1951

 Causes of variation in 154-day weights among Poland China pigs from multiple crosses of inbred lines

 William Thomas Magee
 Iowa State College

 Follow this and additional works at: http://lib.dr.iastate.edu/rtd

 Part of the Animal Sciences Commons, Genetics Commons, Physiology Commons, and the Veterinary Physiology Commons

 Recommended Citation
 Magee, William Thomas, "Causes of variation in 154-day weights among Poland China pigs from multiple crosses of inbred lines" (1951). Retrospective Theses and Dissertations. 14099.
 http://lib.dr.iastate.edu/rtd/14099

 This Dissertation is brought to you for free and open access by Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
NOTE TO USERS

This reproduction is the best copy available.

UMI®
CAUSES OF VARIATION IN 154-DAY
WEIGHTS AMONG POLAND CHINA PIGS FROM MULTIPLE
CROSSES OF INBRED LINES

by

William Thomas Magee

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Animal Breeding

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State College

1951
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. STATISTICAL METHODS</td>
<td>7</td>
</tr>
<tr>
<td>III. ANALYSIS OF EXPERIMENTAL RESULTS</td>
<td>26</td>
</tr>
<tr>
<td>A. Sources and Scope of the Data</td>
<td>28</td>
</tr>
<tr>
<td>B. Factors Considered in the Analysis</td>
<td>30</td>
</tr>
<tr>
<td>C. Estimation of Effects</td>
<td>31</td>
</tr>
<tr>
<td>D. Tests of Hypotheses</td>
<td>36</td>
</tr>
<tr>
<td>E. Estimating Variance Components</td>
<td>44</td>
</tr>
<tr>
<td>F. An Improvement in Procedure</td>
<td>46</td>
</tr>
<tr>
<td>IV. DISCUSSION AND APPLICATION</td>
<td>47</td>
</tr>
<tr>
<td>V. SUMMARY</td>
<td>55</td>
</tr>
<tr>
<td>VI. LITERATURE CITED</td>
<td>57</td>
</tr>
<tr>
<td>VII. ACKNOWLEDGEMENTS</td>
<td>59</td>
</tr>
</tbody>
</table>
Any change observed on the threed lines, double crosses, and probed G1 lines were observed on the three lines, double crosses, and probed G1 lines. The same point in each of the three components increased over 10% when both size and color were increased, but none of the increases were observed on the threed lines, double crosses, and probed G1 lines. There was no additional increase in the threed lines, double crosses, but there was some additional increase in the threed lines, double crosses. Mortality between the two groups decreased about the same extent. Mortality between the two groups decreased about the same extent. Mortality between the two groups decreased about the same extent. Mortality between the two groups decreased about the same extent. Mortality between the two groups decreased about the same extent.

When (1962) reported on the effect of threed lines, double crosses, and probed G1 lines on the threed lines, double crosses, and probed G1 lines. He compared the three threed lines with three threed lines. The three threed lines were compared with three threed lines. He compared the three threed lines with three threed lines. The three threed lines were compared with three threed lines. He compared the three threed lines with three threed lines. The three threed lines were compared with three threed lines.

Introduction

Introduction laid no stress on the results between different lines, nor the results between different lines, nor the results between different lines, nor the results between different lines, nor the results between different lines, nor the results between different lines.
from mating inbred males to crossbred females equalled or exceeded the randombred stock in average size of litter, but only the double cross progeny exceeded the randombred stock in individual weight at 33 days, which was weaning age. In most of the traits studied, the double cross progeny were superior to those from any other type of mating.

Most other experiments on inbreeding and crossbreeding in plants and animals have shown that crossing two inbred lines which are not closely related but are from the same species will produce F₁'s that are more vigorous than either of the inbred parental lines.

Two major explanations of the cause of hybrid vigor have been expounded. Shull (1908) and East (1908) postulated that there was a physiological stimulus from having different alleles at some loci. Thus maximum vigor would be expressed in only those individuals that are heterozygous at these loci. This phenomenon has been described more precisely for the individual locus as overdominance, i.e., the heterozygote is preferred to either homozygote.

East and Jones (1919) proposed the linkage and dominant factor hypothesis. The partial or complete dominance of desirable genes to their less desirable or detrimental alleles and the linkage between desirable and undesirable genes are the two primary components of this hypothesis. Thus hybrid vigor would result from crossing two individuals who each contribute some desirable dominant genes which will cover up their deleterious alleles from the other parent.

Epistatic effects among nonallellic genes might also be an underlying cause of hybrid vigor.
A breeding plan to utilize hybrid vigor would start much the same way regardless of the underlying cause. In each case the first phase of the breeding plan would be to form inbred lines and to test them in crosses. After this, the breeding plan used would depend on the cause of the hybrid vigor. If dominance or epistasis is the underlying cause, the progeny from the lines that cross well together should be selected and interbred at random for several generations to allow crossing over among the different loci. Then new inbred lines would be formed from this stock and the procedure repeated. The ultimate goal of this procedure would be to develop an inbred line as good as any cross between lines.

If overdominance is the underlying cause, the first lines formed would be divided into two groups, each group being composed of lines that cross well with lines in the other group. The individuals within each group would then be random bred and new lines developed from within each group. The new lines would be tested on a tester stock formed from the lines of the other group. This procedure would be repeated until lines were developed that had unlike alleles at all loci which exhibited overdominance. Crosses between these lines would be better than any pure line that could be developed. A more complete discussion of this matter is given by Lush (1946). Whichever plan is followed, the early commercial use of the lines will involve utilizing the better crosses found up to that time.

The most important factor to consider in deciding whether to use crossesline progeny for commercial production is the way in which their values compare with outbred or crossbred stock. Crossline progeny are
pigs produced by mating a boar from one inbred line to a female from another inbred line. Most comparisons between crossline progeny and outbreds have involved single crosses. The better single cross progeny usually have a faster rate of growth than outbred pigs. Some experimental comparisons between three-way cross progeny and outbred stock have been made. In swine, a three-way cross is customarily made by breeding an inbred boar to a female who is a single cross between two other lines.

Robison (1944) reported that mating boars of one line to single cross females of two other lines has resulted in more pigs saved per litter at market age, fewer runts, faster gains, and greater gains per unit of feed than resulted from mating unrelated outbred boars and sows.

On the basis of results with breed crosses of non-inbred stock, Winters et al. (1944) believes that crosses involving three or perhaps four lines will produce pigs which are better than either crossbred or outbred pigs.

Dickerson, Lush and Culbertson (1946) reported that crosses exceeded inbreds by 3.4 pounds at 56 days and by 25 pounds at 154 days. In total litter weight at 154 days, crosses exceeded inbreds by 290 pounds. These facts suggest that using the best of such improved lines in topcrossing on outbred sows or in triple crosses offers opportunity for surpassing the performance of outbreds.
Chambers, and Whatley (1950) found that three-line cross litters contained 1.7 more pigs and were 2.96 pounds heavier at 160 days than single cross litters. Heterosis was expressed to a greater extent in increased viability of the pigs and productivity of the crossline gilts than in the increased growth rate of the pig or its efficiency of gain.

Thus all of the experiments on or related to three-way crosses in swine indicate that litters from this type of mating will be larger and the pigs will be faster growing than is the case with outbred litters. The present study was to determine what factors cause the variation in 15th-day weight of three-way cross pigs. The multiple crosses which involved more than three lines were treated as three-way crosses by designating as a special line all maternal granddams which were themselves crosslines. The sires and maternal grandsires were always inbred boars. Thus all pigs had a sire from one line and a dam which was produced by crossing 2 other lines.

The specific objectives of the study were:

1. To obtain estimates of the general combining ability of the inbred lines, their general combining ability for maternal effects as shown by the single cross females, and their specific combining ability with respect to the performance of the single cross females.

2. To obtain information on the most probable 15th-day weight of progeny of males of each inbred line when mated to a random sample of single cross females from the same population of lines or when the males are used for to perscrossing.
3. To obtain information on the probable 15½-day weight of future progeny of specific three-way crosses which have been tested.

4. To obtain information on the probable 15½-day weight of progeny of specific three-way crosses which have not been tested, but could be made among the lines represented in the present study.
II. STATISTICAL METHODS

Because the numbers of litters produced by the different three-way crosses in different season-place groups were unequal, the method of least squares was used to estimate differences between the effects of the variables studied and to test hypotheses. To use this method it is necessary to formulate a mathematical model that describes the underlying biology reasonably and is amenable to statistical treatment.

The model used in this analysis was as follows:

\[
y_{ijklm} = \mu + e_i + \frac{1}{2}(e_j + e_k + m_j + m_k) + d_{jk} + t_{ijk} + b_1 + s_{ijkl} + r_{ijklm} + e_{ijklm}
\]

\(y_{ijklm}\) denotes the 154-day weight of the \(n\)th pig produced in the \(n\)th litter in the \(l\)th season-place group by mating a boar of the \(i\)th line to a gilt produced by crossing the \(j\)th and \(k\)th lines.

\(\mu\) indicates the effects which are alike for all pigs and hence do not cause any of the observed variation.

\(e_i\) is an effect common to all progeny of the parent belonging to the \(i\)th \((j, k)\) line. \(e_i\) is \(\frac{1}{2}\) the additively genetic (genic) value of line \(i\), as defined by Lush (1949). If the lines did not differ from each other in the \(y\)-values of the pigs their boars sired, all the \(g\)'s would be zero.

\(d_{jk}\) is an effect, in addition to \(\frac{1}{2}e_j\) \((e_k)\), which is common to all progeny that had line \(j(k)\) for a maternal grand parent. \(\frac{1}{2}d_{jk}\) includes both
However, some of the errors that affect the total value of the
outflow from the dam are not transferred from one to the next,
but rather transferred into the next season's record because
these errors increase the material value of the data for their
state and can, therefore, be transferred from one season to
the next. A similar procedure may be followed to determine
the material value of any other data, such as the error of
the precipitation and potential evaporation.
\( t_{ijk} \) is an interaction of line 1 with \( j \) and \( k \). That is, it is an effect, in addition to \( \alpha_i + \frac{1}{2}(\alpha_j + \alpha_k + \mu_j + \mu_k) + d_{jk} \), which is common to all progeny from sires of line 1 mated to single cross females from line \( j \) by line \( k \). \( t_{ijk} \) includes interactions between the sire lines and the separate dam lines or the single crosses between them.

\( b_1 \) is an effect common to all animals produced in the \( 1^{\text{st}} \) season-place group.

\( c_{ijkl} \) is an interaction between the \( ijk^{\text{th}} \) three-way cross and the \( l^{\text{th}} \) season-place group.

\( r_{ijklm} \) is an effect common to all the pigs in the \( n^{\text{th}} \) litter of the \( ijk^{\text{th}} \) three-way cross produced in the \( l^{\text{th}} \) season-place group, but not common to pigs in other litters of the same cross.

\( e_{ijklmn} \) is an effect peculiar to the 15\text{th}-day weight of the \( n^{\text{th}} \) pig of the \( ijkln^{\text{th}} \) litter.

