Ann. Inst. Statist. Math.
Vol. 43, No. 4, 707–714 (1991)

JACKKNIFE VARIANCE ESTIMATORS OF THE LOCATION ESTIMATOR IN THE ONE-WAY RANDOM-EFFECTS MODEL*

YOUNGJO LEE

Department of Statistics, Hallym University, Chunchon, Korea

(Received October 11, 1989; revised July 6, 1990)

Abstract. In a one-way random-effects model, we frequently estimate the variance components by the analysis-of-variance method and then, assuming the estimated values are true values of the variance components, we estimate the population mean. The conventional variance estimator for the estimate of the mean has a bias. This bias can become severe in contaminated data. We can reduce the bias by using the delta method. However, it still suffers from a large bias. We develop a jackknife variance estimator which is robust with respect to data contamination.

Key words and phrases: Jackknife, one-way random-effect models, robustness, Monte Carlo study.

1. Introduction

Consider the one-way random-effects model,

$$y_{ij} = \mu + a_i + e_{ij}, \qquad i = 1, \ldots, k, \quad j = 1, \ldots, n_i,$$

where the a_i and e_{ij} are independent random variables with $a_i \sim N(0, \sigma_a^2)$ and $e_{ij} \sim N(0, \sigma_e^2)$. Let $\pi = \sigma_e^2 + \sigma_a^2$ and $\rho = \sigma_a^2/\pi$. When ρ is known, the best linear unbiased estimator of μ can be written as $\hat{\mu}(\rho) = \sum w_i(\rho)\bar{y}_i$ (Birkes *et al.* (1981)), where

$$w_i(\rho) = [n_i/\{(n_i - 1)\rho + 1\}] \Big/ \left[\sum n_i/\{(n_i - 1)\rho + 1\}\right] \qquad \text{and}$$
$$\bar{y}_i = \sum_j y_{ij}/n_i.$$

Then, the variance of the best linear unbiased estimator is

$$\operatorname{var}(\hat{\mu}(\rho)) = \pi \sum W_i(\rho)^2(\rho + (1 - \rho)/n_i).$$

* This research was supported by the Korea Science and Engineering Foundation.

However, in practice, ρ is usually unknown. In this case, we first estimate ρ by some method. Then, pretending that the estimate $\hat{\rho}$ is the true parameter, we usually use the estimator $\hat{\mu}(\hat{\rho}) = \sum w_i(\hat{\rho})\bar{y}_i$, and estimate its variance as

$$\widehat{\mathrm{var}}(\hat{\mu}(\hat{\rho})) = \hat{\pi} \sum w_i(\hat{\rho})^2(\hat{\rho} + (1-\hat{\rho})/n_i). \tag{1.1}$$

Kackar and Harville (1984) showed that under normality, the following equation holds:

$$\mathrm{var}(\hat{\mu}(\hat{\rho})) = \mathrm{var}(\hat{\mu}(\rho)) + E(\hat{\mu}(\hat{\rho}) - \hat{\mu}(\rho))^2. \tag{1.2}$$

This equation is true in general mixed linear models if the variance components estimators are translation invariant, which is true of most of them. Hence, they argued that the conventional variance estimator (1.1) might suffer from a downward bias since it estimates only the first term of (1.2). They suggested using the estimator (1.1) for estimating the first term of the equation (1.2) and the δ-method for estimating the second term.

However, their discussion is mainly about the cases when ρ is estimated either by the maximum likelihood method or the restricted maximum likelihood method. Furthermore, their numerical studies are confined to balanced designs. Estimating variance components by maximum likelihood requires iteration and the computing would be difficult and might even be infeasible for large data sets.

The most frequently used estimators of variance components have been analysis-of-variance estimators. For the one-way random-effects model, Swallow (1981) and Swallow and Monahan (1984) conducted Monte Carlo comparisons of various variance components estimators and found that the analysis-of-variance estimator is adequate unless the design is severely unbalanced and $\rho > 0.5$. They also have other appealing properties: they are familiar, easy to compute, and are unbiased. Hence, it would be interesting to develop the variance estimator of the conventional location estimator $\hat{\mu}(\hat{\rho})$.

