely illustrated and finely printed books of genetics published by the western co u tries. This book became more popular later on account of its ready availability and low price.

In East Africa where the author went to on a visiting professorship, the book found ready acceptance in the Universities of Dar-as-Salaam and Nairobi. There is now demand of this book in Western Africa. This is partly due to the fact as explained in the preface of the first edition that in this book not only genetics but also the allied branches, namely biometry, population genetics, evolution and cytology have been dealt with.

The first edition of this book was published by the author himself but the Vikas Publishing House, New Delhi has now taken the responsibility of bringing out its second edition. It is expected that the second edition will find market not only in Bangladesh but in India and other developing countries through the network of the agencies of this renowned Publishing House.

Some of the chapters such as 'Apomixis', 'Karyotype' have been omitted because these topics are not dealt with in undergraduate classes and therefore increase the volume of the book without much benefit to these group of students. The last chapter on 'Biometry', namely 'Analysis of Variance in factorial experiment' has been completely rewritten by Professor A.H. Talukder of the Department of Statistics, Dacca University. My heartfelt thanks are due to him who not only went through the portion of Biometry but rewrote the last chapter.

There were many printing mistakes in the first edition. Effort has been made to present the book with minimum of printing errors but still there may be some printing and even factual mistakes. The author will therefore be grateful if the students and other readers point out these mistakes so that in the next edition or reprint these mistakes are taken care of.

If the students for whom the book has been written find the book useful for learning the fascinating subject of genetics, the author will find his effort greatly rewarded.

A. S. ISLAM

PREFACE TO THE FIRST EDITION

I have been teaching genetics in post-graduate classes for a number of years. While delivering lectures, I have often felt that most of the students are unable to follow the lectures. When I assessed their answer books, I found that their answers suffer from lack of understanding of the subject. I have often discussed with my students as to the reasons of their not understanding the subject of genetics fully. The majority of the students are of opinion that although there are a large number of books on genetics, evolution, cytology and plant breeding, they do not have the means to buy them all for lack of funds and secondly they are unable to dig information from all the relevant books necessary to grasp the subject matter and prepare themselves for examination.

It is true that in western countries such a problem does not exist. There are big libraries attached to each department and students have sufficient funds to buy the necessary text-books. Further, because these books are written in their mother tongue, they can easily follow and comprehend them and score high marks in the examination. Not only that, those who come to the University in western countries are students with a good academic background. In those countries, the jobs being many, the mediocre students leave the educational institutions long before they reach the University stage. In our country every one, whether he has aptitude for higher education or not wants to get into the University, obtains a degree in order to secure a job.

This is the reason why the number of post-graduate students, is ever increasing all over Bangladesh and elsewhere where such problems exist. The majority of students choosing biology at the University stage do not have adequate theoretical background.

This book has, therefore, been written with a purpose to helping those students, who for the reasons mentioned overleaf, cannot acquire knowledge from a large number of books. The variety of facts and their interpretation which they need for acquiring knowledge in genetics and allied subjects have been put together in one book. This book does *not* claim any originality. In fact, most of the chapters have been written with the aid of existing text-books on genetics etc. The references to these books have, however, been

cited at appropriate places. In a few chapters, however, such as, on 'Polyploidy', 'Origin of species' 'Gene and development,' 'Genetic basis of apomixis', some new informations, which the author and his students obtained through their experimentation, have been given.

The chapter on population genetics has been written with a little extra care so that the biology students of this part of the world who are weak in mathematics understand fully this important branch of genetics which is fast developing. The chapter on 'How to locate genes in chromosomes' has also been given a little different treatment than in other text-books.

This book has been primarily written for B.Sc. students in Botany, Zoology and Agriculture of the Universities of Bangladesh. Since science has no national boundary, students of other countries as well as those of post-graduate classes may also find the book useful particularly its chapters on 'Population genetics', 'How to locate genes in chromosomes', 'Genetic basis of apomixis', 'Polyploidy'.

An expert, while going through the book, will observe many repetitions not only between different chapters but within the same chapter. The author is not apologetic on this issue because he thinks that repetitions are necessary in a text-book.

