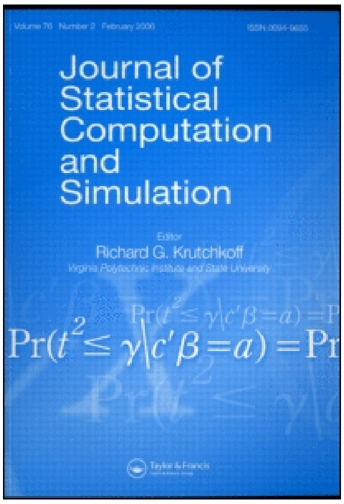

Journal of Statistical Computation and Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713650378

The equivalence between likelihood ratio test and F-test for testing variance component in a balanced one-way random effects model

Yan Lu [a]; Guoyi Zhang [a]
[a] Department of Mathematics and Statistics, University of New Mexico, Albuquerque, New Mexico, USA

First Published on: 03 March 2009

Journal of Statistical Computation and Simulation
Vol. 00, No. 0, Month 2009, 1–8

The equivalence between likelihood ratio test and F-test for testing variance component in a balanced one-way random effects model

Yan Lu* and Guoyi Zhang

Department of Mathematics and Statistics, University of New Mexico, Albuquerque, New Mexico, USA

(*Received 24 July 2008; final version received 16 December 2008*)

In mixed linear models, it is frequently of interest to test hypotheses on the variance components. F-test and likelihood ratio test (LRT) are commonly used for such purposes. Current LRTs available in literature are based on limiting distribution theory. With the development of finite sample distribution theory, it becomes possible to derive the exact test for likelihood ratio statistic. In this paper, we consider the problem of testing null hypotheses on the variance component in a one-way balanced random effects model. We use the exact test for the likelihood ratio statistic and compare the performance of F-test and LRT. Simulations provide strong support of the equivalence between these two tests. Furthermore, we prove the equivalence between these two tests mathematically.

Keywords: finite sample distribution; hypothesis; mixed model; simulations; variance component

1. Introduction

This paper assumes a one-way random effects model

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \tag{1}$$

where $i = 1, \ldots, m$; $j = 1, \ldots, k$. μ is the fixed unknown intercept, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m)'$ is the random effect and $\boldsymbol{\varepsilon} = (\varepsilon_{11}, \varepsilon_{12}, \ldots, \varepsilon_{mk})'$ is the error term. Assume $\boldsymbol{\alpha}$ and $\boldsymbol{\varepsilon}$ are normally and independently distributed with mean $\mathbf{0}$ and variances $\sigma_\alpha^2 \mathbf{I}_m$ and $\sigma^2 \mathbf{I}_{mk}$. A standard test of the variance component σ_α^2 is as the following:

$$\mathbf{H}_0 : \sigma_\alpha^2 = 0 \quad vs. \quad \mathbf{H}_a : \sigma_\alpha^2 > 0. \tag{2}$$

F-test is commonly used for this situation because, in this case, F-test is a uniformly most powerful unbiased test. On the other hand, likelihood ratio test (LRT) is a well-known and widely

*Corresponding author. Email: luyan@math.unm.edu

ISSN 0094-9655 print/ISSN 1563-5163 online

DOI: 10.1080/00949650802695664
http://www.informaworld.com

used statistical test. One problem of test (2) is that zero is at the boundary of the parameter space, so the limiting distribution of the likelihood ratio statistic is not χ^2. Hartly and Rao [1] stated without giving a proof that the asymptotic distribution of $-2\log \mathbf{L}$ is a central χ^2. Other papers related to asymptotic distribution of likelihood ratio statistic include Stram and Lee [2], Shephard and Harvey [3] and Stern and Welsh [4]. χ^2 mixture [2] is another way to approximate the distribution of likelihood ratio statistic and it works well when the number of independent groups is large. There was another test called locally optimal test proposed by Westfall [5,6], which followed papers by Harville and Fenech [7] and Seely and El-Bassinouni [8]. Westfall [5] compared the locally optimal test to the F-test for unbalanced designs. However, "the literature on LRT in the context of linear mixed models is much less extensive" [9, p. 55]. Recently, Crainiceanu and Ruppert [10] derived finite sample distribution of likelihood ratio statistics in linear mixed models, which makes it possible to derive the exact test for likelihood ratio statistics. In this article, we consider the standard test (2) in a one-way balanced random effects model. We discover that F-test and LRT are equivalent by simulation studies. Furthermore, we prove the equivalence between the two tests in theory.

