ON χ^2 AND INDEPENDENCE PROPERTIES OF SUMS OF SQUARES

by

S. R. Searle and C. Berry

Biometrics Unit, Cornell University, Ithaca, New York 14853

BU-763-M

Revised,
September, 1983

Abstract

In a fixed effects analysis of variance it is well known that sums of squares from the method of fitting constants are, under normality assumptions, distributed as multiples of χ^2-distributions, and they and the error sum of squares are generally independent of one another.

In mixed (and random) models with normality assumptions, the error sum of squares is always a multiple of a χ^2 variable but this is no necessarily so for other sums of squares; and although those other sums of squares are independent of the error sum of squares, they are not generally independent of each other.

1. Basic Theorems

Two basic theorems concerning χ^2 and independence properties of quadratic forms in normal variables come from Searle (1971, Section 2.5, Theorems 2 and 4).

When $\mathbf{y}$ has a normal distribution with mean $\boldsymbol{\mu}$ and positive definite dispersion matrix $\mathbf{V}$, i.e., $\mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{V})$,

Theorem B1: $\mathbf{y}'\mathbf{A}\mathbf{y}$ has a non-central χ^2-distribution if and only if $\mathbf{AV}$ is idempotent.

Paper No. BU-763-M in the Biometrics Unit.

Theorem B2: $y'Ay$ and $y'By$ are independent if and only if $AVB = 0$.

In Theorem B1 the degrees of freedom of the χ^2-density equal the rank of A, and the non-centrality parameter is $\frac{1}{2}\mu'A\mu$. In Theorem B2 the condition $AVB = 0$ can be equivalently stated as $BVA = 0$.

2. The Method of Fitting Constants

Represent the familiar linear model for estimating estimable functions of fixed effects β as

$$y = X\beta + e \quad . \tag{1}$$

y is the vector of observations, β is the vector of parameters of the model and X is the associated coefficient matrix, often an incidence matrix. e is the vector of differences $e = y - E(y)$ where $E(y) = X\beta$ is the expected value of y over repeated sampling; and e is assumed to be a vector of random variables with mean 0 and dispersion matrix $\sigma^2 I$:

$$E(e) = 0 \qquad \text{and} \qquad \mathrm{var}(e) = \sigma^2 I_N \quad . \tag{2}$$

The sum of squares due to fitting (1) and (2) by least squares will be denoted by $R(\beta)$ and is

$$R(\beta) = y'X(X'X)^{-}X'y = y'Py \tag{3}$$

for

$$P = X(X'X)^{-}X' \quad = XX^{+} \tag{4}$$

where $(X'X)^{-}$ is any generalized inverse of $X'X$ satisfying $X'X(X'X)^{-}X'X = X'X$ and X^{+} is the Moore-Penrose inverse of X (e.g., Searle, 1982, Chapter 8). Then the error sum of squares after fitting (1) is

$$SSE = y'y - R(\beta) = y'My \tag{5}$$

for

$$\mathbf{M} = \mathbf{I} - \mathbf{P} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}' = \mathbf{I} - \mathbf{X}\mathbf{X}^{+} \quad . \tag{6}$$

Properties of $(\mathbf{X}'\mathbf{X})^{-}$ and $\mathbf{X}^{+}$ ensure that $\mathbf{P}$ amd $\mathbf{M}$ are symmetric and idempotent, i.e.,

$$\mathbf{P}^2 = \mathbf{P} = \mathbf{P}' \quad \text{and} \quad \mathbf{M}^2 = \mathbf{M} = \mathbf{M}' \quad , \tag{7}$$

and that

$$\mathbf{P}\mathbf{X} = \mathbf{X}, \quad \mathbf{M}\mathbf{X} = \mathbf{0} \quad \text{and} \quad \mathbf{M}\mathbf{P} = \mathbf{0} \quad .$$

Now consider a partitioned form of (1), using

$$\mathbf{X} = [\mathbf{X}_1 \;\; \mathbf{X}_2 \;\; \mathbf{X}_3 \;\; \mathbf{X}_4] \quad , \tag{8}$$

namely

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{X}_3\boldsymbol{\beta}_3 + \mathbf{X}_4\boldsymbol{\beta}_4 + \mathbf{e} \quad . \tag{9}$$