The error term, which includes both \( r_{ijklm} \) and \( e_{ijklmn} \), is composed of a wide variety of factors. Factors that could cause differences among the litter effects include environmental differences that cause full-sib litters to differ, differences among sires from line \( i \), differences among the grandsires and granddams from line \( j \) and line \( k \), as well as differences among the litters because they were from different crossline \((j \times k)\) dams. Factors that could cause variation among the pigs within litters as well as between litters include chance in Mendelian sampling, effects of the genes that are not additive and are not included in \( d_{jk} \) or \( f_{jk} \), failure of the mathematical model to fit the biology of the data, and all unanalyzed environmental factors. The litter differences
could be correlated because about one half of the litters produced by
one of the sire lines in a particular season-place group would be by
one sire and the other half by another sire. Also, litter effects
could be correlated because they had the same maternal grandsire and in
some cases the same crossline dam. To ascertain the magnitude and
consequences of these correlations would be a very involved process and
was not attempted here. Random environmental effects as well as chance
in Mendelian sampling would tend to obscure these correlations. There-
fore, these correlations were assumed to be unimportant. The method of
least squares gives the best linear unbiased estimated of the effects,
when the errors are uncorrelated and distributed with mean = 0 and a
constant variance $\sigma^2$. Also, for the tests of significance to be entirely
valid, the errors must be normally distributed.

To get a clearer concept of what the different effects ($\mu$, $g_i$, $m_j$,
d$_{jk}$, and $t_{ijk}$) really represent, we will consider what means would approach
these effects in a hypothetical population where large and equal numbers of
progeny are produced in all the possible three-way crosses among many
lines. In this hypothetical population the different means approach the
following values:

---

1The expected correlation between two litters because they are by
the same sire is

$$1/n (1 + F) \frac{h^2}{\sqrt{1 + (n-1)t}} \sqrt{\frac{1}{1 + (n-1)t}},$$

where $F$ = inbreeding of the sire, $h^2$ = heritability, $n$ = number of
pigs in one litter, $m$ = number of pigs in the other litter, and $t$ =
phenotypic correlation between litter mates. The expected correlation
between two litters because they have the same as above except there
is an additional $1/n (1 + F')$ term, where $F'$ = inbreeding of the
grandsire. The effect of two litters having the same crossline dam
would be equal to having the same sire except there would be an
additional correlation because of common preweaning environment.
\[
\begin{align*}
\bar{y} & \quad \longrightarrow \mu \\
\bar{y}_i & \quad \longrightarrow \mu + \epsilon_i \\
\bar{y}_{.j} & \quad \longrightarrow \mu + \frac{1}{2}(\epsilon_j + m_j) \\
\bar{y}_{.jk} & \quad \longrightarrow \mu + \frac{1}{2}(\epsilon_j + \epsilon_k + m_j + m_k) + d_{jk} \\
\bar{y}_{ijk} & \quad \longrightarrow \mu + \epsilon_i + \frac{1}{2}(\epsilon_j + \epsilon_k + m_j + m_k) + d_{jk} + t_{ijk}
\end{align*}
\]

where a . in a subscript denotes summation over a given class, i.e.,

\[
\frac{1}{n} \sum y_{ij} = \bar{y}_j, \quad \frac{1}{n} \sum y_{ij} = \bar{y}_i.
\]

Therefore, \( \bar{y}_i \ldots \ldots \rightarrow \bar{y} \ldots \ldots \) approaches \( \epsilon_i, \) \( \bar{y}_{.j} \ldots \ldots - \frac{1}{2}(\bar{y}_{.j} \ldots \ldots + \bar{y} \ldots \ldots ) \) approaches \( \frac{1}{2}m_j, \) and \( \bar{y}_{.jk} \ldots \ldots - (\bar{y}_{.j} \ldots \ldots + \bar{y}_{.jk} \ldots \ldots) + \bar{y} \ldots \ldots \) approaches \( d_{jk}, \) \( \bar{y}_{ijk} \ldots \ldots - (\bar{y}_i \ldots \ldots + \bar{y}_{.jk} \ldots \ldots) + \bar{y} \ldots \ldots \) approaches \( t_{ijk}. \)

The factors that could make sample means differ from the above combination of effects include the following: sampling errors could make the sample means deviate from the true effect; unequal numbers in the different subclasses of a sample could make the means contain a complex function of the effects; and sampling from a population with unequal subclass numbers could make the means approach different values than those given above. The only way to reduce the influence of the first factor is to increase the size of sample, but the effect of the last two factors can be circumvented by using the method of least squares instead of the means to obtain estimates of the differences between the effects.

If we want the least squares estimates of the differences between the effects to be useful in making inferences about how these lines will perform in crosses with other lines still untested or in top crosses,
the lines from which we make these estimates must be a random sample
of lines from the population about which we are going to make inferences.
We will assume that the 12 inbred lines of swine at the Iowa Agricultural
Experiment Station are a random sample of lines that could have been
formed from the better purebred Poland China swine during the early
1930's.

The least squares method of estimating effects and testing signifi-
cance has been thoroughly discussed by Yates (1934), Hazel (1946),
Henderson (1956), Kempthorne (1950), and Koch (1950).

Using the model given above, with the assumptions that \( \mu, \varepsilon_1, m_j, \)
d\( t_{ijk}, c_{ijkl}, \varepsilon_{ijklm}, \) and \( b \) are unknown constants and the \( e_{ijklmn} \)
have expectation zero, common variance \( \sigma^2 \) and are uncorrelated, we
obtain the best linear unbiased estimates of these constants (effects)
by applying the method of least squares. The method of least squares is
to obtain the estimates which minimize the sum of the squares of devia-
tions, i.e.,

\[
\sum_{ijklmn}
\left[
Y_{ijklmn} - (\mu + \varepsilon_1 + m_j + t_{ijk} + c_{ijkl} + d_{jk} + t_{ijk} + b + c_{ijkl})
\right]^2
\]

The values for \( \mu, \varepsilon_1, m_j, d_{jk}, t_{ijk}, b, c_{ijkl} \) which minimize
this sum of squares are found by solving the set of simultaneous
equations that are obtained by taking the partial differential of the
sum of squares with respect to each of the constants and setting each
derivative \( = 0 \). The equations obtained in this manner are conventionally
called the normal equations. The normal equations for the above model
contain so many unknowns that it would be very difficult to solve them.
estimate of \( \bar{x} \), where \( \bar{x} \) is the mean of all these.

that parameter decreases from the mean of the parameter, i.e., \( \bar{x} \) is an
are imposed, the estimate obtained for a parameter is an estimate of how
the above restriction that that estimate of the parameter sum to zero
explicit numbers In each subseries will use the above restrictions.
Since the least way to compare the three-way crosses would be to have
parameters for a population which have equal numbers In all subseries.

least square estimate of the parameters are

\[
\hat{\beta} = \sum_{i} \hat{\beta}_i = 0, \quad \text{for all } i
\]

\[
\sum_{i} \hat{\beta}_i = 0, \quad \text{for all } i
\]

The restrictions that these three-way crosses would be to have

regardless of which three-way crosses are applied.

As the effects (or parameters) and also the double interaction terms do to any
parameters (or parameters) and also the double interaction terms do to any

\( \hat{\beta}_i \) have pointed out that the estimate of the double interaction terms do to any

Take note to make them independent, but keep comparing (1940) p. 6 and Koch (1940).

can be solved, a number of restrictions could be applied to the data.

Therefore, some restrictions must be put on the data in the data before they

Three times the number of unknown parameters which are to be estimated

In all cases the number of independent normal equations will be

For the different parameters are given on the following pages.

The normal equations that were used to solve

parameters included in it can be found by taking the different combinations

estimates are obtained for the joint parameters, are

\( \hat{\beta}_i \) can be combined into one parameter, say \( \hat{\beta}_i \), when

To lessen the difficulties, several of the parameters in the model can
The methods used to solve the normal equations are primarily those given by Henderson (1938) and Kempthorne (1950). The tests of hypotheses that will be discussed here are those that will be used in the next section.

To test, first, the hypotheses that $c_{ijkl} = 0$, $t_{ijk} = 0$, and $b_1 = 0$, we will need several sum of squares. $R(\cdot)$ will be used to denote the reduction in sum of squares of $y$ due to fitting the parameters shown in the parentheses.

To obtain $R(\mu, g, m, d, t, b)$, we will use the following model:

$$y_{ijklmn} = \tilde{t}_{ijk} + b_1 + \tilde{r}_{ijklm} + \tilde{e}_{ijklmn}$$

where $\tilde{t}_{ijk} = \mu + g_i + \frac{1}{2}(s_j + s_k + m_j + m_k) + d_{jk} + t_{ijk}$. $e_{ijkl}$ was not included in this model because the sum of squares for the interaction between three-way crosses and season-place groups was found by considering the difference between $R(\mu, g, m, d, t, b)$ and the sum of squares among the three-way cross, season-place subclasses.
Using the above model, the normal equations are:

\[ t_{ijk} \cdot n_{ijkl} + \sum_{1}^{n_{ijkl}} b_{1} = y_{ijk} \]

\[ b_{1} = \sum_{ijkl} n_{ijkl} \cdot t_{ijkl} + n_{ijkl} \cdot b_{1} = y_{ijkl} \]

\( n_{ijkl} \) represents the number of pigs in a particular litter and as before, in the subscript means summation over a given classification, i.e.,

\[ n_{ijkl} = \sum_{m} n_{ijklm} \]. There will be one fewer independent equation than there are unknowns in the equations because the sum of the \( t_{ijkl} \) equations = sum of the \( b_{1} \) equations. One method of solving these equations is to absorb the \( t_{ijkl} \) equations into the \( b_{1} \) equations using the method given by Henderson (1948), strike out the row and column of one of the \( b_{1} \) equations after the \( t_{ijkl} \) equations are absorbed, and solve the remaining equations by iteration. By striking out the row and column of one of the \( b_{1} \) equations, say the \( b_{a} \) equation, the estimates of the \( b \)'s (which will be denoted by \( b' \)) that are obtained by iteration are estimates of the difference between the effect of \( b_{1} \) and \( b_{a} \), i.e., estimates of \( b_{1} - b_{a} \). Then

\[ R(\mu, g, s, d, t, b) = \sum_{ijkl} \frac{(y_{ijkl} - \ldots)^{2}}{n_{ijkl}} + \sum_{1}^{n_{ijkl}} b_{1} \cdot B \ldots 1 \ldots \] where \( B \ldots \ldots \) is the right hand side of the \( b_{1} \) equation after the \( t_{ijkl} \) equations have been absorbed. This reduction in sum of squares is conveniently written as

\[ \sum_{ijkl} t_{ijkl} \cdot y_{ijkl} + \sum_{1}^{n_{ijkl}} b_{1} \cdot y_{ijkl} \ldots 1 \ldots \] However, the first method given is less affected by sampling errors and is easier to calculate. The additional sum of squares due to season-place groups is \( \sum_{1}^{n_{ijkl}} b_{1} \cdot B \ldots 1 \ldots \)

The sum of squares among litters within three-way cross, season-place subclasses is

\[ \sum_{ijkl} \left[ \frac{(y_{ijkl}^{2})}{n} - \frac{(y_{ijkl} \ldots)^{2}}{n_{ijkl}} \cdot \frac{n_{ijkl}}{n} \right] \], which will
be designated by ALSS. The sum of squares within litters is
\[ \sum (y_{ijklm})^2 - \sum (\bar{y}_{ijkl})^2, \]
which will be represented by \( W \). The sum of squares within three-way cross, season-place subclasses is \( = W + \text{ALSS}. \)

Using the above calculations the analysis of variance table can be set up as shown in Table 1.

