SOME CONTRIBUTIONS TO ANOVA IN ONE OR MORE DIMENSIONS: I

By S. N. Roy and R. Gnanadesikan[1]

University of North Carolina

0. Introduction and Summary. Two models are considered in detail which are the Models I and II of ANOVA in the terminology of Eisenhart [2]. The present paper, which deals with the one dimensional or univariate case, and its sequel, which will deal with the multidimensional or multivariate case, seek to give a unified general treatment, using matrix methods, of certain problems under the two models of ANOVA. Section 1 of each paper, which deals with Model I, is of the nature of a résumé giving the main results of a general treatment discussed elsewhere [10, 11, 12] by one of the authors. Section 2 of each paper, which deals with Model II or variance components model, is self-contained, and presents a natural tie-up between the analyses under the two models for a k-way classification. Results in estimation, testing of hypotheses and confidence bounds are presented, although the main emphasis is on the results in confidence bounds (simultaneous and/or separate) on meaningful parametric functions which are physically natural and mathematically convenient measures of departure from customary null hypotheses.

It will be seen that a mixed model, which would include both Models I and II as special cases, can be defined, and the associated problems can be studied by using methods which are, essentially, a combination of the methods given for the separate models in Sections 1 and 2, respectively, of this paper. Since nothing essentially new is involved in such a study, this paper does not explicitly discuss it.

Unless otherwise stated, capital letters will denote matrices and small letters in boldface will denote column vectors. Such letters when primed denote transposes. For instance, $A(p \times q)$ denotes a matrix with p rows and q columns, $A'(q \times p)$ denotes the transpose of A, $\mathbf{a}(p \times 1)$ denotes a column vector with p elements and $\mathbf{a}'(1 \times p)$, the transpose of $\mathbf{a}$, is a row vector. In particular, $I(p)$ will denote the identity matrix of order p and $\mathbf{0}(p \times 1)$ and $0(p \times q)$ will stand, respectively, for the null vector of order p and the null matrix with p rows and q columns.

1. Résumé of problems and results under the univariate Model I of ANOVA.

1.1 *The Model I.* Let $\mathbf{x}'(1 \times n) = (x_1, x_2, \cdots, x_n)$ be a set of n observable stochastic variates such that

$$(1.1.1) \qquad \mathbf{x}(n \times 1) = A(n \times m)\boldsymbol{\xi}(m \times 1) + \boldsymbol{\varepsilon}(n \times 1), \qquad m < n$$

Received July 24, 1957.

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

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

where $A(n \times m)$, to be called the *design* matrix, is a matrix whose elements are constants given by the design of the experiment and is of rank $r \leqq m < n$ and where,

(i) $\boldsymbol{\xi}(m \times 1)$ is a set of unknown parameters;

(ii) $\boldsymbol{\varepsilon}(n \times 1)$, whose elements are physically of the nature of errors, is a random sample from the normal population $N(0, \sigma^2)$.

Under this model it is easily seen that $\mathbf{x}(n \times 1)$ is a set of n normal independent (and hence uncorrelated) variates with a common variance σ^2 and the respective means given by,

$$E(\mathbf{x}) = A(n \times m)\boldsymbol{\xi}(m \times 1). \tag{1.1.2}$$

It may be noted that the assumption of normality in (ii) above is not necessary for problems of linear estimation, and the results presented below on linear estimation are all valid merely under the assumption that $\boldsymbol{\varepsilon}(n \times 1)$ (hence $\mathbf{x}(n \times 1)$) is a set of uncorrelated stochastic variates with a common variance σ^2. Next, since A is of rank $r \leqq m < n$, we can find a basis $A_I(n \times r)$ of A, which, without any loss of generality and by renumbering the columns of A and the elements of $\boldsymbol{\xi}$, can be taken to be the first r columns of A and we may write (1.1.2) as,

$$E(\mathbf{x}) = n\underset{\quad r \qquad m-r}{[A_I \quad A_D\ ]}\underset{1}{\begin{bmatrix}\boldsymbol{\xi}_I \\ \boldsymbol{\xi}_D\end{bmatrix}}\begin{matrix} r. \\ m-r\end{matrix} \tag{1.1.3}$$

1.2 *Linear estimation.* We seek an unbiased minimum variance linear estimate $\mathbf{b}'(1 \times n)\mathbf{x}(n \times 1)$ of a given linear function $\mathbf{c}'(1 \times m)\boldsymbol{\xi}(m \times 1)$ of the unknown parameters $\boldsymbol{\xi}(m \times 1)$. The partitioning of A into A_I and A_D determines the partitioning of $\boldsymbol{\xi}$ into $\boldsymbol{\xi}_I$ and $\boldsymbol{\xi}_D$ and the partitioning into $\boldsymbol{\xi}_I$ and $\boldsymbol{\xi}_D$ determines that of $\mathbf{c}'$ into $\mathbf{c}_I'$ and $\mathbf{c}_D'$, so that we can rewrite $\mathbf{c}'\boldsymbol{\xi}$ as $\mathbf{c}_I'\boldsymbol{\xi}_I + \mathbf{c}_D'\boldsymbol{\xi}_D$. The main results in linear estimation follow. [11, pp. 77–81]

(1.2.1) All the following results are invariant under the choice of a basis A_I of A (with a consequent determination of $\boldsymbol{\xi}_I$ and $\mathbf{c}_I$).

The necessary and sufficient condition on $\mathbf{c}$ that an unbiased linear estimate $\mathbf{a}'\mathbf{x}$ of $\mathbf{c}'\boldsymbol{\xi}$ exists (in which case $\mathbf{c}'\boldsymbol{\xi}$ will be said to be *estimable* and the corresponding condition will be said to be the *estimability* condition) is that,

$$\mathbf{c}_D' = \mathbf{c}_I'(A_I'A_I)^{-1}A_I'A_D, \tag{1.2.2}$$

or, in other words, that $\mathbf{c}_D'$ should be related to $\mathbf{c}_I'$ through the same matrix postfactor through which A_D is related to A_I.