This paper is organized as follows. In Section 2, we review F-test and LRT. In Section 3, we report our simulation results. Finally, a proof of the equivalence between the two tests is given in Section 4.

2. Background

We first introduce some notation. Define $\bar{y}_{..} = \sum_{i=1}^{m}\sum_{j=1}^{k} y_{ij}/mk$, $\bar{y}_{i.} = \sum_{j=1}^{k} y_{ij}/k$ for each i, $\mathrm{SSE} = \sum_{i=1}^{m}\sum_{j=1}^{k}(y_{ij} - \bar{y}_{i.})^2/(k-1)m$, $\mathrm{SSB} = k\sum_{i=1}^{m}(\bar{y}_{i.} - \bar{y}_{..})^2$, $\mathrm{MSB} = \mathrm{SSB}/(m-1)$ and $\mathrm{MSE} = \mathrm{SSE}/(m(k-1))$.

For model (1), the ratio of $\mathrm{MSB}/(\sigma^2 + k\sigma_\alpha^2)$ to MSE/σ^2 has an F-distribution with degrees of freedom $(m-1, m(k-1))$. Under $\mathbf{H}_0$ in Equation (2), MSB/MSE has an F-distribution with degrees of freedom $(m-1, m(k-1))$.

Before we discuss the LRT, we first write model (1) in matrix form,

$$\mathbf{Y} = \mathbf{X}\mu + \mathbf{Z}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \tag{3}$$

where $\mathbf{X}$ is simply an $mk \times 1$ vector of 1s, $\mathbf{Z}$ is an $mk \times m$ matrix with every column containing only 0s with exception of a k-dimensional vector of 1s corresponding to the level parameter, $\mathbf{Y}$ is the response vector and $\boldsymbol{\varepsilon}$ is the random error vector.

Twice the log-likelihood function of Equation (3), we have

$$2\log\{\mathbf{L}(\mu, \sigma_\alpha^2, \sigma^2)\} = -\log(\sigma^2) - \log|\mathbf{V}| - \frac{(\mathbf{Y} - \mathbf{X}\mu)^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\mu)}{\sigma^2},$$

where $\mathbf{V} = \mathbf{I}_{mk} + \lambda\mathbf{Z}\mathbf{Z}^T$, $\lambda = \sigma_\alpha^2/\sigma^2$ and the likelihood ratio statistic is defined as

$$\mathrm{LRT} = 2\sup_{H_a}\{\mathbf{L}(\mu, \sigma_\alpha^2, \sigma^2)\} - 2\sup_{H_0}\{\mathbf{L}(\mu, \sigma_\alpha^2, \sigma^2)\}. \tag{4}$$

The standard maximum likelihood estimators are as follows

$$\hat{\mu} = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{Y},$$

and

$$\hat{\sigma}^2 = \frac{(\mathbf{Y} - \mathbf{X}\hat{\mu})^T \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{X}\hat{\mu})}{mk}.$$

Under the null hypothesis, we obtain the likelihood estimators as follows

$$\hat{\mu} = \bar{y}_{..},$$

and

$$\hat{\sigma}_0^2 = \frac{1}{km} \sum_{i=1}^{m} \sum_{j=1}^{k} (y_{ij} - \bar{y}_{..})^2.$$

Under the alternative hypothesis, we obtain the likelihood estimators as follows

$$\hat{\mu} = \bar{y}_{..},$$

and

$$\hat{\sigma}^2 = \frac{1}{(k-1)m} \sum_{i=1}^{m} \sum_{j=1}^{k} (y_{ij} - \bar{y}_{i.})^2.$$

$$\hat{\sigma}_\alpha^2 = \begin{cases} \dfrac{1}{k}\left(\dfrac{k\sum_{i=1}^{m}(\bar{y}_{i.} - \bar{y}_{..})^2}{m} - \hat{\sigma}^2\right) & \text{if} \quad \hat{\sigma}^2 \leq \dfrac{k\sum_{i=1}^{m}(\bar{y}_{i.} - \bar{y}_{..})^2}{m}, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

