MIXED MODEL VARIANCE ANALYSIS WITH NORMAL ERROR AND POSSIBLY NON-NORMAL OTHER RANDOM EFFECTS: PART I: THE UNIVARIATE CASE[1]

By S. N. Roy and Whitfield Cobb

University of North Carolina

0.1. Introduction and summary. The mixed model with one factor represented by fixed effects, one factor by random effects, and a normal error, has often stipulated that these random effects be a sample drawn from a normally distributed population. In the case of a single response (or univariate) experiment, the variance of this normal distribution is a natural measure of the dispersion of these random effects, and confidence bounds on the ratio of this variance to the error variance [13], [14], [9] and simultaneous confidence bounds on both variances (in the latter case with a confidence coefficient $\geqq$ a specified value) [9] have already been found for certain classes of experimental designs. But when a distribution is not normal—or not assumed at the outset to be normal—the variance may not reveal as much about the distribution as some other measure such as interquartile range. In the present paper we seek confidence bounds on what, in a sense to be explained presently, might be called representations of the interquartile range and of analogous differences between higher order quantiles of the population from which the random effects are drawn. The method of obtaining these bounds involves an element of approximation comparable to grouping continuous data into k classes, since it replaces the actual random-effects variate by a "substitute variate" having k equally probable discrete values. The main idea is this. Let us assume, for simplicity of discussion, that we have a real valued stochastic variate. One comment here might be helpful. If the stochastic variate is observable, it seems natural to attempt to approximate its unknown distribution by introducing unknown probabilities over a finite set of preassigned class intervals, then trying to estimate these probabilities and then (especially for a continuous distribution) increasing the number of class intervals. On the other hand, if the variate is unobservable, as in the present set-up, it seems natural to try to approximate the distribution by replacing the stochastic variate by a "substitute" variate which is supposed to take, as a first approximation, two (unknown) values with equal probabilities, or as a second approximation, three (unknown) values with equal probabilities, or in general k (unknown) values with equal probabilities. We then try to estimate, in terms of our observations, these unknown values, which may be regarded as approximations to the 1st, 3rd, $\cdots$, $(2k - 1)$th quantiles of the unknown distribution. The random effects variate postulated in our model may have either a continuous

Received August 7, 1959; revised May 25, 1960.

[1] This research was sponsored by the United States Air Force through the Office of Scientific Research of the Air Research and Development Command.

or a discrete distribution, provided in the latter case there are enough distinct values to make these k quantiles meaningful parameters of the distribution. From now on, for brevity, we shall refer to these unknown values as the quantiles of the unknown distribution. It will be seen later that it is the differences between these unknown values rather than the unknown values themselves which we can estimate or make inferences about, and the number, $k - 1$, of such differences which can be estimated is restricted by the experiment.

Turning now to the confidence bounds, we observe that in the derivation of these bounds use is made of the same kind of sums of squares as in the normal variance components analysis. Unlike the more familiar confidence statements where the confidence coefficient may be specified at will, here except for the case of *two* blocks, only a lower bound on the confidence coefficient is specifiable, and this includes as a factor a decreasing function of k, the number of discrete values of the substitute variable. For $k = 2, 3, 4, 5$ the geometric shape of a $(k - 1)$-dimensional confidence region has been found.

It is also shown how the usual inference about the fixed effects can be made from this model and then how the above type of confidence bounds can be found for each of *several* random-effects factors in an experiment with orthogonal design.

In the case of a multiresponse—usually called multivariate—experiment, the model frequently stipulates that the random effects be samples from a multivariate normal population. Although the variance matrix of this distribution has a readily available estimator, confidence bounds have presented many difficulties. For the extremely restricted model in which the variance matrix of each random-effects factor is proportional to the variance matrix of the error, Roy and Gnanadesikan [10] obtained simultaneous confidence bounds on the characteristic roots of this latter matrix and on the proportionality constants. In Part II of this paper the authors present (with a confidence coefficient greater than or equal to a preassigned value) confidence bounds on the characteristic roots of the variance-covariance matrix of a random-effects variate without assuming any such relation to the error matrix. The second part will also consider the p-variate extension of the univariate substitute variate and the associated confidence bounds for the case where the p-dimensional distribution of the random effects is not necessarily normal. This development will be only indicated in principle for $p > 2$ but will be discussed in some detail for the case $p = 2$.

0.2. Notation and presuppositions. A general m by n matrix will be denoted by a capital Latin letter from either end of the alphabet, say $\mathbf{A}(m \times n)$, its transpose by $\mathbf{A}'(n \times m)$, but certain special letters will denote special types of matrices. Thus $\mathbf{I}(m)$ denotes the m-rowed identity matrix; $\mathbf{J}(m \times n)$ has every element unity; $\mathbf{K}((m + 1) \times m)$ is $\mathbf{I}(m)$ bordered below by a row of zeros; $\mathbf{O}(m \times n)$ has every element zero. Triangular matrices are denoted by $\mathbf{T}(m)$, orthonormal by $\mathbf{L}(m \times n)$. Repeated use is made of an $(m - 1) \times m$ orthonormal matrix ($m - 1$ mutually orthogonal rows of m elements) all of whose rows are orthogonal to a row of m identical elements. A (but by no means the only) matrix having these properties is readily obtained by removing the row of identical elements from the matrix of the Helmert Transformation [Kendall and Buckland, *A Dictionary of Statistical Terms*, p. 126]. For this reason we denote any such matrix by $\mathbf{H}((m - 1) \times m)$. The maximum and minimum characteristic roots of the matrix $\mathbf{A}$ are denoted by $\text{ch}_{\max}(\mathbf{A})$ and $\text{ch}_{\min}(\mathbf{A})$ respectively. A column vector of m components is denoted by a lower case letter such as $\mathbf{a}(m)$, whereas $\mathbf{a}'(m)$ denotes a row vector. In particular $\mathbf{o}$ and $\mathbf{j}$ denote vectors whose components are all zero and all unity, respectively. After the dimensions of a matrix or vector have been indicated, that part of the symbol may be omitted in subsequent references to the same matrix or vector.

Existence theorems for triangular matrix factors of certain matrices are proved in [12].

The direct sum of $\mathbf{A}(m \times n)$ and $\mathbf{B}(p \times q)$ is defined to be

$$\begin{bmatrix} \mathbf{A}(m \times n), \mathbf{O}(m \times q) \\ \mathbf{O}(p \times n), \mathbf{B}(p \times q) \end{bmatrix}$$

and is denoted by $\mathbf{A} \dotplus \mathbf{B}$. The left direct product (Kronecker product) of $\mathbf{A}(m \times n)$ and $\mathbf{B}(p \times q)$ is defined to be

$$\begin{bmatrix} b_{11}\mathbf{A}, & b_{12}\mathbf{A}, & \cdots, & b_{1q}\mathbf{A} \\ b_{21}\mathbf{A}, & b_{22}\mathbf{A}, & \cdots, & b_{2q}\mathbf{A} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ b_{p1}\mathbf{A}, & b_{p2}\mathbf{A}, & \cdots, & b_{pq}\mathbf{A} \end{bmatrix}$$

and is denoted by $\mathbf{A} \cdot\times \mathbf{B}$. Both the notation and the properties of direct sums and products may be found in [5]. Inverses of partitioned and patterned matrices will be found by the methods of [1] and [11].

Some properties of special matrices are listed in Appendix A and referred to by number, e.g. (A.4), when specifically used.

1. Random Effects in a Two-Factor Mixed Model Uniresponse Problem.

1.1. *The two-factor model and its structure matrix.* We start with a general two-factor model in which each observation is assumed to be the sum of three terms, the first two corresponding to the two factors or criteria by which the experimental units are classified and the third term being of the nature of an error. Included in this model is the postulate that the n errors, in a set of n observations, are independently and identically distributed normal deviates, each with distribution denoted by $N(0, \sigma^2)$. We shall suppose that there are t categories in the first classification and s in the second. The unobservable terms or elements of the model we denote by a_i, $i = 1, 2, \cdots, t$, and b_j, $j = 1, 2, \cdots, s$. For convenience we shall refer to these a's and b's as treatment effects and block effects respectively, but this designation should not limit the application or prejudice the interpretation of the model. The important distinction is that whereas the a's are regarded as unknown but fixed constants, the b's are assumed to be a sample of size s from some unknown (but presumably continuous) distribution, which may or may not be normal but is postulated to be independent of the normal error.