Another way to express (1.2.2) would be to say that

$$\text{Rank}\begin{bmatrix} A \\ \mathbf{c}' \end{bmatrix} = \text{Rank}\ [A], \tag{1.2.21}$$

which means that (1.2.2) is a convenient mathematical test for (1.2.21)

The unbiased minimum variance linear estimate of an *estimable* $\mathbf{c}'\boldsymbol{\xi}$ is given by

$$\mathbf{c}_I'(A_I'A_I)^{-1}A_I'\mathbf{x}. \tag{1.2.3}$$

The variance of this linear estimate is given by

$$\mathbf{c}_I'(A_I'A_I)^{-1}\mathbf{c}_I\sigma^2. \tag{1.2.4}$$

An unbiased estimate of σ^2 is given by

$$\mathbf{x}'[I(n) - A_I(A_I'A_I)^{-1}A_I']\mathbf{x} \,/\, (n - r). \tag{1.2.5}$$

(1.2.6) The least squares linear estimate of an *estimable* linear function $\mathbf{c}'\boldsymbol{\xi}$ is the same as the unbiased minimum variance estimate given by (1.2.3).

1.3 *Testing of linear hypotheses.* The problem is to test, in terms of the customary F-test, which has a number of well-known good properties, the linear hypothesis

$$H_0: C(s \times m)\boldsymbol{\xi}(m \times 1), \quad \text{or,} \quad \underset{s}{}\underset{r\quad m-r}{[C_1 \quad C_2\,]}\begin{bmatrix}\boldsymbol{\xi}_I\\ \boldsymbol{\xi}_D\end{bmatrix}\begin{matrix} r\\ m-r\end{matrix} \underset{1}{} = \mathbf{0} \tag{1.3.1}$$

against the alternative,

$$H_1: [C_1\, C_2]\begin{bmatrix}\boldsymbol{\xi}_I\\ \boldsymbol{\xi}_D\end{bmatrix} = \mathbf{n}(s \times 1) \neq \mathbf{0} \quad \text{(say)},$$

where $C(s \times m)$ is a matrix given by the hypothesis and is called the *hypothesis* matrix and $\mathbf{n}$ is an arbitrary unspecified nonnull vector. Also rank $(C) = s \leqq r \leqq m < n$. In the discussion in [11, 12] a more general C is introduced, but in almost all problems of practical interest C occurs in the relatively simpler form considered above. The main results follow [11, pp. 81–83b].

(1.3.2) All the following results are invariant under the choice of a basis A_I of A (with a consequent determination of $\boldsymbol{\xi}_I$ and C_1).

A sufficient set of conditions (which under certain further restrictions would also be necessary) for the existence of a similar region test for (1.3.1) is given by

$$C_2 = C_1(A_I'A_I)^{-1}A_I'A_D\,, \tag{1.3.3}$$

or, in other words, C_2 should be related to C_1 through the same matrix postfactor through which A_D is related to A_I. In such a case, the hypothesis (1.3.1) will be said to be *testable* and the condition (1.3.3) will be called the *testability* condition. The testability condition (1.3.3) is a close analogue of the estimability condition (1.2.2), and can also be expressed in a form similar to (1.2.21).

The F-statistic for (1.3.1), having, under H_0, the central F-distribution with degrees of freedom s and $(n - r)$, is given by

$$\frac{\mathbf{x}'A_I(A_I'A_I)^{-1}C_1'[C_1(A_I'A_I)^{-1}C_1']^{-1}C_1(A_I'A_I)^{-1}A_I'\mathbf{x}/s}{\mathbf{x}'[I(n) - A_I(A_I'A_I)^{-1}A_I']\mathbf{x}\,/(n - r)} = \frac{\sigma^2\chi^2/s}{\sigma^2\chi_0^2/(n - r)} \quad \text{(say)}, \tag{1.3.4}$$

to indicate that the numerator multiplied by s/σ^2 has the χ^2-distribution with degrees of freedom s, being a central or non-central χ^2 according as H_0 or H_1 is true, and that the denominator multiplied by $(n - r) / \sigma^2$ has an independent central χ^2-distribution with degrees of freedom $(n - r)$, no matter whether H_0 is true or not.

The quadratic form in the numerator of (1.3.4) is sometimes referred to as the sum of squares due to the hypothesis (1.3.1), and the quadratic form in the denominator is called the sum of squares due to error.

Under H_1 the above F-statistic has a non-central F-distribution with degrees of freedom s and $(n - r)$ and a non-centrality parameter δ^2/σ^2 where,

$$\delta^2 = \mathbf{n}'[C_1(A_I'A_I)^{-1}C_1']^{-1}\mathbf{n}, \tag{1.3.5}$$

which, being a positive definite quadratic form, is zero if, and only if, $\mathbf{n} = \mathbf{0}$, i.e., only under H_0.

Suppose we have two different hypotheses H_{01} and H_{02} given by

$$H_{01}: \underset{s_1}{} \underset{r \quad m-r}{[C_{11} \quad C_{12}]} \underset{1}{\begin{bmatrix} \xi_I \\ \xi_D \end{bmatrix}} \begin{matrix} r \\ m-r \end{matrix} = \mathbf{0}(s_1 \times 1)$$

and

$$H_{02}: \underset{s_2}{} \underset{r \quad m-r}{[C_{21} \quad C_{22}]} \underset{1}{\begin{bmatrix} \xi_I \\ \xi_D \end{bmatrix}} \begin{matrix} r \\ m-r \end{matrix} = \mathbf{0}(s_2 \times 1)$$

against respective alternatives, H_1 and H_2, like the one indicated under (1.3.1), and suppose that rank $[C_{11}\ C_{12}] = s_1$, rank $[C_{21}\ C_{22}] = s_2$ such that $s_1 + s_2 \leqq r \leqq m < n$. Then for H_{01} and H_{02} we shall have respectively

$$F_1 = \frac{\sigma^2\chi_1^2/s_1}{\sigma^2\chi_0^2/(n-r)} \quad \text{and} \quad F_2 = \frac{\sigma^2\chi_2^2/s_2}{\sigma^2\chi_0^2/(n-r)},$$

