Approximate posterior quantiles of a parametric function with application to Bayesian tolerance limits

by Rahul Mukerjee
Indian Institute of Management, Post Box No. 16757, Calcutta 700 027, India
rmuk1@hotmail.com
and N. Reid
Department of Statistics, University of Toronto, Toronto Canada M5S 3G3
reid@utstat.utoronto.ca

Summary

We give an explicit higher order asymptotic formula for the posterior quantiles of any smooth parametric function and indicate conditions for the approximate frequentist validity of such quantiles. It is shown that the results can be readily applied to the derivation of accurate Bayesian tolerance limits. Numerical studies suggest that the use of the formula given here can entail a significant gain in computational accuracy.

Some key words: Frequentist validity; Marginal posterior density; Second order.

1 Introduction

The study of posterior quantiles of a canonical parameter or a parametric function has been of substantial current interest. Recent work in this area centers around the derivation of probability matching priors which ensure approximate frequentist validity of posterior quantiles – see for example, Tibshirani (1989), Datta and Ghosh (1995), Sun and Ye (1996), Mukerjee and Ghosh (1997), Mukerjee and Reid (1999), and the references therein. Quite a few of these authors noted that consideration of asymptotics with margin of error $o(n^{-1})$, as opposed to $o(n^{-1/2})$, where n is the sample size, helps in significantly narrowing down the class of probability matching priors. The underlying algebra, however, bypasses the derivation of an explicit asymptotic formula, up to that order of approximation, for the posterior quantiles themselves. In fact, as noted for example in Mukerjee and Ghosh (1997), if one works with the sole objective of characterizing probability matching priors then even an expansion for the relevant marginal posterior density, with explicit description of all the terms, is not required.

On the other hand, from the point of view of the ultimate user, there is an obvious need for a higher order asymptotic formula for posterior quantiles which can be directly employed in the construction of accurate credible sets. In particular, as seen in Section 3 below, such a formula can find an important application in the construction of Bayesian tolerance limits which are of immediate industrial use. The present paper addresses this problem and provides an explicit formula, with margin of error $o(n^{-1})$, for the posterior quantiles of any smooth parametric function. This is done in Section 2. The formula is valid even for priors that are not probability matching and hence applicable even in a purely Bayesian setup. However, in Section 2 we also indicate the conditions for the approximate frequentist validity of such posterior quantiles. Compared to the existing work on probability matching priors, the algebra gets more intricate since now an expansion for the relevant marginal posterior density has to be worked out in full detail. At the expense of heavier algebra, however, we achieve immediate applicability of the results. Thus, in Section 3, it is seen that the results can be readily used in the derivation of accurate Bayesian tolerance limits, an application that has been hitherto unnoticed.

Earlier, Datta and Ghosh (1995) gave a formula for approximate posterior quantiles with margin of error $o(n^{-1/2})$. In keeping with the current trend of research in this area, it would be theoretically satisfying to improve their result to a margin of error $o(n^{-1})$ which is precisely what has been done here. More importantly, the numerical examples in Section 4 demonstrate that the use of our higher order asymptotic formula entails a significant gain in computational accuracy over that of theirs.

2 The main formula

2.1 Setup and some notation

Consider a sequence $\{X_i\}$, $i \geq 1$, of independent and identically distributed random variables with common density $f(x;\theta)$ where $\theta = (\theta_1, \ldots, \theta_p)'$ is an unknown parameter vector lying in $\mathcal{R}^p$ or some open subset thereof. We work essentially under the assumptions of Johnson (1970) and also need the Edgeworth assumptions of Bickel and Ghosh (1990, p. 1078). Let θ have a prior density $\pi(\cdot)$ which is positive and thrice continuously differentiable for all θ. Write $\ell(\theta) = n^{-1}\sum_{i=1}^{n} \log f(X_i;\theta)$ and let $\hat{\theta}$ be the maximum likelihood estimator of θ based on $X = (X_1, \ldots, X_n)'$, where n is the sample size. Also, for $1 \leq j,\ r,\ s,\ u \leq p$, let

$D_j \equiv \partial/\partial\theta_j$,

$$a_{jr} = \{D_j D_r \ell(\theta)\}_{\theta=\hat\theta} \;, \qquad a_{jrs} = \{D_j D_r D_s \ell(\theta)\}_{\theta=\hat\theta} \;,$$
$$a_{jrsu} = \{D_j D_r D_s D_u \ell(\theta)\}_{\theta=\hat\theta} \;, \qquad c_{jr} = -a_{jr} \;,$$
$$\pi_j(\theta) = D_j\pi(\theta) \;, \quad \pi_{jr}(\theta) = D_j D_r \pi(\theta) \;, \quad \hat\pi = \pi(\hat\theta) \;, \quad \hat\pi_j = \pi_j(\hat\theta) \;, \quad \hat\pi_{jr} = \pi_{jr}(\hat\theta) \;.$$

