Clarke, B.R. and Hogan, R. (2005) ***A way of deriving ANOVA for mixed models and variance component models based on an historical representation of component sums of squares.*** **In: Proceedings of the 55th International Statistical Institute Session, Sydney, 2005. The Hague, The Netherlands: ISI**

A Way of Deriving ANOVA for Mixed Models and Variance Component Models Based on an Historical Representation of Component Sums of Squares

Brenton R. Clarke and Rebecca Hogan
Murdoch University, Mathematics and Statistics
South Street
Murdoch 6150, Australia
B.Clarke@murdoch.edu.au

A representation of sums of squares in two way layouts deriving from the history of the discussion of the introduction of the ANOVA method of R.A.Fisher by J.O.Irwin was introduced recently by the first listed author of this paper (see Clarke (2002)). Partitions of Helmert matrices and Kronecker products were used to easily derive the distribution theory of component sums of squares in fixed effect models to do with the two way layout. In this paper we show how the derivation of distribution theory to do with mixed models and variance component models can easily follow from the same representation. We consider the equation that represents the mixed model

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Wu} + \boldsymbol{\epsilon}, \tag{1}$$

where the vector of parameters $\boldsymbol{\beta}$ is a set of parameters representing constants, while $\boldsymbol{u}$ is a vector of random effects. The errors in the vector $\boldsymbol{\epsilon}$ describe the usual independent normal mean zero variance σ_ϵ^2 errors. Thus for example when we consider a mixed model for the two way layout with r rows and s columns then the observation in the (i, j)th cell is given as

$$Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}.$$

For this model μ represents the overall grand mean, and the $\alpha_i's$ are constants satisfying the restraint $\sum_{i=1}^r \alpha_i = 0$. There are no linear restraints placed on the random effect parameters $\beta_j's$ and we assume these are independent normal variables with mean zero and variance σ_β^2. These are also independent of the normal errors ϵ_{ij}. The model describing these data is given by (1) where $\boldsymbol{Y}$ is the vector formed from ordering the observations column by column underneath each other from the two way layout and

$$\boldsymbol{X} = [\mathbf{1_s} \otimes \mathbf{1_r} \;:\; \mathbf{1_s} \otimes \boldsymbol{I_r}],\;\; \boldsymbol{\beta} = (\mu, \alpha_1, \ldots, \alpha_r)'$$

and

$$\boldsymbol{W} = [\boldsymbol{I_s} \otimes \mathbf{1_r}] \text{ and } \boldsymbol{u} = (\beta_1, \ldots, \beta_s)'.$$

Here for example $\mathbf{1_r}$ is the $r \times 1$ vector of ones and $\boldsymbol{I_r}$ is the identity matrix of order r. The sums of squares in the ANOVA for this mixed model are exactly the same as in the fixed effects model described in Clarke (2002) except their distributions and consequent expected mean sum of squares are different. Consequently we may use a similar "canonical decomposition" that is used in Clarke (2002) so that letting for example $\boldsymbol{P_A}$ be the orthogonal projection matrix used to obtain the sum of squares corresponding to the fixed effects factor A, and $\boldsymbol{P_B}$ corresponds to the orthogonal projection matrix used to obtain the sum of squares corresponding to the random factor B, etcetera we have the sums of squares in the ANOVA given by

$$\begin{aligned} SSM &= \boldsymbol{Y'P_MY} &= \boldsymbol{Y'C_M'C_MY} \\ SSA &= \boldsymbol{Y'P_AY} &= \boldsymbol{Y'C_A'C_AY} \\ SSB &= \boldsymbol{Y'P_BY} &= \boldsymbol{Y'C_B'C_BY} \\ SSE &= \boldsymbol{Y'P_EY} &= \boldsymbol{Y'C_E'C_EY} \end{aligned}$$

and $\boldsymbol{C_A} \equiv \boldsymbol{C_T}$ in Clarke (2002), while all the other matrices expressed in defining the sums are as in section 3 of that paper. The difference in the ANOVA for this mixed model comes in the discussion of expected mean squares. These are derived as follows. The vector $\mathbf{z} = (z_1, \ldots, z_{rs})'$ is given here in matrix notation by

