Journal of the Research Institute of Basic Sciences,
Korea Maritime University, Vol. 2, 1992.

AN EXACT TEST PROCEDURE FOR VARIANCE COMPONENTS OF A RANDOM EFFECT LINEAR MODEL

Choon-il Park*
Seung-Chun Li+

*Dept. of Applied Mathematics, Korea Maritime Univ., pusan, Korea
+Dept. of Statistics, Pusan Univ. of Foreign Studies, Pusan, Korea

Abstract

For the unbalanced linear model, exact test procedures for variance components are available only in restricted cases. In this note we derive a simultaneous test procedure for the variance components of a general random effect linear model. The test statistic has central F-distribution on the boundary of the hypotheses and the test is exact. Power of the test is given. We also obtain an explicit form of the test statistic for the one-way random model . Under some restrictions on the design matrix, we give an exact test procedure for each random effects.

Key Words: variance components, exact test procedure, MINQUE,
power of test

1. Introductions

For the balanced fixed and random effect linear models, appropriate F-tests for testing the significance of fixed and random effects are exact and known to be UMPU and UMPIU (Graybill,1976). Mathew and Sinha(1988) recently established the UMPU and UMPIU character of standard F-tests for the balanced mixed effect models. However, with unbalanced data the situation is different. Exact test procedures for variance components are known only for restricted case. For example, an exact test procedure for the one-way random model, is suggested by Spjtvoll(1967), and for the two-way random model, some results can be found in Spjotvoll(1968), Thomsen(1975), and Kuri and Littel(1987). Verdooren(1988) gave some results for the two and three stage nested designs.

in this work, we suggest a simultaneous test procedure for all of the random effects of the model. the resulting test statistic is shown to be distributed as the usual F-distribution on the boundary of hypotheses. We start with a general random effect linear model and give explicit form of the test statistic for one-way random model. We also give some results on the procedures for each random effects when the model has some restrictions on the design matrices.

2. General model case

Consider the following general linear model.

$$\mathbf{y} = 1\mu + X_1\xi_1 + \cdots + X_k\xi_k + \varepsilon \tag{1}$$

where $\mathbf{y}$ is a vector of n observations, μ is a fixed unknown constant, 1 is n-vector of 1's, X_i is $n \times b_i$ design matrix, ξ_i is b_i-vector of uncorrelated random effects, and ε is n-vector of random errors. We assume that ξ_i's and ε are independent and multinormally distributed with zero mean vectors and variance-covariance matrices $\sigma_i^2 I$ for $i=1,\cdots,k$ and $\sigma_{k+1}^2 I$, respectively.

Let C be a full row rank matrix of order $(n-1) \times n$ such that $C1 = 0$, $CC' = I_{n-1}$ and $C'C = I_n - {}^1/_n 11'$. Then multiplying both sides of equation (1) by C yields

$$\mathbf{z} = CX_1\xi_1 + \cdots + CX_k\xi_k + C\varepsilon \tag{2}$$

where $z = Cy$.

It can be shown(see Rao,section 9,1971,for example) that the MINQUE of $\sigma = (\sigma_1^2,\cdots,\sigma_k^2,\sigma_{k+1}^2)$ for the model (1) is the same as that for the model (2), and is given by

$$S\hat{\sigma} = u \tag{3}$$

where $S = \{tr(RV_iRV_j)\}$, $u = \{z'RV_iRz\}$, $V_i = CX_iX_i'C'$ for $i,j = 1,\cdots,k+1$, and $R = \Sigma_{i=1}^{k+1} r_iV_i)^{-1} = (I + \Sigma_{i=1}^{k} r_iV_i)^{-1}$ with r_i denoting the a-priori values of $\rho_i = \sigma_i^2/\sigma_{k+1}^2$, for $i = 1,\cdots,k$, and $r_{k+1} = 1$.

Since $\Sigma_{i=1}^{k} r_iV_i$ is symmetric, we can decompose it by

$$\sum_{i=1}^{k} r_iV_i = PDP' \tag{4}$$

where D is a diagonal matrix of eigen values of $\Sigma_{i=1}^{k} r_iV_i$ and P is the matrix of corresponding eigen vectors. We assume that the first h diagonal elements of D is nonzero where h is the rank of $\Sigma_{i=1}^{k} r_iV_i$.

Since P is orthogonal, We have

$$R = \left[I + \sum_{i=1}^{k} r_i V_i \right]^{-1} = (I + PDP')^{-1} = P(I + D)^{-1}P' \quad (5)$$

and we can decompose

$$(I + D)^{-1} = D_1 + D_2 \quad (6)$$

where $D_1 = diag(1/(1 + \lambda_1),\cdots,1/(1 + \lambda_h),0,\cdots,0)$, $D_2 = diag(0,\cdots,0,1,\cdots,1)$ and λ_i's are nonzero eigen values of $\Sigma_{i=1}^{k} r_i V_i$.

