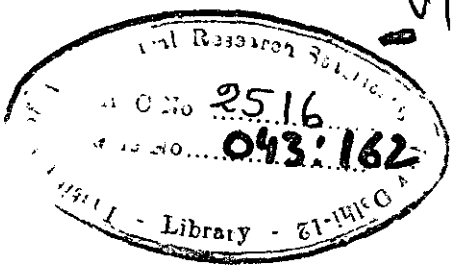
A STUDY OF GENETIC MODELS, WITH LINKAGE IN QUANTITATIVE INHERITANCE

AND

A STUDY ON LINKAGE BETWEEN GENES FOR PIGMENTATION IN VARIOUS PARTS OF THE RICE PLANT

BY

P.S. Puri



Dissertation submitted in partial fulfilment of the requirements for the award of Diploma in Agricultural & Animal Husbandry statistics of the Indian Council of Agricultural Research

New Delhi

-1958-

ACKNOWLEDGEMENT

I have pleasure in expressing my deepest sense of gratitude to Dr. V.G. Panse, M.Sc., Ph.D.(London), F.N.I., F.A.Sc., Statistical Adviser, Indian Council of Agricultural Research, New Delhi, for his valuable guidance, constant encouragement and constructive criticism during the course of investigation and writing this thesis.

P. S. Puri

(Prem Singh Puri)

PART I

"A STUDY OF GENETIC MODELS, WITH LINKAGE IN QUANTITATIVE INHERITANCE"

C O N T E N T S

I. INTRODUCTION

1.1. The importance of the study of inheritance of quantitative characters, particularly its study in the field of plant and animal breeding, has now been very well realised. The progress in this direction has of course been very slow as compared to the genetics of qualitative characters, which is mainly due to the fact, that after the rediscovery of Mendel's principles of heredity in 1900, there was a conflict whether the inheritance of quantitative characters, which exhibit continuous variation, could be explained on the basis of Mendel's theory, which was thought to exhibit a discontinuous heredity mechanism. Actually quantitative inheritance had been studied even before the rediscovery of Mendelism. Francis Galton made extensive studies in this field in the latter half of the nineteenth century. Darwin believed that continuously varying characters were the ones which had great importance for evolution. In his efforts to develop the theory underlying Darwin's belief, Galton made no progress towards the solution of the problems of heredity, but he did devise excellent biometrical tools for the description and analysis of variation.

1.2. When Mendelian heredity came to the fore with its emphasis on sharply discontinuous unit characters, it appeared that this was completely dishermonious with the Galtonian studies and the belief was widespread that quantitative inheritance could not be understood in Mendelian terms. It is to the credit of Nilsson-Ehle & E.M. East, to have proposed independently the multiple factor hypothesis, according to which quantitative inheritance could be analysed in Mendelian terms if it be assumed that there are numerous independent pairs of genes, each pair contributing a small amount to the determination of a quantitative character, these contributions being similar and supplementary in their effect.

In 1909, Johannsen showed by his experiments on beans, that variation in a character, such as bean size was the combined result of genetical agencies and the environmental influences at work and that the effect of these two agencies could be distinguished by breeding tests only. In view of the results obtained by Nilsson-Ehle & Johannsen, it became apparent that through a polygenic system having a large number of genes, with similar and supplementary effects and a larger effect of environment, continuous distributions of the type observed, could be produced. Thus it was the multiple-factor hypothesis together with the idea of there being an effect of environment, which resolved the conflict of Biometry & Mendelism, of continuity and discontinuity, thereby unifying the Mendelian & Biometrical approaches to the study of inheritance of quantitative characters.

1.3. The knowledge that the polygenes are borne on the chromosomes provides us with a base from which to explore the genetical properties of continuous variation. Since the quantitative characters are controlled by a large number of genes, each having a small effect and are largely affected by the environment, the individual genotype cannot be distinguished nor the effect of various genes followed up individually. This made necessary development of statistical techniques by which the resultant effects of a large number of genes could be followed in populations, without attempting the impossible approach of recognising the individual genes or genotypes. With this idea in mind, East and his co-workers, could show satisfactorily that the inheritance of a number of quantitative characters in maize & tobacco can be explained by the multiple factor hypothesis. The multiple factor hypothesis, thus was finally taken to be recognised as providing a reasonable explanation of the inheritance of quantitative characters.

1.4. The full implications of this hypothesis have, however, only gradually become realised, as it brought out many problems relating to the number and

magnitudes of effects of the factors, their mode of action, whether additive or multiplicative, their dominance and linkage relationships. These problems were gradually tackled by a number of research workers. The special types of experiment and statistical analysis necessary for the study of continuous variation have only gradually become available. Nevertheless, though slow, progress has been real and we are now in a position to see not merely how continuous variation can be explained genetically, but also how experiments can be conducted enabling us to understand and to measure the special quantities in terms of which continuous variation can be analysed and its behaviour predicted in some measure. Some of the statistical techniques which have been developed so far will be explained in the next chapter.

1.5. Turning towards plant and animal breeding practices an important instance of the application of the statistical-cum-genetical approach in the development of the plant breeding technique was the switchover from mass selection to progeny row breeding (Hutchinson & Panse, 1937). This approach enabled the plant breeder not only to proceed with selection systematically and efficiently but also to estimate the amount of genetic variability existing in his material. Apart from genetic variability, quantities like number of factors, magnitude of dominance effects, the strength of linkage present among the factors, etc are equally important for the understanding of the genetic situation. The present investigation deals with the development of appropriate statistical methods for elucidation of these quantities, with particular reference to selection in plant breeding.

1.6. The most practical problem in plant breeding i.e. determining the speed of genetic advance, was investigated by Panse (1940 a & b), who introduced the method of genetic models which forms a valuable contribution to the study of relationship between the number of genetic factors and their variability on the one hand and the statistical properties of the selected progenies on

the other. These F_2 - properties are, the genotypic mean, mean genotypic variance within progeny, variance of progeny means, covariance of progeny mean and genotypic variance within progeny and the variance of genotypic variance within progeny. He gave a method for estimating the effective number of factors and by setting up different hypothetical genetic models regarding the systems of genetic factors, magnitude of dominance effects etc, he studied the effect of variation in the number of factors operating to produce a given amount of genotypic variability, on the above five F_2- properties. The results from this approach, can be used to anticipate the response to selection under a particular genetic set up and also to judge from the results of selection, the genetic situation obtaining in any particular material.

1.7. Adopting the same approach as originally given by Panse; Sohil (1954) worked out similar results in the case of 18 other genetic models which covered a sufficiently wide and useful range of the number, magnitude and dominance relationships of genetic factors. In these genetic models, considered separately by Panse & Sohil, independent segregation of the genes concerned was assumed. It may be mentioned that apart from, the rigour of selection, effective number of factors, variation in magnitude of action of the factors, their dominance relations and the also of the contribution of non-heritable agents to F_2 - variation, linkage among factors is also an important factor influencing the speed of genetic advance under selection. Although the broad effect of the linkage could be expressed as ranging between the two limits, namely fifty percent recombination, which is equivalent to independent segregation of the two factors and complete linkage, which would amount to the two factors behaving as a single factor with its effect being equal to the sum of the two, it is of considerable scientific interest to work out explicitly the effect of varying degrees of linkage among the genes concerned on the statistical properties of the segregating population.

I took up the present investigation with the object to study the effect of linkage on the speed of genetic advance due to selection and on various other F_3 properties, following the statistical approach developed by Panse. This is presented in Part I of the thesis.

During the same period, I undertook another problem, namely "A study on linkage between genes for pigmentation in various parts of the rice plant" which is presented in Part II of the thesis.

II. REVIEW OF LITERATURE

2.1. We shall now proceed to give a review of the developments in the various statistical techniques adopted by various research workers in the field of quantitative inheritance to obtain the information of genetic interest, useful to the plant breeder.

The discussion of the genetical basis of continuous variation concerns the relations between gene and character. Broadly two main principles have emerged from studies made so far. The first of these, which we owe to Johannsen (1909) is that the phenotype is the joint product of genotype and environment. The variation in a character may therefore result from variations in genotype and environment. Johannsen showed that the two kinds of variations in the character, heritable and nonheritable, cannot be distinguished by mere inspection. Only a breeding test can distinguish the one from the other. The second principle is that the variation in a character may be caused by alteration in any of a number of genes. The genes controlling the same character may be related in various ways. They may all be members of a polygenic system, or they may all be major genes whose effects on the character are neither small nor supplementary, or lastly the variation may be due partly to a major gene and partly to a polygenic system.

2.2. Unlike discontinuous variation which is produced by genes of major effect, in continuous variation it is necessary to put the biometrical tools and quantities like means, variances etc on some measurable scale. This scale should obviously be one which would facilitate both the analysis of the data, and the interpretation and use of the resulting statistics. The main criteria of scaling, as have been discussed by Mather (1949), are two. One is that the genic effects on the average be simply additive. The second is that the contribution made by nonheritable agents must be independent of the genotypes.

If the chosen scale is deficient in either or both ways, the statistics by which the distributions are described will be distorted in greater or less degree. For this purpose some scaling tests are devised (Mather 1949). Data on the mean measurements of two parents, the F_1 & F_2 generations and the backcrosses of the F_1 to the parents have been given by Powers (1941) for six characters in certain tomato crosses, where he has discussed the adequacy of the scales used in taking the measurements for representation of the characters and for analysis of the variation.

In a later paper (1943), Powers has dealt further with the use of the logarithmic transformation for data on fruit weights to convert the geometric action of genes into the additive one. The use of transformations to remove epistasy is appropriate, but there may be considerable other difficulties in that each set of data may require its own individual transformation which would make generalisation impossible. This comes to our attention forcibly when the data on the same population collected in different years required different transformations as shown by Powers (1950, 1953).

8.3. Fisher's study of correlation between relatives (1918) has been of great value in the study of quantitative inheritance, as it gave a support to the multifactorial hypothesis, thereby helping in filling up a gap between biometrical and Mendelian approaches. Further it showed how the observed variation in quantitative characters could be split up with the help of correlation coefficients, into components assignable to the additive action of genes, to dominance and to environmental fluctuations. Fisher demonstrated this approach on Pearson & Lee's cubit measurement on human data.

The classical paper on various third degree statistics and their genetical interpretation, by Fisher, Immer, & Tedin (1932) is of notable importance. It gave a number of results relating to second degree statistics and also relating to the progenies of the third generation, which are obtained by intercrossing pairs of plants taken at random from F_2 (biparental

progenies) or by exposing one F_2 plant as mother to pollination by the pollen of F_2 as a whole (maternal progenies). Third degree statistics showed how the average degree of dominance manifested by various factors can be assessed.

2.4. After the multi-factorial hypothesis was well recognised by the geneticists, the genetic situation regarding the number of genes producing variability in segregating populations, excited the interest of statistical research workers. In this respect, Student (1934) in his paper on the calculation of the minimum number of genes in Winter's selection experiment, attempted to give an estimate of the number of genes producing variability in oil content in maize. Basing his estimate on the ratio of maximum genotypic range to the genetic standard deviation, assuming equal effects for various factors and also independence of their segregation, he reached the conclusion that there were at least 20 polygenes controlling oil content in maize in Winter's experiment.

The phenomenon of linkage between polygenes and the major genes is also an interesting one. This phenomenon was well demonstrated by Sax (1923) who reported the case of apparent linkage between polygenes governing the character of continuous variation, namely bean-weight and a major gene responsible for its pigmentation, a qualitative character. The results found by him did not give much evidence of linkage as it appeared that the effect on the bean-weight might be a pleiotropic secondary effect of the major gene responsible for pigmentation.

A more convincing evidence on the linkage of polygenes with major genes was provided by Rasmusson (1935) who investigated the variation of flowering time in crosses of garden pea. He confirmed the presence of linkage between polygenes governing the flowering time and a major gene governing flower colour, by obtaining lines, in which the association between two characters was absent and also lines in which it was reversed. The linkage

between polygenes and major genes has been established in several other cases. Mather (1941 & 1942) has demonstrated the presence of linkage between polygenes themselves in his study of abdominal chaetae number in drosophila and its fertility. The two characters governed by two polygenic systems were found to be associated. The possibility of the same ploygenic system governing both of the characters (pleiotropic effect) was ruled out when the linkage was found to be broken when proceeding with breeding and selection alternately.

2.5. / It might be said that the linkage of polygenes and the concept of effective number of factors form a most significant aspect of the study of polygenic variation. In this respect methods of ascertaining the effective number of factors developed by Wright (1934) and Fense (1940) are of special importance. Further, the estimation of the speed of advance corresponding to a given degree of selection is known to depend upon the number of segregating factors which govern the inheritance of the particular character under study. In a polygenic system, it has already been pointed out that it is impracticable to mark the segregation of individual genes. At best combinations of genes could be considered as our ultimate units, which are brought to be so constituted that no recombination occurs within them, while it does occur between these units with 50% frequency. The hypothetical number of these units are known as the effective number of factors, governing the inheritance of the particular character. /

Wright (1934), Goodwin (1944) and others worked out estimates of the number of factors on the assumption that the magnitude of the effects of all the factors is the same and further that allelomorphs are distributed indirectionally. They estimated the number by the ratio

$$\frac{(\text{parental difference})^2}{\text{Genotypic variance in } F_2}$$

The assumption involved is hardly justifiable, since the effect of incomplete

concentration of like allelomorphs will be to reduce the parental difference, thereby underestimating the number. The inequality of magnitude of all the factors further lowers the estimate. Panse (1940 a) derived another estimate given by the ratio

$$\frac{(\text{Mean genotypic variance within } F_3 \text{ progeny})^2}{\text{Variance of genotypic variance within } F_3 \text{ progeny}}$$

which assumes the equality of a certain function of magnitude of effects & of dominance effects. This estimate overcomes the difficulties of incomplete concentration and incomplete reinforcement.

In 1940, Panse introduced in another paper, a technique with a new approach of vital importance particularly for the plant breeding studies. With the use of the F_2 genotypic and environmental variance, observed in a plant breeding experiment and with the help of hypothetical genetic models regarding the number of factors governing the character under study and their dominance relations, he showed how the results of selection under different genetic constitutions can be anticipated, so far as the speed of genetic advance and various F_3 properties are concerned.

The models set up by Panse were based on experimental data on hair-length of F_3 progenies of three cotton crosses. From the analysis of the data it appeared that both genotypic and environmental components of F_2 variance were of the same magnitude and were 1.5 units. The effective number of factors was nearabout three. With these quantities as given, following were the 5 models set up.

1 - An infinite series of factors with no dominance, their magnitudes represented by the series $\sqrt{\frac{3}{2}}, \sqrt{\frac{3}{6}}, \sqrt{\frac{3}{8}}, \ldots$

2 - Three equal factors with no dominance.

3.- Two infinite series of factors with dominance in opposite directions, the magnitudes of factors is represented (for both series) by the series

$\frac{2}{\sqrt{5}}, \frac{2}{5}, \frac{2}{5\sqrt{5}} \ldots$

4 - Three factors with equal variance, one being without dominance and other two having dominance in opposite directions.

5 - Three equal factors, two with dominance in opposite directions and one without dominance, as is in model 4.

In the last model the effective number of factors was $\frac{32}{11}$.

Panse obtained the regressions of five statistical properties of the F_3 on F_2 parental value and thus obtained the mean values for these properties making the selection of the highest 10% of the F_2 population. The five F_3 properties studied were:

(1) Genotypic mean (2) Mean genotypic variance within progeny (3) Variance of the progeny mean (4) Covariance between the progeny mean and genotypic variance within progeny (5) Variance of genotypic variance within progeny.

By this technique, not only the results of selection can be anticipated but also the information regarding the genetic set up, by comparing estimates obtained experimentally with the theoretical values, can be derived.

In 1954, Sokil considered another set of 18 genetic models, with wider applicability. The models were based on F_2 & F_3 data on fibre-length, measured as halo-length, obtained from 16 inter-racial cotton crosses. On the basis of the range of variability observed, models were set up with F_2 genotypic variances of 0.5, 1.5 and 2.5 units, the environmental variance being 2.0 units in each case. The number of factors taken to be 2, 8 and 32 were thought of to cover a sufficiently wide and useful range of the effective number of factors, representing situations involving a small, a moderate and a large number of such factors. There were thus nine combinations of F_2 variability and effective number of factors. Out of 18 models, 9 were studied with no dominance and other 9 with full dominance.

In all the genetic models, so far considered by Panse and Sokil, the factors were assumed to be independently segregating. Studies on the linkage among factors are of a great importance to the plant breeder and

geneticist alike. The present investigation is based on the study of the various genetic models on linked factors, following the technique developed by Fonse, so as to see in what manner the linkage affects the speed of genetic advance and the various F_3 statistical properties, after making selection in F_2.

Before finishing with the review, it is worthwhile to refer to Mather's book on Biometrical Genetics (1949) already referred to, which has brought together the various types of studies in the genetics of quantitative characters in a most compact and valuable form. Also a comprehensive account of various developments in statistical application to genetics, with their full mathematical background, giving the basic statistical concepts and tools which the genetic research worker needs, has recently been given by Kempthorne (1957) in his book "An introduction to Genetic Statistics".

III. EXPERIMENTAL MATERIAL AND MODELS SET UP

3.1. It is necessary to point out in the beginning that following Fisher (1930), the terms 'genetic and 'genotypic' are used in different senses to indicate different quantities. Accordingly, the 'genetic variance' in a population, stands for the part of the variation arising from a strictly additive action of genes, while the 'genotypic variance' also includes, besides the genetic variance, the variation arising from non-additive interaction of genes such as dominance and epistacy. By using the regression method proposed by Panse (1940 a), in the presence of dominance the component of variation in F_2 estimated by the regression can be shown to be a quantity intermediate between those representing the genetic and genotypic variances. This was the basis used by Panse and then Sokil, on which the magnitudes of the genotypic variance were assumed in different models making use of experimental data on the character to be studied.

3.2 In the models set up by Panse, the ratio of genotypic variance to the environmental one was assumed to be constant viz. 1:1 and their common value 1.5 units was taken to be the same for all the models. Sokil (1954) considered the effect of unsymmetrical dominance, constructing models with factors showing dominance only in one direction, this being an extreme case of dominance. A more general situation is where the polygenes exhibiting dominance in opposite directions do not balance it completely.

In setting up his models, Sokil selected three different values of the genotypic variance, namely .5, 1.5 and 2.5 units, the environmental variance being 2.0 units in each case. Consideration of these values was an improvement over the one considered by Panse, as these values cover a very useful range of the ratio of genotypic variance to the environmental one, namely $\frac{1}{4}$ to $\frac{5}{4}$ and as such the models are useful in giving results that can be compared with a wide range of data concerning various quantitative characters that may be obtained experimentally.

3.3. The models studied by Panse took into consideration the effects of variation in the number of factors under the condition that the effective number of factors was three which again was based on experimental data. Owing to this restriction the infinite geometric series of factors comes out to be rapidly convergent and only the first few factors are of large magnitude in comparison to the rest and consequently account for most of the effects observed in segregation. This explains to some extent the small difference between the results for models consisting respectively of only three factors and of an infinite series of factors. In the light of this, Sokil in his paper has dealt with a range of values for the effective number of factors in his 18 models covering the values 2, 8 and 32. This has brought out the effect of variation in number of factors on the genetic advance and F_3 statistical properties more clearly.

The factors considered by Panse and then by Sokil in their models, were taken to be independently segregating whereas the present study is concerned with models involving linked factors so as to study the effect of linkage of various degree on the speed of genetic advance and various F_3 statistical properties, after making 10% selection in F_2. The phase of the linkage has been taken upto be 'coupling' in this work.

3.4. In the models, which have been dealt in this work, the dominance relation among the alleles of a factor is supposed to be completely absent. This is because the algebra which tends to become very cumbersome in dealing with linkage is simplified to some extent in the absence of dominance.

The values of genotypic and environmental variances taken to be 1.5 and 2.0 units respectively in the present dissertation, have been borrowed from Sokil's set of values for the construction of the models dealt here. It has been done for comparison of the results with Sokil's results for models with independent factors.

3.5. Actually the present models can be divided into two classes. The first class deals with the models, where the genotypic variance has been kept fixed to 1.5 units and the magnitudes of the various factors involved is calculated from this, depending upon the amount of linkage being considered in a particular model. In the models belonging to second class the magnitudes of various factors have been calculated on the assumption that they were independently acting without any linkage and the F_2 genotypic variance was 1.5 units. With these fixed magnitudes of the factors, various F_3 statistical properties and the genetic advance have been calculated making 10% selection of F_2 individuals. The factors which have previously been taken to be independent, are now taken as linked in pairs keeping their magnitudes unchanged for the second class of models. The main point of difference between the two classes is, that in the first class the magnitude of various factors changes with the change in the linkage values of the various linked pairs, the F_2 genotypic variance being fixed throughout to 1.5 units, whereas in second class the magnitude of the factors is kept fixed, the genotypic variance being changeable with change in linkage values. It may be pointed out that the factors are kept eq ual in their magnitudes.

The description of the various models is given below:

Models belonging to first class:

(a) Two linked factors equal in magnitudes, producing F_2 genotypic variance equal to 1.5 units. The different values of the recombination fraction denoted by 'p', which have been considered in the models are 0, 0.1, 0.2, 0.25, 0.40, 0.45 and 0.5.

(b) Two independent pairs of linked factors (4 factors in all) producing F_2 genotypic variance equal to 1.5 units.
The different pairs of recombination values say p and p' which have been considered in the models are (p=p'=0), (p=p'=0.2), (p=0.2, p'=0.4), (p=p'=0.4) and (p=p'=0.5).

Models belonging to second class:

(a) The character under study is governed by two linked factors, the genotypic value of each of them being equal to $\sqrt{1.5} = 1.2247432$ units. The various recombination values considered are same as in models (a) of first class.

(b) The character is governed by two independent pairs of linked factors and the magnitude of each factor being equal to $\sqrt{.75} = .8660254$ units. The various pairs of values of p & p' are same as are considered in models (b) of first class.

It may be pointed out that in the present investigation, the various values of p and p' are so chosen as to cover tight, intermediate and loose linkages thereby giving a closer picture of the effect of linkage on the results of selection.

As mentioned earlier, the environmental variance has been kept constant for all the models and is equal to two units. The selection intensity has been fixed at 10% in all the above models.

3.6. Among other assumptions involved in the present investigation is one that the factors are additive in action. It is justified on the consideration that the action of genes can be made additive approximately by suitable transformation of the scale and even if there is small deviation from additivity, it would show only effects similar to those of environments (Fisher, 1918). The effects of interaction of genes at the same locus, namely dominance have been taken to be completely absent as stated earlier. The genotypic values of various factors involved are assumed to be equal. Another assumption, common to most of the investigations in the genetics of quantitative characters and also made in the present work, is that of absence of fertility or viability disturbances. Again, the phase of the linkage among various pairs of linked factors has been taken to be 'coupling' in the present work whereas for repulsion phase, one has to repeat the whole process afresh.

IV. MOMENT GENERATING FUNCTION FOR THE JOINT DISTRIBUTION OF THE WHOLE F_2

4.1. Starting conveniently with a pair of linked factors, say A-a and B-b, the cross AB/ab X AB/ab gives ten different possible genotypes in F_2. Let p be the recombinant fraction which shows the amount of linkage between the two factors. Let X_1, X_2 and X_3 respectively denote the F_2 phenotypic value, F_3-genotypic variance and F_3-genotypic mean for a particular genotype in F_2.

Let d_a and d_b (measured from corresponding mid-parent values) be the genotypic values of the factors A-a and B-b such that in the absence of dominance the contributions of the genotypes AA, Aa and aa are d_a, 0 and $-d_a$ respectively and of BB, Bb and bb are d_b, 0 and $-d_b$ respectively. Following is the table giving various values of the three genotypic properties, namely X_1, X_2 and X_3 as mentioned above for the ten genotypes.

