J. R. Statist. Soc. B (1995)
57, *No.* 4, *pp.* 761–769

On Perturbed Ellipsoidal and Highest Posterior Density Regions with Approximate Frequentist Validity

By J. K. GHOSH†

Purdue University, West Lafayette, USA, and Indian Statistical Institute, Calcutta, India

and

RAHUL MUKERJEE

Indian Institute of Management, Calcutta, India

[Received November 1993. Revised November 1994]

SUMMARY

This paper considers, in the multiparameter case, perturbed ellipsoidal and highest posterior density regions with both Bayesian and frequentist validity up to $o(n^{-1})$.

Keywords: ELLIPSOIDAL REGION; HIGHEST POSTERIOR DENSITY REGION; LOCATION MODEL; LOCATION-SCALE MODEL; MINIMAL PERTURBATION

1. INTRODUCTION AND PRELIMINARIES

In recent years there has been a revival of interest in the characterization of priors ensuring approximate frequentist validity of posterior credible sets—see Tibshirani (1989), Lee (1989), Severini (1991), Ghosh and Mukerjee (1992, 1993) and the references therein. A related problem of finding, for a given prior, Bayesian credible sets with both Bayesian and frequentist validity up to $o(n^{-1})$, where n is the sample size, has recently been considered by Severini (1993) in the one-parameter case. As he discussed, this problem can be of interest if one believes that both the Bayesian and the frequentist points of view are important. Here we consider the same problem in the multiparameter case and give two sets of solutions based on perturbed ellipsoidal and highest posterior density (HPD) regions. Our method of solution, however, is different from that in Severini (1993). In particular, in the multiparameter case, an approach based on inversion of approximate posterior characteristic functions is seen to be helpful; see, for example, Ghosh and Mukerjee (1993). Also, unlike Severini (1993), who considered conditional frequentist validity, we do not require the specification of an ancillary or an approximately ancillary statistic.

Let $\{X_i\}$, $i \geqslant 1$, be a sequence of independent and identically distributed possibly vector-valued random variables each with density $f(x;\ \theta)$ where $\theta = (\theta_1, \ldots, \theta_p)' \in \Theta$, an open subset of $\mathscr{R}^p$. We make the assumptions in Johnson (1970), section 2, with $K = 2$ in his notation. Let θ have a prior density $\pi(\ )$ which is positive and thrice continuously differentiable at all θ. If $\pi(\ )$ is not proper, we shall require that there is an n_0 (>0) such that, for all $X_1, \ldots, X_{n_0}$, the posterior of θ given $X_1, \ldots, X_{n_0}$ is proper. Let $X = (X_1, \ldots, X_n)'$, where n is the sample size,

†*Address for correspondence*: Indian Statistical Institute, 203 Barrackpore Trunk Road, Calcutta 700 035, India.
E-mail: jkg@isical.ernet.in

 0035–9246/95/57761

$$l(\theta) = n^{-1} \sum_{i=1}^{n} \log f(X_i;\ \theta)$$

and $\hat{\theta}$ be the maximum likelihood estimator of θ based on X. Define $\hat{\pi} = \pi(\hat{\theta})$ and, for $1 \leqslant i, j, r, s \leqslant p$, let

$$\pi_i(\theta) = D_i\,\pi(\theta), \qquad \pi_{ij}(\theta) = D_i D_j\,\pi(\theta), \qquad \hat{\pi}_i = \pi_i(\hat{\theta}), \qquad \hat{\pi}_{ij} = \pi_{ij}(\hat{\theta}),$$

$$a_{ij} = \{D_i D_j\, l(\theta)\}_{\theta=\hat{\theta}}, \qquad a_{ijr} = \{D_i D_j D_r\, l(\theta)\}_{\theta=\hat{\theta}}, \qquad a_{ijrs} = \{D_i D_j D_r D_s\, l(\theta)\}_{\theta=\hat{\theta}},$$

$$c_{ij} = -a_{ij}, \qquad V_i = D_i \log f(X_1;\ \theta), \qquad V_{ij} = D_i D_j \log f(X_1;\ \theta),$$

$$V_{ijr} = D_i D_j D_r \log f(X_1;\ \theta),$$

$$I_{ij} = E_\theta(V_i V_j), \qquad L_{ij,r} = E_\theta(V_{ij} V_r), \qquad L_{ijr} = E_\theta(V_{ijr}),$$