where the denominator in F_1 and F_2 (the same for both) is the same as that of (1.3.4), and the respective numerators are obtained by substituting C_{11} and C_{21} for C_1 in the numerator of (1.3.4). χ_1^2 and χ_2^2 are each distributed independently of χ_0^2, but we might seek to know the condition for χ_1^2 and χ_2^2 to be distributed independently. The independence of χ_1^2 and χ_2^2, although it would not by any means imply the independence of F_1 and F_2, would nevertheless simplify the distribution problem connected with the simultaneous testing of H_{01} and H_{02} and the associated simultaneous confidence interval estimation. In this situation we would say that F_1 and F_2 are quasi-independent and H_{01} and H_{02} are testable in a quasi-independent manner. The necessary and sufficient condition for this is that

$$C_{11}(A_I'A_I)^{-1}C_{21}' = 0(s_1 \times s_2). \tag{1.3.6}$$

This could be easily generalized to k hypotheses,

$$H_{0i}: \underset{s_i}{} \underset{r \quad (m-r)}{[C_{i1} \quad C_{i2}]} \underset{1}{\begin{bmatrix} \xi_I \\ \xi_D \end{bmatrix}} \begin{matrix} r \\ (m-r) \end{matrix} = \mathbf{0}(s_i \times 1) \qquad (i = 1, \cdots, k),$$

where rank $[C_{i1}\, C_{i2}] = s_i$ such that $\sum_{i=1}^{k} s_i \leqq r \leqq m < n$. A set of necessary and sufficient conditions for these k hypotheses to be testable in a quasi-independent manner is given by,

$$C_{i1}(A_I'A_I)^{-1}C_{j1}' = 0(s_i \times s_j), \qquad (i \neq j = 1, \cdots, k). \tag{1.3.7}$$

1.4 *The associated confidence bounds.* We observe from (1.3.1) that $\mathbf{n}(\neq \mathbf{0})$ represents a deviation from the null hypothesis H_0. The main results follow [11, 12].

With a joint confidence coefficient $\geqq 1 - \alpha$, for a preassigned α, we have the following simultaneous confidence bounds:

$$[\sigma^2\chi^2]^{\frac{1}{2}} - \left[sF_\alpha(s, n-r)\frac{\sigma^2\chi_0^2}{(n-r)}\right]^{\frac{1}{2}} \leqq \{\mathbf{n}'[C_1(A_I'A_I)^{-1}C_1']^{-1}\mathbf{n}\}^{\frac{1}{2}} \leqq [\sigma^2\chi^2]^{\frac{1}{2}} + \left[sF_\alpha(s, n-r)\frac{\sigma^2\chi_0^2}{(n-r)}\right]^{\frac{1}{2}}, \tag{1.4.1}$$

where $\sigma^2\chi^2$ and $\sigma^2\chi_0^2$ (both independent of σ^2) are just the quantities defined in (1.3.4), and $F_\alpha(s, n-r)$ is the upper α-point of the central F-distribution with degrees of freedom s and $(n-r)$;

$$[\sigma^2\chi^{(i)2}]^{\frac{1}{2}} - \left[sF_\alpha(s, n-r)\frac{\sigma^2\chi_0^2}{(n-r)}\right]^{\frac{1}{2}} \leqq \{\mathbf{n}^{(i)\prime}[C_1^{(i)}(A_I'A_I)^{-1}C_1^{(i)\prime}]^{-1}\mathbf{n}^{(i)}\}^{\frac{1}{2}} \leqq [\sigma^2\chi^{(i)2}]^{\frac{1}{2}} + \left[sF_\alpha(s, n-r)\frac{\sigma^2\chi_0^2}{(n-r)}\right]^{\frac{1}{2}}, \tag{1.4.2}$$

for $i = 1, 2, \cdots, k$, where $\mathbf{n}^{(i)}$, $C_1^{(i)}$ and $\sigma^2\chi^{(i)2}$ denote, respectively, the vector $\mathbf{n}$ with the ith component left out, the matrix C_1 with the ith row left out and the $\sigma^2\chi^2$ defined in (1.3.4) with $C_1^{(i)}(\overline{s-1 \times r})$ in place of $C_1(s \times r)$; and likewise

$$[\sigma^2\chi^{(i,j)2}]^{\frac{1}{2}} - \left[sF_\alpha(s, n-r)\frac{\sigma^2\chi_0^2}{(n-r)}\right]^{\frac{1}{2}} \leqq \{\mathbf{n}^{(i,j)\prime}[C_1^{(i,j)}(A_I'A_I)^{-1}C_1^{(i,j)\prime}]^{-1}\mathbf{n}^{(i,j)}\}^{\frac{1}{2}} \leqq \sigma^2\chi^{(i,j)2}]^{\frac{1}{2}} + \left[sF_\alpha(s, n-r)\frac{\sigma^2\chi^2}{(n-r)}\right]^{\frac{1}{2}}, \tag{1.4.3}$$

for $i \neq j = 1, 2, \cdots, s$, where $\mathbf{n}^{(i,j)}$, $C_1^{(i,j)}$ and $\sigma^2\chi^{(i,j)2}$ denote respectively, the vector $\mathbf{n}$ with the ith and jth components left out, the matrix C_1 with the ith and jth rows left out and the $\sigma^2\chi^2$ defined in (1.3.4) with $C_1^{(i,j)}(\overline{s-2 \times r})$ in place of $C_1(s \times r)$; and so on, till we come down to just any single element of $\mathbf{n}$, a single row of C_1 and a consequent truncation on χ^2. Notice that there are $\binom{s}{1}$ statements like (1.4.2), $\binom{s}{2}$ statements like (1.4.3) and so on till we finally come down to $\binom{s}{s-1}$, or $\binom{s}{1}$ statements at the end. Thus the total number of confidence statements like these would be $2^s - 1$. We observe, from these simultaneous confidence bounds, that we obtain confidence bounds not only on parametric functions which measure departure from the whole set of s linear hypotheses in H_0 but also on parametric functions which measure departures from all possible subsets of the s linear hypotheses in H_0.

