

CREDIBLE AND CONFIDENCE SETS FOR THE
RATIO OF VARIANCE COMPONENTS IN
THE BALANCED ONE-WAY MODEL

By TONGLIN ZHANG
Purdue University, USA
and
MICHAEL WOODROOFE
University of Michigan, USA

SUMMARY. Classical confidence intervals for the ratio of variance components can degenerate to a single point with positive probability, leaving no allowance for error. An alternative method is considered which avoids this problem, but maintains the frequentist coverage probability. The alternative intervals are compared to Bayesian credible intervals, derived from a non-informative prior; and the Bayesian intervals are shown to have quite high frequentist coverage probabilities.

1. Introduction

In the one-way random effects model in the analysis of variance, there are observations of the form

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad (1)$$

for $j = 1, \dots, J_i$ and $i = 1, \dots, I$, where α_i and ϵ_{ij} are independent random variables for which $\alpha_i \sim N(0, \sigma_\alpha^2)$ and $\epsilon_{ij} \sim N(0, \sigma^2)$. The parameters $-\infty < \mu < \infty$, $\sigma_\alpha^2 \geq 0$ and $\sigma^2 > 0$ are unknown here, and confidence intervals for the two variances, σ^2 and σ_α^2 , and their ratio, $\lambda = \sigma_\alpha^2/\sigma^2$, are of interest. In this paper, interest centers on confidence and credible intervals for λ , in the balanced case, when $J_1 = \dots = J_I = J$. Then,

Paper received January 2002.

AMS (1991) subject classification. 62F15, 62J10, 62F25.

Keywords and phrases. Between variance, frequentist coverage probability, prior and posterior distributions, total positivity, unimodality, within variance.

complete sufficient statistics are

$$\begin{aligned}\bar{y}_{..} &= \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J y_{ij} \sim N\left(\mu, \frac{\sigma^2 + J\sigma_\alpha^2}{IJ}\right), \\ SSW &= \sum_{i=1}^I \sum_{j=1}^J (y_{ij} - \bar{y}_{i.})^2 \sim \sigma^2 \chi_{I(J-1)}^2, \\ SSB &= \sum_{i=1}^I J(\bar{y}_{i.} - \bar{y}_{..})^2 \sim (\sigma^2 + J\sigma_\alpha^2) \chi_{I-1}^2\end{aligned}$$

where

$$\bar{y}_{i.} = \frac{1}{J} \sum_{j=1}^J y_{ij}$$

for $i = 1, \dots, I$; and $\bar{y}_{..}$, SSB and SSW are independent for fixed μ , σ^2 and σ_α^2 . Suppose that $I, J \geq 2$ throughout the paper, and let $MSB = SSB/(I-1)$ and $MSW = SSW/[I(J-1)]$. Then

$$X = \frac{MSB}{MSW} \sim (1 + J\lambda)\mathcal{F}, \quad (2)$$

where \mathcal{F} has the F-distribution with $I-1$ degrees of freedom in the numerator and $I(J-1)$ in the denominator, and classical confidence intervals treat $X/(1 + J\lambda)$ as a pivotal quantity. Specifically, letting $0 < a < b < \infty$, intervals of the form

$$[L, U] = \left[\frac{X-b}{Jb}, \frac{X-a}{Ja} \right] \quad (3)$$

have been recommended by Scheffe (1959), Burdick and Graybill (1992), Searle et al. (1992), and *SAS* among others. The coverage probability of the interval is $1 - \alpha$, where $\alpha = F(a) + [1 - F(b)]$ and F is the distribution function of \mathcal{F} ; and conventional choices of a and b are $a = F^{-1}(\alpha/2)$ and $b = F^{-1}(1 - \alpha/2)$ for a given α . An obvious problem with the interval (3) is that the endpoints can be negative with positive probability, and the authors just cited differ on how to handle this problem. Burdick & Graybill (1992) and *SAS* suggest replacing negative values by 0, in which case the interval may degenerate to $[0, 0]$ with positive probability; Scheffe (1959, p.230), however, does not recommend replacing negative values by zero, since “the length of a two-side confidence interval is a measure of the error of some point estimate of \dots .” A less obvious problem is *flip-flopping* in which experimenters report a confidence interval when $L > 0$, but an upper confidence bound is reported

when $L < 0$ (with a replaced by the α^{th} quantile). Such flip-flopping, of course, destroys the exact nature of the confidence intervals.

Similar problems arise when setting confidence intervals for a non-negative normal mean or the signal in a Poisson model with a nuisance background. To address problems of this nature, Feldman and Cousins (1998) proposed confidence regions of high likelihood in the allowable parameter space. This method is called the *unified method* in the physics literature precisely because it makes a natural transition from a confidence bound to a confidence interval. The unified approach is applied to the variance component λ in Section 2 and shown to lead to exact confidence intervals with non-empty interiors. Bayesian methods can also be used to avoid the problem of negative endpoints. Intervals derived by Hill (1965) are reconsidered in Section 3 and compared to the unified and classical intervals. In addition, the frequentist coverage probability of the Bayesian intervals is shown to be quite high.