where $D_i \equiv \partial/\partial\theta_i$. Note that I_{ij}, $L_{ij,r}$ and L_{ijr} are functions of θ and that the per observation information matrix at θ is given by $I \equiv I(\theta) = (I_{ij})$ which is assumed to be positive definite at each θ. All formal expansions for the posterior, as used here, are valid for sample points in a set S, which may be defined along the lines of Bickel and Ghosh (1990), with P_θ-probability $1 + o(n^{-1})$ uniformly over compact sets of θ. The $p \times p$ matrix $C = (c_{ij})$ is positive definite over S. Let $C^{-1} = (c^{ij})$ and $I^{-1} = (I^{ij})$.

Throughout, unless otherwise stated, the summation convention will be followed, i.e. summation will be implied over repeated subscripts or superscripts. For example, $a_{ijr}h_j h_r$ and $c^{ij}\hat{\pi}_{ij}$ will stand for $\Sigma_j\, \Sigma_r\, a_{ijr}h_j h_r$ and $\Sigma_i\, \Sigma_j c^{ij}\hat{\pi}_{ij}$ respectively. For subsequent use, we note from Ghosh and Mukerjee (1993) that the posterior density of $h(\theta) \equiv h = (h_1, \ldots, h_p)' = n^{1/2}(\theta - \hat{\theta})$ under the prior $\pi(\ )$ is given by

$$\tilde{\pi}(h|X) = \phi(h;\ C^{-1})\Bigg(1 + n^{-1/2}\{T_{11}(\pi, h) + \frac{1}{6}\, T_{12}(h)\} + n^{-1}\bigg[\frac{1}{2}\{T_{21}(\pi, h) - G_1(\pi)\} + \frac{1}{24}\{T_{22}(h) - G_2\} + \frac{1}{6}\{T_{11}(\pi, h)\, T_{12}(h) - G_3(\pi)\} + \frac{1}{72}\{T_{12}^2(h) - G_4\}\bigg]\Bigg) + o(n^{-1}), \tag{1.1a}$$

where $\phi(\ ;\ C^{-1})$ is the p-variate normal density with null mean vector and dispersion matrix C^{-1}, and, with $c^{(1)}_{ijrs} = c^{ij}c^{rs} + c^{ir}c^{js} + c^{is}c^{jr}$,

$$T_{11}(\pi, h) = \hat{\pi}^{-1} h_i \hat{\pi}_i, \qquad T_{12}(h) = a_{ijr}h_i h_j h_r, \tag{1.1b}$$

$$T_{21}(\pi, h) = \hat{\pi}^{-1} h_i h_j \hat{\pi}_{ij}, \qquad T_{22}(h) = a_{ijrs}h_i h_j h_r h_s, \tag{1.1c}$$

$$\left.\begin{aligned} G_1(\pi) &= \hat{\pi}^{-1} c^{ij}\hat{\pi}_{ij}, \\ G_2 &= a_{ijrs}c^{(1)}_{ijrs}, \\ G_3(\pi) &= \hat{\pi}^{-1} a_{ijr}\hat{\pi}_s c^{(1)}_{ijrs}, \\ G_4 &= a_{ijr}a_{suv}(9c^{ij}c^{rs}c^{uv} + 6c^{is}c^{ju}c^{rv}), \end{aligned}\right\} \tag{1.1d}$$

each of the implicit summations being over the range from 1 to p. In what follows, for positive integral ν, $K_\nu(\ )$ and $k_\nu(\ )$ denote respectively the cumulative distribution function and the probability density function of a central χ^2-variate with ν degrees of freedom. Also, z^2 denotes the upper α-point of a central χ^2-variate with p degrees of freedom.

