NOTES ON CALCULATING SUMS OF SQUARES IN ANALYSES OF VARIANCE OF UNBALANCED DATA

BU-724-M

by

November, 1980

S. R. Searle
Biometrics Unit, Cornell University, Ithaca, New York

ABSTRACT

Five different methods of calculating sums of squares used in statistical computing packages are outlined and examples given. Presentation is in the form of notes prepared for the session entitled "Tutorial on Analysis of Variance for Unbalanced Data Using Computer Routines SAS GLM and BMDP2V" at the 36th Annual Conference on Applied Statistics, Newark, New Jersey. These notes also represent a summary of "Some computational and model equivalences in analyses of variance of unequal-subclass-numbers data" by Searle, Speed and Henderson, The American Statistician, 1981.

MODELS

A general model: $E(\underset{\sim}{y}) = \underset{\sim\sim}{Xb}$.

Two specific models: two-factors (rows and columns)

(i) without interaction: $E(y_{ijk}) = \mu + \alpha_i + \beta_j$

(ii) with interaction: $E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$

$i = 1, \cdots, a, \quad j = 1, \cdots, b, \quad k = 1, \cdots, n_{ij}$ with $n_{ij} \geq 0$.

FIVE METHODS OF CALCULATING SUMS OF SQUARES

1. Reductions due to sub-models: $R(\underset{\sim}{b}|\underset{\sim}{b}_2) = R(\underset{\sim}{b}_1, \underset{\sim}{b}_2) - R(\underset{\sim}{b}_2)$.

2. Full rank reparameterization — often with Σ-restrictions.

3. Weighted squares of mean — SSA_w and SSB_w.

4. Numerator of the F-statistic for testing a hypothesis.

5. An "indirect" method of calculation for full rank models.

AN ILLUSTRATIVE EXAMPLE

	y_{ijk}			$y_{i..}$
	7,9	6	2	24
	8	4,8	12	32
$y_{.j.}$	24	18	14	$56 = y_{...}$

	n_{ij}			$n_{i.}$
	2	1	1	4
	1	2	1	4
$n_{.j}$	3	3	2	$8 = n_{..}$

METHOD 1: $R(\underset{\sim}{b}_1|\underset{\sim}{b}_2)$

Full model

$$E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{b}, \text{ with normal equations } \underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{b}^o = \underset{\sim}{X}'\underset{\sim}{y} .$$

$$R(\underset{\sim}{b}) = \underset{\sim}{b}^{o'}(\underset{\sim}{X}'\underset{\sim}{y}) = \Sigma\begin{pmatrix} \text{each element of solution vector} \\ \times \text{ corresponding r.h.s. of normal equations} \end{pmatrix} .$$

We call this the R-algorithm.

Sub model

$$E(\underset{\sim}{y}) = \underset{\sim}{X}_1\underset{\sim}{b}_1 + \underset{\sim}{X}_2\underset{\sim}{b}_2 .$$

$$R(\underset{\sim}{b}_1|\underset{\sim}{b}_2) = R(\underset{\sim}{b}_1,\underset{\sim}{b}_2) - R(\underset{\sim}{b}_2) .$$

LM 444
("Linear Models", p. 444)

Example 1: $R(\mu)$

Model: $E(y_{ijk}) = \mu$

Normal equations: $8\hat{\mu} = 56$ Solution: $\hat{\mu} = 7$

Sum of squares: $R(\mu) = 7(56) = 392$.

Example 2: $R(\mu,\alpha)$

Model: $E(y_{ijk}) = \mu + \alpha_i$.

Normal Equations (LM 232)

$$\begin{aligned} 8\mu^o + 4\alpha_1^o + 4\alpha_2^o &= 56 \\ 4\mu^o + 4\alpha_1^o \qquad\quad &= 24 \\ 4\mu^o \qquad\quad + 4\alpha_2^o &= 32 \end{aligned}$$

A Solution (LM 233)

$$\begin{bmatrix} \mu^o \\ \alpha_1^o \\ \alpha_2^o \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \\ 8 \end{bmatrix}$$

Sum of squares: $R(\mu,\alpha) = 0(56) + 6(24) + 8(32) = 400$.

Example 3: $R(\alpha|\mu)$

$$R(\alpha|\mu) = R(\mu,\alpha) - R(\mu) = 400 - 392 = 8 .$$

Example 4: $R(\mu,\alpha,\beta)$

Model: $y_{ijk} = \mu + \alpha_i + \beta_j$

Normal Equations (LM 264)

$$\left[\begin{array}{c|cc|ccc} 8 & 4 & 4 & 3 & 3 & 2 \\ \hline 4 & 4 & . & 2 & 1 & 1 \\ 4 & . & 4 & 1 & 2 & 1 \\ \hline 3 & 2 & 1 & 3 & . & . \\ 3 & 1 & 2 & . & 3 & . \\ 2 & 1 & 1 & . & . & 2 \end{array}\right] \left[\begin{array}{c} \mu^o \\ \hline \alpha_1^o \\ \alpha_2^o \\ \hline \beta_1^o \\ \beta_2^o \\ \beta_3^o \end{array}\right] = \left[\begin{array}{c} 56 \\ \hline 24 \\ 32 \\ \hline 24 \\ 18 \\ 14 \end{array}\right]$$

A Solution (LM 264)

$$\begin{bmatrix} \mu^o \\ \alpha_1^o \\ \alpha_2^o \\ \beta_1^o \\ \beta_2^o \\ \beta_3^o \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 0 \\ 62 \\ 92 \\ 16 \\ -16 \\ 0 \end{bmatrix}$$