Some Notation. The model (1) is invariant under changes of location and scale; X is a maximal invariant in the space of the sufficient statistics; and λ is a maximal invariant in the parameter space. In this paper, inferences about λ are based on X . To simplify some of the formulas, it is convenient to work with the parameter

$$\theta = 1 + J\lambda = 1 + J\frac{\sigma_a^2}{\sigma^2};$$

and since this transformation is linear, there is no problem converting confidence and credible intervals for θ into corresponding intervals for λ . Observe that $1 \leq \theta < \infty$; and let f_θ and F_θ denote the density and distribution function of X for a given θ . Then

$$\begin{aligned} f_\theta(x) &= \frac{1}{\theta} f\left(\frac{x}{\theta}\right), \\ F_\theta(x) &= F\left(\frac{x}{\theta}\right), \end{aligned}$$

where f and F are the density and distribution function of the F distribution with $I - 1$ degrees of freedom for the numerator and $I(J - 1)$ for the denominator. Thus, the likelihood function given x is

$$L(x; \theta) = f_\theta(x) = \frac{\Gamma(r + s)r^r s^s}{\Gamma(r)\Gamma(s)} \frac{\theta^s x^{r-1}}{[s\theta + rx]^{r+s}}, \tag{4}$$

where $r = (I - 1)/2$ and $s = I(J - 1)/2$. Differentiating the logarithm of (4) shows that $L(x; \theta)$ is maximized for fixed x when

$$\theta = \hat{\theta} = \max\{1, x\},$$

that $L(x; \theta)$ decreasing in $\theta \geq \hat{\theta}$, and that $L(x; \theta)$ is increasing in $1 \leq \theta \leq \hat{\theta}$ when $x > 1$.

2. Confidence Interval from Unified Approach

The unified approach was employed by Feldman and Cousins (1998) in the normal and Poisson cases with restrictions on the parameters; but the approach can be applied to any one-parameter family of distributions. Suppose that $X \sim f_\theta$ has a density depending on a real parameter θ taking values in an interval Ω . Write $L(x; \theta) = f_\theta(x)$ for the likelihood; let $\hat{\theta} = \hat{\theta}(x)$ be the maximum likelihood estimator (MLE) of θ ; and let $R_\theta(x) = L(x; \theta)/L(x; \hat{\theta})$ be the likelihood ratio statistic. Then, a level $1 - \alpha$ confidence set is $\{\theta : R_\theta(x) \geq c_\theta\}$ where c_θ are determined by

$$P_\theta\{x : R_\theta(x) \geq c_\theta\} = 1 - \alpha. \quad (5)$$

It is assumed here that the MLE exists for all x and that a solution c_θ exists in (5) for all θ . Thus, the unified approach consists of taking regions of high likelihood. This approach is common in statistics, but often combined with a chi-square approximation.

In the variance components problem, the likelihood ratio is

$$R_\theta(x) = \frac{L(x; \theta)}{L(x; \hat{\theta})} = \begin{cases} \theta^s x^r (r+s)^{r+s} / (s\theta + rx)^{r+s}, & \text{when } x \geq 1 \\ \theta^s (s+rx)^{r+s} / (s\theta + rx)^{r+s}, & \text{when } x \leq 1. \end{cases} \quad (6)$$

Observe that

$$\frac{d}{dx} \log[R_\theta(x)] = \begin{cases} r(r+s)[(s+rx)^{-1} - (s\theta+rx)^{-1}], & \text{when } x < 1 \\ r[x^{-1} - (r+s)(s\theta+rx)^{-1}], & \text{when } x > 1. \end{cases}$$

If $\theta > 1$, then $R_\theta(x)$ is increasing for $0 \leq x < \theta$ and decreasing for $x > \theta$. If $\theta = 1$, then $R_1(x) = 1$ for $0 \leq x \leq 1$ and $R_1(x)$ is decreasing if $x > 1$. Figure 1 displays the graph of $R_\theta(x)$ for selected θ , when $I = 6$ and $J = 5$.

Since $R_\theta(x)$ is unimodal in x for each θ , $\{x : R_\theta(x) \geq c\}$ is an interval for any $\theta \geq 1$ and $0 < c < 1$. So, the unified confidence interval can be obtained by solving

$$F_\theta(b) - F_\theta(a) = 1 - \alpha$$

and

$$[a, b] = \{x : R_\theta(x) \geq c\},$$

for a , b and $0 \leq c \leq 1$ for each θ . If $\theta = 1$, then a solution is $a = 0$, $b = F^{-1}(1 - \alpha)$, and $c = R_1(b)$. When $\alpha < 1 - F(1)$, the following algorithm may be used to find a and b when $\theta > 1$.

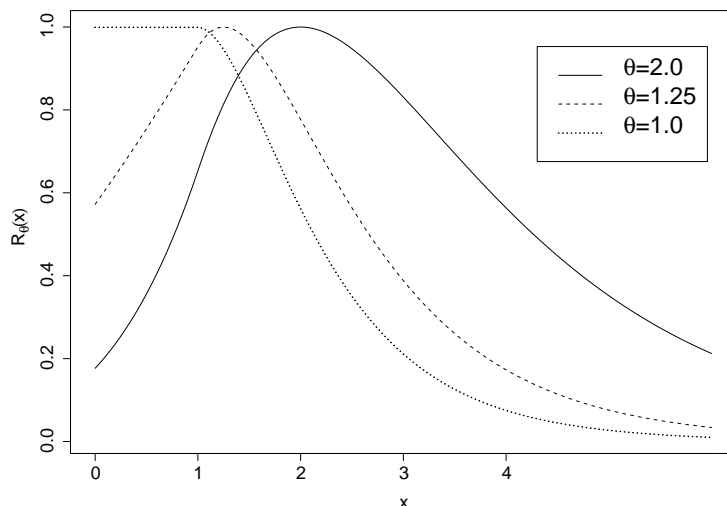


Figure 1: Plots of $R_\theta(x)$ when $I = 6$ and $J = 5$.

Unified Algorithm.