2. PERTURBED ELLIPSOIDAL AND HIGHEST POSTERIOR DENSITY REGIONS

By equations (1.1a), up to the first order of approximation, $h \equiv h(\theta)$ has a null mean vector and a dispersion matrix C^{-1} in the posterior set-up. This motivates us to consider a perturbed ellipsoidal region for θ of the form

$$R_1(\alpha, \pi, X) = \{\theta\colon (h(\theta) - n^{-1/2}d_\pi(\hat\theta))'C(h(\theta) - n^{-1/2}d_\pi(\hat\theta)) \leqslant \lambda_{1n}(\alpha, \pi, X)\}, \tag{2.1}$$

where $d_\pi(\theta) = (d_{1\pi}(\theta), \ldots, d_{p\pi}(\theta))'$ and $\lambda_{1n}(\alpha, \pi, X)$ are to be so chosen that the region has both posterior and frequentist coverage probability $1 - \alpha + o(n^{-1})$, and $0 < \alpha < 1$. It should be made explicit here that, for each j, $d_{j\pi}(\theta)$ is a smooth function with a functional form possibly dependent on $\pi(\ )$ and α but not on n. Let

$$\left.\begin{aligned} F_1(\pi) &= \frac{1}{2}G_1(\pi) + \frac{1}{2}(d_\pi(\hat\theta))'C(d_\pi(\hat\theta)) - \hat\pi^{-1}d_{r\pi}(\hat\theta)\hat\pi_r + \frac{1}{2}d_{r\pi}(\hat\theta)a_{ijr}c^{ij}, \\ F_2(\pi) &= \frac{1}{24}G_2 + \frac{1}{6}G_3(\pi) - \frac{1}{2}d_{r\pi}(\hat\theta)a_{ijr}c^{ij}, \\ F_3 &= \frac{1}{72}G_4, \\ F_0(\pi) &= -F_1(\pi) - F_2(\pi) - F_3. \end{aligned}\right\} \tag{2.2}$$

Then, as discussed in Appendix A, with

$$\lambda_{1n}(\alpha, \pi, X) = z^2 - \{n\,k_p(z^2)\}^{-1}\left\{\sum_{j=0}^{2} F_j(\pi)K_{p+2j}(z^2) + F_3K_{p+6}(z^2)\right\}, \tag{2.3}$$

the relationships

$$P^\pi\{\theta \in R_1(\alpha, \pi, X)|X\} = 1 - \alpha + o(n^{-1}) \tag{2.4}$$

and

$$P_\theta\{\theta \in R_1(\alpha, \pi, X)\} = 1 - \alpha + 2\{n\,\pi(\theta)\}^{-1}k_{p+2}(z^2)\,\Delta_1\{\alpha, \pi, d_\pi(\theta), \theta\} + o(n^{-1}) \tag{2.5}$$

hold, where $P^\pi\{\ |X\}$ is the posterior probability measure for θ under $\pi(\ )$ and

$$\begin{aligned} \Delta_1\{\alpha, \pi, d_\pi(\theta), \theta\} = D_r[\tfrac{1}{2}\{1 + (p+2)^{-1}z^2\}I^{ij}I^{rs}L_{ijs}\,\pi(\theta) + \tfrac{1}{2}D_iI^{ir}\,\pi(\theta) \\ - \pi(\theta)D_iI^{ir} - d_{r\pi}(\theta)\,\pi(\theta)]. \end{aligned} \tag{2.6}$$

By equations (2.4) and (2.5), the perturbed ellipsoidal region (2.1) will have both frequentist and posterior coverage probability $1 - \alpha + o(n^{-1})$ provided that $d_\pi(\theta)$ satisfies the partial differential equation

$$\Delta_1\{\alpha,\ \pi,\ d_\pi(\theta),\ \theta\} = 0, \tag{2.7}$$

and $\lambda_{1n}(\alpha, \pi, X)$ is chosen as in equation (2.3). In particular, by equation (2.6), $d_\pi(\theta) = \bar{d}_\pi(\theta)$, where $\bar{d}_\pi(\theta) = (\bar{d}_{1\pi}(\theta), \ldots, \bar{d}_{p\pi}(\theta))'$ and

$$\bar{d}_{r\pi}(\theta) = \tfrac{1}{2}\{1 + (p+2)^{-1}z^2\}I^{ij}I^{rs}L_{ijs} + \tfrac{1}{2}\pi(\theta)^{-1}D_i I^{ir}\pi(\theta) - D_i I^{ir}, \qquad 1 \leqslant r \leqslant p, \tag{2.8}$$

satisfies condition (2.7).