In any planned experiment there will be n experimental units, say one from each block-treatment cell. This is before we have made any observations at all on any unit. Now according as we make one or several types of observations on each unit, it is a univariate or multivariate (that is, a uniresponse or a multiresponse) problem. The response from each experimental unit would presumably depend on the block-treatment cell the unit came from. For a uniresponse problem we have n observations on the n experimental units. If the observations, treatment effects, block effects, and errors are respectively ordered and written as column vectors—$\mathbf{y}(n)$, $\mathbf{a}(t)$, $\mathbf{b}(s)$, $\boldsymbol{\varepsilon}(n)$—then the two-factor model described above can be represented by

$$(1.1.1) \qquad \mathbf{y}(n) = \mathbf{M}(n \times (s + t)) \begin{bmatrix} \mathbf{a}(t) \\ \mathbf{b}(s) \end{bmatrix} + \boldsymbol{\varepsilon}(n),$$

where

$$(1.1.2) \qquad \mathbf{M}(n \times (s + t)) = [\mathbf{M}_0(n \times t), \mathbf{M}_1(n \times s)].$$

That $\mathbf{M}$ is partitioned into two submatrices and that the rank$(\mathbf{M}) \leqq s + t - 1$ result from the basic assumptions of the two-factor mixed model as stated above. Hence $\mathbf{M}$ may with some justification be called the "model matrix" for the observations $\mathbf{y}$. On the other hand, the actual elements of $\mathbf{M}$ are not determined until a specific experimental design has been selected and each experimental unit uniquely classified according to a field plan consistent with this design. Hence $\mathbf{M}$ has frequently been called the "design matrix." Throughout this investigation, both the particular field plan and the type of design will be left unspecified, but the general pattern or structure of $\mathbf{M}$ will be known. For these reasons we shall call $\mathbf{M}$ by the less specific name, "structure matrix."

We require that the design be connected, but we do not at present specify whether complete or incomplete, balanced or partially balanced. In any of these cases an experimental unit will belong to one and only one category of each factor, and hence each row of each submatrix, $\mathbf{M}_0$ and $\mathbf{M}_1$, will consist of zeros except for unity in some one column. Thus for a two-factor model the structure matrix will always be such that

$$(1.1.3) \qquad \mathbf{M}(n \times (s + t)) \begin{bmatrix} \mathbf{j}(t) \\ -\mathbf{j}(s) \end{bmatrix} = \mathbf{M}_0\mathbf{j} - \mathbf{M}_1\mathbf{j} = \mathbf{o}.$$

In what follows we shall consider this to be the only independent linear relation among the columns of $\mathbf{M}$. Since any useful design will have $n \geqq s + t$, it follows from (1.1.3) that rank$(\mathbf{M}) = s + t - 1 \equiv r^*$, and hence any r^* columns of $\mathbf{M}$ would constitute a basis for $\mathbf{M}$. We agree to use the first r^* columns and to denote this basis by $\mathbf{M}_I(n \times r^*)$. Thus, by (A.4),

$$\begin{aligned} \mathbf{M}_I(n \times r^*) &= \mathbf{M}(n \times (s+t))\mathbf{K}((s+t) \times (s+t-1)) \\ &= [\mathbf{M}_0(n \times t), \mathbf{M}_1(n \times s)\mathbf{K}(s \times (s-1))], \end{aligned} \tag{1.1.4}$$

where $\mathbf{K}$ is the special matrix defined in Section 0.2. Then $\mathbf{M}$ is expressible in terms of its basis,

$$\mathbf{M}(n \times (s+t)) = \mathbf{M}_I(n \times r^*)[\mathbf{I}(r^*), \mathbf{f}(r^*)], \tag{1.1.5}$$

where $\mathbf{f}'(r^*) = [\mathbf{j}'(t), -\mathbf{j}'(s-1)]$.

1.2. *Three orthonormal transformations and the resulting statistics.* Whether the ultimate purpose is estimation, testing hypotheses, or confidence bounds, statistical inference about the block effects will require a set of statistics whose distribution depends upon $\mathbf{b}$ and not upon $\mathbf{a}$; and statistical inference about the treatment effects will require a set of statistics whose distribution depends upon $\mathbf{a}$ and not upon $\mathbf{b}$. To this end we define $\mathbf{u}$ and $\mathbf{v}$ by the following transformation on the observations $\mathbf{y}$:

$$\begin{aligned} &\begin{bmatrix} \mathbf{u}(t-1) \\ \mathbf{v}(s-1) \end{bmatrix} \\ &\equiv [\mathbf{T}_0^{-1}(t-1)\mathbf{H}((t-1) \times t) + \mathbf{T}_1^{-1}(s-1)\mathbf{H}((s-1) \times s) \\ &\qquad \cdot \mathbf{K}(s \times (s-1))]\,[\mathbf{M}_I'\mathbf{M}_I]^{-1}\mathbf{M}_I'\mathbf{y}, \end{aligned} \tag{1.2.1}$$

where $\mathbf{T}_0$ and $\mathbf{T}_1$ are lower triangular matrices defined by

$$\begin{aligned} \mathbf{T}_0\mathbf{T}_0' \equiv [\mathbf{H}((t-1) \times t), \mathbf{O}((t-1) \times (s-1))][\mathbf{M}_I'\mathbf{M}_I]^{-1} \\ \cdot \begin{bmatrix} \mathbf{H}'(t \times (t-1)) \\ \mathbf{O}((s-1) \times (t-1)) \end{bmatrix} \end{aligned} \tag{1.2.2}$$

and

$$\begin{aligned} \mathbf{T}_1\mathbf{T}_1' \equiv [\mathbf{O}((s-1) \times t), \mathbf{H}((s-1) \times s)\mathbf{K}(s \times (s-1))][\mathbf{M}'\mathbf{M}]^{-1} \\ \cdot \begin{bmatrix} \mathbf{O}(t \times (s-1)) \\ \mathbf{K}'((s-1) \times s)\mathbf{H}'(s \times (s-1)) \end{bmatrix}. \end{aligned} \tag{1.2.3}$$

A convention as to sign makes $\mathbf{T}_0$ and $\mathbf{T}_1$ unique.

Although (1.2.1) was just used to define the statistics $\mathbf{u}$ and $\mathbf{v}$, it may be combined with

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_0((t-1) \times n) \\ \mathbf{L}_1((s-1) \times n) \end{bmatrix} \mathbf{y} \tag{1.2.4}$$

to define the matrices $\mathbf{L}_0$ and $\mathbf{L}_1$. For any $\mathbf{M}$ meeting the specifications of Section 1.1, $\mathbf{L}_0$ and $\mathbf{L}_1$ are individually orthonormal, that is, $\mathbf{L}_0\mathbf{L}_0' = \mathbf{I}$ and $\mathbf{L}_1\mathbf{L}_1' = \mathbf{I}$. However, it is only under additional restrictions to be discussed later (defining what are called orthogonal designs) that $\mathbf{L}_0$ would be orthogonal to $\mathbf{L}_1$, that is, $\mathbf{L}_0\mathbf{L}_1' = \mathbf{O}$.

By (A.3.11) of [12], the basis $\mathbf{M}_I$ of the structure matrix $\mathbf{M}$ determines an orthonormal matrix $\mathbf{L}_*$ such that

$$\mathbf{M}_I' = \mathbf{T}(r^*)\mathbf{L}_*(r^* \times n). \tag{1.2.5}$$

By a convention of sign, both $\mathbf{T}$ and $\mathbf{L}_*$ can be made unique. Then any (not unique) orthonormal matrix $\mathbf{L}((n - r^*) \times n)$ such that

$$\begin{bmatrix} \mathbf{L} \\ \mathbf{L}_* \end{bmatrix}\begin{bmatrix} \mathbf{L} \\ \mathbf{L}_* \end{bmatrix}' = \begin{bmatrix} \mathbf{I}, & \mathbf{O} \\ \mathbf{O}, & \mathbf{I} \end{bmatrix} \tag{1.2.6}$$

is used to define a third set of statistics

$$\mathbf{w}(n - r^*) \equiv \mathbf{L}\mathbf{y}. \tag{1.2.7}$$

Under the same general specifications of Section 1.1, the matrix

$$\mathbf{L}((n - r^*) \times n),$$

which has been defined in (1.2.6) as the orthogonal completion of $\mathbf{L}_*$, is easily shown to be orthogonal to both $\mathbf{L}_0$ and $\mathbf{L}_1$. That is, $\mathbf{L}\mathbf{L}_0' = \mathbf{O}$ and $\mathbf{L}\mathbf{L}_1' = \mathbf{O}$. Thus $\mathbf{w}$ is a stochastic variate independent of both $\mathbf{u}$ and $\mathbf{v}$ whether or not the stochastic variates $\mathbf{u}$ and $\mathbf{v}$ are mutually independent. When (1.2.5), (1.1.5), and (1.1.1) are appropriately substituted into (1.2.7), the mutual orthogonality of $\mathbf{L}$ and $\mathbf{L}_*$ permits the simplification

$$\begin{aligned} \mathbf{w} &= \mathbf{L}\left\{\mathbf{L}_*'\mathbf{T}'[\mathbf{I}, \mathbf{f}]\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} + \boldsymbol{\varepsilon}\right\} \\ &= \mathbf{L}\,\boldsymbol{\varepsilon}, \end{aligned} \tag{1.2.8}$$

which shows that $\mathbf{w}$ is entirely free of both treatment effects and block effects.