All formal expansions for the posterior, as used here, are valid for sample points in a set S with P_θ-probability $1 + o(n^{-1})$ uniformly over compact sets of θ. The set S may be defined along the line of Bickel and Ghosh (1990). The matrix $C = ((c_{jr}))$ is positive definite over S. Let $C^{-1} = ((c^{jr}))$.

We shall be interested in the posterior quantiles of a smooth parametric function $g(\theta)$. Assume that the gradient vector $\nabla g(\theta) = (g_1(\theta), \ldots, g_p(\theta))'$, where $g_j(\theta) = D_j g(\theta)$, is nonnull for every θ. For $1 \le j,\, r,\, s \le p$, let

$$g_{jr}(\theta) = D_j D_r g(\theta) \;, \qquad g_{jrs}(\theta) = D_j D_r D_s g(\theta) \;,$$
$$\hat g_j = g_j(\hat\theta) \;, \quad \hat g_{jr} = g_{jr}(\hat\theta) \;, \quad \hat g_{jrs} = g_{jrs}(\hat\theta) \;,$$
$$m_j = c^{jr}\hat g_r \;, \quad \hat g_{j0} = \hat g_{jr} m_r \;, \quad a_{jr0} = a_{jrs} m_s \;,$$
$$a_{j00} = a_{jrs} m_r m_s \;, \quad a_{jr00} = a_{jrsu} m_s m_u \;.$$

In the above and elsewhere, we follow the summation convention with each implicit sum ranging from 1 to p. Let

$$\nabla\hat g = (\hat g_1, \ldots, \hat g_p)' \;, \quad b = \{(\nabla\hat g)' C^{-1} (\nabla\hat g)\}^{1/2} \;, \quad \hat g_{00} = \hat g_{jr} m_j m_r \;,$$
$$a_{000} = a_{jrs} m_j m_r m_s \;, \quad a_{0000} = a_{jrsu} m_j m_r m_s m_u \;.$$

2.2 Approximate posterior quantile

Let

$$Y = n^{1/2}\{g(\theta) - g(\hat\theta)\} \;. \tag{2.1}$$

Then

$$Y = h_j \hat g_j + \frac{1}{2} n^{-1/2} h_j h_r \hat g_{jr} + \frac{1}{6} n^{-1} h_j h_r h_s \hat g_{jrs} + o(n^{-1}) \;, \tag{2.2}$$

where $h = (h_1, \ldots, h_p)' = n^{1/2}(\theta - \hat{\theta})$. Ghosh and Mukerjee (1991) gave an expansion, with margin of error $o(n^{-1})$, for the posterior joint density of h under the prior $\pi(\cdot)$. Using that expansion and (2.2), the approximate posterior characteristic function of Y, under the prior $\pi(\cdot)$, is given by

$$\begin{aligned}\exp\Big(\frac{1}{2}\xi^2 b^2\Big)\Big[1 &+ n^{-1/2}\{\xi G_1(\pi) + \xi^3 G_3\} \\ &+ n^{-1}\{\xi^2 G_2(\pi) + \xi^4 G_4(\pi) + \xi^6 G_6\}\Big] + o(n^{-1}) ,\end{aligned}$$

where $\xi = (-1)^{1/2} t$,

$$G_1(\pi) = (\hat{\pi}_j/\hat{\pi})m_j + \frac{1}{2}v_{jr}c^{jr} , \tag{2.3}$$

$$\begin{aligned}G_2(\pi) &= \frac{1}{2}\{G_1(\pi)\}^2 - \frac{1}{2}\{(\hat{\pi}_j/\hat{\pi})m_j\}^2 + \frac{1}{2}(\hat{\pi}_{jr}/\hat{\pi})m_j m_r + c^{jr}w_j(\hat{\pi}_r/\hat{\pi}) \\ &+ c^{jr}\Big(\frac{1}{2}\hat{g}_{jr0} + \frac{1}{4}a_{jr00}\Big) + c^{js}c^{ru}\Big(\frac{1}{4}v_{jr}v_{su} + \frac{1}{2}w_j a_{rsu}\Big) ,\end{aligned} \tag{2.4}$$