If there are no real differences among litters within three-way cross, season-place subclasses, then \( H/S \) will be distributed as \( F \) with \((h - s) \) and \((n - h) \) degrees of freedom. Therefore, to test the hypothesis that there are no litter differences within subclasses we compare the value of \( H/S \) to the value of \( F \), with corresponding degrees of freedom, given in an \( F \) table.

If the hypothesis that there are no litter differences within subclasses is rejected, the statistical significance of variance among the other classifications cannot be tested exactly. An approximate test can be made by comparing the mean square for each classification to \( H \). The accuracy of this approximation depends on how closely the coefficient for \( \sigma^2_r \) in the expectation of \( H \) approaches the coefficient for \( \sigma^2_r \) in the expectation of the other mean squares, if it is assumed that the \( r_{ijklm} \)'s are normally and independently distributed with mean 0 and variance \( \sigma^2_r \).

If there were many lines in the test, determination of the coefficient for \( \sigma^2_r \) in any of the means squares except \( H \) would require inverting large matrices of coefficients in the normal equations. However, in several small samples the coefficient for \( \sigma^2_r \) was found to be about the same in the expectation of \( H \) as in the expectation of mean squares for different classifications ignoring the litter classification. Therefore,
Table I

Outline of the computational procedure for obtaining certain mean squares

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among B</td>
<td>f-1</td>
<td>( \sum_{b}^{1} b ) ...1..</td>
<td>B</td>
</tr>
<tr>
<td>Among T</td>
<td>( \frac{1}{2}(s-1)(s-2)r-s )</td>
<td>( R(\mu, \sigma, \alpha, d, t, b) - R(\mu, \sigma, m, d, b) )</td>
<td>T</td>
</tr>
<tr>
<td>Among C</td>
<td>( z - \frac{1}{2}r(s-1)(s-2) + a - f + 1 )</td>
<td>( \sum(y_{ijklmn})^2 - (w + ALSS) - R(\mu, \sigma, m, d, t, b) )</td>
<td>I</td>
</tr>
<tr>
<td>Among litters within three-way cross, season-place subclasses</td>
<td>h - z</td>
<td>ALSS</td>
<td>E</td>
</tr>
<tr>
<td>Within litters</td>
<td>a - h</td>
<td>W</td>
<td>S</td>
</tr>
</tbody>
</table>

where:

- \( a \) = the number of possible three-way crosses among \( r \) and \( s \) lines that are not represented in the data.
- \( f \) = the number of season-place groups.
- \( h \) = the number of litters.
- \( z \) = total number of observations.
- \( r \) = the number of sire lines.
- \( s \) = the number of lines that were used to produce single cross females.
- \( u \) = the number of possible types of single cross female, which could be made from \( a \) lines, that are not represented in the data.
- \( z \) = the number of three-way cross, season-place subclasses.

Each of the symbols in the column under Sum of Squares, except \( R(\mu, \sigma, \alpha, d, b) \), is explained in the text before the table, and the meaning of \( R(\mu, \sigma, m, d, b) \) is explained in the part of the text which follows after a discussion of the table.
in making the approximate tests of significance it was assumed that
the coefficient of $\sigma^2_r$ in the mean squares were not greatly different.
All further calculations in this section will be made on the basis that
the hypothesis, $\tau_{ijklm} = 0$, was rejected, as it actually turned out to be.

Variation among $O(\sigma_{ijkl})$ values) includes only the additional varia-
tion among three-way crosses and season-place groups beyond that from
the general differences among three-way crosses and among season-place
groups. $I/E$ can be compared to the tabulated $F$ value, with corresponding
degrees of freedom to get an approximate test of the hypothesis that there is
no interaction of three-way crosses with season-place groups.

Variation among $T$ is that caused by the interaction between sire
lines and the single crosses of females. That is, it is the variation
left among the three-way crosses after subtracting that accounted for by
general differences among the sire lines and among the single crosses of
females. If the hypothesis that there is no interaction between three-way
crosses and season-place groups is accepted, the hypothesis that $t_{ijk} = 0$,
i.e., that there is no interaction between the sire lines and the single
cross females, can be tested approximately by comparing $T/E$ to the tabu-
lated $F$ value, with corresponding degrees of freedom.

Variation among $B$ includes only the additional variation accountable
to the season-place groups after the three-way crosses have accounted
for all the variation they can. If the hypothesis that there is no inter-
action between three-way crosses and season-place groups is accepted,
an approximate test of the hypothesis that $b_{l} = 0$, is given by comparing $B/E$
to the corresponding value in an $F$ table.
All further calculations here will be made on the basis that the hypotheses \( c_{ijkl} = 0 \) and \( t_{ijk} = 0 \) were accepted, as the evidence to be presented later shows that they may.

To obtain \( R(\mu, \varepsilon, m, d_i, b) \) we will use the following model:

\[
Y_{ijklmn} = g_i + \frac{1}{2}(e_j + e_k) + d_{jk} + b_1 + r_{ijklm} + a_{ijklmn}
\]

where \( d_{jk} = \mu + \frac{1}{2}m_j + \frac{1}{2}m_k + d_{jk} \). Using this model, the normal equations are:

\[
s_v: \quad (n_{v...} + 1/4n_{v...}) \varepsilon_v + \sum_q (\frac{1}{2}n_{vq...} + \frac{1}{4}n_{qv...} + n_{.vq...}) \varepsilon_q +
\]

\[
= \sum_q \frac{1}{2}n_{vq...} \varepsilon_{vq...} + \sum_{qp} n_{vqp} \varepsilon_{qp} + \frac{1}{4} \sum (n_{v..1} + \frac{1}{2}n_{..v1})b_1 =
\]

\[
Y_{v...} + \frac{1}{2}Y_{v...}
\]

\[
d_{vw}: \quad \sum_q n_{qvw} \varepsilon_q + n_{.vw} (\frac{1}{2}e_v + \frac{1}{2}e_w) + n_{.vw} d_{vw} + \frac{1}{2} n_{.vwl} b_1 = Y_{vw}.
\]

\[
b_1: \quad \sum_q (n_{q..1} + n_{q..1}) \varepsilon_q + \frac{1}{2} n_{.qpl} d_{qp} + n_{...1} b_1 = Y_{...1}.
\]

By using Henderson's (1943, p. 22) method, the coefficients of the \( d_{vw} \)'s and their associated sums can be absorbed into the other equations. To absorb another set of equations, after the \( d_{vw} \) coefficients are absorbed, it is necessary to get an inverse matrix and perform two matrix multiplications. To illustrate the method, let us consider the following:
This is a set of normal equations written in matrix notation with 3 g unknowns and 2 b unknowns. To obtain the first normal equation, multiply, element by element, the first row of S by the column B as follows:

\[ c_{11} \mathbf{s}_1 + c_{12} \mathbf{s}_2 + c_{13} \mathbf{s}_3 + c_{11} b_1 + c_{12} b_2 \text{ and set it } = \text{ to } y_1. \]

This type of multiplication is called matrix multiplication. A discussion of this multiplication can be found in Aitken (1948). To obtain the second equation, matrix multiply the second row of S by B and set \( = y_2 \). Obtain the third equation by matrix multiplication of the third row of S by B and setting that \( = y_3 \), etc.

The S matrix of coefficients is used to represent a group of equations we would have after absorbing the \( \frac{m}{n} \) equations into the g and b equations. If the original matrix were symmetrical, the matrix that is formed by absorption will be symmetrical. The above S matrix is symmetrical. Also, the sum of each row and of each column for any classification in the S matrix is equal to zero.

To proceed further, we need to use two other properties of matrix multiplication, i.e.,
\[ S^{-1} \cdot S = I \]

\[
I = \begin{pmatrix}
100000 \\
010000 \\
001000 \\
000100 \\
000010 \\
000001
\end{pmatrix}
\]

where \( S^{-1} \) is the inverse matrix of the \( S \) matrix and \( I \) is the unit matrix. Therefore, we can solve for the unknowns by matrix multiplication as follows:

\[ S^{-1} \cdot S \cdot B = S^{-1} \cdot Y \]

\[ I \cdot B = S^{-1} \cdot Y \]

\[ B = S^{-1} \cdot Y \]

However, we shall work with only part of the matrix at a time using a method suggested by Kempthorne (1950). Let us move the \( g \) coefficients of the last two equations to the right hand side, i.e.,

\[
\begin{pmatrix}
d_{11} & d_{12} \\
d_{12} & d_{22}
\end{pmatrix}
\begin{pmatrix}
b_1 \\
b_2
\end{pmatrix}
= \begin{pmatrix}
y_{1} - (c_{11} \xi_1 + c_{21} \xi_2 + c_{31} \xi_3) \\
y_{2} - (c_{12} \xi_1 + c_{22} \xi_2 + c_{32} \xi_3)
\end{pmatrix}
\]

(2)

The \( d \)'s sum to zero so we must make some restriction on the \( b \)'s to make our equations independent. If we make the restriction \( b_1 + b_2 = 0 \), we can use Henderson's (1948, p. 32) method to obtain an inverse of the \( d \) matrix, which will be denoted as

\[
\begin{pmatrix}
d_{11}^{11} & d_{12}^{12} \\
d_{12}^{12} & d_{22}^{22}
\end{pmatrix}
\]

Then, using matrix premultiplication, we have
\[
\begin{align*}
\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} &= \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix} \begin{pmatrix} y_1 - (c_{11} \varepsilon_1 + c_{21} \varepsilon_2 + c_{31} \varepsilon_3) \\ y_2 - (c_{12} \varepsilon_1 + c_{22} \varepsilon_2 + c_{32} \varepsilon_3) \end{pmatrix} \\
&= d_{11} \left[ y_1 - (c_{11} \varepsilon_1 + c_{21} \varepsilon_2 + c_{31} \varepsilon_3) \right] + d_{12} \left[ y_2 - (c_{12} \varepsilon_1 + c_{22} \varepsilon_2 + c_{32} \varepsilon_3) \right] \\
&= d_{12} \left[ y_1 - (c_{11} \varepsilon_1 + c_{21} \varepsilon_2 + c_{31} \varepsilon_3) \right] + d_{22} \left[ y_2 - (c_{12} \varepsilon_1 + c_{22} \varepsilon_2 + c_{32} \varepsilon_3) \right]
\end{align*}
\]

(3)

If we substitute these values for \( b_1 \) and \( b_2 \) in our first normal equation of (1), we obtain

\[
\begin{align*}
\{ \varepsilon_1 - \left[ c_{11}(d_{11} c_{11} + d_{12} c_{12}) + c_{12}(d_{12} c_{11} + d_{22} c_{12}) \right] \} & \ v_1 + \\
\{ \varepsilon_2 - \left[ c_{11}(d_{11} c_{21} + d_{12} c_{22}) + c_{12}(d_{12} c_{21} + d_{22} c_{22}) \right] \} & \ v_2 + \\
\{ \varepsilon_3 - \left[ c_{11}(d_{11} c_{31} + d_{12} c_{32}) + c_{12}(d_{12} c_{31} + d_{22} c_{32}) \right] \} & \ v_3 \\
&= y_1 - (d_{11} y_1 + d_{12} y_2 + d_{12} y_1 + d_{22} y_2),
\end{align*}
\]

Similar substitutions can be made in the next two normal equations to obtain three equations with only the three \( g \) unknowns.

By using matrix notation all the way through we can present this much more neatly. Let us symbolize the normal equations by

\[
\begin{pmatrix} S \ar{c} B \ar{d} Y \end{pmatrix} \quad \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}
\]

Then we represent the above procedure by
db = (I₂ - c'g), which is matrix notation for (2).

b = d⁻¹(I₂ - c'g), which is the same as (3).

