Ann. Inst. Statist. Math.
Vol. 46, No. 1, 21–28 (1994)

CONDITIONAL PROPERTIES OF BAYESIAN INTERVAL ESTIMATES

ANNA NICOLAOU

Economic and Social Sciences, University of Macedonia, Egnatia str. 156, P.O. Box 1591, 54006 Thessaloniki, Greece

(Received April 13, 1992; revised March 23, 1993)

Abstract. Consider the construction of an interval estimate for a scalar parameter of interest in the presence of orthogonal nuisance parameters. A conditional prior density on the parameter of interest that is proportional to the square root of its information element, generates one-sided Bayes intervals that are approximately confidence intervals as well, having coverage error of order $O(1/n)$, where n is the sample size. We show that the frequency property of these intervals also holds conditionally on a locally ancillary statistic near the true parameter value.

Key words and phrases: Bayes intervals, nuisance parameters, orthogonal parameters, local ancillarity.

1. Introduction

Let $Y, Y_1, Y_2, \ldots$ be independent and identically distributed random variables with density function that depends on a p-dimensional parameter $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$. Given a set of observations $\boldsymbol{y}_n = (y_1, \ldots, y_n)$ we write $l(\boldsymbol{\theta}, \boldsymbol{y}_n)$ for the log likelihood function and $L(\boldsymbol{\theta}, \boldsymbol{y}_n) = l(\boldsymbol{\theta}, \boldsymbol{y}_n)/n$ for its standardized version. The derivatives of $l(\boldsymbol{\theta}, Y)$ and the corresponding cumulants are denoted by

$$\begin{aligned} &l_r = \partial l(\boldsymbol{\theta}, Y)/\partial\theta_r, \qquad l_{rs} = \partial^2 l(\boldsymbol{\theta}, Y)/\partial\theta_r\partial\theta_s, \\ &\kappa_r = E(l_r; \boldsymbol{\theta}), \qquad \kappa_{rs} = E(l_{rs}; \boldsymbol{\theta}), \\ &\kappa_{rs,t} = \operatorname{cov}(l_{rs}, l_t; \boldsymbol{\theta}), \qquad \kappa_{r,s,t} = \operatorname{cum}(l_r, l_s, l_t; \boldsymbol{\theta}) \end{aligned}$$

and so on. Interest centers on one particular component of the parameter, θ_1 say, treating the remaining components as a nuisance parameter. We assume that the nuisance parameter $\boldsymbol{\theta}_2 = (\theta_2, \ldots, \theta_p)$ is orthogonal to θ_1 with respect to the expected Fisher information matrix; that is $\kappa_{1,r} = 0$ for $r = 2, \ldots, p$ (Cox and Reid (1987)).

The upper bound for the parameter of interest θ_1 of Bayes size α, that is generated from the prior density

$$\pi(\theta_1, \boldsymbol{\theta}_2) = \sqrt{\kappa_{1,1}(\theta_1, \boldsymbol{\theta}_2)}\, g(\boldsymbol{\theta}_2), \tag{1.1}$$

where $\kappa_{1,1}$ is the information element for θ_1 and $g(\cdot)$ is an arbitrary function of the nuisance parameter, also has probabilistic properties of interest in a frequency approach; it is an approximate confidence bound covering the true parameter value with sampling probability $\alpha+O(1/n)$ (Peers (1965), Stein (1982), Tibshirani (1989), Nicolaou (1991)). The specified part of the prior density (1.1) coincides with the reference prior for θ_1 with θ_2 given, derived by Bernardo (1979) on the grounds of information concepts. In some cases, a specific choice of the marginal prior for the nuisance parameter reduces the coverage error to order $O(1/n\sqrt{n})$. However, the resulting joint prior need not yield sets with good confidence properties for the nuisance parameter as well. In the case of a scalar nuisance parameter a conditional prior for it can be proposed by a similar argument considering vise versa θ_2 as the parameter of interest. Then, if the ratio of the information elements for θ_1 and θ_2 can be written as $h(\theta_1)/g(\theta_2)$ for some functions $h(\cdot)$ and $g(\cdot)$, a unique joint prior good for inferences about both parameters and consistent with their conditional priors is given by $\sqrt{\kappa_{1,1}(\theta_1,\theta_2)g(\theta_2)}$.

