Statistica Sinica **13**(2003), 409-422

OBJECTIVE BAYESIAN INFERENCE FOR RATIOS OF REGRESSION COEFFICIENTS IN LINEAR MODELS

Malay Ghosh, Ming Yin and Yeong-Hwa Kim

University of Florida, Analytical Sciences Incorporated and Chung-Ang University

Abstract: The paper considers the standard linear multiple regression model where the parameter of interest is a ratio of two regression coefficients. The general model includes the calibration model, the Fieller-Creasy problem, slope-ratio assays, parallel-line assays and bioequivalence. We provide a unified objective Bayesian analysis for such problems. Both reference priors and probability matching priors are found. Based on some numerical findings, our recommended prior is the one-at-a-time reference prior. The analysis is greatly facilitated by an orthogonal (Cox and Reid (1987)) reparameterization of the original parameter vector.

Key words and phrases: Calibration, Fieller-Creasy, matching priors, orthogonal transformation, parallel-line assay, reference priors, slope-ratio assay.

1. Introduction

There is a large class of important statistical problems which can be broadly described under the heading of inference about the ratio of regression coefficients in a general linear model. Included in this class are (a) the calibration problem, (b) ratio of two means or the Fieller-Creasy problem, (c) slope-ratio assay, (d) parallel-line assay, and (e) bioequivalence. The frequentist solutions to these problems have typically encountered serious difficulties. As an example, confidence sets for ratios of two normal means based on Fieller's pivot (1954) may be degenerate or a union of two disjoint unbounded intervals and, in extreme situations, may even be the entire real line. The same phenomenon occurs as well in the other problems described above.

The prime objective of this paper is to present a unified Bayesian analysis for this general problem using objective priors. We develop both reference priors as well as matching priors for the general problem described above. As we will see in later sections, this process of development will suggest consideration of two general classes of priors. All the reference priors belong to one class, while Jeffreys' prior as well as a second order probability matching prior (to be derived in Section 3) belong to the other class. The former class of priors is considered in Buonaccorsi and Gatsonis (1988) and Mendoza (1988), but the latter has never been treated before in its full generality. Various priors proposed for the specific problems mentioned in (a)−(d) also belong to one or the other of these two general classes. In particular for the calibration problem, Hoadley's (1970) prior plus all the reference priors considered in Ghosh, Carlin and Srivastava (1995) are members of one class, while the prior considered by Hunter and Lamboy (1981) belongs to the other class.

Liseo (1993), in an elegant article, discussed the general problem of elimination of nuisance parameters by a Bayesian approach. As an example, for the ratio of two normal means (the Fieller-Creasy problem) with known common standard deviation, he derived and compared both Jeffreys' prior and the one-at-a-time reference prior (to be introduced in Section 3) and found the superiority of the latter over the former. Yin and Ghosh (2000) obtained the different reference priors as well as matching priors for the problem of the ratio of two location parameters with common unknown scale parameter. For the normal example, all these priors can be viewed as special cases of the priors proposed in this paper.

The two-group reference prior for the slope-ratio assay, as derived by Mendoza (1990), is a special case of our general class of priors. The priors of Buonaccorsi and Gatsonis (1988) for the same problem are different, but are neither reference priors nor matching priors. For the parallel-line assay, Bayesian analyses of the log-relative potency with proper subjective priors are given in Darby (1980). Kim, Carter and Hubert (1991) and Kim, Carter, Hubert and Hand (1993) assigned a prior to the parameter of interest, and used MLE's for the nuisance parameters. Thus their approach, unlike ours, is only a partial Bayes approach.

The key to the derivation of the different priors is a nontrivial orthogonal reparameterization (cf. Cox and Reid (1987)) of the original parameter vector. This orthogonal reparameterization, introduced in Section 2, unifies the development of the different objective priors.

Section 3 develops the various objective priors. Because of the orthogonality of the parameter of interest with the nuisance parameters, we can appeal to the various available general results on the development of reference priors as well as matching priors. Necessary and sufficient conditions for the propriety of posteriors for the two general classes of priors are also given in this section. Section 4 illustrates the general theory for the specific problems of slope-ratio and parallel-line assays, and obtains the necessary expressions needed for the derivation of the different posteriors. We provide also a numerical illustration of the proposed procedure for a parallel-line assay problem. Some concluding remarks are made in Section 5. The proof of a technical result is deferred to the Appendix.