Sum of squares:

$$R(\mu,\alpha,\beta) = \frac{1}{11}\left[0(56) + 62(24) + 92(32) + 16(24) - 16(18) + 0(14)\right] = 411\frac{7}{11}$$

Example 5: $R(\beta|\mu,\alpha)$

$$R(\beta|\mu,\alpha) = R(\mu,\alpha,\beta) - R(\mu,\alpha) = 411\frac{7}{11} - 400 = 11\frac{7}{11} .$$

Example 6: $R(\mu,\alpha,\beta,\gamma)$

Model: $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij}$

Normal Equations (LM 288) A Solution (LM 291)

$$
\left[\begin{array}{c|cc|ccc|cccccc}
8 & 4 & 4 & 3 & 3 & 2 & 2 & 1 & 1 & 1 & 2 & 1 \\
\hline
4 & 4 & \cdot & 2 & 1 & 1 & 2 & 1 & 1 & \cdot & \cdot & \cdot \\
4 & \cdot & 4 & 1 & 2 & 1 & \cdot & \cdot & \cdot & 1 & 2 & 1 \\
\hline
3 & 2 & 1 & 3 & \cdot & \cdot & 2 & \cdot & \cdot & 1 & \cdot & \cdot \\
3 & 1 & 2 & \cdot & 3 & \cdot & \cdot & 1 & \cdot & \cdot & 2 & \cdot \\
2 & 1 & 1 & \cdot & \cdot & 2 & \cdot & \cdot & 1 & \cdot & \cdot & 1 \\
\hline
2 & 2 & \cdot & 2 & \cdot & \cdot & 2 & \cdot & \cdot & \cdot & \cdot & \cdot \\
1 & 1 & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\
1 & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\
1 & \cdot & 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\
2 & \cdot & 2 & \cdot & 2 & \cdot & \cdot & \cdot & \cdot & \cdot & 2 & \cdot \\
1 & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1
\end{array}\right]
\left[\begin{array}{c}
\mu^o \\ \hline \alpha_1^o \\ \alpha_2^o \\ \hline \beta_1^o \\ \beta_2^o \\ \beta_3^o \\ \hline \gamma_{11}^o \\ \gamma_{12}^o \\ \gamma_{13}^o \\ \gamma_{21}^o \\ \gamma_{22}^o \\ \gamma_{23}^o
\end{array}\right]
=
\left[\begin{array}{c}
56 \\ \hline 24 \\ 32 \\ \hline 24 \\ 18 \\ 14 \\ \hline 16 \\ 6 \\ 2 \\ 8 \\ 12 \\ 12
\end{array}\right]
\qquad
\left[\begin{array}{c}
\mu^o \\ \alpha_1^o \\ \alpha_2^o \\ \beta_1^o \\ \beta_2^o \\ \beta_3^o \\ \gamma_{11}^o \\ \gamma_{12}^o \\ \gamma_{13}^o \\ \gamma_{21}^o \\ \gamma_{22}^o \\ \gamma_{23}^o
\end{array}\right]
=
\left[\begin{array}{c}
0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 8 \\ 6 \\ 2 \\ 8 \\ 6 \\ 12
\end{array}\right]
$$

Sum of squares:

$$R(\mu,\alpha,\beta,\gamma) = \left[0(56) + \cdots + 0(14) + 8(16) + 6(6) + 2(2) + 8(8) + 6(12) + 12(12)\right] = 448 \; .$$

Example 7: $R(\gamma|\mu,\alpha,\beta)$

$$R(\gamma|\mu,\alpha,\beta) = R(\mu,\alpha,\beta,\gamma) - R(\mu,\alpha,\beta) = 448 - 411\tfrac{7}{11} = 36\tfrac{4}{11} \; .$$

PARTITIONINGS OF TOTAL SUMS OF SQUARES

(a) Rows before Columns			(b) Columns before Rows		
Term	d.f.	Sum of Squares	Term	d.f.	Sum of Squares
$R(\mu)$	1	392	$R(\mu)$	1	392
$R(\alpha\|\mu)$	1	8	$R(\beta\|\mu)$	2	6
$R(\beta\|\mu,\alpha)$	2	$11\frac{7}{11}$	$R(\alpha\|\mu,\beta)$	1	$13\frac{7}{11}$
$R(\gamma\|\mu,\alpha,\beta)$	2	$36\frac{4}{11}$	$R(\gamma\|\mu,\alpha,\beta)$	2	$36\frac{4}{11}$
SSE	2	10	SSE	2	10
SST	8	458	SST	8	458

SAS GLM

Order of reading factors: α, β, γ

Type I	Type II
$R(\alpha\|\mu)$	$R(\alpha\|\mu,\beta)$
$R(\beta\|\mu,\alpha)$	$R(\beta\|\mu,\alpha)$
$R(\gamma\|\mu,\alpha,\beta)$	$R(\gamma\|\mu,\alpha,\beta)$