In the single-parameter problem Welch and Peers (1963) have shown that the appropriate choice is the Jeffreys' prior. A normal approximation to the corresponding posterior distribution of θ (Johnson (1970)) yields a pivotal quantity W_n that can be used to construct approximate confidence limits with coverage error of order $O(1/n)$. The statistic W_n is a transformation of the maximum likelihood pivot of θ, $T_n = \sqrt{-n\hat{L}''}(\theta_0-\hat{\theta})$, of the form

$$W_n = T_n + \frac{1}{\sqrt{-n\hat{L}''}}\left(\frac{\hat{L}'''(T_n^2+2)+3\hat{\iota}_2'}{6\hat{L}''}\right),$$

where $\iota_2(\theta) = E((l')^2;\theta)$. The dash denotes differentiation with respect to θ and the hat evaluation at the maximum likelihood estimator $\hat{\theta}$. The asymptotic expansion for W_n in terms of the log-likelihood derivatives and their cumulants gives $W_n = -S/\sqrt{\iota_2}+O_p(1/n)$, where S is the locally sufficient statistic for θ introduced by McCullagh ((1984), Formula 26). His approach is to form a pivotal statistic that is independent of a locally ancillary statistic at θ_0; then, inferences based on the marginal distribution of this statistic are automatically conditional. Hence, the frequency property of the Bayes intervals that are generated from Jeffreys' prior also holds conditionally given a locally ancillary statistic. A parameterization invariant form of S is obtained by expressing it in terms of the signed likelihood ratio statistic $z_{\text{dev}} = \text{sgn}(\theta_0-\hat{\theta})\sqrt{2(l(\hat{\theta},\boldsymbol{y}_n)-l(\theta_0,\boldsymbol{y}_n))}$, adjusted by its mean, as

$$-\frac{S}{\sqrt{\iota_2}} = z_{\text{dev}} - E(z_{\text{dev}}) + O_p\left(\frac{1}{n}\right).$$

Furthermore, $S/\sqrt{\iota_2}$ agrees to the order $O_p(1/n)$ with the modification of the signed likelihood ratio that Barndorff-Nielsen derives ((1986), Section 3.5), using a different approach based on approximations of the conditional distribution of the maximum likelihood estimator given the observed value of an approximately ancillary statistic. In fact, Barndorff-Nielsen's conditional confidence intervals have coverage error of order $O(n^{-3/2})$ and his methods extend to encompass the possibility of nuisance parameters (Barndorff-Nielsen (1991)). However, their derivation requires specification of the approximate ancillary statistic.

In this paper we show that the conditional probability of coverage of the one sided Bayes intervals for the parameter of interest that result from the prior density (1.1), still differs from the nominal probability of coverage by $O(1/n)$, given a statistic that is second order locally ancillary near the true parameter value $\boldsymbol{\theta}_0$. The description of the argument is presented in Section 2. Some technical details are given in the Appendix.

2. Main result

For each set of observations we generate the posterior distribution function of θ_1, $\rho(\cdot \mid \boldsymbol{y}_n)$, corresponding to the prior density (1.1). This evaluated at the true parameter value θ_{01} can be approximated by the standard normal distribution as

$$\rho(\theta_{01} \mid \boldsymbol{y}_n) = \Phi(W_n) + O_p\left(\frac{1}{n}\right),$$

where, summing over any index that is repeated,

$$W_n = T_n - \sqrt{\frac{-\hat{L}^{11}}{n}} \left(\frac{\partial \hat{\kappa}_{11}}{\partial \theta_1} \frac{1}{2\hat{\kappa}_{11}} - (T_n^2 - 1)\frac{\hat{L}^{11}\hat{L}_{111}}{6} - \frac{\hat{L}^{ij}\hat{L}_{1ij}}{2} \right), \tag{2.1}$$