2. The delta method

In the balanced one-way random-effects model, the best linear unbiased estimator $\hat{\mu}(\rho)$ and the conventional estimator $\hat{\mu}(\hat{\rho})$ are identical to the overall mean $\bar{y} = \sum\sum y_{ij}/n$ $(n = \sum n_i)$. Hence, the second term of (1.2) vanishes. However, in unbalanced designs, the second term is no longer equal to zero and the estimator $\hat{\mu}(\hat{\rho})$ depends on the variance ratio estimator $\hat{\rho}$. The analysis-of-variance estimators are obtained by equating observed and expected mean squares and solving the resulting equations. In the one-way random-effects model, they are $s_e^2 = \mathrm{SSE}/(n-k)$ and $s_a^2 = (\mathrm{SSA} - (k-1)s_e^2) \,/\, (n - \sum n_i^2/n)$, where $\mathrm{SSE} = \sum(n_i - 1)s_i^2$, $s_i^2 = \sum_j (y_{ij} - \bar{y}_i)^2/(n_i - 1)$ and $\mathrm{SSA} = \sum n_i(\bar{y}_i - \bar{y})^2$ (Searle (1971), p. 474). Hence, the intergroup correlation ρ can be estimated as $\hat{\rho} = s_a^2/(s_a^2 + s_e^2)$. Using the Taylor series expansion, we have

$$\hat{\mu}(\hat{\rho}) \simeq \hat{\mu}(\rho) + (d\hat{\mu}(\rho)/d\rho)\{(d\rho/d\sigma_e^2)(s_e^2 - \sigma_e^2) + (d\rho/d\sigma_a^2)(s_a^2 - \sigma_a^2)\}. \tag{2.1}$$

Therefore, the second term of (1.2) can be approximated by

$$(2.2)\quad E(\hat{\mu}(\hat{\rho})-\hat{\mu}(\rho))^2 \simeq (d\hat{\mu}(\rho)/d\rho)^2\{(d\rho/d\sigma_e^2)^2\operatorname{var}(s_e^2)+(d\rho/d\sigma_a^2)^2\operatorname{var}(s_a^2) + 2(d\rho/d\sigma_e^2)(d\rho/d\sigma_a^2)\operatorname{cov}(s_e^2, s_a^2)\}.$$

In (2.2), we regard the analysis-of-variance estimator s_a^2 as unbiased. However, we usually enforce non-negativity in the estimation of σ_a^2, setting any negative estimate to zero. In this case, the unbiasedness property of the ANOVA estimator no longer holds. Therefore, one might consider replacing $\operatorname{var}(s_a^2)$ in (2.2) by the mean squared error of s_a^2. However, the problem of deriving the exact mean squared error is very difficult. Hence, we use the equation (2.2) to approximate the second term of (1.2); replacing terms in (2.2) with appropriate estimators, we obtain an estimator of the second term of (1.2). An estimator for (2.2) and its derivation can be found in the Appendix.

3. Jackknife methods

Arvensen and Schmitz (1970) and Arvensen and Layard (1975) applied a standard jackknife method to the one-way random-effects model for testing of the variance components ratio. Hinkley (1977) showed that the infinitesimal jackknife method provides a robust variance estimator of regression coefficients even with a heteroscedastic error structure. Wu (1986) suggested a slight modification in the weights of the infinitesimal jackknife method and extended its application to nonlinear regression models. However, in our case, only Hinkley's method will be considered, since the performance of Wu's jackknife variance estimator is very similar. For nonlinear regression models, the performances of various jackknife methods have been numerically studied by Simonoff and Tsai (1986). These studies show that the jackknife samples do not have to be distributed identically to obtain a good variance estimator.