Example 8: $R(\mu,\beta,\gamma)$

Model: $y_{ijk} = \mu + \beta_j + \gamma_{ij}$

Normal Equations (LM 250) A Solution (LM 251)

$$
\left[\begin{array}{c|ccc|cccccc}
8 & 3 & 3 & 2 & 2 & 1 & 1 & 1 & 2 & 1 \\
\hline
3 & 3 & . & . & 2 & . & . & 1 & . & . \\
3 & . & 3 & . & . & 1 & . & . & 2 & . \\
2 & . & . & 2 & . & . & 1 & . & . & 1 \\
\hline
2 & 2 & . & . & 2 & . & . & . & . & . \\
1 & . & 1 & . & . & 1 & . & . & . & . \\
1 & . & . & 1 & . & . & 1 & . & . & . \\
1 & 1 & . & . & . & . & . & 1 & . & . \\
2 & . & 2 & . & . & . & . & . & 2 & . \\
1 & . & . & 1 & . & . & . & . & . & 1
\end{array}\right]
\left[\begin{array}{c}
\mu^o \\ \hline \beta_1^o \\ \beta_2^o \\ \beta_3^o \\ \hline \gamma_{11}^o \\ \gamma_{12}^o \\ \gamma_{13}^o \\ \gamma_{21}^o \\ \gamma_{22}^o \\ \gamma_{23}^o
\end{array}\right]
=
\left[\begin{array}{c}
56 \\ \hline 24 \\ 18 \\ 14 \\ \hline 16 \\ 6 \\ 2 \\ 8 \\ 12 \\ 12
\end{array}\right]
\qquad
\left[\begin{array}{c}
\mu^o \\ \beta_1^o \\ \beta_2^o \\ \beta_3^o \\ \gamma_{11}^o \\ \gamma_{12}^o \\ \gamma_{13}^o \\ \gamma_{21}^o \\ \gamma_{22}^o \\ \gamma_{23}^o
\end{array}\right]
=
\left[\begin{array}{c}
0 \\ 0 \\ 0 \\ 0 \\ 8 \\ 6 \\ 2 \\ 8 \\ 6 \\ 12
\end{array}\right]
$$

Sum of squares:

$$R(\mu,\beta,\gamma) = \left[0(56) + \cdots + 0(14) + 8(16) + 6(6) + 2(2) + 8(8) + 6(12) + 12(12)\right] = 448 .$$

Example 9: $R(\alpha|\mu,\beta,\gamma)$

Pro forma: $R(\alpha|\mu,\beta,\gamma) = R(\mu,\alpha,\beta,\gamma) - R(\mu,\beta,\gamma)$

$= 448 - 448$

$= 0 .$

METHOD 2: FULL RANK REPARAMETERIZATION

"Usual restrictions" ≡ Σ-restrictions: e.g., $\sum_{i=1}^{a} \alpha_i = 0$.

Use 1: Computational convenience for solving normal equations

Example 2: $R(\mu,\alpha)$

Normal Equations

$$8\mu^o + 4\alpha_1^o + 4\alpha_2^o = 56$$

$$4\mu^o + 4\alpha_1^o = 24$$

$$4\mu^o + 4\alpha_2^o = 32$$

$$\alpha_1^o + \alpha_2^o \Rightarrow$$

Solution

$$\begin{bmatrix} \mu^o \\ \alpha_1^o \\ \alpha_2^o \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \\ 1 \end{bmatrix}$$

Sum of squares:

$$R(\mu,\alpha) = 7(56) - 1(24) + 1(32) = 400$$ (see page 3).

Example 4: $R(\mu,\alpha,\beta)$

$$\alpha_1^o + \alpha_2^o = 0 \quad \text{and} \quad \beta_1^o + \beta_2^o + \beta_3^o = 0$$

Normal Equations

$$\begin{bmatrix} 8 & 0 & 1 & 1 \\ 0 & 8 & 1 & -1 \\ 1 & 1 & 5 & 2 \\ 1 & -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} \mu^o \\ \alpha_1^o \\ \beta_1^o \\ \beta_2^o \end{bmatrix} = \begin{bmatrix} 56 \\ -8 \\ 10 \\ 4 \end{bmatrix}$$

Solution

$$\begin{bmatrix} 7 \\ -15/11 \\ 16/11 \\ -16/11 \end{bmatrix}$$

Sum of squares:

$$R(\mu,\alpha,\beta) = 7(56) + \frac{1}{11}\left[-15(-8) + 16(10) - 16(4)\right] = 411\frac{7}{11}$$ (see page 4).

Use 2: As part of the model (restricted models)

Example 4r: $R^{*}(\dot{\mu},\dot{\alpha},\dot{\beta})_{\Sigma}$

Model: $E(y_{ijk}) = \dot{\mu} + \dot{\alpha}_i + \dot{\beta}_j$ with $\Sigma\dot{\alpha}_i = 0$ and $\Sigma\dot{\beta}_j = 0$

Note: Dots above symbols emphasize restricted model.

Normal Equations | Solution

$$\begin{bmatrix} 8 & 0 & 1 & 1 \\ 0 & 8 & 1 & -1 \\ 1 & 1 & 5 & 2 \\ 1 & -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} \hat{\dot{\mu}} \\ \hat{\dot{\alpha}}_1 \\ \hat{\dot{\beta}}_1 \\ \hat{\dot{\beta}}_2 \end{bmatrix} = \begin{bmatrix} 56 \\ -8 \\ 10 \\ 4 \end{bmatrix} \qquad \begin{bmatrix} \hat{\dot{\mu}} \\ \hat{\dot{\alpha}}_1 \\ \hat{\dot{\beta}}_1 \\ \hat{\dot{\beta}}_2 \end{bmatrix} = \begin{bmatrix} 7 \\ -\frac{15}{11} \\ \frac{16}{11} \\ -\frac{16}{11} \end{bmatrix} .$$

Same equations and solution as on page 8.

Sum of squares:

$$R^{*}(\underbrace{\dot{\mu},\dot{\alpha},\dot{\beta}})_{\Sigma} = 7(56) + \frac{1}{11}\left[-15(-8) + 16(10) - 16(4)\right] = 411\frac{7}{11} .$$

↙ See p.10 ↓ restricted model ↘ Σ-restrictions

$R^{*}(\dot{\mu},\dot{\alpha},\dot{\beta})_{\Sigma} = R(\mu,\alpha,\beta)$ when $(\dot{\mu},\dot{\alpha},\dot{\beta})$ and (μ,α,β) represent a full model, not a sub-model.