2. The Orthogonal Transformation

Consider the general regression model

$$y_i = \sum_{j=1}^{r} \beta_j x_{ij} + e_i, \quad i = 1, \ldots, n, \tag{2.1}$$

where the errors e_i are *i.i.d.* $\mathcal{N}(0, \sigma^2)$. Here $\beta_j \in (-\infty, \infty)$ for $j \neq 2$, while $\beta_2 \in (-\infty, \infty) - \{0\}$. The parameter of interest is $\theta_1 = \beta_1/\beta_2$. We write $\boldsymbol{y} = (y_1, \ldots, y_n)^T$, $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ir})^T$, $i = 1, \ldots, n$, $\boldsymbol{X}^T = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_r)^T$, $\boldsymbol{e} = (e_1, \ldots, e_n)^T$. Thus in matrix notations, the model can be rewritten as $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}$. We assume that $\operatorname{rank}(\boldsymbol{X}) = r < n$.

First we introduce a transformation of the parameter vector $(\boldsymbol{\beta}, \sigma)$ which results in the orthogonality of θ_1 with the remaining parameters. We begin with the Fisher information matrix

$$\boldsymbol{I}(\boldsymbol{\beta}, \sigma) = n\sigma^{-2} \begin{bmatrix} ((s_{jl})) & \boldsymbol{0} \\ \boldsymbol{0}^T & 2 \end{bmatrix}, \tag{2.2}$$

where $s_{jl} = n^{-1} \sum_{i=1}^{n} x_{ij} x_{il}$, j, $l = 1, \ldots, r$. Consider the transformation

$$\beta_1 = \theta_1 \theta_2 h(\theta_1); \quad \beta_2 = \theta_2 h(\theta_1); \quad \beta_j = \theta_j - \theta_2 g_j(\theta_1), \quad j = 3, \ldots, r; \quad \sigma = \theta_{r+1}. \tag{2.3}$$

Then the Jacobian matrix is given by

$$\boldsymbol{J} = \begin{bmatrix} \theta_2\{\theta_1 h'(\theta_1) + h(\theta_1)\} & \theta_2 h'(\theta_1) & -\theta_2 g_3'(\theta_1) & \cdots & -\theta_2 g_r'(\theta_1) & 0 \\ \theta_1 h(\theta_1) & h(\theta_1) & -g_3(\theta_1) & \cdots & -g_r(\theta_1) & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{2.4}$$

Let $\boldsymbol{X}^T\boldsymbol{X} = n \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}$, where

$$\boldsymbol{A}_{11} = \begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix}, \quad \boldsymbol{A}_{12} = \boldsymbol{A}_{21}^T = \begin{pmatrix} s_{13} & \cdots & s_{1r} \\ s_{23} & \cdots & s_{3r} \end{pmatrix}, \quad \text{and} \quad \boldsymbol{A}_{22} = \begin{pmatrix} s_{33} & \cdots & s_{3r} \\ \vdots & \ddots & \vdots \\ s_{r3} & \cdots & s_{rr} \end{pmatrix}.$$

Also, let $\boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix} = \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21} = \boldsymbol{A}_{11.2}$, say. We define $Q(\theta_1) = c_{11}\theta_1^2 + 2c_{12}\theta_1 + c_{22}$. Since $\mathrm{rank}(\boldsymbol{X}^T\boldsymbol{X}) = \mathrm{rank}(\boldsymbol{X}) = r$, $\boldsymbol{X}^T\boldsymbol{X}$ is positive definite so that $\boldsymbol{C}$ is also positive definite. This implies that $Q(\theta_1)$ is positive for all θ_1. Then we have the following theorem whose proof is omitted. The details are available from the authors.

Theorem 1. *Let $g_j(\theta_1) = h(\theta_1)(a_{j1}\theta_1 + a_{j2})$, $j = 3, \ldots, r$, where a_{jt} is the (j,t)th element of $\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}$. Then parametric orthogonality holds by choosing $h(\theta_1) = Q^{-\frac{1}{2}}(\theta_1)$.*

Based on the transformation given in the theorem, it follows from (2.1)−(2.4) that the reparameterized Fisher information matrix is given by

$$\boldsymbol{I}(\boldsymbol{\theta}) = n\theta_{r+1}^{-2} \begin{bmatrix} \frac{\theta_2^2|\boldsymbol{C}|}{Q^2(\theta_1)} & 0 & \mathbf{0}^T & 0 \\ 0 & 1 & \mathbf{0}^T & 0 \\ \mathbf{0} & \mathbf{0} & \boldsymbol{A}_{22} & \mathbf{0} \\ 0 & 0 & \mathbf{0}^T & 2 \end{bmatrix}. \tag{2.5}$$

We make repeated use of this information matrix for the development of various priors.