Similarly when (1.1.5), (1.1.1) and (1.2.4) are appropriately substituted into (1.2.1), the following simplification is made possible by (A.1) and (A.7)

$$\begin{aligned} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} &= \begin{bmatrix} \mathbf{T}_0^{-1}\mathbf{H}, & \mathbf{O} \\ \mathbf{O}, & \mathbf{T}_1^{-1}\mathbf{H}\mathbf{K} \end{bmatrix}[\mathbf{M}_I'\mathbf{M}_I]^{-1}\mathbf{M}_I'\left\{\mathbf{M}_I[\mathbf{I}, \mathbf{f}]\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} + \boldsymbol{\varepsilon}\right\} \\ &= \begin{bmatrix} \mathbf{T}_0^{-1}\mathbf{H}, & \mathbf{O} \\ \mathbf{O}, & \mathbf{T}_1^{-1}\mathbf{H}\mathbf{K} \end{bmatrix}[\mathbf{I}, \mathbf{f}]\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} + \begin{bmatrix} \mathbf{L}_0 \\ \mathbf{L}_1 \end{bmatrix}\boldsymbol{\varepsilon} \\ &= \begin{bmatrix} \mathbf{T}_0^{-1}\mathbf{H}, & \mathbf{O} \\ \mathbf{O}, & \mathbf{T}_1^{-1}\mathbf{H}\mathbf{K} \end{bmatrix}\begin{bmatrix} \mathbf{I}, & \mathbf{O}, & \mathbf{j} \\ \mathbf{O}, & \mathbf{I}, & -\mathbf{j} \end{bmatrix}\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} + \begin{bmatrix} \mathbf{L}_0 \\ \mathbf{L}_1 \end{bmatrix}\boldsymbol{\varepsilon} \\ &= \begin{bmatrix} \mathbf{T}_0^{-1}\mathbf{H}, & \mathbf{O}, & \mathbf{o} \\ \mathbf{O}, & \mathbf{T}_1^{-1}\mathbf{H}\mathbf{K}, & -\mathbf{T}_1^{-1}\mathbf{H}\mathbf{K}\mathbf{j} \end{bmatrix}\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} + \begin{bmatrix} \mathbf{L}_0 \\ \mathbf{L}_1 \end{bmatrix}\boldsymbol{\varepsilon} \\ &= \begin{bmatrix} \mathbf{T}_0^{-1}\mathbf{H}, & \mathbf{O} \\ \mathbf{O}, & \mathbf{T}_1^{-1}\mathbf{H} \end{bmatrix}\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} + \begin{bmatrix} \mathbf{L}_0 \\ \mathbf{L}_1 \end{bmatrix}\boldsymbol{\varepsilon} \\ &= \begin{bmatrix} \mathbf{T}_0^{-1}\mathbf{H}\mathbf{a} + \mathbf{L}_0\boldsymbol{\varepsilon} \\ \mathbf{T}_1^{-1}\mathbf{H}\mathbf{b} + \mathbf{L}_1\boldsymbol{\varepsilon} \end{bmatrix}. \end{aligned} \tag{1.2.9}$$

Thus (1.2.9) shows both the general relevance of the statistics **u** and **v** to the unobservable effects **a** and **b** postulated in the model and also the simple form of this relationship.

The particular merits of **u**, **v**, and **w** are (1) the orthonormality of **L**, $\mathbf{L}_0$, and $\mathbf{L}_1$ has preserved for **w** and the error terms of **u** and **v** the same independent normal distribution postulated for $\boldsymbol{\varepsilon}$; (2) the number of components in **u**, **v**, and **w**, is the number of degrees of freedom belonging to **a**, **b**, and error respectively; (3) any estimable contrast among the components of **a** or **b** can be expressed easily in terms of **u** or **v**; (4) in a two-factor fixed-effects model the vanishing of $\mathcal{E}(\mathbf{u})$ or $\mathcal{E}(\mathbf{v})$ is a necessary and sufficient condition for the equality of all fixed components of **a** or **b** respectively—the condition usually described as "no treatment effect" or "no block effect"; (5) sums of squares for testing such null hypotheses on the assumption of fixed effects may be obtained as inner products, $\mathbf{u}'\mathbf{u}$ and $\mathbf{v}'\mathbf{v}$, or the *same* sums of squares may be obtained without ever finding $\mathbf{T}_0$ and $\mathbf{T}_1$ explicitly; (6) any other testable hypothesis about the components of **a** or **b** can be tested using sums of squares easily obtained from **u**, **v**, and **w**, or expressible in terms of them even when obtained differently. But the chief use to which **v** and **w** will be put in the present paper is making some inferences about an unknown population on the assumption that **b** consists of s independent and identically distributed random samples from that population.

1.3. *Quasi-confidence bounds.* Since $\mathbf{L}_1\boldsymbol{\varepsilon}$ and **w** are both $N(\mathbf{o}, \sigma^2\mathbf{I})$, $\boldsymbol{\varepsilon}'\mathbf{L}_1'\mathbf{L}_1\boldsymbol{\varepsilon}/\sigma^2$ and $\mathbf{w}'\mathbf{w}/\sigma^2$ are both central chi square variates with $s - 1$ and $n - r^*$ d.f. respectively. Moreover, since these two chi squares are independent,

$$(n - r^*)\boldsymbol{\varepsilon}'\mathbf{L}_1'\mathbf{L}_1\boldsymbol{\varepsilon}/(s - 1)\mathbf{w}'\mathbf{w}$$

has the variance ratio distribution. Thus for a chosen $\alpha_1 < 1$, there is a constant F_{α_1} such that

$$\text{Pr}\left\{\frac{(n - r^*)\boldsymbol{\varepsilon}'\mathbf{L}_1'\mathbf{L}_1\boldsymbol{\varepsilon}}{(s - 1)\mathbf{w}'\mathbf{w}} \leqq F_{\alpha_1}\right\} = 1 - \alpha_1. \tag{1.3.1}$$

It is important to note that (1.3.1) is true regardless of the population from which **b** is a sample and regardless of the computed values of **v**.

We now proceed from the probability statement (1.3.1), which involves only normal error components, to a confidence-type statement about the unobservable block effects. First the error vector $\mathbf{L}_1\boldsymbol{\varepsilon}$ is expressed as the difference between the computable statistic **v** and the postulated random vector $\mathbf{T}_1^{-1}\mathbf{Hb}$. Thus the inequality in (1.3.1) becomes

$$[\mathbf{v} - \mathbf{T}_1^{-1}\mathbf{Hb}]'[\mathbf{v} - \mathbf{T}_1^{-1}\mathbf{Hb}] \leqq (s - 1)\mathbf{w}'\mathbf{w}F_{\alpha_1}/(n - r^*)$$

or

$$|\,\mathbf{v} - \mathbf{T}_1^{-1}\mathbf{Hb}\,| \leqq ((s - 1)\mathbf{w}'\mathbf{w}F_{\alpha_1}/(n - r^*))^{\frac{1}{2}}. \tag{1.3.2}$$

For $s = 2$, v and $\mathbf{T}_1^{-1}\mathbf{Hb}$ are scalars such that

$$\begin{aligned} v - ((s - 1)\mathbf{w}'\mathbf{w}F_{\alpha_1}/(n - r^*))^{\frac{1}{2}} &\leqq \mathbf{T}_1^{-1}\mathbf{Hb} \\ &\leqq v + ((s - 1)\mathbf{w}'\mathbf{w}F_{\alpha_1}/(n - r^*))^{\frac{1}{2}} \end{aligned} \tag{1.3.2a}$$

with the same probability, $1 - \alpha_1$, as in (1.3.1). For $s \geqq 2$, (1.3.2) implies, but is not implied by,

$$(1.3.3) \qquad \begin{aligned} |\mathbf{v}| - ((s-1)\mathbf{w}'\mathbf{w}F_{\alpha_1}/(n-r^*))^{\frac{1}{2}} &\leqq |\mathbf{T}_1^{-1}\mathbf{Hb}| \\ &\leqq |\mathbf{v}| + ((s-1)\mathbf{w}'\mathbf{w}F_{\alpha_1}/(n-r^*))^{\frac{1}{2}}. \end{aligned}$$

Hence (1.3.3) is true with probability no less than $1 - \alpha_1$ and possibly greater. Since the middle member of (1.3.3) is necessarily non-negative, the left member may be replaced by a suitably defined non-negative bound. Thus we let

$$(1.3.4) \qquad \begin{aligned} l_1 &\equiv [\mathbf{v}'\mathbf{v}]^{\frac{1}{2}} - [(s-1)\mathbf{w}'\mathbf{w}F_{\alpha_1}/(n-r^*)]^{\frac{1}{2}} && \text{if } > 0, \\ l_1 &\equiv 0 && \text{otherwise,} \end{aligned}$$

and

$$l_2 \equiv [\mathbf{v}'\mathbf{v}]^{\frac{1}{2}} + [(s-1)\mathbf{w}'\mathbf{w}'F_{\alpha_1}/(n-r^*)]^{\frac{1}{2}}.$$

The confidence-type statement obtained from (1.3.1) is

$$(1.3.5) \qquad \Pr\{l_1 \leqq [\mathbf{b}'\mathbf{H}'(\mathbf{T}_1\mathbf{T}_1')^{-1}\mathbf{Hb}]^{\frac{1}{2}} \leqq l_2\} \geqq 1 - \alpha_1 .$$

The bounds l_1 and l_2 are determined by the computable statistics $\mathbf{v}$ and $\mathbf{w}$ and the chosen α_1. But since $[\mathbf{b}'\mathbf{H}'(\mathbf{T}_1\mathbf{T}_1')^{-1}\mathbf{Hb}]^{\frac{1}{2}}$ is not a parameter of a distribution, (1.3.5) is not a genuine confidence statement. It is here called a "quasi-confidence" statement. Because it may be used to obtain a confidence statement ultimately, (1.3.5) may also be called a "preliminary" confidence statement.