(a - cd⁻¹c')g = (I₁ - cd⁻¹Y₂)

We can check for arithmetic mistakes in this procedure by observing that the sum of the rows and the sum of the columns of d⁻¹c', cd⁻¹c', and (a - cd⁻¹c') should equal to zero and the sum of the columns d⁻¹Y₂, cd⁻¹Y₂, and (I₁ - cd⁻¹Y₂) should equal zero. Also, the matrix cd⁻¹c' is symmetrical. The product of the last step of this method may be represented by

<table>
<thead>
<tr>
<th>A</th>
<th>g</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁₁ A₁₂ A₁₃</td>
<td>g₁</td>
<td>q₁</td>
</tr>
<tr>
<td>A₁₂ A₂₂ A₂₃</td>
<td>g₂</td>
<td>q₂</td>
</tr>
<tr>
<td>A₁₃ A₂₃ A₃₃</td>
<td>g₃</td>
<td>q₃</td>
</tr>
</tbody>
</table>

where A₁₁ represents a₁₁ = [c₁₁(d⁻¹c₁₁ + d⁻¹c₁₂) + c₁₂(d⁻¹c₁₁ + d⁻¹c₁₂)]

which is, also, the first row and first column entry of (a - cd⁻¹c'),

and Q₁ = Y₁ - (d⁻¹₁ Y₁₁ + d⁻¹₂ Y₁₂ + d⁻¹₂ Y₁₂) and is, also, the first entry in the column (I₁ - cd⁻¹Y₂).

Again the coefficients in the A matrix sum to zero so we must put a restriction on the equations if we are to obtain an inverse. If we impose the restriction that g₁ + g₂ + g₃ = 0, we can use the same method for obtaining the inverse of A as we did for obtaining the inverse of d.

Let the inverse of A = A⁻¹. Then we can solve for g by:
\[ A^{-1}a_g = A^{-1}q, \quad e = A^{-1}q, \quad \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} A^{11} & A^{12} & A^{13} \\ A^{12} & A^{22} & A^{23} \\ A^{13} & A^{23} & A^{33} \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \]

If by solving the equations we find that \( \hat{e}_1 = -1, \hat{e}_2 = 0, \) and \( \hat{e}_3 = 1, \)

we can substitute the numerical values for \( e_1, e_2, \) and \( e_3 \) in \( A^{-1}(q_2 - q_1^g) \)

to obtain numerical values for \( b_1 \) and \( b_2. \) These in turn can be used along

with the numerical values of the \( g_i \)'s, to substitute in the \( d_{wv} \) equations,

which were already absorbed in this example to obtain numerical estimates

of the \( d_{wv} \)'s.

However, \( A^{-1} \) has a more important use than solving the equations, namely it can be used to get the variance and covariances of the \( g_i \)'s, i.e., the variance of \( \hat{e}_1 \) is \( A^{11} \sigma^2 \) and the covariance of \( \hat{e}_1 \) and \( \hat{e}_2 \) is \( A^{12} \sigma^2. \) Estimates of these variances are obtained by replacing \( \sigma^2 \) by \( E, \) the observed error mean square. We may also use \( A^{-1} \) to determine the

variance of linear functions of the \( g_i \)'s, i.e., the estimated variance

of \( (\hat{e}_1 - \hat{e}_2) = (A^{11} + A^{22} - 2A^{12}) E. \)

Let us now consider how to obtain \( E, \) error mean square, and other mean squares needed in test of significances. \( R( \quad ) \) will be used to denote the reduction in sum of squares due to fitting the parameters shown

in the parentheses.

Using the notation of the normal equations with the joint parameter

\( d \) (page 18), \( R(\mu, e, m, d, b) = \)
\[
\frac{(y_{ab..})^2}{n_{ab..}} + \frac{(y_{ac..})^2}{n_{ac..}} + \ldots + \frac{(y_{sx..})^2}{n_{sx..}} + \\
\frac{b_1 y'_1}{n_{y'_1}} + \frac{b_2 y'_2}{n_{y'_2}} + \ldots + \frac{b_g y'_g}{n_{y'_g}} + \varepsilon_{g} y'_a + \ldots + \varepsilon_{b} y'_b + \\
\ldots + \varepsilon_{x} y'_x, \quad \text{where } y'_1 \ldots 1 \text{ and } y'_v \ldots \text{ are the right hand side of the } b_1 \text{ equation and the } g_v \text{ equation respectively, after the } d_{vw} \text{ equations have been absorbed. The degrees of freedom associated with this sum of squares are } r + \frac{s(s-1)}{2} - u + f - 2, \text{ where the symbols are defined in Table I. The sources of these degrees of freedom are from the following classification: } (r - 1) \text{ from the } g's, \ (s - 1) \text{ from the } m's, \\
\frac{s(s-1)}{2} - s + u \text{ from the } d's, \ (f - 1) \text{ from the } b's \text{ and } l \text{ from } \mu. \]

Also, the among sum of squares = \( \varepsilon_a Q_a + \varepsilon_b Q_b + \ldots + \varepsilon_x Q_x \), with \( r - 1 \) degrees of freedom.

If some of the lines are used to produce the crossline dams but not as sire lines, it is not possible to separate the maternal and genic effects of these lines. In the above absorption procedure any effect of a line used only to produce crossline dams is included in the \( d_{vw} \) parameters for that line and all the coefficients for the \( g_v \) equation of that line go to zero. This drops that \( g_v \) equation and that \( g_v \) out of any further calculations.

To obtain \( R(\mu, g, m, b) \), we will use the model: \( y_{ijklm} = \mu + \frac{1}{2}(g_j + g_k + m_j + m_k) + b_1 + \varepsilon_{ijklm} \), where \( b_1 = \mu + b_1 \).

The normal equations, using this model, are:
\( \varepsilon_v: (n_v + 1/\ln n_v) \varepsilon_v + \sum_q (\rho_{n_v} + \delta_{n_v} + 1/\ln n_v . \rho_{n_v}) \varepsilon_q + \\
1/\ln n_v . m_v + \sum_q (\beta_{n_v} + 1/\ln n_v) m_q + \sum_q (n_v + 1/\ln n_v . \rho_{n_v}) h_1 = \\
y_v + \frac{1}{2} y_v . \\
m_v: \frac{1}{2} \left[ n_v \varepsilon_v + \frac{1}{2} m_v \right] + \sum_q n_{n_v} \varepsilon_q + \sum_q n_{n_v} (\frac{1}{2} \varepsilon_q + \frac{1}{2} m_q) + \\
1/\ln n_v . l \bar{b}_1 = y_v . \\
\bar{b}_1: \sum_q (n_v + 1/\ln n_v . \rho_{n_v}) \varepsilon_q + \sum_q n_v (\frac{1}{2} \varepsilon_q + \frac{1}{2} m_q) + n_v \bar{b}_1 = y_v . \\

The \bar{b}_1 equations can be absorbed so the remaining equations are in terms of \( \varepsilon_i \)'s and \( m_j \)'s. In these data, line \( X \) was not used as a male line. Therefore, we cannot get separate estimates of \( m_x \) and \( \varepsilon_x \). The \( \varepsilon_x \) and the \( m_x \) equations are exactly the same. Also the \( \varepsilon_v \) equations = \( \sum_v \varepsilon_v \) equations = 0. To make the equations independent, we can draw a line through the row and column for \( \varepsilon_k \), \( m_k \), and \( m_x \) equations. However, we are thereby estimating different parameter, i.e., \( \nabla \varepsilon_i \) is an estimate of \( \varepsilon_i - \varepsilon_k \), \( \nabla m_j \) is an estimate of \( m_j - m_k \), and \( x \) is an estimate of \( x - \varepsilon_k - m_k \). \( x \) is used to denote the combined effect of \( \varepsilon_x \) and \( m_x \), which we cannot separate.

\[
\kappa(\mathcal{U}, \varepsilon, m, b) = \sum_{l=1}^{n} \frac{(y_{l-1})^2}{n_{l-1}} + \varepsilon_{a}(y_{a,...} + \frac{1}{2} y_{a,...}) + \varepsilon_{b}(y_{b,...} + \frac{1}{2} y_{b,...}) + \varepsilon_{c}(y_{c,...} + \frac{1}{2} y_{c,...}) + \varepsilon_{d}(y_{d,...} + \frac{1}{2} y_{d,...}) + \varepsilon_{e}(y_{e,...} + \frac{1}{2} y_{e,...}) + \varepsilon_{f}(y_{f,...} + \frac{1}{2} y_{f,...}) + \varepsilon_{g}(y_{g,...} + \frac{1}{2} y_{g,...}) + \varepsilon_{h}(y_{h,...} + \frac{1}{2} y_{h,...}) + \varepsilon_{i}(y_{i,...} + \frac{1}{2} y_{i,...}) + \varepsilon_{j}(y_{j,...} + \frac{1}{2} y_{j,...}) + \varepsilon_{k}(y_{k,...} + \frac{1}{2} y_{k,...}) + \varepsilon_{l}(y_{l,...} + \frac{1}{2} y_{l,...}) + \varepsilon_{m}(y_{m,...} + \frac{1}{2} y_{m,...}) + \varepsilon_{n}(y_{n,...} + \frac{1}{2} y_{n,...}) + \varepsilon_{o}(y_{o,...} + \frac{1}{2} y_{o,...}) + \varepsilon_{p}(y_{p,...} + \frac{1}{2} y_{p,...}) + \varepsilon_{q}(y_{q,...} + \frac{1}{2} y_{q,...}) + \varepsilon_{r}(y_{r,...} + \frac{1}{2} y_{r,...}) + \varepsilon_{s}(y_{s,...} + \frac{1}{2} y_{s,...}) + \varepsilon_{t}(y_{t,...} + \frac{1}{2} y_{t,...}) + \varepsilon_{u}(y_{u,...} + \frac{1}{2} y_{u,...}) + \varepsilon_{v}(y_{v,...} + \frac{1}{2} y_{v,...}) + \varepsilon_{w}(y_{w,...} + \frac{1}{2} y_{w,...}) + \varepsilon_{x}(y_{x,...} + \frac{1}{2} y_{x,...}) + \varepsilon_{y}(y_{y,...} + \frac{1}{2} y_{y,...}) + \varepsilon_{z}(y_{z,...} + \frac{1}{2} y_{z,...})
\]
Partition and squares that are in the column under degrees of freedom are denoted:

\[ x \text{ mean squares} \]
\[ S = \text{sum of squares} \]
\[ d.f. = \text{degrees of freedom} \]

\[ x - \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]

\[ (x - \bar{x})^2 \text{ for } x \text{ and } \bar{x} \]

\[ S = \text{sum of squares} \]

\[ d.f. = \text{degrees of freedom} \]

\[ x \text{ for } x \text{ and } \bar{x} \]

\[ \bar{x} \text{ for } x \text{ and } \bar{x} \]
An approximate test of the hypothesis that \( s_1 = 0 \) is given by comparing \( G/N \) to the corresponding value in an \( F \) table.

The factors that determine the accuracy of these approximate tests were discussed on page 15.

The procedure for making any other tests that were applied to these data will be discussed briefly in the section on analysis of experimental results.