$T_n = \sqrt{n}(\theta_{01} - \hat{\theta}_1)/\sqrt{-\hat{L}^{11}}$ and L^{ij} is the matrix inverse of L_{ij} (Lemma A.1 of the Appendix). The distribution of the adjusted pivot W_n can be approximated by an Edgeworth expansion that has the standard normal distribution as a leading term. By virtue of the chosen prior density the $1/\sqrt{n}$ term of the expansion vanishes, and hence the normal approximation holds up to $O(1/n)$. Consequently, $\rho(\theta_{01} \mid \boldsymbol{y}_n)$ is, under repeated sampling, approximately uniformly distributed over the range $(0,1)$ and the corresponding posterior quantile covers the true parameter value with sampling probability $\alpha + O(1/n)$ (Welch and Peers (1963), Peers (1965)).

This result remains valid conditionally on the observed value of a statistic $\mathcal{A} = \mathcal{A}(\boldsymbol{y}_n, \boldsymbol{\theta}_0)$ that is locally ancillary to the second order near $\boldsymbol{\theta}_0$ i.e. whose distribution at $\boldsymbol{\theta}_0 + \boldsymbol{\delta}/\sqrt{n}$, for some fixed $\boldsymbol{\delta}$, differs from that at $\boldsymbol{\theta}_0$ by terms of order $O(1/n)$. The construction of locally ancillary statistics is discussed by Cox (1980) and McCullagh ((1984, 1987), Chapter 8). To proceed some additional notation is necessary. Consider the linear functions of the log likelihood derivatives $V_r = l_r$, $V_{rs} = l_{rs} - \beta^i_{rs} l_i$ and denote their cumulants by ν's. The coefficients $\beta^r_{st} = \kappa_{i,st}\kappa^{i,r}$, are defined so that $\nu_{r,st} = \mathrm{cov}(V_r, V_{st}) = 0$. In the following, when we suppress the dependence of the cumulants on the parameter value we mean that they are evaluated at $\boldsymbol{\theta}_0$. Define the asymptotically normal random variables of zero mean and constant variance

$$Z_r = \frac{\sum_{i=1}^n l_r(\boldsymbol{\theta}_0, Y_i)}{\sqrt{n}}, \qquad Z_{rs} = \frac{\sum_{i=1}^n (l_{rs}(\boldsymbol{\theta}_0, Y_i) - \kappa_{rs})}{\sqrt{n}}$$

and their transformations $R_r = Z_r$, $R_{rs} = Z_{rs} - \beta^t_{rs} Z_t$. A second order ancillary statistic $\mathcal{A}$ is an array with components of the form

$$\mathcal{A}_{rs} = R_{rs} - \frac{1}{\sqrt{n}} \beta^{i,j}_{rs} R_i R_j - \frac{1}{\sqrt{n}} \beta^{i,jk}_{rs} R_i R_{jk} + O_p\left(\frac{1}{n}\right),$$

where the coefficients

$$\beta^{i,j}_{rs} = (\nu_{rs,kl} + \nu_{rs,k,l}) \nu^{i,k} \nu^{j,l}/2,$$
$$(\beta^{i,jk}_{rs} \nu_{jk,tu} + \beta^{i,jk}_{tu} \nu_{jk,rs}) \nu_{iv} = \nu_{rs,tu,v}$$

are determined by the ancillarity requirement. The solution to the second equation is not unique.

