WAITE INSTITUTE
9.10.78
LIBRARY

USE OF BIOMETRICAL TECHNIQUES IN QUANTITATIVE GENETICS

Submitted for the degree of Master of Agricultural Science
in the University of Adelaide

by

Trevor William Hancock, B. Ag. Sci.

Biometry Section,
Waite Agricultural Research Institute,
GLEN OSMOND,
South Australia 5064.

March, 1977

Awarded April 1978

CONTENTS

APPENDICES

SUMMARY

While the existing theory of quantitative genetics has much to offer the plant or animal breeder it is recognised that many of the statistics used are difficult to understand and interpret. In this thesis several biometrical techniques are investigated which can be used by the breeder. The suitability of these techniques is illustrated using the data from the Roseworthy selection trial. This trial was designed to compare the efficiency of selection for clean fleece weight by direct measurement to the conventional method of visual appraisal.

When analysing such breeding data two aspects are of interest, firstly, what changes have occurred between the populations as a result of the different selection criteria, and secondly, what changes can be expected if a particular method of selection is subsequently employed? In Chapters I, II, III and IV the former aspect is considered while the latter is investigated in Chapters II, V and VI.

The conventional methods of analysing breeding experiments are detailed in Chapters I and II. Thus the two flocks are initially compared using both parametric and non-parametric univariate tests. Genetic parameters (i.e. heritability, genetic, phenotypic and environmental correlations) are subsequently calculated for the various subsets of the data.

In view of the problems encountered in applying these correlations, it is proposed in Chapter III that Hotelling's T^2 (a multivariate technique) provides a simpler, but comprehensive, comparison of the two populations. On applying it to the Roseworthy data it can be seen that the two populations have diverged. In patricular, it can be seen from the simultaneous confidence intervals that staple length, clean scoured yield and secondary follicle number are positively associated with the increase in clean fleece weight while crimps per inch and body weight are negatively associated.

In Chapter IV the analysis of discrete variates such as reproductive performance is considered. An "improved" FORTRAN algorithm for the Exact Test of R x C contingency tables is developed. As this test does not require approximation to a known distribution it can be applied to contingency data irrespective of the minimum size of the expected cell frequencies.

Several methods for recognising the important factors among a multivariate data set are considered in Chapter V. Principal Component Analysis is observed to provide the most effective method, especially from the point of consistency. Using its orthogonal components the breeder can predict the changes which would occur following a particular method of selection. Unlike the genetic parameters of Chapter II which are often used for this purpose, the above method remains manageable as the number of variates increases.

In Chapter VI the effects of a major gene on the frequency distribution of a metric character are investigated. As the segregation of a major gene can be expected to result in non-normality of the frequency distribution, it is proposed that a normality test can be used to screen data. When non-normality is identified, the method of moments (Hawkins, 1975) can be used to estimate the means and variances for the general population and for those carrying the major gene respectively.

DECLARATION

I declare that this thesis contains no material which has been accepted for the award of any other degree or diploma in any university and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference has been made in the text of the thesis.

T. W. Hancock

ACKNOWLEDGEMENTS

I wish to express my gratitude to my supervisor Dr Oliver Mayo for his guidance and encouragement during my studies and also to Dr Tony Rathjen for his assistance whilst Dr Mayo has been on sabbatical.

I would like to thank the Director of Roseworthy Agricultural College for making the data available and acknowledge the splendid work of all those who maintained and recorded the Roseworthy flocks (in particular Mr R.E. Brady).

I wish also to thank Mrs Helen Simpson for her excellent typing of this thesis.

WAITE INSTITUTE
LIBRARY

GENERAL INTRODUCTION

Quantitative genetics has been developed to describe the inheritance of continuous characters by plants, animals and man. While Mendelian inheritance provides an adequate model for handling discontinuous characters, segregating at one or a few loci, it cannot effectively represent the more complex inheritance of continuous characters (e.g. for yield in plants or body weight and growth rate in animals). Therefore it is assumed that many genes of small effect combine additively, and in the presence of environmental variation, give rise to the continuous distributions characteristically analysed in quantitative genetics.

Mather (1941) has proposed the term polygenic variation for these results arising from the simultaneous segregation of many genes, the minor genes concerned being collectively called polygenes. However, this categorization of genes has been further developed so that today's geneticist is confronted with Mendelian genes, major genes, super genes, polygenes and more recently neutral genes or isoalleles (Kimura and Ohta, 1973; Wallace, 1975). These five types of gene describe the level of expression of the gene and probably not any fundamental biochemical differences. Thus, the divisions are not distinct and in fact genes may exist which can be considered as Mendelian for one trait and polygenic for a second trait (Thompson and Thoday, 1974). Alternatively, particular genes may change categories as our understanding of them increases.

This model for quantitative genetics then allows us to represent the behaviour of characters under selection (i.e. progeny phenotypes will reflect their respective parents' genotypes to some extent). In particular, the magnitude of the expected response (R), following selection can be described in terms of the selection differential (S) of the parents and the heritability (h^2)

i.e. $R = h^2S$

where h^2 is the ratio of the additive genetic variance (σ^2_A) to the total phenotypic variance (σ^2_P). Clearly the greater h^2 or S, the larger the expected response.

However, this relationship does not directly indicate that the rate of response will be affected by the number of genes concerned (i.e. fewer genes of larger effect would be expected to respond more rapidly to selection than many genes of small effect, c.f. Waddington and Lewontin, 1967). If the polygenes can be located then the response would be even more direct as it could be selected for in a Mendelian manner.*

Genes controlling quantitative characters can be located using linked "marker" genes (Lee and Pateman, 1959) or mapped directly using the approach of Breese and Mather (1957) with *D. melanogaster*. Thoday (1961) has expanded this latter technique to locate polygenes, however, the method is not yet suitable to domestic animals where much less is known about the genome and the chromosome number is much larger.

Selection, whether natural or artificial, can be broadly classed as one of the following three cases (Mather, 1953):

i) Directional Selection - individuals are selected from one (or both) extreme(s) of the range of phenotype with the expectation that the phenotype of the progeny will reflect this bias.

ii) Disruptive Selection - selection against intermediates.

iii) Stabilizing Selection - intermediate phenotypes are chosen at the expense of both extremes.

Lee and Parsons (1968) suggest that the type of natural selection acting on a particular character will influence its response to artificial selection. Using the argument of Mather (1943), Lee and Parsons discuss the arrangement of polygenes on the chromosome. Assuming stabilizing selection is operating it is shown that the repulsion heterozygote gives

---

* Maximum possible response is also a simple function of gene number.

a higher fitness when compared to the coupling heterozygote. Fisher (1930) previously suggested that when two genes affect the same character for which intermediate values are fittest, then closer linkage between the genes will be favoured by the selection. Sheppard (1953) supported Fisher's view, however, Turner (1967 a & b) has used mathematical models to demonstrate that most genes do not exist in a single tight-linked group.

Since many characters in nature exhibit optimal fitness at intermediate phenotypes (Weldon, 1901; di Cesnola, 1907; Rendel, 1943; Karn and Penrose, 1951; Fraccaro, 1956; and Jayant, 1966) it is likely that stabilizing selection is widespread (e.g. Karn and Penrose (1951) report lower fitness for extreme birth weights in humans; but cf. also Robertson (1956)). It would therefore seem likely that for such characters there are polygenes with both negative and positive effects so that the response to directional selection could be very rapid. In contrast, for characters closely related to fertility and viability it would be expected that continuous directional selection would occur. Thus, most of the related polygenes would be expected to be positively biased, therefore, response to further directional selection would be slow. Indeed, both predicted and actual response of characters belonging to the two distinct categories support the generalisations (i.e. fitness characters have a low heritability while other characters have a relatively high one (Robertson, 1955)).

While disruptive selection is seldom seen in nature (for an example, see Clarke and Sheppard, 1962), bidirectional selection is often used in laboratory animals. Meyer and Enfield (1975), selecting for 21 day pupa weight in *Tribolium castaneum*, observed a marked asymmetry of response to such selection. Similar outcomes have been reported earlier (see Falconer, 1960) with the rate of downwards response exceeding the rate of upward change. Meyer and Enfield concluded "that progress relative to selection intensity is greater for downward than for upward selection, and that the more intense the selection in the downward direction, the larger the realized heritability." Falconer (1960) lists the following possible explanations for the observed asymmetry:

- i) The selection differential may depend on the direction of selection, e.g.
    - (a) natural selection may hinder artificial selection in an upward direction but assist it in a downward direction.
    - (b) The fertility may change.
    - (c) Variance may increase (decrease) as the mean increases (decreases).
- ii) Genetic asymmetry may be present. The dominant alleles may exert their effects predominantly in one direction which would lead to greater response in the direction of the recessive alleles (directional dominance), or the distribution of gene frequencies may be asymmetrical, such that the more frequent alleles affect the character in the same direction. In this situation we would expect a more rapid response in the direction of the less frequent alleles. The first explanation for genetic asymmetry also relates to the possible effects of heterozygotes (i.e. response would be slow in one direction as the gene frequencies approached equilibrium but rapid in the other direction).
- iii) Inbreeding depression may reduce the rate of response in the upward direction but increase it in the downward direction.
- iv) Maternal effects may exhibit directional bias.

These explanations for asymmetry can be seen to relate to the type of natural selection previously affecting the particular character. Thus, greater asymmetry would be more likely for characters associated with fitness, for which directional selection had been occurring, than for characters under stabilizing selection.

One further characteristic of response warrants consideration, namely "accelerated response" (see Mather and Wigan, 1942; Clayton and Robertson, 1957; Thoday and Boam, 1961). It is difficult to account for this behaviour, however, recombination between closely linked loci appears to be the most likely explanation. The result can also be explained by mutation or by the progressive accumulation of interacting genes. Thus, "accelerated response" may be an illustration of non-additivity. Falconer (1960) has described the various components of non-additivity and their possible inter-relationships. However, although they are often important in determining the magnitude of the response to selection I will refrain from giving further details here since the components are generally not distinguishable in selection experiments.

While it is difficult to investigate more than two or three segregating Mendelian loci, the advent of the computer has enabled the effects of artificial selection on simple genetical models to be tested (Gill, 1965; Young, 1967; Qureshi and Kempthorne, 1968; Hedrich, 1970). Use of these models has not revolutionised our understanding of quantitative genetics, mainly because linkage and epistasis still prove awkward to handle. However, computer models have contributed in part to our awareness of the limitations of the additive model and therefore the variance component approach based on it. In noting this lack of new results from computer modelling one should also remember that most of the algebra for simple genetic systems was resolved some time ago by geneticists including Fisher, Haldane and Wright.

While recent progress can be described as modest in comparison to the work of these three earlier geneticists, it has still provided valuable contributions to the theory of quantitative genetics. For example, Harris (1966) and Lewontin and Hubby (1966) have demonstrated that species are not essentially homogeneous, as previously considered, but in fact as many as one locus in five or six may be heterozygous.

It is now well-accepted that selection for a particular character will almost inevitably be accompanied by a fall in fertility and viability (e.g. Wigan and Mather (1942), Mather and Harrison (1949), Nordskoy and Wehrli (1963) to mention but a few cases). The selection for genes affecting the desired character appears to lead to changes in linked fertility genes which generates the observed reduction in fitness. Consideration of the type of natural selection previously applying to the two types of characters further clarifies this observation. That is, as stabilizing selection is normally operative for metric characters, any change in emphasis can be expected to generate response. However, for fitness characters, where strong directional selection has previously been operative, any artificial selection for a metric character will decrease the natural selection and this will lead to a decline in fitness. Clearly any decrease in reproductive fitness is extremely undesirable, particularly in domestic animals, but this is far from easy to overcome. This is especially so for the larger domestic animals which may only have one or two progeny per year. Thus, reproductive fitness can only take one of a few discrete outcomes which gives an extremely ineffective measure of any shift in reproductive fitness. While the reproductive performance of the population under selection can be derived for each year, seasonal variations are sufficiently large to confuse all but gross changes. Also, if selection is to be carried out on the individual dam's record then its effectiveness will be marginal (McGuirk,1976). Measurements of critical sex hormones may improve the accuracy of the score, especially of sires, but so far there have been only limited reports in this area (Land, 1974; Bindon and Turner, 1974). Bindon and Piper (1976) conclude that measurements of plasma LH and FSH appear to be of only limited use, however, they suggest that ovulation rate and number of oestrous cycles may be more promising indicators of reproduction rate in sheep. Performance of sires can be obtained from the records of their daughters but this requires large numbers of animals and an extra generation.

Related secondary characters are sometimes used to select for primary characters, like reproductive fitness, which are by their nature unsuitable for direct selection (see Atkins and McGuirk, 1976). Such selection is to be avoided as, first, the observed correlations may be caused by a third character which has been ignored, and secondly the selection may do little more than break down the correlation without causing any response in the primary character.

McGuirk (1976) has partitioned reproductive performance into its two components, fertility and fecundity. He suggests that selection for fertility will result in only limited improvement as the heritability of the character is low and an upper limit (100%) obviously exists. For fecundity there is no upper limit as such, although an optimum sib number must exist for any species. Also the heritability is slightly higher (Forrest and Bichard, 1974). McGuirk has made no attempt to consider selection for shorter re-breeding interval or for lamb survival both of which affect reproductive performance. Clearly if fecundity and re-breeding interval are inversely related, any selection will be greatly complicated. Also for animals which normally have small litter size (e.g. sheep and cattle), any selection for fecundity will increase fertility (i.e. the more ovum released the greater the chances of at least one fertilizing and implanting). Despite these limitations, response to selection for increased multiple births has been reported for the Australian medium Peppin Merinos (Turner, 1962 and 1966) and for New Zealand Romneys (Wallace, 1958 and 1964). However, McGuirk (1976) summarized the relative benefits of selection for reproductive performance in the following manner:

> "While a high level of reproductive performance may be necessary for efficient and profitable production, greater economic benefits may be obtained by selecting for other characters. By selecting for increased reproductive performance we reduce the selection pressure which can be applied to other characters,...."

While any reduction in selection pressure, for reproductive ability, should clearly be avoided it is equally important to avoid any bias against multiple births. If the performance of progeny from larger litters is temporarily depressed then this must be taken into consideration at selection, otherwise there will be unconscious selection against fecundity.

Hammond (1947) put forward the idea that a character will respond most to selection in the environment in which it is most fully expressed even if this environment does give less than maximal performance. Falconer (1952) and Falconer and Latyozewski (1952) investigated the effect of two levels of nutrition on selection for body weight in mice. These authors are reported by Daday, Binet, Grassia and Peak (1973) to have demonstrated, "...improvement attained by selection under 'good' conditions was not realized when the selected strain was transferred to 'bad' conditions instead of being better than the strain selected under bad conditions it was worse, and showed no evidence of any advance beyond the initial level." I consider that this interpretation by Daday, Binet, Grassia and Peak (1973) may have over-emphasised any difference between the investigations of Hammond and those of Falconer and Latyozewski. Daday, Binet, Grassia and Peak selected for plant height in *Medicago sativa* L. under three natural environments. In their summary they stated:

"It is concluded from these investigations that unfavourable environmental conditions may severely limit response to selection, while unusually favourable ones may do so to a moderate extent. Therefore judiciously applied indirect selection may yield more genetic progress than direct selection in *M. sativa* populations under certain extreme conditions." Here it should be noted that "indirect selection" refers to selection on the basis of observations in one environment for performance in another environment. In a reply to Daday, Binet, Grassia and Peak (1973), Rendel and Binet (1974) expand Hammond's idea to illustrate the two sets of conditions for which they believe it was originally meant to apply. These were:

i) A genotype may be so weak, for the particular character of interest, that it may completely fail to express itself in the desired environment. If this character can be measured in a second environment then the genotype can be selected in this environment until it achieves expression in the first. This was illustrated for natural selection of immunity to myxoma virus in rabbits. While selection is absent for viruses with one hundred percent mortality, exposure to less effective strains of the virus (mortality less than ninety percent) allows the selection of animals which can then survive the original strains.

ii) The character may be expressed in both environments but the degree of expression differs. In this situation the response does not depend solely on the respective heritabilities but on the product of the selection differential by the heritability.

Jinks and Connolly (1973) proposed and later confirmed (Jinks and Connolly, 1975), that "...selection for high mean performance in a good environment or for low mean performance in a poor environment leads to selections that are more sensitive to environmental variations than selections for high mean performance in a poor environment or for low mean performance in a good environment."

Gibson and Bradley (1974) compared artificial stabilizing selection to natural selection in both constant and fluctuating temperatures. For bristle number in *Drosophila melanogaster* they observed a decrease in both genetic and environmental variance over the thirty-nine generations of stabilizing selection. During this period the mean number of bristles remained constant. In contrast, the mean bristle numbers of both control lines increased between generations 0 to 19 but then decreased so that after 39 generations there was a significant decrease in this character. The phenotypic variance remained constant initially and then decreased over the second half of the experiment. Their results demonstrated that fluctuating environments did not affect the process of stabilizing selection. In conclusion, this investigation generally supports previous work by Thoday (1959), Prout (1962) and Scharloo (1964) in showing that phenotypic variance can be reduced under artificial stabilizing selection but unlike Prout and Thoday, Gibson and Bradley found this reduction not only in additive genetic variance but also in environmental variance.

What conclusion if any can be drawn from these related studies? Clearly the answer is not simple. The practical breeder must precisely define his aim and take care that the character he chooses effectively supports this aim. The importance of this simple statement cannot be over-emphasised. The following, rather depressing, statement made by Robertson (1966) may illustrate the situation: "In general it would be fair to say that, until the last decade, genetics has had comparatively little effect in animal improvement." Robertson suggests that this failure has not been due to insufficient understanding of genetics, but resulted from selection for economically unimportant characters.

(Donald (1973) has expressed similar thoughts.) This can be illustrated by considering the selection history of maize. The development of hybrid varieties has enabled yields of up to thirty percent greater than previously obtained by many years of mass selection. However, Gardner (1961) has achieved yields which compare with the best hybrids. This response is believed to have occurred because Gardner selected for total yield per plant, given by the sum of all ears, whereas previous selection had been based on obtaining maximal yield from individual ears. Thus, Gardner's best plants were similar to hybrid plants as they produced two moderately large ears which together gave yields greater than one single large ear.

Interpretation of long term selection trials can be severely complicated by the presence of environmental variation. This environmental variation can occur both between and within generations. While the former makes it difficult to estimate the response to selection it should not affect the accuracy of the selection applied at each generation, unless there is a genotype by environment interaction.

Mention has already been made of the work of Gibson and Bradley (1974) who found that fluctuating environments did not affect stabilizing selection in *Drosophila melanogaster*. It would seem unlikely that their conclusions would hold for all species particularly when directional selection was applied. However, since the particular species will be required to exist in the presence of such environmental variation the method of selection should remain unchanged. It would then be hoped that the selected progeny exhibit general adaptation over the range of environments. Although verification of the last point is difficult, some attempt should be made. This leads back to the problem of how to estimate response when there is considerable environmental variation between generations. Maintenance of inbred lines, while suitable for plants and some laboratory animals, is obviously impracticable for domestic animals. Random bred control populations can prove helpful but care must be taken to avoid both unconscious selection and inbreeding depression. Also, maintenance of a random control population may be beyond the resources available to the breeder.

In comparison, any variation within generations will severely weaken the selection intensity. Such variation is mainly encountered in field crops, as a result of soil variation. Its effect can be reduced by the sowing of control plots which are used to adjust the yield of neighbouring plots. Although the procedure of adjustment is simple in one dimensional designs, it is greatly complicated for two-dimensional grid designs.

It should now be clear that the existing theory of quantitative genetics is inadequate in several areas when applied to selection experiments. Robertson (1966) listed the following three reasons:

i) it cannot predict the limits to selection;

ii) it cannot predict the changes in reproductive fitness, which might be expected to decline on selection in either direction for almost any trait;

iii) it gives little information about the real nature of the gene segregations underlying the observed variability.

To these, Lee and Parsons (1968) have added a further three reasons:

iv) it cannot predict the rate of response to selection, especially once an accelerated response to selection or a plateau has occurred;

v) it cannot predict the nature of correlated responses in general (including reproductive fitness mentioned above);

vi) it is usually based on an underlying genetic architecture of a large number of genes all acting additively, which frequently does not exist.

Together, these six problems provide a fair summary of the present state of quantitative genetics. It should however be appreciated that quantitative genetics, even with its weaknesses, still remains the only effective theory to describe the inheritance of continuous variation. In fact, its short term predictions (up to five generations) are found to be of worthwhile assistance to the breeder (Piper, 1971). The problems as listed above tend to belittle the important contributions made by investigators such as Robertson (1961 and 1970) on the theory of limits. (Robertson showed that the expected limit under selection was proportional to N_i, where N is the effective population size and i is the selection differential in standard deviations. In particular, the half life of the response would be less than 1.4N generations, given that the additive model was applicable. If, however, recessive genes were favoured, the half life could approach 2N. The investigation also considered the role of linkage and although the conclusions are not unexpected the justification of them, by Robertson, is of great significance.

In the following chapters I will consider particular biometrical techniques with reference to how they can be applied, first to explain the effect of artificial selection on continuous variation and secondly to assist the breeder in attaining his aims. The Roseworthy Fleece Selection Experiment will be simultaneously investigated to demonstrate the techniques. This experiment, carried out between 1954 and 1965, was designed to compare two methods of ram selection:

i) selection by visual appraisal only;

ii) selection by clean fleece weight following a preliminary visual appraisal.

These two methods were called the "Visual" and "Index" methods respectively. Basically they provide a comparison of subjective stud practices with objective fleece measurement with particular reference to clean fleece weight. Mayo, Potter, Brady and Hooper (1969) have previously used mainly univariate statistics to analyse the data. However, their analyses left certain aspects of the data untouched (e.g. genetical parameters and multivariate methods).

The present investigation has been divided into six sections. The first of these extends the previous analyses of Mayo, Potter, Brady and Hooper (1969). Separate analyses are provided for "single raised" and "twin raised" progeny whereas the previous authors pooled all data. The two flocks are compared using both parametric and non-parametric tests for two independent samples.

Genetic parameters are presented in the second section (i.e. heritability, genetic, environmental and phenotypic correlations.

Thirdly, Hotelling's T^2 is demonstrated to provide an effective method of comparing two populations when several variables have been measured.

An improved algorithm for Fisher's exact test for R x C contingency tables follows; it is applied to categorical fertility data. This test is especially valuable when the R x C table contains several cells with expected frequencies less than five.

Section five compares methods for choosing the "best subset" of linear variables. This approach provides a means of choosing from amongst a set of variables those which relate most closely to the breeder's interests.

Lastly, methods for the detection of major genes affecting quantitative traits are investigated using computer simulation.

I - DESIGN OF ROSEWORTHY EXPERIMENT AND SUMMARY STATISTICS

INTRODUCTION

In this first chapter the Roseworthy College clean fleece weight selection experiment is introduced. Univariate statistics are presented to enable a preliminary evaluation of the experiment. In later chapters the data will be further investigated using alternative techniques.

As stated earlier the experiment was designed to compare two methods of ram selection for increased wool production in Merino sheep. The methods were:

i) selection of rams on the basis of visual appraisal (the "Visual" method);

ii) preliminary culling of 70% of rams on visual appraisal before final assessment on clean fleece weight (the "Index" method).

The reasons for the choice of these two methods is given by Schinckel (1955). (All replacement ewes were chosen by visual appraisal.) Having observed that the annual increase in fleece weight for Australian sheep had decreased considerably between 1930 and 1950 Schinckel suggested the most likely explanation was the combined effects of (i) stabilization of livestock management and (ii) reduction of response to commercial selection. As the heritability of clean fleece weight was as high as 0.4 Schinckel further proposed that other methods may increase the rate of response. Initially, it was considered that selection on the basis of an index derived from economically important characters would be optimal. (Suggested characters were fleece weight, body size, staple length, quality, yield and absence of undesirable faults, e.g. malformed jaws, faulty hocks, excessive skin development.) However, Schinckel subsequently chose to use method (ii) described above which is similar to the "half classing" method described by Morley (1955). The name "Index" has been retained although the original index of economic characters has not been used. Thus, the two methods were designed to compare commercial visual appraisal with partial selection for clean fleece weight. While this comparison was provided by the design, its interpretation was greatly complicated by the fact of seasonal variation.

A random bred control flock would have done much to overcome this problem although its inclusion would have introduced problems of its own (i.e. inbreeding effects if present will be confounded with temporal effects). Also, this third flock was not included in the experiment as it would have greatly increased management problems and depleted the available resources. While the presence of seasonal variation will be seen later to complicate the drawing of conclusions on the long term trends of the two methods, the experiment still provides worthwhile comparison between the two methods as proposed by Schinckel.

Similar selection experiments on Merinos have been carried out by the New South Wales Department of Agriculture (Dun, 1958; Pattie ,1965; Robards and Pattie, 1967; Saville and Robards, 1972; Robards, Williams and Hunt, 1974; Pattie and Barlow, 1974; Barlow, 1974) and the CSIRO Division of Animal Genetics (Turner, 1958; Turner, Dolling and Kennedy, 1968; Brown and Turner, 1968; Turner, Brooker and Dolling, 1970; Turner, McKay and Guinane, 1972; Jackson, Nay and Turner, 1975). In both of these organisations more than one selection experiment has been carried out concurrently.

Two selection experiments have been investigated by CSIRO. Turner, Dolling and Kennedy (1968) describe the response observed between 1950 and 1964 in wool and body characteristics for the following three selection groups:

i) S - Selection for high clean wool weight rejecting for high fibre diameter and high wrinkle score. During 1950-59 the selection criteria included the performance of ram's half-sibs but from 1961-64, mass selection was practiced.

ii) MS - Mass selection for high clean wool weight rejecting for high fibre diameter and wrinkle score for the years 1950-59. During the later years of the experiment (1961-64) low crimp number replaced high fibre diameter as a criterion for rejection.

iii) C - Random bred control group.

During the period 1950-59 response in clean fleece weight was similar in the S and MS groups. This suggests that examination of half sibs has contributed little to the selection differential, as would be expected from the relatively large heritability for clean fleece weight. Fibre number per unit area was found to make the greatest contribution to the increase in fleece weight but greasy wool weight, percentage clean yield, body weight and staple length exhibited smaller increases. No further response was observed during 1961-64 which led Turner, Dolling and Kennedy to conclude, "It is impossible to say at this stage whether a selection plateau has been reached or whether drought conditions interfered with response."

In the second CSIRO selection experiment, sixteen small flocks have been maintained to allow selection for high and low values of eight single characters (clean wool weight, clean wool weight per unit skin area, body weight, wrinkle score, fibre number per unit skin area, fibre diameter, staple length and percentage clean yield). Divergence of the character under selection was observed for all eight individual characters although for long staple and low fibre number there was only initial response. Response was seldom symmetrical and correlated characters responded as expected. When interpreting these results it should be noted that the annual group sizes were small (i.e. one or two rams with 30-50 ewes).

The New South Wales Department of Agriculture has used two-way selection to investigate four characters (clean wool weight, crimps per inch, weaner body weight and wrinkle score). All four experiments commenced in 1951 but the crimps per inch, weaner body weight and wrinkle score programs were terminated in 1972. Flock sizes were larger than used for the similar CSIRO experiments - i.e. 100 ewes mated to 5 rams. A randomly selected control flock of 100 ewes and 10-25 rams was maintained for comparison. For clean fleece weight Pattie and Barlow (1974) report that the flock selected for increased fleece weight (the 'fleece plus' flock) increased initially in the first two generations by approximately one standard deviation but little further response was observed in the subsequent three generations. In comparison, the 'fleece minus' flock decreased almost linearly over the 5 generations with an overall realized heritability of 0.44 for ewes and 0.38 for rams. Barlow (1974) concludes that, "Most of the response in W [clean fleece weight] in the 'fleece plus' flock arose through increases in fibre density, fibre diameter and staple length. Staple length was the major component associated with the response in W in the 'fleece minus' flock." Robards and Pattie (1967) reported similar divergence between the 'crimps plus' and 'crimps minus' flocks. Greater response was observed in the 'crimps plus' than the 'crimps minus' flock. The 'crimps minus' flock was observed to exhibit greater clean fleece weight than the random bred control which is to be expected from the high negative correlation between these two characters. For weaner body weight Pattie (1965) reports that divergence occurred in ewes but was less obvious for rams.

From this summary of fleece selection experiments in the Australian Merino it can be seen that the Roseworthy experiment, although obviously related to the other experiments, provides the only direct comparison between "conventional stud" practices using visual appraisal and "improved stud" practices based on partial selection for clean fleece weight. It should also be noted that whereas the CSIRO and N.S.W. Department of Agriculture experiments apply to the medium-wool Peppin strain the Roseworthy experiment relates to the South Australian strong-wool "Bungaree" strain of Merino.

MATERIALS AND METHODS

Mayo, Potter, Brady and Hooper (1969) have more than adequately summarized the experimental procedures therefore the following will be limited to a brief summary of important points with specific references to differences in method of analysis, from these previous authors.

(a) Selection methods

During the years 1954 to 1965 inclusive two flocks of approximately 200 ewes were maintained at Roseworthy Agricultural College. All ewe replacements were selected visually from progeny born within the respective flocks. Each year seven rams, six "two-tooth" plus one "four-tooth" ram from the previous year's rams, were joined to each flock. All progeny were classed by visual appraisal at approximately 15 months of age. The four grades were: reserves, studs, flocks and culls. Six reserve rams were then chosen from the "visual" progeny and these, along with one ram from the previous mating, were then joined with the "Visual" flock. Similarly, six plus one rams were selected and joined to the "Index" flock but in this case 70% of the available hogget rams were culled on visual grounds before selecting six on clean fleece weight.

(b) Analysis

The following fifteen quantitative characters have been considered in the present work (except where otherwise stated the character was measured at hogget shearing, i.e. 15-16 months of age):

1. date of birth - expressed as number of days from the 1st of January of the particular year (recorded at birth);

2. birth weight of lamb in kgs (recorded at birth);
3. weaner body weight in kgs (recorded at approximately four months of age);
4. hogget body weight in kgs;
5. lamb fleece weight in kgs (recorded at approximately four months of age);
6. greasy fleece weight in kgs;
7. percentage yield;
8. clean scoured fleece weight in kgs;
9. mean staple length in cm;
10. crimps per inch (or per 2.54 cm);
11. mean fibre diameter in microns;
12. coefficient of variation of fibre diameter;
13. primary follicle number per square cm;
14. secondary follicle number per square cm;
15. skin thickness in cm (this variable was only recorded for the years 1958 to 1965 inclusive).

Characters 1 to 14 were recorded for all progeny born in the years 1954 to 1965 while skin thickness was only measured from 1958 onwards. All progeny records with one or more variables absent have been ignored in the analysis. Although this results in a slight loss of information it removes the added complication of estimating missing values in the multivariate methods of later chapters. All data have been partitioned on year of birth, flock type, sex and birth type. Simple descriptive statistics are presented for each group. In particular, the coefficients of skewness and kurtosis are presented to identify any deviations from normality. The coefficient of skewness is estimated by the statistic g_1, given by the division of the third central moment by the cube of the standard deviation, i.e.

$$g_1 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^3}{n\left[\frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n}\right]^{3/2}}$$

where $y_1, y_2, \ldots, y_n$ is a sample of n observations. The standard error of g_1 is given by

$$\sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3}} .$$

The coefficient of kurtosis is estimated by the statistic g_2, which is three less than the fourth central moment divided by the fourth power of the standard deviation,

i.e.

$$g_2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^4}{n\left[\frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n}\right]^{2}} - 3$$

The standard error of g_2 is

$$\sqrt{\frac{24n(n-1)^2}{(n-2)(n-3)(n+3)(n+5)}}$$

(Both these standard errors are approximations but their accuracy is considered more than sufficient for the present circumstances.)

Parametric significance tests are applied to compare between flocks. Firstly variances are compared using the simple variance ratio test. The corresponding means are then compared using student's t-test for independent samples. This second test has been applied irrespective of the outcome of the first. Obviously the t-test is invalid if the variances are heterogeneous, however, the outcome of the t-test has been presented to demonstrate the behaviour of this statistic under such circumstances. Although in other fields one may be solely interested in comparing means, in quantitative genetics it is also important to identify differences between variances.

In comparison, a non-parametric test, the Kolmogorov-Smirnov two-sample statistic, has been applied to establish whether the two samples differ in central tendency. This test would give more accurate comparison than the parametric t-test if one or both of the samples concerned were distributed non-normally or if the variances were heterogeneous. However, it should be appreciated that the use of similar tests on the same data will increase the probability of rejection. The present study has used this approach for comparative reasons only and it should not be interpreted as recommending in any way such repetitive use of statistics.

RESULTS

Table I-1 parts (1) to (15) list the basic statistics for progeny raised as singles where parts (1) to (15) correspond to the 15 variables listed previously. The data have been partitioned into the four classes given by sex and flock type for the 12 years from 1954 to 1965. The corresponding information has been included for dams born between 1949 and 1953 and used in the early years of the experiment. Table I-2 parts (1) to (15) give similar details for progeny born and raised as twins. Clearly the tables contain an immense amount of information of fundamental importance. The present summary can do little more than highlight a few of the more important aspects. (In later chapters better techniques will be used to extract further details from the data.)

First we note that twin progeny differ from singles for many of the variables. In particular, for clean fleece weight of twins we see that the mean performance can be as much as 0.5 kg less than for singles. Clearly any selection which ignores this aspect may lead to reduction in fecundity. These differences are largest for variables measured at birth (e.g. birth weight approximately 17% reduction) or early in the life of the lamb (e.g. weaner body weight and lamb fleece weight approximately

12% and 20% reduction respectively), while at 18 months many variables show either small or negligible differences (i.e. for greasy fleece weight, clean fleece weight and skin thickness singles have higher mean values than twins, while the opposite occurs for crimps per inch and primary follicle number). No attempt has been made to use statistical tests to compare between singles and twins as this is not of prime importance to the study. To do so would only expand further the large number of tables to be included in this thesis and establish what is obvious for most of the variables. (Tables I-1 and I-2 contain all the information necessary to calculate the appropriate parametric tests if desired.)

Tables I-3 parts (1) to (15) and I-4 parts (1) to (15) summarise the results of both parametric and non-parametric two sample tests, comparing data sets from the Index and Visual flocks. When interpreting such data one must remember that chance alone can result in the rejection of the null hypothesis when it is in fact true (i.e. error type I). Thus, if we choose a 5 percent significance level (α) then we can expect 5 percent of our tests to exhibit significance due to chance alone. This, combined with the large seasonal effects, greatly weakens one's ability to distinguish any divergence between the flocks over the twelve years. Table I-5 indicates the number of significant results (when $\alpha = 0.05$) for the fifteen variables. Thus the variance ratio test, the t-test and the Kolmogorov-Smirnov test, are significant on 12.5%, 24.4% and 18.1% of occasions respectively. More detailed comparison of the t-test and the Kolmogorov-Smirnov test demonstrates that they give consistent results with the latter test leading to rejection on fewer occasions. As the parametric t-test assumes the data are distributed normally with equal variances it is to be expected that it will lead to more frequent rejection when these assumptions hold (as indicated by the values of skewness and kurtosis in Tables I-1 and I-2). Siegel (1956) describes the relative ability of the non-parametric test to reject the null hypothesis, given that the data are distributed normally, as the "power-efficiency" of the particular test. He concludes that the Kolmogorov-Smirnov test, when compared to the t-test, has a "power-efficiency" of near 96% for small samples and this decreases slightly as sample size increases. From Table I-5 we observe, ignoring the variance ratio test, that the Kolmogorov-Smirnov test leads to rejection on 123 occasions whereas the t-test does so on 163 occasions. This would suggest a "power-efficiency" of 75%. However, if the t-tests are restricted to exclude cases where a significant variance ratio has been observed, the "power-efficiency" increases to 90%. It is not clear whether Siegel considered the variance ratio test when calculating his "power-efficiency". Also it should be remembered that he made his calculations on simulated data which had been designed to satisfy the assumptions of the t-test. Clearly the present data support the view that this non-parametric test provides quite a valuable alternative to the more widely used t-test. However, as its calculation is slightly more complicated, for all but small samples, its use may be conditional on the availability of a computer or programmable calculator. In conclusion, as the Kolmogorov-Smirnov test makes far fewer assumptions about the data and since it tests the equality of the distributions of the two samples, it provides quite a valuable univariate test. This is particularly so in the present application to quantitative genetics where one is interested in identifying any differences in the two distributions and not just the means or variances of normal distributions.

So far I have mainly considered the relative performance of the three tests without any reference to the behaviour of particular variates over the twelve years of the trial. As previously stated, Mayo, Potter, Brady and Hooper (1969) have considered this latter aspect. However, their approach differed in several aspects from mine:

- i) they did not separate progeny on birth type;
- ii) the following additional variates are considered here:
  - (a) date of birth,
  - (b) weaner body weight,
  - (c) lamb fleece weight,
  - (d) yield percentage,
  - (e) coefficient of variation of fibre diameter,
  - (f) skin thickness.

  (The inclusion of these variates allows a more complete assessment of the two flocks. Although weaner body weight, lamb fleece weight and coefficient of variation of fibre diameter may be considered slightly less important variates the remaining three - date of birth, yield percentage and skin thickness - are of importance to the sheep breeder.)
- iii) Primary follicle number and secondary follicle number have been considered individually and not as a combined total. Since the number of primary follicles is much less than the number of secondary follicles this separation contributes worthwhile information particularly on the former variate.
- iv) The following additional statistics have been calculated: coefficient of skewness, coefficient of kurtosis, variance ratio test and the Kolmogorov-Smirnov test.

From the above it can be seen that the present analysis follows closely the previous work. However, it has been represented here to provide a more detailed summary which can then be referred to in later chapters. In Figure I-1 parts (1) to (14) and Figure I-2 parts (1) to (14) the mean values from Tables I-1 and I-2 have been plotted. This enables identification of trends over time between the two flocks.

Firstly, from Table I-1(8) and Figure I-1(8), we see that for both sexes, clean fleece weight (for single born progeny) has diverged with the Index progeny significantly exceeding the Visual progeny during most years of the experiment. For this variable, and indeed for most of the variables, it can be seen that variation between seasons is large. (No attempt has been made to remove this variation as it is the author's belief that such techniques often lead to problems in their own right, i.e. genotype x environment interactions often prove exceptionally difficult to identify and understand.)

The 1954 male progeny means appear to be abnormally high when compared to the rest of the data. No explanation could be found for this discrepancy and it can only be concluded that some unknown aspect of the environment was highly favourable for the male progeny during the first year of this trial. If this was an intentional factor of management, it may have been because it was considered advantageous to select animals in a favourable environment. Whatever the reason this biased treatment of male progeny was discontinued during the period from 1955 to 1965. However, as the mean clean fleece weight of males was always less than the 1954 value this greatly detracts from the experiment. The mean values for dams born in years preceding 1954 further support the view that the 1954 males were abnormal. A similar pattern can be seen for the clean fleece weight of progeny born and raised as twins although there is obviously greater variation in the means due to the smaller numbers in the groups.

For date of birth we observe that the variance ratio is frequently significant (i.e. 17 occasions out of 46). Neither flock is solely responsible for this rejection although the variance of the Index flock tends to be larger for the later years of the trial. It would be expected that these larger variances resulted from poor ram performance at mating. Mean date of birth similarly fluctuates between the two flocks over the seasons. In particular, mean date of birth for the Index flock is much later during 1965 (approximately 10 days). Closer inspection of Tables I-1(1) and I-2(1) indicates that both flocks began lambing at about the same time but the lambing of the Index flock extended over a greater period of time. Mayo, Potter, Brady and Hooper (1969) have reported that lambing percentage of the Index flock was markedly lower in 1965. This suggests that problems occurred at joining in the Index flock. However, from the available information it is impossible to determine whether this is related to the selection method or is simply due to chance. Therefore, irrespective of the reason for this disparity in date of birth between the flocks in 1965, it should be remembered that its presence may result in anomalous behaviour in the other variates measured subsequently on these animals.

For birth weight and weaner body weight the two flocks behave similarly. Males outweigh females and singles outweigh twins but within sex by birth type groups no consistent trends can be identified between the flocks. As one would expect, maternal buffering results in birth weight showing far less variation from season to season. However, in all progeny groups birth weight appears to decrease initially to 1957 then slowly recover with a final increase in 1965.

Weaner body weight in comparison shows much more marked variation between seasons. Interestingly, the fluctuations for this character are similar to those observed for date of birth particularly during the years 1956 to 1960 inclusive. During the following growth period to hogget body weight differences between sexes are further accentuated while the opposite is true between birth types. Flocks are similar initially with the Visual flock outweighing the Index flock over the latter stages of the trial. From the frequency of significant differences it can be seen that this difference in weight is larger within female than male progeny. It would be rather foolish to attempt to explain this in terms of the selection practiced and it should be repeated that the two selection methods differed only in the method of ram replacement with the ewes of both flocks being chosen on visual appraisal.

No consistent differences can be seen between the two flocks for the fleece weight of lambs. Even sex differences are non-existent at this early stage. However, the usual differences due to birth type and seasonal fluctuations are quite apparent.

As would be expected, the pattern for greasy fleece weight closely resembles clean fleece weight, which has been described earlier. While males have greater clean fleece weight than females this difference is larger for greasy fleece weight. This change arises as the percentage yield of males is generally lower than females. Within each sex, Index progeny have higher greasy fleece weights than Visual progeny but the divergence between these flocks is less obvious for this variate than for clean fleece weight. Again, percentage yield provides the link as it can be seen that it is signficantly higher for Index progeny for most years from 1957 onwards. Thus, partial selection for clean fleece weight has resulted in greater clean fleece weights in the Index flock than the Visual flock and there has been an opposing change in the level of contamination which together result in less marked changes in greasy fleece weight. The results suggest that the rate of response in male progeny is greater than for female progeny within both flocks, but this has not been established statistically. Even if this sex by genotype interaction had been proven it would still remain unclear whether this was due to sex-linked genes (see Beilharz, 1963; and James, 1973) or simply a scale effect associated with the greater fleece weight of males.

Table I-3(9) indicates that the Index progeny have significantly greater staple length than the Visual progeny. This significance occurs comparatively early in the experiment (i.e. 1955 for the single females and 1956 for the corresponding males). However, while the male progeny continue to exhibit significant differences, the females in comparison fail to show similar differences from 1962 onwards.

The graphical presentation for this variable suffers from two problems:

i) the two sexes overlap - i.e. the females initially (1955 and 1956) have longer staples than males from the same flock, but by the end of the experiment this has been reversed.

ii) There is a remarkable change in mean staple length for all groups during 1962. Why this season should differ so markedly from the others is unclear. Also it is interesting to note that apart from fleece weight (clean and greasy) no other variables show anywhere near the discontinuity observed for staple length during 1962. Further, the clean fleece weight is no greater than observed in other years. All this suggests that the 1962 season differed from other seasons in some manner which specifically affected staple length (and therefore fleece weight) but other variables were unaffected. The simplest explanation would be that the sheep were either shorn early at lamb shearing or late at hogget shearing or both.

For the variate crimps per inch, we observe that the mean for the Index flock becomes significantly less than for the Visual flock. Although the difference occurs for both sexes it becomes significant earlier in the females (1955) than the males (1959). There appears to be a decrease in crimp number over time but the 1954 and 1955 means complicate this conclusion.

Comparison of the plots for hogget body weight and crimps per inch suggests that these two variates are negatively correlated over the 12 years of the experiment. This relationship is of further interest when we look at fibre diameter as this variate behaves similarly to hogget body weight. Like the other variates, fibre diameter is frequently but not always significant, however, the direction of the difference between flocks is inconsistent over the 12 years. This would suggest that genetic differences have developed between the two flocks but unlike the differences for the other variates this character shows complex interactions with the environment. Mean coefficient of variation of fibre diameter is similar to mean fibre diameter in that it shows both positive and negative differences between the Index and Visual flocks but its pattern over the 12 years is almost opposite. Thus, we have body weight and fibre diameter on one hand and crimps per inch and coefficient of variation of fibre diameter on the other. Primary follicle number and secondary follicle number can also be included in the latter group. Also, the fluctuations in clean fleece weight and birth weight bear some resemblance to the former group of fibre diameter and hogget body weight. It should be recognised that this division of variates into those which either respond positively or negatively to seasonal conditions is based on purely subjective assessment of the graphs presented. However, the observation, although speculative, illustrates the large effect of seasonal conditions over the twelve years of the experiment. Thus, when the individual variates are considered, the two groups seem reasonable from a biological point of view. That is, clean fleece weight, birth weight, hogget body weight and fibre diameter respond positively during good seasons whereas crimps per inch, coefficient of variation of fibre diameter, primary follicle number and secondary follicle number are comparatively larger in poorer seasons.

(Note that as follicle numbers are expressed on a per square cm basis, this suggests that the main effect of seasons is on total surface area of the animals.)

For primary follicle number there is no consistent divergence between the two flocks (i.e. of the five significant t-tests three occur for the male progeny in the first three years). In comparison, secondary follicle number appears more favoured by the Index than by the Visual selection method.

As skin thickness was not measured during 1954 to 1957 inclusive, it is difficult to assess how the two selection methods have affected this variate. However, in all cases of significance (as determined by the t-test), the Index flock has greater skin thickness than the Visual flock.

While tables I-1 and I-2 list the values of skewness and kurtosis, along with their respective standard errors, these statistics will not be considered in detail until Chapter VI of this thesis. At this stage, it need only be noted that there was no variate for which the lack of normality was consistent enough to require a transformation.

Clearly the present investigations could have been continued further, however, as the main aim of this chapter is to provide a summary of the Roseworthy Experiment, using univariate statistics, no further tests will be applied. Mayo, Potter, Brady and Hooper (1969) have presented regression coefficients of the difference between the Index and Visual flock means to establish whether the flocks diverged. While this led them to similar conclusions as the present analysis, the use of the regression coefficient overlooks any non-linearity in the divergence, which can be seen from the figures to often be present.

Although the interpretation of these figures remains subjective it was felt that this was the best way to present the data.

SUMMARY

The aim of this chapter has been to use simple statistical techniques to enable an introductory evaluation of the twelve years of selection. It was neither desired nor expected that this approach would resolve all the questions of interest to the breeder. (The data will be considered in greater detail in later chapters.)

For comparative progress, both parametric and non-parametric tests have been used. This comparison illustrates that the Kolmogorov-Smirnov test can make a valuable contribution to quantitative genetics particularly as it tests for any difference in distribution (i.e. it is not constrained to the means and variances of normal distributions).

Comparisons between the two selection techniques over the 12 years of the experiment enable the fifteen variates to be divided into the following four general groups:

i) No consistent differences between the two flocks
- date of birth, birth weight, weaner body weight, lamb fleece weight and primary follicle number.

ii) Index flock consistently greater than Visual flock
- greasy fleece weight, percentage yield, clean fleece weight, staple length, secondary follicle number and skin thickness.

iii) Index flock consistently less than Visual flock
- hogget body weight and crimps per inch.

iv) Fluctuating differences between flocks
- fibre diameter and coefficient of variation of fibre diameter.

The investigation also highlights the effect of seasonal variation over the 12 years.

TABLE I-1(1) Basic statistics for date of birth (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	71	111	146	120.99	44.2141	.7891	.81±.28	1.49±.56
1955	71	106	136	118.06	32.1682	.6731	.84±.28	1.02±.56
1956	66	108	142	121.17	79.9872	1.1009	.69±.29	-.58±.58
1957	56	107	151	117.88	90.6568	1.2723	1.72±.32	2.57±.63
1958	67	116	158	127.70	79.1823	1.0871	1.28±.29	1.59±.58
1959	74	114	151	124.45	54.1957	.8558	.94±.28	.86±.55
1960	59	119	159	136.81	98.4302	1.2916	.43±.31	-.51±.61
1961	70	110	150	124.63	82.0919	1.0829	.36±.29	-.69±.57
1962	80	110	144	122.58	60.3234	.8684	.55±.27	-.44±.53
1963	64	109	145	121.56	73.4881	1.0716	1.01±.30	.57±.59
1964	76	113	146	123.88	43.7325	.7586	.98±.28	1.55±.54
1965	46	114	171	137.96	243.1092	2.2989	.51±.35	-.73±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	111	138	120.94	26.1572	.6444	.99±.30	1.72±.59
1955	79	108	148	119.33	62.7621	.8913	1.14±.27	1.14±.53
1956	66	110	141	120.17	59.2179	.9472	.87±.29	-.03±.58
1957	84	108	152	115.13	47.0067	.7481	2.63±.26	9.36±.52
1958	62	114	162	128.95	117.5550	1.3770	1.55±.30	2.39±.60
1959	71	112	146	122.30	46.0398	.8053	1.09±.28	1.75±.56
1960	56	118	155	130.80	87.1789	1.2477	.91±.32	.29±.63
1961	73	90	146	120.99	58.9581	.8987	-.16±.28	3.39±.56
1962	85	110	150	121.54	72.0370	.9206	1.22±.26	1.29±.52
1963	69	107	154	121.32	85.6616	1.1142	.83±.29	1.12±.57
1964	67	114	136	123.64	30.0516	.6697	.02±.29	-.97±.58
1965	77	111	158	126.31	99.1384	1.1347	1.13±.27	1.09±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	112	143	121.33	55.6316	1.0057	.90±.32	.50±.63
1955	62	106	139	118.34	51.5391	.9117	.72±.30	.30±.60
1956	46	105	143	119.63	74.8604	1.2757	.81±.35	.26±.69
1957	67	109	145	117.58	52.3379	.8838	1.82±.29	3.42±.58
1958	57	116	149	125.91	55.0815	.9830	1.17±.32	1.16±.62
1959	72	114	188	126.46	123.7447	1.3110	2.71±.28	11.39±.56
1960	83	121	162	135.99	76.5242	.9602	.43±.26	-.35±.52
1961	71	111	149	123.10	73.1473	1.0150	.81±.28	.19±.56
1962	72	113	156	124.99	77.7040	1.0389	.97±.28	1.03±.56
1963	43	110	153	120.77	93.6113	1.4755	1.65±.36	2.44±.71
1964	74	112	142	122.34	38.0076	.7167	.80±.28	.34±.55
1965	37	112	183	132.73	346.2027	3.0589	1.04±.39	.12±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	107	150	120.65	45.2028	.7979	1.33±.28	4.05±.56
1955	58	109	141	117.60	61.6470	1.0310	1.26±.31	.80±.62
1956	78	108	140	118.94	38.0867	.6988	.96±.27	.77±.54
1957	62	106	136	114.03	32.2285	.7210	2.04±.30	4.44±.60
1958	60	116	158	128.33	89.1751	1.2191	1.34±.31	1.24±.61
1959	75	114	148	124.03	56.9182	.8712	.88±.28	.40±.55
1960	73	120	159	131.95	60.6636	.9116	.93±.28	1.06±.56
1961	86	110	146	120.48	59.8995	.8346	.85±.26	.52±.51
1962	82	108	141	121.37	64.9756	.8902	.73±.27	-.07±.53
1963	70	105	152	122.13	91.3600	1.1424	.84±.29	.65±.57
1964	67	114	151	125.03	57.4233	.9258	1.04±.29	1.34±.58
1965	70	114	157	124.66	77.2720	1.0507	1.37±.29	2.10±.57
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	40	86	134	103.83	119.8917	1.7313	.57±.37	-.10±.73
1950	58	103	148	117.53	96.3234	1.2887	1.18±.31	1.30±.62
1951	88	104	137	114.76	52.5056	.7724	.89±.26	.41±.51
1952	93	103	148	117.91	102.4273	1.0495	1.14±.25	.76±.50
1953	104	103	139	114.04	54.7364	.7255	.95±.24	.46±.47

TABLE I-1(2) Basic statistics for birth weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	71	2.72	5.90	4.48	.4680	.0812	-.37±.28	-.21±.56
1955	71	2.04	5.90	4.09	.5914	.0913	-.17±.28	.16±.56
1956	66	2.27	5.44	3.92	.5675	.0927	-.22±.29	-.52±.58
1957	56	1.59	5.22	3.91	.5705	.1009	-.97±.32	1.19±.63
1958	67	2.27	5.67	4.12	.5641	.0918	.00±.29	-.47±.58
1959	74	1.36	6.12	4.15	.8804	.1091	-.28±.28	.06±.55
1960	59	2.72	6.35	4.34	.4648	.0888	-.04±.31	.45±.61
1961	70	2.72	6.12	4.37	.5661	.0899	.07±.29	-.61±.57
1962	80	2.72	5.90	4.30	.4101	.0716	.14±.27	.01±.53
1963	64	2.72	5.90	4.31	.5217	.0903	-.04±.30	-.42±.59
1964	76	2.50	5.67	4.43	.4533	.0772	-.43±.28	-.11±.54
1965	46	3.86	6.12	5.01	.2908	.0795	.08±.35	-.28±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	2.27	5.67	4.18	.5256	.0913	-.29±.30	-.00±.59
1955	79	2.50	5.90	4.22	.5675	.0848	-.02±.27	-.51±.53
1956	66	1.81	5.22	3.87	.5055	.0875	-.38±.29	.24±.58
1957	84	2.27	5.90	4.02	.4719	.0750	-.29±.26	.39±.52
1958	62	2.72	5.22	4.06	.3155	.0713	-.35±.30	-.41±.60
1959	71	2.50	6.12	4.25	.5866	.0909	-.24±.28	-.40±.56
1960	56	2.72	5.90	4.46	.3693	.0812	-.34±.32	.46±.63
1961	73	2.04	5.90	4.08	.5544	.0871	-.25±.28	.03±.56
1962	85	2.50	6.12	4.49	.5117	.0776	-.27±.26	.19±.52
1963	69	2.04	5.44	4.19	.4548	.0812	-.40±.29	.20±.57
1964	67	2.50	5.90	4.41	.5818	.0932	-.11±.29	-.40±.58
1965	77	2.04	6.12	4.77	.5157	.0818	-.83±.27	1.74±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	2.50	5.44	3.95	.4682	.0923	-.00±.32	-.68±.63
1955	62	1.81	5.22	3.78	.6084	.0991	-.19±.30	-.28±.60
1956	46	2.50	4.99	3.60	.2931	.0798	.06±.35	-.35±.69
1957	67	1.36	5.22	3.69	.6135	.0957	-.78±.29	.43±.58
1958	57	2.72	5.67	3.92	.5400	.0973	.10±.32	-.57±.62
1959	72	2.04	5.44	3.94	.5105	.0842	-.27±.28	-.46±.56
1960	83	2.72	5.44	3.99	.3964	.0691	.13±.26	-.75±.52
1961	71	2.72	5.67	4.00	.6029	.0921	.21±.28	-.71±.56
1962	72	2.72	5.44	4.09	.4267	.0770	.00±.28	-.42±.56
1963	43	2.72	6.12	4.03	.4454	.1018	.39±.36	.69±.71
1964	74	2.50	5.67	4.16	.5216	.0840	-.26±.28	-.67±.55
1965	37	3.40	5.90	4.58	.3339	.0950	.07±.39	-.15±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	2.72	5.90	4.04	.4526	.0798	.02±.28	-.19±.56
1955	58	2.04	4.99	3.79	.4941	.0923	-.50±.31	-.20±.62
1956	78	2.04	5.44	3.57	.5256	.0821	.09±.27	-.31±.54
1957	62	2.04	5.22	3.61	.3705	.0773	.17±.30	.30±.60
1958	60	2.72	5.22	3.87	.3806	.0796	-.07±.31	-.53±.61
1959	75	1.81	5.22	3.95	.6464	.0928	-.62±.28	-.41±.55
1960	73	1.81	4.99	3.78	.4170	.0756	-.50±.28	.41±.56
1961	86	2.27	5.67	3.96	.4667	.0737	-.14±.26	.11±.51
1962	82	1.59	5.67	4.19	.6264	.0874	-.71±.27	.82±.53
1963	70	2.50	5.44	4.05	.4527	.0804	.12±.29	-.34±.57
1964	67	2.50	5.44	4.29	.3026	.0672	-.58±.29	.39±.58
1965	70	3.40	5.90	4.48	.3797	.0737	.11±.29	-.57±.57
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	40	2.72	5.67	4.09	.3712	.0963	.12±.37	.28±.73
1950	58	2.04	4.99	3.84	.3571	.0785	-.71±.31	.83±.62
1951	88	2.04	5.44	3.88	.5094	.0761	.00±.26	-.05±.51
1952	93	1.59	5.67	3.94	.4558	.0700	-.10±.25	.72±.50
1953	104	1.81	4.99	3.73	.4063	.0625	-.47±.24	.23±.47

TABLE I-1(3) Basic statistics for weaner body weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	71	25.13	39.10	30.68	8.1601	.3390	.77± .28	.71± .56
1955	71	16.24	41.10	32.73	17.4219	.4954	-.62± .28	2.12± .56
1956	66	19.55	38.92	28.47	15.2953	.4814	.12± .29	.34± .58
1957	56	12.38	27.53	21.58	11.7692	.4584	-.48± .32	.19± .63
1958	67	17.96	35.52	27.80	14.0304	.4576	-.41±.29	-.27±.58
1959	74	14.88	37.10	23.80	14.3499	.4404	.45±.28	1.26±.55
1960	59	19.37	41.32	28.15	20.1048	.5837	.34±.31	-.03±.61
1961	70	19.55	41.50	31.24	21.6071	.5556	.06±.29	-.51±.57
1962	80	22.36	42.73	30.85	18.6897	.4833	.06±.27	-.11±.53
1963	64	19.96	43.73	34.80	26.4454	.6428	-.60±.30	-.06±.59
1964	76	23.95	43.50	32.99	18.1681	.4889	.29±.28	-.25±.54
1965	46	23.77	40.91	32.56	12.1985	.5150	.25±.35	.10±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	22.14	41.32	30.59	11.9612	.4357	.04±.30	.61±.59
1955	79	22.54	45.31	33.77	14.3983	.4269	-.15±.27	.96±.53
1956	66	19.14	39.92	28.66	18.1912	.5250	-.03±.29	-.18±.58
1957	84	9.98	29.35	22.32	13.5915	.4022	-.91±.26	1.40±.52
1958	62	17.15	34.93	27.85	15.0561	.4928	-.45±.30	-.22±.60
1959	71	17.96	30.35	24.08	7.1782	.3180	.06±.28	-.09±.56
1960	56	21.55	40.51	31.08	17.9479	.5661	-.12±.32	-.61±.63
1961	73	17.15	39.92	30.60	15.1170	.4551	-.72±.28	1.08±.56
1962	85	23.95	43.73	32.62	15.2611	.4237	.02±.26	.02±.52
1963	69	21.36	53.98	34.39	24.6432	.5976	.70±.29	2.58±.57
1964	67	15.24	43.32	32.94	22.8769	.5843	-.66±.29	1.76±.58
1965	77	22.54	41.91	33.25	17.2273	.4730	-.28±.27	-.21±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	17.78	33.34	26.62	9.8872	.4240	-.28±.32	.13±.63
1955	62	22.54	38.51	29.54	9.7387	.3963	.31±.30	.34±.60
1956	46	17.96	32.34	24.59	9.2461	.4483	.51±.35	.13±.69
1957	67	11.57	25.76	20.04	8.7432	.3612	-.70±.29	.42±.58
1958	57	10.98	31.93	25.11	13.4118	.4851	-1.15±.32	2.66±.62
1959	72	13.56	25.95	21.49	6.9283	.3102	-.67±.28	.53±.56
1960	83	15.38	36.51	25.28	13.2366	.3993	.22±.26	.77±.52
1961	71	21.14	34.52	27.70	8.3097	.3421	.05±.28	.17±.56
1962	72	18.37	31.93	26.42	9.6023	.3652	-.49±.28	-.31±.56
1963	43	19.96	35.11	28.56	10.7586	.5002	-.34±.36	.13±.71
1964	74	18.37	35.34	27.67	12.0263	.4031	-.35±.28	.27±.55
1965	37	17.78	33.93	27.45	12.1234	.5724	-.51±.39	.15±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	19.96	30.35	25.89	4.7015	.2573	-.34±.28	-.01±.56
1955	58	20.37	43.91	28.97	13.2212	.4774	1.06±.31	3.69±.62
1956	78	18.78	31.34	25.52	7.0091	.2998	-.02±.27	-.16 ±.54
1957	62	14.15	28.76	20.75	6.2344	.3171	.44±.30	1.23±.60
1958	60	14.56	31.93	24.62	12.5575	.4575	-.37±.31	.15 ±.61
1959	75	14.79	28.12	21.72	8.6442	.3395	-.32±.28	-.32 ±.55
1960	73	14.56	35.11	25.82	14.4905	.4455	-.48±.28	.42 ±.56
1961	86	15.15	33.52	26.86	11.9865	.3733	-.85±.26	1.49 ±.51
1962	82	23.36	34.11	28.64	8.5194	.3223	.03±.27	-1.00 ±.53
1963	70	21.77	37.10	28.95	12.8231	.4280	.03±.29	-.63 ±.57
1964	67	20.77	33.93	28.38	10.1150	.3885	-.33±.29	-.54 ±.58
1965	70	14.79	35.11	28.25	8.8597	.3558	-1.00±.29	4.62 ±.57
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	40	14.97	27.67	22.36	8.0947	.4499	-.51±.37	-.30 ±.73
1950	58	17.78	28.12	23.47	6.7022	.3399	-.24±.31	-.72 ±.62
1951	88	16.56	34.11	26.53	7.7272	.2963	-.51±.26	1.11 ±.51
1952	93	19.96	33.52	26.14	8.8784	.3090	.19±.25	-.02 ±.50
1953	104	18.14	32.75	26.65	9.0692	.2953	-.54±.24	-.05 ±.47

TABLE I-1(4) Basic statistics for hogget body weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	71	45.90	75.25	63.70	31.1109	.6620	-.30 ±.28	.49±.56
1955	71	42.32	66.45	54.99	22.5855	.5640	-.50 ±.28	.22±.56
1956	66	37.51	59.87	48.63	23.3549	.5949	.07 ±.29	-.13±.58
1957	56	32.34	57.88	47.37	30.6169	.7394	-.51 ±.32	.24±.63
1958	67	41.37	66.95	54.11	22.7325	.5825	-.02 ±.29	.33±.58
1959	74	45.50	66.04	56.30	21.8763	.5437	.15 ±.28	-.55±.55
1960	59	52.30	76.43	62.45	29.6216	.7086	.48 ±.31	-.06±.61
1961	70	48.08	75.84	60.57	40.4512	.7602	.26 ±.29	-.36±.57
1962	80	44.91	74.25	58.09	30.9983	.6225	-.13 ±.27	-.16±.53
1963	64	52.71	76.25	63.45	33.3755	.7221	.26 ±.30	-.87±.59
1964	76	47.90	78.24	62.06	34.1114	.6700	-.14 ±.28	.10±.54
1965	46	52.48	73.66	61.88	22.6688	.7020	.32 ±.35	-.36±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	50.71	75.25	63.14	27.2407	.6576	-.15 ±.30	-.23±.59
1955	79	39.92	71.67	56.17	28.5922	.6016	-.25 ±.27	1.15±.53
1956	66	38.51	60.69	49.05	21.0049	.5641	.17 ±.29	-.20±.58
1957	84	35.52	58.47	47.80	21.5413	.5064	-.47 ±.26	.18±.52
1958	62	41.14	68.72	53.43	26.6349	.6554	.24 ±.30	.45±.60
1959	71	46.49	66.86	57.07	20.0331	.5312	.08 ±.28	-.28±.56
1960	56	54.30	77.25	64.31	33.1923	.7699	.21 ±.32	-.88±.63
1961	73	39.51	73.84	59.40	44.1687	.7779	-.74 ±.28	1.08±.56
1962	85	41.32	72.57	60.36	31.4728	.6085	-.74 ±.26	1.43±.52
1963	69	45.50	73.84	64.06	26.4352	.6190	-.54 ±.29	1.09±.57
1964	67	48.22	75.25	64.20	30.4814	.6745	-.24 ±.29	.16±.58
1965	77	48.31	77.93	62.53	32.6507	.6512	.15 ±.27	.23±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	26.13	58.47	46.90	27.4070	.7059	-1.21 ±.32	3.73±.63
1955	62	32.93	58.47	43.80	15.5826	.5013	.56 ±.30	2.26±.60
1956	46	28.94	52.07	39.13	20.5396	.6682	.17 ±.35	.63±.69
1957	67	26.76	47.31	38.54	16.0882	.4900	-.22±.29	-.02±.58
1958	57	31.75	53.52	44.84	19.0433	.5780	-.46±.32	.20±.62
1959	72	22.36	47.72	39.16	18.7776	.5107	-.91±.28	2.15±.56
1960	83	29.35	47.31	39.82	15.3330	.4298	-.43±.26	-.33±.52
1961	71	32.52	51.71	40.66	15.8966	.4732	.25±.28	.12±.56
1962	72	30.75	52.07	42.63	18.4675	.5065	-.14±.28	.42±.56
1963	43	34.52	53.48	45.77	16.1744	.6133	-.41±.36	.19±.71
1964	74	29.76	54.48	44.24	34.7566	.6853	-.37±.28	-.75±.55
1965	37	35.24	56.47	42.95	26.2500	.8423	.85±.39	.31±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	33.52	59.47	46.66	15.3899	.4656	-.09±.28	1.69±.56
1955	58	36.11	49.31	42.31	11.4250	.4438	.07±.31	-.67±.62
1956	78	28.76	51.89	39.36	18.6746	.4893	-.01±.27	-.04±.54
1957	62	30.94	50.30	39.28	11.2802	.4265	.35±.30	1.05±.60
1958	60	34.56	52.75	43.88	12.8474	.4627	.06±.31	-.02±.61
1959	75	31.93	47.31	39.20	12.3462	.4057	.10±.28	-.86±.55
1960	73	29.94	52.89	41.31	20.5430	.5305	.13±.28	-.45±.56
1961	86	18.55	50.71	39.50	26.6462	.5566	-.87±.26	2.44±.51
1962	82	37.42	53.89	46.23	18.1415	.4704	-.10±.27	-1.09±.53
1963	70	35.92	57.70	45.95	18.3838	.5125	.28±.29	.33±.57
1964	67	35.34	57.70	47.04	19.5368	.5400	-.19±.29	.14±.58
1965	70	35.24	55.47	45.18	20.4125	.5400	.26±.29	-.63±.57
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	40	26.13	44.32	34.58	11.4746	.5356	-.14±.37	1.41±.73
1950	58	31.53	54.07	41.11	16.7359	.5372	.06±.31	.79±.62
1951	88	30.53	51.89	39.41	16.0140	.4266	.64±.26	.82±.51
1952	93	28.94	57.29	40.77	15.2676	.4052	.54±.25	2.68±.50
1953	104	36.51	53.30	44.58	12.4303	.3457	-.28±.24	-.33±.47

TABLE I-1(5) Basic statistics for lamb fleece weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	71	1.00	2.13	1.57	.0550	.0278	.30±.28	-.06±.56
1955	71	1.04	2.54	1.82	.0754	.0326	-.01±.28	.65±.56
1956	66	.73	2.27	1.46	.0976	.0384	.12±.29	.10±.58
1957	56	.50	1.27	.89	.0318	.0238	-.05±.32	-.21±.63
1958	67	.68	2.04	1.26	.0738	.0332	.60±.29	.56±.58
1959	74	.64	1.86	1.20	.0719	.0312	.18±.28	-.47±.55
1960	59	.73	1.81	1.24	.0765	.0360	.14±.31	-.73±.61
1961	70	.68	2.27	1.38	.1202	.0414	.32±.29	-.30±.57
1962	80	.86	2.18	1.39	.0689	.0294	.36±.27	.10±.53
1963	64	.73	3.40	2.25	.1846	.0537	-.11±.30	1.72±.59
1964	76	1.41	3.77	2.35	.2024	.0516	.44±.28	.76±.54
1965	46	1.27	3.04	2.08	.2140	.0682	.13±.35	-.84±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	1.04	2.04	1.52	.0578	.0303	.18±.30	-.59±.59
1955	79	1.09	2.72	1.80	.1085	.0371	.08±.27	-.06±.53
1956	66	.82	2.27	1.46	.0861	.0361	.27±.29	-.12±.58
1957	84	.27	1.36	.97	.0456	.0233	-.73±.26	1.30±.52
1958	62	.64	1.95	1.28	.0785	.0356	-.34±.30	-.40±.60
1959	71	.68	1.68	1.23	.0582	.0286	.26±.28	-.50±.56
1960	56	.73	2.63	1.46	.1117	.0447	.49±.32	1.62±.63
1961	73	.82	2.18	1.37	.0884	.0348	.04±.28	-.24±.56
1962	85	.77	2.18	1.42	.0629	.0272	-.24±.26	.78±.52
1963	69	1.36	3.08	2.26	.1141	.0407	.05±.29	.28±.57
1964	67	1.45	3.04	2.30	.1384	.0454	-.16±.29	-.52±.58
1965	77	1.45	3.08	2.22	1221	.0398	-.08±.27	-.28±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	.91	2.18	1.48	.0747	.0369	.13±.32	-.12±.63
1955	62	1.27	2.72	1.83	.1016	.0405	.42±.30	-.31±.60
1956	46	.77	2.13	1.56	.0866	.0434	-.21±.35	.19±.69
1957	67	.54	1.36	.93	.0321	.0219	-.03±.29	-.06±.58
1958	57	.50	1.91	1.34	.1040	.0427	-.52±.32	-.14±.62
1959	72	.73	1.95	1.22	.0526	.0270	.21±.28	.33±.56
1960	83	.68	2.36	1.24	.0806	.0312	.82±.26	1.57±.52
1961	71	.77	2.18	1.41	.0890	.0354	.19±.28	-.28±.56
1962	72	.91	2.04	1.39	.0612	.0292	.43±.28	-.09±.56
1963	43	1.41	2.68	2.08	.0887	.0454	-.27±.36	-.30±.71
1964	74	1.63	3.54	2.31	.1269	.0414	.64±.28	.58±.55
1965	37	1.09	2.99	2.05	.1488	.0634	-.33±.39	.73±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	.82	1.91	1.49	.0436	.0248	-.34±.28	.25±.56
1955	58	1.04	2.50	1.74	.1003	.0416	.13±.31	.30±.62
1956	78	.95	1.95	1.46	.0492	.0251	.19±.27	-.36±.54
1957	62	.45	1.72	.97	.0398	.0253	.68±.30	2.19±.60
1958	60	.64	1.95	1.28	.0751	.0354	.36±.31	.53±.61
1959	75	.50	1.68	1.17	.0512	.0261	-.56±.28	.43±.55
1960	73	.54	2.09	1.32	.0978	.0366	-.31±.28	.05±.56
1961	86	.77	2.18	1.40	.0886	.0321	.04±.26	-.20±.51
1962	82	1.00	1.91	1.41	.0495	.0246	-.04±.27	-.71±.53
1963	70	1.41	2.86	2.13	.1027	.0383	.04±.29	-.19±.57
1964	67	1.32	2.81	2.19	.1194	.0422	-.12±.29	-.51±.58
1965	70	1.23	2.86	2.14	.1179	.0410	.14±.29	-.22±.57
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	40	.59	2.59	1.58	.1713	.0654	-.12±.37	-.02±.73
1950	58	.86	2.00	1.29	.0487	.0290	.41±.31	.44±.62
1951	88	1.04	2.22	1.65	.0697	.0282	-.10±.26	-.57±.51
1952	93	.77	2.09	1.33	.0746	.0283	.27±.25	-.21±.50
1953	104	.64	1.77	1.22	.0550	.0230	-.45±.24	.02±.47

TABLE I-1(6) Basic statistics for hogget greasy fleece weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	71	5.81	10.66	8.25	.9562	.1161	.27±.28	.23±.56
1955	71	4.35	7.85	5.80	.5430	.0875	.26±.28	-.01±.56
1956	66	4.26	8.66	6.41	.6666	.1005	.05±.29	.02±.58
1957	56	3.27	6.99	5.50	.6009	.1036	-.23±.32	-.03±.63
1958	67	4.72	8.85	6.76	.6278	.0968	.21±.29	.07±.58
1959	74	4.67	8.39	6.53	.5808	.0886	.06±.28	-.07±.55
1960	59	5.81	9.75	7.38	.6074	.1015	.56±.31	.40±.61
1961	70	4.67	8.66	6.71	.7348	.1025	-.17±.29	-.37±.57
1962	80	5.22	9.71	7.28	.6638	.0911	-.09±.27	.82±.53
1963	64	5.35	9.12	7.22	.8341	.1142	.13±.30	-.68±.59
1964	76	4.45	9.16	7.17	.8946	.1085	-.23±.28	.06±.54
1965	46	5.90	8.94	7.54	.4933	.1036	-.41±.35	-.48±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	5.22	10.89	8.16	.9560	.1232	-.17±.30	.73±.59
1955	79	3.95	7.39	5.87	.3857	.0699	-.51±.27	.72±.53
1956	66	4.72	7.76	6.08	.4945	.0866	.28±.29	-.57±.58
1957	84	3.18	8.17	5.42	.6637	.0889	-.10±.26	1.43±.52
1958	62	3.99	8.57	6.58	.5858	.0972	-.56±.30	1.57±.60
1959	71	4.22	7.89	6.22	.5760	.0901	-.15±.28	-.20±.56
1960	56	5.63	8.85	7.24	.5280	.0971	-.32±.32	.01±.63
1961	73	2.81	8.07	6.15	.9473	.1139	-.48±.28	.50±.56
1962	85	5.44	9.12	7.17	.6471	.0873	-.18±.26	-.37±.52
1963	69	5.40	8.71	6.86	.6081	.0939	.32±.29	-.54±.57
1964	67	5.35	8.21	6.68	.4178	.0790	.05±.29	-.47±.58
1965	77	5.49	8.71	6.96	.6664	.0930	.09±.27	-.86±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	4.26	8.07	6.49	.6735	.1107	-.30±.32	.13±.63
1955	62	4.22	7.17	5.74	.3787	.0782	.06±.30	-.07±.60
1956	46	4.72	6.67	5.63	.2810	.0782	.24±.35	-.87±.69
1957	67	3.45	6.49	5.01	.2931	.0661	-.09±.29	.84±.58
1958	57	3.67	7.80	6.26	.5634	.0994	-.54±.32	1.13±.62
1959	72	4.04	6.94	5.40	.4370	.0779	.01±.28	-.61±.56
1960	83	4.76	8.07	6.11	.5311	.0800	.27±.26	-.45±.52
1961	71	3.86	6.31	4.84	.3158	.0667	.33±.28	-.31±.56
1962	72	3.86	7.53	5.95	.4950	.0829	-.44±.28	.58±.56
1963	43	5.04	7.71	6.09	.4247	.0994	.29±.36	-.38±.71
1964	74	3.90	7.44	5.53	.3380	.0676	-.02±.28	.97±.55
1965	37	5.08	7.58	5.79	.3664	.0995	1.28±.39	1.24±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	4.94	8.26	6.50	.4912	.0832	.20±.28	.11±.56
1955	58	4.49	6.85	5.33	.2510	.0658	.49±.31	.27±.62
1956	78	4.17	6.58	5.23	.2976	.0618	.27±.27	-.39±.54
1957	62	3.54	6.40	4.91	.3032	.0699	.29±.30	-.02±.60
1958	60	4.81	7.44	6.00	.3577	.0772	.40±.31	-.58±.61
1959	75	3.90	6.49	5.06	.3651	.0698	.27±.28	-.44±.55
1960	73	3.99	7.17	5.77	.4987	.0827	-.26±.28	-.51±.56
1961	86	2.54	6.21	4.61	.4422	.0717	-.07±.26	.54±.51
1962	82	4.54	7.67	5.86	.4452	.0737	.45±.27	.21±.53
1963	70	4.58	7.21	5.82	.4048	.0760	.05±.29	-.64±.57
1964	67	3.99	7.08	5.22	.4654	.0833	.41±.29	-.27±.58
1965	70	4.45	7.80	5.75	.4482	.0800	.65±.29	.49±.57
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	40	3.13	4.90	4.02	.2378	.0771	-.04±.37	-1.02±.73
1950	58	4.31	7.08	5.62	.3363	.0761	.32±.31	.14±.62
1951	88	4.54	7.76	5.75	.3862	.0662	.11±.26	.25±.51
1952	93	3.99	6.99	5.24	.3149	.0582	.64±.25	.53±.50
1953	104	4.17	7.85	5.75	.4454	.0654	.27±.24	.07±.47

TABLE I-1(7) Basic statistics for percentage yield (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	71	41.40	70.90	62.23	23.3734	.5738	-1.42±.28	3.54±.56
1955	71	55.30	76.20	65.89	15.5893	.4686	-.27±.28	.06±.56
1956	66	41.10	67.90	55.64	28.3409	.6553	-.03±.29	.22±.58
1957	56	50.90	67.30	59.40	19.0202	.5828	-.09±.32	-.90±.63
1958	67	54.90	73.60	63.90	16.2076	.4918	.05±.29	-.31±.58
1959	74	45.20	74.00	61.16	31.5252	.6527	-.54±.28	.50±.55
1960	59	51.50	72.80	63.97	20.9498	.5959	-.66±.31	.45±.61
1961	70	44.80	69.00	58.97	25.7994	.6071	-.43±.29	-.11±.57
1962	80	55.70	73.40	64.48	21.2138	.5149	.31±.27	-.87±.53
1963	64	51.40	74.40	61.56	32.6478	.7142	.14±.30	-.90±.59
1964	76	51.60	73.30	63.84	26.4222	.5896	-.30±.28	-.55±.54
1965	46	48.10	68.60	59.62	28.1261	.7819	-.44±.35	-.48±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	53.50	72.60	61.93	20.6148	.5720	.11±.30	-.37±.59
1955	79	55.60	76.30	64.81	16.1648	.4523	.25±.27	-.07±.53
1956	66	40.80	67.30	55.76	26.3736	.6321	-.25±.29	.12±.58
1957	84	42.50	70.00	58.94	22.2164	.5143	-.45±.26	.85±.52
1958	62	51.10	68.10	60.88	16.3878	.5141	-.19±.30	-.40±.60
1959	71	45.80	72.90	60.78	26.3999	.6098	-.07±.28	.66±.56
1960	56	47.20	69.30	63.04	20.2526	.6014	-1.31±.32	1.88±.63
1961	73	40.50	69.70	56.57	30.0720	.6418	-.18±.28	.28±.56
1962	85	53.30	77.10	62.84	20.1771	.4872	.35±.26	.11±.52
1963	69	51.00	70.60	61.33	19.1348	.5266	-.07±.29	-.55±.57
1964	67	52.10	71.30	62.32	18.6903	.5282	-.19±.29	-.60±.58
1965	77	49.30	76.60	61.19	30.0539	.6247	.03±.27	.00±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	56.10	73.00	64.40	14.9402	.5212	.13±.32	-.21±.63
1955	62	56.20	80.60	66.90	19.5528	.5616	.03±.30	.79±.60
1956	46	49.60	69.70	62.74	15.3749	.5781	-.80±.35	1.45±.69
1957	67	48.30	70.80	61.90	22.6310	.5812	-.46±.29	-.13±.58
1958	57	54.50	78.40	67.59	22.7130	.6312	-.41±.32	.68±.62
1959	72	51.40	77.00	62.51	30.5321	.6512	.32±.28	-.38±.56
1960	83	53.10	76.00	64.79	19.8617	.4892	-.00±.26	.09±.52
1961	71	53.00	73.00	64.52	18.6948	.5131	-.30±.28	-.17±.56
1962	72	46.50	75.10	65.99	22.2934	.5564	-1.05±.28	2.68±.56
1963	43	54.00	70.00	64.07	16.3212	.6161	-.84±.36	.12±.71
1964	74	53.00	73.30	65.42	19.3165	.5109	-.56±.28	-.05±.55
1965	37	58.00	73.50	66.24	14.9568	.6358	-.11±.39	-.44±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	57.40	76.10	65.57	16.0202	.4750	.16±.28	-.12±.56
1955	58	59.70	72.90	67.26	11.6744	.4486	-.34±.31	-.66±.62
1956	78	51.60	71.00	62.03	16.5955	.4613	-.25±.27	-.06±.54
1957	62	50.70	72.20	59.00	18.4911	.5461	.42±.30	-.02±.60
1958	60	51.70	85.80	65.07	27.9264	.6822	.38±.31	3.31±.61
1959	75	50.10	71.40	60.93	19.9423	.5157	-.30±.28	-.44±.55
1960	73	53.00	72.10	63.99	13.6649	.4327	-.45±.28	.32±.56
1961	86	51.00	73.70	62.89	18.2704	.4609	-.40±.26	.84±.51
1962	82	54.50	72.00	63.61	14.5122	.4207	-.21±.27	.06±.53
1963	70	50.50	72.40	63.86	17.6118	.5016	-.20±.29	.40±.57
1964	67	52.80	71.50	64.19	16.8000	.5007	-.66±.29	-.03±.58
1965	70	50.90	77.40	64.63	23.7916	.5830	-.13±.29	.17±.57
DAMS BORN IN 1949, 1950,1951,1952,1953								
1949	40	54.20	71.40	62.72	18.6428	.6827	.24±.37	-.80±.73
1950	58	56.60	70.00	63.99	10.3802	.4230	-.20±.31	-.47±.62
1951	88	54.00	73.50	64.54	15.5239	.4200	-.29±.26	-.27±.51
1952	93	58.30	77.50	67.26	11.9256	.3581	-.03±.25	.27±.51
1953	104	44.80	75.70	67.27	18.2965	.4194	-1.42±.24	5.91±.47

TABLE I-1(8) Basic statistics for clean fleece weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive and used in the initial years of the experiment.)

