

SOME RESULTS ON EXPERIMENTAL DESIGNS WHEN THE
USUAL ASSUMPTIONS ARE INVALID

by

Hale Caterson Sweeny, B.M.E., M.S.
411

Thesis submitted to the Graduate Faculty of the
Virginia Polytechnic Institute
in candidacy for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

APPROVED:

~~_____~~
Chairman, Advisory Committee

May 15, 1956

Blacksburg, Virginia

TABLE OF CONTENTS

I	General Introduction	Page 4
<u>PART I</u>		
<u>TESTS OF HOMOGENEITY FOR EXPERIMENTAL</u>		
<u>DESIGNS WITH CORRELATED AND HETEROGENEOUS</u>		
<u>ERRORS</u>		
II	Introduction to Part I	13
III	Theory of the Test Procedures	17
IV	Application of the Theory to Design of Complete Block Experiments	43
V	Applications and Computational Formulae	69
<u>PART II</u>		
VI	<u>ON BERKSON'S CASE OF LINEAR REGRESSION</u> <u>WITH ERRORS IN THE INDEPENDENT VARIABLES</u>	89
<u>PART III</u>		
VII	<u>SOME CONSEQUENCES OF ANALYZING</u> <u>NON-ADDITIVE DATA ARISING FROM AN</u> <u>UNREPLICATED FACTORIAL EXPERIMENT</u>	98
VIII	Summary	110
	Acknowledgements	113
	Bibliography	114
	Vita	118

	FIGURES	Page
I	Ratio of Power of T^2 Analysis to Power of Usual Analysis of Randomized Block Design	49
TABLES		
5.1	Results of a Reaction Rate Experiment	73
5.2	Transformed Data, Randomized Block Example	74
5.3	Computational Routine, Randomized Block Example	77
5.4	Latin Square Design	78
5.5	Data for Latin Square Example	79
5.6	Summed Data of Latin Square Example	80
5.7	Transformed Data of Latin Square Example	81
5.8	Data for Split Plot Example	83
5.9	Transformed Data for Split Plot Example	85
5.10	Totals Over Formulas, Split Plot Example	87
7.1	Values of Untransformed Data	102
7.2	Linear Contrasts	102
7.3	Calculated Values of F_B	105
7.4	Calculated Values of F_{B^2}	106
7.5	Calculated Values of F_{BC}	107

I. GENERAL INTRODUCTION

The cornerstone of science is direct observation or experimentation. Except in abstract fields such as mathematics, any hypothesis relating to the real world must ultimately be accepted as true or rejected as false on the basis of the agreement between what is observed and what is expected to be observed, given that the hypothesis is true. We are immediately led to experimentation to obtain the necessary data for a study of such agreement.

Unfortunately, earlier scientists were more concerned with the results of their experiments than with the validity of their experimental methods and the philosophy behind these methods. For example, in attempting to locate some physical optimum, earlier physical scientists and--sadly enough--many such scientists today consider the dictum "vary one factor at a time, holding all other factors constant" to be the ultimate in good experimental technique.

The study of a "theory" of the design and analysis of experiments was begun in the early 1920's by R. A. Fisher and others at the Rothamsted Experimental Station in England. The earliest work seems to have been concerned with the analysis of data already collected; it was soon realized that proper initial arrangement of the experiment would

result in more efficient experimental techniques, and work was started on this problem of the design of experiments. This theory, embracing both the design and the analysis of experiments, has grown in importance and today holds a prominent place in science.

For most of the simpler designs based on this theory, four assumptions are made. The first assumption is:

- (a) An observation consists of a sum of component parts, one part being attributed to the particular class of items tested, one part being attributed to the time or place at which the trial occurred, one part being a "random" error term, and so on.

This is generally referred to as the assumption of additivity. In effect, it implies that there is some "true" value, and an observation only deviates from this "true" value by an amount which we term "experimental error". Further, the "true" value is a sum of component parts which individually describe the behavior of different aspects of the experiment.

The experimental error at any one trial is incapable of being predicted exactly; it is a random sort of term. Rather than consider the behavior of a single error term, the behavior of a large number of such terms is described by specifying the functional form of the probability distribution of these terms. The second assumption made is that

- (b) The probability distribution describing the random errors is of the form known as Gaussian or Normal.

The last two assumptions also concern this error term, for it is assumed that

- (c) The possession of information about an error term gives no information about other error terms; they are stochastically independent, and
- (d) The variances of the error terms are constant from trial to trial and from treatment to treatment.

These assumptions concern the analysis of the experimental results. When the treatments to be tested are different levels of independent variables, and it is desired to test for relationships between the responses and these independent variables, another assumption is made:

- (e) The levels of the factors are known without error.

As an example of this last assumption, consider the classical agricultural experiment where various amounts of fertilizer (here considered as various levels of the independent variables) are spread on plots and it is desired to measure the relationship between yield and amount of fertilizer. The assumption implies here that

the amounts of fertilizer on each plot are known exactly.

Much work, both theoretical and practical in nature, has been done on the consequences of the failure of the data to follow the stated assumptions and on methods of "correcting" either the data or the analysis to handle such cases. We shall discuss some of this work below.

When the data does not follow assumption (a), very little can be done to correct for this unless the true nature of the non-additivity is known. Thus, in the physical sciences one finds that the more usual model has component parts which are multiplicative in nature. This follows from the theories of dimensional analysis. Here of course, we only need to transform the data by a linearizing transform such as by use of the logarithm of the observation to correct for this perturbation. Of course, in so transforming the data, other assumptions may be violated and this must be examined in each case.

Tukey [7] * has given a method of testing for non-additivity when the type of non-additivity is unknown. Also, he has given a method of deriving an empirical transformation from the data to allow the observations to be transformed to a set which is additive. There are two features of the work which seem disturbing and would warrant investigation, and these are:

* Numbers in square brackets refer to bibliography.

(a) The test and transformation are designed to be effective against only one particular type of non-additivity and it is not clear what happens when the test and transformation are applied to data which has another type of non-additivity. Also, it is not clear what happens in the analysis when the data is actually additive, but the test is significant by chance alone and a transformation of the type suggested is made prior to analysis.

(b) The transformation is derived from the data itself, and then applied to the data. As such it is a self-contained operation. It would seem that such an operation should place some restrictions on the final analysis and should demand some modification of the density function of the test statistic used in the analysis. No such modification has been suggested in the literature.

This is the only practical suggestion for handling non-additive data which has been found in a review of the literature. The problem of transformation has been discussed in general by Tukey and Anscombe [8] .

The consequences of the failure of the data to follow the second assumption--the error having a Gaussian or Normal distribution--has been investigated by Bradley [1] , [2], Box and Andersen [3], and others. The results of these investigations seem to show that this is a very weak assumption; the actual distribution of the errors can be fairly

far from the normal form with very little effect on the results.

Departures in the data from the third assumption--stochastic independence--are probably the most serious which arise. Good examples of such perturbations are difficult to find; the effect is present when, if an error is low in one trial, it is consistently either high or low in another. Thus, the perturbations will appear if the errors are interrelated.

There is very little work reported in the literature on the analysis of experiments when this assumption is not true. There is one simple important case where usual methods of analysis do handle correlated error and this is when the data are collected in pairs from only two treatments. Here we assume the model

$$x_{ij} = \pi_i + \varepsilon_{ij} \quad , \quad i = 1, 2; \quad j = 1, \dots, N \quad (1.1)$$

where x_{ij} denotes the j^{th} observation on the i^{th} treatment, π_i is the (unknown) true mean for the i^{th} treatment, and ε_{ij} is the random error term associated with x_{ij} , with

$$E(\varepsilon_{ij}) = 0$$

In terms of the assumption previously listed, we invoke assumption (a) in the additive form of (1.1) and invoke a generalized form of assumption (b) in the description of the random error term ε_{ij} . The usual hypothesis that

it is desired to test is that the means of the two treatments are equal; in mathematical notation,

$$H_0 : \pi_1 = \pi_2 \quad (1.2)$$

If we assume that the errors are correlated, such that

$$E(\varepsilon_{1j} \varepsilon_{2k}) = \delta_{jk} \sigma_{12} \quad (1.3)$$

where δ_{jk} is Kronecker's delta;

$$\begin{aligned} \delta_{jk} &= 1 \text{ if } j = k, \\ &= 0 \text{ otherwise,} \end{aligned}$$

then it is known that the proper analysis of the data is made by the paired Student-t test. To perform the analysis, we form the differences between the pairs:

$$y_j = x_{1j} - x_{2j}, \quad (1.4)$$

and test whether the mean difference is zero. In the description of this test in texts, it is usually stated that the assumption made is that the differences are normally distributed. The analysis is valid under a more general model of the form

$$x_{1j} = \pi_i + \beta_j + \varepsilon_{ij} \quad (1.5)$$

where β_j is an additive part associated with the j^{th} pair of observations. It is seen that in forming the differences (1.4), the β_j term will disappear and the test would be carried out exactly as if the model had been (1.1). Now if the data arose from the design known as "randomized blocks design" with only two treatments, the model would be as

given in (1.5) where the β_j term is an effect due to blocks. In this case, it can be shown that the method of analysis of a randomized block experiment is exactly equivalent to the paired Student-t test and thus the design is valid over a wider set of assumptions than usually given.

In fact, Brandt [22] used the method of subtracting observations in cross-over trials for the analysis of data. However, he did not mention the validity of the method under the wider set of conditions given herein.

There has been almost no work reported on the analysis of experimental designs when the errors are correlated. The only results found are due to Hsu [9] and Graybill [15], whose works form the basis of much of the work done herein. Hsu did not apply his results to experimental designs per se, and Graybill, applying Hsu's results to a randomized block design, seems to have overlooked the fact that the method developed is valid when errors are correlated, and presented the method for use when the fourth assumption (d) is invalid.

The fourth assumption concerns the homogeneity of variances of the errors associated with different treatments. In the example just given of the paired Student-t test, the variances of the errors may be different just as the errors may be correlated. If the data do not arise from a paired situation, we are led immediately to the

Behrens-Fisher problem. No really satisfactory exact solution to this problem has been found.

These and other points concerning the assumptions made in the analysis of experiments and the consequences of violating the assumptions have been admirably summarized in the papers by Bartlett [4], Cochran [5], and Eisenhart [6].

There seems to have been no work reported on the consequences of errors arising in the factor levels in factorial experiments. This would seem to be especially important, as such errors are the rule rather than the exception.

The work in this dissertation is concerned with assumptions (a), (c), (d), and (e), and is divided into three parts as follows:

(1) The first part extends the work of Hsu and Graybill; methods are derived for the analysis of complete block designs when the experimental errors are correlated and have heterogeneous variances.

(2) A result due to Geary is extended to show that the estimates of the coefficients of linear and quadratic components in factorial experiments may be consistent when the levels of qualitative factors are subject to errors of certain type.

(3) The consequences of analyzing multiplicative data arising from unreplicated factorial experiments are examined.

PART I

TESTS OF HOMOGENEITY FOR EXPERIMENTAL DESIGNS
WITH CORRELATED AND HETEROGENEOUS ERRORSII. INTRODUCTION TO PART I.

We shall, in this part, develop tests of significance for complete block experimental designs when the errors are correlated and have heterogeneous variances. In all cases, it will be assumed that the models describing the observations are linear and that the errors have a joint multivariate normal distribution.

This part of the dissertation is divided into four chapters:

- (II) Introduction
- (III) Theory of the test procedures, wherein derivations of exact analyses for experimental designs are given based on models without the assumptions of homogeneity of variances and independence of errors.
- (IV) Applications of the theory to complete block experiments
- (V) Numerical examples, wherein the theory presented is applied to problems and worked examples of the various types of designs are given.

Five cases have been studied; four have been satisfactorily solved. The unsolved case has been studied and some of the difficulties discussed. These cases refer to different assumptions concerning the structure of the error variance-covariance matrix. We write this covariance matrix as

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix} .$$

The values on the major diagonal, σ_{ii} , are the measures of variation of the individual error terms; the values off of the major diagonal, σ_{ij} ($i \neq j$), are measures of the relations between the error terms. The five cases are:

Case 1: Σ completely known.

Case 2: It is known that Σ is proportional to a given matrix, the constant of proportionality being estimated from the data.

Case 3: All diagonal terms constant and equal, all off-diagonal terms constant and equal. The matrix is then of the form

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}$$

where σ^2 and ρ must be estimated from the data.

Case 4: All off diagonal terms are known to be zero, all diagonal terms may be different and unknown. The covariance matrix thus has the structure

$$\Sigma = \begin{bmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{pp} \end{bmatrix}$$

Case 5: All terms are unspecified and unknown; every variance and covariance must be estimated from the data. This is the most general case; it implies complete ignorance of any relation between the terms.

In the notation used above, we write the covariance matrix of the errors under assumptions (c) and (d) as

$$\Sigma = q^2 \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

An extensive review of the literature fails to disclose previous work on any of the above five cases, except the last one. Graybill [15] published work on Case 5, for a randomized block design only; it was this work which suggested the present problem. However, Case I is a classical case and it is felt that some discussion of this case must have appeared. With these exceptions, all work presented herein is original.

In the third section of this part of the dissertation, applications of the last case (Case 5 above) only are worked out in detail. The formulas for the other cases are given.

III THEORY OF THE TEST PROCEDURES

3.1 Notation. Let X_i denote the p -variate observation row vector $[x_{i1} \ x_{i2} \ \dots \ x_{ip}]$. A sample of N such observations is denoted by the matrix

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} = \begin{matrix} & \mu_1 & \mu_2 & & \\ \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ x_{N1} & x_{N2} & \dots & x_{Np} \end{bmatrix} & & & & \end{matrix} \quad (3.1)$$

We shall assume throughout that X_i, X_j are independent, for all i, j ($i \neq j$). We also assume that the components of X_i follow the multivariate normal law. That is, if Σ is the p -square variance-covariance matrix associated with the observation vector X_i ,

$$\Sigma = [\sigma_{rs}] \quad (3.2)$$

and M is the row vector of means $[\mu_1, \mu_2, \dots, \mu_p]$ with

$$E[x_{ik}] = \mu_k, \quad i = 1, \dots, N; \quad k = 1, \dots, p, \quad (3.3)$$

then the density of X_i may be written

$$f(X_i) = (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(X_i - M) \Sigma^{-1} (X_i - M)'\right\} \quad (3.4)$$

We denote this by the notation: $X_i \sim \text{MVN}(M, \Sigma)$. We also

$$\text{let } \Sigma^{-1} = \Lambda = [\lambda_{ij}] \quad (3.5)$$

The maximum likelihood estimator of μ_j from a random sample of N observation vectors can be shown to be the particular sample mean

$$\hat{\mu}_j = \bar{x}_j = \frac{\sum_{i=1}^N x_{ij}}{N}; \quad j=1, \dots, p, \quad (3.6)$$

and the maximum-likelihood estimators of the elements of the variance-covariance matrix, when this matrix has an unknown structure as in case 5, are given by

$$\hat{\sigma}_{ij} = \frac{v_{ij}}{N} = \frac{1}{N} \sum_{k=1}^N (x_{ki} - \bar{x}_i) (x_{kj} - \bar{x}_j); \quad i, j = 1, \dots, p. \quad (3.7)$$

The matrix of sample sums of squares and products is denoted by $V = [v_{ij}]$ with v_{ij} implicitly defined by (3.7), and its inverse by $V^{-1} = [v^{-1}ij]$. The unbiased matricial estimator of Σ is given by $\hat{\Sigma} = \frac{1}{N-1} V$.

3.2 Tests of Homogeneity and Hsu's Transformation.

It is desired to construct practical tests of the hypothesis of homogeneity of means,

$$H_0: \mu_1 = \mu_2 = \dots = \mu_p, \quad (3.8)$$

in the multivariate case for applications in the analysis of experimental designs. An important case of this is given when the variance-covariance matrix Σ has the structure

$$\Sigma = I\sigma^2 \quad (3.9)$$

where I is the p -square identity matrix. This case forms the basis for the analysis of variance, wherein the hypothesis of homogeneity (3.8) is the one of main interest.

Hsu [9] showed that in one case to be discussed below, a test of the hypothesis (3.8) could be reduced to a test of a simpler hypothesis by the following transformation. Let

$$y_{ik} = x_{ik} - x_{ip}; \quad k=1, \dots, p-1; \quad i=1, \dots, N. \quad (3.10)$$

As y_{ik} is a linear function of normally distributed variates, it is itself normally distributed with means

$$E(y_{ik}) = E(x_{ik} - x_{ip}) = \mu_k - \mu_p = \dot{\mu}_k, \quad k=1, \dots, p-1 \quad (3.11)$$

and a $(p-1)$ -square covariance matrix $\dot{\Sigma}$ with elements

$$\text{cov}(y_{ik}, y_{im}) = \delta_{ir} [\sigma_{km} - \sigma_{kp} - \sigma_{mp} + \sigma_{pp}] = \dot{\sigma}_{km}, \quad (3.12)$$

where δ_{ir} is a Kronecker delta. The hypothesis of homogeneity (3.8) reduces to the hypothesis,

$$H_0: \dot{\mu}_i = 0, \quad i=1, \dots, p-1, \quad (3.13)$$

for the only solution to the set of equations

$$(\mu_k - \mu_p) = 0, \quad k=1, \dots, p-1,$$

is given by

$$\mu_1 = \mu_2 = \dots = \mu_p.$$

Thus the hypothesis (3.8) can be tested by applying the transformation (3.10) and testing the equivalent hypothesis

(3.13). Denoting the vector of transformed means (3.11) by \dot{M} , we have reduced the problem to one of the development of tests of the hypothesis

$$H_0: \dot{M} = \dot{M}^0 \quad (3.14)$$

where \dot{M}^0 is a specified vector, for the hypothesis (3.8),

$$\dot{M}^0 = (0, 0, \dots, 0).$$

Different test procedures follow from different a priori assumptions concerning the structure of the covariance matrix Σ . We shall consider tests under the following assumptions corresponding to the cases listed in Chapter II:

Case 1 Σ known.