3. Simulations

We first introduce the result from Crainiceanu and Ruppert [10], which we use in our simulation study. Crainiceanu and Ruppert [10] developed a method to get finite sample distributions of likelihood ratio statistics, which follows that

$$\begin{aligned} \text{LRT} &\overset{D}{=} km \log(X_{m-1} + X_{(k-1)m}) \\ &\quad - \inf_{d \geq 0} \left\{ km \log\left(\frac{X_{m-1}}{1+d} + X_{(k-1)m}\right) + m \log(1+d) \right\}, \end{aligned} \tag{6}$$

where notation $\overset{D}{=}$ denotes equivalence in distribution and X_{m-1} and $X_{(k-1)m}$ are independent random variables with distribution χ^2_{m-1} and $\chi^2_{(k-1)m}$.

In this section, we perform a simulation study to investigate the two tests. First, we compare the percentages of samples for which the test statistics exceed the critical value to the nominal level. Then we calculate power of the two tests. The following are the simulation details:

(1) Calculate critical values for both tests. We use Equation (6) to generate finite sample distribution of likelihood ratio statistic and to find corresponding critical values. For each setting, we use a different seed and generate 100,000 samples from Equation (6). Critical value is the $100(1-\gamma)\%$ percentile, where γ is the significant level of test. Critical value for F-test is $F_\gamma[m-1, m(k-1)]$. Results are listed in Table 1.

Table 1. Critical values for the two tests.

Significance level (γ)	k	Tests	m: 2	10	50
0.01	5	F	11.25862	2.887560	1.634977
		LRT	3.062403	4.361615	4.91984
	10	F	8.28542	2.610879	1.576229
		LRT	2.590879	4.137277	4.707984
	20	F	7.352545	2.501878	1.551395
		LRT	2.444884	4.024391	4.670615
0.05	5	F	5.317655	2.124029	1.418051
		LRT	0.9052984	1.877785	2.335454
	10	F	4.413873	1.985595	1.382671
		LRT	0.6955273	1.764155	2.228158
	20	F	4.098172	1.929425	1.367567
		LRT	0.6283923	1.705018	2.225503
0.1	5	F	3.457919	1.792902	1.312488
		LRT	0.2665952	0.990356	1.362770
	10	F	3.006977	1.702053	1.286975
		LRT	0.1779033	0.9079695	1.27152
	20	F	2.842442	1.664704	1.276034
		LRT	0.1358003	0.8746464	1.274283

(2) Compare to the nominal levels. The following model is used to generate samples:

$$y_{ij} = 0.5 + \alpha_i + \varepsilon_{ij}, \tag{7}$$

where $i = 1, \ldots, m$; $j = 1, \ldots, k$. $\boldsymbol{\alpha}$ and $\boldsymbol{\varepsilon}$ are normally and independently distributed with mean $\mathbf{0}$ and variance $\sigma_\alpha^2 \mathbf{I}_m$, $\sigma^2 \mathbf{I}_{mk}$, respectively.

For each setting, we use a different seed and generate 100,000 samples from model (7) with $\sigma_\alpha^2 = 0$ and $\sigma^2 = 1$. For each sample, we apply F-test and LRT and the percentages of samples for which the test statistics exceed the critical value are reported in Table 2.

Table 2. Tests comparison (the numbers in the table are the percentages of samples for which the test statistics exceed the critical value).

Significance level (γ)	k	Tests	m: 2	10	50
0.01	5	F	0.01015	0.01026	0.01009
		LRT	0.01034	0.01006	0.00973
	10	F	0.00958	0.00977	0.01013
		LRT	0.00978	0.0096	0.01061
	20	F	0.00989	0.01041	0.00945
		LRT	0.00970	0.01041	0.00992
0.05	5	F	0.05028	0.04888	0.05037
		LRT	0.05049	0.04959	0.04943
	10	F	0.05022	0.05015	0.05001
		LRT	0.05025	0.05038	0.05107
	20	F	0.04938	0.0496	0.05036
		LRT	0.04840	0.05003	0.05039
0.1	5	F	0.09926	0.10016	0.10066
		LRT	0.09921	0.10123	0.09809
	10	F	0.10153	0.09906	0.10078
		LRT	0.0996	0.10006	0.10221
	20	F	0.10089	0.10019	0.10003
		LRT	0.09994	0.10062	0.0998

Both tests give almost the same results and work very well as we can see that the percentages of samples for which the test statistics exceed the critical value are very close to the nominal level.