Then in the method of fitting constants, a typical sum of squares is of the form $R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)$ for sub-vectors $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ of some partitioning of $\boldsymbol{\beta}$. For some partitionings, $\mathbf{X}_3$ and $\mathbf{X}_4$ may not exist, but whether they exist or not, $R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)$ is defined as

$$R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1) = R(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) - R(\boldsymbol{\beta}_1) \quad . \tag{10}$$

Applying (3) to each term in (10), we use

$$\mathbf{M}_i = \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^{-}\mathbf{X}_i' \qquad \text{for} \qquad i = 1,2 \tag{11}$$

and

$$\mathbf{M}_{12} = \mathbf{X}_{12}(\mathbf{X}_{12}'\mathbf{X}_{12})^{-}\mathbf{X}_{12} \quad \text{for} \quad \mathbf{X}_{12} = [\mathbf{X}_1 \;\; \mathbf{X}_2] \quad , \tag{12}$$

where $\mathbf{M}_i$ and $\mathbf{M}_{12}$ are $\mathbf{M}$ of (6) using $\mathbf{X}_i$ (for $i = 1,2$) and $\mathbf{X}_{12}$ of (12) in place of $\mathbf{X}$. Similarly for $\mathbf{P}_i$ and $\mathbf{P}_{12}$. Then for (10),

$$R(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = \mathbf{y}'\mathbf{P}_{12}\mathbf{y} \quad \text{and} \quad R(\boldsymbol{\beta}_1) = \mathbf{y}'\mathbf{P}_1\mathbf{y}$$

so that (10) is

$$\begin{aligned} R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1) &= \mathbf{y}'(\mathbf{P}_{12} - \mathbf{P}_1)\mathbf{y} \\ &= \mathbf{y}'\mathbf{P}_{2|1}\mathbf{y} \quad \text{for} \quad \mathbf{P}_{2|1} = \mathbf{P}_{12} - \mathbf{P}_1 \quad . \end{aligned} \tag{13}$$

Based on (7), matrices $\underset{\sim}{P}_{12}$ and $\underset{\sim}{P}_1$ are symmetric, and therefore so is $\underset{\sim}{P}_{2|1}$. It is also idempotent because from (7) and (12)

$$\underset{\sim}{P}_{12}\underset{\sim}{X}_{12} = \underset{\sim}{X}_{12} \Rightarrow \underset{\sim}{P}_{12}\underset{\sim}{X}_1 = \underset{\sim}{X}_1 \quad \text{and} \quad \underset{\sim}{X}_1'\underset{\sim}{P}_{12} = \underset{\sim}{X}_1' \tag{14}$$

and so therefore

$$\underset{\sim}{P}_{12}\underset{\sim}{P}_1 = \underset{\sim}{P}_{12}\underset{\sim}{X}_1\underset{\sim}{X}_1^+ = \underset{\sim}{X}_1\underset{\sim}{X}_1^+ = \underset{\sim}{P}_1 \quad .$$

Hence

$$\underset{\sim}{P}_{2|1}\underset{\sim}{P}_1 = \underset{\sim}{P}_{12}\underset{\sim}{P}_1 - \underset{\sim}{P}_1^2 = \underset{\sim}{P}_1 - \underset{\sim}{P}_1 = \underset{\sim}{0} \tag{15}$$

and thus

$$\underset{\sim}{P}_{2|1}^2 = \underset{\sim}{P}_{12}^2 + \underset{\sim}{P}_1^2 - \underset{\sim}{P}_1 - \underset{\sim}{P}_1 = \underset{\sim}{P}_{12} - \underset{\sim}{P}_1 = \underset{\sim}{P}_{2|1} \quad . \tag{16}$$