Case 2 $\Sigma = K\sigma^2$, $K = [k_{ij}]$ and is known.

Case 3 $\sigma_{11} = \sigma_{jj}$ for all i, j ; $\sigma_{ij} = \text{constant}$.

Case 4 $\sigma_{ij} = 0$, $i \neq j$; σ_{11} unspecified.

Case 5 Σ general; no structure postulated.

Tests of homogeneity exist for some of these cases without use of the transformation; in these cases, tests using the transformation are derived to demonstrate equivalence. The likelihood ratio principle is used in all tests developed in this paper.

Case 1. We desire a test of the hypothesis (3.8) when the sample is drawn from a multivariate normal population with known covariance matrix. Applying Hsu's transformation as

in (3.1), the joint density of a sample of N transformed observations may be written as

$$f(Y_1, \dots, Y_N) = (2\pi)^{-\frac{N(p-1)}{2}} |\dot{\Sigma}|^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2} t \dot{N} - \frac{N}{2} (\bar{Y} - \dot{M}) \dot{\Lambda} (\bar{Y} - \dot{M})' \right\}, \quad (3.16)$$

where \bar{Y} is the vector of sample means $[\bar{y}_1, \bar{y}_2, \dots, \bar{y}_{p-1}]$, \dot{M} is the vector of transformed population means $[\dot{\mu}_1, \dot{\mu}_2, \dots, \dot{\mu}_{p-1}]$, \dot{V} is the matrix of sample products and crossproducts defined analogously to (3.7), and $\dot{\Lambda}$ is the inverse of $\dot{\Sigma}$, the transformed covariance matrix. As Σ is assumed to be known, thus $\dot{\Sigma}$ and $\dot{\Lambda}$ are known.

To construct a test of the transformed hypothesis (3.14), we shall derive the likelihood ratio statistic and then find the distribution of this statistic.

To obtain an estimate of \dot{M} by maximum likelihood, it is easily seen that the likelihood under the general alternative (3.16) is maximized for variations in \dot{M} when $\bar{Y} = \dot{M}$, making the quadratic form involving \dot{M} in the exponent vanish. Thus, for testing the hypothesis $H_0: \dot{M} = \dot{M}^0$, we have the likelihood ratio

$$\frac{(2\pi)^{-\frac{N(p-1)}{2}} |\dot{\Sigma}|^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2} t \dot{N} - \frac{N}{2} (\bar{Y} - \dot{M}^0) \dot{\Lambda} (\bar{Y} - \dot{M}^0)' \right\}}{(2\pi)^{-\frac{N(p-1)}{2}} |\dot{\Sigma}|^{-\frac{N}{2}} \exp \left\{ -\frac{1}{2} t \dot{N} \right\}} \quad (3.17)$$

$$= \exp \left\{ -\frac{N}{2} (\bar{Y} - \dot{M}^0) \dot{\Lambda} (\bar{Y} - \dot{M}^0)' \right\} = \exp \left\{ -\frac{1}{2} Q_1 \right\}, \quad (3.18)$$

where \dot{M}^0 is the vector $(0, 0, \dots, 0)$. We use Q_1 as a test statistic. When the null hypothesis is true, $\{-\frac{1}{2} Q_1\}$ is the exponent in the distribution of \bar{Y} , and it is well known that Q_1 thus follows the central chi-square law with $(p-1)$ degrees of freedom.

If the null hypothesis is not true, let the true transformed vector of means be denoted by $[\dot{\mu}_1] = \dot{M}$, and write

$$\phi_1^2 = N \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \lambda_{ij} \dot{\mu}_i \dot{\mu}_j = N \dot{M} \dot{\Lambda} \dot{M}'. \quad (3.19)$$

Then Q_1 may be shown to follow the non-central chi-square law;

$$g(Q_1 | \dot{M}) = \sum_{s=0}^{\infty} \frac{e^{-\frac{\phi_1^2}{2}} \left(\frac{\phi_1^2}{2}\right)^s e^{-\frac{Q_1}{2}} (Q_1)^{\frac{2s+p-3}{2}}}{s! 2^{\frac{p-1}{2}+s} \Gamma\left(\frac{p-1}{2}+s\right)}. \quad (3.20)$$

For a table of the percentage points of the non-central chi-square distribution, see E. Fix, [10].

The usual likelihood ratio test of the hypothesis (3.8) without applying a transformation may be found in an analogous manner. Let μ be the common value of the means specified in (3.8). Maximization of the logarithm of the likelihood function with respect to μ leads to the

equation

$$-\sum_{i=1}^p \bar{x}_i \sum_{j=1}^p \lambda_{ij} + \lambda \sum_{i=1}^p \sum_{j=1}^p \lambda_{ij} = 0$$

yielding the estimate

$$\lambda = \frac{\sum_{i=1}^p \sum_{j=1}^p \lambda_{ij} \bar{x}_i}{\sum_{i=1}^p \sum_{j=1}^p \lambda_{ij}}. \quad (3.21)$$

On forming the likelihood ratio, we have

$$\frac{\exp \left\{ -\frac{1}{2} \text{tr} \Lambda V - \frac{N}{2} (\bar{X} - \hat{M}) \Lambda (\bar{X} - \hat{M})' \right\}}{\exp \left\{ -\frac{1}{2} \text{tr} \Lambda V \right\}} = \exp \left\{ -\frac{N}{2} (\bar{X} - \hat{M}) \Lambda (\bar{X} - \hat{M})' \right\} = \exp \left\{ -\frac{Q_2}{2} \right\}, \quad (3.22)$$

where $\hat{M} = [\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_p]$, $\hat{\mu}_i$ being defined in (3.21).

We take Q_2 as our test statistic. It is easily found that Q_2 follows the χ^2 law with $(p-1)$ degrees of freedom. We now show that Q_1 and Q_2 are identical.

Let the transformation be as before:

$$y_{ij} = x_{ij} - x_{ip} \quad j=1, \dots, p-1 \quad (3.10)*$$

and let

$$y_{ip} = x_{ip} \quad (3.23)$$

* repeated

It should be noted that the covariance matrix of the $[y_{1j}]$ as defined in (3.10)* and (3.23) is not the covariance matrix $\dot{\Sigma}$. However, $\dot{\Sigma}$ is the first $(p-1)^{\text{th}}$ principal minor in the covariance matrix of the $[y_{1j}]$ which we denote by $\ddot{\Sigma}$. It should be further noted that the first $(p-1)^{\text{th}}$ principal minor of $\ddot{\Lambda} = \ddot{\Sigma}^{-1}$ is not given by $\hat{\Lambda}$. If $\ddot{\lambda}_{1j}$ is an element of $\ddot{\Lambda}$, it can be shown [21] that $\hat{\Lambda}$ is related to $\ddot{\Lambda}$ by

$$\hat{\lambda}_{1j} = \ddot{\lambda}_{1j} - \frac{\ddot{\lambda}_{1p} \ddot{\lambda}_{pj}}{\ddot{\lambda}_{pp}}, \quad 1, j = 1, \dots, p-1. \quad (3.24)$$

We may write the transformation in matrix form as

$$Y_1 = X_1 T ; \quad i = 1, \dots, N, \quad (3.25)$$

where Y_1 is the p -columned row vector $[y_{11} \dots y_{1p}]$, X_1 is the p -columned row vector $[x_{11} \dots x_{1p}]$ and T is given by

$$T = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & & 1 & 0 \\ -1 & -1 & \dots & -1 & 1 \end{bmatrix}. \quad (3.26)$$

Let

$$j = [1, 1, \dots, 1], \quad (3.27)$$

a row vector of p unit elements. $\frac{Q_2}{N}$, from (3.22) may then be written as

$$\begin{aligned} \frac{Q_2}{N} &= (\bar{X} - \hat{M}) \wedge (\bar{X} - \hat{M})' \\ &= \bar{X} \wedge \bar{X}' - \hat{M} \wedge \bar{X}' - \bar{X} \wedge \hat{M}' + \hat{M} \wedge \hat{M}'. \end{aligned} \quad (3.28)$$

But $\hat{\mu}$, as defined in (3.21) may be written as

$$\hat{\mu} = \frac{\bar{X} \wedge j'}{j \wedge j'} = \frac{j \wedge \bar{X}'}{j \wedge j'} \quad (3.29)$$

and

$$\hat{M} = \hat{\mu} j. \quad (3.30)$$

Substitution of (3.30) in (3.28) leads to

$$\begin{aligned} \frac{Q_2}{N} &= \bar{X} \wedge \bar{X}' - 2 \hat{\mu} j \wedge \bar{X}' + \hat{\mu}^2 j \wedge j', \\ &= \bar{X} \wedge \bar{X}' - 2 \frac{(j \wedge \bar{X}')^2}{j \wedge j'} + \frac{(j \wedge \bar{X}')^2}{j \wedge j'}, \\ &= \bar{X} \wedge \bar{X}' - \frac{(j \wedge \bar{X}')^2}{j \wedge j'}. \end{aligned} \quad (3.31)$$

The matrix T has a determinant equal to 1 as can be seen from its form and thus possesses an inverse. Denoting the inverse relation for the vectors of sample means by

$$\bar{X} = \bar{Y} T^{-1} \quad (3.32)$$

we have

$$\frac{Q_2}{N} = \bar{Y} T^{-1} \wedge (T^{-1})' \bar{Y}' - \frac{[j \wedge (T')^{-1} \bar{Y}']^2}{j \wedge j'} \quad (3.33)$$

But the inverse of the covariance matrix of the Y's is given by

$$\ddot{\Lambda} = T^{-1} \Lambda (T')^{-1}, \quad (3.34)$$

and (3.33) is written as

$$\frac{Q_2}{N} = \bar{Y} \ddot{\Lambda} \bar{Y}' - \frac{(\mathbf{j}T \ddot{\Lambda} \bar{Y}')^2}{\mathbf{j}T \ddot{\Lambda} T' \mathbf{j}'}. \quad (3.35)$$

But

$$\mathbf{j}T = [0, 0, \dots, 0, 1], \quad (3.36)$$

and

$$\mathbf{j}T \ddot{\Lambda} T' \mathbf{j}' = \ddot{\lambda}_{pp}, \quad (3.37)$$

where $\ddot{\lambda}_{ij}$ is the ij^{th} element in $\ddot{\Lambda}$. Further,

$$\mathbf{j}T \ddot{\Lambda} \bar{Y}' = \sum_{i=1}^p \ddot{\lambda}_{pi} \bar{y}_i \quad (3.38)$$

and

$$\bar{Y} \ddot{\Lambda} \bar{Y}' = \sum_{i=1}^p \sum_{j=1}^p \ddot{\lambda}_{ij} \bar{y}_i \bar{y}_j. \quad (3.39)$$

Thus

$$\begin{aligned} \frac{Q_2}{N} &= \sum_{i=1}^p \sum_{j=1}^p \ddot{\lambda}_{ij} \bar{y}_i \bar{y}_j - \frac{(\sum \ddot{\lambda}_{pi} \bar{y}_i)^2}{\ddot{\lambda}_{pp}} \\ &= \sum_{i=1}^p \sum_{j=1}^p \left(\ddot{\lambda}_{ij} - \frac{\ddot{\lambda}_{ip} \ddot{\lambda}_{pj}}{\ddot{\lambda}_{pp}} \right) \bar{y}_i \bar{y}_j. \end{aligned} \quad (3.40)$$

From the relation (3.24) it is seen that we may write this as

$$\begin{aligned} \frac{Q_2}{N} &= \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \ddot{\lambda}_{ij} \bar{y}_i \bar{y}_j, \quad i, j = 1, \dots, p-1 \\ &= \frac{Q_1}{N}. \end{aligned} \quad (3.41)$$

Thus, in this instance, the application of Hsu's transformation leads to the same result as that obtained by direct application of the likelihood ratio procedure.

Case 2. Let us assume that the variance-covariance matrix has structure $\Sigma = \sigma^2 K$, K known. This is one generalization of the usual assumption of the analysis of variance that the covariance matrix has the form $\Sigma = \sigma^2 I$, where I is the p -square identity matrix. The Hsu transformation maps K into \dot{K} , and $\dot{\Sigma}$ into $\sigma^2 \dot{K}$, where the ij th element of \dot{K} is defined by

$$\dot{k}_{ij} = k_{ij} - k_{ip} - k_{pj} + k_{pp}, \quad i, j = 1, \dots, p-1.$$

The likelihood function in its general form may be written

$$[2\pi\sigma^2]^{-\frac{N(p-1)}{2}} |\dot{K}|^{-\frac{N(p-1)}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \mathbf{t}' \dot{K}^{-1} \mathbf{v} - \frac{N}{2\sigma^2} (\bar{y} - \dot{M}^0)' \dot{K}^{-1} (\bar{y} - \dot{M}^0) \right\}, \quad (3.42)$$

As in (3.16). Given the null hypothesis (3.14) true, maximization of the logarithm of the likelihood leads, with some reduction, to the maximizing equation

$$\frac{-N(p-1)}{2} \frac{1}{\sigma^2} + \frac{1}{2\sigma^2} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \dot{k}^{ij} \sum_{t=1}^N (y_{ti} - \dot{\mu}_i^0) (y_{tj} - \dot{\mu}_j^0) = 0$$

from which we obtain, on setting $\dot{M}^0 = 0$,

$$\hat{\sigma}^2 = \frac{1}{N(p-1)} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \dot{k}^{ij} \sum_{t=1}^N y_{ti} y_{tj}. \quad (3.43)$$

Under the general alternative hypothesis, we estimate \dot{M} by \bar{Y} and on maximizing the logarithm of the right hand member of (3.42) with respect to σ^2 , we obtain

$$\frac{-N(p-1)}{2} \frac{1}{\sigma^2} - \frac{1}{2\sigma^2} \text{tr } \dot{K}^{-1} \dot{V} = 0, \quad (3.44)$$

from which we obtain the estimator

$$s^2 = \frac{1}{N(p-1)} \text{tr } \dot{K}^{-1} \dot{V} = \frac{1}{N(p-1)} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} k^{ij} \dot{v}_{ij}. \quad (3.45)$$

Note that we use $\hat{\sigma}^2$ as the estimator of the variance when the null hypothesis is true and s^2 as the estimator under the general alternative hypothesis. On forming the likelihood ratio we find, after some simplification, the ratio

$$\left[\frac{s^2}{\hat{\sigma}^2} \right]^{\frac{N(p-1)}{2}}. \quad (3.46)$$

When the null hypothesis is true, $\frac{N(p-1)}{\sigma^2} \hat{\sigma}^2$ is the exponent in the density of Y , and is thus distributed as a χ^2 with $N(p-1)$ degrees of freedom. This quadratic form may be partitioned into two parts

$$\begin{aligned} & \frac{1}{\sigma^2} \sum_{t=1}^N \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} k^{ij} y_{ti} y_{tj} \\ &= \frac{1}{\sigma^2} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} k^{ij} \dot{v}_{ij} + \frac{N}{\sigma^2} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} k^{ij} \bar{y}_i \bar{y}_j. \end{aligned} \quad (3.47)$$

Application of Cochran's theorem [19] to (3.47) shows that the two terms on the right-hand side are independent, and further that the first term on the right-hand side is distributed as a χ^2 with $(p-1)(N-1)$ degrees of freedom; the second term on the right-hand side being distributed independently of the first term as a χ^2 with $(p-1)$ degrees of freedom. This can be demonstrated by examining the ranks of these two quadratic forms.

The first form is given by

$$\frac{1}{\sigma^2} \sum_{j=1}^{p-1} \sum_{i=1}^{p-1} k^{ij} \sum_{t=1}^N (y_{ti} - \bar{y}_i)(y_{tj} - \bar{y}_j), \quad (3.48)$$

and, on multiplying it out and collecting the matrices of the component quadratic forms, we obtain a $(p-1)N$ square matrix (with matricial elements) proportional to:

$$\begin{bmatrix} (N-1)K^{-1} & -K^{-1} & -K^{-1} & \dots & -K^{-1} \\ -K^{-1} & (N-1)K^{-1} & -K^{-1} & \dots & -K^{-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -K^{-1} & -K^{-1} & -K^{-1} & \dots & (N-1)K^{-1} \end{bmatrix}. \quad (3.49)$$

If we add to the left column all of the other columns, the first column vanishes. Then, working on the $(p-1)(N-1)$ principal minor we obtain, after some simplification,

$$\begin{bmatrix}
 0 & 0 & 0 & 0 & \dots & 0 & 0 \\
 0 & NK^{-1} & 0 & 0 & \dots & 0 & 0 \\
 0 & 0 & NK^{-1} & 0 & \dots & 0 & 0 \\
 \vdots & \vdots & & & & \vdots & \\
 \vdots & \vdots & & & & \vdots & \\
 \vdots & \vdots & & & & \vdots & \\
 0 & 0 & 0 & 0 & \dots & NK^{-1} & 0 \\
 0 & -NK^{-1} & -NK^{-1} & -NK^{-1} & \dots & -NK^{-1} & K^{-1}
 \end{bmatrix} . \quad (3.50)$$

From an examination of (3.50), it can be shown that it has rank $(N-1)(p-1)$. The other quadratic form in (3.47) is given by

$$\frac{N}{\sigma^2} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} k^{ij} \bar{y}_i \bar{y}_j$$

and it can be seen that the matrix of the quadratic form is proportional to

$$\begin{bmatrix}
 K^{-1} & K^{-1} & \dots & K^{-1} \\
 K^{-1} & K^{-1} & \dots & K^{-1} \\
 \vdots & \vdots & & \vdots \\
 \vdots & \vdots & & \vdots \\
 K^{-1} & K^{-1} & \dots & K^{-1}
 \end{bmatrix} \quad (3.51)$$

This matrix is easily shown to have rank $(p-1)$. The quadratic form on the left-hand side of (3.47) has rank $N(p-1)$. Thus, adding ranks of the forms on the right-

hand side, we obtain

$$(N-1)(p-1) + (p-1) = N(p-1)$$

which, by Cochran's theorem, demonstrates that the two component forms in (3.47) are independent and that they are distributed as χ^2 variates with $(p-1)(N-1)$ and $(p-1)$ degrees of freedom.