Year	N	Min.	Max	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	71	3.08	6.67	5.14	.5615	.0889	-.28±.28	.08±.56
1955	71	2.77	5.35	3.82	.2989	.0649	.29±.28	-.27±.56
1956	66	2.81	4.63	3.55	.1968	.0546	.51±.29	-.40±.58
1957	56	1.86	4.17	3.26	.2141	.0618	-.40±.32	.18±.63
1958	67	2.95	5.58	4.31	.3062	.0676	.09±.29	-.28±.58
1959	74	2.72	5.13	3.99	.3356	.0673	-.22±.28	-.64±.55
1960	59	3.40	6.62	4.71	.3133	.0729	.69±.31	1.52±.61
1961	70	2.72	5.31	3.95	.3557	.0713	.09±.29	-.56±.57
1962	80	3.18	6.26	4.70	.3388	.0651	.07±.27	.18±.53
1963	64	2.95	6.03	4.43	.3894	.0780	.10±.30	-.10±.59
1964	76	3.22	6.03	4.58	.3541	.0683	.21±.28	-.16±.54
1965	46	3.22	5.26	4.49	.2978	.0805	-.26±.35	-.82±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	3.18	6.71	5.05	.5017	.0892	.23±.30	.30±.59
1955	79	2.50	4.76	3.79	.2278	.0537	-.27±.27	-.10±.53
1956	66	2.40	4.26	3.38	.1830	.0527	-.02±.29	-.51±.58
1957	84	1.86	4.45	3.19	.2369	.0531	-.03±.26	.25±.52
1958	62	2.04	5.35	3.97	.3069	.0704	-.76±.30	1.82±.60
1959	71	2.31	5.13	3.78	.3126	.0664	-.09±.28	.05±.56
1960	56	3.08	5.63	4.57	.3158	.0751	-.49±.32	-.13±.63
1961	73	1.72	4.81	3.46	.3196	.0662	-.06±.28	.26±.56
1962	85	3.22	6.17	4.52	.3720	.0662	.27±.26	-.01±.52
1963	69	3.27	5.81	4.20	.2364	.0585	.60±.29	.54±.57
1964	67	2.90	5.13	4.17	.1909	.0534	-.54±.29	.48±.58
1965	77	3.27	5.53	4.24	.2487	.0568	.17±.27	-.38±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	2.59	5.26	4.18	.3313	.0776	-.48±.32	-.05±.63
1955	62	2.90	4.99	3.83	.1659	.0517	.15±.30	-.15±.60
1956	46	2.81	4.35	3.53	.1273	.0526	.07±.35	-.11±.69
1957	67	2.18	3.99	3.09	.1172	.0418	.05±.29	.40±.58
1958	57	2.68	5.26	4.22	.2506	.0663	-.44±.32	.42±.62
1959	72	2.31	4.35	3.37	.2175	.0550	-.06±.28	-.76±.56
1960	83	2.22	5.22	3.93	.2720	.0572	-.20±.26	.58±.52
1961	71	2.31	4.35	3.12	.1682	.0487	.45±.28	.18±.56
1962	72	2.45	5.35	3.94	.2876	.0632	-.27±.28	.21±.56
1963	43	3.27	4.63	3.91	.1176	.0523	.11±.36	-.42±.71
1964	74	2.50	4.76	3.61	.1631	.0469	.18±.28	.54±.55
1965	37	2.99	4.94	3.82	.2158	.0764	.50±.39	-.41±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	3.36	5.40	4.25	.2075	.0541	.46±.28	-.17±.56
1955	58	2.90	4.90	3.58	.1294	.0472	.92±.31	2.10±.62
1956	78	2.45	4.54	3.24	.1544	.0445	.71±.27	.93±.54
1957	62	1.95	4.17	2.90	.1596	.0507	.43±.30	1.31±.60
1958	60	2.77	4.81	3.90	.2015	.0580	-.01±.31	-.48±.61
1959	75	2.22	4.17	3.08	.1742	.0482	.27±.28	-.35±.55
1960	73	2.54	4.94	3.69	.2276	.0558	-.08±.28	-.56±.56
1961	86	1.54	4.22	2.90	.2154	.0500	0.00±.26	.53±.51
1962	82	2.36	4.81	3.71	.2057	.0501	-.19±.27	.31±.53
1963	70	2.86	4.49	3.71	.1755	.0501	-.11±.29	-.55±.57
1964	67	2.45	4.81	3.35	.2237	.0578	.28±.29	.02±.58
1965	70	2.63	5.04	3.72	.2134	.0552	.23±.29	.46±.57
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	40	1.86	3.08	2.51	.0943	.0486	-.05±.37	-1.01±.73
1950	58	2.77	4.54	3.59	.1434	.0497	.14±.31	-.55±.62
1951	88	2.59	4.81	3.71	.1593	.0425	-.11±.26	.18±.51
1952	93	2.72	4.76	3.52	.1339	.0379	.58±.25	.48±.50
1953	104	2.50	5.17	3.87	.2828	.0521	.07±.24	-.20±.47

TABLE I-1(9) Basic statistics for staple length in cms (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	71	10.16	16.51	12.68	1.4178	.1413	.43±.28	.17±.56
1955	71	9.65	14.48	11.73	1.2221	.1312	.50±.28	-.34±.56
1956	66	9.91	14.73	11.82	1.2154	.1357	.54±.29	-.12±.58
1957	56	9.40	13.97	11.52	1.2139	.1472	.26±.32	-.55±.63
1958	67	8.64	13.46	11.68	1.2864	.1386	-.45±.29	-.38±.58
1959	74	9.40	13.72	11.82	.9089	.1108	-.32±.28	-.37±.55
1960	59	10.16	14.73	12.39	1.1198	.1378	.02±.31	-.40±.61
1961	70	9.14	14.22	11.47	1.0428	.1221	-.11±.29	.38±.57
1962	80	9.91	17.27	13.45	2.1133	.1625	.12±.27	.26±.53
1963	64	9.14	15.24	11.86	1.2143	.1377	.18±.30	.66±.59
1964	76	9.65	15.75	12.46	1.6851	.1489	.23±.28	-.43±.54
1965	46	9.91	14.22	11.75	.9876	.1465	.35±.35	-.20±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	9.65	15.24	12.66	1.3532	.1466	-.21±.30	.20±.59
1955	79	8.89	13.97	11.45	1.2379	.1252	.13±.27	-.49±.53
1956	66	9.65	13.46	11.31	.6238	.0972	.22±.29	-.34±.58
1957	84	9.40	12.95	10.99	.8148	.0985	.24±.26	-.77±.52
1958	62	8.13	13.72	11.12	1.1310	.1351	-.25±.30	.32±.60
1959	71	8.38	13.72	11.28	1.1972	.1299	.01±.28	-.06±.56
1960	56	9.91	14.22	12.02	1.0399	.1363	-.05±.32	-.68±.63
1961	73	7.62	13.72	10.88	1.4241	.1397	-.37±.28	.35±.56
1962	85	10.67	15.75	13.16	.9382	.1051	.01±.26	-.06±.52
1963	69	7.87	13.72	11.65	1.3611	.1404	-.30±.29	.35±.57
1964	67	10.16	13.97	11.99	.7928	.1088	.01±.29	-.75±.58
1965	77	9.14	13.46	11.32	.8294	.1038	.03±.27	-.39±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	8.13	15.24	12.59	1.8417	.1830	-.42±.32	.84±.63
1955	62	9.40	13.97	12.41	.8733	.1187	-.61±.30	.32±.60
1956	46	10.16	14.22	11.91	.8293	.1343	.13±.35	-.35±.69
1957	67	8.64	13.97	11.43	.9394	.1184	.05±.29	.25±.58
1958	57	9.65	14.73	11.74	1.2759	.1496	.12±.32	-.47±.62
1959	72	8.89	14.22	11.43	1.0677	.1218	.15±.28	-.13±.56
1960	83	10.16	13.72	11.84	.9313	.1059	.08±.26	-.92±.52
1961	71	8.64	12.95	10.78	.8049	.1065	.06±.28	-.02±.56
1962	72	10.16	15.75	13.05	1.5419	.1463	-.15±.28	-.62±.56
1963	43	9.40	14.22	11.80	1.0178	.1539	.10±.36	-.25±.71
1964	74	9.91	13.46	11.64	.7200	.0986	.14±.28	-.56±.55
1965	37	10.16	13.72	11.44	.7310	.1406	.59±.39	-.00±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	10.41	14.22	12.59	.8161	.1072	-.59±.28	-.19±.56
1955	58	9.91	13.97	11.72	.9125	.1254	.17±.31	-.37±.62
1956	78	9.91	13.46	11.44	.5997	.0877	.44±.27	-.22±.54
1957	62	9.40	12.45	11.11	.5393	.0933	-.23±.30	-.84±.60
1958	60	8.89	13.46	11.33	1.0806	.1342	.18±.31	-.58±.61
1959	75	9.14	13.46	11.02	.9294	.1113	.23±.28	-.57±.55
1960	73	9.14	13.72	11.66	.8415	.1074	-.04±.28	-.21±.56
1961	86	8.38	14.99	10.44	1.2078	.1185	.89±.26	2.11±.51
1962	82	10.67	16.26	12.78	1.1257	.1172	.17±.27	.63±.53
1963	70	9.40	13.46	11.53	.7392	.1028	-.00±.29	-.41±.57
1964	67	8.13	14.22	11.46	1.1694	.1321	-.14±.29	.53±.58
1965	70	9.65	13.46	11.38	.6861	.0990	.12±.29	.25±.57
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	40	9.14	13.46	10.59	.5399	.1162	1.35±.37	4.07±.73
1950	58	11.18	14.99	12.35	.6445	.1054	.52±.31	.34±.62
1951	88	10.41	14.48	12.39	.7519	.0924	.37±.26	-.16±.51
1952	93	9.91	13.72	11.81	.8368	.0949	.05±.25	-.83±.50
1953	104	10.16	14.99	12.57	1.1688	.1060	.23±.24	-.34±.47

TABLE I-1(10) Basic statistics for crimps per inch (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	71	5	12	7.13	1.9694	.1665	.77±.28	1.13±.56
1955	71	4	11	7.20	2.1606	.1744	.53±.28	.14±.56
1956	66	6	16	9.56	6.5270	.3145	.77±.29	-.31±.58
1957	56	6	15	8.71	2.9351	.2289	1.28±.32	2.64±.63
1958	67	5	16	8.18	2.6947	.2005	1.70±.29	7.03±.58
1959	74	4	10	6.65	1.7927	.1556	.35±.28	-.11±.55
1960	59	5	11	7.03	1.6195	.1657	.49±.31	.38±.61
1961	70	6	12	7.83	.8687	.1114	1.32±.29	4.82±.57
1962	80	5	10	7.44	1.5150	.1376	.14±.27	-.62±.53
1963	64	5	10	7.25	1.7778	.1667	.38±.30	-.40±.59
1964	76	5	10	6.83	1.2904	.1303	.39±.28	-.05±.54
1965	46	5	9	6.26	1.2193	.1628	.77±.35	.29±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	4	11	7.32	1.9299	.1750	.15±.30	.27±.59
1955	79	4	11	7.32	2.8088	.1886	.38±.27	-.80±.53
1956	66	6	16	9.23	5.2552	.2822	.78±.29	-.09±.58
1957	84	5	14	8.76	2.2559	.1639	.37±.26	.76±.52
1958	62	5	14	7.95	2.5714	.2037	1.02±.30	2.08±.60
1959	71	4	11	7.20	1.9606	.1662	.31±.28	-.10±.56
1960	56	5	10	7.11	1.4792	.1625	.28±.32	-.83±.63
1961	73	6	13	8.38	1.1842	.1274	1.21±.28	3.54±.56
1962	85	5	11	7.37	1.4249	.1295	.58±.26	.12±.52
1963	69	5	11	7.45	2.0158	.1709	.38±.29	-.16±.57
1964	67	5	10	7.28	1.5396	.1516	.12±.29	-.40±.58
1965	77	5	10	7.07	1.7457	.1506	.19±.27	-.95±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	5	11	7.06	1.5340	.1670	.60±.32	.62±.63
1955	62	4	10	6.82	1.7877	.1698	.41±.30	-.06±.60
1956	46	5	13	8.33	2.6691	.2409	.45±.35	.23±.69
1957	67	5	11	7.79	1.6223	.1556	.13±.29	-.47±.58
1958	57	5	9	7.25	1.2957	.1508	.17±.32	-.92±.62
1959	72	5	10	6.96	1.5053	.1446	.08±.28	-.66±.56
1960	83	5	11	7.84	1.5239	.1355	.10±.26	-.01±.52
1961	71	7	11	8.35	.6885	.0985	.18±.28	.25±.56
1962	72	5	10	7.57	1.0374	.1200	.05±.28	.15±.56
1963	43	5	10	7.58	1.7730	.2031	.07±.36	-.66±.71
1964	74	5	10	6.77	1.1931	.1270	.34±.28	.24±.55
1965	37	5	9	6.41	1.1366	.1753	.25±.39	-.61±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	5	10	7.38	1.3247	.1366	.01±.28	-.27±.56
1955	58	5	14	7.60	2.6295	.2129	1.21±.31	2.63±.62
1956	78	5	16	9.21	4.6587	.2444	.71±.27	.39±.54
1957	62	6	12	8.52	1.6637	.1638	.22±.30	-.39±.60
1958	60	5	13	7.58	3.2302	.2320	.98±.31	1.46±.61
1959	75	5	10	7.51	1.3885	.1361	.03±.28	-.62±.55
1960	73	5	11	7.77	.9311	.1129	.01±.28	1.12±.56
1961	86	7	11	8.44	.6495	.0869	.53±.26	.35±.51
1962	82	5	10	7.76	1.0509	.1132	.16±.27	-.14±.53
1963	70	5	11	7.79	1.5331	.1480	.41±.29	.34±.57
1964	67	5	10	7.34	1.4410	.1467	.48±.29	-.27±.58
1965	70	5	11	7.09	1.6737	.1546	.61±.29	.76±.57
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	40	6	12	8.40	1.6308	.2019	.34±.37	.09±.73
1950	58	5	13	8.36	2.5508	.2097	.41±.31	.12±.62
1951	88	5	12	8.11	1.8950	.1467	.73±.26	.79±.61
1952	93	5	13	7.99	1.9238	.1438	1.00±.25	2.26±.50
1953	104	5	14	8.28	2.8050	.1642	.97±.24	1.24±.47

TABLE I-1(11) Basic statistics for fibre diameter in microns (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	71	24.5	40.1	29.54	7.1575	.3175	.98±.28	2.20±.56
1955	71	22.0	34.6	28.23	4.7121	.2576	.04±.28	.83±.56
1956	66	17.3	27.6	21.32	4.0420	.2475	1.08±.29	1.60±.58
1957	56	20.8	31.5	26.09	5.3982	.3105	.01±.32	-.01±.63
1958	67	18.9	27.9	24.73	3.0744	.2142	-.51±.29	.54±.58
1959	74	21.2	32.0	26.75	3.8420	.2279	.09±.28	.40±.55
1960	59	20.8	29.4	25.30	2.8643	.2203	.10±.31	.27±.61
1961	70	20.2	30.2	24.57	5.3653	.2769	.15±.29	-.46±.57
1962	80	19.6	30.6	25.69	2.9922	.1934	.19±.27	1.65±.53
1963	64	23.4	31.0	27.23	2.4604	.1961	.04±.30	-.11±.59
1964	76	18.3	31.2	25.15	5.8679	.2779	-.13±.28	.15±.54
1965	46	23.5	30.1	26.94	2.2140	.2194	-.06±.35	-.09±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	24.5	35.6	29.91	5.3942	.2926	-.08±.30	-.43±.59
1955	79	22.7	33.3	27.54	3.9826	.2245	.23±.27	.27±.53
1956	66	18.8	28.0	22.51	3.6734	.2359	.79±.29	1.01±.58
1957	84	18.8	30.6	25.51	5.7166	.2609	-.34±.26	.16±.52
1958	62	17.7	29.2	24.67	5.6036	.3006	-.20±.30	.38±.60
1959	71	23.0	31.4	26.00	3.6971	.2282	.68±.28	.18±.56
1960	56	21.2	31.7	25.12	4.3051	.2773	.43±.32	.50±.63
1961	73	18.8	31.4	24.54	8.0369	.3318	.48±.28	-.46±.56
1962	85	23.5	32.8	27.05	3.4373	.2011	.30±.26	-.00±.52
1963	69	24.4	31.8	27.81	2.5860	.1936	-.05±.29	-.16±.57
1964	67	22.5	32.6	26.31	4.7863	.2673	.63±.29	.16±.58
1965	77	21.8	33.7	26.92	4.0483	.2293	.29±.27	1.00±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	24.7	34.8	27.92	4.2163	.2769	.97±.32	1.44±.63
1955	62	22.5	31.1	26.86	4.3043	.2635	-.07±.30	-.66±.60
1956	46	18.5	28.9	23.50	4.3978	.3092	.16±.35	.05±.69
1957	67	19.7	28.9	25.08	3.8224	.2389	-.25±.29	-.16±.58
1958	57	21.7	29.6	25.21	3.7180	.2554	.14±.32	-.49±.62
1959	72	20.7	31.2	25.67	4.4512	.2486	.20±.28	.23±.56
1960	83	20.9	28.3	24.42	2.4494	.1718	.07±.26	-.03±.52
1961	71	18.6	27.7	23.89	3.6052	.2253	-.37±.28	.46±.56
1962	72	20.6	29.0	24.61	3.8324	.2307	.24±.28	-.57±.56
1963	43	23.2	30.5	26.56	3.0915	.2681	.28±.36	-.63±.71
1964	74	20.9	30.3	25.38	3.8816	.2290	-.22±.28	-.02±.55
1965	37	23.0	30.7	25.91	2.8490	.2775	.38±.39	.16±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	24.6	32.6	28.50	4.4833	.2513	-.16±.28	-1.06±.56
1955	58	23.6	30.2	26.46	1.9420	.1830	.22±.31	-.04±.62
1956	78	17.7	28.5	24.14	4.8560	.2495	-.21±.27	.14±.54
1957	62	18.1	29.8	24.37	6.4233	.3219	.01±.30	.05±.60
1958	60	20.2	31.5	26.00	6.0676	.3180	.12±.31	-.48±.61
1959	75	18.6	28.0	24.06	3.4383	.2141	.07±.28	.09±.55
1960	73	20.1	29.9	24.33	5.1678	.2661	.45±.28	-.54±.56
1961	86	19.1	28.4	24.07	4.1639	.2200	.15±.26	-.56±.51
1962	82	20.4	30.7	25.80	3.9154	.2185	.02±.27	.00±.53
1963	70	24.4	32.8	27.72	4.5106	.2538	.55±.29	-.39±.57
1964	67	21.0	29.2	25.34	3.9442	.2426	-.08±.29	-.70±.58
1965	70	22.2	30.2	26.00	2.3946	.1850	-.10±.29	.25±.57
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	40	21.8	29.4	24.97	3.0555	.2764	.49±.37	.03±.73
1950	58	22.9	30.9	26.99	3.7704	.2550	-.12±.31	-.69±.62
1951	88	24.3	33.4	28.30	4.6071	.2288	.24±.26	-.68±.51
1952	93	19.8	30.0	24.16	4.9353	.2304	.41±.25	-.05±.50
1953	104	20.7	30.4	24.65	3.1518	.1741	.32±.24	.21±.47

TABLE I-1(12). Basic statistics for coefficient of variation of fibre diameter (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	71	15.4	30.4	20.22	10.9941	.3935	.70±.28	-.06±.56
1955	71	13.4	31.1	20.58	17.7011	.4993	.56±.28	-.40±.56
1956	66	14.1	34.4	22.98	18.1984	.5251	.43±.29	.02±.58
1957	56	15.6	31.1	21.91	14.1418	.5025	.47±.32	-.27±.63
1958	67	14.0	32.4	19.82	14.4553	.4645	.66±.29	.64±.58
1959	74	13.5	26.0	17.91	5.0043	.2600	.67±.28	1.45±.55
1960	59	13.2	28.8	18.54	10.2558	.4169	1.00±.31	.91±.61
1961	70	12.7	30.0	19.37	13.4691	.4387	.57±.29	.07±.57
1962	80	12.5	26.1	17.34	6.5776	.2867	.97±.27	1.34±.53
1963	64	12.5	28.4	17.75	9.0679	.3764	.83±.30	1.01±.59
1964	76	14.5	29.2	21.60	12.5330	.4061	.12±.28	-.75±.54
1965	46	14.4	23.2	18.36	3.7292	.2847	.21±.35	-.42±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	13.0	32.6	21.41	16.9448	.5186	.48±.30	.08±.59
1955	79	14.0	30.9	20.20	12.8077	.4026	.63±.27	.31±.53
1956	66	16.5	33.8	23.72	12.6085	.4371	.32±.29	.23±.58
1957	84	14.1	35.9	22.28	17.9656	.4625	.66±.26	.30±.52
1958	62	14.8	28.3	21.10	13.7495	.4709	.16±.30	-1.07±.60
1959	71	13.4	29.4	18.42	7.4286	.3235	.93±.28	2.20±.56
1960	56	13.4	30.0	20.03	14.8750	.5154	.56±.32	-.24±.63
1961	73	14.4	27.7	19.80	7.8217	.3273	.37±.28	-.34±.56
1962	85	12.1	25.5	18.35	7.4285	.2956	.33±.26	-.28±.52
1963	69	12.2	25.5	17.98	8.9531	.3602	.42±.29	-.34±.57
1964	67	15.5	28.6	20.41	10.7187	.4000	.71±.29	-.28±.58
1965	77	11.5	24.9	17.94	6.8732	.2988	.11±.27	-.00±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	12.7	33.4	20.92	20.8937	.6163	.87±.32	.40±.63
1955	62	14.6	35.3	22.47	16.5662	.5169	.71±.30	1.11±.60
1956	46	18.9	35.1	26.07	16.3386	.5960	.15±.35	-.68±.69
1957	67	16.9	34.8	23.09	12.6988	.4354	.87±.29	1.06±.58
1958	57	15.7	32.7	20.55	12.2533	.4636	.91±.32	.93±.62
1959	72	14.4	28.1	20.16	11.4274	.3984	.67±.28	-.44±.56
1960	83	13.7	31.8	20.30	11.8776	.3783	.44±.26	.17±.52
1961	71	12.2	32.2	20.00	12.7548	.4238	.60±.28	1.07±.56
1962	72	12.3	26.3	19.04	10.5901	.3835	.22±.28	-.43±.56
1963	43	15.4	28.1	19.43	7.5174	.4181	.88±.36	.76±.71
1964	74	16.4	33.4	22.04	10.5333	.3773	.52±.28	.48±.55
1965	37	14.6	25.7	19.62	6.2825	.4121	.31±.39	-.30±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	13.0	35.3	22.29	20.1914	.5333	.62±.28	.38±.56
1955	58	16.4	29.9	22.75	11.4573	.4445	.02±.31	-.83±.62
1956	78	16.9	36.3	24.53	15.9321	.4519	.57±.27	.01±.54
1957	62	17.1	35.9	25.22	18.5793	.5474	.26±.30	-.09±.60
1958	60	15.2	32.6	21.11	13.9981	.4830	.94±.31	.71±.61
1959	75	15.2	28.6	20.80	10.0514	.3661	.63±.28	-.37±.55
1960	73	14.1	27.8	20.80	9.1489	.3540	.14±.28	-.51±.56
1961	86	12.3	28.2	20.71	11.2014	.3609	.25±.26	-.24±.51
1962	82	13.4	30.0	19.56	9.6339	.3428	.77±.27	.98±.53
1963	70	14.5	33.8	21.10	14.7697	.4593	.88±.29	.74±.57
1964	67	16.9	32.2	22.25	11.4480	.4134	.73±.29	.21±.58
1965	70	13.2	31.6	19.61	9.6974	.3722	1.12±.29	2.18±.57
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	40	14.3	30.1	21.09	13.7796	.5869	.39±.37	-.63±.73
1950	58	14.0	28.3	20.58	13.5814	.4839	.23±.31	-.69±.62
1951	88	13.1	32.4	21.29	13.5100	.3918	.13±.26	-.26±.51
1952	93	13.5	37.7	24.58	19.2624	.4551	.21±.25	-.04±.50
1953	104	12.8	36.2	23.19	15.8282	.3901	.13±.24	.43±.47

TABLE I-1(13) Basic statistics for primary follicle number per sq.cm (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	71	106	278	196.03	1245.7706	4.1888	.01± .28	-.22±.56
1955	71	193	411	266.90	2441.0901	5.8636	.74±.28	-.17±.56
1956	66	182	404	295.32	2254.1280	5.8441	.30±.29	-.15±.58
1957	56	214	544	350.93	3842.6130	8.2836	.42±.32	.87±.63
1958	67	202	404	282.57	2035.2492	5.5115	.73±.29	.28±.58
1959	74	162	364	243.81	1682.5391	4.7683	.30±.28	-.16±.55
1960	59	160	342	234.00	1418.8966	4.9040	.26±.31	.11±.61
1961	70	174	372	263.89	1817.9578	5.0962	.40±.29	-.23±.57
1962	80	172	400	277.45	2907.8456	6.0289	.37±.27	-.46±.53
1963	64	154	330	231.34	1340.8323	4.5772	.13±.30	-.36±.59
1964	76	134	360	254.13	1942.1158	5.0551	.09±.28	-.04±.54
1965	46	154	336	250.87	2199.9382	6.9155	-.07±.35	-.72±.69

MALE PROGENY OF VISUAL FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	63	100	282	183.14	1471.1889	4.8324	.48±.30	-.10±.59
1955	79	166	343	250.89	1378.4099	4.1771	.36±.27	.05±.53
1956	66	193	393	276.97	2060.8914	5.5880	.57±.29	-.36±.58
1957	84	226	530	331.57	3053.7900	6.0295	.59±.26	.89±.52
1958	62	168	442	267.26	2681.8667	6.5769	.70±.30	.61±.60
1959	71	158	342	241.27	1358.0845	4.3736	.24±.28	-.03±.56
1960	56	168	308	230.43	1185.5584	4.6012	.32±.32	-.69±.63
1961	73	176	404	274.58	2481.4977	5.8304	.38±.28	-.40±.56
1962	85	180	560	278.24	2983.1821	5.9242	1.59±.26	6.70±.52
1963	69	172	328	244.87	1244.9974	4.2478	.53±.29	-.36±.57
1964	67	150	366	252.66	1982.7137	5.4399	.31±.29	-.04±.58
1965	77	146	370	239.18	1609.6507	4.5721	.29±.27	.80±.54

FEMALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	55	148	332	223.78	1940.8404	5.9404	.36±.32	-.16±.63
1955	62	174	373	256.92	2456.7639	6.2949	.77±.30	-.01±.60
1956	46	202	434	295.63	2244.1493	6.9847	.41±.35	.39±.69
1957	67	246	484	363.58	3017.3985	6.7109	.08±.29	-.65±.58
1958	57	192	420	279.40	2439.4236	6.5419	.91±.32	.80±.62
1959	72	186	418	299.67	2989.1831	6.4433	.11±.28	-.70±.56
1960	83	162	380	261.18	2075.6133	5.0007	.40±.26	-.07±.52
1961	71	200	376	294.06	1794.2825	5.0271	-.06±.28	-.69±.56
1962	72	206	540	289.10	2369.9482	5.7372	1.93±.28	8.33±.56
1963	43	176	374	256.51	1969.7320	6.7681	.16±.36	-.51±.71
1964	74	160	622	303.60	4697.3128	7.9673	1.59±.28	5.33±.55
1965	37	200	402	276.70	2617.6036	8.4111	.60±.39	-.17±.76

FEMALE PROGENY OF VISUAL FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	71	118	354	222.70	1784.3541	5.0132	.65±.28	.72±.56
1955	58	171	364	252.95	1649.4885	5.3329	.45±.31	-.03±.62
1956	78	206	468	293.67	3040.6667	6.2436	.71±.27	.13±.54
1957	62	226	564	353.13	3530.1798	7.5458	.60±.30	1.41±.60
1958	60	192	390	277.43	2247.6056	6.1205	.50±.31	-.21±.61
1959	75	190	542	279.79	3146.2241	6.4769	1.55±.28	5.07±.55
1960	73	172	414	257.18	1902.6484	5.1053	.75±.28	1.17±.56
1961	86	208	452	301.16	2432.2320	5.3181	.59±.26	.32±.51
1962	82	182	418	286.78	2713.7537	5.7528	.54±.27	-.15±.53
1963	70	152	390	262.14	1865.0807	5.1618	.38±.29	.28±.57
1964	67	154	426	298.39	3533.8471	7.2625	-.02±.29	-.51±.58
1965	70	124	364	267.77	1765.6282	5.0223	-.21±.29	1.00±.57

DAMS BORN IN 1949,1950,1951,1952,1953

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1949	40	225	509	336.68	4484.2763	10.5881	.57±.37	-.15±.73
1950	58	198	462	290.31	3061.3055	7.2651	1.28±.31	1.62±.62
1951	88	205	483	294.56	2308.4565	5.1218	.97±.26	2.39±.51
1952	93	190	484	310.58	3886.2244	6.4643	.53±.25	.18±.50
1953	104	168	386	287.19	2221.8656	4.6221	-.24±.24	-.61±.47

TABLE I-1(14) Basic statistics for secondary follicle number per sq. cm (where g_1 and g_2 are Fisher's coefficient of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	71	3348	7690	4511.72	676775.1195	97.6322	1.13±.28	1.94±.56
1955	71	3780	8538	5683.85	1037381.4757	120.8760	.54±.28	-.08±.56
1956	66	2909	7807	5327.50	964697.3000	120.8992	.06±.29	-.08±.58
1957	56	3980	9370	6256.82	901176.2584	126.8560	.24±.32	1.61±.63
1958	67	2714	7672	4752.54	831646.6160	111.4120	.44±.29	.49±.58
1959	74	2330	6020	4192.62	587176.5398	89.0776	.17±.28	-.00±.55
1960	59	2846	6020	4420.41	475577.0731	89.7810	-.13±.31	-.53±.61
1961	70	2996	7660	4709.03	673140.4919	98.0627	.81±.29	1.16±.57
1962	80	3084	7140	5073.55	680970.7316	92.2612	.29±.27	.26±.53
1963	64	2566	6516	4415.63	674064.2381	102.6268	.57±.30	.51±.59
1964	76	2908	6828	4735.90	801624.0954	102.7019	.14±.28	-.47±.54
1965	46	2868	6522	4550.30	591098.3053	113.3577	.35±.35	.04±.69
MALE PROGENY OF VISUAL FLOCK								
1954	63	2550	7564	4212.48	608585.7696	98.2858	1.13±.30	4.09±.59
1955	79	3485	8928	5355.09	830435.1074	102.5272	.65±.27	1.56±.53
1956	66	3035	8300	5061.02	795148.2613	109.7620	.98±.29	2.35±.58
1957	84	2922	8386	5795.71	1225428.8330	120.7826	-.10±.26	.02±.52
1958	62	2562	6954	4722.26	682412.1290	104.9126	.08±.30	.11±.60
1959	71	2460	5846	4148.17	442532.4282	78.9484	-.12±.28	-.26±.56
1960	56	2310	6076	4293.18	654582.8039	108.1156	-.08±.32	-.28±.63
1961	73	3140	6730	4631.67	635555.1126	93.3072	.50±.28	-.32±.56
1962	85	2822	10042	4557.91	1017516.1815	109.4111	1.98±.26	8.70±.52
1963	69	2796	5930	4261.68	544362.5439	88.8218	.38±.29	-.30±.57
1964	67	2490	5620	4207.18	435856.0583	80.6555	-.15±.29	.11±.58
1965	77	2902	6702	4417.71	555363.0226	84.9265	.16±.27	.22±.54
FEMALE PROGENY OF INDEX FLOCK								
1954	55	2872	6412	4683.78	527738.4700	97.9553	-.30±.32	-.27±.63
1955	62	3294	8847	5229.77	931849.2269	122.5962	1.02±.30	2.29±.60
1956	46	3697	7605	5378.61	982170.1546	146.1216	.55±.35	-.34±.69
1957	67	3548	8012	5643.58	1055922.4894	125.5390	-.04±.29	-.52±.58
1958	57	3136	7960	4959.47	721614.7895	112.5163	.70±.32	1.45±.62
1959	72	3014	7616	4767.06	825610.2504	107.0832	.34±.28	.22±.56
1960	83	3332	7476	4764.15	715858.9057	92.8698	.80±.26	.70±.52
1961	71	3712	7324	5145.18	632087.4946	94.3538	.55±.28	-.18±.56
1962	72	3080	9114	5166.51	1158263.4646	126.8345	.92±.28	1.78±.56
1963	43	3446	6180	4505.72	463560.1107	103.8290	.52±.36	-.39±.71
1964	74	2384	7354	5080.57	873406.1940	108.6406	.10±.28	-.15±.55
1965	37	3372	7166	5138.05	931787.6637	158.6929	.29±.39	-.57±.76
FEMALE PROGENY OF VISUAL FLOCK								
1954	71	3106	6066	4544.73	494365.8559	83.4440	.33±.28	-.69±.56
1955	58	3665	7642	5166.33	643754.2943	105.3528	.92±.31	.78±.62
1956	78	3033	7899	5039.05	908104.4649	107.8998	.54±.27	.35±.54
1957	62	3130	7656	5392.97	883177.4088	119.3516	.10±.30	-.09±.60
1958	60	2854	6926	4775.80	658057.5864	104.7264	.12±.31	-.01±.61
1959	75	3016	6754	4639.97	592060.9182	88.8490	.28±.28	.01±.55
1960	73	3422	6736	4660.55	536250.4734	85.7082	.69±.28	.18±.56
1961	86	3306	7146	4878.56	698578.6966	90.1277	.39±.26	-.26±.51
1962	82	2480	6820	4394.20	689138.9244	91.6741	.38±.27	.33±.53
1963	70	3004	6486	4501.97	538234.8398	87.6873	.55±.29	.04±.57
1964	67	3404	7092	4920.66	669450.3501	99.9590	.50±.29	-.06±.58
1965	70	3184	7696	4986.86	748524.2402	103.4079	.83±.29	1.12±.57
DAMS BORN IN 1949, 1950,1951,1952,1953								
1949	40	3327	7660	5523.15	880997.7718	148.4080	.38±.37	.33±.73
1950	58	3260	9320	5161.48	1432633.3420	157.1640	.94±.31	1.49±.62
1951	88	3072	7247	5157.42	857615.3039	98.7199	.02±.26	-.44±.51
1952	93	2518	8640	5437.85	1070370.7162	107.2817	.27±.25	.34±.50
1953	104	3158	7748	5176.93	873803.0148	91.6622	.40±.24	-.20±.47

TABLE I-1(15) Basic statistics for skin thickness in cm (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as singles and has been partitioned by sex and flock type.

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1958	67	.208	.307	.249	.0004	.0026	.63 .29	.32 .58
1959	74	.216	.325	.269	.0005	.0027	.25 .28	-.26 .55
1960	59	.229	.315	.269	.0003	.0023	-.01 .31	.33 .61
1961	70	.198	.315	.252	.0005	.0026	-.11 .29	.69 .57
1962	80	.198	.290	.237	.0003	.0021	.11 .27	-.33 .53
1963	64	.175	.284	.231	.0004	.0026	.30 .30	.37 .59
1964	76	.183	.279	.231	.0005	.0025	-.06 .28	-.33 .54
1965	46	.216	.290	.259	.0003	.0027	-.28 .35	-.55 .69
MALE PROGENY OF VISUAL FLOCK								
1958	62	.203	.328	.252	.0006	.0032	.78 .30	.94 .60
1959	71	.218	.310	.269	.0004	.0022	-.15 .28	-.02 .56
1960	56	.221	.315	.260	.0004	.0025	.12 .32	-.02 .63
1961	73	.201	.295	.245	.0003	.0021	.36 .28	.15 .56
1962	85	.170	.297	.223	.0006	.0026	.56 .26	.31 .52
1963	69	.173	.279	.232	.0005	.0026	-.14 .29	-.28 .57
1964	67	.183	.287	.230	.0004	.0024	.40 .29	.32 .58
1965	77	.208	.307	.254	.0004	.0022	.43 .27	.08 .54
FEMALE PROGENY OF INDEX FLOCK								
1958	57	.224	.318	.267	.0005	.0031	.41 .32	-.47 .62
1959	72	.211	.310	.266	.0005	.0027	-.36 .28	-.63 .56
1960	83	.198	.302	.243	.0004	.0022	.53 .26	.40 .52
1961	71	.216	.305	.252	.0004	.0023	.34 .28	-.16 .56
1962	72	.157	.279	.215	.0006	.0028	.14 .28	-.18 .56
1963	43	.193	.300	.232	.0005	.0035	.62 .36	.31 .71
1964	74	.175	.279	.218	.0005	.0025	.28 .28	.31 .55
1965	37	.221	.325	.260	.0007	.0043	.74 .39	.15 .76
FEMALE PROGENY OF VISUAL FLOCK								
1958	60	.231	.328	.260	.0004	.0025	1.04 .31	1.59 .61
1959	75	.211	.312	.266	.0004	.0023	.13 .28	-.47 .55
1960	73	.198	.312	.242	.0005	.0026	.36 .28	.36 .56
1961	86	.203	.279	.243	.0002	.0017	.19 .26	-.07 .51
1962	82	.152	.254	.205	.0003	.0020	-.15 .27	.21 .53
1963	70	.185	.282	.226	.0004	.0025	.42 .29	.00 .57
1964	67	.168	.269	.203	.0005	.0028	.64 .29	-.04 .58
1965	70	.208	.310	.244	.0004	.0023	.92 .29	1.50 .57

TABLE I-2(1) Basic statistics for date of birth (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	27	111	138	116.41	35.0199	1.1389	2.22± .45	5.15± .87
1955	30	109	128	116.23	25.9092	.9293	.40± .43	-.58± .83
1956	26	110	144	119.12	72.2662	1.6672	1.53± .46	1.57± .89
1957	18	108	134	115.44	51.3203	1.6885	1.41± .54	1.00±1.04
1958	11	118	131	123.82	18.9636	1.3130	.07± .66	-1.25±1.28
1959	29	115	139	122.62	39.8867	1.1728	1.29± .43	1.00± .85
1960	5	124	144	128.80	75.2000	3.8781	1.36± .91	.04±2.00
1961								
1962	14	113	152	123.57	93.6484	2.5863	1.78± .60	3.44±1.15
1963	16	115	141	123.94	59.2625	1.9246	1.29± .56	.69±1.09
1964	27	114	141	123.52	50.5670	1.3685	.97± .45	.42± .87
1965	12	117	192	141.17	450.8788	6.1297	.95± .64	.77±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	107	141	119.91	93.9913	2.0670	.79± .49	-.39± .95
1955	25	113	151	121.92	112.1600	2.1181	1.59± .46	1.84± .90
1956	10	113	127	121.00	25.1111	1.5846	-.07± .69	-1.25±1.33
1957	12	109	120	113.00	9.2727	.8790	.69± .64	.49±1.23
1958	11	120	155	128.36	105.6545	3.0992	1.85± .66	2.23±1.28
1959	37	113	138	121.49	26.5345	.8468	1.16± .39	2.15± .76
1960	18	118	144	127.33	65.4118	1.9063	.84± .54	-.33±1.04
1961	10	114	130	121.70	43.5667	2.0873	-.05± .69	-1.63±1.33
1962	10	109	126	120.00	36.0000	1.8974	-.64± .69	-1.02±1.33
1963	20	106	142	116.75	117.0395	2.4191	1.04± .51	.39± .99
1964	35	116	147	124.80	37.5176	1.0353	1.20± .40	2.95± .78
1965	11	117	146	129.18	158.1636	3.7919	.43± .66	-1.55±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	111	138	118.00	33.5714	1.0759	1.67± .43	3.16± .85
1955	23	108	121	115.26	16.9289	.8579	-.37± .48	-.92± .93
1956	22	110	144	120.36	57.1948	1.6124	1.45± .49	2.41± .95
1957	21	109	143	117.19	116.4619	2.3550	1.55± .50	.91± .97
1958	9	118	131	124.00	13.5000	1.2247	.19± .72	.06±1.40
1959	40	115	136	121.38	27.3173	.8264	1.26± .37	1.12± .73
1960	6	124	144	133.67	74.2667	3.5182	.03± .85	-1.79±1.74
1961								
1962	19	113	152	122.47	88.4854	2.1580	1.76± .52	3.24±1.01
1963	19	115	137	121.58	45.2573	1.5434	1.24± .52	.62±1.01
1964	24	114	145	125.21	64.2591	1.6363	.95± .47	.89± .92
1965	10	117	192	142.40	624.9333	7.9053	.71± .69	-.55±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	107	141	119.97	78.5345	1.6456	.72± .43	-.17± .85
1955	26	110	151	121.50	133.6200	2.2670	1.30± .46	.71± .89
1956	14	112	140	121.64	75.6319	2.3243	1.26± .60	.58±1.15
1957	14	110	125	116.00	40.7692	1.7065	.62± .60	-1.43±1.15
1958	18	117	148	128.28	101.3889	2.3733	.76± .54	-.58±1.04
1959	24	71	138	118.08	145.9058	2.4656	-2.22± .47	8.08± .92
1960	14	122	144	132.14	40.4396	1.6996	.53± .60	-.50±1.15
1961	17	110	139	121.71	70.7206	2.0396	.36± .55	-1.03±1.06
1962	15	109	127	119.07	25.3524	1.3001	-.16± .58	-.72±1.12
1963	22	107	145	121.41	94.2532	2.0698	.63± .49	.55± .95
1964	35	116	132	122.83	23.3227	.8163	.38± .40	-.93± .78
1965	5	122	146	131.80	108.7000	4.6626	.39± .91	-1.45±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	86	129	106.60	119.6857	2.8247	-.07± .58	-.11±1.12
1950	17	101	130	114.12	49.8603	1.7126	.94± .55	1.31±1.06
1951	17	104	116	111.12	14.1103	0.9111	-.28± .55	-1.08±1.06
1952	19	108	126	116.37	35.3567	1.3641	.28± .52	-1.05±1.01
1953	16	104	119	109.56	24.1292	1.2280	.42± .56	-1.21±1.09

TABLE I-2(2) Basic statistics for birth weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	27	2.95	4.54	3.51	.1874	.0833	.54±.45	-.44± .87
1955	30	2.27	4.54	3.42	.3332	.1054	.21±.43	-.77± .83
1956	26	2.04	4.54	3.20	.5075	.1397	.22±.46	-.99± .89
1957	18	2.04	4.31	3.13	.3301	.1354	.01±.54	-.39±1.04
1958	11	2.27	3.63	3.20	.1898	.1314	-1.02±.66	-.19±1.28
1959	29	2.04	5.67	3.56	.8952	.1757	.60±.43	-.33± .85
1960	5	2.95	4.54	3.67	.3446	.2625	.34±.91	-.80±2.00
1961								
1962	14	2.72	4.54	3.52	.3304	.1536	.03±.60	-1.01±1.15
1963	16	2.27	4.76	3.44	.3581	.1496	.20±.56	.09±1.09
1964	27	1.81	4.31	3.34	.4138	.1238	-.47±.45	-.15± .87
1965	12	3.63	4.99	4.29	.1165	.0985	-.02±.64	.47±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	2.95	4.31	3.54	.1698	.0878	.24±.49	-.90± .95
1955	25	2.50	4.99	3.87	.3235	.1138	-.53±.46	.33± .90
1956	10	1.81	4.99	3.24	.8121	.2850	.29±.69	-.24±1.33
1957	12	2.50	4.08	3.36	.2416	.1419	.05±.64	-.90±1.23
1958	11	2.27	3.63	3.01	.1450	.1148	-.32±.66	-.40±1.28
1959	37	2.27	5.22	3.60	.3794	.1013	.41±.39	.39± .76
1960	18	2.72	4.54	3.63	.3994	.1490	.29±.54	-1.34±1.04
1961	10	2.27	4.54	3.63	.5372	.2318	-.37±.69	-.86±1.33
1962	10	2.95	4.31	3.88	.1766	.1329	-1.04±.69	.31±1.33
1963	20	2.72	4.08	3.38	.2431	.1103	-.15±.51	-1.42± .99
1964	35	2.50	5.44	3.67	.4614	.1148	1.01±.40	.65± .78
1965	11	2.50	6.35	3.96	.9296	.2907	1.18±.66	1.79±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	2.04	4.54	3.21	.4803	.1287	.03±.43	-1.08± .85
1955	23	1.36	4.54	3.18	.6079	.1626	-.56±.48	.15± .93
1956	22	1.59	4.08	2.90	.4503	.1431	-.21±.49	-.87± .95
1957	21	1.81	4.54	3.01	.4277	.1427	.45±.50	-.11± .97
1958	9	2.27	3.86	3.20	.2500	.1667	-.37±.72	-.35±1.40
1959	40	2.18	4.31	3.16	.2390	.0773	.04±.37	-.37± .73
1960	6	2.95	3.86	3.36	.1526	.1595	.02±.85	-1.67±1.74
1961								
1962	19	2.27	4.31	3.31	.2476	.1141	.36±.52	.32±1.01
1963	19	2.04	4.54	3.33	.4804	.1590	-.10±.52	-.32±1.01
1964	24	2.95	4.54	3.56	.3197	.1154	.36±.47	-1.32± .92
1965	10	3.18	4.54	3.81	.1920	.1386	.20±.69	-1.09±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	2.27	4.54	3.51	.4358	.1226	-.46±.43	-.95± .85
1955	26	2.04	4.54	3.41	.3353	.1136	-.16±.46	-.04± .89
1956	14	2.27	3.63	2.95	.2137	.1235	-.28±.60	-1.19±1.15
1957	14	2.27	3.63	2.98	.2284	.1277	.11±.60	-1.41±1.15
1958	18	2.27	4.08	3.05	.2555	.1191	.50±.54	.04±1.04
1959	24	2.27	4.31	3.20	.3816	.1261	.10±.47	-1.03± .92
1960	14	2.27	4.31	3.40	.3561	.1595	-.26±.60	-.85±1.15
1961	17	2.50	4.31	3.31	.3347	.1403	.22±.55	-1.18±1.06
1962	15	2.50	4.08	3.36	.2072	.1175	-.16±.58	-.60±1.12
1963	22	2.27	4.08	3.08	.2335	.1030	.16±.49	-.57± .95
1964	35	2.50	4.54	3.32	.2237	.0800	.25±.40	.03± .78
1965	5	2.72	4.54	3.63	.5144	.3207	.00±.91	-1.30±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	2.50	4.57	3.25	.3341	.1493	.74±.58	-.34±1.12
1950	17	1.59	3.63	2.95	.2315	.1167	-1.38±.55	1.96±1.06
1951	17	2.72	4.08	3.24	.1721	.1006	.44±.55	-.56±1.06
1952	19	2.72	4.54	3.31	.3104	.1274	.76±.52	-.44±1.01
1953	16	2.04	4.08	2.91	.3101	.1392	-.03±.56	-.22±1.09

TABLE I-2(3) Basic statistics for weaner body weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	27	17.37	35.52	26.15	15.1722	.7496	.20±.45	.26± .87
1955	30	20.77	35.52	28.72	12.9553	.6571	-.52±.43	-.21± .83
1956	26	19.96	30.75	24.58	9.5572	.6063	.45±.46	-.63± .89
1957	18	9.98	21.36	17.04	7.6479	.6518	-.67±.54	.70±1.04
1958	11	19.96	30.12	24.04	9.2644	.9177	.79±.66	-.33±1.28
1959	29	13.38	27.76	21.04	13.1555	.6735	-.29±.43	.29± .85
1960	5	22.95	32.52	27.90	13.2869	1.6301	-.11±.91	-1.08±2.00
1961								
1962	14	19.14	30.35	25.35	15.7464	1.0605	-.24±.60	-1.40±1.15
1963	16	24.95	37.33	29.22	9.6852	.7780	.97±.56	.97±1.09
1964	27	18.96	36.51	28.84	15.8079	.7652	-.40±.45	-.13± .87
1965	12	21.36	40.32	28.64	21.0541	1.3246	1.13±.64	1.86±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	18.96	31.84	25.44	12.0581	.7403	-.21±.49	-.99± .95
1955	25	20.77	40.91	29.66	15.7243	.7931	.25±.46	1.63± .90
1956	10	19.78	33.34	24.61	21.0815	1.4519	1.03±.69	-.35±1.33
1957	12	16.15	25.54	18.68	6.0664	.7110	1.85±.64	3.17±1.23
1958	11	12.16	33.75	23.46	28.8881	1.6206	-.30±.66	.74±1.28
1959	37	13.97	26.76	20.17	8.5218	.4799	.09±.39	.05± .76
1960	18	19.37	37.51	26.28	24.7292	1.1721	.73±.54	-.19±1.04
1961	10	22.14	33.52	26.42	12.3132	1.1096	.69±.69	-.28±1.33
1962	10	26.94	37.33	30.55	10.1300	1.0065	.98±.69	.01±1.33
1963	20	21.77	37.92	29.39	19.5276	.9881	.10±.51	-.73± .99
1964	35	22.14	36.33	29.40	12.9347	.6079	.18±.40	-.71± .78
1965	11	27.35	33.93	30.43	4.5825	.6454	.12±.66	-1.23±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	17.96	30.35	23.49	11.5147	.6301	.11±.43	-.99± .85
1955	23	18.55	34.75	25.81	12.4213	.7349	.26±.48	.48± .93
1956	22	17.37	33.11	21.89	10.6083	.6944	1.72±.49	4.48± .95
1957	21	12.16	20.55	17.07	4.2159	.4481	-.38±.50	-.16± .97
1958	9	19.14	26.94	22.74	7.8085	.9315	.19±.72	-1.00±1.40
1959	40	13.79	22.95	19.04	5.3144	.3645	-.35±.37	-.71± .73
1960	6	19.14	24.95	21.58	7.0648	1.0851	.14±.85	-1.79±1.74
1961								
1962	19	18.37	30.53	23.64	9.8472	.7199	.21±.52	-.32±1.01
1963	19	18.78	31.12	25.59	10.9683	.7598	-.38±.52	-.49±1.01
1964	24	11.16	29.76	23.29	17.8177	.8616	-1.02±.47	1.57± .92
1965	10	18.78	27.35	23.56	6.1477	.7841	-.42±.69	-.40±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	15.88	35.11	23.62	13.6632	.6864	.93±.43	2.52± .85
1955	26	20.37	36.51	26.80	17.7144	.8254	1.14±.46	.56± .89
1956	14	16.56	23.95	21.18	5.2775	.6140	-.60±.60	-.62±1.15
1957	14	13.97	20.37	17.20	3.7217	.5156	.04±.60	-1.02±1.15
1958	18	17.78	27.35	21.27	4.0264	.4730	1.29±.54	2.97±1.04
1959	24	14.15	22.14	19.21	3.7107	.3932	-.53±.47	.10± .92
1960	14	16.96	28.35	22.64	12.8859	.9594	-.15±.60	-1.12±1.15
1961	17	19.96	30.12	24.59	10.2618	.7769	.28±.55	-.80±1.06
1962	15	19.78	31.12	24.44	8.3358	.7455	.80±.58	.25±1.12
1963	22	21.86	30.75	25.75	5.6255	.5057	.43±.49	-.31± .95
1964	35	18.96	30.53	25.64	8.9445	.5055	-.40±.40	-.48± .78
1965	5	22.95	27.35	24.74	2.8913	.7604	.64±.91	-.84±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	15.88	21.32	18.51	2.2397	.3864	.02±.58	-.49±1.12
1950	17	16.96	22.36	19.31	3.0264	.4219	.20±.55	-.89±1.06
1951	17	19.96	25.26	22.60	2.5415	.3867	.38±.55	-.79±1.06
1952	19	17.37	30.53	23.01	10.3029	.7364	.21±.52	.08±1.01
1953	16	21.36	32.34	24.09	8.8515	.7438	1.51±.56	1.61±1.09

TABLE I-2(4) Basic statistics for hogget body weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	27	45.50	77.25	60.93	54.9879	1.4271	.33±.45	-.06± .87
1955	30	37.92	63.28	53.32	43.0993	1.1986	-.53±.43	-.48± .83
1956	26	42.09	56.29	47.90	14.5529	.7481	.50±.46	-.42± .89
1957	18	38.92	52.89	45.15	11.3222	.7931	.25±.54	.07±1.04
1958	11	31.57	57.33	50.36	46.8388	2.0635	-2.01±.66	3.42±1.28
1959	29	42.32	66.04	55.03	30.0597	1.0181	-.20±.43	-.19± .85
1960	5	52.89	64.27	59.51	31.8605	2.5243	-.38±.91	-1.79±2.00
1961								
1962	14	41.91	64.27	55.84	36.5506	1.6158	-1.07±.60	.40±1.15
1963	16	41.32	69.45	59.54	51.6063	1.7959	-.96±.56	.68±1.09
1964	27	50.08	67.45	59.14	15.9913	.7696	.12±.45	.39± .87
1965	12	52.80	66.54	59.68	17.4865	1.2071	.16±.64	-1.02±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	53.71	68.45	60.60	14.3594	.8079	.10±.49	-.37± .95
1955	25	38.92	64.68	54.79	40.7875	1.2773	-1.13±.46	.84± .90
1956	10	38.92	59.47	48.47	49.8061	2.2317	.46±.69	-1.03±1.33
1957	12	40.73	53.71	47.55	12.6757	1.0278	.03±.64	-.33±1.23
1958	11	29.57	60.15	52.55	72.9800	2.5758	-1.88±.66	2.90±1.28
1959	37	44.09	65.27	54.38	23.0584	.7894	.25±.39	-.25± .76
1960	18	53.30	73.26	62.53	32.9445	1.3529	.30±.54	-.67±1.04
1961	10	47.90	64.27	57.23	27.2049	1.6494	-.38±.69	-.85±1.33
1962	10	51.71	66.27	59.12	16.4962	1.2844	-.06±.69	-.30±1.33
1963	20	47.08	68.67	59.15	42.0417	1.4499	-.21±.51	-1.05± .99
1964	35	53.07	76.84	62.28	26.5436	.8709	.45±.40	.47± .78
1965	11	55.97	71.26	61.83	18.2357	1.2876	.77±.66	.18±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	30.12	54.88	45.79	22.9257	.8891	-.93±.43	2.20± .85
1955	23	32.34	48.72	41.11	21.0303	.9562	-.44±.48	-.52± .93
1956	22	24.95	45.72	36.78	23.4846	1.0332	-.45±.49	.16± .95
1957	21	25.54	50.71	37.15	27.1084	1.1362	.15±.50	1.25± .97
1958	9	36.97	55.93	43.51	32.1434	1.8898	1.16±.72	.61±1.40
1959	40	31.93	51.30	38.95	17.8613	.6682	.86±.37	.59± .73
1960	6	33.11	39.92	35.97	6.3731	1.0306	.44±.85	-1.03±1.74
1961								
1962	19	37.33	48.90	43.03	15.8182	.9124	.22±.52	-1.20±1.01
1963	19	33.75	47.72	43.12	15.8923	.9146	-.88±.52	-.12±1.01
1964	24	36.51	49.08	41.81	10.2608	.6539	.64±.47	-.21±..92
1965	10	39.10	51.48	43.04	17.3147	1.3159	.86±.69	-.44±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	33.34	55.70	46.19	18.4377	.7974	-.77±.43	1.50± .85
1955	26	34.11	52.48	41.62	15.6800	.7766	.52±.46	.68± .89
1956	14	31.53	42.91	37.75	11.9967	.9257	-.38±.60	-.85±1.15
1957	14	29.12	44.09	37.03	18.8318	1.1598	-.57±.60	-.43±1.15
1958	18	36.15	46.54	41.21	7.8701	.6612	.02±.54	-.70±1.04
1959	24	28.76	42.73	36.68	14.9270	.7886	-.30±.47	-.60± .92
1960	14	21.73	48.31	39.59	36.8512	1.6224	-1.75±.60	3.67±1.15
1961	17	32.34	49.49	39.35	22.6000	1.1530	.54±.55	-.54±1.06
1962	15	37.92	50.71	43.12	15.3058	1.0101	.50±.58	-.90±1.12
1963	22	38.10	48.49	44.04	9.6611	.6627	-.29±.49	-1.14± .95
1964	35	32.34	52.57	43.42	20.9070	.7729	-.11±.40	.57± .78
1965	5	39.92	46.40	43.53	5.5238	1.0511	-.49±.91	-.55±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	28.13	36.92	32.99	5.0706	.5814	-.21±.58	-.12±1.12
1950	17	31.12	45.09	37.74	15.9675	.9692	-.26±.55	-.78±1.06
1951	17	31.53	47.08	37.03	14.8225	.9338	1.04±.55	.97±1.06
1952	19	30.75	45.72	39.28	16.7677	.9394	-.21±.52	-.76±1.01
1953	16	39.51	49.71	43.38	7.7151	.6944	.79±.56	.13±1.09

TABLE I-2(5) Basic statistics for lamb fleece weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	27	.86	1.68	1.30	.0513	.0436	-.37±.45	-.07± .87
1955	30	1.04	2.18	1.55	.0844	.0530	.13±.43	-.49± .83
1956	26	.68	1.68	1.21	.0639	.0496	-.33±.46	-.85± .89
1957	18	.32	.91	.65	.0183	.0319	-.29±.54	.51±1.04
1958	11	.82	1.23	.98	.0174	.0398	1.04±.66	-.13±1.28
1959	29	.73	1.41	.97	.0237	.0286	.93±.43	.92± .85
1960	5	1.18	1.36	1.26	.0045	.0301	.37±.91	-.78±2.00
1961								
1962	14	.45	1.50	1.00	.0639	.0676	-.31±.60	.33±1.15
1963	16	1.13	2.18	1.76	.0807	.0710	-.49±.56	-.46±1.09
1964	27	1.32	2.95	1.85	.1372	.0713	.86±.45	1.10± .87
1965	12	.82	2.27	1.65	.1550	.1136	-.42±.64	-.14±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	.73	1.86	1.21	.0830	.0614	.66±.49	.12± .95
1955	25	.73	2.09	1.47	.0823	.0574	-.38±.46	1.03± .90
1956	10	.77	1.54	1.09	.0654	.0809	.24±.69	-1.02±1.33
1957	12	.50	1.09	.74	.0199	.0407	.97±.64	1.72±1.23
1958	11	.64	1.23	.96	.0302	.0524	-.30±.66	-.72±1.28
1959	37	.54	1.18	.96	.0328	.0298	-.52±.39	-.92± .76
1960	18	.64	1.59	1.17	.0856	.0690	-.27±.54	-.78±1.04
1961	10	.82	1.45	1.13	.0649	.0806	.11±.69	-1.67±1.33
1962	10	.77	1.91	1.21	.1180	.1086	.67±.69	-.24±1.33
1963	20	1.04	2.50	1.75	.1166	.0764	-.24±.51	.17± .99
1964	35	1.13	2.50	1.89	.1115	.0564	-.12±.40	-.36± .78
1965	11	1.32	2.22	1.90	.0876	.0892	-.86±.66	-.47±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	.64	1.72	1.26	.0722	.0499	-.26±.43	-.44± .85
1955	23	1.04	1.86	1.52	.0538	.0484	-.42±.48	-.66± .93
1956	22	.91	1.63	1.21	.0449	.0452	.35±.49	-.82± .95
1957	21	.45	.95	.68	.0198	.0307	.34±.50	-.86± .97
1958	9	.91	1.23	1.09	.0098	.0330	-.15±.72	-.37±1.40
1959	40	.68	1.23	.94	.0143	.0189	.02±.37	.41± .73
1960	6	.82	1.32	1.02	.0385	.0801	.29±.85	-1.18±1.74
1961								
1962	19	.64	1.32	1.03	.0359	.0435	-.16±.52	-.66±1.01
1963	19	1.18	2.09	1.65	.1004	.0727	.01±.52	-1.39±1.01
1964	24	.54	2.50	1.68	.2763	.1073	-.25±.47	-.33± .92
1965	10	.73	2.13	1.48	.1560	.1249	-.06±.69	-.05±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	.73	1.72	1.23	.0562	.0440	.41±.43	-.19± .85
1955	26	.86	2.04	1.47	.0692	.0516	-.23±.46	.33± .89
1956	14	.64	1.41	1.13	.0509	.0603	-.87±.60	-.14±1.15
1957	14	.50	.91	.69	.0194	.0372	.37±.60	-1.22±1.15
1958	18	.59	1.36	1.00	.0493	.0523	-.16±.54	-.70±1.04
1959	24	.54	1.36	1.02	.0490	.0452	-.33±.47	-.42± .92
1960	14	.68	1.59	1.09	.0995	.0843	.14±.60	-1.39±1.15
1961	17	.73	1.63	1.13	.0754	.0666	.26±.55	-.86±1.06
1962	15	.86	1.41	1.11	.0293	.0442	.40±.58	-1.13±1.12
1963	22	1.27	2.31	1.73	.0720	.0572	.38±.49	-.73± .95
1964	35	1.18	2.72	1.85	.1044	.0546	.46±.40	.20± .78
1965	5	1.23	2.18	1.64	.1619	.1800	.39±.91	-1.50±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	.77	1.50	1.11	.0499	.0577	.20±.58	-.96±1.12
1950	17	.73	1.23	.97	.0182	.0328	.23±.55	-.51±1.06
1951	17	1.09	1.54	1.36	.0182	.0327	-.41±.55	-.87±1.06
1952	19	.68	1.50	1.02	.0405	.0462	.65±.52	1.44±1.01
1953	16	.73	1.41	.99	.0416	.0510	.57±.56	-.87±1.09

TABLE I-2(6) Basic statistics for greasy fleece weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	27	5.94	9.48	7.50	.7034	.1614	.26±.45	-.24± .87
1955	30	4.40	7.03	5.69	.5620	.1369	-.04±.43	-1.00± .83
1956	26	4.76	7.94	5.98	.5075	.1397	.88±.46	.76± .89
1957	18	3.27	5.58	4.53	.3733	.1440	-.34±.54	-.52±1.04
1958	11	4.85	6.53	5.75	.4505	.2024	-.09±.66	-1.53±1.28
1959	29	4.49	8.57	5.99	.5668	.1398	1.11±.43	3.37± .85
1960	5	6.26	8.98	7.29	1.0682	.4622	.90±.91	-.45±2.00
1961								
1962	14	5.22	8.39	7.07	.5933	.2059	-.61±.60	.84±1.15
1963	16	4.45	8.12	6.94	.8557	.2313	-1.18±.56	1.35±1.09
1964	27	5.72	7.58	6.63	.2460	.0955	.37±.45	-.82± .87
1965	12	6.21	8.35	6.88	.3301	.1658	1.29±.64	1.61±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	6.03	9.07	7.38	.6664	.1740	.43±.49	-.48± .95
1955	25	4.04	7.08	5.49	.4492	.1340	.25±.46	.30± .90
1956	10	4.17	6.49	5.65	.6690	.2587	-.92±.69	-.55±1.33
1957	12	4.35	6.31	5.16	.2857	.1543	.37±.64	.04±1.23
1958	11	2.95	7.62	5.97	1.4387	.3616	-1.33±.66	1.91±1.28
1959	37	3.81	7.17	5.75	.5589	.1229	-.45±.39	.06± .76
1960	18	4.72	7.98	6.54	.9149	.2255	-.23±.54	-.92±1.04
1961	10	4.85	7.58	6.10	.7120	.2668	.07±.69	-.81±1.33
1962	10	5.53	9.25	6.59	1.2188	.3491	1.41±.69	1.47±1.33
1963	20	4.63	7.26	6.40	.5345	.1635	-.72±.51	-.25± .99
1964	35	5.04	8.30	6.50	.6129	.1323	.28±.40	-.50± .78
1965	11	5.35	7.58	6.77	.5711	.2278	-.76±.66	-.88±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	4.94	7.67	6.19	.5039	.1318	.51±.43	-.49± .85
1955	23	3.72	6.62	5.22	.4839	.1451	-.19±.48	-.37± .93
1956	22	4.45	6.21	5.28	.3009	.1170	-.04±.49	-1.24± .95
1957	21	3.40	5.49	4.53	.2657	.1125	-.31±.50	-.41± .97
1958	9	5.17	6.76	6.01	.2490	.1663	.01±.72	-.77±1.40
1959	40	3.86	6.17	4.94	.2508	.0792	-.02±.37	-.29± .73
1960	6	4.99	6.31	5.65	.2673	.2111	.22±.85	-1.36±1.74
1961								
1962	19	5.04	7.39	6.00	.5087	.1636	.50±.52	-.96±1.01
1963	19	4.22	7.03	5.89	.4907	.1607	-.53±.52	.22±1.01
1964	24	3.49	6.35	5.14	.5772	.1551	-.37±.47	-.37± .92
1965	10	5.13	6.08	5.59	.1069	.1034	-.32±.69	-1.07±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	5.26	7.44	6.14	.3391	.1081	.18±.43	-.94± .85
1955	26	4.22	6.49	5.43	.2634	.1007	.11±.46	.29± .89
1956	14	3.86	6.71	4.92	.5282	.1942	.75±.60	.91±1.15
1957	14	3.81	5.63	4.58	.2406	.1311	.85±.60	.30±1.15
1958	18	4.49	7.17	5.39	.4760	.1626	1.02±.56	.51±1.04
1959	24	3.36	5.58	4.72	.3123	.1141	-.39±.47	-.21± .92
1960	14	3.54	6.80	5.25	.7968	.2386	-.03±.60	-.62±1.15
1961	17	3.45	5.99	4.45	.4823	.1684	.72±.55	-.21±1.06
1962	15	4.72	6.49	5.55	.3135	.1446	.31±.58	-.93±1.12
1963	22	4.54	6.99	5.48	.3724	.1301	.40±.49	-.12± .95
1964	35	4.13	6.03	4.89	.2031	.0762	.52±.40	.03± .78
1965	5	4.58	5.58	4.94	.1687	.1837	.65±.91	-.83±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	2.99	4.22	3.69	.1353	.0950	.08±.58	-.85±1.12
1950	17	4.35	5.90	5.12	.1695	.0998	.05±.55	-.80±1.06
1951	17	4.40	5.85	5.09	.2365	.1180	-.10±.55	-1.42±1.06
1952	19	3.90	6.94	4.74	.5051	.1630	1.52±.52	3.03±1.01
1953	16	4.40	6.08	5.26	.2583	.1271	.07±.56	-.97±1.09

TABLE I-2(7) Basic statistics for percentage yield (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	27	51.90	70.00	62.87	21.5574	.8935	-.72±.45	-.20± .87
1955	30	60.50	75.70	66.39	13.6675	.6750	.37±.43	-.19± .83
1956	26	46.40	71.50	56.54	45.8161	1.3275	.32±.46	-.75± .89
1957	18	50.30	64.60	58.76	17.3450	.9816	-.11±.54	-.88±1.04
1958	11	58.10	69.70	63.43	17.5382	1.2627	.10±.66	-1.40±1.28
1959	29	51.30	73.30	63.77	24.6685	.9223	.17±.43	.23± .85
1960	5	54.90	67.80	63.40	32.8350	2.5626	-.69±.91	-1.22±2.00
1961								
1962	14	53.50	74.50	64.49	41.4629	1.7209	.15±.60	-.99±1.15
1963	16	55.70	67.10	61.35	10.3947	.8060	-.12±.56	-.92±1.09
1964	27	53.80	71.70	64.98	25.1231	.9646	-.75±.45	-.08± .87
1965	12	57.50	68.30	63.47	11.8042	.9918	-.08±.64	-1.12±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	52.40	76.60	64.90	31.4995	1.1966	-.11±.49	-.17± .95
1955	25	54.90	72.80	64.74	15.2133	.7801	-.34±.46	.47± .90
1956	10	50.20	65.00	54.85	23.8472	1.5443	.92±.69	-.31±1.33
1957	12	53.60	65.10	59.20	13.1055	1.0450	.03±.64	-1.18±1.23
1958	11	43.80	67.20	58.36	39.5745	1.8968	-.86±.66	.79±1.28
1959	37	44.90	72.40	62.17	29.3134	.8901	-.70±.39	1.33± .76
1960	18	55.20	69.80	62.69	17.0182	.9723	.13±.54	-.60±1.04
1961	10	47.80	67.20	54.97	29.4423	1.7159	1.18±.69	.78±1.33
1962	10	58.70	67.40	63.12	10.5729	1.0282	-.04±.69	-1.35±1.33
1963	20	53.60	70.20	61.16	27.8373	1.1798	.35±.51	-1.09± .99
1964	35	53.30	70.80	62.08	19.3989	.7445	-.34±.40	-.59± .78
1965	11	43.70	63.30	56.43	37.1862	1.8386	-.91±.66	-.30±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	58.00	73.30	65.97	14.8486	.7156	-.02±.43	-.48± .85
1955	23	59.90	74.40	67.94	12.2571	.7300	-.24±.48	-.19± .93
1956	22	55.70	76.70	62.56	21.8224	.9960	1.32±.49	1.95± .95
1957	21	58.50	70.10	63.30	10.3140	.7008	.36±.50	-.74± .97
1958	9	49.90	73.40	63.87	46.4900	2.2728	-.75±.72	.13±1.40
1959	40	56.00	75.70	64.64	25.4665	.7979	.48±.37	-.57± .73
1960	6	51.80	69.80	60.13	33.9187	2.3776	.34±.85	-.17±1.74
1961								
1962	19	58.00	79.70	68.24	20.3292	1.0344	.23±.52	1.35±1.01
1963	19	56.70	70.70	64.28	16.7584	.9392	-.38±.52	-.78±1.01
1964	24	57.30	72.40	65.45	11.7009	.6982	-.23±.47	.27± .92
1965	10	44.10	71.10	65.27	63.8134	2.5261	-2.04±.69	3.19±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	55.50	86.10	67.37	29.6894	1.0118	1.12±.43	3.44± .85
1955	26	57.30	74.90	67.42	22.3808	.9278	-.55±.46	.03± .89
1956	14	58.40	70.70	63.27	12.5068	.9452	.95±.60	.31±1.15
1957	14	49.20	67.90	58.29	35.9438	1.6023	.05±.60	-1.10±1.15
1958	18	57.20	75.20	66.84	29.9285	1.2895	-.22±.54	-.89±1.04
1959	24	47.10	69.00	62.69	22.5182	.9686	-1.48±.47	3.04± .92
1960	14	59.20	74.20	65.73	17.3960	1.1147	.48±.60	-.05±1.15
1961	17	56.20	71.20	63.89	16.0061	.9703	-.22±.55	-.69±1.06
1962	15	55.30	71.40	64.20	18.8700	1.1216	-.47±.58	-.41±1.12
1963	22	54.50	72.80	63.56	21.1226	.9799	.39±.49	-.28± .95
1964	35	55.50	71.40	63.15	15.8302	.6725	.05±.40	-.75± .78
1965	5	61.60	67.40	64.28	4.9570	.9957	.21±.91	-1.07±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	56.00	70.50	61.58	13.4160	.9457	.57±.58	.61±1.12
1950	17	59.10	69.90	62.80	9.3037	.7398	.66±.55	-.32±1.06
1951	17	58.60	72.70	66.72	16.9819	.9995	-.50±.55	-.62±1.06
1952	19	57.70	72.90	66.14	20.3604	1.0352	-.23±.52	-.90±1.01
1953	16	62.70	76.30	68.83	10.7103	.8182	.21±.56	.28±1.09

TABLE I-2(8) Basic statistics for clean fleece weight in kgs (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	27	3.49	5.85	4.72	.3882	.1199	-.25±.45	-.91± .87
1955	30	2.99	4.40	3.77	.2117	.0840	-.46±.43	-1.20± .83
1956	26	2.54	4.49	3.37	.2165	.0913	.40±.46	-.11± .89
1957	18	1.63	3.58	2.67	.2223	.1111	-.39±.54	-.03±1.04
1958	11	3.04	4.45	3.65	.2339	.1458	.19±.66	-1.35±1.28
1959	29	2.77	4.99	3.81	.2613	.0949	.02±.43	.03± .85
1960	5	3.77	6.03	4.65	.8699	.4171	.50±.91	-1.07±2.00
1961								
1962	14	3.49	5.35	4.56	.3073	.1481	-.25±.60	-.64±1.15
1963	16	2.77	5.08	4.26	.4418	.1662	-.74±.56	-.17±1.09
1964	27	3.63	5.31	4.27	.1715	.0797	.55±.45	-.30± .87
1965	12	3.72	5.04	4.36	.1569	.1143	-.02±.64	-.88±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	3.77	6.08	4.78	.3052	.1178	.03±.49	.09± .95
1955	25	2.59	5.17	3.57	.3095	.1113	.74±.46	.91± .90
1956	10	2.13	3.95	3.10	.2900	.1703	-.25±.69	-.46±1.33
1957	12	2.54	3.81	3.05	.1086	.0951	.69±.64	.55±1.23
1958	11	1.63	4.90	3.51	.8338	.2753	-.35±.66	-.05±1.28
1959	37	2.54	4.49	3.56	.2296	.0788	-.43±.39	-.56± .76
1960	18	2.77	5.35	4.10	.4510	.1583	-.07±.54	-.46±1.04
1961	10	2.59	4.31	3.35	.3084	.1756	.29±.69	-.72±1.33
1962	10	3.58	6.03	4.15	.5026	.2242	2.07±.69	3.30±1.33
1963	20	3.04	4.67	3.90	.2458	.1109	-.39±.51	-1.01± .99
1964	35	3.27	4.94	4.02	.2296	.0810	.11±.40	-1.17± .78
1965	11	3.31	4.45	3.79	.1586	.1201	.36±.66	-1.04±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	3.18	5.40	4.09	.2957	.1010	.52±.43	-.14± .85
1955	23	2.54	4.63	3.55	.2464	.1035	-.10±.48	-.31± .93
1956	22	2.54	3.81	3.30	.1235	.0749	-.51±.49	-.48± .95
1957	21	2.18	3.45	2.87	.1299	.0786	-.01±.50	-.89± .97
1958	9	3.36	4.54	3.83	.1548	.1312	.37±.72	-.83±1.40
1959	40	2.36	4.08	3.19	.1300	.0570	.44±.37	.18± .73
1960	6	2.77	3.77	3.38	.1224	.1428	-.80±.85	-.33±1.74
1961								
1962	19	3.40	5.17	4.09	.2508	.1149	.70±.52	-.35±1.01
1963	19	2.40	4.85	3.79	.2986	.1254	-.37±.52	.85±1.01
1964	24	2.22	4.22	3.36	.2633	.1047	-.40±.47	-.36± .92
1965	10	2.54	4.22	3.64	.2250	.1500	-1.13±.69	1.02±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	3.13	4.76	4.13	.1428	.0702	-.38±.43	-.03± .85
1955	26	2.81	4.85	3.67	.2161	.0912	.46±.46	.28± .89
1956	14	2.36	3.90	3.10	.2015	.1200	-.00±.60	-.84±1.15
1957	14	1.95	3.40	2.69	.2290	.1279	.06±.60	-1.30±1.15
1958	18	2.86	5.04	3.59	.2405	.1156	1.21±.54	2.24±1.04
1959	24	2.00	3.86	2.96	.2027	.0919	-.40±.47	-.09± .92
1960	14	2.09	4.58	3.46	.3960	.1682	-.21±.60	.14±1.15
1961	17	2.22	3.63	2.84	.1793	.1027	.39±.55	-.87±1.06
1962	15	2.86	4.35	3.56	.1816	.1100	.14±.58	-.82±1.12
1963	22	2.68	4.35	3.48	.2237	.1008	.13±.49	-.87± .95
1964	35	2.59	3.72	3.08	.0950	.0521	.41±.40	-.82± .78
1965	5	2.86	3.77	3.17	.1269	.1593	1.08±.91	-.29±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	1.68	2.72	2.28	.0803	.0732	-.32±.58	-.51±1.12
1950	17	2.77	3.86	3.22	.1125	.0814	.29±.55	-1.01±1.06
1951	17	2.77	3.95	3.39	.0848	.0706	-.04±.55	-.24±1.06
1952	19	2.31	3.99	3.13	.1888	.0997	-.35±.52	-.24±1.01
1953	16	3.04	4.17	3.62	.1121	.0837	.10±.56	-.84±1.09

TABLE I-2(9) Basic statistics for staple length in cm (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 + S.E.
MALE PROGENY OF INDEX FLOCK								
1954	27	10.41	15.49	12.75	1.2310	.2135	-.01±.45	.55± .87
1955	30	9.40	14.73	12.08	1.8317	.2471	.23±.43	-.75± .83
1956	26	9.91	14.22	12.10	1.5584	.2448	-.08±.46	-.73± .89
1957	18	9.40	12.95	11.22	1.0721	.2441	.08±.54	-.94±1.04
1958	11	10.16	12.95	11.38	.6170	.2368	.33±.66	-.24±1.28
1959	29	9.91	13.97	11.82	1.1042	.1951	.04±.43	-.65± .85
1960	5	10.92	14.99	12.85	2.0839	.6456	.23±.91	-.51±2.00
1961								
1962	14	10.92	15.24	13.32	1.7639	.3550	-.06±.60	-1.01±1.15
1963	16	10.41	14.48	12.16	1.1172	.2642	.27±.56	-.17±1.09
1964	27	10.41	15.24	12.86	1.0578	.1979	.03±.45	.53± .87
1965	12	10.41	13.72	12.02	.9892	.2871	-.06±.64	-.96±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	9.91	15.75	12.84	2.5851	.3428	.04±.49	-.74± .95
1955	25	10.16	13.46	11.51	.7404	.1721	.64±.46	-.37± .90
1956	10	9.91	12.70	11.63	.7427	.2725	-.59±.69	-.42±1.33
1957	12	9.40	13.21	11.26	1.0596	.2972	.33±.64	-.16±1.23
1958	11	8.38	14.22	10.99	2.0270	.4293	.56±.66	1.21±1.28
1959	37	8.13	13.46	11.40	1.4844	.2003	-.30±.39	.21± .76
1960	18	9.91	13.72	11.85	1.3662	.2755	-.52±.54	-.98±1.04
1961	10	9.14	11.94	10.36	.8717	.2952	.46±.69	-.94±1.33
1962	10	8.64	13.21	12.07	1.9247	.4387	-1.60±.69	1.79±1.33
1963	20	9.40	14.48	11.70	1.5991	.2828	.23±.51	-.23± .99
1964	35	9.14	14.99	12.11	1.8518	.2300	.30±.40	-.14± .78
1965	11	9.40	12.70	11.11	.7367	.2588	-.13±.66	.20±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	10.41	14.99	12.38	1.0963	.1944	-.09±.43	-.07± .85
1955	23	8.89	14.22	12.05	1.5591	.2604	-.57±.48	.15± .93
1956	22	9.91	14.48	11.86	1.0285	.2162	.27±.49	.71± .95
1957	21	9.14	12.95	11.37	1.0187	.2203	-.36±.50	-.15± .97
1958	9	9.65	13.21	11.18	1.3871	.3926	.51±.72	-.98±1.40
1959	40	9.91	14.73	11.68	.7345	.1355	1.25±.37	2.75± .73
1960	6	11.68	13.72	12.49	.5527	.3035	.66±.85	-.77±1.74
1961								
1962	19	11.94	17.78	13.61	1.7227	.3011	1.66±.52	3.47±1.01
1963	19	10.16	13.72	12.05	1.1636	.2475	-.16±.52	-1.01±1.01
1964	24	9.40	14.73	11.63	1.5034	.2503	.64±.47	.26± .92
1965	10	10.41	14.22	12.01	1.2337	.3512	.39±.69	-.21±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	11.18	14.73	12.71	.5921	.1429	.24±.43	.54± .85
1955	26	9.65	14.22	11.92	1.0422	.2002	.08±.46	-.07± .89
1956	14	10.67	12.95	11.72	.4750	.1842	.16±.60	-.94±1.15
1957	14	9.91	12.95	11.52	.7008	.2237	-.35±.60	-.64±1.15
1958	18	9.65	14.48	11.47	1.5958	.2978	.63±.54	-.13±1.04
1959	24	9.40	12.70	11.06	.7348	.1750	-.18±.47	-.62± .92
1960	14	9.65	13.21	11.76	1.0365	.2721	-.57±.60	-.60±1.15
1961	17	9.14	12.95	10.50	1.1044	.2549	.93±.55	-.02±1.06
1962	15	10.67	15.75	12.77	2.0504	.3697	.37±.58	-.43±1.12
1963	22	9.91	14.22	11.77	.9117	.2036	.31±.49	.65± .95
1964	35	9.40	13.46	11.18	.9981	.1689	.34±.40	-.53± .78
1965	5	9.40	12.70	11.13	1.7548	.5924	-.03±.91	-1.33±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	9.40	12.70	10.82	.8922	.2439	.21±.58	-.68±1.12
1950	17	11.43	14.22	12.90	.5745	.1838	-.14±.55	-.67±1.06
1951	17	11.43	13.97	12.40	.7600	.2114	.68±.55	-.85±1.06
1952	19	7.37	14.48	11.24	1.9128	.3178	-.28±.52	2.82±1.01
1953	16	11.43	14.99	12.72	.9589	.2448	.66±.56	-.10±1.09

TABLE I-2(10) Basic statistics for crimps per inch (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK								
Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	27	5	11	7.37	2.3960	.2979	.44±.45	-.24± .87
1955	30	4	8	6.33	1.0575	.1877	-.12±.43	-.50± .83
1956	26	6	16	9.46	6.0985	.4843	.97±.46	.48± .89
1957	18	6	14	8.50	4.2647	.4868	.99±.54	.87±1.04
1958	11	5	12	8.64	4.8545	.6643	.01±.66	-1.14±1.28
1959	29	5	9	6.59	1.6084	.2355	.28±.43	-1.04± .85
1960	5	4	9	7.00	4.5000	.9487	-.35±.91	-1.24±2.00
1961								
1962	14	5	9	7.14	1.2088	.2938	-.29±.60	-.69±1.15
1963	16	5	9	6.75	1.5333	.3096	.05±.56	-1.04±1.09
1964	27	5	8	6.41	.6353	.1534	.08±.45	-.41± .87
1965	12	5	7	5.83	.6970	.2410	.31±.64	-1.37±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	5	10	7.55	1.8788	.2922	-.15±.49	-.41± .95
1955	25	5	11	7.16	2.3900	.3092	.69±.46	-.00± .90
1956	10	7	18	10.20	10.6222	1.0306	1.42±.69	1.37±1.33
1957	12	7	12	8.25	1.6591	.3718	2.18±.64	4.31±1.23
1958	11	6	11	8.36	2.4545	.4724	-.14±.66	-.77±1.28
1959	37	5	9	6.89	1.9324	.2285	.26±.39	-1.23± .76
1960	18	5	10	7.11	2.3399	.3605	.22±.54	-1.00±1.04
1961	10	7	10	8.80	.8444	.2906	-.51±.69	-.32±1.33
1962	10	6	12	8.40	2.9333	.5416	.74±.69	.10±1.33
1963	20	5	11	7.20	1.8526	.3044	.79±.51	1.43± .99
1964	35	5	10	6.89	1.5160	.2081	.22±.40	-.39± .78
1965	11	5	8	6.82	.9636	.2960	-.30±.66	-.85±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	5	8	6.62	.8867	.1749	-.49±.43	-.65± .85
1955	23	5	9	6.74	1.2016	.2286	-.10±.48	-.56± .93
1956	22	6	15	8.14	4.6948	.4620	1.72±.49	2.79± .95
1957	21	6	9	7.71	1.1143	.2304	-.19±.50	-1.14± .97
1958	9	7	9	7.89	.6111	.2606	.18±.72	-1.15±1.40
1959	40	4	9	6.80	1.3949	.1867	-.08±.37	-.37± .73
1960	6	7	8	7.67	.2667	.2108	-.71±.85	-1.50±1.74
1961								
1962	19	5	8	6.90	.6550	.1857	-.46±.52	-.07±1.01
1963	19	6	9	7.11	.7661	.2008	.31±.52	-.65±1.01
1964	24	5	10	6.79	1.3025	.2330	.95±.47	1.22± .92
1965	10	5	8	5.90	.7667	.2769	1.24±.69	1.50±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	5	11	7.48	2.1158	.2701	.46±.43	-.10± .85
1955	26	6	9	7.27	.8446	.1802	.08±.46	-.89± .89
1956	14	5	13	8.07	3.1484	.4742	1.26±.60	2.74±1.15
1957	14	7	10	8.29	.8352	.2442	.04±.60	-.84±1.15
1958	18	6	12	7.67	2.1176	.3430	1.55±.54	2.65±1.04
1959	24	5	9	7.21	1.7373	.2691	-.16±.47	-1.00± .92
1960	14	6	9	7.79	.9505	.2606	-.08±.60	-1.09±1.15
1961	17	7	10	8.41	.8824	.2278	.03±.55	-.82±1.06
1962	15	6	9	7.47	1.4095	.3065	.08±.58	-1.42±1.12
1963	22	5	11	7.41	2.2532	.3200	.83±.49	.11± .95
1964	35	5	10	7.11	1.3983	.1999	.21±.40	-.17± .78
1965	5	5	8	6.60	1.8000	.6000	.11±.91	-1.60±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	6	10	7.67	2.2381	.3803	.20±.58	-1.35±1.12
1950	17	6	10	7.53	1.8897	.3334	.46±.55	-.92±1.06
1951	17	5	9	7.41	1.2574	.2720	-.32±.55	-.41±1.06
1952	19	5	13	8.32	4.1170	.4655	.25±.52	-.02±1.01
1953	16	6	9	7.81	.8292	.2276	-.17±.56	-.83±1.09

