Jour. Ind. Soc. Ag. Statistics
Vol. XXXIX, No. 1 (1987), pp. 1-8

CONFIDENCE INTERVALS ON FUNCTIONS OF COMPONENTS OF VARIANCE IN SAMPLES FROM MODERATE NON-NORMAL POPULATIONS

R. A. SINGHAL
Indian Veterinary Research Institute, Izatnagar

(Received : December, 1981)

SUMMARY

Confidence intervals for some functions of variance components associated with non-normal balanced one-way Model II are obtained using two moment approximation.

Keywords : Confidence intervals; Non-normal populations; Random model; Variance components.

Introduction

Approximate confidence intervals on variance components, with normality, have been obtained by several workers. Anderson and Bancroft [1], Scheffee [8], Graybill [7], Bogyo and Becker [4] and Searle [9], among others, however, derive exact confidence intervals for functions of variance components in a balanced random effects model relying on a property of classical F-distribution. Atiqullah [2] derives asymptotic confidence intervals for a ratio of variance components in samples from symmetrical populations and infers that the intervals are insensitive to kurtosis of random effects. Snedecor and Cochran [11] surmise that for positive kurtosis of random effects the confidence intervals for inter-group variance component are too narrow. The inferences of Atiqullah and of Snedecor and Cochran are not commensurate with the power of variance ratio test in random models (Tan and Wong [12], and Singhal and Singh [10]).

This paper considers approximate confidence intervals for certain functions of variance components σ_a^2 and σ_e^2 associated, respectively, with the non-normal group-effects (a_i) and error-effects (e_{ij}) in the model

$$y_{ij} = m + a_i + e_{ij}, \tag{1.1}$$

$(i = 1, 2, \ldots, g \quad \text{and } j = 1, 2, \ldots, r)$
where m is an additive constant and a_i and e_{ij} are independent random samples from infinite moderate non-normal populations. To arrive at the confidence intervals, the non-normal sampling distribution of the ratio of mean squares associated with (1.1) has been approximated to the F-distribution using first two moments. Computed results in Section 4 characterise the behaviour of the violation of normality assumption on the confidence intervals.

2. Moment Approximation to the Sampling Distribution of Variance Ratio

$$\text{For } F' = s_g^2/s_e^2, \tag{2.1}$$

where s_g^2 and s_e^2 are 'between-groups' and 'within-groups' mean squares for (1.1) with $p = (g - 1)$ and $q = (N - g)$ degrees of freedom, respectively, the mean and variance are directly derivable from (2.1) of Singhal and Singh [10], after heavy and lengthy simplifications, as :

$$\mu_1' = \frac{q(1 + B)}{(q - 2)} + \frac{q(qB - 2)}{N(q^2 - 4)} \lambda_{4e} + \frac{4q((g + 4) - B(q - g)) \lambda_{3e}^2}{N(q + 4)(q^2 - 4)} \tag{2.2}$$

$$\begin{aligned}\mu_2 = {} & \frac{2(p + q - 2) q^2 (1 + B)^2}{p(q - 4)(q - 2)^2} + \frac{q^2 B^2 \lambda_{4a}}{g(q - 2)(q - 4)} \\ & + \frac{q^2(p(q^2 - 4) - 2(pq + 4q - 8)(q + 2)(1 + B) + q(pq + 6q + 2p - 12)(1 + B)^2) \lambda_{4e}}{pN(q + 2)(q - 4)(q - 2)^2} \\ & + \frac{8q^2(2(pq + 3q - p - 6)(q + 4)(1 + B) - (q - p - 1)(qp + 4q - 8)(1 + B)^2) \lambda_{3e}^2}{pN(q + 2)(q - 2)^2 (q^2 - 16)}\end{aligned} \tag{2.3}$$

where $B = r\sigma_a^2/\sigma_e^2$, $N = gr$, and λ_3 and λ_4 are standardized third and fourth cumulants associated with a_i and e_{ij}.