Considering now the HPD region, as noted in Ghosh and Mukerjee (1993), up to $o(n^{-1})$, this is approximable as $R_2(\alpha, \pi, X) = \{\theta\colon W\{\pi, X, h(\theta)\} \leqslant \lambda_{2n}(\alpha, \pi, X)\}$, where, with $h \equiv h(\theta)$,

$$W\{\pi, X, h(\theta)\} = h'Ch - n^{-1/2}\left\{2\,T_{11}(\pi, h) + \frac{1}{3}\,T_{12}(h)\right\} + n^{-1}\left\{\hat{\pi}^{-2}\hat{\pi}_i\hat{\pi}_j c^{ij} - T_{21}(\pi, h) - \frac{1}{12}\,T_{22}(h) + T_{11}^2(\pi, h)\right\}, \tag{2.9}$$

and $\lambda_{2n}(\alpha, \pi, X)$ is such that $R_2(\alpha, \pi, X)$ has posterior coverage probability $1 - \alpha + o(n^{-1})$. In the present context, the above motivates us to consider a perturbed HPD region of the form

$$R_2^*(\alpha, \pi, X) = \{\theta\colon W^*\{\pi, X, h(\theta)\} \leqslant \lambda_{2n}^*(\alpha, \pi, X)\}, \tag{2.10}$$

where

$$W^*\{\pi, X, h(\theta)\} = W\{\pi, X, h(\theta)\} - h'Ch + (h - n^{-1/2}b_\pi(\hat{\theta}))'C(h - n^{-1/2}b_\pi(\hat{\theta})) \tag{2.11}$$

is obtained by perturbing the leading term in equation (2.9) in a manner similar to equation (2.1). Here $b_\pi(\theta) = (b_{1\pi}(\theta), \ldots, b_{p\pi}(\theta))'$ and, for each j, $b_{j\pi}(\theta)$ is a smooth function with a functional form that is possibly dependent on $\pi(\ )$ and α but not on n; actually, as we shall see later, the appropriate choice of $b_\pi(\theta)$ does not depend even on α.

Let

$$F(\pi) = \frac{1}{24}G_2 + \frac{1}{72}G_4 + \frac{1}{2}(b_\pi(\hat{\theta}))'C(b_\pi(\hat{\theta})) - \frac{1}{2}b_{r\pi}(\hat{\theta})a_{ijr}c^{ij}.$$

Then with

$$\lambda_{2n}^*(\alpha, \pi, X) = z^2 - \{n\,k_p(z^2)\}^{-1}F(\pi)\{K_{p+2}(z^2) - K_p(z^2)\} - 2n^{-1}\,b_{r\pi}(\hat{\theta})\hat{\pi}_r/\hat{\pi}, \tag{2.12}$$

as indicated in Appendix A, we obtain

$$P^\pi\{\theta \in R_2^*(\alpha, \pi, X)\,|\,X\} = 1 - \alpha + o(n^{-1}), \tag{2.13}$$

$$P_\theta\{\theta \in R_2^*(\alpha, \pi, X)\} = 1 - \alpha - 2\{np\,\pi(\theta)\}^{-1}z^2\,k_p(z^2)\,\Delta_2\{\pi,\ b_\pi(\theta),\ \theta\} + o(n^{-1}), \tag{2.14}$$

where

$$\Delta_2\{\pi,\ b_\pi(\theta),\ \theta\} = D_r\left\{b_{r\pi}(\theta)\,\pi(\theta) + \frac{1}{2}\,I^{ir}D_i\,\pi(\theta) + \frac{1}{2}\,I^{ir}I^{js}L_{is,j}\,\pi(\theta)\right\}. \tag{2.15}$$

By equations (2.13) and (2.14), the perturbed HPD region (2.10) will have both frequentist and posterior coverage probability $1 - \alpha + o(n^{-1})$ provided that $b_\pi(\theta)$ satisfies

$$\Delta_2\{\pi, b_\pi(\theta), \theta\} = 0, \tag{2.16}$$

and $\lambda_{2n}^*(\alpha, \pi, X)$ is chosen as in equation (2.12). Note that the partial differential equation (2.16) does not involve α. In particular, by equation (2.15), $b_\pi(\theta) = \bar{b}_\pi(\theta)$, where

$$\bar{b}_\pi(\theta) = (\bar{b}_{1\pi}(\theta), \ldots, \bar{b}_{p\pi}(\theta))'$$

and

$$\bar{b}_{r\pi}(\theta) = -\frac{1}{2} I^{ir} I^{js} L_{is,j} - \frac{1}{2} \pi(\theta)^{-1} I^{ir} D_i \pi(\theta), \qquad 1 \leqslant r \leqslant p, \tag{2.17}$$

satisfies condition (2.16).