While writing this book I have often asked this question to me. Were I a student, would I have understood this book or struggled for understanding the subject matter dealt in it? I don't know how far I have succeeded in achieving this objective. It is left to the students using this book to judge.

My sincere thanks are due to Dr. M.A. Hannan of Bangladesh Atomic Energy Commission who helped me in writing up certain portions of the book relating to genetics of micro-organisms. Mr. A.M. Mojmadar of the department extended his help in writing up portions of biochemical genetics and genetic map in *Neurospora* for which I sincerely thank him. My heartfelt gratitude is due to Dr. M.A. Zaman of the department and Dr. M.M. Mia of AEC, Dacca who read critically a few chapters and offered their valuable criticism. My sincere thanks are also due to Mr. Mahtabuddin Ahmed of JRI, Dacca for his criticism in the portion on biometry.

I also thank most sincerely my student Mr. Satyajit Bhadra of final year M.Sc. who supplied and worked almost all the numerical sums given in the book and partly prepared the index. I also express my sincere gratefulness to Mr. Abdul Matin of the department who drew the illustrations and made the cover design of the book.

Last but not the least my sincerest gratitude is due to my student and colleague Dr. Mozammel Haque of the department but for whose untiring effort and steadfast devotion to the completion of this task, the book will have never seen the light of the day. Mr. A. Hai, Proprietor, Asiatic Press and his entire staff members particularly, A Moslim, G. Mustafa, A. Malek Sarker, A. Salam, Sirajul Islam, A. Latif also deserve may sincere thanks who worked very hard to get the book published.

Lastly I must point out humbly that many mistakes have crept in the book (both printing and factual) in spite of my best efforts. Besides many mistakes will be detected by the readers when they will go through the book. I shall deem it a great privilege if the readers point out these mistakes to me in writing so that in the second edition or second impression I get an opportunity to correct them.

To the teachers I have one special request to make, that is, they may kindly impress upon the students the importance of going through the exercises given at the end of chapters and discuss with them the answers question by question.

If my students for whom the book has been written are benefited out of this book, I shall deem my humble effort amply rewarded.

A.S. ISLAM

Appendix I

BIOMETRY

1. TEST OF HOMOGENEITY OF PROPORTIONS

In the study of inheritance pattern, one is often confronted with a situation in which the different families from the product of a single cross do not give exactly the same ratio. For instance, in the F_2 of the cross, round seed × angular seed (round dominant), Mendel obtained 3 : 1 ratio in the majority of families but in some he got 19 : 10 *i.e.*, roughly 2 : 1 ratio in one family, 26 : 6 *i.e.* approximately 4.3 : 1 ratio in another. Now the question is whether the pooled sum of the entire population is to be regarded as an illustration of monohybrid ratio. For this, we do a test called the test for homogeneity which tells us whether the samples *i.e.* the different families are homogenous. The test is done using the following formula:

$$\frac{\Sigma_i(n_{ij}^2/n_{i.}) - n_{.j}^2/n_{..}}{p(1-p)} \; ; j=1 \text{ or } 2.$$

where i is treatment (family), j is column *i.e.*, the total number of plants in one family, p is the expected ratio of plants in one class and $(1-p)$ in the other.

Let us write the composition of 10 F_2 families in terms of round and angular seeds as was recorded by Mendel.

Plant	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Total*
Round seed	45	27	24	19	32	26	88	22	28	25	336
Angular seed	12	8	7	10	11	6	24	10	6	7	101
Total treats $n_{i.}$	57	35	31	29	43	32	112	32	34	32	437

$p=0.75$; $1-p=0.25$ because the assumption is that the seeds of the ten plants are segregating for 3 : 1 ratio.