The expressions (2.2) and (2.3) agree exactly with the corresponding expressions of Gayen [5]* for the balanced one-way fixed effects model when $B = 0$ (i.e. $\sigma_a^2 = 0$).

*The factor 2 $(2p + 7)$ appearing in his variance expression for the coefficient of λ_{3e}^2 should be 8 $(2p + 7)$.

The sampling distribution of (2.1) for non-normal random effects is approximated by considering the ratio F'/k to follow the central F-distribution with s and q degrees of freedom. On equating μ_1'/k and μ_2/k^2 to the mean and variance of the Snedecor's F-distribution we get, after simplification,

$$k = (q - 2)\, \mu_1'/q$$

and

$$s = \frac{2(q-2)\, \mu_1'^2}{(q-4)\, \mu_2 - 2\, \mu_1'^2} . \tag{2.4}$$

The approximation (see Tiku, [13] gives power of the test (see Singhal and Singh, [10]) approximately equal to

$$\int_0^{F'/k} f(s, q; F)\, dF = I_x\left(\tfrac{1}{2}s, \tfrac{1}{2}q\right), \tag{2.5}$$

where $x = q(1 + B)/(q(1 + B) + sF\alpha/k)$, k and s are provided by (2.4), $I(a, b)$ is the Incomplete B-function and $F\alpha$ is the $\alpha\%$ critical point of the traditional F-distribution with s and q degrees of freedom. To obtain critical points of the F-distribution either interpolation or next higher integral value for fractional values of s may be considered. The agreement between the approximation (2.5) and the exact power of the test is found close even for small samples e.g. when $g = 3$, $r = 5$, $\sigma_a^2/\sigma_e^2 = 0.25$, $\alpha = 0.05$, $\lambda_{4a} = 1.0$ and for $s = 1.8134$, as the integral on evaluation yielded a value 0.2125 whereas the exact power of this test was 0.2159.

The next section provides confidence limits for different functions considered by various workers.

3. Confidence Limits on Functions of Variance Components

The Table 1 provides $(1 - \alpha)\%$ confidence limits for σ_a^2 and its functions with σ_e^2 for normal behaviour of random effects.

Here α is the level of significance, F_1 and F_2, respectively, are lower and upper $(1 - \frac{1}{2}\alpha)\%$ and $(\frac{1}{2}\alpha)\%$ critical points of the F-distribution with p and q degrees of freedom.

In order to investigate the effects of deviation from normality assumption on the confidence limits for the parameters in Table 1 it would be necessary to interpolate F_1 and F_2 for s and q degrees of freedom, since s will, generally, be fractional.

The magnitude and direction of change on the confidence intervals due to non-normality in random effects is illustrated through numerical results in what follows,

TABLE 1—(1 − α)% CONFIDENCE LIMITS

Parameter	*Lower limit*	*Upper limit*
σ_a^2	$(s_g^2 - s_e^2 F_2)/rF_2$	$(s_g^2 - s_e^2 F_1)/rF_1$
	(Anderson and Bancroft, [1], p. 322)	
$w(= \sigma_a^2/\sigma_e^2)$	$(s_g^2 - s_e^2 F_2)/rs_e^2 F_2$	$(s_g^2 - s_e^2 F_1)/rs_e^2 F_1$
	(Scheffee, [8], p. 229)	
$R(= \sigma_e^2 / (\sigma_a^2 + \sigma_e^2)$	$\dfrac{r\sigma_e^2 F_1}{s_g^2 + (r-1) s_e^2 F_1}$	$\dfrac{r s_e^2 F_2}{s_g^2 + (r-1) s_e^2 F_2}$
	(Graybill, [7], p. 379)	
$t(= \sigma_a^2 / (\sigma_a^2 + \sigma_e^2)$	$\dfrac{s_g^2 - s_e^2 F_2}{s_g^2 + (r-1) s_e^2 F_2}$	$\dfrac{s_g^2 - s_e^2 F_1}{s_g^2 + (r-1) s_e^2 F_1}$
	(Searle, (9)	