Remark 1. The margins of error in equations (2.4), (2.13) and also in the approximations for the posterior characteristic functions used in Appendix A are at most of the order $O(n^{-3/2})$; see theorem 2.1 in Johnson (1970). The same holds for the frequentist approximations (2.5) and (2.14) as well under appropriate Edgeworth assumptions (see Bhattacharya and Ghosh (1978)). In fact, if we work under the assumptions of Johnson (1970) (with $K = 3$ in his notation) together with suitable Edgeworth assumptions, then it should be possible to show that these errors are of the order $O(n^{-2})$; see Barndorff-Nielsen and Hall (1988).

Remark 2. The solutions for $d_\pi(\theta)$ and $b_\pi(\theta)$, as shown in equations (2.8) and (2.17) respectively, can be interpreted in terms of the first-order biases of estimators given by

(a) the maximum likelihood estimator,
(b) the posterior mean of θ under $\pi(\ )$ and
(c) the posterior mode of θ under $\pi(\ )$,

Denoting the first-order biases of these estimators by $n^{-1}\beta_i(\theta)$, $i = 1, 2, 3$ respectively, these solutions can be expressed as

$$\begin{aligned} \bar{d}_\pi(\theta) &= \tfrac{1}{2}\beta_2(\theta) - \beta_1(\theta) + \{z^2/(p+2)\}\{\beta_2(\theta) - \beta_3(\theta)\}, \\ \bar{b}_\pi(\theta) &= \tfrac{1}{2}\beta_2(\theta) - \beta_3(\theta). \end{aligned} \tag{2.18}$$

Remark 3. To make a choice between rival solutions of equation (2.7) or (2.16), we propose a principle of minimal perturbation which seems to be sensible from a Bayesian point of view. This is discussed with reference to equation (2.16). Thus, given $\pi(\ )$, we should first check whether $b_\pi(\theta) = 0$ satisfies equation (2.16). If not, then using $I \equiv I(\theta)$ as a Riemannian metric a solution with a smaller value of

$$\int (b_\pi(\theta))' I(\theta) (b_\pi(\theta))\, \pi(\theta)\, \mathrm{d}\theta$$

will be preferred to another with a larger value of the same quantity. For $p > 1$, it is difficult to characterize all the solutions of equation (2.16) and we suggest the use of a solution which at least satisfies

$$\int (b_\pi(\theta))' I(\theta)(b_\pi(\theta))\, \pi(\theta)\, d\theta < \infty. \tag{2.19}$$

3. EXAMPLES AND DISCUSSION

3.1. *Example 1*

Consider the multiparameter location model with $f(x; \theta)$ of the form $f(x; \theta) = f^*(x^{(1)} - \theta_1, \ldots, x^{(p)} - \theta_p)$, where $\theta = (\theta_1, \ldots, \theta_p)' \in \mathscr{R}^p$ and $x = (x^{(1)}, \ldots, x^{(p)})'$. Here for $1 \leqslant i, j, s \leqslant p$, I_{ij}, I^{ij}, $L_{is,j}$ and L_{ijs} are all constants, independent of θ, provided that they exist. Hence, if $\pi(\theta) =$ constant, then $d_\pi(\theta) = 0$ and $b_\pi(\theta) = 0$ satisfy equations (2.7) and (2.16) respectively, i.e. no perturbation is required at all to achieve our aim with ellipsoidal and HPD regions. However, this does not happen under a p-variate normal prior but then, by equations (2.17) and (2.8), both the solutions $\bar{b}_\pi(\theta)$ and $\bar{d}_\pi(\theta)$ are linear in θ and satisfy respectively condition (2.19) and the analogous condition for $d_\pi(\theta)$. In fact, most of the location models arising in practice (e.g. the multivariate normal or Cauchy location models) are sufficiently symmetric to ensure that $L_{is,j} = L_{ijs} \equiv 0$, $1 \leqslant i, j, s \leqslant p$. For such models, equations (2.18) can be simplified further to $\bar{d}_\pi(\theta) = -\bar{b}_\pi(\theta) = \frac{1}{2}\beta_2(\theta) = \frac{1}{2}\beta_3(\theta)$ and, specifically, under a p-variate normal prior with mean vector μ and a positive definite dispersion matrix Ω, equations (2.8) and (2.17) yield $\bar{d}_\pi(\theta) = -\bar{b}_\pi(\theta) = -\frac{1}{2}(\Omega I)^{-1}(\theta - \mu)$.