Let the jackknife samples $X_1, X_2, \ldots, X_k$ be such that $X_i = (\bar{y}_i, s_i^2)$. Note that X_i's are independent but not identically distributed unless the design is balanced. Arvensen and Layard (1975) used these jackknife samples for deriving a test on the variance components ratio. The standard delete-one jackknife variance estimator is

$$\widehat{\operatorname{var}}_{SJ}(\hat{\mu}(\hat{\rho})) = \sum [J_i - J]^2/k(k-1),$$

where $J_i = k\hat{\mu}(\hat{\rho}) - (k-1)\hat{\mu}(\hat{\rho})_{(i)}$, $J = \sum J_i/k$ and $\hat{\mu}(\hat{\rho})_{(i)}$ is the estimator $\hat{\mu}(\hat{\rho})$ when the i-th sample X_i is deleted. Following the notation of Efron (1982), let P_0 and Q_i be the resampling vectors such that $P_0 = (1/k, 1/k, \ldots, 1/k)$ and $Q_i = P_0 + \epsilon(\delta_i - P_0)$, where δ_i is the i-th coordinate vector. For any statistics $\hat{\theta}$, define $\hat{\theta}(P)$ to be the resampled value of $\hat{\theta}$ based on the resampling scheme P. Note that $\hat{\theta}(P_0)$ is $\hat{\theta}$. Then, the estimated influence function of $\hat{\mu}(\hat{\rho})$ at X_i can be shown to be

$$I(\hat{\mu}(\hat{\rho}), X_i) = \lim_{\epsilon\to 0}\{\hat{\mu}(Q_i, \hat{\rho}(Q_i)) - \hat{\mu}(P_0, \hat{\rho}(P_0))\}/\epsilon = I_{1i} + I_{2i},$$

where
$$I_{1i} = \lim_{\epsilon\to 0}\{\hat\mu(Q_i,\, \hat\rho(Q_i)) - \hat\mu(P_0,\, \hat\rho(Q_i))\}/\epsilon$$
$$= \lim_{\epsilon\to 0}\{\hat\mu(Q_i,\, \hat\rho(P_0)) - \hat\mu(P_0,\, \hat\rho(P_0))\}/\epsilon,$$
$$I_{2i} = \lim_{\epsilon\to 0}\{\hat\mu(P_0,\, \hat\rho(Q_i)) - \hat\mu(P_0,\, \hat\rho(P_0))\}/\epsilon \quad \text{and}$$
$$\hat\mu(P_1,\, \hat\rho(P_2)) = \sum w_j(\hat\rho(P_2))\bar y_j(P_1).$$

Note that I_{1i} can be viewed as $I_{1i} = I(\hat\mu(\rho),\, X_i)\mid_{\rho=\hat\rho}$, the estimated influence function for $\hat\mu(\rho)$ at X_i. Since $\hat\mu(\rho)$ is the simple regression estimator of the linear model
$$w(\rho)_i^{1/2}\bar y_i = w(\rho)_i^{1/2}\mu + f_i,$$
where $f_i \sim N\big(0,\, \pi \,/\, \left[\sum n_i/\{(n_i-1)\rho+1\}\right]\big)$, we can use Hinkley's method (1977). His estimated influence function of $\hat\mu(\rho)$ at X_i for the above simple regression model is $I(\hat\mu(\rho),\, X_i) = kw_i(\rho)(\bar y_i - \hat\mu(\rho))$. Hence, we use
$$I_{1i} = I(\hat\mu(\rho),\, X_i)\mid_{\rho=\hat\rho} \; = kw_i(\hat\rho)(\bar y_i - \hat\mu(\hat\rho)).$$