4. Computations of Results and Discussion

Considering certain *a priori* values of parameters computations of (2.2) and (2.3) and whence of (2.4) were carried out. To obtain 90% confidence limits for the parameters in Table 1, the lower and upper critical points of F-distribution for fractional s in combination with q were interpolated using four point Lagrangian interpolation formula. Incorporating these critical values confidence intervals for the parameters σ_a^2, w and R (the limits for t may be obtained using the relation $t = 1 - R$) were computed. The values of s_g^2 and s_e^2 were obtained from the corresponding values of $\hat{\sigma}_a^2$ and $\hat{\sigma}_e^2$ where $\hat{\sigma}_a^2 = (s_g^2 - s_e^2)/r$ and $\hat{\sigma}_e^2 = s_e^2$. To save space detailed results have not been given here but a few, for (i) $g = 5$, $r = 5$, (ii) $g = 9$, $r = 3$ and (iii) $g = 7$, $r = 5$ and for varying s_g^2 and s_e^2, are summarised in Table 2. The values of non-normality parameters were considered within Barton and Dennis [3] limits so that the Edgeworth series representation of the density function of random effects is unimodal and positive definite.

Based on the computed results, the pattern of change in confidence limits due to the non-normality parameters may be put to record as under.

(*a*) The confidence limits are insensitive to the skewness in error effects since there is no appreciable change in s when compared to p due to λ_{3e} (Table 2). The expressions (2.2) and (2.3) are independent of λ_{3a}.

(*b*) For $\lambda_{4a} > 0$, the confidence intervals are remarkably wider and for $\lambda_{4a} < 0$, these are appreciably narrow in comparison to when $\lambda_{4a} = 0$. For example, when $g = 9$, $r = 3$, $s_g^2 = 13.5$ and $s_e^2 = 1.5$, the 90% limits for σ_a^2, w and R are (0.765, 35.529), (0.510, 23.686) and (0.040, 0.662) for $\lambda_{4a} = 2.4$. Correspondingly for $\lambda_{4a} = -1.0$ these are (1.146, 12.981), (0.764, 8.654) and (0.104, 0.567) in comparison to the normal limits (0.997, 17.630), (0.665, 11.754) and (0.078, 0.601). For non-normal error effects, though the direction of the effect was, in general, similar to that of the group effects but magnitude of displacement in the limits was of lesser degree. Further it was interesting to note that for small values of s_g^2 and s_e^2, and for certain combinations of r and g the direction of change was in reverse order for kurtotic error effects e.g. when $g = 5$, $r = 5$, $s_g^2 = 3.5$ and $s_e^2 = 2.0$. The intervals for positive parameter are negative. For a discussion on this aspect the reader may refer to Searle [9].

Results show disagreement with Snedecor and Cochran [11], for the limits on σ_a^2 where they conjecture that for positive kurtosis of random effects the limits would be 'too narrow' and with Atiqullah [2] with regard to the sensitiveness of the limits on σ_a^2/σ_e to the kurtosis. The observations of these workers are at variance with the power of the test (see Tan and Wong, [12] and Singhal and Singh, [10]).

Snedecor and Cochran's analytical inference on confidence intervals is correct only in respect of the direction of variance of $\hat{\sigma}_a^2$ for positive kurtosis but is erroneous when it comes to the size of the confidence intervals. Atiqullah arrived at confidence limits of the σ_a^2/σ_e^2 through asymptotic results and used smaller numerical values of the estimators for computations which are not, probably, sufficiently large to reveal the magnitude of the effect of kurtosis on the confidence limits.

