Optimal Exact F Tests

Theresa Utlaut
Intel Corporation, Technology Development
2111 NE 25th Street
Hillsboro OR 97124 USA
theresa.l.utlaut@intel.com

David Birkes
Oregon State University, Statistics Department
44 Kidder Hall
Corvallis OR 97331 USA
birkes@stat.orst.edu

Dawn VanLeeuwen
New Mexico State University, Agricultural and Extension Education
P.O. Box 30003, MSC 3130
Las Cruces NM 88003 USA
statinu4@nmsuvm1.nmsu.edu

In balanced mixed-effects analysis-of-variance models, there are optimal exact F tests for most of the usual hypotheses about the fixed effects and the variance components. In such models Seifert (1979) showed that the usual ANOVA F tests of the fixed effects are uniformly most powerful invariant unbiased (UMPIU). ElBassiouni and Seely (1980) and Mathew and Sinha (1988) established UMPU F tests for linear combinations of certain canonical linear combinations of the variance components. In this paper we generalize these results to models that are "partially balanced" and illustrate them in a split-plot model.

VanLeeuwen, Seely and Birkes (1998) introduced the concept of an error-orthogonal (EO) design. A linear model has an EO design if (i) the least-squares estimator of its mean vector is a uniformly best linear unbiased estimator and (ii) the linear space spanned by the possible covariance matrices of the vector of least-squares residuals is a commutative quadratic subspace (Seely, 1971). The structure of an EO linear model allows optimal exact F tests for certain canonical hypotheses (Utlaut, 1999).

Here we will focus on simple error-orthogonal (SEO) designs, defined as follows. Consider a variance-components model

$$E(y) = X\boldsymbol{b}, \quad \mathrm{Cov}(y) = \sigma_1^2 V_1 + \cdots + \sigma_r^2 V_r + \sigma_{r+1}^2 I .$$

The model is SEO if (i) $V_j = m_j P_j$ for an orthogonal projection (o.p.) matrix P_j , (ii) $P_X P_j = P_j P_X$ where P_X is the o.p. matrix on $\varepsilon(X)$, the range space (or column space) of X, and (iii) $Q_j Q_k \neq 0$ implies $Q_j Q_k = Q_l$ for some l where $Q_j = MP_j$ and $M = I - P_X$.

In an ANOVA model, a type of partial balance called b&r-balance (VanLeeuwen, Birkes and Seely, 1999) implies SEO. A design is r-balanced if, for every pair of random effects, the marginal incidence matrix, obtained by summing the cell counts over all indices not involved in either of the two effects, is balanced. A design is b-balanced if, for every random effect and fixed effect, the marginal incidence matrix is conditionally balanced in the sense that for any given vector of indices for the fixed effect, the entries are equal. A design is b&r-balanced if it is b-balanced and r-balanced.

In an SEO model the o.p. matrix M can be decomposed as $M = \Sigma_h M_h$ where (a) the M_h's are mutually orthogonal o.p. matrices and (b) every Q_j can be written as a sum of some of the M_h's. That is, $Q_j = \Sigma_h w_{jh} M_h$ where $w_{jh} = 0$ or 1. The matrices M_h can be constructed as follows. Start with the Q_j's, including $Q_{r+1} = M$. For pairs of matrices for which $\varepsilon(Q_j) \subset \varepsilon(Q_k)$, replace Q_k by $Q_k - Q_j$.

Continue in this way until all the matrices are mutually orthogonal. This leads to a spectral decomposition of the covariance matrix of the vector of least-squares residuals as

$$\mathrm{Cov}(My) = \pi_1(\psi)M_1 + \cdots + \pi_s(\psi)M_s$$

where $\pi_h(\psi) = \Sigma_j w_{jh} m_j \sigma_j^2$ and $\psi = (\sigma_1^2,\ldots,\sigma_{r+1}^2)$. Let $V_\psi = \mathrm{Cov}(y)$ and $\mathrm{MS}_h = y'M_h y/\mathrm{rank}(M_h)$.