Table 4.1: Values of X_1, X_2 and X_3 for various F_2 genotypes from cross AB/ab X AB/ab

S.No.	Genotype	Frequency	X_1 (F_2-phenotypic value)	X_2 (F_3-genotypic variance)	X_3 (F_3-genotypic mean)
1.	AB/Ab	$q^2/4$	d_a+d_b+e	0	d_a+d_b
2.	Ab/Ab	$pq/2$	d_a+e	V_b	d_a
3.	Ab/Ab	$p^2/4$	d_a-d_b+e	0	d_a-d_b
4.	AB/aB	$pq/2$	d_b+e	V_a	d_b
5.	AB/ab (c)	$q^2/2$	e	V_c	0
6.	Ab/aB (R)	$p^2/2$	e	V_R	0
7.	Ab/ab	$pq/2$	$-d_b+e$	V_a	$-d_b$
8.	aB/aB	$p^2/4$	$-d_a-d_b+e$	0	$-d_a+d_b$
9.	aB/ab	$pq/2$	$-d_a+e$	V_b	$-d_a$
10.	ab/ab	$q^2/4$	$-d_a-d_b+e$	0	$-d_a-d_b$
Sum of the proportions:			0	$\frac{1}{2}D_a$	0

Where (c) and (R) denote coupling and repulsion phases respectively, being the contribution due to environment to the phenotypic value and q being equal to (1-p). Also $V_a = \frac{d_a^2}{2}$, $V_b = \frac{d_b^2}{2}$, $V_c = \frac{d_a^2+d_b^2}{2} + d_a d_b(1-2p)$, $V_R = \frac{d_a^2+d_b^2}{2} + d_a d_b(1-2q)$ and $D_c = d_a^2 + d_b^2 + 2(1-2p)^2 d_a d_b$.

It is convenient to treat the values for the individual genotypes as deviations from the corresponding F_2 means, in order to obtain the moments about the mean directly and thus saving troublesome adjustments later for transforming the moments about the origin into those about the mean. Table 4.2 shows these deviations denoted as small x_1, x_2 and x_3.

Table 4.2: Deviations of values of x_1, x_2 and x_3 from their means

S.No.	Genotype	Frequency	x_1 (F_2-phenotypic value)	x_2 (F_2-genotypic variance)	x_3 (F_2-genotypic mean)
1.	AB/AB	$q^2/4$	d_a+d_b+e	$-\frac{1}{4}D_c$	d_a+d_b
2.	AB/Ab	$pq/2$	d_a+e	$V_b-\frac{1}{4}D_c$	d_a
3.	Ab/Ab	$p^2/4$	d_a-d_b+e	$-\frac{1}{4}D_c$	d_a-d_b
4.	AB/aB	$pq/2$	d_b+e	$V_a-\frac{1}{4}D_c$	d_b
5.	AB/ab(c)	$q^2/2$	e	$V_c-\frac{1}{4}D_c$	0
6.	Ab/aB(R)	$p^2/2$	e	$V_R-\frac{1}{4}D_c$	0
7.	Ab/ab	$pq/2$	$-d_b+e$	$V_a-\frac{1}{4}D_c$	$-d_b$
8.	aB/aB	$p^2/4$	$-d_a+d_b+e$	$-\frac{1}{4}D_c$	$-d_a+d_b$
9.	aB/ab	$pq/2$	$-d_a+e$	$V_b-\frac{1}{4}D_c$	$-d_a$
10.	ab/ab	$q^2/4$	$-d_a-d_b+e$	$-\frac{1}{4}D_c$	$-d_a-d_b$

4.2. Consider now the distribution of each of the three variates X_1, X_2 and X_3 in all the ten genotypes seperately. The phenotypic value X_1 can be considered to be normally distributed with an environmental variance'v'about the corresponding genotypic means, given in table 4.1 for various genotypes. The other two variates X_2 and X_3 assume constant values, since within a genotype all values remain constant. Since for a normal distribution $K_1 = u_1$, $K_2 = u_2$ and K_3 and higher = 0, the cumulants of the distributions of X_1, X_2 and X_3 can be written for every genotype seperately. They have been shown in table 4.3 in the last column of which, the joint cumulant function $K_{x_1x_2x_3}(t)$ for all the three variables has been given for all genotypes seperately.

Table 6.3: Cumulants for distributions of variables X_1, X_2 and X_3 and their joint cumulant function for different genotypes.

No.	Genotype	Frequency	K	X_1 (F_2-genotypic value)	X_2 (F_3 genotypic variance)	X_3 (F_3 genotypic mean)	$K_{x_1x_2x_3}(t)$
1.	AB/AB	$q^2/4$	K_1	d_a+d_b	$-\frac{1}{2}D_c$	d_a+d_b	$(d_a+d_b)(t_1+t_3)$
			K_2	v	0	0	$+(-D_c/4)t_2+v\frac{t_1^2}{2}$.
			K_3 & higher	0	0	0	
2.	AB/Ab	$pq/2$	K_1	d_a	$V_b-\frac{1}{2}D_c$	d_a	$d_a(t_1+t_3)+(V_b-D_c/4)t_2$
			K_2	v	0	0	$+v\frac{t_1^2}{2}$.
			K_3 & higher	0	0	0	
3.	Ab/Ab	$p^2/4$	K_1	d_a-d_b	$-\frac{1}{2}D_c$	d_a-d_b	$(d_a-d_b)(t_1+t_3)-$
			K_2	v	0	0	$\frac{1}{2}D_c t_2+v\frac{t_1^2}{2}$.
			K_3 & higher	0	0	0	
4.	AB/aB	$pq/2$	K_1	d_b	$V_a-\frac{1}{2}D_c$	d_b	$d_b(t_1+t_3)+(V_a-\frac{1}{2}D_c)t_2$
			K_2	v	0	0	$+v\frac{t_1^2}{2}$.
			K_3 & higher	0	0	0	
5.	AA/ab(c)	$q^2/2$	K_1	0	$V_c-\frac{1}{2}D_c$	0	$(V_c-\frac{1}{2}D_c)t_2+v\frac{t_1^2}{2}$.
			K_2	v	0	0	
			K_3 & higher	0	0	0	
6.	Ab/aB(R)	$p^2/2$	K_1	0	$V_R-\frac{1}{2}D_c$	0	$(V_R-\frac{1}{2}D_c)t_3+v\frac{t_1^2}{2}$.
			K_2	v	0	0	
			K_3 & higher	0	0	0	
7.	Ab/ab	$pq/2$	K_1	$-d_b$	$V_a-\frac{1}{2}D_c$	$-d_b$	$-d_b(t_1+t_3)+(V_a-\frac{1}{2}D_c)t_2$
			K_2	v	0	0	$+v\frac{t_1^2}{2}$.
			K_3 & higher	0	0	0	
8.	aB/aB	$p^2/4$	K_1	$-d_a+d_b$	$-\frac{1}{2}D_c$	$-d_a+d_b$	$(-d_a+d_b)(t_1+t_3)+$
			K_2	v	0	0	$(-\frac{1}{2}D_c)t_2+v\frac{t_1^2}{2}$.
			K_3 & higher	0	0	0	
9.	aB/ab	$pq/2$	K_1	$-d_a$	$V_b-\frac{1}{2}D_c$	$-d_a$	$-d_a(t_1+t_3)+(V_b-\frac{1}{2}D_c)t_2$
			K_2	v	0	0	$+v\frac{t_1^2}{2}$.
			K_3 & higher	0	0	0	
10.	ab/ab	$q^2/4$	K_1	$-d_a-d_b$	$-\frac{1}{2}D_c$	$-d_a-d_b$	$(-d_a-d_b)(t_1+t_3)$
			K_2	v	0	0	$-\frac{1}{2}D_c t_2+v\frac{t_1^2}{2}$.
			K_3 & higher	0	0	0	

t_1, t_2, and t_3 are the arbitrary quantities involved in the cumulant function $K_{x_1x_2x_3}(t)$.

4.3. The joint moment generating functions for the variables X_1, X_2 and X_3 corresponding to various genotypes are given by the 'e' raised to the power, by the expressions given in the last column of the table 4.3. The moment generating function for the joint distribution over the whole F_2 is the sum of the corresponding functions for the ten genotypes multiplied by the relative frequencies of the genotypes. It is thus given by

$$e^{v\frac{t_1}{2}^2}\cdot\Big\{\frac{q^2}{4}\cdot e^{(d_a+d_b)(t_1+t_2)-\frac{1}{2}D_c t_2}+\frac{pq}{2}\cdot e^{d_a(t_1+t_3)+(V_b-\frac{1}{2}D_c)t_2}+$$

$$\frac{p^2}{4}\cdot e^{(d_a-d_b)(t_1+t_3)-\frac{1}{2}D_c t_2}+\frac{pq}{2}\cdot e^{d_b(t_1+t_3)+(V_a-\frac{1}{2}D_c)t_3}+$$

$$\frac{q^2}{2}\cdot e^{(V_b-\frac{1}{2}D_c)t_2}+\frac{p^2}{2}\cdot e^{(V_R-\frac{1}{2}D_c)t_2}+\frac{pq}{2}\cdot e^{-(d_b)(t_1+t_3)+(V_a+\frac{1}{2}D_c)t_2}+$$

$$\frac{p^2}{4}\cdot e^{(-d_a+d_b)(t_1+t_3)-\frac{1}{2}D_c t_2}+\frac{pq}{2}e^{-d_a(t_1+t_3)+(V_b-\frac{1}{2}D_c)t_2}+$$

$$\frac{q^2}{4}\cdot e^{(-d_a-d_b)(t_1+t_3)-\frac{1}{2}D_c t_2}\Big\}.$$

On expansion of various exponential terms, in the bracket of this expression and on simplification it reduces to

$$e^{v\frac{t_1}{2}^2}\cdot\Big\{\frac{q^2}{4}(1-\tfrac{1}{2}D_c t_2+\frac{D_a^2 t_2^2}{16\,2!}-\ldots)(1+\frac{(d_a+d_b)^2}{2!}(t_1+t_3)^2+\frac{(d_a+d_b)^4(t_1+t_3)^4}{4!}+\ldots)+$$

$$pq(1+\frac{d_a^2(t_1+t_3)^2}{2!}+\ldots)\Big(1+t_2(V_b-\tfrac{1}{2}D_c)+\frac{(V_b-\frac{1}{2}D_c)^2t_2^2}{2!}+\ldots\Big)+$$

$$\frac{p^2}{4}(1-\tfrac{1}{2}D_c t_2+\frac{D_a^2t_2^2}{16\,2}+\ldots)\Big(1+\frac{(d_a-d_b)^2(t_1+t_3)^2}{2!}+\ldots\Big)+$$

$$pq\Big\{1+(V_a-\tfrac{1}{2}D_c)t_2+(V_a-\tfrac{1}{2}D_c)^2t_2^2+\ldots\Big\}\Big(1+\frac{d_b^2(t_1+t_3)^2}{2!}+\ldots\Big)+$$

$$\frac{q^2}{2}\Big\{1+(V_b-\tfrac{1}{2}D_c)t_2+(V_b-\tfrac{1}{2}D_c)^2t_2^2+\ldots\Big\}+\frac{p^2}{2}\Big\{1+(V_R-\tfrac{1}{2}D_c)t_2+(V_R-\tfrac{1}{2}D_c)^2t_2^2+\ldots\Big\}\Big]$$

This is the moment generating function for a model consisting of a pair of linked factors, the moments and product-moments being given by coefficients of the various terms in t_1, t_2 and t_3. On further expansion, the expression in the bracket can be put in a tabular form showing the coefficients of the powers of (t_1+t_3) on the one hand and the coefficients of the powers of t_2 on the other. This has been given in table 4.4. The various terms in each cell of the table are so lengthy that it is very difficult to put them to further algebraic treatment. To avoid this the terms have been denoted by the letters (A_{ij}) as shown in each cell.

4.4. Using the notations A_{ij}'s, the moment generating function given in table 4.4 can be represented by

$$e^{v\frac{t_1^2}{2}}\Big\{1+A_{20}(t_1+t_3)^2+A_{40}(t_1+t_3)^4+\ldots\ldots$$
$$+A_{21}(t_1+t_3)^2t_2+A_{41}(t_1+t_3)^4t_2+\ldots\ldots$$
$$+A_{02}t_2^2+A_{22}(t_1+t_3)^2t_2^2+A_{42}(t_1+t_3)^4t_2^2+\ldots\ldots\Big\}.$$

The corresponding cumulant function for the joint distribution over the whole F_3 is obviously equal to

$$v\frac{t_1^2}{2}+\log\Big\{1+A_{20}(t_1+t_3)^2+\ldots+A_{21}(t_1+t_3)^2t_2+\ldots+A_{02}t_2^2+A_{22}(t_1+t_3)^2t_2^2+\ldots\Big\}.$$

The term in the bracket can be expanded as $\log(1+x)$ giving the cumulant function in terms of the powers of t_2 and (t_1+t_3) which has been presented in table 4.5, in a form similar to table 4.4.

The cumulant $K_{p,q,r}$ is given by the coefficient of the term $\frac{t_1^p.t_2^q.t_3^r}{p!\,q!\,r!}$ in the expansion of the cumulant function given in table 4.5.

Table 4.4: Moment generating function for pair of linked factors (Linked in coupling phase)

Powers of $(t_1 + t_3)$

Powers of t_2	0	1	2	3	4	5	6	7	8
0	1	0	$\left\{\left(\frac{d_a^2+d_b^2}{4}+\frac{d_ad_b}{4}(1-2p)\right)\right\}$	0	$\left\{\left(\frac{d_a^4+d_b^4}{48}+\frac{d_ad_b}{12}\right)(d_a^2+d_b^2)(1-2p)+\frac{d_a^2d_b^2}{8}(1-2p+2p^2)\right\}$	0	$\left\{\left(\frac{d_a^6+d_b^6}{1440}+\frac{d_ad_b}{720}\right)(3d_a^4+3d_b^4+10d_a^2d_b^2)(1-2p)+\frac{1}{96}d_a^2d_b^2(d_a^2+d_b^2)(1-2p+2p^2)\right\}$	0	$\left\{\left(\frac{d_a^8+d_b^8}{80640}+\frac{d_ad_b}{10080}\right)(d_a^6+7d_a^4d_b^2+7d_a^2d_b^4+d_b^6)(1-2p)+\frac{d_a^2d_b^2}{5760}(24d_a^4+50d_a^2d_b^2+24d_b^4)(1-2p+2p^2)\right\}$
			(A_{20})		(A_{40})		(A_{60})		(A_{80})
1	0	0	$\left\{\left(\frac{-d_a^2}{4}\left(\frac{D_a}{4}-2pqV_b\right)-\frac{d_b^2}{4}\left(\frac{D_a}{4}-2pqV_a\right)-\frac{d_ad_b}{2}\frac{D_a}{4}(1-2p)\right)\right\}$	0	$\left\{\left(\frac{-d_a^4}{48}\left(\frac{D_a}{4}-2pqV_b\right)-\frac{d_b^4}{48}\left(\frac{D_a}{4}-2pqV_a\right)-\frac{d_ad_b}{12}\frac{D_a}{4}(d_a^2+d_b^2)(1-2p)-\frac{d_a^2d_b^2}{8}\frac{D_a}{4}(1-2p+2p^2)\right)\right\}$	0	$\left\{\left(\frac{-d_a^6}{1440}\left(\frac{D_a}{4}-2pqV_b\right)-\frac{d_b^6}{1440}\left(\frac{D_a}{4}-2pqV_a\right)-\frac{d_ad_b}{720}\frac{D_a}{4}(3d_a^6+10d_a^2d_b^2+3d_b^4)(1-2p)-\frac{d_a^2d_b^2}{96}\frac{D_a}{4}(d_a^2+d_b^2)(1-2p+2p^2)\right)\right\}$	0
			(A_{21})		(A_{41})		(A_{61})		
2	$\left\{\frac{pq}{2}(V_a^2+V_b^2)+\frac{q^2}{2}+2pqV_aV_b-\frac{D_a^2}{32}\right)$	0	$\left\{\frac{q^2}{8}\left(\frac{D_a}{4}\right)^2(d_a+d_b)^2+\frac{pq}{4}d_a^2\left(V_b-\frac{D_a}{4}\right)^2+\frac{p^2}{8}\left(\frac{D_a}{4}\right)^2(d_a-d_b)^2+\frac{p^2}{4}V_a^2d_b^2\right)$	0	$\left\{\frac{q^2}{96}\left(\frac{D_a}{4}\right)^2(d_a+d_b)^4+\frac{pq}{48}\left(V_b-\frac{D_a}{4}\right)^2d_a^4+\frac{p^2}{96}\left(\frac{D_a}{4}\right)^2(d_a-d_b)^4+\frac{p^2}{4}V_a^2d_b^4\right)$	0	

23

23

Table 4.5 gives the cumulant function of the F_2 distribution for a model consisting of only one pair of linked factors. In case there are more than one independent pairs of linked factors, then the cumulant function of the F_2 distribution for that model will be the sum of cumulant functions of the individual linked pairs. This is only in the case, where the individual linked pairs are independently distributed over different chromosomes. In that case table 4.5 can be modified so as to give a cumulant function in the case where there are a given number of independently linked pairs involved in the model. The terms in each cell of the table will have to be replaced by the sum of the corresponding terms related to different linked pairs, which can be obtained by calculating different values of A_{ij}'s involved for each linked pair seperately. In table 4.5, the letters B_{ij}'s shown in bracket in each cell denote the sum of the corresponding expressions of that cell in A_{ij}'s, over all linked pairs. Thus the modified table 4.5 in B_{ij}'s will denote the cumulant function of the F_2 distribution for the model involving various independent pairs of linked factors. In the case of one linked pair only, B_{ij}'s will denote the various expressions in A_{ij}'s of the corresponding cells as they are.

4.8. While using Panse's technique as mentioned earlier one has to fit regression lines of F_3 characters on F_2 - parental values X_1. This makes use of the various moments and product moments of the F_2 distribution jointly found over all the independent linked pairs involved. It thus becomes necessary to transform the modified table of B_{ij}'s into a table giving the moment generating function directly in terms of the coefficients of various terms in t_1, t_2 and t_3 from which any moment or product moment can be readily found for use. Table 4.6 shows these coefficients in terms of B_{ij}'s. It may be borne in mind that the notations A_{ij}'s and B_{ij}'s had to be used only to avoid the lengthy algebraic expressions and to shorten the various algebraic operations involved.

Table 4.S_2 Expansion of cumulant function for pair of linked factors

Powers of t_2 \ Powers of (t_1+t_3)	0	1	2	3	4	5	6	7	8
0	0	0	Δ_{20} (B_{20})	0	$\Delta_{40}=\frac{\Delta_{20}^2}{2}$ (B_{40})	0	$\Delta_{60}=\Delta_{20}\Delta_{40}+\frac{\Delta_{20}^3}{3}$ (B_{60})	0	$\Delta_{80}=\Delta_{20}\Delta_{60}-\frac{\Delta_{40}^2}{2}-\frac{\Delta_{20}^4}{4}+\Delta_{20}^2\Delta_{40}$ (B_{80})
1	0	0	Δ_{21} (B_{21})	0	$\Delta_{41}=\Delta_{20}\Delta_{21}$ (B_{41})	0	$\Delta_{61}=\Delta_{20}\Delta_{41}+\Delta_{40}\Delta_{21}+\Delta_{20}^2\Delta_{21}$ (B_{61})	0	$+\frac{\Delta_{21}^2}{2}$
2	Δ_{02} (B_{02})	0	$\Delta_{22}=\Delta_{20}\Delta_{02}$ (B_{22})	0	$\Delta_{42}=\Delta_{20}\Delta_{22}+\Delta_{40}\Delta_{02}+\frac{\Delta_{21}^2}{2}+\Delta_{20}^2\Delta_{02}$ (B_{42})	0			

Table 4.6: Moment generating Function for joint distribution of whole F_2 for three statistical properties

Powers of t_1	Powers of t_2 and t_3: 0	t_2	t_2^2	t_3	t_2t_3	t_3^2
0	1	0	B_{02}	0	0	B_{20}
1	0	0	0	$2B_{20}$	$2B_{21}$	0
2	$B_{20}+\frac{V}{2}$	B_{21}	$B_{22}+\frac{V}{2}B_{02}+B_{20}B_{02}$	0	0	$6B_{40}+\frac{V}{2}B_{20}+3B_{20}^2$
3	0	0	0	$3B_{40}+2B_{20}^2+B_{20}V$	$4(B_{41}+B_{20}B_{21})+B_{21}V$	0
4	$(B_{40}+\frac{B_{20}^2}{2})+\frac{B_{20}}{2}V+\frac{V^2}{8}$	$(B_{41}+B_{20}B_{21})+\frac{B_{21}}{2}V$	$(B_{42}+\frac{B_{21}^2}{2}+B_{20}B_{22}+B_{40}B_{02}+\frac{B_{20}^2B_{02}}{2})+(B_{23}+B_{20}B_{02})\frac{V}{2}+\frac{B_{02}}{8}V^2$	0	0	$15(B_{60}+B_{20}B_{40}+\frac{B_{20}^3}{6})+(6B_{40}+3B_{20}^2)\frac{V}{2}+\frac{B_{20}}{2}V^2$
5	0					
6	$(B_{60}+B_{20}B_{40}+\frac{B_{20}^3}{6})+(B_{40}+\frac{B_{20}^2}{2})\frac{V}{2}+\frac{B_{20}}{8}V^2+V^3/48$					
7	0					
8	$(B_{80}+\frac{B_{40}^2}{2}+B_{20}B_{60}+\frac{B_{20}^2B_{40}}{2}+\frac{B_{20}^4}{24})+(B_{60}+B_{20}B_{40}+\frac{B_{20}^3}{6})\frac{V}{2}+(B_{40}+\frac{B_{20}^2}{2})\frac{V^2}{8}+\frac{B_{20}}{48}V^3+V^4/384$					

V. CONSIDERATION OF MODELS WITH ONE PAIR OF LINKED FACTORS

5.1. We now consider models with a single pair of linked factors. These models are being dealt with in the present chapter in view of the fact that for the models involving one pair of linked factors it is possible to enumerate all the ten possible genotypes and calculate the results of selection directly. Not only less computations are involved in this method of enumeration but also it is more accurate owing to fewer approximations involved as compared to the general regression method given by Panse.

Denoting the factors involved by A-a and B-b, the ten genotypes and their relative frequencies in terms of p and q ($=1-p$) in the F_2 can be set down as follows:

AB/AB	$q^2/4$	Ab/aB(R)	$p^2/2$
AB/Ab	$pq/2$	Ab/ab	$pq/2$
Ab/Ab	$p^2/4$	aB/aB	$p^2/4$
AB/aB	$pq/2$	aB/ab	$pq/2$
AB/ab(c)	$q^2/2$	ab/ab	$q^2/4$

In all the models discussed in the present paper, the genotypic values of the various factors are assumed to be equal. Hereafter their common genotypic value will be denoted by the letter 'a' so that

$d_a = d_b = \cdots = a$.

In the models belonging to class II, this value of a is given whereas in the case of class I this can be calculated with the help of the given value of the F_2 genotypic variance. With the knowledge of 'a', the genotypic value of the F_3 progeny mean and the genotypic variance within progeny are known for each of the ten genotypes. The phenotypic values of these genotypes will be scattered about their genotypic values with variance represented by the environmental component 'v' of F_2 variance. Let ξ be the deviate of the F_2 distribution for the required level of selection

i.e. 10%. Corresponding to each model one can calculate Σ using the procedure given in the next paragraph. Knowing Σ and hence $\Sigma V^{\frac{1}{2}}$, the actual deviate beyond which the F_2 individuals are selected in different cases, the proportions of different genotypes in the selected population can be obtained and from these the mean values for various properties can be easily calculated; thereby giving directly the results of selection. V, as used here, is total F_2 variance i.e. the sum of the F_2 genotypic variance and the environmental variance.

3.2. The deviate Σ, as mentioned above is expressible in terms of the normal deviate 'x' corresponding to the same level of selection as that of Σ and coefficients a, b, c, d, e, etc which depend upon the cumulants K_1, K_2, etc, of the distribution under study, by the relations $K_1 = a$, $K_2 = V = bV$, $K_3 = cV^{3/2}$ etc, a and V being respectively the mean and variance of the normal distribution (Cornish & Fisher, 1937). The various cumulants K_1, K_2, ... can be found from the table 4.5, for different models. For convenience the values of a and V have been taken to be K_1(=0) and K_2 respectively, which reduces the expression for Σ to give

$$= x + \frac{c}{6}(x^2-1) + \frac{d}{24}(x^3-3x) - \frac{c^2}{36}(2x^3-5x) + \frac{e}{120}(x^4-6x^2+3)$$
$$- \frac{cd}{24}(x^4-5x^2+2) + \frac{c^3}{324}(12x^4-53x^2+17) + \frac{f}{720}(x^5-10x^3+15x) -$$
$$\frac{d^2}{384}(8x^5-34x^3+29x) - \frac{de}{180}(2x^5-17x^3+21x) + \frac{c^2d}{288}(14x^5-103x^3+107x)$$
$$+ \frac{c^4}{7776}(352x^5-1638x^3+1511x).$$

In our models there being no dominance, the coefficients c and e are zero and hence on substituting the value of the normal deviate x equal to 1.28155 for 10% level of selection, the expression for Σ reduces to

$$= 1.28155 - .07249d + .00227f + .00776d^2$$

By substituting the numerical values of the coefficients d and f corresponding to each model, the values of Σ have been obtained for all models. The various values of Σ for the models with one linked pair are given in table 5.1.