Further to provide computational ease for applications, the expressions (2.2) and (2 3) may be approximated to $0(N^{-2})$ as:

$$\mu_1' \doteq (1 + B)\left(1 + \frac{2}{N} + \frac{2(g+2)}{N^2}\right) + \left(B + \frac{2}{N}\right)\frac{\lambda_{4e}}{N} - 4B\,\frac{\lambda_{3e}^2}{N^2} \tag{2.2 bis}$$

TABLE 2—THE 90% CONFIDENCE LIMITS FOR FUNCTIONS OF VARIANCE COMPONENTS FOR NORMAL AND NON-NORMAL UNIVERSES

s_g^2	s_e^2	*Parameter*	*Normal*	λ_{4e}		λ_{4a}		λ_{3e}^2*
				−1.0	*2.4*	*−1.0*	*2.4*	*0.25*
				Case (i) g = 5 and r = 5				
3.5	2.0	σ_a^2	(−0.201, 5.592)	(−0.201, 5.608)	(−0.200, 5.542)	(−0.196, 4.960)	(−0.211,7. 412)	3.978
		W	(−0.100, 2.796)	(−0.100, 2.804)	(−0.100, 2.771)	(−0.098, 2.480)	(−0.105, 3.706)	—
		R	(0.263, 1.112)	(0.263, 1.112)	(0.265, 1.111)	(0.287, 1.109)	(0.212, 1.118)	—
15.0	2.5	σ_a^2	(0.354, 25.180)	(0.375, 22.348)	(0.318, 31.347)	(0.440, 16.373)	(0.210, 70.088)	4.013
		W	(0.741, 10.072)	(0.150, 0.894)	(0.127, 12.539)	(0.176, 6.549)	(0.088, 28.035)	—
		R	(0.090, 0.876)	(0.101, 0.870)	(0.074, 0.887)	(0.132, 0.850)	(0.034, 0.919)	—
21.5	1.5	σ_a^2	(0.923, 36.509)	(0.966, 31.041)	(0.857, 48.899)	(1.091, 20.945)	(0.639, 138.410)	4.034
		W	(0.616, 24.339)	(0.644, 20.694)	(0.571, 32.599)	(0.727, 13.963)	(0.459, 92.273)	—
		R	(0.039, 0.619)	(0.046, 0.608)	(0.030,0 .636)	(0.067, 0.579)	(0.011, 0.685)	—
				Case (ii) g = 9 and r = .3				
2.9	2.0	σ_a^2	(−0.345, 3.228)	(−0.346, 3.284)	(−0.342, 3.102)	(−0.342, 3.102)	(−0.352, 3.547)	7.859
		W	(−0.172, 1.614)	(−0.173, 1.642)	(−0.171, 1.551)	(−0.171, 1.551)	(−0.176, 1.774)	—
		R	(0.383, 1.208)	(0.378, 1.209)	(0.392, 1.206)	(0.392, 1.206)	(0.360, 1.214)	—

10.0	2.5	σ_a^2	(0.276, 12.597)	(0.299, 11.712)	(0.235, 14.514)	(0.345, 10.222)	(0.152, 33.045)	7.986
		W	(0.110, 5.039)	(0.119, 4.685)	(0.094, 5.805)	(0.138, 4.089)	(0.061, 13.218)	—
		R	(0.166, 0.901)	(0.176, 0.893)	(0.147, 0.914)	(0.196, 0.879)	(0.070, 0.943)	—
13.5	1.5	σ_a^2	(0.997, 17.630)	(1.050, 15.652)	(0.911, 22.000)	(1.146, 12.981)	(0.765, 35.529)	8.038
		W	(0.665, 11.754)	(0.700, 10.435)	(0.607, 14.667)	(0.764, 8.654)	(0.510,23. 686)	—
		R	(0.078, 0.601)	(0.087, 0.588)	(0.064, 0.622)	(0.104, 0.567)	(0.040, 0.662)	—
					Case (iii) $g = 7$ and $r = 5$			
3.5	2.0	σ_a^2	(−0.159, 3.151)	(−0.158, 3.135)	(−0.159, 3.182)	(−0.153, 2.873)	(−0.171 ,3.913)	5.978
		W	(−0.079, 1.576)	(−0,079, 1.568)	(−0.080, 1.591)	(−0.077, 1.436)	(−0.085, 1.956)	—
		R	(0.388, 1.086)	(0.389, 1.086)	(0.386, 1.087)	(0.410, 1.083)	(0.338, 1.093)	—
15.0	2.5	σ_a^2	(0.534, 14.719)	(0.556, 13.578)	(0.492, 17.272)	(0.635, 10.632)	(0.373, 31.013)	6.014
		W	(0.213, 5.888)	(0.222, 5.431)	(0.197, 6.909)	(0.254, 4.253)	(0.149, 12.405)	—
		R	(0.145, 0.824)	(0.155, 0.818)	(0.126, 0.836)	(0.190, 0.798)	(0.075. 0.870)	—
21.5	1.5	σ_a^2	(1.181, 21.514)	(1.226, 19.317)	(1.104, 26.525)	(1.381, 14.163)	(0.900, 56.130)	6.026
		W	(0.788, 14.343)	(0.818, 12.878)	(0.736, 17.683)	(0.921, 9.442)	(0.600, 37.420)	—
		R	(0.065, 0.559)	(0.072, 0.550)	(0.054, 0.576)	(0.096, 0.521)	(0.026, 0.625)	—