The $\frac{2}{N(p-1)}$ th root of our likelihood ratio (3.42) may be written as

$$\frac{N(p-1)s^2}{N(p-1)\hat{\sigma}^2} \cdot \frac{1}{\frac{1 + N \sum \sum k^{ij} \bar{y}_i \bar{y}_j}{\sum \sum k^{ij} v_{ij}}} \quad (3.52)$$

We take

$$Q_3 = (N-1) \frac{\left(\frac{N}{\sigma^2} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} k^{ij} \bar{y}_i \bar{y}_j \right)}{\left(\frac{1}{\sigma^2} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} k^{ij} v_{ij} \right)} \quad (3.53)$$

as our test statistic; it is easily seen that this has the F-distribution with $(p-1)$ and $(N-1)(p-1)$ degrees of freedom when the null hypothesis is true.

When the null hypothesis is not true, it can be shown that (3.53) follows the non-central F-distribution which here depends on the parameter

$$\phi_3 = \frac{N}{\sigma^2} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \hat{k}^{ij} \hat{\mu}_i \hat{\mu}_j \quad (3.54)$$

The non-central F-distribution has been tabled by P. C. Tang [11] .

We may interpret $\hat{\mu}_i$ in (3.54) as the deviation of the population mean of the i^{th} variate from the population mean of the p^{th} variate.

Consider the special case $K=I$. Then K has elements: $\hat{k}_{11} = k_{11} + k_{pp} = 2$; $\hat{k}_{ij} = k_{pp} = 1$. The elements of \hat{k}^{ij} may be shown to be $\hat{k}^{ii} = \frac{p-1}{p}$, $\hat{k}^{ij} = \frac{-1}{p}$ for $i \neq j$; $i, j = 1, \dots, p-1$. In this case, we obtain the statistic

$$Q_4 = \frac{N(N-1) \left[\sum_{i=1}^{p-1} \bar{y}_i^2 - \frac{1}{p} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \bar{y}_i \bar{y}_j \right]}{\sum_{i=1}^{p-1} \hat{v}_{ii} - \frac{1}{p} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \hat{v}_{ij}} \quad (3.55)$$

which may be written in terms of the original observations as

$$Q_4 = \frac{N(N-1) \left[\sum_{i=1}^p \bar{x}_i^2 - \frac{1}{p} \sum_{i=1}^p \sum_{j=1}^p \bar{x}_i \bar{x}_j \right]}{\sum_{i=1}^p v_{ii} - \frac{1}{p} \sum_{i=1}^p \sum_{j=1}^p v_{ij}}$$

or

$$Q_4 = \frac{N(N-1) \left\{ \sum_1^p \bar{x}_i^2 - \frac{1}{p} \left(\sum_{i=1}^p \bar{x}_i \right)^2 \right\}}{\sum_1^p v_{ii} - \frac{1}{p} \sum_1^p \sum_1^p v_{ij}}. \quad (3.56)$$

When $p = 2$, the statistic reduces to

$$F = \frac{N(N-1) (\bar{x}_1 - \bar{x}_2)^2}{v_{11} + v_{22} - 2v_{12}} \quad (3.57)$$

which is distributed according to the F-distribution with 1 and $N-1$ degrees of freedom. It is well known that the square of the Student-t statistic, with ϕ degrees of freedom, follows the F-distribution with 1 and ϕ degrees of freedom. We may identify (3.57) as the square of the Student-t statistic for testing equality of means in the paired sample case by noting that $\frac{v_{11} + v_{22} - 2v_{12}}{N-1}$ estimates $\text{Var}(x_1 - x_2)$, and is the usual estimate of that variance.

The test presented was based on Hsu's transformation, which has two desirable properties. First, the composite null hypothesis (3.8) is transformed (at least in the means space) into a simple hypothesis. The second advantage will become clear in section (3.1). However, from a strict multivariate standpoint, we may examine the effects of

making this transformation by constructing the best test of the hypothesis of homogeneity directly, and comparing powers of the two methods. It is not patently obvious, at this point, that such a test would be either the same or different from the test developed above. By the likelihood ratio test procedure, we find a statistic

$$Q_5 = \frac{N_P (N-1) \left\{ \sum_{i=1}^P \sum_{j=1}^P k^{ij} \bar{x}_i \bar{x}_j - \frac{\sum_{i=1}^P \sum_{j=1}^P k^{ij} \bar{x}_i^2}{\sum_{i=1}^P \sum_{j=1}^P k^{ij}} \right\}}{(P-1) \sum_{i=1}^P \sum_{j=1}^P k^{ij} v_{ij}}, \quad (3.58)$$

which has the F-distribution with $(p-1)$ and $p(N-1)$ degrees of freedom. When $K=I$, this reduces to

$$Q_6 = \frac{N_P (N-1)}{P-1} \frac{\left\{ \sum_{i=1}^P \bar{x}_i^2 - \frac{(\sum_{i=1}^P \bar{x}_i)^2}{P} \right\}}{\sum_{i=1}^P v_{ii}}, \quad (3.59)$$

and for $p=2$, we have

$$\frac{N(N-1) (\bar{x}_1 - \bar{x}_2)^2}{v_{11} + v_{22}}, \quad (3.60)$$

which has the F-distribution with 1 and $2N-2$ degrees of freedom. This may be identified as the square of the Student-t statistic based on two non-paired samples with the same number of observations in each sample. Thus, application of the transformation leads, in these special cases, to generalizations of the paired t-statistic, while straightforward application of the likelihood ratio procedure leads to a generalization of the non-paired t-statistic. In terms of the analysis of variance, the statistic (3.59) is the statistic used in a one-way classification, while (3.56) is the statistic used in a two-way classification.

Case 3. Consider now the case wherein the variance-covariance matrix has a structure of the form: $\sigma_{11} = \sigma^2$, $i = 1, \dots, p$; $\sigma_{1j} = \rho \sigma^2$, $i \neq j$, $i, j = 1, \dots, p$. The application of Hsu's transformation yields a covariance matrix with elements: $\dot{\sigma}_{11} = 2 \sigma^2(1-\rho)$, $i = 1, \dots, p-1$; $\dot{\sigma}_{1j} = \sigma^2(1-\rho)$, $i \neq j$, $i, j = 1, \dots, p-1$. Thus this case reduces to Case 2, with $\dot{k}_{11} = 2$, $\dot{k}_{1j} = 1$ ($i \neq j$). As these constants are the same as the ones developed in the last section under the assumption that $K = I$, we may use the expression (3.56) as the test statistic. It is a rather interesting fact that this should be so, for the statistic (3.56) was based on the model $\Sigma = I \sigma^2$, and we now find the same statistic arising when we use the model postulated above.

The distribution under the alternative hypothesis is given by the non-central F-distribution, which depends upon the parameter (3.54). As the constants k^{ij} are known in this case, we write (3.54) as

$$\begin{aligned} \phi_3^2 &= \frac{N}{\sigma^2(1-p)} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} k^{ij} \mu_i \mu_j \\ &= \frac{N}{\sigma^2(1-p)} \left\{ \sum_{i=1}^p \mu_i^2 - \left(\frac{\sum_{i=1}^p \mu_i}{p} \right)^2 \right\} \end{aligned} \quad (3.61)$$

Wilks [12] developed a test under the same hypothesis, without transformation, and termed the statistic so found L_m . In our notation,

$$L_m = \frac{\frac{1}{Np} \left\{ \sum_{i=1}^p v_{ii} - \frac{1}{p-1} \sum_{i \neq j}^p v_{ij} \right\}}{\frac{1}{Np} \left\{ \sum_{i=1}^p v_{ii} - \frac{1}{p-1} \sum_{i \neq j}^p v_{ij} \right\} + \frac{1}{p-1} \sum_{i=1}^p (\bar{x}_i - \bar{\bar{x}})^2} \quad (3.62)$$

where

$$\bar{\bar{x}} = \frac{1}{p} \sum_{i=1}^p \bar{x}_i .$$

The range of L_m is from 0 to 1, and L_m , given the null hypothesis true, has the β -distribution with parameters $\frac{1}{2}(N-1)(p-1)$ and $\frac{1}{2}(p-1)$. As noted by Wilks, the L_m

criteria may be rewritten as an F-statistic by considering

$$F = \frac{\frac{1}{2}(N-1)(p-1)(1-L_m)}{\frac{1}{2}(p-1)L_m} \quad (3.63)$$

and then F has the F -distribution with $(p-1)$ and $(N-1)(p-1)$ degrees of freedom. Wilks concludes that... "The use of L_m as a criterion for testing H_m (the hypothesis of homogeneity of means) is equivalent to an analysis of variance test for testing 'row' effects in a $p \times N$ rectangular layout when rows are associated with the p variables in the multivariate population, and columns are associated with the N individuals in the sample." Writing the statistic obtained by Wilks in the form of F (3.63), it is easily seen that this statistic is identical with the statistic (3.56).

Case 4. We now consider the case where the variance-covariance matrix has the structure: σ_{ii} possibly different for different i , $\sigma_{ij} = 0$ ($i \neq j$). The application of Hsu's transformation leads to a transformed variance-covariance matrix of size $(p-1) \times (p-1)$ with elements: $\dot{\sigma}_{ii} = \sigma_{ii} + \sigma_{pp}$ ($i = 1, \dots, p-1$) and $\dot{\sigma}_{ij} = \sigma_{pp}$ ($i \neq j$), $i, j = 1, \dots, p-1$. Thus, the transformed variance-covariance matrix has possibly different diagonal elements, and constant covariance elements.

The straightforward application of the likelihood ratio test procedure leads immediately to a basic difficulty

in the estimation of the variance-covariance elements by the method of maximum likelihood. Under the null hypothesis, we write the partial derivative of the logarithm of the likelihood function with respect to $\dot{\sigma}_{kk}$ as

$$\frac{\partial \ln f}{\partial \dot{\sigma}_{kk}} = \frac{-N}{2} \dot{\lambda}_{kk} - \frac{1}{2} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \frac{\partial \dot{\lambda}_{ij}}{\partial \dot{\sigma}_{kk}} \sum_{r=1}^N x_{ri} x_{rj} = 0 \quad (3.64)$$

Since

$$\dot{\lambda}_{ij} = \frac{\delta_{ij}}{\dot{\sigma}_{ii} - \dot{\sigma}^2} - \frac{\dot{\sigma}^2}{(\dot{\sigma}_{ii} - \dot{\sigma}^2)(\dot{\sigma}_{jj} - \dot{\sigma}^2) \left[1 + \dot{\sigma}^2 \sum_{k=1}^{p-1} (\dot{\sigma}_{kk} - \dot{\sigma}^2)^{-1} \right]} \quad (3.65)$$

where δ_{ij} is Kronecker's delta and $\dot{\sigma}^2 = \dot{\sigma}_{ij}$ for all $i \neq j$ ($i, j = 1, \dots, p-1$),

$$\frac{\partial \dot{\lambda}_{ij}}{\partial \dot{\sigma}_{kk}} = -\dot{\lambda}_{ik} \dot{\lambda}_{jk} \quad i, j = 1, \dots, p-1. \quad (3.66)$$

Then we are led to the set of $p-1$ estimating equations

$$-N \dot{\lambda}_{kk} + \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \dot{\lambda}_{ik} \dot{\lambda}_{jk} \sum_{r=1}^N x_{ri} x_{rj} = 0; \quad k=1, \dots, (p-1) \quad (3.67)$$

The estimation equation obtained when the logarithm of the likelihood function is partially differentiated with respect

to $\dot{\sigma}^2$ is quite involved. The author has not been able to clear the expression $\frac{\partial \lambda_{ij}}{\partial \dot{\sigma}^2}$ of terms in $\dot{\sigma}_{11}$ and $\dot{\sigma}^2$. This set of equations was studied by Russell [13] who has also failed to find a solution for $p \geq 4$. The solutions found in the case $p = 4$ are in closed form, but are rather involved from a computational standpoint.

case 5. Finally, we assume that the variance-covariance matrix must be estimated completely from the sample; no a priori information or assumptions about possible relationships within the matrix are postulated. Application of Hsu's transformation to the problem reduces the hypothesis of homogeneity of means to an hypothesis that the transformed means are zero. In this situation, as pointed out by Hsu [9], we may use the well known likelihood ratio statistic known as Hotelling's T^2 statistic [14], where

$$T^2 = N \bar{Y} \dot{V}^{-1} \bar{Y}, \quad (3.68)$$

\bar{Y} being the row vector of transformed means,

$$\bar{Y} = [\bar{y}_1] = [\bar{x}_1 - \bar{x}_p] ; \quad i = 1, \dots, p-1$$

and \dot{V} denoting the matrix of sample products and cross products, with

$$\begin{aligned} \dot{v}_{ij} &= \sum_{\alpha=1}^n (y_{\alpha i} - \bar{y}_i)(y_{\alpha j} - \bar{y}_j) \\ &= v_{ij} - v_{ip} - v_{jp} + v_{pp} ; \quad i, j = 1, \dots, p-1. \end{aligned}$$

It is well known that $\frac{N-p+1}{p-1} T^2$ follows the F-distribution with $p-1$ and $N-p+1$ degrees of freedom, given the null hypothesis true. When the null hypothesis is not true, $\frac{N-p+1}{p-1} T^2$ has the non-central F-distribution depending upon a parameter

$$\phi^2 = \frac{N}{p} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \lambda_{ij} \dot{\mu}_i \dot{\mu}_j, \quad (3.69)$$

with $p-1$ and $N-p+1$ degrees of freedom.

In the other cases, we have derived the best test of the hypothesis of homogeneity without performing the transformation, where possible. It does not seem possible to do so in this case; in fact, Hsu's transformation was introduced in this particular case to give a method of testing the hypothesis of homogeneity (3.8) when the covariance matrix is perfectly general.

We should note that T^2 , as defined by (3.68) is not the T^2 as developed by Hotelling. If \dot{T}^2 is the statistic developed by Hotelling, and T^2 is the statistic defined in (3.68), then

$$\dot{T}^2 = (N-1)T^2. \quad (3.70)$$

When $p = 2$, we have

$$T^2 = N(\bar{x}_1 - \bar{x}_2) \left[v_{11} + v_{22} - 2v_{12} \right]^{-1} (\bar{x}_1 - \bar{x}_2)$$

$$= \frac{N (\bar{x}_1 - \bar{x}_2)^2}{v_{11} + v_{22} - 2v_{12}} \quad (3.71)$$

And $(N-1)T^2$ can be seen to be the square of the paired two-sample t-statistic, following the F-distribution with 1 and $N-1$ degrees of freedom.

3.3 Discussion

It should be noted that the tests developed herein were developed for an additive model of the form

$$x_{ik} = \mu_i + \varepsilon_{ik} \quad , \quad i=1, \dots, p; \quad k=1, \dots, N \quad (3.72)$$

In the transformation, we subtract values of x with k held constant; thus, the same final form would be obtained if another component, only depending on k (perhaps a block effect) were added to the model. Denoting this component by β_k , the expanded model would be

$$x_{ik} = \mu_i + \beta_k + \varepsilon_{ik} \quad (3.73)$$

This holds true if β_k is either a fixed or random component.

In case 2 with $\Sigma = K\sigma^2$, we obtained the test statistic Q_3 (3.53) derived by use of the transformation and found that when $K=I$ and $p=2$, Q_3 reduced to Q_4 (3.56), the paired t-statistic. We then obtained the statistic Q_5 (3.58) derived without the transformation and found that, when $t=2$,

and $K = I$, Q_5 reduced to the two sample t statistic. In case 4, with $\sigma_{11} = \sigma^2$; $\sigma_{1j} = \rho\sigma^2$, we derived the test statistic by use of the transformation and found that it was identical with Q_4 . Wilks's test statistic (3.62), derived under these same covariance assumptions, but without use of the transformation was found to be identical with Q_4 . Thus, since the use of the transformation produces statistics which are valid when a block effect is present, we see that the usual test of homogeneity of treatments in an analysis of variance of a two-way layout is valid also when all treatments are equally correlated.

For the case $p = 2$, when $\Sigma = I\sigma^2$ and when block effects are absent, the assumption of possible correlations leads to the paired t-test when the appropriate one is the two sample t-test. When block effects are present and $\Sigma = I\sigma^2$, the use of the Hsu transformation is still correct and leads to a test equivalent to the paired sample t-test. Again, when Σ is general, the paired sample t-test is appropriate whether block effects are present or absent.

IV APPLICATION OF THE THEORY TO DESIGN OF COMPLETE
BLOCK EXPERIMENTS

4.1 Graybill [15] seems to have been the first to note that Hsu's transformation could be applied to the Randomized Block Design when conditions arise which make the usual assumptions untenable. The model for fixed effects is usually given as

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}, \quad i=1, \dots, p; \quad j=1, \dots, r, \quad (4.1)$$

where y_{ij} is the observed value on the i^{th} treatment in the j^{th} block,

μ is the overall mean,

τ_i is the (fixed) added effect of the i^{th} treatment,

$$\text{with } \sum_{i=1}^p \tau_i = 0,$$

β_j is the (fixed) added effect of the j^{th} block, with

$$\sum_{j=1}^r \beta_j = 0,$$

and ε_{ij} is the random normal error, with

$$E [\varepsilon_{ij}] = 0, \quad E [\varepsilon_{ij} \varepsilon_{km}] = \delta_{ik} \delta_{jm} \sigma^2, \quad (4.2)$$

where δ_{ik} , δ_{jm} are Kronecker deltas.

To test the hypothesis

$$H_0: \tau_1 = \tau_2 = \dots = \tau_p, \quad (4.3)$$

we proceed as in forming the variate

$$x_{ij} = y_{ij} - y_{pj}, \quad i=1, \dots, p-1. \quad (4.4)$$

It is seen that, under this transformation, the model for the new variables is

$$\begin{aligned}
 x_{ij} &= \mu + \tau_i + \beta_j + \varepsilon_{ij} - \mu - \tau_p - \beta_j - \varepsilon_{pj} \\
 &= (\tau_i - \tau_p) + (\varepsilon_{ij} - \varepsilon_{pj}), \quad (4.5)
 \end{aligned}$$

and this is of the same form as given in (3.9). It is seen in (4.5) that the transformation both reduces the hypothesis to a simpler one, and removes block effects from the model.