3.2. *Example 2*

Consider the location–scale model with $f(x; \theta)$ of the form $f(x; \theta) = \theta_1^{-1} f^*\{(x - \theta_2)/\theta_1\}$, with $\theta_1 > 0$ and $\theta_2 \in \mathscr{R}^1$. Here $p = 2$ and, for each i, j and s, I_{ij} is proportional to θ_1^{-2} whereas $L_{is,j}$ and L_{ijs} are proportional to θ_1^{-3} provided that they exist. Hence the solutions shown in equations (2.8) and (2.17) are of the forms

$$\bar{d}_{r\pi}(\theta) = \tau_{1r}\theta_1 + \frac{1}{2}\pi(\theta)^{-1}\theta_1^2 g^{ir} D_i \pi(\theta),$$

$$\bar{b}_{r\pi}(\theta) = \tau_{2r}\theta_1 - \frac{1}{2}\pi(\theta)^{-1}\theta_1^2 g^{ir} D_i \pi(\theta),$$

where $I^{-1} = (\theta_1^2 g^{ir})$ and g^{ir}, τ_{1r} and τ_{2r} are constants ($i, r = 1, 2$). These solutions satisfy condition (2.19) and the analogous condition for $d_\pi(\theta)$ under commonly used priors like that given by the product of a gamma density in θ_1 and a univariate normal density in θ_2.

Combining our techniques with those in Mukerjee and Dey (1993), the present results can, in principle, be extended to models involving nuisance parameters. This is because then the marginal posterior density of $h^{(1)} \equiv h^{(1)}(\theta^{(1)}) = n^{1/2}(\theta^{(1)} - \hat{\theta}^{(1)})$, where $\theta^{(1)} = (\theta_1, \ldots, \theta_q)'$ is the parameter of interest ($1 \leqslant q < p$) and $\hat{\theta}^{(1)}$ is the maximum likelihood estimator of $\theta^{(1)}$, is expressible in a form similar to equation (1.1a). For example, analogously to equation (2.1), we may consider a perturbed ellipsoidal region

$$\{\theta^{(1)}: (h^{(1)}(\theta^{(1)}) - n^{-1/2}d_\pi(\hat{\theta}))'(C^{11})^{-1}(h^{(1)}(\theta^{(1)}) - n^{-1/2}d_\pi(\hat{\theta})) \leqslant \lambda_n(\alpha, \pi, X)\},$$

where C^{11} is the principal submatrix of C^{-1} given by its first q rows and columns and $d_\pi(\theta) = (d_{1\pi}(\theta), \ldots, d_{q\pi}(\theta))'$, the functional form of $d_{j\pi}(\theta)$ being independent of n ($1 \leqslant j \leqslant q$). After considerable algebra, it can be shown that to meet the twin objectives regarding correct posterior and frequentist coverage, up to $o(n^{-1})$, $d_\pi(\theta)$ must satisfy the partial differential equation

$$\begin{aligned} &I^{ij}\{D_iD_j\pi(\theta) + 2e_{vj}D_iD_v\pi(\theta) + e_{vi}e_{wj}D_vD_w\pi(\theta)\} - \pi(\theta)(D_iD_jI^{ij} + 2D_iD_ve_{vj}I^{ij} \\ &\quad + D_vD_we_{vi}e_{wj}I^{ij}) + D_iI^{ij}M_j\pi(\theta) + D_vI^{ij}M_je_{vi}\pi(\theta) + D_vI^{ij}Q_{ijv}\pi(\theta) \\ &\quad + \tfrac{1}{3}\{1 + (q+2)^{-1}\omega^2\}\{D_iI^{(1)}_{ijrs}B_{sjr}\pi(\theta) + D_vI^{(1)}_{ijrs}B_{sjr}e_{vi}\pi(\theta)\} \\ &\quad - 2\{D_id_{i\pi}(\theta)\pi(\theta) + D_vd_{i\pi}(\theta)e_{vi}\pi(\theta)\} = 0, \end{aligned} \tag{3.1}$$