For the α-quantile of $\rho(\cdot \mid \boldsymbol{y}_n)$ to be an approximate conditional confidence limit for θ_1, it suffices to show that the statistic W_n is independent of the ancillary $\mathcal{A}$ to second order, locally at $\boldsymbol{\theta}_0$. The second order independence follows if the third-order joint cumulants of W_n and $\mathcal{A}$ satisfy

$$\begin{aligned} \text{(2.2)} \qquad \operatorname{cov}(W_n, \mathcal{A}_{rs}) &= O\left(\frac{1}{n}\right) \operatorname{cum}(W_n, W_n, \mathcal{A}_{rs}) \\ &= O\left(\frac{1}{n}\right) \operatorname{cum}(W_n, \mathcal{A}_{rs}, \mathcal{A}_{tu}) = O\left(\frac{1}{n}\right). \end{aligned}$$

We now proceed deriving approximations for these cumulants to show that they vanish up to the required order. An asymptotic expansion of W_n in terms of the $O_p(1)$ quantities R's is obtained as

$$\text{(2.3)} \qquad W_n = W_n^* + \frac{C}{\sqrt{n}} + O_p\left(\frac{1}{n}\right),$$

where

$$\begin{aligned} W_n^* = -\sqrt{\nu^{1,1}} \Bigg\{ R_1 + \frac{1}{2\sqrt{n}} \Bigg(& 2\nu^{i,j}(R_{1i} + \beta^t_{1i}) R_j + \nu^{i,r} \nu^{j,s} \nu_{1rs} R_i R_j \\ & - \nu^{1,1} R_1 (R_{11} + \beta^t_{11} R_t + \nu_{i11} \nu^{i,j} R_j) + (\nu^{1,1} R_1)^2 \frac{\nu_{111}}{3} \Bigg) \Bigg\}, \end{aligned}$$

and C is a constant function of $\boldsymbol{\theta}_0$, not depending on n

$$C = \sqrt{\nu^{1,1}} \left(-\frac{1}{2\nu_{11}} \frac{\partial \nu_{11}}{\partial \theta_1} - \frac{\nu^{i,j} \nu_{1ij}}{2} + \frac{\nu^{1,1} \nu_{111}}{6} \right),$$

by substituting into (2.1) the following expansions

$$\sqrt{n}(\theta_{01} - \hat{\theta}_1) = -\kappa^{1,1} \left(Z_1 + \frac{2\kappa^{i,j} Z_{1i} Z_j + \kappa^{i,r} \kappa^{j,s} \kappa_{1rs} Z_i Z_j}{2\sqrt{n}} \right) + O_p\left(\frac{1}{n}\right),$$

$$(-\hat{L}^{11})^{-1/2} = (\kappa^{1,1})^{-1/2} \left\{ 1 - \frac{1}{2\sqrt{n}} \kappa^{1,1} (Z_{11} + \kappa_{i11} \kappa^{i,j} Z_j) \right\} + O_p\left(\frac{1}{n}\right),$$

$$\hat{L}_{ijk} = \kappa_{ijk} + O_p\left(\frac{1}{\sqrt{n}}\right).$$

Combining the orthogonality condition with the cumulant formulae, derived in Lemma A.2 (see the Appendix), we obtain, after some algebra, that

$$\begin{aligned}
\operatorname{cov}(W_n, \mathcal{A}_{rs}) &= -\sqrt{\nu^{1,1}}\left(\nu_{1,rs} + \frac{1}{\sqrt{n}}\nu^{i,j}\beta^t_{1i}\nu_{j,rs}\right) + O\left(\frac{1}{n}\right), \\
\operatorname{cum}&(W_n, W_n, \mathcal{A}_{rs}) \\
&= \frac{\nu^{1,1}}{\sqrt{n}}(\nu_{1,1,rs} + 2\nu^{i,j}\nu_{1,j}\nu_{rs,1i} - \nu^{1,1}\nu_{1,1}\nu_{rs,11} - 2\beta^{i,j}_{rs}\nu_{1,i}\nu_{1,j}) + O\left(\frac{1}{n}\right), \\
\operatorname{cum}&(W_n, \mathcal{A}_{rs}, \mathcal{A}_{tu}) \\
&= -\sqrt{\frac{\nu^{1,1}}{n}}(\nu_{1,rs,tu} - \nu_{1,i}(\beta^{i,jk}_{rs}\nu_{tu,jk} + \beta^{i,jk}_{tu}\nu_{jk,rs})) + O\left(\frac{1}{n}\right).
\end{aligned}$$