Consider the linear combination $\Sigma_{i=1}^{k+1} r_i u_i$ where u_i is the i-th element of u of equation (3). Applying the results of (5) and (6), we can decompose $\Sigma_{i=1}^{k=1} r_i u_i$ into two parts as follow :

$$\sum_{i=1}^{k+1} r_i u_i = z'R(I + \sum_{i=1}^{k} r_i V_i)Rz = z'Rz = z'PD_1P'z + z'PD_2P'z. \quad (7)$$

Suppose that r_i and $\rho_i > 0$ for all i. Then $\Sigma_{i=1}^{k} \rho_i V_i$ and $\Sigma_{i=1}^{k} r_i V_i$ are commuting and there exists an orthogonal matrix P such that each $D = P'(\Sigma_{i=1}^{k} r_i V_i)P$ and $D^* = P'(\Sigma_{i=1}^{k} \rho_i V_i)P$ are diagonal matrices with diagonal elements being eigen values of $\Sigma_{i=1}^{k} r_i V_i$ and $\Sigma_{i=1}^{k} \rho_i V_i$, respectively.
Note that when the diagonal element in D is nonzero, corresponding element in D^* is nonzero as well, since the column space of $\Sigma_{i=1}^{k} \rho_i V_i$ is equal to that of $\Sigma_{i=1}^{k} r_i V_i$. Hence $I + \Sigma_{i=1}^{k} \rho_i V_i$ can be written as

$$I + \sum_{i=1}^{k} \rho_i V_i = P(I + D^*)P'$$
$$= PD_1^*P' + PD_2P'$$

where $D_1^* = diag(1 + \lambda_1^*,\cdots, 1 + \lambda_h^*,0,\cdots,0)$ with λ_i^*'s being nonzero eigenvalues of $\Sigma_{i=1}^{k} \rho_i V_i$.

Let $D_i = D_i^{1/2} D_i^{1/2}$, for i=1,2. Since z/ρ_{k+1} $\tilde{N}(o, I + \Sigma_{i=1}^{k} \rho_i V_i)$, It follows that

$$Var(P'z/\sigma_{k+1}) = D_1^* + D_2, \quad (8)$$
$$D_1 Var(P'z) D_2 = 0, \quad (9)$$
$$D_1^{1/2}P'z/\sigma_{k+1} \quad \tilde{N}(0, D_1 D_1^*) \quad (10)$$

and

$$D_1^{1/2}P'z/\sigma_{k+1} \quad \tilde{N}(0, D_2). \quad (11)$$

Hence $z'PD_1P'z/\rho_{k+1}^2$ and $z'PD_2P'z/\rho_{k+1}^2$ are independent for $\rho_i > 0$, $i = 1,\cdots,k$,and distributed as $\Sigma_{i=1}^{k}(1+\lambda_i^*/1+\lambda_i)X_i^2$ and X^2, respectively, where X_i^2's and X^2 are independent chi-square random variables with 1 and n-h-1 degrees of freedom.
These suggest us to consider $F = ((n-h-1)/h)z'PD_1P'z/z'PD_2P'z$ as a test statistic to test H_0 : $\rho_i \le r_i$ for all $i = 1,2,\ldots,k$ against H_1 : $\rho_i > r_i$ for at least one i. On the boundary of hypotheres, the distribution of F is

F(h,n-h-1). Preumably we reject the null hypothesis if F is too large.
Consider the eigen values of $\Sigma_{i=1}^{k}(\rho_i + \varepsilon_i)V_i$ where $\varepsilon_i \geq 0$ for all i. Since $\Sigma_{i=1}^{k}\rho_i V_i$ and $\Sigma_{i=1}^{k}(\rho_i + \varepsilon_i)V_i$ are commuting, we have

$$P'\left(\sum_{i=1}^{k}(\rho_i + \varepsilon_i)V_i\right)P = \tilde{D} \tag{12}$$

where $\tilde{D} = diag(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_h, 0, \ldots, 0)$ with $\tilde{\lambda}_i$'s being nonzero eigen values of $\Sigma_{i=1}^{k}(\rho_i + \varepsilon_i)V_i$. Thus

$$P\left(\sum_{i=1}^{k}\varepsilon_i V_i\right)P = \tilde{D} - D^* = diag(\tilde{\lambda}_1 - \lambda^*_1, \ldots, \tilde{\lambda}_h - \lambda^*_h, 0, \ldots, 0). \tag{13}$$