From the Appendix, we have
$$I_{2i} = d\hat\mu(\rho^*)/d\rho^* \mid_{\rho^*=\hat\rho} \{(d\hat\rho/ds_e^2)I(s_e^2,\, X_i) + (d\hat\rho/ds_a^2)I(s_a^2,\, X_i)\},$$
where
$$I(s_e^2,\, X_i) = \{k(n_i-1)(s_i^2 - s_e^2)\}/(n-k),$$
$$I(s_a^2,\, X_i) = I_{(s_a^2\ge 0)}\Big[-(k-1)I(s_e^2,\, X_i) + kn_i\Big\{(\bar y_i - \bar y)^2 - \Big(1 - 2n_i/n + \sum n_j^2/n^2\Big)s_a^2 - (1/n_i - 1/n)s_e^2\Big\}\Big] \Big/ \Big(n - \sum n_i^2/n\Big)$$
and $I_{()}$ is an indicator function. $I(s_e^2,\, X_i)$ and $I(s_a^2,\, X_i)$ are the estimated influence functions of s_e^2 and s_a^2, respectively. When s_a^2 is negative, we usually set it to zero. In this case, a little perturbation in the data does not affect the sign of s_a^2 and hence, $I(s_a^2,\, X_i) = 0$. Note that I_{2i} is equivalent to
$$I_{2i} = d\hat\mu(\rho)/d\rho\{(d\rho/d\sigma_e^2)I(s_e^2,\, X_i) + (d\rho/d\sigma_a^2)I(s_a^2,\, X_i)\}\mid_{\rho=\hat\rho,\sigma_e^2=s_e^2,\sigma_a^2=s_a^2}.$$

Hence, if we use $\sum I(s_e^2,\, X_i)^2/k(k-1)$, $\sum I(s_a^2,\, X_i)^2/k(k-1)$ and $\sum I(s_e^2,\, X_i)I(s_a^2,\, X_i)/k(k-1)$ as estimates of $\mathrm{var}(s_e^2)$, $\mathrm{var}(s_a^2)$ and $\mathrm{cov}(s_e^2,\, s_a^2)$, respectively, $\sum I_{2i}^2/k(k-1)$ can be considered as an estimator for (2.2). As Efron (1982) pointed out, there seems to be a real similarity in derivation between the infinitesimal jackknife and the δ-method. However, as we can see in the numerical study in the next section, the performances of these two estimators are quite different in a contaminated data set.

Equation (1.2) holds when the error assumptions are true. However, in general, the cross-product term would not vanish in contaminated data. Hence, there seem to be two possible infinitesimal jackknife estimators for the variance of the conventional location estimator $\hat{\mu}(\hat{\rho})$,

$$\widehat{\text{var}}_{IJ1}(\hat{\mu}(\hat{\rho})) = \sum (I_{1i} + I_{2i})^2/k(k-1) \quad \text{and}$$
$$\widehat{\text{var}}_{IJ2}(\hat{\mu}(\hat{\rho})) = \sum (I_{1i}^2 + I_{2i}^2)/k(k-1).$$

4. Monte Carlo comparisons

To compare variance estimators in a variety of designs, we generate models with $n_1 = \cdots = n_{k-2} = m_1$ and $n_{k-1} = n_k = m_2$, so that $(k,\, m_1,\, m_2)$ determines a design. Note that if $m_1 = m_2$, the resulting design is balanced. For a selected design, random variables X_i's are generated from distributions, $\bar{y}_i \sim N(0,\, \pi\{\rho + (1-\rho)/n_i\})$ and $(n_i - 1)s_i^2 \sim \pi(1-\rho)\chi^2_{n_i-1}$, using the IMSL subroutines GGNML and GGCHS. For the contaminated data, we generate $\bar{y}_1 \sim N(0,\, \pi\{100\rho + (1-\rho)/n_1\})$. This is equivalent to $a_1 \sim N(0,\, 100\sigma_a^2)$ and $a_i \sim N(0,\, \sigma_a^2)$ for $i = 2, \ldots, k$. Hence, in this case one group has a larger between-group variance than the rest. We have included π for completeness, but without loss of generality, in this paper, we take $\pi = 100$. In Table 1, the expected value of various variance estimators for $\hat{\mu}(\hat{\rho})$ and their standard deviations (in parentheses) are computed by the Monte Carlo method using 100,000 pseudo-random samples.