Graybill applied the transformation only to Case 5 for a randomized block design. In making this transformation, he seemed to be mainly concerned with variance heterogeneity and not with correlated errors. Thus, he was mainly concerned with Case 4, but operated as though it were a Case 5 problem. Thus, Hotelling's T^2 as defined in (3.68) is used as the test statistic. It has been seen here that T^2 follows the F-distribution, with $p-1$ and $N-p+1$ degrees of freedom. It should be noted that the restriction that $N > p$ requires as many or more blocks as treatments.

The hypothesis of equality of fixed block effects is usually checked in the analysis of variance of a randomized block design. It can not be done by these methods, for if a variate from any other observation vector is subtracted from a variate in the first observation vector, then the two vectors are correlated, and independence no longer holds between observations.

The usual analysis of the randomized block design assumes a variance-covariance matrix of the form given by

Case 2, when the known matrix K is the identity matrix I . The statistic in that case is given by (3.56), and it has been easily verified that (3.56) is the same statistic usually used in testing homogeneity of means in a randomized block design. Thus, the application of Hsu's transformation under the usual assumptions leads to the same results as the straightforward method of derivation. Also it should be noted that, when all variates are equally correlated as in Case 3, the statistic (3.56) is also applicable for the analysis of a randomized block design.

Graybill, in his original paper, stressed the point that this analysis was valid for heterogeneous variances. The work by Box [16] and others would seem to indicate that the fact that this method of analysis is also valid under the presence of correlations in the errors is of much more importance.

4.2 Power of the Analysis. It has been seen that the assumptions associated with the usual methods of analysis are not made in general in the tests presented herein. As reductions in the assumptions in tests of significance usually lead to loss of power and efficiency, it is natural to inquire as to the size of the loss incurred.

The first comparison^{and only (V.S.)} we shall make concerns the loss in power incurred by assuming that the covariance matrix is perfectly general (Case 5) when actually the covariance

matrix has the form $\Sigma = I\sigma^2$ (as in the usual analysis).

The parameter of non-centrality for Case 5, as given by (3.69), is

$$\phi_4^2 = \frac{N}{p} \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} \lambda_{ij} \mu_i \mu_j = \frac{N}{p} \dot{M} \dot{\Lambda} \dot{M}' \quad (4.6)$$

where \dot{M} is the row vector of true means $[\mu_i]$ of the transformed variates. It shall now be shown that the parameter of non-centrality (4.6) is the same as the one usually associated with the randomized block design, when $\Sigma = I\sigma^2$. Hsu's transformation may be written in the matrix notation

$$Y = X T \quad (4.7)$$

where Y is the $(p-1)$ columned row vector of transformed variates, X is the (p) columned row vector of original variates, and

$$T = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -1 & -1 & -1 & \dots & -1 \end{bmatrix}, \quad (4.8)$$

a p by $(p-1)$ matrix. This form of the transformation is different from (3.26).

If J denotes a $(p-1)$ columned row vector $J = [1, 1, \dots, 1]$, then T may be written

$$T' = [I' ; -J'] \quad (4.9)$$

where I is the $(p-1)$ square identity matrix. If the untransformed covariance matrix is Σ , then (4.6) may be written as

$$\frac{N}{p} \dot{M} \dot{\Lambda} \dot{M}' = \frac{N}{p} M T (T' \Sigma T)^{-1} T' M' \quad (4.10)$$

where M is now the row vector of original means $[\gamma_1]$.

Under the assumption that $\Sigma = I \sigma^2$, we have

$$\begin{aligned} (T' \Sigma T)^{-1} &= \frac{1}{\sigma^2} (T' T)^{-1} = \frac{1}{\sigma^2} (I + J' J)^{-1} \\ &= \frac{1}{\sigma^2} \left(I - \frac{1}{p} J' J \right) \end{aligned} \quad (4.11)$$

Thus

$$\begin{aligned} \frac{N}{p} \dot{M} \dot{\Lambda} \dot{M}' &= \frac{N}{p \sigma^2} M T \left(I - \frac{1}{p} J' J \right) T' M' \\ &= \frac{N}{p \sigma^2} M \left(T T' - \frac{1}{p} T J' J T' \right) M \end{aligned} \quad (4.12)$$

But

$$T T' - \frac{1}{p} T J' J T' = \begin{bmatrix} I - \frac{1}{p} J' J & -\frac{1}{p} J' \\ -\frac{1}{p} J & \frac{p-1}{p} \end{bmatrix} \quad (4.13)$$

If we partition M between the $(p-1)^{\text{th}}$ and p^{th} element, denoting this by

$$M = [\gamma_m \quad \gamma_p]$$

then

$$\frac{N}{p} \dot{M} \dot{\Lambda} \dot{M}' = \frac{N}{p \sigma^2} (m m' - \frac{1}{p} m J' J m' - \frac{1}{p} \gamma_p J m' - \frac{1}{p} m J' \gamma_p + \gamma_p^2 - \frac{1}{p} \gamma_p^2). \quad (4.14)$$

Since $\sum_{1}^p \gamma_1 = 0$, then $J m' = \gamma_p$, and (4.14) reduces to

$$\begin{aligned} \frac{N}{p} \dot{M} \dot{\wedge} \dot{M}' &= \frac{N}{p \sigma^2} \left\{ [mm' + \gamma_p] - \frac{1}{p} [mJ' + \gamma_p]^2 \right\} \\ &= \frac{N}{p \sigma^2} \sum_{i=1}^p \gamma_i^2. \end{aligned} \quad (4.15)$$

But $\frac{N}{p \sigma^2} \sum_{i=1}^p \gamma_i^2$ is the parameter of non-centrality associated with the usual analysis of variance test. Thus, the test using T^2 and the usual test have the same parameters of non-centrality when $\Sigma = I \sigma^2$, and comparisons may be made in terms of this parameter. Both statistics are distributed as F , the T^2 having $(p-1)$ and $(N-p+1)$ degrees of freedom and the usual statistic having $(p-1)$ and $(N-1)(p-1)$ degrees of freedom. The ratio of the power of the T^2 test to the power of the analysis of variance test for a randomized block experiment with five treatments is shown in Figure 1 for an $\alpha = .05$ significance level test, and in Figure 2 for an $\alpha = .01$ significance level test. Several points should be noted:

- (1) The power ratio is less than unity, implying that under conditions when $\Sigma = I \sigma^2$, the T^2 test (Case 5) will not detect differences as often as the analysis of variance test.
- (2) As the number of blocks increases, the power ratio approaches unity, and is asymptotically equal to unity.

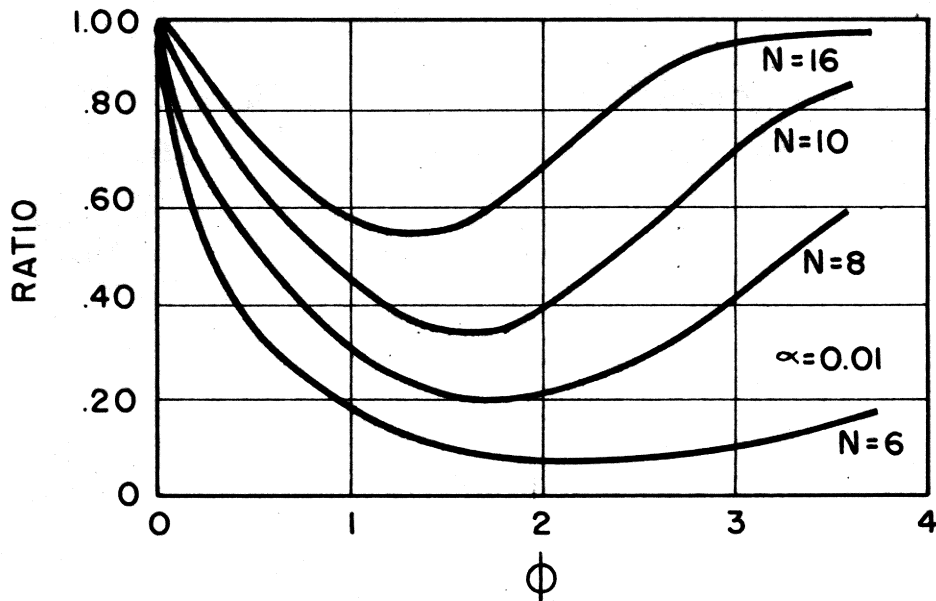
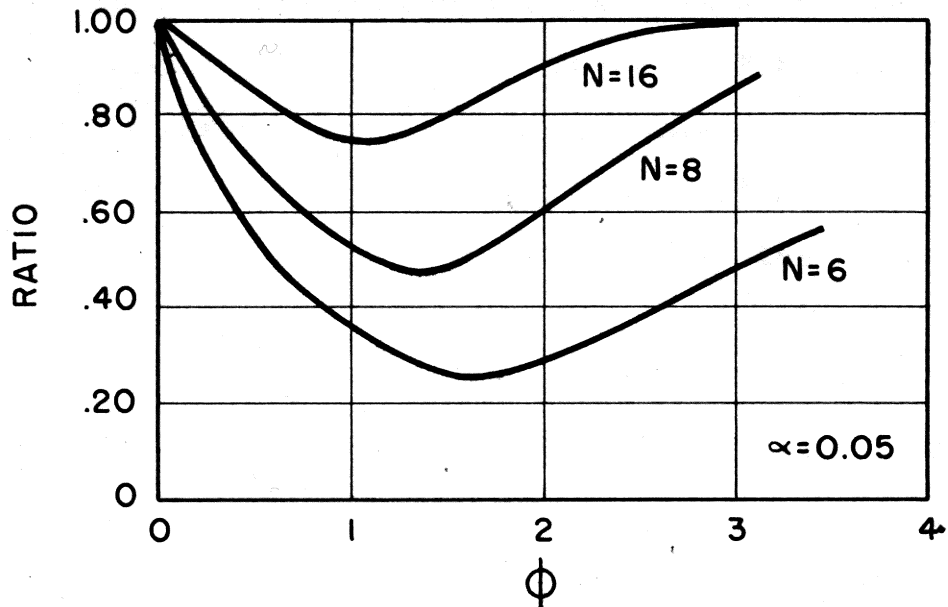


FIG. I—RATIO OF POWER OF T^2 ANALYSIS TO POWER OF USUAL ANALYSIS OF RANDOMIZED BLOCK DESIGN, FOR $P=5$

- (3) The minimum ratio of the power in both cases occurs generally between $\phi = 1$ and $\phi = 2$. For 6 blocks, 5 treatments, $\alpha = .01$, the minimum ratio is .08 at $\phi = 2$. This is not surprising considering that the degrees of freedom associated with the T^2 statistic are 2 and for the analysis of variance, 20.
- (4) Further examinations show that less power would be lost for a decreased number of treatments, and more power would be lost for an increased number of treatments, given equal numbers of blocks and equal size parameter ϕ . When $p = 2$, the two methods have equal power. This is shown by the equality of the statistics (3.57) and (3.71).

4.3 Single Degree of Freedom Comparisons

Single degree of freedom contrasts may be tested in the randomized blocks case as follows. Suppose that the experimenter is interested in some contrast

$$C_1 = \sum_1^p k_1 \tau_1 = KM' \quad (4.16)$$

where K is the row vector $[k_1]$ obeying the restriction $\sum k_1 = KJ' = 0$, with J defined in (4.9). M is the row vector $[\tau_1]$. For each block, we may form the statistic

$$\hat{C}_1 = K\bar{Y}' \quad (4.17)$$

where \bar{Y} is the row vector of sample means of the p treat-

ments. It is easily seen that

$$\begin{aligned} E [\hat{C}_1] &= KE [\bar{Y}'] = K[\mu J' + M' + \sum_1^N \beta_j J'] \\ &= \mu KJ' + M' \\ &= M' \end{aligned} \quad (4.18)$$

since $\sum_1^N \beta_j = 0$. The variance of \hat{C}_1 is

$$\text{Var} [\hat{C}_1] = K \sum K' \quad (4.19)$$

which we estimate by

$$\text{Var} [\hat{C}_1] = K \hat{\sum} K'. \quad (4.20)$$

The hypothesis $H_0: KM' = 0$ may be tested by forming the statistic

$$\frac{\sqrt{N} K \bar{Y}'}{[K \hat{\sum} K']^{1/2}} = t_1 \quad (4.21)$$

which follows the Student-t distribution with $(N-1)$ degrees of freedom.

Of course, if we have another contrast

$$C_2 = H M' \quad (4.22)$$

with $H J' = 0$, then

$$\text{Cov} (C_1 C_2) = K \sum H' \quad (4.23)$$

which will not, in general, be zero even if $K H' = 0$. In general, orthogonal contrasts do not produce independent tests, and without a knowledge of the true covariance matrix it would seem impossible to construct independent single degree of freedom contrasts. Thus, for cases 1 and 2, where Σ is either completely known or a matrix K proportional to Σ is known, we may find orthogonal contrasts which are uncorrelated and, due to normality, independent. An important exception to this is where all variance elements are identical, and all covariance elements are identical. Then, we may write the covariance matrix as

$$\Sigma = \sigma^2 [(1 - \rho) I + \rho J' J] \quad (4.24)$$

where ρ is the intraclass correlation coefficient, and σ^2 is the common variance term. Then, if $K H' = 0$, we have

$$\begin{aligned} K \Sigma H' &= \sigma^2 K [(1 - \rho) I + \rho J' J] H' \\ &= \sigma^2 [(1 - \rho) K H' + \rho (K J')(J H')] \\ &= 0. \end{aligned} \quad (4.25)$$

Since this case is analogous to the usual analysis of variance, we may compare the square of the linear contrast with the usual error term obtained in the analysis of variance.

4.4 Latin Square. A Latin Square design is one in which every treatment appears once and only once at each level

of two factors, usually called rows and columns. As an example of such a design, consider the following 3 x 3 design, where A, B, C represent the three different treatments:

	Col. 1	Col.2	Col. 3
Row 1	A	B	C
Row 2	B	C	A
Row 3	C	A	B

The methods used for analyzing such a design must depend upon the assumptions made. Two examples will be given herein.

If x_{ijklm} is the observed value of the treatment (i) which appears in the j^{th} row and k^{th} column of the m^{th} square, then the model is

$$x_{ijklm} = \mu + \tau_i + \beta_j + \alpha_k + \gamma_m + \varepsilon_{ijklm}; \quad i, j, k = 1, \dots, r \\ m = 1, \dots, N \quad (4.26)$$

where

μ is the grand mean over the whole experiment,

τ_i is the added effect of the i^{th} treatment with

$$\sum_{i=1}^r \tau_i = 0,$$

β_j is the added effect of the j^{th} row, with $\sum_{j=1}^r \beta_j = 0,$

α_k is the added effect of the k^{th} column, with

$$\sum_1^r \alpha_k = 0,$$

ν_m is the added effect of the m^{th} square, with

$$\sum_{m=1}^N \nu_m = 0.$$

Associated with the vector of errors $[\varepsilon_{ijkl}]$ we shall assume a general covariance matrix Σ . In the usual models, $\Sigma = \sigma^2 I$.

To test the hypothesis $H_0: \tau_1 = \tau_2 = \dots = \tau_r$, we first form the set $[x_{i..m}]$, where

$$\begin{aligned} x_{i..m} &= \sum_{\substack{j \\ k \\ (jk)\text{with } i}} x_{ijklm} \\ &= r\mu + r\tau_i + r\nu_m + \sum_{\substack{j \\ k \\ (jk)\text{with } i}} \varepsilon_{ijklm}. \end{aligned} \quad (4.27)$$

The notation, $[(jk)\text{with } i]$, means summing over all j and k where i appears. On forming the set $[x_{i..m} - x_{r..m}] = [y_{i..m}]$, we see that

$$y_{i..m} = r(\tau_i - \tau_r) + \sum_{\substack{j \\ k \\ (jk)\text{with } i}} \varepsilon_{ijklm} - \sum_{\substack{j \\ k \\ (jk)\text{with } r}} \varepsilon_{rjkm} \quad (4.28)$$

thus, the test of the hypothesis of treatment means is made

by testing the hypothesis that the $[y_{1..m}]$ have a set of means all zero.

\sum_y is formed in the usual way from the matrix of $[y_{1..m}]$ under Case 5. For the other cases, \sum_y is formed from the known forms of \sum_x .

It is obvious that row and column effects may be tested in an analogous manner, summing over all of the factors except the one being tested, and performing a subtraction to reparametrize the model.

If the original vector X_m has associated with it a variance-covariance matrix with constant covariance and constant variance elements, then it can be shown that the tests of equality of row means, column means and treatment means are uncorrelated pairwise, as in the case of single degree of freedom contrasts.

It is, of course, often possible to design an experiment so that the restriction that the number of squares exceeds the size of a square may be dropped. Thus, if it is possible to arrange the square so that the correlations between the major classifications are zero, tests of the homogeneity of the major classification means may be made by means of the usual analysis of variance, retaining the advantages of the Latin square. As an example, suppose that it is desired to test the effectiveness of an attachment to some machine tool, and it has been determined that

the errors associated with each attachment are correlated in time. Then, a Latin square design could be used, as shown below:

Let M_i denote the i^{th} machine tool, r machine tools being available in all. Let T_j represent the different time periods, $j = 1, \dots, r$, and let A_k denote the r different attachments. We could then use (for $r = 3$, say) the following design.

	T_1	T_2	T_3
A_1	M_1	M_2	M_3
A_2	M_2	M_3	M_1
A_3	M_3	M_1	M_2

The totals for A_1, A_2 and A_3 will be uncorrelated under these conditions, and only two squares would be required, as a minimum, analyzing the totals by a two way classification (considering squares as blocks).