Since in practice α_1 would be chosen to be small, the increase in the probability of (1.3.5) over the probability of (1.3.1) will be very small in comparison with $1 - \alpha_1$. This increase, by itself, need not be disturbing. What is disturbing is that the quasi-confidence interval, from l_1 to l_2, may be too wide. It is quite possible that even in the near future quasi-confidence intervals narrower than that of (1.3.5) may be found by others if not by us. In the meantime, in an application the interval of (1.3.5) may not be too wide from a practical standpoint. We have not tried to improve upon (1.3.5) since the main emphasis of this paper is on the inferences which can be made without assuming a distribution for b, but working in terms of the substitute variate in the sense briefly explained in the introduction and to be developed in Section 1.5. For $s = 2$, that is, for the case of two blocks, it is always open to us to go back to (1.3.2a) and use, instead of (1.3.5),

$$(1.3.6) \qquad \Pr\{l_1^* \leqq \mathbf{T}_1^{-1}\mathbf{Hb} \leqq l_2^*\} = 1 - \alpha_1 ,$$

where l_1^* and l_2^* are the left and right sides respectively of the inequality (1.3.2a). It is also open to us to make corresponding changes at all subsequent stages, keeping in mind the limitations stated after (1.5.2) and (1.6.1) imposed by an s so small. We believe that the idea of a substitute variate and the kind of use which will be made of it in this paper could have much wider application, especially when dealing with unobservable stochastic variates with distributions about which it might be wise to make as few assumptions as possible.

1.4. *Simplifications and restrictions.* Although $\mathbf{T}_1$ appears in the formal definition of $\mathbf{v}$, (1.2.2) and (1.2.1) may be combined to obtain the sum of squares

$$(1.4.1)\quad \begin{aligned} \mathbf{v}'\mathbf{v} = \mathbf{y}'\mathbf{M}_I[\mathbf{M}_I'\mathbf{M}_I]^{-1}\begin{bmatrix}\mathbf{O}\\ \mathbf{K}'\mathbf{H}'\end{bmatrix}\Big\{[\mathbf{O}, \mathbf{HK}][\mathbf{M}_I'\mathbf{M}_I]^{-1} \\ \times\begin{bmatrix}\mathbf{O}\\ \mathbf{K}'\mathbf{H}'\end{bmatrix}\Big\}^{-1}[\mathbf{O}, \mathbf{HK}][\mathbf{M}_I'\mathbf{M}_I]^{-1}\mathbf{M}_I'\mathbf{y}\end{aligned}$$

in which $\mathbf{T}_1$ no longer appears explicitly. Moreover, $\mathbf{T}_1$ enters into the middle member of (1.3.5) only in the symmetric matrix

$$(1.4.2)\qquad \mathbf{H}'(\mathbf{T}_1\mathbf{T}_1')^{-1}\mathbf{H}.$$

Using the definition of $\mathbf{T}_1$ in (1.2.2), (1.4.2) can be expressed in terms of $\mathbf{M}_I$. Also $\mathbf{M}_I'\mathbf{M}_I$ can by means of (1.1.4) be expressed as

$$(1.4.3)\qquad \begin{bmatrix}\mathbf{M}_0'\mathbf{M}_0, & \mathbf{M}_0'\mathbf{M}_1\mathbf{K}\\ \mathbf{K}'\mathbf{M}_1'\mathbf{M}_0, & \mathbf{K}'\mathbf{M}_1'\mathbf{M}_1\mathbf{K}\end{bmatrix},$$

and then by (A.9) the inverse of this $(s + t - 1)$th order matrix can be expressed as a partitioned matrix whose submatrices involve inverses of only tth order and $(s - 1)$th order matrices. Using this result, (1.4.1) reduces to

$$(1.4.4)\quad \begin{aligned}\mathbf{v}'\mathbf{v} = \mathbf{y}'\{\mathbf{I} - \mathbf{M}_0[\mathbf{M}_0'\mathbf{M}_0]^{-1}\mathbf{M}_0'\}\mathbf{M}_1\mathbf{K}\{\mathbf{K}'\mathbf{M}_1'\mathbf{M}_1\mathbf{K} - \mathbf{K}'\mathbf{M}_1'\mathbf{M}_0 \\ \times [\mathbf{M}_0'\mathbf{M}_0]^{-1}\mathbf{M}_0'\mathbf{M}_1\mathbf{K}\}^{-1}\mathbf{K}'\mathbf{M}_1'\{\mathbf{I} - \mathbf{M}_0[\mathbf{M}_0'\mathbf{M}_0]^{-1}\mathbf{M}_0'\}\mathbf{y}\end{aligned}$$

and (1.4.2) becomes

$$(1.4.5)\qquad \mathbf{H}'[\mathbf{K}'\mathbf{H}']^{-1}\mathbf{K}'\{\mathbf{M}_1'\mathbf{M}_1 - \mathbf{M}_1'\mathbf{M}_0[\mathbf{M}_0'\mathbf{M}_0]^{-1}\mathbf{M}_0'\mathbf{M}_1\}\mathbf{K}[\mathbf{HK}]^{-1}\mathbf{H}.$$

By means of (A.8), whose conditions are satisfied, (1.4.5) is further reduced to

$$(1.4.6)\qquad \mathbf{M}_1'\mathbf{M}_1 - \mathbf{M}_1'\mathbf{M}_0[\mathbf{M}_0'\mathbf{M}_0]^{-1}\mathbf{M}_0'\mathbf{M}_1.$$

These simplifications have not required any more restrictive assumptions about the structure matrix than those set forth in Section 1.1 above. But even greater simplification is possible for certain types of experimental design.

Suppose (1) every treatment is applied to r experimental units (in r distinct blocks). Suppose (2) every block contains q experimental units (to which q distinct treatments are applied). Then, for any design satisfying these two specifications, (1.4.6) reduces to

$$(1.4.7)\qquad q\mathbf{I}(s) - r^{-1}[\mathbf{M}_0'\mathbf{M}_1]'\mathbf{M}_0'\mathbf{M}_1.$$

As a further restriction, suppose (3) for every pair of blocks the number of treatments-in-common is $q(r - 1)/(s - 1)$; then (1.4.7) becomes

$$(1.4.8)\qquad \delta^2[\mathbf{I}(s) - s^{-1}\mathbf{J}(s)],$$

where $\delta^2 \equiv qs(r - 1)/r(s - 1)$. The δ^2 is thus a positive, rational number determined by the size of the experiment and the type of experimental design. It is accordingly called the "design constant" for a particular experiment. Designs which satisfy all three of these restrictions have been called "linked block" by Youden [16]. They include randomized block (for which $\delta^2 = q = t$), symmetric balanced incomplete block designs, and those partially balanced incomplete block designs which are duals of balanced incomplete block designs. Henceforth it is to be understood that the experimental design is one of these types which satisfy the linked block conditions.

The underlying motivation for these three restrictions on the design is to reduce (1.4.2) to (1.4.8) or, in other words to make $\mathbf{b}'\mathbf{H}'(\mathbf{T}_1\mathbf{T}_1')^{-1}\mathbf{H}\mathbf{b}$, which appears in the quasi-confidence statement (1.3.5), proportional to the corrected sum of squares of the block effects,

$$\sum_{i=1}^{s} b_i^2 - s^{-1}\left(\sum_{i=1}^{s} b_i\right)^2. \tag{1.4.9}$$

But when these three restrictions do hold, (1.4.4) can be further simplified to

$$\mathbf{v}'\mathbf{v} = \delta^{-2}\mathbf{y}'[\mathbf{I} - r^{-1}\mathbf{M}_0\mathbf{M}_0']\mathbf{M}_1\mathbf{K}[\mathbf{I} + \mathbf{J}]\mathbf{K}'\mathbf{M}_1'[\mathbf{I} - r^{-1}\mathbf{M}_0\mathbf{M}_0']\mathbf{y}, \tag{1.4.10}$$

which requires no matrix inversion whatever. Moreover, when (1.4.8) is set equal to (1.4.2), it follows by (A.2) and (A.1) that

$$\mathbf{T}_1 = \delta^{-1}\mathbf{I}, \tag{1.4.11}$$

where the positive root of δ^2 will make $\mathbf{T}_1$ agree with the convention of having positive diagonal terms.