As mentioned in Section 2, for balanced cases the conventional variance estimator $\widehat{\text{var}}(\hat{\mu}(\hat{\rho}))$ is identical to the estimator based on the δ-method. Even in the balanced case, the conventional variance estimator $\widehat{\text{var}}(\hat{\mu}(\hat{\rho}))$ has upward bias due to enforcing nonnegativity in estimating σ_a^2. In Table 1, this bias gets relatively larger as the value of ρ becomes smaller, in which case the chance of getting a negative analysis of variance estimate for σ_a^2 increases. The bias becomes severe when data are contaminated in the unbalanced design. The δ-method only trivially improves the conventional variance estimator in these cases since the δ-method itself is essentially based on the conventional method. These two methods provide variance estimators with the smallest standard deviation when ρ is small. However, they suffer large bias.

When the design is balanced or $\rho = 1$, X_i's have identical distributions. In these cases, the standard and the infinitesimal jackknife methods provide the same variance estimators. Also, $\hat{\mu}(\hat{\rho}) = \sum \bar{y}_i/k$ and the jackknife variance estimator becomes $\sum(\bar{y}_i - \bar{y})^2/k(k-1)$, which is unbiased. Hence, we can expect that when ρ is large or the design is quite balanced, the biases of jackknife variance estimators are small and their performances are quite similar. It seems that the infinitesimal jackknife method automatically modifies itself to give an estimator identical to the standard jackknife estimator when data become balanced. In Table 1, we can see that these jackknife estimators are quite robust to contaminated data. When ρ is small, jackknife estimators have some bias. Among these jackknife estimators, the infinitesimal jackknife variance estimator $\widehat{\text{var}}_{IJ2}(\hat{\mu}(\hat{\rho}))$ has generally the smallest bias and smallest standard deviation. This bias seems to become negligible as the number of groups k increases, as we can see in the case $(k = 10,\, m_1 = 2,\, m_2 = 19)$. A Monte Carlo study for some other cases showed similar results.

Table 1. The expected value of various variance estimators.