Table 5.1: Values of Σ for class I and II models with one linked pair of factors.

S.No.	Value of recombinant fraction 'p'	for class I models	for class II models
1.	0.0	1.2953175	1.3086019
2.	0.10	1.2943421	1.3049571
3.	0.20	1.2935732	1.3010755
4.	0.25	1.2929897	1.2990696
5.	0.40	1.2906579	1.2927485
6.	0.45	1.2895957	1.2905066
7.	0.50	1.2883205	1.2883205

5.3. For illustration of the enumeration method, let us take the model with p = .25 from amongst models of class I. The F_2 genotypic variance $H^V F_2$ is given to be equal to 1.5 units. Again $H^V F_2 = \frac{1}{2}(d_a^2 + d_b^2 + 2(1-2p)d_a d_b)$ in the absence of dominance. For $d_a = d_b = a$, $H^V F_2 = 2a^2(1+p)$ which gives the genotypic value of a factor $= \sqrt{\frac{H^V F_2}{2(1+p)}}$. On substituting the values of $H^V F_2 = 1.5$ and of $p = .25$, we have $a = 1.0$ unit. The F_2 genotypic values of the various genotypes can then be known. For instance the value for the genotype AB/AB is 2.0 units and so on. The actual deviate for 10% selection of the highest phenotypes is $\Sigma V^{\frac{1}{2}} = 1.2929897 \times \sqrt{5.5} = 2.4189618$, measured from the F_2 mean which is zero here, on the scale of genotypic values. To find the proportion of the genotype AB/AB, which is selected, we calculate

the normal deviate given by $\frac{(2.4189613-2.0000000)}{\sqrt{2}}$, the factor $\sqrt{2}$ in the denominator being the square root of the value of the environmental component of the F_2 - variance assumed i.e. 'v'.

The normal deviate = $\frac{.4189613}{\sqrt{3}}$ = .296250

By entering the table of the normal probability integral (Tables for Statisticians & Biometricians by K. Pearson, 1930) against the above value of the normal deviate and using interpolation it is found that (1-.6164791) = .3835209 of the individuals belonging to this genotype are selected. This constitutes $\frac{9}{64}$ (.3835209) = .0539326 of the whole population, 9/64 being equal to $q^2/4$, the frequency corresponding to the genotype AB/AB. Similarly the proportions of the other genotypes can be found out. These proportions total upto .1001034 which is quite close to the fraction selected as might be expected. Making use of these proportions, we can calculate the genetic advance and various other F_2- properties.

Various results of selection have been presented in tables 5.2 and 5.3 separately for models of class I and of class II with one pair of linked factors. As stated already, the value of F_2- genotypic variance $H^V F_2$ is fixed at 1.5 units for class I models whereas for class II models it is changing. Various values of $H^V F_2$ for latter models, found with the help of fixed magnitude of the factors and the various corresponding values of the recombination fraction 'p', have been shown in the last column of the table 5.3.

Table 5.3: Results of selection for models of class I with one pair of linked factors with $H^V F_2$ = 1.5 units and environmental variance 'v' = 2.0 units.

S.No. of the model	(1) 'p'	(2) Limit of selective advance	(3) Advance achieved in F_3	(4) F_3 advance as proportion of that possible	(5) $\bar{H}^V F_3$	(6) $H^V \bar{F}_3$	(7) $H^V \bar{F}_3 / V_{F_3}$	(8) $V_{V_{F_3}}$	(9) $\bar{H}^V F_3$ without selection
1.	0.0	1.7320309	1.3434363	77.56%	.8241953	.5467832	=.4355360	.3811906	.7500000
2.	0.10	1.8257418	1.3476542	73.81%	.8205916	.5758577	=.3863153	.3159045	.6833333
3.	0.20	1.9364916	1.3529121	69.86%	.8282071	.6109903	=.3469388	.2602169	.6375000
4.	0.25	2.0000000	1.3560598	67.80%	.8456576	.6314731	=.3282207	.2370561	.6250000
5.	0.40	2.2360680	1.3686474	61.21%	.8321522	.7093118	=.3003501	.2001033	.6500000
6.	0.45	2.3354963	1.3743758	58.85%	.8334737	.7427841	=.3026449	.2073725	.6836363
7.	0.50	2.4494864	1.3812731	56.39%	.8517576	.7815837	=.3124799	.2333560	.7500000

Table 5.3: Results of selection for models of class II with one pair of linked factors with genotypic value's'= 1.3247432 units fixed for each factor & environmental variance 'v' = 2.0 units.

S.No. of the model	(1) 'p'	(2) Limit of selective advance	(3) Advance achieved in F_3	(4) F_3 advance as proportion of that possible	(5) $\bar{H}^V F_3$	(6) $H^V \bar{F}_3$	(7) $H^V \bar{F}_3 / V_{F_3}$	(8) $V_{V_{F_3}}$	(9) $H^V F_2$ before selection	(10) $\bar{H}^V F_3$ without selection
1.	0.0	2.4494864	2.2164781	90.49%	.2343137	.5185973	=.6301751	.7721069	3.00	1.5000
2.	0.10	2.4494864	2.0780770	84.63%	.2334496	.6196709	=.5380246	.6183273	2.70	1.2300
3.	0.20	2.4494864	1.9180741	78.31%	.2326561	.6971651	=.5276657	.4780653	2.40	1.0300
4.	0.25	2.4494864	1.8363418	74.97%	.2075316	.7266237	=.4925061	.4150962	2.25	0.9375
5.	0.40	2.4494864	1.5794821	64.24%	.2376612	.7774751	=.3300445	.2711416	1.80	0.7800
6.	0.45	2.4494864	1.4792813	60.39%	.2180918	.7824912	=.3446146	.2646878	1.69	0.7575
7.	0.50	2.4494864	1.3812731	56.39%	.2517576	.7815837	=.3124799	.2333560	1.50	0.7500

5.4. To extract more information and to make the results of tables 5.2 and 5.3 handy, they have been transformed to the basis of unit F_2- genotypic variance and are put in tables 5.4 and 5.5 respectively. This transformation is done by dividing the results for (1) Genetic advance achieved (2) Mean genotypic variance within progeny ($H^V F_2$), (3) variance of progeny mean ($H^V \overline{F}_2$), (4) covariance between the progeny mean and genotypic variance within progeny ($H^V \overline{F}_2 / V_{F_2}$) and (5) variance of genotypic variance within progeny ($V_{V_{F_2}}$), by $\sqrt{H^V F_2}$, $H^V F_2$, $H\sqrt{V_{F_2}}$, $(H^V F_2)^{3/2}$ and $(H^V F_2)^2$ respectively. The value of the F_2 genotypic variance may change from model to model particularly in the case of models where genotypic value of the factor is kept fixed. These transformations should enable the various tables to be used more easily with other genetical data, since if $H^V F_2 = b$ the various columns of the tables can be multiplied by the factors shown in their bottom rows to give the values appropriate for the cases under consideration. Also they can be used to compare with the F_2 values actually observed so as to enable the plant breeder to fit a suitable genetic model to his situation, if possible.

In the last column of table 5.5, the corresponding changed values of the environmental variance as against the unit value of $H^V F_2$, have also been given which enable us to judge how various results are affected with the simultaneous change in the linkage value and the environmental variance. As stressed earlier, it is not the environmental or the genotypic variance but their ratio which is important.

As a special case table 5.3 can be put in the form obtained after transforming the results to the basis of fixed genotypic value equal to unity, which will change the enviromental variance from 2.0 units to 1.3333333 units for all the models presented in that table. The $H^V F_2$ will naturally be different for different models. The transformed values have

Table 5.4: Results of selection for models of class I (Table 5.2) as transformed to basis of unit F_2 genotypic variance, with transformed value of environmental variance 'v' = 1.3333333 units.

S.No. of the model	(1) 'p'	(2) Limit of selective advance	(3) Advance achieved in F_3	(4) F_3 advance as proportion of that possible	(5) $H^{\bar{V}}F_3$	(6) $H^{V_{\bar{F}}}F_3$	(7) $H^{U_{\bar{F}}}F_3/V_{F_3}$	(8) $V_{V_{F_3}}$	(9) $\bar{V}$ H_{F_2} without selection
1.	0.0	1.4142136	1.0969112	77.96%	.3161303	.3645255	-.2370757	.1696180	.5000000
2.	0.10	1.4907120	1.1003351	73.81%	.2137277	.2839051	-.2102834	.1404020	.6333336
3.	0.20	1.5311389	1.1046431	69.86%	.2214714	.4078268	-.1877881	.1186518	.4250000
4.	0.25	1.6329933	1.1072182	67.80%	.2304384	.4209981	-.1786607	.1058593	.4166667
5.	0.40	1.8357420	1.1174960	61.21%	.2391015	.4728745	-.1637680	.0989948	.4333333
6.	0.45	1.9069253	1.1221732	58.85%	.2288153	.4951894	-.1647890	.0921656	.4590909
7.	0.50	2.0000000	1.1278043	56.39%	.2673334	.5210218	-.1700925	.1033027	.5000000
Where $H^{V}F_2=b$ multiplied by		$\sqrt{b}$	$\sqrt{b}$	1	b	b	$b^{3/2}$	b^2	b

Table 5.5: Results of selection for models of class II (Table 5.3) as transformed to basis of unit F_2 genotypic variance.

S.No. of the model	(1) 'p'	(2) Limit of selective advance	(3) Advance achieved in F_3	(4) F_3 advance as proportion of that possible	(5) $H^{\bar{V}}F_3$	(6) $H^{V}F_3$	(7) $H^{U_{\bar{F}}}F_3/V_{F_3}$	(8) $V_{V_{F_3}}$	(9) Environmental variance 'v'
1.	0.0	1.4142136	1.2796842	90.49%	.0967718	.1728659	-.1212772	.0357896	0.667
2.	0.10	1.4907120	1.2616345	84.63%	.1334999	.2295077	-.1328409	.0848460	0.741
3.	0.20	1.5311389	1.2331115	78.31%	.1594400	.2904855	-.1419195	.0839974	0.833
4.	0.25	1.6329933	1.2243279	74.97%	.1811474	.3229439	-.1458277	.0819943	0.889
5.	0.40	1.8357420	1.1728043	64.24%	.2709229	.4819306	-.1573714	.0936957	1.111
6.	0.45	1.9069253	1.1516190	60.39%	.3139899	.4742007	-.1625952	.0998761	1.212
7.	0.50	2.0000000	1.1278043	56.39%	.3673334	.5210218	-.1700925	.1033027	1.333
Where $H^{V}F_2=b$ multiplied by		$\sqrt{b}$	$\sqrt{b}$	1	b	b	$b^{3/2}$	b^2	b

been given in table 5.6 with the corresponding changed values of F_2- genotypic variance in various cases. The last row gives the correction factor with which to multiply different values so as to give results of various properties for the models with genotypic value =1= unit, fixed for each factor.

Table 5.6: Results of selection for models of class II (Table 5.3) as transformed to basis of unit genotypic value for each factor, with transformed value of environmental variance 'v' = 1.8333333 units.

S.No. of the model	(1) 'p'	(2) Limit of selective advance	(3) Advance achieved in F_3	(4) F_3 advance as proportion of that possible	(5) $H^{\bar{V}}F_3$	(6) $H^{V}F_3$	(7) $H^{\bar{V}}F_3/V_{F_3}$	(8) $V_{V_{F_3}}$	(9) $H^{V}F_2$ before selection	(10) $H^{\bar{V}}F_3$ without selection
1.	0.0	2.0000000	1.8097493	90.49%	.1893425	.3457315	-.3430243	.3031586	2.0	1.000
2.	0.10	2.0000000	1.6926626	84.63%	.2223997	.4131139	-.3200805	.2749010	1.8	0.820
3.	0.20	2.0000000	1.5661081	78.31%	.2551041	.4647767	-.2872252	.2124735	1.6	0.680
4.	0.25	2.0000000	1.4993688	74.97%	.2717211	.4844153	-.2680367	.1846872	1.5	0.625
5.	0.40	2.0000000	1.2847445	64.24%	.3251075	.5183167	-.2069703	.1205074	1.2	0.520
6.	0.45	2.0000000	1.2073293	60.39%	.3453879	.5216203	-.1875847	.1037501	1.1	0.505
7.	0.50	2.0000000	1.1278068	56.39%	.3678394	.5210218	-.1709923	.1038027	1.0	0.500
Where genotypic values multiplied by		a	a	1	a^3	a^2	a^3	a^4	a^2	a^2

The various results given in the above tables have been discussed in chapter VIII in the light of linkage and other aspects.

It may be mentioned here that the enumeration method has also been used in getting the results in the case of models of class II with two independent pairs of linked factors, apart from the general regression method. It has been done in order to see, how much approximation is involved in the general method particularly in the presence of linkage as compared to the enumeration method. There is no doubt that the results given by the enumeration method are more exact as compared to the general method.

VI. REGRESSION OF F_3 CHARACTERS ON F_2 PARENTAL VALUES FOR THE MODELS INVOLVING TWO INDEPENDENT PAIRS OF LINKED FACTORS.

6.1. It has been mentioned earlier, that in order to use the statistical techniques given by Panse, one has to fit regression lines of F_3 characters on F_2 parental values X_i, which makes use of the various moments and product-moments of the F_3 distribution, jointly found over all the independent pairs of linked factors involved. Table 4.6 gives directly the various moments and product-moments for different models (with no dominance) after knowing the values of various A_{ij}'s and B_{ij}'s involved therein, A_{ij}'s being the function of the linkage values of the two independent pairs of linked factors (p and p', say in the present case) and of the genotypic values d_a, d_b, ---- etc and hence of 'a', since $d_a = d_b = d_c = \text{-----} = a$'a' for all the models.

In class II models, 'a' has a fixed value, whereas in class I models, 'a' can be estimated from the given value of $H^{\prime}F_3$, the F_3 genotypic variance in different models depending upon the values of p and p' taken. After the genotypic value 'a' is known for a particular model, the moments and the product-moments for that model can be obtained easily from table 4.6. These have been grouped in tables 6.1 to 6.6 for convenience of later calculations. All the moments are about the means of X_1, X_2 and X_3. As the title of this chapter shows, the models involving two independent pairs of linked factors belonging to both classes will be dealt here.

Table 6.1: Moments of X_1, F_2- phenotypic value for models of class I & II

(All odd moments = 0)

Values of p & p' for various models	Powers of x_1 for models of class I				Powers of x_1 for models of class II			
	x_1^2	x_1^4	x_1^6	x_1^8	x_1^2	x_1^4	x_1^6	x_1^8
p=p'=0	3.5	35.6250000	567.4375000	18216.0312500	5.0	70.5000000	1564.5000000	46144.8000000
p=p'=.3	3.5	35.7656250	594.1479492	19307.0932770	4.4	53.3600000	1122.5100000	30591.8345000
p=.2, p'=.4	3.5	35.8816327	597.6444242	19656.1477314	4.1	40.6300000	930.0675000	24167.4011850
p=p'=.4	3.5	36.0000000	605.3906250	19999.7167969	3.8	42.2400000	764. 3550000	18944.9407500
p=p'=.5	3.5	36.1875000	614.4375000	19300.6035156	3.5	36.1875000	614.4375000	19300.6035156

Table 6.2: Product-moments: X_3, Genotypic mean of F_3 progeny and powers of X_1, F_2- phenotypic value, for models of class I & II.

Values of p & p' for various models	Models of class I					Models of class II				
	x_3	x_3x_1	$x_3x_1^2$	$x_3x_1^3$	$x_3x_1^4$	x_3	x_3x_1	$x_3x_1^2$	$x_3x_1^3$	$x_3x_1^4$
p=p'=0	0	1.5000000	0	14.6250000	0	0	3.0000000	0	40.5000000	0
p=p'=.3	0	1.5000000	0	14.7656250	0	0	3.4000000	0	29.1600000	0
p=.2, p'=.4	0	1.5000000	0	14.8316326	0	0	3.1000000	0	24.0300000	0
p=p'=.4	0	1.5000000	0	15.0000000	0	0	1.8000000	0	19.4400000	0
p=p'=.5	0	1.5000000	0	15.1875000	0	0	1.5000000	0	15.1875000	0

Table 6.3: Product moments: X_3, Genotypic variance within progeny and powers of X_1, F_2 phenotypic value for models of class I & II.

Values of p & p' for various models	Models of class I					Models of class II				
	x_2	x_2x_1	$x_2x_1^2$	$x_2x_1^3$	$x_2x_1^4$	x_2	x_2x_1	$x_2x_1^2$	$x_2x_1^3$	$x_2x_1^4$
p=p'=0	0	0	-.5625000	0	+10.1250000	0	0	-2.2500000	0	+34.0000000
p=p'=.3	0	0	-.4078125	0	+ 7.6579589	0	0	+1.0440000	0	+23.0310000
p=.2, p'=.4	0	0	-.3765306	0	+ 6.8882471	0	0	+0.7330000	0	+15.3539500
p=p'=.4	0	0	-.3000000	0	+ 5.6359375	0	0	+0.4320000	0	+ 8.7885000
p=p'=.5	0	0	-.2812500	0	+ 5.4843750	0	0	+0.2812500	0	+ 5.4843750

Table 6.4: Product-moments: $\bar{x}_3^2$, square of Genotypic mean of F_3 progeny and the powers of $\bar{x}_1$, F_2- phenotypic value for models of Class I & II.

Values of p & p' for various models	Models of class I					Models of class II				
	$\bar{x}_3^2$	$\lvert\bar{x}_3^2\bar{x}_1\rvert$	$\bar{x}_3^2\bar{x}_1^2$	$\lvert\bar{x}_3^2\bar{x}_1^3\rvert$	$\bar{x}_3^2\bar{x}_1^4$	$\bar{x}_3^2$	$\lvert\bar{x}_3^2\bar{x}_1\rvert$	$\bar{x}_3^2\bar{x}_1^2$	$\lvert\bar{x}_3^2\bar{x}_1^3\rvert$	$\bar{x}_3^2\bar{x}_1^4$
p=p'=0	1.5	0	8.6250000	0	114.1875000	3.0	0	23.0000000	0	535.5000000
p=p'=.2	1.3	0	8.7656250	0	118.3666922	2.4	0	19.5600000	0	383.6300000
p=.2,p'=.4	1.3	0	9.8316336	0	120.4750364	2.1	0	15.6300000	0	251.9275000
p=p'=.4	1.5	0	9.0000000	0	125.3906250	1.8	0	12.2400000	0	186.6850000
p=p'=.5	1.5	0	9.1875000	0	131.0625000	1.5	0	9.1875000	0	131.0625000

Table 6.5: Product-moments, $\bar{x}_3\,\bar{x}_{3s}$ product of F_3 progeny mean and genotypic variance within progeny and powers of $\bar{x}_1$, F_2- phenotypic value for models of class I & II.

Values of p & p' for various models	Models of class I						Models of class II					
	$\lvert\bar{x}_2\bar{x}_3\rvert$	$\bar{x}_2\bar{x}_3\bar{x}_1$	$\lvert\bar{x}_2\bar{x}_3\bar{x}_1^2\rvert$	$\bar{x}_2\bar{x}_3\bar{x}_1^3\rvert$	$\lvert\bar{x}_2\bar{x}_3\bar{x}_1^4$		$\lvert\bar{x}_2\bar{x}_3\rvert$	$\bar{x}_2\bar{x}_3\bar{x}_1$	$\lvert\bar{x}_2\bar{x}_3\bar{x}_1^2\rvert$	$\bar{x}_2\bar{x}_3\bar{x}_1^3$	$\lvert\bar{x}_2\bar{x}_3\bar{x}_1^4$	
p=p'=0	0	-.5625000	0	-6.7500000	0	0	-2.2500000	0	-40.5000000	0		
p=p'=.2	0	-.6078125	0	-8.0110839	0	0	-1.9440000	0	-16.7670000	0		
p=.2,p'=.4	0	-.8765306	0	-4.6290634	0	0	-0.7300000	0	-10.9309300	0		
p=p'=.4	0	-.3000000	0	-3.9359375	0	0	-0.9320000	0	-6.1965000	0		
p=p'=.5	0	-.2812500	0	-3.7968750	0	0	-0.2812500	0	-3.7968750	0		

Table 6.6: Product-moments, $\bar{x}_2^2$, squares of Genotypic variance within F_3 progeny and powers of $\bar{x}_1$, F_2 phenotypic value for models of class I & II.

Values of p & p' for various models	Models of class I					Models of class II				
	$\bar{x}_2^2$	$\lvert\bar{x}_2^2\bar{x}_1\rvert$	$\bar{x}_2^2\bar{x}_1^2$	$\lvert\bar{x}_2^2\bar{x}_1^3\rvert$	$\bar{x}_2^2\bar{x}_1^4$	$\bar{x}_2^2$	$\lvert\bar{x}_2^2\bar{x}_1\rvert$	$\bar{x}_2^2\bar{x}_1^2$	$\lvert\bar{x}_2^2\bar{x}_1^3\rvert$	$\bar{x}_2^2\bar{x}_1^4$
p=p'=0	.2812500	0	.9843750	0	1.0968750	1.1250000	0	3.9375000	0	96.5000000
p=p'=.2	.1933594	0	.6557959	0	7.0671602	0.4950000	0	2.0921400	0	29.1135330
p=.2,p'=.4	.1721939	0	.5976763	0	6.8003770	0.3375000	0	1.8420350	0	17.1133763
p=p'=.4	.1250000	0	.4389063	0	4.8668556	0.1800000	0	0.6864300	0	8.8580765
p=p'=.5	.1406250	0	.4921875	0	5.4843263	0.1406250	0	0.4921875	0	5.4843263

6.2. From the moments and product-moments obtained above, fourth degree regression equations of five F_3 properties on F_2 parental value (X_1) were calculated for each model. The five F_3 properties are (1) X_3, the genotypic mean of F_3 progeny (2) X_3, the genotypic variance within F_3 progeny (3) X_2X_3, the product of progeny mean and variance within progeny (4) X_3^2 the square of the F_3 progeny mean and (5) X_3^2, the square of the genotypic variance within F_3 progeny respectively. The last three quantities are useful for calculating the covariance of F_3 progeny mean and variance within progeny, the variance of progeny mean and the variance of variance within progeny, respectively. Since regressions of five dependent variates on the same independent variate are to be calculated, it is convenient to obtain the covariance matrix for each of the models from the moments of X_1, the independent variate. This was done for all the models involving two independent pairs of linked factors. Since the models are without dominance, the odd moments are all zero and consequently the alternate c-coefficients in the matrices obtained are also zero. Thus in models with $pq'=0$, belonging to the models of class I, we have the following set of equations for the calculation of the covariance matrix:

$$A + 0.B + 3.5C + 0.D + 35.6250000E = 1\ 0\ 0\ 0\ 0$$
$$0.A + 3.5B + 0.C + 35.6250D + 0.E = 0\ 1\ 0\ 0\ 0$$
$$3.5A + 0.B + 35.6250C + 0.D + 387.43750E = 0\ 0\ 1\ 0\ 0$$
$$0.A + 35.6250B + 0.C + 587.4375D + 0.E = 0\ 0\ 0\ 1\ 0$$
$$35.6250A + 0.B + 587.43750C + 0.D + 13216.031250E = 0\ 0\ 0\ 0\ 1$$

solving these sets of equations we obtain the following matrix:

$$\begin{pmatrix} 1.9803610 & 0 & -.3389075 & 0 & .0120336 \\ 0 & .7463300 & 0 & -.0452731 & 0 \\ -.3389075 & 0 & .1936568 & 0 & -.0071060 \\ 0 & -.0452731 & 0 & .0044479 & 0 \\ .0120336 & 0 & -.0071060 & 0 & .0003589 \end{pmatrix}$$

By taking product moments corresponding to this model from the appropriate tables (6.1 to 6.6), we obtain the coefficients A, B, C, D and E in regression equations of the type $y = A+Bx_1+Cx_1^2+Dx_1^3+Ex_1^4$, where y stands for one of the five properties studied. Thus we have for $y = x_3$,

$A = 1.9303410\ (0) + 0.\ (1.5) + \ldots\ldots = 0$

$B = 0.(0) + .74653\ (1.5) + \ldots\ldots\ldots = .4376759$

$C = -.3339473\ (0) + 0\ (1.5) + \ldots\ldots = 0$

$D = 0.\ (0) - .0452731\ (1.5) + \ldots\ldots = -.0028593$

$E = .0120936\ (0) + 0.\ (1.5) + \ldots\ldots = 0$

In the present models the coefficients A, C, E turn out to be zero for properties X_3 and X_2X_3, while coefficients B and D are zero in equations for the remaining three properties. Therefore the equations are of the form $y = Bx_1 + Dx_1^3$ and $y = A+Cx_1^3 + Ex_1^4$ in the two cases respectively.