The details of the above statistical procedure were worked out using a small set of data composed in the following way. Numerical values for \( s_1 \), \( m_j \), and \( d_{jk} \) and \( b_1 \) were assigned to four lines and to two season-place groups. Three-way crosses among these 4 lines were made up to include all the crosses possible, with the restrictions that one of the lines was not used as a sire line and one possible type of crossline dam was omitted. Unequal numbers of pigs were assigned to the different crosses. The weight of each of the pigs in a particular cross at a particular time was determined by taking the sum of \( s_1 \) for the sire line, \( \frac{3}{2}(s_j + s_k + m_j + m_k) + d_{jk} \) for the two dam lines, and \( b_1 \) for the season-place group. The method of least squares was applied to the data formed in this way. The differences among the least squares estimates were the same, except for small rounding errors, as the differences among the values that were used to form the data.
III. ANALYSIS OF EXPERIMENTAL RESULTS

A. Source and Scope of the Data

The data analyzed in this study are the three-way cross pigs produced by crossing the twelve inbred lines of Poland China swine maintained by the Iowa Agricultural Experiment Station. The first of these lines (Scoutmaster) was started in 1930; the others were started in 1937 or 1938. Some of the latter lines were formed by branching off from the Scoutmaster line, while others were formed from animals which were not related to those in the Scoutmaster and related lines. The Scoutmaster has been maintained as a four-sire line. The others are one-sire or two-sire lines. An extensive history of the lines is given by Vernon (1948).

In the fall of 1942, all single cross litters possible were produced, within the limits of the physical facilities and availability of females above the requirements for maintaining the lines. In each of the following falls, as many crossline litters (single crosses, three-way crosses, and multiple crosses) as possible were produced. In making the crosses, the general procedure was to select two boars from each line. These were mated to their proportional part of the females available for crossing. The purpose of the crossing program was as follows:

1. To compare the performance of single cross and multiple cross (including three-way cross) pigs, and compare the performance of each to inbred pigs.
2. To measure the general combining ability of the lines in crosses.
3. To measure the maternal ability of the lines.
4. To obtain information on the value of specific crosses.

Until 1948, all the inbred and crossline animals were maintained at the farm near Napier. Starting with the spring of 1948, the facilities at the Ankeny farm were used to produce crossline litters in both the spring and the fall season. The rations at the two farms are slightly different, but the management practices differ only to the extent that the farm personnel methods differ.

When a group of pigs reach weaning age, most of them are put together in a pen that contains some pasture and are fed a growing-fattening ration. The grouping is done according to age without regard to pedigree. From some litters, two or four representative pigs were selected to go on "record of performance" tests. These pigs are fed a fattening ration in concrete-floored pens with four pigs per pen.

All the pigs produced by crossing three or more lines from the fall of 1944 to the fall of 1949, inclusive, are included in this study. Crosses which involved more than three lines were included in this study and classified as three-way crosses according to sire and maternal grandsire. In all crosses involving more than three lines, the sire and the maternal grandsire were inbreds but the maternal granddam was always a cross of two or more lines. Since any of the lines represented in the crossline maternal granddam of a litter will, on the average, contribute 1/8 or less of the total number of genes any pig in the litter receives, the differences between the various crosses in the maternal granddam would have but small effects on the pigs. Consequently all crossbred
maternal granddams were denoted as merely belonging to line X.

Of the 78 possible types of crossbred dams, 43 were represented in this study. One hundred ninety-seven of the 802 possible three-way crosses were represented. The number of pigs in each cross varied from one to 41.

B. Factors Considered in the Analysis

The 15\(\frac{1}{4}\)-day weight of individual pigs was the object of this analysis. The weight of the litters would have been more easily analyzed and, at a slightly older age, is the primary characteristic which interests the commercial producer. However, the differences in number of pigs in the litter tend to obscure the differences in weight of the individual pigs. Also, when a trait is composed of two parts, more progress can sometimes be made by considering the parts separately than by considering the trait as a whole. Therefore, the weights of the individual pigs were used in this study. As it was not practical to weigh all the pigs when they were exactly 15\(\frac{1}{4}\) days old, the 15\(\frac{1}{4}\)-day weight of each pig was obtained by weighing the pigs twice, once within two weeks before and once within two weeks after 15\(\frac{1}{4}\) days. Each weight was corrected to 15\(\frac{1}{4}\) days, and these two corrected weights were averaged. This averaging of weights taken on different days doubtless reduced somewhat the errors of weighing and of temporary conditions, such as "fill", which are in single weights.

The three-way cross pigs were raised in five different places - Section A, Section B, and the record of performance (R.O.P.) at the
farm near Napier; the farm on Beech Avenue in the edge of Ames; and
the farm near Ankeny. To those in charge, it seemed that the pigs did
better at one place one season and at another place the next season.
Therefore, season-place groups were used to correct simultaneously for
seasons, places, and season-place interaction. The seasons were 1944
spring, 1944 fall, 1945 spring, ..., 1949 fall. There were nineteen
of these season-place groups.

No correction was made for age of the dams because all females
were about one year old, except that only 1½ year old sows were used at
Ankeny in the spring of 1948 and 1949. Thus, any differences due to
the dams being gilts or sows were included in the constants fitted for
season-place groups.

Since all the dams were crosslines, no corrections were needed for
inbreeding.

C. Estimation of Effects

The tests of significance in the next section indicate that the
hypothesis, \( d_{jk} = 0 \) is acceptable. However, the least squares estimates
of the values of \( d_{jk} \) are better estimates than zero. Therefore the
model \( y_{ijklm} = \mu + \frac{1}{2} \gamma_j + \frac{1}{2} \gamma_k + d_{jk} + b_1 + r_{ijklm} + e_{ijklm} \) was used to
obtain estimates of the different effects, where \( d_{jk} = \mu + \frac{1}{2} \gamma_j + \frac{1}{2} \gamma_k \).
For the least squares estimates to be the best linear unbiased estimates
of the effects the errors, both \( r_{ijklm} \) and \( e_{ijklm} \) must be uncorrelated and
distributed with mean = 0 and constant variance \( \sigma^2_r \) and \( \sigma^2_e \), respectively.
\( t_{ijk} \) was left out of the model because the test of significance indicates
that the hypothesis, \( t_{ijk} = 0 \), is acceptable, and the estimates of many of the \( t_{ijk} \)'s would be based on so few pigs that the sampling errors associated with the estimates would be very large. \( a_{ijkl} \) was left out of the model because the mean square associated with this classification was about the same as the mean square among litters within three-way cross, season-place subclasses.

The method of obtaining estimates of the parameters in the above model is given on page 18. Using this method, the following estimates of the \( g \)'s were obtained:

\[
\hat{\xi}_a = 1.8, \quad \hat{\xi}_b = -3.5, \quad \hat{\xi}_c = -2.3, \quad \hat{\xi}_d = 1.3, \quad \hat{\xi}_e = -3.7, \quad \hat{\xi}_f = -4.4, \\
\hat{\xi}_g = 10.3, \quad \hat{\xi}_h = 12.1, \quad \hat{\xi}_i = -2.9, \quad \hat{\xi}_j = 5.4, \quad \hat{\xi}_k = -5.5, \quad \hat{\xi}_l = -2.4.
\]

The subscripts for the \( g \)'s represent the inbred lines used in this study, i.e., \( \text{a, b, c, d, e, f, g, h, i, j, k} \) and \( s \) represent the lines Alliance, Blackbird, Charmer, Defender, Enterprise, Fearless, Grandee, Hawkeye, Informer, Jumbo, Kickoff, and Scoutmaster, respectively.

The \( g \)'s estimated in this case, or in any case where both \( g \) and \( m \) are in the model, are determined by how the lines perform as sire lines. The difference between how the lines perform as sire lines and as dam lines is included in the estimates of \( m \). The estimated variances and covariances associated with these estimates are given in Table 2. The
value 1453 in Table 2 is the mean square for variation among litters within three-way cross, season-place subclasses. This estimate of the error variance is only approximately correct for testing significance of other mean squares because the coefficient of $\sigma^2_r$ in mean square among litters within subclasses may be slightly different from what it is in the mean squares for which this is used as an error term. From Table 2 the variance of any estimate or any linear combination of the estimates is readily obtained, i.e., the estimated variance of $\widehat{\xi}_A$ is $(.0048) (1453) = 7.0$ and the estimated variance of $(\widehat{\xi}_g - \widehat{\xi}_c)$ is 

\[ (.0048 + .0062 - 2(.0007)) (1453) = 13.9. \]

From his work with crossbreeding and farm testing, Dr. Hazel (1950) stated that he considered line G to be one of the best and line C one of the poorest lines for obtaining rapid growth rate in crossbred litters. To test the hypothesis that $\xi_g = \xi_c$, we compare $\frac{10.3 - (-2.3)}{\sqrt{13.9}} = 5.0$ to the values of $t$, with 39 d.f., in a $t$ table. With 39 d.f., the chances of getting a $t = 5.0$ is much less than 1 in 100 when really $\xi_g = \xi_c$. Therefore, the hypothesis that the genic values of lines C and G are the same was rejected. Similarly, hypotheses that $\xi_g = \xi_B$, $\xi_g = \xi_e$, $\xi_g = \xi_h$, would each be rejected at the .01 level of significance and the hypotheses that $\xi_g = \xi_a$ and $\xi_g = \xi_d$ would be accepted. Since part of Dr. Hazel’s opinion of the lines was probably based on observations of the pigs in this study, the true level of significance would not be as high as is indicated here. In each case the estimated genic value of line G was higher than that of the other lines. Comparisons with lines F, H, I, J
Table 2

Variance and Covariance Matrix for $\varepsilon_1$
(multiplied by 10,000)

<table>
<thead>
<tr>
<th>$\varepsilon_a$</th>
<th>$\varepsilon_b$</th>
<th>$\varepsilon_c$</th>
<th>$\varepsilon_d$</th>
<th>$\varepsilon_e$</th>
<th>$\varepsilon_f$</th>
<th>$\varepsilon_g$</th>
<th>$\varepsilon_h$</th>
<th>$\varepsilon_i$</th>
<th>$\varepsilon_j$</th>
<th>$\varepsilon_k$</th>
<th>$\varepsilon_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>48</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>-27</td>
<td>-21</td>
<td>-15</td>
<td>-24</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>-28</td>
<td>-20</td>
<td>-12</td>
<td>-30</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>11</td>
<td>9</td>
<td>5</td>
<td>7</td>
<td>-24</td>
<td>-24</td>
<td>-21</td>
<td>-32</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>10</td>
<td>3</td>
<td>5</td>
<td>-24</td>
<td>-26</td>
<td>-20</td>
<td>-30</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>3</td>
<td>7</td>
<td>-29</td>
<td>-19</td>
<td>-16</td>
<td>-30</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>3</td>
<td>-30</td>
<td>-18</td>
<td>-16</td>
<td>-25</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>-28</td>
<td>-17</td>
<td>-15</td>
<td>-25</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>261</td>
<td>-12</td>
<td>-22</td>
<td>10</td>
<td>-27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>198</td>
<td>-2</td>
<td>-19</td>
<td>-20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>178</td>
<td>-21</td>
<td>-18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>255</td>
<td>-29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
errors. Since it is probable one of the reasons for the increase of the estimate over the two

estimates are subject to large sampling errors because they are estimates of

vines with a procedure are shown in Table 4

The estimates in \( m - x \) are an unbiased estimate of \( m \),

\( m = 0 \) then we can solve for \( m \) in the theory. Since

The different lines to obtain equation in only and

\( \Phi (r) = 0 \) we can sum the data over

(2) \( E (x) \) the unbiased estimates of the differences between the \( d \) the

information is lost because of the three. Comparisons between the \( d \) the

cannot be compared to a \( \Phi (r) \) because the \( x \) is a component

where the other \( d \), do not contain any \( r \) value. Therefore, a

any \( r \) that has an \( x \) for a support, is placed because to continue,

The estimates that were obtained for the \( d \) as the shown in Table 4.