- i) Let $z = \theta F^{-1}(1 - \alpha)$, so that $P_\theta[X \leq z] = 1 - \alpha$.
- ii) If $R_\theta(z) \leq R_\theta(0)$, let $a = 0$ and $b = z$.
- ii') Otherwise, let $c_0 = R_\theta(0)$, $c_1 = R_\theta(z)$, and iterate step iii) to convergence.
- iii) Let $c = (c_0 + c_1)/2$; solve the equations $R_\theta(a) = R_\theta(b) = c$ for $0 < a < \theta < b$; if $P_\theta[a \leq X \leq b] < 1 - \alpha$, let $c_1 = c$; otherwise let $c_0 = c$.

The condition $\alpha < 1 - F(1)$ insures that $z > \theta$ in i). In iii), solving $R_\theta(a) = c = R_\theta(b)$ numerically for $0 < a < \theta < b < \infty$ is straightforward, since $R_\theta(x)$ is increasing in $0 < x < \theta$ and decreasing in $\theta < x < \infty$; and $P_\theta[a \leq X \leq b] = F_\theta(b) - F_\theta(a)$.

Write $a = a_\theta$ and $b = b_\theta$ to emphasize the dependence of these quantities on θ . Figure 2 displays the graph of a_θ and b_θ when $I = 6$ and $J = 5$. It is shown below that a_θ/θ and b_θ/θ are non-decreasing. Thus, a_θ and b_θ are strictly increasing where $a_\theta > 0$. So, for any $x > 0$, the equation $a_\theta = x$ has a unique solution $\theta = u(x)$. Similarly, for any $x > b_1$, the equation $b_\theta = x$ has a unique solution $\theta = \ell(x)$. Let $\ell(x) = 1$ for $x \leq b_1$. Then

$$\{\theta : R_\theta(x) \geq c_\theta\} = [\ell(x), u(x)].$$

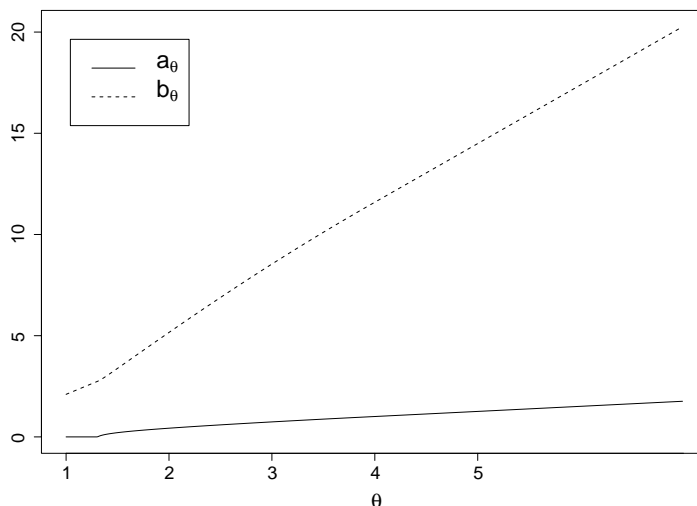


Figure 2: Plots of a_θ and b_θ when $I = 6$ and $J = 5$ when $\alpha = 0.1$.

The unified intervals have exact coverage probability $1 - \alpha$ for all θ by construction. Graphs of $\ell(x)$ and $u(x)$ are included in Figure 3.

Monotonicity. It remains to show that $\tilde{a}_\theta := a_\theta/\theta$ and $\tilde{b}_\theta := b_\theta/\theta$ are non-decreasing in θ . Let

$$\tilde{R}_\theta(x) = R_\theta[\theta x].$$

Then

$$\tilde{R}_\theta(x) = \frac{[s + r\theta x]^{r+s}}{\theta^r [s + rx]^{r+s}} \text{ if } \theta x \leq 1,$$

$$\tilde{R}_\theta(x) = \frac{[s + r]^{r+s} x^r}{[s + rx]^{r+s}} \text{ if } \theta x > 1,$$

and

$$[\tilde{a}_\theta, \tilde{b}_\theta] = \{x : \tilde{R}_\theta(x) \geq c_\theta\},$$

where

$$1 - \alpha = P_1\{x : \tilde{R}_\theta(x) \geq c_\theta\} = F[\tilde{b}_\theta] - F[\tilde{a}_\theta]$$

and F is the F-distribution with $I - 1$ and $I(J - 1)$ degrees of freedom.

LEMMA 1 $\tilde{R}_\theta(x)$ is strictly decreasing in θ on the interval where $\theta x \leq 1$. Further, if $1 \leq \theta_1 < \theta_2 < \infty$, then $R_{\theta_2}(x) \leq R_{\theta_1}(x)$ for all $0 \leq x < \infty$ with strict inequality if $x \leq 1/\theta_1$.

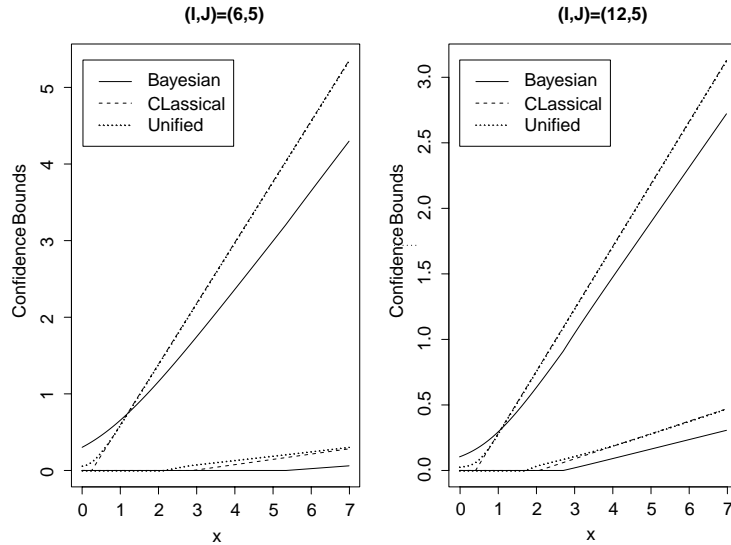


Figure 3: Confidence and Credible intervals for λ when $\alpha = 0.1$.