where ω^2 is the upper α-point of a central χ^2-variate with q degrees of freedom,

$$I^{(1)}_{ijrs} = I^{ij}I^{rs} + I^{ir}I^{js} + I^{is}I^{jr},$$

$$M_j = \sigma_{wv}(L_{jwv} + L_{wvv'}e_{v'j}),$$

$$Q_{ijv} = \sigma_{wv}(L_{ijw} + 2L_{iwv'}e_{v'j} + L_{wv'v''}e_{v'i}e_{v''j}),$$

$$B_{sjr} = L_{sjr} + 3L_{sjw}e_{wr} + 3L_{swv}e_{wj}e_{vr} + L_{wvv'}e_{ws}e_{vj}e_{v'r}$$

and the implicit sums range over 1 to q for i, j, r and s and over $q+1$ to p for w, v, v' and v''. In the above, with the submatrix of $I \equiv I(\theta)$ given by its last $p-q$ rows partitioned as $(I_{(21)} \quad I_{(22)})$ where $I_{(22)}$ is square of order $p-q$, σ_{wv} is the (w, v)th element of $I^{-1}_{(22)}$ and e_{vi} is the (v, i)th element of $-I^{-1}_{(22)}I_{(21)}$. It is not difficult to see that equation (3.1) is in agreement with equations (2.6) and (2.7).

To illustrate an application of equation (3.1), we consider the univariate normal or Cauchy location-scale model where the scale parameter θ_1 is of interest and the location parameter θ_2 is the nuisance parameter. Then $q = 1$, $d_\pi(\theta)$ is a scalar and it can be shown that equation (3.1) admits a solution of the form $d_\pi(\theta) = y_1\theta_1 + y_2\theta_1^2\{\pi(\theta)\}^{-1}D_1\pi(\theta)$, where y_1 and y_2 are suitable constants.

ACKNOWLEDGEMENTS

We thank the referees for very constructive suggestions. The work of RM was supported by a grant from the Centre for Management and Development Studies, Indian Institute of Management, Calcutta.

APPENDIX A: PROOFS OF CERTAIN RELATIONSHIPS

A.1. *Proofs of Equations (2.4) and (2.13)*

Let

$$W^{**}\{\pi, X, h(\theta)\} = W^*\{\pi, X, h(\theta)\} + 2n^{-1}b_{r\pi}(\hat{\theta})\hat{\pi}_r/\hat{\pi}.$$

From equations (1.1) it can be shown that the approximate posterior characteristic functions of $(h - n^{-1/2}d_\pi(\hat{\theta}))'C(h - n^{-1/2}d_\pi(\hat{\theta}))$ and $W^{**}\{\pi, X, h(\theta)\}$, under $\pi(\;)$, are

$$(1-2\xi)^{-p/2}\left[1+n^{-1}\left\{\sum_{j=0}^{2}F_j(\pi)(1-2\xi)^{-j}+F_3(1-2\xi)^{-3}\right\}\right]+o(n^{-1}) \tag{A.1}$$

and

$$(1-2\xi)^{-p/2}[1+n^{-1}\{(1-2\xi)^{-1}-1\}F(\pi)]+o(n^{-1}) \tag{A.2}$$

respectively, where $\xi=(-1)^{1/2}t$, $F_j(\pi)$ ($j=0, 1, 2$) and F_3 are as in equations (2.2) and $F(\pi)$ is as in the context of equation (2.12); see Ghosh and Mukerjee (1993) for more details on similar results. Inversion of equations (A.1) and (A.2), which can be justified as in Chandra and Ghosh (1979), yields equations (2.4) and (2.13). Incidentally, equation (A.2) implies posterior Bartlett adjustability of $W^{**}\{\pi, X, h(\theta)\}$ (see DiCiccio and Stern (1993)) and, in a sense, explains why equation (2.16) does not involve α.