ρ	0.05	0.25	0.50	0.75	0.95
			Design = ($k = 6$, $m_1 = 2$, $m_2 = 2$)		
Uncontaminated					
$\text{var}(\hat{\mu}(\hat{\rho}))$	8.78	10.46	12.55	14.64	16.31
$\widehat{\text{var}}(\hat{\mu}(\hat{\rho}))$	11.24 (.017)	11.65 (.019)	12.97 (.025)	14.88 (.030)	16.63 (.033)
$\widehat{\text{var}}_{SJ}(\hat{\mu}(\hat{\rho}))$	8.74 (.017)	10.41 (.021)	12.49 (.025)	14.57 (.029)	16.23 (.032)
Contaminated					
$\text{var}(\hat{\mu}(\hat{\rho}))$	22.6	79.4	150.3	221.2	277.9
$\widehat{\text{var}}(\hat{\mu}(\hat{\rho}))$	27.5 (.076)	79.8 (.321)	151.7 (.640)	224.3 (.960)	282.9 (1.218)
$\widehat{\text{var}}_{SJ}(\hat{\mu}(\hat{\rho}))$	22.4 (.072)	78.9 (.320)	149.6 (.631)	220.2 (.942)	276.7 (1.192)
			Design = ($k = 6$, $m_1 = 2$, $m_2 = 19$)		
Uncontaminated					
$\text{var}(\hat{\mu}(\hat{\rho}))$	4.51	7.91	11.11	13.97	16.17
$\widehat{\text{var}}(\hat{\mu}(\hat{\rho}))$	4.09 (.008)	6.98 (.016)	10.40 (.027)	13.70 (.038)	16.29 (.048)
δ-method	4.20 (.008)	7.06 (.016)	10.43 (.027)	13.70 (.038)	16.29 (.048)
$\widehat{\text{var}}_{SJ}(\hat{\mu}(\hat{\rho}))$	5.69 (.014)	8.61 (.018)	11.37 (.029)	13.98 (.028)	16.10 (.032)
$\widehat{\text{var}}_{IJ1}(\hat{\mu}(\hat{\rho}))$	5.47 (.014)	8.31 (.018)	11.26 (.029)	13.99 (.028)	16.11 (.032)
$\widehat{\text{var}}_{IJ2}(\hat{\mu}(\hat{\rho}))$	4.84 (.012)	7.68 (.017)	10.74 (.023)	13.69 (.028)	16.01 (.032)
Contaminated					
$\text{var}(\hat{\mu}(\hat{\rho}))$	16.5	74.1	146.5	219.0	277.0
$\widehat{\text{var}}(\hat{\mu}(\hat{\rho}))$	10.7 (.030)	33.7 (.124)	63.9 (.246)	94.1 (.369)	118.5 (.468)
δ-method	10.8 (.030)	33.8 (.124)	63.9 (.246)	94.1 (.369)	118.5 (.468)
$\widehat{\text{var}}_{SJ}(\hat{\mu}(\hat{\rho}))$	17.9 (.068)	75.0 (.316)	147.0 (.627)	219.1 (.937)	276.8 (1.185)
$\widehat{\text{var}}_{IJ1}(\mu(\rho))$	19.4 (.071)	77.1 (.319)	148.6 (.628)	220.0 (.938)	277.0 (1.186)
$\widehat{\text{var}}_{IJ2}(\hat{\mu}(\hat{\rho}))$	17.4 (.067)	74.7 (.315)	146.8 (.626)	219.0 (.937)	276.8 (1.185)
			Design = ($k = 10$, $m_1 = 2$, $m_2 = 19$)		
Uncontaminated					
$\text{var}(\hat{\mu}(\hat{\rho}))$	3.13	5.28	6.98	8.52	9.71
$\widehat{\text{var}}(\hat{\mu}(\hat{\rho}))$	3.07 (.005)	4.86 (.009)	6.79 (.013)	8.48 (.018)	9.75 (.022)
δ-method	3.16 (.005)	4.96 (.008)	6.83 (.013)	8.49 (.018)	9.75 (.022)
$\widehat{\text{var}}_{SJ}(\hat{\mu}(\hat{\rho}))$	3.72 (.009)	5.66 (.010)	7.11 (.011)	8.53 (.013)	9.70 (.014)
$\widehat{\text{var}}_{IJ1}(\hat{\mu}(\hat{\rho}))$	3.40 (.008)	5.32 (.009)	7.03 (.011)	8.53 (.013)	9.70 (.014)
$\widehat{\text{var}}_{IJ2}(\hat{\mu}(\hat{\rho}))$	3.05 (.007)	4.91 (.008)	6.69 (.010)	8.37 (.013)	9.67 (.014)
Contaminated					
$\text{var}(\hat{\mu}(\hat{\rho}))$	7.19	28.5	55.1	81.6	102.8
$\widehat{\text{var}}(\hat{\mu}(\hat{\rho}))$	5.87 (.014)	17.0 (.056)	30.7 (.110)	44.2 (.164)	55.0 (.207)
δ-method	5.96 (.014)	17.1 (.056)	30.7 (.110)	44.2 (.164)	55.0 (.207)
$\widehat{\text{var}}_{SJ}(\hat{\mu}(\hat{\rho}))$	7.96 (.027)	29.1 (.115)	55.4 (.225)	81.8 (.335)	103.0 (.424)
$\widehat{\text{var}}_{IJ1}(\mu(\rho))$	8.40 (.026)	29.8 (.114)	56.1 (.225)	82.2 (.336)	103.0 (.425)
$\widehat{\text{var}}_{IJ2}(\hat{\mu}(\hat{\rho}))$	7.45 (.024)	28.7 (.113)	55.2 (.224)	81.8 (.335)	103.0 (.424)

The infinitesimal jackknife estimator can be extended for variance estimators of the fixed effects in the general mixed linear model, and would be expected to have as good a performance as in the one-way random-effects model.