2. Univariate Variance Components.

2.1 *The Model II of ANOVA.* Let $\mathbf{x}'(1 \times n) = (x_1, x_2, \cdots, x_n)$ be a set of n observable stochastic variates such that

$$\text{(2.1.1)} \qquad \mathbf{x}(n \times 1) = A(n \times m)\,\boldsymbol{\xi}(m \times 1) + \boldsymbol{\varepsilon}(n \times 1), \qquad m < n,$$

$$= n\begin{bmatrix} A_1 & A_2 & \cdots & A_k \end{bmatrix}_{\;m_1 \;\; m_2 \;\cdots\; m_k} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{bmatrix}\begin{matrix} m_1 \\ m_2 \\ \vdots \\ m_k \end{matrix}_{\,1} + \boldsymbol{\varepsilon}, \quad \sum_{i=1}^{k} m_i = m, \quad \text{(say)},$$

where $A(n \times m)$, to be called the *design* matrix, is a matrix whose elements are constants given by the design of the experiment and is of rank $r \leqq m < n$, and where

(i) $\boldsymbol{\xi}_i(m_i \times 1)$ is a random sample of size m_i from the normal population $N(\mu_i, \sigma_i^2)$ for $i = 1, 2, \cdots, k$, and $\boldsymbol{\varepsilon}(n \times 1)$ and $\boldsymbol{\xi}_i$'s (for $i = 1, 2, \cdots, k$) are mutually independent;

(ii) $\boldsymbol{\varepsilon}(n \times 1)$, whose elements are physically of the nature of errors, is a random sample from $N(0, \sigma^2)$.

Under this model it is seen that $\mathbf{x}(n \times 1)$ is n-variate normal

$$N[E(\mathbf{x}), \quad \Sigma(n \times n)],$$

where

$$\text{(2.1.2)} \qquad E(\mathbf{x})(n \times n) = A\;(n \times m)\begin{bmatrix} \mu_1 & \cdot & \mathbf{1} \\ \mu_2 & \cdot & \mathbf{1} \\ \vdots & & \vdots \\ \mu_k & \cdot & \mathbf{1} \end{bmatrix}\begin{matrix} m_1, \\ m_2 \\ \vdots \\ m_k \end{matrix}_{\,1}$$

$\mathbf{1}(m_i \times 1)$ denoting a vector of m_i unities ($i = 1, \cdots, k$), and

$$\text{(2.1.3)} \qquad \Sigma(n \times n) = E(\mathbf{x}\,\mathbf{x}') - E(\mathbf{x})E(\mathbf{x}')$$

$$= A(n \times m)\begin{bmatrix} \sigma_1^2\, I(m_1) & 0 & \cdots & 0 \\ 0 & \sigma_2^2 I(m_2) & \cdots & 0 \\ \cdots & \cdots & & \cdots \\ 0 & 0 & \cdots & \sigma_k^2 I(m_k) \end{bmatrix} A'(m \times n) + \sigma^2 I(n)$$

$$= \sum_{i=1}^{k} \sigma_i^2 A_i\, A_i' + \sigma^2 I(n).$$

As in section 1.1, it is to be observed that for purposes of point estimation the assumption of normality of the distributions made in (i) and (ii) of the Model II is not necessary. In fact, this assumption is even, probably, not as realistic in practice as the assumption of sampling from finite populations with certain known probabilities, which will be discussed in later papers. The assumption of normality, however, simplifies the distribution problems connected with testing of hypotheses (simultaneous and/or separate) and confidence interval estimation.

We shall consider, in greater detail, that restricted type of k-way classification for which the design matrix $A(n \times m)$ is such that each row of the submatrix $A_i(n \times m_i)(i = 1, 2, \cdots, k)$ has one and only one non-zero element which is equal to unity and rank $(A) = m - k + 1$. [Notice that in general for such A, rank $(A) \leqq (m - k + 1)$.] It can be seen that the usual complete and incomplete *connected* designs are included in this general case. For one thing, these designs are relatively simple to discuss from the standpoint of either testing of hypotheses or confidence interval estimation under Model II and, for another thing, this discussion will prepare the ground for the relatively more difficult problem of unconnected designs which will be discussed in subsequent papers.

Our objectives are: (i) to estimate any *estimable* linear function of $\mu_1, \mu_2, \cdots \mu_k$ and to test *testable* linear hypotheses on $\mu_1, \mu_2, \cdots, \mu_k$; (ii) to obtain estimates of, and test hypotheses on, the variance components $\sigma_1^2, \sigma_2^2, \cdots, \sigma_k^2, \sigma^2$; and (iii) to obtain confidence bounds on the parameters or certain parametric functions [e.g., ratios like σ_i^2/σ^2] which are meaningful measures of departure from certain customary null hypotheses.

2.2 *Linear estimation and testing of linear hypotheses.* Using a result given in [11] we can establish the following lemma [3, pp. 59–63].

LEMMA 1: *For the restricted k-way classification defined above, the necessary and sufficient condition for the estimability of* $\mathbf{c}'(1 \times m)\ E(\xi)(m \times 1)$, *a linear function of* $\mu_1, \cdots, \mu_k$, *is that coefficient of* μ_1 = *coefficient of* $\mu_2 = \cdots =$ *coefficient of* μ_k.

This lemma establishes that, for the restricted k-way classification defined above, the only independent linear function of $\mu_1, \cdots, \mu_k$ which is *estimable* and hypotheses on which are *testable* is the sum $\mu = [\mu_1 + \mu_2 + \cdots + \mu_k]$, all other such functions being merely multiples of μ.

2.3 *Estimation of the variance components.* We shall seek $(k + 1)$ quadratic forms, $q_i = \mathbf{x}'Q_i\mathbf{x}'(i = 0, 1, \cdots, k)$, of the observations to be utilized in the point estimation of, testing of hypotheses on and the confidence interval estimation of the variance components. We shall impose on the q_i's the following restrictions, which will be justified presently:

(2.3.1) q_i, of rank n_i, is distributed as $\lambda_i\chi^2_{(n_i)}$, where $\chi^2_{(n_i)}$ denotes the central χ^2 variate with degrees of freedom n_i and where $\lambda_i = E(q_i/n_i)$ $(i = 0, 1, \cdots, k)$.

(2.3.2) q_i's are mutually independent.