TABLE I-2(11) Basic statistics for fibre diameter (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	27	24.20	35.50	28.89	8.6736	.5668	.33±.45	-.74± .87
1955	30	24.00	32.90	27.74	3.8479	.3581	.41±.43	.51± .83
1956	26	19.40	26.00	22.09	2.7487	.3251	.79±.46	.07± .89
1957	18	23.70	30.40	26.98	4.0312	.4732	.19±.54	-1.11±1.04
1958	11	23.40	27.80	25.59	1.7109	.3944	-.11±.66	-.83±1.28
1959	29	24.80	30.20	27.08	2.1112	.2698	.29±.43	-.67± .85
1960	5	25.20	29.40	26.88	3.0870	.7857	.49±.91	-1.26±2.00
1961								
1962	14	23.30	27.80	25.29	1.4398	.3207	.53±.60	-.34±1.15
1963	16	26.40	31.60	28.29	1.5305	.3093	1.17±.56	1.38±1.09
1964	27	21.90	30.70	26.06	3.1570	.3419	.26±.45	.84± .87
1965	12	25.90	29.60	27.42	1.1233	.3060	.58±.64	-.30±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	23.90	33.80	29.48	5.3466	.4930	-.42±.49	.20± .95
1955	25	24.00	33.00	28.34	6.2566	.5003	.15±.46	-.71± .90
1956	10	19.80	24.60	22.32	2.4907	.4991	-.11±.69	-1.09±1.33
1957	12	22.90	29.50	25.54	3.1754	.5144	.77±.64	.19±1.23
1958	11	20.20	27.40	24.90	4.6740	.6519	-.77±.66	-.10±1.28
1959	37	23.10	33.10	26.33	3.8798	.3238	1.23±.39	2.13± .76
1960	18	21.60	28.70	24.74	4.4461	.4970	.53±.54	-.83±1.04
1961	10	21.30	27.90	25.13	4.1268	.6424	-.41±.69	-.66±1.33
1962	10	22.70	27.10	25.33	1.9268	.4390	-.60±.69	-.59±1.33
1963	20	24.00	34.10	28.00	5.5194	.5253	.65±.51	.52± .99
1964	35	22.50	30.90	26.22	5.0056	.3782	.35±.40	-.63± .78
1965	11	23.80	31.10	27.94	5.1725	.6857	-.23±.66	-.98±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	24.30	34.30	28.72	6.8879	.4874	.34±.43	-.36± .85
1955	23	22.70	31.70	26.84	4.1515	.4249	.19±.48	.16± .93
1956	22	19.90	29.60	23.32	4.2990	.4421	1.28±.49	2.47± .95
1957	21	20.10	29.80	24.68	6.4859	.5557	.12±.50	-.47± .97
1958	9	23.30	28.10	25.81	2.4036	.5168	-.23±.72	-.88±1.40
1959	40	21.40	28.50	25.18	3.3934	.2913	-.06±.37	-.92± .73
1960	6	21.20	26.10	24.43	3.1427	.7237	-1.00±.85	-.01±1.74
1961								
1962	19	22.60	29.00	25.33	3.5720	.4336	.46±.52	-.41±1.01
1963	19	24.80	30.10	27.10	2.4461	.3588	.35±.52	-.73±1.01
1964	24	22.90	29.30	25.56	3.3825	.3754	.25±.47	-.83± .92
1965	10	25.10	30.10	26.65	2.0739	.4554	1.43±.69	1.34±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	24.70	33.80	28.35	4.2247	.3817	.33±.43	.09± .85
1955	26	24.50	30.30	27.36	2.5673	.3142	.12±.46	-1.04± .89
1956	14	20.40	29.80	25.79	6.0413	.6569	-.67±.60	.09±1.15
1957	14	21.40	30.10	25.04	6.6857	.6911	.61±.60	-.55±1.15
1958	18	20.80	29.10	25.79	4.9069	.5221	-.44±.54	-.32±1.04
1959	24	20.40	27.30	24.26	3.3990	.3763	-.33±.47	-.51± .92
1960	14	19.40	25.90	23.24	5.8949	.6489	-.52±.60	-1.12±1.15
1961	17	20.60	29.70	24.34	4.6374	.5223	.42±.55	.62±1.06
1962	15	23.50	28.70	25.56	2.7511	.4283	.65±.58	-.73±1.12
1963	22	23.90	30.20	27.55	2.1969	.3160	-.23±.49	.06± .95
1964	35	20.70	29.40	25.21	4.1455	.3442	.08±.40	-.44± .78
1965	5	24.80	27.90	26.56	1.4330	.5354	-.48±.91	-1.01±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	21.60	27.90	24.92	3.1603	.4590	.19±.58	-.64±1.12
1950	17	24.60	30.60	27.59	2.5386	.3864	-.30±.55	-.36±1.06
1951	17	20.20	33.00	27.81	8.6761	.7144	-.53±.55	1.09±1.06
1952	19	20.60	27.10	23.38	2.6018	.3700	.46±.52	-.08±1.01
1953	16	21.40	28.60	25.18	3.4043	.4613	-.07±.56	-.16±1.09

TABLE I-2(12) Basic statistics for coefficient of variation (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	27	13.50	26.20	19.45	13.0687	.6957	-.14±.45	-.87± .87
1955	30	12.40	28.40	20.35	15.0564	.7084	.16±.43	-.46± .83
1956	26	14.20	31.10	23.15	19.9154	.8752	-.01±.46	-.82± .89
1957	18	14.60	25.50	20.61	9.1076	.7113	-.39±.54	-.37±1.04
1958	11	14.40	25.30	20.38	10.7356	.9879	-.18±.66	-.80±1.28
1959	29	14.50	20.30	17.79	2.7884	.3101	-.42±.43	-.73± .85
1960	5	20.40	32.50	25.04	20.6930	2.0344	.90±.91	-.38±2.00
1961								
1962	14	15.20	24.00	18.97	9.8914	.8406	.29±.60	-1.30±1.15
1963	16	14.50	22.50	19.06	9.1358	.7556	-.30±.56	-1.51±1.09
1964	27	17.30	27.70	21.90	8.6858	.5672	.13±.45	-.65± .87
1965	12	15.70	20.50	18.17	3.3570	.5289	-.11±.64	-1.59±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	16.10	29.20	21.42	13.3444	.7788	.40±.49	-.77± .95
1955	25	14.90	29.10	21.26	14.3509	.7577	.61±.46	-.43± .90
1956	10	16.50	30.00	22.45	18.3228	1.3536	.44±.69	-.74±1.33
1957	12	19.10	32.50	24.19	20.0536	1.2927	.63±.64	-.93±1.23
1958	11	15.80	24.80	21.12	8.6096	.8847	-.31±.66	-1.01±1.28
1959	37	13.20	29.40	18.94	11.4419	.5561	.88±.39	.90± .76
1960	18	15.50	25.20	20.02	5.7948	.5674	.48±.54	-.20±1.04
1961	10	15.40	27.90	20.59	20.0188	1.4149	.23±.69	-1.43±1.33
1962	10	15.40	20.20	17.84	1.8693	.4324	-.02±.69	-.37±1.33
1963	20	14.80	32.30	19.55	12.6753	.7961	2.32±.51	6.46± .99
1964	35	16.10	27.40	22.17	8.2449	.4854	-.01±.40	-.46± .78
1965	11	16.10	26.40	18.96	8.4305	.8755	1.61±.66	1.98±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	15.20	28.60	22.08	11.4890	.6294	.12±.43	-.79± .85
1955	23	17.40	32.20	24.22	16.6333	.8504	.45±.48	-.59± .93
1956	22	19.50	40.80	25.43	22.3192	1.0072	1.47±.49	3.13± .95
1957	21	17.30	28.10	22.99	7.5883	.6011	-.36±.50	-.29± .97
1958	9	18.20	24.90	21.00	4.7400	.7257	.36±.72	-.80±1.40
1959	40	14.00	28.40	21.32	12.3615	.5559	.04±.37	-.67± .73
1960	6	18.30	27.70	22.22	18.0777	1.7358	.25±.85	-1.69±1.74
1961								
1962	19	13.20	27.60	20.95	19.6493	1.0169	-.03±.52	-.98±1.01
1963	19	15.40	28.10	20.64	10.0081	.7258	.44±.52	.21±1.01
1964	24	17.10	29.60	23.10	12.3674	.7178	.29±.47	-.78± .92
1965	10	16.60	22.30	20.17	3.3068	.5750	-.74±.69	-.55±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	15.60	30.50	21.89	15.4414	.7297	.37±.43	-.64± .85
1955	26	17.00	38.50	21.92	22.8392	.9372	1.87±.46	3.81± .89
1956	14	16.00	29.60	23.32	14.2018	1.0072	-.15±.60	-.51±1.15
1957	14	19.90	32.60	25.04	16.9826	1.1014	.23±.60	-1.14±1.15
1958	18	15.40	30.30	20.31	13.0034	.8499	1.09±.54	1.25±1.04
1959	24	14.50	32.40	22.16	22.3842	.9658	.51±.47	-.17± .92
1960	14	17.30	30.10	22.90	12.7169	.9531	.36±.60	-.43±1.15
1961	17	13.50	24.20	19.60	11.9362	.8379	-.20±.55	-1.22±1.06
1962	15	15.30	30.40	21.57	21.6007	1.2000	.48±.58	-.88±1.12
1963	22	16.90	28.20	22.63	10.2668	.6831	.22±.49	-1.01± .95
1964	35	17.30	30.50	23.38	11.6477	.5769	.28±.40	-.65± .78
1965	5	15.80	21.90	19.62	5.7170	1.0693	-.85±.91	-.67±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	15.30	33.30	22.97	24.0978	1.2675	.55±.58	-.35±1.12
1950	17	16.70	35.00	22.41	20.4474	1.0967	1.52±.55	1.77±1.06
1951	17	17.00	27.40	22.04	7.1251	.6474	.09±.55	-.58±1.06
1952	19	15.90	34.30	24.09	25.3165	1.1543	.37±.52	-.16±1.01
1953	16	19.40	34.80	23.49	14.9425	.9664	1.82±.56	2.87±1.09

TABLE I-2(13) Basic statistics for primary follicle number per sq. cm. (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

Year	N.	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
MALE PROGENY OF INDEX FLOCK								
1954	27	110	302	213.48	2102.0285	8.8234	-.05±.45	-.33± .87
1955	30	186	351	254.17	1588.7644	7.2773	.59±.43	-.16± .83
1956	26	238	414	317.50	2483.0600	9.7725	.15±.46	-.93± .89
1957	18	260	488	355.56	3790.3791	14.5113	.22±.54	-.45±1.04
1958	11	206	360	283.82	2457.9636	14.9483	-.12±.66	-1.23±1.28
1959	29	146	398	228.14	3005.4089	10.1801	1.06±.43	1.58± .85
1960	5	176	222	200.00	298.0000	7.7201	-.17±.91	-.98±2.00
1961								
1962	14	192	392	280.14	2185.5165	12.4943	.54±.60	.96±1.15
1963	16	144	354	215.13	2800.5167	13.2300	1.03±.56	1.04±1.09
1964	27	200	384	264.74	2197.2764	9.0211	.68±.45	-.02± .87
1965	12	156	338	231.83	3363.2424	16.7413	.50±.64	-.91±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	148	306	209.00	1332.0952	7.7814	1.06±.49	.93± .95
1955	25	167	434	263.28	3494.9600	11.8236	1.02±.46	1.32± .90
1956	10	206	411	286.80	4433.9556	21.0570	.63±.69	-.76±1.33
1957	12	252	416	351.67	2064.6061	13.1168	-.71±.64	.01±1.23
1958	11	216	382	290.00	3364.8000	17.4897	.34±.66	-1.17±1.28
1959	37	180	320	243.14	1396.1201	6.1427	.30±.39	-.73± .76
1960	18	166	330	224.78	1987.3595	10.5076	.73±.54	-.09±1.04
1961	10	242	342	290.20	1275.0667	11.2919	-.06±.69	-1.40±1.33
1962	10	222	322	281.80	1245.7333	11.1612	-.62±.69	-1.05±1.33
1963	20	194	326	267.60	1958.9895	9.8969	-.48±.51	-1.15± .99
1964	35	186	338	259.77	1304.6521	6.1054	.23±.40	-.56± .78
1965	11	146	314	245.46	2333.6727	14.5654	-.37±.66	-.14±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	154	348	227.52	1965.3300	8.2323	.73±.43	.33± .85
1955	23	195	376	285.13	2409.1186	10.2345	.35±.48	-.98± .93
1956	22	220	478	340.59	4717.3961	14.6433	.17±.49	-.45± .95
1957	21	282	526	388.38	4374.6476	14.4332	.53±.50	-.37± .97
1958	9	222	302	268.89	792.1111	9.3815	-.44±.72	-1.12±1.40
1959	40	180	470	300.25	2954.9103	8.5949	.43±.37	1.03± .73
1960	6	264	372	303.67	1463.0667	15.6155	.98±.85	-.24±1.74
1961								
1962	19	232	416	303.16	1357.4737	8.4526	1.14±.52	3.24±1.01
1963	19	176	366	263.90	3548.2105	13.6656	.24±.52	-.94±1.01
1964	24	244	446	320.79	2841.4764	10.8809	.58±.47	-.49± .92
1965	10	226	330	284.40	1344.7111	11.5962	-.36±.69	-1.32±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	166	330	229.17	1756.1478	7.7818	.35±.43	-.49± .85
1955	26	192	336	270.08	1882.2338	8.5084	-.47±.46	-.92± .89
1956	14	216	409	315.57	3161.1868	15.0266	.11±.60	-.66±1.15
1957	14	254	448	352.71	2673.6044	13.8193	.15±.60	-.45±1.15
1958	18	226	388	292.22	2041.3595	10.6494	.60±.54	-.55±1.04
1959	24	182	374	285.42	1591.1232	8.1423	-.37±.47	.99± .92
1960	14	194	370	261.86	3683.3626	16.2203	.56±.60	-1.13±1.15
1961	17	202	374	290.00	2198.5000	11.3721	-.05±.55	-.67±1.06
1962	15	214	364	291.20	1806.1714	10.9732	.23±.58	-.76±1.12
1963	22	222	326	265.73	1102.5887	7.0794	.42±.49	-1.12± .95
1964	35	200	410	304.97	3199.8521	9.5616	.03±.40	-.81± .78
1965	5	228	336	291.60	1768.8000	18.8085	-.60±.91	-.93±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	280	534	369.60	4405.9714	17.1386	.91±.58	.50±1.12
1950	17	203	431	311.82	4408.1544	15.7334	.26±.55	-.84±1.06
1951	17	229	427	309.41	2829.1324	12.9004	.83±.55	.17±1.06
1952	19	226	438	323.47	3024.3743	12.6166	-.03±.52	-.52±1.01
1953	16	222	406	299.00	2612.8000	12.7789	.42±.56	-.35±1.09

TABLE I-2(14) Basic statistics for secondary follicle number per sq. cm. (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type. (The corresponding information is included for the dams born from 1949 to 1953 inclusive, and used in the initial years of the experiment.)

MALE PROGENY OF INDEX FLOCK								
Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1954	27	2944	7400	4706.96	887227.8063	181.2741	.52±.45	.94± .87
1955	30	3748	9064	5587.73	1156941.2368	196.3790	1.05±.43	2.12± .83
1956	26	3635	7944	5149.58	902037.8538	186.2626	.69±.46	1.15± .89
1957	18	3722	7400	5563.44	1207696.8497	259.0256	.14±.54	-.80±1.04
1958	11	3800	6502	4833.46	777904.8727	265.9297	.80±.66	-.58±1.28
1959	29	2078	5776	3745.10	591467.8818	142.8126	.29±.43	.60± .85
1960	5	2936	4898	4018.80	497645.2000	315.4822	-.46±.91	-.51±2.00
1961								
1962	14	2676	6804	4720.86	864684.1319	248.5219	.01±.60	1.14±1.15
1963	16	2544	5332	3817.13	528233.0500	181.6991	.50±.56	-.12±1.09
1964	27	2720	5454	4194.59	483383.6353	133.8025	-.33±.45	-.40± .87
1965	12	2892	5552	4101.33	626244.6061	228.4449	.43±.64	-.35±1.23
MALE PROGENY OF VISUAL FLOCK								
1954	22	3214	5518	4284.91	532304.2771	155.5495	.05±.49	-1.21± .95
1955	25	3603	7055	5291.20	846440.9167	184.0044	.08±.46	-.82± .90
1956	10	3106	6247	4968.80	1129049.2889	336.0133	-.46±.69	-1.12±1.33
1957	12	4726	7700	6076.00	870160.0000	269.2830	.50±.64	-.62±1.23
1958	11	3768	6324	4547.27	556967.4182	225.0187	1.08±.66	.92±1.28
1959	37	2568	6044	3986.05	765294.5526	143.8181	.52±.39	-.18± .76
1960	18	2776	5064	4206.67	436119.0588	155.6561	-.79±.54	-.26±1.04
1961	10	4082	6948	4897.60	673617.6000	259.5414	1.59±.69	2.05±1.33
1962	10	3310	5800	4526.40	677150.0444	260.2211	.24±.69	-1.02±1.33
1963	20	3182	5740	4180.90	564305.0421	167.9740	.75±.51	-.56± .99
1964	35	2438	5554	4090.23	515726.5345	121.3880	.17±.40	-.22± .78
1965	11	2512	5332	4011.64	928458.2545	290.5259	-.14±.66	-1.00±1.28
FEMALE PROGENY OF INDEX FLOCK								
1954	29	3430	6438	4400.00	366279.1429	112.3847	1.16±.43	2.62± .85
1955	23	3623	7346	5353.70	957322.8577	204.0165	.24±.48	-.42± .93
1956	22	2789	9098	5615.73	2181121.4459	314.8680	.97±.49	1.01± .95
1957	21	3544	7838	5678.19	1106451.1619	229.5390	-.06±.50	.04± .97
1958	9	3498	4948	4430.22	255080.4444	168.3516	-.73±.72	-.82±1.40
1959	40	3416	7128	4617.35	589273.9256	121.3748	1.28±.37	1.56± .73
1960	6	3642	6846	4816.00	1169132.8000	441.4244	1.14±.85	.33±1.74
1961								
1962	19	3382	6182	5027.37	581447.5789	174.9357	-.09±.52	-.58±1.01
1963	19	2554	4858	4173.05	361272.6082	137.8925	-1.18±.52	.95±1.01
1964	24	3822	7572	5175.25	1199391.2391	223.5501	.85±.47	-.27± .92
1965	10	3566	6572	4742.20	753081.2889	274.4233	.65±.69	.18±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1954	29	3366	6036	4476.83	455874.4335	125.3787	.64±.43	-.10± .85
1955	26	3551	7290	5152.89	828927.3062	178.5548	.42±.46	-.18± .89
1956	14	3355	6704	4904.29	1122477.7582	283.1554	.30±.60	-1.00±1.15
1957	14	3264	7636	5584.43	1353675.1868	310.9519	-.15±.60	-.48±1.15
1958	18	3038	6732	4624.44	689378.1438	195.7008	.67±.54	.80±1.04
1959	24	2852	7076	4451.58	1048095.2971	208.9752	.51±.47	.29± .92
1960	14	3432	7368	4809.29	1040355.4505	272.6006	1.03±.60	.92±1.15
1961	17	2436	7162	5048.59	1017606.8824	244.6615	-.43±.55	1.44±1.06
1962	15	3082	5700	4415.87	669460.8381	211.2598	-.06±.58	-1.25±1.12
1963	22	3160	5714	4425.36	523203.1948	154.2140	.10±.49	-.81± .95
1964	35	2994	6914	4818.00	627756.0000	133.9249	.16±.40	.47± .78
1965	5	3570	5366	4491.20	674831.2000	367.3775	.16±.91	-1.68±2.00
DAMS BORN IN 1949,1950,1951,1952,1953								
1949	15	3756	8888	5379.40	1494211.9714	315.6171	1.59±.58	2.62±1.12
1950	17	3020	6250	4615.88	804750.7353	217.5736	.07±.55	-.81±1.06
1951	17	3992	7519	5359.88	1100079.3603	254.3827	.56±.55	-.78±1.06
1952	19	4028	7242	5514.63	1068473.3567	237.1401	.51±.52	-.86±1.01
1953	16	3290	6914	4828.88	743500.7833	215.5662	.61±.56	.49±1.09

TABLE I-2(15) Basic statistics for skin thickness in cm. (where g_1 and g_2 are Fisher's coefficients of skewness and kurtosis respectively). The data applies only to the progeny born and raised as twins and has been partitioned by sex and flock type.

MALE PROGENY OF INDEX FLOCK

Year	N	Min.	Max.	Mean	Variance	S.E.(Mean)	g_1 ± S.E.	g_2 ± S.E.
1958	11	.216	.318	.247	.0007	.0081	1.57±.66	2.46±1.28
1959	29	.226	.300	.260	.0003	.0032	.07±.43	.32± .05
1960	5	.251	.302	.270	.0004	.0091	.86±.91	-.74±2.00
1961								
1962	14	.196	.267	.233	.0004	.0056	-.00±.60	-.80±1.15
1963	16	.213	.269	.246	.0003	.0045	-.34±.56	-1.11±1.09
1964	27	.201	.267	.230	.0003	.0036	.23±.45	-.71± .87
1965	12	.241	.287	.260	.0003	.0046	.45±.64	-1.05±1.23
MALE PROGENY OF VISUAL FLOCK								
1958	11	.211	.267	.240	.0002	.0046	-.20±.66	-.27±1.28
1959	37	.203	.312	.262	.0005	.0036	-.05±.39	.34± .76
1960	18	.234	.284	.259	.0002	.0036	.23±.54	-.68±1.04
1961	10	.224	.274	.249	.0003	.0052	.03±.69	-1.03±1.33
1962	10	.173	.257	.218	.0007	.0084	-.19±.69	-.97±1.33
1963	20	.198	.272	.241	.0004	.0045	-.35±.51	-.64± .99
1964	35	.100	.264	.226	.0003	.0032	.05±.40	-.27± .78
1965	11	.226	.282	.247	.0003	.0049	.65±.66	-.07±1.28
FEMALE PROGENY OF INDEX FLOCK								
1958	9	.239	.305	.264	.0004	.0070	.75±.72	-.51±1.40
1959	40	.211	.323	.263	.0006	.0039	.42±.37	.19± .73
1960	6	.208	.274	.245	.0006	.0099	-.27±.85	-1.05±1.74
1961								
1962	19	.175	.259	.225	.0005	.0049	-.43±.52	.00±1.01
1963	19	.198	.269	.232	.0003	.0040	.02±.52	.51±1.01
1964	24	.168	.246	.211	.0004	.0042	-.31±.47	-.40± .92
1965	10	.203	.282	.248	.0005	.0068	-.50±.69	.20±1.33
FEMALE PROGENY OF VISUAL FLOCK								
1958	18	.221	.295	.251	.0003	.0043	.96±.54	.86±1.04
1959	24	.211	.307	.251	.0007	.0053	.43±.47	-.71± .92
1960	14	.203	.262	.235	.0003	.0048	-.17±.60	-.99±1.15
1961	17	.206	.295	.237	.0006	.0060	.79±.55	-.11±1.06
1962	15	.160	.274	.212	.0009	.0078	.46±.50	-.40±1.12
1963	22	.188	.287	.232	.0006	.0053	.23±.49	-.29± .95
1964	35	.165	.229	.198	.0003	.0028	-.16±.40	-.78± .78
1965	5	.203	.249	.239	.0004	.0090	-1.49±.91	.23±2.00

Table I-3(1) Univariate comparison of Index and Visual flocks for the variate date of birth. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		MALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.69	.036*	.05	.962	.133	.514
	T	2.68	.018*	-1.56	.126	.306	.146
1955	S	1.95	.005**	-1.12	.265	-.134	.448
	T	4.33	.000***	-2.61	.012*	.320	.086
1956	S	1.35	.228	.69	.492	-.091	.859
	T	2.88	.102	-.66	.517	.438	.085
1957	S	1.93	.007**	1.98	.049*	.173	.226
	T	5.53	.006**	1.11	.276	-.194	.842
1958	S	1.48	.117	-.72	.474	.104	.813
	T	5.57	.012*	-1.35	.192	.273	.480
1959	S	1.18	.494	1.83	.070	-.158	.270
	T	1.50	.248	.80	.425	-.131	.870
1960	S	1.13	.652	3.34	.001**	-.283	.014*
	T	1.15	.734	.35	.727	.389	.445
1961	S	1.39	.166	2.60	.010*	.249	.018*
	T						
1962	S	1.19	.427	.81	.416	.135	.376
	T	2.60	.156	1.03	.314	-.286	.549
1963	S	1.17	.540	.16	.875	-.115	.697
	T	1.97	.186	2.24	.032*	.450	.033*
1964	S	1.46	.121	.23	.815	.064	.993
	T	1.35	.410	-.76	.449	-.210	.413
1965	S	2.45	.001**	5.06	.000***	.362	.001**
	T	2.85	.110	1.63	.119	.311	.451

Table I-3(2) Univariate comparison of Index and Visual flocks for the variate birth weight in kg. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

Year of birth		MALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.12	.635	2.45	.016*	-.214	.074
	T	1.10	.826	-.20	.839	.150	.866
1955	S	1.04	.856	-1.05	.294	-.087	.893
	T	1.03	.950	-2.88	.006**	.447	.005**
1956	S	1.12	.643	.40	.687	.091	.859
	T	1.60	.339	-.15	.884	.100	.999
1957	S	1.21	.430	-.86	.389	.060	.998
	T	1.37	.607	-1.18	.247	.250	.602
1958	S	1.79	.023*	.56	.574	-.175	.225
	T	1.31	.678	1.06	.300	-.455	.081
1959	S	1.50	.089	-.67	.504	.110	.691
	T	2.36	.016*	-.25	.800	-.195	.464
1960	S	1.26	.392	-.99	.327	.134	.592
	T	1.16	.988	.14	.887	.189	.981
1961	S	1.02	.929	2.28	.024*	.163	.248
	T						
1962	S	1.25	.322	-1.82	.071	-.150	.263
	T	1.87	.350	-1.69	.104	.400	.200
1963	S	1.15	.578	1.01	.316	.086	.926
	T	1.47	.421	.36	.722	.125	.983
1964	S	1.28	.294	.15	.882	.088	.900
	T	1.11	.783	-1.91	.061	-.247	.235
1965	S	1.77	.040*	2.00	.047*	.224	.089
	T	7.98	.002**	1.12	.276	-.386	.229

Table I-3(3) Univariate comparison of Index and Visual flocks for the variate weaner body weight in kg. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

Year of birth		MALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.47	.121	.17	.866	-.132	.525
	T	1.26	.596	.66	.510	-.161	.808
1955	S	1.21	.412	-1.59	.114	.178	.153
	T	1.21	.613	-.91	.365	.153	.808
1956	S	1.19	.487	-.27	.788	.121	.580
	T	2.21	.115	-.02	.982	.234	.714
1957	S	1.15	.573	-1.20	.233	.161	.298
	T	1.26	.710	-1.67	.107	.333	.282
1958	S	1.07	.777	-.07	.945	.090	.911
	T	3.12	.087	.31	.761	-.182	.826
1959	S	2.00	.004**	-.50	.618	.171	.197
	T	1.54	.219	1.07	.286	.250	.200
1960	S	1.12	.613	-3.54	.000***	.319	.004**
	T	1.86	.579	.67	.508	.411	.382
1961	S	1.43	.135	.89	.374	.181	.157
	T						
1962	S	1.22	.361	-2.76	.006**	-.215	.034*
	T	1.55	.513	-3.43	.002**	.614	.013*
1963	S	1.07	.774	.47	.637	.166	.267
	T	2.02	.173	-.13	.899	.213	.673
1964	S	1.26	.333	.07	.944	-.106	.750
	T	1.22	.577	-.59	.560	.156	.752
1965	S	1.41	.212	-.95	.344	-.215	.112
	T	4.59	.023*	-1.18	.251	.417	.167

Table I-3(4) Univariate comparison of Index and Visual flocks for the variate hogget body weight in kg. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		MALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.14	.596	.60	.552	-.130	.548
	T	3.83	.003**	.19	.849	-.214	.510
1955	S	1.27	.316	-1.42	.157	-.153	.292
	T	1.06	.899	-.84	.406	.240	.316
1956	S	1.11	.670	-.52	.604	.091	.859
	T	3.42	.014*	-.32	.755	.223	.748
1957	S	1.42	.146	-.50	.621	.113	.713
	T	1.12	.808	-1.87	.072	.389	.150
1958	S	1.17	.527	.78	.437	-.141	.464
	T	1.56	.496	-.66	.515	.364	.217
1959	S	1.09	.712	-1.02	.310	.141	.397
	T	1.30	.450	.51	.612	.235	.257
1960	S	1.12	.669	-1.78	.077	.199	.165
	T	1.03	.838	-1.04	.309	-.389	.445
1961	S	1.09	.714	1.08	.282	.087	.906
	T						
1962	S	1.02	.947	-2.61	.010*	-.211	.040*
	T	2.22	.235	-1.49	.150	.329	.392
1963	S	1.26	.346	-.65	.516	-.143	.431
	T	1.23	.664	.17	.865	.163	.897
1964	S	1.12	.643	-2.24	.027*	.198	.099
	T	1.66	.185	-2.61	.011	-.415	.007**
1965	S	1.44	.187	-.65	.518	-.142	.534
	T	1.04	.939	-1.22	.236	.409	.181

Table I-3(5) Univariate comparison of Index and Visual flocks for the variate lamb fleece weight in kg. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

Year of birth		MALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.05	.837	1.20	.234	-.124	.603
	T	1.62	.243	1.28	.206	-.423	.017*
1955	S	1.44	.122	.32	.752	.104	.743
	T	1.03	.959	1.06	.293	-.233	.347
1956	S	1.13	.616	-.07	.948	-.061	.993
	T	1.02	.897	1.28	.211	.362	.220
1957	S	1.43	.156	-2.24	.027*	.232	.042*
	T	1.09	.850	-1.79	.084	.361	.208
1958	S	1.06	.806	-.28	.783	.149	.397
	T	1.74	.397	.38	.711	-.273	.480
1959	S	1.24	.373	-.70	.487	.127	.531
	T	1.38	.381	.24	.810	-.226	.298
1960	S	1.46	.157	-3.97	.000***	.395	.000***
	T	18.92	.011*	.70	.489	.444	.300
1961	S	1.36	.198	.20	.841	.090	.885
	T						
1962	S	1.10	.676	-.59	.555	.139	.346
	T	1.85	.305	-1.70	.103	.457	.106
1963	S	1.62	.053	-.08	.938	.117	.671
	T	1.44	.474	.13	.895	.125	.983
1964	S	1.46	.116	.72	.476	-.103	.773
	T	1.23	.565	-.43	.665	-.141	.841
1965	S	1.75	.031*	-1.94	.055	-.232	.072
	T	1.77	.377	-1.70	.104	.386	.229

Table I-3(6) Univariate comparison of Index and Visual flocks for the variate hogget greasy fleece weight in kg. Parametric tests (variance ratio and t-test) and a non-parametric test (Kolmogorov-Smirnov) are presented for the 12 years 1954 to 1965.

		MALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.00	1.00	.51	.609	-.090	.902
	T	1.06	.910	.51	.612	-.163	.799
1955	S	1.41	.142	-.62	.539	-.137	.418
	T	1.25	.580	1.03	.309	-.253	.261
1956	S	1.35	.231	2.47	.015*	-.242	.026*
	T	1.32	.554	1.19	.241	-.200	.843
1957	S	1.10	.700	.55	.584	.095	.870
	T	1.31	.664	-2.93	.007**	.500	.032*
1958	S	1.07	.786	1.30	.196	-.129	.577
	T	3.19	.081	-.53	.604	.273	.480
1959	S	1.01	.973	2.46	.015*	-.182	.149
	T	1.01	.957	1.28	.206	.167	.647
1960	S	1.15	.602	.99	.324	-.097	.898
	T	1.17	.720	1.53	.142	.389	.445
1961	S	1.29	.290	3.63	.000***	.266	.009**
	T						
1962	S	1.03	.907	.88	.382	.101	.728
	T	2.05	.231	1.28	.215	.457	.106
1963	S	1.37	.202	2.45	.016*	.233	.042*
	T	1.60	.330	1.96	.058	.438	.041*
1964	S	2.14	.002**	3.57	.000***	-.296	.003**
	T	2.49	.019*	.77	.443	.277	.142
1965	S	1.35	.277	4.03	.000***	.370	.001**
	T	1.73	.382	.40	.690	.288	.536

Table I-3(7) Univariate comparison of Index and Visual flocks for the variate percentage yield. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		MALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.13	.616	.36	.719	-.158	.316
	T	1.46	.356	-1.38	.174	.281	.215
1955	S	1.04	.880	1.65	.100	.188	.115
	T	1.11	.776	1.61	.113	-.240	.316
1956	S	1.07	.775	-.13	.900	.121	.580
	T	1.92	.308	.72	.478	.292	.444
1957	S	1.17	.542	.59	.559	.119	.655
	T	1.32	.647	-.30	.766	-.194	.842
1958	S	1.01	.962	4.24	.000***	.331	.001**
	T	2.26	.215	2.22	.038*	-.455	.081
1959	S	1.19	.456	.42	.675	-.151	.323
	T	1.19	.643	1.24	.220	.186	.516
1960	S	1.03	.901	1.10	.275	-.162	.364
	T	1.93	.303	.31	.758	.433	.326
1961	S	1.17	.523	2.71	.008**	.239	.026*
	T						
1962	S	1.05	.820	2.32	.021*	.182	.105
	T	3.92	.046*	.61	.545	-.357	.304
1963	S	1.71	.032	.26	.795	.154	.344
	T	2.68	.058	.13	.900	-.238	.548
1964	S	1.41	.152	1.90	.060	-.234	.031*
	T	1.30	.474	2.42	.018*	.347	.035*
1965	S	1.07	.822	-1.55	.123	-.150	.463
	T	3.15	.073	3.45	.002**	-.644	.008**

Table I-3(8) Univariate comparison of Index and Visual flocks for the variate clean fleece weight in kg. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		MALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.12	.653	.74	.462	-.142	.439
	T	1.27	.570	-.36	.718	.160	.817
1955	S	1.31	.243	.33	.743	-.110	.682
	T	1.46	.320	1.50	.140	-.280	.172
1956	S	1.08	.769	2.23	.028*	-.136	.442
	T	1.34	.534	1.46	.154	-.277	.508
1957	S	1.11	.695	.87	.388	.148	.384
	T	2.05	.229	-2.42	.022*	.472	.049*
1958	S	1.00	.990	3.53	.001**	-.276	.011**
	T	3.57	.057	.45	.658	-.182	.826
1959	S	1.07	.766	2.25	.026*	-.185	.135
	T	1.14	.707	2.05	.045*	.264	.156
1960	S	1.01	.975	1.40	.166	-.109	.813
	T	1.93	.304	1.50	.149	.378	.478
1961	S	1.11	.653	5.03	.000***	.378	.000***
	T						
1962	S	1.10	.676	1.87	.064	.174	.134
	T	1.64	.407	1.61	.121	.586	.020*
1963	S	1.65	.044*	2.43	.017*	.284	.007**
	T	1.80	.227	1.86	.071	.413	.062
1964	S	1.86	.011*	4.71	.000***	-.365	.000***
	T	1.34	.445	2.20	.032*	.294	.098
1965	S	1.20	.482	2.65	.009**	.314	.005**
	T	1.01	.979	3.43	.003**	-.561	.028*

Table I-3(9) Univariate comparison of Index and Visual flocks for the variate staple length in cm. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

Year of birth		MALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.05	.855	.07	.940	.142	.437
	T	2.10	.074	.24	.015	.162	.808
1955	S	1.01	.959	1.55	.124	.129	.488
	T	2.47	.026*	1.82	.074	-.260	.236
1956	S	1.95	.008**	3.04	.003**	-.257	.015*
	T	2.10	.247	1.09	.282	-.331	.307
1957	S	1.49	.099	3.13	.002**	.220	.061
	T	1.01	1.015	-.11	.913	.139	.981
1958	S	1.14	.613	2.90	.004**	-.319	.002**
	T	3.29	.074	.80	.433	-.273	.480
1959	S	1.32	.245	3.20	.002**	-.251	.015*
	T	1.34	.422	1.47	.145	.227	.289
1960	S	1.08	.784	1.86	.064	-.139	.552
	T	1.53	.479	1.61	.122	.578	.094
1961	S	1.37	.194	3.15	.002**	.260	.012*
	T						
1962	S	2.25	.000***	1.53	.129	.153	.244
	T	1.09	.860	2.24	.036*	-.500	.063
1963	S	1.12	.648	1.07	.285	.139	.466
	T	1.43	.486	1.17	.249	.300	.288
1964	S	2.13	.002**	2.46	.015*	-.223	.046*
	T	1.75	.143	2.40	.019*	.358	.028*
1965	S	1.19	.496	2.46	.015*	.241	.055
	T	1.34	.650	2.35	.028*	-.409	.181

Table I-3(10) Univariate comparison of Index and Visual flocks for the variate crimps per inch. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		MALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.02	.939	-.79	.432	.133	.520
	T	1.28	.574	-.41	.681	.146	.881
1955	S	1.30	.265	-.46	.645	-.097	.812
	T	2.26	.037*	-2.37	.022*	.207	.487
1956	S	1.24	.385	.79	.432	-.076	.952
	T	1.74	.263	-.73	.468	.215	.781
1957	S	1.30	.275	-.17	.862	.107	.769
	T	2.57	.115	.37	.712	-.278	.481
1958	S	1.05	.855	.80	.428	-.106	.792
	T	1.98	.297	.33	.741	-.182	.826
1959	S	1.09	.705	-2.41	.017*	.190	.117
	T	1.20	.621	-.92	.360	-.120	.922
1960	S	1.09	.737	-.32	.753	.058	.999
	T	1.92	.306	-.13	.896	-.200	.970
1961	S	1.36	.197	-3.27	.001**	-.269	.008**
	T						
1962	S	1.06	.781	.39	.700	.087	.864
	T	2.43	.143	-2.19	.039*	.329	.392
1963	S	1.13	.615	-.83	.407	-.089	.913
	T	1.21	.718	-1.03	.312	-.188	.794
1964	S	1.19	.457	-2.29	.024*	.209	.070
	T	2.39	.025*	-1.75	.085	-.269	.165
1965	S	1.43	.195	-3.47	.001**	-.264	.028*
	T	1.38	.602	-2.60	.017*	.386	.229

Table I-3(11) Univariate comparison of Index and Visual flocks for the variate fibre diameter in microns. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		MALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.33	.258	-.86	.390	.230	.046*
	T	1.62	.261	-.77	.446	.274	.237
1955	S	1.18	.469	2.03	.044*	.219	.044*
	T	1.63	.211	-1.00	.324	.200	.525
1956	S	1.10	.701	-3.50	.001**	.394	.000***
	T	1.10	.930	-.37	.711	.238	.679
1957	S	1.06	.830	1.43	.155	.143	.433
	T	1.27	.700	2.00	.055	.361	.208
1958	S	1.82	.018*	.19	.852	.156	.347
	T	2.73	.129	.91	.375	-.364	.217
1959	S	1.04	.873	2.32	.022*	-.209	.066
	T	1.84	.100	1.71	.092	.338	.035*
1960	S	1.50	.128	.51	.609	-.195	.180
	T	1.44	.788	2.06	.052	.667	.037*
1961	S	1.50	.093	.07	.944	.139	.425
	T						
1962	S	1.15	.535	-4.88	.000***	-.360	.000***
	T	1.34	.613	-.08	.934	.243	.722
1963	S	1.05	.844	-2.13	.035*	-.218	.066
	T	3.61	.015*	.45	.656	.338	.181
1964	S	1.23	.399	-2.98	.003**	.227	.040*
	T	1.59	.228	-.29	.770	-.166	.684
1965	S	1.83	.030*	.07	.948	.108	.833
	T	4.60	.019*	-.17	.484	.364	.285

Table I-3(12) Univariate comparison of Index and Visual flocks for the variate coefficient of variation of fibre diameter. Parametric tests (variance ratio and t-test) and a non-parametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		MALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.54	.080	-1.86	.065	.209	.087
	T	1.02	.948	-1.89	.065	.278	.226
1955	S	1.38	.164	.60	.548	.144	.359
	T	1.05	.913	-.87	.387	.207	.487
1956	S	1.44	.142	-1.08	.281	.197	.105
	T	1.09	.951	.42	.674	.262	.575
1957	S	1.27	.345	-.52	.603	-.119	.655
	T	2.20	.139	-2.63	.014*	.333	.282
1958	S	1.05	.845	-1.93	.056	.189	.160
	T	1.25	.734	-.56	.585	.182	.826
1959	S	1.48	.096	-1.23	.220	.171	.198
	T	4.10	.000***	-1.68	.097	-.270	.138
1960	S	1.45	.164	-2.26	.026*	.240	.056
	T	3.57	.055	3.38	.003**	.667	.037*
1961	S	1.72	.023*	-.79	.431	-.146	.363
	T						
1962	S	1.13	.586	-2.45	.015*	-.228	.021*
	T	5.29	.017*	1.06	.299	.429	.147
1963	S	1.01	.957	-.45	.656	.093	.884
	T	1.39	.525	-.44	.666	.288	.332
1964	S	1.17	.518	2.07	.040*	-.224	.043*
	T	1.05	.875	-.36	.717	-.105	.977
1965	S	1.84	.028*	.94	.349	.198	.170
	T	2.51	.147	-.79	.436	.242	.717

Table I-3(13) Univariate comparison of Index and Visual flocks for the variate primary follicle number per sq. cm. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		MALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.18	.498	2.02	.045*	-.293	.005**
	T	1.58	.289	.37	.711	-.226	.447
1955	S	1.77	.014*	2.26	.025*	.202	.076
	T	2.20	.044*	-.68	.500	.173	.689
1956	S	1.09	.719	2.27	.025*	-.227	.043*
	T	1.79	.243	1.51	.141	-.384	.168
1957	S	1.26	.340	1.93	.055	.226	.051
	T	1.84	.307	.19	.853	.194	.842
1958	S	1.32	.272	1.79	.075	-.229	.053
	T	1.37	.629	-.27	.791	-.182	.826
1959	S	1.24	.369	.39	.696	-.116	.633
	T	2.15	.031*	-1.32	.192	-.271	.136
1960	S	1.20	.504	.53	.597	.122	.701
	T	6.67	.079	-1.20	.243	-.500	.191
1961	S	1.36	.195	-1.38	.171	-.169	.212
	T						
1962	S	1.03	.910	-.09	.926	.077	.933
	T	1.75	.402	-.89	.926	.243	.722
1963	S	1.08	.763	-2.17	.032*	.240	.033*
	T	1.43	.457	-3.24	.003**	-.500	.014*
1964	S	1.02	.927	.20	.843	-.107	.739
	T	1.68	.153	.47	.639	.130	.895
1965	S	1.37	.228	1.47	.145	.179	.265
	T	1.44	.572	-.61	.549	.326	.398

Table I-3(14) Univariate comparison of Index and Visual flocks for the variate secondary follicle number per sq. cm. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for the 12 years 1954 to 1965.

Year of birth		MALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.11	.672	2.15	.033*	-.216	.069
	T	1.69	.235	1.72	.092	-.241	.372
1955	S	1.25	.338	2.09	.039*	.175	.167
	T	1.37	.438	1.09	.282	-.193	.565
1956	S	1.21	.438	1.63	.105	-.227	.043*
	T	1.25	.620	.50	.624	-.208	.812
1957	S	1.36	.225	2.55	.012*	.286	.006**
	T	1.39	.589	-1.33	.195	.306	.373
1958	S	1.22	.436	.20	.844	-.135	.517
	T	1.40	.607	.82	.421	-.273	.480
1959	S	1.33	.235	.37	.710	.077	.958
	T	1.29	.485	-1.17	.246	-.199	.441
1960	S	1.38	.232	.91	.365	-.111	.796
	T	1.14	.742	-.56	.585	-.356	.548
1961	S	1.06	.809	.57	.568	.096	.838
	T						
1962	S	1.49	.073	3.58	.000***	.344	.000***
	T	1.28	.727	.53	.602	-.329	.392
1963	S	1.24	.387	1.14	.257	.154	.347
	T	1.03	.909	-1.46	.152	.300	.288
1964	S	1.84	.012*	3.97	.000***	-.303	.002**
	T	1.07	.875	.58	.567	.267	.171
1965	S	1.06	.797	.94	.347	.110	.817
	T	1.48	.527	.24	.809	-.303	.478

Table I-3(15) Univariate comparison of Index and Visual flocks for the variate skin thickness in cm. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for the 8 years 1958 to 1965.

Year of birth		MALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1958	S	1.44	.148	-.87	.387	.132	.550
	T	3.09	.089	.76	.453	.182	.826
1959	S	1.55	.065	.01	.990	.096	.829
	T	1.63	.185	-.41	.684	-.140	.823
1960	S	1.20	.485	2.59	.011*	-.242	.052
	T	1.83	.331	1.28	.214	.300	.732
1961	S	1.41	.153	2.12	.036*	.234	.031*
	T						
1962	S	1.71	.017*	4.18	.000***	.307	.001**
	T	1.58	.440	1.63	.110	-.329	.392
1963	S	1.08	.767	-.45	.654	-.194	.133
	T	1.25	.667	.82	.415	.163	.897
1964	S	1.15	.555	.44	.661	-.139	.426
	T	1.01	.993	.74	.466	.184	.566
1965	S	1.10	.750	1.17	.244	.219	.101
	T	1.06	.920	2.01	.058	.455	.110

Table I-4(1) Univariate comparison of Index and Visual flocks for the variate date of birth. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

Year of birth		FEMALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.23	.412	.54	.593	.117	.711
	T	2.34	.028*	-1.00	.322	.276	.127
1955	S	1.20	.492	.54	.593	-.215	.097
	T	7.89	.000***	-2.45	.018*	-.308	.140
1956	S	1.97	.009**	.52	.604	.101	.880
	T	1.32	.551	-.47	.644	.221	.659
1957	S	1.62	.057	3.08	.003**	-.322	.002**
	T	2.86	.057	.37	.713	-.190	.808
1958	S	1.62	.071	-1.54	.127	-.117	.740
	T	7.51	.007**	-1.22	.232	-.333	.299
1959	S	2.17	.001**	1.56	.122	.118	.605
	T	5.34	.000***	1.51	.136	-.275	.157
1960	S	1.26	.315	3.03	.003**	-.239	.018*
	T	1.84	.350	.44	.663	.286	.724
1961	S	1.22	.378	2.01	.046*	.160	.228
	T						
1962	S	1.20	.434	2.66	.009**	.230	.027*
	T	3.49	.022*	1.26	.215	-.302	.309
1963	S	1.02	.912	-.73	.466	-.221	.119
	T	2.08	.121	.06	.949	-.261	.368
1964	S	1.51	.086	-2.32	.022*	.241	.025
	T	2.76	.007**	1.42	.160	.269	.192
1965	S	4.48	.000***	3.05	.003**	.242	.094
	T	5.75	.108	.90	.386	-.300	.613

Table I-4(2) Univariate comparison of Index and Visual flocks for the variate birth weight in kg. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		FEMALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.03	.887	-.74	.464	-.081	.963
	T	1.10	.799	-1.72	.092	.241	.227
1955	S	1.23	.429	-.08	.938	-.078	.975
	T	1.81	.152	-1.21	.233	-.161	.808
1956	S	1.79	.036*	.27	.786	.094	.924
	T	2.11	.169	-.25	.803	.156	.931
1957	S	1.66	.048*	.67	.506	.243	.033*
	T	1.87	.248	.16	.875	.095	.999
1958	S	1.42	.186	.39	.698	-.066	.996
	T	1.02	.831	.74	.469	.278	.484
1959	S	1.27	.318	-.07	.944	.066	.988
	T	1.60	.194	-.30	.765	.117	.953
1960	S	1.05	.821	2.04	.044*	-.164	.208
	T	2.33	.358	-.14	.889	-.214	.927
1961	S	1.29	.259	.37	.709	-.086	.887
	T						
1962	S	1.47	.099	-.89	.376	-.112	.643
	T	1.19	.745	-.30	.764	.179	.851
1963	S	1.02	.971	-.13	.896	-.057	.999
	T	2.06	.115	1.34	.187	.208	.627
1964	S	1.72	.026*	-1.19	.237	.140	.424
	T	1.43	.337	1.76	.084	.248	.268
1965	S	1.14	.684	.83	.411	.085	.981
	T	2.68	.202	.61	.550	-.300	.613

Table I-4(3) Univariate comparison of Index and Visual flocks for the variate weaner body weight in kg. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		FEMALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	2.10	.004**	1.55	.123	.211	.101
	T	1.19	.654	-.14	.887	.207	.377
1955	S	1.36	.242	.91	.362	.145	.477
	T	1.43	.404	-.89	.379	-.145	.881
1956	S	1.32	.283	-1.80	.075	-.235	.065
	T	2.01	.197	.71	.485	.156	.931
1957	S	1.40	.183	-1.46	.146	.132	.548
	T	1.13	.836	-.19	.850	.095	.999
1958	S	1.07	.802	.74	.463	.146	.482
	T	1.94	.239	1.57	.128	.500	.044*
1959	S	1.25	.350	-.49	.625	-.133	.464
	T	1.43	.363	-.30	.764	.158	.750
1960	S	1.09	.689	-.91	.363	.195	.083
	T	1.82	.526	-.65	.527	.286	.724
1961	S	1.44	.114	1.63	.105	.207	.058
	T						
1962	S	1.13	.600	-4.59	.000***	-.284	.003**
	T	1.10	.762	-.77	.449	.182	.837
1963	S	1.19	.546	-.59	.559	-.169	.363
	T	1.95	.144	-.17	.862	.201	.666
1964	S	1.19	.476	-1.27	.208	.151	.337
	T	1.99	.066	-2.51	.015*	-.365	.030*
1965	S	1.37	.263	-1.23	.220	-.227	.131
	T	2.13	.486	-.95	.362	.300	.613

Table I-4(4) Univariate comparison of Index and Visual flocks for the variate hogget body weight in kg. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		FEMALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.78	.023*	.30	.768	.165	.310
	T	1.24	.568	-.34	.739	.172	.576
1955	S	1.36	.238	2.21	.029*	.207	.120
	T	1.34	.476	-.42	.677	-.130	.938
1956	S	1.10	.703	-.28	.783	-.092	.932
	T	1.96	.214	-.65	.521	.182	.841
1957	S	1.43	.162	-1.13	.259	.168	.266
	T	1.44	.506	.07	.945	.143	.963
1958	S	1.48	.137	1.30	.195	.203	.144
	T	4.08	.014*	1.43	.166	.333	.299
1959	S	1.52	.076	-.06	.953	-.072	.973
	T	1.20	.658	2.15	.036*	-.242	.271
1960	S	1.34	.199	-2.20	.029*	.179	.137
	T	5.78	.064	-1.39	.180	-.690	.019*
1961	S	1.68	.027*	1.55	.124	.140	.372
	T						
1962	S	1.02	.934	-5.21	.000***	-.344	.000***
	T	1.03	.966	-.06	.951	-.165	.902
1963	S	1.14	.663	-.22	.827	.068	.997
	T	1.64	.274	-.82	.416	.165	.848
1964	S	1.78	.018*	-3.15	.002**	.264	.011
	T	2.04	.078	-1.49	.141	-.277	.167
1965	S	1.29	.367	-2.32	.022*	-.302	.018*
	T	3.13	.283	-.24	.814	.400	.339

Table I-4(5) Univariate comparison of Index and Visual flocks for the variate lamb fleece weight in kg. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		FEMALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.71	.034*	-.16	.874	-.117	.711
	T	1.28	.512	.40	.691	-.241	.227
1955	S	1.01	.961	1.57	.118	-.160	.360
	T	1.28	.556	.67	.503	.189	.646
1956	S	1.76	.029*	2.24	.027*	.226	.082
	T	1.13	.774	1.11	.273	.175	.867
1957	S	1.24	.391	-1.33	.187	.133	.535
	T	1.02	.998	-.20	.842	.071	.999
1958	S	1.39	.218	1.05	.297	.289	.011*
	T	5.04	.026*	1.19	.244	.389	.170
1959	S	1.03	.906	1.46	.146	.147	.346
	T	3.43	.001**	-1.68	.098	.342	.043*
1960	S	1.21	.394	-1.51	.133	.222	.034*
	T	2.59	.301	-.46	.650	-.357	.482
1961	S	1.01	.976	.28	.782	-.089	.866
	T						
1962	S	1.24	.353	-.54	.588	-.091	.850
	T	1.22	.710	-1.24	.224	.196	.775
1963	S	1.16	.616	-.75	.456	-.125	.726
	T	1.39	.463	-.93	.360	-.232	.502
1964	S	1.06	.804	2.02	.046*	-.131	.506
	T	2.65	.010*	-1.52	.134	-.344	.048*
1965	S	1.26	.403	-1.28	.204	-.163	.466
	T	1.04	.878	-.75	.466	.200	.894

Table I-4(6) Univariate comparison of Index and Visual flocks for the variate hogget greasy fleece weight in kg. Parametric tests (variance ratio and t-test) and a non-parametric test (Kolmogorov-Smirnov) are presented for the 12 years 1954 to 1965.

Year of birth		FEMALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.37	.213	-.11	.909	-.151	.413
	T	1.49	.301	.27	.791	-.138	.794
1955	S	1.51	.119	4.03	.000***	-.302	.006**
	T	1.84	.143	-1.21	.231	-.314	.125
1956	S	1.06	.847	3.91	.000***	.275	.019*
	T	1.76	.243	1.73	.092	-.377	.119
1957	S	1.03	.891	1.08	.284	-.153	.365
	T	1.10	.875	-.27	.787	-.262	.471
1958	S	1.57	.087	2.08	.039*	.204	.138
	T	1.91	.354	2.39	.025*	.611	.008**
1959	S	1.20	.445	3.22	.002**	.246	.018*
	T	1.25	.534	1.63	.107	-.250	.238
1960	S	1.07	.787	2.89	.004**	-.200	.071
	T	2.98	.235	1.01	.328	.429	.286
1961	S	1.40	.147	2.34	.021*	.212	.048*
	T						
1962	S	1.11	.642	.85	.395	.159	.237
	T	1.62	.362	2.04	.050*	-.312	.274
1963	S	1.05	.845	2.14	.035*	.199	.198
	T	1.32	.540	2.01	.051	.344	.120
1964	S	1.38	.183	2.90	.004**	-.289	.004**
	T	2.84	.006**	1.58	.121	.315	.084
1965	S	1.22	.515	.30	.766	.117	.829
	T	1.58	.522	3.31	.006**	-.800	.007**

Table I-4(7) Univariate comparison of Index and Visual flocks for the variate percentage yield. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		FEMALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.07	.795	-1.65	.102	-.168	.289
	T	2.00	.072	-1.13	.263	.241	.227
1955	S	1.67	.051	-.50	.620	.130	.605
	T	1.83	.158	.44	.665	.157	.826
1956	S	1.08	.793	.95	.345	.171	.308
	T	1.74	.303	-.48	.631	.260	.473
1957	S	1.22	.426	3.61	.000***	-.319	.002**
	T	3.48	.012*	3.21	.003**	-.500	.018*
1958	S	1.23	.438	2.71	.008**	.260	.029*
	T	1.55	.423	-1.23	.231	-.222	.709
1959	S	1.53	.071	1.91	.058	.171	.191
	T	1.13	.769	1.53	.132	-.225	.347
1960	S	1.45	.106	1.21	.227	-.164	.207
	T	1.95	.308	-2.45	.025*	-.690	.019*
1961	S	1.02	.914	2.36	.020*	.202	.068
	T						
1962	S	1.54	.062	3.46	.001**	.329	.000***
	T	1.08	.901	2.64	.013*	-.428	.058
1963	S	1.08	.802	.27	.790	.133	.655
	T	1.26	.625	.53	.600	.201	.666
1964	S	1.15	.566	1.71	.089	-.221	.050*
	T	1.35	.453	2.30	.025*	.324	.072
1965	S	1.59	.130	1.73	.086	.251	.075
	T	12.87	.025*	.27	.793	-.500	.160

Table I-4(8) Univariate comparison of Index and Visual flocks for the variate clean fleece weight in kg. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

Year of birth		FEMALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.60	.066	-.84	.403	-.134	.558
	T	2.07	.059	-.33	.742	.172	.576
1955	S	1.28	.344	3.53	.001**	-.317	.003**
	T	1.14	.746	-.84	.406	-.204	.559
1956	S	1.21	.487	4.03	.000***	.400	.000***
	T	1.63	.308	1.46	.154	.286	.364
1957	S	1.36	.219	3.00	.003**	-.252	.025*
	T	1.76	.246	1.28	.210	-.310	.287
1958	S	1.24	.410	3.62	.000***	.356	.001**
	T	1.55	.537	1.23	.230	.389	.170
1959	S	1.25	.346	3.95	.000***	.301	.002**
	T	1.56	.216	2.23	.030*	-.225	.347
1960	S	1.20	.441	2.99	.003**	-.190	.097
	T	3.24	.202	-.28	.781	-.357	.482
1961	S	1.28	.286	3.07	.003**	.284	.003**
	T						
1962	S	1.40	.144	2.85	.005**	.259	.009**
	T	1.38	.546	3.27	.003**	-.456	.037*
1963	S	1.49	.165	2.62	.010*	.312	.008**
	T	1.33	.522	1.93	.061	.342	.125
1964	S	1.37	.188	3.52	.001**	-.289	.004**
	T	2.77	.007**	2.55	.013*	.396	.015**
1965	S	1.01	.944	1.07	.288	.113	.858
	T	1.77	.610	1.97	.071	-.700	.024*

Table I-4(9) Univariate comparison of Index and Visual flocks for the variate staple length in cm. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

Year of birth		FEMALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	2.26	.001*	.00	1.00	.172	.266
	T	1.85	.109	-1.38	.173	.241	.227
1955	S	1.04	.864	4.02	.000***	-.353	.001**
	T	1.50	.330	.40	.690	.161	.808
1956	S	1.38	.210	3.04	.003**	.296	.009**
	T	2.17	.154	.44	.661	-.175	.867
1957	S	1.74	.030*	2.07	.040*	-.208	.098
	T	1.45	.494	-.46	.646	.310	.287
1958	S	1.18	.529	2.02	.046*	.165	.336
	T	1.15	.880	-.59	.562	.167	.907
1959	S	1.15	.555	2.49	.014*	.179	.153
	T	1.00	.973	2.82	.006**	-.300	.099
1960	S	1.11	.662	1.22	.226	-.115	.606
	T	1.88	.504	1.58	.132	.429	.286
1961	S	1.50	.081	2.08	.039*	.245	.014*
	T						
1962	S	1.37	.170	1.46	.145	.176	.154
	T	1.19	.717	1.78	.084	-.389	.102
1963	S	1.38	.236	1.51	.135	.169	.367
	T	1.28	.587	.88	.383	.285	.276
1964	S	1.62	.044*	1.12	.264	-.100	.808
	T	1.51	.272	1.57	.123	.188	.581
1965	S	1.07	.804	.40	.690	.144	.620
	T	1.42	.605	1.37	.192	-.400	.339

Table I-4(10) Univariate comparison of Index and Visual flocks for the variate crimps per inch. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

Year of birth		FEMALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.16	.560	-1.52	.130	.164	.316
	T	2.39	.025*	-2.68	.010*	.310	.066
1955	S	1.47	.140	-2.89	.005**	.209	.116
	T	1.42	.393	-1.84	.072	-.206	.549
1956	S	1.75	.045*	-2.39	.019*	-.173	.298
	T	1.49	.462	.09	.926	.188	.813
1957	S	1.03	.918	-3.21	.002**	.233	.047*
	T	1.33	.602	-1.65	.108	.214	.697
1958	S	2.49	.001*	-1.21	.230	-.148	.461
	T	3.47	.079	.43	.674	.167	.907
1959	S	1.08	.731	-2.76	.000**	-.173	.183
	T	1.25	.533	-1.28	.205	.142	.848
1960	S	1.64	.034*	.43	.671	-.095	.816
	T	3.56	.169	-.28	.783	-.286	.724
1961	S	1.06	.793	-.69	.494	-.074	.963
	T						
1962	S	1.01	.959	-1.13	.260	.107	.703
	T	2.15	.127	-1.67	.105	.267	.445
1963	S	1.16	.584	-.83	.410	-.133	.658
	T	2.94	.024*	-.78	.443	-.136	.949
1964	S	1.21	.431	-2.97	.004**	.184	.150
	T	1.07	.874	-1.04	.301	-.176	.657
1965	S	1.47	.206	-2.74	.007**	-.208	.203
	T	2.35	.265	-1.23	.242	.300	.613

Table I-4(11) Univariate comparison of Index and Visual flocks for the variate fibre diameter in microns. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		FEMALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.06	.820	-1.56	.121	-.273	.015*
	T	1.63	.202	-.59	.557	-.172	.576
1955	S	2.22	.003**	1.23	.222	-.201	.140
	T	1.62	.246	-1.00	.320	-.186	.666
1956	S	1.10	.729	-1.59	.116	-.167	.334
	T	1.41	.472	-3.23	.003**	.695	.000***
1957	S	1.68	.040*	1.79	.075	-.201	.116
	T	1.03	.924	-.41	.681	-.167	.901
1958	S	1.63	.067	-1.91	.058	-.211	.118
	T	2.04	.306	.03	.979	-.167	.907
1959	S	1.29	.273	4.91	.000***	.416	.000***
	T	1.00	.971	1.93	.058	-.233	.307
1960	S	2.11	.001**	.27	.784	-.202	.065
	T	1.88	.504	1.08	.296	.333	.560
1961	S	1.15	.535	-.57	.568	-.136	.401
	T						
1962	S	1.02	.930	-3.74	.000***	.311	.001**
	T	1.30	.627	-.38	.708	.158	.924
1963	S	1.46	.189	-3.01	.003**	-.272	.029*
	T	1.11	.806	-.95	.350	-.285	.276
1964	S	1.02	.944	.14	.892	-.096	.839
	T	1.23	.617	.67	.503	.137	.885
1965	S	1.19	.529	-.28	.780	-.113	.858
	T	1.45	.767	.12	.906	.300	.613

Table I-4(12) Univariate comparison of Index and Visual flocks for the variate coefficient of variation of fibre diameter. Parametric tests (variance ratio and t-test) and a non-parametric test (Kolmogorov-Smirnov) are presented for single born (S) and twin born (T) progeny for the 12 years 1954 to 1965.

		FEMALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.03	.885	-1.68	.096	-.231	.056
	T	1.34	.439	.19	.850	-.138	.794
1955	S	1.45	.162	-.42	.676	.141	.510
	T	1.37	.456	1.80	.079	.375	.043*
1956	S	1.03	.906	2.07	.040*	.261	.030*
	T	1.57	.404	1.41	.169	-.279	.390
1957	S	1.46	.131	-3.07	.003**	.343	.001**
	T	2.24	.102	-1.77	.085	.357	.161
1958	S	1.14	.617	-.83	.408	-.151	.439
	T	2.74	.149	.52	.605	.333	.299
1959	S	1.14	.585	-1.18	.241	-.151	.317
	T	1.81	.100	-.82	.418	-.125	.925
1960	S	1.30	.259	-.95	.343	.135	.412
	T	1.42	.561	-.37	.715	-.357	.482
1961	S	1.14	.565	-1.28	.203	.143	.348
	T						
1962	S	1.10	.678	-1.00	.319	.119	.576
	T	1.10	.836	-.40	.692	.165	.902
1963	S	1.96	.020*	-2.49	.014*	-.239	.075
	T	1.03	.965	-1.99	.054	-.349	.112
1964	S	1.09	.727	-.39	.698	.076	.966
	T	1.06	.857	-.30	.767	-.155	.790
1965	S	1.54	.156	.01	.995	.107	.895
	T	1.73	.454	.30	.626	-.200	.894

Table I-4(13) Univariate comparison of Index and Visual flocks for the variate primary follicle number per sq. cm. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for the 12 years 1954 to 1965.

		FEMALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.09	.735	.14	.889	.162	.331
	T	1.12	.768	-.15	.884	.172	.576
1955	S	1.49	.131	.48	.633	-.093	.910
	T	1.28	.548	1.14	.260	.261	.281
1956	S	1.35	.271	.20	.841	.219	.102
	T	1.49	.462	1.14	.262	-.279	.390
1957	S	1.17	.531	1.04	.301	-.166	.281
	T	1.64	.365	1.70	.099	-.286	.372
1958	S	1.09	.755	.22	.826	-.080	.974
	T	2.50	.175	-1.41	.171	-.444	.090
1959	S	1.05	.829	2.17	.031*	.203	.077
	T	1.86	.117	1.16	.250	-.267	.181
1960	S	1.09	.708	.56	.577	-.111	.647
	T	2.52	.316	1.55	.139	.571	.076
1961	S	1.36	.189	-.96	.340	.081	.922
	T						
1962	S	1.15	.561	.28	.777	.121	.559
	T	1.33	.561	.88	.386	-.375	.124
1963	S	1.06	.826	-.67	.507	-.138	.609
	T	3.22	.012*	-.12	.902	-.263	.358
1964	S	1.33	.241	.48	.632	.109	.714
	T	1.13	.778	1.08	.285	.202	.494
1965	S	1.48	.161	.97	.335	-.143	.623
	T	1.32	.670	-.34	.738	.200	.894

Table I-4(14) Univariate comparison of Index and Visual flocks for the variate secondary follicle number per sq. cm. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for the 12 years 1954 to 1965.

		FEMALE PROGENY					
Year of birth		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1954	S	1.07	.791	1.09	.280	.256	.026*
	T	1.24	.566	-.46	.650	.103	.951
1955	S	1.45	.160	.39	.697	-.108	.805
	T	1.15	.723	.74	.461	.229	.423
1956	S	1.08	.750	1.89	.061	.243	.052
	T	1.94	.219	1.56	.128	.338	.200
1957	S	1.20	.482	2.59	.011*	.264	.016*
	T	1.22	.665	.25	.806	-.214	.697
1958	S	1.10	.726	1.20	.234	.162	.354
	T	2.70	.155	-.64	.527	-.278	.484
1959	S	1.39	.158	.92	.361	.138	.419
	T	1.78	.110	.74	.464	-.225	.347
1960	S	1.33	.211	.81	.418	-.119	.566
	T	1.12	.791	.01	.990	-.262	.802
1961	S	1.11	.668	2.03	.044*	.171	.169
	T						
1962	S	1.68	.024*	5.02	.000***	.374	.000***
	T	1.15	.766	2.25	.032*	-.414	.071
1963	S	1.16	.609	.03	.978	.092	.947
	T	1.45	.431	-1.20	.236	.273	.320
1964	S	1.30	.273	1.08	.284	-.140	.424
	T	1.91	.084	1.45	.151	.193	.552
1965	S	1.24	.431	.83	.411	.159	.493
	T	1.12	.990	.54	.601	-.400	.339

Table I-4(15) Univariate comparison of Index and Visual flocks for the variate skin thickness in cm. Parametric tests (variance ratio and t-test) and a nonparametric test (Kolmogorov-Smirnov) are presented for the 8 years 1958 to 1965.

Year of birth		FEMALE PROGENY					
		F Value	2-tail Prob.	T Value	2-tail Prob.	Kol-Smir.	2-tail Prob.
1958	S	1.45	.163	1.69	.094	.180	.245
	T	1.30	.612	1.10	.101	.333	.299
1959	S	1.27	.314	-.00	.998	.125	.538
	T	1.11	.753	1.96	.054	.317	.072
1960	S	1.23	.367	.08	.939	-.065	.988
	T	1.81	.358	1.02	.319	.333	.560
1961	S	1.51	.069	3.47	.001**	.281	.003**
	T						
1962	S	1.79	.011*	2.78	.006**	.265	.007**
	T	1.96	.181	1.55	.132	-.442	.046*
1963	S	1.20	.486	1.39	.168	.207	.167
	T	1.99	.146	.13	.898	.213	.602
1964	S	1.08	.736	3.98	.000***	-.345	.000***
	T	1.56	.233	2.68	.010*	.352	.040*
1965	S	1.95	.017*	3.74	.000***	.336	.006**
	T	1.13	.980	.73	.481	-.400	.339

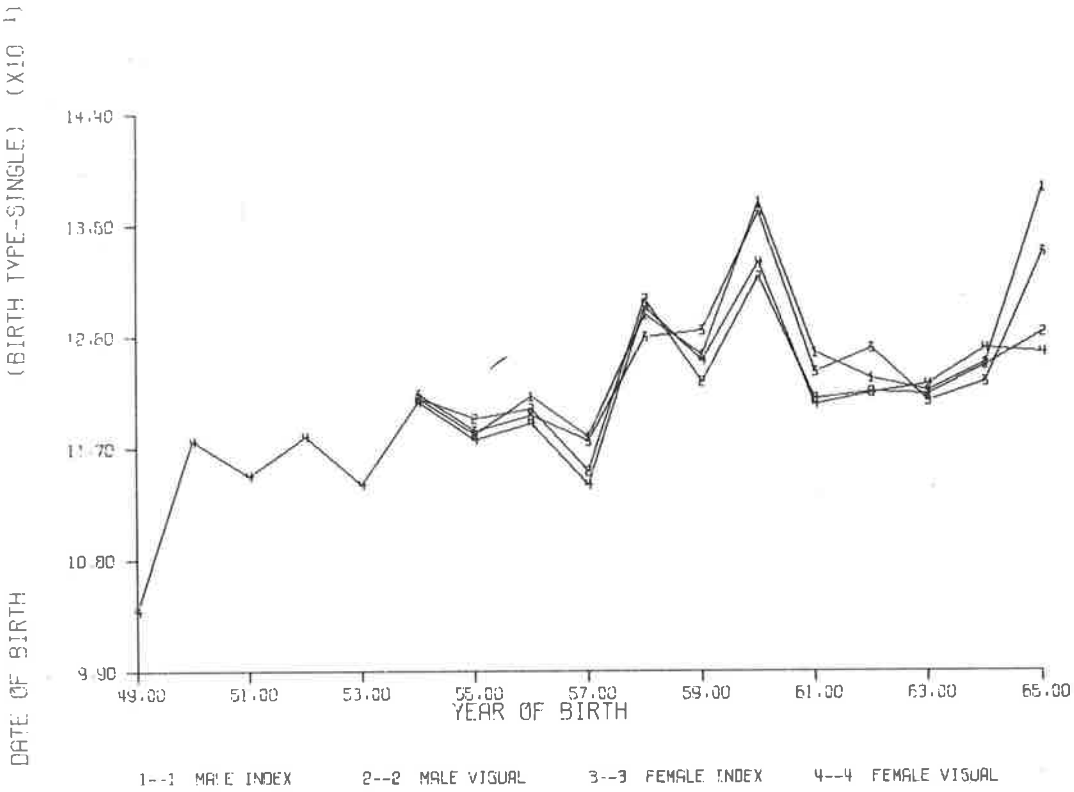
Table I-5: Frequency of rejection (when $\alpha = 0.05$) for the three two sample tests considered in Tables I-3(1) to (15) and I-4(1) to (15).