Acknowledgements

Thanks to Dr. Nelder, Imperial College and referees for their careful readings and helpful comments.

Appendix

Derivation of the delta method. From Searle ((1971) p. 474), we have

$$\text{(A.1)}\quad \mathrm{var}(s_e^2) = a\sigma_e^4, \quad \mathrm{cov}(s_e^2, s_a^2) = b\sigma_e^4 \quad \text{and} \quad \mathrm{var}(s_a^2) = c\sigma_e^4 + d\sigma_e^2\sigma_a^2 + e\sigma_a^4,$$

where $a = 2/(n-k)$, $b = -2n(k-1) \,/\, \{(n-k)(n^2 - \sum n_i^2)\}$, $c = 2n^2(n-1)(k-1) \,/\, \{(n-k)(n^2 - \sum n_i^2)^2\}$, $d = 4n/(n^2 - \sum n_i^2)$ and $e = 2\{n^2 \sum n_i^2 + (\sum n_i^2)^2 - 2n \sum n_i^3\} \,/\, (n^2 - \sum n_i^2)^2$. For any parameter θ, let $U(f(\theta))$ be an unbiased estimator for $f(\theta)$. Since $E\{U(f(\theta))^2\} = f(\theta)^2 + \mathrm{var}(U(f(\theta)))$, we can have unbiased estimators

$$U(\sigma_e^4) = s_e^4/(1+a), \quad U(\sigma_e^2\sigma_a^2) = s_e^2 s_a^2 - bU(\sigma_e^4) \quad \text{and}$$
$$U(\sigma_a^4) = \{s_a^4 - cU(\sigma_e^4) - dU(\sigma_e^2\sigma_a^2)\}/(1+e).$$

Note that $U(\sigma_a^4)$ can take negative values. In this case, even though the negative estimates must be retained for unbiasedness, negative estimates are, in practice, often set to zero. By replacing these estimators in equations (A.1), we can obtain the estimators for (A.1).

Derivation for $I(s_a^2, X_i)$. Let $\bar{y}^i = \{(1-\epsilon)\sum n_j \bar{y}_j + \epsilon k n_i \bar{y}_i\} \,/\, \{(1-\epsilon)n + \epsilon k n_i\} = \bar{y} + c_{\epsilon,i}(\bar{y}_i - \bar{y})$, where $c_{\epsilon,i} = \epsilon k n_i/\{(1-\epsilon)n + \epsilon k n_i\}$. Then, $\mathrm{SSA}(Q_i) = (1-\epsilon)\sum n_j(\bar{y}_j - \bar{y}^i)^2 + \epsilon k n_i(\bar{y}_i - \bar{y}^i)^2 = (1-\epsilon)\sum n_j(\bar{y}_j - \bar{y})^2 + d_{\epsilon,i}(\bar{y}_i - \bar{y})^2$, where $d_{\epsilon,i} = n(1-\epsilon)c_{\epsilon,i}^2 + k\epsilon n_i(1-c_{\epsilon,i})^2$. Note that $E\{(\bar{y}_j - \bar{y})^2\} = (1 - 2n_i/n + \sum n_j^2/n^2)\sigma_a^2 + (1/n_i - 1/n)\sigma_e^2$. Hence, $E\{\mathrm{SSA}(Q_i)\} = a_{\epsilon,i}\sigma_e^2 + b_{\epsilon,i}\sigma_a^2$, where $a_{\epsilon,i} = (1-\epsilon)(k-1) + d_{\epsilon,i}(1/n_i - 1/n)$ and $b_{\epsilon,i} = (1-\epsilon)(n - \sum n_j^2/n) + d_{\epsilon,i}(1 - 2n_i/n + \sum n_j^2/n^2)$. Therefore, $s_a^2(Q_i) = \{\mathrm{SSA}(Q_i) - a_{\epsilon,i}s_e^2(Q_i)\}/b_{\epsilon,i}$. Let $a = \lim_{\epsilon\to 0} a_{\epsilon,i} = k-1$ and $b = \lim_{\epsilon\to 0} b_{\epsilon,i} = n - \sum n_j^2/n$. Then,