Step 1. Square each number of all the ten classes either of the round seeded or angular seeded class and divide each by the total number of seeds in that class. Since the number of seeds in the angular seeded class is small, let us use the numbers in the ten classes of this category. Subtract correction factor ($n_{.j}^2/n_{..}$) from this value.

$$\frac{12^2}{57}+\frac{8^2}{35}+\frac{7^2}{31}+\frac{10^2}{29}+\frac{11^2}{43}+\frac{6^2}{32}+\frac{24^2}{112}+\frac{10^2}{32}+\frac{6^2}{34}+\frac{7^2}{32}-\frac{101^2}{437}=$$

$$2.5263+1.8285+1.5806+3.4482+2.8139+1.1250+5.1428+$$

$$3.1250+1.0588+1.5406-10201/437$$

$$=24.1897-23.3432=0.8465.$$

Step 2. Divide 0.8465 by the product of $p \times (1-p)$ *i.e.* $0.75 \times 0.25=0.1875$. This is chi-square (χ^2) :

$$\chi^2=\frac{0.8465}{0.1875}=4.514.$$

Conclusion. The observed χ^2 value is 4.514 whereas the tabular χ^2 value for 9 degrees of freedom is 4.17 at 0.1% and 5.9 at 0.25% probability. There is good evidence therefore to suppose that the ten families do not differ significantly in the proportion of round seed vs. angular seed.

2. 't' TEST

The two varieties (*IRRI* and irradiated) of rice have been grown in 10 different locations under same set of conditions. The size of each plot is 1/10th of an acre and the yield per plot is expressed in terms of mds (roughly 40 kg). We want to know whether there is any real difference in their yield. For this we do '*t*' test, the formula of which is shown below.

$t=\frac{d}{s_d}$, where d is the difference between treatment means and s_d is the standard error of difference between treatment means. s_d is determined by taking square root of twice the error variance divided by the total number of observations *i.e.* $\sqrt{\frac{2s^2}{n}}$. For comparison of two treatments '*t*' test is made; for three or more treatments '*F*' test is done.

The data for yield of rice of two varieties are given below:

TABLE I. DATA OF YIELD IN MDS IN 1/10TH ACRE OF LAND

Location no.	*IRRI* X_1	*Local 1 irradiated* X_2	*Location no.*	*IRRI* X_1	*Local 1 irradiated* X_2
1	5.1	4.4	6	4.0	3.5
2	6.0	5.0	7	4.5	3.0
3	4.6	4.2	8	5.2	3.3
4	5.5	4.4	9	5.6	4.6
5	4.5	4.3	10	5.0	4.3

Calculation. Total $IRRI - \Sigma X_1 = 50$; Local 1 irradiated $- \Sigma X_2 = 41$
Sum of squares: $\Sigma X_1^2 = 253.32$; $\Sigma X_2^2 = 171.64$.

Mean $\bar{x}_1 = 5$ lbs. Mean $\bar{x}_2 = 4.1$ lbs.

$n_1 = n_2 = n = 10$.

$$\Sigma x_1^2 = \Sigma X_1^2 - (\Sigma X_1)^2/10 = 3.32.$$

$$\Sigma x_2^2 = \Sigma X_2^2 - (\Sigma X_2)^2/10 = 3.54.$$

$$s^2 = \frac{\Sigma x_1^2 + \Sigma x_2^2}{2(n-1)} = 0\ 381.$$

$$s_d = \sqrt{\frac{2s^2}{n}} = 0.276.$$

$$t = \frac{x_1 - x_2}{s_d} = 3.26.$$

Procedure

Step 1. Determine the two means. Add separately all values; under each variety total is represented symbolically by ΣX_1 and ΣX_2 and divide each total by the number of observations *i e.* 10. These values here are $\frac{50}{10} = 5$ mds and $\frac{41}{10} = 4.1$ mds.

Step 2. Square each and every observation under one variety, say IRRI. Add them all. This is ΣX_1^2 (a). Similarly square all observations under the other variety and add them. This is ΣX_2^2(b). These values in the present examples are 253.32 and 171.64.

Step 3. (*a*) Square the two totals $(\Sigma X_1)^2$ and $(\Sigma X_2)^2$ obtained in step 1, divide each by 10, the number of observations taken per variety.