	Male Progeny			Female Progeny			
Variates	Variance ratio	t-test	Kolmogorov-Smirnov	Variance ratio	t-test	Kolmogorov-Smirnov	Total No. of tests
Date of birth	8	6 (3)†	4	9	7 (5)†	4	23
Birth weight in kg	4	4 (3)	1	3	1 (1)	1	23
Weaner body weight in kg	2	3 (3)	3	1	2 (2)	3	23
Hogget body weight in kg	2	2 (2)	2	4	6 (5)	4	23
Lamb fleece weight in kg	2	2 (2)	3	5	2 (1)	4	23
Hogget greasy fleece weight in kg	2	7 (6)	7	1	11 (11)	7	23
Percentage yield	1	6 (6)	5	2	8 (7)	6	23
Clean fleece weight in kg	2	11 (9)	8	1	13 (12)	12	23
Staple length in cm	4	10 (8)	6	3	7 (6)	3	23
Crimps per inch	2	7 (6)	2	5	7 (5)	1	23
Fibre diameter in microns	4	6 (6)	7	3	4 (4)	5	23
Coefficient of variation of fibre diameter	4	5 (5)	3	1	3 (2)	3	23
Primary follicle number per sq cm	3	5 (4)	4	1	1 (1)	0	23
Secondary follicle number per sq cm	1	5 (4)	4	1	4 (3)	3	23
Skin thickness in cm	1	3 (2)	2	2	5 (3)	6	15
Total	42	82 (69)	61	42	81 (68)	62	337 (295)
Percentage significance	12.5	24.3(23.4)	18.1	12.5	24.4(23.1)	18.2	

† Numbers in brackets indicate the number of significant t-tests given homogeneity of variances

Figure I-1(1) Plot of mean date of birth, for the period 1949 to 1965, with the four flocks by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-1(2) Plot of mean birth weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


DATE OF BIRTH (BIRTH TYPE-SINGLE) (X10 1)
14.40
13.50
12.60
11.70
10.80
9.90
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

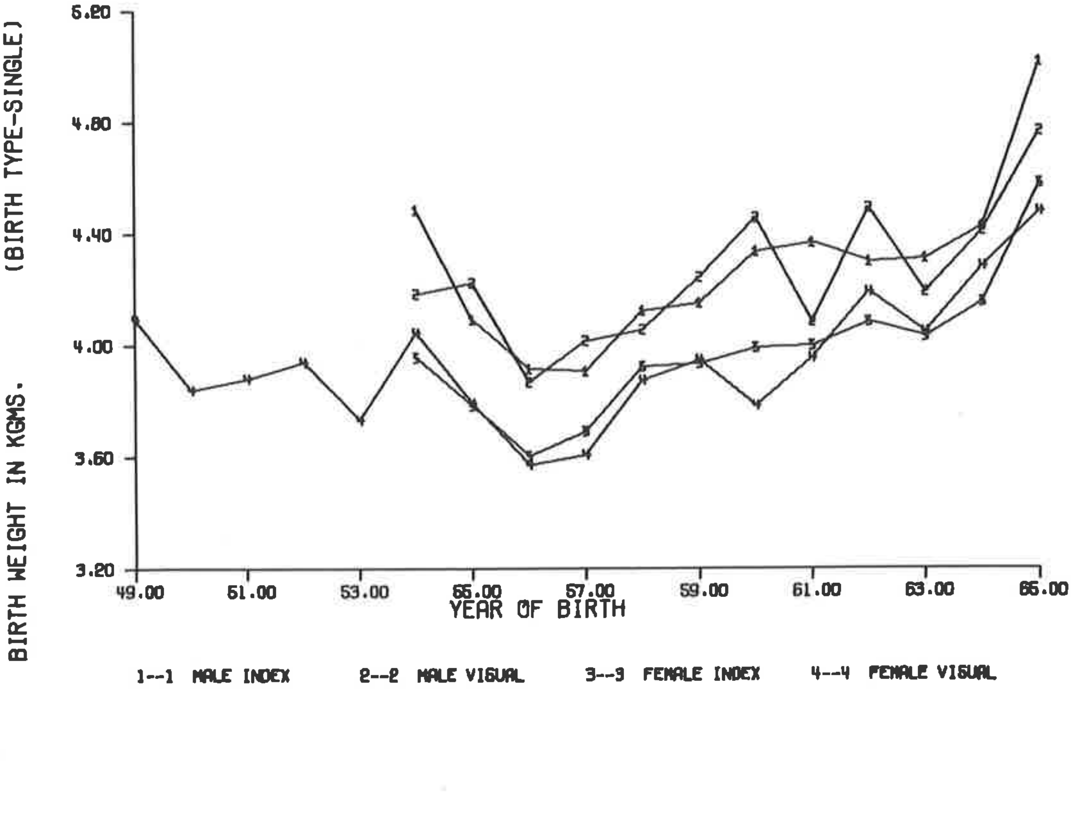
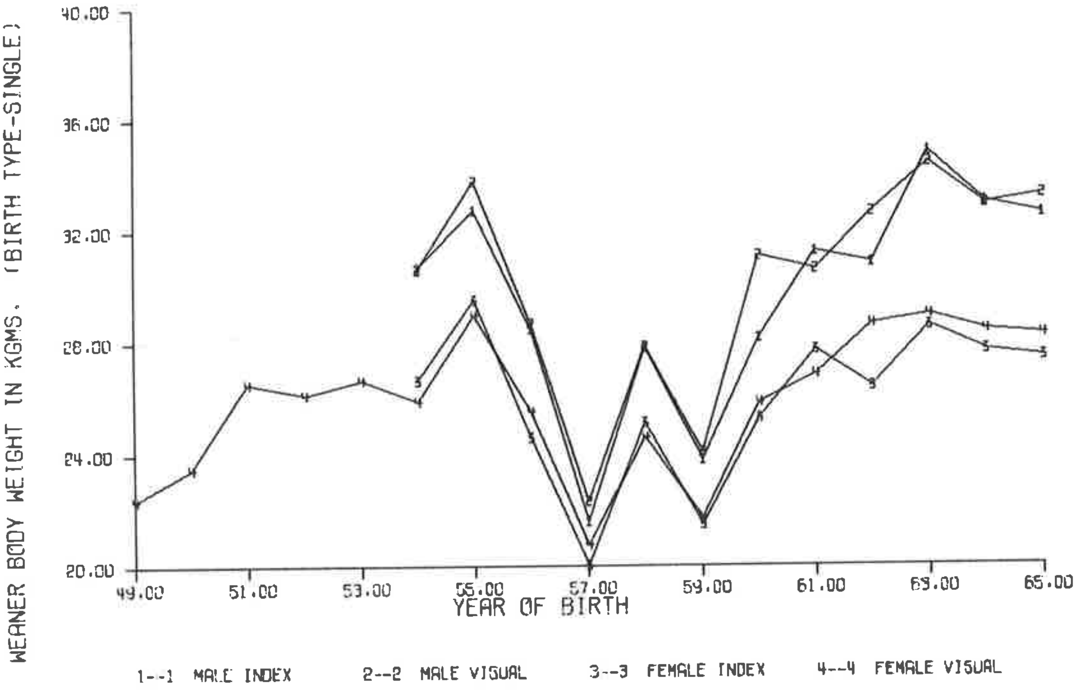

BIRTH WEIGHT IN KGMS. (BIRTH TYPE-SINGLE)
5.20
4.80
4.40
4.00
3.60
3.20
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-1(3) Plot of weaner body weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-1(4) Plot of mean hogget body weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


WEANER BODY WEIGHT IN KGMS. (BIRTH TYPE-SINGLE)
40.00
36.00
32.00
28.00
24.00
20.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

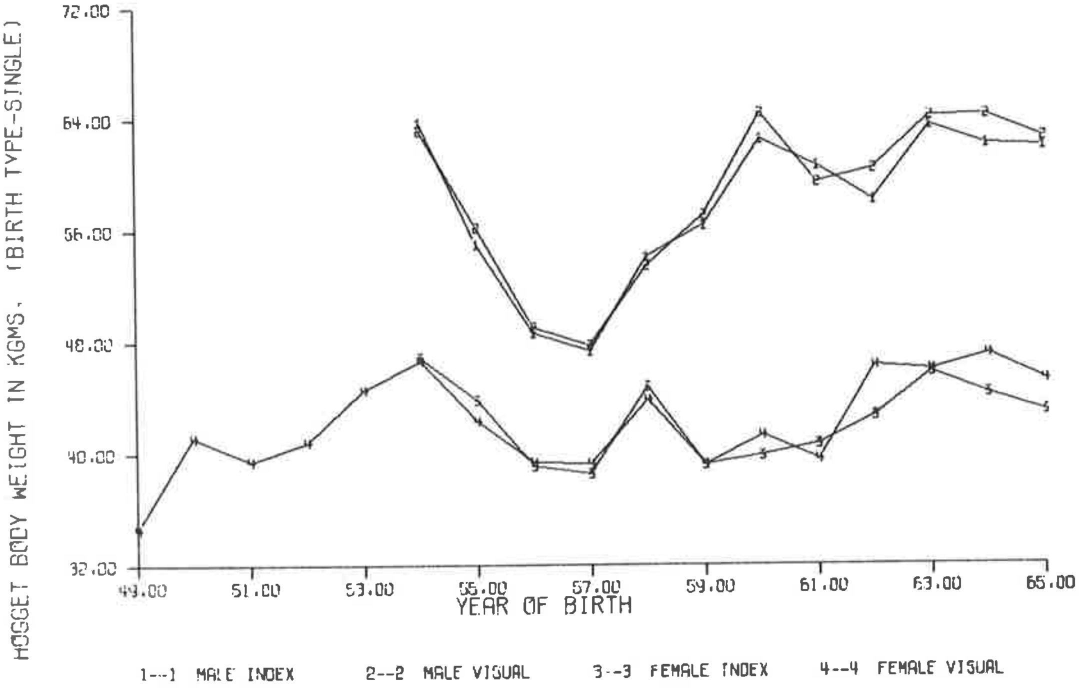
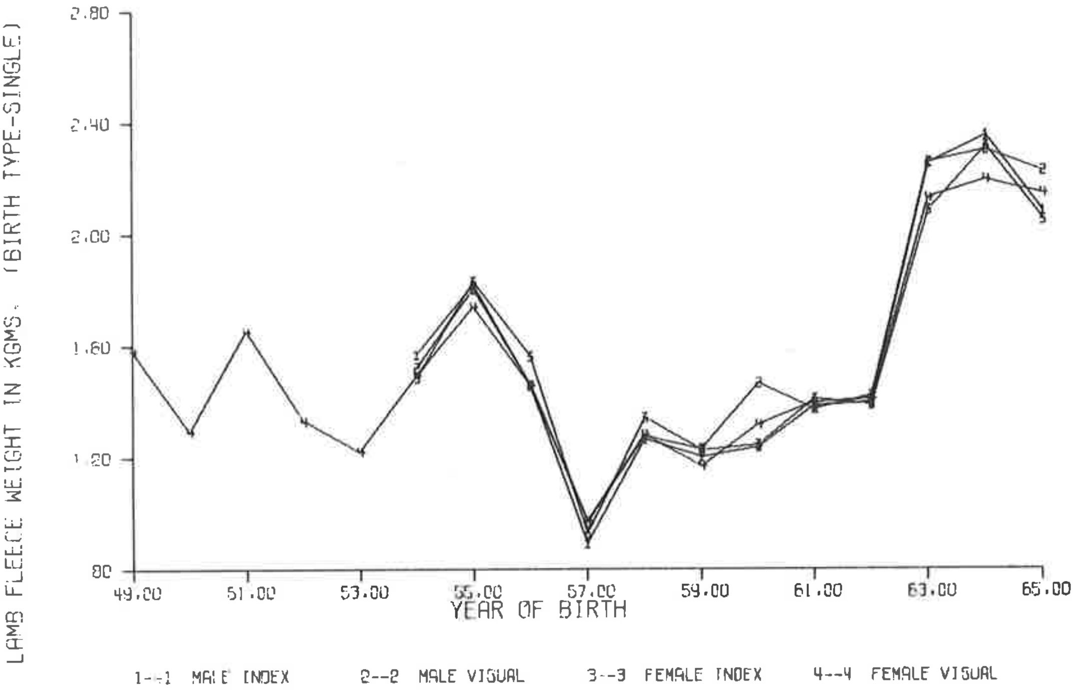

HOGGET BODY WEIGHT IN KGMS. (BIRTH TYPE-SINGLE)
72.00
64.00
56.00
48.00
40.00
32.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-1(5) Plot of mean lamb fleece weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-1(6) Plot of mean hogget greasy fleece weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


LAMB FLEECE WEIGHT IN KGMS. (BIRTH TYPE-SINGLE)
2.80
2.40
2.00
1.60
1.20
80
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

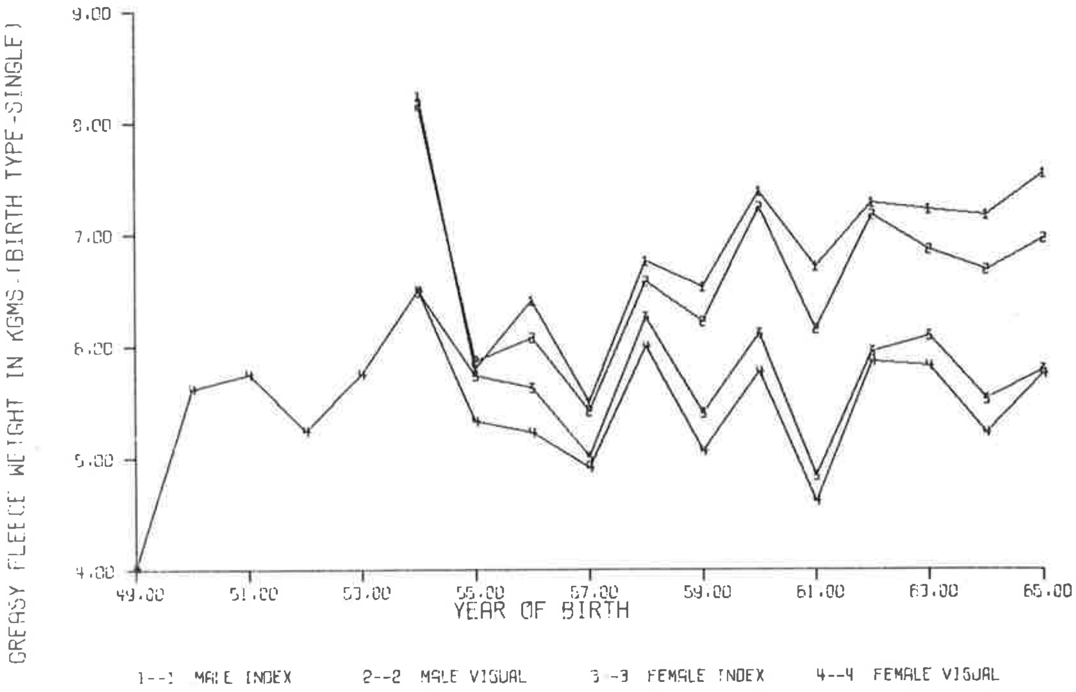
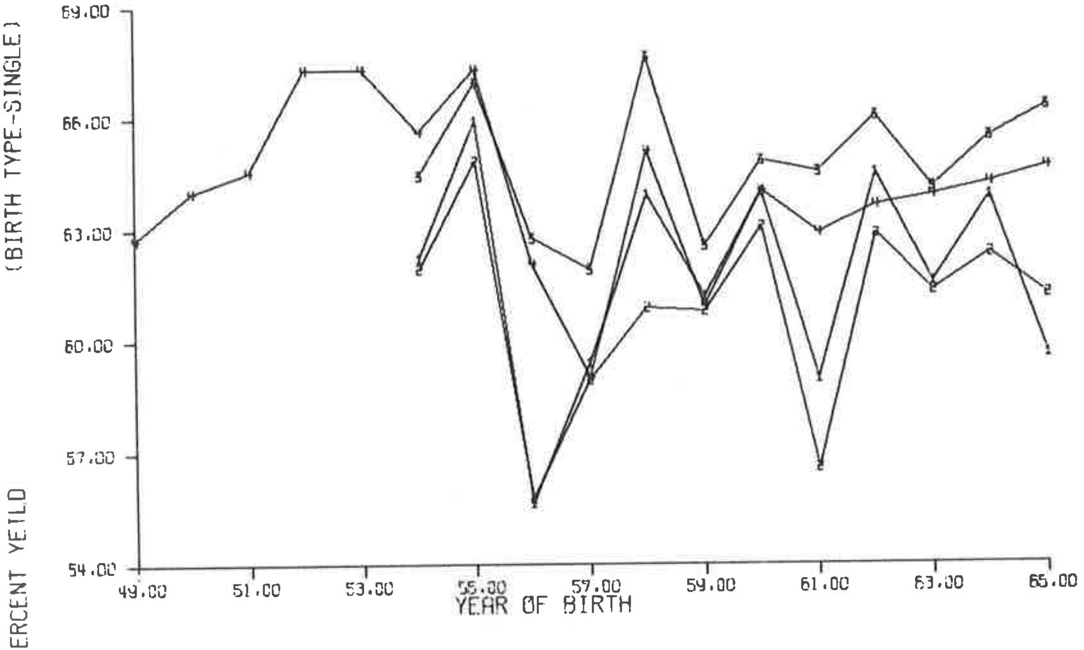

GREASY FLEECE WEIGHT IN KGMS.(BIRTH TYPE-SINGLE)
9.00
8.00
7.00
6.00
5.00
4.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-1(7) Plot of mean percentage yield for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-1(8) Plot of mean clean fleece weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


69.00
66.00
63.00
60.00
57.00
54.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
PERCENT YEILD (BIRTH TYPE-SINGLE)
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

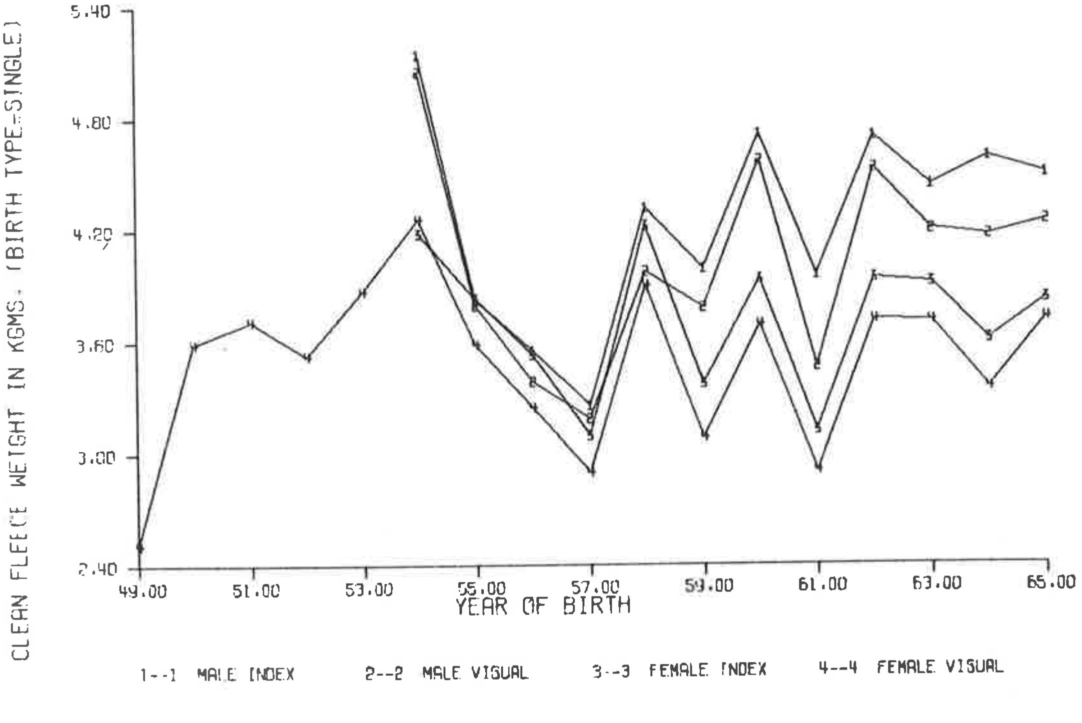
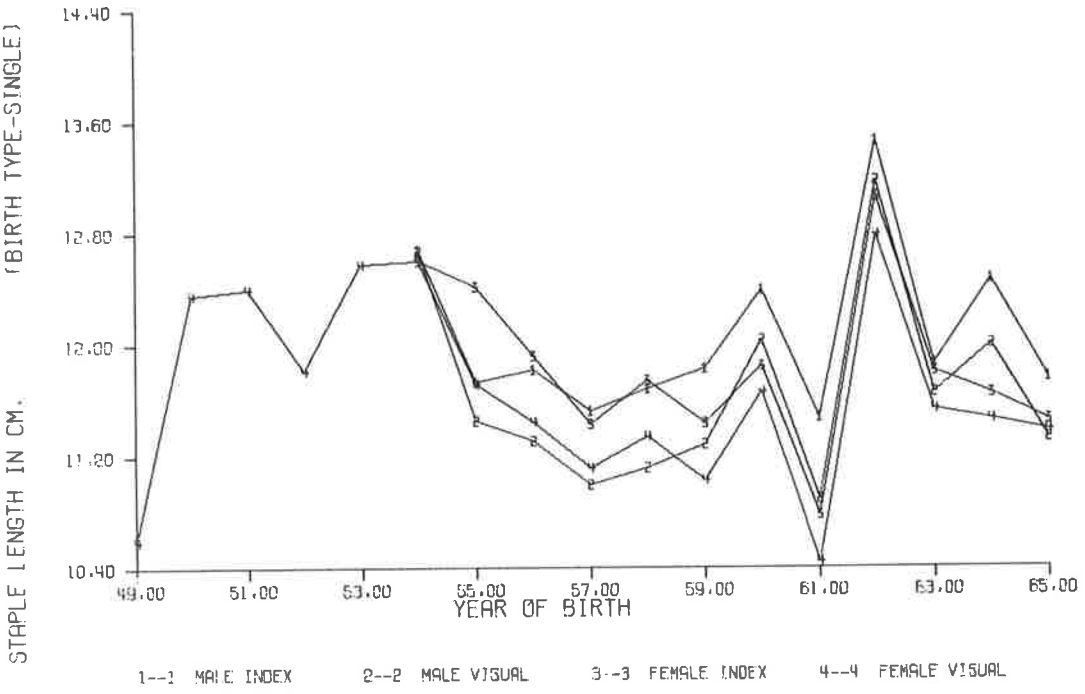

5.40
4.80
4.20
3.60
3.00
2.40
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
CLEAN FLEECE WEIGHT IN KGMS. (BIRTH TYPE-SINGLE)
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-1(9) Plot of mean staple length in cm, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-1(10) Plot of mean crimps per inch, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


STAPLE LENGTH IN CM.
(BIRTH TYPE-SINGLE)
14.40
13.60
12.80
12.00
11.20
10.40
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

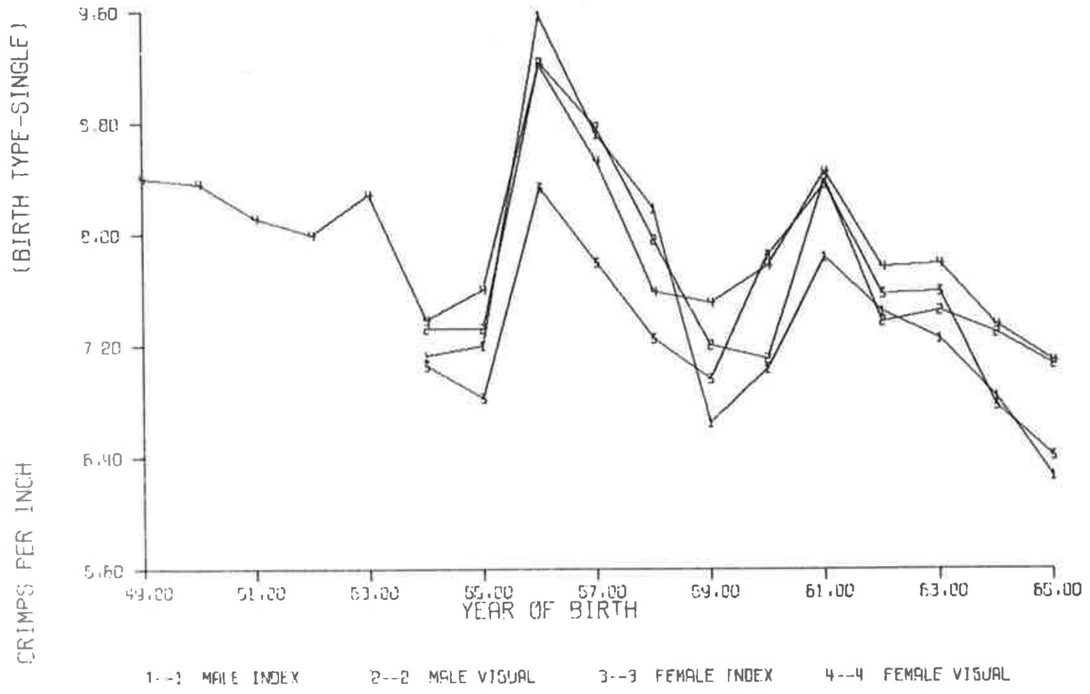
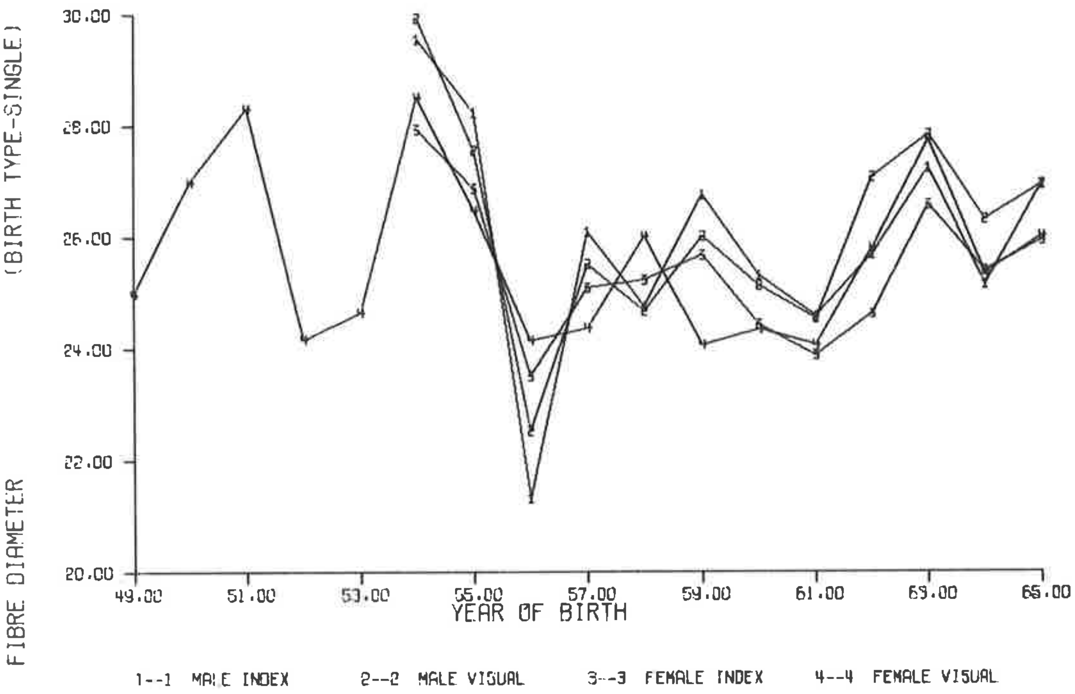
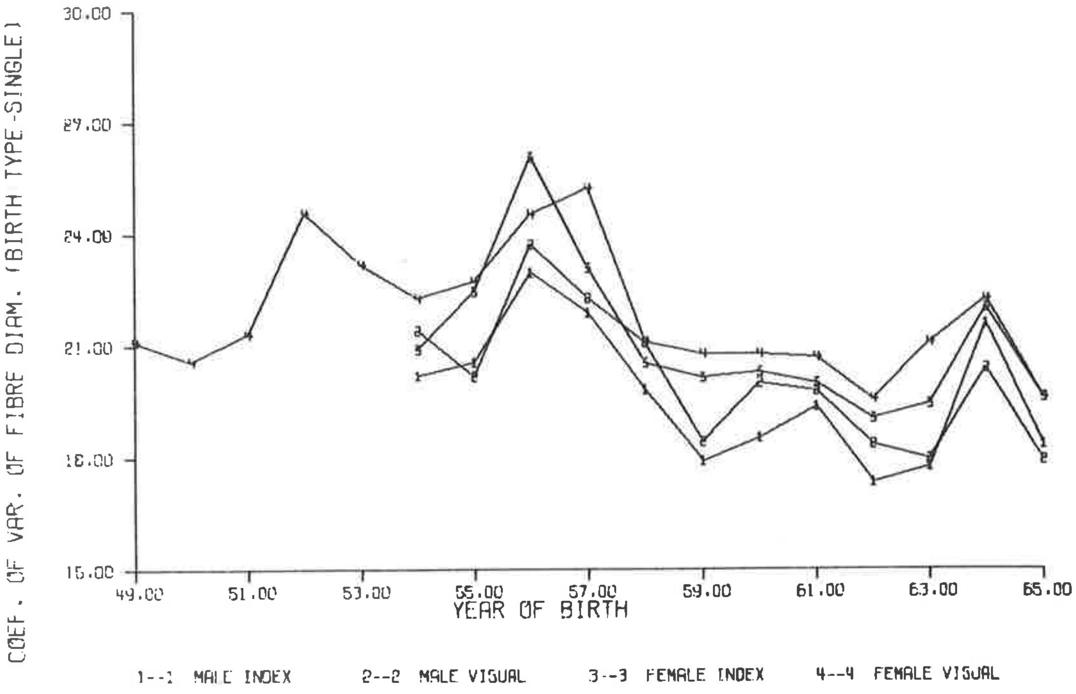

CRIMPS PER INCH
(BIRTH TYPE-SINGLE)
9.60
8.80
8.00
7.20
6.40
5.60
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-1(11) Plot of mean fibre diameter in microns, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-1(12) Plot of mean coefficient of variation of fibre diameter, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)



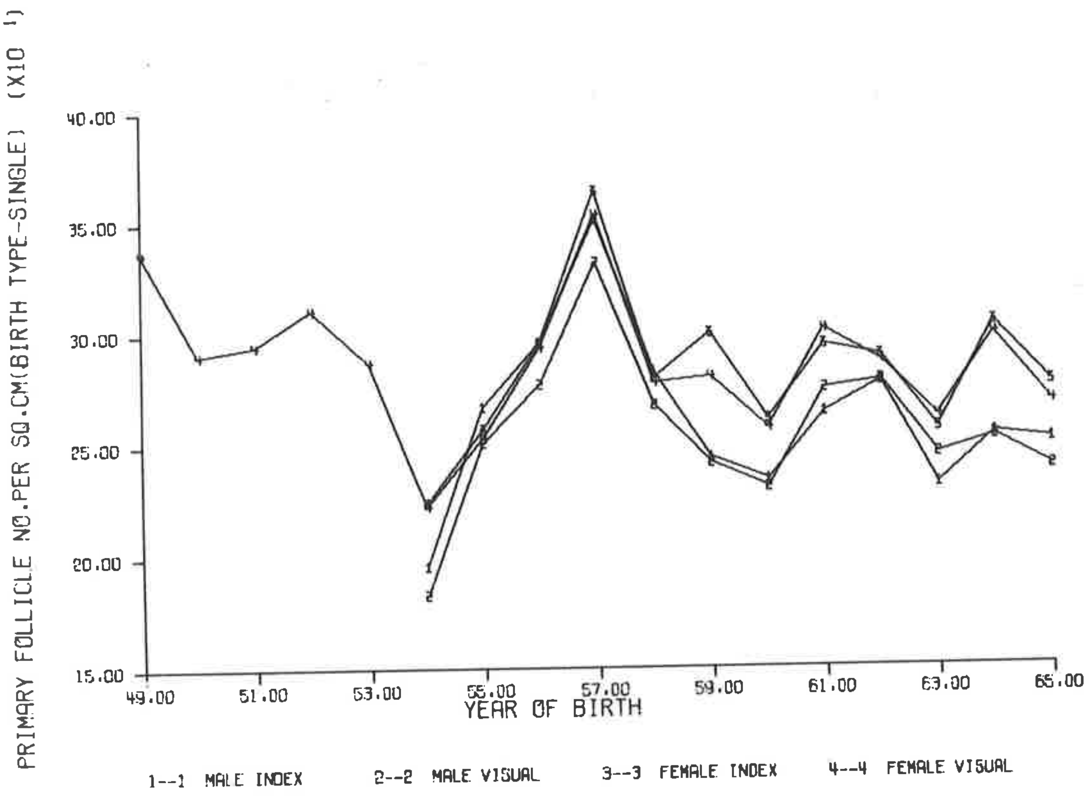
FIBRE DIAMETER (BIRTH TYPE-SINGLE)
30.00
28.00
26.00
24.00
22.00
20.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL



COEF. OF VAR. OF FIBRE DIAM. (BIRTH TYPE-SINGLE)
30.00
27.00
24.00
21.00
18.00
15.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-1(13) Plot of mean primary follicle number per sq. cm, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-1(14) Plot of mean secondary follicle number per sq. cm, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


PRIMARY FOLLICLE NO.PER SQ.CM(BIRTH TYPE-SINGLE) (X10 1)
40.00
35.00
30.00
25.00
20.00
15.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

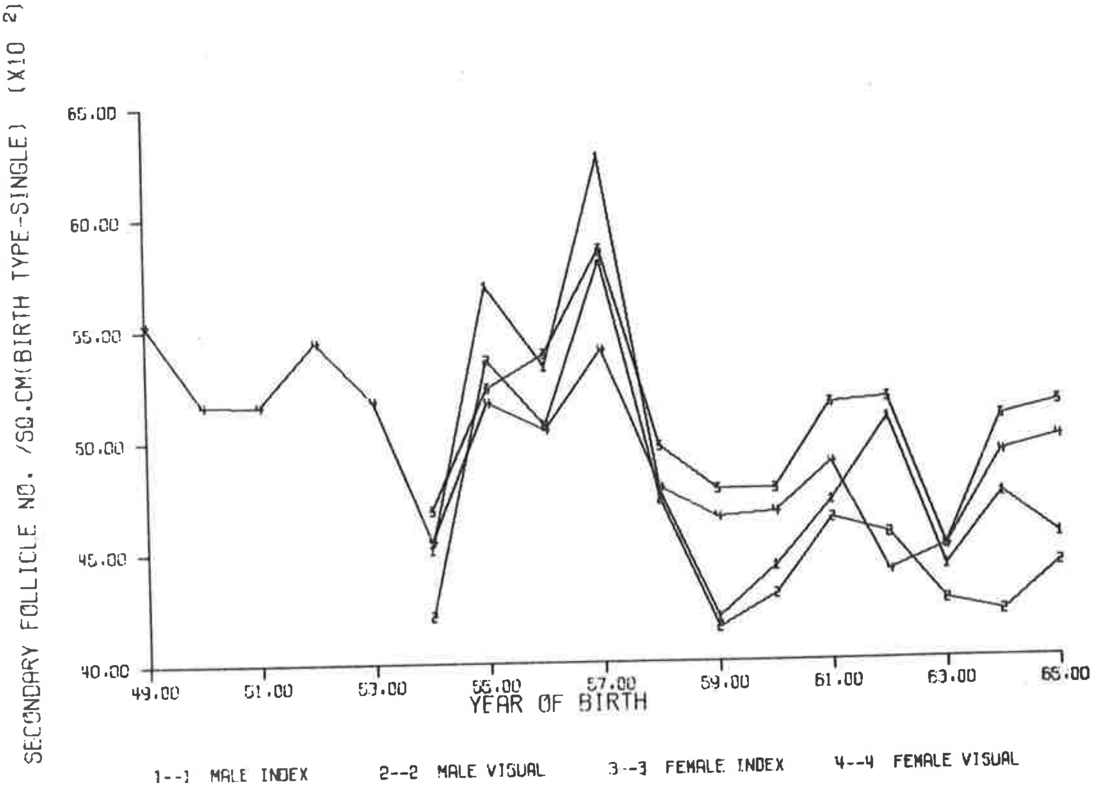
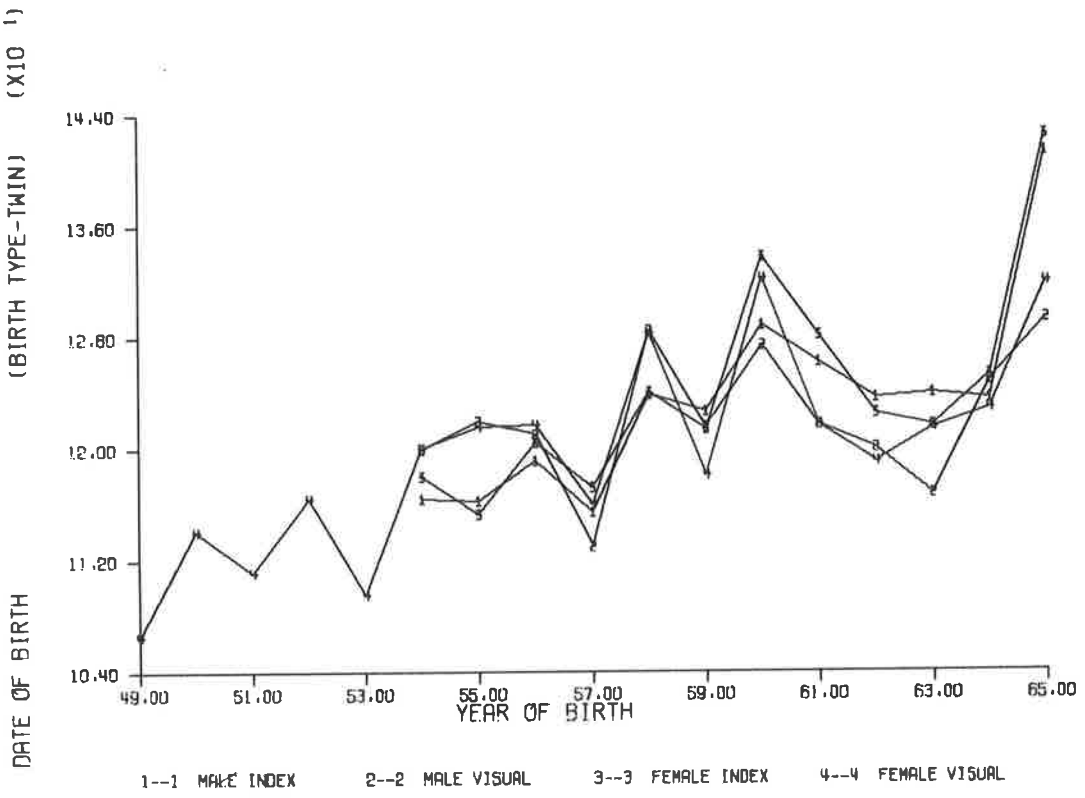

SECONDARY FOLLICLE NO. /SQ.CM(BIRTH TYPE-SINGLE) (X10 2)
65.00
60.00
55.00
50.00
45.00
40.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-2(1) Plot of mean date of birth, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-2(2) Plot of mean birth weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


DATE OF BIRTH (BIRTH TYPE-TWIN) (X10 ¹)
14.40
13.60
12.80
12.00
11.20
10.40
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

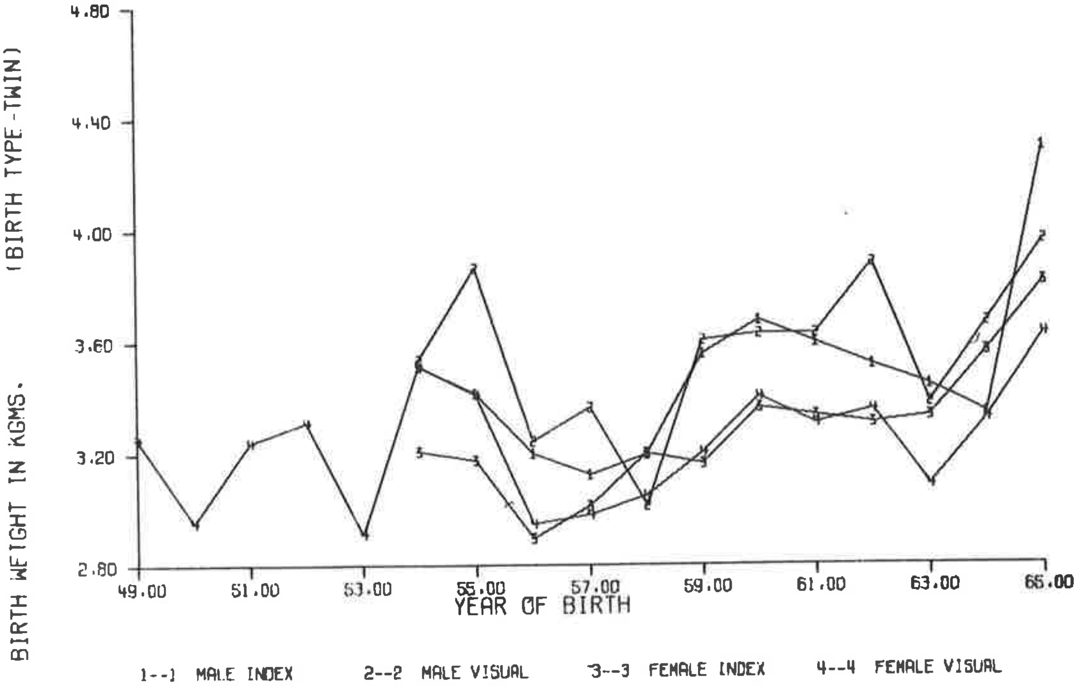
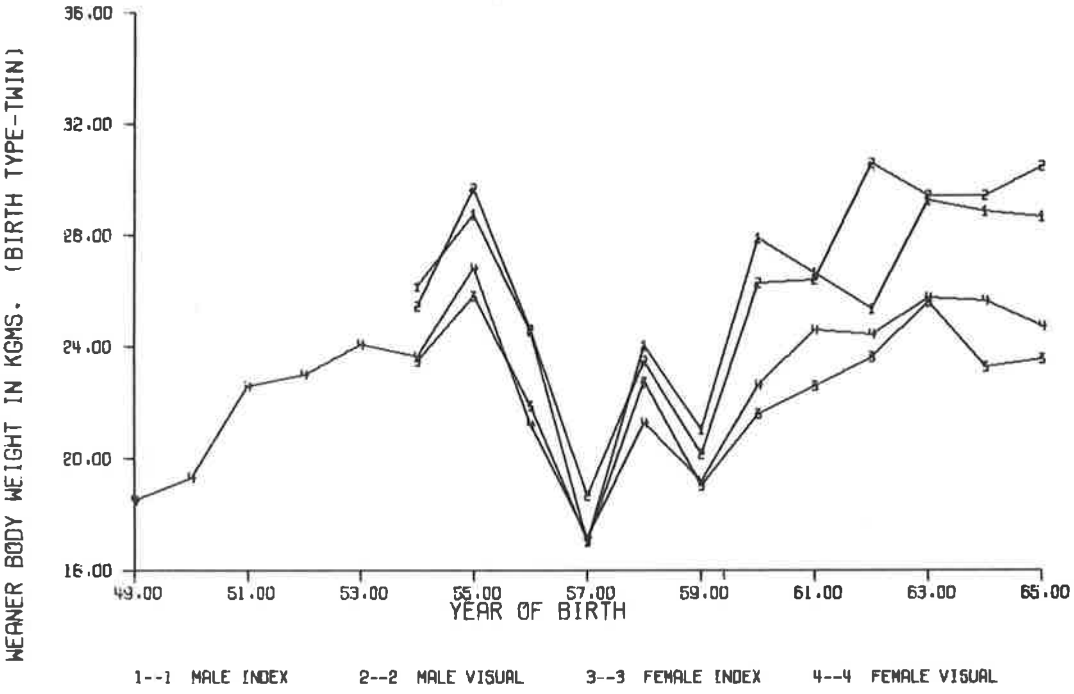

BIRTH WEIGHT IN KGMS. (BIRTH TYPE-TWIN)
4.80
4.40
4.00
3.60
3.20
2.80
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-2(3) Plot of mean weaner body weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-2(4) Plot of mean hogget body weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


36.00
32.00
28.00
24.00
20.00
16.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
WEANER BODY WEIGHT IN KGMS. (BIRTH TYPE-TWIN)
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

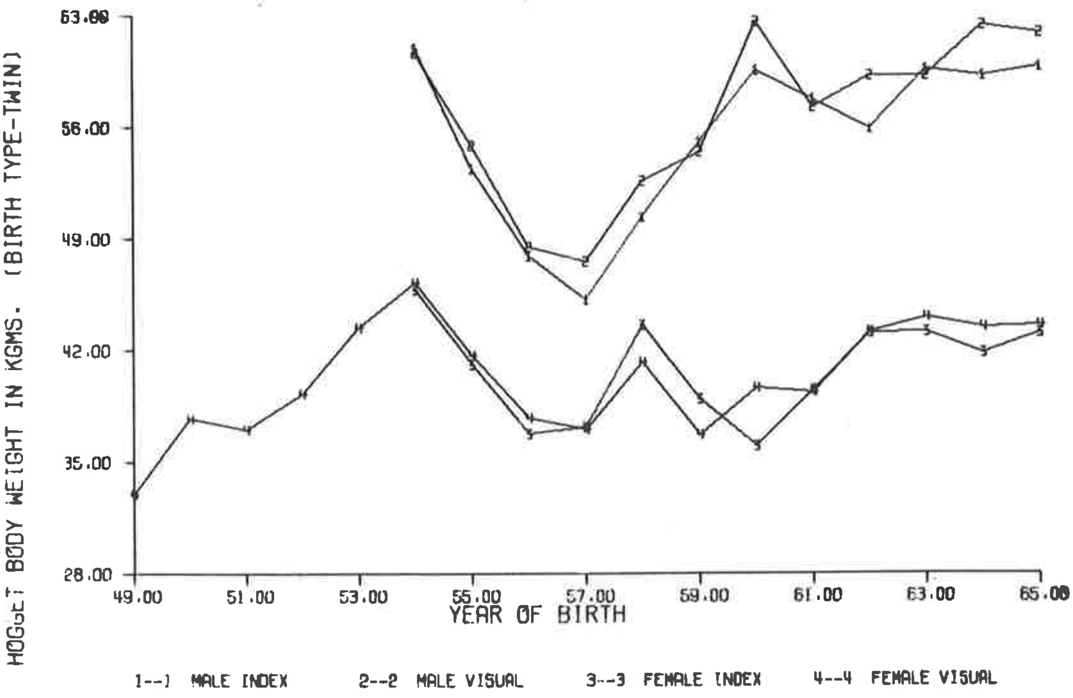
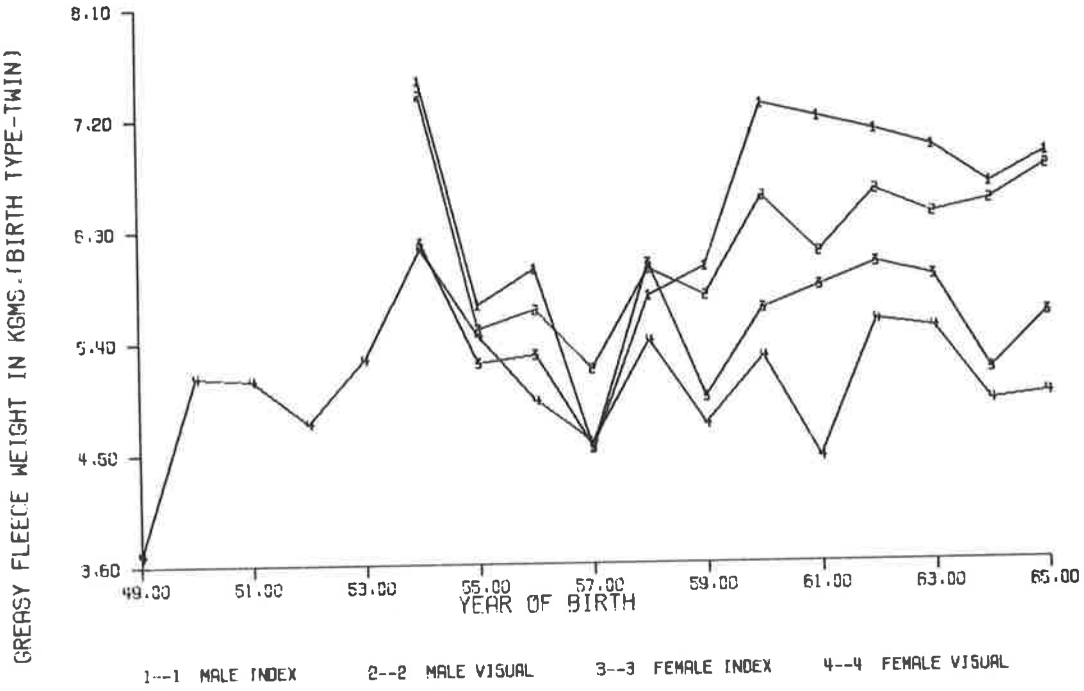

63.00
56.00
49.00
42.00
35.00
28.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
HOGGET BODY WEIGHT IN KGMS. (BIRTH TYPE-TWIN)
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-2(6) Plot of mean hogget greasy fleece weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-2(5) Plot of mean lamb fleece weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


GREASY FLEECE WEIGHT IN KGMS.(BIRTH TYPE-TWIN)
8.10
7.20
6.30
5.40
4.50
3.60
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

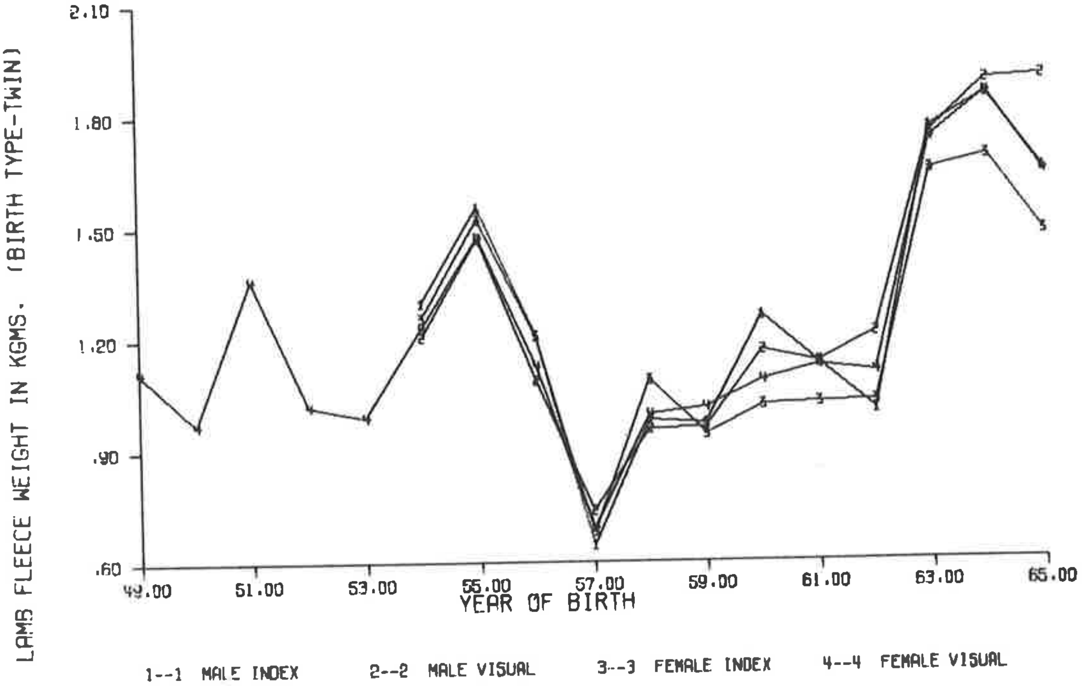
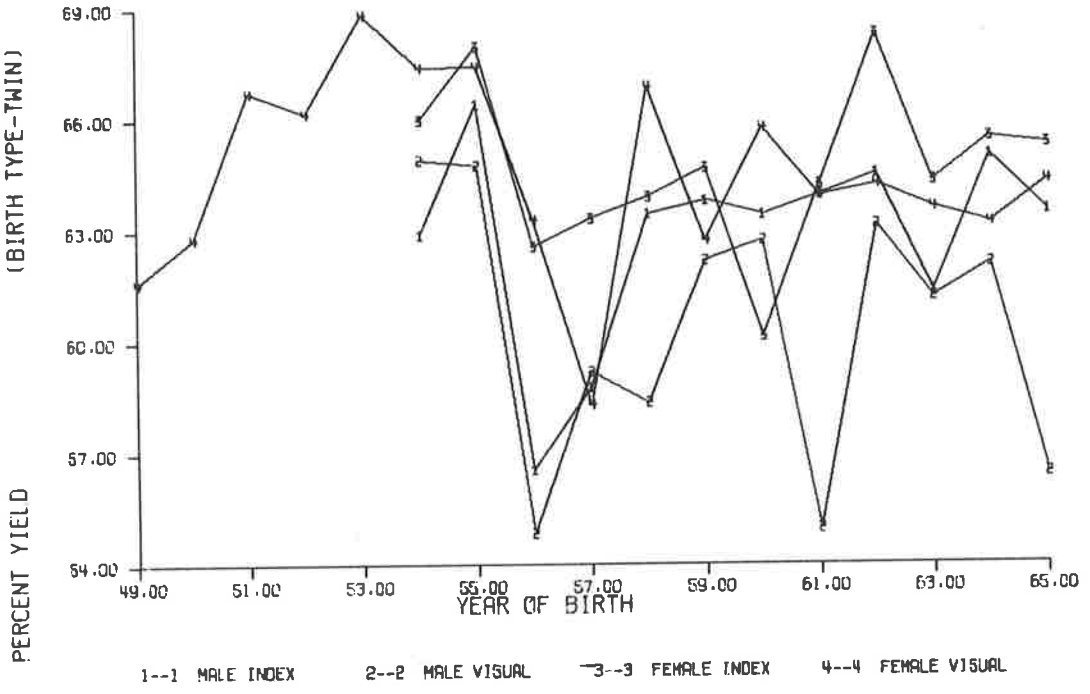

LAMB FLEECE WEIGHT IN KGMS. (BIRTH TYPE-TWIN)
2.10
1.80
1.50
1.20
.90
.60
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-2(7) Plot of mean percentage yield, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-2(8) Plot of mean clean fleece weight in kg, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


69.00
66.00
63.00
60.00
57.00
54.00
PERCENT YIELD (BIRTH TYPE-TWIN)
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

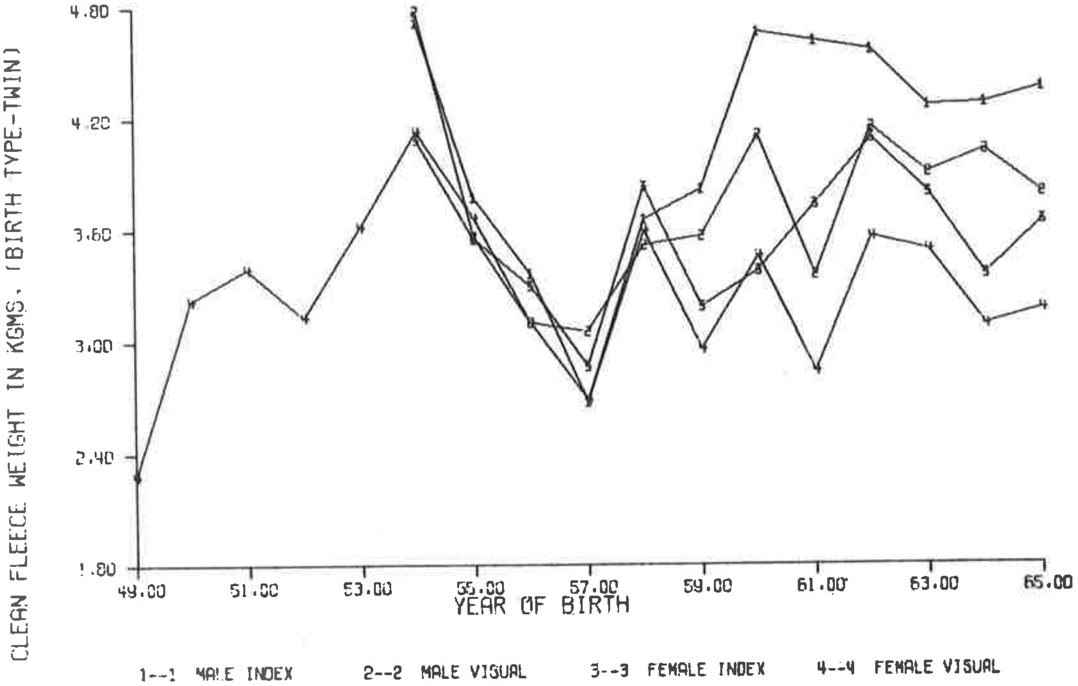
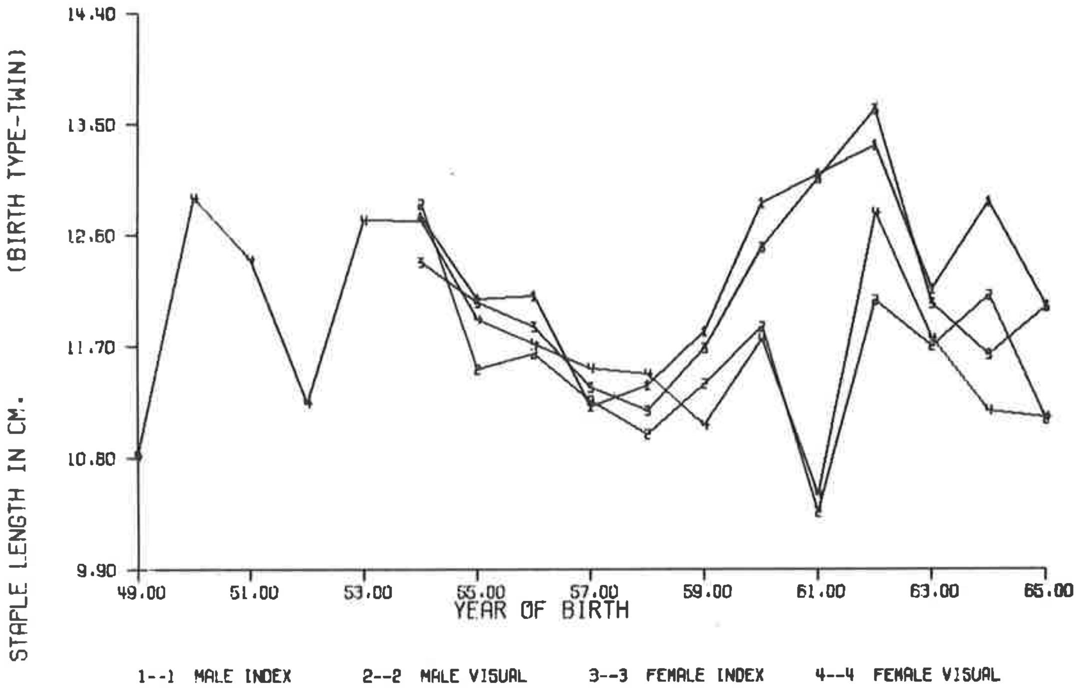

4.80
4.20
3.60
3.00
2.40
1.80
CLEAN FLEECE WEIGHT IN KGMS. (BIRTH TYPE-TWIN)
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-2(9) Plot of mean staple length in cm, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-2(10) Plot of mean crimps per inch, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


14.40
13.50
12.60
11.70
10.80
9.90
STAPLE LENGTH IN CM.
(BIRTH TYPE-TWIN)
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

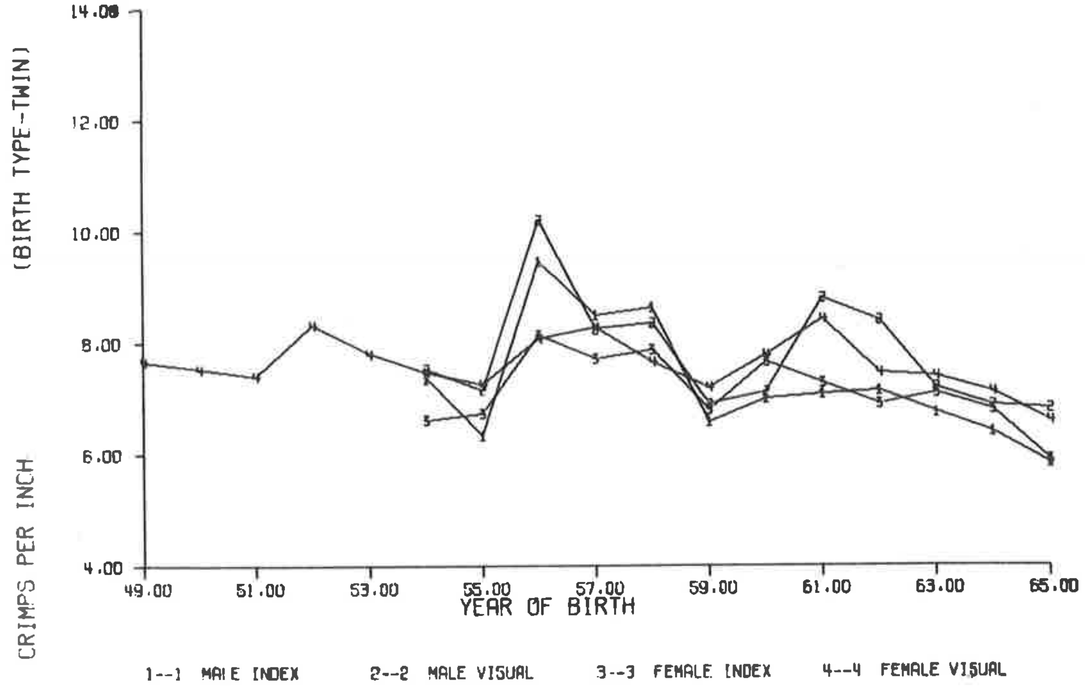
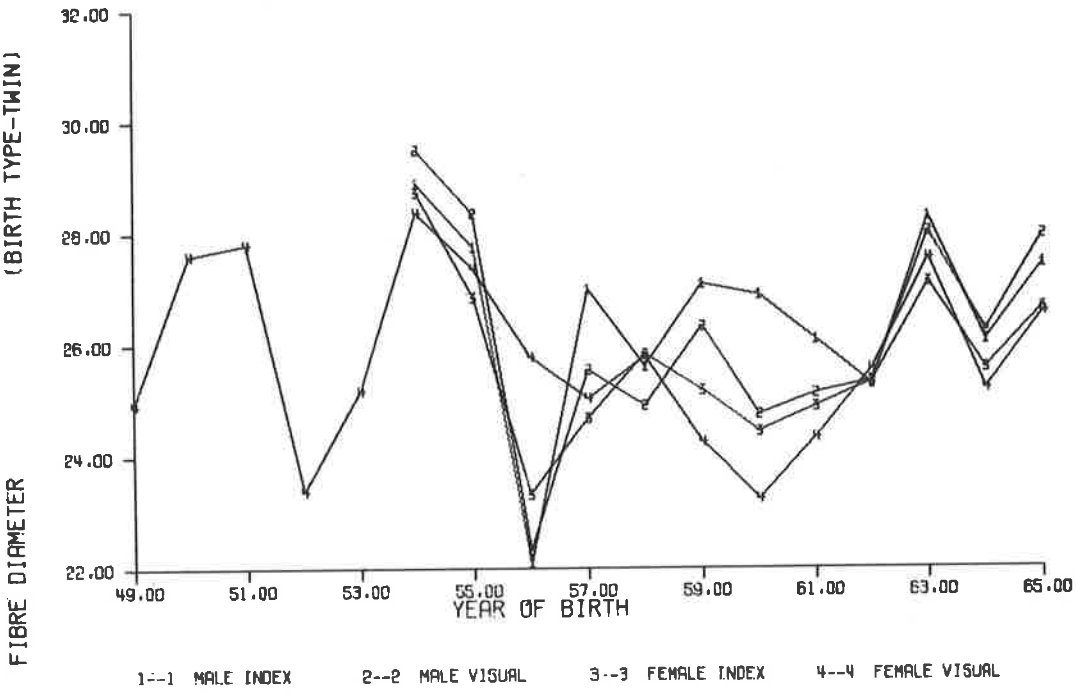

14.00
12.00
10.00
8.00
6.00
4.00
CRIMPS PER INCH
(BIRTH TYPE-TWIN)
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-2(11) Plot of mean fibre diameter in microns, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-2(12) Plot of mean coefficient of variation of fibre diameter, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


FIBRE DIAMETER (BIRTH TYPE-TWIN)
32.00
30.00
28.00
26.00
24.00
22.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

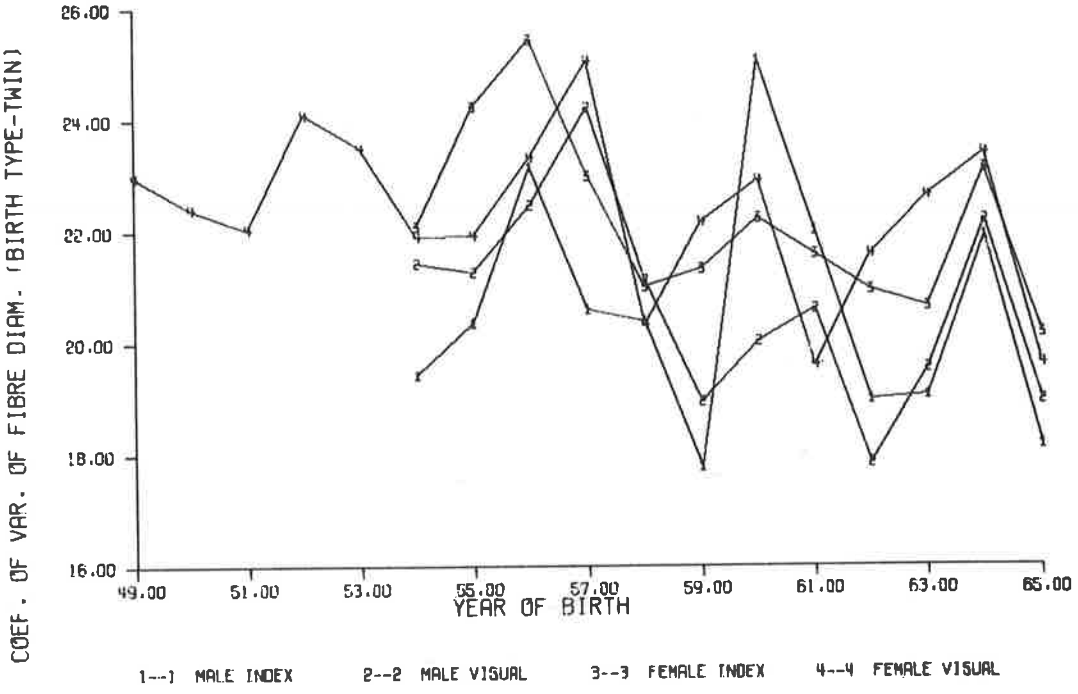
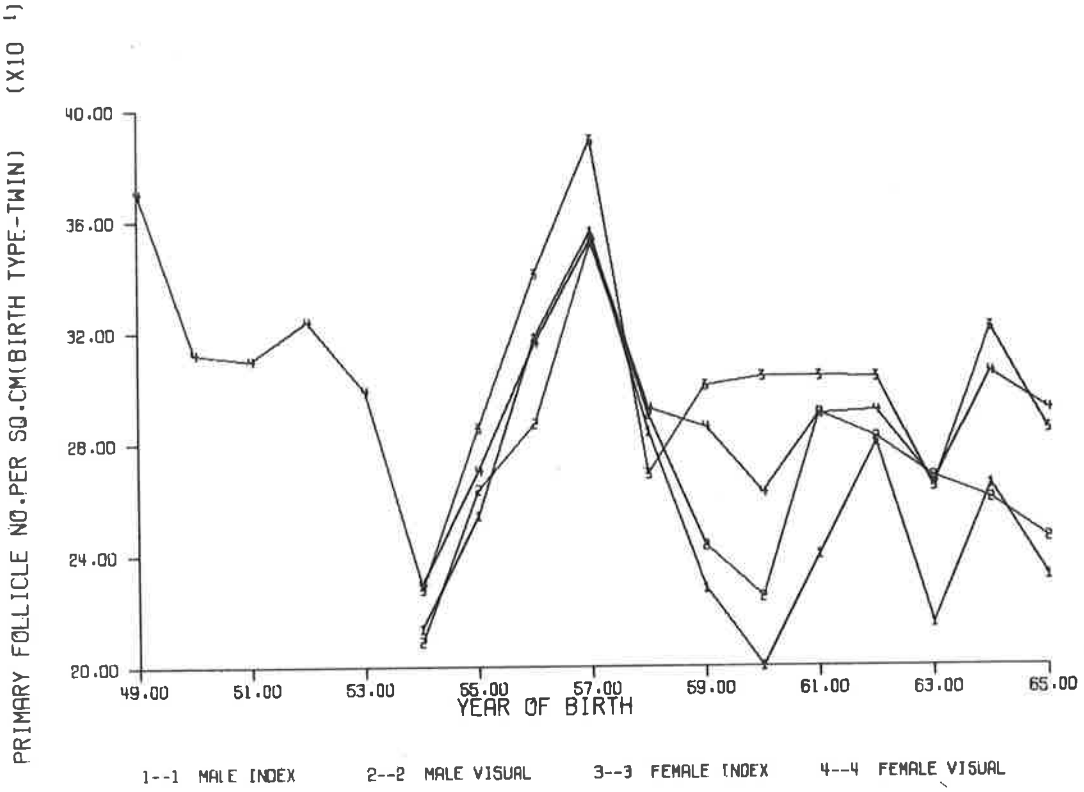
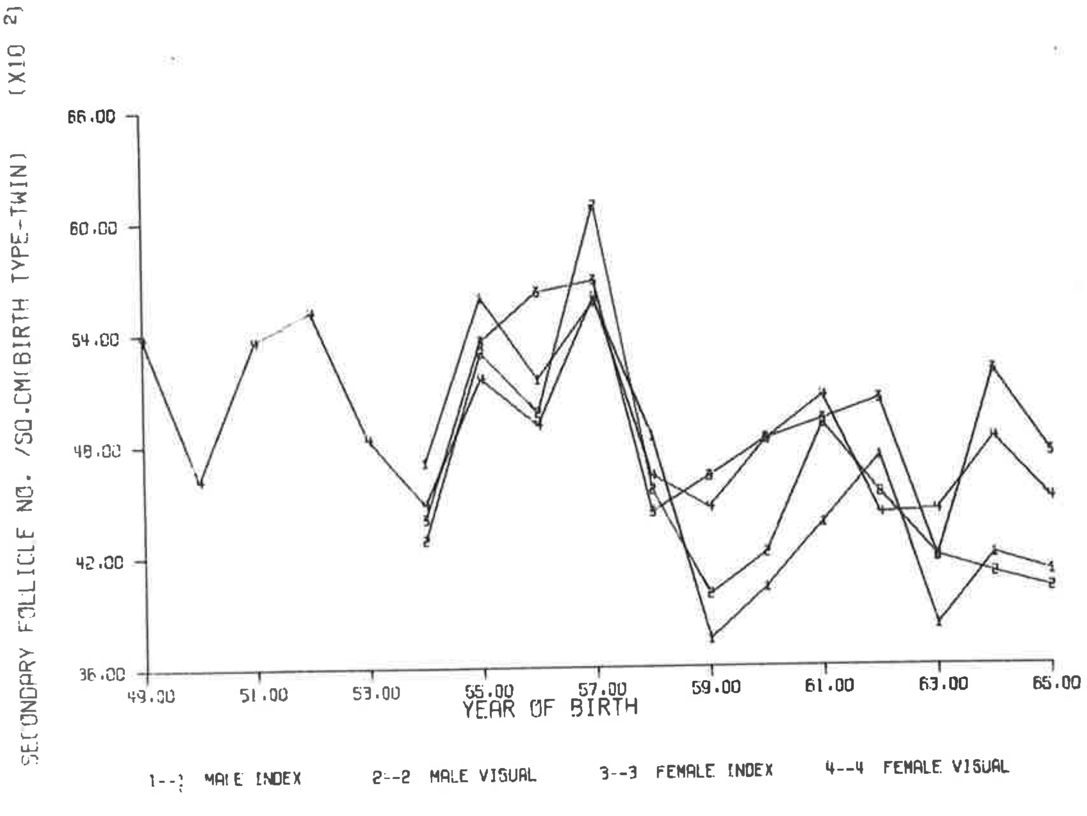

COEF. OF VAR. OF FIBRE DIAM. (BIRTH TYPE-TWIN)
26.00
24.00
22.00
20.00
18.00
16.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

Figure I-2(13) Plot of mean primary follicle number per sq. cm, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)

Figure I-2(14) Plot of mean secondary follicle number per sq. cm, for the period 1949 to 1965, with the four flock by sex combinations indicated. (For ease of plotting the dams born in 1949-53 have been given the same identification as the female Visual group.)


PRIMARY FOLLICLE NO.PER SQ.CM(BIRTH TYPE-TWIN) (X10 1)
40.00
36.00
32.00
28.00
24.00
20.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL


SECONDARY FOLLICLE NO. /SQ.CM(BIRTH TYPE-TWIN) (X10 2)
66.00
60.00
54.00
48.00
42.00
36.00
49.00
51.00
53.00
55.00
57.00
59.00
61.00
63.00
65.00
YEAR OF BIRTH
1--1 MALE INDEX
2--2 MALE VISUAL
3--3 FEMALE INDEX
4--4 FEMALE VISUAL

II.- GENETIC PARAMETRS

INTRODUCTION

The objective of quantitative genetics is to determine from observation of the individual members of a population, how the population should respond to artificial selection. That is measurements on the individual phenotypes are used to contribute information about the net effect of the polygenes underlying the particular variate. This information is generally summarized by the calculation of the following four statistics:

heritability, phenotypic correlation, genetic correlation and environmental correlation.

Falconer (1960), Kempthorne (1957) and others have shown that the partitioning of variances and covariances can be extended beyond these four statistics. The details of their argument will not be repeated here although mention will be made of some of the components especially from the point of view of how they may affect the four statistics considered. It is important however to realize that estimation of these more complicated components requires much more complicated experimental design than are generally encountered in breeding programs. (In particular, the Roseworthy experiment is unsuited to such complex analyses.)

In order to establish a successful selection program, the breeder must be able to make adequate predictions of the genetic worth of each member of the population to be selected. This is best accomplished in terms of the breeding value which Falconer (1960) defines as follows:

> "The breeding value of an individual is equal to the sum of the average effects of the genes it carries, the summation being made over the pair of alleles at each locus and over all loci."

Thus, the mean deviation of an individual's progeny estimates half the breeding value of that individual. It is important to note that the breeding value is a property of the population of potential mates and the individual considered.

In simple terms one can imagine that the phenotypic value (P) can be divided into two parts, the genotypic value (G) and the environmental deviation (E), i.e. $P = G + E$. Here, the genotypic value denotes the proportion of the phenotype which results from the combined effect of the genes carried by the individual, and the environmental deviation is the remaining non-genetic component. This latter component can be either positive or negative. This model can then be extended to cover the corresponding variance components,

i.e. $$V_P = V_G + V_E \qquad 2\text{-}1$$

where V_P is the phenotypic variance

V_G is the genotypic variance

V_E is the environmental variance.

The genotypic value can however be partitioned into the following form:

$$G = A + D + I$$

where A is the breeding value

D is the dominance deviation

and I is the interaction deviation or epistatic deviation.

Thus, the genotypic variance can be extended to give:

$$V_G = V_A + V_D + V_I \qquad 2\text{-}2$$

where V_A is the additive variance derived from the breeding values

V_D is the dominance variance

V_I is the interaction (or epistatic) variance.

If we consider a single locus with two alleles then the dominance deviation is given by the differences between the breeding values and the genotypic values. Summing this over all loci gives the dominance variance (V_D). Thus the dominance variance reflects the non-additivity between alleles within each loci. In comparison, the interaction variance (V_I) arises from non-additivity between loci. The interactions are summed over all loci and may involve two or more loci and both deviations from additive and dominance effects. Thus, V_I may be expressed as the following summation of components:

$$V_I = V_{AA} + V_{DD} + V_{AD} + V_{AAA} + \ldots$$

Together V_D and V_I are often referred to as the non-additive variance, however, as previously noted, the lack of additivity arises in different manners. (That is, while the former relates to the presence of dominance between alleles at the same locus, the latter refers to the presence of epistasis between different loci.) It should be noted that the additive variance is an important component in breeding as it is the main cause of resemblance between relatives. Not surprisingly it therefore forms the main avenue of attack during selection.

Equations 2-1 and 2-2 can be combined to give

$$V_P = V_A + V_D + V_I + V_E \qquad 2\text{-}3$$

When referring to this partitioning of the phenotypic variance, it should be appreciated that V_D and V_I like V_A are dependent on the gene frequencies, and therefore are properties of the particular population measured. In practice V_E also varies between populations or over time, however, while this is intuitively reasonable it is more difficult to justify in terms of the simple genetic models which explain V_A, V_D and V_I. This can best be illustrated by the often quoted behaviour of plant communities where it is observed that V_E measured in a genotypically uniform stand is often much larger than for a corresponding mixed stand. Although this is far from completely understood, it is generally accepted that the inbreds are more sensitive or less "well buffered" to environmental factors than the outbred community.

Equation 2-3 is further complicated by the possibility of genotype-environment correlations or interactions (i.e. in developing 2-1 we have assumed that the environmental deviations and the genotypic values are independent of each other). Genotype-environment correlations may arise if, for example, the best genotypes are allocated to the best environment. Thus, 2-1 becomes

$$V_P = V_G + V_E + 2\mathrm{cov}_{GE}$$

where on practical grounds the covariance is generally inseparable from the genotypic variance.

Genotype-environment interaction reflects failure of the assumption that different genotypes respond similarly to specific environmental stimuli. Again, measurement of this interaction variance can only be accomplished under rather artificial circumstances. In default, the genotypic-environmental interaction becomes part of the environmental variance (V_E).

Use of multiple observations enables an alternative method of partitioning the phenotypic variance, or more particularly the environmental variance component of V_P

i.e. $$V_P = V_G + V_{E_g} + V_{E_s} \qquad 2\text{-}4$$

where V_{E_g} is called the general environmental variance and refers to the portion of the environmental variance contributing to the between-individual component. The special environmental variance (V_{E_s}) which is the remaining portion of the environmental variance, is then the within-individual variance. Thus, use of repeated observations allows the V_E to be partitioned into a variance between individuals (V_{E_g}) and a within individual variance (V_{E_s}). The repeated measurements can be over space or time but care must be taken that the measurements are a repeated expression of the same genes and not resulting from expression of different genes (e.g. Falconer (1960) suggested different genes are believed to affect milk yield during successive lactations). This partitioning of V_P (2-4) enables calculation of the repeatability,

i.e. $$r = \frac{V_G + V_{E_g}}{V_P}$$

The repeatability of a character first indicates the gain in accuracy to be expected from repeated measurements and second provides an upper limit of the heritability of the particular character. Thus, although it has some merit it is not widely used in selection experiments unless the heritability is unobtainable.

Before considering heritabilities, phenotypic, genetic and environmental correlations, it should be noted that maternal environment is a further important contributor to V_E in animal breeding. While post-natal effects can be minimised, prenatal effects are much more difficult to overcome. More importantly, common environment (V_{E_c}), which is basically maternal effects plus any other environmental factors leading to resemblances between relatives (e.g. nutritional status), is often inseparable from V_A. This is particularly the case for full-sibs where similarities between sibs may arise from both genetic and environmental causes. As these two are confounded, any estimate of V_A or any estimate containing V_A in its numerator will be inflated. In particular, the heritability (h^2) which is defined as the ratio of the additive genetic variance (V_A) to the phenotypic variance (V_P) may be inflated by the presence of V_{E_c}. Difficulties associated with V_{E_c} can be shown to be potentially greatest for h^2 calculated on full sibs but can also arise when intra-sire regression of offspring on dam is used to estimate h^2. (In fact, it may even be a contributing factor to the high h^2 for variates like intelligence quotient in man when regression of progeny on father has been used to estimate h^2, the resemblance here being partly due to environmental factors such as nutritional level and parental example. Clearly this occurrence is unique to man and would be absent for measurements on the lower vertebrates such as sheep.)

Heritability as defined above is sometimes referred to as h^2 in the narrow sense and can be represented as follows:

$$h^2_{(\text{narrow sense})} = V_A/V_P$$

(Alternatively, it can be calculated from the regression of breeding value on the phenotypic value.) The above definition serves to distinguish "narrow" h^2 from h^2 in the broad sense which is calculated using the genotypic variance,

i.e. $$h^2_{(\text{broad sense})} = V_G/V_P$$

As this latter ratio includes the non-additive variance it would be expected to be greater in magnitude, thus any predictions based on it would over-estimate response.

Generally, variates with low h^2 are related to reproductive fitness while those with high h^2 occur for variates which have little if any relationship to natural fitness. Table II-1, which is based on work presented by Falconer (1960) and Robertson (1959), indicates the covariance, h^2, and variance of h^2 for four of the more frequently encountered kinships. For the first two types of relatives h^2 is estimated using the regression coefficient (b), obtained following regression of offspring on parent, while for the last two cases h^2 is estimated using intra-class correlation (t) obtained following partitioning of components by analysis of variance. In assessing the meaning of any estimate there are two points to consider:

i) the statistical precision or accuracy, and

ii) the possibility of bias.

While the former provides valuable information about the estimate, the latter, when present, invalidates the estimate. That is, while the variance of the estimate can be allowed for either in using that estimate or reduced by larger sample sizes, the bias once present cannot be overcome by statistical procedures. Thus, apart from *a priori* information, there is no means of identifying bias in an estimate and even if there were, there is generally no valid means of removing it. Returning to Table II-1, we note that for full sibs the presence of V_D and V_{E_C} components will yield a biased estimate of h^2. Maternal effects can also cause bias when h^2 is estimated using regression of offspring on mother, although this is not indicated in the table. Thus, we see that common environment generally is responsible for bias but it can also result from other components. Therefore, half sib analysis or regression of offspring on father would be expected to provide the most reliable estimates of h^2 from the point of view of bias. Although much less emphasis has been given to what determines the statistical precision or accuracy of the estimate, this point should still be considered when designing experiments. On investigating the optimum design for estimating h^2, Robertson (1959) concluded that if h^2 is less than 0.25 then sib analysis provides a more accurate estimate than regression analysis. In particular, for the half-sib method of estimation, he suggested one offspring per dam with not less than five progeny per sire.

Intra-sire regression of offspring on dam, which is often used to estimate h^2, is one slight variant of the types of relationships presented in Table II-1. For this method separate regression of offspring on dam are calculated for each group of dams mated to an individual sire. The regression coefficients are then pooled over sires using a weighted average to give $b_{pooled} = \frac{1}{2}h^2$. Accuracy of the estimate is dependent on the absence of maternal effects. Also, if male offspring are regressed on dams any inequality of variance between the two sexes may affect h^2. To overcome this problem the regression coefficient should be multiplied

Table II-1 Covariance, heritability and variance of heritability for four types of relatives

Relatives	Covariance	Heritability (calculated from regression coefficient (b) or correlation coefficient (t)	Approximate Variance (h^2) †
Offspring on one parent	$\frac{1}{2}V_A$	$h^2 = 2b$	$\sigma^2_{h^2} = 4.\frac{(1+(n-1)t)}{nN}$
Offspring on mid-parent	$\frac{1}{2}V_A$	$h^2 = b$	$\sigma^2_{h^2} = \frac{2.(1+(n-1)t)}{n.N}$
Half sibs	$\frac{1}{4}V_A$	$h^2 = 4t$	$\sigma^2_{h^2} = \frac{32[1+(n-1)t]^2(1-t)}{n(n-1)(N-1)}$
Full sibs	$\frac{1}{2}V_A+\frac{1}{4}V_D+V_{E_C}$	$h^2 < 2t$	$\sigma^2_{h^2} = \frac{8[1+(n-1)t]^2(1-t)^2}{n(n-1)(N-1)}$

† Information presented here is derived from a paper by Robertson (1959)

b = regression of offspring on parent

N = number of families

n = number of offspring

t = intra-class correlation between members of families by the ratio of the phenotypic standard deviation of females to that of males.

So far, the approach has been to consider variates in isolation with little reference to the possibility of effects on the rest of the organism. Clearly, selection on one variate generally leads to changes in other variates therefore the breeder must focus on the whole, not just on one particular aspect of the population. Traditionally, this has been done by considering the phenotypic, genetic and environmental correlations. In later parts of this thesis I will consider the application of multivariate methods to this area, however, for the present I will concentrate on the use of correlations. Previously it has been proposed that the phenotypic value can be partitioned into a genotypic and an environmental component. Thus, we can calculate a genetic correlation (r_A) and an environmental correlation (r_E) or simply a phenotypic correlation (r_p). The genetic correlation then is the correlation of breeding values while the environmental correlation is the correlation of the environmental deviations plus any non-additive genetic deviations. The three correlations can be shown to be related by the following relationship:

$$r_p = h_X h_Y r_A + \sqrt{1 - h_X^2} \sqrt{1 - h_Y^2} \quad r_E$$

where h_X is the square root of the heritability of X,

and h_Y is the square root of the heritability of Y.

The interpretation of observed correlations and their use in selection programs should take into account the following four points:

i) the contribution of the pleiotropic action of genes to the correlation of variates.

ii) The effect of linkage on the correlation of variates.

iii) The effect of artificial selection on related variates.

iv) The possible side-effects on variates previously exposed to directional pressures associated with natural selection when the selection intensity for these is reduced (i.e. selection often leads to a reduction in fitness which is clearly undesirable).

As genetic correlation can arise from both pleiotropy and linkage we cannot separately consider points i) and ii), but it is often suggested that linkage contributes much less to the genetic correlation than does pleiotropy. In fact, given large outbreeding populations, any effect of linkage should be temporary. (However, this will not be the case for a plant breeder using a self-crossing species or an animal breeder with small population sizes.) Although inspection of r_A over generations would be expected to provide information on the declining effect of linkage, it must be remembered that the pleiotropic contribution may also change over time as loci become fixed. Assuming r_A results from pleiotropy then the response in a correlated variate (CR_Y) can be predicted by

$$CR_Y = i\, h_X\, h_Y\, r_A\, \sigma_{P_Y}$$

where i = intensity of selection

and σ_{P_Y} = phenotypic standard deviation of Y.

When the secondary character (Y) is related to fitness then it is likely that natural selection itself is already acting on Y. If the artificial selection is in the opposite direction to the natural selection then the artificial selection will be weakened or the selection may result in progeny with high values of the variate of interest but low fitness.

When using such statistics as V_A, h^2, r_p, r_A and r_E, one should appreciate that each describes the net effect of many segregating genes. Thus, alleles at the individual loci may differ in both magnitude and even direction. For example, while r_A may be observed to be 0.4 which suggests that the individual genes would act to increase both characters, there may also exist genes which increase one character and reduce the other. Alternatively, the correlation may be zero as there exists a balance between genes with positive effects and negative pleiotropic effects. As selection may change this "balance" between the individual polygenes it would be misleading to use these statistics to predict response over more than one generation. (Instead the statistics should be recalculated after each generation.) Also, these statistics may vary between populations therefore extreme care should be taken when extrapolating outside the particular population.

Schinckel (1958) used data from the early years of the Roseworthy experiment to estimate heritabilities and genetic correlations. These values, presented in Table II-2, appear to be the only estimates for the South Australian strong-wool Merino. The estimates which were calculated using parent-offspring regression will be compared later to estimates obtained in the present study. (It should be noted that Schinckel used body weight to the power 0.6 and total number of follicles per square cm whereas the present study considers body weight, primary follicle number per square cm and secondary follicle number per square cm.)

Table II-2 Estimate of genetic correlations and heritabilities (shown in brackets) presented by Schinckel (1958)

BWT$^{0.6}$†	(.76)					
CFW	.47	(.28)				
STL	.04	.37	(.65)			
CRP	.06	-.22	-.54	(.40)		
FBD	-.21	.24	.44	-.17	(.52)	
PFN + SFN	-.13	.14	-.36	.06	-.70	(.62)
	BWT$^{0.6}$†	CFW	STL	CRP	FBD	PFN+SFN

† Schinckel uses body weight to the power 0.6

for explanation of the abbreviations used see later text

Similar estimates have been presented for medium Peppin Merinos by Brown and Turner (1968). Their article also summarises estimates previously reported by Morley (1955), Beattie (1962) and Young, Turner and Dolling (1960). (These estimates have not been reproduced here as this would be redundant, but comparisons with Brown and Turner's tabulations will be made where appropriate.) Barlow (1974) calculated realised genetic correlations for the medium Peppin Merino experimental flocks of the N.S.W. Department of Agriculture.

METHODS

All genetic parameters have been calculated using the computer package, NESREG, which has been developed by Hammond, Jackson and Miller (1972). This program package enables estimation of genetic parameters by nested analysis of variance or offspring-parent regression. As the Roseworthy data is compatible with either form of analysis, both will be applied to the data. In particular, the offspring parent regression itself has been considered under two alternate designs. Thus, genetic correlations and heritabilities have been obtained by the following three methods:

i) half-sib analysis on sire groups;

ii) offspring-sire regression;

iii) intra-sire regression of offspring on dam.

The three methods have been included to enable their consistency and accuracy to be compared.

The following seven variates have been analysed:

i) hogget body weight in kg (BWT);

ii) clean fleece weight in kg (CFW);

iii) staple length in cm (STL);

iv) crimps per inch (2.54 cm) (CRP);

v) fibre diameter in microns (FBD);

vi) primary follicle number (PFN);

vii) secondary follicle number (SFN).

Unfortunately, due to central memory constraints, the computer was unable to analyse more than seven variates at a time, therefore, several variates of interest (e.g. birth weight, percentage yield and skin thickness) had to be excluded from the present analysis. In all three methods the data was divided into its four sex by flock groups and separate parameters estimated. All progeny resulting from multiple births, or where the record was incomplete, have been excluded from the analysis.

A special FORTRAN program was written to present the data in a form suitable for input to NESREG. Basically NESREG requires the parent and progeny information in a contiguous block. Therefore, the existing progeny records, which contained the sire and dam identity, had to be expanded to include the appropriate sire and dam information. As there were only fourteen potential sires per year, the sire information could be stored in an array and retrieved as required. In comparison, the pairing of dam and progeny was somewhat more difficult as the large number of dam records prohibited any in-core storage. Thus, the dam and progeny information had to be read in from tape (or disc) as required and then paired and written out onto a tape (or disc). To increase the efficiency of this pairing, the progeny was ranked in order of its dam's identity number. The dam information was similarly ranked, using its own identity. This then enabled the two records to be conveniently united without necessitating any backward searching.

As mentioned earlier, the analyses were carried out using the package NESREG. Therefore, as the manual to this package contains considerable details on the method of analysis I will not repeat them here.

RESULTS

Table II-3 lists the heritabilities and genetic correlations for the seven variates considered. (Standard errors are included for each h^2 and r_A.) The statistics have been calculated using the entire data set (i.e. over the twelve years 1954 to 1965), but partitioned into four groups corresponding to the two flock types by two sexes. (Only the single born progeny have been included.) Thus, four values appear as a block for each h^2 and r_A, their order of appearance being:

i) Index flock male progeny,

ii) Visual flock male progeny,

iii) Index flock female progeny,

iv) Visual flock female progeny.

A lower triangular display has been used, thus the seven h^2 appear in brackets along the diagonal with the 21(= $P\frac{(P - 1)}{2}$) r_A forming the off-diagonal elements. As stated above the values presented have been calculated over the twelve years of the trial. Similar calculations have been made for each individual year. Extensive investigation of this data failed to reveal any trends over time and as the values were for the most part homogeneous it was felt that presentation of the summary tables was preferable (as well as much more concise). The failure to observe temporal effects associated with the selection applied may be due to the large amount of variation associated with the estimates within years. This mainly resulted from the moderate size of the individual groups which ranged from 37 to 86. Thus, although the observations have failed to establish any trends in the statistics this is far from conclusive demonstration that such changes did not occur. The values of h^2 and r_A presented in Table II-3 have been calculated using half-sib analysis based on sire groups. Tables II-4 and II-5 present similar information, from the same progeny, but using offspring on sire regression and intra-sire regression of offspring on dam respectively. In Tables II-6 and II-7 environmental correlations (r_E) and phenotypic correlations (r_P) appear using a similar layout (note in this case no diagonal elements have been included where h^2 was shown in the previous tables). Here, r_E and r_P have been calculated using half-sib analysis only.

Clearly, Tables II-3 to II-7 contain a large amount of information which the following can do little more than introduce and summarise. This difficulty stems from the presence of the many inter-related questions which are of potential importance, e.g.

i) are the three methods of estimating h^2 and r_A consistent?

ii) Are there differences between the two flocks?

iii) Are there differences between the two sexes?

iv) What do r_A, r_E and r_P and also h^2 tell us about the variate measured?
i.e., has the selection had any effect on the underlying parameters and how would the flocks respond to continued selection?

To give some order, I will consider these questions in the sequence given above, however, it is inevitable that some overlap will remain. All comparisons will be made using the standard errors to assess the significance of the estimates.

Inspection of Tables II-3, II-4 and II-5 indicates that the estimates of h^2 are generally comparable, but the value for fibre diameter is markedly larger from the half-sib analysis (0.51 to 1.03) than from the offspring-sire regression (-.01 to .27) or the intra-sire regression of offspring on dam (.33 to .56). As half-sib analysis is expected to provide one of the more accurate methods of estimating h^2 (see Robertson, 1959), this result should be given most emphasis, however, the observations remain confusing as one expects the bias in the regression estimates to increase not decrease the estimate of h^2. Linkage could be responsible for inflating the half-sib estimate but not the regression estimates, but even this seems a little unlikely. Seasonal differences and the selection process could also make some contribution. While all these explanations are far from satisfactory, the observation, when related to the fluctuating behaviour of fibre diameter in the previous chapter, highlights the need for further investigations of this variate. The genetic correlations (r_A) of fibre diameter with the other six variates all appear consistent over the three methods. (I have already referred to a negative heritability which theoretically one would not expect though in practice variation may lead to such negative components. As r_A cannot be calculated in such situations all values of it have been set to zero.)

The value of r_A remains reasonably consistent for all other combinations of variates except for the correlation between primary follicle number and secondary follicle number. (Follicle development will be considered in greater detail later in this section.) Once again, the half-sib values exceed the values from the two parent offspring regressions but there is less consistency within the four sex by flock types. Seasonal variation may again cause weaknesses in the regression methods due to presence of genotype by environment interactions.