LEMMA 2: *If* $\mathbf{x}(n \times 1)$ *is distributed as* $N[E(\mathbf{x}), \Sigma(n \times n)]$ *and* $q_i = \mathbf{x}'Q_i\mathbf{x}$ $(i = 0, 1, \cdots, k)$ *is a quadratic form of rank* n_i, $(\sum_{i=0}^{k} n_i \leqq n)$, *then, a set of* *necessary and sufficient conditions for* $q_0, q_1, \cdots, q_k$ *to satisfy the above restrictions is given by*

$$(\alpha)\ Q_i \Sigma Q_i = \lambda_i Q_i, \qquad i = 0, 1, \cdots, k;$$

$$(\beta)\ E(\mathbf{x}')Q_i E(\mathbf{x}) = 0, \qquad i = 0, 1, \cdots, k;$$

$$(\gamma)\ Q_i \Sigma Q_j = 0(n \times n), \qquad i \neq j = 0, 1, \cdots, k,$$

where (α) *ensures* χ^2*-distributions (in general non-central),* (β) *ensures the centrality of the* χ^2*-distributions and* (γ) *which actually ensures pair-wise independence but is easily checked to ensure mutual independence of the distributions as well in this case. For a proof of this lemma see* [1, 6].

LEMMA 3: *If* $E[\mathbf{x}(p \times 1)\mathbf{y}'(1 \times p)] = \mathcal{E}(p \times p)$ *then* $E(\mathbf{x}'Q\mathbf{y})$*, where* $Q(p \times p)$ *is symmetric, is* tr $\mathcal{E}Q$ *where "*tr A*" denotes the trace (the sum of the diagonal elements) of the square matrix* A.

PROOF:

$$\begin{aligned} E(\mathbf{x}'Q\mathbf{y}) &= E[\sum_{i,j=1}^{p} q_{ij}x_i y_j], && \text{if } Q(p \times p) = ((q_{ij})), \\ &= \sum_{i,,j=1}^{p} q_{ij}\epsilon_{ij}, && \text{if } \mathcal{E}(p \times p) = ((\epsilon_{ij})), \\ &= \sum_{i,j=1}^{p} \epsilon_{ij} q_{ji}, && \text{since } q_{ij} = q_{ji}, \\ &= \text{tr } \mathcal{E}Q \end{aligned}$$

COROLLARY: *For the Model* II *of ANOVA, we have*

$$(2.3.3)\quad \begin{aligned} \lambda_i = E\left(\frac{q_i}{n_i}\right) &= \frac{1}{n_i} E(\mathbf{x}'\, Q_i\, \mathbf{x}) = \frac{1}{n_i} \text{tr } \{[\Sigma + E(\mathbf{x})E(\mathbf{x}')]Q_i\}, && \text{where } E(\mathbf{x}) \text{ and } \Sigma \text{ are defined in (2.1.2) and (2.1.3)}, \\ &= \frac{1}{n_i} \{\text{tr } \Sigma\, Q_i + E(\mathbf{x}')Q_i\, E(\mathbf{x})\}, && \text{since tr } (AB) = \text{tr } (BA) \text{ and tr (scalar) = scalar ([11], p. } A\text{-1)}, \\ &= \frac{1}{n_i} \text{tr } \Sigma\, Q_i, && \text{if } q_i \text{ satisfies } (\beta) \text{ of Lemma 2}, \\ &= \frac{1}{n_i}\left[\sum_{j=1}^{k} \sigma_j^2 \text{ tr } A_j\, A_j'\, Q_i + \sigma^2 \text{ tr } Q_i\right], && \text{using (2.1.3)}. \end{aligned}$$

This holds for $i = 0, 1, \cdots, k$.

Next, suppose that in Lemma 2 $\Sigma(n \times n)$ is unknown, and we require, as a further condition on the q_i's, that $q_0, q_1, \cdots, q_k$ satisfy (α) and (γ) of Lemma 2 for all symmetric positive definite matrices $\Sigma(n \times n)$. Under Model II, this means that, for all $\sigma_1^2, \cdots, \sigma_k^2$ and σ^2, we require the quadratic forms $q_0, \cdots, q_k$ to satisfy (α) and (γ) in addition to (β). Using (2.1.3) and (2.3.3), this means that, for Model II, (α) and (γ) reduce respectively to

$$\left.\begin{aligned} Q_i A_l A_l' Q_i &= \left[\frac{1}{n_i} \operatorname{tr} A_l A_l' Q_i\right] Q_i, l = 1, 2, \cdots, k; \\ Q_i^2 &= \left[\frac{1}{n_i} \operatorname{tr} Q_i\right] Q_i \end{aligned}\right\} \quad (i = 0, 1, \cdots, k), \tag{2.3.4}$$

and

$$\left.\begin{aligned} Q_i A_l A_l' Q_j &= 0(n \times n), l = 1, \cdots, k; \\ Q_i Q_j &= 0(n \times n) \end{aligned}\right\} \quad (i \neq j = 0, 1, \cdots, k). \tag{2.3.5}$$

Before we proceed further we shall justify the restrictions (2.3.1) and (2.3.2). These restrictions provide a set of sufficient (though not necessary) conditions for ensuring certain good properties of the solutions to the problems of point estimation, testing of hypotheses and confidence interval estimation.

From the standpoint of point estimation, we have the following lemma:

LEMMA 4: *Under Model II, if $q_0, q_1, \cdots, q_k$ are $(k + 1)$ quadratic forms, of ranks $n_0, n_1, \cdots, n_k$ respectively, satisfying* (2.3.1) *and* (2.3.2) *and* $\lambda_i = E(q_i/n_i)$, *then, the unbiased estimate with uniformly least variance of the estimable linear function* $\sum_{i=0}^k l_i \lambda_i$ *is given by* $\sum_{i=0}^k l_i q_i/n_i$ *and this estimate is a unique (except on a set measure zero) function of* $q_0, q_1, \cdots, q_k$. This lemma is essentially the same as the result given by the authors of [4] and the proof follows from a theorem of Lehmann and Scheffé [5] when we notice that, under Model II, if $q_0, \cdots, q_k$ satisfy (2.3.1) and (2.3.2) then $q_0, \cdots, q_k$ from a set of sufficient statistics for $\lambda_0, \lambda_1, \cdots, \lambda_k$ (also for $\sigma^2, \sigma_1^2, \cdots, \sigma_k^2$), and that they can also be shown to satisfy the completeness condition of Lehmann and Scheffé [5].