FITTING SUB-MODELS OF A RESTRICTED MODEL

Example 1r: $R(\dot{\mu}|\dot{\alpha},\dot{\beta})_\Sigma$

$$R^{\#}(\dot{\mu}|\dot{\alpha},\dot{\beta})_\Sigma = R^{\#}(\dot{\mu},\dot{\alpha},\dot{\beta})_\Sigma - R^{\#}(\dot{\alpha},\dot{\beta})_\Sigma$$

$^{\#}$ indicates $R^{\#}$ is a sum of squares for restricted model and/or sub-model thereof, in this case the latter

$$R^{\#}(\dot{\mu}|\dot{\alpha},\dot{\beta})_\Sigma = 411\tfrac{7}{11} - R^{\#}(\dot{\alpha},\dot{\beta})_\Sigma \ .$$

$R^{\#}(\dot{\alpha},\dot{\beta})_\Sigma$ is for the $(\dot{\alpha},\dot{\beta})$-sub-model of the Σ-restricted model $E(y_{ijk}) = \dot{\mu} + \dot{\alpha}_i + \dot{\beta}_j$ with $\Sigma\dot{\alpha}_i = 0$ and $\Sigma\dot{\beta}_j = 0$.

This implies deleting $\dot{\mu}$ and the $\dot{\mu}$-equation from the normal equations for that restricted model (see page 9):

$$\begin{bmatrix} 8 & 1 & -1 \\ 1 & 5 & 2 \\ -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} \tilde{\dot{\alpha}}_1 \\ \tilde{\dot{\beta}}_1 \\ \tilde{\dot{\beta}}_2 \end{bmatrix} = \begin{bmatrix} -8 \\ 10 \\ 4 \end{bmatrix} \quad \text{with solution} \quad \begin{bmatrix} \tilde{\dot{\alpha}}_1 \\ \tilde{\dot{\beta}}_1 \\ \tilde{\dot{\beta}}_2 \end{bmatrix} = \begin{bmatrix} -15/11 \\ 27/11 \\ -5/11 \end{bmatrix} .$$

The R-algorithm applied to this gives

$$R^{\#}(\dot{\alpha},\dot{\beta})_\Sigma = \frac{-15}{11}(-8) + \frac{27}{11}(10) + \frac{-5}{11}(4) = 33\,\frac{7}{11}$$

and so

$$R^{\#}(\dot{\mu}|\dot{\alpha},\dot{\beta})_\Sigma = 411\tfrac{7}{11} - 33\tfrac{7}{11} = 378 \ .$$

NOTE 1: This is not $R(\mu) = 392$ of page 3.

NOTE 2: $R^{\#}(\dot{\alpha},\dot{\beta})_{\Sigma}$ is not the same as $R(\alpha,\beta)$ that comes from the (α,β)-sub-model of the unrestricted model $E(y_{ijk}) = \alpha_i + \beta_j$ using either $\Sigma\alpha_i^o = 0$ or $\Sigma\beta_j^o = 0$ (but not both) for computational convenience. That leads to $R(\alpha,\beta) = R(\mu,\alpha,\beta)$. With $\alpha_1^o + \alpha_2^o = 0$, the calculations are as follows.

Normal Equations

$$\begin{bmatrix} 4 & 2 & 1 & 1 \\ -4 & 1 & 2 & 1 \\ 1 & 3 & \cdot & \cdot \\ -1 & \cdot & 3 & \cdot \\ \cdot & \cdot & \cdot & 2 \end{bmatrix} \begin{bmatrix} \alpha_1^o \\ \beta_1^o \\ \beta_2^o \\ \beta_3^o \end{bmatrix} = \begin{bmatrix} 24 \\ 32 \\ 24 \\ 18 \\ 14 \end{bmatrix}$$

Solution

$$\begin{bmatrix} \alpha_1^o \\ \alpha_2^o \\ \beta_1^o \\ \beta_2^o \\ \beta_3^o \end{bmatrix} = \begin{bmatrix} -15/11 \\ 15/11 \\ 93/11 \\ 61/11 \\ 7 \end{bmatrix}$$

and

$$\alpha_1^o + \alpha_2^o = 0 .$$

The R-algorithm gives

$$R(\alpha^o,\beta^o) = \frac{1}{11}\Big[-15(24) + 15(32) + 93(24) + 61(18)\Big] + 7(14)$$

$$= 411\frac{7}{11}$$

$$= R(\mu,\alpha,\beta) .$$

INTERACTION MODEL WITH Σ-RESTRICTIONS

Σ-restrictions

$$
\left.\begin{array}{l}
\dot{\alpha}_1 + \dot{\alpha}_2 = 0 \\
\dot{\beta}_1 + \dot{\beta}_2 + \dot{\beta}_3 = 0
\end{array}\right.
\quad
\begin{array}{l}
\Rightarrow \dot{\alpha}_2 = -\dot{\alpha}_1 \\
\Rightarrow \dot{\beta}_3 = -\dot{\beta}_1 - \dot{\beta}_2
\end{array}
$$