The generality of (14) and (15) is to be appreciated. The subscripts 1 and 2 represent any (mutually exclusive) sub-vectors $\underset{\sim}{\beta}_1$ and $\underset{\sim}{\beta}_2$ of $\underset{\sim}{\beta}$. Therefore (14) also includes, for example, $\underset{\sim}{P}_{123}\underset{\sim}{P}_1 = \underset{\sim}{P}_1$, and $\underset{\sim}{P}_{123}\underset{\sim}{P}_{12} = \underset{\sim}{P}_{12}$, and so on. Likewise, (15) is easily extended: e.g., $\underset{\sim}{P}_{3|12}\underset{\sim}{P}_{12} = \underset{\sim}{0}$ and $\underset{\sim}{P}_{4|123}\underset{\sim}{P}_{123} = \underset{\sim}{0}$. But more than this, by using the principle of (14) repeatedly we also have

$$\underset{\sim}{P}_{3|12}\underset{\sim}{P}_1 = (\underset{\sim}{P}_{123} - \underset{\sim}{P}_{12})\underset{\sim}{P}_1 = \underset{\sim}{P}_{123}\underset{\sim}{P}_1 - \underset{\sim}{P}_{12}\underset{\sim}{P}_1 = \underset{\sim}{P}_1 - \underset{\sim}{P}_1 - \underset{\sim}{0} \quad ,$$

$$\begin{aligned} \underset{\sim}{P}_{3|12}\underset{\sim}{P}_{2|1} &= (\underset{\sim}{P}_{123} - \underset{\sim}{P}_{12})(\underset{\sim}{P}_{12} - \underset{\sim}{P}_1) \\ &= \underset{\sim}{P}_{123}\underset{\sim}{P}_{12} - \underset{\sim}{P}_{12}^2 - \underset{\sim}{P}_{123}\underset{\sim}{P}_1 + \underset{\sim}{P}_{12}\underset{\sim}{P}_1 \\ &= \underset{\sim}{P}_{12} - \underset{\sim}{P}_{12} - \underset{\sim}{P}_1 + \underset{\sim}{P}_1 = \underset{\sim}{0} \end{aligned}$$

and

$$\begin{aligned} \underset{\sim}{P}_{4|123}\underset{\sim}{P}_{2|1} &= (\underset{\sim}{P}_{1234} - \underset{\sim}{P}_{123})(\underset{\sim}{P}_{12} - \underset{\sim}{P}_1) \\ &= \underset{\sim}{P}_{1234}\underset{\sim}{P}_{12} - \underset{\sim}{P}_{123}\underset{\sim}{P}_{12} - \underset{\sim}{P}_{1234}\underset{\sim}{P}_1 + \underset{\sim}{P}_{123}\underset{\sim}{P}_1 \\ &= \underset{\sim}{P}_{12} - \underset{\sim}{P}_{12} - \underset{\sim}{P}_1 + \underset{\sim}{P}_1 = \underset{\sim}{0} \quad . \end{aligned} \tag{17}$$

It is null products of this nature which give rise to the independence of sums of squares in certain models.

Another consequence of (7) is that

$$M_{12}X_{12} = 0 \quad \Rightarrow \quad M_{12}X_1 = 0 \quad \text{and} \quad M_{12}X_2 = 0 \tag{18}$$

or, more generally that

$$MX_j = 0 \quad \text{and hence} \quad MP_j = 0 \tag{19}$$

where X_j is any subset of the columns of X for which $M = XX^+$. Results (14)–(19) are the basis for the following theorems concerning χ^2-distributions and independence of quadratic forms in fixed and mixed models.

3. Fixed Effects Models

The dispersion matrix of y in fixed effects models is usually taken as $\sigma^2 I$, i.e., $V = \sigma^2 I$. Using this in Theorems B1 and B2 leads to well-known results stated in the following theorem.

Theorem 1: In fixed effects models, with $\text{var}(y) = \sigma^2 I$,

(a) SSE/σ^2, and $R(\beta_2|\beta_1)/\sigma^2$ for any sub-vectors β_1 and β_2 of β, are χ^2-variables,

(b) SSE and $R(\beta_2|\beta_1)$ are independent,

and

(c) In any sequential fitting of sub-vectors of β, such that sums of squares like $R(\beta_2|\beta_1)$, $R(\beta_3|\beta_1\ \beta_2)$ and $R(\beta_4|\beta_1\ \beta_3)$ are calculated, those sums of squares are independent.