*Value of s for $\lambda_{se} = 0.5$.

and

$$\mu_2 \doteq \frac{2(1+B)^2}{p}\left(1+\frac{(g+5)}{N}+\frac{(g^2+13g+20)}{N^2}\right)+\left(1+\frac{6}{N}\right.$$
$$\left.+\frac{2(3g+14)}{N^2}\right)\frac{B^2\lambda_{4a}}{g}+\left(p\left(1+\frac{6}{N}\right)-2(1+B)\left((g+3)\right.\right.$$
$$\left.\left.+\frac{8(g+2)}{N}\right)+(1+B)^2\left((g+5)+\frac{8(g+2)}{N}\right)\right)\frac{\lambda_{4e}}{pN}$$
$$+8(1+B)\left(2(g+2)-(g+3)(1+B)\right)\frac{\lambda_{3e}^2}{pN^2} \qquad (2.3\ bis)$$

Besides, the estimates of λ_3 and λ_4, as suggested by Geary [6], in terms of Fisher's k-statistics can be used for the calculation of confidence limits on a parameter.

REFERENCES

[1] Anderson, R. L. and Bancroft, T. A. (1952) : *Statistical Theory in Research.* McGarw Hill, New York.

[2] Atiqullah, M. (1962) : On the effect of non-normality on the estimation of components of variance, *J. Roy. Statist. Soc.*, B, **24** : 140-47.

[3] Barton, D. E. and Dennis, K. E. (1952) : The conditions under which Gram Charlier and Edgeworth curves are positive definite and unimodel, *Biometrika*, 39 : 425-7.

[4) Bogyo, T. P. and Becker, W. A. (1963) : Exact confidence intervals for genetic heritability estimated from parental halfsib-correlations, *Biometrics*, **19** : 494-96.

[5] Gayen, A. K. (1950) : The distribution of the variance ratio in random samples of any size drawn from non-normal universe, *Biometrika*, **37** : 236-55.

[6] Geary, R. C. (1947) : Testing for normality, *Biometrika*, **34** : 209-41.

[7] Graybill, F. A. (1961) : *An Introduction to Linear Statistical Models*, Vol. I. McGraw Hill, New York.

[8] Scheffee, H. (1959) : *The Analysis of Variance*. Wiley. New York.

[9] Searle, S. R. (1971) : Topics in variance component estimation, *Biometrics*, **27** : 1-76.

[10] inghal, R. A. and Singh, C. (1984) : Distribution of variance ratio test in a non-normal random effects model, *Sankhya*, Sr. B., **46** : 29-35.

[11] Snedecor, G. W. and Cochran, W. G. (1967) : *Statistical Methods*, 6th edn. IBH, New Delhi.

[12] Tan, W. Y. and Wong, S. P. (1980) : On approximating the null and non-null distributions of the *F*-ratio in unbalanced random effect models from non-normal universes, *J. Amer. Statist. Assoc.*, **75** : 655-62.

[13] Tiku, M. L. (1971) : Power function of the *F*-test under non-normal situations, *J. Amer. Statist. Assoc.*, **66** : 913-16.