4.5 Analysis of Factorial Arrangements. We will consider the analysis of factorial arrangements in two sections. First, the analysis of a $p \times q$ arrangement in randomized blocks will be discussed, and extensions to $p \times q \times r \times s \times \dots$ arrangements will be given. We will next consider confounded arrangements, fractional replicates, and balanced confounded arrangements. The analyses of these last three types of arrangements are given in terms of single degree of freedom

contrasts, for in the majority of the cases where factorial arrangements are used, the analysis of single degree of freedom contrasts seems most valuable.

The p x q Factorial Arrangement. We are concerned herein with the relations between variations in response and changes in the levels of two factors, say A and B, in an experiment; the experiment being run to p levels of A and q levels of B in each block, giving pq treatment combinations per block. We assume the usual model

$$x_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + \varepsilon_{ijk}; \quad i = 1, \dots, p \quad (4.29)$$

$$j = 1, \dots, q$$

$$k = 1, \dots, N$$

where μ = grand mean of experiment,

α_i = added effect of the i^{th} level of A, with $\sum_{i=1}^p \alpha_i = 0$,

β_j = added effect of the j^{th} level of B, with

$$\sum_{j=1}^q \beta_j = 0,$$

γ_k = added effect of the k^{th} block, with $\sum_{k=1}^N \gamma_k = 0$,

$(\alpha\beta)_{ij}$ = interaction at the i^{th} level of A and j^{th} level of B, with $\sum_i (\alpha\beta)_{ij} = \sum_j (\alpha\beta)_{ij} = 0$

ε_{ijk} = random error associated with the combination (ij) in the k^{th} block.

As before, we make special assumptions concerning the random error term. The usual assumption is that

$$\text{Cov} [\varepsilon_{ijk} \varepsilon_{rst}] = \delta_{kt} \delta_{ir} \delta_{js} \sigma^2, \quad (4.30)$$

where δ_{ij} is the Kronecker delta; $\delta_{ij} = 0$ if $i \neq j$;

$\delta_{ij} = 1$ if $i = j$. In the most general case (Case 5), we assume

$$\text{Cov} [\varepsilon_{ijk} \varepsilon_{rst}] = \delta_{kt} \sigma_{ijrs}, \quad (4.31)$$

and in the case of equally correlated variates (Case 3)

we assume

$$\text{Cov} [\varepsilon_{ijk} \varepsilon_{rst}] = \delta_{kt} [\delta_{ir} \delta_{js} \sigma^2 + (1 - \delta_{ir} - \delta_{js}) \rho \sigma^2]. \quad (4.32)$$

*Should be a two, I think
W.A. Lawrence*

We consider tests of significance of the A-factor, the B-factor and possible interaction between them. The hypotheses to be tested are:

- (a) $H_{01}: \alpha_1 = \alpha_2 = \dots = \alpha_p$
- (b) $H_{02}: \beta_1 = \beta_2 = \dots = \beta_q$
- (c) $H_{03}: (\alpha\beta)_{11} = (\alpha\beta)_{12} = \dots = (\alpha\beta)_{1p} = \dots = (\alpha\beta)_{pq}$

To test the first hypothesis, we first sum over the different levels of the B-factor for each A-factor. Letting $X_{i.k}$ denote the sum of x_{ijk} over j , we have:

$$\begin{aligned} X_{i.k} &= \sum_{j=1}^q x_{ijk} = q\mu + q\alpha_i + \sum_j \beta_j + \sum_j (\alpha\beta)_{ij} + q\gamma_k + \sum_j \varepsilon_{ijk} \\ &= q\mu + q\alpha_i + q\gamma_k + \sum_j \varepsilon_{ijk}, \end{aligned} \quad (4.33)$$

due to the restrictions in the model. We then perform Hsu's transformation, which yields:

$$Y_{1,k} = X_{1,k} - X_{p,k} = q(\alpha_1 - \alpha_p) + \sum_j \epsilon_{ijk} - \sum_j \epsilon_{pjk}; \quad i = 1, \dots, p-1. \quad (4.34)$$

As we assume the vector $[\epsilon_{ijk}]$ to be distributed as a multivariate normal vector, then the row vector of $(p-1)$ elements $[\sum_j \epsilon_{ijk} - \sum_j \epsilon_{pjk}]$ also has the multivariate normal density, and the vector $[Y_{1,k}]$ may be subjected to all of the tests given in Chapter 3 for the different assumptions concerning the covariance matrix of the original observations.

It should perhaps be noted at this point that all of the restrictions imposed on the model are not required for the validity of the test procedure. Thus, in equation (4.33), we only need impose the restriction that

$\sum_j (\alpha \beta)_{ij} = 0$. They are left in the model for two reasons:

- (1) If the interaction effects are not all zero, we could not reparameterize the model to obtain the test desired;
- (2) the restriction aids in interpretation of the model, giving operational clarity to the different components.

A test of significance of the effects due to variations in the B-factor may be made in an analogous manner, first summing over the A-levels, performing Hsu's transformation

on the resulting sums, and applying the tests appropriate to the assumptions made concerning the error.

To test for the significance of the interaction terms, we define

$$x_{i..k} = \frac{1}{q} \sum_j x_{ijk} = \mu + \alpha_i + \gamma_k + \varepsilon_{i..k}, \quad (4.35)$$

$$x_{.jk} = \frac{1}{p} \sum_i x_{ijk} = \mu + \beta_j + \gamma_k + \varepsilon_{.jk}, \quad (4.36)$$

$$x_{...k} = \frac{1}{pq} \sum_i \sum_j x_{ijk} = \mu + \gamma_k + \varepsilon_{...k}$$

If we define

$$z_{ijk} = x_{ijk} - x_{.jk} - x_{i..k} + x_{...k}, \quad (4.37)$$

then it is seen that

$$z_{ijk} = (\alpha\beta)_{ij} + \varepsilon_{ijk} - \varepsilon_{i..k} - \varepsilon_{.jk} + \varepsilon_{...k} \quad (4.38)$$

and, on applying Hsu's transformation to the z_{ijk} , we obtain $(pq-1)$ values y_{ijk} , where

$$\begin{aligned} y_{ijk} = z_{ijk} - z_{pqk} &= (\alpha\beta)_{ij} - (\alpha\beta)_{pq} + \varepsilon_{ijk} - \varepsilon_{i..k} \\ &\quad - \varepsilon_{.jk} - \varepsilon_{pqk} + \varepsilon_{p..k} + \varepsilon_{.qk}. \end{aligned} \quad (4.39)$$

The tests given in Chapter 3 may then be applied to these transformed values.

It is seen that, in the most general case concerning the assumptions made about the error variances and

covariances, there must be more observation vectors than treatment combinations. This puts severe limitations on the utilization of these results, as (1) for a test of the A-effects, there must be at least p blocks; (2) for a test of the B-effects, there must be at least q blocks; and (3) for a test of the interactions, there must be at least pq blocks. These restrictions would seem to seriously limit the applicability of these results to practical problems. However, if the assumptions concerning the covariance matrix are encountered in practice, it would seem that the methods given herein are the only valid ones available.

A second objection to the tests in this section is that the tests are correlated, the degree of the correlation depending upon the particular covariance matrix met in practice.

aren't they always correlated?

The applications of the methods presented herein lead to a third difficulty, for a new covariance matrix would have to be computed for each of the tests of main effects and interactions. This can be done by either recomputing the covariance matrix from the transformed data or transforming the original covariance matrix.

Extensions of these methods to $p \times q \times r \times s \times \dots$ factorials is obvious, at least for the main effects. As given herein, the transformation required for a test of a main effect is obtained by

- (1) Summing over all factors except the one being tested, and
- (2) Applying Hsu's transformation to the resulting means from such sums.

Thus, in a $p \times q \times r \times s$ factorial arrangement, with the factors A, B, C and D, if x_{ijklmn} represents the observation on the i^{th} level of A, the j^{th} level of B, the k^{th} level of C, the m^{th} level of D in the n^{th} block, we define

$$x_{i\dots n} = \frac{1}{qrs} \sum_j \sum_k \sum_m x_{ijklmn} \quad (4.40)$$

The dots indicate the factors over which the observations are averaged. A test of the levels of the factor A is made by applying Hsu's transformation to the set $[x_{i\dots n}]$; a test of the levels of the factor B is made by applying Hsu's transformation to the set $[x_{.j\dots n}]$, and so on.

To test for two factor interactions, we form linear functions of the average values, just as in the $p \times q$ case. To test the interaction between the A and B factors, form

$$y_{ij\dots n} = x_{ij\dots n} - x_{i\dots n} - x_{.j\dots n} + x_{\dots\dots n} \quad (4.41)$$

and apply Hsu's transformation to the pq resulting values. In general, to test the two-factor interaction between the factor indicated by the subscript u , say U, and the factor indicated by the subscript v , say V, we subtract the mean over all factors except U and the mean over all factors except V from the mean over all factors except U and V,

giving a set, say

$$y_{..u..v..} = x_{..u..v..} - x_{..u.....} - x_{.....v..} + x_{.....} \quad (4.42)$$

Then, Hsu's transformation is applied to the set $[y_{..u..v..}]$, and the resulting values are tested by the appropriate statistic as given in Chapter 3.

Higher order interactions are tested in an analogous manner, the rules of combination being similar. If there are k factors in all, and a test of a particular h -factor interaction is desired, first select the mean obtained by averaging over all of the $(k-h)$ other factors. From this mean, subtract the sum of the $(h-1)$ means obtained by averaging over all possible combinations of $(h-1)$ of the h chosen factors. Add to this the sum of the $(h-2)$ means obtained by averaging over all possible combinations of $(h-2)$ of the h chosen factors. The process is repeated until the grand average is either added to or subtracted from the sum. Hsu's transformation is applied to the resulting set and the test of significance made. This process is valid if all interaction terms are included and if the usual restrictions are imposed on the model.

Confounded and Fractional Factorial Arrangements.

A factorial arrangement is said to be confounded if one or more contrasts among the treatment effects are confounded.

Let $\tau_1, \tau_2, \dots, \tau_p$ be the treatment effects in an experiment. A contrast is then defined by $\sum_{i=1}^p \alpha_i \tau_i$, with $\sum_{i=1}^p \alpha_i = 0$. Let \sum^k indicate summation over the treatments which appear together in the k^{th} block of each repetition. Then, a contrast is said to be confounded if and only if $\sum^k \alpha_i \neq 0$, for any k . It is this property of confounded arrangements which allows single degree of freedom contrasts to be used. In the unconfounded contrasts, each block effect disappears since $\sum^k \alpha_i = 0$, the block effect being constant for all α_i in this sum. In confounded contrasts, the corresponding statistic measures both differences in block effects and the confounded comparison. One estimate of each appropriate single degree of freedom contrast may be calculated for each repetition of the basic design, and a Student-t statistic calculated from the resulting estimates.

A balanced confounded arrangement is handled in exactly the same way; effects not confounded are calculated from all blocks in a repetition, effects which are confounded are calculated from all blocks except those which were generated by the confounded contrast. As before, we only require that the covariance matrix associated with a block be constant over all repetitions.

Fractional factorial arrangements can be derived from the principal block of a confounded arrangement; thus, the

contrast corresponding to either alias may be constructed by considering the appropriate linear function for that contrast in the principal block.

4.6 Split Plot Designs and Incomplete Block Designs. The split plot design, in its most elementary form, may be described as follows. One set of treatments is arranged in a randomized block or Latin square design. Each plot of this basic design is then subdivided into a number of sub-plots and a second treatment is applied at random to the sub-plots. A careful examination of the literature reveals a multiplicity of models adopted by the various authors to describe this design. A composite of the models presented might be

$$x_{ijk} = \mu + \alpha_i + \tau_j + \nu_k + (\tau\nu)_{jk} + \eta_{ij} + \epsilon_{ijk} \quad (4.43)$$

$$i = 1, \dots, r$$

$$j = 1, \dots, t$$

$$k = 1, \dots, s$$

where

x_{ijk} is the observed yield of the k^{th} sub-treatment and the j^{th} main plot treatment in the i^{th} block;

μ is the grand average;

α_i is the added effect of the i^{th} block, with $\sum \alpha_i = 0$;

β_j is the added effect of the j^{th} whole plot treatment, with $\sum \beta_j = 0$;

η_{ij} is the interaction between the whole plot treatments

and the blocks, usually assumed to be a random variable with variance σ_w^2 .

V_k is the added effect of the k^{th} sub-plot treatment,

$$\text{with } \sum_k V_k = 0;$$

$(\tau V)_{jk}$ is the whole-plot - sub-plot interaction term,

and ε_{ijk} is the error term.

For homogeneous independent errors, if the whole-plots were not divided into sub-plots the η_{ij} terms would be the true experimental error terms. Likewise, if the whole-plot treatments did not exist, the whole-plots would become blocks and the proper experimental error term would be ε_{ijk} .

We shall be concerned with two cases:

$$(a) \text{ Variance } (\eta_{ij}) = \sigma_j^2; \quad ; \text{ Cov}(\eta_{ij} \eta_{im}) = \sigma_{jm} \\ \text{Cov}(\eta_{ij} \eta_{pm}) = 0 \text{ for } i \neq p$$

$$(b) \text{ Variance } (\varepsilon_{ijk}) = \sigma_k^2; \quad \text{Cov}(\varepsilon_{ijk} \varepsilon_{ijm}) = \sigma_{km}; \\ \text{Cov}(\varepsilon_{ijk} \varepsilon_{rsk}) = 0 \text{ for } i \neq r, j \neq s.$$

In either of these two cases, the analysis is straightforward and may be carried through as if the split-plot design were a $p \times q$ factorial design (Section 4.5). This is not true under the usual assumptions of regular analysis of variance, due to the covariance terms arising from the η_{ij} . However, in our case we may handle these covariances as shown below.

To analyze the whole-plot treatments, sum over the

sub-plots and proceed as if the design were a randomized block.

Thus,

$$x_{ij.} = \sum_{k=1}^s x_{ijk} = s\mu + s\alpha_i + s\tau_j + s\eta_{ij} + \sum_k \epsilon_{ijk} \quad (4.44)$$

On performing Hsu's transformation on these totals, we obtain

$$\begin{aligned} y_{ij.} &= x_{ij.} - x_{it.} \\ &= s(\tau_j - \tau_t) + s(\eta_{ij} - \eta_{it}) + \sum_k \epsilon_{ijk} - \sum_k \epsilon_{itk}. \end{aligned} \quad (4.45)$$

The last three terms are error terms associated with the treatment differences. This is the same form as (4.5) under either assumption (a) or (b) and we may analyze the set $y_{ij.}$ as before.

To test for homogeneity of sub-plot treatment means, form the sums

$$x_{i.k} = \sum_{j=1}^t x_{ijk} = t\mu + t\alpha_i + t\tau_k + \sum_j \eta_{ij} + \sum_j \epsilon_{ijk}. \quad (4.46)$$

If we now form

$$\begin{aligned} y_{i.k} &= x_{i.k} - x_{i.s} \\ &= t(\tau_k - \tau_s) + \sum_j \epsilon_{ijk} - \sum_j \epsilon_{ijs} \end{aligned} \quad (4.47)$$

we may also analyze the data as in (4.5).

Note that if assumption (a) holds, but the ϵ_{ijk} are homogeneous independent errors, then the analysis of (4.47) would reduce to the usual analysis of variance case. This, however, would be a rather unusual case. More likely, we might find assumption (b) holding, and the η_{ij} as homogeneous independent errors. Then, the test of whole plot treatments would proceed as in the regular analysis of variance and the test of sub-plot treatments would require the methods presented herein.

An example of a split-plot design is given in section 6.4.

Incomplete Block Designs. The subject of incomplete block designs comprises the largest class of designs; indeed, the complete block designs already discussed form a subset of the incomplete block designs. However, satisfactory methods for analyzing incomplete block designs have not been found. This is due to the incompleteness of these designs which prevents the forming of functions of the observations having the necessary expected values when the hypothesis of equal treatment effects is true and which are free of block effects.

It does not seem possible to extend the theory to this class.

V. APPLICATIONS AND COMPUTATIONAL FORMULAE

5.1 The application of formula (3.68) requires that the transformed product matrix be inverted and then the quadratic form

$$N\bar{Y} (\dot{V}^{-1}) \bar{Y}',$$

where \bar{Y} is the vector of transformed means, be evaluated. This may be done simply by the following computational method.

Hsu [9] has shown that (3.68) may be written as the determinantal ratio of the untransformed variates as

$$T^2 = N\bar{Y} \dot{V}^{-1} \bar{Y}' = \frac{-N \begin{vmatrix} V & J' & \bar{X}' \\ J & 0 & 0 \\ \bar{X} & 0 & 0 \end{vmatrix}}{\begin{vmatrix} V & J' \\ J & 0 \end{vmatrix}} \quad (5.1)$$

where V is the untransformed product matrix, J is the $(1 \times p)$ unit vector, and \bar{X} is the $1 \times p$ vector of untransformed means. It is noted that the elements of the determinant in the denominator are the same elements as in the upper left corner of the determinant in the numerator. We use this fact to allow us to compute the two determinants simultaneously.

The forward solution of the abbreviated Doolittle method for inverting matrices [18] consists essentially of

a factoring of the matrix to be inverted into the product of two auxiliary matrices. Thus, if A is the matrix to be inverted, the forward solution yields

$$A = B \cdot C \quad (5.2)$$

where B is a triangular matrix with zeroes above the main diagonal and C is a triangular matrix with zeroes below the main diagonal and unit elements along the main diagonal.

Having factored A in this fashion, we may find the determinant of A as

$$|A| = |B| \cdot |C| \quad (5.3)$$

and $|C| = 1$, due to the triangularity and unit elements along the diagonal. Thus

$$|A| = |B| \quad (5.4)$$

But B is also triangular, so the determinant may be written

$$|A| = \prod_i b_{ii} \quad (5.5)$$

where b_{ii} is the element in the i^{th} row, i^{th} column.