The regression coefficients for the various properties are given below, for models of class I & II in tables 6.7, 6.8 and 6.9.

Table 6.7: Regression coefficients for equations of properties X_3, Genotypic mean of F_3 progeny and X_2X_3, product of F_3 genotypic mean and genotypic variance within progeny.

(equations of the type $y = Bx_1 + Dx_1^3$)

Values of p & p' for various models	Models of class I				Models of class II			
	$y = x_3$		$y = x_2x_3$		$y = x_3$		$y = x_2x_3$	
	B	D	B	D	B	D	B	D
p=p'=0	.4376759	-.0028593	-.1143297	-.0065572	.6044928	-.0031554	-.8381063	-.0193326
p=p'=.2	.4537079	-.0024593	-.0738120	-.0036901	.5798158	-.0027285	-.1297562	-.0085139
p=.2,p'=.4	.4518681	-.0022769	-.0732073	-.0033573	.5416752	-.0024858	-.1068887	-.0061638
p=p'=.4	.4478207	-.0018339	-.0507043	-.0034035	.4951159	-.0019271	-.0611081	-.0047302
p=p'=.5	.4424027	-.0013332	-.0321063	-.0036994	.4424027	-.0013332	-.0321063	-.0036994

Table 6.8: Regression coefficients for equations of properties X_2, genotypic variance within progeny, X_2^2 and X_3^2, square of Genotypic mean of F_3 progeny for class I models.

(equations of the type $y = A + Cx_1^2 + Ex_1^4$)

Values of p & p' for various models	$y = x_2$			$y = x_2^2$			$y = x_3^2$		
	A	C	E	A	C	E	A	C	E
p=p'=0	.0964366	-.0312462	.0003628	.8927207	-.0067451	.0003408	.9218353	.1874772	-.0021819
p=p'=.2	.0686867	-.0220235	.0002346	.2032215	-.0048420	.0001982	.9129124	.1869056	-.0018733
p=.2, p'=.4	.0634804	-.0204093	.0002222	.1795383	-.0036203	.0001484	.9102724	.1858845	-.0017043
p=p'=.4	.0484399	-.0150952	.0001222	.1266664	-.0021978	.0001116	.8981204	.1865783	-.0014254
p=p'=.5	.0438698	-.0133447	.0000784	.1445969	-.0023018	.0001129	.8858594	.1868249	-.0010977

Table 6.9: Regression coefficients for equations of properties X_2, genotypic variance within progeny, X_2^2 and X_3^2, square of genotypic mean of F_3 progeny for class II models.

(equations of the type $y = A + Cx_1^2 + Ex_1^4$)

Values of p & p' for various models	$y = x_2$			$y = x_2^2$			$y = x_3^2$		
	A	C	E	A	C	E	A	C	E
p=p'=0	.2890955	-.0672318	.0006669	1.2320988	-.0455073	.0017096	1.2654338	.4033923	-.0039989
p=p'=.2	.1457861	-.0377954	.0003693	.5359829	-.0163069	.0009949	1.1575337	.3190147	-.0028041
p=.2, p'=.4	.1094778	-.0304502	.0003153	.3579035	-.0088928	.0003310	1.0885910	.2745293	-.0023368
p=p'=.4	.0649780	-.0187344	.0001475	.1866636	-.0036833	.0001733	.9932251	.2312422	-.0017111
p=p'=.5	.0438698	-.0133447	.0000784	.1445969	-.0023018	.0001129	.8858594	.1868249	-.0010977

6.3: These equations help us to study the effects of selection as shown in the next chapter. These also help in interpreting the relations between the various F_3- properties and X_1, the F_2- parental value. In this respect the first two properties, namely X_3, the genotypic mean of F_3 progeny and X_2, the genotypic variance within F_3 progeny are of particular interest in that they represent directly the quantities relevant to selection. The remaining three quantities which are the means of the quantities like the variance of the means of F_3 progenies etc, also represent important properties of the material under selection. Equations for the first two properties are represented graphically in Figures 6.1 to 6.6 for models of both classes for different pairs of values of p and p'. Figures 6.1 and 6.2 represent the graphs for X_3, the genotypic mean of F_3 progeny for models of class I and class II respectively. The remaining figures correspond to the graphs for X_2, the genotypic variance within F_3 progeny.

6.4: From figures 6.1 and 6.2 it appears that the regression of progeny mean on parental value is approximately linear over a wide range of F_2- phenotypic values for both classes of models. The curves being symmetrical in first and third quadrants tend to turn back towards X_1- axis after covering a linear path in each case. With the fixation of magnitude of each factor in class II models the effect of linkage on the regression is distinctly visible from figure 6.2, as the different curves are sufficiently distinct from one an other, unlike curves shown in Figure 6.1 for class I models with fixed F_2 genotypic variance where the curves are very close to one an other. This distinction between curves shown in Figures 6.1 and 6.2, may be traced to the fact that where F_2- genotypic variance is fixed curves show no difference and where factors are fixed in magnitudes, but the F_2 genotypic variances are different, the curves arrange themselves in the order of linkage, the slope being steeper with tighter linkage.

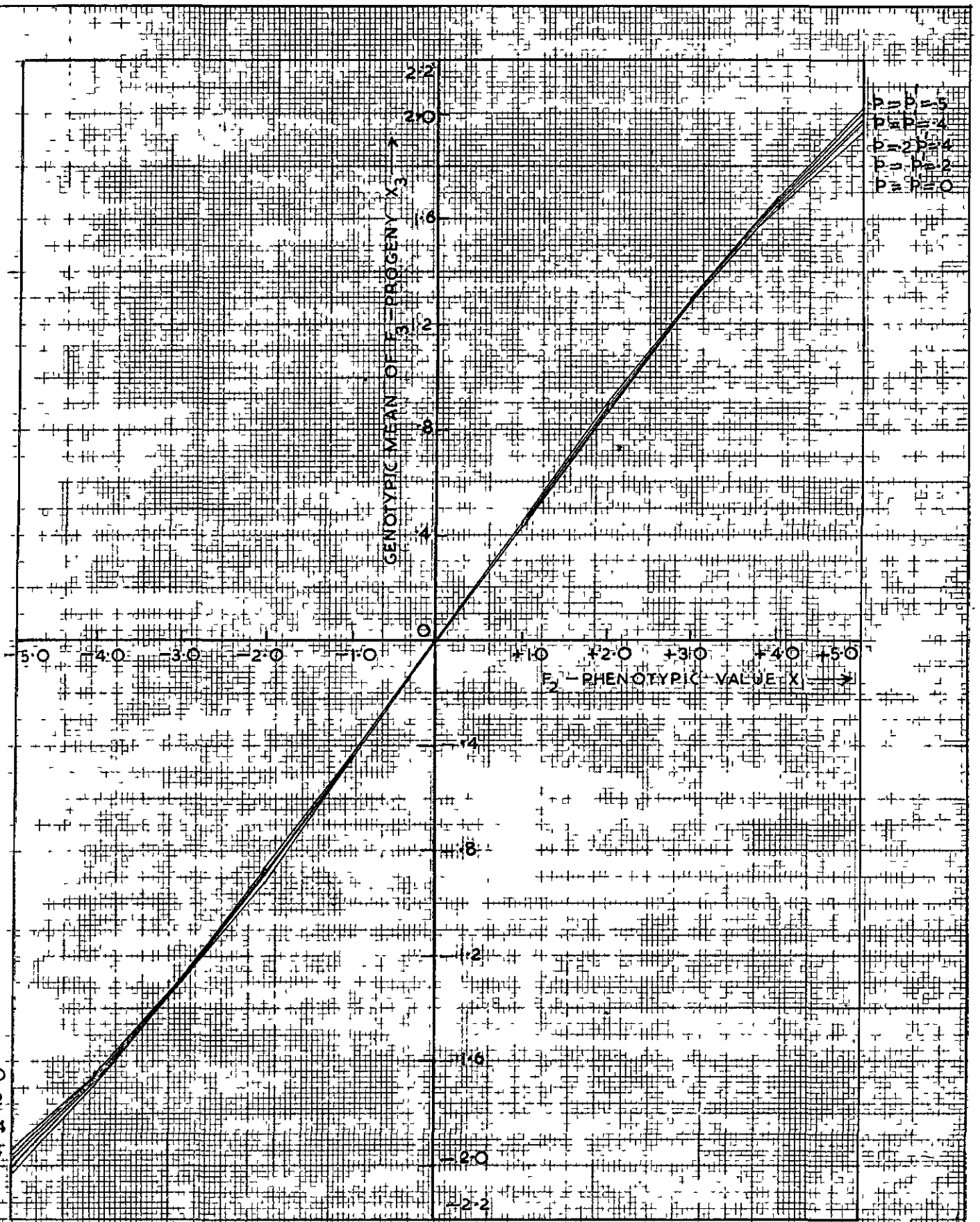


FIG 6·1 RELATION BETWEEN X_3, THE GENOTYPIC MEAN OF F_3 PROGENY AND X_1, THE F_2 PHENOTYPIC VALUE FOR CLASS I MODELS.

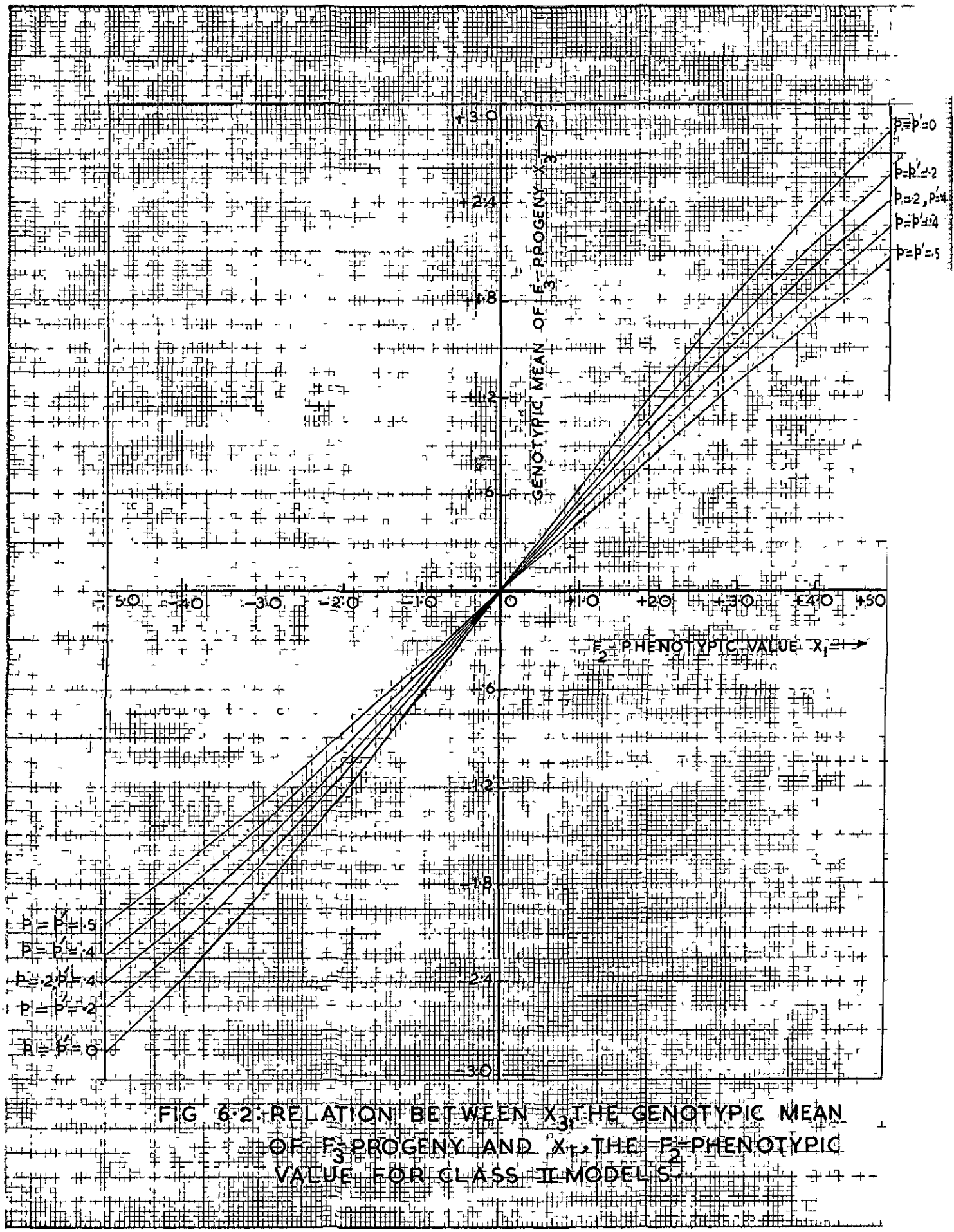


FIG 6·2: RELATION BETWEEN X_3, THE GENOTYPIC MEAN OF F_3-PROGENY AND X_1, THE F_2-PHENOTYPIC VALUE FOR CLASS II MODELS

This confirms the observation that it is F_2- genotypic variance and not the magnitude of individual factors, which predominantly determines the relationship between F_2- values and means of their F_3 progenies. Figure 6.1 shows a very slight opposite trend as compared to figure 6.2, so far as the effect of linkage is concerned. For a fixed F_2- genotypic variance, the slight difference in the slope of the curves (Figure 6.1) is determined by the change in the linkage.

6.3: We next switch on to the relation between X_3, the genotypic variance within F_3 progeny and X_1, the F_2- phenotypic value. Actually this relationship can be visualised as composed of two types of effects, as follows:

(1) The relation between F_2 genotypic variance and the mean variance within F_3 progenies and

(2) The relation of variance within individual F_3 progenies and F_2 parental values.

The second effect can be studied from figures 6.3 and 6.4 where the curves have been plotted between various values of X_1, the F_2 phenotypic value and the deviation of the genotypic variance within F_3 progeny from its mean say small x_3 as directly obtained in the regression method, for the models of two classes respectively, so that the comparisons are not affected by different values of mean variance within F_3 progenies for the different models.

Again the sum total of the above two effects on genotypic variance within F_3 progeny can be studied from the graphs between X_1, the F_2 parental value and the absolute value of the genotypic variance within F_3 progeny i.e. X_3 for the two classes (Figures 6.5 and 6.6). These absolute values are obtained by adding the value of $H'F_3$ (without selection) corresponding to each model to the value of the deviation of X_3 from its mean as directly obtained in the regression method.

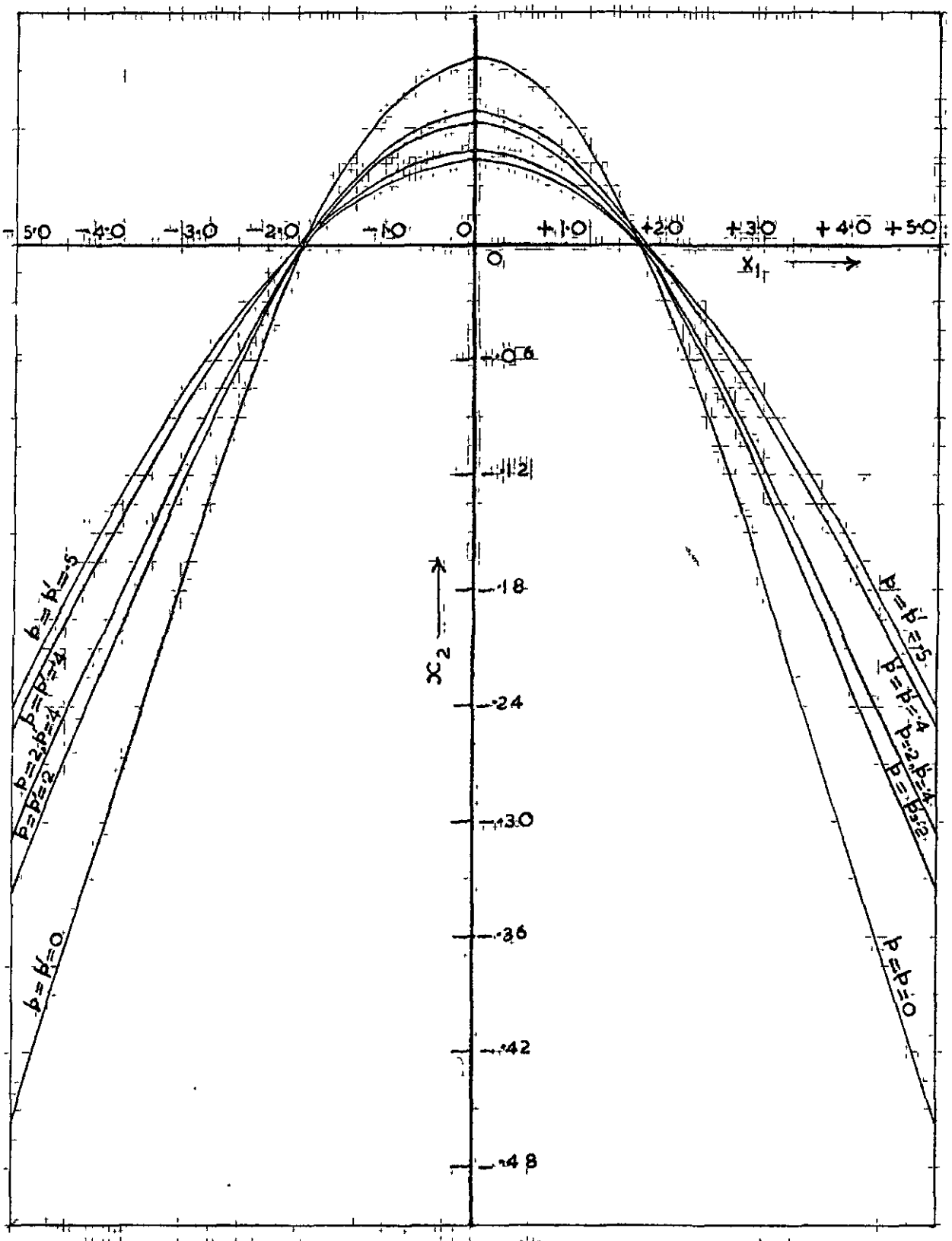


FIG 6·3: RELATION BETWEEN X_2, THE DEVIATION OF GENOTYPIC VARIANCE WITHIN F_3 PROGENY FROM ITS MEAN AND X_1, THE F_2^3 PHENOTYPIC VALUE FOR CLASS I MODELS.

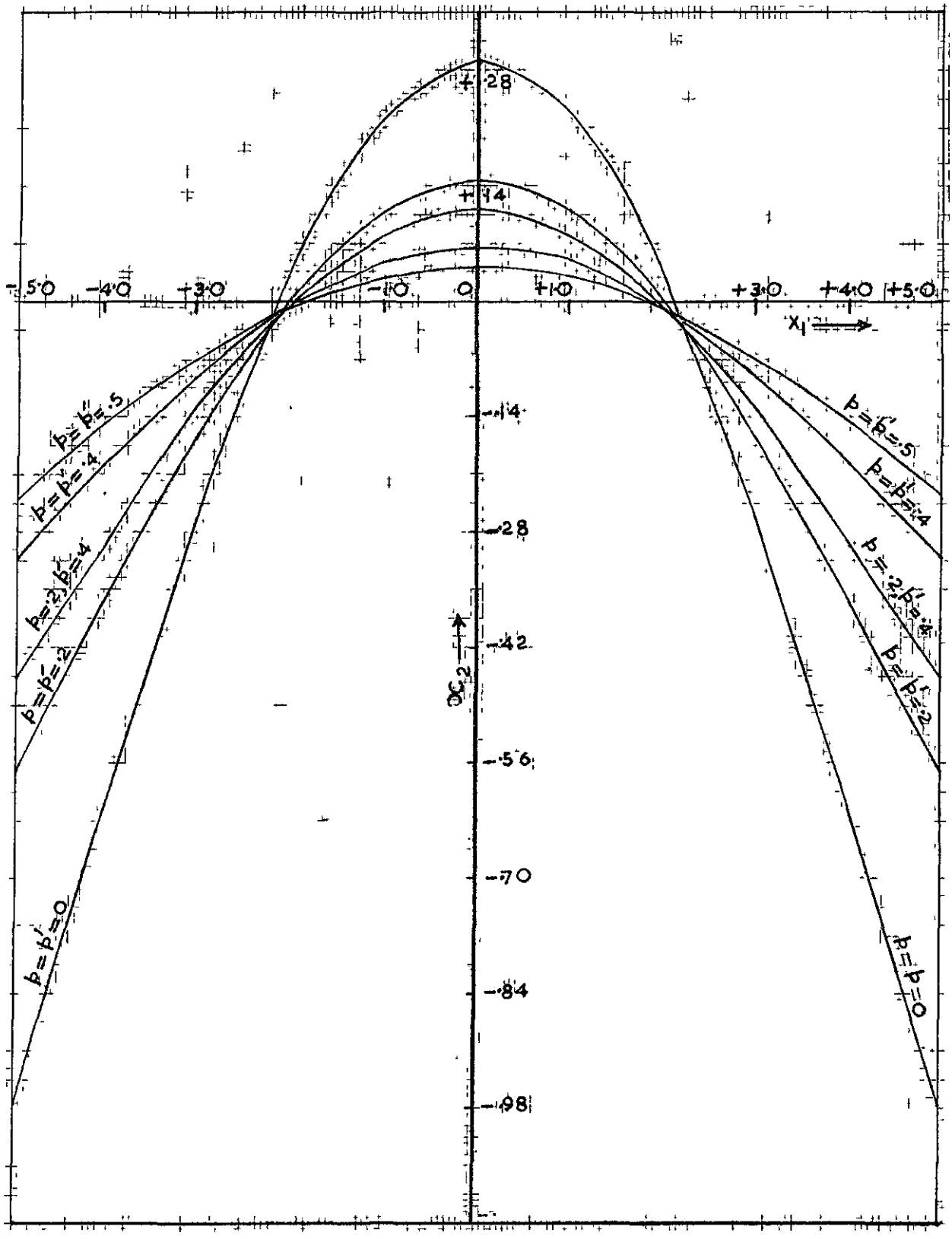


FIG 6·4: RELATION BETWEEN DC_2, THE DEVIATION OF GENOTYPIC VARIANCE WITHIN F_3 PROGENY FROM ITS MEAN AND X_1, THE F_2 PHENOTYPIC VALUE FOR CLASS III MODELS.

We shall discuss the conclusions drawn from the above two studies separately as follows:

Referring to the curves shown in figures 6.3 and 6.4 for the study based on the second effect, it may be pointed out that these curves explain how the variance within individual F_3 progenies is distributed around its mean. This distribution behaves in the same fashion for the two classes of models, except that the curves for class I models are more sharply peaked. It may be a point of some special importance that for a particular value of X_1, the F_3 parental value, the difference between the values of X_2 for models with p=p'=.5 and p=p'=.4, the cases of extremely and moderately loose linkage respectively is far less than the difference of the values for the models with p=p'=0 and p=p'=.2, the cases of extremely and moderately tight linkage respectively. This shows that the disturbance in the variance within an F_3 progeny is very much greater for models with close linkage. Further the maximum value of various curves increases as the linkage is tightened. This fact is very well demonstrated in the above graphs (Figures 6.3 and 6.4) where it is observed that the shape of the curve changes from platykurtic to Leptokurtic as the linkage gets tightened.

We now pass on to study the curves presented in figures 6.5 and 6.6 for the two classes of models. These curves explain the distribution of the absolute value of the genotypic variance within F_3 progeny i.e. X_2, which actually accounts for the sum total of the above two effects.