At this stage we substitute the explicit formulae for the constants β's and take into account that $\nu_{r,st} = 0$, to deduce that

$$\begin{aligned}
\operatorname{cov}(W_n, \mathcal{A}_{rs}) &= O\left(\frac{1}{n}\right), \\
\operatorname{cum}(W_n, W_n, \mathcal{A}_{rs}) & \\
&= \frac{\nu^{1,1}}{\sqrt{n}}(\nu_{1,1,rs} + 2\nu_{rs,11} - \nu_{rs,11} - (\nu_{rs,kl} + \nu_{rs,k,l})\nu^{i,k}\nu^{j,l}\nu_{1,i}\nu_{1,j}) \\
&\quad + O\left(\frac{1}{n}\right) \\
&= \frac{\nu^{1,1}}{\sqrt{n}}(\nu_{1,1,rs} + \nu_{rs,11} - (\nu_{rs,11} + \nu_{rs,1,1})) + O\left(\frac{1}{n}\right) \\
&= O\left(\frac{1}{n}\right), \\
\operatorname{cum}(W_n, \mathcal{A}_{rs}, \mathcal{A}_{tu}) &= -\sqrt{\frac{\nu^{1,1}}{\sqrt{n}}}(\nu_{1,rs,tu} - \nu_{1,rs,tu}) + O\left(\frac{1}{n}\right) = O\left(\frac{1}{n}\right)
\end{aligned}$$

as had to be proved.

Remark 1. Assume for simplicity that the nuisance parameter is one dimensional and consider the signed log likelihood ratio statistic for testing a given value of θ_1

$$z_{\text{dev}} = \operatorname{sgn}(\theta_{01} - \hat{\theta}_1)\sqrt{2(l(\hat{\theta}_1, \hat{\theta}_2, \boldsymbol{y}_n) - l(\theta_{01}, \tilde{\theta}_2, \boldsymbol{y}_n))},$$

where $\tilde{\theta}_2$ is the restricted maximum likelihood estimator of θ_2 with θ_1 fixed at θ_{01}. Regrouping the terms of the modified pivot W_n we obtain

$$\begin{aligned}
W_n = &\left[T_n - \frac{1}{6\sqrt{n}}T_n^2\hat{L}_{111}(-\hat{L}^{11})^{3/2}\right] \\
&- \left[\frac{(-\hat{L}^{11})^{1/2}}{\sqrt{n}}\left(\frac{-\hat{L}^{11}\hat{L}_{111}}{3} + \frac{\partial\hat{\kappa}_{11}}{\partial\theta_1}\frac{1}{2\hat{\kappa}_{11}} - \frac{\hat{L}_{122}\hat{L}^{22}}{2}\right)\right].
\end{aligned}$$

The espressions in the square brackets are the asymptotic expansions in terms of the derivatives of L evaluated at $\hat{\theta}$, of z_{dev} and its posterior mean $E_n(z_{\text{dev}})$ respectively (see Ghosh and Mukerjee (1992), for the derivation of the posterior distribution of z_{dev}). Therefore, W_n has a parametrization-independent expression as $W_n = z_{\text{dev}} - E_n(z_{\text{dev}}) + O_p(1/n)$. Note that in proving the local sufficiency of W_n we used an alternative expansion in terms of expected rather than observed quantities; W_n^* in (2.3) is the asymptotic expansion of z_{dev} about the true parameter point, and $-C/\sqrt{n}$ is its approximate expected value.

Remark 2. The assumption that θ_1 is orthogonal to the nuisance parameter is nontrivial; for a general non-orthogonal parameter the frequency property of the Bayes intervals for θ_1 does not hold conditionally on the locally ancillary statistic.