1.5. *Class marks and "substitute variates."* The model specifies that each component of $\mathbf{b}$ is a sample from an unknown distribution that is presumably either continuous or discrete with a sufficient number of values. Some measure of the dispersion of that distribution is sought, but the variance, however suitable for this purpose in the case of normal distributions, may be inappropriate for other distributions. However, any continuous distribution has uniquely determined quantiles, for which we introduce the following notation: ${}_m\beta_n$ will denote the nth m-tile of the distribution of the variate b. Thus ${}_2\beta_1$ denotes the median; ${}_4\beta_1$ and ${}_4\beta_3$ denote the odd quartiles, etc. These quantiles may be used as class boundaries and class marks for approximating the unknown continuous distribution by a discrete distribution as follows. Suppose the quartiles of b were known: ${}_4\beta_0$, ${}_4\beta_1$, ${}_4\beta_2$, ${}_4\beta_3$, ${}_4\beta_4$. The even quartiles ${}_4\beta_0$, ${}_4\beta_2$, and ${}_4\beta_4$ used as class boundaries would lump all possible values of b into two classes. Instead of the midrange of each class, the median of each class could be used as the class mark, and a new variate taking on only these discrete class marks as its values would be a rough approximation to the original variate b. Since this new variate in a sense replaces the original, it will be called the "substitute variate" and will be denoted by ${}_4b$ if its two possible values are the odd quartiles of b. The same remarks would apply roughly for a discrete distribution with a sufficient number of values.

This example of the quartiles may be extended to sextiles, octiles, $\cdots$, or

$2k$-tiles. The even quantiles—${}_{2k}\beta_0$, ${}_{2k}\beta_2$, $\cdots$, ${}_{2k}\beta_{2k}$—could be made class boundaries, and the odd quantiles—${}_{2k}\beta_1$, ${}_{2k}\beta_3$, $\cdots$, ${}_{2k}\beta_{2k-1}$—taken as the class marks. These k discrete values would thus be the only values of a substitute variate denoted by ${}_{2k}b$. Since these class boundaries are not arbitrarily chosen but are the even $2k$-tiles of the population of b, each class has the same *a priori* probability, viz., $1/k$, regardless of the shape of the density curve of b. Thus the substitute variate ${}_{2k}b$ has k equally-probable values no matter what the unknown distribution of b. In other words,

$$\Pr\{{}_{2k}b = {}_{2k}\beta_{2n-1}\} = 1/k, \qquad n = 1, 2, \cdots, k. \tag{1.5.1}$$

Of course we are not interested in the probabilities $1/k$, which are known, but in the unknown values ${}_{2k}\beta_{2n-1}$. The actual values of the quantiles of b would certainly be desirable but seem no more inferable than the actual block effects or treatment effects. On the other hand, just as contrasts among fixed effects may be estimated and confidence bounds placed on them, so may differences between values of the substitute variate be estimated and have confidence bounds placed on them. The interquartile range, ${}_4\beta_3 - {}_4\beta_1$, is one such difference. And for k an integer greater than 2,

$${}_{2k}\beta_{2m+1} - {}_{2k}\beta_{2m-1}, \qquad m = 1, 2, \cdots, k-1, \tag{1.5.2}$$

constitute a set of $k - 1$ interquantile differences, which would reveal more and more about the distribution of b as k is increased. But it seems reasonable and is also easy to check that from s blocks we cannot estimate these $k - 1$ differences unless $k \leqq s$.

1.6. *Bounding loci in the space of $2k$-tile differences.* The quasi-confidence statement (1.3.5) contained a quadratic form in the components of $\mathbf{b}$, a quadratic form whose matrix was expressible as (1.4.2) or (1.4.6) and then for a large class of useful designs was further simplified to (1.4.8). Temporarily ignoring the scalar design constant, δ^2, we shall now investigate the quadratic form

$$[{}_{2k}\mathbf{b}]'[\mathbf{I}(s) - s^{-1}\mathbf{J}(s)][{}_{2k}\mathbf{b}]. \tag{1.6.1}$$

In this expression the s (unknown) values of the original variate b have been replaced by the k (unknown) values of the substitute variate ${}_{2k}b$. For $s > k$, some of these discrete values must occur more than once. Hence we denote by s_{2n-1} the frequency of the value ${}_{2k}\beta_{2n-1}$, for $n = 1, 2, \cdots, k$, among the s components of ${}_{2k}b$. Of course $\sum_{n=1}^{k} s_{2n-1} = s$. Using (A.10), where y_i is replaced by ${}_{2k}\beta_{2n-1}$ and x_i by s_{2n-1} with summation running from $n = 1$ to $n = k$, and using (A.11), where $z_n = {}_{2k}\beta_{2n+1} - {}_{2k}\beta_{2n-1}$, we can express (1.6.1) as the quadratic form

$$[{}_{2k}\mathbf{d}]'\mathbf{G}[{}_{2k}\mathbf{d}], \tag{1.6.2}$$

where the vector ${}_{2k}\mathbf{d}$ has its $k - 1$ components defined by

$${}_{2k}d_n \equiv {}_{2k}\beta_{2n+1} - {}_{2k}\beta_{2n-1} \qquad \text{for } n = 1, 2, \cdots, k-1, \tag{1.6.3}$$

and where the symmetric matrix $\mathbf{G}$ has in its ith row and jth column for $i \geqq j$,

$$(1.6.4) \qquad G_{ij} \equiv s^{-1}\left(\sum_{n=1}^{j} s_{2n-1}\right)\left(\sum_{n=i+1}^{k} s_{2n-1}\right).$$

In terms of the original model, the vector ${}_{2k}\mathbf{d}$ consists of parameter-like unknown constants, viz., interquantile differences of the distribution from which the components of $\mathbf{b}$ are a sample. The set of integers $s_1, s_3, \cdots, s_{2k-1}$, hereafter denoted by $\mathbf{s}$, represents the hypothetical frequencies of the distinct values of the substitute variate ${}_{2k}b$ corresponding to the s (different) values of the original variate b. But for the purpose of obtaining confidence bounds on the components of ${}_{2k}\mathbf{d}$, we shall think of the components of $\mathbf{s}$ as given (constants) and of the components of ${}_{2k}\mathbf{d}$ as (variable) coordinates of a point in a $(k - 1)$-dimensional parameter space. When (1.6.1) is thus regarded as a function of the $k - 1$ components of ${}_{2k}\mathbf{d}$, a function whose coefficients are determined by a given partition of s into $\mathbf{s}(k)$, (1.6.1) will be denoted by $\Delta_s(\mathbf{d})$. Under this interpretation

$$(1.6.5) \qquad \Delta_s(\mathbf{d}) = \text{constant}$$

determines a locus in the $(k - 1)$-dimensional space whose coordinates are the components of ${}_{2k}\mathbf{d}$. For a given constant term and all possible partitions of a given s, (1.6.5) would determine a discrete family of such loci. The several loci in the family may be classified as bounding or not according to whether they would inclose a bounded region of points having non-negative coordinates. (In all subsequent use of "bounded region" it will be understood to exclude points with any negative coordinate. Thus the bounded region has for its boundary not only the locus of (1.6.5) but also, for $k = 2$, the origin; for $k = 3$, the $d_1 = 0$ and $d_2 = 0$ axes; for $k = 4$, the $d_1 = 0$, $d_2 = 0$, $d_3 = 0$ planes; for $k = 5$, the $d_1 = 0$, $d_2 = 0$, $d_3 = 0$, $d_4 = 0$ hyperplanes.) For a given k, a given s, and given positive constant, there is a locus, denoted by $\Delta_2(\mathbf{d}) =$ constant, such that it incloses a region which is the union of the regions inclosed by all the bounding loci in the family determined by (1.6.5). This locus is called the outer boundary. Similarly there is an inner boundary, denoted by $\Delta_1(\mathbf{d}) =$ constant, which incloses a region which is the intersection of all regions inclosed by the bounding loci. For some values of k, the outer boundary and the inner boundary are themselves members of the family determined by (1.6.5), whereas for other values of k, these two boundaries are composites of more than one bounding locus.

From (1.6.4) and (A.12) it follows that if $s_{2n-1} \neq 0$ for all $n = 1, 2, \cdots, k$, then (1.6.2) is positive definite, and the locus of (1.6.5) would be two points if $k = 2$, an ellipse if $k = 3$, and ellipsoid if $k = 4$, etc.,—all bounding loci. But if $s_1 = 0$ or $s_{2k-1} = 0$, the coefficient of ${}_{2k}\beta_1$ or ${}_{2k}\beta_{2k-1}$ is zero in (1.6.1), and hence, by (A.10), ${}_{2k}d_1$ or ${}_{2k}d_{k-1}$ vanishes from the quadratic form (1.6.2). When such is the case, the locus of (1.6.5) provides no bound on the corresponding coordinate.