The behaviour of the distribution of the absolute value X_2, for class II models where the magnitude of the factors is fixed but the F_2 genotypic variance is changing from model to model, is similar to the distribution of X_2 around its mean (discussed above) in that the various curves arrange themselves in the order of linkage values (Fig. 6.6). But

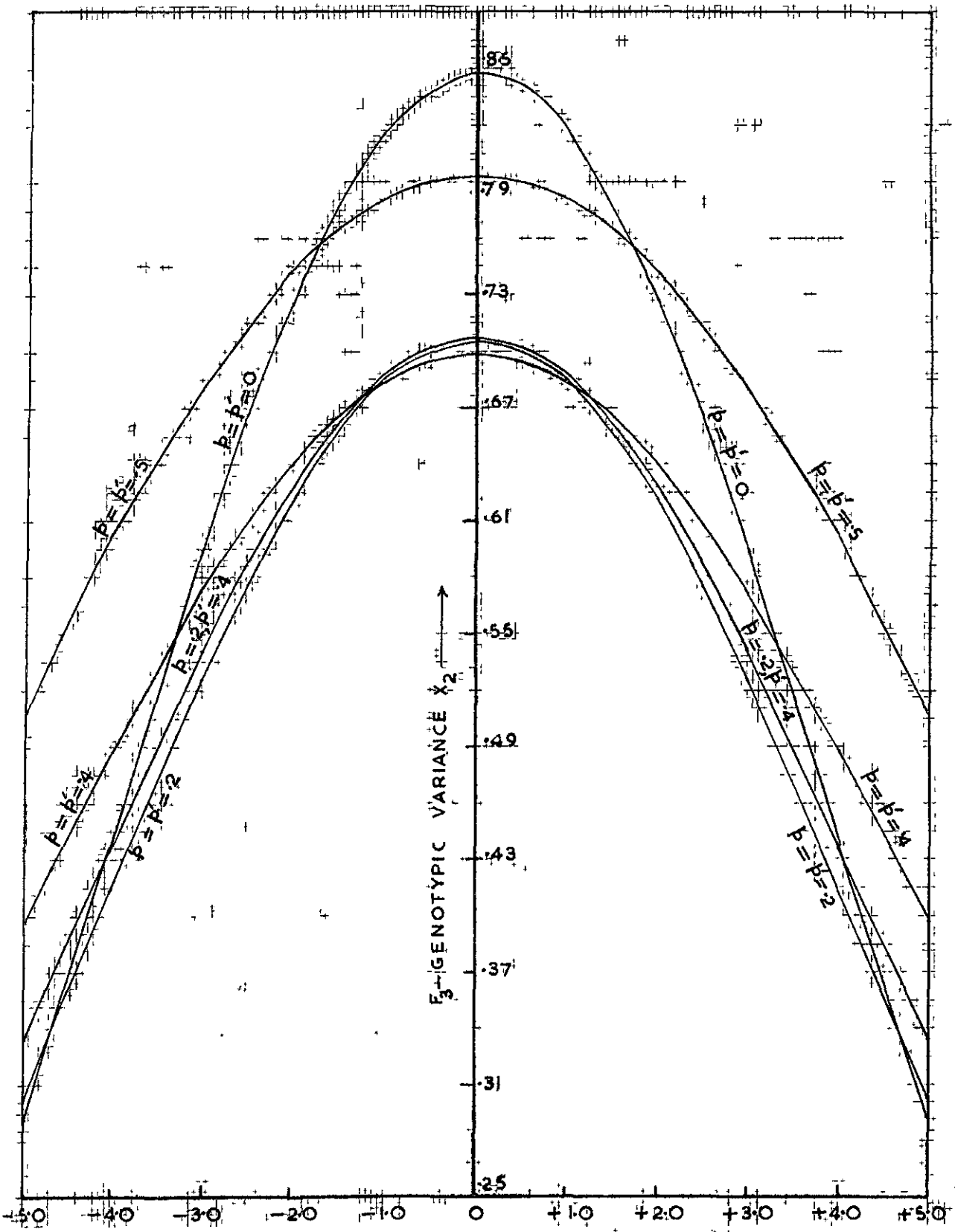


FIG. 6.5: RELATION BETWEEN X_2, THE GENOTYPIC VARIANCE WITHIN F_3 PROGENIES AND X_1, THE F_2 PHENOTYPIC VALUE FOR CLASS I MODELS.

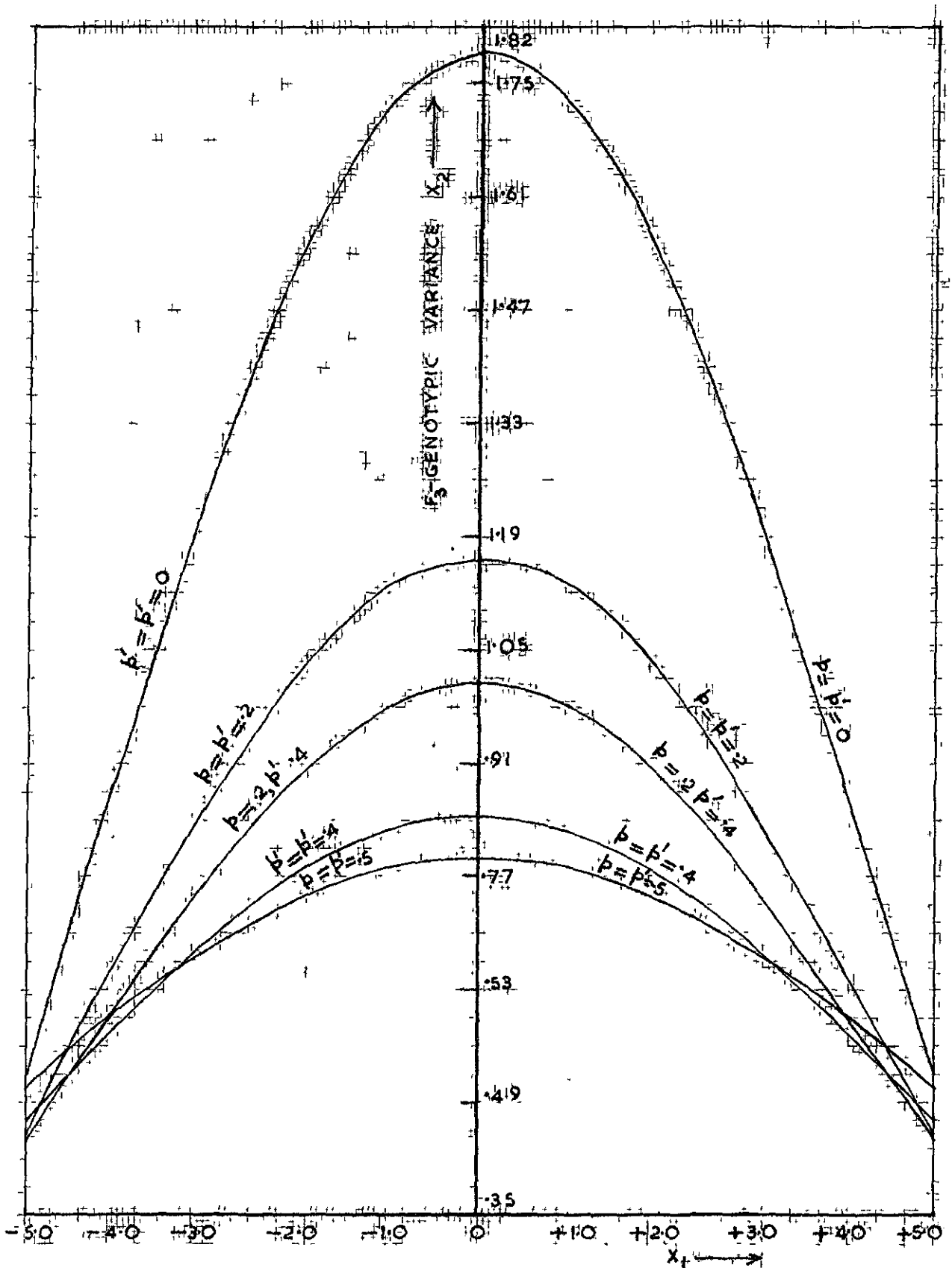


FIG 6·6. RELATION BETWEEN X_2, THE GENOTYPIC VARIANCE WITHIN F_3 PROGENIES AND X_1, THE F_2 PHENOTYPIC VALUE, FOR CLASS III MODELS.

for class I models where F_2- genotypic variance is fixed, the distribution of the absolute value K_2 as shown by curves in Fig. 6.5, is not keeping with the regular trend shown in other graphs. Here the various curves do not arrange themselves in the order of linkage values. This is because of the special behaviour of the first effect, namely, the relation between F_3 genotypic variance and the mean variance within F_3 progenies, for the class I models explained below. When p=p'=.5, all the four factors behave independently and when p=p'=0, the four factors behave as two independent factors each with double the magnitude of a factor. In both these cases, the mean genotypic variance within F_3 progeny turns out to be half of the F_2- genotypic variance irrespective of the number of independently behaving factors involved therein. When the factors are partially linked, this does not hold true. Values of $\bar{H}^{V}F_3$ for various models of class I and II are given in the next chapter in tables 7.2 and 7.4 respectively. It can be seen that for class II models where $H^{V}F_2$ is changing from model to model, the value of $\bar{H}^{V}F_3$ is not exactly half of F_2- genotypic variance for all values of p's, except at p=p'=0 and p=p'=.5, though the trend of continuous decrease in the value of $\bar{H}^{V}F_3$ with the loosening of linkage, is same as that of $H^{V}F_2$ (Table 7.4). On the contrary for class I models where the F_2- genotypic variance is fixed to 1.5 units, the value of $\bar{H}^{V}F_3$ is half of this i.e. .75 unit only for models with p=p'=0 and p=p'=.5 and is less than this for other models where the effect of linkage on $\bar{H}^{V}F_3$ is significant (Table 7.2). This means that for those models the value of $\bar{H}^{V}F_3$ first decreases from the value .75 unit at p=p'=0 and then increases to give value .75 unit again at p=p'=.5. This trend of $\bar{H}^{V}F_3$ for class I models is not consistent with that for class II models. It is because of this reason that while graphs showing the distribution of variance within

individual F_3 progenies behave in a normal fashion for both class of models (Figures 6.3 and 6.4), it is not so in the case of graphs corresponding to the absolute value of the genotypic variance within F_3 progeny, specially in the case of class I models where the behaviour of $H^{\bar{V}}F_3$ is not regular. Obviously linkage affects the genotypic variance within F_3 progeny profoundly even where the initial F_2 genotypic variance is the same.

VII. F_3 STATISTICAL PROPERTIES OF SELECTED POPULATIONS FOR THE MODELS INVOLVING TWO INDEPENDENT PAIRS OF LINKED FACTORS

7.1. After obtaining the regression equations of the five F_3 statistical properties on X_1, the F_2 phenotypic value, as described in the previous chapter, the mean values of those properties corresponding to the selected portion of the F_2 population were obtained by the technique given by Panse. A brief sketch of the method will be given here. The selection intensity was kept same i.e. 10% as was in the case of a single pair of linked factors. For this level of selection, the deviate of the F_2 distribution ξ, was calculated for the models corresponding to each pair of values of p and p', for both classes of models by using the procedure adopted in chapter V. The values of ξ obtained in various cases have been given in table 7.1.

Table 7.1: Values of ξ for 10% selection for models of class I and class II.

Values of p and p'	values of for class I models	Values of for class II models
p=p'=0	1.2883205	1.2948371
p=p'=.2	1.2874681	1.2911492
p=.2, p'=.4	1.2870753	1.2894523
p=p'=.4	1.2860460	1.2872403
p=p'=.5	1.2849069	1.2849069

7.2. The values of ξ given in table 7.1 are in standard measure. The values of the actual deviates of the F_3 distribution will be given by $\xi\sigma$, where σ is the standard deviation of the distribution. The mean value of any F_3 property 'y' for the selected portion of the F_2 population is

obtained from the integral

$\int_{\xi\sigma}^{+\infty} y.f(x_1)dx_1$ where $f(x_1)dx_1$ is the frequency element, by dividing this integral by the fraction of the F_2- population selected, i.e. by Q.i. $\xi\sigma$ stands for x_1 and represents the lower limit of F_2- phenotypic values selected, the upper limit theoritically being $+\infty$. The function 'y' takes different values for different F_3 proportion and has been expressed in terms of F_2- phenotypic values by means of the regression equations of the type (1) $y = Bx_1 + Dx_1^3$ and (2) $y = A+Cx_1^2 + Ex_1^4$. The required integrals are therefore

(1) $\int (Bx_1 + Dx_1^3)\, f(x_1)dx_1$

(2) $\int (A + Cx_1^2 + Ex_1^4)\, f(x_1)dx_1$

The frequency element $f(x_1)dx_1$ can be represented by

$$\left\{ e^{-(K_1-m)\frac{d}{dx_1} + \frac{1}{2!}(K_2-V)\frac{d^2}{dx_1^2} + \frac{1}{3!}K_3\frac{d^3}{dx_1^3} + \ldots.} \right\} \cdot \frac{1}{\sqrt{2\pi V}} \cdot e^{-\frac{(x-m)^2}{2V}} \cdot dx_1$$

(Cornish and Fisher, 1937).

As has been explained in chapter V, we take $m = K_1$ and $V = K_2$. Also the odd cumulants are zero in our case.

Therefore (1) $\int (Bx_1+Dx_1^3)\ f(x_1)\, dx_1$

$$= \int (Bx_1+Dx_1^3).\left(e^{\frac{1}{4!}K_4.\frac{d^4}{dx_1^4} + \frac{1}{6!}K_6\frac{d^6}{dx_1^6}+\ldots.}\right)\frac{1}{\sqrt{2\pi V}}.e^{-\frac{x_1^2}{2V}}.dx_1$$

Also since $x_1 = \xi V^{\frac{1}{2}}$, $dx_1 = V^{\frac{1}{2}}\, d\xi$, the integral (2) can be expressed as

$$\int_{\xi}^{\infty} (B\xi V^{\frac{1}{2}} + D\xi^3 V^{3/2})\left(e^{\frac{1}{4!}+\frac{K_4}{V^2}\frac{d^4}{d\xi^4}+\frac{1}{6!}\frac{K_6}{V^3}\frac{d^6}{d\xi^6}+\ldots.}\right)\frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\xi^2}.d\xi.$$

Substituting further, $d = \frac{K_4}{V^2}$, $f = \frac{K_6}{V^3}$, $h = \frac{K_8}{V^4}$, ... and retaining upto the terms involving h for an approximation, we have this integral as(1) as

$$= \int_{\xi}^{\infty} (B\xi V^{\frac{1}{2}} + D\xi^{3}V^{3/2}) \cdot \left(0 \frac{d}{4!}\frac{d^4}{d\xi^4} + \frac{f}{6!}\frac{d^6}{d\xi^6} + \frac{h}{8!}\frac{d^8}{d\xi^8}\right) \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi^2} \cdot d\xi .$$

On expanding the exponential and putting z for $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\xi^2}$ and the hermite polynominals ξ_r given by $\frac{d^r}{d\xi^r} \cdot z = \xi_r z$, the integral (1) is reduced to

$$\int_{\xi}^{\infty} (B\xi V^{\frac{1}{2}} + D\xi^{3}V^{3/2}) \left(1 + \frac{d}{24}\xi_4 + \frac{f}{720}\xi_6 + \left(\frac{h}{40320} + \frac{d^2}{1152}\right)\xi_8\right) z\,d\xi$$

Multiplying out the brackets and substituting in the product $\xi_{r+1} - r\xi_{r-1}$ for $\xi \cdot \xi_r$, the integral further reduces to

$$\int_{\xi}^{\infty} \left[+BV^{\frac{1}{2}} \left\{ \xi_1 + \frac{d}{24}(\xi_5 + 4\xi_3) + \frac{f}{720}(\xi_7 + 6\xi_5) + \left(\frac{h}{40320} + \frac{d^2}{1152}\right)(\xi_9 + 8\xi_7) \right\} \right.$$

$$-DV^{3/2} \left\{ 3\xi_1 + \xi_3 + \frac{d}{24}(\xi_7 + 13\xi_5 + 48\xi_3 + 24\xi_1) + \frac{f}{720}(\xi_9 + 21\xi_7 + 103\xi_5 + 120\xi_3) \right.$$

$$\left. \left. + \left(\frac{h}{40320} + \frac{d^2}{1152}\right)(\xi_{11} + 37\xi_9 + 193\xi_7 + 336\xi_5) \right\} \right] z\,d\xi .$$

Now since $\int \xi_r z = \xi_{r-1} z$ and therefore on integration, it becomes equal to

$$z \cdot \left[BV^{\frac{1}{2}} \left\{ 1 + \frac{d}{24}(\xi_4 + 6\xi_2) + \frac{f}{720}(\xi_6 + 6\xi_4) + \left(\frac{h}{40320} + \frac{d^2}{1152}\right)(\xi_8 + 8\xi_6) \right\} \right.$$

$$+DV^{3/2} \left\{ \xi_3 + 3 + \frac{d}{24}(\xi_6 + 15\xi_4 + 60\xi_2 + 24) + \frac{f}{720}(\xi_8 + 21\xi_6 + 108\xi_4 + 120\xi_2) \right.$$

$$\left. \left. + \left(\frac{h}{40320} + \frac{d^2}{1152}\right)(\xi_{10} + 37\xi_8 + 192\xi_6 + 336\xi_4) \right\} \right] \quad \text{——— (1)}$$

The second integeral can also be written down as

$$\int (A + C\xi^2 V + E\xi^4 V^2) \left(1 + \frac{d}{24}\frac{d^4}{d\xi^4} + \frac{f}{720}\frac{d^6}{d\xi^6} + \left(\frac{h}{40320} + \frac{d^2}{1152}\right)\frac{d^8}{d\xi^8}\right) z\,d\xi ,$$

which by adopting a procedure similar to that used for the first integeral and since $\int 24\xi = P$ which is here equal to 0.1, can be expressed in the form

$$0.1(\Delta+\sigma V+3EV^2+EV^2{}_0 d)$$

$$+ z\left[A\left\{ \frac{d}{24}\xi_3 + \frac{f}{720}\xi_5 + \left(\frac{h}{40320} + \frac{d^2}{1152}\right)\xi_7 \right\} + \right.$$

$$\sigma V\left\{ 1 + \frac{d}{24}(\xi_5+9\xi_3+12\xi_1) + \frac{f}{720}(\xi_7+15\xi_5+30\xi_3)+\left(\frac{h}{40320}+\frac{d^2}{1152}\right)(\xi_9+17\xi_7+56\xi_5)\right.$$

$$+EV^2\left\{\xi_3+6\xi_1+\frac{d}{24}(\xi_7+23\xi_5+138\xi_3+168\xi_1) + \frac{f}{720}(\xi_9+30\xi_7+255\xi_5+\right.$$

$$\left.\left.\left.660\xi_3+360\xi_1) + \left(\frac{h}{40320}+\frac{d^2}{1152}\right)(\xi_{11}+38\xi_9+435\xi_7+1680\xi_5+1680\xi_3)\right\}\right\}\right] \text{-----(3).}$$

7.3. The two expressions given in 7.2 represent y between the limits $\xi V^{\frac{1}{2}}$ and $+\infty$ and are therefore equivalent to $\bar{y}x0.1$, where $\bar{y}$ is the mean value of y between those limits.

By making numerical substitutions for the hermite polynomials ξ_1, ξ_2, ... etc and for the coefficients d, f, h appropriate to different models, the first expression can be used to give the mean value of y when $y = x_3$ or x_2x_3, and the second when $y = x_2$, x_2^2 or x_3^2.

Table 7.2: Results of selection for models of class I with two independent pairs of linked factors with $H^V F_2 = 1.8$ units and environmental variance $V = 2.0$ units (By Regression Method)

S.No. of the model	(1) Value of p & p'	(2) Limit of selective advance	(3) Advance achieved in F_3	(4) F_3 advance as proportion of that possible	(5) $H^{\bar{V}}F_3$	(6) $H^{\bar{V}}\bar{F}_3$	(7) $H^{\bar{V}}\bar{F}_3/V_{F_3}$	(8) V_{F_3}	(9) $H^{\bar{V}}F_3$ without selection
1.	p=p'=0	2.4494364	1.3818803	56.39%	0.3519471	0.7975836	+.2875281	0.2347644	0.7500000
2.	p=p'=.2	2.7336128	1.3849045	50.67%	0.4951502	0.8201718	+.2126276	0.1619178	0.6375000
3.	p=.2, p'=.4	2.9277024	1.3861070	47.34%	0.5115346	0.8288063	+.1958051	0.1467948	0.6428571
4.	p=p'=.4	3.1622788	1.3906311	43.98%	0.5474609	0.8559070	+.1640591	0.1122159	0.6300000
5.	p=p'=.5	3.4661016	1.3956637	40.28%	0.6554360	0.8849089	+.1594499	0.1286670	0.7500000

Table 7.3: Results of selection for models of class I (Table 7.2) on transformed to basis of unit F_2- genotypic variance, with transformed value of environmental variance $V = 1.3333333$ units.

No. of the model	(1) Values of $p \Delta p'$	(2) Limit of selective advance	(3) Advance achieved in F_3	(4) F_3 advance as proportion of that possible	(5) $H^V F_3$	(6) $H^V \bar{F}_3$	(7) $H^V \bar{F}_3 / V_{F_3}$	(8) $V_{V_{F_3}}$	(9) $H^V F_3$ without selection
1.	$p=p'=0$	2.0000000	1.1277718	56.39%	0.3679697	0.5317217	-.1565107	0.1063397	0.0000000
2.	$p=p'=0.2$	2.2360710	1.1807718	52.57%	0.3901001	0.5467812	-.1157899	0.0719635	0.0350000
3.	$p=0.8, p'=0.4$	2.3904628	1.1317532	47.34%	0.3610231	0.5325375	-.1065829	0.0652421	0.0285714
4.	$p=p'=0.4$	2.5319934	1.1354471	43.93%	0.3649739	0.5706046	-.0393026	0.0293737	0.0333333
5.	$p=p'=0.5$	2.6284309	1.1395562	40.89%	0.4369573	0.5899393	-.0367937	0.0571853	0.0000000
Where $H^V F_2$ ab multiplied by		$\sqrt{b}$	$\sqrt{b}$	1	b	b	$b^{3/2}$	b^2	b

The properties, genotypic mean of selected F_3 progeny and the variance within F_3 progeny, i.e. $\bar{X}_3$ and $\bar{X}_2$ are directly calculated, while for the properties, variance of the F_3- progeny mean; the covariance of F_3 progeny mean and genotypic variance within progeny; and variance of the genotypic variance within progeny, the mean values are obtained by calculating the mean values of the quantities x_3^2, $x_2 x_3$ and x_2^2 and subtracting from them respectively the quantities $(\bar{x}_3)^2$, $(\bar{x}_2)(\bar{x}_3)$ and $(\bar{x}_2)^2$. The various properties so calculated have been presented in tables 7.3 and 7.4 for models of class I & II respectively. Table 7.3 gives the results of table 7.2 transformed to the basis of F_2- genotypic variance as unit and consequently with the changed environmental variance 1.3333333 units instead of 2.0 units. This table helps in transforming the results to the basis of any F_2 genotypic variance say b units, by making use of the suitable

Table 7.4: Results of selection for models of class II with two independent pairs of linked factors with magnitude of each factor, 'a' = .0660254 (fixed) and environmental variance v = 2.0 units (By Regression Method).

S.No. of the model	(1) Values of p Δ p'	(2) Limit of selective advance	(3) Advance achieved in F_3	(4) F_3 advance as proportion that possible	(5) $H^V F_3$	(6) $H^V \overline{F}_3$ I	(7) $H^V \overline{F}_3 / V_{F_3}$	(8) $V_{V_{F_3}}$	(9) $H^V F_3$ before selection	(10) $H^V \overline{F}_3$ without selection
1.	p=p'=0	3.4641016	2.2988938	66.36%	0.8249734	1.2363021	-.6418293	0.7342921	3.00	1.50
2.	p=p'=.2	3.4641016	1.9702155	56.88%	0.7236749	1.1003477	-.3794304	0.3659277	2.60	1.02
3.	p=.2, p'=.4	3.4641016	1.7897683	51.67%	0.6764418	1.0269017	-.2978967	0.2669533	2.10	0.90
4.	p=p'=.4	3.4641016	1.5986433	46.15%	0.6434285	0.9643320	-.2099346	0.1573103	1.80	0.73
5.	p=p'=.5	3.4641016	1.3956637	40.29%	0.6554360	0.8349039	-.1594499	0.1236670	1.50	0.75

Table 7.5: Results of selection for models of class II (Table 7.4) as transformed to basis of unit F_2- genotypic variance.