Note that $\tilde{\lambda}_i - \lambda^*_i$ is an eigen value of $\Sigma_{i=1}^{k}\varepsilon_i V_i$. since $\Sigma_{i=1}^{k}\varepsilon_i V_i$ is a non-negative definite, it follows that

$$\tilde{\lambda}_i - \lambda^*_i \geq 0 \text{ for all } i = 1,2,\ldots,h. \tag{14}$$

Consequently, the unmerator of F increases with ρ_i's. while the denomination is independent of ρ_i's.Hence the rejection region of our test should be upper tail of F-distribution. The power of test is given by

$$Pr\left[\sum_{i=1}^{h}\frac{1+\lambda^*_i}{1-\lambda_i}X_i^2 > \frac{h}{n-h-1}cX^2\right] \tag{14}$$

where c is an appropriate constant such as $F\alpha(h,n-h-1)$ for size α test

Remark 2.1: The power of the test can be calculated by the algorithm given by Farebrother(1984).

3. Special case 1.

We will derive an explicit form of the test statistic for the following one-way random model,

$$y_{ij} = \mu + \xi_i + \varepsilon_i \tag{16}$$

where $j = 1,\ldots,n_i$, $i = 1,\ldots,b$. Here μ is a constant, ξ_i and ε_{ij} are independent normally distributed random variables with means 0 and variance σ_1^2 and σ^2, respectively. In matrix notation, the model can be written as

$$y = 1\mu + X\xi + \varepsilon \tag{17}$$

where 1 is n-vector of 1's, X is $n \times b$ design matrix with $n = \Sigma_{i=1}^{b} n_i$ and ξ and ε are multivariate normal random variables with mean vector 0 and variance-covariance matrices $\sigma_1^2 I$ and $\sigma^2 I$, respectively.

Let $rXX' = QEQ'$, where $E = diag(rn1,\ldots,rnb,0\ldots,0)$, diagonal matrix of eigen values of rXX' and Q is the matrix of corresponding eigen vectors. Then

$$\begin{aligned}(I + rCXX'C')^{-1} &= (I + CQEQ'C')^{-1}\\ &= I - CQE^{1/2}(I + E^{1/2}Q'C'CQE^{1/2})^{-1}E^{1/2}Q'C'\\ &= I - CQE^{1/2}(I + E^{1/2}Q'(I - 1/n\ 11')QE^{1/2})^{-1}E^{1/2}Q'C'\\ &= I - CQE^{1/2}(I + E - 1/n\ E^{1/2}Q'11'QE^{1/2})^{-1}E^{1/2}Q'C' \qquad (18)\end{aligned}$$

Nothing

$$(I + E - 1/nE^{1/2}Q'11'QE^{1/2})^{-1} = (I + E)^{-1} + \frac{(I + E)^{-1}E^{1/2}Q'11'QE^{1/2}(I + E)^{-1}}{n - 1'QE^{1/2}(I + E)^{-1}E^{1/2}Q'1}$$

and

$$n-1'QE^{1/2}(I + E)^{-1}E^{1/2}Q'1 = \sum_{i=1}^{b} n_i/(1 + rn_i),$$

we have the following result some algebra:

$$z'Rz = \sum_{i=1}^{b}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2 + \sum_{i=1}^{b}\frac{n_i(\bar{y}_{i.} - \bar{y}_{..})^2}{1 + rn_i} - \frac{\left[\sum_{i=1}^{b}\frac{rn_i^2(\bar{y}_{i.} - \bar{y}_{..})}{1 + rn_i}\right]^2}{\sum_{i=1}^{b}\frac{n_i}{1 + rn_i}} \qquad (19)$$

where $\bar{y}_{i.} = 1/n_i\sum_{j=1}^{n_i} y_{ij}$ and $\bar{y}_{..} = 1/n\sum_{i=1}^{b}\sum_{j=1}^{n_i} y_{ij}$. It can be shown that the first term of (19) is equal to $z'PD_2P'z$. Thus our test statistic can be written as

$$F = \frac{n - b}{b - 1}\;\frac{\sum_{i=1}^{b}\frac{n_i(\bar{y}_{i.} - \bar{y}_{..})^2}{1 + rn_i} - \frac{\left[\sum_{i=1}^{b}\frac{rn_i^2(\bar{y}_{i.} - \bar{y}_{..})}{1 + rn_i}\right]^2}{\sum_{i=1}^{b}\frac{n_i}{1 + rn_i}}}{\sum_{i=1}^{b}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2}. \qquad (20)$$

Remark 3.1 : Spjøtvoll(1967) considered the hypothesis $H_0 : \rho \le r$ against $H_1 : \rho = r_1$ where $r_1 > r$. For this testing problem, he derived test statistic $W(r,r_1)$ which depend on r_1 and is known to be most powerful similar invariant test. For testing $H_0 : \rho \le r$ against $H_1 : \rho > r$, he used $W(r,\infty)$ which was denoted by $T(r)$ and is equal to our test statistic.