Example. Let $y_1, \ldots, y_n$ be a sample from the gamma distribution $f(y, \kappa, \mu) = (\mu/\kappa)^\kappa y^{\kappa-1} \exp(-\kappa y/\mu)/\Gamma(\kappa)$ and suppose that the parameter of interest is the mean μ, the shape κ being a nuisance parameter. The joint prior density for which the posterior quantiles are approximate confidence limits with coverage error $O(1/n)$ conditionally on a locally ancillary statistic as well as unconditionally is

$$\pi(\kappa, \mu) \propto \frac{\sqrt{\Psi'(\kappa) - 1/\kappa}}{\mu},$$

where $\Psi'(\kappa) = \sum_{i=0}^{\infty}(\kappa + i)^{-2}$ is the trigamma function. An approximation to the α-posterior quantile for the mean, derived by inverting the asymptotic series of Lemma A.1, is

$$\hat{\mu}\left(1 + \frac{z_\alpha}{\sqrt{n\hat{\kappa}}} + \frac{2z_\alpha^2 + 1}{3n\hat{\kappa}}\right), \tag{2.4}$$

where z_α is the α-quantile of the standard normal distribution.

Alternatively, Barndorff-Nielsen's ((1986), Example 3.3) modified signed likelihood ratio statistic for μ, given to the order concerned by

$$r^*(\mu) = \operatorname{sgn}(\mu - \hat{\mu})\sqrt{2(l(\hat{\kappa}, \hat{\mu}, \boldsymbol{y}_n) - l(\hat{\kappa}_\mu, \mu, \boldsymbol{y}_n))} + \frac{1}{3\sqrt{n\hat{\kappa}_\mu}},$$

follows the standard normal distribution to $O(1/n)$ conditionally on an approximate ancillary statistic. Therefore, an approximate conditional confidence bound may be determined by solving $r^*(\mu) \leq z_\alpha$. Taking into account that the orthogonality of the parameters implies that $\hat{\kappa}_\mu = \hat{\kappa} + O_p(1/n)$ and expanding

$$l(\hat{\kappa}, \hat{\mu}, \boldsymbol{y}_n) = l(\hat{\kappa}_\mu, \mu, \boldsymbol{y}_n) - \frac{\hat{\kappa}}{2\hat{\mu}^2}(\mu - \hat{\mu})^2 + \frac{4\hat{\kappa}}{6\hat{\mu}^3}(\mu - \hat{\mu})^3 + O_p\left(\frac{1}{n}\right),$$

we find that the asymptotic expansion of the resulting bound agrees with expression (2.4).

Appendix

The following lemma gives the asymptotic behavior of $\rho(\theta_{01} \mid \boldsymbol{y}_n)$ (Peers (1965), Johnson (1970), DeBruijn (1981)).

LEMMA A.1. *Under certain regularity conditions (Nicolaou (1991)) if* $T_n = \sqrt{n}(\theta_{01} - \hat{\theta}_1)/\sqrt{-\hat{L}^{11}}$ *the posterior distribution function of* θ_1 *evaluated at the true parameter value* θ_{01} *is approximated as*

$$\rho(\theta_{01} \mid \boldsymbol{y}_n) = \Phi(T_n) - \sqrt{\frac{-\hat{L}^{11}}{n}}\left(\frac{\partial}{\partial\theta_1}\log\hat{\pi} - (T_n^2 - 1)\frac{\hat{L}^{11}\hat{L}_{111}}{6} - \frac{\hat{L}^{ij}\hat{L}_{1ij}}{2}\right)\phi(T_n) + O\left(\frac{1}{n}\right),$$

where π *is the corresponding prior density and* Φ *and* ϕ *are the standard normal distribution and density function. The hat denotes evaluation of the functions at the maximum likelihood estimator* $\hat{\boldsymbol{\theta}}$.