This is correction factor for each variety. The values are

$$\frac{50^2}{10} = 250 \text{ and } \frac{41^2}{10} = 168.1$$

(*b*) Subtract $(\Sigma X_1)^2/10$ from ΣX_1^2 and $(\Sigma X_2)^2/10$ from ΣX_2^2. This is the corrected sum of squares for each variety expressed as Σx_1^2 and Σx_2^2. These values here are:

$$\Sigma X_1^2 - (\Sigma X_1)^2/10 = 253.32 - 250 = 3.32$$

$$\Sigma X_2^2 - (\Sigma X_2)^2/10 = 171.64 - 168.1 = 3.54.$$

Step 4. An estimate of the common variance (σ^2) is obtained by adding the two sum of squares and divide them by combined degrees of freedom for two varieties *i.e.*

$$s^2 = \frac{\Sigma x_1^2 + \Sigma x_2^2}{2(n-1)} = \frac{3.32 + 3.54}{2 \times 9} = \frac{6.86}{18} = 0.381.$$

Step 5. Determine s_d, the standard deviation on account of the difference between sample means

$s_d = \sqrt{2s^2/n} = \sqrt{2 \times 0.381/10} = \sqrt{0.762/10} = \sqrt{0.0762} = 0.276.$

Step 6. Find the value of t by dividing the difference between two means by s_d. In the present example this is,

$$t = \frac{5 - 4.1}{0.276} = \frac{0.9}{0.276} = 3.26 \ (t_{.01} = 2.878).$$

Conclusion. Consult tabular value for t at 1% level of significance for 18 degrees of freedom. This value is 2.878. Since the observed value 3.26 is greater than the theoretical value 2.878, the difference between the two is very significant.

3. ANALYSIS OF VARIANCE (AOV): RANDOMISED BLOCK DESIGN

AOV is a statistical method by which analysis of more than two groups of measuring data is made. Fundamentally it is a test of significance of more than two treatments. Let us now see it in an example of a particular variety of *IRRI* rice. Supposing we want to know whether a particular variety of *IRRI* rice gives a better yield compared with four local varieties. For this comparison, five varieties including the one to be tested are grown in a statistically designed plot with five replications for each variety. The design may be many, such as randomized block design, latin square, lattice square, split plot, factorial, etc., and such design should be used which could best serve the purpose. Similarly the number of replications may also vary according to precision wanted and land availability.

Here we shall use the most common design—randomized block design with five replications for each variety of treatment. This design removes heterogeneity of plots in one direction. The size of each replicate was one tenth of an acre. Varieties were laid out in a random order in each block/replication and the blocks were also randomized. The yields in different blocks/replications are given in

AOV table with its working formula

Source (*Source of variation*)	*df* (*degrees of freedom*)	*SS* (*Sum of squares*)	*MS* (*Mean square*)	*F*
Treatment	$t-1$	*SST*	*SST*/*df*	$\frac{MST}{MSE}$
Block/Rep.	$r-1$	*SSR*	*SSR*/*df*	$\frac{MSR}{MSE}$
Error	$(t-1)(r-1)$	*SSE*	*SSE*/*df*	
Total	$rt-1$	*SS* Total	*SST*/*df*	

table II. *Analysis*: An *AOV* table is given on page *iv* with its working formula. Proceed analysis stepwise:

TABLE II: DATA ON YIELD OF FIVE VARIETIES OF RICE
YIELD IN MAUNDS (*Md*=NEARLY 40 KG)

Replication no.	*IRRI*	*Local 1 (Irradiated)*	*Local 2*	*Local 3*	*Local 4*	*Total X.j*
1	5.0	4.0	3.2	1.6	2.0	15.8
2	6.0	4.5	4.4	2.7	2.2	19.8
3	4.0	5.0	2.7	2.5	4.5	18.7
4	5.5	3.0	3.4	4.2	3.0	19.1
5	4.5	3.5	1.8	3.5	3.3	16.6
Total X_i.	25.0	20.0	15.5	14.5	15.0	$90=X..$
X_1	$\bar{x}_1=5.0$	$\bar{x}_2=4.0$	$\bar{x}_3=3.1$	$\bar{x}_4=2.9$	$\bar{x}_5=3.1$	$18=\bar{x}..$
ΣX_{ij}^2	127.50	82.50	51.69	45.99	48.98	356.66
$(X_i.)^2/r$	125.00	80.00	48.05	42.05	45.00	340.10
ΣX_{ij}^2	2.50	2.50	3.64	3.94	3.98	16.56

Step 1. Determine the correction term or factor C as shown here. Add all the observational values (=90). Square this number and divide it by the total number of observations made (=25).