Remark 3.2 : When the design is balanced, the last term of (19) is zero and our test procedure is UMPU and UMPIU.

4. Special case 2.

By a similar argument we can obtain a test procedure for each variance components under certain conditions on design matrices. Suppose $V_iV_j = 0$ for all $i \neq j$, for example, balanced factorial designs. Then there exists an orthogonal matrix P such that

$$V_i = P\Lambda_i P' \text{ for } i = 1,2,\ldots,k \tag{21}$$

with Λ_i being the diagonal matrix of eiegen values of V_i. Since V_i's are orthogonal to each other, if s-th diagonal element of Λ_i is nonzero, then the corresponding diagonal element of $\Lambda_j, j \neq i$ is equal to zero.

Suppose that we wish to test $H_0 : \rho_1 \le r$ against $H_1 : \rho_1 > r$. We may assume that $\Lambda_1 = diag(\lambda_1,\ldots,\lambda_q,0,\ldots,0)$, where q is the rank of V_1. Then

$$(I + rP\Lambda_1P')^{-1} = P(I + r\Lambda_1)^{-1}P' \tag{22}$$

and we can decompose $(I + r\Lambda_1)^{-1}$ into D_1, and D_2 as follows:

$$(I + r\Lambda_1)^{-1} = D_1 + D2 \tag{23}$$

where $D_1 = diag({}^1/_1 + r\lambda_1,\ldots,{}^1/_1 + r\lambda_q,0,\ldots,0)$, and $D_2 = diag(0,\ldots,0,1,\ldots,1)$.

Writing $D_1 = D_1^{1/2}D_1^{1/2}$, we have

$$Var(D_1^{1/2}P'z/\sigma_{k+1}) = D_1^{1/2}P'(I + \sum_{i=1}^{k} \rho_i V_i)PD_1^{1/2}$$

$$= D_1^{1/2}P'(I + \sum_{i=1}^{k} \rho_i P\Lambda_i P')PD_1^{1/2}$$

$$= D^{1/2}{}_1P'P(I + \sum_{i=1}^{k}\rho_i\Lambda_i)P'PD^{1/2}{}_1$$

$$= D^{1/2}{}_1(I + \rho_1\Lambda_1)D^{1/2}{}_1$$

$$= diag\left\{ \frac{1+\rho_1\lambda_1}{1+r\lambda_1}, \ldots, \frac{1+\rho_1\lambda_q}{1+r\lambda_{1q}}, 0, \ldots, 0 \right\} \tag{24}$$

and

$$D_1 Var(P'z)D_2 = 0 \tag{25}$$

Thus $z'PD_1P'z/\sigma_{k+1}$ is distributed as $\sum^q_{i=1}(1+\rho_1\lambda_1)X^2{}_i/(1+r\lambda_1)$ and independent of $z'PD_2P'z/\sigma_{k+1}$. On the boundary of hypotheses, the distribution of $F = (n-h-1)z'/q_iPD_{i1}P'z/z'PD_2P'z$ is $F(q,n-h-1)$. thus we reject the null hypothesis when $F > F_\alpha(q, n-h-1)$.

The test is unbiased level α test and power of test is given by

$$P\left[\sum_{i=1}^{q}\{(1+\rho_1\lambda_i)/(1+r\lambda_i)\} X^2_i > F_\alpha(q,n-h-1)q/(n-h-1)\ X^2\right] \qquad (26)$$

References

Farebrother, R. W.(1984), The distribution of a linear combination of central χ^2 random variables, Applied Statistics, **33**, 363-366.

Graybill, F. A.(1976), Theory and Application of the Linear Model. Duxbury Press. Belmont, CA.

Khuri, A. I. and R. C. Littel (1987), Exact tests for the main effects of variance components in an unbalanced random two-way model, Biometrics, **43**,545-560.

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.

Rao, C. R. (1971), Estimation of Variance and Covariance Components-MINQUE Theory, Journal of Multivariate Analysis 1, 257-275.

Spjøtvoll, E. (1967), Optimum Invariant Tests in Unbalanced Variance Components Models, Ann. Math. Statist. **38**, 422-428

Spjøtvoll, E.(1968), Confidence Intervals and Tests for Variance ratios in unbalanced variance components models, Rev. Internat. Statist. Inst., 37-42

Thomsen. I. (1975), Testing Hypothesis in Unbalanced Variance Components Models for Two-way Layouts, The Annals of Statist., **3**, 257-265.

Verdooren, L. R. (1988), Exact Test and Confidence Intervals for ratio of Variance Components in Unbalanced Two- and Three-stage Nested Designs, Comm. Statist.- Theory Math., 17,1197-1230