It may be noted that Lemma 4 holds, not only for a linear function of λ_i's, but also, in general, for any real valued *estimable* function $f(\lambda_0, \lambda_1, \cdots, \lambda_k)$.

Next, from the standpoint of testing of hypotheses, we observe that hypotheses on variance components are usually composite and that a legitimate quest might be to obtain similar region tests for these hypotheses. From the properties of sufficiency and completeness mentioned above for $q_0, q_1, \cdots, q_k$ satisfying (2.3.1) and (2.3.2), it follows, from another theorem of Lehmann and Scheffé [5], that the class of all similar tests of hypotheses on the σ^2's will be of Neyman structure, or Neyman mechanism regions [8], with respect to $q_0, q_1, \cdots, q_k$.

Finally, if the quadratic forms $q_0, q_1, \cdots, q_k$ satisfy these restrictions then, as will be seen later in this paper, we can obtain simultaneous confidence intervals on $\sigma_1^2, \sigma_2^2, \cdots, \sigma_k^2, \sigma^2$ and on ratios like σ_i^2/σ^2 without running into intractable distribution problems or nuisance parameters, although it is not said that this would be impossible except under these restrictions.

We shall now proceed to obtain a set of quadratic forms $q_0, q_1, \cdots, q_k$ for the restricted k-way classification and present a tie-up between the analysis under Model I, considered under Section 1, and the analysis under Model II.

2.4 *Tie-up between the analysis under Models I and II for the restricted k-way classification.* We recall from Section 1.3 that, under Model I, we can obtain k sums of squares due to the k *testable* hypotheses of equality of the elements of ξ_i $(i = 1, 2, \cdots, k)$, which can, by analogy with (1.3.1), be written as

$$H_{0i}: C_i(\overline{m_i - 1} \times m)\xi(m \times 1) = [C_{i1}\ C_{i2}]\xi, \quad \text{where } r = \text{rank}(A) = (m - k + 1) \text{ and } m = \sum_{i=1}^{k} m_i,$$

$$(2.4.1) \qquad = (m_i - 1)\left[\begin{array}{c|c|c|cccccc|c|c|c} 0 & 0 & \cdots & 1 & 0 & \cdots & 0 & -1 & 0 & \cdots & 0 \\ & & & 0 & 1 & \cdots & 0 & -1 & & & \\ & & & \cdot & \cdot & \cdots & \cdot & \cdot & & & \\ & & & 0 & 0 & \cdots & 1 & -1 & & & \end{array}\right]\xi$$

$$\qquad\qquad m_1\ m_2\ \cdots \qquad m_i \qquad \cdots \qquad m_k$$

$$= \mathbf{0}(\overline{m_i - 1} \times 1),$$

so that rank $(C_i) = (m_i - 1)(i = 1, 2, \cdots, k)$. It is easily verified that, if $\xi_i' = (\xi_{i1}, \xi_{i2}, \cdots, \xi_{im_i})$, then, the hypothesis $\xi_{i1} = \cdots = \xi_{im_i}$ is exactly equivalent to (2.4.1). As in Section 1.3, we obtain k sums of squares due to the k hypotheses $H_{01}, H_{02}, \cdots, H_{0k}$, viz.,

$$(2.4.2) \qquad \mathbf{x}'A_I(A_I'A_I)^{-1}C_{i1}'[C_{i1}(A_I'A_I)^{-1}C_{i1}']^{-1}C_{i1}(A_I'A_I)^{-1}A_I'\mathbf{x}$$

for $i = 1, 2, \cdots, k$ with $(m_i - 1)(= n_i$, say) degrees of freedom. We further have the sum of squares due to error,

$$(2.4.3) \qquad \mathbf{x}'[I(n) - A_I(A_I'A_I)^{-1}A_I']\mathbf{x}$$

with degrees of freedom $= n - r = (n - m + k - 1)(= n_0$, say). Notice that $\sum_{i=0}^{k} n_i = (n - 1) < n$.

Now, under Model II, in the notation of section 2.3, we take q_0 to be the sum of squares due to the error given by (2.4.3) and q_i, for $i = 1, 2, \cdots, k$, to be the sums of squares due to the hypothesis given by (2.4.2). If then we apply the conditions (α), (β) and (γ) of Lemma 2 to this set of quadratic forms, we can verify that $q_0, q_1, \cdots, q_k$ all satisfy (β) so that centrality of the distribution (if it is χ^2 at all) is assured for each $q_i(i = 0, 1, \cdots, k)$. Using the fact that the matrices of these quadratic forms are all idempotent, and by repeated application of Lemma 3, we can also obtain that

$$(2.4.4) \qquad \lambda_0 = \sigma^2; \quad \lambda_i = \nu_i\sigma_i^2 + \sigma^2, \quad \text{for } i = 1, 2, \cdots, k,$$

where

$$\nu_i = \frac{2}{(m_i - 1)}\left\{\begin{array}{r}\text{sum of all the elements in and below the diagonal} \\ \text{of the symmetric matrix } [C_{i1}(A_I'A_I)^{-1}C_{i1}']^{-1}\end{array}\right\}$$