$$G_4(\pi) = G_1(\pi)G_3 + \frac{1}{24}a_{0000} + \frac{1}{6}\hat{g}_{000} + \frac{1}{2}c^{jr}w_j w_r , \tag{2.5}$$

$$G_3 = \frac{1}{6}a_{000} + \frac{1}{2}\hat{g}_{00} , \quad G_6 = \frac{1}{2}G_3^2 , \tag{2.6}$$

with

$$v_{jr} = \hat{g}_{jr} + a_{jr0} , \quad w_j = \hat{g}_{j0} + \frac{1}{2}a_{j00} . \tag{2.7}$$

The above follows after a long process of integration which is being omitted here to save space. The invariance of quantities like $\hat{g}_{jr}$, a_{jrs}, c^{jr} etc. with respect to permutations of their sub- or superscripts helps in simplifying the algebra to some extent.

We now invert the approximate posterior characteristic function of Y as stated above, a step which can be justified as in Chandra and Ghosh (1979). This yields an expansion for the marginal posterior density of Y, under the prior $\pi(\cdot)$, as

$$\begin{aligned}\pi^*(y|X) &= b^{-1}\phi(y/b)\Big[1 + n^{-1/2}\{b^{-1}G_1(\pi)J_1(y/b) + b^{-3}G_3J_3(y/b)\} \\ &+ n^{-1}\Big\{b^{-2}G_2(\pi)J_2(y/b) + b^{-4}G_4(\pi)J_4(y/b) + b^{-6}G_6J_6(y/b)\}\Big] \\ &+ o(n^{-1}) ,\end{aligned} \tag{2.8}$$

where $\phi(\cdot)$ is the standard univariate normal density and $J_i(\cdot)$ is the ith degree Hermite polynomial.

For $0 < \alpha < 1$, let

$$Q(\pi, 1-\alpha) = g(\hat{\theta}) + n^{-1/2} b\{z_\alpha + n^{-1/2} K_1(\pi, \alpha) + n^{-1} K_2(\pi, \alpha)\} , \tag{2.9}$$

where z_α is the $(1-\alpha)$th quantile of a standard normal variate,

$$\begin{aligned} K_1(\pi, \alpha) &= b^{-1} G_1(\pi) + b^{-3} G_3 J_2(z_\alpha) , & (2.10) \\ K_2(\pi, \alpha) &= b^{-2} G_2(\pi) J_1(z_\alpha) + b^{-4} G_4(\pi) J_3(z_\alpha) + b^{-6} G_6 J_5(z_\alpha) \\ &+ 2b^{-3} G_3 z_\alpha K_1(\pi, \alpha) - \frac{1}{2} z_\alpha \{K_1(\pi, \alpha)\}^2 . & (2.11) \end{aligned}$$

By (2.1) and (2.9), we have $g(\theta) \leq Q(\pi, 1-\alpha)$ if and only if

$$Y \leq b\{z_\alpha + n^{-1/2} K_1(\pi, \alpha) + n^{-1} K_2(\pi, \alpha)\} .$$

Hence, writing $P^\pi\{\cdot|X\}$ for the posterior probability measure under $\pi(\cdot)$ given X, by (2.8), (2.10) and (2.11) one can check that

$$P^\pi\Big\{g(\theta) \leq Q(\pi, 1-\alpha)|X\Big\} = 1 - \alpha + o(n^{-1}) . \tag{2.12}$$

In consideration of (2.12), $Q(\pi, 1-\alpha)$ gives an approximation to the $(1-\alpha)$th posterior quantile of $g(\theta)$, under $\pi(\cdot)$, with margin of error $o(n^{-1})$. Equations (2.9)-(2.11), which, in conjunction with (2.3)-(2.6), describe $Q(\pi, 1-\alpha)$ explicitly, represent the key formula of this paper.

Remark 1. Datta and Ghosh (1995) gave an approximation to the $(1-\alpha)$th posterior quantile of $g(\theta)$, under $\pi(\cdot)$, with margin of error $o(n^{-1/2})$. Under our notation, their approximation can be expressed as

$$Q^*(\pi, 1-\alpha) = g(\hat{\theta}) + n^{-1/2} b\{z_\alpha + n^{-1/2} K_1(\pi, \alpha)\} . \tag{2.13}$$

Our approximation $Q(\pi, 1-\alpha)$ adds an extra term involving $K_2(\pi, \alpha)$ to the above and thus reduces the margin of error to $o(n^{-1})$.