were not included because those lines have been omitted.

and \( X \) were not included because those lines have been omitted.
Table 3
Least Squares Estimates of $\hat{d}_{jk} = \mu + \frac{1}{2}(m_j + m_k) + d_{jk}$
and $d_{jk}$ (in pounds)

<table>
<thead>
<tr>
<th>dam lines</th>
<th>$\hat{d}_{jk}$</th>
<th>$d_{jk}$</th>
<th>dam lines</th>
<th>$\hat{d}_{jk}$</th>
<th>$d_{jk}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ab</td>
<td>150</td>
<td>-2</td>
<td>es</td>
<td>151</td>
<td>2</td>
</tr>
<tr>
<td>ac</td>
<td>141</td>
<td>-5</td>
<td>ex</td>
<td>139</td>
<td>0</td>
</tr>
<tr>
<td>ad</td>
<td>156</td>
<td>9</td>
<td>fj</td>
<td>170</td>
<td>14</td>
</tr>
<tr>
<td>ae</td>
<td>160</td>
<td>8</td>
<td>fs</td>
<td>143</td>
<td>-8</td>
</tr>
<tr>
<td>af</td>
<td>141</td>
<td>-12</td>
<td>gh</td>
<td>116</td>
<td>-24</td>
</tr>
<tr>
<td>aj</td>
<td>174</td>
<td>9</td>
<td>gj</td>
<td>158</td>
<td>5</td>
</tr>
<tr>
<td>as</td>
<td>162</td>
<td>3</td>
<td>ga</td>
<td>151</td>
<td>3</td>
</tr>
<tr>
<td>bc</td>
<td>132</td>
<td>-5</td>
<td>hs</td>
<td>168</td>
<td>19</td>
</tr>
<tr>
<td>bd</td>
<td>140</td>
<td>2</td>
<td>ik</td>
<td>152</td>
<td>9</td>
</tr>
<tr>
<td>be</td>
<td>132</td>
<td>-11</td>
<td>is</td>
<td>139</td>
<td>-5</td>
</tr>
<tr>
<td>bg</td>
<td>156</td>
<td>15</td>
<td>ax</td>
<td>139</td>
<td>-10</td>
</tr>
<tr>
<td>bs</td>
<td>146</td>
<td>-4</td>
<td>bx</td>
<td>145</td>
<td>5</td>
</tr>
<tr>
<td>cd</td>
<td>140</td>
<td>9</td>
<td>ox</td>
<td>139</td>
<td>7</td>
</tr>
<tr>
<td>ce</td>
<td>137</td>
<td>1</td>
<td>dx</td>
<td>141</td>
<td>7</td>
</tr>
<tr>
<td>cf</td>
<td>139</td>
<td>1</td>
<td>ex</td>
<td>139</td>
<td>0</td>
</tr>
<tr>
<td>ch</td>
<td>141</td>
<td>5</td>
<td>fx</td>
<td>145</td>
<td>4</td>
</tr>
<tr>
<td>cs</td>
<td>130</td>
<td>-13</td>
<td>gx</td>
<td>138</td>
<td>0</td>
</tr>
<tr>
<td>de</td>
<td>136</td>
<td>-1</td>
<td>hx</td>
<td>138</td>
<td>0</td>
</tr>
<tr>
<td>di</td>
<td>118</td>
<td>-14</td>
<td>ix</td>
<td>144</td>
<td>10</td>
</tr>
<tr>
<td>dj</td>
<td>139</td>
<td>-10</td>
<td>jx</td>
<td>133</td>
<td>-19</td>
</tr>
<tr>
<td>dk</td>
<td>139</td>
<td>-4</td>
<td>kx</td>
<td>142</td>
<td>-10</td>
</tr>
<tr>
<td>ds</td>
<td>148</td>
<td>3</td>
<td>sx</td>
<td>145</td>
<td>-1</td>
</tr>
</tbody>
</table>
### Table 4

Estimates of \( m_j \), obtained by using linear functions \( \hat{d}_{jk} \) (in pounds)

<table>
<thead>
<tr>
<th>line</th>
<th>( \hat{m}_j )</th>
<th>line</th>
<th>( \hat{m}_j )</th>
<th>line</th>
<th>( \hat{m}_j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>17</td>
<td>F</td>
<td>1</td>
<td>J</td>
<td>22</td>
</tr>
<tr>
<td>B</td>
<td>-1</td>
<td>G</td>
<td>-6</td>
<td>K</td>
<td>10</td>
</tr>
<tr>
<td>C</td>
<td>-15</td>
<td>H</td>
<td>-3</td>
<td>S</td>
<td>12</td>
</tr>
<tr>
<td>D</td>
<td>-12</td>
<td>I</td>
<td>-12</td>
<td>X</td>
<td>-9</td>
</tr>
<tr>
<td>E</td>
<td>-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
those found by Henderson (1948).

Estimates of $d_{jk}$ are obtained by subtracting $\hat{y} + \frac{1}{2}(\bar{y}_j + \bar{y}_k)$ from $d_{jk}$. These estimates are shown in Table 3.

D. Tests of Hypotheses

The tests of hypotheses are exact only if the errors are normally and independently distributed. However, Cochran (1947, p. 24) states that

...no serious error is introduced by non-normality in significance levels of the $F$-test or of the two-tailed $t$-test. While it is difficult to generalize about the range of populations that were investigated, this appears to cover most cases encountered in practice.

The procedure for testing the hypotheses that $r_{ijklm} = 0$, $c_{ijkl} = 0$, $t_{ijk} = 0$ and $b_1 = 0$, was outlined in section II. The values of the mean squares for making these tests are given in Table 6. To test the hypothesis that there are no litter effects on the 15th-day weights of pigs within a three-way cross, season-place subclass, the calculated $F(=3.23)$ was compared with 1.62, which is the tabulated .01 value for 39 and 1655 d.f. Since the probability is 1 in 100 that an $F$ calculated in the above manner from a group of data where there were no litter effects within subclasses will be as large as or larger than 1.62, the hypothesis that there are no litter effects within subclasses, i.e., the hypothesis that $r_{ijklm} = 0$, was rejected. Therefore, the accuracy of the following tests is only approximate, although that approximation is probably good, as was discussed on page 15 and above.

Each of the following tests are for the additional variation that
can be accounted for by the classification being tested after the other parameters in the model have accounted for all the variation that they can.

The hypothesis that \( a_{ijkl} = 0 \) is tested by comparing 1.05 to the tabulated value of \( F \) for 228 and 39 d.f., which = 1.55 at the .05 level. This indicates that the hypothesis, \( a_{ijkl} = 0 \), is acceptable.

The hypothesis that \( t_{ijk} = 0 \), when we accept the hypothesis that \( a_{ijkl} = 0 \), is tested by comparing 1.32 to the tabulated value of \( F \) with 143 and 39 d.f., which = 1.57 at the .05 level. This indicates that the hypothesis, \( t_{ijk} = 0 \), is acceptable. This among T mean square is probably biased upward because it would contain some of the differences among litters when there were only a few litters in a particular cross.

The hypothesis that \( b_i = 0 \) was rejected because the calculated value of \( F \) for an approximate test of the hypothesis was 10.07 as compared to the tabulated value of \( F \) with 18 and 39, which = 1.57 at the .05 level.

To test the hypotheses that \( d_{jk} = 0 \) and \( e_i = 0 \), assuming that \( a_{ijkl} = 0 \) and \( t_{ijk} = 0 \), we need to calculate the following sum of squares: \( R(\mu, g, m, d, b) \), \( R(\mu, g, m, b) \) and \( \sum_{i} e_{i} q_{i} \).

To obtain \( R(\mu, g, m, d, b) \) and \( \sum_{i} e_{i} q_{i} \), the estimates of \( e_i \) and \( b_i \) that were calculated in section III G were used. The method for obtaining the sum of squares using these estimates was given in section II. Since \( X \) is not used as a sire line, no estimate for \( g_x \) is used in these
calculations. All of the effects of \( X \) are included in the \( \frac{(y_{jx..})^2}{n_{jx..}} \) terms. The numerical values of \( R(\mu, g, m, b) \), and \( \Xi \Sigma g_i Q_j \), and their associated degrees of freedom are given in Table 6.

To calculate \( R(\mu, g, m, b) \), estimates of \( g_i \) and \( m_j \) must be obtained by using the model given on page 24. The method for calculating these \( g_i \) 's and \( m_j \) 's is given on page 25. The \( g_i \) 's and \( m_j \) 's are least squares estimates of \( (g_i - g_k) \) and \( (m_j - m_1) \) respectively, after the normal equations have been made independent by deleting the \( g_k \), \( m_1 \) and \( m_x \) equations. \( g_k \) and \( m_1 \) are the \( g \) and \( m \) values for lines \( k \) and \( l \), while \( g_i \) and \( m_j \) are general terms used to represent the \( g \) values for lines other than \( k \) and the \( m \) values for lines other than \( l \). These estimates may be expressed as deviations from zero, by assuming that \( g_a + g_b + \ldots + g_k + g_l + x = m_a + m_b + \ldots + m_k + m_l + x = 0 \), and will be designated as \( g_i \) and \( m_j \). The values for \( g_i \) and \( m_j \) are given in Table 5. \( R(\mu, g, m, b) \) was obtained by taking the sum of the \( g_i \) 's and \( m_j \) 's multiplied by their respective adjusted right hand sides plus \( \Xi \Sigma \frac{(y_{jx..})^2}{n_{jx..}} \) as outlined on page 25. The value of \( R(\mu, g, m, b) \) is shown in Table 6.

The F values in Table 6 indicate that the hypothesis, \( g_i = 0 \), should be rejected and the hypothesis, \( d_{jk} = 0 \), is acceptable at the .05 level of significance.

To test the hypothesis that \( m_j = 0 \), assuming that \( a_{ijk} = 0 \), \( t_{ijk} = 0 \), and \( d_{jk} = 0 \), \( R(\mu, g, m, b) \), which has already been calculated, and \( R(\mu, g, b) \) are needed.

The calculation of \( R(\mu, g, b) \) requires the use of the following model:
\[ y_{ijklm} = \beta_i + \frac{1}{n} (\beta_j + \beta_k) + b_m + e_{ijklm} \quad \text{where} \quad b_m = \mu + b_1. \]

Using this model the normal equations are:

\[ g_v = (n_{vv} + 1/\sum_{vq} vq^2) g_v + \sum_q (n_{qv} + 1/\sum_{vq} vq^2) g_q + \sum_1 (n_{v1} + 1/\sum_{vq} vq^2) b_1 = y_{vv} + \hat{y}_{vv}. \]

\[ b_1 = \sum_q (n_{q1} + 1/\sum_{q} q) g_q + n_{..1} b_1 = y_{..1}. \]

To solve these equations, the \( b_1 \) equations were absorbed into the \( g_v \) equations. The resulting \( g_v \) equations were solved by iteration, after they had been made independent by assuming that \( \sum_i g_i = 0 \).