Thus, if we perform the forward solution on the numerator of (5.1), given that V is a $p \times p$ matrix, we obtain the determinant

$$\prod_{i=1}^{p+2} b_{ii} ; \quad (5.6)$$

likewise the determinant of the denominator may be written as

$$\prod_{i=1}^{p+1} b_{ii} \quad (5.7)$$

and, from the nature of the Doolittle method, it is seen that

$$\dot{b}_{ii} = b_{ii} \quad , \quad i = 1, \dots, p+1 \quad . \quad (5.8)$$

Thus, we may write (5.1) as

$$T^2 = -N \frac{\prod_{i=1}^{p+2} b_{ii}}{\prod_{i=1}^{p+1} b_{ii}} = -N b_{p+2, p+2} \quad (5.9)$$

As an illustration of this, consider the case for $p = 2$:

$$\bar{X} = [8, 6] \quad ; \quad V = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} \quad ; \quad N = 10 \quad .$$

For use in the regular formula (3.68), it is seen that the transformed values (3.11), (3.12) would be

$$\bar{Y} = [2] \quad , \quad \dot{V} = [4-2-2+3] = [3] \quad , \quad \text{and thus by} \quad (3.68)$$

$$T^2 = 10 [2] \left[\frac{1}{3}\right] [2] = \frac{40}{3}$$

By the method given above in (6.9), we would carry out the calculations as shown below using the forward Doolittle

method on

$$\begin{bmatrix} 4 & 2 & 1 & 8 \\ 2 & 3 & 1 & 6 \\ 1 & 1 & 0 & 0 \\ 8 & 6 & 0 & 0 \end{bmatrix} .$$

We thus obtain the calculation

$$\begin{array}{cccc} 4 & 2 & 1 & 8 \\ 2 & 3 & 1 & 6 \\ 1 & 1 & 0 & 0 \\ 8 & 6 & 0 & 0 \\ \hline 4 & 2 & 1 & 8 \\ 1 & \frac{1}{2} & \frac{1}{4} & 2 \\ \hline & 2 & \frac{1}{2} & 2 \\ & 1 & \frac{1}{4} & 1 \\ & & -\frac{3}{8} & -\frac{5}{2} \\ & & 1 & \frac{12}{3} \\ & & & -\frac{4}{3} \end{array} \quad (5.10)$$

Using the last value, which is b_{44} and substituting in (5.9), we obtain

$$T^2 = - (10) \left(-\frac{4}{3} \right) = \frac{40}{3}$$

which agrees with the value obtained above.

Experience with this method tends to indicate that a saving in time occurs if there are four or more original treatments.

5.2 A Randomized Block Example. Suppose a randomized block type experiment has been run on the effect of three treatments on the rate of some chemical reaction, and it is desired to test for equality of treatment effects. The data for 10 blocks, suitably scaled, is given in Table 5.1 .

TABLE 5.1 Results of a Reaction Rate Experiment

Block No.	Treat. No. 1	Treat. No. 2	Treat. No. 3
1	10.4	16.1	22.2
2	8.7	18.5	21.8
3	9.9	18.2	20.6
4	9.0	16.7	21.3
5	10.4	20.2	21.4
6	10.2	17.8	21.5
7	11.0	17.6	21.5
8	10.5	18.2	21.1
9	9.2	18.0	22.2
10	11.0	17.7	22.4

We will analyze the data by two methods. The first method will be to transform the basic data as in (4.4) and calculate the test statistic for equality of means as

given in (3.68); the second method will utilize the routine given in Section 5.1.

We first form a new set of data by subtracting the data for one treatment from the rest of the data, row by row. The first treatment results will be subtracted from the other two to obtain non-negative results, as given in Table 5.2.

TABLE 5.2 Transformed Data

Block No.	(Treat.No.2 - Treat.No.1)	(Treat.No.3 - Treat.No.1)
1	5.7	11.8
2	9.8	13.1
3	8.3	10.7
4	7.7	12.3
5	9.8	11.0
6	7.6	11.3
7	6.6	10.5
8	7.7	10.6
9	8.8	13.0
10	6.7	11.4

If all three columns had been substantially equal, the above numbers would all be around zero. If the last two

columns were equal, with smaller numbers in the first column, the above two columns would be approximately equal.

To perform the test, we must calculate the sums, sum of squares, and sum of cross products of the above two columns. These we find to be:

$$\text{Totals: } 5.7 + 9.8 + \dots + 6.7 = 78.7$$

$$11.8 + 13.1 + \dots + 11.4 = 115.7$$

$$\text{Sums of Squares: } (5.7)^2 + (9.8)^2 + \dots + (6.7)^2 = 635.69$$

$$(11.8)^2 + (13.1)^2 + \dots + (11.4)^2 = 1346.89$$

Sum of Cross Products:

$$(5.7)(11.8) + (9.8)(13.1) + \dots + (6.7)(11.4) = 914.54$$

We next calculate the so-called corrected sums of squares and cross products as follows:

$$\text{Corrected Sum of Squares: } 635.69 - \frac{(78.7)^2}{10} = 16.321$$

$$1346.89 - \frac{(115.7)^2}{10} = 8.241$$

$$\text{Corrected Sum of Products: } 914.54 - \frac{(78.7)(115.7)}{10} = 3.980$$

The number ten in the denominator arose because there were ten sets of observations. The numbers subtracted were the squares or product of the totals. From the totals, we calculate the averages:

$$\frac{78.7}{10} = 7.87 \quad , \quad \frac{115.7}{10} = 11.57.$$

The matrix of products and cross-products is then inverted to obtain

$$\begin{pmatrix} 16.321 & 3.980 \\ 3.980 & 8.241 \end{pmatrix}^{-1} = \frac{1}{118.6610} \begin{pmatrix} 8.241 & -3.980 \\ -3.980 & 16.321 \end{pmatrix} = \hat{v}^{-1}$$

The quadratic form is then evaluated as

$$\begin{aligned} T^2 &= N\bar{Y}\hat{v}^{-1}\bar{Y}' = \frac{10}{118.661} (7.87 \quad 11.57) \begin{pmatrix} 8.241 & -3.980 \\ -3.980 & 16.321 \end{pmatrix} \begin{pmatrix} 7.87 \\ 11.57 \end{pmatrix} \\ &= 10 \times 16.605 = 166.05 \end{aligned}$$

We may convert this value to an F-statistic by multiplication by the constant

$$\frac{N - p + 1}{p - 1} = \frac{10 - 3 + 1}{3 - 1} = \frac{8}{2} = 4 .$$

As the tabulated value of the F-distribution for (3 - 1) and (10 - 3 + 1) degrees of freedom at the .05 level is 4.46, we reject the hypothesis of equality of treatment effects, since

$$4 \times 166.05 \gg 4.46 .$$

We now calculate the result using the computational routine. The corrected sum of squares and product matrix and vector of means of the untransformed data is found in an analogous manner to the preceding calculations to be

$$\bar{X} = (10.03 \quad 17.90 \quad 21.60)$$

and

$$V = \begin{bmatrix} 5.941 & 0.140 & 0.250 \\ & 10.660 & -1.570 \\ & & 2.800 \end{bmatrix}$$

The heterogeneity of error terms can be seen in this and in the transformed matrix \dot{V} . In accordance with the method given in Section 5.1, we have the calculations shown in Table 5.3.

TABLE 5.3 Computational Routine

5.941	0.140	0.250	1.000	10.030
	10.660	-1.570	1.000	17.900
		2.800	1.000	21.600
			0.000	0.000
				0.000
5.941	0.140	0.250	1.000	10.030
1.00000	0.02356	0.04208	0.16832	1.68827
	10.6567	-1.57589	0.97643	17.6636
	1.0000	-0.14788	0.09163	1.65751
		2.55644	1.10232	23.78999
		1.00000	0.43119	9.30590
			-0.73310	-13.5648
			1.00000	18.50326
				-16.60525

Taking the negative of the last number and multiplying by 10 [see (5.9)], we obtain, as before

$$T^2 = 166.0525$$

5.3 A Latin Square Example. Suppose that in the last example three reactors were used and it was felt that differences existed between the reactors which would have to be eliminated. This could be done by performing a Latin square type experiment, as shown below. Let 1, 2, 3, denote the different reactors; A, B, C, denote the three trials, and I, II, III, denote the three treatments. The design would then be

TABLE 5.4 Latin Square Design

		Reactors		
		1	2	3
Trial	A	I	II	III
	B	II	III	I
	C	III	I	II

This design allows the effects of different reactors and the effects due to different trials to be removed from the data. Suppose the test were run ten times, again

measuring the reaction rate. The data, suitably coded, might appear as in Table 5.5.

TABLE 5.5 Data for Latin Square Example

Test No.	A.1.I	A.2.II	A.3.III	B.1.II	B.2.III	B.3.I	C.1.III	C.2.I	C.3.II
1	10.7	9.0	11.7	10.0	12.6	15.3	17.8	20.9	17.6
2	10.3	9.1	12.5	10.0	15.6	17.5	18.3	22.5	20.0
3	11.3	7.1	13.0	7.9	14.2	18.2	15.9	19.4	19.3
4	10.3	9.1	13.2	9.5	15.0	16.2	17.0	20.2	19.1
5	11.0	10.5	11.2	10.6	13.0	18.0	17.2	20.2	20.0
6	11.4	9.2	11.5	10.7	13.6	16.4	16.2	20.0	19.1
7	11.2	9.9	13.7	11.1	13.9	18.2	18.8	20.5	18.9
8	14.2	8.9	12.7	9.0	15.3	18.1	16.8	20.5	17.9
9	10.6	10.8	14.0	10.9	14.5	17.6	16.6	19.6	19.4
10	12.0	8.7	14.3	11.1	15.2	17.9	16.8	20.9	19.8

We wish to test whether, over all reactors and trials, there is a difference in reaction rates. We first sum together the results for each of the three treatments as in (4.27). Thus, for the first test, the total for treatment I is found to be

$$10.7 + 15.3 + 20.9 = 46.9,$$

while the total for treatment II is

$$9.0 + 10.0 + 17.6 = 36.6 .$$

Proceeding in this fashion, we find the totals as in Table 5.6.

TABLE 5.6 Summed Data of Latin Square Example

Test No.	Treat. 1	Treat. 2	Treat. 3
1	46.9	36.6	42.0
2	49.9	39.1	46.4
3	48.9	34.2	44.0
4	46.7	37.7	45.0
5	49.2	41.1	46.2
6	47.8	39.0	41.3
7	49.9	39.9	46.4
8	52.8	35.9	44.9
9	47.8	41.0	45.0
10	50.8	39.5	46.3

To begin the analysis, first subtract any one column from the other two, as in (4.28); we subtract the data for treatment II from the other two to produce positive numbers. Proceeding in this fashion, we obtain the results in Table 5.7.

TABLE 5.7 Transformed Data

Test No.	(Treat. 1 - Treat. 2)	(Treat. 3 - Treat. 2)
1	10.3	5.4
2	10.8	7.3
3	14.7	9.8
4	9.0	7.4
5	8.1	5.1
6	8.8	2.3
7	9.9	8.7
8	16.9	8.9
9	6.7	4.0
10	11.3	6.8

We now subject these numbers to the same analysis as in the last section. We find the row sum, sums of squares and products to be:

$$\begin{aligned} \text{Sums:} \quad & 10.3 + \dots + 11.3 = 106.7 \\ & 5.4 + \dots + 6.8 = 65.9 \end{aligned}$$

$$\begin{aligned} \text{Sums of Squares:} \quad & (10.3)^2 + \dots + (11.3)^2 = 1242.5 \\ & (5.4)^2 + \dots + (6.8)^2 = 458.9 \end{aligned}$$

$$\text{Sums of Products:} \quad (10.3)(5.4) + \dots + (11.3)(6.8) = 750.4$$

The corrected sums are found to be

$$1242.5 - \frac{(106.7)^2}{10} = 104.01$$

$$458.9 - \frac{(65.9)^2}{10} = 24.62$$

$$750.4 - \frac{(106.7)(65.9)}{10} = 47.25$$

On evaluating the quadratic form (3.68), we obtain

$$\begin{aligned} T^2 = \bar{N}\bar{Y}\bar{V}^{-1}\bar{Y}' &= \frac{10}{328.164} (10.67 \quad 6.59) \begin{pmatrix} 24.62 & -47.25 \\ -47.25 & 104.01 \end{pmatrix} \begin{pmatrix} 10.67 \\ 6.59 \end{pmatrix} \\ &= 20.5726 \end{aligned}$$

and

$$F = \frac{N-p+1}{p-1} T^2 = \frac{8}{2} T^2 = 82.2904$$

which is highly significant.

To test whether there is a difference between the reactors, we would have summed over trials and treatments and performed the analysis on the resulting totals.

5.4 A Split Plot Example. Herein, we shall consider as a split plot example a hypothetical test of three special crankcase oils used in a standard motor generator set. Suppose the test is made in a cold chamber at six different temperatures. The oil was of a special nature, and only enough oil could be made at one time to test at the six

temperatures. The variable used as a test criterion was the power loss:

(indicated horsepower - brake horsepower).

Suitably coded, the data might appear as in Table 5.8 .

Each test was repeated fifteen times.

TABLE 5.8 Data for Split Plot Example*

Rep.	Formula I						Totals
	Temperature (°F)						
	30°	25°	20°	15°	10°	5°	
1	31	35	36	28	42	31	203
2	36	18	24	36	46	34	194
3	21	21	26	32	34	34	168
4	15	21	24	13	28	15	116
5	17	19	20	26	30	36	148
6	13	11	11	18	24	15	92
7	15	12	14	16	22	24	103
8	13	22	12	21	20	23	111
9	13	16	17	22	23	12	103
10	13	22	16	20	19	22	112
11	22	28	22	17	22	19	130
12	17	20	16	28	24	32	137
13	18	17	20	18	26	22	121
14	13	29	18	29	29	20	138
15	15	17	21	14	26	22	115
	272	308	297	338	415	361	1991

*Raw data coded from Cochran, W. G., and Cox, G. M. 23

TABLE 5.8 (Continued)

Formula II							
Rep.	Temperature (°F)						Totals
	30°	25°	20°	15°	10°	5°	
1	28	35	40	38	44	31	216
2	24	35	36	28	41	50	214
3	23	19	31	24	31	24	152
4	14	15	17	35	26	26	133
5	20	19	18	24	29	25	135
6	13	18	18	18	13	24	104
7	11	14	15	15	18	25	98
8	15	12	13	20	16	26	102
9	16	15	21	17	21	22	112
10	10	13	13	16	26	19	97
11	9	16	22	20	17	22	106
12	12	17	20	23	20	18	110
13	21	24	19	16	24	19	123
14	12	14	11	8	10	24	79
15	10	10	17	15	16	9	77
	<u>238</u>	<u>276</u>	<u>311</u>	<u>317</u>	<u>352</u>	<u>364</u>	<u>1858</u>
Formula III							
1	35	33	34	35	37	52	226
2	32	32	32	35	36	47	214
3	22	13	29	26	30	27	147
4	27	30	27	19	25	24	152
5	10	14	20	24	22	12	102
6	13	22	19	19	26	24	123
7	9	10	20	13	19	22	93
8	13	12	10	13	10	24	82
9	13	7	10	15	17	17	79
10	15	17	16	16	24	24	112
11	17	14	15	14	27	17	104
12	13	19	17	24	22	17	112
13	17	18	32	17	22	26	132
14	8	11	16	14	14	24	87
15	10	17	14	14	20	14	89
	<u>254</u>	<u>269</u>	<u>311</u>	<u>298</u>	<u>351</u>	<u>371</u>	<u>1854</u>

Suppose it is first desired to test whether there is a difference in the horsepower losses between the three formulas. We first form the totals for each replicate and formula, summing over the various temperatures. These totals are given in the last column of Table 5.8. In this way, we effectively average over the results for different temperatures.

We now perform Hsu's transformation on these sums, subtracting the sums of formula III from the sums of formulas I and II. The results are given in Table 5.9.

TABLE 5.9 Transformed Data

Replicate	I - III	II - III
1	-23	-10
2	-20	0
3	21	5
4	-36	-19
5	46	33
6	-31	-19
7	10	5
8	29	20
9	24	33
10	0	-15
11	26	2
12	25	-2
13	-11	-9
14	51	8
15	26	12

For example, the first value obtained (-23) is found from

$$203 - 226 = -23 .$$

As before, we form the sums, sums of squares and sums of products as:

$$\text{Sums: } (-23)+(-20)+ \dots +(26) = 137$$

$$(-10) + (0) + \dots + (-12) = 4$$

$$\text{Sums of squares: } (-23)^2 + \dots + (26)^2 = 11,959$$

$$(-10)^2 + \dots + (-12)^2 = 3882$$

$$\text{Sums of products: } (-23)(-10)+\dots+(26)(-12) = 3929$$

We now form the corrected sum of squares and cross products:

$$11,959 - \frac{(137)^2}{15} = 10,708$$

$$3882 - \frac{(4)^2}{15} = 3881$$

$$3929 - \frac{(137)(4)}{15} = 3893$$

The averages, taken from the above totals, give a vector of means:

$$\left[9.133, 0.267 \right]$$

The matrix of corrected sums of squares and products is formed as

$$V = \begin{bmatrix} 10,708 & 3893 \\ 3893 & 3881 \end{bmatrix}$$

and its inverse is found to be

$$V^{-1} = \begin{bmatrix} .0001238 & -.0001255 \\ -.0001255 & .0003729 \end{bmatrix}$$

On forming the quadratic form, we find

$$(9.133 \quad 0.267) \begin{bmatrix} .0001238 & -.0001255 \\ -.0001255 & .0003729 \end{bmatrix} \begin{pmatrix} 9.133 \\ 0.267 \end{pmatrix} = .00976$$

Then,

$$T^2 = 15 \times .00976 = 0.1464$$

The value of T^2 must be multiplied by the constant

$$\frac{(\text{No. of reps.} - \text{No. of formulas} + 1)}{(\text{No. of formulas} - 1)} = \frac{15 - 3 + 1}{3 - 1} = 6\frac{1}{2} .$$

Then $6\frac{1}{2} \times 0.1464 = 0.9516$ has the F-distribution with $3-1=2$ and $15-3+1=13$ degrees of freedom, which is clearly not significant. Thus, it is concluded that there is no difference between the loss in horsepower due to the three formulas.