2.3 Approximate frequentist validity

We now briefly indicate the necessary and sufficient conditions on the prior $\pi(\cdot)$ which ensure approximate frequentist validity of the posterior quantiles of $g(\theta)$ in the sense that

$$P_\theta\{g(\theta) \leq Q(\pi, 1-\alpha)\} = 1 - \alpha + o(n^{-1}) , \tag{2.14}$$

for every α. To that effect, we introduce some more notation. For $1 \le j,\ r,\ s \le p$, let

$$V_j = D_j \log f(X_1;\theta)\ ,\quad I_{jr} = E_\theta(V_j V_r)\ ,\quad I = ((I_{jr}))\ ,\quad I^{-1} = ((I^{jr}))\ ,$$

$$V_{jrs} = D_j D_r D_s \log f(X_1;\theta)\ ,\quad L_{jrs} = E_\theta(V_{jrs})\ ,\quad \mu_j = I^{jr} g_r(\theta)\ ,$$

$$g_{j0} = g_{jr}(\theta)\mu_r\ ,\quad L_{j00} = L_{jrs}\mu_r\mu_s\ .$$

Also, let

$$\lambda = \{I^{jr} g_j(\theta) g_r(\theta)\}^{1/2}\ ,\quad g_{00} = g_{jr}(\theta)\mu_j\mu_r\ ,\quad L_{000} = L_{jrs}\mu_j\mu_r\mu_s\ .$$

The quantities listed above are all functions of θ.

Then a prior $\pi(\cdot)$ satisfies (2.14) for every α if and only if

$$\Delta_1(\pi,\theta) = 0 \qquad \text{and} \qquad \Delta_2(\pi,\theta) = 0\ , \tag{2.15}$$

where

$$\Delta_1(\pi,\theta) = D_j\left\{\lambda^{-1}\mu_j\pi(\theta)\right\}\ , \tag{2.16}$$

$$\begin{aligned}\Delta_2(\pi,\theta) &= D_r\Big[\{\lambda^{-4}\mu_r\Big(\frac{1}{3}L_{000} + g_{00}\Big) - \lambda^{-2} I^{jr}\Big(\frac{1}{2}L_{j00} + g_{j0}\Big)\Big\}\pi(\theta)\Big] \\ &+ \frac{1}{2} D_j D_r\{\lambda^{-2}\mu_j\mu_r\pi(\theta)\}\ . \end{aligned} \tag{2.17}$$

The derivation of (2.15)-(2.17), which proceeds essentially along the line of Mukerjee and Reid (1999) and involves the use of an auxiliary prior and a shrinkage argument, is omitted here. A prior $\pi(\cdot)$ satisfying (2.15) and hence (2.14) will be called second order probability matching for $g(\theta)$. So far, the characterization of such a prior was known only for a canonical parameter like the one given by $g(\theta) = \theta_1$; see Mukerjee and Ghosh (1997). The conditions in (2.15) can be seen to cover their findings as a special case.

We emphasize at this stage that the explicit formula for approximate posterior quantiles, as given in (2.9)-(2.11), is valid even for priors that are not second order probability matching. Thus the formula can be used even in a purely Bayesian setting where the prior $\pi(\cdot)$ is chosen without any concern for probability matching in the sense described above.

3 Bayesian tolerance limits

Now suppose that random variables $\{X_i\}$, $i \ge 1$, are one-dimensional and let $F(x,\theta)$ denote their common cumulative distribution function where θ continues to be possibly multidimensional. As before, let $X = (X_1, \ldots, X_n)'$ be the observation vector and $\pi(\cdot)$ be a prior which is positive and thrice continuously differentiable for all θ. We investigate a Bayesian tolerance interval of the form $[T, \infty)$ attaining the target posterior coverage probability with margin of error $o(n^{-1})$. Here the tolerance limit T depends on both $\pi(\cdot)$ and X and we intend to give an explicit expression for T. Consideration of a one-sided tolerance interval, as done here, is often natural in practice. For example, with a life distribution, commonly one has reason to be concerned only when life is too short and not when life is too long.