S.No. of the model	(1) Values of p Δ p'	(2) Limit of selective advance	(3) Advance achieved in F_3	(4) F_3 advance as proportion of that possible	(5) $H^V F_3$	(6) $H^V \overline{F}_3$	(7) $H^V \overline{F}_3 / V_{F_3}$	(8) $V_{V_{F_3}}$	(9) Environmental variance 'v'
1.	p=p'=0	2.0000000	1.3273053	66.36%	0.3033245	.4121007	-.1234303	0.0315330	.6666666
2.	p=p'=.2	2.2360710	1.2718452	56.88%	0.3015312	.4333615	-.1020366	0.0633291	.8333333
3.	p=.2, p'=.4	2.3904623	1.2330361	51.67%	0.3321151	.4387151	-.0977903	0.0603337	.9523809
4.	p=p'=.4	3.5819934	1.1932344	46.15%	0.3574608	.5356844	-.0369273	0.0483526	1.1111111
5.	p=p'=.5	2.8284309	1.1395562	40.29%	0.4369573	.5899393	-.0867037	0.0571853	1.3333333
Where $H^V F_2$=b multiplied by		$\sqrt{b}$	$\sqrt{b}$	1	b	b	$b^{3/2}$	b^2	b

Table 7.6: Results of selection for models of class II(Table 7.4) as transformed to basis of unit magnitude of each factor with transformed value of environmental variance v = 2.6666667 units

No. of the model	(1) Values of p & p'	(2) Limit of selective advance	(3) Advance achieved in F_3	(4) F_3 advance as proportion that possible	(5) $\bar{H}^V F_3$	(6) $H^V \bar{F}_3$	(7) $H^V \bar{F}_3 / V_{F_3}$	(8) $V_{V_{F_3}}$	(9) $H^V F_2$ before selection	(10) $\bar{H}^V F_3$ without selection
4.	p=p'=0	4.0	2.6545339	66.36%	1.2382978	1.6434028	-.9881608	1.2054031	4.0	2.00
4.	p=p'=.3	4.0	2.2750089	56.88%	0.9643909	1.4673969	-.5341713	0.6503381	3.2	1.36
4.	p=.2,p'=.4	4.0	2.0666407	51.67%	0.9019224	1.3634028	-.4331646	0.4743843	2.8	1.20
4.	p=p'=.4	4.0	1.8461852	46.15%	0.8579046	1.2856426	-.3282154	0.2796627	2.4	1.04
4.	p=p'=.5	4.0	1.6115736	40.29%	0.8789147	1.1793785	-.2454391	0.2287413	2.0	1.00

Here magnitude of each factor is multiplied by	a	a	1	a^3	a^2	a^3	a^4	a^3	a^2

Table 7.7: Results of selection for models of class II with two independent pairs of linked factors with magnitude of each factor 'a'=.8660254 unit (fixed) and environmental variance v = 2.0 units

(By Enumeration Method).

No. of the model	(1) Values of P & p'	(2) Limit of selective advance	(3) Advance achieved in F_3	(4) F_3 advance as proportion of that possible	(5) $\bar{H}^V F_3$	(6) $H^V \bar{F}_3$	(7) $H^V \bar{F}_3 / V_{F_3}$	(8) $V_{V_{F_3}}$	(9) $H^V F_2$ before selection	(10) $\bar{H}^V F_3$ without selection
4.	p=p'=0	3.4641016	2.3052286	66.55%	0.9235071	1.1648925	-.7998711	0.7606933	3.00	1.50
4.	p=p'=.2	3.4641016	1.9728062	56.94%	0.7833801	1.0549262	-.4374305	0.3647649	2.40	1.03
4.	p=.2,p'=.4	3.4641016	1.7779163	51.32%	0.6761783	1.0101768	-.3279038	0.2628812	2.10	0.90
4.	p=p'=.4	3.4641016	1.6004893	46.20%	0.6830030	0.9456496	-.2851622	0.1559891	1.80	0.78
4.	p=p'=.5	3.4641016	1.3932854	40.23%	0.6554416	0.8771788	-.1664021	0.1278017	1.50	0.75

power of b, which when multiplied with the results of this table given the transformed results. Likewise, the environmental variance will also become 'b' times to what it was before. The corresponding powers of 'b' have been given in the last row of the table 7.3. Table 7.6 gives the results of table 7.4 transformed to the basis of unit magnitude of each factor, with the changed environmental variance as 2.6666667(e3/3) units. Various values of F_2- genotypic variance calculated with the given magnitude of the factor and various pairs of p and p', the linkage fractions involved, have also been included in tables 7.4 and 7.6. Also Table 7.5 gives the results of selection corresponding to table 7.4, as transformed to the basis of unit F_2- genotypic variance, this table being similar to table 5.5 in chapter V.

As pointed out earlier the enumeration method was also simultaneously tried in the case of class II models with two pairs of linked factors, to bring out how much additional approximation is involved in the regression method. The results found by this method have been presented in table 7.7. The last row of table 7.6 gives the usual multipliers suitable to different F_2 proportion so as to bring the results to the basis of the genotypic value equal to 'a' units.

All these tables include also the maximum limits of the advance possible in different cases and also with the percentage of this limit of advance as has actually been achieved due to selection. The significance of these results has been discussed in the next chapter.

VIII. DISCUSSION OF THE RESULTS OF SELECTION

8.1. Having obtained the results of selection in different models, as given in tables 5.2 and 5.3 of chapter V for the models with one pair of linked factors and in tables 7.2, 7.4 and 7.7 of chapter VII for the models involving two independent pairs of linked factors, consideration is now given in this chapter to the influence of various factors and the intensity of linkage on these results. The main object of this dissertation as pointed out earlier is to investigate the effect of linkage on the results of selection. An attempt has been made in representing these results graphically against the intensity of linkage. For the results of the models with one pair of linked factors, where only one value of p, the linkage fraction is involved it is not difficult to put these results in a simple graph against various values of 'p'. On the other hand for the cases where two independent pairs of linked factors with two linkage fractions p and p' are involved it is not possible to represent the results graphically in a simple manner. The approach of representing them in three dimensional space is not easily generalisable to the nth dimensional space for the case where a large number of p's are involved. However an attempt was made to find a suitable function of p's which can be used to represent a measure of the overall linkage intensity for the cases where a number of independent p's are involved.

In the present models where independent pairs of linked factors are considered, there is no interference among the various pairs when the process of crossing over is taking place. The various cross-overs occur completely at random and independently. Owen (1950) defines p, the recombinant fraction as the probability of occurrence of an odd number of

exchange points on the interval between the two loci. This idea can be extended to the present case, where occurrence of an odd number of exchange points takes place simultaneously in more than one independent pair of linked factors. The probability in this case becomes the simple product of various p's, the recombinant fractions and can be represented by Πp_i. The main use of this expression lies in the fact that it gives a single measure of the amount of linkage present among the various factors. This function may not work in case anyone of the p's happens to be zero in which case the corresponding two completely linked factors can be taken in the model to be acting as a single factor with double the magnitude for other factors which may enable us to reduce this expression for the remaining independent linked pairs of factors. Actually there is a need of formulating of such a function, which may work even for such situations also.

8.2. It is interesting to note that the models with no linkage (i.e. $p = \frac{1}{2}$ for one linked pair and $p = p' = \frac{1}{2}$ for two independent linked pairs) in the two classes of models i.e. for models with fixed F_2-genotypic variance and models with fixed magnitude of the factors, are identical. It is this common point which has been used here as the basis for comparing the results of selection for the two classes of models graphically. For each F_2- statistical property the graphs have been plotted against the recombinant fraction p for the models with one linked pair (Figures 8.1 to 8.5) and against Πp_i for the models with two independent linked pairs (Figures 8.6 to 8.10) separately. The corresponding graphs for the models of both classes, have been plotted on the same graph with the same scale in order to make comparison on the basis of the common point mentioned above.

It is important to note that in nearly all the graphs for models involving two independent pairs of factors and with fixed F_2-genotypic variance, there are slight irregularities near the points for the models with unequal values of p and p'. This may be due to the unequal effect of linkage values p and p' on the magnitude 'a' of the factors. However for the second class of models where the magnitude 'a' has been kept fixed, such irregularities in the graphs are absent and the graphs are quite smooth.

3.3. The results of selection for each of the five F_3- properties will be discussed separately, one by one in the following paragraphs in the light of how linkage with its varying intensities affects them.

The first property we shall consider is $\bar{x}_3$, the genotypic mean of the selected progenies i.e. of progenies resulting from F_2 individuals in the selected portion of the F_2 population. This mean also represents the expected value of the phenotypic mean of those progenies and is a property of considerable practical interest in that it gives the immediate gain from selection. Various tables in chapters V and VII give results for this property namely, the mean of F_3 population measured as deviation from the mean of the unselected F_3 population which in the present case is zero for all models, due to the absence of dominance. Thus the mean of the F_3 population will itself give the genetic advance due to selection. These tables along with the results for other F_3- properties, also give the maximum limit of selective advance for various models.

Starting first with the discussion of the results of selection for models of class I (Tables 5.2 and 7.2), it is observed that with same F_2 genotypic variance, limits of selective advance are greater for all cases of linkage when the variance results from four factors

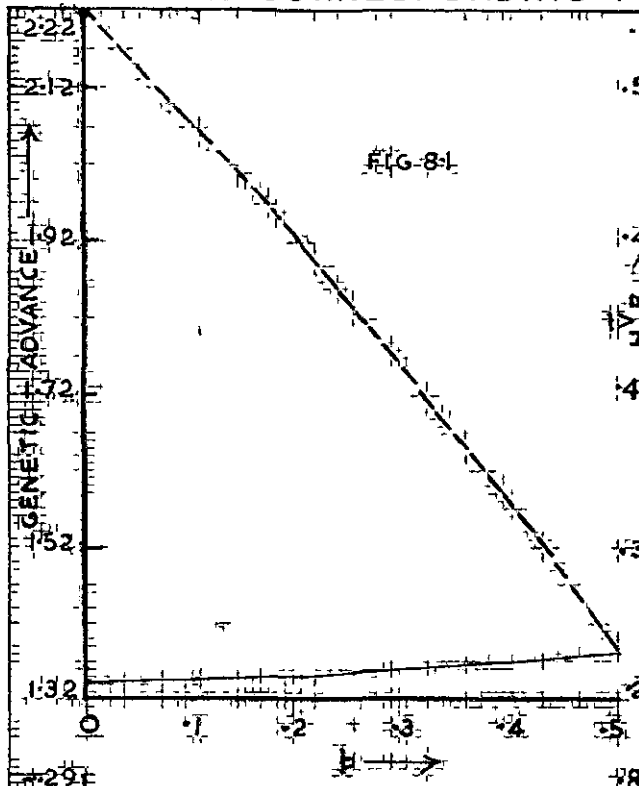
instead of two. In each case the limit increases with the loosening of linkage as is expected.

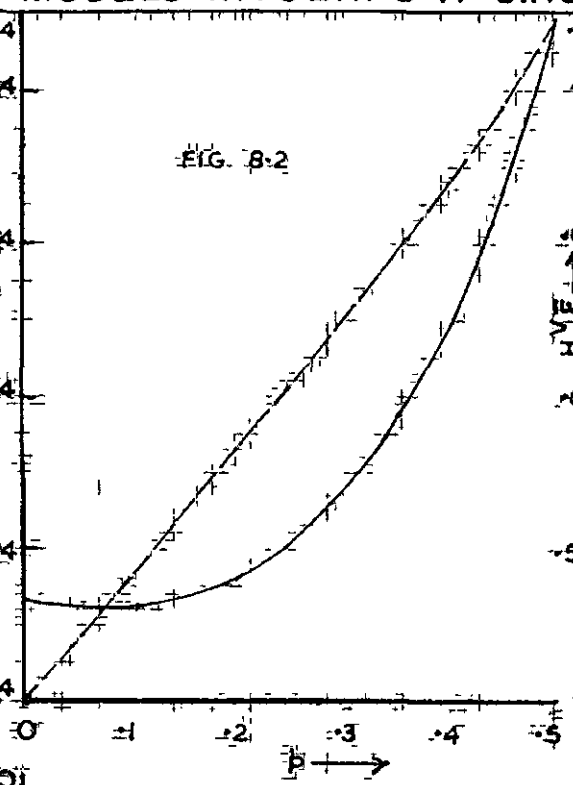
Although the advance achieved in F_3 is nearly same, yet with four factors it is slightly higher than that with two factors, the difference being least with no linkage in each case. On complete loosening of linkage the percentage increase in the advance achieved, over the case with completely tight linkage is almost 1.4% for models with two pairs of linked factors (Table 7.2) as compared to 2.6% for models with a single pair (Table 5.2). It appears from this that with the increase in number of factors the rate of increase in genetic advance goes on becoming inappreciable for this class of models. This fact has been very well depicted by the slopes of the corresponding lines in Figures 8.1 and 8.6, where there is a slightly more perceptible increase in advance achieved with two factors than with four factors. The result is that with only two factors advance achieved is greater percentage of limits to be achieved than with four factors and shows a wider range (56 to 77%) in the former than in the latter case (40 to 56%).

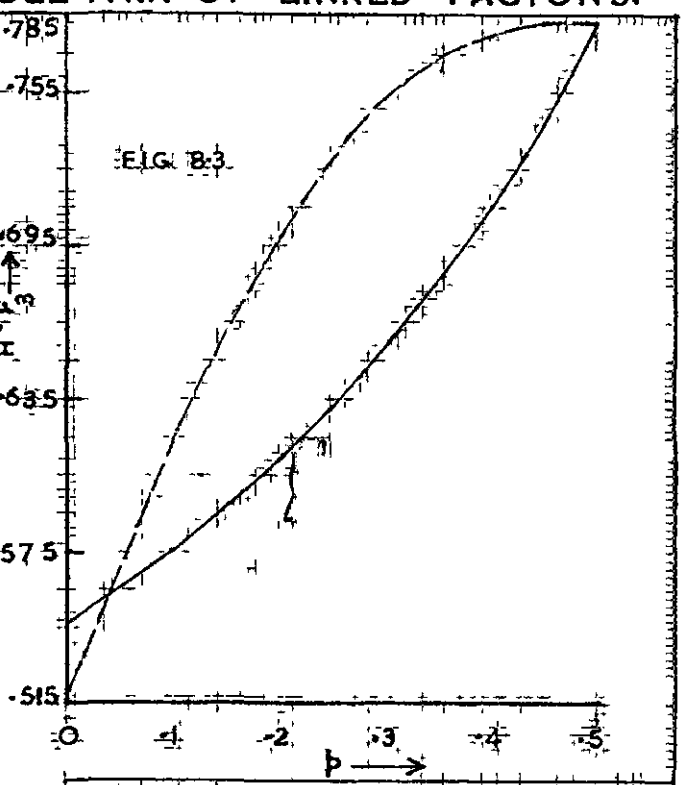
Again as expected the advance in F_3 for a pair of factors with no linkage is identical with advance with two pair of factors with complete linkage in each pair. This is true for the results of other F_3- properties also.

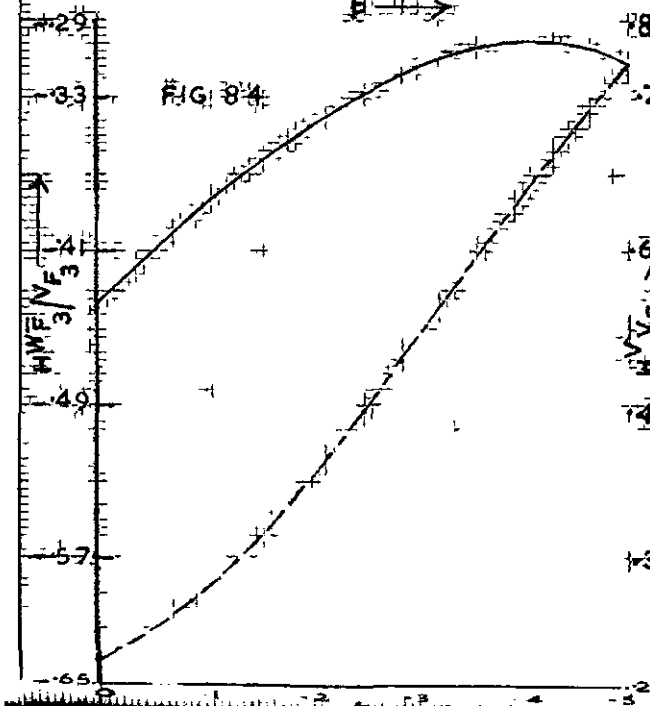
We now proceed to study the results for class II models (Tables 5.3 and 7.4). Here the limits of selective advance are fixed when the size of factors is fixed irrespective of F_2 genotypic variance or strength of linkage. The value of the limit is lower with two factors each of larger size than with four factors each of smaller size.

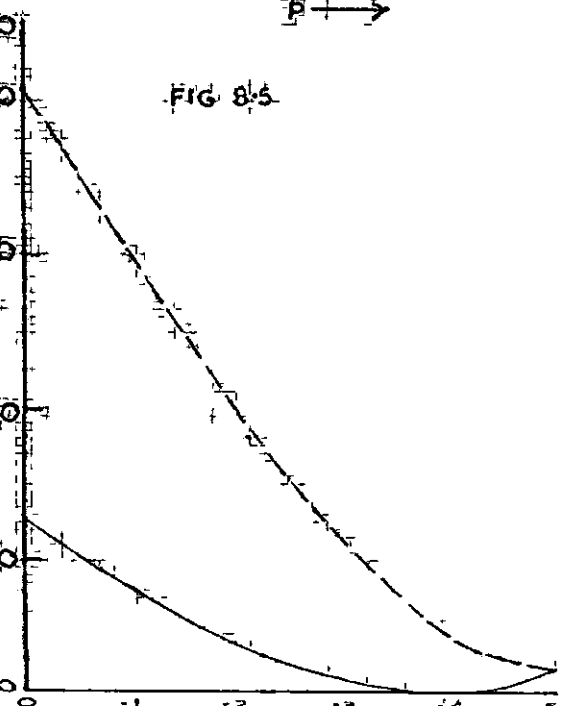
FIGURES CORRESPONDING TO MODELS INVOLVING A SINGLE PAIR OF LINKED FACTORS.











NOTE: UNDOTTED GRAPHS CORRESPOND TO CLASS I MODELS & DOTTED CORRESPOND TO CLASS II MODELS

FIGURES: 8·1 TO 8·5 SHOW THE RELATIONS BETWEEN RECOMBINANT FRACTION 'p' AND F_3 PROPERTIES NAMELY;

1. GENETIC ADVANCE ACHIEVED;
2. MEAN GENOTYPIC VARIANCE WITHIN F_3 PROGENIES;
3. VARIANCE OF PROGENY MEAN;
4. COVARIANCE OF F_3 PROGENY MEAN AND GENOTYPIC VARIANCE WITHIN F_3 PROGENIES;
5. VARIANCE OF GENOTYPIC VARIANCE WITHIN F_3 PROGENIES RESPECTIVELY.

With factor also fixed to correspond to F_2 variance of 1.5 units for independent segregation in either case, the F_2 genotypic variance in models, where linkage between these factors gets closer also steadily increases and so does the genetic advance achieved in F_2, so that with closer linkage the advance steadily goes up. The relation is linear where only a single pair of linked factors is involved (Figure 8.1) and slightly curved showing lesser rate of increase for intermediate values of linkage, when two pairs of factors operate.

The most important contrast with class I models in both cases is the steep change in advance achieved as degree of linkage and consequently F_2- variance changes in class II models as compared to almost a fixed value of advances for all class I models corresponding to fixed F_2 genotypic variance. Advance is much greater for all models in class II corresponding to larger F_2 variance in class I. As with class I models advance achieved with four factors is slightly higher than with two factors and as before percentage advance achieved is greater and has a wider range (56 to 90%) with two factors than with four (40 to 66%).

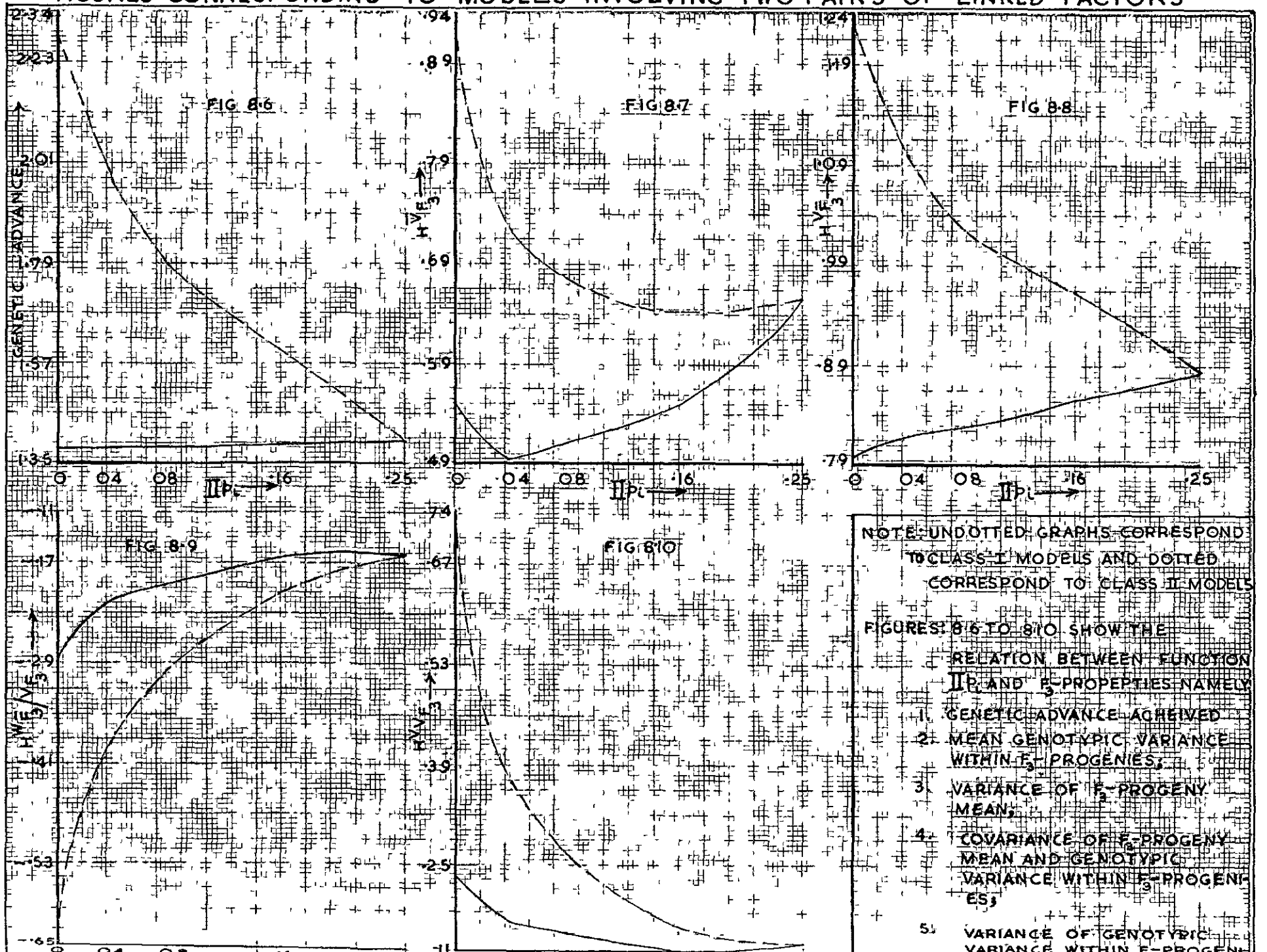
It may be concluded that for the models with fixed F_2 variance, actual advance is also fixed and the maximum attainable advance depends on the magnitude 'a' of each factor, whereas for the models with magnitude as fixed, the maximum attainable advance is fixed and the advance achieved depends upon the F_2 variance where it increases with the increase in the F_2 genotypic variance. This is also confirmed from the various values of F_2 genotypic variance and the corresponding values of genetic advance achieved, presented in the tables 5.6 and 7.5 showing the transformed results. Further, the effect of environmental variance on the advance as can be seen from the tables 5.5 and 7.5 is that the advance decreases with the increase in environmental variance.

8.4. We now switch on to the next two properties namely, the mean variance within selected F_3 progenies and variance between means of these progenies. These properties are of vital interest for plant breeder in that the variance within selected F_3 progenies serves as an index of scope for improvement through single plant selection within progeny, and that between means of these progenies indicates the scope for improvement at the stage of progeny selection.

From the results for class I models presented in tables 5.2 and 7.2, it is observed that the mean variance within selected F_3 progenies is appreciably greater with four factors than with two factors, showing that number of segregating factors influence this property even when starting with identical F_2 variance. With loosening of linkage the change is however much faster with two factors and covers a wider range than with four factors (Figures 8.2 and 8.7). In other words when number of independently segregating factors changes from one to two the rate of change in within F_3 variance is greater than when it changes from two to four. The downward dip, which is somewhat greater with four factors in both curves (Figures 8.2 and 8.7) is to be accounted for by the fact that while in both cases, the variance within F_3 progenies at both extremes of linkage values ($p = 0$ and $p = .5$) is .75 units in the unselected population being half the initial F_2 variance, this variance within F_3 in the unselected population is lower in all cases of partial linkage (columns 9 of tables 5.2 and 7.2) and counteracts to some extent the trend towards increased variance within selected F_3 with low linkage values.