Therefore $C=\dfrac{90^2}{25}=324.$

Step 2. Determine the total sum of squares. It is done by squaring each and every observation (there are 25 such observations) adding them up and then subtracting from this value the correction factor as shown below:

$$\text{Tatal } ss=5.0^2+6.0^2+\ldots\ldots+3.0^2+3.3^2-\frac{90^2}{25},$$

$$=356.66-324=32.66.$$

Step 3. Determine the sum of squares due to the treatments. Add up all 5 observations under each treatment. Since there are five treatments, there will be 5 numbers. Square summation of each treatment, divide it by 5 and then subtract from it the correction factor as shown below:

$$SST=\frac{25^2+20^2+15.5^2+14.5^2+15.0^2}{5}-C$$

$$=340.10-324.00=16.10.$$

Step 4. Calculate block sum of squares. Add all the 5 observa-

tions for each block, square these totals and add them. Divide sum of these squares by 5, the number of observations in each block; subtract from it the correction term. This is given below:

$SSBl = 1/5\ (15.8^2 + 19.8^2 + 18.7^2 + 19.1^2 + 16.6^2) - C$

Step 5. Determine the error sum of squares. This is done by subtracting treatments sum of squares and block sum of squares from

Thus error sum of squares = Total sum of squares—treatment *SS* —block *SS*

$$14.212 = 32.66 - 16.10 - 2.348$$

Now let us enter the results obtained so far in an analysis of variance table (Table III).

TABLE III. ANALYSIS OF VARIANCE TABLE FOR COMPARISON OF YIELD BETWEEN AN IRRI STRAIN AND FOUR LOCAL VARIETIES

Variation source	*df*	*Sum of squares (SS)*	*Mean square (MS)*	*F*
Varieties	4	16.10	4.025	4.025/0.888=4.53
Block	4	2.348	0.587	0 587/0.888=0.66
Error	16	14.212	0.888	
Total	24	32.66		

Step 6. Determine mean square value (*MS*) for treatment sum of squares by dividing 16.10 by 4 (16.10/4=4.025), *MS* due to block by dividing 2.348 by 4 (2.348/4=0.587) and *MS* of error sum of squares by dividing 14.212÷ by 16 as shown above (14.212÷16 =0.888).

Step 7. Determine the *F* value by dividing *MS* value due to treatment by *MS* value due to error. This value here is 4.025/0.888=4.53.

Step 8. Consult the tabular value of *F** for 4 and 16 degrees of freedom (4 refers to numerator and 20 to denominator).

The tabular *F* values at 5% and 1% levels of probability are 3.01 and 4.77 as against calculated value of 4.53. Since the latter value exceeds 5% tabular value *i.e.*, 3.01, the inference is that real difference exists among different varieties of rice.

Supposing the value was less than 2.33 which is the tabular value at the above degrees of freedom at 10% level, then the conclusion would have been that no significant difference exists between the

*F values given in any standard text-book on statistics.

varieties. If F is significant, proceed for further tests: individuals, paired or set some measure of limits by orthogonal 't' test, Duncan's new multiple range or Tukey's procedure or *lsd* method.

Reliability of estimated effects as indicated by standard errors (S.E) which is also used to set limits within which true values fall and these limits are confidence limits.

Standard error of treatment means (*a*) *and that of difference between treatment means* (*b*). To determine (a) the formula $\sqrt{\frac{s^2}{r}}$ is used, where s^2=residual error or error mean sq. and r=number of observations per treatment. In the present example the answer will be, $\sqrt{\frac{0.888}{5}}$ 0.421 md. The formula, $\sqrt{\frac{2s^2}{r}}$ gives a measure of difference between treatment means. This value will be, $\sqrt{\frac{2\times 0.888}{5}}$ =0.596 md.