The tolerance interval $[T, \infty)$ is required to be such that the relation

$$P^{\pi}\{1 - F(T; \theta) \geq \beta | X\} = \alpha + o(n^{-1}) \ , \tag{3.1}$$

holds where α and β $(0 < \alpha, \beta < 1)$ are preassigned. Thus the interval $[T, \infty)$ covers at least a proportion β of the population given by $F(x; \theta)$ with posterior probability $\alpha + o(n^{-1})$. In this sense, it is called a β-content tolerance interval with posterior credibility level $\alpha + o(n^{-1})$; compare Guttman (1970; Ch.9). The interval has frequentist validity up to the same order of approximation provided

$$P_{\theta}\{1 - F(T; \theta) \geq \beta\} = \alpha + o(n^{-1}) \ . \tag{3.2}$$

If the underlying prior $\pi(\cdot)$ ensures (3.2) for every α and β then it will be called second order probability matching for tolerance limits.

Observe that

$$1 - F(T; \theta) \geq \beta \Leftrightarrow g_{\beta}(\theta) \geq T \ ,$$

where $g_{\beta}(\theta)$ is the $(1 - \beta)$th quantile of the population represented by $F(x; \theta)$. Hence by (3.1) T may be interpreted as the $(1 - \alpha)$th posterior quantile of $g_{\beta}(\theta)$ with margin of error $o(n^{-1})$. Consequently, one may obtain T explicitly from (2.9)-(2.11) taking $g(\theta) = g_{\beta}(\theta)$ there. Illustrative examples follow in the next section. Thus the formula given by (2.9)-(2.11) has immediate applicability in the study of Bayesian tolerance limits.

The interpretation of T as a posterior quantile of $g_{\beta}(\theta)$ also implies that the prior $\pi(\cdot)$ is second order probability matching for tolerance limits if and only if, with $g(\theta) = g_{\beta}(\theta)$, it satisfies (2.15) for every β. The following example illustrates this point with reference to a wide-ranging class of models.

Example 1. Consider the location-scale type family of models given by

$$f(x;\theta) = \theta_2^{-1}\psi'(x)f^*[\theta_2^{-1}\{\psi(x)-\theta_1\}] \ , \quad x \in \mathcal{X} \ , \tag{3.3}$$

where $\theta = (\theta_1, \theta_2)'$, $-\infty < \theta_1 < \infty$ and $\theta_2 > 0$. In the above, $\mathcal{X}(\subset \mathcal{R}^1)$ is an open interval which does not depend on θ and $\psi(x)$ is a strictly increasing smooth function which tends to $-\infty$ and $+\infty$ respectively as x tends to the lower and upper boundaries of $\mathcal{X}$. The function f^* is everywhere positive and satisfies $\int_{-\infty}^{\infty} f^*(u)du = 1$. Since ψ is strictly increasing, so is the inverse function ψ^{-1}. Hence denoting the ith derivative function of ψ^{-1} by ψ_i $(i = 1, 2)$, the function ψ_1 is always positive.

The location-scale type family (3.3) includes many commonly arising models. For example, if $\mathcal{X} = \mathcal{R}^1$ and $\psi(x) = x$ then it reduces to the usual location-scale model and covers, in particular, the univariate normal or Cauchy models. The family (3.3) also includes the log-normal and Weibull models. This can be seen taking $\mathcal{X} = (0,\infty)$, $\psi(x) = \log x$ and

$$f^*(u) = \phi(u) \quad \text{and} \quad f^*(u) = \exp(u - e^u)$$

respectively.

Under (3.3), $g_\beta(\theta) = \psi^{-1}(\theta_1 + \theta_2 q_\beta)$ where the constant q_β satisfies

$$\int_{-\infty}^{q_\beta} f^*(u)du = 1 - \beta \ .$$

Taking $g(\theta) = g_\beta(\theta)$ in the setup of Subsection 2.3,

$$g_1(\theta) = \psi_1 \ , \quad g_2(\theta) = q_\beta\psi_1 \ , \tag{3.4}$$

$$g_{11}(\theta) = \psi_2 \ , \quad g_{12}(\theta) = q_\beta\psi_2 \ , \quad g_{22}(\theta) = q_\beta^2\psi_2 \tag{3.5}$$

where $\psi_i = \psi_i(\theta_1+\theta_2 q_\beta)$ $(i = 1, 2)$. By (3.4), the gradient vector of $g_\beta(\theta)$ is always nonnull as the function ψ_1 is always positive. Under (3.3), $I^{-1} = ((k^{jr}\theta_2^2))$ where the k^{jr} are constants free from θ. Hence by (3.4), continuing with the notation of Subsection 2.3,

$$\lambda = k\theta_2\psi_1 \ , \quad \mu_j = (k^{j1} + k^{j2}q_\beta)\theta_2^2\psi_1 \quad (j = 1, 2) \ , \tag{3.6}$$