To test whether there is any difference in horsepower lost over the different temperatures, we form the sum over the three formulas obtaining the results in Table 5.10.

TABLE 5.10 Totals Over Formulas, Split Plot Example

Rep.	30°	25°	20°	15°	10°	5°
1	94	103	110	101	123	114
2	92	85	92	99	123	131
3	66	53	86	82	95	85
4	56	66	68	67	79	65
5	47	52	58	74	81	73
6	39	51	48	55	63	63
7	35	36	49	44	59	71
8	41	46	35	54	46	73
9	42	38	48	54	61	51
10	38	52	45	52	69	65
11	48	58	59	51	66	58
12	42	56	53	75	76	67
13	56	59	71	51	72	67
14	33	54	45	51	53	68
15	35	44	52	43	62	45

The first value is found from

$$31 + 28 + 35 = 94.$$

If we subtract the first column from all of the others, row by row, and then form the matrix of corrected sums of squares and products, we obtain the matrix

$$\dot{V} = \begin{bmatrix} 1104.93 & 208.33 & 281.33 & -111.40 & 44.13 \\ & 617.33 & -49.67 & 290.00 & -425.67 \\ & & 1609.33 & 378.00 & 1161.00 \\ & & & 757.60 & 129.80 \\ & & & & 1567.73 \end{bmatrix} .$$

The vector of means is

$$\bar{Y} = [5.933, 10.333, 11.600, 23.600, 22.133] ,$$

from which we find that

$$N\bar{Y}\dot{V}\bar{Y}' = 32.019405$$

and, since $\frac{(\text{No. of replicates} - \text{No. of variables} + 1)}{(\text{No. of variables} - 1)} T^2$

follows the F-distribution with $15 - 13 + 1 = 13$ and $3 - 1 = 2$ degrees of freedom, then

$$F = \frac{13}{2} \times 32.019 = 208.123$$

is highly significant and it is concluded that horsepower loss is different at different temperatures.

PART II

VI. ON BERKSON'S CASE OF LINEAR REGRESSION WITH
ERRORS IN THE INDEPENDENT VARIABLES

6.1 Introduction. Historically, estimators of regression coefficients in linear and multiple regression have been derived under two different assumptions concerning the nature of the independent variable. For a dependent variable y observed at the vector valued independent variable X , the first derivations were made under the assumption that y and X were jointly distributed according to some probability law; any realized set (y, X) constituted a sample drawn from this multivariate population. Later, around 1920, Fisher [4] substituted the assumption that the values of the independent variable were fixed and without error, and showed that the two assumptions led to the same estimating equations.

However, in the application of the theory, at least to the physical sciences where some functional relationship connects the dependent and independent variables, the more usual case arises of errors in both variables. The independent variable cannot be considered to be a random sample from some population as it is usually under the control of the experimenter, and thus has some approximate value. Investigations concerning application of the theory under

these conditions showed that additional information, in the form of the variance of the error in the independent variable or ratio of the variance of the independent to the variance of the dependent variable, is necessary for the solution of the problem [1] , [6] , [7] .

Berkson [2] has shown that there is one case, at least for simple linear regression, where a solution may be effected without this additional information. In so doing, he differentiated between two types of independent variables which may arise. For example, in performing chemical experiments, the experimenter may be faced with two different independent variables: (a) The weight of the reactants and (b) the temperature of reaction. Let us assume that the reactants are to be weighed out prior to the experiment on a laboratory balance by balancing some amount against weights with values x_1 . If we had, in addition, an error free balance, the true weights of the reactants could also be determined as z_1 , say. Then the error would be $d_1 = z_1 - x_1$. However, in measuring out reactants on our laboratory balance again, we would still observe the weight x_1 ; the true value of the weight z_1' would now be different from z_1 , and a different error $d_1' = z_1' - x_1$ would have been committed. Thus, our error is independent of the observed value x_1 . This type of independent variable is termed a Controlled Variable.

The second independent variable--temperature of reaction--is termed an Uncontrolled Variable, as the observed value x_2 randomly varies about some fixed value \bar{z}_2 . Thus our error is $d_2 = \bar{z}_2 - x_2$, and as \bar{z}_2 is a constant, the realized value of the error is perfectly correlated with the observed value x_2 .

Berkson found that, in the simple model

$$y = \alpha + \beta x + \epsilon,$$

if x is a controlled variable, the usual least squares estimators of α and β are still unbiased estimators as long as the errors d and ϵ are unbiased.

Geary [5] extended Berkson's case to non-linear regression and considered some simple sampling variances of the estimates. In the development of the non-linear case, a third degree model was assumed and it was found that unbiased estimates of the intercept and coefficient of the linear term were not obtainable. This is generalized below.

6.2 Consider the estimation of the parameters in the model

$$y_{jk} = \sum_{i=0}^m \beta_i z_j^i + f_{jk} \quad ; \quad j=1, \dots, r; \quad k=1, \dots, N_j, \quad (6.1)$$

where x_j is the j^{th} realized value of the controlled variable

\exists_j , y_{jk} is the k^{th} observation on y at x_j , f_{jk} is the error associated with y_{jk} and e_{jk} is the error associated with the k^{th} trial at x_j . We select the values of x_j so that

$$\sum_{j=1}^r x_j^{2a-1} = 0, \quad a = 1, 2, \dots, \quad (6.2)$$

and assume

$$\sum_k [e_{jk}^{2a-1}] = \epsilon_{2a-1} = 0 \quad \text{for each } j; \quad a = 1, 2, \dots, \quad (6.3)$$

$$\sum_k [e_{jk}^{2a}] = \epsilon_{2a} \quad \text{for all } j; \quad a = 0, 1, 2, \dots \quad (6.4)$$

Also, we use the notation

$$\sum_{j=1}^r x_j^a = \mu_a, \quad a = 0, 1, 2, \dots \quad (6.5)$$

Note that $\mu_0 = r$ and from (6.2) we have $\mu_a = 0$ if a is odd.

Theorem 1: In the estimation of the parameters in the model (6.1) by Geary's method, subject to the conditions (6.2), (6.3), (6.4),

- (a) Two and only two coefficients are estimable, and these are β_{m-1} and β_m ,
- (b) The two unbiased estimates obtained are identical with the estimates which would be obtained from a least squares approach ignoring the errors in the independent variables.

- (c) If the values of ϵ_{2a} are known, all other coefficients are estimable.

This theorem will be proved by examination of the estimates by generalizations of Geary's method.

We may write (1) as

$$y_{jtk} = \sum_{i=0}^m \beta_i (x_j + e_{jtk})^i + f_{jtk}; \quad j=1, \dots, r; \quad k=1, \dots, N_j. \quad (6.6)$$

Expanding, we obtain

$$y_{jtk} = \sum_{i=0}^m \beta_i \sum_{p=0}^i \binom{i}{p} x_j^{i-p} e_{jtk}^p + f_{jtk}. \quad (6.7)$$

Averaging over the subscript k for each x_j and denoting averages by a horizontal bar, we obtain

$$\bar{y}_j = \sum_{i=0}^m \sum_{p=0}^i \beta_i \binom{i}{p} x_j^{i-p} \bar{e}_j^p + \bar{f}_j. \quad (6.8)$$

As we seek only unbiased estimates, we equate the average error terms to their expected values and obtain

$$\bar{y}_j = \sum_{i=0}^m \sum_{p=0}^i \beta_i \binom{i}{p} x_j^{i-p} \epsilon_{2a} \delta_{2a,p} \quad (6.9)$$

where $\delta_{2a,p} = 1$ when $2a = p$
 $= 0$ otherwise

Due to the oddness-evenness conditions on \underline{i} , we find from (6.9) after a few lines that

$$\sum_{p=0}^i \binom{i}{p} x_j^{i-p} \epsilon_{2a} \delta_{2a,p} = \sum_{t=0}^{\left[\frac{i}{2} \right]} \binom{i}{2t} x_j^{i-2t} \epsilon_{2t} \quad (6.10)$$

where $\left[\frac{i}{2} \right]$ denotes the largest integer in the number $\frac{i}{2}$.

(For i odd, the exact upper limit is $\frac{i-1}{2}$ and for i even, the limit is $\frac{i}{2}$.).

Examination of the indices of summation and condition (6.3) indicates that the system will best be examined separately for odd or even values of m . We consider now the case of even m (the case where m is odd follows almost identically with obvious modifications). Any polynomial of degree m may be divided into two parts

$$\sum_{i=0}^m \beta_i \gamma^i = \sum_{s=0}^{\frac{m}{2}} \beta_{2s} \gamma^{2s} + \sum_{s=0}^{\frac{m-1}{2}} \beta_{2s+1} \gamma^{2s+1} \quad (6.11)$$

such that the first part contains even powers of the variate and the second contains odd powers. Thus, we may write (6.9) with m even as

$$\bar{y}_j = \sum_{s=0}^{\frac{m}{2}} \beta_{2s} \sum_{t=0}^s \binom{2s}{2t} x_j^{2(s-t)} \epsilon_{2t} + \sum_{s=0}^{\frac{m-1}{2}} \beta_{2s+1} \sum_{t=0}^s \binom{2s+1}{2t} x_j^{2(s-t)+1} \epsilon_{2t}, \quad (6.12)$$

eliminating all non-zero items. Now consider the product moments; using the notation in (6.5), we have

$$\sum_{j=1}^r x_j^a \bar{y}_j = \sum_{s=0}^{\frac{m}{2}} \sum_{t=0}^s \beta_{2s} \binom{2s}{2t} \epsilon_{2t} \mu_{2(s-t)+a} + \sum_{s=0}^{\frac{m-1}{2}} \sum_{t=0}^s \beta_{2s+1} \binom{2s+1}{2t} \epsilon_{2t} \mu_{2(s-t)+a+1}. \quad (6.13)$$

It is seen that $2(s-t)$ is always even and thus for odd values of a , the first part in (6.13) disappears. Likewise, when a is even, the second part vanishes. Replacing s as an index of summation with $u = s - t$, we rewrite (6.13)

after some reduction as

$$\sum_{j=1}^r x_j^a \bar{y}_j = \sum_{u=0}^{\frac{m}{2}} \mu_{2u+a} \sum_{t=0}^{\frac{m-2u}{2}} \beta_{2u+2t} \binom{2u+2t}{2t} \epsilon_{2t} \quad \text{for even } a, \quad (6.14)$$

$$= \sum_{u=0}^{\frac{m-1}{2}} \mu_{2u+a} \sum_{t=0}^{\frac{m-2u-1}{2}} \beta_{2u+2t+1} \binom{2u+2t+1}{2t} \epsilon_{2t} \quad \text{for odd } a.$$

Let us define

$$\alpha_{2u+k} = \sum_{t=0}^{\frac{m-2u-k}{2}} \beta_{2u+2t+k} \binom{2u+2t+k}{2t} \epsilon_{2t}; \quad k=0,1. \quad (6.15)$$

Let A denote the column vector $[\alpha_1]$, B denote the column vector $[\beta_1]$, X denote the matrix of controlled values of the observed independent variables $[x_{1j}]$, and let Y denote the column vector of the observed values of the dependent variable $[\bar{y}_j]$. Examination of the structure of (6.14) shows that the set of equations generated by putting $a = 0, 1, 2, \dots, m$ may be expressed matrixially as

$$X'Y = X'XA \quad (6.16)$$

If the values of x had been fixed and error free, we would have obtained the set of normal equations

$$X'Y = X'XB \quad (6.17)$$

Thus, we need only examine the elements in A; if any of these are identical with the corresponding elements in B, it follows that those estimators are identical with the least squares estimators under error free conditions.

It is seen from (6.4) that $\epsilon_0 = 1$, thus we must examine (6.15) for the cases where the upper limit of summation is zero. When $k = 0$, $m - 2u = 0$ or $u = \frac{m}{2}$. Hence (6.15)

gives

$$\alpha_m = \begin{pmatrix} m \\ 0 \end{pmatrix} \beta_m \equiv \beta_m. \quad (6.18)$$

When $k = 1$, $m - 2u - 2 = 0$, or $u = (m-2)/2$. Here (6.15)

gives

$$\alpha_{m-1} = \begin{pmatrix} m-1 \\ 0 \end{pmatrix} \beta_{m-1} \equiv \beta_{m-1} \quad (6.19)$$

For all other cases, α is a bilinear form in β and ϵ_{2t} and thus the other coefficients β_i ($i = 0, \dots, m-2$) are not estimable by this method. This completes the proof of Theorem 1.

6.3 Extensions of the theorem to multivariable situations are immediate, but tedious. For example, in the model

$$y_{jkl} = \beta_{00} + \beta_{10} z_{je} + \beta_{20} z_{je}^2 + \beta_{01} z_{je} v_{kl} + \beta_{01} v_{kl} + \beta_{02} v_{kl}^2 + f_{jkl}, \quad (6.20)$$

if the realized values of the independent variables are

x, v , and are controlled so that

$$\begin{aligned} \text{a) } z_{je} &= x_j + d_{je}, \quad j=1, \dots, m; \quad l=1, \dots, N_{jk}, \\ v_{kl} &= v_k + e_{kl}, \quad k=1, \dots, p; \quad l=1, \dots, N_{jk}, \end{aligned} \quad (6.21)$$

$$\text{b) } \sum_l (f_{jkl}) = \sum_l (d_{je}) = \sum_l (e_{kl}) = 0, \quad (6.22)$$

$$\sum_l (d_{je} e_{kl}) = 0.$$

and

c) the values of x, v are chosen so that

$$\sum_{j=1}^m \bar{x}_j = \sum_{k=1}^m \bar{v}_k = 0 \quad (6.23)$$

then all parameters except β_{00} may be estimated, and the estimates are identical with those derived by least square methods. In general, with an orthogonal array of values of the independent variables and under conditions analogous to (6.21) and (6.22), the coefficients which may be estimated are determined as follows: Let the order of a coefficient be defined as the sum of the powers of the independent variables associated with that coefficient. Then all of the coefficients having the highest order and all of the coefficients having one less than the highest order may be estimated, and the estimates are identical with the least squares estimates.

These extensions are fairly important in those problems concerned with investigations into the shape of response surfaces. For second order models (which are usually assumed) and for the second order designs of Box and Wilson [3], all necessary parameters for location of the maxima, and studies of the shape of the surface may be estimated by the theory of least squares, as long as the independent variables are of the class defined as controlled variables.

PART III

VII SOME CONSEQUENCES OF ANALYZING NON-ADDITIVE DATA
ARISING FROM AN UNREPLICATED FACTORIAL
EXPERIMENT

7.1 The two basic complete block experimental designs are the randomized block and the Latin square designs. In both of these designs, the validity of the error terms is based on proper randomization within blocks or randomization of the Latin squares. It is generally felt that only through randomization may we estimate valid measures of experimental error. In some cases, however, experimenters are faced with the necessity of performing only one block, especially when the treatments constitute a set of factorial treatments. Thus, for example, testing three factors at three levels requires a block of twenty-seven treatments. Limitations of time or money may restrict the experimenter to only one complete set, in effect producing an experiment with only one block.

It has been suggested [1], [2], [3] that, in these cases, an estimate of error might be obtained from the higher order interactions. If the interactions are really non-existent, this is a valid procedure, as the contrasts in this case measure the random error term. However, if the higher order interactions are not zero, the estimate of error will be inflated and the values of

the F-statistic of real main effects will be deflated, giving too few significant results. This assumes that all of the usual assumptions of additivity, normality, and independence of error terms are met.

The effects of this procedure when the data arises from a non-additive model will be examined in this part.

One of the difficulties encountered in studies of the effects of non-additivity is the choice of a non-additive form to use in the alternative case. Of the possible choices, two non-additive forms are of special importance in applications of statistics to the physical sciences, and these are

$$y = \beta_0 \prod_{i=1}^r \beta_i^{x_i} \quad , \quad (7.1)$$

and

$$y = \beta_0 \prod_{i=1}^r x_i^{\beta_i} \quad , \quad (7.2)$$

where y is the observed dependent variable, x_i is the i^{th} observed independent variable, and β_i is the coefficient associated with the variable x_i .

In both of these cases, we may transform to a linear form by considering the logarithmic transformation.

Thus for (7.1) we obtain

$$\log y = \log \beta_0 + \sum_{i=1}^r x_i \log \beta_i \quad , \quad (7.3)$$

and for (7.2),

$$\log y = \log \beta_0 + \sum_{i=1}^r \beta_i \log x_i \quad (7.4)$$

It is seen that the assumption of additive normal independent error terms made in the analysis of variance implies that the error terms associated with (7.3) and (7.4) are of the logarithmic-normal form. Further, after applying such a transformation to (7.1) the analysis is straightforward, and we see that the use of the interaction terms as an estimator of error is a valid procedure, as interactions are non-existent after transformation. However, the analysis of data arising from a transformation of (7.2) into (7.4) is not so straightforward, but requires that consideration be given to the spacings of the independent variable. For example, suppose three equally spaced levels of one factor had been run in the experiment. If the levels were $[(p-1), p, (p+1)]$, we would have to analyze the set $[\text{Log}(p-1), \text{Log } p, \text{Log}(p+1)]$, introducing a non-orthogonal case. Of course, proper design of such an experiment would have resulted in equal spacings of the logarithms. These considerations apply if the experimenter knows that the data arises from such a model.

We now inquire into the effects of analyzing data which really arose from a model of the multiplicative type (7.1), but which has not been transformed. If several randomized blocks, or if a Latin square had been run, the use of Tukey's single degree of freedom test would

have given an indication that the true model was non-linear; if only one block had been run, we would have had no such indication. We consider herein the effects of the use of higher order interactions as an estimator of error for one special case--a 3^2 factorial arrangement with error free observations. We assume the model

$$y_1 = a b^u c^v \quad ; b, c \neq 1 \quad (7.5)$$

in the work which follows.