On comparing the two flocks within each sex, the following observarions can be made:

i) h^2 for clean fleece weight are lower for the Index progeny than for the Visual progeny. This holds for all three methods of calculation and for both male and female offspring although the difference is greatest from the half-sib analysis. The values of the Index progeny are generally low (0.12 - 0.32) whereas those for the Visual progeny (0.31 - 0.38) more closely resemble the level obtained by Schinckel (1958), i.e. 0.28, see Table II-2. As h^2 is the pooled value for the twelve years of the experiment the lower value for the Index flock could result from the greater selection intensity for clean fleece weight in this flock. However, as indicated earlier when the 12 individual h^2 (for the 12 years) are plotted, no evidence could be found to support a gradual decline in the proportion of additive variance for this variate.

ii) For fibre diameter the Index flock appears to have a lower h^2 than the Visual flock but the magnitude of the difference is smaller and less consistent than for clean fleece weight.

iii) In contrast to the above two variates, the h^2 for staple length is greater for the Index progeny (0.36 - 0.57) than for the Visual progeny (0.24 - 0.40). As Schinckel (1958) has reported a h^2 of 0.65, it would appear that the Visual progeny is showing less genetic variability for this variate at the completion of the experiment. As this is opposite to the case of h^2 for clean fleece weight, this observation appears somewhat contradictory (i.e. selection for clean fleece weight would be expected to increase staple length indirectly resulting in a reduction in h^2). It would therefore appear that Visual appraisal has selected heavily against long staples and this has caused a reduction in the h^2 of the Visual flock.

iv) For the remaining variates (i.e. hogget body weight, crimp number, primary follicle number and secondary follicle number) there are no consistent differences in h^2 between the flocks.

v) The following list summarizes the occasions when r_A differs between the two flocks. (Careful inspection of the relevant tables indicates that the differences are seldom consistent over all three methods of calculation, however, on this occasion no attempt will be made to relate this to the weaknesses of particular methods.)

(a) The genetic correlation between staple length and body weight (i.e. r_A(staple length, body weight)) is larger for Visual progeny than for Index progeny in all cases except the male progeny under offspring-dam regression.

(b) r_A(crimp, staple length) is more strongly negative for Visual progeny than for the Index progeny except for the male progeny under offspring-dam regression.

(c) r_A(secondary follicle number, staple length) is negative for Index progeny and near zero for Visual progeny except for the female progeny under offspring-dam regression.

(d) r_A(primary follicle number, fibre diameter) has a negative value between -0.32 to -0.67 for the Visual progeny, while for the Index progeny r_A is near zero (except for the half-sib analysis). For all four cases the value of the genetic correlation differed for one of the methods of calculation, therefore, the points made are far from conclusive.

Reversing the role of flock and sex we can look for evidence of differences in h^2 and r_A between sexes within the same flock. Examination of the three tables indicates only two cases where the males differ consistently from the females in both the flocks. These occur for the r_A firstly between clean fleece weight and staple length and secondly between clean fleece weight and crimp number. In the first case males have lower positive correlations than females while in the second case all correlations are negative with the males having the greater negative association. In both cases the estimates calculated from the intra-sire regression of offspring on dam deviate from the above mentioned pattern.

In the two previous sections I have reported cases where the estimate differs for one factor while using a second factor to provide replication (i.e. compared between flocks using males and females as replicates). While this method provides useful information it should be acknowledged that any interactions between the two factors may be overlooked. In particular, the problem of sex-linkage (see Beilharz (1963) and James (1973)) has been ignored. Although this may be of importance, it is difficult to obtain any effective comparison as the data are not well suited for its consideration (i.e. after partitioning the numbers are small such that only large effects would be observed).

However, it should be noted that in Table II-3 (half-sib analysis) and less so in Table II-4 (offspring-sire regression), the Index female r_A often differs greatly from the values for the other three progeny groups. Thus, looking at Table II-3 r_A(bodyweight, crimp), r_A(clean fleece weight, crimp), r_A(staple length, crimp), r_A(clean fleece weight, fibre diameter), r_A(staple length, fibre diameter), r_A(crimp, fibre diameter), r_A(body weight, primary follicle number), r_A(clean fleece weight, primary follicle number) and r_A(staple length, secondary follicle number) all show this unusual pattern. In comparison, the values of r_A calculated by intra-sire regression of offspring on dam, given in Table II-5 show no such inconsistency. As the half-sib analysis is expected to be more accurate than the intra-sire regression of offspring on dam these values for the Index female progeny in the former are difficult to explain. Clearly if the situation had been reversed it could be explained as maternal effects, however, as this is not the case maternal effects would seem to be "above suspicion". The only plausible explanation would seem to be that the objective selection method applied to the Index flock has favoured selection of rams carrying sex-linked characters which results in the unusual correlations observed among their female progeny. Obviously this explanation is somewhat speculative, however, in view of its implications to sheep breeding it clearly warrants closer investigation.

Given the above summary of h^2 and r_A it is now possible to include the environmental correlations (r_E) and the phenotypic correlation (r_P) to enable a more complete consideration of all aspects relevant to the sheep breeder. As clean fleece weight is the variate of greatest interest, I will consider its relationship with the other variates first and then look at any important features among the rest of the variates.

Previously, I have noted that for clean fleece weight, the Index progeny show a lower h^2 than the Visual progeny. Comparing these to Schinckel's earlier estimates we note that the final value of the Index flock is lower. Therefore, it appears that selection has reduced the additive variance for the Index flock where the selection intensity was larger. Response should continue under further selection but in the absence of further recombination or introduction of external genetic material the response would be expected to decrease in later generations.

Turning to the relationship between body weight and clean fleece weight we observe:

i) r_A between -.24 and .34
ii) r_E " .43 and .51
iii) r_P " .27 and .38
iv) $h^2_{(bodyweight)}$ " .25 and .58.

Together, these estimates indicate that although body weight and clean fleece weight both increase in favourable environments, the selection for the latter variate would not be expected to lead to increased body weight. However, as h^2 for body weight is moderately large, direct selection for this variate would be expected to lead to response.

For staple length we observe:

i) r_A between .09 and .80
ii) r_E " .12 and .37
iii) r_P " .29 and .36
iv) $h^2_{\text{(staple length)}}$ " .24 and .57

(c.f. Schinckel, 1958, $r_A = .37$ and $h^2 = .65$)

With respect to r_A it has already been noted that there is some tendency for the males to have lower values than females. Also, the h^2 for Index progeny is greater than for Visual progeny. In view of the r_A and r_E one would expect that selection for clean fleece weight will result in increased staple length. However, the higher h^2 for the Index progeny complicates the situation (i.e. it is unclear how this should have arisen). Possibly these differences are due to chance or, as suggested earlier, it may be that Visual appraisal selects strongly against long staples. Again, it appears that the h^2 has decreased from Schinckel's earlier value.

The situation for crimp number differs from the previous two variates in that all correlations are negative, i.e.

i) r_A between -.08 and -.65
ii) r_E " -.18 and -.33
iii) r_P " -.27 and -.33
iv) $h^2_{\text{(crimp number)}}$ " .26 and .57.

(c.f. Schinckel, 1958, $r_A = -.22$ and $h^2 = .40$)

As the environmental correlation is smaller in magnitude than the genetic correlation, one would expect that crimp number would decrease on selecting for increased fleece weight. Alternatively, response in crimp number would be expected following direct selection. In comparison with the earlier values it appears that h^2 is unchanged but r_A has become more negative.

It has already been acknowledged that fibre diameter is inconsistent in expression as can be seen from the wide range of its h^2 in the following:

i) r_A between $-.46$ and $.31$
ii) r_E " $.13$ and $.26$
iii) r_P " $.13$ and $.16$
iv) h^2 (fibre diameter) " $-.01$ and 1.03

(c.f. Schinckel, 1958, $r_A = .24$, $h^2 = .52$)

The wide range for h^2 and r_A make it impossible to predict how the variate would respond to either direct selection or indirect selection via clean fleece weight. Clearly this is a very unsatisfactory situation especially as the variate plays such an important part in the marketing of wool.

As the estimates for primary follicle number and secondary follicle number are quite similar, they will be considered simultaneously.

For primary follicle number:

i) r_A between $-.40$ and $.36$
ii) r_E " $-.16$ and $.00$
iii) r_P " $-.08$ and $-.05$
iv) h^2 (primary follicle number) between $.12$ and $.39$.

While for secondary follicle number:

i) r_A between $-.08$ and $.53$
ii) r_E " $-.03$ and $.11$
iii) r_P " $.02$ and $.15$
iv) h^2 (secondary follicle number) between $.16$ and $.46$.

As h^2 is of moderate size, both variates should respond to direct selection. However, for primary follicle number, little, if any, response would be expected to indirect selection associated with clean fleece weight as the value of r_A is low. For secondary follicle number a small positive response may occur. The values of h^2 differ markedly from Schinckel's h^2 for the sum of primary and secondary follicles (0.62). As it seems unlikely that the summation could be responsible for this difference, it must be concluded that either h^2 for both follicle scores has decreased during the selection or Schinckel's estimate was somewhat high.

For medium Peppin Merino ewes, Young, Turner and Dolling (1960) estimated h^2 for fibre density at 0.31, while Brown and Turner (1968) obtained a similar value of 0.42. As both these values are similar to the present estimates it would seem that Schinckel's original estimate was too large, but it should be appreciated that h^2 is directly affected by the magnitude of the environmental variance. This last point is particularly relevant here as follicle number is expressed as the count per square cm, thus, variation in the surface area will affect follicle number. As follicles occur as clusters, the total number per animal can be considered as comprising two components:

i) the number of follicle groups (indicated by the primary follicle number);

and ii) the size of these groups (which in the past has been indicated by the S/P ratio).

Schinckel (1955a) reports that total primary follicle number is complete by birth and the only changes thereafter merely reflect the normal skin expansion during growth. This constancy of the primary follicle number, over the lifespan of the individual, has led to the ratio of secondary to primary follicles (S/P) being widely used as a measure of the development of secondary follicles. Although all secondary follicles are initiated prior to birth, not all necessarily reach maturity. Schinckel (1955b) reports that maturation of secondary follicles was significantly affected by birth weight and growth during the first month following birth. The importance of the post-natal period stems from the observation that 70-80% of the secondary follicles which finally attain maturity have done so by the age of one month. Thus, it is not surprising that later growth has negligible effect on total follicle number although it will have considerable effect on follicle density via body size. As the milk production of the ewe mainly determines the lamb's growth during the first months (Wallace, 1948), then total secondary follicle number would be expected to be positively related to mothering ability. This then implies that, along with birth weight, secondary follicle number provides a useful measure of the maternal environment. In particular, care should be taken that selection is not causing reductions in total secondary follicle number as this may be associated with poorer mothering ability. (For both flocks considered here, there is no evidence of any decrease in total follicle number as indicated by the number of secondary follicles per square cm.)

In the preceding I have mentioned that the ratio S/P has frequently been used to remove the effect of body size. Intuitively this seems quite reasonable but it has not been used here for the following reasons. Firstly, most of the statistical methods used in this thesis assume normality and this is far less likely to be true for such a ratio. Secondly, larger values of S/P may be associated with higher total secondary follicle number but they may also arise from low primary follicle number. Clearly these two situations should be distinguishable. (In fact, Turner (1956) reports that disruptive selection for follicle number resulted in changes in both primary and secondary follicle number.)

So far I have looked at the relationship between clean fleece weight and each of the other six individual variates with special reference to

Table II-3 Heritabilities and genetic correlations (with standard errors) calculated using half-sib analysis based on sire groups. (h^2 and its S.E. in brackets.)

†	I-M	Index male progeny
	V-M	Visual male progeny
	I-F	Index female progeny
	V-F	Visual female progeny

BWT	I-M†	(.48)	(.14)												
	V-M	(.33)	(.12)												
	I-F	(.25)	(.12)												
	V-F	(.53)	(.15)												
CFW	I-M	.08	.33	(.12)	(.09)										
	V-M	-.22	.26	(.33)	(.12)										
	I-F	.06	.36	(.20)	(.11)										
	V-F	.31	.20	(.32)	(.12)										
STL	I-M	-.12	.23	.17	.34	(.37)	(.13)								
	V-M	.20	.26	.54	.21	(.28)	(.11)								
	I-F	.07	.26	.61	.22	(.54)	(.16)								
	V-F	.40	.18	.80	.15	(.40)	(.13)								
CRP	I-M	.07	.25	-.61	.33	-.08	.27	(.26)	(.11)						
	V-M	.10	.24	-.58	.18	-.61	.19	(.38)	(.12)						
	I-F	.48	.30	-.08	.33	.30	.25	(.31)	(.12)						
	V-F	.06	.19	-.37	.19	-.55	.16	(.57)	(.15)						
FBD	I-M	.03	.20	.06	.31	-.10	.21	-.23	.22	(.66)	(.17)				
	V-M	-.03	.21	.04	.21	-.12	.22	-.33	.19	(.66)	(.16)				
	I-F	-.15	.27	.23	.28	.48	.18	-.51	.21	(.51)	(.15)				
	V-F	.01	.17	.08	.19	-.11	.18	-.24	.16	(1.03)	(.22)				
PFN	I-M	-.12	.24	.01	.38	-.07	.26	-.01	.29	-.55	.19	(.29)	(.12)		
	V-M	-.10	.28	-.19	.28	.47	.29	.00	.27	-.32	.23	(.23)	(.10)		
	I-F	.59	.39	.36	.41	-.25	.28	.15	.34	-.36	.28	(.19)	(.11)		
	V-F	-.35	.21	-.22	.25	-.00	.24	.23	.21	-.67	.15	(.33)	(.12)		
SFN	I-M	-.34	.22	.17	.36	-.17	.24	-.11	.28	-.44	.18	.69	.17	(.33)	(.12)
	V-M	-.23	.25	.53	.22	.08	.27	-.25	.24	-.20	.20	.49	.22	(.31)	(.11)
	I-F	-.20	.38	.05	.42	-.58	.26	.04	.36	-.63	.24	.37	.36	(.16)	(.10)
	V-F	-.41	.18	-.08	.23	.02	.22	.05	.20	-.82	.10	.51	.17	(.46)	(.14)
		BWT		CFW		STL		CRP		FBD		PFN		SFN	

Table II-4 Heritabilities and genetic correlations (with standard errors) calculated using offspring-sire regression. (h^2 and its S.E. in brackets.)

BWT	I-M†	(.58)	(.07)												
	V-M	(.38)	(.07)												
	I-F	(.31)	(.07)												
	V-F	(.30)	(.06)												
CFW	I-M	-.04	.23	(.22)	(.11)										
	V-M	.05	.20	(.41)	(.08)										
	I-F	-.24	.32	(.23)	(.09)										
	V-F	.34	.21	(.31)	(.06)										
STL	I-M	.02	.17	.22	.21	(.57)	(.09)								
	V-M	.67	.23	.31	.22	(.24)	(.09)								
	I-F	-.04	.21	.58	.19	(.54)	(.08)								
	V-F	.61	.21	.45	.21	(.28)	(.08)								
CRP	I-M	.18	.18	-.49	.22	-.07	.19	(.38)	(.10)						
	V-M	.07	.22	-.59	.17	-.61	.21	(.31)	(.09)						
	I-F	.44	.28	-.13	.31	.02	.23	(.23)	(.09)						
	V-F	.14	.22	-.28	.21	-.32	.21	(.38)	(.08)						
FBD	I-M	.00	.00	.00	.00	.00	.00	.00	.00	(-.01)	(.07)				
	V-M	.10	.24	.12	.23	.02	.28	-.58	.23	(.25)	(.06)				
	I-F	-.35	.35	-.46	.42	.27	.26	.09	.39	(.16)	(.07)				
	V-F	-.20	.26	.31	.24	.06	.26	-.43	.21	(.27)	(.06)				
PFN	I-M	-.11	.21	-.26	.29	-.30	.20	.36	.24	.00	.00	(.26)	(.06)		
	V-M	.41	.30	-.18	.27	.53	.35	-.25	.29	-.52	.32	(.17)	(.08)		
	I-F	.15	.28	.03	.31	-.01	.23	-.25	.31	.15	.38	(.23)	(.08)		
	V-F	.16	.23	-.40	.21	-.31	.23	.06	.21	-.50	.22	(.39)	(.09)		
SFN	I-M	-.20	.18	.04	.25	-.51	.15	-.05	.21	.00	.00	.55	.18	(.38)	(.08)
	V-M	-.06	.20	.39	.18	-.01	.24	.04	.22	-.71	.19	.04	.28	(.41)	(.08)
	I-F	.01	.25	.05	.28	-.54	.18	-.19	.29	-1.04	.26	.19	.27	(.29)	(.09)
	V-F	-.51	.19	-.08	.21	.04	.22	.14	.20	-.73	.16	.49	.15	(.45)	(.08)
		BWT		CFW		STL		CRP		FBD		PFN		SFN	

†I-M	Index male progeny
V-M	Visual male progeny
I-F	Index female progeny
V-F	Visual female progeny

Table II-5 Heritabilities and genetic correlations (with standard errors) calculated using intra-sire regression of offspring on dam. (h^2 and its S.E. in brackets.)

BWT	I-M†	(.43)	(.07)												
	V-M	(.44)	(.07)												
	I-F	(.45)	(.08)												
	V-F	(.54)	(.07)												
CFW	I-M	.14	.07	(.32)	(.08)										
	V-M	.14	.06	(.34)	(.07)										
	I-F	.20	.07	(.29)	(.08)										
	V-F	.14	.06	(.38)	(.07)										
STL	I-M	-.03	.06	.44	.05	(.54)	(.07)								
	V-M	-.02	.06	.22	.06	(.36)	(.07)								
	I-F	-.24	.06	.09	.07	(.50)	(.08)								
	V-F	.24	.06	.55	.05	(.35)	(.07)								
CRP	I-M	.00	.06	-.65	.05	-.38	.05	(.46)	(.07)						
	V-M	.05	.06	-.52	.05	-.27	.06	(.40)	(.07)						
	I-F	-.01	.06	-.62	.06	-.12	.06	(.49)	(.08)						
	V-F	.07	.06	-.53	.05	-.69	.04	(.44	(.07)						
FBD	I-M	-.09	.07	.02	.08	.25	.06	-.60	.06	(.33)	(.08)				
	V-M	.04	.05	.23	.06	.28	.06	-.39	.05	(.56)	(.07)				
	I-F	.04	.06	.19	.07	.14	.06	-.55	.05	(.44)	(.08)				
	V-F	.10	.06	.30	.06	.38	.06	-.45	.05	(.38)	(.07)				
PFN	I-M	.29	.09	.07	.10	-.07	.08	.05	.09	-.00	.10	(.15)	(.08)		
	V-M	.06	.10	-.32	.10	.07	.10	-.08	.10	-.55	.09	(.12)	(.07)		
	I-F	-.00	.09	-.14	.10	-.01	.08	-.17	.09	.02	.09	(.19)	(.08)		
	V-F	-.24	.08	-.06	.09	-.15	.09	.23	.08	-.44	.08	(.16)	(.07)		
SFN	I-M	-.06	.06	.15	.07	-.43	.05	.06	.06	-.73	.05	.02	.10	(.40)	(.07)
	V-M	-.33	.06	.17	.07	-.10	.07	-.03	.07	-.81	.05	.16	.10	(.30)	(.07)
	I-F	-.10	.07	.11	.08	-.15	.06	.06	.07	-.58	.05	.38	.08	(.38)	(.08)
	V-F	-.17	.06	.18	.07	-.23	.07	.17	.06	-.95	.04	.43	.07	(.35)	(.07)
		BWT		CFW		STL		CRP		FBD		PFN		SFN	

† I-M Index male progeny
V-M Visual male progeny
I-F Index female progeny
V-F Visual female progeny

Table II-6 Environmental correlations (with standard errors) calculated using half-sib analysis based on sire groups.

† I-M Index male progeny
V-M Visual male progeny
I-F Index female progeny
V-F Visual female progeny

CFW	I-M†	.43	.06										
	V-M	.51	.07										
	I-F	.46	.04										
	V-F	.43	.07										
STL	I-M	.34	.09	.37	.05								
	V-M	.08	.06	.19	.05								
	I-F	.09	.08	.21	.07								
	V-F	.08	.09	.12	.07								
CRP	I-M	-.05	.07	-.24	.04	-.31	.06						
	V-M	.01	.07	-.18	.06	-.18	.06						
	I-F	-.14	.06	-.33	.05	-.52	.10						
	V-F	-.02	.12	-.26	.08	-.07	.10						
FBD	I-M	.09	.13	.22	.08	.42	.12	-.20	.09				
	V-M	.13	.10	.26	.10	.28	.10	.05	.12				
	I-F	.22	.08	.13	.07	-.18	.14	-.10	.08				
	V-F	.00	.00	.00	.00	.00	.00	.00	.00				
PFN	I-M	.02	.08	-.06	.04	-.15	.06	.02	.06	.20	.12		
	V-M	-.12	.05	-.04	.05	-.17	.05	-.10	.06	.06	.09		
	I-F	-.26	.05	-.16	.05	-.06	.07	-.06	.05	-.02	.07		
	V-F	.07	.09	-.00	.06	.00	.07	-.19	.10	.00	.00		
SFN	I-M	.08	.08	.11	.05	-.33	.06	.05	.06	-.34	.09	.35	.05
	V-M	-.03	.06	-.03	.06	-.10	.06	.04	.07	-.38	.09	.40	.04
	I-F	-.09	.05	.02	.05	-.14	.07	.02	.05	-.31	.06	.52	.03
	V-F	.05	.11	.08	.07	-.19	.08	-.06	.11	.00	.00	.46	.06
		BWT		CFW		STL		CRP		FBD		PFN	

Table II-7 Phenotypic correlations (with standard errors) calculated using half-sib analysis based on sire groups.

CFW	I-M[†]	.31	.03										
	V-M	.27	.03										
	I-F	.38	.03										
	V-F	.37	.03										
STL	I-M	.15	.04	.31	.03								
	V-M	.11	.04	.29	.03								
	I-F	.08	.04	.32	.03								
	V-F	.23	.04	.36	.03								
CRP	I-M	-.00	.04	-.30	.03	-.24	.04						
	V-M	.04	.04	-.33	.03	-.32	.03						
	I-F	.03	.04	-.27	.04	-.17	.04						
	V-F	.02	.04	-.30	.03	-.30	.03						
FBD	I-M	.06	.04	.14	.04	.15	.04	-.20	.04				
	V-M	.05	.04	.14	.04	.09	.04	-.14	.04				
	I-F	.08	.04	.15	.04	.17	.04	-.26	.04				
	V-F	.06	.04	.16	.04	.14	.04	-.23	.04				
PFN	I-M	-.03	.04	-.05	.04	-.12	.04	.01	.04	-.14	.04		
	V-M	-.11	.04	-.08	.04	-.01	.04	-.07	.04	-.09	.04		
	I-F	-.07	.04	-.06	.04	-.11	.04	-.01	.04	-.13	.04		
	V-F	-.11	.04	-.07	.04	-.00	.04	-.00	.04	-.12	.04		
SFN	I-M	-.09	.04	.12	.04	-.27	.04	.00	.04	-.37	.03	.46	.03
	V-M	-.09	.04	.15	.04	-.05	.04	-.06	.04	-.27	.03	.42	.03
	I-F	-.11	.04	.03	.04	-.25	.04	.02	.04	-.38	.03	.50	.03
	V-F	-.18	.04	.02	.04	-.10	.04	-.00	.04	-.36	.03	.48	.03
		BWT		CFW		STL		CRP		FBD		PFN	

[†] I-M Index male progeny
V-M Visual male progeny
I-F Index female progeny
V-F Visual female progeny

the effect of further selection for clean fleece weight. This approach has been adopted to describe what would have occurred if the Index flock had been maintained for further years of selection. In view of the greater fleece weight of the Index flock this would appear the most likely avenue of further selection. However, if selection is carried out for any one of the other six variates, the h^2 in all cases is large enough to expect positive response for the particular variate. I will not attempt to elaborate on the behaviour of the remaining variates, if such selection were to be applied, as this would prove cumbersome and can be obtained from inspection of the relevant correlations in the tables. The principles involved are the same as indicated for clean fleece weight.

The large r_A between primary follicle number and secondary follicle number and the large negative r_A between fibre diameter and secondary follicle number (and less so between fibre diameter and primary follicle number) warrant mention. Schinckel (1958) reported an even larger negative r_A between fibre diameter and the sum of primary and secondary follicle number. Clearly, if selection was applied to any one of these characters, care should be taken to monitor for changes in the remaining two characters.

SUMMARY

The estimates presented for the genetic parameters indicate that changes have occurred since Schinckel's earlier investigation (Schinckel, 1958). The h^2 for clean fleece weight and fibre diameter are both lower in the Index than in the Visual flocks. However, for staple length the h^2 for the Index flock exceeds that of the Visual flock. Differences were also observed in r_A between the two flocks - r_A(body weight, staple length), r_A(crimp number, staple length), r_A(staple length, secondary follicle number), r_A(fibre diameter, primary follicle number).

When the estimates were compared between sexes, it was noted that the r_A differed between clean fleece weight and each of staple length and crimp number.

Closer comparison of the estimates, from the three types of kinship, showed that the half-sib estimate of r_A for the female progeny of the Index flock was often atypical. If, as suggested, sex-linkage was responsible for this deviation, it would clearly be of importance to the breeder. Howevever, regardless of the explanation, the observation illustrates the value to be gained by obtaining estimates from the alternative methods. That is, although standard errors can be calculated, for any individual method, these give no information on the possible bias which may be incorporated, and this latter aspect is clearly of far greater importance.

Thus, this study indicates that the frequently used genetical parameters are often not as accurate as is widely believed but even so they provide useful information for the breeder. (For further experimental evidence on the limitation of h^2 see Meyer and Enfield (1975).)

III - ASSESSMENT OF FLEECE COMPONENTS USING MULTIVARIATE STATISTICS

INTRODUCTION

In this chapter the approaches used for resolving the relationships among the components of clean fleece will be considered. In view of the statistical difficulties encountered in many of these an alternative multivariate method will be suggested. The utility of this method is demonstrated on data from the fleece weight selection experiment, carried out at Roseworthy Agricultural College, which has previously been analysed, using univariate and bivariate statistics, by Mayo, Potter, Brady and Hooper (1969) and further analysed in Chapters I and II of this thesis. (A summary of this chapter has already been published by myself (see Hancock, 1976).)

When analysing selection experiments, one is often confronted with the situation where many variates have been measured on samples from each of two populations. While multivariate techniques have been available for some time most investigators have preferred to use simpler statistics. This may be due to such reasons as:

i) unfamiliarity with multivariate statistics;

ii) limited computing facilities;

iii) cautiousness over assumptions.

However, while the last of these reasons is valid for multivariate methods in general, the technique that I am suggesting here (Hotelling's T^2) has been shown to be far more robust to deviations from multinormality (Mardia, 1975). Also, the convenience of modern statistical packages has greatly reduced the difficulties involved in finding suitable transformations, when required, for poorly distributed variates. (Although this is usually carried out on each variate in turn, recent work by Andrews, Gnanadesikan and Warner (1973), and Mardia (1975), on the assessment of multivariate normality, should further assist this area.)

Turner (1958) considered the following three methods for assessing the influence of each component on clean fleece weight:

i) gross correlation of each component;

ii) apportioning of the variance;

iii) percentage deviation technique.

For the first method the author clearly demonstrated that changes in the correlation between a particular variate and fleece weight can greatly affect the observed correlation between any other variate and clean fleece weight (Turner, 1958). (Pattie and Barlow (1974) have used similar arguments to demonstrate that the heritability of clean fleece weight can change markedly depending on the magnitude of the genetic correlation between fibre diameter and follicle density.) Also, the second method was shown to have limitations especially with respect to sampling errors. Turner concludes that while the third technique was "...by no means perfect, [it] has proved to be a powerful tool in analysing the source of differences in clean fleece weight between groups of sheep." In particular, this technique suggested that fibre number and staple length were most closely associated with clean fleece weight. Dun (1958) using this third approach found fibre density and cross-sectional area to be most important. Recently, this technique was applied by Barlow (1974) who supported it with the calculation of realized correlated responses and realized genetic correlations. Barlow calculated his realized correlated responses by two methods:

i) regression of cumulated correlated response on cumulative selection differential.

ii) Regression of cumulative correlated response on cumulative response.

These two responses were used subsequently to calculate two realized genetic correlations. By these methods Barlow concluded that the response in clean fleece weight for the Fleece Plus flock was due to fibre density, fibre diameter and staple length, while for the Fleece Minus flock, staple length was the major contributor.

Similarly, Robards, Williams and Hunt (1974) used correlations when reporting that crimp frequency was related positively to live weight and negatively to clean fleece weight.

In the two previous chapters, similar univariate and bivariate statistics have been reviewed and subsequently applied to the Roseworthy data. Both the two sample tests of Chapter I and the genetic parameters of Chapter II have proved informative, but at best these approaches are somewhat cumbersome where more than a few variates have been measured. (Mayo, Potter, Brady and Hooper (1969) also encountered similar difficulties when using repeated t-tests to compare two types of selection for increased wool production.) Thus, if p variates have been recorded, use of such statistics will result in a minimum of p two-sample tests and $\frac{p(p - 1)}{2}$ correlations which for p larger than 4 or 5 becomes difficult to handle. In most cases these numbers are far exceeded as one generally considers p two-sample tests, p heritabilities and $\frac{3p(p - 1)}{2}$ correlations (i.e. phenotypic, genetic and environmental). While the sheer number of statistics to be considered is a problem in its own right, the picture is further complicated by the chance variation associated with such repeated sampling (i.e. the frequency of error type I increases as the number of tests increases).

The Roseworthy experiment, like most large-scale breeding trials, is also weakened by seasonal variation. Thus, for individual variates the presence of such genotype by environment interactions means that it is much more difficult to summarize the outcome of the selection applied. Here, it should be noted that these genotype by environment interactions should not be disregarded as they are an important property of the genotype. That is, the observed value of a particular variate is not necessarily the expression of the same part of the genome during all seasons but may reflect expression of different segments of the genome which are activated on account of the particular seasonal conditions. Where several variates have been recorded, such interaction can result in changes in the inter-relationships between the variates. To illustrate this, one can imagine an irregular p-dimensional surface, corresponding to the p variates measured, which varies its shape slightly as the environment changes (i.e. shrinkage on one axis may be associated with expansion on another axis). While this surface can be described by considering one or two variates at a time, clearly it would be far superior to consider all p dimensions simultaneously.

In summary, it can be seen that repeated t statistics, correlations (phenotypic, genetic and environmental), regressions and Turner's percent deviation have been used extensively to resolve the responses to selection, especially with regard to the behaviour of the fleece components during alternative breeding programs. Although informative, these approaches give no protection against either the effects of correlations among the subsets or the tendency for individual differences to be significant merely by chance as the number of variates increases. Multivariate statistical techniques should overcome these difficulties.

METHODS

(a) Background

Complete details of the sheep used, the selection methods and characters recorded have been given by Mayo, Potter, Brady and Hooper (1969) and in Chapters I and II of this thesis.

Briefly, the two flocks were raised at Roseworthy Agricultural College, between 1954 and 1965, either selecting rams by (i) visual appraisal alone (Visual method) or (ii) clean fleece weight after initial visual appraisal (Index method). The divergence in clean fleece weight of the Index over the Visual animals was previously established using t-tests. The following eight variates will be considered here: clean fleece weight, clean scoured yield percentage, body weight, staple length, crimp number, fibre diameter, primary follicle number and secondary follicle number. The other variates recorded were not considered of sufficient direct importance to fleece weight to warrant their inclusion. Only data from single born animals, for which all eight variates had been recorded, were used. All comparisons are made within sexes.

(b) Statistical analysis

The two samples are compared using Hotelling's T^2 as described in Morrison (1967). This test is basically a multivariate analogue of the square of the univariate t-statistic. Thus, two samples can be compared using

$$T^2 = \frac{N_1 N_2}{N_1 + N_2} (\bar{\underset{\sim}{y}}_1 - \bar{\underset{\sim}{y}}_2)' S^{-1} (\bar{\underset{\sim}{y}}_1 - \bar{\underset{\sim}{y}}_2)$$

where N_1 and N_2 are the number of observations in samples 1 and 2 respectively; $\bar{\underset{\sim}{y}}_1$ and $\bar{\underset{\sim}{y}}_2$ the corresponding (p x 1) mean vectors where p variates have been measured; and S is the p x p pooled estimate of the variance-covariance matrix. The critical region is given by

$$|T^2| \geq \frac{(N_1 + N_2 - 2)p}{(N_1 + N_2 - p-1)} \cdot F_{(p,N_1 + N_2 - p-1)}^{\alpha} = T^2\alpha(p,N_1 + N_2 - p-1)$$

where α is the significance level of the F statistic with p and $(N_1 + N_2 - p-1)$ degrees of freedom.

The mere significance of the T^2 statistic does not indicate which variates are likely to have led to the rejection of equality of the two mean vectors. Further, it would be erroneous to use univariate t-tests as the number of tests and correlations among the variates would distort the critical value chosen for the t-statistic. However, use of T^2 enables calculation of simultaneous confidence intervals for linear functions of the differences. That is, for any vector

$\underset{\sim}{a}' = [a_1, a_2, \ldots, a_p]$ the probability that *all* intervals,

$$\underset{\sim}{a}'(\underset{\sim}{\bar{y}}_1 - \underset{\sim}{\bar{y}}_2) - \sqrt{\underset{\sim}{a}'S\underset{\sim}{a}\,\frac{N_1 + N_2}{N_1 N_2}\,T^2\alpha(p, N_1 + N_2 - p-1)} \leq \underset{\sim}{a}'\underset{\sim}{\delta}$$

$$\leq \underset{\sim}{a}'(\underset{\sim}{\bar{y}}_1 - \underset{\sim}{\bar{y}}_2) + \sqrt{\underset{\sim}{a}'S\underset{\sim}{a}\,\frac{N_1 + N_2}{N_1 N_2}\,T^2\alpha(p, N_1 + N_2 - p-1)}$$

generated by different choices of the elements of $\underset{\sim}{a}$ are simultaneously true, is $(1-\alpha)$ (where $\underset{\sim}{\delta}$ is the vector of mean differences and $T^2\alpha(p, N_1 + N_2 - p-1)$ is the critical value defined in the preceding equation). By varying the form of $\underset{\sim}{a}$, a confidence interval can be calculated for each variate which indicates the magnitude of the differences between flocks. That is, for the vector $\underset{\sim}{a}'_i = [0,0,\ldots,0,1,0\ldots,0]$ with a one in position i and zeroes elsewhere, the asymmetry of these two bounds about zero indicates the direction of divergence for the ith variate. If zero is outside the interval we conclude at the $(1-\alpha).100$ percent joint significance level that the ith variate differs significantly between the two samples. Thus, by calculating bounds for each of the p variates we can establish the direction and magnitude of the differences between the two populations.

Two generalized FORTRAN subroutines have been written to apply this technique to large data sets. These are listed with calling instructions in Appendix A.

RESULTS AND DISCUSSION

In Tables III-1 and III-2, the value of Hotelling's T^2, associated value of F, and level of significance are presented along with the lower and upper bounds of the 95% simultaneous confidence intervals.

Inspection of Table III-1, which refers to the male progeny born during the twelve years of the experiment, indicates that T^2 is not significant for the first year (1954), but is significant thereafter, except for 1960. A similar pattern exists for the corresponding female progeny, shown in Table III-2, with the 1965 value being non-significant. (This latter anomalous result illustrates the well known inadequacy of discrete cut-off probabilities as the observed value (2.01) is extremely close to the critical value (2.03). As one would expect chance to lead to occasional acceptance of the null hypothesis when it is false (i.e. error of the second kind), it seems reasonable to conclude that the two flocks have diverged.

Assessment of the 95% simultaneous confidence interval indicates that the difference between flocks can seldom be associated with one character. (When considering these intervals it should be noted that these simultaneous confidence intervals have been expressed as the difference of the Index flock over the Visual flock.) Thus, for each variate the position of zero in the interval gives a good indication of the difference in response between the two methods of selection. If we look at the intervals for clean fleece weight we notice that the interval has shifted towards the positive end of the scale, indicating that the clean fleece weight of the Index flock exceeds that of the Visual flock (just as has previously been shown in the analysis of the single variate itself, see Mayo, Potter, Brady and Hooper (1969) and Chapter I).

The association between clean fleece weight and the other variates can be established by comparing the behaviour of the confidence intervals for the remaining variates. Thus, if we restrict our attention to staple length and crimps per inch we can see that staple length shows a similar pattern to clean fleece weight while for crimps per inch the confidence intervals become negatively biased indicating that for this variate the Visual flock exceeds the Index flock.

When all the variates are considered similarly, it can be seen that, for both sexes, staple length, clean scoured yield and secondary follicle number are positively associated with the increase in clean fleece weight, while crimps per inch and body weight are negatively associated with this increase in fleece weight of the Index flock over the Visual.

In contrast, the trend for fibre diameter was unclear, considerable variation between seasons being observed.

Comparing this to the literature, we recall:

i) Turner (1958) suggested fibre number and staple length were closely associated with clean fleece weight;

ii) Dun (1958) found fibre density and cross-sectional area were most important;

iii) more recently, Barlow (1974) concluded that the response in clean fleece weight observed in his fleece plus flock was due to fibre density, fibre diameter and staple length;

iv) Robards, Williams and Hunt (1974) reported that crimp frequency was related positively to live weight and negatively to clean fleece weight.

Thus, the present study literally summarizes, in one analysis, all the previous work but at the same time it illustrates the complex behaviour of fibre diameter. Since both Turner (1958) and Barlow (1974) have observed similar behaviour for this variate, and bearing in mind its importance in quality, further research on the effect of seasonal factors on fibre diameter could be rewarding.

In conclusion, Hotelling's T^2 provides a useful method for analysing selection experiments where two populations are involved. Clearly it has many advantages over previous methods, including those used in Chapters I and II of this thesis. Although it is not specifically designed to extract genetic information, as heritabilities and genetic correlations are, it seems difficult to imagine that in the present usage it is not essentially doing so, particularly since both flocks have been raised under the same environmental conditions. In particular, the output is quite compact and not complicated if the number of variates is increased.

Although the present demonstration has involved clean fleece weight in sheep the technique can be used for any selection experiment where several variates have been recorded.

In the present case Hotelling's T^2 has been used to compare two populations but it can be used equally well to compare one population against a particular mean vector. Thus, if only one flock had been selected we could test to see if it differed significantly from a particular set of mean values for the variates measured. Alternatively, if we wanted to test the equality of mean vectors for three or more populations, it would be best to use the multivariate analysis of variance described in Chapter 5 of Morrison (1967).

An extension of the above method of analysis has been published by Mayo, Murdoch and Hancock (1976). In this paper (which is included in Appendix B) the linear discriminant function, as suggested by Fisher (1936) has been used to examine the effect of paternal age, maternal age and birth rank on mutation.

TABLE III-1 Hotelling's T^2 and 95% simultaneous confidence interval for rams

Year	T^2	F	Sig	Body wt.		Yield		Clean fleece wt		Staple length		Crimps per inch		Fibre diameter		Primary foll. no.		Secondary foll. no.	
1954	8.05	0.95	NS	-3.30†	4.42†	-3.06	3.65	-.43	.61	-.83	.86	-1.19	.81	-2.17	1.42	-13.4	39.1	-274	872
1955	23.01	2.74	**	-4.59	2.22	-1.60	3.75	-.32	.37	-.46	1.03	-1.18	.94	- .71	2.09	-13.1	45.1	-318	975
1956	37.16	4.40	***	-3.81	2.96	-3.87	3.64	-.14	.48	-.18	1.20	-1.41	2.08	-2.61	.22	-15.0	51.7	-407	940
1957	25.31	3.00	**	-3.99	3.13	-2.79	3.71	-.27	.41	-.17	1.23	-1.18	1.08	-1.10	2.26	-21.8	60.6	-282	1204
1958	48.71	5.75	***	-2.93	4.29	.08	5.96	-.06	.75	-.24	1.36	- .95	1.41	-1.44	1.58	-20.0	50.6	-604	665
1959	21.60	2.57	*	-3.90	2.35	-3.30	4.05	-.18	.60	-.15	1.24	-1.48	.39	- .58	2.07	-24.1	29.2	-446	535
1960	10.16	1.19	NS	-6.21	2.48	-2.59	4.45	-.29	.58	-.44	1.17	-1.04	.89	-1.28	1.65	-24.5	31.6	-455	709
1961	40.26	4.78	***	-3.30	5.65	-1.24	6.04	.09	.89	-.18	1.35	-1.25	.14	-1.75	1.81	-42.6	21.3	-479	634
1962	59.76	7.15	***	-5.83	1.29	-1.25	4.54	-.21	.55	-.49	1.07	- .70	.84	-2.51	-.22	-35.3	33.8	- 73	1104
1963	24.41	2.89	**	-4.52	3.29	-3.40	3.85	-.16	.63	-.60	1.02	-1.19	.79	-1.73	.55	-39.3	12.2	-404	712
1964	55.12	6.55	***	-6.06	1.79	-1.77	4.81	.05	.78	-.31	1.24	-1.27	.36	-2.75	.44	-29.0	32.0	- 18	1076
1965	34.97	4.12	***	-4.80	3.50	-5.75	2.61	-.14	.65	-.30	1.16	-1.77	.16	-1.39	1.44	-21.3	44.7	-449	715

† Values shown are lower and upper limits of interval respectively.

TABLE III-2 Hotelling's T^2 and 95% simultaneous confidence interval for ewes

Year	T^2	F	Sig	Body wt.		Yield		Clean fleece wt		Staple length		Crimps per inch		Fibre diameter		Primary foll. no.		Secondary foll. no.	
1954	12.62	1.49	NS	-3.13†	3.62†	-4.10	1.76	-.46	.30	-.84	.84	-1.21	.56	-2.14	.97	-30.9	33.1	-391	669
1955	29.50	3.47	**	-1.30	4.28	-3.37	2.65	-.04	.54	-.02	1.41	-1.90	.34	- .95	1.75	-30.5	38.4	-611	738
1956	31.84	3.75	***	-3.61	3.16	-2.38	3.80	-.01	.58	-.17	1.10	-2.40	.65	-2.30	1.03	-38.4	42.3	-405	1084
1957	39.96	4.72	***	-3.44	1.96	- .42	6.20	-.07	.47	-.31	.95	-1.66	.21	- .93	2.35	-31.1	52.1	-268	1169
1958	22.16	2.60	*	-2.10	4.02	-1.35	6.39	-.05	.68	-.43	1.24	-1.50	.82	-2.49	.92	-35.2	39.1	-454	821
1959	57.72	6.87	***	-2.71	2.63	-1.82	4.98	-.01	.59	-.27	1.09	-1.36	.27	.26	2.95	-17.7	57.4	-442	697
1960	28.22	3.37	**	-4.26	1.28	-1.91	3.51	-.09	.57	-.44	.80	- .66	.81	-1.18	1.35	-25.4	33.4	-419	626
1961	26.38	3.15	**	-1.91	4.22	-1.20	4.45	-.07	.51	-.33	1.00	- .63	.45	-1.48	1.12	-37.5	23.3	-270	804
1962	89.53	10.68	***	-6.43	-.77	- .44	5.20	-.10	.55	-.49	1.03	- .86	.49	-2.49	.11	-31.1	35.8	142	1403
1963	28.61	3.35	**	-3.56	3.20	-3.12	3.55	-.12	.52	-.47	1.01	-1.23	.82	-2.77	.45	-40.8	29.6	572	580
1964	48.73	5.78	***	-6.43	.85	-1.73	4.18	-.04	.56	-.49	.85	-1.37	.22	-1.33	1.42	-39.5	49.9	-451	771
1965	17.26	2.01	NS	-6.25	1.79	-2.26	5.47	-.29	.49	-.64	.78	-1.72	.36	-1.45	1.26	-29.6	47.4	-613	916

† Values shown are lower and upper limits of interval respectively

IV - EXACT TEST FOR R x C CONTINGENCY TABLES

INTRODUCTION AND ALGORITHM

In the previous chapters the statistical methods considered have mainly been based on the assumption that the data is distributed over a continuous scale. While quantitative genetics generally relate to such continuous variation, researchers are often confronted with the analysis of data where the possible outcomes form a set of discrete classes (e.g. number of progeny). Thus, for variates like fecundity, while the inheritance may be considered polygenic, the outcome can only take one of several possible values.

The analysis of such variates often involves the interpretation of contingency tables with particular emphasis being directed at whether there is independence between pairs of attributes. (For example, if we wish to compare the lambing performance between two flocks, one of which has been selected for twinning, we can present the data in a 2 x 4 table of flock type by increasing levels of fecundity and test whether fecundity is independent of the previous selection history. If selection for multiple births had been successful we would expect to observe a greater proportion of these in the selected flock than in the normal flock. This shift in fecundity would result in a lack of independence between flock type and fecundity.) While there exists several approximate methods for analysing contingency tables, this approximation becomes poorer as the sample sizes decrease. Sugiura and Otake (1968) have compared eleven methods, or variations of methods, in an effort to resolve which, if any, provide the most accurate estimates of the significance levels. Although this investigation contributes information on the particular advantages or disadvantages of these methods, it cannot be denied that calculation of the exact probability would be far superior. (The description "approximate" used above refers to the fact that the actual distribution of the test statistic remains unknown. However, as its distribution can be shown to tend towards a known distribution for large sample sizes, this can be used to estimate the probability level of the test. This approximation clearly becomes poorer as the expected cell frequencies become smaller and therefore the use of the exact test is preferable as it does not require any such approximation to a sampling distribution. That is, the probability level for the exact test is found by calculating the probability of each possible combination of cell frequencies, given the marginal totals and comparing these with the probability of the observed frequencies.)

Fisher (1925), as shown in the 14th edition of his *Statistical Methods for Research Workers*, 1970, describes the method for the calculation of the exact probability of 2 x 2 tables. Fisher demonstrated that:

the probability of the 2 x 2 tables

a	b	a+b
c	d	c+d
a+c	b+d	N

where N is the sum of a + b + c + d, is given by the expression

$$P = \frac{(a+b)!\,(c+d)!\,(a+c)!\,(b+d)!}{a!\; b!\; c!\; d!\; N!}$$

This formula is then used to calculate the probability of all possible tables (i.e. for the various values of a, b, c and d) which could occur with the same marginal totals. The total probability of observing values as extreme or more so if the two attributes are unrelated is then found by summing all the probabilities which are less than or equal to the probability of the original table. If this probability is less than the chosen significance level it is concluded that the two attributes of interest are related.

Freeman and Halton (1951) have extended this method to enable the calculation of exact probabilities of R x C contingency tables. Thus, if we denote an R x C table as

$$\begin{array}{lll|l}
x_{11} & x_{12}\cdots\cdots\cdots & x_{1C} & X_{1\cdot} \\
x_{21} & x_{22}\cdots\cdots\cdots & x_{2C} & X_{2\cdot} \\
\vdots & \vdots & \vdots & \vdots \\
x_{R1} & x_{R2}\cdots\cdots\cdots & x_{RC} & X_{R\cdot} \\
\hline
X_{\cdot 1} & X_{\cdot 2}\cdots\cdots\cdots & X_{\cdot C} & X_{\cdot\cdot}
\end{array}$$

its probability is given by

$$P = \frac{\prod_{i=1}^{R} X_{i\cdot}! \prod_{j=1}^{C} X_{\cdot j}!}{X_{\cdot\cdot}! \prod_{i=1}^{R} \prod_{j=1}^{C} x_{ij}!}$$

The tail probability is found in a manner analogous to the above. (This fundamental similarity between the simple case for 2 x 2 tables and the more general case of R x C tables often leads to the latter being denoted as "Fisher's exact test for R x C contingency tables". Although there appears to be no evidence to suggest that Fisher was aware of the more general expression it seems unlikely that he would not have realized the possibility.)

While the theoretical value of the test has never been denied, it has not been used widely due to the practical difficulties involved in the identification of all the possible tables, given the observed row and column totals (also the subsequent difficulty of the above probability calculation due to the presence of the factorial terms has further limited its use). Clearly the number of possible tables increases markedly with the dimensions of the table and the magnitude of the individual cell frequencies (x_{ij}'s). Thus, manual calculation is impossible for anything larger than 2 x 2 or 2 x 3 tables with small cell frequencies.

Finney (1948); Pearson and Hartley (1958); Finney, Latscha, Bennett and Hsu (1963); and Bennett and Nahamura (1963), to name just a few, have presented tables which indicate the significance (or not) of a particular contingency table. The presentation of these tables extended the feasibility of the test, however, it still remained limited to 2 x 2 or 2 x 3 tables with small cell frequencies. The development of the computer has further extended the test's capabilities. Thus, Robertson (1960) and Sokal and Rohlf (1969) both present programs for the calculation of Fisher's exact test for 2 x 2 tables. While these first two attempts provided additional convenience it was not till March (1972) attempted to provide a general algorithm for R x C contingency tables that the potential of the computer was established. March's algorithm, which was based on the previously mentioned method of Freeman and Halton (1951), used a subtraction process to generate all the possible R x C arrays given the marginal totals. Although modern computers are often considered to be exceptionally fast per operation the cumulative effect of the generation of the large number of possible arrays is more than enough to over-tax the capabilities of even the most advanced computer. That is, the nature of the sorting process is such that while it would be foolish to attempt it without the aid of a computer, it nevertheless remains a formidable task even in the presence of a computer (for details on the limitations of modern computers see Dahl, Dijkstra and Hoare (1972) or Knuth (1973). Therefore, there remains an upper limit beyond which it is impracticable to calculate this test. Improvements in machine capabilities may raise this upper limit but will do little towards actually removing it. The presence of this "hardware ceiling" then leaves the researcher with three possible strategies:

i) if the expected cell frequencies are "sufficiently large" the approximate tests can be used.

ii) The rows and/or columns of the table can be pooled.

iii) The method of computation (algorithm) may be improved.

Clearly options i) and ii) need only be contemplated when the algorithm fails to handle, within a reasonable time, the particular contingency table. Under these circumstances, either or both of these two options may be applied, the choice depending on the given table. However, as use of options i) and ii) will weaken the inferences, improvement of the algorithm should remain of first priority to the prospective researcher.

Implementation of March's algorithm demonstrates that it can effectively cope with 3 x 4 arrays with small cell frequencies but for anything larger it is hopelessly slow. Inspection of the algorithm shows that while much of it is efficiently programmed the method of generating all the possible arrays contains much room for improvement. As mentioned earlier, March's algorithm uses a method of subtraction. This results in the generation of many arrays which later prove inappropriate when compared to the row and column totals of the original array.

Two independent attempts have subsequently been reported which enable considerable improvements over March's procedure. As these two improved algorithms were developed simultaneously and use differing procedures, I will consider both in turn before looking at their respective advantages and disadvantages.

The first of the improved algorithms to be published was that presented by Boulton (1974). The author makes the following comment with respect to March's original algorithm:

> "The method used to generate all the cell frequency combinations is rather inefficient as it operates by generating all combinations which satisfy a weakened set of constraints and then rej-

> ecting those combinations which violate the actual marginal sum constraints. As the number of combinations rejected very often far exceeds the actual number accepted, the process is very wasteful."

To overcome this weakness Boulton uses a more efficient generation process previously described by Boulton and Wallace (1973). This method generates only the arrays which satisfy the marginal totals. Table IV-1, reproduced from Boulton (1974), indicates the modified algorithm was always faster than March's algorithm. (The times shown enable comparison between the two algorithms but clearly their actual value will depend on the computer used.) The full listing of Boulton's algorithm was not published but has been obtained by personal communication with Boulton.

The second improved algorithm has been published by Hancock (1975). As the content of this paper directly relates to the present thesis it has been included in full in Appendix C. This paper presents details of the method of computation and also a listing of the associated algorithm.

Like Boulton's algorithm, the generation procedure avoids any arrays which are incompatible with the row and column totals. This is accomplished by calling the subroutine INIT which generates the next 'legal' array which conforms to the searching pattern and satisfies the marginal totals. While the successive 'legal' arrays can be quite dissimilar, Freeman and Halton (1951) point out that occasionally sequences arise in which there are only minor differences (i.e. only two of the cell frequencies change). Clearly the probability calculation can be greatly simplified for such cases. The algorithm has been written to recognise these instances and therefore avoid calling the more complicated generation subroutine (INIT) while exploiting the simpler

Table IV-1 Times† for evaluating probabilities
(reproduced from Boulton, 1974)

Contingency table					Probability	Time (secs)	
						Original	Improved
8	12	(20)			.05767116	.026	.013
8	2	(10)					
(16)	(14)	(30)					
5	3	3	0	(11)	.35262364	.290	.095
2	3	1	2	(8)			
(7)	(6)	(4)	(2)	(19)			
5	1	0	0	(6)	.10625089	3.31	.510
1	1	2	1	(5)			
0	1	1	1	(3)			
(6)	(3)	(3)	(2)	(14)			
2	0	0	0	(2)	.12380952	13.9	.693
0	1	0	1	(2)			
0	0	2	0	(2)			
0	1	0	1	(2)			
(2)	(2)	(2)	(2)	(8)			

† times shown were obtained when the two programs were run on a CDC 3200 machine.

probability calculation. The efficiency of this algorithm over March's earlier version is shown in Table IV-2. (This table has been reproduced from Hancock (1975). To simplify the comparison, all R x C arrays have been chosen with all cell frequencies equal to one, i.e. $X_{ij} = 1$, where $i = 1,2,\ldots,R$ and $j = 1,2,\ldots,C$.) The modified algorithm can be seen to become increasingly more valuable as the complexity of the array increases. This advantage can be increased further by the insertion of a statement of the form

```
IF(PS.GT.0.1.AND.PC.LT.0.9)RETURN
```

which results in the termination of the algorithm when the tail probability (PS) exceeds ten percent provided PC, the cumulative probability of all arrays processed, is less than ninety percent. This latter condition has been included to avoid termination of the algorithm when it is close to completion.

Clearly both new algorithms provide worthwhile improvements on March's earlier algorithm. It then remains to establish whether either of these algorithms can be shown to be superior to the other. To answer this question I have compared the two algorithms using a CDC 6400 machine. Table IV-3 lists the respective CP (Central Processor) times required for the arrays shown. (These arrays have been taken from Boulton (1974) and Hancock (1975).)

Boulton's algorithm can be seen to be the faster method. However, extended use of the algorithms has shown that cases exist where this order is reversed, e.g. for the 2 x 5 array (which arose in a consulting problem in another context).

286	14	10	17	2	329
731	41	28	72	9	881
1017	55	38	89	11	1210

Table IV-2 Comparison of Subroutines (CP time required in seconds to compute exact probabilities for R x C contingency tables; where all cell frequencies are chosen equal to one. These are presented to illustrate the relative improvement of R x CPRB over CONP. Obviously the actual times will depend on the machine used).

R x C	CONP (by March)	R x CPRB
2 x 2	.019 (3†)	.018 (3†)
2 x 3	.012 (9)	.010 (7)
3 x 2	.018 (9)	.016 (7)
2 x 5	.073 (81)	.054 (51)
5 x 2	.093 (81)	.055 (51)
3 x 3	.110 (256)	.055 (55)
3 x 4	1.279 (4096)	.509 (415)
4 x 3	1.344 (4096)	.514 (415)
4 x 4	Unknown*	15.495 (10147)

† Number of matrices attempted in the calculation

* Computation was still incomplete after 500 seconds

(Reproduced from Hancock, 1975)

Table IV-3 Comparison of Subroutines. (CP time in seconds required to compute exact probabilities for R x C contingency tables.)

Contingency table				R x C	RXCPRB (by Hancock)	CONP (by Boulton)
8	12			2 x 2	0.014	0.016
8	2					
5	3	3	0	2 x 4	0.032	0.031
2	3	1	2			
5	1	0	0	3 x 4	0.308	0.154
1	1	2	1			
0	1	1	1			
2	0	0	0	4 x 4	0.554	0.175
0	1	0	1			
0	0	2	0			
0	1	0	1			
	(1)†			2 x 2	0.005	0.003
	(1)†			2 x 3	0.008	0.006
	(1)†			3 x 2	0.007	0.006
	(1)†			2 x 5	0.049	0.024
	(1)†			5 x 2	0.054	0.024
	(1)†			3 x 3	0.054	0.026
	(1)†			3 x 4	0.509	0.187
	(1)†			4 x 4	15.118	4.616

† All cell frequencies of the R x C tables have been set as 1, i.e. $X_{ij} = 1$, where i = 1,2,..,R and j = 1,2,...,C

Boulton's algorithm is unable to find a solution in over 500 seconds CP time whereas the alternative method does so in 277.6 seconds. Therefore, it must be concluded that neither method can be described as better than the other as this depends on the contingency table under consideration.

APPLICATION

Clearly the extended capabilities of the exact test enable it to be used widely to analyse data presented in the form of contingency tables. The Roseworthy data contains many such comparisons. However, in the following I will restrict myself to consideration of only one of the more relevant aspects of the breeding program to illustrate the method.

Mayo, Potter, Brady and Hooper (1969) concluded that the fertility of the Index flock was markedly lower than for the Visual flock. This conclusion was based on the fertility statistics, which were presented in their table I, however, no statistical justification was attempted. As mentioned earlier in this thesis, reduction in fitness is often associated with directional selection for a non-reproductive variate. Clearly any lowering in fertility would be extremely undesirable in the breeding of farm animals. If such an effect was operating in the Roseworthy experiment then the greater selection intensity applied for clean fleece weight in the Index flock would be expected to result in the fitness of this flock being lower than for the Visual flock. Use of the exact test provides a convenient method for establishing whether this has occurred. That is, if the reproductive data is presented in a 2 x 5 contingency table, formed from the two flocks by five offspring classes, analysis of this will identify any fertility differences between the two flocks. Thus, if we look at the 1954 lambing we observe the following contingency table of ewe performance:

	Reproductive performance					
	0	1	2	3	4	
Index flock	41	126	9	28	0	204
Visual flock	34	134	11	26	1	206
	75	260	20	54	1	410

where 0 denotes a dry ewe,

1 denotes the birth and rearing of one offspring,

2 denotes a twin birth but subsequent rearing of one offspring,

3 denotes the birth and rearing of twins,

4 denotes any other type of birth.

As the calculated tail probability from the exact test is 0.7627 this indicates that there is no measurable difference in reproductive performance between the two flocks. (It will be noticed that the birth of triplets in the 1954 Visual flock has necessitated a 2 x 5 contingency table, however, for most years only a 2 x 4 table will be required.)

Similar analyses have been carried out on the remaining eleven years of the experiment and these are summarised in Table IV-4. Inspection of the tail probability (i.e. the probability of obtaining an array as deviant or more so than the observed table) indicates that the flocks differed significantly in 1956, 1960, 1961 and 1965. Although the significance of several of these tables supports the hypothesis that the Index ewes had inferior reproductive performance compared to the Visual ewes, it can be seen that in 1960 the number of Index ewes giving birth to a single offspring is larger than expected. However, if the total number of ewes lambing is considered the 2 x 4 table reduces to a 2 x 2 with the following values:

Table IV-4 Analysis of Reproductive Data

Year	Contingency table†							Probability of given table	Tail probability
		0	1	2	3	4			
1954	I	41	126	9	28	0	204	.000619	.762717
	V	34	134	11	26	1	206		
		75	260	20	54	1	410		
		0	1	2	3				
1955	I	48	133	11	27		219	.001493	.894207
	V	51	137	8	26		222		
		99	270	19	53		441		
		0	1	2	3	4			
1956	I	83	112	4	24	1	224	.000010	.015535
	V	65	144	4	12	0	225		
		148	256	8	36	1	449		
		0	1	2	3				
1957	I	82	123	5	20		230	.000078	.064269
	V	60	146	7	13		226		
		142	269	12	33		456		
		0	1	2	3				
1958	I	73	124	17	10		224	.000302	.317049
	V	66	122	27	15		230		
		139	246	44	25		454		
		0	1	2	3				
1959	I	53	146	6	35		240	.000703	.568101
	V	46	146	11	32		235		
		99	292	17	67		475		

Table IV-4 - continued

Year	Contingency table†	0	1	2	3		Probability of given table	Tail probability
1960	I	41	142	6	3	192	.000044	.026579
	V	36	129	17	9	191		
		77	271	23	12	383		
1961	I	65	141	1	1	208	.000001	.000081
	V	41	159	0	14	214		
		106	300	1	15	422		
1962	I	56	152	3	17	228	.001906	.560389
	V	62	167	1	13	243		
		118	319	4	30	471		
1963	I	91	107	4	18	220	.000612	.413344
	V	87	139	7	21	254		
		178	246	11	39	474		
1964	I	42	150	6	26	224	.000630	.460355
	V	45	134	8	35	222		
		87	284	14	61	446		
1965	I	90	83	7	11	191	.000000	.000000
	V	35	147	4	8	194		
		125	230	11	19	385		

† See text for explanation of row and column classes

	Ewes without lambs	Ewes with lambs	
Index	41	151	192
Visual	36	155	191
	77	306	383

which supports the null hypothesis of no interaction.

McGuirk (1976) has divided reproductive performance into its two fundamental components, i.e. fertility and fecundity. Such an approach could have been adopted here, by dividing the information into 2 x 2 tables for fertility and 2 x 3 tables for fecundity and testing each using the exact test. It was felt that this was unnecessary for this set of data although it can be seen that for the 1960 data this may have been of marginal advantage. In fact, when all data were analysed separately for fertility and fecundity, the results were consistent with those shown for reproductive performance (except, that is, for the 1960 data explained above). Although the data have been considered as relating to the reproductive performance of ewes it should be remembered that it may also reflect ram performance. Thus, the low lambing percentage for the Index flock in 1965 may be associated with failure of one or more rams. As the number of rams used per flock is small (i.e. 6 two-toothed and 1 four-toothed) it is difficult to determine whether such failure is indicative of this flock or is simply due to chance.

As the cumulative probability of all the possible combinations of cell frequencies which satisfy the marginal totals should sum to one, checking of this provides a simple test of the accuracy of the algorithm. In the present application this probability was accurate to at least nine significant figures for all the tables analysed. Thus, the test can be seen to provide a convenient method for analysing contingency tables.

In particular, because it is an exact test, the user does not have to be concerned with any of the limitations which plague the use of approximate tests (e.g. expected sample size and Yate's correction for continuity).

V - SELECTION OF THE "BEST" SUBSET OF LINEAR VARIABLES

INTRODUCTION

The problem of selecting the best subset (or subsets) of independent variables during multiple regression has not only received much attention in the past but still demands a good deal of interest (see Cox and Snell, 1974; Kendall, 1975; Draper and Smith, 1966; Hoerl and Kennard, 1970a & b; and Hocking, 1976). While this interest has occurred mainly in the field of statistics the approach also has a valuable contribution to make in quantitative genetics. Today's plant or animal breeder is well aware that he cannot select for a single character as this will inevitably lead to changes in other characters (i.e. continued selection not only modifies the heritability of the particular character undergoing selection but also leads to changes in the genetic correlations with others). Therefore, one must consider the effect of selection on all relevant characters not just the character being selected. The establishment of which characters are relevant is, and will remain, a vital responsibility of the plant or animal breeder.

Although selection for several characters simultaneously (i.e. using either independent culling, tandem selection or selection index - see Young (1961) for discussion of these three methods) may provide a worthwhile advance over univariate selection, the result will still be exposed to the limitations outlined above. Thus, the selection of the 'best' subset of variates may be of value in indicating which variates are going to change following selection for another character, and therefore should be measured at each interval of selection, or alternatively, which variates should be selected for simultaneously using a selection index. Traditionally this has been accomplished by the use of heritabilities and genetic correlations combined with the breeder's knowledge of the plant or animal under consideration. However, it would appear that 'best' subset techniques may improve this existing approach. (Note that heritabilities and genetic correlations are treated earlier in Chapter II.) In particular, the Roseworthy data provide an ideal example on which the value of the various techniques can be assessed.

During most large scale selection programs, like the Roseworthy experiment, one character is often of prime importance while several other related characters warrant consideration. More specifically, we have clean fleece weight which has been selected for but we are also interested in the behaviour of other characters which we expect to be related to clean fleece weight. Thus, in terms of the multiple regression model, clean fleece weight is the dependent variate y and the other characters the independent variates $x_1, x_2, \ldots, x_k$. Then the multiple regression model is

$$y_i = \alpha + \beta_1(x_{1i} - \bar{x}_{1\cdot}) + \beta_2(x_{2i} - \bar{x}_{2\cdot}) + \ldots + \beta_7(x_{7i} - \bar{x}_{7\cdot})$$

where $i = 1, \ldots, n$ (i.e. the number of observations)

and x_1 = body weight

x_2 = percent yield

x_3 = staple length

x_4 = crimps per inch

x_5 = fibre diameter

x_6 = primary follicle number

x_7 = secondary follicle number

(although the use of the word independent in regression statistics implies that the variates can be measured without error it will be used here, in a more general sense, to denote the variables which describe or predict the behaviour of the dependent variate (y).) Similar multiple regression models occur in other selection programs, e.g. carcase weight may be related to body dimensions, or plant yield to plant height, shape, number of tillers, etc.

Obviously, all the independent variates can be retained but this often results in a multiple regression equation which is far more complicated than necessary. (Although the Roseworthy sheep data have been confined to the seven continuous variates which are well behaved with respect to the underlying assumptions of normality, one can readily imagine cases where twenty or more variates may be relevant.) Apart from the sheer bulk of such multiple regression equations it is important to appreciate the following comment made by Allen (1974) on the hazards of using too many variables:

> "...the addition of a variable to a least squares prediction equation almost always increases (never decreases) the variance of a predicted response."

Clearly this limitation had been recognised long before Allen's statement as can be seen in the following (remarkably similar) statement by Walls and Weeks (1969):

> "...the addition of a variable to a regression equation can never reduce (and in fact usually increases) the variance of a predicted response."

Thus, there is a good deal to be gained by discarding those variates which are unimportant (where unimportant can mean that either they show insignificant association with the dependent variate or their relationship is adequately described by one or more other variates which have been retained). Clearly to do this one requires a stopping rule to choose a subset of p important variates from the k variates given.

The following list contains the more noteworthy methods for identifying the important aspects from among a set of related variates:

i) Forward Selection (or Step-up).

ii) Backward Elimination (or Step-down).

iii) Step-wise Regression

iv) Optimum Regression (or Best Subset) - minimization of the residual sum of squares for each p where $1 < p < k$ (or maximization of R^2).

v) Regression by Leaps and Bounds - using Mallow's C_p statistic.

vi) Regression by Leaps and Bounds - using adjusted R^2.

vii) Mean Square Error of Prediction (MSEP).

viii) Ridge Regression.

ix) Principal Component Analysis.

x) Factor Analysis.

The first three methods all test the partial correlation coefficient before deciding whether a variate should be included or excluded. Forward selection begins with the empty set and at each successive step includes that variate, from among those not previously included, which is most closely associated with the dependent variate (i.e. having the largest partial correlation coefficient) until all the variates demonstrating significant association are included. Backward Elimination, as the name implies, starts with the full-set and excludes at each step the non-significant variate showing least association, until only significant variates remain. The third (Step-wise Regression) is simply a combination of these two methods in that Forward Selection is applied with backward reflection (i.e. following inclusion of a new variate the subset is checked to see if any variates have become obsolete). All three of these methods suffer from the following faults:

i) the final subset depends on the choice of significance level (α). As α decreases, so does the size of the subset. The Type I error rates resulting from the sequences of dependent F tests used in Forward Selection, Backward Elimination and Step-wise Regression remain unknown, despite much effort to find solutions (see Draper, Guttman and Kanemasu, 1971; Pope and Webster, 1972; and Aitken, 1974).

ii) When equally valuable subsets of variates exist these procedures will select one arbitrarily. If two subsets are equally informative, or nearly so, then it is best to use both, not one alone.

In practice, Forward Selection and Step-wise Regression give similar results whereas Backward Elimination, which begins with the full set, is likely to arrive at a different final subset from the other two. However, despite these limitations, these techniques, widely used in the past, seem to remain popular.

The fourth method listed above (Selection of the Best Subset) was considered by Hocking and Leslie (1967) for cases where the total number of independent variates (k) is less than 10. The value of each of the $2^k - 1$ possible subsets is then assessed by comparing the residual mean squares, firstly within each subset size (p) and secondly between subset sizes.

The best subsets are those whose residual mean squares are minimal. (Maximization of the multiple correlation coefficient (R^2) where

$$R^2 = 1 - \frac{\text{Residual sum of squares}}{\text{Total sum of squares}}$$

is often used to give similar results.) To assist the computation, Hocking and Leslie (1967) suggested an algorithm, but Lamotte and Hocking (1970) later concluded that, "Subsequent experience with the Hocking-Leslie algorithm revealed that the amount of computation required was unacceptable for $k > 15$." Further extensions to the Hocking-Leslie (1967) algorithm were subsequently proposed (by Lamotte and Hocking, 1970) to minimise the computation for moderately large problems where 'moderately large' was considered by Lamotte to be for k between 40 and 50 (although these authors acknowledged that they had only applied it to cases where $k \leq 30$). I find these claims (of $k \leq 50$) rather excessive but like the above mentioned authors lack specific evidence to support my belief.

It is important to remember that the minimum residual mean squares will always occur for the full-set with the values for the 'good subsets' approaching this minimum value but always exceeding it. Thus, this technique will choose the better subsets for each value of p (where $1 < p \leq k$) but fails to give an upper limit to the size of the residual mean square (RMS) which might be used to partition the 'good subsets' from the 'rest'. Furthermore, given two subsets having the same RMS then the subset with the smaller value of p will generally be of greater interest, but no adequate weighting has yet been found to clarify the case when the RMS for the larger subset is just smaller than the RMS for the smaller subset. Although the hope of finding a truncation rule and/or a weighting for subset size may appear reasonable at first, I believe it is an intuitive attempt to over-simplify a difficult statistical situation. Thus, in biology it is generally far more informative to consider all 'good' subsets than to restrict one's attention to one subset only, especially as this latter choice may be due far more to chance than to any property of the subset.

Recently, Diehr and Hoflin (1974) used Monte Carlo techniques to provide a method for testing whether an observed R^2 (for p variates chosen from k) is significant. However, the author admits this does not resolve the stopping rule as it provides no means for comparing R^2 from different subsets. Also, as the simulation was based on M + 1 independent variates, it is unclear how the method extrapolates to the situation when the variates are correlated.

Many scientists imply that R^2 provides a valuable measure of the 'usefulness' or 'goodness of fit' of a regression equation but Barrett (1974) has pointed out an important limitation. The previous definition of R^2 can be re-written as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_.)^2}$$

where y_i, $i = 1,\ldots.,n$, are the n observations of the dependent variates and $\hat{y}_i$, $i = 1,\ldots.,n$, are the corresponding predicted values on fitting the regression equation. If the vertical distance of the data points to the regression surface (i.e. the 'goodness of fit' $\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$) is held constant, then the steeper the regression the larger $\sum_{i=1}^{n} (y_i - \bar{y}_.)^2$ will become, thus increasing the value of R^2.

Clearly this limitation can only complicate the use of R^2 for comparison between the regression equations for different subsets of the independent variates.

Methods (v) and (vi) are extensions of method (iv) in which further attempts are made to improve the stopping rule. Both Mallow's C_p, of method (v), and the adjusted R^2, used in method (vi), are functions of the residual mean square. However, while the RMS decreases monotonically as p increases from 1 to k, these two statistics exhibit maxima (or minima) associated with the 'best' subsets.

When fitting p parameters to n data points, the adjusted R^2 statistic is defined as

$$R_a^2 = 1 - \frac{[1 - R_p^2]}{[n - p]} . n$$

where R_p^2 is the multiple R^2 based on p parameters (i.e. $R_p^2 = 1 - \frac{RSS_p}{TSS}$ where RSS_p = Residual sums of squares when p parameters fitted, and TSS = Total sum of squares). Replacing p/n, the ratio of parameters fitted to total number of observations, by f gives:

$$R_a^2 = 1 - [1-R_p^2][1/(1-f)]$$

which provides a means of comparing subsets both within and between values of p. The 'best' subsets are naturally those which have the largest R_a^2.

Mallow's C_p statistic, as published by Gorman and Toman (1966), can be presented as

$$C_p = (RSS_p/\hat{\sigma}^2) - (n - 2p),$$

where $\hat{\sigma}^2$ is the best estimate of σ^2. (Generally the residual mean square, when all k parameters are included, is used to estimate σ^2.) Alternatively, C_p can be written as

$$C_p = (n-p) \left[\frac{RSS_p/(n-p)}{\hat{\sigma}^2} - 1\right] + p$$

from which it can be seen that for the 'best' subsets C_p will approach p from above.

While it is advantageous to choose subsets with p small it should be remembered that such sets of variables will be biased if the omitted variables are at all relevant (Narula, 1974). However, Mallow's C_p statistic, unlike many of the other tests, indicates the extent of the bias and thus would be expected to lead to a better choice of subsets (see Gorman and Toman, 1966 or Cox and Snell, 1974). Indeed, Hocking (1972) points out that in practice Mallow's C_p tends to give larger subsets than other techniques. Kennard (1971) has shown that there is a one-to-one correspondence between Mallow's C_p and the adjusted R^2. Thus, these two statistics would be expected to 'choose' similar subsets.

As well as providing an improved criteria for assessment of subsets, methods (v) and (vi) have been further developed by Furnival and Wilson (1974) to minimize the amount of computation required to find the 'best' subsets. This has been accomplished by (1) minimising the amount of computation involved in examining each subset and (2) by finding procedures which establish the 'best' subsets without examining all possible subsets. Furnival and Wilson (1974a) use the name "Regression by Leaps and Bounds" to describe the algorithm they have developed to scan over the subsets. This algorithm can use either the multiple R^2, the adjusted R^2 or Mallow's

C_p as its selection criterion.

The mean square error of prediction (MSEP - method (vii)) has been proposed by Allen (1971) as a criterion for selecting variables. For any set of data the multiple regression model can be written as

$$\underset{n\times 1}{\underset{\sim}{y}} = \underset{(n\times k)}{X'} \underset{(k\times 1)}{\underset{\sim}{\beta}} + \underset{\sim}{\varepsilon}$$

where $\underset{\sim}{y}$ is a vector of n observations of the dependent variate; X is a (k × n) matrix formed from the k independent variates observed at the n observations; $\underset{\sim}{\beta}$ is the vector of k regression coefficients; $\underset{\sim}{\varepsilon}$ is an N-dimensional random normal vector with mean zero and covariance matrix $\sigma^2 I_N$. Using the n observations we can estimate $\underset{\sim}{\beta}$ by $\underset{\sim}{b} = (X'X)^{-1}X'\underset{\sim}{y}$ and this can be used to predict a future value $\hat{y}$,

$$\hat{y} = \underset{(1\times k)}{\underset{\sim}{x}'} \underset{(k\times 1)}{\underset{\sim}{b}}$$

where $\underset{\sim}{x}'$ is a vector containing the values of the k independent variates for which the predicted value is required. Then the mean square error of prediction (MSEP) is given by

$$E(\hat{y} - y)^2 = \sigma^2 + \mathrm{Var}(\hat{y}) + [E(\hat{y}) - x\beta]^2$$

where y is the actual value of the dependent variate. The 'best' subset is obtained by finding the p independent variates out of k which minimises the MSEP. Thus, given a set of n observations and one subsequent observation, the MSEP can be used to 'choose' p independent variates; however, for the prediction of more than one subsequent observation, Allen suggests repeating the process for each further observation. Clearly this method would be tedious for such a case therefore the MSEP will generally be less desirable than the other methods mentioned above.

Method (viii), Ridge Regression, has been included in the above list as it provides a reasonable procedure for obtaining a better multiple regression equation than is obtained by ordinary least squares. Hoerl and Kennard (1970a & b) have shown that the parameter estimates in multiple regression "have a high probability of being unsatisfactory, if not incorrect", if the independent variates are not orthogonal (i.e. when the X'X matrix approaches singularity). The Ridge Regression procedure is based on the addition of small quantities to the diagonal elements of X'X. Thus the matrix $[X'X + cI_k]$ (where c is a small constant, and $c \geq 0$) is used in place of X'X in the estimation of the parameters by least squares. By repeating this estimation for several different values of c a two dimensional plot ("the ridge trace") can be generated which provides a simple means of assessing the stability of the k parameters. Thus, this method may be usefully applied to all k variates, but it does not provide a convenient method for choosing the 'best' subsets. In fact, for this latter case it is unlikely that Ridge Regression would be of advantage as the p 'best' independent variates should in any case lead to an X'X matrix which is far from singular.

The ninth method to be considered is Principal Component Analysis (PCA) which originated with K. Pearson (1901) as a method of fitting planes by orthogonal least squares. Hotelling (1933 and 1936) later proposed the method for the particular purpose of analysing correlation structures. The value of this method can be briefly summarised in the words of Dubzinski (1975):

> "When the causal relationships between the dependent variable Y and the independent variables X need to be explained and interpreted, multiple regression is frequently unsatisfactory if the X variables are, as is often the case, highly correlated. PCA may transform the independent variables into fewer biologically meaningful variables created from linear combinations of the original variables."

Thus, the method has been specifically proposed for identifying the dependence structure of a multivariate population. That is, by using it one can establish the complex of factors which have generated the dependence or variation in the data, so that the observable or manifest variates can be represented as functions of a smaller number of latent factor variates.

Continuing the notation of the previous section where x is the $n \times k$ matrix with mean vector μ and covariance matrix Σ then the k characteristic roots (or eigen values) can be found

i.e. $\lambda_1 > \lambda_2 > \ldots > \lambda_k$.

If the k response variates have been recorded in similar units then the $(k \times k)$ sample covariance matrix $\underset{\sim}{S}$ is used to estimate Σ, where

$$\underset{\sim}{S} = \frac{1}{(n-1)} \sum_{i=1}^{n} (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})'$$

However, if the variates have been recorded in dissimilar units then it is best to use the sample correlation matrix $\underset{\sim}{R}$ which is obtained by the transformation

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

Then associated with the estimates of the eigen values $\ell_1 > \ell_2 > \ldots > \ell_k$ there exist the eigen vectors (or characteristic vectors) $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \ldots, \underset{\sim}{a}_k$ (i.e. $\underset{\sim}{a}_i$ is a $k \times 1$ column vector). It can be shown that these two together constitute the fundamental structure of PCA. Thus the first principal component (1st PC) is the following linear function of the responses:

$$Y_1 = a_{11}X_1 + \ldots + a_{k1}X_k = \underset{\sim}{a}_1' \underset{\sim}{x}$$

with sample variance

$$S_{y1}^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} a_{i1}\, a_{j1}\, S_{ij} = \underset{\sim}{a}_1' \underset{\sim}{S}\underset{\sim}{a} = \ell_1$$

Here we note that ℓ_1 is the largest eigen value and the eigen vector $\underset{\sim}{a}_1$ is chosen such that $\underset{\sim}{a}_1' \underset{\sim}{a}_1 = 1$ (i.e. $\underset{\sim}{a}_1$ is normalised).

The second principal component is

$$Y_2 = a_{12}X_1 + \dots + a_{k2}X_k = \underset{\sim}{a}_2' \underset{\sim}{x}$$

where the coefficients a_{j2} (the elements of the second eigen vector $\underset{\sim}{a}_2$) have been chosen such that

i) $\underset{\sim}{a}_2' \underset{\sim}{a}_2 = 1$ (i.e. $\underset{\sim}{a}_2$ is normalised),

and ii) $\underset{\sim}{a}_1' \underset{\sim}{a}_2 = 0$ (i.e. $\underset{\sim}{a}_1$ and $\underset{\sim}{a}_2$ are orthogonal).

The remaining k - 2 principal components are defined in a similar fashion. Thus the jth principal component is

$$Y_j = a_{1j}X_1 + \dots + a_{kj}X_k = \underset{\sim}{a}_j' \underset{\sim}{x}$$

where

$$\underset{\sim}{a}_i' \underset{\sim}{a}_j \begin{cases} = 1 \text{ if } i = j \\ = 0 \text{ if } i \neq j. \end{cases}$$

Further, the algebraic sign and magnitude of the a_{ij} indicate the direction and importance of ith variate to the jth component. As the eigen vectors are orthogonal, the importance of the jth principal component is given by

$$\frac{\ell_j}{\mathrm{tr}S}, \text{ where the total variance } \mathrm{tr}S = \sum_{i=1}^{k} \ell_i$$

The sample covariance of the variates with the jth component are given by the column vector,

$$S \underset{\sim}{a}_j$$

but by definition $(S - \ell_j I)\underset{\sim}{a}_j = \underset{\sim}{0}$

$$\therefore \quad S\underset{\sim}{a}_j = \ell_j \underset{\sim}{a}_j$$

and the covariance of the ith variate with Y_j is $\ell_j a_{ij}$. The product moment correlation (or component correlation coefficient) of the ith variate and the jth component is obtained on dividing by the component and response standard deviations i.e. component correlation coefficient $= \dfrac{a_{ij}\sqrt{\ell_j}}{S_i}$

However, if the principal components have been calculated from the correlation matrix (R) then the correlation of the variates with the jth component is given by

$$\sqrt{\ell_i}\, \underset{\sim}{a}_j \; .$$

This form of weighting is generally used when presenting components.

In conclusion, principal component analysis partitions the total variance into successively smaller orthogonal proportions, and provided these components are distinct (i.e. $\ell_i \neq \ell_j$), the resulting set of coefficient vectors is unique.

Some workers ignore this important property of the component coefficients and proceed to use transformations to simplify their interpretation (e.g. Daling and Tamura, 1970). Indeed, Morrison (1967) gives the following warning on this extension of the technique: "While the ability of the vectors to generate the original S (or R) is unimpaired their components no longer have the maximum variance property," (i.e. the jth component no longer explains $\dfrac{\ell_j}{\sum_{j=1}^{k} \ell_j}$ of the variation). As such subsequent transformations (or rotations) of the components (which are generally denoted by the name Factor Analysis and included above as method (x)) may in fact distort the interpretation, no further consideration will be given here to their application to the selection of the 'best' subset although some investigations may still maintain their suitability.

From this outline of the theory of Principal components it can be seen that it provides information relevant to the selection of the 'best' subset but that this information is obtained in a fashion which differs markedly from the other methods previously examined. Instead of choosing one subset (or possibly several subsets) of the variables, we here arrive at orthogonal components which describe the factors underlying the covariance or correlation matrix. As the k components explain progressively less of the total sample variance we can initially test to find the first r components which make meaningful contributions, and then test that the r components are in fact distinct. If components i and i + 1 (where i = 1,....,r-1) are not significantly different then their associated eigen vectors $\underset{\sim}{a}_i$ and $\underset{\sim}{a}_{i+1}$ would not be unique and therefore should not be considered further. Although I shall give no more specific details of these tests, it is important to realize that although they can be applied when the covariance matrix S has been used in the calculation their derivation is more limited when the correlation matrix R has been used. Once the r important components have been identified they are sometimes used in place of the original variates in subsequent analysis. For components extracted from the covariance matrix the component scores of the ith observation become

$$y_{i1} = \underset{\sim}{a}_1' \; (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}}), \ldots, y_{ir} = \underset{\sim}{a}_r' \; (\underset{\sim}{x}_i - \bar{\underset{\sim}{x}})$$

where $\underset{\sim}{x}_i$ is the ith observation vector and $\bar{\underset{\sim}{x}}$ is the sample mean vector. Correspondingly, if R has been used the component values for the ith observation are

$$y_{i1} = \underset{\sim}{a}_1' \, \underset{\sim}{z}_i, \ldots, y_{ir} = \underset{\sim}{a}_r' \, \underset{\sim}{z}_i$$

Using this approach the regression on k variables can be simplified to a regression on r component values which may enable greater understanding of the system. However, this situation is not the same as reducing the k original variables to their 'best' subset of p variables.