Proof: (a) $SSE/\sigma^2 = y'(M/\sigma^2)y$, and so AV of Theorem B1 is $(M/\sigma^2)\sigma^2 I = M$ which, by (7), is idempotent. Similarly, $R(\beta_2|\beta_1)/\sigma^2 = y'(P_{2|1}/\sigma^2)y$ for which AV of Theorem B1 is $P_{2|1}$ and by (16) this too is idempotent; and idempotency implies the χ^2-distribution.

(b) For SSE and $R(\beta_2|\beta_1)$, the product AVB of Theorem B2 is

$$\mathbf{M}\sigma^2\mathbf{I}\mathbf{P}_{2|1} = \sigma^2(\mathbf{M}\mathbf{P}_{12} - \mathbf{M}\mathbf{P}_1) = \mathbf{0} \quad , \tag{20}$$

on using (13) and (19); this implies independence.

(c) In any sequential fitting of factors (including factors that are interactions of main effects), a typical pair of sums of squares is $\mathbf{y}'\mathbf{P}_{4|123}\mathbf{y}$ and $\mathbf{y}'\mathbf{P}_{2|1}\mathbf{y}$. In one form or another these represent all possible pairs of sums of squares in a sequential fitting of factors. And for Theorem B2 the corresponding $\mathbf{AVB}$ is $\mathbf{P}_{4|123}\sigma^2\mathbf{I}\mathbf{P}_{2|1} = \sigma^2\mathbf{P}_{4|123}\mathbf{P}_{2|1} = \mathbf{0}$, by (17). Thus (c) is proved. Q.E.D.

4. Mixed Models

Partition $\mathbf{X}\boldsymbol{\beta}$ of (1) as

$$\mathbf{X}\boldsymbol{\beta} = [\mathbf{X}_0 \quad \mathbf{Z}] \begin{bmatrix} \mathbf{b} \\ \mathbf{u} \end{bmatrix} \quad .$$

Then take $\mathbf{u}$ to represent the random effects in the model with

$$\mathrm{var}(\mathbf{u}) = \mathbf{D} \quad \text{and} \quad \mathrm{cov}(\mathbf{u}, \mathbf{e}) = \mathbf{0} \quad .$$

(In many applications $\mathbf{D}$ is a block diagonal matrix of matrices $\sigma_i^2\mathbf{I}_{n_i}$ with n_i being the number of levels of the random effect that has variance component σ_i^2.) Thus the model equation is

$$\mathbf{y} = \mathbf{X}_0\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e} \tag{21}$$

where some $\mathbf{X}$'s of a partitioning like (8) constitute $\mathbf{X}_0$ and some constitute $\mathbf{Z}$; i.e., $\mathbf{X} = [\mathbf{X}_0 \quad \mathbf{Z}]$. The dispersion matrix of $\mathbf{y}$ is then

$$\mathrm{var}(\mathbf{y}) = \mathbf{V} = \mathbf{Z}\mathbf{D}\mathbf{Z}' + \sigma^2\mathbf{I} \tag{22}$$

$$= \mathbf{X}\boldsymbol{\Delta}\mathbf{X}' + \sigma^2\mathbf{I}, \quad \text{for } \boldsymbol{\Delta} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \tag{23}$$

Although $\underset{\sim}{V}$ of (23) is not the same as $\sigma^2\underset{\sim}{I}$ used in fixed effects models, the sums of squares SSE and $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)$ from the method of fitting constants used for fixed effects models are often used in analyzing data from mixed models (e.g., analysis of variance of balanced data, and estimation of variance components from unbalanced data using Henderson' Method 3). It is therefore of interest to have the following theorem, analogous to Theorem 1 but applicable to the mixed model.