A.2. *Proof of Equation (2.5)*

Proceeding as in Ghosh and Mukerjee (1993), we take an auxiliary prior $\bar{\pi}(\ )$ satisfying the regularity conditions in Bickel and Ghosh (1990), section 2, with $m=3$, which are slightly stronger than those in Johnson (1970), and make Edgeworth assumptions as in Bickel and Ghosh (1990), p. 1078. Then, as in the derivation of equation (2.4), inverting the approximate posterior characteristic function of $(h-n^{-1/2}d_\pi(\hat{\theta}))'C(h-n^{-1/2}d_\pi(\hat{\theta}))$ under $\bar{\pi}(\ )$ and using equations (2.1)–(2.3),

$$P^{\bar{\pi}}\{\theta\in R_1(\alpha, \pi, X)|X\}=1-\alpha+n^{-1}\left[\frac{1}{6}\{K_p(z^2)-K_{p+4}(z^2)\}\{G_3(\pi)-G_3(\bar{\pi})\}\right.$$
$$+\{K_p(z^2)-K_{p+2}(z^2)\}\left\{\frac{1}{2}G_1(\pi)-\frac{1}{2}G_1(\bar{\pi})-d_{r\pi}(\hat{\theta})\left(\frac{\hat{\pi}_r}{\hat{\pi}}\right.\right.$$
$$\left.\left.\left.-\frac{\hat{\bar{\pi}}_r}{\hat{\bar{\pi}}}\right)\right\}\right]+o(n^{-1}), \tag{A.3}$$

where $\hat{\bar{\pi}}=\bar{\pi}(\hat{\theta})$ and $\hat{\bar{\pi}}_r=\bar{\pi}_r(\hat{\theta})$, with $\bar{\pi}_r(\theta)=D_r\bar{\pi}(\theta)$, $1\leqslant r\leqslant p$.

We now choose $\bar{\pi}(\ )$ such that $\bar{\pi}(\ )$ and its first-order partial derivatives vanish on the boundaries of a rectangle containing θ as an interior point. We then integrate $E_\theta[P^{\bar{\pi}}\{\theta\in R_1(\alpha, \pi, X)|X\}]$, which can be obtained from equations (1.1d) and (A.3) up to $o(n^{-1})$, with respect to such a $\bar{\pi}(\ )$ and finally allow $\bar{\pi}(\ )$ to converge weakly to the degenerate measure at θ. After some simplification, this yields equation (2.5); see Ghosh and Mukerjee (1991) for more details on this technique.

The proof of equation (2.14) is similar.

REFERENCES

Barndorff-Nielsen, O. E. and Hall, P. (1988) On the level error after Bartlett adjustment of the likelihood ratio statistic. *Biometrika*, **75**, 374–378.

Bhattacharya, R. N. and Ghosh, J. K. (1978) On the validity of the formal Edgeworth expansion. *Ann. Statist.*, **6**, 434–451.

Bickel, P. J. and Ghosh, J. K. (1990) A decomposition of 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. *Sankhya* A, **41**, 22–47.

DiCiccio, T. J. and Stern, S. E. (1993) On Bartlett adjustments for approximate Bayesian inference. *Biometrika*, **80**, 731–740.

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

——(1992) Bayesian and frequentist Bartlett corrections for likelihood ratio and conditional likelihood ratio tests. *J. R. Statist. Soc.* B, **54**, 867-875.

——(1993) Frequentist validity of highest posterior density regions in the multiparameter case. *Ann. Inst. Statist. Math.*, **45**, 293-302; correction, 602.

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

Lee, C. B. (1989) Comparison of frequentist coverage probability and Bayesian posterior coverage probability, and applications. *PhD Thesis*. Purdue University, West Lafayette.

Mukerjee, R. and Dey, D. K. (1993) Frequentist validity of posterior quantiles in the presence of a nuisance parameter: higher order asymptotics. *Biometrika*, **80**, 499-505.

Severini, T. A. (1991) On the relationship between Bayesian and non-Bayesian interval estimates. *J. R. Statist. Soc.* B, **53**, 611-618.

——(1993) Bayesian interval estimates which are also confidence intervals. *J. R. Statist. Soc.* B, **55**, 533-540.

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