Proof. On the interval where $\theta x < 1$,

$$\frac{d}{d\theta} \log \tilde{R}_\theta(x) = (r + s) \frac{rx}{s + \theta rx} - \frac{r}{\theta} < 0,$$

so that $\tilde{R}_\theta(x)$ strictly decreasing in θ . Next, if $1 \leq \theta_1 < \theta_2 < \infty$, then clearly $\tilde{R}_{\theta_2}(x) < \tilde{R}_{\theta_1}(x)$ for $x < 1/\theta_2$ and $\tilde{R}_{\theta_2}(x) = \tilde{R}_{\theta_1}(x)$ for $x > 1/\theta_1$. If $1/\theta_2 < x < 1/\theta_1$, let $\theta_0 = 1/x$. $\theta_1 < \theta_0 < \theta_2$ and $\tilde{R}_{\theta_2}(x) = \tilde{R}_{\theta_0}(x) > \tilde{R}_{\theta_1}(x)$. □

PROPOSITION 1 *If $\alpha \leq 1 - F(1)$, then \tilde{a}_θ and \tilde{b}_θ non-decreasing in θ .*

PROOF. First observe that if $1 \leq \theta_1 < \theta_2$, then

$$P_1\{x : \tilde{R}_{\theta_2}(x) \geq c_{\theta_1}\} \leq P_1\{x : \tilde{R}_{\theta_1}(x) \geq c_{\theta_1}\} = 1 - \alpha,$$

so that $c_{\theta_2} \leq c_{\theta_1}$; that is, c_θ is non-increasing in θ . Next, $\tilde{b}_\theta \geq 1$ for all θ , since $\alpha \leq 1 - F(1)$. So,

$$c_\theta = \tilde{R}_\theta[\tilde{b}_\theta] = \frac{[s + r]^{r+s} \tilde{b}_\theta^r}{[s + r\tilde{b}_\theta]^{r+s}},$$

and, therefore, \tilde{b}_θ is non-decreasing in θ . That \tilde{a}_θ is non-decreasing in θ then follows from $F_0[\tilde{a}_\theta] = F_0[\tilde{b}_\theta] - (1 - \alpha)$. □

3. Bayesian Credible Intervals

Bayesian analyses of variance component models have been provided by several authors of which Hill (1965) and Tiao and Tan (1965) provide early examples, and Natarajan and Kass (2000) a recent one. The discussion here loosely follows that of Hill (1965). If θ is given the improper prior distribution $d\theta/\theta$, $1 \leq \theta < \infty$, then the marginal density of X and the posterior density of θ given x are

$$\bar{f}(x) = \int_1^\infty f_\theta(x) \frac{d\theta}{\theta} = \int_1^\infty f\left(\frac{x}{\theta}\right) \frac{d\theta}{\theta^2} = \frac{1}{x} F(x),$$

and

$$g(\theta|x) = \frac{f_\theta(x)}{f(x)} = \frac{xf(x/\theta)}{\theta^2 F(x)} \quad (7)$$

for $1 \leq \theta < \infty$, where f and F are the density and distribution function of \mathcal{F} . Alternatively, x/θ has a posterior F-distribution, conditioned not to exceed x . It follows that the posterior distribution function of θ given x is

$$G(\theta|x) = \frac{F(x) - F(x/\theta)}{F(x)}.$$

By differentiating (7), it is easily seen that the posterior mode of $g(\theta|x)$ is

$$\tilde{\theta} = \max\left\{1, \frac{(s-1)r}{s(r+1)}x\right\}.$$

Moreover, $g(\theta|x)$ is decreasing in $\theta > \tilde{\theta}$ and increasing in $1 < \theta < \tilde{\theta}$ when $\tilde{\theta} > 1$. So, $1 - \alpha$ Bayesian credible intervals for θ of minimal length have the form $[\ell, u]$, where $\ell = \ell(x)$ and $u = u(x)$ solve the equations

$$G(u|x) - G(\ell|x) = 1 - \alpha$$

and

$$[\ell, u] = \{\theta : g(\theta|x) \geq c\}$$

for some $c > 0$. (See Berger, 1980, Page 266). Let

$$x^o = \frac{(r+1)s}{r(s-1)} \leq \infty.$$

If $0 \leq x \leq x^o$, then $\tilde{\theta}(x) = 1$ and $g(\theta|x)$ is decreasing in $1 \leq \theta < \infty$. In this case $\ell = 1$ and u must solve $G(u|x) = 1 - \alpha$; that is, $u(x) = x/F^{-1}[\alpha F(x)]$ when $\tilde{\theta}(x) = 1$. If $\alpha < F(x^o)$, then the following algorithm can be used to find ℓ and u for large values of x :

Bayesian Algorithm.

- i) Let $z = x/F^{-1}[\alpha F(x)]$, so that $G(z|x) = 1 - \alpha$, as above.
- ii) If $g(1|x) \geq g(z|x)$, then let $u = z$ and $\ell = 1$.
- ii') Otherwise: let $c_0 = g(1|x)$, $c_1 = g(z|x)$, and iterate iii) to convergence.
- iii) Let $c = (c_0 + c_1)/2$; solve the equations $g(\ell|x) = c = g(u|x)$ for ℓ and u in the range $1 < \ell < \tilde{\theta} < u$. If $G(u|x) - G(\ell|x) \leq 1 - \alpha$, let $c_1 = c$ otherwise, let $c_0 = c$ and iterate.