Thus $[1/\nu_i][q_i/n_i - q_0/n_0]$ will be an unbiased estimate of σ_i^2, while q_0/n_0 will be an unbiased estimate of σ^2, so that, if these particular $q_0, q_1, \cdots, q_k$ satisfy (α) and (γ) of Lemma 2 as well, then, by Lemma 4, these estimates will also have uniformly least variance in the class of unbiased estimates. It is easily verified that q_0 always satisfies (α) and (γ) of Lemma 2, so that, the error sum of squares obtained under Model I is always, under Models I and II, distributed as $\sigma^2\chi^2_{(n_0)}$, where $\chi^2_{(n_0)}$ is the central χ^2 variate with $n_0 (= n - m + k - 1)$ degrees of freedom, independently of $q_1, q_2, \cdots, q_k$ and therefore, always provides (when divided by n_0) an unbiased estimate, with uniformly least variance, of σ^2. Next, applying the conditions (α) and (γ) of Lemma 2 to q_i $(i = 1, 2, \cdots, k)$ and simplifying the conditions, we get them respectively in the forms

$$(2.4.5) \qquad C_{i1}(A_I'A_I)^{-1}C_{i1}' = \frac{1}{\nu_i}[I(m_i - 1) + J(\overline{m_i - 1} \times \overline{m_i - 1})], \quad \text{for } i = 1, 2, \cdots, k,$$

where J $(p \times q)$ stands for a matrix all of whose elements are equal to unity, and

$$(2.4.6) \qquad C_{i1}(A_I'A_I)^{-1}C_{j1}' = 0(\overline{m_i - 1} \times \overline{m_j - 1})$$

for $i \neq j = 1, 2, \cdots, k$. Note that (2.4.5) and (2.4.6) are independent of the unknown variance components $\sigma_1^2, \cdots, \sigma_k^2$ and σ^2.

The conditions (2.4.5) and (2.4.6) are both satisfied by the usual complete designs like Randomized Block, Latin Square, Factorial designs under a strictly additive model with no interactions. However, the incomplete designs, like Balanced Incomplete Block designs, do not, in general, satisfy (2.4.6) while they do satisfy (2.4.5). Thus the restrictions (2.3.1) and (2.3.2), taken together, are not too restrictive in that the usual complete designs have sums of squares like (2.4.2) which are useful in the analysis (to answer customary questions) under *both* Models I and II. However, they are restrictive in that the incomplete designs do not have sums of squares like (2.4.2) that can be used directly and conveniently in the analyses of *both* the models.

In Sections 2.5 and 2.6, we shall discuss some simple situations by assuming that the k-way classification under consideration has sums of squares like (2.4.2) which satisfy both (2.4.5) and (2.4.6), thus rendering relatively easy simultaneous testing and simultaneous confidence interval estimation of $\sigma_1^2/\sigma^2, \cdots, \sigma_k^2/\sigma^2$. If, however, (2.4.6) were not satisfied but merely (2.4.5) as, for example, in incomplete block designs, then *simultaneous* testing or *simultaneous* interval estimation would be far more difficult (and will be discussed in later papers) but the *separate* tests and *separate* confidence interval estimation can be obtained in exactly the same way as in the following discussion. Problems involving interactions in factorial designs will also be discussed in later communications.

2.5 *Tests of hypotheses on the variance components.* The usual hypotheses tested are $H_{0i}:\sigma_i^2 = 0$ against respective alternatives $H_{1i}:\sigma_i^2 > 0$ for $i = 1, 2, \cdots, k$. Working in terms of the sums of squares (2.4.2), (2.4.3), it can be seen from (2.4.4) that these hypotheses are equivalent to $H_{0i}':\lambda_i = \lambda_0$ against $H_{1i}':\lambda_i > \lambda_0$ for $i = 1, 2, \cdots, k$. For each i, therefore, under Model II, we can test H_{0i}' against H_{1i}' by taking as the critical region the region defined by

$$(2.5.1) \qquad F_i = \frac{q_i/n_i}{q_0/n_0} > F_\alpha(n_i, n_0)$$

where F_i, under H_{0i}', has a central F-distribution with degrees of freedom n_i and n_0 and $F_\alpha(n_i, n_0)$ is the upper $100\alpha\%$ point of the central F-distribution with degrees of freedom n_i and n_0. From (1.3.4), notice that the critical regions (2.5.1) for the individual hypotheses H_{0i}, under Model II, are identical with those obtained for the individual hypotheses (2.4.1) under Model I.

The critical regions under Models I and II have an identical nature even when we consider the simultaneous hypotheses $H_0:\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 = 0$ against the alternative H_1: at least one $\sigma_i^2 > 0$, which is equivalent to considering $H_0':\lambda_1/\lambda_0 = \cdots = \lambda_k/\lambda_0 = 1$ against H_1': at least one $\lambda_i/\lambda_0 > 1$. The critical region of the simultaneous test obtained by the heuristic union-intersection principle [9] is,

$$(2.5.2) \qquad F_1 > a_1, \quad F_2 > a_2, \quad \cdots, \quad F_k > a_k,$$

where $F_i = (q_i/n_i) / (q_0/n_0)$, F_i's are, in the terminology of section 1.3, quasi-independent variance ratios and a_i's are such that the region (2.5.2) is of size α for a preassigned α. It is easily seen that the critical region (2.5.2) is identical with that of the simultaneous ANOVA test of Ghosh for Model I [7].

2.6 *Simultaneous confidence statements.* When the q_i's, given by (2.4.2), (2.4.3), satisfy the restrictions (2.3.1) and (2.3.2) we can find constants $\chi_{1\alpha_j}^2(n_j) = \chi_{1\alpha_j}^2$ (say) and $\chi_{2\alpha_j}^2(n_j) = \chi_{2\alpha_j}^2$ (say) for $j = 0, 1, \cdots, k$, such that the simultaneous statements,

$$(2.6.1) \quad \chi_{1\alpha_0}^2 \leqq q_0/\lambda_0 \leqq \chi_{2\alpha_0}^2, \quad \chi_{1\alpha_1}^2 \leqq q_1/\lambda_1 \leqq \chi_{2\alpha_1}^2, \cdots, \chi_{1\alpha_k}^2 \leqq q_k/\lambda_k \leqq \chi_{2\alpha_k}^2$$

have a joint confidence coefficient $(1 - \alpha) = \prod_{j=0}^k (1 - \alpha_j)$, where

$$P(\chi_{1\alpha_j}^2 \leqq \chi_{(n_j)}^2 \leqq \chi_{2\alpha_j}^2) = (1 - \alpha_j)$$

and $\chi_{(n_j)}^2$ denotes the central χ^2 variate with n_j degrees of freedom, $(j = 0, 1, \cdots, k)$. By inverting the statements (2.6.1), we obtain, with a joint confidence coefficient $(1 - \alpha)$, the simultaneous confidence statements,