On the other hand, if $s_{2m-1} = 0$ for $1 \neq m \neq k$, then ${}_{2k}\beta_{2m-1}$ vanishes from the quadratic form (1.6.1), and it might seem that both d_{m-1} and d_m would vanish from the quadratic form (1.6.2), so that neither d_{m-1} nor d_m would be bounded by the locus of (1.6.5). However, if $s_{2m+1} \neq 0$ and $s_{2m-3} \neq 0$, we know by (A.10) that (1.6.1) can be expressed in terms of

$$({}_{2k}\beta_3 - {}_{2k}\beta_1),\ ({}_{2k}\beta_5 - {}_{2k}\beta_3),\ \cdots\ ({}_{2k}\beta_{2m+1} - {}_{2k}\beta_{2m-3}),\ \cdots\ ({}_{2k}\beta_{2k-1} - {}_{2k}\beta_{2k-3}),$$

and by (1.6.3) these are

$${}_{2k}d_1\,,\ {}_{2k}d_2\,,\ \cdots,\ {}_{2k}d_{m-1} + {}_{2k}d_m\,,\ \cdots,\ {}_{2k}d_{k-1}\,.$$

Note that here the sum ${}_{2k}d_{m-1} + {}_{2k}d_m$, rather than ${}_{2k}d_{m-1}$ and ${}_{2k}d_m$ separately, appears in (1.6.1). Hence the locus of (1.6.5) is "flattened" if $s_{2m-1} = 0$ but is still bounding. By an obvious extension of this argument, it follows that the locus of (1.6.5) will be bounding when $s_{2m-1} = 0$ for more than one value of m, provided only that $s_1 \neq 0$ and $s_{2k-1} \neq 0$. This is seen to be plausible even on the rough consideration that a bound on $({}_{2k}\beta_{k-1} - {}_{2k}\beta_1)$ necessarily imposes bounds on all intermediate $2k$-tile differences, and basically this is what is formalized in the above argument.

From the set of bounding loci for each value of k, an inner and outer boundary must be found. As defined above, "inner" and "outer" are designations applied to loci by virtue of extreme properties of the matrix of coefficients regardless of the constant term in the equation. In Appendix B the matrices of the quadratic forms $\Delta_1(\mathbf{d})$ and $\Delta_2(\mathbf{d})$ are derived from the matrix $\mathbf{G}$. For any k,

$$\Delta_1(\mathbf{d}) = \begin{cases} \dfrac{s^2-1}{4s}\left(\sum {}_{2k}d_n\right)^2 & \text{for } s \text{ odd}, \\[2ex] \dfrac{s}{4}\left(\sum {}_{2k}d_n\right)^2 & \text{for } s \text{ even}; \end{cases} \tag{1.6.6}$$

$$\Delta_2(\mathbf{d}) = \frac{s-1}{s}\sum ({}_{2k}d_n)^2 + \frac{1}{s}\sum_{n\neq m}\sum ({}_{2k}d_n)({}_{2k}d_n)\,. \tag{1.6.7}$$

Thus the inner boundary is a point for $k = 2$, a straight line for $k = 3$, a plane for $k = 4$, and a hyperplane for $k = 5$. The outer boundary is a point for $k = 2$, an ellipse for $k = 3$, an ellipsoid for $k = 4$, and a hyperellipsoid for $k = 5$.

Since, for $2 \leqq k \leqq s$, the k non-negative integers into which s is partitioned are hypothetical frequencies of the k equiprobable values ${}_{2k}\beta_{2n-1}$, each partition of s has an *a priori* probability

$$\frac{s!}{k^s \prod_{n=1}^{k} (s_{2n-1})!}\,. \tag{1.6.8}$$

But all that is needed now is the *a priori* probability, say γ, that a locus obtained from (1.6.5) be bounding. Since $s_1 \neq 0$ and $s_{2k-1} \neq 0$ are necessary and sufficient conditions for bounding loci,

$$\begin{aligned} \gamma &= 1 - \Pr\{s_1 = 0\} - \Pr\{s_{2k-1} = 0\} + \Pr\{s_1 = s_{2k-1} = 0\} \\ &= \begin{cases} 1 - 2^{-s} - 2^{-s} + 0 = 1 - 2^{1-s} & \text{for } k = 2, \\ 1 - 2[(k-1)/k]^s + [(k-2)/k]^s & \text{for } k > 2. \end{cases} \end{aligned} \tag{1.6.9}$$

This γ is thus the *a priori* probability that a point whose Cartesian coordinates are the components of $\mathbf{d}$ should lie on any one of the bounding loci. Hence the probability is not less than γ that such a point should lie between the inner and outer boundaries or on one of these.

1.7. *The final confidence statement.* The quasi-confidence statement (1.3.5) asserts that with probability $\geqq 1 - \alpha_1$ two computable quantities, l_1^2 and l_2^2, are bounds for a certain quadratic form in the unobservable random effects $\mathbf{b}$, which form can be reduced, for many designs, to $\delta^2\mathbf{b}'[\mathbf{I} - s^{-1}\mathbf{J}]\mathbf{b}$. In Section 1.6 it has been shown that with the same degree of approximation which results from grouping continuous data into k classes, the continuous variate b may be replaced by the k-valued substitute variate ${}_{2k}b$. Thus, the quasi-confidence bounds would apply to $\delta^2[{}_{2k}\mathbf{b}]'[\mathbf{I} - s^{-1}\mathbf{J}][{}_{2k}\mathbf{b}]$ or the equivalent form, $\delta^2[{}_{2k}\mathbf{d}]'\mathbf{G}[{}_{2k}\mathbf{d}]$. With *a priori* probability $\geqq\gamma$ given by (1.6.9), the points whose coordinates, ${}_{2k}\mathbf{d}$, satisfy the equation $[{}_{2k}\mathbf{d}]'\mathbf{G}[{}_{2k}\mathbf{d}] = c$ lie in a region bounded by $\Delta_1(\mathbf{d}) = c$ and $\Delta_2(\mathbf{d}) = c$. Because of the postulated independence of $\mathbf{b}$ and $\boldsymbol{\varepsilon}$, we can combine this *a priori* probability statement about $\Delta_1(\mathbf{d})$ and $\Delta_2(\mathbf{d})$ with the quasi-confidence statement (1.3.5) to state with confidence coefficient $\geqq (1 - \alpha_1)\gamma$, that the $k - 1$ differences between successive odd $2k$-tiles of the distribution of b are coordinates of a point lying in the region bounded by $\delta^2\Delta_1(\mathbf{d}) = l_1^2$ and $\delta^2\Delta_2(\mathbf{d}) = l_2^2$.

To make these final confidence bounds more explicit, let us consider in detail $k = 2$ and $k = 3$. For $k = 2$, we refer to (1.6.6) and (1.6.7), supposing s even. Combining the two equations giving inner and outer boundaries, we obtain a confidence interval for the interquartile range:

$$\delta^{-1}(4/s)^{\frac{1}{2}}l_1 \leqq {}_4\beta_3 - {}_4\beta_1 \leqq \delta^{-1}[s/(s-1)]^{\frac{1}{2}}l_2 . \tag{1.7.1}$$

For $k = 3$, we refer to (1.6.6) and (1.6.7), again supposing s to be even. This time we obtain simultaneous confidence bounds on two interquantile ranges, ${}_6\beta_3 - {}_6\beta_1$ and ${}_6\beta_5 - {}_6\beta_3$. If these two parameters are represented respectively by the abscissa and ordinate of a point, then the confidence region is given by

$$\begin{aligned} (s-1)x^2 + 2xy + (s-1)y^2 &\leqq \delta^{-2}sl_2^2 , \\ x + y &\geqq \delta^{-1}(4/s)^{\frac{1}{2}}l_1 , \\ x &\geqq 0, \\ y &\geqq 0. \end{aligned} \tag{1.7.2}$$

2. The General Uniresponse Mixed-Model Problem with Two or More Factors.

2.1. *Fixed effects in a two-factor mixed model.* The previous sections have been almost exclusively concerned with the random effects in a uniresponse model with one factor represented by fixed effects and one factor by random effects; but the fixed effects may be of greater interest to the experimenter. With this in mind, we pointed out that the statistic $\mathbf{u}$ defined in (1.2.1), is entirely independent of the random effects **b** (though not of the statistic **v**) and is quite simply related to the fixed effects **a**. Suppose now it is desired to estimate a particular contrast among the fixed effects, say $\mathbf{c}'\mathbf{a}$ where $\mathbf{c}'\mathbf{j} = 0$. Then

$$\mathbf{c}'\mathbf{H}'\mathbf{T}_0\mathbf{u} \tag{2.1.1}$$

is an unbiased estimator of $\mathbf{c}'\mathbf{a}$. Just as $\mathbf{T}_1$ did not need to be determined explicitly in order to obtain $\mathbf{v}'\mathbf{v}$, so here (2.1.1) can be computed without finding $\mathbf{T}_0$ explicitly. Now regardless of whether **b** (hence also **y**) is normally distributed, **u** is normal and hence (2.1.1) is also. The variance of (2.1.1) is the same whether **b** consists of fixed effects or random effects. Thus if t_α is the upper $\frac{1}{2}\alpha$ point of the t-distribution with $n - r^*$ d.f., we can assert with confidence coefficient $1 - \alpha$ that

$$\begin{aligned}\mathbf{c}'\mathbf{H}'\mathbf{T}_0\mathbf{u} - t_\alpha[\mathbf{c}'\mathbf{R}\mathbf{c}\mathbf{w}'\mathbf{w}/(n - r^*)]^{\frac{1}{2}} \leqq \mathbf{c}'\mathbf{a} \leqq \mathbf{c}'\mathbf{H}'\mathbf{T}_0\mathbf{u} \\ + t_\alpha[\mathbf{c}'\mathbf{R}\mathbf{c}\mathbf{w}'\mathbf{w}/(n - r^*)]^{\frac{1}{2}},\end{aligned} \tag{2.1.2}$$

where

$$\mathbf{R} \equiv [\mathbf{M}_0'\mathbf{M}_0 - \mathbf{M}_0'\mathbf{M}_1\mathbf{K}(\mathbf{K}'\mathbf{M}_1'\mathbf{M}_1\mathbf{K})^{-1}\mathbf{K}'\mathbf{M}_1'\mathbf{M}_0]^{-1}. \tag{2.1.3}$$

For neither the point estimate (2.1.1) nor the confidence interval (2.1.2) is any further restriction on design necessary. But for some designs, especially randomized block, **R** will be somewhat simpler than (2.1.3).