The condition $\alpha < F(x^o)$ insures that $z > \tilde{\theta}$ in i). In iii), solving $g(\ell|x) = c = g(u|x)$ is straightforward since $g(\theta|x)$ is increasing in $1 < \theta < \tilde{\theta}$ and decreasing in $\tilde{\theta} < \theta < \infty$.

Figure 3 displays the classical and unified confidence intervals along with the Bayesian credible intervals for selected I and J . Observe that the unified intervals do have a non-empty interior for all x in contrast to the classical ones. The Bayesian upper confidence bound is larger than the the classical and unified ones for small x , but smaller moderate for large x .

Frequentist Coverage Probability. It is shown below that $u(x)$ is strictly increasing and that $\ell(x)$ is strictly increasing where it exceeds one. Thus, for any $\theta > 1$, the equation $\ell(x) = \theta$ has a unique solution $b = b_\theta$. Similarly, for $\theta \geq u(0)$, the equation $u(x) = \theta$ has a unique solution $a = a_\theta$. Let $a_\theta = 0$ for $1 \leq \theta < u(0)$. Then $\ell(x) \leq \theta \leq u(x)$ if and only if $a_\theta \leq x \leq b_\theta$, and the frequentist coverage probability of the Bayesian credible intervals is

$$P_\theta[\ell(X) \leq \theta \leq u(X)] = \int_{a_\theta}^{b_\theta} f_\theta(x)dx = F\left[\frac{b_\theta}{\theta}\right] - F\left[\frac{a_\theta}{\theta}\right] \tag{8}$$

for $1 \leq \theta < \infty$ (see Figure 4). Figure 5 displays the coverage probability of the Bayesian method as a function of θ for selected I and J when $\alpha = 0.1$. Observe that the minimum frequentist coverage probability is only slightly less than the credible level $1 - \alpha$. For the selected cases, the minimum coverage probabilities are: 0.8983 when $(I, J) = (6, 5)$, 0.8999 when $(I, J) = (6, 200)$, 0.8934 when $(I, J) = (12, 5)$ and 0.8980 when $(I, J) = (12, 200)$.

Together Figures 3 and 5 reveal the following stark contrast between the Bayesian and frequentist viewpoints: From a frequentist perspective, the Bayesian intervals are too long (on the average) for small λ , since the actual coverage probability greatly exceeds the nominal value. From a Bayesian perspective, the classical intervals are not believable for small x , since they

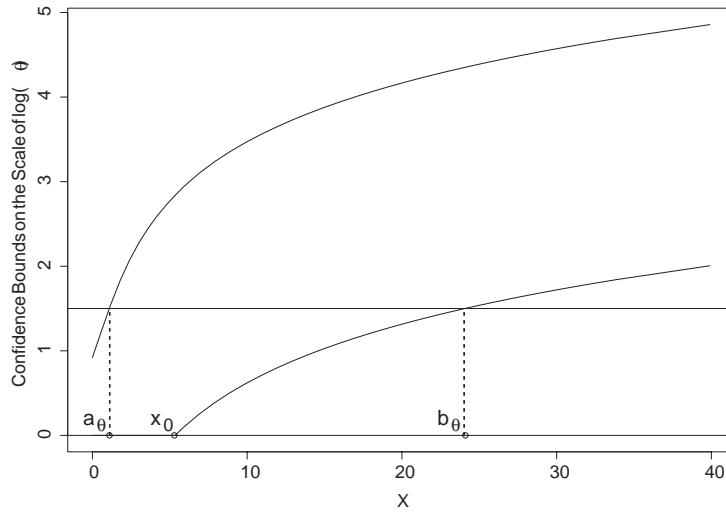


Figure 4: Solutions for a_θ & b_θ in (8) and x_0 in (9).

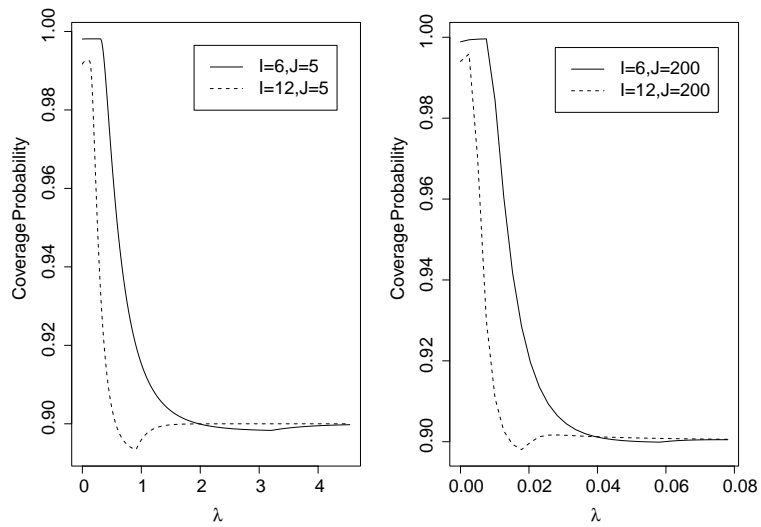


Figure 5: Frequentist coverage probability of Bayesian intervals for selected I and J when $\alpha = 0.1$.

degenerate to a point. The unified intervals offer a compromise, but still seem too short from a Bayesian perspective.