LEMMA A.2. *The ordinary and generalized cumulants of* R_r, R_{rs} *ignoring terms of order* $O(1/n)$ *are*

$$ER_r = 0, \quad \operatorname{cov}(R_r, R_s) = \nu_{r,s}, \quad \operatorname{cum}(R_r, R_s, R_t) = \frac{\nu_{r,s,t}}{\sqrt{n}},$$
$$\operatorname{cov}(R_r, R_{st}) = 0, \quad \operatorname{cov}(R_r, R_s R_t) = \frac{\nu_{r,s,t}}{\sqrt{n}}, \quad \operatorname{cov}(R_{rs}, R_t R_u) = \frac{\nu_{rs,t,u}}{\sqrt{n}},$$
$$\operatorname{cov}(R_r, R_s R_{tu}) = \frac{\nu_{r,s,tu}}{\sqrt{n}}, \quad \operatorname{cov}(R_{rs}, R_t R_{uv}) = \frac{\nu_{rs,t,uv}}{\sqrt{n}},$$
$$\operatorname{cum}(R_{rs}, R_t, R_u R_v) = 0,$$
$$\operatorname{cov}(R_r R_s, R_t R_u) = \nu_{r,t}\nu_{s,u} + \nu_{r,u}\nu_{s,t}, \quad \operatorname{cum}(R_{rs}, R_t, R_u R_{vw}) = \nu_{t,u}\nu_{rs,vw},$$
$$\operatorname{cum}(R_r, R_s, R_{tu}) = \frac{\nu_{r,s,tu}}{\sqrt{n}}, \quad \operatorname{cum}(R_{rs}, R_{uv}, R_t R_{wz}) = 0.$$

The proof of Lemma A.2 follows from tedious, elementary calculations using the formulae (3.2) of McCullagh (1987) to express generalized cumulants as combinations of ordinary cumulants.

References

Barndorff-Nielsen, O. E. (1986). Inference on full or partial parameters based on the standardized signed log likelihood ratio, *Biometrika*, **73**, 307–322.

Barndorff-Nielsen, O. E. (1991). Modified signed log likelihood ratio, *Biometrika*, **78**, 557–563.

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference (with discussion), *J. Roy. Statist. Soc. Ser. B*, **41**, 113–147.

Cox, D. R. (1980). Local ancillarity, *Biometrika*, **67**, 279–286.

Cox, D. R. and Reid, R. (1987). Orthogonal parameters and approximate conditional inference (with discussion), *J. Roy. Statist. Soc. Ser. B*, **49**, 1–39.

DeBruijn, N. C. (1981). *Asymptotic Methods in Analysis*, Dover, New York.

Ghosh, J. K. and Mukerjee, R. (1992). Bayesian and frequentist Bartlett corrections for likelihood ratio and conditional likelihood ratio tests, *J. Roy. Statist. Soc. Ser. B*, **54**, 867–875.

Johnson, R. A. (1970). An asymptotic expansion for posterior distributions, *Ann. Math. Statist.*, **38**, 1899–1907.

McCullagh, P. (1984). Local sufficiency, *Biometrika*, **71**, 233–244.

McCullagh, P. (1987). *Tensor Methods in Statistics*, Chapman and Hall, London.

Nicolaou, A. (1991). Bayesian intervals with good frequentist behavior in the presence of nuisance parameters, *J. Roy. Statist. Soc. Ser. B*, **55**(2), 377–390.

Peers, H. W. (1965). On confidence points and Bayesian probability points in the case of several parameters, *J. Roy. Statist. Soc. Ser. B*, **27**, 9–16.

Stein, C. (1982). On the coverage probability of confidence sets based on a prior distribution, Tech. Report, 180, Dept. of Statistics, Stanford University.

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

Welch, B. L. and Peers, H. W. (1963). On formulae for confidence points based on integrals of weighted likelihoods, *J. Roy. Statist. Soc. Ser. B*, **25**, 318–329.