$$
\left.\begin{array}{r}
\dot{\gamma}_{11} + \dot{\gamma}_{12} + \dot{\gamma}_{13} = 0 \\
\dot{\gamma}_{21} + \dot{\gamma}_{22} + \dot{\gamma}_{23} = 0 \\
\dot{\gamma}_{11} + \dot{\gamma}_{21} = 0 \\
\dot{\gamma}_{12} + \dot{\gamma}_{22} = 0 \\
\dot{\gamma}_{13} + \dot{\gamma}_{23} = 0
\end{array}\right\}
\Rightarrow
\begin{array}{l}
\dot{\gamma}_{11} = \dot{\gamma}_{11} \\
\dot{\gamma}_{12} = \dot{\gamma}_{12} \\
\dot{\gamma}_{13} = -\dot{\gamma}_{11} - \dot{\gamma}_{12} \\
\dot{\gamma}_{21} = -\dot{\gamma}_{11} \\
\dot{\gamma}_{22} = -\dot{\gamma}_{12} \\
\dot{\gamma}_{23} = \dot{\gamma}_{11} + \dot{\gamma}_{12}
\end{array}
$$

Model Equations

$$
E\begin{bmatrix} 7 \\ 9 \\ 6 \\ 2 \\ 8 \\ 4 \\ 8 \\ 12 \end{bmatrix}
=
\begin{bmatrix}
1 & 1 & 1 & . & 1 & . \\
1 & 1 & 1 & . & 1 & . \\
1 & 1 & . & 1 & . & 1 \\
1 & 1 & -1 & -1 & -1 & -1 \\
1 & -1 & 1 & . & -1 & . \\
1 & -1 & . & 1 & . & -1 \\
1 & -1 & . & 1 & . & -1 \\
1 & -1 & -1 & -1 & 1 & 1
\end{bmatrix}
\begin{bmatrix} \dot{\mu} \\ \dot{\alpha}_1 \\ \dot{\beta}_1 \\ \dot{\beta}_2 \\ \dot{\gamma}_{11} \\ \dot{\gamma}_{12} \end{bmatrix}
$$

Normal Equations

$$
\left[\begin{array}{c:c:cc:cc}
8 & 0 & 1 & 1 & 1 & -1 \\ \hdashline
0 & 8 & 1 & -1 & 1 & 1 \\ \hdashline
1 & 1 & 5 & 2 & 1 & 0 \\
1 & -1 & 2 & 5 & 0 & -1 \\ \hdashline
1 & 1 & 1 & 0 & 5 & 2 \\
-1 & 1 & 0 & -1 & 2 & 5
\end{array}\right]
\left[\begin{array}{c}
\hat{\dot{\mu}} \\ \hdashline \hat{\dot{\alpha}}_1 \\ \hdashline \hat{\dot{\beta}}_1 \\ \hat{\dot{\beta}}_2 \\ \hdashline \hat{\dot{\gamma}}_{11} \\ \hat{\dot{\gamma}}_{12}
\end{array}\right]
=
\left[\begin{array}{c}
56 \\ \hdashline -8 \\ \hdashline 10 \\ 4 \\ \hdashline 18 \\ 4
\end{array}\right]
$$

Solution

$$
\begin{bmatrix} \hat{\dot{\mu}} \\ \hat{\dot{\alpha}}_1 \\ \hat{\dot{\beta}}_1 \\ \hat{\dot{\beta}}_2 \\ \hat{\dot{\gamma}}_{11} \\ \hat{\dot{\gamma}}_{12} \end{bmatrix}
= \frac{1}{72}
\begin{bmatrix}
10 & 0 & -1 & -1 & -3 & 3 \\
0 & 10 & -3 & 3 & -1 & -1 \\
-1 & -3 & 19 & -8 & -3 & 0 \\
-1 & 3 & -8 & 19 & 0 & 3 \\
-3 & -1 & -3 & 0 & 19 & -8 \\
3 & -1 & 0 & 3 & -8 & 19
\end{bmatrix}
\begin{bmatrix} 56 \\ -8 \\ 10 \\ 4 \\ 18 \\ 4 \end{bmatrix}
=
\begin{bmatrix} 7 \\ -10/6 \\ 1 \\ -1 \\ 10/6 \\ 10/6 \end{bmatrix}
$$

<u>Example 6r</u>: $R^*(\dot{\mu},\dot{\alpha},\dot{\beta},\dot{\gamma})_\Sigma$

$$R^*(\dot{\mu},\dot{\alpha},\dot{\beta},\dot{\gamma})_\Sigma = \frac{1}{72}\left[7(56) - (10/6)(-8) + 1(10) - 1(4) + (10/6)(18+4)\right]$$

$$= 448 = R(\mu,\alpha,\beta,\gamma) \text{ of page 5}$$

and $(\mu,\alpha,\beta,\gamma)$ is the full model.

<u>Example 1r</u>: $R^*(\dot{\mu}|\dot{\alpha},\dot{\beta},\dot{\gamma})_\Sigma$ See pages 3, 10 and 12.

$$R^*(\dot{\mu}|\dot{\alpha},\dot{\beta},\dot{\gamma})_\Sigma = R^*(\dot{\mu},\dot{\alpha},\dot{\beta},\dot{\gamma})_\Sigma - R^*(\dot{\alpha},\dot{\beta},\dot{\gamma})_\Sigma$$