3. Developemtn of Objective Priors

3.1. Jeffreys' and reference priors

We begin with Jeffreys' prior given by

$$\pi^J(\boldsymbol{\theta}) \propto |\boldsymbol{I}(\boldsymbol{\theta})|^{\frac{1}{2}} \propto \theta_{r+1}^{-(r+1)}|\theta_2|Q^{-1}(\theta_1). \tag{3.1}$$

This is a reference prior when all parameters are treated as equally important. The two-group reference prior of Bernardo (1979) with $\boldsymbol{\theta}_{(1)} = \{\theta_1\}$ and $\boldsymbol{\theta}_{(2)} = \{\theta_2, \ldots, \theta_{r+1}\}$ is given by $\pi^{2R}(\boldsymbol{\theta}) \propto \theta_{r+1}^{-r}Q^{-1}(\theta_1)$. The three-group reference prior with $\boldsymbol{\theta}_{(1)} = \{\theta_1\}$, $\boldsymbol{\theta}_{(2)} = \{\theta_2, \ldots, \theta_r\}$, and $\boldsymbol{\theta}_{(3)} = \{\theta_{r+1}\}$, arranged according to their order of importance, is given by $\pi^{3R}(\boldsymbol{\theta}) \propto \theta_{r+1}^{-1}Q^{-1}(\theta_1)$. This is different from the three-group reference prior with $\boldsymbol{\theta}_{(1)} = \{\theta_1\}$, $\boldsymbol{\theta}_{(2)} = \{\theta_2\}$, and $\boldsymbol{\theta}_{(3)} = \{\theta_3, \ldots, \theta_{r+1}\}$: $\pi_*^{3R}(\boldsymbol{\theta}) \propto \theta_{r+1}^{-(r-1)}Q^{-1}(\theta_1)$. Finally the one-at-a-time reference prior with the partition $\{\theta_1\}, \ldots, \{\theta_{r+1}\}$, where θ_1 is the parameter of interest and the remaining parameters are considered in an arbitrary order of importance, is given by $\theta_{r+1}^{-1}Q^{-1}(\theta_1)$, which is the same as $\pi^{3R}(\boldsymbol{\theta})$. All these reference priors can be derived from Berger and Bernardo (1992), and more directly from Theorem 1 of Datta and Ghosh (1995).

Based on the above calculations, it follows that the reference priors belong to a general class of priors of the form $\pi_a^{(1)}(\boldsymbol{\theta}) \propto \theta_{r+1}^{-a}Q^{-1}(\theta_1)$. In particular, the choices $a = r$, 1 and $r-1$ lead respectively to π^{2R}, π^{3R} and π_*^{3R}. In the original parameterization, the prior $\pi_a^{(1)}(\boldsymbol{\theta})$ reduces to

$$\pi_a^{(1)}(\beta_1,\ldots,\beta_r,\sigma) \propto \sigma^{-a}\Big(c_{11}\beta_1^2 + 2c_{12}\beta_1\beta_2 + c_{22}\beta_2^2\Big)^{-\frac{1}{2}}. \tag{3.2}$$

Also, one can generalize $\pi_a^{(1)}(\boldsymbol{\theta})$ further by replacing $Q^{-1}(\theta_1)$ with an arbitrary function $k(\theta_1)$ as done for example in Mendoza (1988) and Buonaccorsi and Gatsonis (1988). This will not be pursued here.