7.2 As levels of the two factors, we choose $(-1, 0, 1)$. That this may always be done may be seen by considering the actual levels $(p-1, p, p+1)$. Then the values of the independent variable at these three levels would be

$$\{ ab^{p-1}, ab^p, ab^{p+1} \} \quad (7.6)$$

and if we let ab^p be a constant term, then we may write the three levels as

$$\{ (ab^p) b^{-1}, (ab^p) b^0, (ab^p) b^{+1} \} . \quad (7.7)$$

Denoting the constant term so obtained by a , we obtain the values of the dependent variable for the 3^2 factorial as shown in Table 7.1.

TABLE 7.1 Values of Untransformed Data

Level of Factor B	Level of Factor C			Total
	-1	0	1	
-1	abc^{-1-1}	abc^{-1^0}	abc^{-1}	$ab(c^{-1-1}+c)$
0	abc^{0-1}	abc^{0^0}	abc^0	$ab(c^{0-1}+c)$
1	abc^{1-1}	abc^{1^0}	abc^1	$ab(c^{1-1}+c)$
Totals	$ac(b^{-1-1}+b)$	$ac(b^{-1}+b)$	$ac(b^{-1}+b)$	$a(b^{-1}+b)(c^{-1}+c)$

We may compute eight contrasts from the nine observations above. Except for the constant divisor associated with these contrasts, the values of the contrasts are given in Table 7.2.

TABLE 7.2 Linear Contrasts

$$\begin{aligned}
 B &= \frac{a}{bc} (b^2 - 1)(c + c + 1) \\
 B^2 &= \frac{a}{bc} (b^2 - 2b + 1)(c^2 + c + 1) \\
 C &= \frac{a}{bc} (b + b + 1)(c^2 - 1) \\
 C^2 &= \frac{a}{bc} (b + b + 1)(c^2 - 2c + 1) \\
 BC &= \frac{a}{bc} (b^2 - 1)(c^2 - 1) \\
 BC^2 &= \frac{a}{bc} (b^2 - 1)(c^2 - 2c + 1) \\
 B^2C &= \frac{a}{bc} (b^2 - 2b + 1)(c^2 - 1) \\
 B^2C^2 &= \frac{a}{bc} (b^2 - 2b + 1)(c^2 - 2c + 1)
 \end{aligned}$$

These estimators were arrived at by the process usually employed in analyzing a 3^2 factorial arrangement into single degree of freedom comparisons.

Now consider the resulting analysis of variance. For the estimate of error, we shall pool the interaction terms which involve squares; that is, we take

$$\begin{aligned} SS_{\text{error}} &= \left[\frac{\widehat{B^2 C}}{12} \right]^2 + \left[\frac{\widehat{B C^2}}{12} \right]^2 + \left[\frac{\widehat{B^2 C^2}}{36} \right]^2 \\ &= \frac{a}{12bc} \left[(b-1)^2(c-1)^2 \left\{ (b-1)^2(c+1)^2 + (b+1)^2(c-1)^2 + \frac{1}{3} (b-1)^2(c-1)^2 \right\} \right]. \end{aligned} \quad (7.8)$$

Examination of the form of the contrasts in Table 7.2 shows that an interchange of the letters b, c results in an interchange of the letters B, C . Since (7.8) is not affected by the interchange of the letters b, c , we need only consider tests of significance for B, B^2, BC , eliminating C, C^2 through symmetry.

The contribution of B to the total sum of squares is given by

$$SS_B = \left[\frac{\widehat{B}}{6} \right]^2 = \frac{a^2}{6b^2c} (b-1)^2 (b+1)^2 (c^2 + c + 1)^2 \quad (7.9)$$

and the corresponding test of significance of B would estimate

$$\begin{aligned} F_B &= \frac{3SS_B}{SS_E} \\ &= \frac{6(b+1)^2(c^2+c+1)^2}{(c-1)^2 \left\{ (b-1)^2(c+1)^2 + (b+1)^2(c-1)^2 + \frac{1}{3}(b-1)^2(c-1)^2 \right\}}. \end{aligned} \quad (7.10)$$

Some values of F_B for various combinations of values of b and c have been calculated and are given in Table 7.3. In practice we would deem such an effect significant at the $\alpha = .05$ level if the value of F were greater than 10.13.

An examination of Table 7.3 shows that when b is small (and thus has little effect on y in (7.5)), the test of significance would indicate that B is significant at the $\alpha = .05$ level, given that $c < 10$. Also, when b is large and c is large, the test would indicate that B is non-significant. Taking the limit as b increases, it is seen by interpolation that significance would be obtained if c were smaller than 4.2.

In an analogous manner, we find the expression for the variance ratio of B^2 to the estimate of error as

$$F_B^2 = \frac{2(c^2 + c + 1)^2}{(c-1)^2 \left[(c+1)^2 \left[\frac{b+1}{b-1} \right]^2 \cdot (c-1)^2 + \frac{(c-1)^2}{3} \right]}, \quad (7.11)$$

Some numerical values of F_B^2 for various values of b and c are shown in Table 7.4. It would appear from an examination of Table 7.4 that the analysis would indicate that B^2 is a significant term whenever c is small. Likewise, it would appear that the analysis would show the B^2 term to be too small if b is small and c is large.

In a similar manner, we find the variance ratio of

TABLE 7.3 Calculated Values of F_B

c	Lim F_B $b \rightarrow 1$	b						Lim F_B $b \rightarrow \infty$
		1.1	1.5	2	3	5	10	
1.1	657,366	328,678	35,241	----	----	----	2,217.	1,486.
1.5	2166	2,048	----	----	----	----	120.6	82.25
2	294	----	----	----	88.20	57.10	40.56	28.45
3	69.62	----	----	46.99	33.42	23.79	17.84	13.05
5	22.52	----	20.41	17.50	13.68	10.48	8.25	6.25
10	11.26	11.26	10.50	9.36	7.73	6.21	5.06	3.98

TABLE 7.4 Calculated Values of F_B^2

c	b									
	1.1	1.2	1.5	2	3	5	10			
1.1	248	389	469	486	492	493	494			
1.2	29.45	68.34	113.18	127.07	---	---	---			
1.5	1.54	4.93	14.34	21.02	24.61	26.17	26.91			
2	0.217	0.752	2.85	5.34	7.35	8.71	9.05			
3	0.047	0.168	0.720	1.58	2.98	3.20	3.62			
5	0.017	0.061	0.272	0.648	1.14	1.55	1.84			
10	0.0084	0.030	0.140	0.347	0.642	0.922	1.13			

TABLE 7.5 Values of FBC

c	Lim FBC $b \rightarrow 1$	b						Lim FBC $b \rightarrow \infty$
		1.5	2	3	5	10		
1.5	225	111.75	58.58	30.68	18.35	12.53	8.54	
2	61	58.58	39.76	24.30	15.73	11.17	7.84	
3	36	30.68	24.30	17.28	12.30	9.23	6.75	
5	20.25	18.35	15.73	12.30	9.43	7.42	5.65	
10	13.45	12.53	11.18	9.23	7.42	6.05	4.75	

BC to the error term to be an estimate of

$$F_{BC} = 9 \left[\left(\frac{b-1}{b+1} \right) + \left(\frac{c-1}{c+1} \right) + \frac{1}{3} \left(\frac{b-1}{b+1} \right) + \left(\frac{c-1}{c+1} \right) \right]^{-1} \quad (7.12)$$

Some calculated values of F_{BC} are given in Table 7.5. It is seen that the BC term will tend to significance if both b and c are small.

As a whole, it is seen that very wrong conclusions may be drawn from an analysis of the type studied herein. If the transformation had been made, then the variance ratio for B would tend to be large if b were large, the variance ratios for B^2 and BC would be approximately equal to 1, as these effects are non-existent in the model. However, when the transformation was not made, the variance ratio for B tended to be large if b was small. Thus, for the B effect, we could conclude significance when the effect was really small and non-significance when the effect was large. The conclusions on the B^2 term would depend on c, but in general, the value of F_{B^2} will be small if b is small. The value of F_{BC} will tend to be large if b or c is small.

We have considered, herein, a three level factorial experiment. The same type results would apply in the two level case; interactions there would tend to be inflated if the coefficients were different.

All of the results obtained so far have been based

on an error free model. General statements concerning the effects of introduction of errors cannot be made. However, in one case calculated with random logarithmic-normal errors, the values did not depart far from those shown above.

VIII. GENERAL SUMMARY

Tests of significance for the hypothesis of homogeneity of means have been developed in the first part of this thesis for all of the usual common complete block experimental designs. It is seen that when the covariance matrix of errors associated with a complete block is completely unspecified or completely specified, valid tests may be easily derived. Between these two extremes, there are cases where the application of a few restrictions on the covariance matrix leads to intractible equations. Thus, it seems impossible to use, in these cases, this added knowledge in the derivation of tests for experimental designs. Such a case is the assumption of independent but heterogeneous errors as discussed on page 37. Of course, in this case, we might be able to develop iterative computational methods for obtaining estimates and use the large sample distribution of the likelihood ratio statistic in a test of significance. This, however, would not seem to be of value in the more interesting and important studies on the validity of the usual methods when heterogeneity of error is present.

It has been shown that the tests in the usual analysis of variance are relatively insensitive to some departures from the common assumptions. In the case of correlated errors, this may be due to the fact that the usual tests

are valid under the more general set of conditions of equal variances and equal covariances. Thus, the effect of the presence of covariance terms must be measured from some "average" covariance term rather than from zero.

We may always use the tests developed for Case 5. They are valid for any set of values of the variances and covariances. However, such use will result in tests of poorer power than tests based on some added knowledge concerning the variances and covariances. Consider, for example, the case cited previously of independent but heterogeneous errors. If we analyzed data having such errors by the standard methods of the analysis of variance, we would arrive at significance levels which are not true values. Using the tests developed herein, for Case V, we would find valid significance levels, but the power would be less than that obtained by use of the theoretically correct test.

In Part II, it was shown that in some cases, estimates of regression coefficients are unbiased when factor levels contain errors. No statements have been made concerning the effects on the analysis of variance of such errors. This would seem to be a fruitful area for research. Another fruitful area for future research is on further effects of non-additivity. The multiplicative models, especially the one drawn from the Theory of Dimensional

Analysis, seem ideally suited for alternate cases of study.

ACKNOWLEDGEMENTS

With a deep sense of gratitude I would like to acknowledge

- _____, for his aid and encouragement throughout the author's graduate studies,
- _____, for the inspiration and guidance he gave throughout the execution of this work, and
- my wife, _____, for the typing of this thesis and the many other loving encouragements which she has given over the past few years.

BIBLIOGRAPHY

Part I

1. Bradley, R. A., "The Distribution of the t and F Statistics for a Class of Non-normal Populations", Va. Journal of Science., Vol. 3 (1952) pp. 1-32.
2. Bradley, R. A., "Corrections for Normality in the Use of the Two-Sample t - and F - Tests at High Significance Levels", Annals of Math. Stat., Vol. 23 (1952) pp.103-113.
3. Box, G. E. P., and Andersen, S. L., "Robust Tests for Variances and Effect of Non-normality and Variance Heterogeneity on Standard Tests, Tech. Report 7, Institute of Statistics, Raleigh, N. C., 1954.
4. Bartlett, M. S., "The Use of Transformations". Biometrics, Vol. 3 (1947) pp. 39-52.
5. Cochran, W. G., "Some Consequences when the Assumptions for the Analysis of Variance are not Satisfied". Biometrics, Vol. 3 (1947) pp. 22-38.
6. Eisenhart, C., "The Assumptions Underlying the Analysis of Variance". Biometrics, Vol. 3 (1941) pp. 1-21.
7. Tukey, J. W., "One Degree of Freedom for Non-additivity", Biometrics, Vol. 5 (1949) pp. 232-242.
8. Tukey, J. W., and Anscombe, F. J., "Criticism of Transformations", unpublished manuscript.

9. Hsu, P. L., "Notes on Hotelling's Generalized T," Annals of Mathematical Statistics, Vol. 9 (1938) pp. 231-243.
10. Fix, E., Tables of the Non-central χ^2 Distribution, Berkeley, University of California Press, (1949).
11. Tang, P. C., "The Power Function of the Analysis of Variance Tests with Tables and Illustrations of Their Use," Stat. Res. Mem., 2 (1938) pp. 126-149.
12. Wilks, S. S., "Sample Criteria for Testing Equality of Means, Equality of Variances, and Equality of Covariances in a Normal Multivariate Distribution", Annals of Math. Stat., Vol. 17 (1946) pp. 257-281.
13. Russell, T. S., Progress Report No. 6, "Comparison of the Consistency of Technicians' Subjective Ratings", Virginia Polytechnic Institute, March, 1955.
14. Hotelling, H., "The Generalization of Student's Ratio," Annals of Math. Stat., 2 (1931) pp. 360-378.
15. Graybill, F., "Variance Heterogeneity in a Randomized Block Design," Biometrics, Vol. 10 No. 4, (1954) pp. 516-521.
16. Box, G. E. P., "Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems". Annals of Math. Stat., Vol. 25 (1954) pp. 290-302 and Vol. 25 (1954) pp. 484-498.
17. Kempthorne, O., The Design and Analysis of Experiments, (1952) John Wiley, New York.

18. Anderson, R. L., and Bancroft, T. A., Statistical Theory in Research, (1952) McGraw-Hill, New York.
19. Rao, C. R., Advanced Statistical Methods in Biometric Research, (1952) John Wiley, New York, p. 55.
20. Federer, W. T., Experimental Design, (1955) Macmillan, New York, p. 271.
21. Mood, A. M., Introduction to the Theory of Statistics, (1950) McGraw-Hill, New York, p. 181.
22. Brandt, A. E., "Tests of Significance in Reversal or Switchback Trials", Iowa. Agr. Exp. Sta. Res. Bul. 234 (1938) pp. 60-87.
23. Cochran, W. G. and Cox, G., Experimental Designs, (1950) John Wiley, New York.

Part II

1. Bartlett, M. S., "The Fitting of Straight Lines if Both Variables Are Subject to Error", Biometrics 5 (1948) p. 207.
2. Berkson, J., "Are There Two Regressions?", Journal of Am. Stat. Ass'n., 45 (1950) p. 250.
3. Box, G. E. P., and Wilson, K. B., "On the Experimental Attainment of Optimum Conditions," J. Roy. Stat. Soc., B, 13 (1951) p. 1.
4. Fisher, R. A., "The Goodness of Fit of Regression Formulae and the Distribution of Regression Coefficients" J. Roy. Stat. Soc., 85 (1922) 597.

5. Geary, R. C., "Non-linear Functional Relationship Between Two Variables When One Variable is Controlled", Journal of Am. Stat. Assn., 48 (1953) 94.
6. Kendall, M. G., "Regression, Structure, and Functional Relationship", Biometrika, 38 (1951) p. 11.
7. Lindley, D. V., "Regression Lines and Linear Functional Relationships", J. Roy. Stat. Soc. Suppl., 9 (1947) p. 21.

Part III

1. Yates, F., The Design and Analysis of Factorial Experiments, (1937) Imp. Bu. Soil Sci., Harpenden, England, p. 27.
2. Goulden, C. H., Methods of Statistical Analysis, (1952) John Wiley, New York, p. 245.
3. Quenouille, M. H., The Design and Analysis of Experiments, (1953) Hafner, New York, p. 63.
4. Tukey, J., (See Ref. 7, Part I).

**The vita has been removed from
the scanned document**

Some Results on Experimental Designs
when the Usual Assumptions are
Invalid

by

Hale Caterson Sweeny

ABSTRACT

SOME RESULTS ON EXPERIMENTAL DESIGNS WHEN THE USUAL ASSUMPTIONS ARE INVALID

Hale Caterson Sweeny

In the derivation of the theory of the analysis of variance, as related to complete block experimental designs, several assumptions are made: (1) the various effects contributing to the size of an observation act in an additive manner, (2) the errors associated with a complete block have jointly a multivariate normal distribution with zero means and variance-covariance matrix Σ , and (3) Σ has the form $\Sigma = I \sigma^2$, where I is the identity matrix. This last assumption is usually considered as two separate assumptions--independence of error terms and homogeneity of error variances. When the treatments to be tested constitute a set of factorial treatments, it is further assumed that, (4) the levels of the treatments are known without error. The work reported herein is divided into three parts. In the first part, tests of significance for the hypothesis of homogeneity of means are developed for four alternate cases concerning Σ . These four cases are: (1) Σ may be of any form, but Σ is known. (2) A matrix proportional to Σ is known; the constant of proportionality being unknown. (3) All

STATISTICAL INSTITUTE
LIBRARY
MAY 19 1964

diagonal terms in Σ are equal and all off-diagonal terms in Σ are equal, the values of the terms being unknown, and (4) Σ has an unknown general form.

To derive these tests, the multivariate sample is first subjected to a transformation affected by subtracting one of the variates in the vector of observations from the rest of the variates. This reduces the composite hypothesis of homogeneity of means to a simple hypothesis. The tests are then derived by the likelihood ratio approach. It is shown that the resulting tests are more general than first supposed, being applicable when block effects are present. In the first three cases, the tests are compared with the tests which can be derived in the absence of block effects to show power equalities.

The tests are then applied to the randomized block design, the Latin square design, and the split plot design. For the randomized block design, a study is made of the loss in power incurred in assuming case (4), when the assumptions of the usual analysis of variance should have been made.

The second part of the dissertation is concerned with assumption (4). Berkson has shown that if the independent variables are of the type known as "controlled" variables, then coefficients in a simple linear regression model may be estimated without bias by the usual least squares

formulae. This work was extended by Geary to polynomial regression. A theorem is proved to the effect that only two of the coefficients in a univariate polynomial regression may be estimated in an unbiased way by the use of least squares. Extensions to multivariate situations and to factorial designs are noted, especially in connection with the estimation of minima or maxima points in a response surface study.

In the third part of this work, the effects of a non-additive model in an unreplicated factorial experiment are studied. It has been suggested in the literature that in analyzing unreplicated factorial experiments, estimates of error may be obtained from higher order interactions. It is shown that if the data arises from a multiplicative model, this procedure may lead to entirely erroneous conclusions being drawn from a straightforward analysis of the data.