Monotonicity. It remains to show that ℓ and u are monotone. Recall that $\tilde{\theta}(x) = 1$ for $0 \leq x \leq x^o = (r + 1)s/[r(s - 1)]$, and let

$$x_0 = \inf\{x : \ell(x) > 1\}. \tag{9}$$

Then $x^o < x_0$, since $g(\theta|x)$ is decreasing in θ when $x \leq x^o$. It is also relevant that $x^2f(x)$ is increasing in $0 \leq x \leq x^o$ and decreasing in $x^o \leq x < \infty$. Let

$$\tilde{\ell}(x) = \frac{x}{u(x)} \quad \text{and} \quad \tilde{u}(x) = \frac{x}{\ell(x)}$$

and note the reversal. Then $[\tilde{\ell}(x), \tilde{u}(x)] = \{z : z^2f(z) \geq c, 0 \leq z \leq x\}$, where $c = c_x$ is uniquely determined by

$$F[\tilde{u}(x)] - F[\tilde{\ell}(x)] = (1 - \alpha)F(x); \tag{10}$$

and $c_x = \tilde{\ell}(x)^2f[\tilde{\ell}(x)] \leq \tilde{u}(x)^2f[\tilde{u}(x)]$ with equality if $\tilde{u}(x) < x$.

LEMMA 2 . *If $\alpha < F(x^o)$, then $\tilde{\ell}(x) < x^o$ for all x , and $\ell(x) > 1$ for all $x_0 < x < \infty$.*

PROOF. Since $\tilde{u}(x) \leq x$ for all x , $F[\tilde{\ell}(x)] = F[\tilde{u}(x)] - (1 - \alpha)F(x) \leq \alpha F(x) \leq \alpha$ and, therefore, $\tilde{\ell}(x) \leq F^{-1}(\alpha) < x^o$. Next, let $v(x) = F^{-1}[\alpha F(x)]$ and $z = x/v(x)$. Then, from the algorithm for ℓ and u , $\ell(x) = 1$ iff $g(1|x) \geq g(z|x)$ iff $x^2f(x) \geq v(x)^2f[v(x)^2]$. So, if $x^o < x_1 < x_2$ and $\ell(x_2) = 1$, then $x_1^2f(x_1) \geq x_2^2f(x_2) \geq v(x_2)^2f[v(x_2)] \geq v(x_1)^2f[v(x_1)]$, so that $\ell(x_1) = 1$. □

PROPOSITION 2 *$\ell(x)$ is strictly increasing in $x_0 < x < \infty$; and $u(x)$ is strictly increasing in $0 \leq x < \infty$. Furthermore, $\tilde{\ell}(x)$ is strictly decreasing in $x_0 \leq x < \infty$; $\tilde{u}(x)$ is strictly increasing over the same range; and $\tilde{\ell}(x)$ and $\tilde{u}(x)$ are both strictly increasing in $0 \leq x < x_0$.*

PROOF. The proof uses elementary properties of totally positive functions. Relevant background can be found in Karlin (1968) and Lehmann (1986). If $0 \leq x \leq x_0$, then $\ell(x) = 1$ and $G[u(x)|x] = 1 - \alpha$. Now $f_\theta(x)$ is TP_2 (totally positive of order two) in (θ, x) . So, $g(\theta|x)$ is TP_2 in (x, θ) and, therefore $G(\theta|x_1) > G(\theta|x_2)$ for all $\theta > 1$ when $0 < x_1 < x_2 < \infty$. See, for example, Lehmann (1986, p.84). So, if $0 < x_1 < x_2 \leq x_0$, then

$G[u(x_1)|x_2] < G[u(x_1)|x_1] = 1 - \alpha$ and, therefore, $u(x_2) > u(x_1)$. That is, $u(x)$ is strictly increasing in $0 \leq x \leq x_0$.

It is shown next that $\tilde{\ell}(x)$ is strictly decreasing in $x_0 \leq x < \infty$ and that $\tilde{u}(x)$ is strictly increasing over the same range. To see this, let $x_0 \leq x_1 < x_2 < \infty$. If $\tilde{u}(x_2) \leq \tilde{u}(x_1)$, then $\tilde{\ell}(x_1)^2 f[\tilde{\ell}(x_1)] = \tilde{u}(x_1)^2 f[\tilde{u}(x_1)] \leq \tilde{u}(x_2)^2 f[\tilde{u}(x_2)] = \tilde{\ell}(x_2)^2 f[\tilde{\ell}(x_2)]$, so that $\tilde{\ell}(x_1) \leq \tilde{\ell}(x_2)$; but then

$$F[\tilde{u}(x_2)] - F[\tilde{\ell}(x_2)] \leq F[\tilde{u}(x_1)] - F[\tilde{\ell}(x_1)] = (1 - \alpha)F(x_1) < (1 - \alpha)F(x_2),$$

contradicting (10). So, $\tilde{u}(x_1) < \tilde{u}(x_2)$ and, therefore, $\tilde{\ell}(x_1) > \tilde{\ell}(x_2)$, since $\tilde{\ell}^2 f(\tilde{\ell}) = \tilde{u}^2 f(\tilde{u})$, as above.

If $0 \leq x < x_0$, then $\tilde{u}(x) = x$ and (10) becomes $F[\tilde{\ell}(x)] = \alpha F(x)$. Thus $\tilde{\ell}(x)$ increases in x over this range.