Comparison between paired means for their level of significance. (*a*) Least significant difference (*lsd*) method: The formula for this is *lsd* (.05)=$t_{.05}s_d$ where .05 means 5% probability level $t_{.05}$ is the tabular value given for Student's t distribution (Student is the pseudo- name of the scientist who derived this distribution at 5% level and s_d s $\sqrt{\frac{2s^2}{r}}$, the standard error of the difference between two treatment means. To determine *lsd* therefore at a particular level of significance for a particular error degrees of freedom the 't' value is obtained from the table and is multiplied by s_d, the standard error between treatment means. In the present example the *lsd* at 0.1, 0.05 and 0.01 level of significance will be:

$$lsd\ (0.1) = 1.746\times 0.596 = 1.041$$
$$lsd\ (0.05) = 2.12\times 0.596 = 1.264$$
$$lsd\ (0.01) = 2.921\times 0.596 = 1.741$$

Now calculate the observed difference between two means under test. If this value is more than 1.741, the difference is very significant and if it is below 1.741 but more than 1.264 then the difference is significant. If it is 1.041 or above but below 1.264 it is just significant.

The differences between *IRRI* and local 1 and other paired means are:

$x_1-x_2=1.0$, $x_2-x_3=0.9$, $x_3-x_2=0.2$, $x_4-x_5=0.1$. The conclusion is that difference between *IRRI* and local 1 is significant at 10% level and other differences are not significant. However, the difference

of mean between x_1-x_3 (1.9), x_1-x_4 (2.1), x_1-x_5 (2.0) is very significant because each one of these differences is more than 1.741, the tabular value at 1% significance.

Tukey's w procedure. For comparing all pairs of treatments, *lsd* method is not suitable. Tukey's w procedure necessary to judge significance of each pair of observed difference is, therefore, used to cover all cases of real differences (Tukey's 'w' test is also known as honestly significant differences) (*hsd*) and is determined by the following formula.

$$w=q\alpha\ (p, n_2)\ s_x$$

where q is the tabular value of studentised range for this particular test being the probability level (0.5 and 0.01) and p, n_2 represent the number of treatments and error degrees of freedom respectively. s_x is the standard error of treatment means. In the present example w value at 0.05 and 0.01 level of significance, for $p=5$ and $n_2=16$ will be:

$$w_{.05}(5,16)=4.33\times 0.421=1.823$$
$$w_{.01}(5,16)=5.49\times 0.421=2.31$$

The means showing no significant difference are grouped together bg underscoring them. In other words, any two means not underscored are significantly different.

5.0 4.0 3.1 3.0 2.9

4. DOUBLE GROUPING : LATIN SQUARE

Double grouping provide more opportunity than randomized block design for reduction of errors. It is meant to eliminate error due to

TABLE IV. FIELD LAYOUT ACCORDING TO LATIN SQUARE DESIGN FOR FOUR VARIETIES. A=IRRI, B, C AND D ALL LOCAL VARIETIES. YIELD PER 1/10TH ACRE IS EXPRESSED IN MDS (MD=APPROX. 40 KG).

					Row total	
	A	*B*	*D*	*C*		
	5.0	4.4	3.0	3.5	15.9	
	C	*D*	*A*	*B*		
	2.7	4.5	4 0	3.4	14.6	
	B	*A*	*C*	*D*		
	3.2	6.0	4.2	3 3	16.7	
	D	*C*	*B*	*A*		
	2.2	2.5	2.7	5.5	12.9	
Column total	13.1	17.4	13.9	15.7	69.1	Grand total

difference among rows and also errors due to difference among columns. But this design is not suitable for a large number of treatments, say more than 12. Number of replications equals or is multiple of number of treatments. In the following example we shall compare the yield of four varieties of rice and subject the data for a more rigorous analysis by taking into account also the error factors on account of blocks and rows. For such an analysis, the layout of the field is made in a design called latin square design (Table IV) in which four varieties, four replications are made with one variety occurring only once both in a row and a column. For three varieties it will be 3×3 *i.e.*, replications will be three and with five, 5 replications with each variety occurring only once in a row and a column.