where $k = (k^{11} + 2q_\beta k^{12} + q_\beta^2 k^{22})^{1/2}$ is a positive constant free from θ. Using the positive definiteness of I^{-1} and hence that of $((k^{jr}))$, from (2.16) and (3.6) one can check that $\Delta_1(\pi, \theta)$

equals 0 for every β if and only if

$$D_1\{\theta_2\pi(\theta)\} = 0 \quad \text{and} \quad D_2\{\theta_2\pi(\theta)\} = 0 \ ,$$

that is, if and only if

$$\pi(\theta) \propto \theta_2^{-1} \ . \tag{3.7}$$

We next examine if the above prior satisfies $\Delta_2(\pi,\theta) = 0$ as well for every β. Under (3.3), L_{jrs} is proportional to θ_2^{-3} for each j, r, s. Hence by (3.6), L_{000}/λ^3 is a constant free from θ, while L_{j00}/λ^2 is proportional to θ_2^{-1} for each j. Therefore, by (3.6) and (3.7),

$$D_r\{\lambda^{-4}\mu_r L_{000}\pi(\theta)\} = 0 \ , \quad D_r\{\lambda^{-2}I^{jr}L_{j00}\pi(\theta)\} = 0 \ . \tag{3.8}$$

$$D_j D_r\{\lambda^{-2}\mu_j\mu_r\pi(\theta)\} = 0 \ . \tag{3.9}$$

Also by (3.5) and (3.6), after a little algebra,

$$\lambda^{-2}g_{00} = k^2\theta_2^2\psi_2 \ , \quad \lambda^{-1}g_{10} = k\theta_2\psi_2 \ , \quad \lambda^{-1}g_{20} = kq_\beta\theta_2\psi_2 \ ,$$

where g_{00}, g_{10} and g_{20} are as defined in Subsection 2.3. Hence by (3.6),

$$\lambda^{-4}\mu_r g_{00} - \lambda^{-2}I^{jr}g_{j0} = 0 \quad (r = 1, 2) \ . \tag{3.10}$$

From (2.17) and (3.8)-(3.10), it follows that the prior in (3.7) satisfies $\Delta_2(\pi,\theta) = 0$ for every β. Hence this is the unique second order probability matching prior for tolerance limits under the location-scale type family of models (3.3).

4 Further examples

We continue with Bayesian tolerance intervals of the form $[T, \infty)$ and present two examples illustrating the determination of T, via (2.9)-(2.11) as indicated in Section 3, so as to attain the target coverage probability with margin of error $o(n^{-1})$. Following Guttman (1970, Ch.9), we deliberately choose examples where exact computation of probability is possible. This allows us to examine how well the exact coverage probability of $[T, \infty)$ agrees with the target for small to moderate values of the sample size n. It is of interest to explore the gain in computational accuracy through the present asymptotics which have margin of error

$o(n^{-1})$ rather than $o(n^{-1/2})$. For this purpose, in the examples we also compare the exact coverage of $[T, \infty)$ with that of a tolerance interval $[T^*, \infty)$, where T^* is obtained taking $g(\theta) = g_\beta(\theta)$ in (2.13) which corresponds to asymptotics with margin of error $o(n^{-1/2})$.

Example 2. Let $f(x; \theta)$ represent the univariate normal model with mean θ_1 and standard deviation θ_2, where $-\infty < \theta_1 < \infty$ and $\theta_2 > 0$. Then $g_\beta(\theta) = \theta_1 + z_\beta \theta_2$. We work with the prior $\pi(\theta) \propto \theta_2^{-1}$ which, as noted in Example 1, is second order probability matching for tolerance limits.

Here $\hat{\theta} = (\bar{X}, s)'$, where $\bar{X} = n^{-1} \sum_{i=1}^n X_i$ and $s = \{n^{-1} \sum_{i=1}^n (X_i - \bar{X})^2\}^{1/2}$ are the sample mean and the sample standard deviation respectively. Then, following the notation introduced in Section 2 with $g(\theta) = g_\beta(\theta)$ and using (2.3)-(2.6),

$$b = A_\beta s \ , \quad b^{-i} G_i(\pi) = M_i \ (i = 1, 2, 4) \ , \quad b^{-i} G_i = M_i \ (i = 3, 6) \ ,$$

where $A_\beta = (1 + \frac{1}{2} z_\beta^2)^{1/2}$, and

$$M_1 = \frac{5}{4} A_\beta^{-1} z_\beta \ , \quad M_2 = A_\beta^{-2} \Big(\frac{3}{2} + \frac{71}{32} z_\beta^2 \Big) \ , \quad M_3 = \frac{1}{6} A_\beta^{-3} z_\beta \Big(3 + \frac{5}{4} z_\beta^2 \Big) \ ,$$
$$M_4 = A_\beta^{-4} \Big(\frac{1}{4} + \frac{11}{8} z_\beta^2 + \frac{49}{96} z_\beta^4 \Big) \ , \quad M_6 = \frac{1}{2} M_3^2 \ .$$