Suppose now that the hypothesis of equality of all fixed effects **a** is to be tested. Then the test statistic, $[\mathbf{u}'\mathbf{u}/(t - 1)]/[\mathbf{w}'\mathbf{w}/(n - r^*)]$, has the central F distribution if and only if this hypothesis is true, quite irrespective of the distribution of **b**. On the other hand, it may be desired to test a more general hypothesis, $\mathbf{C}\mathbf{a} = \mathbf{o}$, where $\mathbf{C}(g \times t)$ is of rank $g \leqq t - 1$. The hypothesis is untestable if the rank of $\begin{bmatrix}\mathbf{M}', \mathbf{C}' \\ \mathbf{O}\end{bmatrix}$ is $g + n - r^*$. The hypothesis is said to be completely testable or weakly testable according as the rank of $\begin{bmatrix}\mathbf{M}', \mathbf{C}' \\ \mathbf{O}\end{bmatrix}$ is equal to or greater than $n - r^*$ (and less than $g + n - r^*$). In either testable case the test statistic, $(\mathbf{u}'\mathbf{T}_0'\mathbf{H}\mathbf{C}'[\mathbf{C}\mathbf{R}\mathbf{C}']^{-1}\mathbf{C}\mathbf{H}'\mathbf{T}_0\mathbf{u}/g)/(\mathbf{w}'\mathbf{w}/(n - r^*))$, has the central F distribution, with g and $n - r^*$ d.f., if and only if $\mathbf{C}\mathbf{a} = \mathbf{o}$. But if the hypothesis is weakly testable, the test will have the same power for a possibly weaker hypothesis as for the hypothesis being tested.

2.2. *A mixed model with more than two factors.* From the two-factor mixed model of Section 1.1, it is an easy extension to a multifactor mixed model in which the observed response of each experimental unit is the sum of one (fixed) treatment effect, m (random) block effects, and a normally distributed error. Instead of a random sample of size s from a single distribution of block effects, we now postulate m distributions of block effects, mutually independent and independent of the normal error. The formal relation of the observations **y** to these unobservables of the model may still be expressed as in (1.1.1), but now the structure matrix has $m + 1$ submatrices instead of just two, each consisting of all zero elements except for a single unity in each row. Thus

$$\mathbf{M}(n \times (s + t)) = [\mathbf{M}_0(n \times t), \mathbf{M}_1(n \times g_1), \mathbf{M}_2(n \times g_2), \cdots, \mathbf{M}_m(n \times g_m)]$$

and

$$\mathbf{b}'(s) = [\mathbf{b}_1'(g_1), \mathbf{b}_2'(g_2), \cdots, \mathbf{b}_m'(g_m)],$$

where $\mathbf{b}_i$ is a sample of size g_i from the ith distribution of block effects.

Although this is a stronger restriction than is necessary, when $m > 1$ we have restricted our investigation to orthogonal designs of two well-known types—the complete factorial type in which $n = t \cdot g_1 \cdot g_2 \cdots g_m$ and the Latin Square type in which $t = g_1 = g_2 = \cdots = g_m$ and $n = t^2$. For both types the structure matrix will have a basis $\mathbf{M}_I(n \times (s + t - m))$ such that $\mathbf{M}_I'\mathbf{M}_I$ and $[\mathbf{M}_I'\mathbf{M}_I]^{-1}$ are formally expressible in terms of $r, t, g_1, \cdots, g_m$. Then premultiplying $[\mathbf{M}_I'\mathbf{M}_I]^{-1}\mathbf{M}_1'\mathbf{y}$ by

$$\left[r^{\frac{1}{2}}\mathbf{H}((t-1) \times t) \dotplus \left(\frac{rt}{g_1}\right)^{\frac{1}{2}} \mathbf{H}((g_1 - 1) \times g_1)\mathbf{K}(g_1 \times (g_1 - 1)) \dotplus \cdots \dotplus \left(\frac{rt}{g_m}\right)^{\frac{1}{2}} \mathbf{H}((g_m - 1) \times g_m)\,\mathbf{K}(g_m \times (g_m - 1))\right]$$

effects an orthonormal transformation on the observations $\mathbf{y}$ and defines $m + 1$ vectors of statistics, which vectors are mutually independent, each vector involving the effects of only one of the $m + 1$ factors and, of course, the normal error. Thus analogous to (1.2.9) we now have

$$\begin{aligned}
\mathbf{u}(t-1) &= r^{\frac{1}{2}}\mathbf{H}((t-1) \times t)\mathbf{a}(t) + \mathbf{L}_0\boldsymbol{\varepsilon}, \\
\mathbf{v}_1(g_1 - 1) &= \delta_1\mathbf{H}((g_1 - 1) \times g_1)\mathbf{b}_1(g_1) + \mathbf{L}_1\boldsymbol{\varepsilon}, \\
&\vdots \\
\mathbf{v}_m(g_m - 1) &= \delta_m\mathbf{H}((g_m - 1) \times g_m)\mathbf{b}_m(g_m) + \mathbf{L}_m\boldsymbol{\varepsilon},
\end{aligned}$$

where $\delta_i^2 = rt/g_i$ and $\mathbf{L}_i((g_i - 1) \times n)$ is orthonormal. There is also a statistic $\mathbf{w}(n - r^*)$, relevant to error only, defined as in (1.2.7) except that r^*, the rank of the structure matrix, is now $s + t - m$. Each $\mathbf{v}_i$ can be used, just as $\mathbf{v}$ was used in Section 1.3 and Section 1.7, with $\mathbf{w}$ to obtain quasi-confidence bounds and then final confidence bounds on the $2k$-tile differences of each random-effects factor. Analogous extension to a model with more than one fixed-effects factor is also possible.

APPENDIX A

Miscellaneous Lemmas

Proofs will not be given where direct verification is straightforward.

(A.1) $\mathbf{H}((m-1) \times m)\mathbf{j}(m) = \mathbf{o}(m-1)$.

(A.2) $\mathbf{H}((m-1) \times m)\mathbf{H}'(m \times (m-1)) = \mathbf{I}(m-1)$.

(A.3) $\mathbf{H}'(m \times (m-1))\mathbf{H}((m-1) \times m) = \mathbf{I}(m) - m^{-1}\mathbf{J}(m)$.

(A.4) Postmultiplying $\mathbf{A}(m \times n)$ by $\mathbf{K}(n \times (n-1))$ removes the nth column of $\mathbf{A}$.
Premultiplying $\mathbf{A}(m \times n)$ by $\mathbf{K}'((m-1) \times m)$ removes the mth row of $\mathbf{A}$.

(A.5) Postmultiplying $\mathbf{A}(m \times n)$ by $\mathbf{K}'(n \times (n+1))$ adjoins $\mathbf{o}(m)$ as an $(n+1)$th column.
Premultiplying $\mathbf{A}(m \times n)$ by $\mathbf{K}((m+1) \times m)$ adjoins $\mathbf{o}'(n)$ as an $(m+1)$th row.

(A.6) Any matrix $\mathbf{A}$ can be partitioned into $[\mathbf{AK}, (\mathbf{A} - \mathbf{AKK}')\mathbf{j}]$.

(A.7) If $\mathbf{Aj} = \mathbf{o}$, $\mathbf{A} = [\mathbf{AK}, -\mathbf{AKj}]$.

(A.8) If $\mathbf{Cj} = \mathbf{o}$ where $\mathbf{C}(m)$ is symmetric and of rank $m - 1$, then $\mathbf{C} = \mathbf{H}'\{\mathbf{HK}[\mathbf{K}'\mathbf{CK}]^{-1}\mathbf{K}'\mathbf{H}'\}^{-1}\mathbf{H} = \mathbf{H}'(\mathbf{K}'\mathbf{H}')^{-1}\mathbf{K}'\mathbf{CK}(\mathbf{HK})^{-1}\mathbf{H}$.

(A.9) If $\mathbf{A}(m)$ and $\mathbf{D}(n)$ are both nonsingular,

$$\begin{bmatrix} \mathbf{A}, \mathbf{B} \\ \mathbf{C}, \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1} & (\mathbf{BD}^{-1}\mathbf{C} - \mathbf{A})^{-1}\mathbf{BD}^{-1} \\ (\mathbf{CA}^{-1}\mathbf{B} - \mathbf{D})^{-1}\mathbf{CA}^{-1} & (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1} \end{bmatrix}.$$

This is stated as an exercise in [1]. It may be derived by solving four simultaneous matrix equations or simply verified by postmultiplying the two members by $\begin{bmatrix} \mathbf{A}, \mathbf{B} \\ \mathbf{C}, \mathbf{D} \end{bmatrix}$ and then factoring each of the four combinations so obtained.