Similar remarks apply to various features of variance between F_3 progenies in the selected portion of the population, except that there is a steady increase in variance throughout as linkage becomes

FIGURES CORRESPONDING TO MODELS INVOLVING TWO PAIRS OF LINKED FACTORS



looser both in models with two and four factors without an intermediate depression (Figures 8.3 and 8.8).

Refering to tables 8.3 and 7.6, giving the results for class II models, it is observed that for these models also where the factor size is fixed, the variance within selected F_3 progenies as well as between means of these progenies, is dominated by the number of factors, both variances being substantially larger where four factors of smaller size operate as compared to two factors of larger size. In two factors models both variances increase steadily with loosening of linkage, the relationship being linear for within progenies variance (Figure 8.2) and somewhat curved showing looser rate of increase with very loose linkage for between F_3 variance (Figure 8.3). The general trend is similar to that in two factors models with F_2 genotypic variance fixed (class I models). In four factor models of class II however the trend is completely reversed for both variances, the variances steadily decreasing with loosening of linkage (Figures 8.7 and 8.8). This contrast can be explained by the relative influence of number of factors and mean within F_3 progeny variance in the unselected population in the two sets of models. In two factors models the change from one to two factors with loosening of linkage dominates over lowered mean within progeny variance in unselected F_3, while the change from two to four factors is not sufficient to counteract the lowered mean variance (columns 10 of tables 8.3 and 7.4).

Since variance within selected F_3 progenies and variance between means of these progenies is index of potentiality for further response to selection in this material, it is clear that looser linkage provides greater scope for further selection in all cases considered except the last one namely, four factor models with fixed factor size.

Again as is seen from tables 5.5 and 7.5, greater environmental variability in relation to the genotypic leads to a greater variance both within and between F_3 progenies. This is to be expected, for greater environmental variability would bring together more diverse genotypes into the selected portion of the population.

8.3. The results for covariance of the progeny mean and genotypic variance within F_3 progeny may be considered next. The importance of this property lies in that it explains the relation between the advance already achieved and the potentiality for further advance. For all the models considered here, the covariance is negative which shows that higher the advance is achieved in F_3, the less is left the scope for further advance.

For class I models, the covariance is numerically less with four factors than with two factors (Tables 5.2 and 7.2) showing that the number of segregating factors influence this property even when starting with same F_2 genotypic variance. However, in both cases, it decreases numerically with loosening of linkage, except that it slightly increases with no linkage for case where two factors are involved (Figures 8.4 & 8.9).

For class II models, where the factor size is fixed, the covariance is considerably affected by the increase in the number of factors, it being numerically less where four factors of smaller magnitude operate as compared to two factors each of larger size, except when the linkage is completely tight for the two cases where the covariance is almost same (Tables 5.3 and 7.4). In two factor models the covariance decreases (numerically) steadily with loosening of linkage and the relation being linear (Figure 8.4). The same trend is maintained for the four factor models except that the relation is curved showing lesser rate of numerical decrease with loosening of linkage (Figure 8.9).

It appears from the above models, that with increase in the number of factors, the covariance decreases numerically irrespective of whether the magnitude of the factors is fixed or the F_2 genotypic variance is kept same. It is thus clear that the correlation between progeny mean and variance within progeny would serve as a good index for the number of factors operating in F_2 segregation. With the small value of this correlation whether positive or negative, it might be inferred that the number of factors is rather large. On the other hand this number will be indicated as small if the correlation is significant.

8.6. We now proceed to the last property namely, variance of genotypic variance within F_3 progeny. It was first thought of to study the various results of selection with respect to the effective number of factors, the estimate of which is given by Panse (1940 a) where the use of the above property has been made. It is mainly for this reason, that the property namely, variance of variance within F_3 progeny has been studied. But later it was found that the results of selection for various other properties do not show a smooth relationship with this property, particularly it is so far the class I models with four factors. The idea of studying the results of selection against this property was thus dropped. An attempt was then made to construct a suitable function of p's which can be used to represent a measure of the overall linkage intensity for the present case. The utility of this function has been discussed in the first article of this chapter.

Another use of variance of genotypic variance within F_3 progeny is that it enables us to calculate the correlation between progeny mean and variance within progeny in order to give some idea about the relation between the advance achieved and the scope of further advance. Also it will give some idea about the number of factors involved, as discussed in the previous section. The results for this propety have been tabulated

along with those for other proportion and have been presented graphically also (Figures 8.5 and 8.10).

As mentioned earlier, the results of selection for class II models with four factors have been obtained by enumeration method apart from the general regression method given by Panse (Tables 7.4 and 7.7). It was done so in order to see how much approximation is involved in the general method as compared to the enumeration method. A close study of the results shows that the results for the properties (i) Genotypic mean of F_3 progeny (ii) Mean genotypic variance within F_3 progeny and (iii) Variance of genotypic variance within F_3 progeny are almost the same for both methods. The difference is shown for the remaining two properties namely, variance of F_3 progeny mean and the covariance of the F_3 mean and variance within F_3 progeny. These proportion are slightly over-estimated in the regression method as compared to the enumeration method. It is justifiable to say that one may have to work out the regression polynomial to a higher degree with the help of moments of higher order, for these proportion.

SUMMARY

The most important problem in plant breeding is to determine the speed of genetic advance under selection. This was investigated by Fenco, who introduced the method of genetic models which he considered with independently segregating factors. In the present investigation following the statistical approach given by him a similar study has been made of genetic models with linked factors. The effect of linkage on the speed of genetic advance and various F_3- statistical properties after making selection in F_2, has been studied for models with one linked pair and with two independent pairs of linked factors. The various factors have been taken to be equal in magnitude. In one set of these models effect of linkage in coupling phase has been studied keeping the magnitude of each factor fixed, while in the other set the effect of keeping the F_2 genotypic variance in the start (before selection) fixed, has been studied. The various values of the recombinent fractions have been so chosen, so as to cover the whole range of linkage .

The various F_3 statistical properties studied were - (1) the genotypic mean of F_3 progeny (2) mean genotypic variance within F_3 progeny (3) the variance of the genotypic mean of F_3 progeny (4) the covariance of the F_3 progeny mean and variance within progeny and (5) the variance of genotypic variance within F_3 progeny. For the models with a single linked pair of factors corresponding to the both sets of models, since the number of genotypic classes was not large, the results were calculated using the direct method of enumerating the genotypes, whereas for the models involving two independent linked pairs the regression method given by Fenco was adopted. For the latter type of models, the moments and

product-moments of the three variates, x_1, F_3 phenotypic value, x_2, genotypic variance within F_3 progeny and x_3, the genotypic mean of F_3 progeny, were obtained for each model and the regression equations of five F_3 - properties on the F_2 phenotypic value were calculated. These equations enable the study of the effects of various characteristics of the models on the relation between F_3 properties and F_3 parental value.

With the help of these regression equations the values of each of the five F_3 properties were calculated for 10% level of selection. To obtain these results the mean value for an F_3 property was expressed in terms of the regression coefficients, corresponding to that property, cumulants of the F_2 distribution and hermite polynomials corresponding to the level of selection. The effects of various aspects of the genetic set up on the results of selection were considered.

It is concluded that for the models with fixed F_2 genotypic variance, actual advance is also practically fixed and the maximum attainable advance depends on the value of each factor, whereas for the models with magnitudes of factors as fixed, the maximum attainable advance is fixed and the advance achieved depends upon the F_2- genotypic variance where it increases with the increase in F_2 genotypic variance.

In models with fixed magnitude of the factors, the variance between as well as within F_3 progenies decreases with closer linkage, when only one pair of linked factors is considered, while it increases when two pairs of linked factors are segregating. In models with the initial F_2- genotypic variance fixed, the results are consistent for one and two pairs of linked factors. Here the variance between and within F_3 progenies decreases with closer linkage. Further, the covariance of progeny mean and genotypic variance within progeny numerically increases with closer linkage, irrespective of whether the initial F_2 genotypic

variance or the magnitude of the factors is kept fixed. However, this increase is in general, more in the latter case than in the former. These were the broad conclusions derived from the results.

In the present work the possibility of using a single index for linkage when more than one pair of linked factors is segregating has also been considered. Again, the results of selection for models with two independent linked pairs of factors, where the genetic magnitude of the factors is kept fixed, have been obtained by the direct method of enumerating the genotypes, apart from the general regression method given by Panse. It was done so in order to see how much approximation is involved in the general method as compared to the enumeration method.

REFERENCES

Bokil, S.D. (1954) "A statistical study of quantitative characters with special reference to plant breeding". Jour. Ind. Soc. Agric. Stat. Vol VI. No.2.

Cornish, E.A. and Fisher, R.A. (1937) "Moments and cumulants in the specification of distributions". Revue de l'Institute international de statistique, 5, 1.

Fisher, R.A. (1918) "The correlation between relatives on the supposition of Mendelian inheritance". Trans. Roy. Soc. Edin. 52.

_____ (1930) "The genetical theory of natural selection". Clarendon Press, Oxford.

_____, (1932) "The genetical interpretation of statistics of the third degree in the study of quantitative inheritance". Genetics, 17.
Imer, F.R. and Todin, C.

Goodwin, R.H. (1944) "The inheritance of flowering time in a short-day species, Solidago sempervirens L". Genetics, 29.

Hutchinson, J.B. and Panse, V.G. (1937) "The design of field tests of plant breeding material". Ind. Jour. Agric. Sci. 7, 331.

Johannsen, W. (1909) "Elemente der exakten Erblichkeitslehre". Fischer Jena.

Kempthorne, O. (1957) "An introduction to Genetic Statistics". John Wiley & Sons, New York.

Mather, K. (1941) "Variation and selection of polygenic characters", Jour. Genetics, 41.

_____ (1943) "The balance of polygenic combinations". Jour. Genetics, 43.

_____ (1949) "Biometrical Genetics". Methuen & Co. Ltd London.

Owen, A.R.G. (1950) "The theory of genetical recombination". Advances in Genetics, Vol. III.

Panse, V.G. (1940a) "The inheritance of quantitative characters and plant breeding". Jour. Genetics, 40.

_____ (1940b) "A statistical study of quantitative inheritance". Ann. Eug. 10.

Pearson, K. (1930) "Tables for Statisticians and Biometricians" Part I, 3rd edition, Cam. Univ. Press.

Powers, L. (1941) "Inheritance of quantitative characters in crosses involving two species of Lycopersicon". Jour. Agric. Res. 63.

_____ (1942) "The nature of the series of environmental variances and the estimation of the genetic variances and the geometric means in crosses involving species of Lycopersicon". Genetics, 27.

_____ (1950) "Determining scales a nd the use of transformations in studies of weight per locule of tomato fruit". Biometrics, 6.

_____ (1953) "Components of variance method and partitioning method of genetic analysis applied to weight per fruit of tomato hybrid and parental populations". Tech. Bull. U.S. Dep. Agric., 1131.

Rasmusson, J.M.	(1935)	"Studies on the inheritance of quantitative characters in Pisum-I. Preliminary note on the genetics of flowering". Hereditas, 20.
Sax, K.	(1923)	"The association of size differences with seed-coat pattern and pigmentation in Phaseolus vulgaris". Genetics, 8.
'Student'	(1934)	"A calculation of the minimum number of genes in Winter's selection experiment". Ann. Eugenics, 6.
Wright, S.	(1934)	"The results of crosses between inbred strains of Guinea pigs, differing in number of digits". Genetics, 19.

PART II

" A STUDY ON LINKAGE BETWEEN GENES FOR PIGMENTATION IN VARIOUS PARTS OF THE RICE PLANT "

A STUDY ON LINKAGE BETWEEN GENES FOR PIGMENTATION IN VARIOUS PARTS OF THE RICE PLANT.

1.1. This work deals with the data, received from Central Rice Research Institute, Cuttack, on linkage between genes for pigmentation in various parts of the rice plant. It is the F_2 - data showing segregation for individual characters and joint segregation for pairs of characters in one cross viz Ruiho X Dular. It has been stated that some absurd cross-over values have been obtained, by using both (i) Maximum likelihood method and (ii) Minimum discrepancy method, given by Haldane (1953). In the second method, the expressions for the estimates given by Murty (1954) are stated to have been used.

1.2. It has been found from a careful scrutiny of the original data received, given in table I that some of the observations are inconsistent, so far as the individual segregation ratios are concerned. The individual segregation ratios found from the observed joint segregation given in table I, do not tally with the individual segregation ratios, given in the data (table 2) for a particular part of the plant. For example, for the individual segregation derived from the joint segregation, for the part tip and margin, it is seen from the observations, that in 6 cases out of 8 cases, 169 plants show pigmentation while in the remaining two cases, 171 plants show this character. The only explanation which can be given is the misclassification of two of plants in the latter cases. The two plants have been wrongly placed in the pigmented class. Under these circumstances the observations for the joint segregation have been adjusted so as to give an identical ratio for each individual segregation in all the cases, as forming more reliable basic data for the test of linkage

and estimation of its value. The cases involving the part internode have not been considered at all, as there is no possibility of arriving at a definite individual segregation ratio from the data, which is obvious from the variation in the number of plants with internode pigmented obtained in different cases of the joint segregation, namely 213, 203, 163, 218, 214, 199 and 169. These numbers are not at all mutually consistent. Thus the part of data wherever internode is concerned has been omitted from consideration. The remaining data, after making appropriate adjustments have been given in table 6. There were in all 36 cases out of which 8 cases for internode have been omitted and hence 28 cases are to be investigated for linkage.

2.1. The various hypothetical F_2 ratios suggested are 3P: 1G; 9P: 7G; 27P: 37G; and 162P: 94G where P stands for pigmented and G for green. These have been explained on the following Mendelian hypotheses.

AA is a basic gene pair essential for production of pigmentation.

Cc is another basic gene called the chromogen, complementary to AA, and absence of either of them will withhold production of pigment.

For the production of pigment in any plant part, besides the two basic genes, one or more localisation genes are essential before any pigment can develop in that particular part.

The hypothetical genotypes of the two parents are given below:

Zuiho: AA; cc; $\underline{lep_1\ lep_1}$; $\underline{lep_2\ lep_2}$; $\underline{lep_3\ lep_3}$; $\underline{lmp\ lmp}$; $\underline{lgp\ lgp}$; $\underline{jbp\ jbp}$; $\underline{lxp\ lxp}$; $\underline{ntp_1\ ntp_1}$; $\underline{ntp_2\ ntp_2}$; $\underline{ntp_3\ ntp_3}$; $\underline{smp\ smp}$; $\underline{Ap\ Ap}$; $\underline{sp\ sp}$

Dular: $\underline{AA\ CC}$ $\underline{lep_1\ lep_1}$ $\underline{lep_2\ lep_2}$ $\underline{lep_3\ lep_3}$ $\underline{lmp\ lmp}$ $\underline{lgp\ lgp}$ $\underline{jbp\ jbp}$ $\underline{lxp\ lxp}$ $\underline{ntp_1\ ntp_1}$ $\underline{ntp_2\ ntp_2}$ $\underline{ntp_3\ ntp_3}$ $\underline{smp\ smp}$ $\underline{A_p\ A_p}$ $\underline{s_p\ s_p}$

Where lep lep is the gene pair governing the character of pigmentation for the part leafsheath, lmp lmp for tip Margin, lgp lgp for

ligule, Jbp Jbp for juncture back, Lap Lap for leaf axil, Nt_p Nt_p for Internode, Sp Sp for septum, Ap Ap for apiculus and Gap Gap for Stigma. The localisation genes Lap Lap and Ap Ap are not segregating in either of the two parents and they have been shown above in the genotypes of the two parents, as they were suggested. It has been assumed that there is complete dominance among two alleles of any of allelic pairs.

The following hypothetical ratios have been set up for various pigmentation characters.

(i) ratio 3P : 1G for the parts Apiculus and Leaf Axil:- The two parents differ only in one gene pair which in this case is the C-c pair.

(ii) Ratio 9P : 7G for the parts tip Margin, ligule, Juncture back & stigma:-

The parent differs in C-c gene pair and one localisation gene pair.

(iii) Ratio 27P : 37G for the part septum:- The parents differ in C-c and two localisation gene pairs.

(iv) Ratio 162P : 94G for the parts leaf sheath & internode:- The two parents differ in C-c gene pair and three localisation gene pairs. Pigmentation is produced when both the basic genes are present plus at least two of the localisation genes. Thus for leaf sheath the F_1 will be AA Cc Lap_1 lap_1 Lap_2 lap_2 Lap_3 lap_3 .....

These hypothetical ratios can be verified in the different cases.

On the basis of these ratios, the individual segregations of different parts have been tested and the results have been shown in table 3. The data do not show any evidence through χ^2- test against these hypothetical ratios for different parts.

3.2. It is important to note that no two of the above hypothetical ratios are independent of each other, as the gene pair C-c is common to

every ratio. Thus for getting the joint hypothetical ratio for a joint segregation it is not correct simply to multiply the two ratios but this joint hypothetical ratio has to be calculated seperately, taking the common C-c pair into account. So the expressions given by Murty (1954) for the estimate of the cross-over value in minimum discrepancy method, are not applicable here, as the mutual independence of a pair of ratios has been assumed there. The correct joint expected ratios have been shown in table 6 for every pair of parts. As shown in the same table a test for linkage has been carried out in each of the 28 cases. It has been found out that out of 28 cases, 15 cases do not give any evidence of linkage. The cross over value has therefore to be estimated in the remaining 13 cases as shown in the same table. The different cases have been dealt in the following paragraphs for estimation of the cross-over values, with their standard errors, there being in all four types of combinations.

(A) Ratio 9:7 versus ratio 9:7

(B) Ratio 9:7 versus ratio 27:37

(C) Ratio 9:7 versus ratio 162:94

and lastly (D) Ratio 27:37 versus ratio 162:94

3.1. Ratio 9:7 versus ratio 9:7 = type (A):-

Considering the case of tip Margin (9:7) versus ligule (9:7) following are the observed frequencies and their expected ratio for the combined segregation.

Frequencies or ratio	Both pigmented	Only T.M. Pigmented	Only ligule pigmented	None pigmented	Total
Observed frequencies	n_1=138	n_2=11	n_3=14	n_4=134	317=n
Expected	$\frac{3}{16}(2+\theta)=m_1$	$\frac{3}{16}(1-\theta)=m_2$	$\frac{3}{16}(1+\theta)=m_3$	$\frac{1}{16}(4+3\theta)=m_4$	1

where $\theta = (1-p)^2$, p being the cross over value between the two linked factors, ni and mi are the observed and expected frequencies respectively.

(i) Estimation of θ :- following are the two methods used for estimation:

(a) Maximum likelihood method:-

The logarithm of likelihood expression is given by $L = n_1 \log \frac{2}{16}(2+\theta) + n_2 \log \frac{2}{16}(1-\theta) + n_3 \log \frac{2}{16}(1-\theta) + n_4 \log \frac{1}{16}(4+3\theta)$. Maximising it, w.r. to θ, we have

$$\frac{dL}{d\theta} = \frac{n_1}{2+\theta} + \frac{n_2+n_3}{1-\theta} + \frac{3n_4}{4+3\theta} = 0 \quad \cdots\cdots\cdots (1)$$

which reduces to the quadratic

$$3n\theta^2 + (n_1+10n_2+10n_3+3n_4)\theta - 4n_1 + 8n_2 + 8n_3 - 6n_4 = 0.$$

Substituting the numerical values of n's we have

$951\,\theta^2 + 810\,\theta - 1236 = 0$, giving $\hat{\theta} = .9911146$

$\therefore \hat{p} = 1 - \sqrt{\theta} = 1-\sqrt{.9911146} = 1-.8894 = .111$

(b) Haldane's minimum discrepancy method:-

The expression given by Haldane (1933), which is to be minimised for estimation of θ, is given as

$$\sum_p \frac{[f_p(\theta)]^2}{n_p+1} - \sum_p \frac{(nm_p)^2}{n_p+1}$$ In this case, this expression is equal to $\left\{\frac{3n}{16}(2+\theta)\right\}^2 \Big/ (n_1+1) + \left\{\frac{3n}{16}(1-\theta)\right\}^2 \Big/ (n_2+1) + \left\{\frac{3n}{16}(1-\theta)\right\}^2 \Big/ (n_3+1) + \left\{\frac{n}{16}(4+3\theta)\right\}^2 \Big/ (n_4+1)$, Differentiating it w.r. to θ and equating it to zero, we have on solving it for θ,

$$\hat{\theta} = \frac{\left(\frac{1}{n_2+1} + \frac{1}{n_3+1} - \frac{4}{3(n_4+1)} + \frac{2}{n_1+1}\right)}{\sum_p \frac{1}{(n_p+1)}}$$

substituting the numerical values of n_p's we have

$\hat{\theta} = .7791533$

$\therefore \quad \hat{p} = 1-\sqrt{\theta} = 1-\sqrt{.7791533} = 1-.882 = .118$

(ii) Calculations for variance of (p) in the above two cases:-

(a) In Maximum likilihood method, we know $-\frac{1}{V_\theta} = E\left(m_p \frac{d^2 \log m_p}{d\theta^2}\right)$.

Differentiating left hand expression of (1), w.r.t. to θ and substituting expected values of n_p i.e m_p, we have

$-\frac{1}{V_\theta} = -\frac{3n}{16}\left(\frac{1}{2+\theta} + \frac{2}{1-\theta} + \frac{3}{4+3\theta}\right)$ which gives $V_\theta = \frac{8}{3n} \frac{(2+\theta)(1-\theta)(4+3\theta)}{(3\theta+13)}$

Also since $\theta = (1-p)^2 \quad \therefore \quad V_p = \frac{V_\theta}{\left(\frac{d\theta}{dp}\right)^2} = \frac{V_\theta}{4\theta}$

$\therefore \quad V_p = \frac{2(2+\theta)(1-\theta)(4+3\theta)}{3n\theta\ (3\theta+13)}$ As estimated in (i) a

$\hat{\theta} = .791, \quad \therefore V_{\hat{p}} = \frac{2(2.791)(.209)(6.373)}{3x317(.791)(19.328)} = .00051137$

which gives S.E. $(\hat{p}) = \sqrt{V_{\hat{p}}} = .0226$

(b) Using minimum discrepancy method, it has been found above that

$$\hat{\theta} = \frac{\frac{1}{n_2+1} + \frac{1}{n_3+1} + \frac{4}{3(n_4+1)} + \frac{2}{n_1+1}}{\sum_r \left[\frac{1}{n_r+1}\right]}$$

let $F = \log \hat{\theta} = \log \left\{ \frac{1}{n_2+1} + \frac{1}{n_3+1} + \frac{4}{3(n_4+1)} + \frac{2}{n_1+1} \right\} - \log \left(\sum_r \frac{1}{(n_r+1)} \right)$

Again as given by Nathor (1951)

$$\frac{1}{n} V_p = S\left\{ n_p E \left(\frac{\partial F}{\partial n_p}\right)^2 \right\} - E\left(\frac{\partial F}{\partial n}\right)^2$$

$$= S\left\{ n_p E \left(\frac{\partial F}{\partial n_p}\right)^2 \right\} \qquad \text{---------(2)}$$

as $\frac{\partial F}{\partial n} = 0$ in the case under consideration.

In this case $n_1 E\left(\frac{\partial F}{\partial n_1}\right)^2 = \frac{3}{16}(2+\theta)\ E \left\{ \frac{\frac{3}{(n_1+1)^2}}{\frac{1}{n_2+1}+\frac{1}{n_3+1}+\frac{4}{5(n_4+1)}+\frac{3}{n_1+1}} + \frac{\frac{1}{(n_1+1)^2}}{\sum_r \frac{1}{n_r+1}} \right\}^2$

$$= \frac{3}{16}(2+\theta) \left\{ \frac{\frac{3}{\left\{1+\frac{3n}{16}(2+\theta)\right\}^2}}{\left(\frac{1}{1+\frac{3n}{16}(1+\theta)}+\frac{1}{1+\frac{3n}{16}(1+\theta)}+\frac{4}{5\left\{1+\frac{n}{16}(4+3\theta)\right\}}+\frac{3}{1+\frac{3n}{16}(2+\theta)}\right.} + \right.$$

$$\left. \left. \frac{\frac{1}{\left\{1+\frac{3n}{16}(2+\theta)\right\}^2}}{\frac{1}{1+\frac{3n}{16}(2+\theta)}+\frac{2}{1+\frac{3n}{16}(1+\theta)}+\frac{1}{1+\frac{n}{16}(4+3\theta)}} \right\}^2 \right\}$$

Similarly the expressions for the remaining terms like $n_p E\left(\frac{\partial F}{\partial n_p}\right)^2$ can be found out giving thereby the value of V_p.