$$s_a^2(Q_i) - s_a^2(P_0) = A1 + A2 + A3 + A4,$$

where $A1 = \{\mathrm{SSA}(Q_i) - a_{\epsilon,i}s_e^2(Q_i)\}/b_{\epsilon,i} - \{\mathrm{SSA}(Q_i) - a_{\epsilon,i}s_e^2(P_0)\}/b_{\epsilon,i}$, $A2 = \{\mathrm{SSA}(Q_i) - a_{\epsilon,i}s_e^2(P_0)\}/b_{\epsilon,i} - \{\mathrm{SSA}(Q_i) - as_e^2(P_0)\}/b_{\epsilon,i}$, $A3 = \{\mathrm{SSA}(Q_i) - as_e^2(P_0)\}/b_{\epsilon,i} - \{\mathrm{SSA}(Q_i) - as_e^2(P_0)\}/b$ and $A4 = \{\mathrm{SSA}(Q_i) - as_e^2(P_0)\}/b - \{\mathrm{SSA}(P_0) - as_e^2(P_0)\}/b$. Since $\lim_{\epsilon\to 0}(a_{\epsilon,i} - a)/\epsilon = (n - n_i k)/n$ and $\lim_{\epsilon\to 0}(b_{\epsilon,i} - b)/\epsilon = kn_i(1 - 2n_i/n + \sum n_j^2/n^2) - b$, it can be shown that $\lim_{\epsilon\to 0} A1/\epsilon = -aI(s_e^2, X_i)/b$, $\lim_{\epsilon\to 0} A2/\epsilon = -(n - n_i k)s_e^2/(nb)$, $\lim_{\epsilon\to 0} A3/\epsilon = -\{kn_i(1 - 2n_i/n + \sum n_j^2/n^2) - b\}s_a^2/b$ and $\lim_{\epsilon\to 0} A4/\epsilon = \{kn_i(\bar{y}_i - \bar{y})^2 - bs_a^2 - as_e^2\}/b$. By rearranging these terms appropriately, we have the desired results.

REFERENCES

Arvensen, J. N. and Layard, M. W. J. (1975). Asymptotically robust tests in unbalanced variance components models, *Ann. Statist.*, **3**, 1122–1134.

Arvensen, J. N. and Schmitz, T. (1970). Robust procedures for variance components problems using the jackknife, *Biometrics*, **26**, 677–686.

Birkes, D., Seely, J. and Azzam, A. (1981). An efficient estimator of the mean in a two-stage nested model, *Technometrics*, **23**, 143–148.

Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*, SIAM, Philadelphia.

Hinkley, D. (1977). Jackknifing in unbalanced situations, *Technometrics*, **19**, 285–292.

Kackar, R. N. and Harville, D. A. (1984). Approximations for standard errors of estimations of fixed and random effects in mixed linear models, *J. Amer. Statist. Assoc.*, **79**, 853–862.

Searle, S. R. (1971). *Linear Models*, Wiley, New York.

Simonoff, J. S. and Tsai, C. (1986). Jackknife-based estimators and confidence regions in nonlinear regression, *Technometrics*, **28**, 103–122.

Swallow, W. H. (1981). Variances of locally minimum variance quadratic unbiased estimators ("MIVQUE's") of variance components, *Technometrics*, **23**, 271–283.

Swallow, W. H. and Monahan, J. F. (1984). Monte Carlo comparison of ANOVA, MIVQUE, REML, and ML estimators of variance components, *Technometrics*, **26**, 47–57.

Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis, *Ann. Statist.*, **14**, 1261–1295.