\( g_i \) is the least squares estimate of \( g_i \), using the above model. These estimates of the \( g_i \)'s are: \( g_1 = 2.5, g_2 = \ldots = g_9 = 0, g_9 = 1.6, g_9 = 11.9, g_9 = 7.7, g_9 = -1.4, g_9 = .3, g_9 = -6.3, g_9 = -1.7, \) and \( g_9 = 5.4 \).

\[ R(\mu, g, b) = \sum_1 \frac{(y_{..1} - x_{..1})^2}{n_{..1} - 1}, \]

where \( o_i \) is the right hand side of the \( g_i \) equation after the \( b_1 \) equations have been absorbed.

The calculated value of \( R(\mu, g, b) \) is shown in Table 6.

To test the hypothesis that \( m_j = 0 \), assuming that \( c_{ijkl} = 0, \)
\( t_{ijk} = 0, \) and \( d_{jk} = 0, \) we consider the \( F \) value, .55, from Table 6. Since it is less than one we would certainly accept the hypothesis that \( m_j = 0 \).

Another test of the hypothesis that \( g_i = 0 \) is available if we assume that \( c_{ijkl} = 0, t_{ijk} = 0, d_{jk} = 0 \) and \( m_j = 0. \) To make this test, we compare \( \sum_i g_i o_i / 12 \) to the mean square among litters within three-way cross, season-place subclasses. This ratio is 4.75 compared to 2.01 from the tabulated \( F \) table with corresponding degrees of freedom. This
Table 5
Least squares estimates of $e_{ij}$ and $m_i$, obtained with $d_{jk}$ dropped out of the model

<table>
<thead>
<tr>
<th>line</th>
<th>$\check{e}_{ij}$</th>
<th>$\check{m}_i$</th>
<th>$\check{e}_{ij} + \check{m}_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>2.2</td>
<td>5.0</td>
<td>7.2</td>
</tr>
<tr>
<td>b</td>
<td>-3.5</td>
<td>6.6</td>
<td>-2.9</td>
</tr>
<tr>
<td>c</td>
<td>-6.3</td>
<td>-7.5</td>
<td>-13.8</td>
</tr>
<tr>
<td>d</td>
<td>1.7</td>
<td>1.3</td>
<td>3.0</td>
</tr>
<tr>
<td>e</td>
<td>-2.7</td>
<td>-4.3</td>
<td>-7.0</td>
</tr>
<tr>
<td>f</td>
<td>-2.3</td>
<td>5.1</td>
<td>2.8</td>
</tr>
<tr>
<td>g</td>
<td>12.2</td>
<td>-4.8</td>
<td>7.4</td>
</tr>
<tr>
<td>h</td>
<td>9.3</td>
<td>-4.8</td>
<td>4.5</td>
</tr>
<tr>
<td>i</td>
<td>-6.3</td>
<td>4.5</td>
<td>-1.8</td>
</tr>
<tr>
<td>j</td>
<td>1.1</td>
<td>-3.0</td>
<td>-1.9</td>
</tr>
<tr>
<td>k</td>
<td>-6.2</td>
<td>1.1</td>
<td>-5.1</td>
</tr>
<tr>
<td>s</td>
<td>-2.4</td>
<td>3.2</td>
<td>-5.6</td>
</tr>
<tr>
<td>x</td>
<td></td>
<td></td>
<td>3.3</td>
</tr>
</tbody>
</table>
### Table 6

Analysis of variance for testing the different effects

<table>
<thead>
<tr>
<th>Sources of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Total</td>
<td>2137</td>
<td>46,454,748</td>
<td></td>
</tr>
<tr>
<td>2. R(μ,ε,m,d,t,b)</td>
<td>215</td>
<td>45,306,028</td>
<td></td>
</tr>
<tr>
<td>3. R(μ,ε,ε,m,d,t,b)</td>
<td>72</td>
<td>45,032,595</td>
<td></td>
</tr>
<tr>
<td>4. R(μ,ε,m,b)</td>
<td>42</td>
<td>44,978,521</td>
<td></td>
</tr>
<tr>
<td>5. R(μ,ε,b)</td>
<td>31</td>
<td>44,969,656</td>
<td></td>
</tr>
<tr>
<td>6. R(μ,ε)</td>
<td>19</td>
<td>44,886,276</td>
<td></td>
</tr>
<tr>
<td>7. Between three-way cross, season-place subclasses</td>
<td>443</td>
<td>45,653,602</td>
<td></td>
</tr>
<tr>
<td>1 - (15 + 16)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Interaction between three-way crosses and season-place groups (7 - 2)</td>
<td>228</td>
<td>347,574</td>
<td>1,524</td>
</tr>
<tr>
<td>9. Among B</td>
<td>18</td>
<td>263,398</td>
<td>14,633</td>
</tr>
<tr>
<td>10. Among T(2-3)</td>
<td>143</td>
<td>273,433</td>
<td>1,912</td>
</tr>
<tr>
<td>11. Among Ω</td>
<td>11</td>
<td>57,877</td>
<td>5,262</td>
</tr>
<tr>
<td>12. Among D(3-4)</td>
<td>30</td>
<td>54,074</td>
<td>1,802</td>
</tr>
<tr>
<td>13. Among Κ(4-5)</td>
<td>14</td>
<td>8,865</td>
<td>606</td>
</tr>
<tr>
<td>14. Among Κ(5)</td>
<td>12</td>
<td>83,380</td>
<td>6,948</td>
</tr>
<tr>
<td>15. Among litters within three-way cross season-place subclasses</td>
<td>39</td>
<td>56,682</td>
<td>1,453</td>
</tr>
<tr>
<td>16. Within litters</td>
<td>1655</td>
<td>744,464</td>
<td>450</td>
</tr>
</tbody>
</table>

\[
\frac{\text{SST}}{\text{SSE}} = 1.05, \frac{1653}{1453} = 10.07, \frac{1912}{1453} = 1.32, \frac{5262}{1453} = 3.62, \frac{1802}{1453} = 1.24,
\]

\[
\frac{\text{SST}}{\text{SSE}} = .55, \frac{606}{1453} = 4.78, \frac{1453}{1450} = 3.23
\]

*Among G* was calculated with Ω's, m's, d's, b's and e's included in the model, while "Among G₂" was calculated without m's and d's in the model.
again indicates the hypothesis that $g_1 = 0$ should be rejected. The reason that the mean square among $G$ had 12 degrees of freedom this time and only 11 in the other test is that the effect of line $x$ was attributed to the maternal effect the first time, but included as a genic effect in this test.

Thus the tests indicate that we have evidence for rejecting only the hypotheses that $r_{ijklm} = 0$, $e_1 = 0$ and $b_1 = 0$.

E. Estimating Variance Components

To obtain estimates of $\sigma^2_w$, $\sigma^2_r$, and $\sigma^2_g$, we shall assume that the $e_{ijklm}$'s, $r_{ijklm}$'s and $e_1$'s are each independently distributed around mean = 0, with variances $\sigma^2_w$, $\sigma^2_r$, and $\sigma^2_g$ respectively. Using these assumptions we can set up Table 7. By considering the mean squares and expectations in Table 6, we obtain the following estimates of the variance components: $\hat{\sigma}^2_w = 450$, $\hat{\sigma}^2_r = 228$, $\hat{\sigma}^2_g = 27$ (from Among G) or $= 30$ (from among $G_2$). Thus the estimate of $\sigma^2_g$ is only about .04 as large as the expected variance between two pigs chosen at random, with the restriction that they must be from different litters within a three-way cross, season-place subclass ($\hat{\sigma}^2_w + \hat{\sigma}^2_r$).
Table 7

Mean squares and expectations needed for estimating \( \sigma_w^2 \), \( \sigma_r^2 \) and \( \sigma_g^2 \)

<table>
<thead>
<tr>
<th>Sources of Variation</th>
<th>d.f.</th>
<th>Mean Square</th>
<th>Expectation of the mean square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among G</td>
<td>11</td>
<td>5.262</td>
<td>( \sigma_w^2 + k_3 \sigma_r^2 + k_5 \sigma_g^2 )</td>
</tr>
<tr>
<td>Among G_2</td>
<td>12</td>
<td>6.948</td>
<td>( \sigma_w^2 + k_2 \sigma_r^2 + k_4 \sigma_g^2 )</td>
</tr>
<tr>
<td>Between litters within subclasses</td>
<td>39</td>
<td>1.453</td>
<td>( \sigma_w^2 + k_1 \sigma_r^2 )</td>
</tr>
<tr>
<td>Within litters</td>
<td>1655</td>
<td>150</td>
<td>( \sigma_w^2 )</td>
</tr>
</tbody>
</table>

The \( k \)'s in the above expectation have the following values:

\[
k_1 = n - \sum_{i,j}^{k} \frac{(n_{i,j,\ldots})^2}{n_{i,j,\ldots}} = k, h
\]

Exact estimates of \( k_2 \) and \( k_3 \) are not readily available with this number of lines, but working with examples that contain only 4 lines indicates that values of \( k_2 \) and \( k_3 \) will not be widely different from \( k_1 \)

\[
k_4 = \frac{\text{sum of the diagonals of the g equation after the } \mu, m, d \text{ and } b \text{ equation have been absorbed}}{11} = 14.36
\]

\[
k_5 = \frac{\text{sum of the diagonals of the g equations after the } \mu \text{ and } b \text{ equations have been absorbed}}{12} \text{ (or) } \frac{1}{n} \sum_{i=1}^{n} \left( \frac{m_{i,1,\ldots,1}}{n_{i,1,\ldots,1}} \right)^2 = 181.4
\]

\(^1\text{This method is given by Cockerham (1949).}\)
In the model

be the same, except for rounding errors. Whichever of $f$ or $g$ we need

the estimates of the differences between the estimates would

both in the model. Then $f$ is determined as the information between

respect to $f$ and combining $E_0$ as $E_0$ as necessary when $E_0$ and $E_f$

instead of $E_f$ with respect to $E_0$ and then with

$f(x)$ can be formed by differencing the estimate between

equations because $x$ is used only at a later time. Also, the $E_f$

$E_f$ and $E_f$ would avoid the computations that were in the

$m$ could be obtained by extracting the $f$ differences between the estimate of

how the losses performed at the times and at the times, the estimate of

and from the same times and $m$ was determined as the $f$ differences between

the computations more straightforward. Since the $E_f$ values were obtained

of a material dependence an $f$ rather than $E_f$ could have made

looking back over the analyses, it appears that with the $f$

...
The least squares estimates were used for comparing the different

14. DISCUSSION AND APPLICATION

-7-
even though the tests of hypotheses in section III D indicate the hypotheses that \( m_j = 0 \) and \( d_{jk} = 0 \) could be accepted. \( t_{ijk} \) and \( a_{ijkl} \) are not included in the model because estimates of the individual \( t \)'s and \( c \)'s would in most cases be based on only a few observations.

In considering which line should give the best results when used to top cross on outbred herds, the genic values of the different lines should receive primary consideration. Although the maternal values should receive secondary consideration, they should be considered in most herds because some of the females by sires used for top crossing will be used as sows the next year. To obtain precise estimates of the worth of the lines in top crossing with regard to 154-day weight, a selection index would have to be set up. However, in this study \( \hat{g}_g \) is so much larger than the \( \hat{g} \) (the \( \hat{g} \)'s are estimates of \( \hat{g} \) the genic differences between the inbred lines) value for any other line still in the herd that line \( G \) would certainly be the choice of the lines to use in top crossing. Lines \( I, K, L, J \) and \( X \) have been culled from the herd. Since \( \hat{m}_a \) is high and \( \hat{g}_a \) is above average, line \( A \) would be the second choice for use in top crossing.