Theorem 2: In mixed models, with $\underset{\sim}{V} = \underset{\sim}{X}\underset{\sim}{\Delta}\underset{\sim}{X}' + \sigma^2\underset{\sim}{I}$ of (22),

(a) SSE/σ^2 has a χ^2-distribution,

(b) $R(\beta_2|\beta_1)/\lambda$ has a χ^2-distribution if and only if $\underset{\sim}{P}_{2|1}\underset{\sim}{V}\underset{\sim}{P}_{2|1} = \lambda\underset{\sim}{P}_{2|1}$; or equivalently, if and only if $\underset{\sim}{P}_{2|1}\underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{Z}'\underset{\sim}{P}_{2|1} = (\lambda - \sigma^2)\underset{\sim}{P}_{2|1}$,

(c) SSE and $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)$ are independent,

and

(d) sums of squares $R(\underset{\sim}{\beta}_4|\underset{\sim}{\beta}_1,\underset{\sim}{\beta}_2,\underset{\sim}{\beta}_3)$ and $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)$ are independent if and only if $\underset{\sim}{P}_{4|123}\underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{Z}'\underset{\sim}{P}_{2|1} = \underset{\sim}{0}$.

Proof: (a) $SSE/\sigma^2 = \underset{\sim}{y}'(\underset{\sim}{M}/\sigma^2)\underset{\sim}{y}$ and $\underset{\sim}{A}\underset{\sim}{V}$ of Theorem B1 is

$$(\underset{\sim}{M}/\sigma^2)\ \underset{\sim}{X}\Delta\underset{\sim}{X}\ + \sigma^2\underset{\sim}{I}) = \underset{\sim}{M}\ ,$$

because $\underset{\sim}{M}\underset{\sim}{X} = \underset{\sim}{0}$ by (7). Furthermore, $\underset{\sim}{M}$ is idempotent, also by (7), and so SSE/σ^2 has a χ^2-distribution.

(b) $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)/\lambda = \underset{\sim}{y}'(\underset{\sim}{P}_{2|1}/\lambda)\underset{\sim}{y}$ and $\underset{\sim}{A}\underset{\sim}{V}$ of Theorem B1 is $(\underset{\sim}{P}_{2|1}/\lambda)\underset{\sim}{V} = \underset{\sim}{P}_{2|1}\underset{\sim}{V}/\lambda$. This is idempotent if $\underset{\sim}{P}_{2|1}\underset{\sim}{V}\underset{\sim}{P}_{2|1}\underset{\sim}{V} = \lambda\underset{\sim}{P}_{2|1}\underset{\sim}{V}$ in which, because $\underset{\sim}{V}$ as a dispersion matrix is positive semi-definite, there is no great loss of generality in taking it as non-singular. Therefore for nonsingular $\underset{\sim}{V}$ the condition is $\underset{\sim}{P}_{2|1}\underset{\sim}{V}\underset{\sim}{P}_{2|1} = \lambda\underset{\sim}{P}_{2|1}$, and on using (22) and (16) this is $\underset{\sim}{P}_{2|1}\underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{Z}'\underset{\sim}{P}_{2|1} = (\lambda - \sigma^2)\underset{\sim}{P}_{2|1}$.

(c) Theorem B2 applied to SSE and $R(\beta_2|\beta_1)$ has AVB as

$$M(X\Delta X' + \sigma^2 I)P_{2|1} = 0$$

from (7) and (20). Independence follows.

(d) The condition $AVB = 0$ of Theorem B2 is

$$P_{4|123}(X\Delta X' + \sigma^2 I)P_{2|1} = 0$$

which, because of (22) and (17), reduces to

$$P_{4|123}ZDZ'P_{2|1} = 0 \quad . \qquad \text{QED}$$

Corollaries

[1] If in Theorem 2(b), β_1 includes u, i.e., if β includes all the random effects, then from (7) and (14)

$$P_{2|1}Z = P_{12}Z - P_1Z = 0$$

and $R(\beta_2|\beta_1)/\sigma^2$ is a χ^2-variable. This means, in referring to $R(\beta_2|\beta_1)$ as the sum of squares due to β_2 adjusted for β_1, that in a mixed model any sum of squares which is "adjusted for all random effects" is distributed as the σ^2 multiple of a χ^2-variable.