Theorem 1: Consider a normal mixed-effects linear model that is SEO. Suppose P_0 is an o.p. matrix such that (i) $\varepsilon(P_0) \subset \varepsilon(X)$ and (ii) $P_0 V_\psi = \pi_h(\psi)M_h$ for a particular h. Let $\mathrm{MS}_0 = y'P_0 y/\mathrm{rank}(P_0)$. Then $\mathrm{MS}_0/\mathrm{MS}_h$ is the test statistic for an exact F test of H_0: $P_0X\beta = 0$ versus H_a: $P_0X\beta \neq 0$. The test is UMPIU.

Theorem 2: Consider a normal mixed-effects linear model that is SEO. Suppose $\pi_f(\psi) \geq \pi_g(\psi)$. Then $\mathrm{MS}_f/\mathrm{MS}_g$ is the test statistic for an exact F test of H_0: $\pi_f(\psi) = \pi_g(\psi)$ versus H_a: $\pi_f(\psi) > \pi_g(\psi)$. The test is UMPIU.

Let us apply these results to a split-plot model with random blocks,

$$y_{ijklm} = \mu + \alpha_i + \beta_j + \theta_{ij} + c_k + d_{ijl} + e_{ijklm}$$

for $m = 1,\ldots,n_{ijkl}$, where α_i is a fixed whole-plot treatment effect, β_j is a fixed subplot treatment effect , θ_{ij} is a fixed interaction effect, c_k is a random block effect, d_{ijl} is random whole-plot error, and e_{ijklm} is random subplot error. The design of the model is determined by the values of n_{ijkl}, the number of subplots in whole plot ikl to which the j-th subplot treatment is applied. To check whether or not a design has the SEO property it is helpful to reparameterize it by dropping the α_i and β_j terms. Then the design is b&r-balanced if and only if (1) n_{ijkl} depends only on ij (i.e., for each given treatment combination ij, in all whole plots in all blocks, the number of subplots to which the ij treatment combination is applied is the same) and (2) $n_{i\cdot kl}$ (which is the number of subplots in whole plot ikl) is the same in all whole plots in all blocks. For such designs, all the usual ANOVA tests are optimal exact F tests. Such a design would be useful, for instance, in a situation in which all the whole plots contained 3 subplots and there were 2 subplot treatments. Then the subplot treatments could not balanced within each whole plot, but optimal exact F tests could still be obtained if the number of subplots to which subplot treatment 1 is assigned is the same among all the whole plots having any given whole-plot treatment.

REFERENCES

ElBassiouni, M.Y. and Seely, J. (1980). Optimal tests for certain functions of the parameters in a covariance matrix with linear structure. *Sankhya* Ser. A 42, 64-77.

Mathew, T. and Sinha, B.K. (1988). Optimum tests for fixed effects and variance components in balanced models. *J. Amer. Statist. Assoc.* 83, 133-135.

Seely, J. (1971). Quadratic subspaces and completeness. *Ann. Math. Statist.* 42, 710-721.

Seifert, B. (1979). Optimal testing for fixed effects in general balanced mixed classification models. *Math. Oper. Statist.* 10, 237-255.

Utlaut, T.L. (1999). F-tests in partially balanced and unbalanced mixed linear models. Unpublished Ph.D. thesis, Statistics Dept., Oregon State Univ.

VanLeeuwen, D.M., Birkes, D.S. and Seely, J.F. (1999). Balance and orthogonality in designs for mixed classification models. *Ann. Statist.* 27, 1927-1947.

VanLeeuwen, D.M., Seely, J.F. and Birkes, D.S. (1998). Sufficient conditions for orthogonal designs in mixed linear models'. *J. Statist. Plann. and Inference* 73, 373-389

RÉSUMÉ

En modèles mixtes de l' analyse de variance, tests F exacts optimaux peuvent être obtenus avec ne pas les plans balancés que des plans de balance partielle qui s' apellent de balance b&r. De plus généralité, en modèles linéaires mixtes, tests F exacts optimaux peuvent être obtenus avec des plans simples orthogonaux en ce qui concerne l' erreur. Les resultants sont appliqué à un modèle split-plot de blocs aléatoires.