From (2.9)-(2.11), an explicit expression for T now follows as

$$T = \bar{X} + Bs \ , \tag{4.1}$$

where

$$\begin{aligned} B = B^* \ &+ \ n^{-3/2} A_\beta \Big[M_2 J_1(z_\alpha) + M_4 J_3(z_\alpha) + M_6 J_5(z_\alpha) \\ &+ \ 2 M_3 z_\alpha \{ M_1 + M_3 J_2(z_\alpha) \} - \frac{1}{2} z_\alpha \{ M_1 + M_3 J_2(z_\alpha) \}^2 \Big] \ , \end{aligned}$$

with

$$B^* = z_\beta + n^{-1/2} A_\beta z_\alpha + n^{-1} A_\beta \{ M_1 + M_3 J_2(z_\alpha) \} \ ,$$

the M_i being as shown above. Similarly, by (2.13),

$$T^* = \bar{X} + B^* s \ . \tag{4.2}$$

Exact computation is possible in this example. With $\pi(\theta) \propto \theta_2^{-1}$, one can check that for any constant k,

$$P^{\pi}\{1 - F(\bar{X} + ks; \theta) \geq \beta | X\} = P\{t' \leq -k(n-1)^{1/2}\} \ , \tag{4.3}$$

where t' is a noncentral t-variate with degrees of freedom $n-1$ and noncentrality parameter $n^{1/2} z_{1-\beta}$. Let ρ and ρ^* denote the values of the right-hand side of (4.3) for $k = B$ and $k = B^*$ respectively. Then by (4.1) and (4.2), ρ and ρ^* represent the exact posterior coverage probabilities of the Bayesian tolerance intervals $[T, \infty)$ and $[T^*, \infty)$ respectively. In fact, in this example the prior $\pi(\theta) \propto \theta_2^{-1}$ is exact probability matching in the sense that

$$P^{\pi}\{1 - F(\bar{X} + ks; \theta) \geq \beta | X\} = P_{\theta}\{1 - F(\bar{X} + ks; \theta) \geq \beta\}$$

for any constant k and, as such, ρ and ρ^* represent the corresponding exact frequentist coverage probabilities as well.

Table 1 shows the values of ρ and ρ^* for $\beta = 0.90$, $\alpha = 0.90$, 0.95, and $n = 5(3)20$. The tables of the noncentral t-distribution given in Resnikoff and Lieberman (1957) help in the construction of this table. It is satisfying to note from Table 1 that ρ is quite close to the target α for $n \geq 8$. Furthermore, the difference $|\rho - \alpha|$ is always about half to one-thirds of the difference $|\rho^* - \alpha|$. Hence use of asymptotics with margin of error $o(n^{-1})$ as against $o(n^{-1/2})$ entails a significant gain in computational accuracy.

Table 1. Exact coverage probabilities of Bayesian tolerance intervals under the normal model ($\beta = 0.90$)

α	n	5	8	11	14	17	20
0.90	ρ	0.854	0.877	0.886	0.890	0.893	0.894
	ρ^*	0.811	0.848	0.863	0.872	0.877	0.880
0.95	ρ	0.909	0.930	0.938	0.942	0.943	0.945
	ρ^*	0.869	0.904	0.918	0.925	0.930	0.933

Example 3. Consider the exponential model given by $f(x; \theta) = \theta^{-1} \exp(-x/\theta)$, $0 < x < \infty$, where $\theta > 0$. Then $g_{\beta}(\theta) = (-\log \beta)\theta$. We work with the prior $\pi(\theta) \propto \theta^{-1}$ which can be

seen to be second order probability matching for tolerance limits. Here $\hat{\theta} = \bar{X} = n^{-1}\sum_{i=1}^{n} X_i$ and, as in Example 2, we get

$$b = (-\log\beta)\bar{X}\ , \quad b^{-1}G_1(\pi) = 1\ , \quad b^{-2}G_2(\pi) = \frac{5}{2}\ ,$$
$$b^{-3}G_3 = \frac{2}{3}\ , \quad b^{-4}G_4(\pi) = \frac{23}{12}\ , \quad b^{-6}G_6 = \frac{2}{9}\ .$$