[2] Similarly in Theorem 2(d), if β_1 includes u or if $[\beta_2' \;\; B_3']'$ includes u the condition is satisfied and the sums of squares are independent. This means that in a mixed model two sums of squares are independent if at least one of them is "adjusted for all random effects."

5. Random Models

Almost all random models include a term μ for the overall mean. Generally speaking a random model is therefore a mixed model with a single fixed effect μ, and with X_0 of (21) being

$$\underset{\sim}{X}_0 = \underset{\sim}{1}_N = [1 \ \ 1 \cdots 1]' \tag{24}$$

a vector of ones, N being the number of observations in y . In this framework, Theorem 2 is therefore applicable to random models, and so no special theorem is needed for those models.

6. Example

Consider the 1-way classification with a classes and n_i observations in the i'th class, with $N = n_. = \sum_{i=1}^{a} n_i$. Equivalent forms of the model equation are

$$y_{ij} = \mu + \alpha_i + e_{ij}, \quad \text{for } i = 1, \cdots, a \text{ and } j = 1, \cdots, n_i$$

and

$$\underset{\sim}{y} = \mu \underset{\sim}{1}_N + D\{\underset{\sim}{1}_{n_i}\}\underset{\sim}{\alpha} + \underset{\sim}{e} \quad , \tag{25}$$

where $D\{\underset{\sim}{1}_{n_i}\}$ is a block diagonal matrix of vectors $\underset{\sim}{1}_{n_i}$ for $i = 1, \cdots, a$, and where $\underset{\sim}{1}_{n_i}$ is a vector of n_i ones similar to $\underset{\sim}{1}_N$ of (24).

For the randon model

$$\underset{\sim}{D} = \sigma_\alpha^2 \underset{\sim}{I}_a \quad \text{and } \underset{\sim}{V} = \sigma_\alpha^2 D\{\underset{\sim}{J}_{n_i}\} + \sigma_e^2 \underset{\sim}{I}_N \tag{26}$$

where J_{n_i} is square, of order n_i, with every element unity. For notational convience we write

$$\underset{\sim}{1}_i \equiv \underset{\sim}{1}_{n_i} \quad \text{and} \quad \underset{\sim}{J}_i \equiv \underset{\sim}{J}_{n_i}$$

and then note that

$$\underset{\sim}{J}_i = \underset{\sim}{1}_i \underset{\sim}{1}_i' \quad \text{and} \quad \underset{\sim}{J}_i^2 = n_i \underset{\sim}{J}_i \quad . \tag{27}$$

Furthermore on defining

$$\underset{\sim}{J}_i = \underset{\sim}{J}_i / n_i \quad \text{and} \quad \bar{\underset{\sim}{J}}_N = \underset{\sim}{J}_N / N$$

with

$$\underset{\sim}{J}_i \bar{\underset{\sim}{J}}_i = \underset{\sim}{J}_i \quad \text{and} \quad D\{\bar{J}_i\}\bar{J}_N = \bar{J}_N \tag{28}$$

it is easily verified that for

$$\underset{\sim}{n}' = [n_1\underset{\sim}{1}_1' \;\; n_2\underset{\sim}{1}_2' \cdots n_a\underset{\sim}{1}_a'],\; \underset{\sim}{n}'D\{\bar{\underset{\sim}{J}}_i\} = \underset{\sim}{n}' \quad \text{and} \quad \bar{\underset{\sim}{J}}_N D\{\underset{\sim}{J}_i\} = \underset{\sim}{1}_N\underset{\sim}{n}'/N; \tag{29}$$

and all J-matrices are symmetric.

The sums of squares involved in fitting (25) are

$$R(\mu) = \underset{\sim}{y}'\bar{\underset{\sim}{J}}_N\underset{\sim}{y} \quad \text{and } R(\underset{\sim}{\alpha}|\mu) = \underset{\sim}{y}'[D\{\bar{\underset{\sim}{J}}_i\} - \bar{\underset{\sim}{J}}_N]\underset{\sim}{y} \tag{30}$$

with

$$SSE = \underset{\sim}{y}'[\underset{\sim}{I} - D\{\bar{\underset{\sim}{J}}_i\}]\underset{\sim}{y} \quad . \tag{31}$$

We apply Theorem 2 to these.