If the animals in these lines possess a random sample of the genes in the better Poland China herds of Iowa, the \( g \) values estimate how the different lines should compare with an average boar from these outbred Poland China herds. In his farm testing program Hasek (1950) has made estimates of how boars from several of the inbred lines compare with outbred Poland China boars and boars of other breeds. To estimate the difference between the boars of an inbred line and outbred boars,
This does not indicate an estimate of the interaction between
the data or power needed to be done other than close
the data at 45 degrees or power of the data and some of
the data, we would expect a group between the data of the

Furthermore, we use an estimate of the product of the data

The following quantities of how prices from different times will

be different estimates of the E, and from the results indicate that the

Table 6

| Time | 
|------|---|
| 1    | 5.4 |
| 13   | 5.2 |
| 3.5  | 6.7 |
| 5    | 9   |

These data are taken from Table 6.

| Description | 
|-------------|---|

The corresponding bore and the weight of the data from the product of the data on an

In the same sense, then he used the average difference

he calculated the difference between the weight of the data on the data from an
In the presence of the line, if the line is one of the lines in the crossing, changes, and
may cease, because the line is still in the herd. Since these were only
the lines which produce heterosis at 12.5 days than any other three-
the same indicates that more than the breed to cross the line. An or-
position of the line for the site line and the line are given from in the table. Com-
tines which are still in the herd are given from in the table. The
and the line are given on page 26.
the three end season-peace groups, the line are given on page 26.
are the means of the different crossing per cent change. Corrected for
line and the line are given from in the table. The line are given from in the table
To compare different three-year classes, we need for the three-year class.

Inbreed Females.

Self, because crossing females appear to make better mothers than
operations will probably use three-year classes to produce the three-
some manner that crossing seed can be made to. The field is
be possible to sell crossing pigs to farmers in much the
permanently established as an important part of the entire industry.
not fix for their young pigs next season. Usually, if the field is too hot to
usually without all of the young pigs. And so back to the field is too hot to
young crossing pigs to farmers. The farmers who use the crossing pigs
which will sell well wanting pigs to farmers have opened the way to sell the
in the past few years the development of a new pig heteroresis
across was not included in Henderson's analysis.
and estimating the productivity of single crosses. The G x G method
the line, which Henderson (1949) found to be an important factor when

---
Table 9
\( \hat{d}_{jk} + \frac{1}{2} \hat{g}_{j} + \frac{1}{2} \hat{g}_{k} \) values for crossline dams. Crosses which involve lines still in the herd are listed first (in pounds)

<table>
<thead>
<tr>
<th>Crossline dams</th>
<th>( \hat{d}<em>{jk} + \frac{1}{2} \hat{g}</em>{j} + \frac{1}{2} \hat{g}_{k} )</th>
<th>Crossline dams</th>
<th>( \hat{d}<em>{jk} + \frac{1}{2} \hat{g}</em>{j} + \frac{1}{2} \hat{g}_{k} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ab</td>
<td>149.1</td>
<td>di</td>
<td>117.5</td>
</tr>
<tr>
<td>ac</td>
<td>137.4</td>
<td>dj</td>
<td>142.7</td>
</tr>
<tr>
<td>ad</td>
<td>157.4</td>
<td>dk</td>
<td>137.2</td>
</tr>
<tr>
<td>ae</td>
<td>159.1</td>
<td>fj</td>
<td>170.9</td>
</tr>
<tr>
<td>as</td>
<td>161.7</td>
<td>fs</td>
<td>139.7</td>
</tr>
<tr>
<td>bc</td>
<td>125.5</td>
<td>gh</td>
<td>127.2</td>
</tr>
<tr>
<td>bd</td>
<td>138.6</td>
<td>gj</td>
<td>165.5</td>
</tr>
<tr>
<td>be</td>
<td>127.9</td>
<td>hs</td>
<td>172.6</td>
</tr>
<tr>
<td>bg</td>
<td>159.4</td>
<td>ik</td>
<td>148.0</td>
</tr>
<tr>
<td>bs</td>
<td>142.4</td>
<td>is</td>
<td>136.7</td>
</tr>
<tr>
<td>cd</td>
<td>136.1</td>
<td>ax</td>
<td>140.0</td>
</tr>
<tr>
<td>ce</td>
<td>130.9</td>
<td>bx</td>
<td>143.2</td>
</tr>
<tr>
<td>cs</td>
<td>124.7</td>
<td>cx</td>
<td>135.3</td>
</tr>
<tr>
<td>de</td>
<td>134.7</td>
<td>dx</td>
<td>141.7</td>
</tr>
<tr>
<td>ds</td>
<td>146.7</td>
<td>ex</td>
<td>137.4</td>
</tr>
<tr>
<td>es</td>
<td>148.1</td>
<td>fx</td>
<td>142.8</td>
</tr>
<tr>
<td>gs</td>
<td>154.7</td>
<td>gx</td>
<td>142.9</td>
</tr>
<tr>
<td>af</td>
<td>139.8</td>
<td>hx</td>
<td>144.5</td>
</tr>
<tr>
<td>aj</td>
<td>177.2</td>
<td>ix</td>
<td>143.0</td>
</tr>
<tr>
<td>of</td>
<td>132.2</td>
<td>jx</td>
<td>135.4</td>
</tr>
<tr>
<td>ch</td>
<td>142.7</td>
<td>kx</td>
<td>138.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ex</td>
<td>143.7</td>
</tr>
</tbody>
</table>
AX dams, and none of these 58 pigs had line G as the sire line, further
testing should be done to determine just which cross will produce the
heaviest pigs at 15\(\frac{1}{2}\) days.

To obtain estimates of which of the three-way crosses, that involve
crossline dams that were not tested, would produce the heaviest pigs,
we could consider \(\hat{e}_i\) for the possible sire line and \(\hat{\beta}(\hat{e}_j + \hat{e}_k + \hat{m}_j + \hat{m}_k)\)
for the possible crossline dams. However, all the crossline dams
that could be made from the lines still in the herd were used in this
study as female parents, except those involving line G. Line G has
such a large \(\hat{e}_i\) value that it normally would be used as the sire line.

The estimates of the \(e_i\)'s which are obtained when \(m_j\) and \(d_{jk}\) are
left out of the model correspond very closely to those that were obtained
with \(m_j\) and \(d_{jk}\) in the model. The correlation between them is .97.
Therefore, the estimates of the \(e_i\)'s were not greatly influenced by
the values assigned to \(m_j\) and \(d_{jk}\). When only the effects \(\epsilon_i\, g\) and \(b\)
are included in the model, the G X A single cross is again expected
to produce heavier pigs at 15\(\frac{1}{2}\) days than any other single cross between
lines still in the herd. Also, a three-way cross which involves line G
as the sire line and line A as one of the female lines is expected
to be the best three-way cross.

Before these estimates can be used to predict which three-way cross
will produce the heaviest litters at 15\(\frac{1}{2}\) days, estimates of the litter
size for the crosses must be obtained.

Since the \(g\) values for the lines appear to be more important than
the interactions among them, the performance of the inbred pigs should
give information on how the lines will perform in crosses. The
correlation between the weights of the inbred pigs in the lines and the
estimates of the $g$ values for the lines was .62 when $s_j$ and $d_{jk}$ were
considered in the model and .65 when they were not. The average 154-
day weight of the pigs in each of the inbred lines produced during
the years 1940 fall to 1945 spring, as calculated by Hazle (1950), was
used as the weights of the inbred pigs in each of the lines.

The large mean square among litters within crosses and season-
place groups as compared to the mean square within litters indicates
that a breeder may improve his lines by selecting within the lines as
the inbreeding of the lines increases. For this selection to be effec-
tive, there must be genic variation within lines. If the mean square
among litters had been about the same as the mean square within litters,
an estimate of the genic variation within lines would necessarily be near
zero. The relatively large mean square among litters could be completely
caused by common pre-weaning environment; so it does not necessarily
mean that there is genic variation left within the lines, but it indicates
there could be, and if there is, a breeder will improve the lines by
selecting within them.

These considerations indicate that the best way to form inbred
lines which will produce heavy linecross pigs at 154 days is to:

1. Start many lines, probably by breeding a boar to several of
   his sisters.

2. Cull a large percentage of the lines on the basis of the 154-
day weight of the pigs produced in the first and second
generations of inbreeding.
3. Select within lines.

4. Make as many of the comparisons as possible within season-place groups.

5. Test the more promising lines in crosses to determine the best cross combination.
V. SUMMARY

The differences among the genic values of the lines, season-place groups, litters within three-way cross season-place subclasses, and pigs within litters were important causes of the variation in the 154-day weights of the 2137 three-way cross pigs in this study. The maternal effects of the lines and the interactions between the different effects were not statistically significant.

Variance components were estimated as: $\hat{\sigma}_g^2 = 27$, $\hat{\sigma}_r^2 = 228$, $\hat{\sigma}_w^2 = 450$, where the $g$'s (the genic values of the lines), $r$'s (litter effects within three-way cross season-place subclasses) and the effects peculiar to individual pigs within litters are assumed to be independently distributed around mean = zero with variance $\hat{\sigma}_g^2$, $\hat{\sigma}_r^2$ and $\hat{\sigma}_w^2$ respectively.

The estimates of the $g$ values of the lines were $\hat{\varepsilon}_a = 1.5$, $\hat{\varepsilon}_b = -3.5$, $\hat{\varepsilon}_c = -8.3$, $\hat{\varepsilon}_d = 1.3$, $\hat{\varepsilon}_e = -3.7$, $\hat{\varepsilon}_f = 4.4$, $\hat{\varepsilon}_g = 10.3$, $\hat{\varepsilon}_h = 12.1$, $\hat{\varepsilon}_i = -2.9$, $\hat{\varepsilon}_j = 5.4$, $\hat{\varepsilon}_k = -5.5$, $\hat{\varepsilon}_l = -2.4$. $\hat{\varepsilon}_1$ represents the least squares estimate of the $g$ value of line 1, obtained using $\mu$, $g$, $m$ (maternal effect of a line), $d$ (interaction between the lines used to produce the single cross females), and $s$ (season-place effect) in the model.

The estimates of the $g$'s were affected only slightly by leaving $m$ and $d$ out of the model.

There was a close correspondence between the way lines A, B, C, and S (only lines tested on farms) performed in the farm testing program
and the estimates of their $g$ values.

The correlation between the estimates of the $g$ values of the lines and the weight of the pigs in the inbred lines was above .6.

The large mean square among season-place groups as compared to the mean square among litters within three-way cross season-place subclasses indicates that comparisons among pigs should be on a within season-place group basis. In other words, the effects of the season and place sometimes differ so importantly that they can hide or confuse differences among lines, if the pigs in the lines are produced in different seasons or places.

Estimates of the different effects indicate that a cross involving line 6 as a sire line and line 4 as one of the dam lines will produce heavier pigs at 154 days than any other cross among the lines still in the herd.
VI. LITERATURE CITED


---------- 1950. Information on crossbreeding tests. Private communication.


VII. ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to Dr. J. L. Lush for his encouragement, counsel, and criticism throughout the investigation. Sincere thanks also are expressed to Dr. L. N. Hazel for his many helpful suggestions in the analysis and interpretation of the data. Also, the author is deeply grateful to his wife, Eileen Magee, for her inspiration and aid in preparing this manuscript.