$$(2.6.2) \quad \begin{array}{c} c_{1\alpha_0}q_0 \leqq \lambda_0 \leqq c_{2\alpha_0}q_0, \quad c_{1\alpha_1}q_1 \leqq \lambda_1 \leqq c_{2\alpha_1}q_1, \quad \cdots, \\ c_{1\alpha_k}q_k \leqq \lambda_k \leqq c_{2\alpha_k}q_k, \end{array}$$

where $c_{1\alpha_j} = [\chi^2_{2\alpha_j}]^{-1}$ and $c_{2\alpha_j} = [\chi^2_{1\alpha_j}]^{-1}$ for $j = 0, 1, \cdots, k$. Recalling (2.4.4), we can obtain the following set of simultaneous confidence interval statements, which are implied by (2.6.2), on the variance components:

$$\text{(2.6.3)}\quad \begin{array}{c} c_{1\alpha_0} q_0 \leqq \sigma^2 \leqq c_{2\alpha_0} q_0, \quad \frac{1}{\nu_1}[c_{1\alpha_1} q_1 - c_{2\alpha_0} q_0] \leqq \sigma_1^2 \leqq \frac{1}{\nu_1}[c_{2\alpha_1} q_1 - c_{1\alpha_0} q_0], \\ \cdots, \quad \frac{1}{\nu_k}[c_{1\alpha_k} q_k - c_{2\alpha_0} q_0] \leqq \sigma_k^2 \leqq \frac{1}{\nu_k}[c_{2\alpha_k} q_k - c_{1\alpha_0} q_0]. \end{array}$$

Since (2.6.2) implies (2.6.3) it follows that the confidence coefficient associated with the statements (2.6.3) is $\geqq (1 - \alpha)$. In order to be non-trivial, of course, the constants $c_{1\alpha_j}$, $c_{2\alpha_j}$ (for $j = 0, 1, \cdots, k$) must be such that all the bounds in (2.6.3) are non-negative.

As a simple extension we can also obtain simultaneous confidence interval statements on $\mu = \mu_1 + \mu_2 + \cdots + \mu_k$ and the variance components.

Next we shall obtain simultaneous confidence bounds on the ratios σ_1^2/σ^2, $\sigma_2^2/\sigma^2, \cdots, \sigma_k^2/\sigma^2$. When the q_i's satisfy the restrictions (2.3.1) and (2.3.2) then, for $i = 1, 2, \cdots, k$, $F_i = (q_i/n_i\lambda_i) / (q_0/n_0\lambda_0)$ are quasi-independent in the sense of section 1.3, each having a central F-distribution with degrees of freedom n_i and n_0. The joint distribution of these quasi-independent F_i's is known [7, 3] and we can determine constants F_{i1}, F_{i2}, for $i = 1, 2, \cdots, k$, such that the simultaneous statements,

$$\text{(2.6.4)}\quad F_{11} \leqq \frac{q_1/n_1\lambda_1}{q_0/n_0\lambda_0} \leqq F_{12}, \cdots, F_{k1} \leqq \frac{q_k/n_k\lambda_k}{q_0/n_0\lambda_0} \leqq F_{k2},$$

have a joint probability $= (1 - \alpha)$, for a preassigned α. Recalling (2.4.4), we can invert (2.6.4) to obtain the simultaneous confidence interval statements,

$$\text{(2.6.5)}\quad \begin{array}{c} \frac{1}{\nu_1}\left[\frac{n_0}{n_1 F_{12}}\frac{q_1}{q_0} - 1\right] \leqq \frac{\sigma_1^2}{\sigma^2} \leqq \frac{1}{\nu_1}\left[\frac{n_0}{n_1 F_{11}}\frac{q_1}{q_0} - 1\right], \\ \cdots \qquad \cdots \qquad \cdots \\ \frac{1}{\nu_k}\left[\frac{n_0}{n_k F_{k2}}\frac{q_k}{q_0} - 1\right] \leqq \frac{\sigma_k^2}{\sigma^2} \leqq \frac{1}{\nu_k}\left[\frac{n_0}{n_k F_{k1}}\frac{q_k}{q_0} - 1\right], \end{array}$$

with a joint confidence coefficient $= (1 - \alpha)$. Here, again, for non-triviality the bounds should all be non-negative.

REFERENCES

[1] O. CARPENTER, "Note on the extension of Craig's theorem to non-central variates," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 455–457.

[2] C. EISENHART, "The assumptions underlylng the analysis of variance," *Biometrics*, Vol. 3 (1947), pp. 1–21.

[3] R. GNANADESIKAN, "Contributions to multivariate analysis including univariate and multivariate variance components analysis and factor analysis," Institute of Statistics, University of North Carolina, *Mimeo. Series, No.* 158 (1956).

[4] F. A. GRAYBILL AND A. W. WORTHAM, "A note on uniformly best unbiased estimators for variance components," *J. Amer. Stat. Assn.*, Vol. 51 (1956), pp. 266–268.

[5] E. Lehmann and H. Scheffé, "Completeness, similar regions and unbiased estimation, Part I," *Sankhya*, Vol. 10 (1950), pp. 305–340.
[6] J. Ogawa, "On the independence of quadratic forms in a non-central normal system," *Osaka Math. Jour.*, Vol. 2 (1950), pp. 151–159.
[7] K. V. Ramachandran, "On the simultaneous analysis of variance test," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 521–528.
[8] S. N. Roy, "Univariate and multivariate analysis as problems in testing of composite hypotheses I," *Sankhya*, Vol. 10 (1950), pp. 29–80.
[9] S. N. Roy, "On a heuristic method of test construction and its uses in multivariate analysis," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 220–238.
[10] S. N. Roy and R. C. Bose, "Simultaneous confidence interval estimation," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 513–536.
[11] S. N. Roy, "A report on some aspects of multivariate analysis," Institute of Statistics, University of North Carolina, *Mimeo. Series, No.* 121 (1954).
[12] S. N. Roy and R. Gnanadesikan, "Further contributions to multivariate confidence bounds," Institute of Statistics, University of North Carolina, *Mimeo. Series, No.* 155 (1956). Also *Biometrika*, Vol. 44 (1957), pp. 399–410.