(a) SSE/σ^2 is a χ^2-variable.

(b) $R(\alpha|\mu)/\lambda$ is a χ^2-variable if

$$\underset{\sim}{P}_{2|1}\underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{Z}'\underset{\sim}{P}_{2|1} = (\lambda - \sigma^2)\underset{\sim}{P}_{2|1} \quad . \tag{32}$$

Using $D\{\underset{\sim}{J}_i\} - \bar{\underset{\sim}{J}}_N$ from (30) for $\underset{\sim}{P}_{2|1}$, and $D\{\underset{\sim}{1}_i\}$ from (25) for $\underset{\sim}{Z}$ with $\underset{\sim}{D}$ of (26), the left-hand side of (32) is

$$\begin{aligned}
&(D\{\bar{\underset{\sim}{J}}_i\} - \bar{\underset{\sim}{J}}_N)D\{\underset{\sim}{1}_i\}\sigma_\alpha^2 D\{\underset{\sim}{1}_i'\}(D\{D\{\bar{\underset{\sim}{J}}_i\} - \bar{\underset{\sim}{J}}_N)\\
&= \sigma_\alpha^2(D\{\bar{\underset{\sim}{J}}_i\} - \bar{\underset{\sim}{J}}_N)D\{\underset{\sim}{J}_i\}(D\{\bar{\underset{\sim}{J}}_i\} - \bar{\underset{\sim}{J}}_N), \text{ using (27)}\\
&= \sigma_\alpha^2(D\{\bar{\underset{\sim}{J}}_i\} - \underset{\sim}{1}_N\underset{\sim}{n}'/N)(D\{\bar{\underset{\sim}{J}}_i\} - \bar{\underset{\sim}{J}}_N), \text{ using (28) and (29)}\\
&= \sigma_\alpha^2(D\{\bar{\underset{\sim}{J}}_i\} - \underset{\sim}{1}_N\underset{\sim}{n}'/N - \underset{\sim}{n}\underset{\sim}{1}_N/N + \underset{\sim}{J}\Sigma_i n_i^2/N^2) \quad ,
\end{aligned} \tag{33}$$

using (29) again. And the right-hand side of (32) is $(\lambda - \sigma^2)(\underset{\sim}{D}\{\bar{\underset{\sim}{J}}_i\} - \bar{\underset{\sim}{J}}_N)$. Clearly, this does not in general equal (33). Therefore $R(\underset{\sim}{\alpha}|\mu)$ in the random model with unbalanced data (having unequal numbers of observations in the subclasses) does not have a χ^2-distribution. But for balanced data, with $n_i = n$ for all i, we have $n' = n\underset{\sim}{1}_N'$ in (29) and $N = an$ and so (33) becomes

$$\Sigma_\alpha^2(D\{\underset{\sim}{J}_n\} - \underset{\sim}{J}_N/a - \underset{\sim}{J}_N/a + \underset{\sim}{J}_N/a) = \sigma_\alpha^2(D\{\underset{\sim}{J}_n\} - \underset{\sim}{J}_N/a) \tag{34}$$

where $D\{\underset{\sim}{J}_n\}$ is block diagonal with a matrices $\underset{\sim}{J}_n$ on the diagonal. The right-hand side of (32) is

$$(\lambda - \sigma^2)(D\{\bar{\underset{\sim}{J}}_n\} - \underset{\sim}{J}_N/an) = [(\lambda - \sigma^2)/n](D\{\underset{\sim}{J}_n\} - \underset{\sim}{J}_N/a)$$

which equals (34) for

$$(\lambda - \sigma^2)/n = \sigma^2_\alpha, \quad \text{i.e., for } \lambda = n\sigma^2_\alpha + \sigma^2 \quad .$$

Hence with balanced data, $R(\underset{\sim}{\alpha}|\mu)/(n\sigma^2_\alpha + \sigma^2)$ has a χ^2-distribution, as is well known. But as already shown, with unbalanced data, $R(\underset{\sim}{\alpha}|\mu)$ does not have a χ^2-distribution.