$R^*(\dot{\alpha},\dot{\beta},\dot{\gamma})_\Sigma$ comes from deleting $\dot{\mu}$ and the $\dot{\mu}$-equation from the normal equations on page 12:

$$\begin{bmatrix} 8 & 1 & -1 & 1 & 1 \\ 1 & 5 & 2 & 1 & 0 \\ -1 & 2 & 5 & 0 & -1 \\ 1 & 1 & 0 & 5 & 2 \\ 1 & 0 & -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} \tilde{\dot{\alpha}}_1 \\ \tilde{\dot{\beta}}_1 \\ \tilde{\dot{\beta}}_2 \\ \tilde{\dot{\gamma}}_{11} \\ \tilde{\dot{\gamma}}_{12} \end{bmatrix} = \begin{bmatrix} -8 \\ 10 \\ 4 \\ 18 \\ 4 \end{bmatrix} \quad \text{with solution} \quad \frac{1}{30}\begin{bmatrix} -50 \\ 51 \\ -9 \\ 113 \\ -13 \end{bmatrix} .$$

The R-algorithm applied to this gives

$$R^*(\dot{\alpha},\dot{\beta},\dot{\gamma})_\Sigma = \frac{1}{30}\left[-50(-8) + 51(10) - 9(4) + 113(18) - 13(4)\right] = 95.2$$

and so

$$R^*(\dot{\mu}|\dot{\alpha},\dot{\beta},\dot{\gamma})_\Sigma = 448 - 95.2 = 352.8 . \qquad \text{[BMDP2V]}$$

<u>NOTE</u>: $R^*(\dot{\mu}|\dot{\alpha},\dot{\beta},\dot{\gamma})_\Sigma = 352.8$ in the with-interaction model

$R^*(\dot{\mu}|\dot{\alpha},\dot{\beta})_\Sigma = 378$ in the no-interaction model of page 10

$R(\mu) = 392$ of page 3.

<u>Example 9r</u>: $R^{*}(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_{\Sigma}$

$$R^{*}(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_{\Sigma} = R^{*}(\dot{\mu},\dot{\alpha},\dot{\beta},\dot{\gamma})_{\Sigma} - R^{*}(\dot{\mu},\dot{\beta},\dot{\gamma})_{\Sigma}$$

$$= 448 - R^{*}(\dot{\mu},\dot{\beta},\dot{\gamma})_{\Sigma} \quad \text{(See page 13.)}$$

$R^{*}(\dot{\mu},\dot{\beta},\dot{\gamma})_{\Sigma}$ comes from deleting $\dot{\alpha}_1$ and the $\dot{\alpha}_1$-equation from the normal equations on page 12:

$$\begin{bmatrix} 8 & 1 & 1 & 1 & -1 \\ 1 & 5 & 2 & 1 & 0 \\ 1 & 2 & 5 & 0 & 1 \\ 1 & 1 & 0 & 5 & 2 \\ -1 & 0 & 1 & 2 & 5 \end{bmatrix} \begin{bmatrix} \tilde{\dot{\mu}} \\ \tilde{\dot{\beta}}_1 \\ \tilde{\dot{\beta}}_2 \\ \tilde{\dot{\gamma}}_{11} \\ \tilde{\dot{\gamma}}_{12} \end{bmatrix} = \begin{bmatrix} 56 \\ 10 \\ 4 \\ 18 \\ 4 \end{bmatrix} \quad \text{with solution} \quad \begin{bmatrix} 7 \\ \frac{1}{2} \\ -\frac{1}{2} \\ 1\frac{1}{2} \\ 1\frac{1}{2} \end{bmatrix}.$$

The R-algorithm applied to this gives

$$R^{*}(\dot{\mu},\dot{\beta},\dot{\gamma})_{\Sigma} = 7(56) + \tfrac{1}{2}(10) - \tfrac{1}{2}(4) + 1\tfrac{1}{2}(18) + 1\tfrac{1}{2}(4) = 428$$

and so

$$R^{*}(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_{\Sigma} = 448 - 428 = 20$$

[SAS GLM Type III and BMDP2V]

<u>NOTE</u>: $R(\alpha|\mu,\beta,\gamma) \equiv 0$ on page 7

$R^{*}(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_{\Sigma} = 20 \neq 0$ here.

METHOD 3: WEIGHTED SQUARES OF MEANS ANALYSIS (LM 369)

Only for all cells filled.

Define: $x_{ij} = \bar{y}_{ij.}$, $\quad w_i = b^2/\sum_j(1/n_{ij})$, $\quad$ and $\quad \bar{x}_{[1]} = \Sigma w_i \bar{x}_{i.}/\Sigma w_i$.

$$SSA_w = \sum_{i=1}^{a} w_i(\bar{x}_{i.} - \bar{x}_{[1]})^2$$ = sum of squares for rows in weighted squares of means analysis.

Example:

x_{ij}			$\bar{x}_{i.}$
8	6	2	16/3
8	6	12	26/3

n_{ij}		
2	1	1
1	2	2

$w_1 = 9/(\frac{1}{2} + 1 + 1) = 18/5$

$w_2 = 9/(\frac{1}{2} + 1 + 1) = 18/5$

$$x_{[1]} = \tfrac{1}{2}(16/3 + 26/3) = 7$$

$$SSA_w = \frac{18}{5}\left[(16/3 - 7)^2 + (26/3 - 7)^2\right] = 20 .$$

NOTE: $R^*(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma = 20$, on page 14.

For the all-cells-filled case

$R^*(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma \equiv SSA_w$ is an identity.

[See The American Statistician, February, 1981.]