But since $F = \log \theta$, $V_{\hat{\theta}} = \frac{V_F}{\left(\frac{dF}{d\theta}\right)^2} = \theta^2 V_F$, Also $\theta = (1-p)^2$

$$\therefore\ V_{\hat{p}} = \frac{V_\theta}{d\theta} = \frac{\theta^2 V_F}{d\theta} = \frac{\theta}{4} V_F$$

substituting the numerical values of θ and n in (2), we have

$V_F = .004265$ and hence $V_{\hat{p}} = \frac{(1-.118)^2}{4} \times .004365 = .00032946$

$\therefore$ S.E. $(\hat{p}) = .0288$.

Similarly in other similar cases of 9:7 ratio versus 9:7 ratio, such as <u>tip margin versus juncture back</u>, <u>tip margin versus stigma</u>, <u>ligule versus juncture back</u>, <u>ligule versus stigma</u> and <u>J_3 back versus stigma</u>, the estimates of the cross over values and their standard errors can be worked out, by the two methods given above.

8.2. Ratio 9:7 versus ratio 27:37 - (B) type:-

Considering the case of Juncture back (9:7) versus septum (27:37) following are the observed frequencies and the expected ratio, taking any one of the two localisation genes in the hypothesis for the ratio 27:37 to be linked with the one localisation gene in the hypothesis for ratio 9:7. The expected ratio has been worked out, taking the basic factor C-c, which is common to both of the ratios, into consideration.

	Both pigmented	Only J. Back pigmented	Only septum pigmented	None pigmented	Total
Expected	$n_1 = \frac{9}{64}(2+\theta)$	$\frac{9}{64}(2-\theta) = n_2$	$\frac{9}{64}(1-\theta) = n_3$	$\frac{1}{64}(9\theta+19) = n_4$	1
Observed frequencies	$n_1 = 124$	$n_2 = 52$	$n_3 = 18$	$n_4 = 123$	$317 = n$

(i) Estimation of θ :- Following are the two methods used for estimation of θ.

(a) Maximum likelihood method:- The logrithim of the likelihood expression is given by:

$$L = n_1 \log \frac{9}{64}(2+\theta) + n_2 \log \frac{9}{64}(2-\theta) + n_3 \log \frac{9}{64}(1-\theta) + n_4 \log \frac{1}{64}(9\theta+19).$$

Maximising it, w.r. to θ, we have

$$\frac{dL}{d\theta} = \frac{n_1}{2+\theta} - \frac{n_2}{2-\theta} - \frac{n_3}{1-\theta} + \frac{9n_4}{9\theta+19} = 0 \text{ ----- (3).}$$

which reduces to cubic...

$$9n\theta^3 - (8n_1 - 28n_2 - 19n_3 + 9n_4)\theta^2 - (89n_1 - n_2 + 36n_3 + 36n_4)\theta + (38n_1 - 38n_2 - 76n_3 + 36n_4) = 0$$

Putting the numerical values of n_r's for the above case, it gives

$$2853\,\theta^3 - 301\,\theta^2 - 9860\,\theta + 5796 = 0 \text{ Solving this for } \theta,$$

we have $\hat{\theta} = .656$. $\therefore \hat{p} = 1 - \sqrt{\hat{\theta}} = 1 - .809 = .191$.

(b) Minimum discrepancy method:- as before, the expression

$$\sum_r \frac{(f_r(\theta))^2}{n_r + 1} = \frac{\left(\frac{9n}{64}(2+\theta)\right)^2}{n_1 + 1} + \frac{\left(\frac{9n}{64}(2-\theta)\right)^2}{n_2 + 1} + \frac{\left(\frac{9n}{64}(1-\theta)\right)^2}{n_3 + 1} + \frac{\left(\frac{n}{64}(9\theta+19)\right)^2}{n_4 + 1}$$

minimising this expression w.r. to θ, it gives

$\frac{9(2+\theta)}{n_1+1} - \frac{9(2-\theta)}{n_2+1} + \frac{9(1-\theta)}{n_3+1} + \frac{2\theta+19}{n_4+1} = 0$ which gives

$$\hat{\theta} = \frac{\frac{1}{n_3+1} + \frac{2}{n_2+1} - \frac{19}{9(n_4+1)} - \frac{2}{n_1+1}}{\sum_r \left[\frac{1}{n_r+1}\right]}$$, Putting the numerical

values of n_r's, it gives $\hat{\theta} = .654962$

$\therefore \quad \hat{p} = 1 - \sqrt{\theta} = 1 - .809 = .191$

(ii) Calculations for variance of (p) in the above two cases:-

(a) In maximum likihood method, it is known that $\frac{1}{V_\theta} = S\left(n_r \frac{d^2 \log n_r}{d\theta^2}\right)$.

Differentiating left hand side of equation (3), w.r.to θ, and substituting nn_r for n_r's we have

$$\frac{1}{V_{\hat{\theta}}} = \frac{9n}{64}\left\{\frac{4}{4-\theta^2} + \frac{1}{1-\theta} + \frac{9}{2\theta+19}\right\}$$

which gives $V_{\hat{\theta}} = \frac{64}{9n} \frac{(4-\theta^2)(1-\theta)(2\theta+19)}{(152-40\theta-61\theta^2)}$

$\therefore \quad V_{\hat{p}} = \frac{V_\theta}{4\theta} = \frac{16}{9n} \frac{(4-\theta^2)(1-\theta)(2\theta+19)}{\theta(152-40\theta-61\theta^2)}$

substituting the value of θ, we have

$V_{\hat{p}} = .00201694$, which gives S.E. $(\hat{p}) = .0449$.

(b) Using minimum discrepancy method, we found above that

$$\hat{\theta} = \frac{\frac{1}{n_3+1} + \frac{2}{n_2+1} - \frac{19}{9(n_4+1)} - \frac{2}{n_1+1}}{\sum_r \left[\frac{1}{n_r+1}\right]}$$, as before

$$F = \log \theta = \log \left\{ \frac{1}{n_3+1} + \frac{2}{n_2+1} - \frac{19}{9(n_4+1)} - \frac{2}{n_1+1} \right\} - \log \left(\sum_r \frac{1}{n_r+1} \right)$$

Using $\frac{1}{n} V_F = S\left\{ n_r E \left(\frac{\partial F}{\partial n_r}\right)^2 \right\}$, we have

$$V_{\hat{p}} = \frac{9n}{64}(2+\theta)\left\{\frac{\left\{\frac{9n}{64}(2+\theta)+1\right\}^{-2}}{\frac{1}{1+\frac{9n}{64}(1-\theta)}+\frac{2}{1+\frac{9n}{64}(2-\theta)}-\frac{19}{9\left\{1+\frac{n}{64}(98+19)\right\}}-\frac{2}{1+\frac{9n}{64}(2+\theta)}}\right.$$

$$\left.+\frac{\left\{\frac{9n}{64}(2+\theta)+1\right\}^{-2}}{\frac{1}{\left\{1+\frac{9n}{64}(2+\theta)\right\}}+\frac{1}{1+\frac{9n}{64}(2-\theta)}+\frac{1}{1+\frac{9n}{64}(1-\theta)}+\frac{1}{1+\frac{9n}{64}(98+19)}}\right\}$$

$$+\frac{9n}{64}(2+\theta)\left\{\frac{\left\{1+\frac{9n}{64}(2-\theta)\right\}^{-2}}{(\text{same expression as in in first bracket})}+\frac{\left\{1+\frac{9n}{64}(2+\theta)\right\}^{-2}}{(\text{same expression as in first bracket})}\right\}$$

$$+\frac{9n}{64}(1-\theta)\left\{\frac{\left\{1+\frac{9n}{64}(1-\theta)\right\}^{-2}}{(\text{same expression})}+\frac{\left\{1+\frac{9n}{64}(1-\theta)\right\}^{-2}}{(\text{same expression})}\right\}$$

$$+\frac{n}{64}(98+19)\left\{\frac{9\left\{1+\frac{n}{64}(98+19)\right\}^{-2}}{(\text{same expression})}+\frac{\left\{1+\frac{n}{64}(98+19)\right\}^{-2}}{(\text{same expression})}\right\} \quad \ldots\ldots\ldots (4)$$

Substituting the numerical values of n and θ in (4), it gives on simplification, $V_{\hat{p}} = .018103$. Now making use of the relation $V_{\hat{p}} = \frac{\theta}{4}V_{\hat{p}}$, we have $V_{\hat{p}} = .001981$ $\therefore$ $S.E.(p) = .0445$.

Similarly in the remaining cases of 9:7 ratio versus 27:37 ratio, such as tip versus septum, ligule vs septum and stigma versus septum, the estimates of the cross over values and their standard errors can be worked out, for the above two methods.

3.3. Ratio 163:94 versus ratio 9:7 -(C) type:- For illustration the case of leaf sheath (163:94) versus ligule (9:7) has been completely dealt with here. Let any one of the three independent localisation genes, governing pigmentation or leaf sheath be linked with the single localisation gene for ligule.

Taking the common gene C-c, into consideration, following is the expected ratio, which has been calculated.

	Both pigmented	Only leaf sheath pigmented	Only ligule pigmented	None pigmented	Total
Expected Ratio	$\frac{9}{256}(13+2\theta)$	$\frac{9}{256}(5-2\theta)$	$\frac{9}{256}(3-2\theta)$	$\frac{1}{4}+\frac{3}{256}(1+6\theta)$	1
Observed Freq.	$169=n_1$	$38=n_2$	$3=n_3$	$107=n_4$	317=n

Using the maximum liklihood method, we have the logrithim of the liklihood function as

$$n_1 \log\frac{9}{256}(13+2\theta) + n_2 \log\frac{9}{256}(5-2\theta)+n3 \log\frac{9}{256}(3-2\theta)+ n_4 \log \left\{\frac{1}{4} + \frac{3}{256}(1+6\theta)\right\}$$

Maximising it, w.r. to θ we have

$$\frac{n_1}{13+2\theta} - \frac{n_2}{5-2\theta} - \frac{n_3}{3-2\theta} + \frac{6n_4}{67+12\theta} = 0$$ which simplifies to a cubic

$$72n\theta^3-\theta^2(20n_1+628n_2-556n_3-160n_4)- \theta(802n_1+638n_2+92n_3+1602n_4)$$
$$+(1005n_1-2613n_2+4355n_3+1755n_4) = 0$$

Substituting the numerical values of n_i's the observed frequencies, it reduces to $22824\,\theta^3 + 41412\,\theta^2 - 283002\,\theta + 265271 = 0$

which does not give any solution for θ, lying between 0 and unity and hence the estimate of p is impossible to get. Still the value of $\theta = 1.26$ satisfies this equation. This value of θ, if taken to be not different from unity, gives $p = 0$, which reveals the presence of complete linkage or pleiotropic effect.

Consideration of pleiotropic effect:

Let one of three localisation genes, governing pigmentation in leaf sheath, be the same localisation gene which is responsible for pigmentation, in ligule, in the presence of the basic gene C-c. Thus on

putting $\theta = 1.0$ in the above expected ratio, the new ratio is given as 135:27:9:85.

	Both pigmented	Only leaf sheath pigmented	Only ligule pigmented	None pigmented	Total
Expected Ratio	135/256	27/256	9/256	85/256	1.0
Expected frequencies	167.8	33.4	11.1	105.3	317
Observed frequencies	169	33	8	107	317

Carrying out the χ^2 test, for testing whether this ratio fits the observed frequencies, we have $\chi_3^2 = 6.5$ which has the probability between .1 and .2. The non-significance of χ_3^2 shows that the above ratio, under the pleiotropic effect, fits the observed frequencies well.

The remaining three cases, leaf sheath versus juncture hook, leaf sheath versus stigma and leaf sheath versus tip margin, behave exactly as the above case has done, except in the last case where the probability of χ_3^2 was slightly less than .05 showing its significance. But since in this case also, the above equation does not give any solution for θ, lying between 0 and unity, this case may also be taken as showing the pleiotropic effect.

3.4. Ratio 163:94 versus ratio 27:37. (D) type:-

Taking the case of leaf sheath (163:94) versus septum (27:37) and assuming that one of the three localisation genes of the former case, is linked with one of the two localisation genes of the latter case, we have the expected ratio as:

	Both pigmented	Only leaf sheath pigmented	Only septum pigmented	None pigmented	Total
Expected Ratio	$\frac{27}{1024}(13+29)$	$\frac{27}{1024}(11+29)$	$\frac{27}{1024}(3+29)$	$\frac{1}{1024}(295+549)$	1.0
Observed frequencies	$139 = n_1$	$68 = n_2$	$8 = n_3$	$107 = n_4$	$317=n$

As usual, using the Maximum liklihood method, we have the equation for θ, as

$216n\theta^3-(332n_1+2260n_2+1896n_3+108n_4)\theta^2+(6478n_1+3794n_2+6542n_3+8046n_4)\theta-41222=0$

substituting the values of the observed frequencies n_r, it gives

$674728\theta^3 + 993648\theta^2 - 15229988 + 41222 = 0$, which like the type (C), does not give any solution of θ, lying between Zero and unity.

Investigating the presence of pliotropic effect, as before, we have the new ratio, on putting $\theta = 1$, as 403: 243: 87: 349. The usual X_3^2 in this case comes out to be 3.59, which has the probability between .1 and .2 and hence confirms the presence of pliotropic effect.

4.1. Following are the summarised results in tabulated form for the estimates of cross over values in different cases, with their standard errors, as found by using the two methods, mentioned earlier.

S. No.	Type	Combinations	Maximum liklihood method		Minimum discrepancy method	
			Estimate (p)	S.E. (p)	Estimate (p)	S.E. (p)
1.		Tip Margin(9:7) Vs. Ligule(9:7)	.111	.0226	.118	.0283
2.		" " Vs. J.Book(9:7)	.125	.0243	.131	.0250
3.		" " Vs. Stigma(9:7)	.180	.0299	.186	.0364
4.	(A)	Ligule(9:7) Vs. J.Book(9:7)	.071	.0178	.083	.0194
5.		" Vs.Stigma (9:7)	.204	.0323	.219	.0340
6.		J.Book(9:7) Vs. "	.113	.0228	.123	.0240
7.		Tip margin(9:7) Vs.Septum(27:37)	.270	.0533	.259	.0529
8.	(B)	Ligule(9:7) Vs. "	.212	.0368	.186	.0438
9.		J.Book(9:7) Vs. "	.191	.0449	.191	.0445
10.		Stigma (9:7) Vs. "	.187	.0353	.191	.0445
11.		Tip margin(9:7)Vs. Leaf sheath(162:94)				
12.		Ligule (9:7) Vs. " "				
13.	(C)	J.Book (9:7) Vs. " "	Pliotropic effect is present.			
14.		Stigma (9:7) Vs. " "				
15.	(d)	Septum(27:37) Vs. " %				

4.3. Efficiency: From the standard errors of the two types of estimates of the recombination fraction p, as given in the above results, it is obvious that the maximum likelihood method is on the average more efficient than the minimum discrepancy method, as originally given by Prof. Haldane. In the present case the efficiency of minimum discrepancy method ranges from 61.0 to 103.0 per cent and on the average comes out to be about 83 per cent as compared with the maximum likelihood method.

REFERENCES

Fisher, R.A. (1938) Statistical methods for research workers, 2nd Ed. Edinburgh, Oliver and Boyd.

Haldane, J.B.S. (1953a) A class of efficient estimates of a parameter, Proc. Int. Stat. Conf. India, 1951.

_____ (1953b) The estimation of two parameters from a sample, Sankhya, 12.

Mather, K. (1951) The measurement of linkage in heridity, 2nd Ed. London, Methuen & Co. Limited.

Murty, V.N. (1954) Estimation of linkage by the method of minimum discrepancy, Genetics, Vol. 39, No. 5.

Table.1: Joint segregation of pigment in different parts of rice plant and individual segregation, got from joint segregation, for different parts seperately.
(Data received from C.R.R.I. Cuttack)

S. No.	Pair of parts	Observed frequencies PD	PG	GP	GG	Total	First part	Individual segregation got from joint segregation freq. Pigmented	Green	2nd part	Individual segregation got from joint segregation freq. Pigmented	Green
1.	Leaf sheath & T margin	167	40	2	108	317	L.sheath	207	110	T.margin	109	148
2.	" & ligule	169	38	3	107	317	"	207	110	Ligule	172	145
3.	" &Juncture Back	173	34	3	107	317	"	207	110	J. Back	176	141
4.	" & leaf axil	204	3	31	79	317	"	207	110	Leaf axil	235	82
5.	"&Internode	201	6	12	98	317	"	207	110	Internode	213	104
6.	" & septum	139	68	3	107	317	"	207	110	septum	142	175
7.	" & stigma	177	30	3	107	317	"	207	110	stigma	180	137
8.	" & apiculus	205	2	39	71	317	"	207	110	apiculus	244	73
9.	Tip margin & ligule	158	13	14	132	317	Tip margin	171	146	ligule	172	145
10.	" & J. Back	158	11	18	130	317	"	169	148	J. Back	176	141
11.	" & leaf axil	168	1	67	81	317	"	169	148	leaf axil	235	82
12.	" & internode	168	1	35	113	317	"	169	148	internode	203	114
13.	" & septum	116	53	26	122	317	"	169	148	septum	142	175
14.	" & apiculus	168	1	76	72	317	"	169	148	apiculus	244	73
15.	" & stigma	134	17	26	120	317	"	171	146	stigma	180	137
16.	Ligule & J. Back	165	7	11	134	317	Ligule	172	145	J.Back	176	141
17.	" & leaf axil	167	5	65	80	317	"	172	145	leaf axil	232	85
18.	" & internode	122	50	21	124	317	"	172	145	internode	143	174
19.	" & septum	122	50	21	124	317	"	172	145	septum	143	174
20.	" & apiculus	170	2	74	71	317	"	172	145	apiculus	244	73
21.	" & stigma	154	19	27	117	317	"	173	144	stigma	181	136
22.	Juncture back &leaf axil	174	6	57	80	317	J.back	180	137	leaf axil	231	86
23.	" & internode	169	7	43	104	323*	"					
24.	" & septum	124	53	18	122	317	"	177	140	septum	142	175
25.	" & apiculus	170	2	74	71	317	"	172	145	apiculus	244	73
26.	" & stigma	165	17	18	117	317	"	182	135	stigma	283	134
27.	Leaf axil & internode	210	25	3	79	317	Leaf axil	235	82	internode	213	104
28.	" & septum	139	97	3	78	317	"	236	81	septum	142	175
29.	" & apiculus	231	4	18	69	317	"	235	82	apiculus	244	73
30.	" & stigma	176	53	4	84	317	"	229	88	stigma	180	137
31.	Internode & septum	140	74	2	101	317	Internode	214	103	septum	142	175
32.	" & apiculus	198	1	46	72	317	"	199	118	apiculus	244	73
33.	" & stigma	169	20	47	81	317	"	189	128	stigma	216	101
34.	Septum & apiculus	140	2	104	71	317	Septum	142	175	apiculus	244	73
35.	" & stigma	125	27	54	111	317	"	152	165	stigma	179	138
36.	Apiculus & stigma	179	65	1	72	317	Apiculus	244	73	stigma	180	137

Table 2: F_2 segregation for pigmentation in different parts of rice plant in the cross Zuiho x Dular, as given in data received.

S.No.	Plant part	Pigmented	Non-pigmented	Total
1.	Leaf sheath	207	110	317
2.	Tip Margin	169	148	317
3.	Ligule	172	145	317
4.	Juncture back	176	141	317
5.	Leaf axil	235	82	317
6.	Internode	213	104	317
7.	Septum	142	175	317
8.	Apiculus	244	73	317
9.	Stigma	179	138	317

Table 3: Test for the hypothetical ratios to fit observed frequencies for F_2 segregation for each part separately, excluding internode.

S. No.	Plant part	Expected ratio	Frequencies Observed pigmented	Frequencies Non-pigmented	Expected pigmented	Frequencies Non-pigmented	Total	X_1^2	Probability value
1.	Leaf sheath	162P:94G	207	110	200.6	116.4	317	0.86	.3-.5
2.	Tip margin	9P:7G	169	148	178.3	138.7	317	1.11	.2-.3
3.	Ligule	9P:7G	172	145	178.3	138.7	317	0.50	.3-.5
4.	Juncture back	9P:7G	176	141	178.3	138.7	317	0.063	0.7-.8
5.	Leaf axil	3P:1G	235	82	237.75	79.25	317	0.127	0.7-.8
6.	Septum	27:37G	142	175	133.7	183.3	317	0.88	0.3-.5
7.	Apiculus	3P:1G	244	73	237.75	79.75	317	0.50	0.3-0.5
8.	Stigma	9:7G	180	137	178.3	138.7	317	0.04	0.9

Table. 4: Adjusted data for the joint segregation of pigment in different parts with test of linkage.
(excluding internode)

S. No.	Plant part	Expected ratio	Adjusted observed frequencies				Expected frequencies				Total	X_n^2	Probability Significant or not
			PP	PG	GP	GG	PP	PG	GP	GG			
1.	L. sheath & T. margin	486:162:90:286	167	40	3	108	150.5	50.2	27.8	88.5	317	X_3^2=32.11	Significant
2.	" & ligule	486:162:90:286	169	38	3	107	150.5	50.2	27.8	88.5	317	X_3^2=31.23	"
3.	" & J.back	486:162:90:286	173	34	3	107	150.5	50.2	27.8	88.5	317		"
4.	" & leaf axil	162: 0 :30: 64	204	3	31	79	200.6	0	37.1	79.3	317	X_2^2= 1.25	.5-.7
5.	" & septum	1458:1134:270:1284	139	68	3	107	112.8	87.8	20.9	95.5	317	X_3^2=27.2	Significant
6.	" & apiculus	162: 0 :80: 64	205	2	39	71	200.6	0	37.1	79.3	317	X_3^2= 1.165	.5-.7
7.	" & stigma	486:162:90:286	177	30	3	107	150.5	50.2	27.8	88.5	317		Significant
8.	Tip margin & ligule	27: 9 : 9: 19	158	11	14	134	133.7	44.6	44.6	94.1	317	X_3^2=67.6	Significant
9.	" & J.back	27: 9 : 9: 19	159	11	18	130	133.7	44.6	44.6	94.1	317		Significant
10.	" & leaf axil	9: 0 : 3: 4	168	1	67	81	178.3	0	59.4	79.3	317	X_2^2= 1.49	.3-.5
11.	" & septum	81: 63:27: 85	116	53	26	122	100.3	78.3	88.4	105.3	317	X_3^2= 9.63	Significant
12.	" & apiculus	9: 0 : 3: 4	168	1	76	72	178.3	0	59.4	79.3	317	X_2^2= 5.79	.05-.1
13.	" & stigma	27: 9: 9: 19	134	15	26	122	133.7	44.6	44.6	94.1	317	X_3^2=38.75	Significant
14.	Ligule & J.back	27: 9: 9: 19	165	7	11	134	133.7	44.6	44.6	94.1	317		"
15.	" & leaf axil	9: 0: 3: 4	170	2	65	80	178.3	0	59.4	79.3	317	X_2^2=.75	.5-.7
16.	" & septum	81: 63:27: 85	122	50	20	125	100.3	78.3	33.4	105.3	317		Significant
17.	" & apiculus	9: 0: 3: 4	170	2	74	71	178.3	0	59.4	79.3	317	X_2^2= 4.66	.05-.1
18.	" & stigma	27: 9: 9: 19	153	19	27	118	133.7	44.6	44.6	94.1	317		Significant
19.	J.back & L.axil	9: 0: 3: 4	174	2	61	80	178.3	0	59.4	79.3	317	X_2^2= 0.7	.95-.98
20.	" & septum	81: 63:27: 85	124	52	18	123	110.3	78.0	33.4	105.3	317		Significant
21.	" & apiculus	9: 0: 3: 4	174	2	70	71	178.3	0	59.4	79.3	317	X_2^2= 3.78	.2-.3
22.	" & stigma	27: 9: 9: 19	163	11	15	126	133.7	44.6	44.6	94.1	317		Significant
23.	Leaf axil & septum	27: 21: 0: 16	139	96	3	79	183.7	104.0	0	79.3	317	X_2^2= 4.57	.1-.2
24.	" & apiculus	3: 0: 0: 1	231	4	13	69	237.8	0	0	79.2	317	X_1^2= 1.75	.1-.2
25.	" & stigma	9: 3: 0: 4	180	55	0	82	178.3	59.4	0	79.3	317	X_2^2= .43	.8-.9
26.	Septum & apiculus	27: 0:21: 16	140	2	104	71	183.7	0	104.0	79.3	317	X_2^2= 1.39	.3-.5
27.	" & stigma	81: 27:63: 85	125	17	54	121	100.3	33.4	78.0	105.3	317		Significant
28.	Apiculus & stigma	9: 3: 0: 4	179	65	1	72	178.3	59.4	0	79.3	317	X_2^2= 1.21	.5-.7