Step 1. From the Table IV the variety totals under each treatment are shown below:

$A=5.0+6.0+4.0+5.5=20.5$ mds.
$B=3.2+4.4+2.7+3.4=13.7$,,
$C=2.7+2.5+4.2+3.5=12.9$,,
$D=2.2+4.5+3.0+3.3=13.0$,,

Step 2. Calculate the correction factor by squaring the grand total and dividing it by the number of observations *i.e.* $\frac{X^2}{16}$ .. $=\frac{60.1^2}{16}$
$=225.75=C$.

Step 3. Determine total sum of squares by squaring each of the 16 observations and adding them. Subtract correction factor from this number.

Total $SS=\sum_{ij} X_{ij}^2=5.0^2+2.7^2+3.2^2.....3.3^2+5.5^2-C$
$=244.11-225.75=18.36$ (a).

Step 4. Square four treatment totals (each obtained by adding 4 replicates of the same treatment) and add these four squares. Divide this value by 4. Subtract correction factor from this *i.e.*

$$SST=\frac{\Sigma_t X_t^2}{r}-C=\frac{20.5^2+13.7^2+12.9^2+13.0^2}{4}-C$$
$=235.835-225.75$ or 10.085 (b).

Step 5. Square four row totals (each obtained by adding 4 observations in each row), add these four squares. Divide this number by 4. Subtract from it the correction factor *i.e.*

$$\text{Row } SS=\Sigma_i X_i^2-C=\frac{15.9^2+14.6^2+16.7^2+12.9^2}{4}-C$$
$=228.5175-225.75=2.7675$ (d).

Step 6. Square four column totals (each obtained by adding 4 observations in each column), add these four squares. Divide this

number by 4 and subtract correction factor from the quotient *i.e.*

$$\text{Column } SS=\Sigma_j X_j^2.-C=\frac{13.1^2+17.1^2+13.9^2+15.7^2}{4}-C$$

$$=227.8175-225.75=2.0675\ (e)$$

Step 7. Determine error variance by subtracting from the total sums of squares (a), the combined sums of ($b+d+e$) *i.e.* $a-(b+d+e)$ or $SSE=18.36-(10.085+2.0675+2.0675)=3.44$.

Step 8. Enter the results in the analysis of variance Table V as shown below:

TABLE V. ANALYSIS OF VARIANCE

Source of variation	*df*	*SS*	*MS*	*F*
Rows	3	2.7675	0.9225	0.9225/0.573=1.609
Column	3	2.0675	0.6892	0.6892/0.573=1.203
Varieties	3	10.0850	3.3600	0.362/0.573=5.867
Error	6	3.4400	0.5730	
Total	15	18.3600		

Significant at 5%. Tabular value at 5% is 4.76.

Conclusion. Since the observed value 5.867 is greater than the tabular value at 5% level, the difference between the varieties is significant. The other two F values are insignificant indicating that the errors due to rows and blocks are small.

Some other values for this example

$$s=\sqrt{0.573}=0.76$$

$$s_{\bar{x}}=\sqrt{\frac{s^2}{r}}=\sqrt{\frac{0.573}{4}}=\sqrt{0.14325}=0.38$$

$$s_d=\sqrt{2s^2/r}=\sqrt{2\times 0.573}=\sqrt{0.2865}=0.536$$

Coefficient of variability $CV=\sqrt{s^2}/\bar{x}^*\times 100=\sqrt{.573}/3.75\times 100$

$$=\frac{0.76\times 100}{3.75}=\frac{76}{3.75}=20.26\%$$

5. LINEAR CORRELATION

In a set of experiments carried out to determine whether there is any proportionate increase in jute fibre yield (Y) corresponding to the dosage increase in N fertilizers (X), the following data (Table VI)

*Mean $\bar{x}=60.1/16=3.75$.