(3) Calculate power of the tests. We generate 100,000 samples from model (7) with $\sigma_\alpha^2 = 0.09, 1, 9$ and $\sigma^2 = 1$ for each setting using a different seed. The results are reported in Tables 3–5. We can see that the two tests almost have the same power. Equation (8) can also be used to calculate power.

$$P\left(\frac{\mathrm{MSB}}{\mathrm{MSE}} > F_\gamma | \sigma_\alpha^2 > 0\right) = P\left(\frac{\mathrm{MSB}/(\sigma^2 + k\sigma_\alpha^2)}{\mathrm{MSE}/\sigma^2} > \frac{\sigma^2}{\sigma^2 + k\sigma_\alpha^2} F_\gamma\right), \tag{8}$$

where γ is the significance level, $F_\gamma[m-1, m(k-1)]$ is the critical value.

Table 3. Power of the tests $\lambda = 0.09$.

			m		
Significance level (γ)	k	Tests	2	10	50
0.01	5	F	0.02286	0.06682	0.27836
		LRT	0.02332	0.06565	0.27426
	10	F	0.05208	0.21231	0.78583
		LRT	0.05301	0.21117	0.78968
	20	F	0.11299	0.53116	0.99473
		LRT	0.11186	0.53112	0.99490
0.05	5	F	0.09213	0.19207	0.52209
		LRT	0.09244	0.19354	0.51860
	10	F	0.14611	0.41039	0.91525
		LRT	0.14613	0.41108	0.91625
	20	F	0.23492	0.72211	0.99865
		LRT	0.23280	0.72282	0.99865
0.1	5	F	0.16075	0.30051	0.65041
		LRT	0.16056	0.30251	0.64618
	10	F	0.22725	0.53221	0.95468
		LRT	0.22403	0.53373	0.95526
	20	F	0.31735	0.79945	0.99968
		LRT	0.31589	0.79985	0.99966

Table 4. Power of the tests $\lambda = 1$.

			m		
Significance level (γ)	k	Tests	2	10	50
0.01	5	F	0.20632	0.87851	1.00000
		LRT	0.20821	0.87713	1.00000
	10	F	0.39500	0.98768	1.00000
		LRT	0.39680	0.98756	1.00000
	20	F	0.55872	0.99917	1.00000
		LRT	0.55755	0.99917	1.00000
0.05	5	F	0.37424	0.94984	1.00000
		LRT	0.37472	0.95021	1.00000
	10	F	0.53460	0.99604	1.00000
		LRT	0.53461	0.99605	1.00000
	20	F	0.65973	0.99965	1.00000
		LRT	0.65829	0.99965	1.00000
0.1	5	F	0.47081	0.97072	1.00000
		LRT	0.47068	0.97096	1.00000
	10	F	0.60843	0.99737	1.00000
		LRT	0.60605	0.99739	1.00000
	20	F	0.71258	0.99984	1.00000
		LRT	0.71187	0.99984	1.00000

Table 5. Power of the tests λ = 9.

			m		
Significance level (γ)	k	Tests	2	10	50
0.01	5	F	0.63433	0.99994	1.00000
		LRT	0.63573	0.99994	1.00000
	10	F	0.76535	1.00000	1.00000
		LRT	0.76617	1.00000	1.00000
	20	F	0.84207	1.00000	1.00000
		LRT	0.84153	1.00000	1.00000
0.05	5	F	0.74275	0.99998	1.00000
		LRT	0.74298	0.99998	1.00000
	10	F	0.82701	1.00000	1.00000
		LRT	0.82702	1.00000	1.00000
	20	F	0.88116	1.00000	1.00000
		LRT	0.88054	1.00000	1.00000
0.1	5	F	0.79241	0.99999	1.00000
		LRT	0.79236	0.99999	1.00000
	10	F	0.85743	1.00000	1.00000
		LRT	0.85654	1.00000	1.00000
	20	F	0.90209	1.00000	1.00000
		LRT	0.90180	1.00000	1.00000

For example, let $m = 2, k = 5, \gamma = 0.01, \lambda = 0.09$, we have $\sigma^2 F_\gamma/(\sigma^2 + k\sigma_\alpha^2) = 7.764566$ and power of the test from Equation (8) is 0.02368, which is close to the empirical power reported in Table 3.