Hence by (2.9)-(2.11),

$$T = (-\log\beta)B\bar{X}\ , \tag{4.4}$$

where

$$B = B^* + \frac{1}{36}n^{-3/2}z_\alpha(17 + 13z_\alpha^2)\ ,$$

with $B^* = 1 + n^{-1/2}z_\alpha + \frac{1}{3}n^{-1}(2z_\alpha^2 + 1)$. Similarly, by (2.13),

$$T^* = (-\log\beta)B^*\bar{X}\ , \tag{4.5}$$

B^* being as just stated.

This example also allows exact computation. With $\pi(\theta) \propto \theta^{-1}$, for any constant k, one has

$$P^\pi\{1 - F(k\bar{X};\theta) \geq \beta | X\} = 1 - \sum_{j=0}^{n-1} e^{-\tau}\frac{\tau^j}{j!}\ , \tag{4.6}$$

where $\tau \equiv \tau(k) = n(-\log\beta)/k$. Let ρ and ρ^* be the values of the right-hand side of (4.6) for $k = (-\log\beta)B$ and $k = (-\log\beta)B^*$ respectively. Then by (4.4) and (4.5), ρ and ρ^* represent the exact posterior coverage probabilities of the Bayesian tolerance intervals $[T,\infty)$ and $[T^*,\infty)$ respectively. The prior $\pi(\theta) \propto \theta^{-1}$ is exact probability matching in this example and, therefore, ρ and ρ^* also represent the corresponding exact frequentist coverage probabilities.

From (4.4)-(4.6) observe that both ρ and ρ^* depend only on α and not on β. Table 2 shows the values of ρ and ρ^* for $\alpha = 0.90$, 0.95 and $n = 5(3)20$. The same conclusions as in Example 2 hold. The gain in computational accuracy through the use of asymptotics with margin of error $o(n^{-1})$ rather than $o(n^{-1/2})$ is now even more pronounced.

Table 2. Exact coverage probabilities of Bayesian tolerance intervals under the exponential model

α	n	5	8	11	14	17	20
0.90	ρ	0.924	0.912	0.907	0.905	0.904	0.903
	ρ^*	0.829	0.859	0.871	0.878	0.882	0.885
0.95	ρ	0.978	0.964	0.959	0.956	0.955	0.954
	ρ^*	0.847	0.895	0.914	0.923	0.929	0.933

Acknowledgement

This work was supported by research grants from the Natural Sciences and Engineering Research Council of Canada and the Centre for Management and Development Studies, Indian Institute of Management Calcutta.

References

Bickel, P.J. and Ghosh, J.K. (1990). A decomposition for the likelihood ratio statistic and the Bartlett correction – a Bayesian argument. *Ann. Statist.* **18**, 1070-1090.

Chandra, T.K. and Ghosh, J.K. (1979). Valid asymptotic expansions for the likelihood ratio statistic and other perturbed chi-square variables. *Sankhyā A* **41**, 22-47.

Datta, G.S. and Ghosh, J.K. (1995). On priors providing frequentist validity for Bayesian inference. *Biometrika* **82**, 37-45.

Ghosh, J.K. and Mukerjee, R. (1991). Characterization of priors under which Bayesian and frequentist Bartlett corrections are equivalent in the multiparameter case. *J. Mult. Anal.* **38**, 385-393.

Guttman, I. (1970). *Statistical Tolerance Regions: Classical and Bayesian.* Charles Griffin, London.

Johnson, R.A. (1970). Asymptotic expansions associated with posterior distributions. *Ann. Math. Statist.* **41**, 851-864.

Mukerjee, R. and Ghosh, M. (1997). Second order probability matching priors. *Biometrika* **84**, 970-975.

Mukerjee, R. and Reid, N. (1999). On a property of probability matching priors: matching the alternative coverage probabilities. *Biometrika* **86**, 333-340.

Resnikoff, G.J. and Lieberman, G.J. (1957). *Tables of the Non-central t-Distribution.* Stanford University Press, Stanford, CA.

Sun, D. and Ye, K. (1996). Frequentist validity of posterior quantiles for a two parameter exponential family. *Biometrika* **83**, 55-65.

Tibshirani, R. (1989). Noninformative priors for one parameter of many. *Biometrika* **76**, 604-608.