Although principal component analysis has reduced the dimensions of the data it still retains some information from each variate. However, the technique can be expanded to test whether the weightings associated with the ith variate are trivial for the r components considered (i.e. $a_{ij} = 0$ where $j = 1,\ldots,r$). This test can be used to eliminate one or more variables from the analysis but in practice it is seldom used as the analysis itself proves quite efficient in handling such variables. Only if further observations are required as part of a continuing experiment would there be any value in excluding the variates which contribute little if anything to our understanding of the system, even then such exclusion should consider other factors like the economic value, heritability, genetic and phenotypic correlation, and anything else which may increase the importance (whether genetic, agronomic or commercial) of the particular variate.

METHODS

The Roseworthy data set provides a valuable opportunity to compare the consistency of the methods described while simultaneously providing further information on the relationships between the variates. Obviously consideration of all nine methods would prove excessively time consuming; therefore it is proposed that only the following five most promising methods (as indicated in the introduction) should be considered:

i) Forward Selection;
ii) Backward Elimination;
iii) Regression by Leaps and Bounds - using the adjusted R^2 (R_a^2);
iv) Regression by Leaps and Bounds - using Mallow's C_p statistic;
v) Principal Components.

Also, since it would be rather trivial to reproduce further details on the methods of calculation for each of these tests, I shall simply refer the reader to the most relevant publication and briefly mention any computer programs which I have used. In order to correspond to the basic structure of the data set each method will be applied in turn to the data for the

Index and Visual flocks, within each sex, for the twelve years from 1954 to 1965, for the progeny born and raised as singles only.

The Forward Selection and the Backward Elimination techniques were applied using the appropriate SPSS program (i.e. Statistical Package for Social Sciences as developed by Nie, Hadlaihull, Jenkins, Steinbrenner and Bent (1975)). (An exemplary listing of the SPSS program used for Forward Selection is given in Appendix D.) This program closely follows the approach described by Draper and Smith (1966).

Since the method of Stepwise Regression is very similar to Forward Selection I have not applied it to the data.

Regression by Leaps and Bounds has been applied via a FORTRAN program which incorporates subroutines (SCREEN, COEF, PIVOT, STORE and BACK) as referred to by Furnival and Wilson (1974). The selection criterion is stipulated by the value of the parameter IBIT used when calling subroutine SCREEN, i.e. CALL SCREEN (RR,KX,NR,NDEF,IBIT,MBST) where if IBIT=2 the adjusted R^2 is used or if IBIT=3 Mallow's C_p statistic is used.

Calculation of Principal Components has similarly been by means of a FORTRAN program, specially written for the task, but using in this case an IMSL (International Mathematical and Statistical Libraries - 1974) subroutine OPRINC to calculate the eigen values, eigen vectors and component correlation matrix. This information is then evaluated using the procedure outlined by Morrison (1967, pp. 222-244). As there is considerable variation in scale between the eight variates the correlation matrix R has been used in preference to the covariance matrix S.

RESULTS AND DISCUSSION

While to analyse and compare one data set only, using the five methods of interest would in itself be a formidable task, I have considered below what amounts to 48 related data sets. Clearly to summarize such an immense amount of material requires several large tables even when one eliminates much of the less important information (e.g. the lower order principal components and the partial correlation coefficients have not been included). Although this approach has some disadvantages, these are far outweighed by the fact that evaluation over repeated data sets allows the consistency of the methods to be observed.

Tables V-1 to V-4 summarise the five methods for the Male Index, Male Visual, Female Index and Female Visual portions of the data respectively for the twelve years of the experiment. Thus, for forward selection the numbers indicate the order of inclusion of the variate into the final equation (where α the probability of inclusion has been set at 5%). Then in Table V-1 (i.e. for the male progeny born in the Index flock) we see for the year 1954 that percent yield was the first variate included, body weight the second and crimps frequency the third, while the four in brackets for secondary follicle number indicates that this variate has the next largest partial correlation coefficient but as the associated F test is less than the critical value this variate has not been included in the final equation. Thus, using Forward Selection we identify a subset of size 3 for male Index progeny born in 1954. Similar subsets have been established for the rest of the data.

The corresponding results obtained following analysis by Backward Elimination are presented in the table immediately below the results for Forward Selection. A similar numbering system has been used, however, the order of merit shown indicates the significance of the F ratio in the final subset (i.e. when all variates with significance greater than 5 percent have been excluded). The number in brackets indicates the last variate to be excluded.

A slightly different method of representation has been used for the two Regression by Leaps and Bounds techniques. Here the three most important subsets (as established by the statistic of interest - i.e. C_p or R_a) are indicated, in their order of subset size (p). The members of each subset have again been ranked from one to p, according to the magnitude of the F ratio, but it should be appreciated that the procedure of testing differs from that used in Forward Selection and Backward Elimination (i.e. Regression by Leaps and Bounds tests over all p members of each of the $(2^k - 1)$ possible subsets whereas Forward Selection tests progressively each of the remaining variates before including the variate which has the largest partial correlation coefficient providing the associated probability is less than the critical value). Thus, for Forward Selection the numbers shown can truly be associated with the order of inclusion but for Regression by Leaps and Bounds the numbers simply indicate the order of the magnitude of the F-ratios. The value of R_a^2 and C_p is indicated for each of the three subsets and a "†" has then been used to indicate the "most important" subset among each group of three (i.e. the subset having the largest value of R_a^2, or the value of C_p which approaches closest to p). It should be appreciated that this symbol has been included to show which subset would be obtained if the program was set to find one subset only instead of the three shown, and not to suggest this subset is "measurably" superior to the other two. As mentioned earlier it is a marked advantage of Regression by Leaps and Bounds that it allows identification of more than one subset.

Tables V-1 to V-4 indicate that the four methods considered so far do not lead to exactly the same subset. However, the outcomes of the methods are nevertheless related. Thus, we see that Forward Selection generally chooses similar variates although the size of the subset (p) is often a good deal smaller. This latter feature arises mainly from the choice of significance level (i.e. $\alpha = 0.05$). Increasing α would naturally increase p but as $\alpha = 0.05$ is widely used the generalisation would seem appropriate. Generally Backward Elimination identifies the same subset as Forward Selection, however, it can be seen that in the cases where this is not so it terminates with a larger subset. Use of C_p or R_a^2 give subsets which are much more similar but there is again a tendency for the former statistic (C_p) to give slightly smaller subsets than R_a but larger than the two previous stepping methods. This is not intended to imply that the value of a method is determined by the size of the resulting subset but simply to indicate the relative behaviour of the three methods, though it should be remembered that it is in principle far worse to exclude a relevant variate than to include an irrelevant one (i.e. the inconvenience of an extra variate is preferable to the bias resulting from exclusion of a relevant variate).

Since the present study, like any other investigation of real data, provides no direct means of distinguishing between these two types of variates the superiority, or otherwise, of any of the methods cannot be established beyond doubt. While simulation may appear a valuable alternative I believe such an approach would encounter serious difficulties in the generation of correlated data sets, as any such correlated data set would (like real data) be specific and therefore generalization would prove difficult. Given these constraints it should prove worthwhile to compare the performance of the various methods when they are applied in turn to similar data. Only by this means can some insight be obtained on their relative value and consistency. Before doing so, it is best to pause and consider the nature of the present data set. Here we have two flocks, which differ in method of ram selection, thus it seems appropriate to compare these within sexes over the twelve years of the trial. Intuitively one might expect the relationship between clean fleece weight and the variates measured to differ only between sexes (if at all) at the beginning of the experiment. This initial pattern could then change slowly over the years of the experiment if the two flocks diverge. In particular, since this is a field experiment conducted over seasons, it would not be surprising to observe temporal fluctuations in the relatioships or even interactions between time and some of these relationships. However, if the technique is to be of much value such fluctuations or interactions should be minimal.

Looking at Tables V-1 to V-4 again it can be seen that there is some consistency within sexes and over seasons but it would be extremely bold to suggest that they fit closely the expected results, since numerous examples can be found where variates important one year are absent the next and *vice versa* while in addition the subset size varies considerably. That is, while the four methods are often consistent within one set of data they vary considerably when compared over flocks, sexes and seasons. Clearly if one were to analyse only one of these sets of data, using any of the four methods, any inferences would be very misleading if extended to the sheep population in general. Here I must acknowledge that this data set may be itself atypical of data in general and thus do the four methods some injustice, though I have no reason to suppose that this is the case. However, unless one has strong evidence of the oddity of a particular data set (and I fail to see how one can in fact obtain such evidence without gathering many such data sets) then inferences drawn from the use of such methods may be extremely ill-founded.

In deriving Tables V-1 to V-4 I have used 48 data sets to compare the final subset obtained by either Forward Selection or Backward Elimination to the three best subsets of the Leaps and Bounds technique, demonstrating that the former techniques generally choose fewer variates, but it can be shown that this final subset obtained by Forward Selection or Backward Elimination may not in fact be optimal. This further drawback of Forward Selection (and to a lesser extent of Backward Elimination) can best be illustrated by the following examples, originally presented by Oosterhoff, and represented in Kendall (1975). (The Roseworthy data could be used to demonstrate this point but as it would require a much more lengthy presentation of the analyses shown in Tables V-1 to V-4 I have chosen to use the simpler and more concise examples of Oosterhoff.

In Oosterhoff's first example he gives a set of artificial observations from which the following multiple correlation coefficients (R^2) were calculated for the subsets shown:

x_1	0.6397	x_2,x_4	0.8138	x_1,x_2,x_3	0.9644
x_4	0.5608	x_1,x_2	0.7627	x_2,x_3,x_4	0.9144
x_2	0.2528	x_1,x_3	0.0899	x_1,x_2,x_4	0.8179
x_3	0.0906	x_1,x_4	0.6439	x_1,x_3,x_4	0.6906
		x_3,x_4	0.5608		
		x_2,x_3	0.2563	x_1,x_2,x_3,x_4	0.9737

(i.e. the regression of y on x_1 yields an R^2 of 0.6397, while the regression on x_1 and x_2 yields an R^2 of 0.7627). Then Forward Selection would proceed choosing x_1; then add x_2 to give x_1,x_2; then add x_3 to give x_1,x_2,x_3; and finally include x_4.

For a given significance level this process may terminate at the subset of size two, but from the table it can be seen that the pair x_2,x_4 would provide the best possible subset of size two not x_1,x_2. Backward Elimination can be seen to lead to the same outcome and Stepwise Regression is unlikely to change the result although this may be a little more difficult to follow from the above table.

The second example relates to the following set of ten observations relating the dependent variate y to two independent variates x_1 and x_2 (again the data are artificial but this in no way reduces the force of the argument)

y	x_1	x_2
29	7	7
-48	-19	-12
18	38	39
-12	45	49
44	- 5	- 7
57	15	12
47	-38	-40
10	38	39
86	59	59
46	-27	-29

The correlation between y and x_1 is 0.104 while R^2 is 0.011. Similarly the correlation of y and x_2 is -0.00635 and $R^2 = 0.000$. Clearly it would appear pointless to regress y on x_1 and x_2, since y is nearly independent of both the x's, and in fact Forward Selection would not attempt this regression (unless α was very large), but when we calculate this regression we find $R^2 = 0.999$ for the following regression equation:

$$y = 27.7 + 8.948(x_1 - 11.3) - 8.898(x_2 - 11.1)$$

Thus we see that y is very highly dependent on the difference between x_1 and x_2. In this case Backward Elimination would indicate the true situation but Forward Selection and Stepwise Regression would not. This ability of Backward Elimination to correctly handle such correlated variates has led to it being widely accepted as the preferable method of the three stepping procedures mentioned.

In the preceding results and discussion I have considered the first four methods only, leaving till last the method of Principal Components. This approach has been characteristic of the previous sections as well and stems mainly from the fact that this last method differs markedly from the previous four. Thus, we find that while the first four are variations on multiple regression, and are therefore univariate techniques depending on functions of the residual mean square, Principal Components in contrast is recognised as a true multi-variate technique. Not surprisingly, the results are presented in a form which also differs a good deal from the earlier methods.

As described previously Principal Components arrives at r orthogonal linear contrasts which indicate the contribution of each variate to the underlying factors of the data. Although the data can provide k components, only the first three have been included in Tables V-1 to V-4, as the lower order components contribute far less information (as the decreasing size of their eigen values indicates). Thus we see in the tables that the first three component correlation coefficients (indicated as ith CCC where i = 1,2,3) have been presented along with their eigen value (ℓ_i) and the proportion of the variance explained by each CCC (i.e. $\frac{\ell_i}{\mathrm{tr}\mathcal{S}} = \frac{\ell_i}{k}$). Turning to the 1954 male Index progeny (in Table V-1), we observe the following weights:

.09, .77, .77, .47, −.29, −.10, .48, .55

(associated with body weight, percentage yield, clean fleece weight, staple length, crimps, fibre diameter, primary follicle number and secondary follicle number respectively) explaining 25.4 percent of the variation as the eigen value is 2.03. It should be noted that the sign associated with the k elements of the component correlation coefficient bears only local importance (i.e. within the component) and use of a different algorithm or solution of a similar data set may result in the sign of all the k elements being reversed,

i.e. −.09, −.77, −.77, −.47, +.29, +.10, −.48, −.55

would be interpreted in exactly the same manner as the actual component discussed above). Thus, to ease interpretation of Tables V-1 to V-4, and also to assist later plotting, the sign of the component as a whole has been reversed in a number of cases before presentation. The interpretation of Principal Components is greatly simplified by the use of two dimensional plots in which values of the first component are plotted against the corresponding values of the second. This then allows a position in two dimensional space to be associated with each of the k variates which is much more comprehensible to the human mind than a mass of numbers in a table (e.g. the first plot in Figure V-1 shows the two dimensional pattern of the eight variates for the 1954 Index male progeny, where the numbers 1 to 8 correspond to the variates body weight, percent yield, clean fleece weight, staple length, crimps, fibre diameter, primary follicle number and secondary follicle number respectively). The plots of the first and second component correlation coefficients for the male and female progeny are presented in Figures V-1 and V-2 respectively. To assist interpretation the two flocks have been presented alongside each other for each of the twelve years of the experiment.

Previously I have explained how the signs of all k values making up the component correlation coefficients have been reversed for some of the forty-eight data sets considered. It has been proposed that this in no way affects or biases the analysis. However, a second constraint has been applied to the components which may lead to the plots being considered biased. This constraint arises from the fact that the order of the components shown in Tables V-1 to V-4 has in several cases been changed before plotting. When the components are considered carefully, it is seen that among the three shown for each set of data, there are two patterns of coefficients which occur repeatedly, these are:

i) a component having relatively large positive coefficients for clean fleece weight and percentage yield (and a positive but smaller coefficient for staple length) and an equally large but negative coefficient for crimps, all coefficients for the remaining variates being close to zero. This component then describes the inverse relationship between fleece weight and crimps which has been reported in previous studies (e.g. Robards, Williams and Hunt (1974)). As this component most frequently occurred first in the set of three components, shown in Tables V-1 to V-4 and summarised in Table V-5, I will refer to it as the first component correlation coefficient (1st CCC).

ii) Secondly, we observe a component having relatively large positive coefficients for secondary follicle number and primary follicle number and often, but by no means consistently, a moderately large negative coefficient for fibre diameter. (The coefficients for the remaining variates again being distributed about zero.) This component associated with fleece density can be seen in Table V-5 to occur most frequently as the second observed component and therefore I will refer to it as the second component correlation coefficient (2nd CCC).

Table V-5 compares the original order of these first and second components for both the male and female progeny groups. That is, for the male progeny the 1st CCC was observed in position one on twenty occasions and position two on four occasions, while the 2nd CCC occurred 4, 19 and 1 times in positions one, two and three respectively. A similar pattern was observed for the female progeny. Thus, it would appear that there exist two consistent components, over the forty-eight data sets considered, but the order of these components varies occasionally. Whether any meaning can or should be attached to these odd occurrences is unknown but it is my opinion that this behaviour does not greatly detract from the apparent consistency of the components. That is, analysis of any one of the data sets (or any group of them) using Principal Component

Analysis would result in basically similar conclusions, which is a markedly different outcome from that observed for the previous four methods.

Previously I have indicated that analysis of the Roseworthy data provides an opportunity both to compare five representative methods of analysis and also to expand our understanding of the behaviour of the two Roseworthy flocks under different selection programs. Since the preceding has mainly dealt with the first of these points I will now consider the second in more detail.

Naturally, as the interpretation of any data depends to a large extent on the appropriateness of the method of analysis then this second aspect will depend largely on the conclusions of the first. Bearing this in mind, one is immediately led to ask - What is the biological meaning or implication of the two component correlation coefficients? To answer this we recall that the first CCC was characterised by two relatively large positive coefficients for clean fleece weight and percentage yield and a large negative coefficient for crimps per inch (all the remaining variates being associated with inconsistent coefficients much smaller in magnitude, except staple length which exhibited some evidence of a positive coefficient of intermediate magnitude). This first CCC suggests that there is one factor which results in the simultaneous expression of high fleece weight and low crimp frequency (where percentage yield and clean fleece weight behave similarly but individually). Therefore it would be expected that selection for extremes of either fleece weight or crimps would result in an opposite response in the other variate. While the behaviour of staple length may be similar to clean fleece weight or percentage yield any response to selection for clean fleece weight would be comparatively smaller.

Interestingly, when we compare the outcome of the two selection techniques, over the twelve years of the trial, we see that the Index method, as would be expected, has resulted in greater clean fleece weights than the Visual method. Secondly, this increase (or divergence) has been associated with a reduction in number of crimps and an increase in percent yield and staple length. However, two other variates, namely secondary follicle number and body weight have also exhibited divergence which would not be expected from the first component correlation coefficients.

Turning to the second CCC we observe large positive coefficients for both primary and secondary follicle number which suggest a second factor, orthogonal to the first. This second factor can be identified as responsible for the variation in follicle number. The fact that it is orthogonal to the first CCC would suggest that it is controlled by a separate mechanism, therefore, while we would expect these two components to be unrelated, in terms of their expression, it must be remembered that the selection has been applied to the phenotypic values (clean fleece weight) which would result in indirect selection in favour of follicle number. Unless there was a large negative genetic correlation between these two variates and this was not supported by the estimates obtained in Chapter II, we could expect that an increase in clean fleece weight would be accompanied by an increase in follicle number.

The behaviour of body weight in comparison cannot be explained in terms of the first two components as the coefficient is small and inconsistent in both cases. It could be that later components may relate to body weight but as no consistent pattern could be seen and since these remaining components explain proportionally less of the variation anyway, this aspect was not pursued. It would appear that response in body weight was not due to the two factors observed but simply a direct response to selection common to both flocks, body size being favoured more by the subjective method of visual appraisal than by the objective measurement of clean fleece weight.

Lastly, the observed consistency of the two components, over the twelve years of the experiment, would suggest first that the relationships between the eight variates have not been measurably affected by either of the two selection methods, and secondly, that the flocks involved should both respond to further selection. Together, these two points answer questions which are of prime importance to the breeder, i.e.

i) has the response observed for one (or more) variates been to the detriment of other variates?

ii) Will further generations of selection give continued response in the variate (or variates) of interest?

However, while Principal Components provides a useful tool for looking at these two problems it must be remembered that the results only extend over the variates considered and therefore its value depends largely on the ability of the breeder to choose and accurately measure the appropriate variates. In particular, the analysis of Chapter IV suggests that a fertility problem may have arisen in the Index flock in the later years of the experiment. This variate has not been included in the above analysis as:

i) its distribution is discrete;

ii) its form is markedly different between sexes;

iii) not all progeny have been measured (especially males) and those which have, also vary in the number of years of recording;

iv) for ewes its assessment extends over up to six years' records and is therefore difficult to accumulate and it only becomes available years after the other variates;

v) it is known to vary with the age of the ewe (McGuirk, 1976) and so comparison over time may well be almost impossible.

Clearly consideration of this variate is, and will remain, a major problem to the animal breeder. While it can be included in this and other methods of analysis, all suffer from some, or all, of the above limitations and the methods are therefore only marginally effective in their handling of the problem. This problem warrants further consideration but since this is beyond the bounds of the present Chapter no further attempt will be made here. Advances in this area will depend jointly on the statistical methods used and advances in the biologist's understanding of fertility.

In conclusion, the four most widely used methods for choosing the 'best' subset, from a correlated set of variates, have all demonstrated an extreme lack of consistency when applied over similar groups of data. It is proposed that Principal Component Analysis can provide an effective means for reducing the dimensions of correlated data sets, by identifying the factors which contribute to the correlation structure. It is implied that these factors relate to the genetical expression of related variates and therefore indicate the response which can be expected following selection. Although the behaviour of the variates over 12 years of selection supports this hypothesis, it cannot be established conclusively by my analysis. However, the genetic correlations, which are presented in Chapter II of this study, further support the idea. While the method here has not reduced the variates to a smaller (or best) subset which may then be selected for jointly, using a selection index (as may have been hoped by some breeders), it has provided a form which may be more valuable. In fact, the first two component correlation coefficients may be used to transform the data to two orthogonal scores which can then be used individually or simultaneously to select progeny.

Table V-1: Summary of four methods of analysis for the male progeny of the Index Flock

	Forward selection [Backward elimination]							Regression by Leaps and Bounds Using R^2_a								Regression by Leaps and Bounds Using Mallow's Cp								Principal Components									
Year	BWT	YLD	STL	CRP	FBD	PFN	SFN	BWT	YLD	STL	CRP	FBD	PFN	SFN	R^2_a	BWT	YLD	STL	CRP	FBD	PFN	SFN	Cp	BWT	YLD	CFW	STL	CRP	FBD	PFN	SFN	Eigen value	Modified order*
1954	2	1		3			(4)	2	1		3			4	49.6 †	2	1		3				4.27	.09	.77	.77	.47	-.29	-.10	.48	.55	2.03	(1)
	[2	1		3			(4)]	2	1		3	5		4	49.4	2	1		3			4	3.88†	.55	.13	.32	-.69	.46	-.30	-.36	.37	1.47	
								2	1		3		5	4	49.3	2	1		3	4			4.66	-.42	-.13	-.37	.15	.48	-.71	.36	.45	1.41	(2)
1955	1	2		3		(4)		1	2		3		4		41.5 †	1	2		3				2.52	.62	.66	.81	.54	-.63	.20	.10	.18	2.25	(1)
	[1	2		3		(4)	]	1	2		3	5	4		40.9	1	2		3		4		2.49†	.13	.20	-.05	-.25	.22	-.68	.68	.90	1.88	(2)
								1	2		3		4	5	40.6	1	2		3	4			3.90	.50	-.18	.10	.41	.53	-.47	-.31	-.16	1.09	
1956		1	(3)	2					1	2	3		4		21.2 †		1	3	2				3.97	-.42	.27	.42	-.15	-.55	-.46	.61	.80	1.97	(2)
	[	1	(3)	2			]		1	2	3		4	5	20.3		1	2			3		4.93	.28	.57	.69	.59	-.37	.44	.04	-.34	1.69	(1)
									2	1	3	6	5	4	20.5		1	2	3		4		3.54†	-.03	.07	.03	.57	.09	-.69	-.50	.10	1.07	
1957	(2)		1					3		1			4	2	33.3	2		1					4.23	.39	.58	.69	.81	-.52	.71	-.30	-.43	2.66	(1)
	[(2)		1				]	4		1		5	3	2	34.4 †			1		3		2	4.32	.49	-.01	.33	.12	-.29	-.19	.68	.73	1.48	(2)
								4	6	1		5	3	2	33.6	3		1			4	2	4.23†	-.65	.69	-.08	.09	-.07	-.05	.27	.17	1.01	
1958		1	2				(3)		1	2			4	3	31.2 †		1	2			4	3	3.49†	.29	.24	.08	-.36	.39	-.71	.76	.85	2.23	(2)
	[	1	2				(3)]	5	1	2			4	3	31.0	5	1	2			4	3	4.73	.29	.63	.80	.65	-.63	-.14	-.05	.14	1.98	(1)
									1	2		5	3	4	30.5		1	2		5	3	4	5.11	-.78	.42	-.07	-.26	-.33	.15	.25	.10	1.07	
1959	3	1		2	(4)			3	1	5		2		4	50.3	3	1	5		2		4	7.53	.13	.56	.74	.66	-.66	.59	-.39	-.27	2.34	(1)
	[3	1	5	(6)	2		4]	2	1	5	6	3		4	51.9 †	2	1		4	3		5	7.65	.32	.47	.49	-.13	-.17	-.44	.55	.84	1.81	(2)
								2	1	5	6	3	7	4	51.4	2	1	5	6	3		4	6.42†	-.85	.08	-.04	-.02	-.16	.31	.35	.20	1.02	
1960		1			2	(3)		4	1			3	2	5	28.6		1			2	3		4.91	-.34	-.24	-.58	-.64	.29	-.70	.65	.70	2.39	(2)
	[	1			2	(3)	]	4	1	6		3	2	5	28.8 †	4	1			2	3		5.56	.15	.77	.55	-.07	-.39	.12	.51	.53	1.62	(1)
								4	1		6	3	2	5	28.3	4	1			3	2	5	5.65†	.72	.17	.16	-.38	.64	-.10	-.06	-.11	1.15	
1961	1	2			(3)			1	2			3			53.2 †	1	2						1.75	.67	.52	.86	.55	-.17	.32	-.33	-.21	2.03	(1)
	[1	2			(3)		]	1	2	4		3			53.2	1	2			3			1.72†	.13	.56	.28	-.12	.29	-.52	.57	.75	1.65	(2)
								1	2			3	4		52.8	1	2	3					2.41	-.48	.31	-.07	.24	-.75	.45	.47	.17	1.40	

Table V-1:Contd.

Year	BWT	YLD	STL	CRP	FBD	PFN	SFN	BWT	YLD	STL	CRP	FBD	PFN	SFN	R^2_a	BWT	YLD	STL	CRP	FBD	PFN	SFN	Cp	BWT	YLD	CFW	STL	CRP	FBD	PFN	SFN	Eigen value	Modified order*
1962	3	1	2				(4)	4	1	2		6	5	3	41.1	4	1	2		6	5	3	7.94	.30	.52	.72	.59	-.73	.49	-.36	-.37	2.27	(1)
	[4	1	2	(7)	6	5	3]	4	1	3	6		5	2	39.3	4	1	3	6		5	2	10.21	-.12	.63	.38	-.36	-.19	.06	.64	.77	1.72	(2)
								4	1	3	7	6	5	2	41.9†	4	1	3	7	6	5	2	8.00†	.74	-.08	.23	.25	.13	-.60	.09	.19	1.08	
1963	3	4	2	1	5	7	6	1	5	4		3	6	2	61.0	1	5	4		3	6	2	10.36	.23	.69	.87	.69	-.75	.32	-.20	.07	2.46	(1)
	[1	6	4	7	3	5	2]	2		3	5	4	6	1	60.2	2		3	5	4	6	1	11.57	-.40	.31	.06	-.33	-.14	.10	.75	.88	1.73	(2)
								1	6	4	7	3	5	2	63.2†	1	6	4	7	3	5	2	8.00†	-.77	.22	-.21	.03	.07	.63	-.24	-.25	1.21	
1964	3	5	1	4	(6)		2	3	5	2	4			1	51.6	3	5	2	4			1	7.23	.69	.47	.58	.75	-.51	.60	-.33	-.68	2.79	(1)
	[3	5	2	4	(6)		1]	3	4	2	5	6		1	52.8†	3	4	2	5	6		1	6.58†	-.19	.46	.63	.14	-.41	-.42	.47	.52	1.49	(2)
								3	4	2	5	6	7	1	52.5	3	4	2	5	6	7	1	8.00	.02	.55	-.10	-.06	.19	-.53	-.64	-.04	1.04	
1965	3	1	5	2		(6)	4	3	1	5	4			2	62.6	3	1	5	4			2	6.77	.19	-.47	-.34	-.73	.48	-.55	.75	.63	2.40	(2)
	[3	1	5	4		(6)	2]	3	1	4	5		6	2	64.3†	3	1	4	5		6	2	6.00†	.08	.65	.85	.10	-.43	-.47	.29	.61	2.01	(1)
								2	1	4	5	7	6	3	63.4	2	1	4	5	7	6	3	8.00	.89	.02	.19	.34	.40	-.11	.15	-.34	1.25	

*See text for explanation

Table V-2: Summary of the four methods of analysis for the male progeny of the Visual Flock

	Forward selection [Backward elimination]							Regression by Leaps and Bounds Using R_a^2								Regression by Leaps and Bounds Using Mallow's Cp								Principal Components									
Year	BWT	YLD	STL	CRP	FBD	PFN	SFN	BWT	YLD	STL	CRP	FBD	PFN	SFN	R_a^2	BWT	YLD	STL	CRP	FBD	PFN	SFN	Cp	BWT	YLD	CFW	STL	CRP	FBD	PFN	SFN	Eigen value	Modified order*
1954	2	1					(3)	2	1		5	4		3	34.4 †	2	1					3	5.85	.04	.50	.48	.66	-.54	-.55	.51	.69	2.26	(1)
	[2	1					(3)]	2	1		5	4	6	3	34.2	2	1			4		3	5.07	-.39	-.61	-.66	-.00	.18	-.40	.59	.31	1.61	(2)
								2	1	6	5	4		3	33.3	2	1		5	4		3	4.94†	.53	.07	.29	-.48	.67	-.16	.20	.44	1.31	
1955	3	1	2	4		(5)		2	1	4	3		5		46.0	2	1	4	3		5		7.64	.19	.60	.80	.70	-.75	.52	.18	.12	2.40	(1)
	[2	1	4	3		(5)	]	2	1	5	3		4	6	46.7	2	1	5	3		4	6	7.61	-.28	.46	.07	-.21	.06	-.52	.49	.89	1.63	(2)
								2	1	5	4	7	3	6	47.2 †	2	1	5	4	7	3	6	8.00†	.69	-.41	-.10	.14	-.09	-.17	.59	.05	1.06	
1956		1		(2)				3	1		2				29.1		1		2				4.70	-.17	.81	.64	.31	-.63	-.45	.35	.70	2.39	(1)
	[	1		(2)			]	3	1		2	4		5	29.8	3	1		2				3.60	-.02	-.18	-.39	-.52	.39	-.60	.50	.49	1.45	(2)
								4	1		2	5	6	3	29.9 †	3	1		2	4			4.70†	.93	.07	.28	.18	.38	-.09	-.10	.16	1.15	
1957	1	(6)		2	5	3	4	1	6		3	5	2	4	46.3	1			3	5	2	4	7.96	.65	.48	.85	.53	-.38	.45	-.18	.12	2.05	(1)
	[1	(6)		3	5	2	4]	1		6	4	5	2	3	45.5	1	6		3	5	2	4	7.33	-.15	.51	-.07	-.07	-.23	-.22	.77	.81	1.63	(2)
								1	6	7	3	5	2	4	46.5 †	1	6	7	3	5	2	4	8.00†	.49	-.09	.27	-.38	.22	-.65	-.26	.28	1.08	
1958	2	(4)		1	3			2			1	3			21.0	2			1	3			2.08	.23	.77	.53	.34	-.77	-.06	.45	.61	2.23	(1)
	[2	(4)		1	3		]	1	4		2	3			21.8 †	1	4		2	3			2.61†	-.34	-.13	-.49	-.56	.13	-.68	.66	.63	2.10	(2)
								1	3		2	4		5	21.0	1	2		3			4	3.48	.85	-.27	.24	-.23	.22	-.14	-.01	.20	1.02	
1959		1		(2)				5	1		2		4	3	40.0 †		1		2		4	3	5.50	.44	.70	.71	.72	-.64	.24	-.32	-.29	2.36	(1)
	[	1		(2)			]	5	1		3	6	4	2	39.0	4	1		2			3	5.88	.13	.43	.40	.09	.47	-.65	.14	.69	1.50	(2)
								5	1	6	3		4	2	39.0	5	1		2		4	3	4.01†	-.16	.07	-.06	-.02	.32	-.41	-.82	-.42	1.15	

Table V-2: contd.

Year	B W T	Y L D	S T L	C R P	F B D	P F N	S F N	B W T	Y L D	S T L	C R P	F B D	P F N	S F N	R_a^2	B W T	Y L D	S T L	C R P	F B D	P F N	S F N	Cp	B W T	Y L D	C F W	S T L	C R P	F B D	P F N	S F N	Eigen value	Modified order*
1960		1	2	(3)					1	2					38.6 †		1	2					-.17†	-.30	.81	.70	.38	-.53	-.24	.55	.59	2.36	(1)
	[	1	2	(3)			]		1	2	3				38.5		1	2	3				.89	-.61	-.12	-.43	-.63	.30	-.62	.20	.60	1.83	(2)
									1	2		3			38.2		1	2		3			1.14	-.54	-.10	-.07	-.24	-.34	.60	.21	-.16	0.90	
1961	1	2	(4)	3				1	2	4	3				49.3 †	1	2		3				3.34	.69	.21	.86	.58	-.35	.42	-.08	.01	1.90	(1)
	[1	2	(4)	3			]	1	2	4	3		5		49.2	1	2	4	3				3.29†	-.41	.57	.02	.26	-.11	.02	.64	.77	1.59	(2)
								1	2	4	3	5			48.6	1	2	4	3		5		4.48	-.31	.62	.21	.08	-.18	-.60	-.59	-.18	1.30	
1962	(2)	1						3	1		6	5	4	2	36.9 †	3	1			4	5	2	6.79	-.12	.76	.74	.54	-.72	.51	-.04	.20	2.23	(1)
	[3	1		(6)	4	5	2]	3	1	6		4	5	2	36.6	3	1		6	5	4	2	6.67†	-.45	.04	-.10	-.18	-.07	-.21	.84	.84	1.69	(2)
								3	1	7	6	5	4	2	36.6	3	1	6		4	5	2	7.04	-.71	-.08	-.29	-.25	-.08	.60	-.11	-.21	1.08	
1963	1		(4)	3			2	2	4	5		3	6	1	44.2	2	4	5		3	6	1	7.32	.14	.69	.61	.72	-.80	.45	.05	-.14	2.24	(1)
	[2	(5)	4		3		1]	2	3		6	4	5	1	43.7	2	3		6	4	5	1	7.86	.27	.34	.44	.05	.25	-.65	.55	.82	1.83	(2)
								2	4	6	7	3	5	1	44.5 †	2	4	6	7	3	5	1	8.00†	.80	-.35	.42	.03	.17	-.01	-.45	-.10	1.17	
1964	3	1		2		(4)		3	2	6	1		4	5	41.1	3	1		2				6.93	.14	.64	.78	.61	-.58	.40	-.24	-.34	2.08	(1)
	[3	1		2		(4)	]	3	1		2	6	4	5	40.9	3	2		1		4		5.78	-.16	.29	.33	-.14	-.46	-.53	.48	.80	1.59	(2)
								3	1	6	2	7	4	5	41.1 †	3	2		1		4	5	6.43†	-.85	-.06	-.20	-.10	-.35	.47	.14	-.06	1.13	
1965	3	(4)	2	1				2	3	4	1			5	43.3	2	3	4	1				5.40	.02	.69	.81	.71	-.79	.14	.25	.24	2.39	(1)
	[2	3	(4)	1			]	2	5	3	1	6		4	43.4 †	2	3	4	1			5	5.17	-.49	.08	-.16	-.19	.12	-.49	.82	.72	1.74	(2)
								2	4	3	1	6	7	5	42.6	2	5	3	1	6		4	6.00†	.76	-.29	.22	.08	.02	-.62	-.09	.30	1.20	

*See text for explanation.
†"most important" subset among each group of three shown

Table V-3: Summary of the four methods of analysis for the female progeny of the Index Flock

	Forward selection [Backward elimination]							Regression by Leaps and Bounds Using R^2_a								Regression by Leaps and Bounds Using Mallow's Cp								Principal Components									
Year	BWT	YLD	STL	CRP	FBD	PFN	SFN	BWT	YLD	STL	CRP	FBD	PFN	SFN	R^2_a	BWT	YLD	STL	CRP	FBD	PFN	SFN	Cp	BWT	YLD	CFW	STL	CRP	FBD	PFN	SFN	Eigen value	Modified order*
1954	3	2	1			(4)		3	1	2		5	4	6	53.1	3	1	2		5	4	6	8.31	.38	.49	.80	.70	-.49	-.31	.54	.30	2.24	(1)
	[3	2	1			(4)	]	1	3	2	6	5	4		53.0	1	3	2	6	5	4		8.32	-.58	-.30	-.41	.06	.12	-.54	.57	.76	1.81	(2)
								2	3	1	7	5	4	6	54.3 †	2	3	1	7	5	4	6	8.00†	-.45	.57	-.08	-.44	-.60	.13	-.03	.09	1.11	
1955	1	3		2		(4)		1	3		2		4	5	39.9	1	3		2				6.29	.43	.06	.57	.72	-.64	.76	-.44	-.69	2.70	(1)
	[1	3		2		(4)	]	1	3	6	2		4	5	41.5 †	1	3		2		4		6.07	.46	.59	.56	-.01	-.30	-.08	.67	.56	1.73	(2)
								1	3	5	2	7	4	6	40.8	1	3	6	2		4	5	6.38†	-.58	.08	-.36	.13	-.50	.36	.36	.12	1.01	
1956		1	3	2			(4)		3	2	1				26.4			2	1				3.44	-.31	.68	.59	.38	-.54	-.37	.64	.67	2.34	(1)
	[		2	1			(3)]			1	2		4	3	26.3		3	2	1				2.92†	-.40	-.31	-.54	-.58	.35	-.38	.48	.56	1.69	(2)
									5	1	2		4	3	26.8 †			1	2			3	3.58	-.73	.01	-.01	-.06	-.13	.72	.09	-.10	1.09	
1957	1	2				(3)		1	2				3		43.4	1	2				3		1.92†	.11	-.15	-.23	-.81	.20	-.47	.69	.77	2.06	(2)
	[1	2				(3)	]	1	2	4			3		43.7 †	1	2	3					2.56	.65	.53	.86	.13	.40	-.17	.09	.11	1.66	(1)
								1	2	4		5	3		43.2	1	2	4			3		2.70	.66	-.75	.08	-.29	.03	.18	-.26	-.18	1.23	
1958	2			1		(3)		2	5		1	4	3		37.2 †	2			1	4	3		5.01	.37	.37	.79	.70	-.48	.34	-.31	-.47	2.05	(1)
	[2			1		(3)	]	1	5	6	2	4	3		36.6	2	4		1		3		5.54	-.24	.45	.12	.20	-.65	-.37	.67	.63	1.71	(2)
								2	4		1	5	3	6	35.9	2	5		1	4	3		4.55†	-.74	-.13	-.29	.25	-.20	.47	.23	-.40	1.18	
1959	(5)	1	3			2	4	5	3	4	6		1	2	46.6 †		3	2			1	4	5.87	.25	.54	.68	.73	-.59	.46	-.57	-.51	2.50	(1)
	[(5)	3	2			1	4]	5	3	4		6	1	2	46.6	5	3	2			1	4	6.18	.11	.59	.51	.31	.28	-.61	.22	.69	1.69	(2)
								5	3	4	6	7	1	2	46.5	5	3	4	6		1	2	6.81†	.76	-.20	.08	-.17	.47	-.27	-.47	-.23	1.21	

Table V-3: Contd.

Year	BWT	YLD	STL	CRP	FBD	PFN	SFN	BWT	YLD	STL	CRP	FBD	PFN	SFN	R^2_a	BWT	YLD	STL	CRP	FBD	PFN	SFN	Cp	BWT	YLD	CFW	STL	CRP	FBD	PFN	SFN	Eigen value	Modified order*
1960	2			1			(3)	1		4	2	5		3	28.7	2			1			3	5.33	.48	.31	.61	.32	-.65	.68	-.42	-.46	2.07	(1)
	[2			1			(3)]	1		5	2	4	6	3	28.9 †	1		4	2			3	5.45†	.42	.64	.48	-.45	-.29	-.26	.56	.71	1.99	(2)
								1	7	4	2	5	6	3	28.7	1			2	4		3	5.69	.52	.06	-.20	-.64	.36	.13	-.09	-.37	1.00	
1961	2	3	1		(4)			2	1	3		4			37.2 †	1	2	3					2.88	.49	.57	.72	.71	-.28	.65	-.46	-.67	2.77	(1)
	[1	2	3		(4)		]	1	2	3	4				36.9	2	1	3		4			3.23†	-.00	.49	.37	.25	-.33	-.18	.72	.54	1.39	(2)
								1	2	3	5	4			36.9	1	2	3	4				3.62	.45	.04	.21	.25	.74	-.46	-.12	.16	1.11	
1962	3	1			2	(4)		3	1			2	4	5	51.3	3	1			2	4	5	7.22	.49	.39	.76	.59	-.52	.72	-.31	-.34	2.33	(1)
	[3	1	(6)		2	4	5]	3	1	6		2	4	5	52.7 †	3	1	6		2	4	5	6.30†	.14	.60	.36	-.16	-.20	-.13	.68	.83	1.73	(2)
								2	1	6	7	5	3	4	52.2	2	1	6	7	5	3	4	8.00	.72	-.44	.06	.05	.53	.22	.31	.09	1.14	
1963			1				(2)	4		1		3	5	2	37.6 †	4		1		3		2	5.59	-.44	.45	.38	.63	-.74	.50	-.43	-.29	2.00	(1)
	[4		1		3	(5)	2]	4	6	1		3	5	2	36.3	4		1		3	5	2	4.33†	.02	.32	.52	.18	-.08	-.01	.61	.83	1.47	(2)
								3		2	6	4	5	1	36.1	4	6	1		3	5	2	6.13	.73	-.53	.61	.15	-.05	.26	-.02	-.21	1.32	
1964	1	3		2			(4)	1	3		2				46.2	1	3		2				1.46†	.60	.55	.66	.56	-.45	.66	-.10	-.58	2.39	(1)
	[1	3		2			(4)]	1	3		2			4	46.2 †	1	3		2			4	2.47	-.23	.41	.22	-.19	-.65	.00	.74	.59	1.61	(2)
								1	3		2	4			45.8	1	3		2	4			2.95	-.62	-.11	-.59	.29	-.31	.52	.09	-.31	1.28	
1965	4	1	2		(5)		3	3	4	1		5		2	48.3	3	4	1				2	5.38	-.03	.41	.20	-.45	.60	-.71	.65	.76	2.27	(2)
	[3	(4)	1				2]	2	4	1		5	6	3	48.7 †	3	4	1		5		2	5.20	.02	.58	.85	.51	-.50	.34	.19	.31	1.81	(1)
								2	4	1	7	5	6	3	46.9	2	4	1		5	6	3	6.00†	-.81	-.19	-.27	-.20	-.34	.38	.50	.21	1.36	

*See text for explanation

†"most important" subset among each group of three shown

Table V-4: Summary of the four methods of analysis for the female progeny of the Visual Flock

Year	Forward selection [Backward elimination]							Regression by Leaps and Bounds Using R_a^2								Regression by Leaps and Bounds Using Mallow's Cp								Principal Components									
	BWT	YLD	STL	CRP	FBD	PFN	SFN	BWT	YLD	STL	CRP	FBD	PFN	SFN	R_a^2	BWT	YLD	STL	CRP	FBD	PFN	SFN	Cp	BWT	YLD	CFW	STL	CRP	FBD	PFN	SFN	Eigen value	Modified order*
1954	(2)			1				3	5		1	6	4	2	18.5	2	3	4	1				8.59	-.36	.69	.38	.44	-.47	-.44	.44	.75	2.11	(1)
	[(2)			1			]	4		6	2	5	3	1	18.1	3	5		1	6	4	2	8.39	-.28	-.01	-.70	-.42	.55	-.53	.33	.30	1.52	(2)
								1	6	7	3	5	4	2	20.2 †	1	6	7	3	5	4	2	8.00†	.71	-.22	.17	-.03	.18	.20	.62	.34	1.14	
1955	2	(4)		1			3	1	3		4	5		2	34.3 †	1	4		3			2	4.82	.55	.66	.72	.80	-.78	.25	.08	.02	2.58	(1)
	[1	(4)		2			3]	1	4		3	5	6	2	34.0	1	2			4		3	5.72	-.06	-.03	.31	-.05	.21	-.14	.70	.79	1.28	(2)
								1	3	6	4	5		2	33.1	1	3		4	5		2	4.83†	-.59	.16	-.02	-.10	-.12	.83	.11	.04	1.11	
1956	2	1		3		(4)		3	1		2		4		38.2	2	1		3				5.52	.61	.61	.80	.44	-.55	.39	-.61	-.36	2.42	(1)
	[2	1		3		(4)	]	3	1	5	2		4		38.8 †	3	1		2			4	3.94	.03	.45	.23	.20	-.31	-.44	.55	.80	1.53	(2)
								3	1	5	2	6	4		38.0	3	1	5	2			4	4.32†	.41	-.27	-.12	.71	.36	.15	.31	-.00	1.01	
1957	1	2	(3)					1	2	3					56.4 †	1	2						.86	.66	.85	.80	.54	-.34	-.03	.38	.38	2.49	(1)
	[1	2	(3)				]	1	2	3		4			56.0	1	2	3					.65†	-.40	.14	-.34	-.26	-.04	-.59	.64	.77	1.71	(2)
								1	2	3				4	55.9	1	2			3			1.58	.46	.06	.26	-.58	.48	-.55	-.13	-.13	1.18	
1958	(3)	2	1					3	1	2					41.0		2	1					2.89	-.31	.84	.67	.69	-.72	.32	.33	.07	2.45	(1)
	[(3)	2	1				]	3	1	2			4		41.3 †	3	1	2					2.19†	-.43	.11	-.21	-.26	-.05	-.65	.65	.91	1.98	(2)
								3	1	2		5	4		40.9	3	1	2			4		2.93	.69	.06	.51	.36	.47	-.46	.14	.14	1.34	
1959	2	1		3		(4)		1	2		3				38.7 †	1	2		3				1.89†	.39	.57	.69	.73	-.70	.52	-.20	.37	2.42	(1)
	[1	2		3		(4)	]	1	2		3		4		38.4	1	2		3		4		3.26	.11	.46	.38	.14	-.03	-.61	.52	.74	1.57	(2)
								1	2		3		4	5	38.6	1	2		3			4	3.38	-.77	.04	-.24	.08	-.54	.31	.24	.27	1.17	

Table V-4: Cont'd.

Year	BWT	YLD	STL	CRP	FBD	PFN	SFN	BWT	YLD	STL	CRP	FBD	PFN	SFN	R^2_a	BWT	YLD	STL	CRP	FBD	PFN	SFN	Cp	BWT	YLD	CFW	STL	CRP	FBD	PFN	SFN	Eigen value	Modified order*
1960	1	3	4	2	(7)	5	6	4		2	5		1	3	40.6	4		2	5		1	3	7.65	.61	.37	.76	.50	-.52	.43	-.59	-.48	2.37	(1)
	[4	(6)	2	5		1	3]	5	6	2	3		1	4	42.3 †	5	6	2	3		1	4	6.72†	.28	.65	.37	-.00	.07	-.62	.14	.63	1.44	(2)
								4	6	2	5	7	1	3	42.1	4	6	2	5	7	1	3	8.00	-.12	.32	-.12	-.66	.16	-.31	-.68	-.39	1.31	
1961	1	2			3	(4)		1	2			3	4		60.8	1	2			3			3.96	.65	.55	.86	.64	-.48	.48	.12	-.27	2.40	(1)
	[1	2			3	(4)	]	1	2	5		3	4		61.0 †	1	2			3	4		4.32†	.01	.28	.02	.16	.12	-.19	.81	.87	1.57	(2)
								1	2	5		3	4	6	60.9	1	2	4		3			4.64	-.66	.47	-.26	.02	-.62	.18	-.02	.01	1.15	
1962	2			1		(4)	3	2	4		5		3	1	32.2 †	2	4		5		3	1	4.86†	-.67	-.20	-.35	-.54	.49	-.63	.70	.69	2.51	(2)
	[2	4		5	(6)	3	1]	2	4		5	6	3	1	32.0	2	4		5	6	3	1	6.10	.07	.65	.69	-.11	-.50	.11	.43	.56	1.68	(1)
								1	5	6	4		3	2	31.4	1	5	6	4		3	2	6.75	.58	-.06	.38	-.30	.43	-.32	-.04	-.05	0.86	
1963	2	3		1		(4)		2	3		1				38.8 †	2	3		1				1.97†	.08	.49	.66	.67	-.57	.64	-.35	-.57	2.30	(1)
	[2	3		1	(4)		]	2	3		1		4		38.3	2	3		1		4		3.53	-.62	.12	.10	-.22	-.64	.27	.64	.42	1.53	(2)
								1	3		2	4			38.2	1	3		2	4			3.59	-.52	-.33	-.57	-.15	-.02	.44	-.27	-.51	1.24	
1964	2	3	1	(5)	4			2	3	1		4			56.3 †	2	3	1		4			3.06†	.63	.22	.80	.66	-.25	.56	-.52	-.66	2.61	(1)
	[2	3	1	(5)	4		]	2	3	1	5	4			55.9	2	3	1	5	4			4.69	-.43	.61	.29	.40	-.67	.26	.67	.49	2.00	(2)
								2	3	1		4	5		55.7	2	3	1		4	5		4.90	-.36	-.49	-.30	-.23	-.55	.59	-.03	-.37	1.30	
1965	2	1		3		(4)		2	1		3		4		29.5	2	1		3		4		4.14†	.15	.71	.76	.72	-.77	.11	.14	.24	2.30	(1)
	[2	1		3			(4)]	2	1	5	4			3	30.0 †	2	1		3			4	4.29	.25	-.29	.18	-.26	.07	-.22	.74	.81	1.50	(2)
								2	1	5	3		6	4	29.6	3	1	4				2	4.36	.71	-.04	.17	.17	.23	-.70	-.35	-.11	1.24	

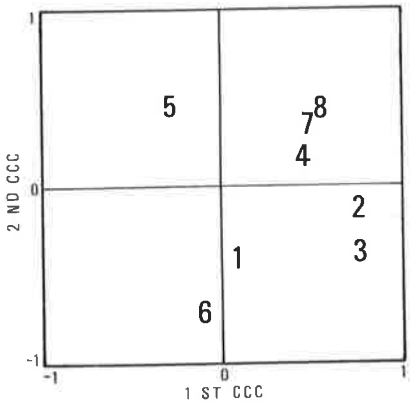
*See text for explanation

†"most important" subset among each group of three shown

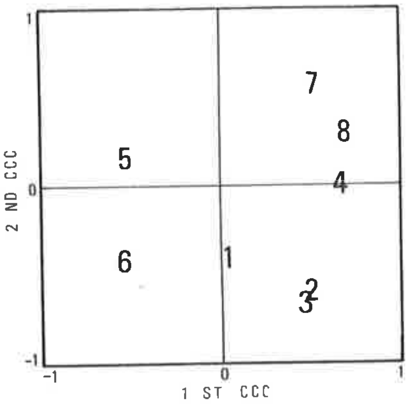
Figure V-1 Plots of the first and second component correlation coefficients (CCC) for the 1954 to 1965 male progeny

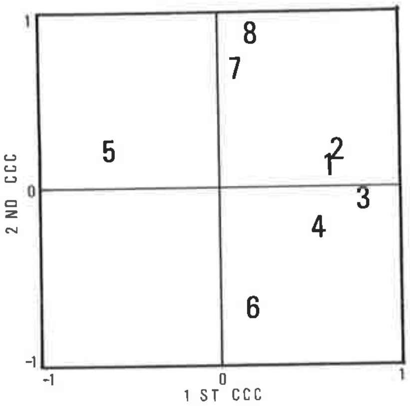
Where 1 denotes body weight
2 denotes percentage yield
3 denotes clean fleece weight
4 denotes staple length
5 denotes crimps
6 denotes fibre diameter
7 denotes primary follicle number
8 denotes secondary follicle number

INDEX MALE VISUAL

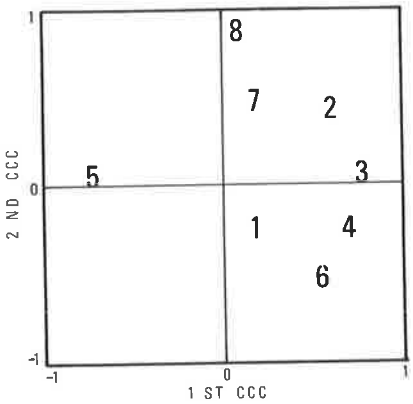


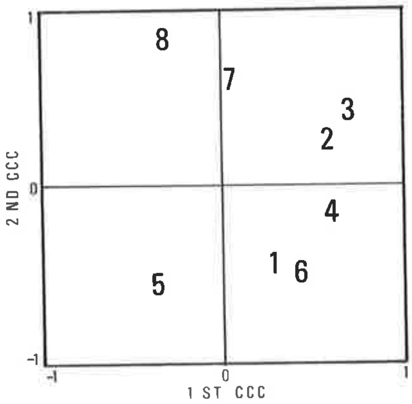
1954



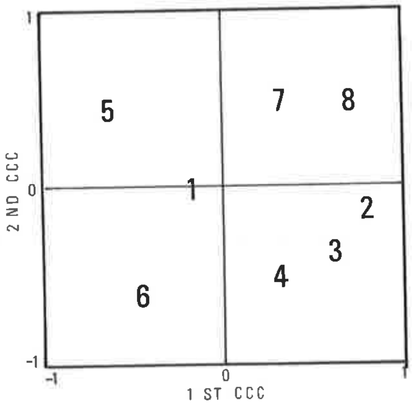


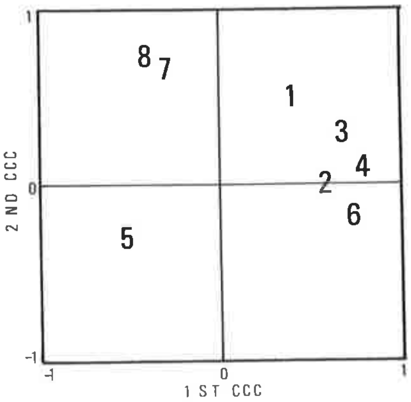
1955





1956





1957

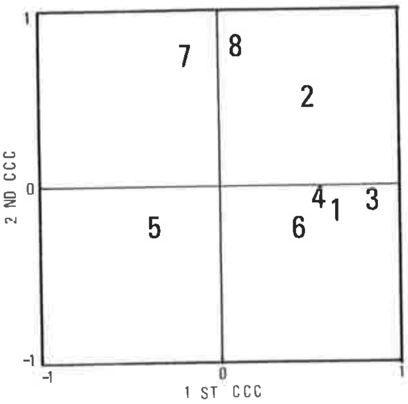


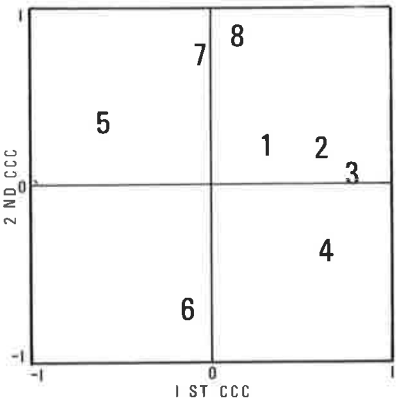
Figure V-1 (continued)

Where 1 denotes body weight
2 denotes percentage yield
3 denotes clean fleece weight
4 denotes staple length
5 denotes crimps
6 denotes fibre diameter
7 denotes primary follicle number
8 denotes secondary follicle number

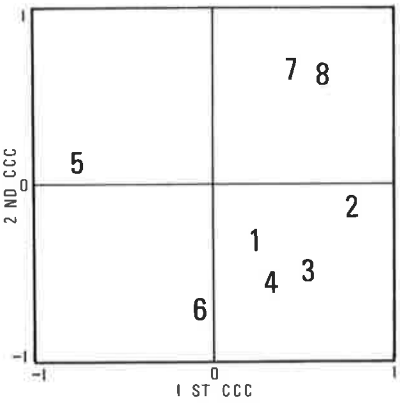
INDEX

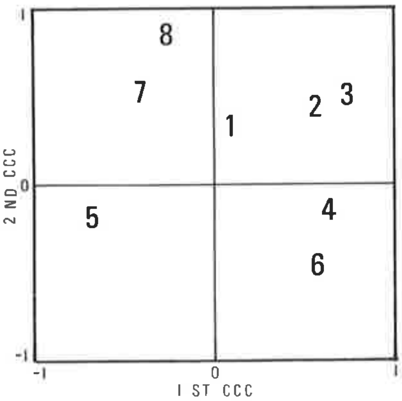
MALE

VISUAL

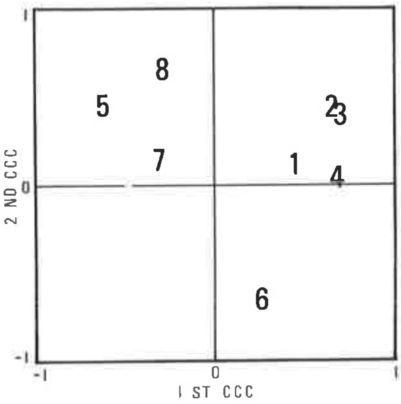


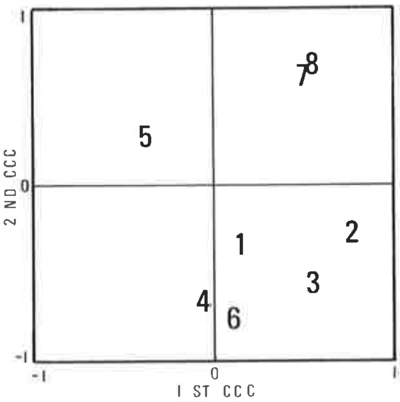
1958



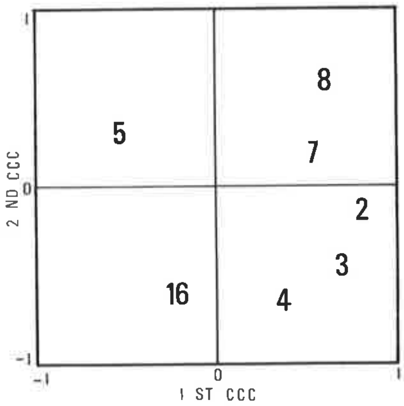


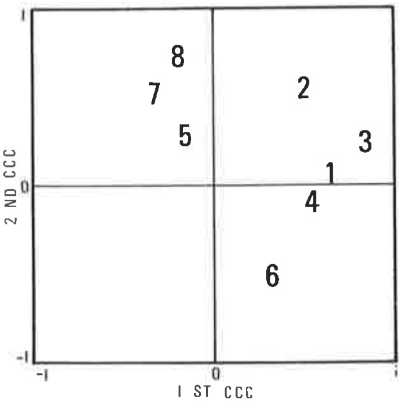
1959





1960





1961

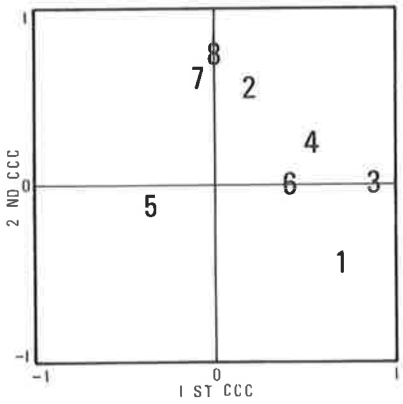
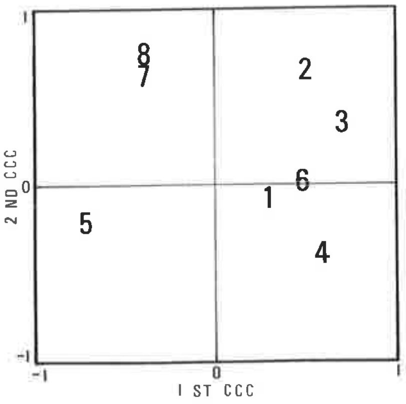


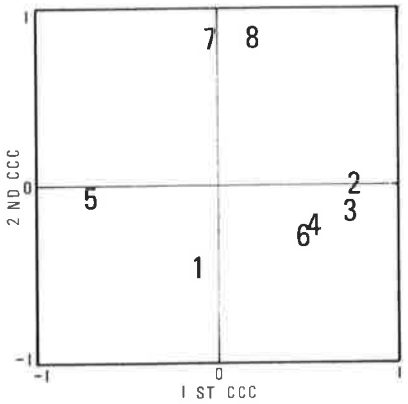
Figure V-1 (continued)

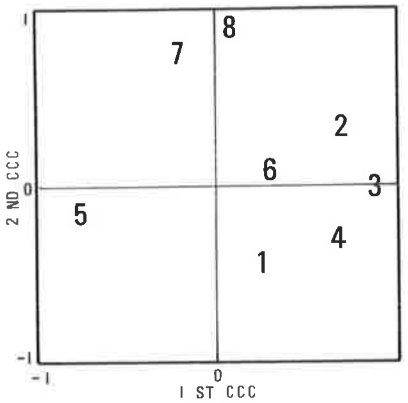
Where 1 denotes body weight
2 denotes percentage yield
3 denotes clean fleece weight
4 denotes staple length
5 denotes crimps
6 denotes fibre diameter
7 denotes primary follicle number
8 denotes secondary follicle number

INDEX MALE VISUAL

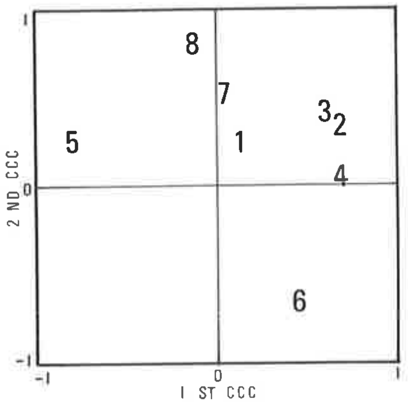


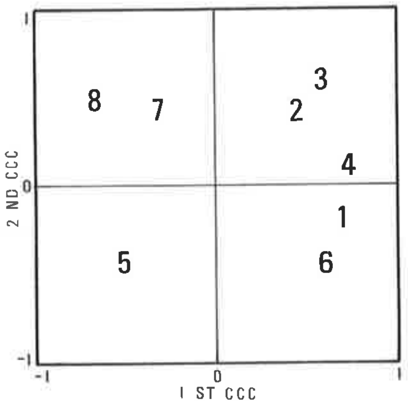
1962



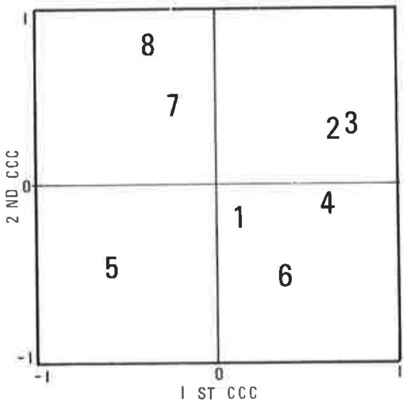


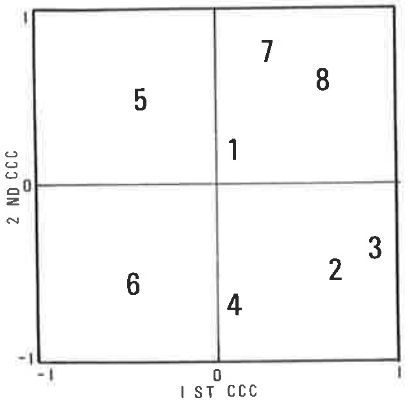
1963





1964





1965

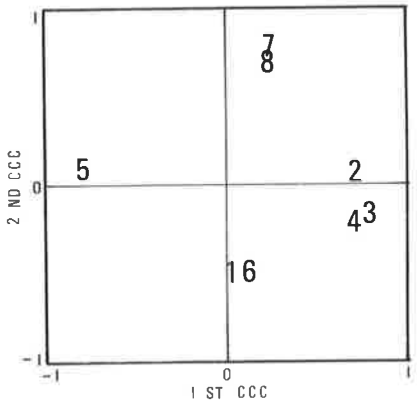
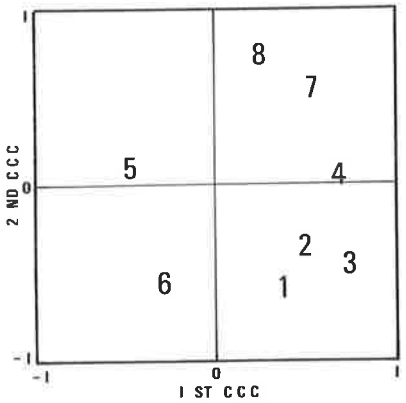


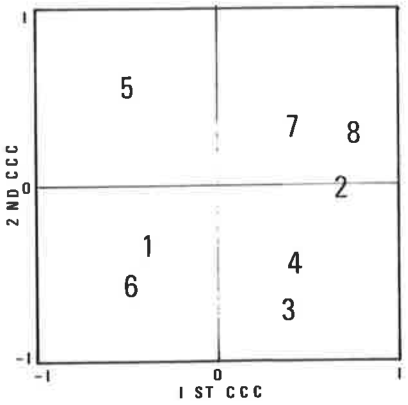
Figure V-2 Plots of the first and second component correlation coefficients (CCC) for the 1954 to 1965 female progeny

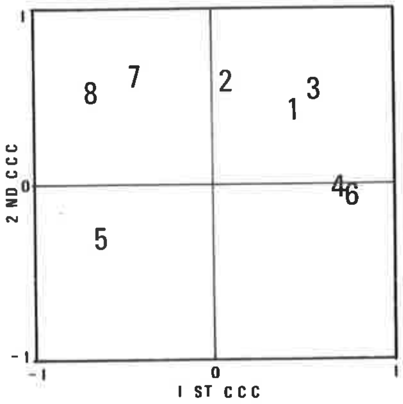
Where 1 denotes body weight
2 denotes percentage yield
3 denotes clean fleece weight
4 denotes staple length
5 denotes crimps
6 denotes fibre diameter
7 denotes primary follicle number
8 denotes secondary follicle number

INDEX FEMALE VISUAL

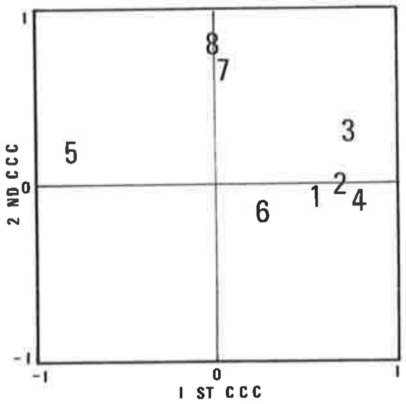


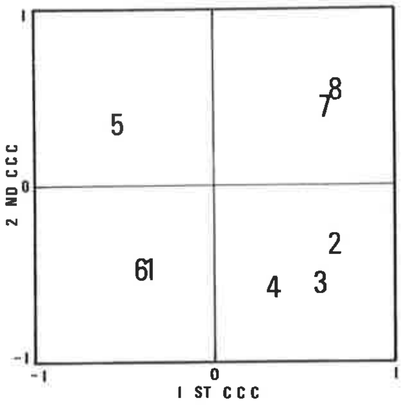
1954



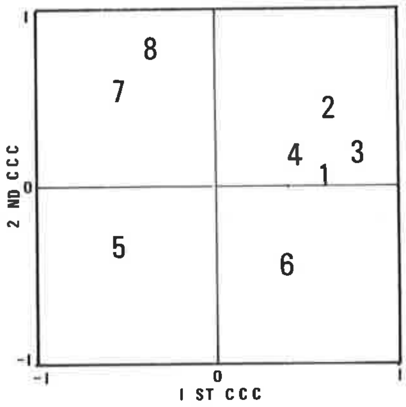


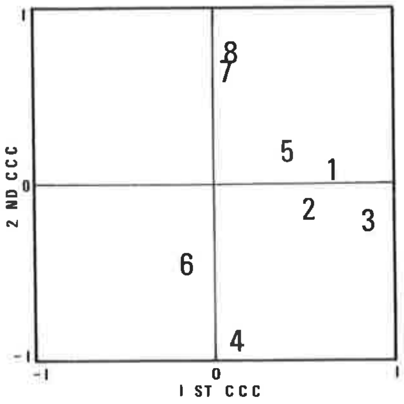
1955





1956





1957

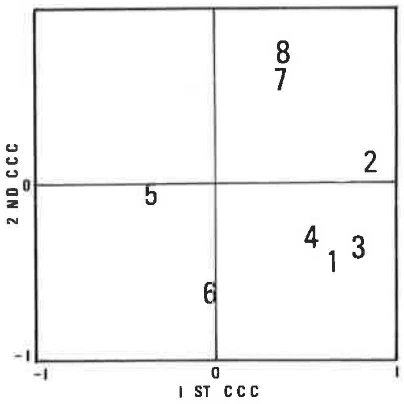
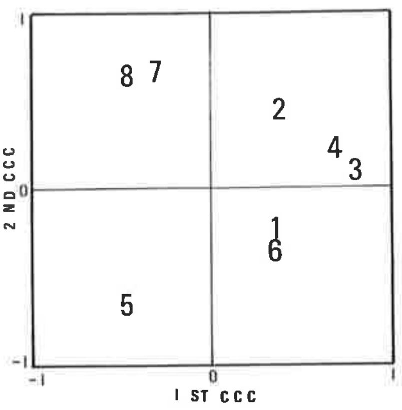


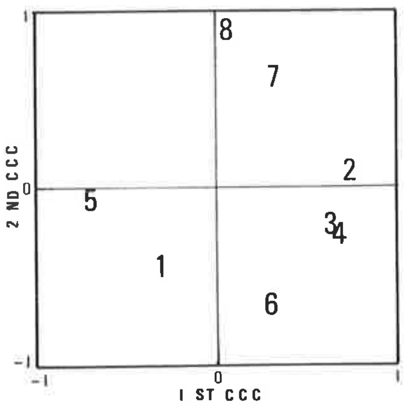
Figure V-2 (continued)

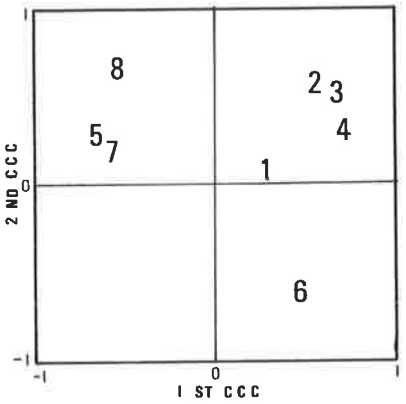
Where 1 denotes body weight
2 denotes percentage yield
3 denotes clean fleece weight
4 denotes staple length
5 denotes crimps
6 denotes fibre diameter
7 denotes primary follicle number
8 denotes secondary follicle number

INDEX FEMALE VISUAL

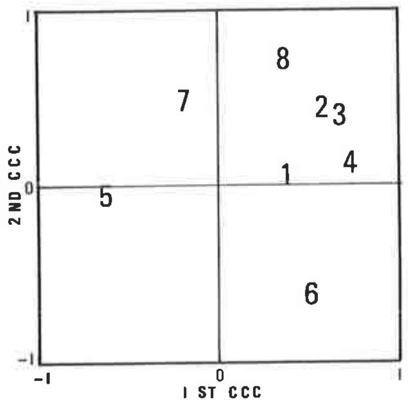


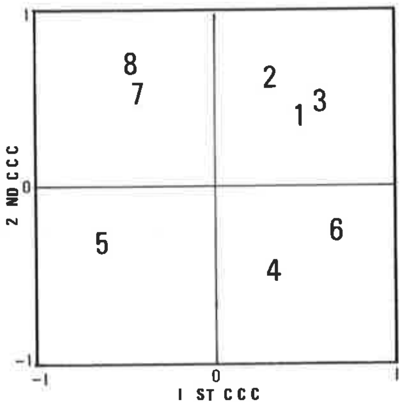
1958



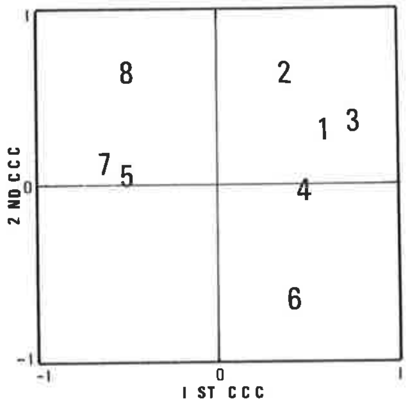


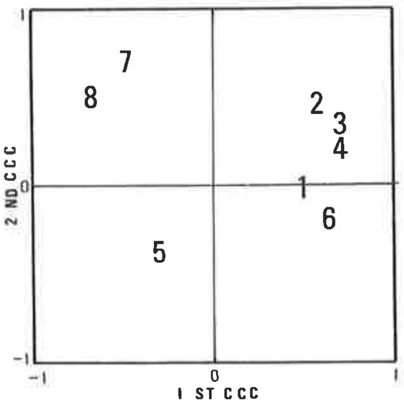
1959





1960





1961

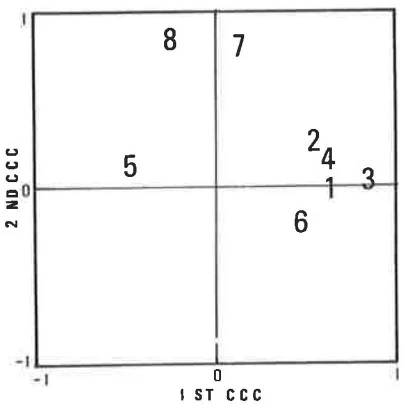
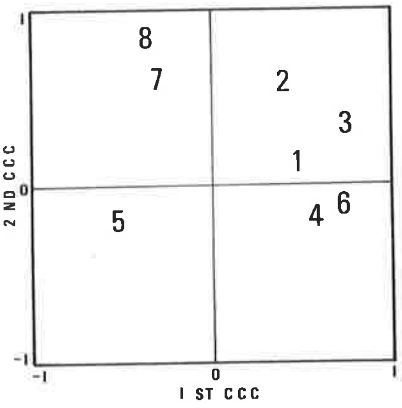


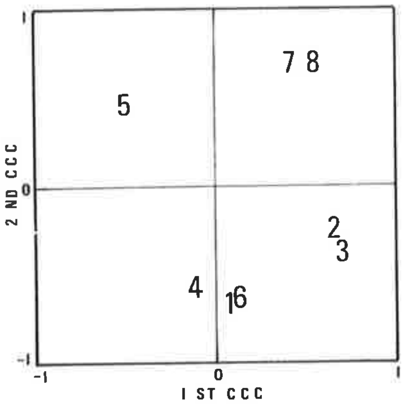
Figure V-2 (continued)

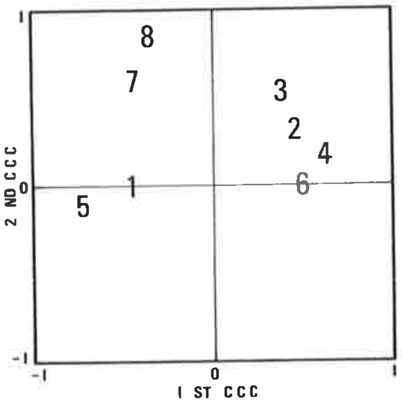
Where 1 denotes body weight
2 denotes percentage yield
3 denotes clean fleece weight
4 denotes staple length
5 denotes crimps
6 denotes fibre diameter
7 denotes primary follicle number
8 denotes secondary follicle number

INDEX FEMALE VISUAL

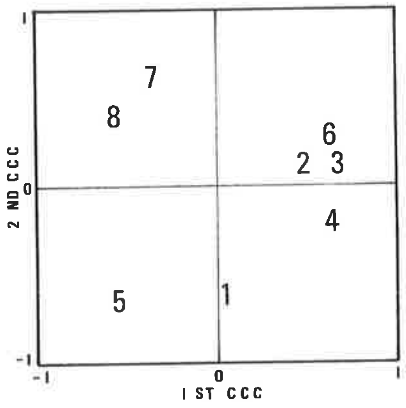


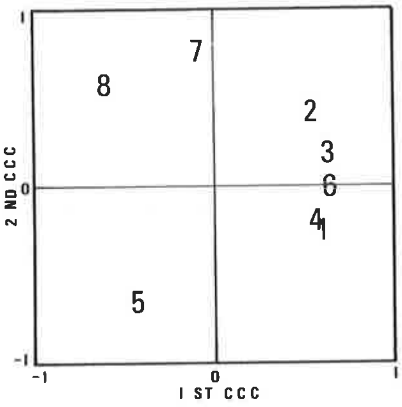
1962



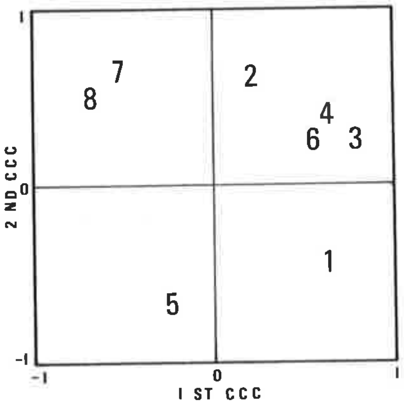


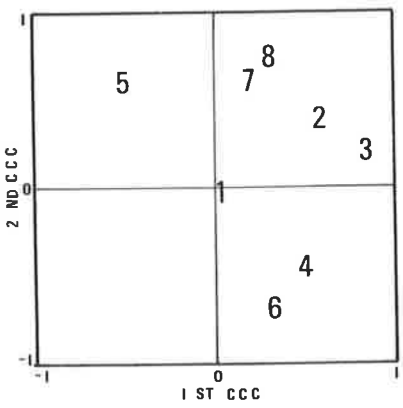
1963





1964





1965

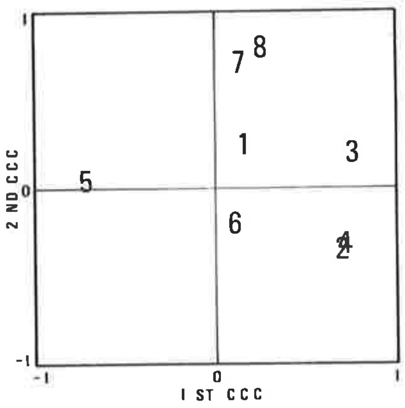


Table V-5 Comparison of modified order to original order of components for both sexes

		Original order of components		
Modified order of components*		1	2	3
Male	1	20	4	0
	2	4	19	1
Female	1	21	3	0
	2	3	21	0

* See text for explanation.

VI - DETECTION OF MAJOR GENES USING TESTS FOR NORMALITY

INTRODUCTION

It is generally accepted that the frequency distribution of a quantitative character will be non-normal if a major gene is involved in the inheritance of the character (Hammond and James, 1970). Departure from normality, as a result of contamination, has been recognised since early this century (Pearson, 1894, 1902). Fisher, Immer and Tedin (1932), however, were the first to attempt to apply normality tests in quantitative genetics when they investigated the genetical interpretation of statistics of the third degree.

More recently, Mérat (1968) has proposed the use of third and fourth moment statistics to detect such genes. After a preliminary test for heterogeneity of variances Mérat suggests comparing the departure from normality of the families with the highest and lowest variances. Although g_1 (the Skewness Coefficient), can be used to identify asymmetry the main interest lies in the use of g_2 (the Kurtosis Coefficient) to identify families showing platykurtic distributions, which Mérat concludes are likely to be associated with the segregation of major genes. (The derivation of both the Skewness Coefficient (g_1) and the Kurtosis Coefficient (g_2) have been given earlier in Chapter I.)

Hammond and James (1970) applied four methods involving higher order statistics (including Mérat's method) to detect genes of large effect. Two characters of *Drosophila* (abdominal chaetae number and progeny number) were considered, but little evidence of non-normality was observed. Indeed, as other authors (Jones, 1967; Robertson, 1967; and Frankham, 1969) had previously reported indirect evidence of major genes associated with the inheritance of abdominal chaetae number, Hammond concluded: "The sensitivity of the methods used is somewhat uncertain for the traits studied." However, while several relevant factors (i.e. the magnitude of the heritability, non-random environmental component of the variation, non-additive variation, linkage and natural selection) were discussed, the power of the tests was not pursued. (Piper, 1971 also noted the need for establishing the size of the deviation which would be required before a major gene could be detected.)

In a later article, Hammond and James (1972) evaluated O'Donald's method (O'Donald, 1971) for estimating the number of loci, using higher degree statistics. O'Donald had estimated the number of loci ($\hat{n}$) using the formula

$$\hat{n} = \frac{1}{2}(g_1^2 - g_2)$$

but when Hammond applied this to his abdominal chaetae data, within sexes, he found all estimates were extremely low, demonstrating that O'Donald's formula was biased downwards.

It can be seen that these authors have relied heavily on the use of the coefficients of skewness and kurtosis to test for the presence of major genes. However, after searching the literature the following tests of normality were considered to provide possible alternatives worthy of comparison to the tests for skewness and kurtosis:

i) Haldane's Test for Bimodality (H) - this technique was proposed by Haldane (1951) as a means of identifying the peaks and troughs occurring in the frequency distribution of a metric trait. Basically the test consists of comparing the observed frequencies in adjacent cells to establish the cases where there has been a significant change in frequency. (The Standard Normal Distribution is used to establish significance.) If more than one such change is located, these are assessed to see if they are indicative of an antimode (trough) or a pair of modes (peaks).

ii) Studentized Range (U) which is the ratio of the range to the standard deviation as suggested by David, Hartley and Pearson

(1954). (This statistic is sometimes confused with one proposed by Geary (1935) formed from the ratio of the mean deviation to the standard deviation.)

iii) W statistic, as proposed by Shapiro and Wilk (1965) which basically compares the slope, for the regression of the ordered observations on the expected values of the order statistics, with the usual symmetric sample sum of squares about the mean. Calculation of W in practice proves rather difficult as it involves the covariance matrix of the normal order statistics. Thus, Shapiro and Wilk (1965) only presented coefficients and percentage points for sample sizes of 50 or less.

However, Shapiro and Francia (1972) presented a simplified form of the test (W') which compared favourably, in terms of sensitivity, with W. A table of the "Empirical Percentage Points of W' Test" was included by these authors for n = 35, 50, 51(2)99 (where n = sample size). Later work by Weisberg (1974) further verified that for samples ≥ 35, W' has approximately the same accuracy as W. In view of this I have considered only W' in the following work.

iv) Joint Tests using both g_1 and g_2. The following brief summary illustrates several of the joint tests which have been considered, often with reservations.

The first possibility is to test for both skewness and kurtosis and reject if either or both show evidence of non-normality. (This I will call the 'simple joint test' of normality, and use J to represent it.)

Secondly, we can consider the P statistic suggested by Pearson, 1902, where

$$P = 2b_2 - 3b_1 - 6$$

(note $g_2 = b_2 - 3$, and $g_1 = \sqrt{b_1}$).