$$(2) \qquad \mathbf{z} = \mathbf{CY} = \begin{bmatrix} \mathbf{C}_M \\ \mathbf{C}_A \\ \mathbf{C}_B \\ \mathbf{C}_E \end{bmatrix} \mathbf{Y} = \begin{bmatrix} (rs)^{-\frac{1}{2}} \mathbf{1_s'} \otimes \mathbf{1_r'} \\ s^{-\frac{1}{2}} \mathbf{1_s'} \otimes \boldsymbol{B_r} \\ r^{-\frac{1}{2}} \boldsymbol{B_s} \otimes \mathbf{1_r'} \\ \boldsymbol{B_s} \otimes \boldsymbol{B_r} \end{bmatrix} \mathbf{Y}$$

The matrix $\boldsymbol{B_s}$ can for instance be the $s-1$ rows of the Helmert matrix that are orthogonal to the unit vector. Hence we have for example using the rules of Kronecker product

$$E[(z_2, \ldots, z_r)'] = \boldsymbol{C_A X \beta} = \sqrt{s} \boldsymbol{B_r} (\alpha_1, \ldots, \alpha_r)'$$

Also

$$\begin{aligned} \text{cov}((z_2, \ldots, z_r)') &= \text{cov}(\boldsymbol{C_A Y}) \\ &= \boldsymbol{C_A} \text{cov}(\boldsymbol{Y}) \boldsymbol{C_A'} \\ &= \boldsymbol{C_A} \{ (\boldsymbol{I_s} \otimes \mathbf{1_r 1_r'}) \sigma_\beta^2 + (\boldsymbol{I_s} \otimes \boldsymbol{I_r}) \sigma_\epsilon^2 \} \boldsymbol{C_A'} \end{aligned}$$

which after a little algebra yields

$$\text{cov}((z_2, \ldots, z_r)') = \sigma_\epsilon^2 \boldsymbol{I_{r-1}}.$$

Consequently

$$\boldsymbol{C_A Y} \sim N(\sqrt{s} \boldsymbol{B_r} (\alpha_1, \ldots, \alpha_r)', \ \sigma_\epsilon^2 \boldsymbol{I_{r-1}})$$

which now easily leads to the noncentral chi-squared distribution $SSA \sim \sigma_\epsilon^2 \chi^2_{r-1;\delta}$, where $\delta = \frac{s}{\sigma_\epsilon^2} \sum_{i=1}^r \alpha_i^2$ and the expected mean sum of squares for factor A is then $E[MSA] = \frac{s}{r-1} \sum_{i=1}^r \alpha_i^2 + \sigma_e^2$ Now turning our attention to the expected mean squares for factor B consider

$$\begin{aligned} \text{cov}((z_{r+1}, \ldots, z_{r+s-1})') &= \text{cov}(\boldsymbol{C_B Y}) \\ &= \boldsymbol{C_B} \text{cov}(\boldsymbol{Y}) \boldsymbol{C_B'} \\ &= \boldsymbol{C_B} (\boldsymbol{I_s} \otimes \mathbf{1_r 1_r'}) \sigma_\beta^2 + (\boldsymbol{I_s} \otimes \boldsymbol{I_r}) \sigma_\epsilon^2) \boldsymbol{C_B'} \end{aligned}$$

which simplifies using Kronecker product algebra and substituting $\boldsymbol{C_B} = \frac{1}{\sqrt{r}} (\boldsymbol{B_s} \otimes \mathbf{1_r'})$ to $(r\sigma_\beta^2 + \sigma_\epsilon^2) \boldsymbol{I_{s-1}}$. Therefore $\boldsymbol{C_B Y} \sim N(\mathbf{0}, (r\sigma_\beta^2 + \sigma_\epsilon^2) \boldsymbol{I_{s-1}})$ and so $E[MSB] = r\sigma_\beta^2 + \sigma_\epsilon^2$.

REFERENCES

Clarke, B.R. (2002) A representation of orthogonal components in analysis of variance. *Intern. Math. Journal*, Vol. 1, 133-147.

RÉSUMÉ

On généralise une représentation historique de composantes orthogonales dans l'analyse de la variance en utilisant des techniques modernes d'algèbre matricielle développées dans Clarke(2002). Ceci inclut les modèles de composantes de la variance et les modèles mixtes. Ces représentations sont utiles dans le cadre de l'enseignement de l'ANOVA.