METHOD 4: NUMERATOR OF AN F-STATISTIC (LM 190)

$H: \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m}$, $\quad \underset{\sim}{K}'\underset{\sim}{b}$ estimable, $\quad \underset{\sim}{K}'$ full row rank, s say.

$$F = \frac{Q}{s\hat{\sigma}^2}, \text{ with } Q = (\underset{\sim}{K}'\underset{\sim}{b}^{o} - \underset{\sim}{m})[\underset{\sim}{K}'(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{K}]^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^{o} - \underset{\sim}{m}), \text{ for } \underset{\sim}{b}^{o} = (\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{y} .$$

Example:

The hypothesis

$$H: 6\mu + 3(\alpha_1 + \alpha_2) + 2(\beta_1 + \beta_2) + (\gamma_{11} + \gamma_{12} + \gamma_{13} + \gamma_{21} + \gamma_{22} + \gamma_{23}) = 0$$

has

$$\underset{\sim}{K}' = [6 \;\; 3 \;\; 3 \;\; 2 \;\; 2 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1] .$$

With

$$\underset{\sim}{b}^{o\prime} = [0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 8 \;\; 6 \;\; 2 \;\; 8 \;\; 6 \;\; 12], \text{ of page 7}$$

and, implicitly,

$$(\underset{\sim}{X}'\underset{\sim}{X})^{-} = \text{diag}\{0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; \tfrac{1}{2} \;\; 1 \;\; 1 \;\; 1 \;\; \tfrac{1}{2} \;\; 1\} ,$$

$$Q = \frac{(8 + 6 + 2 + 8 + 6 + 12)^2}{\tfrac{1}{2} + 1 + 1 + 1 + \tfrac{1}{2} + 1} = \frac{42^2}{5} = 352.8 .$$

$Q = 352.8 = R^{*}(\dot{\mu}|\dot{\alpha},\dot{\beta},\dot{\gamma})_{\Sigma}$ because the hypothesis is equivalent to $H: \dot{\mu} = 0$.

METHOD 5: "INDIRECT" METHOD, FOR FULL RANK MODELS

Model: $E(\underset{\sim}{y}) = \underset{\sim}{X}_1\underset{\sim}{b}_1 + \underset{\sim}{X}_2\underset{\sim}{b}_2$

Normal Equations — Solution — Definition

$$\begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{X}_1 & \underset{\sim}{X}_1'\underset{\sim}{X}_2 \\ \underset{\sim}{X}_2'\underset{\sim}{X}_1 & \underset{\sim}{X}_2'\underset{\sim}{X}_2 \end{bmatrix} \begin{bmatrix} \hat{\underset{\sim}{b}}_1 \\ \hat{\underset{\sim}{b}}_2 \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{y} \\ \underset{\sim}{X}_2'\underset{\sim}{y} \end{bmatrix}, \quad \begin{bmatrix} \hat{\underset{\sim}{b}}_1 \\ \hat{\underset{\sim}{b}}_2 \end{bmatrix} = \begin{bmatrix} \underset{\sim}{T}_{11} & \underset{\sim}{T}_{12} \\ \underset{\sim}{T}_{21} & \underset{\sim}{T}_{22} \end{bmatrix} \begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{y} \\ \underset{\sim}{X}_2'\underset{\sim}{y} \end{bmatrix}, \quad \begin{bmatrix} \underset{\sim}{T}_{11} & \underset{\sim}{T}_{12} \\ \underset{\sim}{T}_{21} & \underset{\sim}{T}_{22} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{X}_1 & \underset{\sim}{X}_1'\underset{\sim}{X}_2 \\ \underset{\sim}{X}_2'\underset{\sim}{X}_1 & \underset{\sim}{X}_2'\underset{\sim}{X}_2 \end{bmatrix}^{-1}.$$

"Indirect" Method (SAS HARVEY)
"Invert part of the inverse" (LM 115)
} Define: $Q_{\underset{\sim}{b}_1} = \hat{\underset{\sim}{b}}_1'\underset{\sim}{T}_{11}^{-1}\hat{\underset{\sim}{b}}_1$.

Example 1r:

No-interaction model (see page 9).

Normal Equations — Solution

$$\begin{bmatrix} 8 & 0 & 1 & 1 \\ 0 & 8 & 1 & -1 \\ 1 & 1 & 5 & 2 \\ 1 & -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 56 \\ -8 \\ 10 \\ 4 \end{bmatrix} \qquad \begin{bmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \frac{1}{22(27)} \begin{bmatrix} 77 & 0 & -11 & -11 \\ 0 & 81 & -27 & 27 \\ -11 & -27 & 152 & -64 \\ -11 & 27 & -64 & 152 \end{bmatrix} \begin{bmatrix} 56 \\ -8 \\ 10 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ -\frac{15}{11} \\ \frac{16}{11} \\ -\frac{16}{11} \end{bmatrix}.$$

$$Q_{\dot{\mu}} = \hat{\mu}T_{\dot{\mu}\dot{\mu}}^{-1}\hat{\mu} = 7\left[\frac{77}{22(27)}\right]^{-1}7 = 378 = R^*(\dot{\mu}|\dot{\alpha},\dot{\beta})_\Sigma \text{ of page 10.}$$

Interaction model (See page 12)

$$Q_{\dot{\mu}} = \hat{\mu}T_{\mu\mu}^{-1}\hat{\mu} = 7(10/72)^{-1}(7) = 352.8 = R^*(\dot{\mu}|\dot{\alpha},\dot{\beta},\dot{\gamma})_\Sigma .$$

RELATIONSHIPS BETWEEN THE METHODS

Relationships depend upon model and data.

Full rank models

The "indirect" $Q_{\underset{\sim}{b}_1} = R(\underset{\sim}{b}_1|\underset{\sim}{b}_2)$ = the numerator s.s. Q for testing $H: \underset{\sim}{b}_1 = 0$.

Non full rank models (e.g., 2-way crossed classification)

After making $E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$ full rank by using Σ-restrictions, the "indirect" $Q_{\dot{\alpha}} = R^{*}(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_{\Sigma}$ = the numerator s.s. Q for testing $H: \dot{\alpha}_i$'s all zero.