That $u(x) = x/\tilde{\ell}(x)$ is strictly increasing in $x_0 \leq x < \infty$ follows directly. For ℓ , first observe that $\tilde{\ell}(x)$ and $\tilde{u}(x)$ are differentiable in $x_0 < x < \infty$, by the Implicit Function Theorem. It then follows that ℓ and u are differentiable and that $\tilde{\ell}' = [1 - \tilde{\ell}u']/u$, $\tilde{u}' = [1 - \tilde{u}\tilde{\ell}']/\tilde{\ell}$, and

$$\tilde{u}'f(\tilde{u}) - \tilde{\ell}'f(\tilde{\ell}) = (1 - \alpha)f(x)$$

for $x_0 < x < \infty$. Combining these relations with $\tilde{\ell}^2 f(\tilde{\ell}) = \tilde{u}^2 f(\tilde{u})$ and $\tilde{u} \leq x$, then leads to

$$(\ell' - u')\tilde{\ell}^2 f(\tilde{\ell}) = \tilde{u}f(\tilde{u}) - \tilde{\ell}f(\tilde{\ell}) - (1 - \alpha)xf(x) > -\tilde{\ell}f(\tilde{\ell});$$

that is,

$$\ell' > u' - \frac{1}{\tilde{\ell}} = -u\frac{\tilde{\ell}'}{\tilde{\ell}} \geq 0$$

to complete the proof. □

4. Minimum Frequentist Coverage Probability

Analytical lower bounds for the coverage probability of the Bayesian intervals are considered in this section.

PROPOSITION 3 *If $\alpha < F(x_0)$, then the coverage probability of the Bayesian credible interval is at least $(1 - \alpha)F(x_0)$, where x_0 is defined in (9).*

PROOF. The values a_θ and b_θ are uniquely determined in equation (8). For any $\theta \geq 1$, $b_\theta/\theta = \tilde{u}(b_\theta) \geq \tilde{u}(x_0)$, since $\tilde{u}(x)$ is strictly increasing in

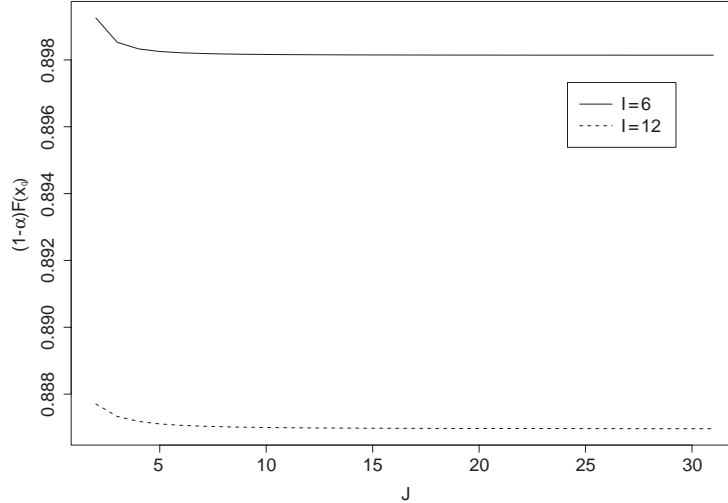


Figure 6: Graphs of $(1 - \alpha)F(x_0)$ as a function of J for selected I when $\alpha = 0.1$.

$x_0 \leq x < \infty$ and $b_\theta \geq x_0$. When $\theta > u(0)$, $a_\theta/\theta = \tilde{\ell}(a_\theta) \leq \tilde{\ell}(x_0)$, since the maximum of $\tilde{\ell}(x)$ is x_0 . Thus,

$$\begin{aligned} P_\theta[\ell(X) \leq \theta \leq u(X)] &= F[\tilde{u}(b_\theta)] - F[\tilde{\ell}(a_\theta)] \\ &\geq F[\tilde{u}(x_0)] - F[\tilde{\ell}(x_0)] = (1 - \alpha)F(x_0). \end{aligned}$$

When $\theta \leq u(0)$, $a_\theta = 0$. Thus

$$P_\theta[\ell(X) \leq \theta \leq u(X)] = F[\tilde{u}(b_\theta)] \geq F[\tilde{u}(x_0)] = (1 - \alpha)F(x_0).$$

Hence, for all $\theta \geq 1$, the coverage probability is not less than $(1 - \alpha)F(x_0)$. \square

Figure 6 displays the values of $(1 - \alpha)F(x_0)$ as a function of $J \geq 2$ for selected I . Observe that $(1 - \alpha)F(x_0)$ exceeds .89 for $I = 6$, and .88 for $I = 12$. Observe too that $(1 - \alpha)F(x_0)$ decreases when J increases. Figure 7 displays the values of $\min_{J \geq 2}(1 - \alpha)F(x_0)$ as a function of I . Observe that this quantity decreases slowly as I increase to a minimum of about .837 when $I = 600$.

When $s > 1$, it is possible to obtain an explicit lower bound for $F(x_0)$, leading to the lower bound $(1 - \alpha)/(1 + \alpha)$ for the coverage probability. To

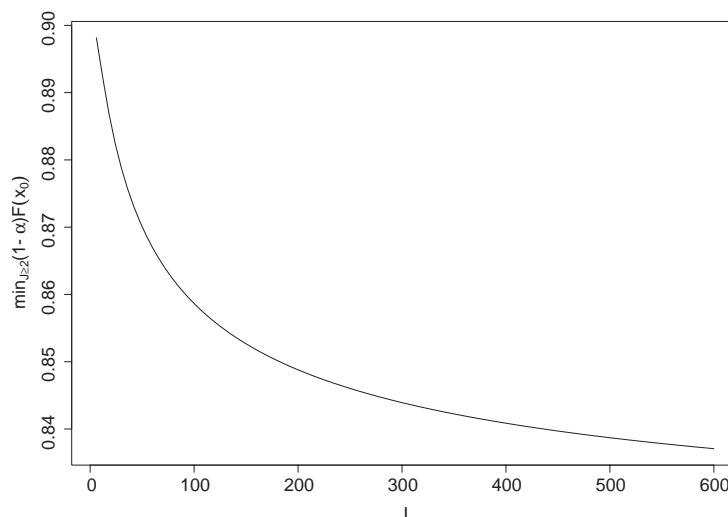


Figure 7: Graph of $\min_{J \geq 2} (1 - \alpha)F(x_0)$ as a function of I when $\alpha = 0.1$.

derive it, let $h(x) = \log[x^2 f(x)]$, so that

$$h'(x) = \frac{r+1}{x} - \frac{(r+s)r}{s+rx} = \frac{(r+1)s - r(s-1)x}{x(s+rx)}$$

and h attains its maximum at $x^o = (r+1)s/[r(s-1)]$. Then, for each $0 < z < x^o$, there is a unique $y = y(z)$ for which $x^o < y < \infty$ and $h(y) = h(z)$. Further, y is continuously differentiable and

$$h'(z) = h'[y(z)]y'(z)$$

for $0 < z < x^o$.