Recently, Bowman and Shenton (1973) have considered P using Monte Carlo simulations of the normal and gamma densities to obtain the percentiles of P.

Thirdly, two 'omnibus' tests involving the joint use of g_1 and g_2 have been put forward by D'Agostino and Pearson (1973). Both these tests result in a χ^2 statistic, after assuming g_1 and g_2 are independent, under the null hypothesis. However, D'Agostino and Pearson (1974) withdrew the tests after Professor Frank Anscombe had pointed out that while g_1 and g_2 were uncorrelated, they were not independent. This criticism must surely apply equally to the previous two joint tests, but while I have not considered D'Agostino's two 'omnibus' tests further, I have continued with the 'simple joint test' and the P statistic as they are frequently used. This criticism is especially true for the 'simple joint test' which is basically the procedure most people adopt when testing for normality (i.e. few people consider either g_1 or g_2 alone - most look at both in turn rejecting normality if either indicates deviation from normality).

v) "Lower Tailed" Kurtosis Statistic (C) - Deviation associated with Kurtosis - can be described as either leptokurtic (where $g_2 > 0$) or platykurtic (where $g_2 < 0$). Leptokurtosis arises if the distribution is much more peaked than the normal distribution while platykurtosis indicates the distribution is unusually flattened. As the presence of a major gene would be expected to result in a platykurtic distribution the Kurtosis Statistic will be applied as a one-tailed test and this would be expected to provide a more efficient test than the corresponding two-tailed test.

Several other tests of normality appeared in the literature. However, they have not been included in this study as previous work had shown them to be comparatively insensitive to deviations from normality (see Shapiro, Wilk and Chen, 1968). Among those rejected was a group called the 'distance tests', because they depend on use of the sample estimates of the mean ($\bar{x}$) and variance (s^2) for the population mean (μ) and variance (σ^2). (Whereas W', g_1, g_2, u and H are scale and origin invariant.) The actual members of the 'distance tests' were:

a) Cramer-Von Mises statistic - Cramer, 1928.

b) Kolmogorov-Smirnov test - Kolmogorov, 1933.

c) Weighted Cramer-Von Mises statistic - Anderson and Darling, 1954.

d) Modified Kolmogorov-Smirnov test - Durbin, 1961 (see also Lilliefors (1967) for more recent approach)

e) Simple Chi-squared test.

The aim of the present study is to first, investigate alternative tests for normality to determine which are more suited to identifying the type of non-normality expected when a major gene was present in the population and second, to establish the minimum deviation which a major gene would have to cause before its presence could be consistently detected.

Although subsets of the above tests have been compared previously for a range of alternative distributions, their sensitivity has not been assessed when the alternative distribution is bimodal (i.e. Shapiro, Wilk and Chen (1968), used the following unimodal distributions, Beta (p,q), Binomial ($k,.5$), Chi-squared (ν), Double Chi-squared (β), Johnson SB (γ,δ), Logistic (α,β), Log normal (μ,σ^2), Non-central Chi-squared (ν,λ), Poisson (λ), Student T (ν), Tukey (a,λ), and Weibull (k,λ), while Dyer (1974) considered the Uniform, Exponential, Double Exponential and Cauchy distributions). Thus, while these comparisons have provided sound foundations on which to choose the most promising tests, I will refrain from giving further details as these would only be loosely related to the area of interest.

Clearly the deviation from normality in the presence of a major gene will depend on the frequency of the major gene and its manner of expression. The latter of these two factors may involve a deviation in mean and/or variance.

If the frequency is sufficiently large the overall distribution will be bimodal. However, this pattern will become less distinct as the frequency of the major gene decreases, until at low frequencies only a skewed distribution may be seen. Since skewed distributions could also be due to the presence of "outliers" or simply the "natural distribution" of the character, care should be practiced when making any suggestions with respect to major genes for such data.

Chen (1971) investigated the power of the W statistic for samples drawn from a contaminated normal distribution. For sample sizes between 3 and 50 he included up to 7 observations differing in either scale or location and found W was sensitive to such contamination. However, as only the very small sample sizes could be considered bimodal and only one statistic was assessed, this work gives but a preliminary glimpse in the direction of interest here.

Before considering the relative efficiency of the above tests, it should be noted that this investigation deals with the first part of what is in practice a two-stage procedure. That is, for a given set of observations one would first test for departure from normality, and if this was supported, then look more closely to ascertain what is responsible for this discrepancy. (This second aspect will be considered in more detail later in this chapter.)

SIMULATION

As the expected frequency and effect of a major gene can vary, simulation was used to compare the sensitivity of the following eight test statistics over a range of frequencies and effects:

i) Haldane's Test for Bimodality (H).

ii) Studentized Range (U).

iii) Shapiro-Francia Statistic (W') - often called the Modified Shapiro-Wilk Statistic.

iv) Simple Joint Test (J).

v) Pearson's P Statistic (P).

vi) Skewness Statistic (S).

vii) Kurtosis Statistic (K) (as a two-tailed test).

viii) "Lower Tailed" Kurtosis Statistic (C).

A uniform random number generator, which returns a random value between 0 and 1, was used to generate two normal populations, $N(0,1)$ and $N(\lambda,1)$, where λ is the difference between the means of the two populations, which both have a standard deviation of one. These populations were mixed in 51 different proportions - from 0(2)100 percent of the second population $N(\lambda,1)$ (i.e. 0,2,4,6,....,98,100).

For a given significance level (α) the power of the 8 statistics, at each of the 51 proportions, was obtained using repeated sampling. Graphs of the power plotted against proportion provide an effective means of comparing the 8 test statistics for a given difference (λ).

As the power of a test is the probability of accepting the alternate hypothesis given the null hypothesis is false,

i.e. Power = $1-\beta$, where β is the probability of accepting the null hypothesis when it is false,

the power measures the sensitivity or efficiency of a test for a level of significance α.

The following parameters were used for the simulation:

i) α (the significance level) = 0.10;

ii) n (the number of observations in each mixture) = 50, 100;

iii) λ (the distance in standard deviations) = 0, 1, 2, 3, 4;

iv) n_s (the number of repeated samples generated) = 100.

Although $\alpha = 0.05$ would have been a more useful level of significance, the more conservative value (0.10) was accepted as the tables of percentiles for many of the two-tailed tests did not include the upper and lower 2.5% critical values. These could have been found by Monte Carlo sampling of the standard normal population (i.e. N(0,1)) but it was felt that this would be an unnecessary burden on the project. In order to check both the critical values used and the simulation technique, the significance level was calculated for all the cases where the null hypothesis was known to be true (i.e. where no mixing had occurred or $\lambda = 0$).

Two sample sizes (i.e. n = 50 or 100) were chosen to investigate the power of the tests, though it was of course accepted that smaller or larger numbers may be encountered in data.

The distance obviously represents the deviation of the major gene from the population in general and for simplicity unit increments of the standard deviation have been used.

The number of samples (n_s = 100) generated for each mixture could be considered low, but it was found to provide a reasonable estimate of power without using ridiculous amounts of computer time. When the value of the power for each test sample is plotted for each proportion, one not only gets an idea of the power of the test statistics, but also their "stability" or "repeatability". (Consequently, if the points are joined by a straight line, this gives a "zig-zag" effect which may detract from the appearance but, to my mind, not from the interpretation.)

RESULTS

In Figures VI-1(1) to VI-1(7), the power of each of the seven tests has been plotted against the proportion 'contamination' of the N(0,1) population by the N(λ,1) population, when λ takes the values 0, 1, 2, 3, 4; $\alpha = 0.10$ and $n = 50$. (Haldane's Test has not been simulated for the smaller sample size as it performed so poorly at $N = 100$.) Similar power curves are presented for all eight tests when $n = 100$ in figures VI-2(1) to VI-2(8). Clearly, all the tests exhibit limited power when the deviation of the major gene is less than two standard deviations. This limitation, although not unexpected, indicates that many major genes of genetical significance will remain undetectable among sample sizes of 50 to 100. This situation is unlikely to improve greatly even for much larger sample sizes.

The power curves of the eight tests for $\lambda = 3$ and $\alpha = 0.10$ have been combined in Figures VI-3 and VI-4, for $n = 50$ and 100 respectively. It can be seen that the proportion of contamination has quite marked effects between and within tests - as the following points illustrate:

i) Although all the tests behave symmetrically, and thus could have been plotted from 0 to 50 percent 'contamination', it was felt that inclusion of the full scale would be beneficial to the reader.

ii) For both the extremes of p (the proportion of contamination), the power is seen to be close to the limiting value of 0.10 for all the tests (i.e. when the null hypothesis is known to be true the power $= 1-\beta = \alpha$).

iii) Although the power of the Skewness Statistic (S) rises rapidly initially in the presence of positive or negative skewness, it naturally lacks power for the symmetric distributions observed between $p = 0.4$ and 0.6.

iv) Kurtosis Statistic (K) lacks power about $p = 0.2$ and $p = 0.8$ where the mixed distribution changes from a leptokurtic to a platykurtic distribution. This results in the power being trimodal having a high central peak when the mixture is symmetric and bimodal, and two lower peaks when the mixture is leptokurtic due to a low level of contamination.

v) The Studentized Range (U) behaves in a similar fashion to K and since its power is much lower it warrants little further attention.

vi) Pearson's P Statistic also behaves like K but, as its central peak is broader and its two side peaks lower, would appear marginally more valuable. The latter of these two points is considered advantageous as it would mean P is less likely to include deviations due to 'outliers' amongst its rejected distributions.

vii) All tests (except C, U and H) fail to distinguish 'outliers' from other forms of non-normality. Although the figure may suggest that in excess of 4 percent outliers is required to give reasonable levels of power it should be remembered that the model has constrained the magnitude of outliers to levels far below what may occur in practice. Thus, even lower frequencies of outliers may hinder the technique.

viii) The behaviour of the Simple Joint Test (J) can be seen to be a conservative combination of K and S. This results from the lack of independence and the fact that a significance level of $\alpha/2 = 0.05$ has been chosen for the two components (K and S). The observed significance level for the joint test was 0.0868 and 0.0897 for N equal to 50 and 100 respectively. (This approach has been used as it provides a simple method for approximating the two critical values corresponding to $\alpha = 0.10$.)

ix) Haldane's Test for Bimodality (H) demonstrates extremely poor power at any level of mixing and since the test requires much larger amounts of computer time it must be concluded that the test will rarely be used in practice.

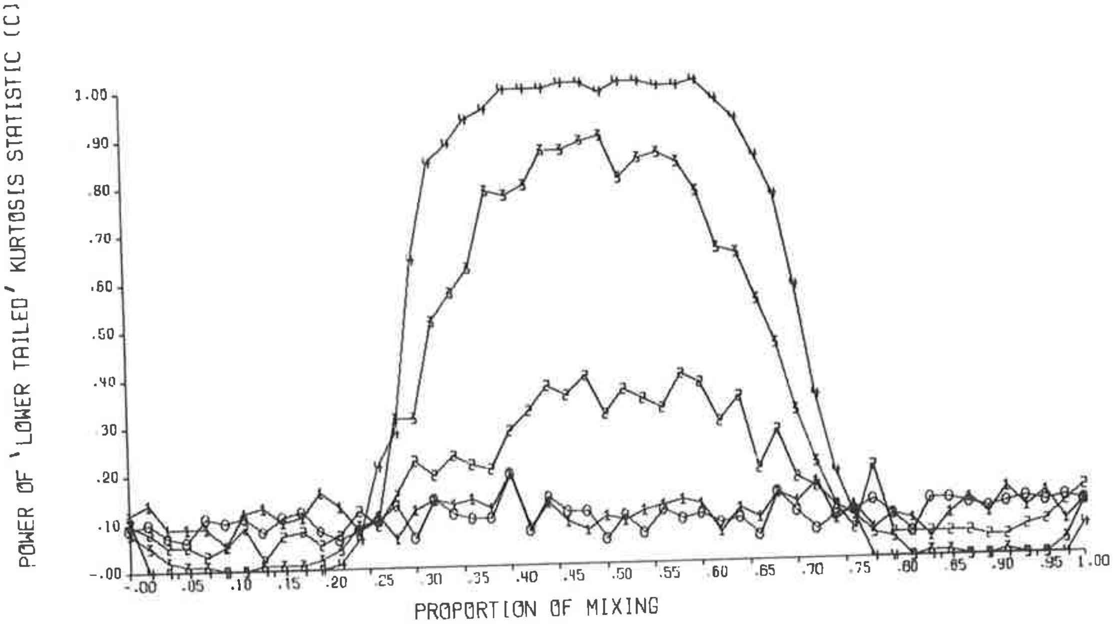
x) When Kurtosis was applied as a one-tailed test, the power increased for the platykurtic region between $p = 0.25$ and 0.75 and reduced elsewhere. Thus, if there was strong prior evidence suggesting that the major gene occurred with a frequency between 0.4 and 0.6 then the "lower-tailed" Kurtosis Test (C) would provide the best test for non-normality.

xi) The modified Shapiro-Wilk Statistic (W') can be seen to be as powerful as any of the statistics for frequencies up to 0.3 (or above 0.7). For the intermediate frequencies its power is comparatively lower. However, although this reduction in power becomes larger the fewer the number of observations (Figures IV-3, IV-4), this statistic does not demonstrate the extreme lack of power at any proportion which contrasts with that shown by most of the other tests.

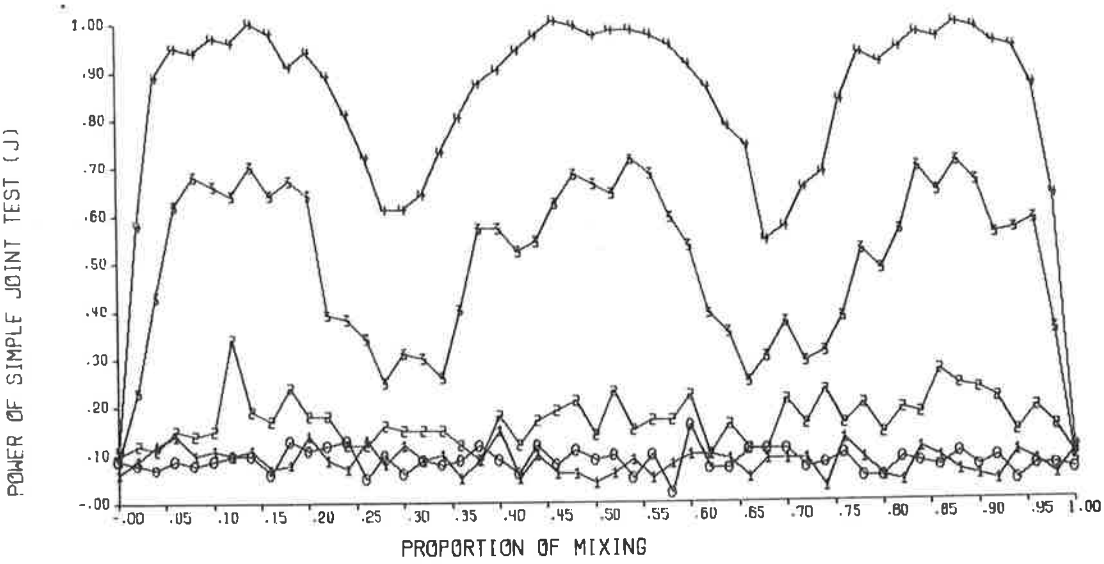
As major genes could also cause variation in scale this could have been investigated using an approach similar to the above (e.g. mixing 2 populations $N(0,1)$ and $N(0,\lambda)$ where $\lambda = 1,2,3,4$). Since such genes would be of limited importance, this type of model has been ignored. It is however more likely that major genes would exhibit both scale and location differences and this more complex situation may be worth consideration, although it would be expected that the additional effect of variance would increase the power of the test for non-normality.

Figure VI-1(1) Power of the "Lower Tailed" kurtosis statistic (C) plotted against the level of contamination for $n = 50$ and $\alpha = 0.10$.

Figure VI-1(2) Power of the Simple Joint Test (J) plotted against the level of contamination for $n = 50$ and $\alpha = 0.10$.



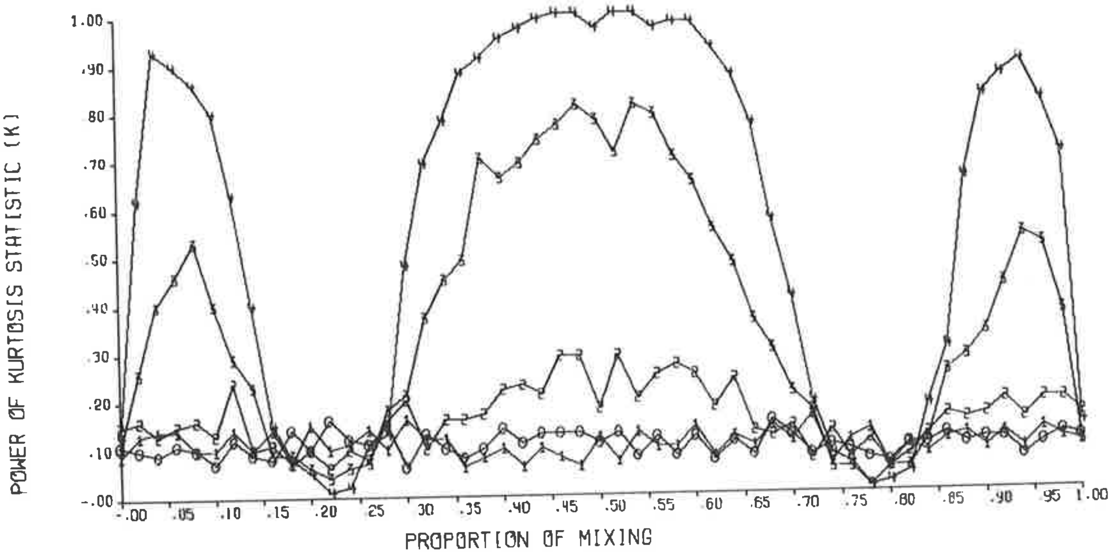
WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS



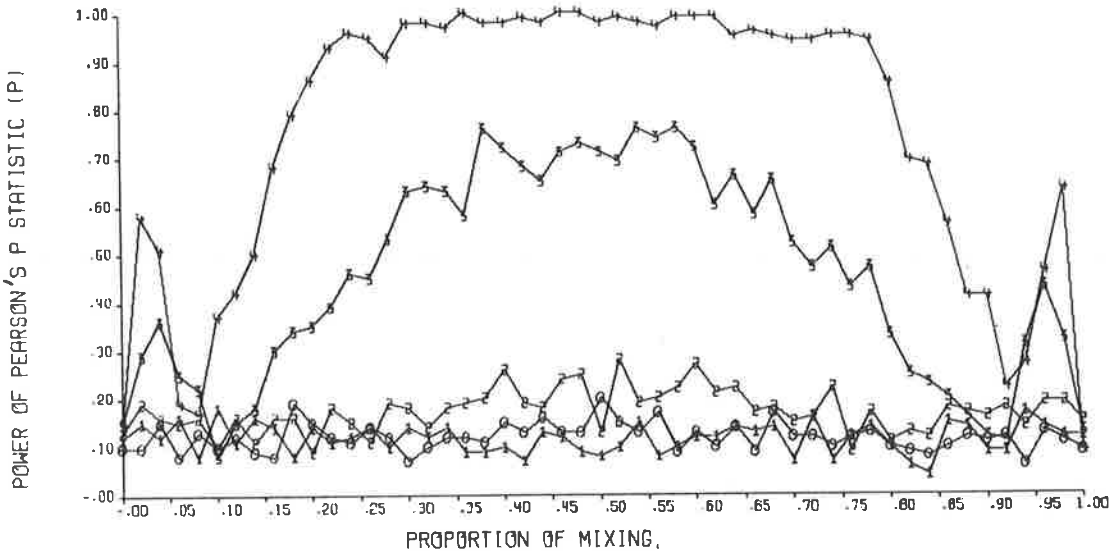
WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS

Figure VI-1(3) Power of the Kurtosis Statistic (K) plotted against the level of contamination for $n = 50$ and $\alpha = 0.10$.

Figure VI-1(4) Power of the Pearson's P Statistic plotted against the level of contamination for $n = 50$ and $\alpha = 0.10$.



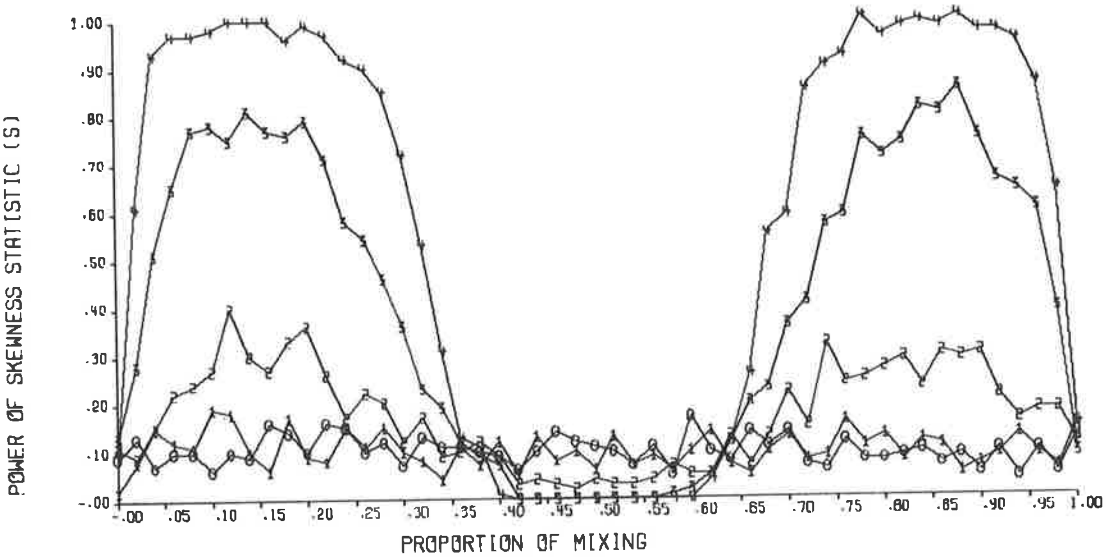
WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS



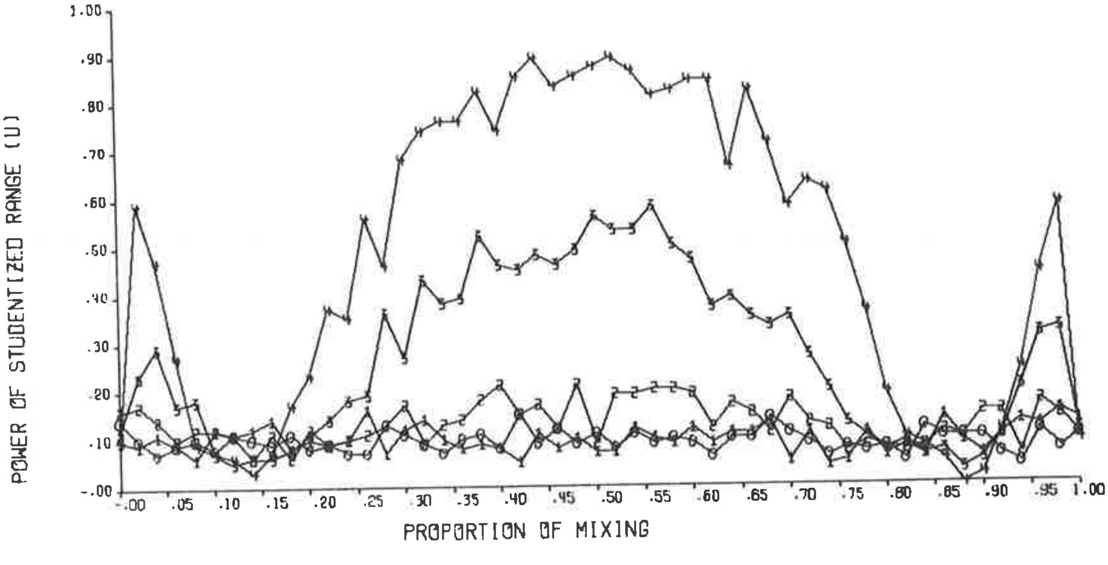
WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS

Figure VI-1(5) Power of the Skewness Statistic (S) plotted against the level of contamination for $n = 50$ and $\alpha = 0.10$.

Figure VI-1(6) Power of the Studentized Range (U) plotted against the level of contamination for $n = 50$ and $\alpha = 0.10$.

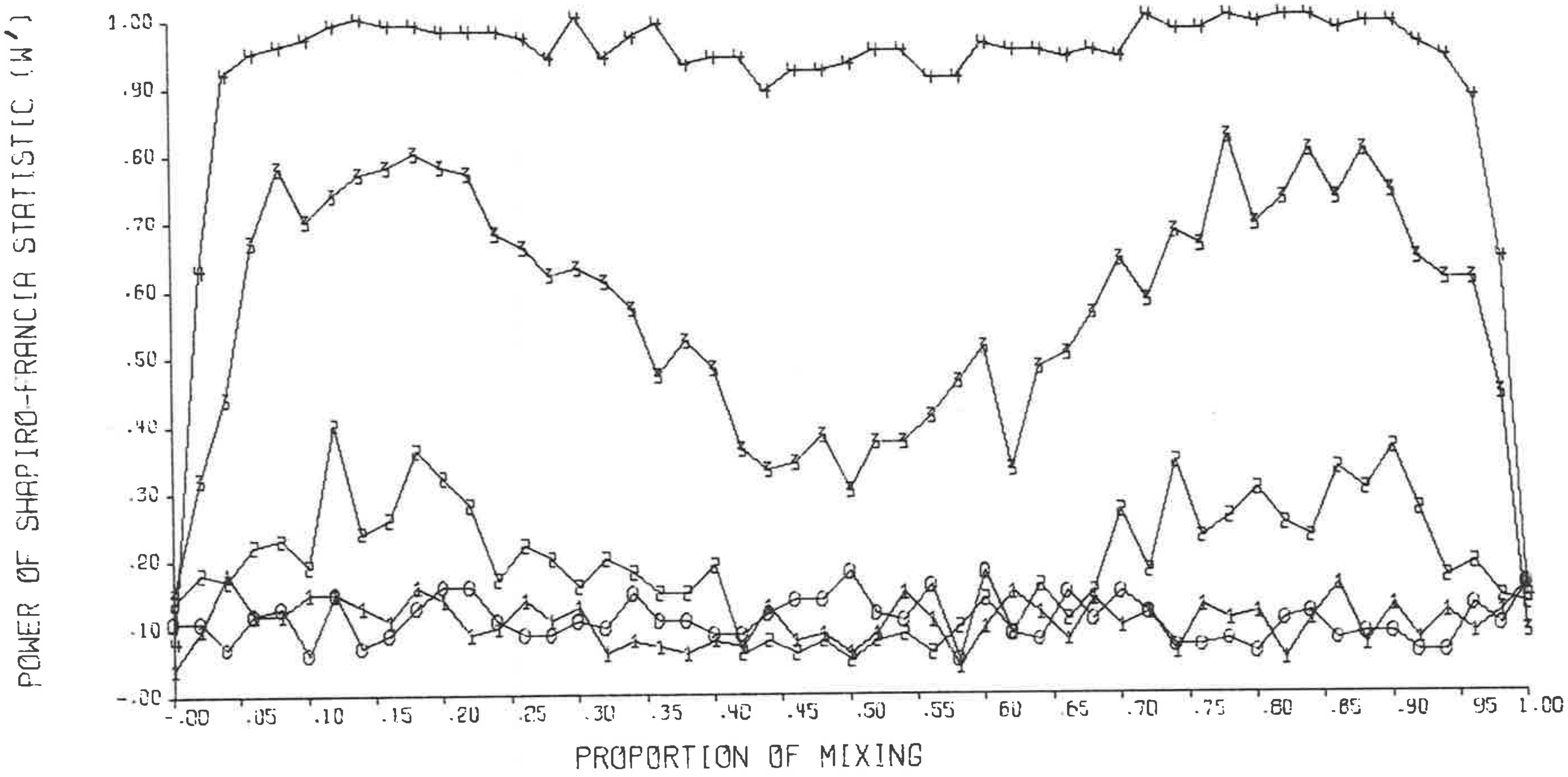


WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS



WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS

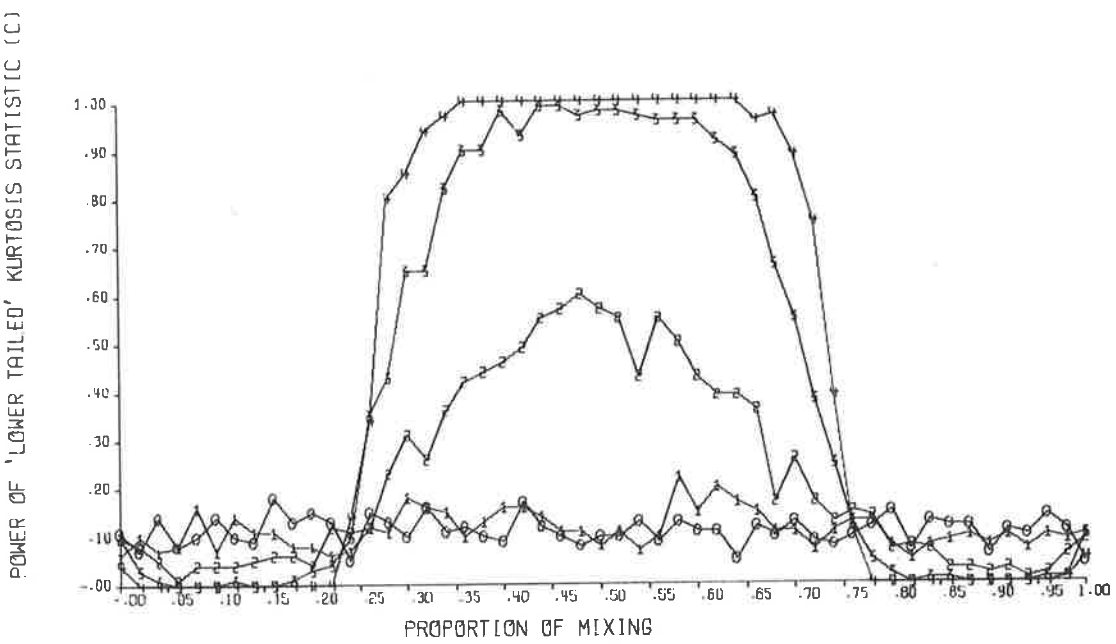
Figure VI-1(7) Power of the Shapiro-Francia Statistic (W') plotted against the level of contamination for n = 50 and $\alpha = 0.10$.



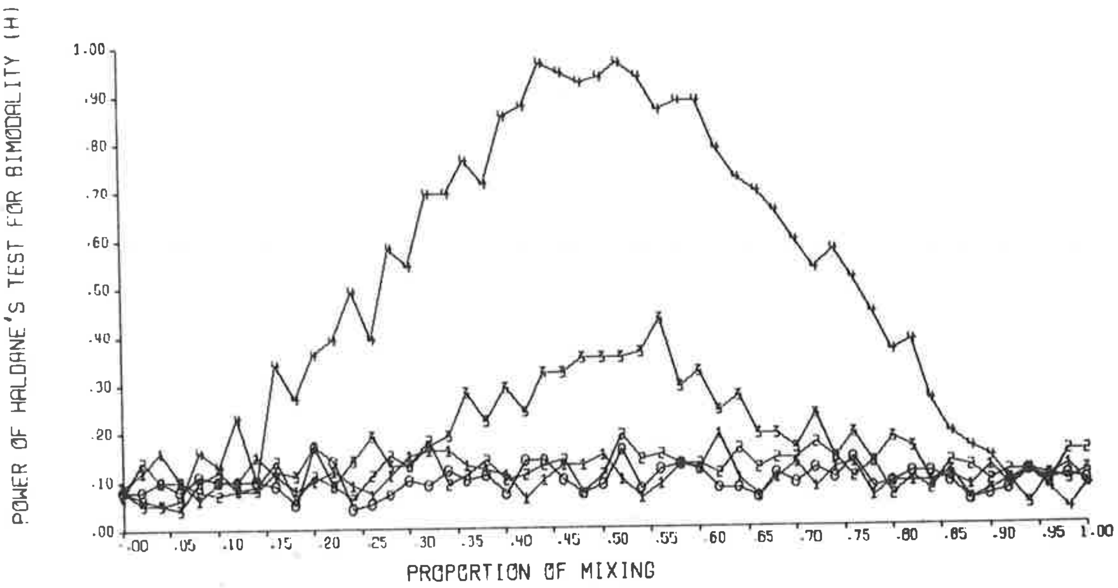
WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS

Figure VI-2(1) Power of the "Lower Tailed" Kurtosis Statistic (C) plotted against the level of contamination for $n = 100$ and $\alpha = 0.10$.

Figure VI-2(2) Power of Haldane's Test (H) plotted against the level of contamination for $n = 100$ and $\alpha = 0.10$.



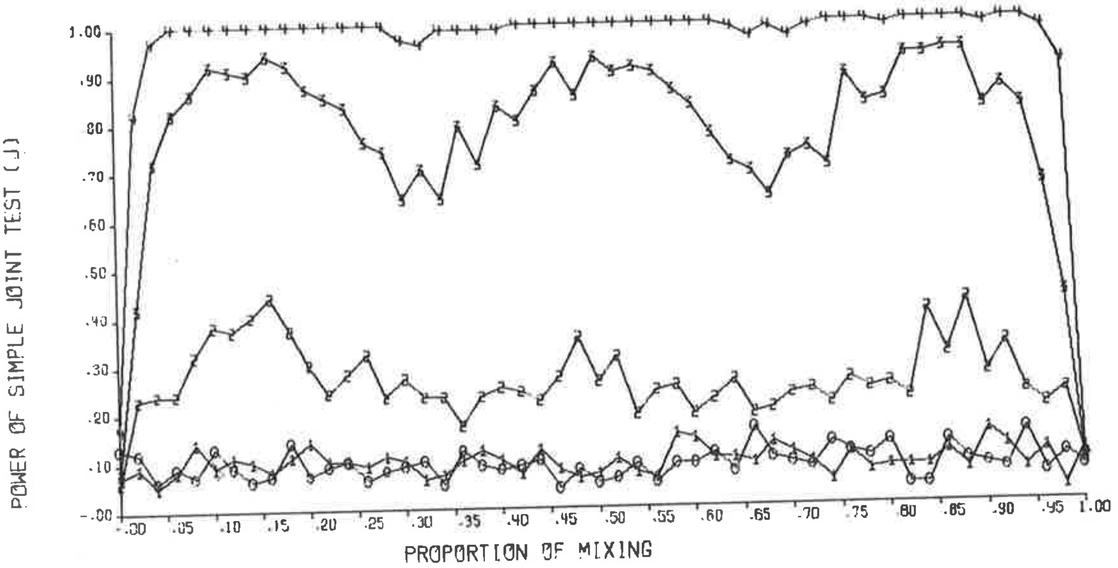
WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS



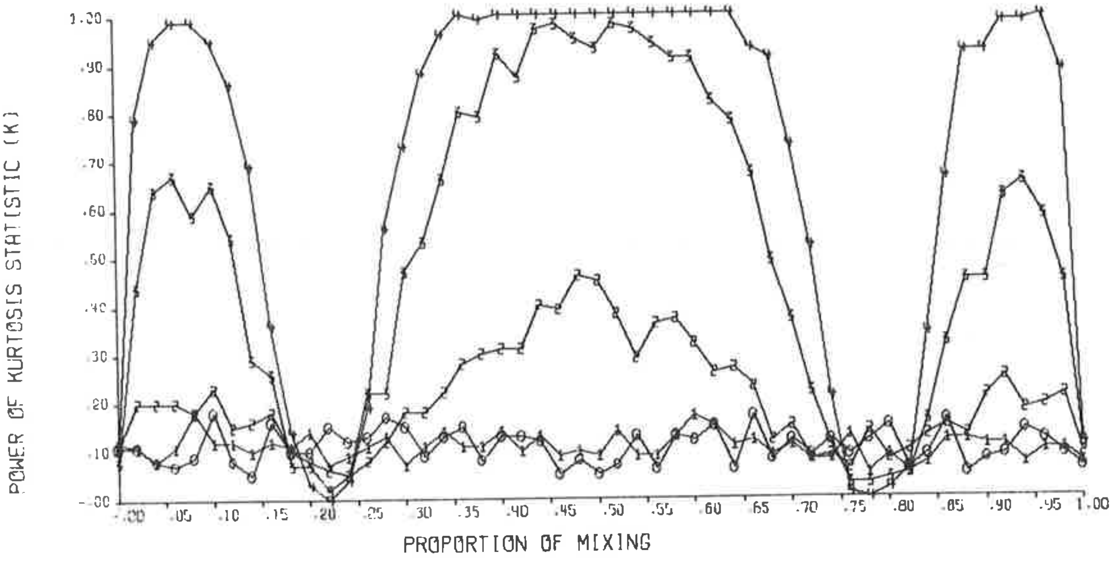
WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS

Figure VI-2(3) Power of the Simple Joint Test (J) plotted against the level of contamination for $n = 100$ and $\alpha = 0.10$.

Figure VI-2(4) Power of the Kurtosis Statistic (K) plotted against the level of contamination for $n = 100$ and $\alpha = 0.10$.



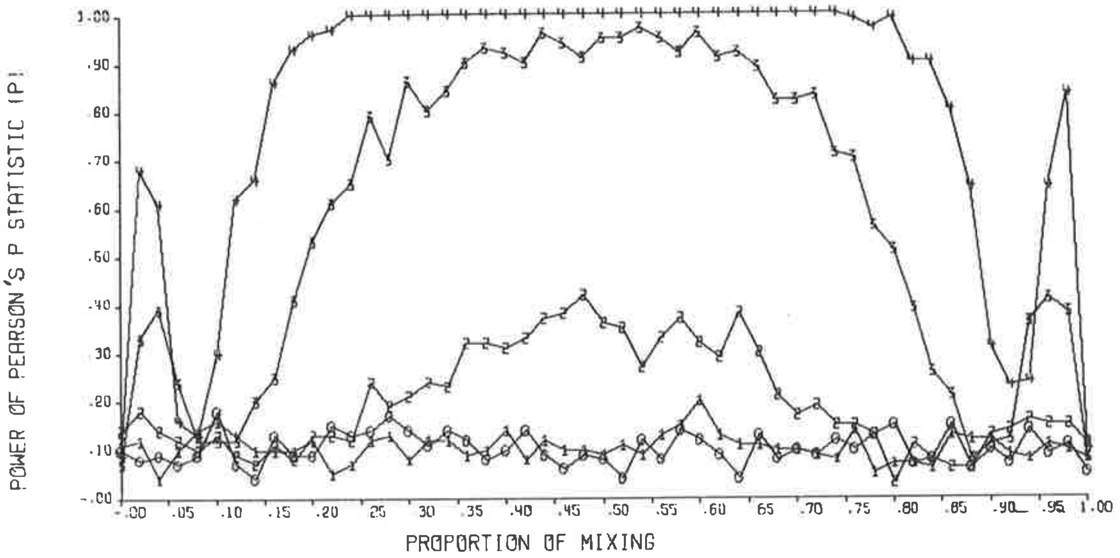
WHERE 0,1,2,3,4 INDICATE THE DIFFERENCE BETWEEN THE TWO POPULATION MEANS



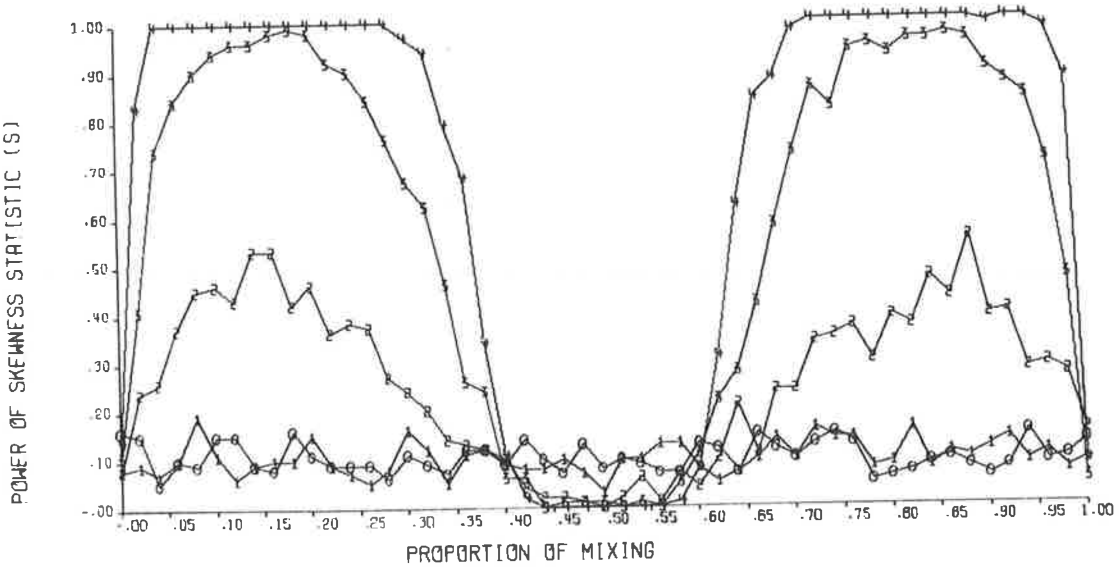
WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS

Figure VI-2(5) Power of the Pearson's P Statistic plotted against the level of contamination for $n = 100$ and $\alpha = 0.10$.

Figure VI-2(6) Power of the Skewness Statistic (S) plotted against the level of contamination for $n = 100$ and $\alpha = 0.10$.



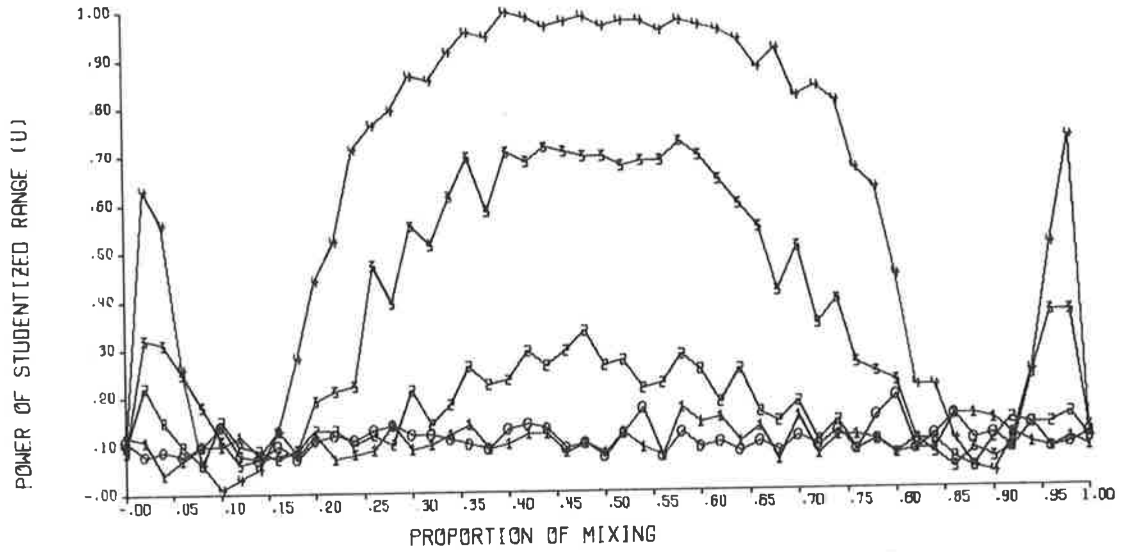
WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS



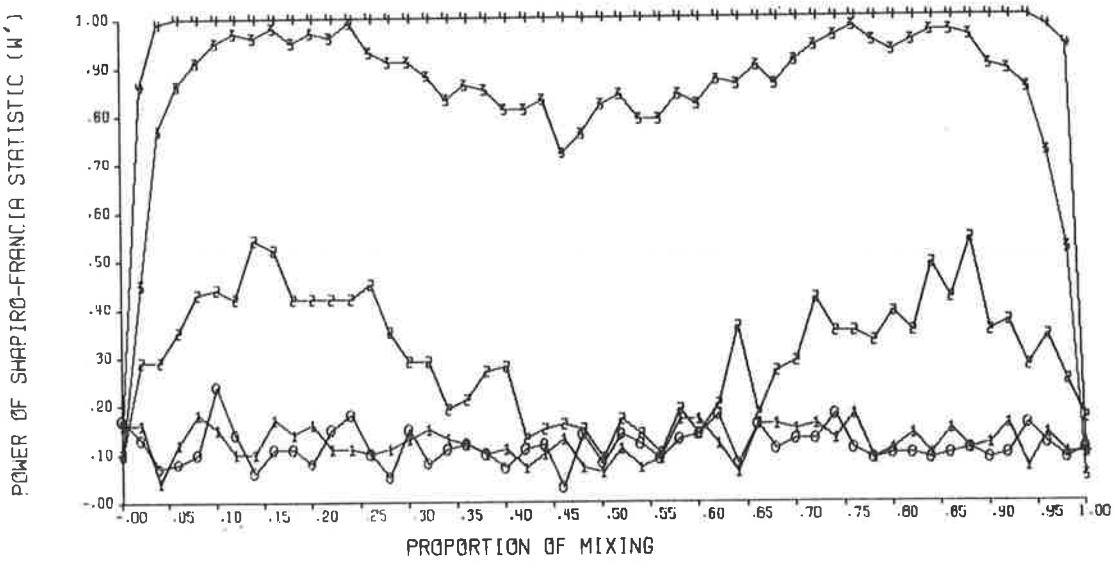
WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS**

Figure VI-2(7) Power of the Studentized Range (U) plotted against the level of contamination for $n = 100$ and $\alpha = 0.10$.

Figure VI-2(8) Power of the Shapiro-Francia Statistic (W') plotted against the level of contamination for $n = 100$ and $\alpha = 0.10$.



WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS



WHERE 0,1,2,3,4 INDICATE THE DISTANCE BETWEEN THE TWO POPULATION MEANS

Figure VI-3 Comparison of the power curves for the following seven tests (when $\lambda = 3$, $\alpha = 0.10$, $n = 50$)

(i) "Lower Tailed" Kurtosis Statistic - C

(ii) Simple Joint Test - J

(iii) Kurtosis Statistic - K

(iv) Pearson's P Statistic - P

(v) Skewness Statistic - S

(vi) Studentized Range - U

(vii) Shapiro-Francia Statistic (W') - W

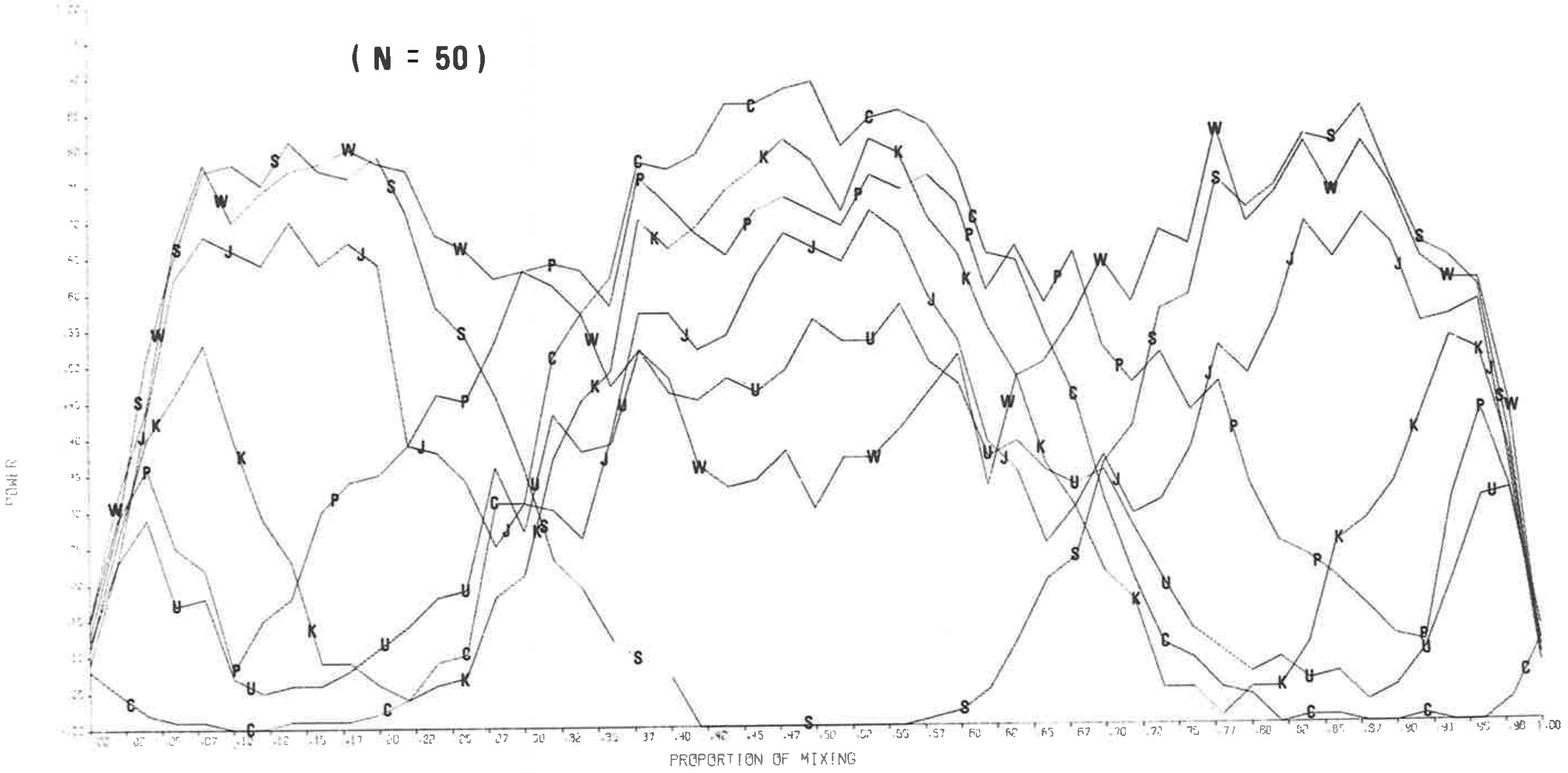
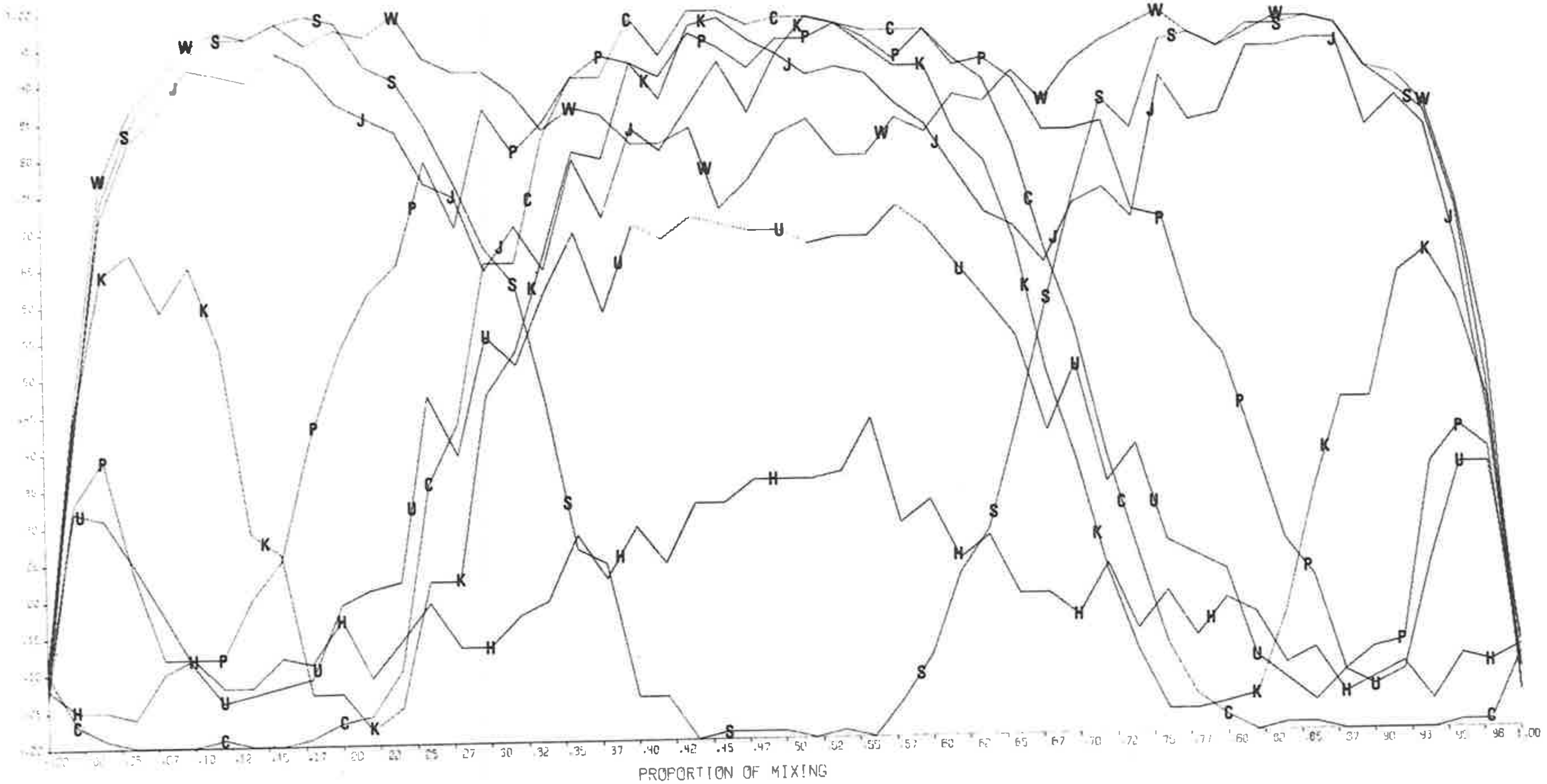

(N = 50)
POWER
PROPORTION OF MIXING
.60
.55
.50
.45
.40
.02
.05
.07
.10
.12
.15
.17
.20
.22
.25
.27
.30
.32
.35
.37
.40
.42
.45
.47
.50
.52
.55
.57
.60
.62
.65
.67
.70
.72
.75
.77
.80
.82
.85
.87
.90
.93
.95
.98
1.00
C
J
K
P
S
U
W

Figure VI-4 Comparison of the power curves for the following eight tests (when $\lambda = 3$, $\alpha = 0.10$, $n = 100$)

(i) "Lower Tailed" Kurtosis Statistic - C
(ii) Haldane's Test - H
(iii) Simple Joint Test - J
(iv) Kurtosis Statistic - K
(v) Pearson's P Statistic - P
(vi) Skewness Statistic - S
(vii) Studentized Range - U
(viii) Shapiro-Francia Statistic (W') - W

(N =100)


POWER
PROPORTION OF MIXING
1.00
.00
.02
.05
.07
.10
.12
.15
.17
.20
.22
.25
.27
.30
.32
.35
.37
.40
.42
.45
.47
.50
.52
.55
.57
.60
.62
.65
.67
.70
.72
.75
.77
.80
.82
.85
.87
.90
.93
.95
.98
1.00
W
S
J
K
P
U
H
C

DISCUSSION

Clearly, one should not use several or all of the eight tests simultaneously on a given set of data, but should choose the most appropriate on the basis of prior knowledge. Of the eight tests for non-normality investigated, W' (modified Shapiro-Wilk Statistic) provides the most versatile means for testing the type of deviations from normality which would be expected in the presence of a major gene. However, if prior knowledge suggested that the frequency of the major gene was near 0.5 then the 'lower-tail' test of kurtosis (C) would be preferable. Thus, for a limited range of frequencies, C provides a useful test unaffected by the presence of outliers. In comparison, W' is far less robust to the presence of outliers but it can be used effectively over a much wider range of contamination. Such bimodal distributions may be extremely important in other fields of genetics, and indeed in science generally, but it is likely that major genes would occur over a wider range of frequencies.

While both W' and C are best attempted with the use of a computer, calculation of the former requires somewhat more elaborate programming to generate the expected values of the normal order statistics and to rank the observations. To assist this computation a set of FORTRAN subroutines has been written following the method outlined by Shapiro and Francia (1972). (See Appendix E, where these subroutines form part of a program which will be discussed later.)

While the discussion so far has centred on the identification of mixed distributions, it is worth noting that Subrahmaniam, Subrahmaniam and Messeri (1975) have looked at the behaviour of three tests of significance when sampling is from such mixtures of two normal populations. The tests considered were:

i) One sample t-test.

ii) Analysis of variance test for equality of two or more means.

iii) Analysis of variance test for regression coefficient equal to zero.

Their investigation indicates that the first two tests are robust when the contamination is small, while the third exhibits only a minor effect with respect to the significance level. From this investigation it would seem reasonable to conclude that inability to recognise mixed distributions will not greatly weaken any analysis from a statistical point of view, however, for the quantitative geneticist, the loss may be considerable.

Previously, I indicated that the investigation had concentrated on the first of two related problems. This approach was adopted, as the establishment of non-normality was considered as being a distinct preliminary problem from understanding or identifying the cause. Although this approach has not been widely adopted in the past, it was considered to be preferable, as the user is provided with a convenient test to screen for departure from normality and so provide statistical justification for proceeding with the more complicated second stage of the examination.

Attempts to unravel the inheritance of quantitative characters may be broadly divided into two categories. First, Mather and Harrison (1949); Thoday (1961); Cooke and Mather (1962); Gibson and Thoday (1962); Spickett (1963); Wolstenholme and Thoday (1963); Thoday, Gibson and Spickett (1964); Spickett and Thoday (1966); and Law (1967) have developed methods based on chromosome assay while secondly, Fisher, Immer and Tedin (1932); Wright (1934); Panse (1940); Mather (1949); Kalmus and Maynard Smith (1965); Mérat (1968); O'Donald (1971); Hammond and James (1972); and Jinks and Towey (1976) have attempted to use the statistical properties of the distribution to estimate the number of "effective factors". As the former approach relies on a detailed knowledge of the genome and the availability of suitable marker genes, for progeny testing, its use will remain limited to all but a few species. This has been supported by Piper (1971) who carried out a detailed assessment of Thoday's (1961) method. Piper concluded that the method was unsuitable for domestic animal breeding where the linkage map was poorly documented and suppression of crossing-over was impossible. Even if these facilities were available the analyses would require much care to avoid producing unreliable answers. Recently, McMillan and Robertson (1974) have questioned the accuracy of Thoday's procedure on the following two grounds:

i) it may detect loci which do not exist.

ii) The estimated effect of those major loci which do exist will almost inevitably be magnified by the accumulation of effects from closely linked undetected loci.

Also, any detection will depend on the standard (tester) stock used (Thoday, 1973). Thus, while its application to *Drosophila*, by Gibson and Thoday (1962); Wolstenholme and Thoday (1963); Spickett (1963); Thoday, Gibson and Spickett (1964); Spickett and Thoday (1966); has proven informative, it is unlikely that it will be suitable for more general use.

In comparison, the latter methods, based on properties of the distributions, must be handled with care as they rely on assumptions whose failure can greatly distort the estimates (see Hammond and James, 1972; Hopkins, 1974). Also, both approaches are directed more at the estimation of polygenic variation (except for Kalmus and Maynard Smith, 1965), whereas the interest of the present investigation is in major genes. Although the division between polygenes and major genes is far from distinct (Thompson and Thoday, 1974), it was felt that concentration on the latter would provide a more sensible starting point as the possible gains for the breeder are potentially greater, particularly in the short term.

Piper (1971) also studied the feasibility of Wright's (1952) backcrossing method for isolation of genes of large effect. He concluded that the number of generations required to identify a major gene was sufficiently large to make the "technique impractical for most species." Secondly, the technique was unable to distingiush between a closely linked group of polygenes and a major gene. This second criticism may be directed, equally well, at the method developed in the present study, however, the effect of recombination, although important in the long term, would be expected to be of far less significance in domestic animals where the generation interval is comparatively long.

As mentioned briefly above, Kalmus and Maynard Smith (1965) developed formulae for the estimation of gene frequencies,means and variances associated with incomplete dominance at a single loci. As the method relies on the existence of a distinct antimode it would seem of only limited value. Also, three more general methods have recently been reported in the statistical literature for estimating the respective means, variances and frequencies of the distinct populations for data from mixed populations. These methods are:

i) moments estimates;

ii) maximum likelihood;

iii) minimum χ^2.

Although these methods were not specifically developed with their genetic use in mind, they would nevertheless appear to be of considerable potential for locating major genes. Fryer and Robertson (1972) compared these three methods and concluded:

> "The methods do not differ essentially with regard to bias but for the mean squared error the grouped estimates are shown to be more

> accurate than the moments estimates for most distributions, though the moment estimates seem preferable for distributions which are particularly difficult to estimate."

These authors further concluded that the methods of maximum likelihood and minimum χ^2 were of similar accuracy. In view of this last conclusion, the lesser known method of minimum χ^2 will not be considered further in this investigation.

Although the method of moments has only recently emerged as a method for estimating the five parameters associated with the mixture of two normal distributions, the method was first considered by Pearson (1894). Hasselblad (1966) reconsidered the problem while Cohen (1967) showed that the computation could be simplified from Pearson's original method which required the solution of a ninth degree equation. The problem has attracted further attention from Robertson and Fryer (1970); Behboodian (1970); John (1970); Hawkins (1971); Fryer and Robertson (1972); Rayment (1972); and Hawkins (1972) and others. The last author has made available listings of a FORTRAN program (Hawkins, 1975) which, following some minor modifications, has provided a valuable method for estimating the five parameters associated with the mixture of two normal distributions. The procedure has been checked against mixtures of known means, variances and proportions and found to provide a reasonable level of accuracy. However, it did show some tendency to identify falsely two normal distributions when the data had been generated from only one normal distribution. Incorrect estimates also arose when the level of contamination was low or the respective means and variances were extremely similar. To check the validity of the two sets of estimated means and variances, Hawkins had included in his program a variance ratio test and an independent t-test. However, such an approach was unsatisfactory as these estimates were tested on the data which generated them. In comparison, the preliminary testing, using the W' or C statistics as proposed above, would seem to provide a less biased method for protecting against such false estimates. This was in fact supported by analysis of simulated data.

Hawkins (1972) points out that the method of moments may sometimes generate more than one set of solutions for a given set of data. As previous algorithms failed to identify such additional solutions, their continued use would be ill-advised since the solution omitted could be the more suitable one for the data. Clearly the identification of multiple solutions requires their relative merit to be assessed, however, this situation is preferable to choosing, arbitrarily, only one solution.

The following three alternatives were proposed by Hawkins as providing a possible decision rule:

i) comparison of the observed and theoretical sixth moments.

ii) Checking of the medians for each solution.

iii) Use of a χ^2 test to compare the observed classed data to the expected frequencies for each possible solution.

While all three of the above may form useful criteria, there will always remain some situations where it is impossible to identify one solution as being better than the others.

A general FORTRAN program has been written to apply first the Shapiro-Francia Statistic (W') and subsequently the method of moments if there is evidence of non-normality. (The full listing is included in Appendix E.) This program has been tested with simulated data and found to be capable of correctly locating the two populations when the differences between means are sufficient to cause W' to be significant. While this latter condition is seldom satisfied for means which differ by less than two standard deviations, the procedure would still seem worthy of consideration. However, it should also be remembered that any bimodality identified may sometimes result from non-genetic factors. Therefore, the procedure provides the geneticist with a convenient screening method but follow-up investigations should be carried out to check that a major gene is in fact segregating.

As noted above, the alternative method of maximum likelihood has been demonstrated by Fryer and Robertson (1972) to provide a comparable technique for estimating the five parameters associated with mixed distributions. The method has been demonstrated by Lester, Elston and Graham (1972) and Elston, Namboodiri, Nino and Pollitzer (1974) to identify factors which indicate the presence of a major gene. Although a program listing of the maximum likelihood method has been obtained (Morton and Yee, personal communication to O. Mayo and T.W. Hancock, 1976), it has not as yet been successfully implemented on the local computer. However, this program has since been used extensively by its authors who have developed it to remove the effect of skewness before estimating the parameters using the maximum likelihood method (see Maclean, Morton, Elston and Yee (1976)).

In the preceding it has been assumed that one is either dealing with one or a mixture of two normal distributions. Obviously the presence of non-normality due to any other reasons will weaken all of the above methods of analysis. The use of power transformations by Maclean, Morton, Elston and Yee to remove skewness before attempting to estimate the required parameters appears reasonable, however, it remains to be shown that a single transformation determined on all of the data will provide the most suitable transformation for situations where there are two or more dissimilar distributions. Also, before adopting this approach, it would seem worthwhile to investigate the robustness of the above two estimation techniques in the presence of non-normality. Also, as neither method determines which distribution a particular observation actually arose from, it is difficult to imagine how this problem of non-normality will ever be effectively resolved.

The present approach has assumed normality, however, the possible effects of the failure of this assumption should not be ignored.

APPLICATION

The procedure developed above has been applied to the Roseworthy data using the subroutines listed in Appendix E. (The main program BIMOD was replaced by code which was better suited to handle the intricacies of the Roseworthy data.) Only single born progeny were considered and the test was applied within the sex by flock by year subgroups. Tables VI-1 to VI-4 summarise the results for the variates body weight, clean fleece weight, fibre diameter and secondary follicle number respectively, where parts (1) and (2) of each refer to the two sexes. In each table the number of observations (n) is given with the value of W' and its probability. If W' is significant, at the 5 percent level, the method of moments is then applied. Where this latter method is unable to find any solutions which satisfy the moment equations the statement "NO VALID SOLUTIONS" is printed by the program. It would appear reasonable to conclude that for these cases the non-normality detected is due to reasons other than bimodality. However, in most cases the method obtains estimates for the two means ($\bar{x}_1$, $\bar{x}_2$) and standard deviations (s_1, s_2) and the proportion of mixing (p_1). Occasionally two valid solutions have been detected for the one set of data (see Hawkins, 1972). In this application no attempt has been made to determine which set of solutions is better although this could have been done (see earlier discussion).

Inspection of the four tables shows that there is scattered evidence of bimodality, however, there is little consistency over sexes within flock type. This could either be associated with the small sample size or an indication of sex-linkage. While Bielharz (1963) has suggested that sex-linkage is more widespread than generally believed, this is not supported in the tables where the bimodality is observed to have a similar frequency in the two sexes. For body weight in 1961 and 1962, both sexes of the Visual flock are observed to suggest the presence of a major gene.

From the results it can be concluded that the technique provides a workable method for estimating the components of a mixed distribution. The preliminary test for normality (W') reduces the possibility of falsely identifying two populations when the distribution can be explained in terms of one. This also means that the computer time required for scanning data is greatly reduced as the program for the methods of moments estimation is much more time consuming than the preliminary testing of normality.

It should be remembered that the test as applied assumes we are dealing with a population which is uniform except for the segregation of a major gene. Thus, if we are in fact looking at data which shows heterogeneity, for other reasons, the test may be misleading. For example, if we had not partitioned the sheep progeny on sex (or birth type) we would more than likely have observed bimodality. Or, if the ram fertility had been poor, necessitating a return to service (as occurred in 1965), this may produce a bimodal distribution for time of birth which may still be reflected in measurements made up to a year later in the progeny. Therefore, when evidence of bimodality occurs the user should check that a non-genetic factor is not responsible.

Table VI-1(1) Results of bimodality analysis on the body weights (in kg) of the male progeny. The two flocks (I - Index and V - Visual) are analysed separately for each of the twelve years of the trial.

		$n^{\dagger}$	W'	Sig	p_1	$\bar{x}_1$	s_1	p_2	$\bar{x}_2$	s_2
1954	I	71	.983	ns						
	V	63	.990	ns						
1955	I	71	.970	ns						
	V	79	.969	*	.53	55.5	6.67	.47	56.9	3.02
1956	I	66	.991	ns						
	V	66	.989	ns						
1957	I	56	.977	ns						
	V	84	.977	ns						
1958	I	67	.988	ns						
	V	62	.982	ns						
1959	I	74	.989	ns						
	V	71	.982	ns						
1960	I	59	.976	ns						
	V	56	.976	ns						
1961	I	70	.987	ns						
	V	73	.958	*	.37	55.6	8.03	.62	61.6	4.17
1962	I	80	.969	*		No valid solutions				
	V	85	.967	*	.31	57.1	7.30	.69	61.8	3.77
1963	I	64	.975	ns						
	V	69	.966	*		No valid solutions				
1964	I	76	.985	ns						
	V	67	.984	ns						
1965	I	46	.977	ns						
	V	77	.985	ns						

† n = sample size

W' = Shapiro-Francia statistic with associated significance

p_1, $\bar{x}_1$ and s_1 = the estimates of the relative proportion, mean and standard deviation of population one ,respectively

$p_2(=1-p_1)$, $\bar{x}_2$ and s_2 = similar estimates for population two

Table VI-1(2) Results of bimodality analysis on the body weights (in kg) of the female progeny. The two flocks (I - Index and V - Visual) are analysed separately for each of the twelve years of the trial.

		$n^{\dagger}$	W'	Sig	p_1	$\bar{x}_1$	s_1	p_2	$\bar{x}_2$	s_2
1954	I	55	.901	**	.18	42.3	8.41	.82	47.9	3.34
	V	71	.964	*		No valid solutions				
1955	I	62	.949	*	.83	43.4	3.03	.17	45.8	6.40
	V	58	.986	ns						
1956	I	46	.974	ns						
	V	78	.989	ns						
1957	I	67	.988	ns						
	V	62	.973	ns						
1958	I	57	.981	ns						
	V	60	.990	ns						
1959	I	72	.948	**		No valid solutions				
	V	75	.981	ns						
1960	I	83	.980	ns						
	V	73	.985	ns						
1961	I	71	.988	ns						
	V	86	.949	**	.13	34.2	7.89	.87	40.3	3.98
1962	I	72	.976	ns						
	V	82	.970	*	.53	42.8	2.43	.47	50.1	1.86
1963	I	43	.980	ns						
	V	70	.983	ns						
1964	I	74	.972	ns						
	V	67	.990	ns						
1965	I	37	.936	*	.46	39.6	1.61	.54	45.7	5.31
					.69	40.4	2.67	.31	48.8	4.34
	V	70	.981	ns						

† n = sample size

W' = Shapiro-Francia statistic with associated significance

p_1, $\bar{x}_1$ and s_1 = the estimates of the relative proportion, mean and standard deviation of population one, respectively

$p_2 (=1-p_1)$, $\bar{x}_2$ and s_2 = similar estimates for population two

Table VI-2(1) Results of bimodality analysis on the clean fleece weight (in kg) of the male progeny. The two flocks (I - Index and V - Visual) are analysed separately for each of the twelve years of the trial.

		$n^{\dagger}$	W'	Sig	p_1	$\bar{x}_1$	s_1	p_2	$\bar{x}_2$	s_2
1954	I	71	.986	ns						
	V	63	.978	ns						
1955	I	71	.987	ns						
	V	79	.982	ns						
1956	I	66	.968	ns						
	V	66	.990	ns						
1957	I	56	.980	ns						
	V	84	.986	ns						
1958	I	67	.989	ns						
	V	62	.947	**	.31	3.68	.76	.69	4.10	.35
1959	I	74	.985	ns						
	V	71	.992	ns						
1960	I	59	.959	*	.76	4.60	.41	.24	5.07	.76
	V	56	.975	ns						
1961	I	70	.990	ns						
	V	73	.979	ns						
1962	I	80	.991	ns						
	V	85	.981	ns						
1963	I	64	.991	ns						
	V	69	.971	ns						
1964	I	76	.992	ns						
	V	67	.968	ns						
1965	I	46	.957	ns						
	V	77	.989	ns						

† n = sample size

W' = Shapiro-Francia statistic with associated significance

p_1, $\bar{x}_1$ and s_1 = the estimates of the relative proportion, mean and standard deviation of population one, respectively

$p_2 (=1-p_1)$, $\bar{x}_2$ and s_2 = similar estimates for population two

Table VI-2(2) Results of bimodality analysis on the clean fleece weight (in kg) of the female progeny. The two flocks (I - Index and V - Visual) are analysed separately for each of the twelve years of the trial.

		n[†]	W'	Sig	p_1	$\bar{x}_1$	s_1	p_2	$\bar{x}_2$	s_2
1954	I	55	.977	ns						
	V	71	.972	ns						
1955	I	62	.983	ns						
	V	58	.935	**	No valid solutions					
1956	I	46	.984	ns						
	V	78	.960	*	.69	3.12	.27	.31	3.53	.46
					.85	3.14	.29	.15	3.82	.38
1957	I	67	.984	ns						
	V	62	.948	*	.84	2.86	.33	.16	3.12	.59
1958	I	57	.965	ns						
	V	60	.981	ns						
1959	I	72	.985	ns						
	V	75	.991	ns						
1960	I	83	.988	ns						
	V	73	.978	ns						
1961	I	71	.978	ns						
	V	86	.987	ns						
1962	I	72	.983	ns						
	V	82	.987	ns						
1963	I	43	.979	ns						
	V	70	.979	ns						
1964	I	74	.983	ns						
	V	67	.981	ns						
1965	I	37	.972	ns						
	V	70	.984	ns						

[†] n = sample size

W' = Shapiro-Francia statistic with associated significance

p_1, $\bar{x}_1$ and s_1 = the estimates of the relative proportion, mean and standard deviation of population one, respectively

$p_2(=1-p_1)$, $\bar{x}_2$ and s_2 = similar estimates for population two

Table VI-3(1) Results of bimodality analysis on the fibre diameter (in microns) of the male progeny. The two flocks (I - Index and V - Visual) are analysed separately for each of the twelve years of the trial.

		n†	W'	Sig	p_1	$\bar{x}_1$	s_1	p_2	$\bar{x}_2$	s_2
1954	I	71	.941	**	No valid solutions					
	V	63	.986	ns						
1955	I	71	.980	ns						
	V	79	.985	ns						
1956	I	66	.919	**	.56	20.4	.72	.44	22.5	2.47
	V	66	.944	**	.47	21.6	.76	.53	23.3	2.25
1957	I	56	.990	ns						
	V	84	.985	ns						
1958	I	67	.967	ns						
	V	62	.969	ns						
1959	I	74	.990	ns						
	V	71	.960	*	No valid solutions					
1960	I	59	.984	ns						
	V	56	.976	ns						
1961	I	70	.989	ns						
	V	73	.968	*	.68	23.0	1.63	.32	27.8	1.82
1962	I	80	.951	**	No valid solutions					
	V	85	.986	ns						
1963	I	64	.987	ns						
	V	69	.990	ns						
1964	I	76	.990	ns						
	V	67	.968	ns						
1965	I	46	.984	ns						
	V	77	.976	ns						

† n = sample size

W' = Shapiro-Francia statistic with associated significance

p_1, $\bar{x}_1$ and s_1 = the estimates of the relative proportion, mean and standard deviation of population one, respectively

$p_2 (=1-p_1)$, $\bar{x}_2$ and s_2 = similar estimates for population two

Table VI-3(2) Results of bimodality analysis on the fibre diameter (in microns) of the female progeny. The two flocks (I - Index and V - Visual) are analysed separately for each of the twelve years of the trial.

		$n^{\dagger}$	W'	Sig	p_1	$\bar{x}_1$	s_1	p_2	$\bar{x}_2$	s_2
1954	I	55	.937	**		No valid solutions				
	V	71	.969	ns						
1955	I	62	.990	ns						
	V	58	.988	ns						
1956	I	46	.985	ns						
	V	78	.981	ns						
1957	I	67	.991	ns						
	V	62	.984	ns						
1958	I	57	.984	ns						
	V	60	.989	ns						
1959	I	72	.983	ns						
	V	75	.981	ns						
1960	I	83	.992	ns						
	V	73	.974	ns						
1961	I	71	.971	ns						
	V	86	.982	ns						
1962	I	72	.985	ns						
	V	82	.993	ns						
1963	I	43	.979	ns						
	V	70	.965	*	.61	26.4	1.10	.39	29.8	1.58
1964	I	74	.984	ns						
	V	67	.989	ns						
1965	I	37	.973	ns						
	V	70	.986	ns						

† n = sample size

W' = Shapiro-Francia statistic with associated significance

p_1, $\bar{x}_1$ and s_1 = the estimates of the relative proportion, mean and standard deviation of population one, respectively

$p_2(=1-p_1)$, $\bar{x}_2$ and s_2 = similar estimates for population two

Table VI-4(1) Results of bimodality analysis on the secondary follicle number per sq cm of the male progeny. The two flocks (I - Index and V - Visual) are analysed separately for each of the twelve years of the trial.

		n†	W'	Sig	p_1	$\bar{x}_1$	s_1	p_2	$\bar{x}_2$	s_2
1954	I	71	.922	**		No valid solutions				
	V	63	.916	**	.89	4106.7	579.9	.11	5105.8	1387.1
1955	I	71	.972	ns						
	V	79	.960	*		No valid solutions				
1956	I	66	.990	ns						
	V	66	.928	**	.73	4835.8	544.8	.27	5659.9	1258.1
1957	I	56	.947	*		No valid solutions				
	V	84	.990	ns						
1958	I	67	.978	ns						
	V	62	.987	ns						
1959	I	74	.991	ns						
	V	71	.993	ns						
1960	I	59	.991	ns						
	V	56	.995	ns						
1961	I	70	.946	**		No valid solutions				
	V	73	.973	ns						
1962	I	80	.981	ns						
	V	85	.857	**		No valid solutions				
1963	I	64	.957	*		No valid solutions				
	V	69	.980	ns						
1964	I	76	.993	ns						
	V	67	.984	ns						
1965	I	46	.984	ns						
	V	77	.984	ns						

† n = sample size

W' = Shapiro-Francia statistic with associated significance

p_1, $\bar{x}_1$ and s_1 = the estimates of the relative proportion, mean and standard deviation of population one, respectively

$p_2 (=1-p_1)$, $\bar{x}_2$ and s_2 = similar estimates for population two

Table VI-4(2) Results of bimodality analysis on the secondary follicle number per sq cm of the female progeny. The two flocks (I - Index and V - Visual) are analysed separately for each of the twelve years of the trial.

		$n^{\dagger}$	W'	Sig	p_1	$\bar{x}_1$	s_1	p_2	$\bar{x}_2$	s_2
1954	I	55	.975	ns						
	V	71	.975	ns						
1955	I	62	.937	**	.81	5002.7	665.9	.19	6188.3	1333.8
	V	58	.942	**	.58	4767.5	395.1	.42	5727.8	874.3
1956	I	46	.961	ns						
	V	78	.974	ns						
1957	I	67	.989	ns						
	V	62	.991	ns						
1958	I	57	.958	*		No valid solutions				
	V	60	.991	ns						
1959	I	72	.984	ns						
	V	75	.988	ns						
1960	I	83	.958	**		No valid solutions				
	V	73	.962	*		No valid solutions				
1961	I	71	.972	ns						
	V	86	.982	ns						
1962	I	72	.948	**	.83 .86	4907.3 4919.6	779.4 791.2	.17 .14	6456.6 6685.9	1343.2 1278.3
	V	82	.984	ns						
1963	I	43	.961	ns						
	V	70	.976	ns						
1964	I	74	.980	ns						
	V	67	.978	ns						
1965	I	37	.983	ns						
	V	70	.950	**	.56	4637.0	446.9	.44	5435.0	1034.2

† n = sample size

W' = Shapiro-Francia statistic with associated significance

p_1, $\bar{x}_1$ and s_1 = the estimates of the relative proportion, mean and standard deviation of population one, respectively

$p_2 (=1-p_1)$, $\bar{x}_2$ and s_2 = similar estimates for population two

GENERAL CONCLUSIONS

As breeding of domestic animals is both expensive and time consuming, we should strive both for an efficient deployment of the resources and for the maximum rate of genetic gain.

The Roseworthy program was initiated to test whether selecting rams by direct measurement of the important metric character clean fleece weight was more effective than the conventional method of visual assessment. While the outcome of the trial is of considerable interest to practical breeders, the full appreciation of it depends upon the interpretation of a large number of statistics.

This thesis has been concerned with the clear and unambiguous assessment of breeding data. The biometrical techniques used here have been developed to investigate two aspects of interest. Firstly, to measure the changes in the populations under different methods of selection, and here Chapters I, II, III and IV are particularly appropriate. Secondly, to predict the changes which would occur if various methods of selection were to be employed (Chapters II, V, VI).

Chapters I and II dealt with the conventional methods of analysing breeding experiments, and the detailed statistics presented demonstrate the difficulties involved in interpreting the interrelationships between the characters, under direct and indirect selection. Unfortunately, these difficulties are not readily appreciated by either the breeder or the scientist and in practice they have usually been ignored.

In comparison, Hotelling's T^2, presented in Chapter III, provides a much simpler, but comprehensive answer to the same questions. Its application to the Roseworthy data has indicated that the two populations have diverged. In particular, staple length, clean scoured yield and secondary follicle number are positively associated with the increase in clean fleece weight while crimps per inch and body weight are negatively associated. The tabulation of the simultaneous confidence intervals over a number of years provides a direct and meaningful summary of the trends in these characters which are so important to the strategy of the breeder.

The improved algorithm for the Exact Test of R x C Contingency Tables, developed in Chapter IV, provides a valuable method for the analysis of discrete data such as reproductive records which cannot be handled by Hotelling's T^2. As the test is exact (i.e. it does not rely on approximation to a sampling distribution), it can be applied to tables irrespective of the minimum size of the expected cell frequencies. Thus it can be used to analyse tables where the approximate tests, such as Chi-squared, become inaccurate. Although an upper limit remains with respect to the size of table which can be analysed, this is seldom likely to be an important restriction.

In Chapter V, Principal Component Analysis is seen to provide a useful means for identifying the complex interrelationships between the observed variates. In particular, this method does not suffer from the extreme inconsistency which is seen to occur for the other techniques based on multiple regression analysis. The orthogonal nature of the components enables the breeder to predict the changes which would occur if a particular selection method is adopted. It is further suggested that the first few components may be used to transform the data to a reduced number of new scores which may then be used to form a selection index. (Although little reference has been made with respect to the value of a selection index in the preceding chapters, the author recognises its value (see Hazel and Lush, 1942; Young, 1961; Finney, 1962; Henderson, 1963; Tallis, 1962 and 1968; Harville, 1975; Van Vleck, 1976; and Sales and Hill, 1976 a & b).) As the components are orthogonal, a selection index based on them should provide a much more effective criterion for selection.

In Chapter VI statistical methods are developed which enable the breeder firstly, to test for the presence of a major gene and secondly, to estimate the parameters of interest. While unimodal distributions approaching normality occur in the presence of polygenic variation, the segregation of a major gene can be expected to distort the frequency distribution. If the effect of a major gene is large and its frequency sufficient this distortion may cause the frequency distribution to become distinctly bimodal. Simulation is used to investigate the relative power of eight tests of normality to recognise such distortion. It is concluded that the Shapiro-Francia W' Statistic provides the most versatile test, although the "Lower-Tailed" Kurtosis Statistic may be more powerful if the frequency of the major gene is near 0.5. However, all tests remain virtually powerless if the effect of the major gene is less than two standard deviations from the general population mean. Once non-normality is established, the method of moments can be applied to estimate the respective means and variances for the two component populations (i.e. associated with the presence and absence of the major gene). The frequency of the major gene is also estimated.