Further interpretation depends upon model and data.

1. No interaction model (usually $n_{ij} = 0$ or 1)

The "indirect" $Q_{\dot{\alpha}} = R(\alpha|\mu,\beta)$ = the numerator s.s. Q for testing $H: \alpha_i$'s all equal.

Example: From page 16, $Q_{\dot{\alpha}} = (-15/11)\left(\frac{81}{22(27)}\right)^{-1}(-15/11) = \frac{150}{11}$

$$= 13\tfrac{7}{11} = R(\alpha|\mu,\beta) \text{ of page 6.}$$

2. With interaction model (all cells filled)

The "indirect" $Q_{\dot{\alpha}} = SSA_w$ = the numerator s.s. Q for testing $H: (\alpha_i + \bar{\gamma}_{i.})$'s all equal.

Example: From page 12, $Q_{\dot{\alpha}} = (-10/6)(10/72)^{-1}(-10/6)$

$$= 20 = R^{*}(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_{\Sigma} \text{ of page 14}$$

$$= SSA_w \text{ of page 15.}$$

Using page 16, $H: \alpha_1 + \frac{1}{3}(\gamma_{11}+\gamma_{12}+\gamma_{13}) = \alpha_2 + \frac{1}{3}(\gamma_{21}+\gamma_{22}+\gamma_{23})$

$$\text{has} \quad Q = \frac{[\frac{1}{3}(8+6+2) - \frac{1}{3}(8+6+12)]^2}{\frac{1}{9}(\frac{1}{2}+1+1+1+\frac{1}{2}+1)} = 20.$$

3. With interaction model (empty cells)

$Q_{\dot{\alpha}} = R^*(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma$ = the numerator s.s. Q for testing $H:\dot{\alpha}_i$'s all equal.

The $\dot{\alpha}_i$'s, as functions of α_i's and other parameters of the over-parameterized model, depend upon the pattern of empty cells.

Example (a):

✓	✓	✓
✓	✓	

✓ = data

$$\dot{\alpha}_1 = \frac{1}{2}(\alpha_1 - \alpha_2) + \frac{1}{4}(\gamma_{11} + \gamma_{12} - \gamma_{21} - \gamma_{22}) .$$

Example (b):

✓	✓	✓
✓	✓	
✓		✓

$$\dot{\alpha}_1 = \alpha_1 - \frac{1}{3}(\alpha_1 + \alpha_2 + \alpha_3) + \frac{1}{15}(4\gamma_{11} + 3\gamma_{12} + 3\gamma_{13} - 2\gamma_{21} - 3\gamma_{22} - 2\gamma_{31} - 3\gamma_{33})$$

$$\dot{\alpha}_2 = \alpha_2 - \frac{1}{3}(\alpha_1 + \alpha_2 + \alpha_3) + \frac{1}{15}(-2\gamma_{11} - 4\gamma_{12} + \gamma_{13} + 6\gamma_{21} + 4\gamma_{22} - 4\gamma_{31} - \gamma_{33}) .$$

THE BEST PROCEDURE FOR EMPTY CELL CASES

The cell means model:

$$E(y_{ijk}) = \mu_{ijk} .$$

ESTIMABLE FUNCTIONS

A general estimable function

Model: $E(\tilde{y}) = \tilde{X}\tilde{b}$

Normal equations: $\tilde{X}'\tilde{X}\tilde{b}^{o} = \tilde{X}'\tilde{y}$

Generalized inverse: $\tilde{G} = (\tilde{X}'\tilde{X})^{-}$, $\tilde{X}'\tilde{X}\tilde{G}\tilde{X}'\tilde{X} = \tilde{X}'\tilde{X}$

Sometimes use $\tilde{G}^{*} = \tilde{G}\tilde{X}'\tilde{X}\tilde{G}' = \tilde{G}^{*}{}' = \tilde{G}^{*}\tilde{X}'\tilde{X}\tilde{G}^{*}$

Define $\tilde{H} = \tilde{G}\tilde{X}'\tilde{X} = \tilde{G}^{*}\tilde{X}'\tilde{X}$

A general estimable function: $\tilde{\ell}'\tilde{b}$ for $\tilde{\ell}' = \tilde{\ell}^{*}{}'\tilde{H}$ for any $\tilde{\ell}^{*}{}'$

Testable hypotheses

$H: \tilde{K}'\tilde{b} = \tilde{m}$ with $\tilde{K}'\tilde{b}$ estimable, $\tilde{K}'$ of full row rank s.

$F = Q/s\hat{\sigma}^2$, $Q = (\tilde{K}'\tilde{b}^{o} - \tilde{m})'(\tilde{K}'\tilde{G}\tilde{K})^{-1}(\tilde{K}'\tilde{b}^{o} - \tilde{m})$.

$\tilde{K}'\tilde{b}$ is $\tilde{k}'\tilde{b}$ for s LIN vectors $\tilde{k}'$.

These $\tilde{k}'$ vectors have different forms for different hypotheses.

Example: $H: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$

$$H: \begin{cases} \alpha_1 - \alpha_4 = 0 \\ \alpha_2 - \alpha_4 = 0 \\ \alpha_3 - \alpha_4 = 0 \end{cases}$$

For any 3 linearly independent vectors $(\ell_1 \;\; \ell_2 \;\; \ell_3)$

$$H: \ell_1\alpha_1 + \ell_2\alpha_2 + \ell_3\alpha_3 - (\ell_1 + \ell_2 + \ell_3)\alpha_4 = 0 .$$