PROPOSITION 4 For $0 < z < x^o$, $1 - F[y(z)] \leq F(z)$.

PROOF. If $0 < z < x^o$, then

$$F(z) = \int_0^z f(x) dx$$

and

$$1 - F[y(z)] = \int_{y(z)}^{\infty} f(x) dx = \int_0^z f[y(x)]|y'(x)| dx = \int_0^z f(x) \frac{x^2 |y'(x)|}{y(x)^2} dx.$$

So, it suffices to show that

$$\frac{x^2|y'(x)|}{y(x)^2} \leq 1 \tag{11}$$

for $0 < x < x^o$. Observe that

$$\frac{x|y'|}{y} = \left(\frac{s + ry}{s + rx}\right) \frac{x^o - x}{y - x^o}.$$

Here

$$\frac{s + ry}{s + rx} \leq \frac{y}{x},$$

since $x \leq y$. So, (11) would be implied by $y - x^o \geq x^o - x$, which would be implied by

$$h(x^o - w) \leq h(x^o + w) \tag{12}$$

for $0 < w < x^o$. To establish (12), let

$$\begin{aligned} k(w) &= h(x^o - w) - h(x^o + w) \\ &= (r + 1)[\log(x^o - w) - \log(x^o + w)] \\ &\quad - (r + s)\{\log[s + r(x^o - w)] - \log[s + r(x^o + w)]\}. \end{aligned}$$

Then $k(0) = 0$, $k'(0) = 0$, and

$$k'(w) = -(r + 1)\frac{2x^o}{(x^o)^2 - w^2} + (r + s)\frac{2t}{t^2 - w^2},$$

where $t = s/r + x^o$. So,

$$\begin{aligned} ((x^o)^2 - w^2)(t^2 - w^2)k'(w) &= 2(r + s)t((x^o)^2 - w^2) - 2(r + 1)x^o(t^2 - w^2) \\ &= 2[(r + 1)x^o - (r + s)t]w^2 \\ &< 0, \end{aligned}$$

and the proposition follows. □

COROLLARY 1 $F(x_0) \geq 1/(1 + \alpha)$.

PROOF. From (9), $\tilde{u}(x_0) = x_0$ and $\tilde{\ell}(x_0)^2 f[\tilde{\ell}(x_0)] = \tilde{u}(x_0)^2 f[\tilde{u}(x_0)]$. So, $x_0 = y[\tilde{\ell}(x_0)]$ in the notation of Proposition 4 and, therefore, $1 - F(x_0) \leq F[\tilde{\ell}(x_0)]$; and $F[\tilde{\ell}(x_0)] = \alpha F(x_0)$ from the definition of the credible intervals. So, $1 - F(x_0) \leq \alpha F(x_0)$, as asserted. □

References

- BERGER, J. (1980). *Statistical Decision Theory: Foundations, Concepts, and Methods*, Springer-Verlag, New York.
- BURDICK, R. and GRAYBILL, F. (1992). *Confidence Intervals on Variance Components*. Marcel Dekker, New York.
- FELDMAN, G.J. and ROBERT, C. (1998). Unified approach to the classical statistical analysis of small signals. *Physical Review D*, **57**, 3873-3889.
- HILL, B.M. (1965). Inference about variance components in the one-way model. *J. Amer. Statist. Assoc.*, **60**, 806-825.
- KARLIN, S. (1968). *Total Positivity*. Stanford University Press, Stanford, CA.
- LEHMANN, E.L. (1986). *Testing Statistical Hypothesis*. Wiley, New York.
- NATARAJAN, R. and KASS, R. (2000). Reference Bayesian methods for generalized linear models. *J. Amer. Statist. Assoc.*, **95**, 227-237.
- ROE, B. and WOODROOFE, M. (1999). Improved probability method for estimating signal in the presence of background. *Physical Review D*, **60**, 053009/(1-5).
- ROE, B. and WOODROOFE, M. (2000). Setting confidence belts. *Physical Review D*, **63**, 013009/(1-9).
- SCHEFFÉ, H. (1959). *The Analysis of Variance*. Wiley, New York.
- SEARLE, SHAYLE R., CASELLA, G. and MCCULLOCH, C.E. (1992). *Variance Components*. Wiley, New York.
- TIAO, G.C. and TAN, W.Y. (1965). Bayesian analysis of random-effect models in the analysis of variance I. Posterior distribution of variance components. *Biometrika*, **52**, 37-53.

TONGLIN ZHANG
DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
WEST LAFAYETTE, IN 47907-1399
USA
E-mail: tlzhang@stat.purdue.edu

MICHAEL WOODROOFE
DEPARTMENT OF STATISTICS
UNIVERSITY OF MICHIGAN
462 WEST HALL, EAST UNIVERSITY
ANN ARBOR, MI 48109-109
USA
E-mail: michaelw@stat.lsa.umich.edu