Application of the above procedure to the Roseworthy data fails to establish any consistent evidence of the segregation of a major gene for any of the four metric characters considered over the twelve years of the trial. (This may be partly due to the small number of rams used.)

While the fundamental theory of quantitative genetics has not changed markedly since the work of Fisher, Wright and Haldane, this cannot be said with respect to our general understanding of its application. For example, herd recording, artificial insemination and more recently computers have all greatly increased the effectiveness of breeding programs.

(In the preceding chapters the role of the computer has clearly been of central importance.) Also, investigations such as those of Robertson (1961 and 1970) on the theory of limits in artificial selection have done much to increase our knowledge.

Although quantitative genetics has much to offer the breeder, it should be recalled that the following six points have been recognised earlier (in the General Introduction to this thesis) as the main weaknesses in the current theory:

i) inability to predict limits to response;

ii) inability to predict changes in reproductive fitness following selection for a desired character;

iii) insufficient information about the nature of the underlying gene loci;

iv) inability to predict accurately the rate of response, especially following a plateau or a period of accelerated response;

v) inability to predict the response in correlated characters;

vi) reservations associated with the additive model.

While Ewens (1969) made the following comment on the theory of Population Genetics it can be applied equally well to summarize the use of additive genetic models in Quantitative Genetics:

> "...because the biological world is infinitely more complex than our mathematical models, it is impossible to expect that mathematics can play in the biological sciences the fundamental and ubiquitous role which it plays in the physical sciences."

Although the biometrical methods developed in Chapters III to VI do not overcome all the weaknesses listed, it is the author's belief that their use will assist the breeder particularly as they rely less heavily on the additive genetic models.

In conclusion, the present state of quantitative genetics can still be effectively summarized by the following statement of Falconer (1960);

> "The importance of this branch of genetics need hardly be stressed; most of the characters of economic value to plant and animal breeders are metric characters, and most of the changes concerned in micro-evolution are changes in metric characters. It is therefore in this branch that genetics has its most important application to practical problems and also its most direct bearing on evolutionary theory."

APPENDIX A

Generalized FORTRAN subroutines for applying Hotelling's T^2.

The following two generalized FORTRAN subroutines have been written to apply Hotelling's T^2 to large data sets. To use them a simple main program is required which reads the data for the two samples into two data arrays, which I will call Y1 and Y2, of size N1 x P and N2 x P respectively (where N1 and N2 are the number of observations in sample 1 and sample 2 respectively and P is the number of variates measured). Subroutine SWEEP is then called to calculate the mean vector and variance-covariance matrix for these two samples. Thus, from the (N1 x P) data matrix Y1 we obtain a (P x 1) mean vector X1, and a (P x P) variance-covariance matrix A1. Similarly, X2 and A2 are calculated from Y2. Subroutine HOTEL T2 is then called to calculate Hotelling's T^2. This subroutine prints out details of the test and also returns relevant information via formal parameters to the main program if required.

For the data matrix Y1, subroutine SWEEP is called as follows:

```
CALL SWEEP (N1,P,Y1,NY,IN,K,X1,NX1,NI,A1,NA1)
```

where ON INPUT

N1 = the number of observations in sample 1

P = the number of variates

Y1 = (N1 x P) data matrix

NY = the column size of Y1

K = number of subclasses, labelled 1 to K in array IN

IN = the indicator array, IN(1),...,IN(N) gives the subclass of the corresponding observations

NI = array for subclass numbers

NX1 = column size of X1

NA1 = column size of A1.

ON OUTPUT

X1 = (P x 1) vector of means

A1 = (P x P) cariance-covariance matrix.

A similar call is made for the second data matrix Y2,

i.e. `CALL SWEEP (N2,P,Y2,NY,IN,K,X2,NX2,NI,A2,NA2)`

subroutine HOTEL T2 is then called as follows:

`CALL HOTEL T2 (N1,X1,A1,N2,X2,A2,P,KP,TT,FT,F,SI)`

where ON INPUT

N1,X1,A1,N2,X2,A2,P are as explained above

KP is the column size of A1 and A2

ON OUTPUT

TT = the value of Hotelling's T^2

FT = the associated test value of F

F = a (3 x 1) vector

where F(1) is the 95% critical value

F(2) is the 99% critical value

F(3) is the 99.9% critical value

SI = the inverse of the pooled variance-covariance matrix.

Complete listings of subroutines SWEEP and HOTEL T2 are listed later with an example output (from the 1954 male progeny). This output should be self-explanatory but it should be pointed out that, in addition to the previously mentioned information, the linear discriminant function between the two samples is presented. It should also be noted that subroutine HOTEL T2 calls two subroutines from the International Mathematical and Statistical Library (IMSL) package (1975) available through the local system software. If this package is not available one would expect that similar subroutines could be substituted. The two subroutines are called LINV2F and MDFI. Subroutine LINV2F is called using

`CALL LINV2F(S,P,KP,SI,IDGT,WK,IER)`

This subroutine calculates the inverse of the P x P matrix S, where S is stored in full storage mode. On output the inverse is stored in the P x P matrix SI (S and SI must be mutually exclusive).

<u>ON INPUT</u>

S = input matrix of dimension P x P containing the matrix to be inverted

P = the size or order of S

KP = the column size of S

IDGT = set to 4 and denotes the accuracy of the elements of S

WK = work area of dimension greater than or equal to P**2 + 3*N.

<u>ON OUTPUT</u>

SI = output matrix of dimension P x P containing the inverse of S

IER = error parameter

Subroutine MDFI is called three times to calculate the 95%, 99% and 99.9% critical values of the F distribution,

i.e. `CALL MDFI(PROB,D1,D2,X,IER)`

MDFI inverts the F probability distribution function. That is, an X is found such that the probability of an F(D1,D2) distributed random variable being not greater than X is given by the probability PROB.

```
      SUBROUTINE HOTEL T2(N1,X1,A1,N2,X2,A2,P,KP,TT,FT,F,SI)
C****     COMPARING TWO MULTI-VARIATE-NORMAL SAMPLES
C         USING HOTELLING'S T-SQUARED STATISTIC
C              (SEE P.125 OF MORRISON)
C        INPUT PARAMETERS
C                   N1,N2 SAMPLE SIZES RESPECTIVELY
C                   X1,X2 MEAN VECTORS RESPECTIVELY
C                   A1,A2 MATRIX OF SUMS OF SQUARES AND PRODUCTS RESPECTIV
C                   P     NUMBER OF VARIATES
C                   KP    COLUMN SIZE(=ROW SIZE) OF A1 AND A2
      INTEGER P
      DIMENSION X1(1),X2(1),A1(KP,KP),A2(KP,KP),S(20,20),
     + SI(20,20),Y(20),Z(20),WK(500),F(5),A(20),D(20)
      M=1
      PRINT 10,N1,P,M,(X1(I),I=1,P)
   10 FORMAT(1H0,*NO. OF OBSERVATIONS =*,I10,20X,*NO. OF VARIATES =*,
     1  /,* MEAN VECTOR FOR SAMPLE*,I3,* =*,/,1X,15F9.3)
   15 FORMAT(1X,15E9.3)
      M=2
      PRINT 10,N2,P,M,(X2(I),I=1,P)
C****      CHECK NO. OF VARIATES
      IF(P.GT.20) STOP "NO. OF VARIATES EXCEEDS LIMIT(20)"
C****      CALCULATE POOLED COVARIANCE MATRIX S
      C=FLOAT(N1+N2-2)
      DO 20 I=1,P
      DO 20 J=1,P
   20 S(I,J)=(A1(I,J)+A2(I,J))/C
      PRINT 21
   21 FORMAT(2X,*POOLED COVARIANCE MATRIX S*)
      DO 22 I=1,P
   22 PRINT 15,(S(I,J),J=1,P)
C****      CALCULATE INVERSE OF S (STORE IN MATRIX SI) USING IMSL
C               SUBROUTINE LINV2F
      CALL LINV2F(S,P,KP,SI,4,WK,IER)
C****      CALCULATE T-SQUARED
      CT=FLOAT(N1*N2)/FLOAT(N1+N2)
      DO 40 I=1,P
   40 Y(I)=X1(I)-X2(I)
      DO 44 I=1,P
      Z(I)=0.
      DO 44 J=1,P
   44 Z(I)=Z(I)+SI(I,J)*Y(J)
      TT=0.
      DO 48 I=1,P
      D(I)=Z(I)
   48 TT=TT+Y(I)*Z(I)
      TT=TT*CT
C****      OBTAIN CRITICAL F VALUES FROM IMSL SUBROUTINE MDFI
      DF1=FLOAT(P)
      DF2=FLOAT(N1+N2-P-1)
      CALL MDFI(.95,DF1,DF2,F(1),IER)
      CALL MDFI(.99,DF1,DF2,F(2),IER)
      CALL MDFI(.999,DF1,DF2,F(3),IER)
      FT=TT*DF2/DF1/C
      PRINT 50,TT,FT,DF1,DF2,(F(I),I=1,3) .
   50 FORMAT(1H0,*T-SQUARED=*,F10.3,*    ASSOCIATED F=*,F10.3,*    WIT
     1DEGREES OF FREEDOM*,F5.0,F10.0,/,50X,*CRITICAL F VALUES ARE*
```

```
     23F10.2)
      DO 60 I=1,P
      DO 62 J=1,P
62    A(J)=0.
      A(I)=1
      DO 64 K=1,P
      Z(K)=0.
      DO 64 J=1,P
64    Z(K)=Z(K)+S(K,J)*A(J)
      ASA=0.
      AD=0.
      DO 68 K=1,P
      AD=AD+A(K)*Y(K)
68    ASA=ASA+A(K)*Z(K)
      ASA=SQRT(ASA*F(1)*C*DF1/DF2/CT)
      ALB=AD-ASA
      AUB=AD+ASA
      PRINT 70,ALB,I,I,AUB
70    FORMAT(10X,F13.4,5X,*.LE.(X1*,I2,* - X2*,I2,*).LE.*,F13.4)
60    CONTINUE
      PRINT 52,(Y(I),I=1,P)
52    FORMAT(1H0,* VECTOR OF MEAN DIFFERENCES IS*,/,(* *,10F13.4))
      PRINT 80,(D(I),I,I=1,P)
80    FORMAT(* DISCRIMINANT FUNCTION IS *,/,5X,*Y =*,8(F8.4,*  X*,I2,*
     +*))
      DX1=DX2=0.
      DO 84 I=1,P
      DX1=DX1 + X1(I)*D(I)
84    DX2=DX2 + X2(I)*D(I)
      PRINT 86,DX1,DX2
86    FORMAT(1H0,*MEAN DISCRIMINANT FOR SAMPLE 1 =*,F13.4,10X,*MEAN DI
     +RIMINANT FOR SAMPLE 2 =*,F13.4)
      RETURN
      END
```

```
      SUBROUTINE  SWEEP(N,R,Y,NY,IN,K,YB,NYB,NI,SS,NSS)
      DIMENSION  Y(NY,1),YB(NYB,1),NI(1),SS(NSS,1),IN(1)
      INTEGER  R
C     Y, DATA MATRIX ARRAY OF ROW DIMENSION NY.
C     R  = NO. OF VARIABLES.
C     N  = NO  OF OBSERVATIONS.
C     K  = NO. OF SUBCLASSES LABELLED 1 TO K IN ARRAY IN
C     IN = INDICATOR ARRAY,IN(1),   ...,IN(N) GIVES THE SUBCLASS OF
C                     CORRESPONDING OBSERVATIONS.
C     NI = ARRAY FOR SUBCLASS NUMBERS.
C     YB = ARRAY OF ROW DIMENSION NYB, FOR HOLDING SUBCLASS SUMS ,AND
C                     THEN MEANS.
C     SS = ARRAY OF ROW DIMENSION NSS, FOR ACCUMULATING SS AND SP MATRI
C           SET ARRAYS YB,NI AND SS INITIALLY TO ZERO
      DO  10  J=1,R
      DO  11  I=1,K
   11 YB(I,J)=0.
      DO  10  I=1,R
   10 SS(I,J)=0
      DO  12  I=1,K
   12 NI(I)=0
C           ACCUMULATE SUMS FOR CALCULATION OF MEANS
      DO  13  I=1,N
      J=IN(I)
      DO  14  L=1,R
   14 YB(J,L)=YB(J,L)+Y(I,L)
   13 NI(J)=NI(J)+1
C           CALCULATE MEANS
      DO  15  J=1,R
      DO  15  I=1,K
      IF(NI(I).EQ.0) GO TO 15
      YB(I,J)=YB(I,J)/NI(I)
   15 CONTINUE
C           SUBTRACT MEANS AND CALCULATE TOP HALF OF SS AND SP MATRIX
      DO  17  I=1,N
      J=IN(I)
      DO  18  L=1,R
   18 Y(I,L)=Y(I,L)-YB(J,L)
      DO  17  L=1,R
      DO  17  M=L,R
   17 SS(L,M)=SS(L,M) +Y(I,L)*Y(I,M)
C           CALCULATE BOTTOM HALF OF SS AND SP MATRIX
      DO  16  I=1,R
      DO  16  J=1,I
   16 SS(I,J)=SS(J,I)
      RETURN
      END
```

APPENDIX B

Analysis of parental age effects on mutation

published by

Mayo, O., Murdoch, J.L. and Hancock, T.W. (1976)

Mayo, O., Murdoch, J. L., & Hancock, T. W. (1976). On the estimation of parental age effects on mutation. *Annals of Human Genetics*, *39*(4), 427-431.

NOTE:
This publication is included in the print copy
of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:
https://doi.org/10.1111/j.1469-1809.1976.tb00147.x

APPENDIX C

Remark on algorithm 434[G2]

Exact probabilities for R x C contingency tables

Hancock (1975)

Hancock, T.W. (1975). Remark on algorithm 434 [G2] exact probabilities for R×C contingency tables. *Communications of the ACM, 18*(2), [n.p.].

NOTE:

This publication is included in the print copy
of the thesis held in the University of Adelaide Library.

Algorithm

```
      SUBROUTINE RXCPRB(MATRIX, NRD, NCD, NR, NC,
     * PT, PS, PC)
C *        THIS SUBROUTINE COMPUTES EXACT
C *        PROBABILITIES FOR R X C CONTINGENCY TABLES
C *INPUT VIA FORMAL PARAMETERS
C *   NRD = THE ROW DIMENSION
C *   NCD = THE COLUMN DIMENSION
C      NR = THE NUMBER OF ROWS IN MATRIX (R=NR-1).
C      NC = THE NUMBER OF COLUMNS IN MATRIX (C=NC-1).
C      MATRIX = SPECIFICATION OF THE CONTINGENCY
C           TABLE. THIS MATRIX IS PARTITIONED AS
C           FOLLOWS
C *        X(1,1),X(1,2),.............,X(1,C)     X(1,NC)
C *           .       .   ,............,    .           .
C *           .       .   ,............,    .           .
C *        X(R,1),X(R,2),.............,X(R,C)     X(R,NC)
C *        X(NR,1),X(NR,2),...........,X(NR,C)    X(NR,NC)
C *         WHERE X(I,J) ARE THE OBSERVED CELL
C *         FREQUENCIES, X(I,NC) ARE THE ROW TOTALS,
C *         X(NR,J) ARE THE COLUMN TOTALS, AND X(NR,NC)
C *         IS THE TOTAL SAMPLE SIZE.
C           NOTE THAT THE ORIGINAL CELL FREQUENCIES ARE
C           DESTROYED BY THIS SUBROUTINE.
C OUTPUT ARGUMENTS.
C      PT = THE PROBABILITY OF OBTAINING THE GIVEN
C      TABLE.
C      PS = THE PROBABILITY OF OBTAINING A TABLE AS
C           PROBABLE AS, OR LESS PROBABLE THAN, THE
C           GIVEN TABLE.
C *    PC = THE PROBABILITY OF OBTAINING ALL OF THE
C           TABLES POSSIBLE WITHIN THE CONSTRAINTS OF
C           THE MARGINAL TOTALS. (THIS SHOULD BE 1.0.
C           DEVIATIONS FROM 1.0 REFLECT THE ACCURACY OF
C           THE COMPUTATION.)
C EXTERNALS.
C *    INIT(MATRIX,NRD,NCD,NR,NC) = SUBROUTINE WHICH
C *              RETURNS THE *NEXT* MATRIX TO SATISFY
C *              THE MARGINALS.
C      FACLOG(N) = FUNCTION TO RETURN THE FLOATING
C           POINT VALUE OF LOG BASE 10 OF N FACTORIAL.
      DIMENSION MATRIX(NRD,NCD)
      INTEGER R, C
      R = NR - 1
      C = NC - 1
C COMPUTE LOG OF CONSTANT NUMERATOR
      QXLOG = -FACLOG(MATRIX(NR,NC))
      DO 10 I=1,R
        QXLOG = QXLOG + FACLOG(MATRIX(I,NC))
   10 CONTINUE
      DO 20 J=1,C
        QXLOG = QXLOG + FACLOG(MATRIX(NR,J))
   20 CONTINUE
C COMPUTE PROBABILITY OF GIVEN TABLE
      RXLOG = 0.0
      DO 40 I=1,R
        DO 30 J=1,C
          RXLOG = RXLOG + FACLOG(MATRIX(I,J))
   30   CONTINUE
   40 CONTINUE
      PT = 10.0**(QXLOG-RXLOG)
      PS = 0.0
      PC = 0.0
C * 1. ALL CELL VALUES INITIALLY SET TO ZERO
      DO 60 I=1,R
        DO 50 J=1,C
          MATRIX(I,J) = 0
   50   CONTINUE
   60 CONTINUE
C * 2. EACH CYCLE STARTS HERE
   70 KEY = 1
      MATRIX(2,2) = -1
C * 3. GENERATING SET OF FREQUENCIES PROGRESSIVELY IN
C *    LOWER RIGHT (R-1) * (C-1) CELLS.
      DO 160 I=2,R
        DO 150 J=2,C
          MATRIX(I,J) = MATRIX(I,J) + 1
C * 4. CHECKING SUMMATIONS .LE. RESPECTIVE MARGINALS
C *      I.E. (SUM OF ELTS. J TO C IN ROW I) .LE.
C *      MATRIX(I,NC) AND  (SUM OF ELTS. I TO R IN COL.
C *      J).LE. MATRIX(NR,J)
          ISUM = 0
          JSUM = 0
          DO 80 M=J,C
            ISUM = ISUM + MATRIX(I,M)
   80     CONTINUE
          IF (ISUM.GT.MATRIX(I,NC)) GO TO 130
          DO 90 K=I,R
            JSUM = JSUM + MATRIX(K,J)
   90     CONTINUE
          IF (JSUM.GT.MATRIX(NR,J)) GO TO 130
C * 5. JUMP TO STATEMENT 170 WHERE ALL CELLS PRIOR TO
C *    MATRIX(I,J) ARE SET TO ZERO.
          IF (KEY.EQ.2) GO TO 170
          IP = I
          JP = J
C * 6. CALL SUBROUTINE INIT TO FIND THE NEXT BALANCED
C *    MATRIX
          CALL INIT(MATRIX, NRD, NCD, NR, NC)
C COMPUTE LOG OF THE DENOMINATOR
          RXLOG = 0.0
          DO 110 K=1,R
            DO 100 M=1,C
```

```
            RXLOG = RXLOG + FACLOG(MATRIX(K,M))
 100        CONTINUE
 110      CONTINUE
C * COMPUTE PX. ADD TO PC AND ALSO PS IF PX .LE. PT
C     (ALLOW FOR ROUND-OFF ERROR)
          PX = 10.0**(QXLOG-RXLOG)
          PC = PC + PX
          IF ((PT/PX).GT.0.99999) PS = PS + PX
C * 7. IF POSSIBLE A SEQUENCE OF MATRICES AND
C *    ASSOCIATED PROBABILITIES (PX,PC AND PS) ARE
C *    GENERATED BY MANIPULATING CELLS (1,1), (1,2),
C *    (2,1) AND (2,2)  (SIMILARLY ALLOWING
C *    FOR ROUND-OFF ERROR)
 120      IF (MATRIX(1,2).LT.1 .OR.
     *     MATRIX(2,1).LT.1) GO TO 140
          MATRIX(1,1) = MATRIX(1,1) + 1
          MATRIX(2,2) = MATRIX(2,2) + 1
          PX = PX*FLOAT(MATRIX(1,2))*FLOAT(MATRIX(2,1))
     *         /FLOAT(MATRIX(1,1))/FLOAT(MATRIX(2,2))
          PC = PC + PX
          IF ((PT/PX).GT.0.99999) PS = PS + PX
          MATRIX(1,2) = MATRIX(1,2) - 1
          MATRIX(2,1) = MATRIX(2,1) - 1
          GO TO 120
 130      IP = I
          JP = J
C * 8. KEY SET TO 2 AS CYCLE COMPLETED
 140      KEY = 2
 150    CONTINUE
 160 CONTINUE
     RETURN
C * ALL CELLS OF MATRIX PRIOR TO THE (I,J)TH. ARE
C * SET TO ZERO.
 170 DO 180 M=2,JP
       MATRIX(IP,M) = 0
 180 CONTINUE
     IP = IP - 1
     DO 200 K=1,IP
       DO 190 M=2,C
         MATRIX(K,M) = 0
 190   CONTINUE
 200 CONTINUE
     GO TO 70
     END


     SUBROUTINE INIT(MATRIX, NRD, NCD, NR, NC)
C * THIS SUBROUTINE RETURNS THE *NEXT* MATRIX TO
C *  SATISFY (1) THE MARGINALS AND (2) THE SEQUENCE
C *  OF GENERATION DEFINED IN SUBROUTINE RXCPRB.
     DIMENSION MATRIX(NRD,NCD), MROW(50), MCOL(50)
     INTEGER R, C
     R = NR - 1
     C = NC - 1
C * EQUIVALENCE MROW AND MCOL TO ROW AND COLUMN
C * MARGINALS RESPECITVELY.
     DO 10 K=1,R
       MATRIX(K,1) = 0
       MROW(K) = MATRIX(K,NC)
  10 CONTINUE
     DO 20 M=1,C
       MCOL(M) = MATRIX(NR,M)
  20 CONTINUE
C * FOR EACH ROW, SUBTRACT ELEMENTS 2 TO C FROM MROW
     DO 40 K=2,R
       DO 30 M=2,C
         MROW(K) = MROW(K) - MATRIX(K,M)
  30   CONTINUE
  40 CONTINUE
C * FOR EACH COLUMN, SUBTRACT ELEMENTS 2 TO R FROM
C * MCOL
     DO 60 M=2,C
       DO 50 K=2,R
         MCOL(M) = MCOL(M) - MATRIX(K,M)
  50   CONTINUE
  60 CONTINUE
C * FORMING *NEXT BALANCED* ARRAY
     DO 90 I=1,R
       IR = NR - I
       DO 80 J=1,C
         MIN = MIN0(MROW(IR),MCOL(J))
         IF (MIN.EQ.0) GO TO 70
         MATRIX(IR,J) = MATRIX(IR,J) + MIN
         MROW(IR) = MROW(IR) - MIN
         MCOL(J) = MCOL(J) - MIN
  70     IF (MROW(IR).EQ.0) GO TO 90
  80   CONTINUE
  90 CONTINUE
     RETURN
     END


     FUNCTION FACLOG(N)
C INPUT ARGUMENT.
C     N = AN INTEGER GREATER THAN OR EQUAL TO ZERO.
C FUNCTION RESULT.
C     FACLOG = THE LOG TO THE BASE 10 OF N FACTORIAL.
      DIMENSION TABLE(101)
      DATA TPILOG /0.3990899342/
      DATA ELOG /0.4342944819/
      DATA IFLAG /0/
C USE STIRLINGS APPROXIMATION IF N GT 100
```

```
      IF (N.GT.100) GO TO 20
C LOOK UP ANSWER IF TABLE WAS GENERATED
      IF (IFLAG.EQ.0) GO TO 30
   10 FACLOG = TABLE(N+1)
      RETURN
C HERE FOR STIRLINGS APPROXIMATION
   20 X = FLOAT(N)
      FACLOG = (X+0.5)*ALOG10(X) - X*ELOG + TPILOG +
     * ELOG/(12.0*X) - ELOG/(360.0*X*X*X)
      RETURN
C HERE TO GENERATE LOG FACTORIAL TABLE
   30 TABLE(1) = 0.0
      DO 40 I=2,101
        X = FLOAT(I-1)
        TABLE(I) = TABLE(I-1) + ALOG10(X)
   40 CONTINUE
      IFLAG = 1
      GO TO 10
      END
```

APPENDIX D

Exemplary listing of SPSS program for Forward Selection

```
ATTACH(DATA,HDATA,ID=W,MR=1)
UPDATE(P=DATA,Q,K=TAPE,D)
ATTACH(SPSS,SPSS)
RFL(75000)
SPSS(D=TAPE)
---------- END OF SECTION ----------
*D PB54.2
*C PB54
---------- END OF SECTION ----------
RUN NAME          FORWARD SELECTION FOR ROSEWORTHY SHEEP DATA
VARIABLE LIST       FLK,SEX,BTH,BWT,YLD,CFW,STL,CRP,FBD,PFK,SFN
INPUT MEDIUM        TAPE
N OF CASES          ESTIMATED 500
INPUT FORMAT        FIXED(F1.0,8X,F1.0,F1.0,20X,F5.1,9X,F4.1,
                    F4.1,F3.1,F2.0,F4.1,4X,F3.0,F5.0)
COMPUTE             BWT=BWT*0.45359
COMPUTE             CFW=CFW*0.45359
COMPUTE             STL=STL*2.54
READ INPUT DATA
COMMENT
*SELECT IF        (SEX EQ 0)
COMMENT           FLOCK BY BIRTH TYPE-MALES ONLY
CROSSTABS         FLK BY BTH
STATISTICS        ALL
COMMENT
*SELECT IF        (SEX EQ 1)
COMMENT           FLOCK BY BIRTH TYPE-FEMALES ONLY
CROSSTABS         FLK BY BTH
STATISTICS        ALL
COMMENT
*SELECT IF          (SEX EQ 0 AND FLK EQ 1 AND BTH EQ 0)
COMMENT             SEX=0=MALE   FLOCK=1=INDEX   BIRTH TYPE=0=SINGLE
REGRESSION        METHOD=FORWARD/
                    VARIABLES=CFW,BWT,YLD,STL,CRP,FBD,PFN,SFN/
                    REGRESSION=CFW,(80,4.0,.500) WITH BWT TO SFN(5)
COMMENT
*SELECT IF           (SEX EQ 0 AND FLK EQ 2 AND BTH EQ 0)
COMMENT             SEX=0=MALE   FLOCK=2=VISUAL   BIRTH TYPE=0=SINGLE
REGRESSION        METHOD=FORWARD/
                    VARIABLES=CFW,BWT,YLD,STL,CRP,FBD,PFN,SFN/
                    REGRESSION=CFW,(80,4.0,.500) WITH BWT TO SFN(5)
COMMENT
*SELECT IF          (SEX EQ 1 AND FLK EQ 1 AND BTH EQ 0)
COMMENT             SEX=1=FEMALE   FLOCK=1=INDEX   BIRTH TYPE=0=SINGLE
REGRESSION        METHOD=FORWARD/
                    VARIABLES=CFW,BWT,YLD,STL,CRP,FBD,PFN,SFN/
                    REGRESSION=CFW,(80,4.0,.500) WITH BWT TO SFN(5)
COMMENT
*SELECT IF          (SEX EQ 1 AND FLK EQ 2 AND BTH EQ 0)
COMMENT             SEX=1=FEMALE   FLOCK=2=VISUAL   BIRTH TYPE=0=SINGLE
REGRESSION        METHOD=FORWARD/
                    VARIABLES=CFW,BWT,YLD,STL,CRP,FBD,PFN,SFN/
                    REGRESSION=CFW,(80,4.0,.500) WITH BWT TO SFN(5)
FINISH
---------- END OF INFORMATION ----------
```

APPENDIX E

Generalized FORTRAN program for testing for major genes.

(The subroutines DECOM, SIGMA and COMPUT listed below were supplied by personal communication from Hawkins (1975). Several changes,however, were required to Hawkins's listing to enable it to operate successfully. These were justified algebraically and produced correct answers for the example data set supplied.)

```
      PROGRAM BIMOD(INPUT,OUTPUT,TAPE6=OUTPUT)
C****      THIS PROGRAM TESTS FOR THE PRESENCE OF A MAJOR GENE
C****      (I.E. NON-NORMALITY IN THE DATA)
      DIMENSION TITLE(10),X(500),C(10),IRAY(6)
      COMMON B(500)
      DATA IRAY/6 * (-0)/
C****      SET PRINT LIMIT TO ZERO
      IRAY(4)=0
      CALL SYSTEMC(115,IRAY)
C****      READ TITLE CARD
      READ 10,(TITLE(I),I=1,8)
   10 FORMAT(8A10)
      PRINT 11,(TITLE(I),I=1,8)
   11 FORMAT(1X,8A10)
C****      READ SAMPLE SIZE(N)AND CRITICAL VALUE FOR
C****      SHAPIRO-FRANCIA W"TEST (CW)
      READ*,N,CW
      READ *,A1B,DA1,AUL
C****      READ N OBSERVATIONS
      READ *,(X(I),I=1,N)
      PRINT 20,N,CW,A18,DA1,AUL
   20 FORMAT(10X,I10,4F10.3)
      PRINT 22,(X(I),I=1,N)
   22 FORMAT(10X,10F12.2)
      IF(N.GE.500) STOP "DATA LENGTH GE 500"
C****      CALCULATE EXPECTED VALUES OF NORMAL ORDER STATISTICS FOR
C****      N OBSERVATIONS
      CALL EVNOS(N)
      PRINT 30,(B(I),I=1,N)
   30 FORMAT(5X,10F12.4)
      Q=0
      DO 35 I=1,N
   35 Q=Q+B(I)*B(I)
      Q=SQRT(Q)
      DO 36 I=1,N
   36 B(I)=B(I)/Q
      PRINT 30,(B(I),I=1,N)
C****      CALL SUBROUTINE WD TO CALCULATE SHAPIRO-FRANCIA W"
      CALL WD(N,X,W,ISIG)
      IF(W.LT.CW)GO TO 42
      PRINT 40,W,CW
   40 FORMAT(5X,*AS OBSERVED VALUE OF W"  =*,F12.5,*  WHICH IS GT CRI
     +AL VALUE (*,F7.3,*  ) THERE IS NO EVIDENCE OFNON-NORMALITY*)
      STOP
   42 PRINT 44,W,CW
   44 FORMAT(5X,*AS OBSERVED VALUE OF W"  =*,F12.5,*  WHICH IS LT CRI
     +AL VALUE (*,F7.3,* ) NON-NORMALITY IS INDICATED *,/,40X,*METHOD
     1 MOMENTS WILL BE APPLIED TO LOCATE POSSIBLE MAJOR GENE*)
      DO 48 I=1,5
   48 C(I)=0.
      DO 50 I=1,N
      C(1)=C(1)+X(I)
      S=X(I)*X(I)
      C(2)=C(2)+S
      S=S*X(I)
      C(3)=C(3)+S
      S=S*X(I)
```

```
      C(4)=C(4)+S
      S=S*X(I)
   50 C(5)=C(5)+S
      XN=FLOAT(N)
      DO 55 I=1,5
   55 C(I)=C(I)/XN
      PRINT 56,(C(I),I=1,5)
   56 FORMAT(10X,5G15.5)
      CALL DECOM(TITLE,C,A1B,DA1,AUL,N)
      STOP
      END
```

```
      SUBROUTINE WD(N,Y,W,ISIG)
      DIMENSION Y(1)
      COMMON B(500)
      XN=FLOAT(N)
      ISIG=0
      S=0
      SS=0. $ SBY=0.
      CALL RANK(Y,N)
      DO 30 I=1,N
   30 S=S+Y(I)
      S=S/XN
      DO 40 I=1,N
      SBY=SBY+B(I)*Y(I)
      D=Y(I)-S
   40 SS=D*D+SS
      W=SBY*SBY/SS
      IF(W.LT.0.980) ISIG=1
      RETURN
      END
```

```
      SUBROUTINE EVNOS(N)
C****    THIS SUBROUTINE CALCULATES THE EXPECTED VALUES OF NORMAL
C****       ORDER STATISTICS FOR A SAMPLE OF SIZE N.
C****    REFERENCE - EXPECTED VALUES OF NORMAL ORDER STATISTICS
C****    BY  H.LEON HARTER - BIOMETRIKA(1961),48,P.151
      DIMENSION DX(500),PXL10(500)
      COMMON EV(500)
      DATA IFLAG/0/
      COEFL10=SQRT(2.*3.1415926536)
      COEFL10=ALOG10(COEFL10)
      DO 20 I=1,N
   20 EV(I)=0.
C****    CHECK DISTRIBUTION FUNCTION HAS BEEN GENERATED U
      IF(IFLAG.EQ.1) GO TO 36
      DO 30 I=1,160
      X=.05*FLOAT(I)
      CALL MDNOR(X,D)
      DX(I)=D
   30 PXL10(I)=-.5*X*X*.4342944819 - COEFL10 + ALOG10(X)
      IFLAG=1
C****    NUMERICAL INTEGRATION BEGINS HERE
   36 K=N/2
      DO 40 J=1,K
      AA=FACLOG(N)
      AB=FACLOG(N-J)
      AC=FACLOG(J-1)
      S=0.
      DO 50 I=1,152
      AD=ALOG10(1.-DX(I))
      AE=ALOG10(DX(I))
      ADN=AD*FLOAT(N-J)
      AEN=AE*FLOAT(J-1)
      AD=AD*FLOAT(J-1)
      AE=AE*FLOAT(N-J)
      AF=AA-AB-AC+PXL10(I)
      AP=AF+AD+AE
      AN=AF+ADN+AEN
   50 S=S + 10.**AP*.05 - 10.**AN*.05
      EV(J)=S
   40 EV(N+1-J)=-S
      RETURN
      END
```

```
      FUNCTION FACLOG(N)
C   INPUT ARGUMENT.
C     N = AN INTEGER GREATER THAN OR EQUAL TO ZERO.
C   FUNCTION RESULT.
C     FACLOG = THE LOG TO THE BASE 10 OF N FACTORIAL.
      DIMENSION  TABLE(101)
      DATA  TPILOG/0.39908 99342/
      DATA  ELOG  /0.43429 44819/
      DATA  IFLAG/0/
C   USE STIRLINGS APPROXIMATION IF N GT 100
      IF(N.GT.100)  GO  TO  50
C   LOOK UP ANSWER IF TABLE WAS GENERATED
      IF(IFLAG.EQ.0)  GO  TO  100
   10 FACLOG=TABLE(N+1)
      RETURN
C   HERE FOR STIRLINGS APPROXIMATION
   50 X=FLOAT(N)
      FACLOG=(X+0.5) * ALOG10(X) - X*ELOG + TPILOG
     1    + ELOG/(12.0*X) - ELOG/(360.0*X*X*X)
      RETURN
C   HERE TO GENERATE LOG FACTORIAL TABLE
  100 TABLE(1)=0.0
      DO  120  I=2,101
        X=FLOAT(I-1)
  120   TABLE(I)=TABLE(I-1) + ALOG10(X)
      IFLAG=1
      GO  TO  10
      END
```

```
      SUBROUTINE RANK(V,N)
      DIMENSION V(1)
      M=N
   10 M=M/2
      IF(M)70,70,20
   20 K=N-M
      J=1
   30 I=J
   40 IM=I+M
      L=IDSWCH(I,IM,V,N)
      IF(L.EQ.0)60,50
   50 I=I-M
      IF(I.LT.1)60,40
   60 J=J+1
      IF(J.GT.K)10,30
   70 RETURN
      END
```

```
      FUNCTION IDSWCH(I1,I2,V,N)
      DIMENSION V(1)
      I=0
      X=V(I1)
      Y=V(I2)
      IF(X.LT.Y) GO TO 21
      V(I1)=V(I2)
      V(I2)=X
      I=1
   21 IDSWCH=I
      RETURN
      END
```

```
      SUBROUTINE DECOM(AID,C,A1B,DA1,AUL,NUM)
C$    DEBUG
C$    ARRAYS
      DIMENSION AID(10),KT(120),U4(120),U5(120),D45(120),AM0(120)
     +  ,A(120),C(6),B(5),D(5)
      COMMON/BLK20/E1,E2,E3,E4,E5
       G(A1,U1,S1,A2,U2,S2)=
     1  A1*(5.*U1*S1*  (3.*S1  +2.*U1*U1)+U1**5)+
     1A2*(5.*U2*S2*  (3.*S2  +2.*U2*U2)+U2**5)
       H(A1,U1,S1,A2,U2,S2)=
     1  A1*(3.*S1*  (2.*U1*U1+S1  )+U1**4)+A2*(3.*S2*  (2.*U2*U2+S2
     1  )+U2**4)
      E1=C(1)  $E2=C(2)
      E3=C(3)  $  E4=C(4)
      E5=C(5)
CC READ IN IDENTIFICATION, DATA, AND SPECIFICATIONS
      WRITE(6,108)
      WRITE(6,111)(AID(J),J=1,8)
  111 FORMAT(1X,8A10)
C A1B IS THE INITIAL STARTING POSITION
C DA1 IS THE INTERVAL SIZE
C AUL IS THE END OF THE SEARCH RANGE
C NUM IS THE NUMBER OF OBSERVATIONS USED IN FINDING THE MOMENTS
    2 WRITE(6,207)   A1B,DA1,AUL  ,NUM
      WRITE(6,601)
      WRITE(6,602)
C  CALCULATE AND PRINT MASTER MOMENT DATA
C
      VARX=E2-E1*E1
      SDV=SQRT(VARX)
      EX3=E3-3.*E2*E1+2.*E1*E1*E1
      SKEW=EX3/(VARX*SDV)
      D(1)=E1
      D(2)=VARX
      D(3)=EX3
      B(1)=0.0
      B(2)=1.0
      B(3)=SKEW
      D(4)=E4-4.0*E3*E1+6.0*E2*E1*E1-3.0*E1**4
      B(4)=D(4)/(VARX*VARX)
      D(5)=E5-5.0*E4*E1+10.0*E3*E1*E1-10.0*E2*E1**3+4.0*E1**5
      B(5)=D(5)/(VARX*VARX*SDV)
      DO 710 I=1,5
      WRITE(6,603) I,C(I),B(I),D(I)
  710 CONTINUE
  601 FORMAT(*0 K  ORIG.MOM.  STAND.MOM.  MEAN MOM.*)
  602 FORMAT(*  ==================================  */)
  603 FORMAT(3X,I1,G14.7,2X,G14.7,G14.7)
      U=E1
      ENC=((AUL-A1B)/DA1)+0.1
      NC=ENC
      IF(NC.GE.120) STOP "INTERVAL SIZE (DA1) TOO SMALL"
      KAT=0
      KN=1
      KM=1
      KOOK=0
      DO 3 K=1,120
```

```
      KT(K)=0
      A(K)=0.
      U4(K)=0.
      U5(K)=0.
      D45(K)=0.
    3 CONTINUE
C
CCC SATISFACTION OF E4
      A1=A1B
      SIG=SQRT(E2-E1*E1)
      DU1=0.04*SIG
      ISTART=1
      ITEX=0
      SIG=SDV
      U1=U-3.*SDV
      ANSWER=U1
    5 A2=1.-A1
      ITEM=0
      J=0
      IF(AUL-A1)80,80,10
CC PRELIMINARY REFINEMENT
CC
   10 U2=(E1-A1*U1)/A2
      CALL SIGMA(A1,A2,U1,U2,  SIG12,SIG22)
C-C  CHECK FOR ANY NEGATIVE OR ZERO VARIANCES
      IF(U-U1)70,11,11
   11 IF(SIG12)14,12,12
   12 IF(SIG22)14,20,20
   14 U1=U1+DU1
      GO TO 10
   20 E4C=H(A1,U1,SIG12,A2,U2,SIG22)
      D4=E4C-E4
      IF(J)25,25,30
   25 J=J+1
      D4P=D4
      GO TO 14
   30 IF(D4P*D4) 35,35,25
CC  SECONDARY REFINEMENT
   35 D41P=D4P
      ITEM=1
      ITEX=ITEX+1
      AMO(ITEX)=A1
      IF(ITEX-1)37,36,37
   36 ISTART=KN
      IF(D4P*D4) 37,50,37
   37 D4P=D4
      U11=U1-DU1
      DU11=0.02*DU1
   40 U21=(E1-A1*U11)/A2
      CALL SIGMA(A1,A2,U11,U21,  SIG12,SIG22)
      E4C1=H(A1,U11,SIG12,A2,U21,SIG22)
      D41=E4C1-E4
      IF(D41*D41P)55,51,45
   45 D41P=D41
      U11=U11+DU11
      GO TO 40
CC  SECONDARY REFINEMENT COMPLETE - INTERPOLATE AND PRINT
```

```
   50 ANSWER=U1
      ITEM=1
      KT(KN)=1
      U4(KN)=ANSWER
      A(KN)=A1
      GO TO 60
   51 ANSWER=U11
      KT(KN)=1
      U4(KN)=ANSWER
      A(KN)=A1
      GO TO 60
   55 ANSWER=(U11-DU11)+((ABS(D41P))/(ABS(D41)+ABS(D41P)))*DU11
      A(KN)=A1
      KT(KN)=1
      U4(KN)=ANSWER
CC  PRINT AND CONTINUE THE SEARCH OR RECYCLE
CC
   60 CONTINUE
      IF(E1-ANSWER)70,70,65
   65 J=0
      GO TO 14
   70 IF(AUL-A1)80,80,75
   75 IF(ITEM)77,77,76
   76 CONTINUE
      GO TO 78
   77 CONTINUE
   78 KN=KN+1
      U1=ANSWER
      A1=A1+DA1
      GO TO 5
   80 CONTINUE
CCC  SATISFACTION OF E5
C
      IF(ITEX-1)85,85,90
   85 WRITE(6,120) ITEX
      GO TO 502
   90 A1=AMO(1)
      KM=ISTART
      WRITE(6,109)
      ITEM=0
      U1=U-3.*SIG
      ANSWER=U1
  305 A2=1.-A1
      J=0
CC  PRELIMINARY REFINEMENT
      IF(AUL-A1)380,380,310
  310 U2=(E1-A1*U1)/A2
      CALL SIGMA(A1,A2,U1,U2,  SIG12,SIG22)
C-C  CHECK FOR ANY NEGATIVE OR ZERO VARIANCES
      IF(U-U1)370,311,311
  311 IF(SIG12)314,312,312
  312 IF(SIG22)314,320,320
  314 U1=U1+DU1
      GO TO 310
  320 E5C=G(A1,U1,SIG12,A2,U2,SIG22)
      D5=E5C-E5
      IF(J)325,325,330
```

```
  325 J=J+1
      D5P=D5
      GO TO 314
  330 IF(D5P*D5)335,350,325
CC  SECONDARY REFINEMENT
  335 D51P=D5P
      ITEM=1
      D5P=D5
      U11=U1-DU1
      DU11=0.02*DU1
  340 U21=(E1-A1*U11)/A2
      CALL SIGMA(A1,A2,U11,U21, SIG12,SIG22)
      E5C1=G(A1,U11,SIG12,A2,U21,SIG22)
      D51=E5C1-E5
      IF(D51*D51P)355,351,345
  345 D51P=D51
      U11=U11+DU11
      GO TO 340
CC  SECONDARY REFINEMENT COMPLETE - INTERPOLATE AND PRINT
  350 ANSWER=U1
      KT(KM)=KT(KM)+1
      U5(KM)=ANSWER
      GO TO 360
  351 ANSWER=U11
      KT(KM)=KT(KM)+1
      U5(KM)=ANSWER
      GO TO 360
  355 ANSWER=(U11-DU11)+((ABS(D51P))/(ABS(D51)+ABS(D51P)))*DU11
      KT(KM)=KT(KM)+1
      U5(KM)=ANSWER
CC  PRINT AND CONTINUE THE SEARCH OR RECYCLE
  360 CONTINUE
  365 J=0
      GO TO 314
  370 IF(AUL-A1)380,380,375
  375 IF(ITEM)377,377,376
  376 CONTINUE
      GO TO 378
  377 CONTINUE
  378 KM=KM+1
      A1=A1+DA1
      U1=ANSWER
      GO TO 305
  380 CONTINUE
CC  CHECH FOR PAIRED DATA IN U4 AND U5 AND CALCULATE DIFFERENCES
C
      DO 410 KK=1,NC
      IF(KT(KK)-2)410,405,410
  405 D45(KK)=U4(KK)-U5(KK)
  410 CONTINUE
CC PRINT ANSWER TABLE
      DO 420 KK=1,NC
      IF(KT(KK)-2)420,414,420
  414 KOOK=KOOK+1
      IF(KOOK-1)417,416,417
  416 WRITE(6,411)
      WRITE(6,412)
```

```
  417 WRITE(6,415)  A(KK),U4(KK),U5(KK),D45(KK)
  420 CONTINUE
CC LOCATE ESTIMATE OF A1
      KK=0
  421 KLIP=1
  422 KK=KK+1
      IF(KK-NC)423,423,500
  423 IF(KT(KK)-2)422,424,422
  424 IF(KLIP-1)426,425,426
  425 KLIP=KLIP+1
      GO TO 422
  426 KKM=KK-1
      IF(KT(KKM)-2)422,429,422
  429 KLIP=KLIP+1
      IF(D45(KKM)*D45(KK))427,427,422
  427 KGO=KK
CC INTERPOLATE VALUE OF A1 AND CALCULATE OTHER PARAMETERS
  440 KLS=KGO-1
      KAT=1
      ZZ=(ABS(D45(KLS)))/(ABS(D45(KLS))+ABS(D45(KGO)))
      A1F=A(KLS)+DA1*ZZ
      A2F=1.-A1F
      U1F=U4(KLS)+(U4(KGO)-U4(KLS))*ZZ
      U2F=(E1-U1F*A1F)/A2F
      CALL SIGMA(A1F,A2F,U1F,U2F, SIG12,SIG22)
      SIG1=SQRT(SIG12)
      SIG2=SQRT(SIG22)
      N1=A1F*NUM+.5
      N2=A2F*NUM+.5
      DFT=N1+N2-2
      NFT=DFT
      DT1=SQRT(((N1-1)*SIG12+(N2-1)*SIG22)/DFT)
      DT2=SQRT(1./FLOAT(N1)+1./FLOAT(N2))
      T=(U1F-U2F)/(DT1*DT2)
      N1=N1-1
      N2=N2-1
      F=SIG12/SIG22
      IF(F.GE.1.) GO TO 700
      F=1./F
      NTEMP=N1
      N1=N2
      N2=NTEMP
  700 WRITE(6,111) (AID(J),J=1,8)
      WRITE(6,113)
      WRITE(6,701)
      WRITE(6,114)
      WRITE(6,702)
      WRITE(6,115)  A1F,U1F,SIG1
      WRITE(6,703)   F,N1,N2
      WRITE(6,116)  A2F,U2F,SIG2
      WRITE(6,704)   T,NFT
      CALL COMPUT(A1F,U1F,U2F,SIG12,SIG22)
      GO TO 422
  500 IF(KAT-1)502,504,502
  502 WRITE(6,111)   (AID(J),J=1,8)
      WRITE(6,428)
  504 CONTINUE
```

```
100 FORMAT(1H ,F12.5)
101 FORMAT(1H ,2F12.5)
108 FORMAT(1H1,//)
109 FORMAT(1H ,/)
113 FORMAT(1H0,35H  INDEX  A       MEAN          ST DEV)
114 FORMAT(1H ,35H================================   )
115 FORMAT(1H ,3X,1H1,2X,F7.5,F11.5,F12.5)
116 FORMAT(1H ,3X,1H2,2X,F7.5,F11.5,F12.5    )
120 FORMAT(1H ,4X,43HINSUFFICIENT SOLUTIONS FOR A1 + U1, NUMB = ,I2,
205 FORMAT(F26.4)
206 FORMAT(3F7.5,I5)
207 FORMAT(* *,3F7.4,3X,I5)
251 FORMAT(*+*, 40X,*MEAN *, G15.7)
252 FORMAT(*+*, 40X,*STDV *, G15.7)
253 FORMAT(*+*, 40X,*SKEW *, G15.7)
411 FORMAT(* *,*    A1           U1(4)            U1(5)             D45*
412 FORMAT(* *,*    ================================================*)
415 FORMAT(* *,3X,F6.4,3G15.7 )
428 FORMAT(1H ,37H NO VALID SOLUTIONS FOR RANGE STUDIED,//)
701 FORMAT(*+*,40X,*   TESTS  *,10X,*DF* )
702 FORMAT(*+*,40X,*=============*,7X,*========* )
703 FORMAT(*+*,35X, *  F  *,G15.7,6X,2(I3,2X))
704 FORMAT(*+*,35X,*  T  *,G15.7,6X,I4)
600 CONTINUE
    END
```

```
      SUBROUTINE SIGMA(A1,A2,F1,F2, SIG12,SIG22)
      COMMON/BLK20/E1,E2,E3,E4,E5
      B1=E2 - (A1*F1*F1 + A2*F2*F2)
      B2=(E3-A1*F1**3-A2*F2**3)/3.
      D=A1*(F2-F1)
      IF(ABS(D).LE.1.E-10) GO TO 20
      SIG12=(B1*F2-B2)/D
      SIG22=(B1-A1*SIG12)/A2
      RETURN
   20 SIG12=-1
      SIG22=-1
      RETURN
      END
```

```
      SUBROUTINE COMPUT(AI,UI,UJ,SI,SJ)
      DIMENSION A(5),B(5),C(5)
      COMMON/BLK20/E1,E2,E3,E4,E5
      EQUIVALENCE (E1,C(1))
      AJ=1.-AI
      A(1)=AI*UI+AJ*UJ
      A(2)=AI*(SI+UI*UI)+AJ*(SJ+UJ*UJ)
      A(3)=AI*UI*(3*SI+UI*UI)+AJ*UJ*(3*SJ  +UJ*UJ)
      A(4)=AI*(3*SI*(2*UI*UI+SI)+UI**4)+AJ*(3*SJ*(2*UJ*UJ+SJ)+UJ**4)
      A(5)=AI*(5*UI*SI*(3*SI+2*UI*UI)+UI**5)+AJ*(5*UJ*SJ*(3*SJ+2*UJ*UJ
     1+UJ**5)
      DO 5 I=1,5
    5 B(I)=C(I)-A(I)
      WRITE(6,3)
      WRITE(6,4)
    4 FORMAT(16X,*======================================================
     1=== *)
      WRITE(6,2) (C(I),I=1,5)
      WRITE(6,6) (A(I),I=1,5)
      WRITE(6,7) (B(I),I=1,5)
    3 FORMAT(*0*,17X,*1*,11X,*2*,11X,*3*,11X,*4*,11X,*5* )
    2 FORMAT(*0INPUT NOM. *,F10.5,F12.4,F12.3,F12.2,F12.1 )
    6 FORMAT(*0CALCULATED *,F10.5,F12.4,F12.3,F12.2,F12.1 )
    7 FORMAT(*0DIFFERENCE *,F10.5,F12.4,F12.3,F12.2,F12.1 )
      RETURN
      END
```

BIBLIOGRAPHY

Aitken, M.A. (1974) Simultaneous inference and the choice of variable subsets in multiple regression. *Technometrics* 16(2), 221-227.

Allen, D.M. (1971) Mean square error of prediction as a criterion for selecting variables. *Technometrics* 13(3), 469-481.

Allen, D.M. (1974) The relationship between variable selection and data augmentation and a method for prediction. *Technometrics* 16(1), 125-127.

Anderson, T.W. and Darling, D.A. (1954) A test of goodness of fit. *J. Amer. Stat. Assoc.* 49, 765-69.

Andrews, D.F., Gnanadesikan, R. and Warner, J.L. (1973) Methods for assessing multivariate normality. *Proc. International Symposium Multivariate Analysis* (P. Krishnaiah ed.), Vol. 3, 95-116, Academic Press, New York.

Atkins, K.D. and McGuirk, B.J. (1976) Genetic improvements of fertility and fly resistence. In *Sheep Breeding* (Tomes, G.J., Robertson,D.E., and Lightfoot, R.J. eds) 271-275.

Barlow, R. (1974) Selection for clean fleece weight in Merino sheep. II Correlated responses to selection. *Aust. J. Agric. Res.* 25, 973-94.

Barrett, J.P. (1974) The coefficient of determination - some limitations. *The American Statistician* 28/1, 19-20.

Beattie, A.W. (1962) Relationships among productive characters of Merino sheep in north-western Queensland. 2. Estimates of genetic parameters with particular reference to selection for wool weight and crimp frequency. *Qld J. of Agric. Sci.* 19, 17-26.

Behboodian, J. (1970) On a mixture of normal distributions. *Biometrika* 57, 215-7.

Beilharz, R.G. (1963) On the possibility that sex-chromosomes have a greater effect than autosomes on inheritance. *Journal of Genetics* 58(3), 441-449.

Bennett, B.M. and Nakamura, E. (1963) Tables for testing significance in a 2 x 3 contingency table. *Technometrics* 5(4), 501-11.

Bindon, B.M. and Piper, L.R. (1976) Assessment of new and traditional techniques of selection for reproduction rate. In *Sheep Breeding* (Tomes,G.J., Robertson,D.E. and Lightfoot,R.J. eds) 357-371.

Bindon, B.M. and Turner, H.N. (1974) Plasma LH of the prepubertal lamb: a possible early indicator of fecundity. *Journal of Reproduction and Fertility* 39, 85-88.

Boulton, D.M. (1974) Remark on Algorithm 434(G2). Communications of the A.C.M. 17(6), 326.

Boulton, D.M. and Wallace, C.S. (1973) Occupancy of a rectangular array. *Computer J.* 16(1), 57-63.

Bowman, K.O. and Shenton, L.R. (1973) Remarks on the distribution of Karl Pearson's k-statistic, $k = 2b_2 - 3b_1 - 6$. Contributed paper *39th Session of the International Statistical Institute* 134-142.

Breese, E.L. and Mather, K. (1957) The organization of polygenic activity within a chromosome in *Drosophila*. I. Hair characters. *Heredity* 11, 373-95.

Brown, G.H. and Turner, H.N. (1968) Response to selection in Australian Merino sheep. II. Estimates of phenotypic and genetic parameters for some production traits in Merino ewes and an analysis of the possible effects of selection on them. *Aust. J. Agric. Res.* 19, 303-22.

Chen, E.H. (1971) The power of the Shapiro-Wilk W Test for normality in samples from contaminated normal distribution. *J. Amer. Stat. Assoc.* 66(336), 760-62.

Clarke, C.A. and Sheppard, P.M. (1962) Disruptive selection and its effects on a metrical character in the butterfly *Papilio dandanus*. *Evolution* 16, 214-226.

Clayton, G.A. and Robertson, A. (1957) An experimental check on quantitative genetical theory. II. The long term effects of selection. *J. Genet.* 55, 152-170.

Cohen, A.C. (1967) Estimation of mixtures of two normal distributions. *Technometrics* 9, 15-28.

Cooke, P. and Mather, K. (1962) Estimating the components of continuous variation. II. Genetical. *Heredity* 17, 211-36.

Cox, D.R. and Snell, E.J. (1974) The choice of variables in observational studies. *Applied Statist.* 23(1), 51.

Cramer, H. (1928) On the composition of elementary errors. *Skand. Aktuar.* 11, 141-80.

Daday, H., Binet, F.E., Grassia, A. and Peak, J.W. (1973) The effect of environment on heritability and predicted selection response in *Medicago Sativa*. *Heredity* 31(3), 293-308.

D'Agostino, R. and Pearson, E.S. (1973) Tests for departure from normality. Empirical results for the distribution of b_2 and $\sqrt{b_1}$. *Biometrika* 60(3), 613-622.

D'Agostino, R. and Pearson, E.S. (1974) Corrections and amendments: Tests for departure from normality. Empirical results for the distribution of b_2 and $\sqrt{b_1}$ *Biometrika* 61(3), 647.

Dahl, O.-J., Dijkstra, E.W. and Hoare, C.A.R. (1972) *Structured Programming* Academic Press, London and New York.

Daling, J.R. and Tamura, H. (1970) Use of orthogonal factors for selection of variables in a regression equation - an illustration. *Applied Statistics* 19, 260-268.

David, H.A., Hartley, H.O. and Pearson, E.S. (1954) The distribution of the ratio, in a single normal sample, of range to standard deviation. *Biometrika* 41, 482-93.

di Cesnola, A.P. (1907) A first study of natural selection in *Helix arbustorum* (Helicogena). *Biometrika* 5, 387-99.

Diehr, G. and Hoflin, D.R. (1974) Approximating the distribution of the sample R^2 in best subset regressions. *Technometrics* 16(2), 317-320.

Donald, H.P. (1973) Improvement in biological efficiency. Animal Breeding: contributions to the efficiency of livestock production. *Philisophical Transactions of the Royal Society of London* (B) 267(882), 131-144.

Draper, N.R., Guttman, I. and Kanemasu, H. (1971) The distribution of certain regression statistics. *Biometrika* 58, 295-298.

Draper, N.R. and Smith, H. (1966) *Applied Regression Analysis.* Wiley, New York.

Dubzinski, M.L. (1975) From *Developments in Field Experiment Design and Analysis.* (V.J. Bofinger, ed.)

Dun, R.B. (1958) The influence of selection and plane of nutrition on the components of fleece weight in Merino sheep. *Aust. J. Agric. Res.* 9, 802.

Durbin, J. (1961) Some methods of constructing exact tests. *Biometrika* 48, 41-55.

Dyer, A.R. (1974) Comparisons of tests for normality with a cautionary note. *Biometrika* 61, 185-189.

Elston, R.C., Namboodiri, K.K., Nino, H.V. and Pollitzer, W.S. (1974) Studies on blood and urine glucose in Seminole Indians: Indications for segregation of a major gene. *Amer. J. hum. Genet.* 26, 13-34.

Ewens, W.J. (1969) *Population Genetics.* Methuen & Co. Ltd, London.

Falconer, D.S. (1952) The problem of environment and selection. *Amer. Nat.* 86, 293-298.

Falconer, D.S. (1960) *Introduction to Quantitative Genetics.* Oliver and Boyd, Edinburgh.

Falconer, D.S. and Latyszewski, M. (1952) The environment in relation to selection for size in mice. *J. Genet.* 51, 67-80.

Finney, D.J. (1948) The Fisher-Yates test of significance in 2 x 2 contingency tables. *Biometrika* 35, 149-154.

Finney, D.J. (1962) Genetic gains under three methods of selection. *Genet. Res.* 3, 417-23.

Finney, D.J., Latscha, R., Bennett, B.M. and Hsu, P. (1963) Tables for testing significance in a 2 x 2 contingency table. Cambridge University Press, New York.

Fisher, R.A. (1925) *Statistical Methods for Research Workers.* Oliver and Boyd, Edinburgh.

Fisher, R.A. (1930) *The Genetical Theory of Natural Selection.* Oxford, Clarendon Press.

Fisher, R.A. (1936) The use of multiple measurements in taxonomic problems. *Annals of Eugenics* 7, 179-188.

Fisher, R.A., Immer, F.R. and Tedin, O. (1932) The genetical interpretation of statistics of the third degree in the study of quantitative inheritance. *Genetics, Princeton.* 17, 107-24.

Forrest, P.A. and Bichard, M. (1974) Analysis of production records from a Lowland sheep flock. 3. Phenotypic and genetic parameters from reproductive performance. *Animal Production* 19, 33-45.

Fraccaro, M. (1956) A contribution to the study of birth weight based on an Italian sample. *Ann. hum. Genet.* 20, 282-97.

Frankham, R. (1969) Genetical analysis of two abdominal bristle selection lines. *Aust. J. biol. Sci.* 22, 1485-95.

Freeman, G.H. and Halton, J.H. (1951) Note on an exact treatment of contingency, goodness of fit and other problems of significance. *Biometrika* 38, 141-49.

Fryer, J.G. and Robertson, C.A. (1972) A comparison of some methods for estimating mixed normal distributions. *Biometrika* 59(3), 639-648.

Furnival, G.M. and Wilson, R.W. (1974a) Regression by leaps and bounds. *Technometrics* 16(4), 499-511.

Furnival, G.M. and Wilson, R.W. (1974b) Regression by leaps and bounds. A program for finding the best subset regressions. Personal communication.

Gardiner, C.O. (1961) Estimates of genetic parameters in cross-fertilizing plants and their implication in plant breeding. In *Statistical Genetics and Plant Breeding*. *Natn. Acad. Sci., Natn. Res. Council* Publ. No. 982, 225-52.

Geary, R.C. (1935) The ratio of the mean deviation to the standard deviation as a test of normality. *Biometrika* 27, 310-32.

Gibson, J.B. and Bradley, B.P. (1974) Stabilising selection in constant and fluctuating environments. *Heredity* 33, 293-302.

Gibson, J.B. and Thoday, J.M. (1962) Effects of disruptive selection. VI. A second chromosome polymorphism. *Heredity* 17, 1-26.

Gill, J.L. (1965) Selection and linkage in simulated genetic populations. *Aust. J. biol. Sci.* 18, 1171-1187.

Gorman, J.W. and Toman, R.J. (1966) Selection of variables for fitting equations to data. *Technometrics* 8(1), 27-51.

Haldane, J.B.S. (1951) Simple tests for bimodality and bitangentiality. *Annals Eugenics* 16, 359-64.

Hammond, J. (1947) Animal breeding in relation to nutrition and environmental conditions. *Biol. Rev.* 22, 195-213.

Hammond, K., Jackson, N. and Miller, D.H. (1972) Analysis of quantitative genetic variation by nested analysis of variance and offspring-parent regression. A program package - NESREG. *Dept. of Animal Husbandry Research Report No. 1*, University of Sydney, Sydney, Australia.

Hammond, K., and James, J.W. (1970) Genes of large effect and the shape of the distribution of a quantitative character. *Aust. J. biol. Sci.* 23, 867-876.

Hammond, K. and James, J.W. (1972) The use of higher degree statistics to estimate the number of loci which contribute to a quantitative character. *Heredity* 28(1), 146-7.

Hancock, T.W. (1975) Remark on algorithm 434(G2).
Communications of the A.C.M. 18(2), 117-119.

Hancock, T.W. (1976) Assessment of fleece components using multivariate statistics.
Proc. Aust. Soc. Anim. Prod. 11, 177-80.

Harris, H. (1966) Enzyme polymorphisms in man.
Proc. Roy. Soc., Lond. B 164, 298-310.

Harville, D.A. (1975) Index selection with proportionality constraints.
Biometrics 31, 223-25.

Hasselblad, V. (1966) Estimation of parameters for a mixture of normal distributions.
Technometrics 8, 431-44.

Hawkins, R.H. (1971) A note on mixed distributions in hydrology.
In *Proceedings of Symposium on Statistical Hydrology*, Tucson, Arizona.

Hawkins, R.H. (1972) A note on multiple solutions to the mixed distribution problem.
Technometrics 14(4), 973-76.

Hawkins, R.H. (1975) Mixed normal distributions. A moment solution.
Personal communication, Dept. of Forest Science, Utah State Uni., Logan, Utah, 84322, U.S.A.

Hazel, L.N. and Lush, J.L. (1942) The efficiency of three methods of selection.
J. Hered. 33, 393-99.

Hedrick, P.W. (1970) Selection on finite populations. II. The selection limit and rate on response for a Monte Carlo simulation model.
Genetics 65, 175-186.

Henderson, C.R. (1963) Selection index and expected genetic advance.
Statistical Genetics and Plant Breeding No. 982, 141-63.

Hocking, R.R. (1972) Criteria for selection of a subset regression: Which one should be used?
Technometrics 14(4), 967-70.

Hocking, R.R. (1976) The analysis and selection of variables in linear regression. *Biometrics* 32, 1-49.

Hocking, R.R. and Leslie, R.N. (1967). Selection of the best subset in regression analysis.
Technometrics 9(4), 531-40.

Hoerl, A.E. and Kennard, R.W. (1970a) Ridge regression. Biased estimation for nonorthogonal problems.
Technometrics 12(1), 55-67.

Hoerl, A.E. and Kennard, R.W. (1970b) Ridge regression. Applications to nonorthogonal problems.
Technometrics 12(1), 69-82.

Hopkins, A.M. (1974) Personal communication from Honours Thesis submitted to Waite Agricultural Research Institute, University of Adelaide.

Hotelling, H. (1933) Analysis of a complex of statistical variables in principal components.
J. educational Psychology 24, 417-441 and 498-520.

Hotelling, H. (1936) Simplified calculation of principal components.
Psychometrika 1, 27-35.

IMSL Library (1974) International Mathematical and Statistical Libraries Inc., Houston, Texas.

Jackson, N., Nay, T. and Turner, H.N. (1975) Response to selection in Australian Merino sheep. VII. Phenotypic and genetic parameters for some wool follicle characteristics and their correlation with wool and body traits.
Aust. J. agric. Res. 26, 937-57.

James, J.W. (1973) Covariances between relatives due to sex-linked genes.
Biometrics 29, 584-88.

Jayant, K. (1966) Birth weight and survival; a hospital survey repeat after 15 years.
Ann. hum. Genet. 29, 367-375.

Jinks, J.L. and Connolly, V. (1973) Selection for specific and general response to environmental differences.
Heredity 30(1), 33-40.

Jinks, J.L. and Connolly, V. (1975) Determination of the environmental sensitivity of selection lines by the selection environment.
Heredity 34(3), 401-406.

Jinks, J.L. and Towey, Philomena (1976) Estimating the number of genes in a polygenic system by genotype assay.
Heredity 37(1), 69-81.

John, S. (1970) On identifying the population of origin of each observation in a mixture of observations from two normal populations.
Technometrics 12(3) 553-63.

Jones, L.P. (1967) Effects of X-rays on response to selection for a quantitative character of *Drosophila melanogaster*.
Genet. Res. 9, 221-31.

Kalmus, H. and Maynard Smith, S. (1965) The antimode and lines of optimal separation in a genetically determined bimodal distribution, with particular reference to phenylthiocarbamide sensitivity.
Ann. hum. Genet. 29, 127-138.

Karn, M.N. and Penrose, L.S. (1951) Birth weight and gestation time in relation to maternal age, parity and infant survival.
Ann. Eugen. 17, 147-64.

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

Kendall, M.G. (1975) *Multivariate Analysis.* Charles Griffin & Co., London.

Kennard, R.W. (1971) A note on the C_p statistic. *Technometrics* 13(4), 899-900.

Kimura, M. and Ohta, T. (1973) Mutation and evolution at the molecular level. *Genetics Supplement* 73, 19-35.

Knuth, D.E. (1973) *The Art of Computer Programming Vol. 3/ Sorting and Searching.* Wesley Publ. Co., London.

Kolmogorov, A.N. (1933) Sulla determinazione empirica di una legge di distribuzione. *G. Inst. Ital. Attuari* 4, 83-91.

Lamotte, L.R. and Hocking, R.R. (1970) Computational efficiency in the selection of regression variables. *Technometrics* 12(1), 83-93.

Land, R.B. (1974) Physiological studies and genetic selection for sheep fertility. *Animal Breeding Abstracts* 42, 155-58.

Law, C.N. (1967) The location of genetic factors controlling a number of quantitative characters in wheat. *Genetics* 56, 445-61.

Lee, B.T.O. and Parsons, P.A. (1968) Selection prediction and response. *Biological Review* 43, 139-174.

Lee, B.T.O. and Patemen, J.A. (1959) Linkage of polygenes controlling size of ascospores in *Neurospora crassa.* *Nature, Lond.* 183, 698-9.

Lerner, I.M. and Donald, H.P. (1966) *Modern Developments in Animal Breeding.* Academic Press, London and New York.

Lester, R.H., Elston, R.C. and Graham, J.B. (1972) Variations in levels of blood clotting factors IX and X in a population of normal men: possible genetic polymorphisms. *Amer. J. hum. Genet.* 24, 168-81.

Lewontin, R.C. and Hubby, J.L. (1966) A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura.* *Genetics* 54, 595-609.

Lilliefors, H.W. (1967) On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J. Amer. Stat. Assoc.* 62(318), 399-402.

Maclean, C.J., Morton, N.E., Elston, R.C. and Yee, S. (1976) Skewness in commingled distributions. *Biometrics* 32, 695-699.

McGuirk, B.J. (1976) Relevance to producers of research into the genetic improvement of reproductive performance. *Proc. Aust. Soc. Anim. Prod.* XI, 93-100.

McMillan, I. and Robertson, A. (1974) The power of methods for the detection of major genes affecting quantitative characters. *Heredity* 32(3), 349-356.

March, D.L. (1972) Exact probabilities for R x C contingency tables(G2). Communications of the A.C.M. 15(11), 991-992.

Mardia, K.V. (1975) Assessment of multinormality and the robustness of Hotelling's T^2 test. *Applied Statistics* 24(2), 163-171.

Mather, K. (1941) Variation and selection of polygenic characters. *J. Genet.* 41, 159-93.

Mather, K. (1943) Polygenic inheritance and natural selection. *Biological Review* 18, 32-64.

Mather, K. (1949) *Biometrical Genetics*. Methuen & Co., London.

Mather, K. (1953) The genetical structures of populations. *Symp. Soc. Exp. Biol.* I, 66-95.

Mather, K. and Harrison, B.J. (1949) The manifold effect of selection. *Heredity* 3, 1-52, 131-62.

Mather, K. and Wigan, L.G. (1942) Correlated response to the selection of polygenic characters. *Ann. Eugen.* 11, 354-64.

Mayo, O., Murdoch, J.L. and Hancock, T.W. (1976) On the estimation of parental age effects on mutation. *Ann. hum. Genet., Lond.* 39, 427-31.

Mayo, O., Potter, J.C., Brady, R.E. and Hooper, C.W. (1969) Response to partial selection on clean fleece weight in South Australian strong-wool Merino sheep. I. Results of the experiment. *Aust. J. Agric. Res.* 20, 151-67.

Mérat, P. (1968) Distributions de frequences, interpretation du determinisme genetique des characteres quantitatifs et recherche de génes majeurs. *Biometrics* 24(2), 277-293.

Meyer, H.H. and Enfield, F.D. (1975) Experimental evidence on limitations of the heritability parameter. *Theoretical and Applied Genetics* 45, 268-73.

Morley, F.H.W. (1955) Genetic improvement of Australian Merino sheep. *Agric. Gaz., N.S.W.* 66, 400-11.

Morley, F.H.W. (1955) Selection for economic characters in Merino sheep. V. Further estimates of phenotypic and genetic parameters. *Aust. J. Agric. Res.* 6, 77-90.

Morrison, D.F. (1967) *Multivariate Statistical Methods.* McGraw-Hill, New York.

Narula, S.C. (1974) Predictive mean square error and stochastic regressor variables. *Applied Statist.* 23(1), 11-17.

Nie, N.H., Hadlaihull, C., Jenkins, J.G., Steinbrenner, K. and Bent, D.H. (1975) *Statistical Package for the Social Sciences.* McGraw-Hill, New York, second edition.

Nordskøg, A.W. and Wehrli, M. (1963) Effects of selection on reproductive fitness in the fowl. *Proc. Xth Int. Congr. Genet.* 1, 259.

O'Donald, P. (1971) The distribution of genotypes produced by alleles segregating at a number of loci. *Heredity* 26, 233-41.

Panse, V.G. (1940) The application of genetics to plant breeding. II. The inheritance of quantitative characters and plant breeding. *J. Genet.* 40, 283-302.

Pattie, W.A. (1965) Selection for weaning weight in Merino sheep. 1. Direct approach to selection. *Aust. J. Exper. Agric. and Anim. Husb.* 5, 353-60 and 361-8.

Pattie, W.A. and Barlow, R. (1974) Selection for clean fleece weight in Merino sheep. I. Direct response to selection. *Aust. J. Agric. Res.* 25, 643-55.

Pearson, E.S. and Hartley, H.O. (1958) *Biometrika Tables for Statisticians.* Cambridge University Press, London.

Pearson, K. (1894) Contributions to the mathematical theory of evolution. *Philosophical Transaction of the Royal Society* Series A. 185, 71-110.

Pearson, K. (1901) On lines and planes of closest fit to systems of points in space. *Philosophical Magazine Ser. 6*, Vol. 2, 559-72.

Pearson, K. (1902) On the mathematical theory of errors of judgement with special reference to the personal equation. *Philosophical Transactions of the Royal Society* Series A. 198,235-99.

Piper, L.R. (1971) Ph.D. Thesis submitted to the Institute of Animal Genetics, University of Edinburgh.

Pope, P.T. and Webster, J.T. (1972) The use of an F-statistic in stepwise regression procedures. *Technometrics* 14, 327-40.

Prout, T. (1962) The effects of stabilising selection on the time of development in *Drosophila melanogaster.* *Genet. Res.* 3, 364-82.

Qureshi, A.W. and Kempthorne, O. (1968) On the fixation of genes of large effects due to continued transaction selection in small populations of polygenic systems with linkage. *Theoret. Appl. Genet.* 38, 249-55.

Rayment, P.R. (1972) The identification problem for a mixture of observations from two normal populations. *Technometrics* 14(4), 911-918.

Rendel, J.M. (1943) Variations in the weights of hatched and unhatched duck's eggs. *Biometrika* 33, 48-52.

Rendel, J.M. and Binet, F.E. (1974) The effect of environment on heritability and predicted selection response: A Reply. *Heredity* 33(1), 106-108.

Robards, G.E. and Pattie, W.A. (1967) Selection for crimp frequency in the wool of Merino sheep. 1. Direct response to selection. *Aust. J. Agric. Res. and Anim. Husb.* I, 552-558.

Robards, G.E., Williams, A.J. and Hunt, M.H.R. (1974) Selection for crimp frequency in the wool of Merino sheep. 2. Efficiency of conversion of food to wool. *Aust. J. Exp. Agric. and Anim. Husb.* 14, 441-48.

Robertson, A. (1955) Selection in animals: synthesis. *Cold Spring Harb. Symp. Quant. Biol.* 20, 225-9.

Robertson, A. (1956) The effect of selection against extreme deviants based on deviation or on homozygosis. *J. of Genetics* 54, 236-248.

Robertson, A. (1959) The sampling variance of the genetic correlation coefficient. *Biometrics* 15, 469-85.

Robertson, A. (1961) A theory of limits in artificial selection. *Proc. of the Royal Society B.* 153, 234-49.

Robertson, A. (1966) Artificial selection in plants and animals. *Proc. of the Royal Society B.* 164, 341-9.

Robertson, A. (1967) The nature of quantitative genetic variation. In *Heritage of Mendel* (R.A. Brink, ed.) 265-80. (University of Wisconsin Press, Madison, Wisconsin, U.S.A.)

Robertson, A. (1970) A theory of limits in artificial selection with many linked loci. In *Mathematical Topics in Population Genetics* (K. Kojima, ed.) 246-88.

Robertson, C.A. and Fryer, J.C. (1970) The bias and accuracy of moment estimators. *Biometrics* 57, 57-65.

Robertson, W.H. (1960) Programming Fisher's exact method of comparing two percentages. *Technometrics* 2(1), 103-107.

Sales, J. and Hill, W.G. (1976a) Effect of sampling error in efficiency of selection indices. 1. Use of information from relatives for single trait improvement. *Anim. Prod.* 22, 1-17.

Sales, J. and Hill, W.G. (1976b) Effect of sampling errors on efficiency of selection indices. 2. Use of information on associated traits for improvement of a single important trait.
Anim. Prod. 23, 1-14.

Saville, D.G. and Robards, G.E. (1972) Efficiency of conversion of food to wool in selected and unselected Merino types.
Aust. J. Agric. Res. 23, 117-30.

Scharloo, W. (1964) The effect of disruptive and stabilising selection on the expression of a *cubitus interruptus* mutant in *Drosophila*.
Genetics 50, 553-62.

Schinckel, P.G. (1955) Fleece measurement and flock improvement.
Aust. Vet. J. 32, 274-7.

Schinckel, P.G. (1955a) The post-natal development of the skin follicle population in a strain of Merino sheep.
Aust. J. Agric. Res. 6, 68-76.

Schinckel, P.G. (1955b) The relationship of skin follicle development to growth rate in sheep.
Aust. J. Agric. Res. 6, 308-23.

Schinckel, P.G. (1958) The relationship of lamb birthcoat to adult fleece structure in a strain of Merino sheep.
Aust. J. Agric. Res. 9, 567-78.

Shapiro, S.S. and Francia, R.S. (1972) An approximate analysis of variance test for normality.
J. Amer. Stat. Assoc. 67(337), 215-6.

Shapiro, S.S. and Wilk, M.B. (1965) An analysis of variance test for normality (complete samples).
Biometrika 52(3 & 4), 591-611.

Shapiro, S.S., Wilk, M.B. and Chen, H.J. (1968) Comparative study of various tests for normality.
J. Amer. Stat. Assoc. 63, 1343-72.

Sheppard, P.M. (1953) Polymorphism, linkage and the blood groups.
Amer. Naturalist 87, 283-94.

Siegel, S. (1956) *Nonparametric Statistics for Behavioural Sciences.*
McGraw-Hill Book Co., Inc., New York.

Sokal, R.R. and Rohlf, F.J. (1969) *Biometry.* W.H. Freeman & Co.

Spickett, S.G. (1963) Genetic and developmental studies of a quantitative character.
Nature, Lond. 199, 870-3.

Spickett, S.G. and Thoday, J.M. (1966) Regular responses to selection. 3. Interaction between located polygenes.
Genetical Research 7, 96-121.

Subrahmaniam, K., Subrahmaniam, K. and Messeri, J.Y. (1975) On the robustness of some tests of significance in sampling from a compound normal population. *J. Amer. Stat. Assoc.* 70(350), 435-8.

Sugiura, N. and Ôtake, M. (1968) Numerical comparison of improved methods of testing in contingency tables with small frequencies. *Ann. Inst. of Stat. Maths.* 20, 505-17.

Tallis, G.M. (1962) A selection index for optimum genotype. *Biometrics* 18, 120-3.

Tallis, G.M. (1968) Selection for an optimum growth curve. *Biometrics* 24, 169-77.

Thoday, J.M. (1959) Effects of disruptive selection. I. Genetic flexibility. *Heredity* 13, 187-203.

Thoday, J.M. (1961) The location of polygenes. *Nature, Lond.* 191, 368-70.

Thoday, J.M. (1973) The origin of genes found in selected lines. Atti della Accademia delle Scienze dell 'Instituto di Bologna, Classe di Scienze Fisiche, Anno 261, Memorie, Serie III, N.1,15-25.

Thoday, J.M. and Boam, T.B. (1961) Regular responses to selection. I. Description of responses. *Genet. Res.* 2, 161-76.

Thoday, J.M., Gibson, J.B. and Spickett, S.G. (1964) Regular responses to selection. 2. *Genet. Res.* 5, 1-19.

Thompson, J.N. and Thoday, J.M. (1974) A definition of standard nomenclature for "polygenic loci". *Heredity* 33, 430-7.

Turner, H.N. (1956) Selection for fibre number in Merino sheep. *Nature, Lond.* 178, 330.

Turner, H.N. (1958) Relationships among clean wool weight and its components. I. Changes in clean wool weight related to changes in the compoments. *Aust. J. Agric. Res.* 9(4), 521-52.

Turner, H.N. (1966) Selection for increased reproduction rate. *Wool Technol. Sheep Breed.* 13(1), 69-79.

Turner, H.N., Brooker, M.G. and Dolling, C.H.S. (1970) Response to selection in Australian Merino sheep. III. Single character selection for high and low values of wool weight and its components. *Aust. J. Agric. Res.* 21, 995-84.

Turner, H.N., Dolling, C.H.S. and Kennedy, J.F. (1968) Response to selection in Australian Merino sheep. 1. Selection for high clean wool weight, with a ceiling on fibre diameter and degree of skin wrinkle. Response in wool and body characteristics. *Aust. J. Agric. Res.* 19, 79-112.

Turner, H.N., Hayman, R.H., Triffitt, L.K. and Prunster, R.W. (1962) Response to selection for multiple birth in Australian Merino: a progress report.
Animal Production 4, 165.

Turner, H.N., McKay, E. and Guinane, F. (1972) Response to selection in Australian Merino sheep. IV. Reproduction rate in groups selected for high clean wool weight with a ceiling on degree of skin wrinkle and either fibre diameter or crimp frequency.
Aust. J. Agric. Res. 23, 131-48.

Turner, J.R.G. (1967a) Why does the genotype not congeal?
Evolution 21, 645-656.

Turner, J.R.G. (1967b) On supergenes. I. The evolution of supergenes.
Amer. Naturalist 101, 195-221.

Van Vleck, L.D. (1976) Selection for direct, maternal and grand-maternal genetic components of economic traits.
Biometrics 32, 173-181.

Waddington, H. and Lewontin, R.C. (1967) A note on evolution and changes in the quantity of genetic information in C.H. Waddington (ed.) *Towards a Theoretical Biology* Vol. 1, Edinburgh.

Wallace, B. (1975) Gene control mechanisms and their possible bearing on the neutralist-selectionist controversy.
Evolution 29(2), 193-202.

Wallace, L.R. (1948) Growth of lambs before and after birth in relation to the level of nutrition. Parts I and III.
J. Agric. Sci. 38, 93, 367.

Wallace, L.R. (1958) Breeding Romneys for better lambing percentages.
New Zealand J. Agric. 97, 545-50.

Wallace, L.R. (1964) Breeding performance of Romneys improves in long term experiment.
New Zealand J. Agric. 100, 417-24.

Walls, R.C. and Weeks, D.L. (1969) A note on the variance of a predicted response in regression.
Amer. Statistician 23/3, 24-26.

Wehrhahn, C. and Allard, R.W. (1965) The detection and measurement of the effect of individual genes involved in the inheritance of a quantitative character in wheat.
Genetics 51, 109-19.

Weisberg, S. (1974) An empirical comparison of the percentage points of W and W'.
Biometrika 61(3), 644-46.

Weldon, W.F.R. (1901) A first study of natural selection in *Clausilia laminata* (Montagu).
Biometrika 1, 109-24.

Wigan, L.G. and Mather, K. (1942) Correlated response to the selection of polygenic characters.
Ann. Eugen. II, 354-64.

Wolstenholme, D.R. and Thoday, J.M. (1963) Effects of disruptive selection. VII. A third chromosome polymorphism.
Heredity 18, 413-31.

Wright, S. (1934) The results of crosses between inbred strains of guinea pigs, differing in number of digits.
Genetics 19, 537-51.

Wright, S. (1952) The genetics of quantitative variability.
In *Quantitative Inheritance* (eds Reeve, E.C.R. and Waddington, C.H.) pp. 5-41.

Young, S.S.Y. (1961) A further examination of the relative efficiency of three methods of selection for genetic gains under less-restricted conditions.
Genet. Res. 2, 106-21.

Young, S.S.Y. (1967) Computer simulation of directional selection in large populations. II. The additive x additive and mixed models.
Genetics 56, 73-87.

Young, S.S.Y., Turner, H.N. and Dolling, C.H.S. (1960) Comparison of estimates of repeatability and heritability for some production traits in Merino rams and ewes. II. Heritability.
Aust. J. Agric. Res. 11, 604-17.