(A.10) $$\sum_{i=1}^{m} x_i y_i^2 - \left(\sum_{i=1}^{m} x_i y_i\right)^2 \Big/ \sum_{i=1}^{m} x_i = \sum_{i=2}^{m}\sum_{j=1}^{i-1} x_i x_j (y_i - y_j)^2 \Big/ \sum_{i=1}^{m} x_i .$$

Proof: For $\sum_{i=1}^{m} x_i \neq 0$, the above is equivalent to

$$\left(\sum_{i=1}^{m} x_i\right)\left(\sum_{i=1}^{m} x_i y_i^2\right) = \left(\sum_{i=1}^{m} x_i y_i\right)^2 + \sum_{i=2}^{m}\sum_{j=1}^{i-1} x_i x_j (y_i - y_j)^2,$$

whose right member can be rearranged as follows

$$\begin{aligned}
&\left(\sum_{i=1}^{m} x_i y_i\right)^2 + \sum_{i=2}^{m}\sum_{j=1}^{i-1} x_i x_j (y_i^2 - 2y_i y_j + y_j^2) \\
&= \left(\sum_{i=1}^{m} x_i y_i\right)^2 + \sum_{i=1}^{m}\sum_{j=1}^{m} (1 - \delta_{ij}) x_i x_j y_j^2 - \sum_{i=1}^{m}\sum_{j=1}^{m} (1 - \delta_{ij}) x_i x_j y_i y_j \\
&= \left(\sum_{i=1}^{m} x_i y_i\right)^2 + \sum_{i=1}^{m} x_i \sum_{j=1}^{m} x_j y_j^2 - \sum_{i=1}^{m} x_i^2 y_i^2 - \sum_{i=1}^{m} x_i y_i \sum_{j=1}^{m} x_j y_j + \sum_{i=1}^{m} x_i^2 y_i^2 \\
&= \sum_{i=1}^{m} x_i \sum_{j=1}^{m} x_j y_j^2 ,
\end{aligned}$$

which is obviously the same as the left member. (The Kronecker delta $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.)

(A.11) $\sum_{i=2}^{m} \sum_{j=1}^{i-1} x_i x_j (\sum_{n=j}^{i-1} z_n)^2$ is a quadratic form in $z_1, \cdots, z_{m-1}$ whose (symmetric) matrix has, for $p \geqq q$, $(\sum_{i=1}^{q} x_i)(\sum_{j=p+1}^{m} x_j)$ as the element in the pth row and qth column.

PROOF: The $\binom{m}{2}$ terms of the given double sum can be arranged in a triangular array, and for a given n, $1 \leqq n \leqq m - 1$, z_n^2 will occur in a rectangular subarray of the first n columns (say) and nth through $(m - 1)$th rows. The sum of these $n(m - n)$ terms is

$$\sum_{i=n}^{m-1} x_{i+1} \sum_{j=1}^{n} x_j z_n^2 = \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{j=n+1}^{m} x_j\right) z_n^2 .$$

Also from the same triangular array it is apparent that for a given n and l, $1 \leqq n \leqq l \leqq m - 1$, the factor $2z_n z_l$ occurs in $n(m - l)$ terms in the first n columns and the lth through $(m - 1)$th rows. The sum of these terms is

$$2\sum_{i=l}^{m-1} x_{i+1} \sum_{j=1}^{n} x_j z_n z_l = \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{j=l+1}^{m} x_j\right) 2z_n z_l .$$

These same two sums would be obtained from the diagonal and the off-diagonal elements respectively of the quadratic form whose matrix is defined above. (A.12) If $\sum_{k=1}^{m} x_k > 0$ and no $x_k < 0$, the real symmetric $(m - 1)$th order matrix whose element in the pth row and qth column, $p \geqq q$, is $(\sum_{k=1}^{q} x_k) \times (\sum_{k=p+1}^{m} x_k)$ is at least positive semidefinite with its vacuity equal to the number of values of k for which $x_k = 0$.

PROOF: $x_k \neq 0$ for at least one k. Without loss of generality suppose $x_k \neq 0$ for $k = 1, \cdots, n$ and $x_k = 0$ for $k = n + 1, \cdots, m$. Adding $(\sum_{k=j+1}^{m} x_k)/x_1$ times the first column to the jth column makes $\sum_{k=1}^{m} x_k$ a common factor of all elements in the jth column for $j = 2, 3, \cdots, m - 1$. Then subtracting the jth row from the $(j - 1)$th row for $j = 2, 3, \cdots, m - 1$ leaves only zeros above the principal diagonal. The elements in this diagonal are x_1x_2 in the first row and $(\sum_{k=1}^{m} x_k)x_{j+1}$ in the jth row for $j = 2, 3, \cdots, m - 1$. Considering the sequence of lower triangular submatrices consisting of the first k rows, for $k = 2, \cdots, n - 1$, each is nonsingular with determinant equal to

$$\left(\sum_{i=1}^{m} x_i\right)^{k-1} \prod_{j=1}^{k+1} x_j .$$

For $k \geqq n$, each is singular. It follows from this and Gundelfinger's rule [2] that the original matrix has $n - 1$ positive characteristic roots and $m - n$ zeros.

APPENDIX B

Inner and Outer Boundaries

Either to pick out the outer boundary from the family of loci given by (1.6.5) or, failing that, to construct an outer boundary for the family, we consider the matrix $\mathbf{G}$ defined by (1.6.4) with the restriction that $s_1 \neq 0$ and $s_{2k-1} \neq 0$. If g_{ij} denotes the element in the matrix of the outer boundary, corresponding to G_{ij} in (1.6.4), then, since ${}_{2k}d_{2n-1} \geqq 0$ for all n, it is sufficient that

(B.1) $$g_{ij} \leqq G_{ij} \quad \text{for all } i, j = 1, \cdots, k - 1.$$

From (B.1) and (1.6.4) it follows that

$$\text{(B.2)} \qquad g_{ij} = \begin{cases} (s-1)/s & \text{for } i = j, \\ 1/s & \text{for } i \neq j. \end{cases}$$

For $k = 2, 3, 5$ the outer boundary thus determined belongs to the family (1.6.5); for $k = 4$, it does not.

On the other hand, to pick out the inner boundary from the family of loci, we want the largest possible G_{ij} for all i, j. Inspection of (1.6.4) leads to the conclusion that for the inner boundary

$$\text{(B.3)} \qquad G_{ij} = \begin{cases} (s^2-1)/4s & \text{for } s \text{ odd}, \\ s/4 & \text{for } s \text{ even}, \end{cases}$$

for all i, j. Thus it appears that for all values of k, the inner boundary is that member of the family of (1.6.5) which is completely flat, i.e., with only one nonvanishing characteristic root.

REFERENCES

[1] Aitken, A. C., *Determinants and Matrices*, 8th ed., Oliver and Boyd, Edinburgh, 1954.

[2] Browne, E. T., *Introduction to the Theory of Determinants and Matrices*, University of North Carolina Press, Chapel Hill, c. 1958.

[3] Ghosh, M. N., "Simultaneous tests of linear hypotheses," *Biometrika*, Vol. 42 (1955), pp. 441–449.

[4] Heck, D. L., "Some Uses of the Distribution of the Largest Root in Multivariate Analysis." Institute of Statistics, University of North Carolina, Mimeograph Series No. 194, 1958.

[5] MacDuffee, C. C., *The Theory of Matrices*. J. Springer, Berlin, 1933.

[6] Ramachandran, K. V., "On the simultaneous analysis of variance test," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 521–528.

[7] Ramachandran, K. V., "Contribution to simultaneous confidence interval estimation," *Biometrics*, Vol. 12 (1956), pp. 51–56.

[8] Roy, S. N. and Gnanadesikan, R., "Further contributions to multivariate confidence bounds," *Biometrika*, Vol. 44 (1957), pp. 399–410.

[9] Roy, S. N. and Gnanadesikan, R., "Some contributions to ANOVA in one or more dimensions: I," *Ann. Math. Stat.*, Vol. 30 (1959), pp. 304–317.

[10] Roy, S. N. and Gnanadesikan, R., "Some contributions to ANOVA in one or more dimensions: II," *Ann. Math. Stat.*, Vol. 30 (1959), pp. 318–340.

[11] Roy, S. N. and Sarhan, A. E., "On inverting a class of patterned matrices," *Biometrika*, Vol. 43 (1956), pp. 227–231.

[12] Roy, S. N., *Some Aspects of Multivariate Analysis*, John Wiley and Sons, New York, 1957.

[13] Thompson, W. A. Jr., "On the ratio of variances in the mixed incomplete block model," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 721–733.

[14] Thompson, W. A. Jr., "The ratio of variances in a variance components model," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 325–329.

[15] Wilks, S. S., *Mathematical Statistics*, Princeton University Press, Princeton, 1943.

[16] Youden, W. J., "Linked blocks: a new class of incomplete block designs" (abstract), *Biometrics*, Vol. 7 (1951), p. 124.