4. Proof of the equivalence

THEOREM 4.1 *For model* (1), *the LRT statistic Equation* (4) *is a one to one function of* MSB/MSE. *Hence LRT is equivalent to F-test.*

Proof We first prove that the likelihood ratio statistic is a one-to-one function of F statistic by two cases: $\hat{\sigma}_\alpha^2 > 0$ and $\hat{\sigma}_\alpha^2 = 0$.

Case 1 $\hat{\sigma}_\alpha^2 > 0$

Since $\hat{\sigma}^2 = \text{MSE}$, $\hat{\sigma}_\alpha^2 = 1/k(((m-1)/m)\text{MSB} - \text{MSE})$ and $\text{SSE} + \text{SSB} = \sum_{i=1}^m \sum_{j=1}^k (y_{ij} - \bar{y}_{..})^2$,

$$\text{LRT} = -(mk - m)\log(\hat{\sigma}^2) - \sum_{i=1}^m \log(\hat{\sigma}^2 + k\hat{\sigma}_\alpha^2) - \frac{1}{\hat{\sigma}^2}\sum_{i=1}^m\sum_{j=1}^k (y_{ij} - \hat{\mu})^2$$

$$= -(mk - m)\log(\hat{\sigma}^2) - m\log(\hat{\sigma}^2 + k\hat{\sigma}_\alpha^2) - \frac{mk\hat{\sigma}_0^2}{\hat{\sigma}^2}$$

$$+ \frac{\hat{\sigma}_\alpha^2}{\hat{\sigma}^2}\left(\frac{k^2}{\hat{\sigma}^2 + k\hat{\sigma}_\alpha^2}\right)\sum_{i=1}^m (\bar{y}_{i.} - \bar{y}_{..})^2 + mk\log(\hat{\sigma}_0^2) + mk$$

$$= -(mk - m)\log\left(\frac{\text{SSE}}{(k-1)m}\right) - m\log\left(\frac{\text{SSB}}{m}\right) - \frac{(k-1)m(\text{SSE}+\text{SSB})}{\text{SSE}}$$
$$+ \frac{m^2(k-1)(\text{SSB}/m - \text{SSE}/(k-1)m)}{\text{SSE}} + mk\log\left(\frac{1}{mk}(\text{SSE}+\text{SSB})\right) + mk$$
$$= c + mk\log(\text{SSE}+\text{SSB}) - mk\log(\text{SSE}) + m\log(\text{SSE}) - m\log(\text{SSB})$$
$$= c + m\left[\log\left(1 + c^*\frac{\text{MSB}}{\text{MSE}}\right)^k - \log\left(c^*\frac{\text{MSB}}{\text{MSE}}\right)\right]$$
$$= c - m\log(c^*) + m\log\left(\frac{(1+c^*\text{MSB}/\text{MSE})^k}{\text{MSB}/\text{MSE}}\right) \tag{9}$$

where c and $c^* = (m-1)/((k-1)m)$ are constants. Since $\hat{\sigma}_\alpha^2 > 0$, we have $\text{MSB}/\text{MSE} > m/(m-1)$. If $x > m/(m-1)$ and $k \geq 2$, derivative of function $f(x) = (1+c^*x)^k/x$ is positive. So LRT is a strictly increasing function.