(c) SSE and $R(\underset{\sim}{\alpha}|\mu)$ are independent.

(d) $R(\mu)$ and $R(\underset{\sim}{\alpha}|\mu)$ are independent if and only if the following product is null:

$$(D\{\underset{\sim}{J}_i\} - \bar{\underset{\sim}{J}}_N)D\{\underset{\sim}{\mathbf{1}}_i\}\sigma^2_\alpha D\{\underset{\sim}{\mathbf{1}}_i'\}\bar{\underset{\sim}{J}}_N = \sigma^2_\alpha(D\{\bar{\underset{\sim}{J}}_i\} - \bar{\underset{\sim}{J}}_N)D\{\underset{\sim}{J}_i\}\bar{\underset{\sim}{J}}_N$$

$$= \sigma^2_\alpha(\underset{\sim}{n}\underset{\sim}{\mathbf{1}}_N'/N - \underset{\sim}{J}_N\Sigma_i n_i^2/N^2) \quad . \qquad (35)$$

In general, this is clearly not null, a typical element being $n_i/N - \Sigma_i n_i^2/N^2$. Therefore for unbalanced data $R(\mu) = N\bar{y}^2_{..}$ and $R(\underset{\sim}{\alpha}|\mu)$ are not independent. But for balanced data each element of (35) is zero, i.e., (35) is null, and so $R(\mu)$ and $R(\underset{\sim}{\alpha}|\mu)$ are then independent.

References

Searle, S. R. (1971). Linear Models. Wiley, New York.

Searle, S. R. (1982). Matrix Algebra Useful for Statistics, Wiley, New York.

③

STATISTICS 416/417 MATRIX ALGEBRA Fall, 1983

Course Organization

416 MATRIX ALGEBRA I.

Dates: August 31 – October 21

Lectures: M,W,F: 8 – 8:50 a.m., Warren 345

Exams ($1\frac{1}{2}$ hours, closed book)

Prelim: Thurs., September 29, 7:00 – 8:30 p.m., Warren 101 and 201

Final: Thurs., October 20, 7:00 – 8:30 p.m., Warren 101 and 201

Auditing: Auditing is allowed only with instructor's permission.

Dropping: The course cannot be dropped after September 21, 1983.

417 MATRIX ALGEBRA II.

Dates: October 24 – December 9

Lectures: M,W,F: 8 – 8:50 a.m., Warren 345

Exams

Prelim: Thurs., November 17, 7:00 – 8:30 p.m., Warren 201

Final: $2\frac{1}{2}$ hours, open book, as scheduled during exam week, December 16–22

Auditing: No auditors permitted.

Dropping: The course cannot be dropped after November 9, 1983.

FOR EACH COURSE:

Instructor: S. R. Searle, Warren 339

Office hours: 9–10 a.m. and 4–5 p.m., Monday, or by appointment.

Assistant: Walter Piegorsch, Emerson 262, phone 6-4498

Text: MATRIX ALGEBRA USEFUL FOR STATISTICS, S. R. Searle, Wiley, 1982.

Homework Assignments: Assignments will be given every Wednesday. They are to be returned one week later. Each week's assignment will be graded 2, 1, or 0, representing, in an approximate manner, that an assignment is essentially correct, has serious deficiencies, or is either late or mostly wrong. In each course, all assignments must be handed in in order to get a grade other than F.

Discussion Period:

Monday, 1:25 – 3:30 p.m., Warren 245 (245, not 345), starting September 12. Homework assignments will be discussed and assistance offered; there will be no lecture. All activity will center on questions asked by students.

Composition of Final Grade Assessment:

Homework, 10%; Preliminary Exam, 40%; Final Exam, 50%.

Reading Assignment: To begin the semester, read hand-outs ④ and ⑤, namely "Proof", and its Appendix.