Case 2 $\hat{\sigma}_\alpha^2 = 0$

In this case,

$$\text{LRT} = -(mk-m)\log(\hat{\sigma}^2) - \sum_{i=1}^{m}\log(\hat{\sigma}^2) - \frac{1}{\hat{\sigma}^2}\sum_{i=1}^{m}\sum_{j=1}^{k}(y_{ij}-\hat{\mu})^2$$
$$+ mk\log(\hat{\sigma}_0^2) + \frac{1}{\hat{\sigma}_0^2}\sum_{i=1}^{m}\sum_{j=1}^{k}(y_{ij}-\hat{\mu})^2$$
$$= -(mk-m)\log\left(\frac{\text{SSE}}{(k-1)m}\right) - m\log\left(\frac{\text{SSE}}{m(k-1)}\right) - \frac{(k-1)m(\text{SSE}+\text{SSB})}{\text{SSE}}$$
$$+ mk\log\left(\frac{1}{mk}(\text{SSE}+\text{SSB})\right) + mk$$
$$= c' - (k-1)m\frac{\text{SSB}}{\text{SSE}} + mk\log\left(1 + \frac{\text{SSB}}{\text{SSE}}\right)$$
$$= c' - (m-1)\frac{\text{MSB}}{\text{MSE}} + mk\log\left(1 + c^*\frac{\text{MSB}}{\text{MSE}}\right)$$

where c' and $c^* = (m-1)/((k-1)m)$ are constants. Since $\hat{\sigma}_\alpha^2 = 0$, we have $\text{MSB}/\text{MSE} \leq m/(m-1)$. If $x \leq m/(m-1)$ and $m, k \geq 2$, derivative of function $f(x) = -(m-1)x + mk\log(1+c^*x)$ is positive. So LRT in Case 2 is also a strictly increasing function.

From the above proof, we conclude that the likelihood ratio statistic is a one-to-one function of F statistic MSB/MSE under both cases $\hat{\sigma}_\alpha^2 > 0$ and $\hat{\sigma}_\alpha^2 = 0$.

The proof of equivalence can be obtained by proving that the two tests have the same results of rejection or acceptance. Consider Case 1: $\hat{\sigma}_\alpha^2 > 0$. Given an arbitrary significance level γ, let F_γ be the critical value of F-test and L_γ be the critical value of the LRT. Let $\text{LRT} = g(\text{MSB}/\text{MSE})$, where g represents the one-to-one continuous increasing function. Clearly $L_\gamma = g(F_\gamma)$ and the statement $\text{MSB}/\text{MSE} > F_\gamma$ is equivalent to the statement $g(\text{MSB}/\text{MSE}) > g(F_\gamma)$ i.e. $\text{LRT} > L_\gamma$. Proof of equivalence in Case 2: $\hat{\sigma}_\alpha^2 = 0$ can be obtained similarly. ■

References

[1] H.O. Hartley and J.N.K. Rao, *Maximum-likelihood estimation for mixed analysis of varince model*, Biometrika 54 (1967), pp. 93–108.
[2] D.O. Stram and J.W. Lee, *Variance components tesing in the longitudinal mixed effects model*, Biometrics 50 (1994), pp. 1171–1177.
[3] N.G. Shephard and A.C. Harvey, *On the probability of estimating a deterministic component in the local level model*, J. Time Ser. Anal 4 (1990), pp. 339–347.
[4] S.E. Stern and A.H. Welsh, *Likelihood inference for small variance components*, Can. J. Statist. 28 (2000), pp. 517–532.
[5] P.H. Westfall, *Robustness and power of tests for a null variance ratio*, Biometrika 75 (1988), pp. 207–214.
[6] P.H. Westfall, *Power comparisons for invariant variance ratio tests in mixed anova models*, Ann. Statist. I (1989), pp. 318–326.
[7] D.A. Harville and A.P. Fenech, *Confidence intervals for a varinace ratio, or for heritability, in an unbalanced mixed linear model*, Biometrics 41 (1985), pp. 137–152.
[8] J.F. Seely and Y. El-Bassinouni, *Applying Wald's variance component test*, Ann. Statist. II (1983), pp. 197–201.
[9] J. Jiang, *Linear and Generalized Linear Mixed Models and Their Applications*, Springer, New York, 2007.
[10] C.M. Crainiceanu and D. Ruppert, *Likelihood ratio tests in linear mixed models with one variance component*, J. R. Statist. Soc. B 66 (2004), pp. 165–185.