3.2. Probability matching priors

We first describe the probability matching criterion based on posterior quantiles in general notations. Let $\{Z_i;\, i \geq 1\}$ be a sequence of i.i.d. (possibly vector-valued) random variables with common pdf $f_\theta(z)$, where $\theta = (\theta_1,\ldots,\theta_p)^T$ belongs to some open subset of R^p, and θ_1 is the parameter of interest. We write $Z = (Z_1,\ldots,Z_N)^T$. Suppose $\theta_1^{1-\alpha}(\Pi, Z)$ is the $(1-\alpha)$th posterior quantile that satisfies $P^{\Pi}(\theta_1 \leq \theta_1^{1-\alpha}(\Pi,Z)|Z) = 1-\alpha+o(N^{-u})$ for some $u > 0$. We seek to characterize priors Π such that $P(\theta_1 \leq \theta_1^{1-\alpha}(\Pi,Z)|\theta) = 1-\alpha+o(N^{-u})$. Priors Π satisfying this property with $u = 1/2$ are called first order probability matching priors, while those satisfying this property with $u = 1$ are called second order probability matching priors. In typical applications, $o(N^{-1/2})$ and $o(N^{-1})$ are actually $O(N^{-1})$ and $O(N^{-3/2})$ respectively. As shown in Welch and Peers (1963), Stein (1985), Tibshirani (1989), Datta and J. K. Ghosh (1995a, b), Datta (1996), Mukerjee and Dey (1993), Mukerjee and Ghosh (1997), Sun and Ye (1996), and many others, such priors are obtained by solving certain differential equations.

For the problem at hand, asymptotics is based on N independent replications of the set-up given in (2.1). Since θ_1 is orthogonal to $(\theta_2,\ldots,\theta_{r+1})$ (see (2.5)), from Tibshirani (1989), the class of first order probability matching priors is characterized by

$$\pi^F(\boldsymbol{\theta}) \propto \theta_{r+1}^{-1}|\theta_2|Q^{-1}(\theta_1)q(\theta_2,\ldots,\theta_{r+1}), \tag{3.3}$$

where q is an arbitrary positive-valued function differentiable in its arguments. Jeffreys' prior as well as the other reference priors are all first order probability matching priors. Indeed, there are infinitely many such priors. To narrow down the selection of priors within this class, we now consider second order probability matching priors.

To this end, we first need to find the reparameterized likelihood $L(\theta_1,\ldots,\theta_r,\theta_{r+1})$. The original likelihood is given by

$$L(\beta_1,\ldots,\beta_r,\sigma) \propto \sigma^{-n}\exp[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \sum_{j=1}^{r} x_{ij}\beta_j)^2]. \tag{3.4}$$

By (2.3) and the definition of $g_j(\theta_1)$, the reparameterized likelihood is given by

$$L(\theta_1,\ldots,\theta_r,\theta_{r+1}) \propto \theta_{r+1}^{-n}\exp\Big[-\frac{1}{2\theta_{r+1}^2}\sum_{i=1}^n(y_i-\theta_2 h(\theta_1)\{\theta_1 x_{i1}+x_{i2} - \sum_{j=3}^r x_{ij}(a_{j1}\theta_1+a_{j2})\} - \sum_{j=3}^r x_{ij}\theta_j)^2\Big]. \quad (3.5)$$

From Mukerjee and Ghosh (1997), due to the orthogonality of θ_1 with $(\theta_2,\ldots,\theta_{r+1})$, the class of second order probability matching priors is characterized by solving

$$\frac{1}{6}q(\theta_2,\ldots,\theta_{r+1})\frac{\partial}{\partial\theta_1}(I_{11}^{-3/2}L_{1,1,1})+\sum_{v=2}^{r+1}\sum_{s=2}^{r+1}\frac{\partial}{\partial\theta_v}\{I_{11}^{-1/2}L_{11s}I^{sv}q(\theta_2,\ldots,\theta_{r+1})\}=0, \quad (3.6)$$

where $L_{1,1,1}=E\Big[\frac{\partial\log L}{\partial\theta_1}\Big]^3$, $L_{11s}=E\Big[\frac{\partial^3\log L}{\partial\theta_1^2\,\partial\theta_s}\Big]$, $s=2,\ldots,r+1$, and I^{sv} is the (s,v)th element of $\boldsymbol{I}^{-1}(\boldsymbol{\theta})$, the inverse of the Fisher Information matrix.

After much algebra, (3.6) simplifies to $\theta_2^{-1}\frac{\partial q}{\partial\theta_2}=\theta_{r+1}^{-1}\frac{\partial q}{\partial\theta_{r+1}}$ so that a general class of solutions is given by $q(\theta_2,\ldots,\theta_r,\theta_{r+1}) \propto h_1(\theta_2^2+\theta_{r+1}^2)h_2(\theta_3,\ldots,\theta_r)$, where h_1 and h_2 are both positive functions, differentiable in their arguments, but are otherwise arbitrary. It appears though from our limited simulation study in related problems that the choice of the second order probability matching prior does not matter much in practice, and we have therefore decided to take q as a constant. Thus we propose the second order probability matching prior

$$\pi^S(\boldsymbol{\theta}) \propto \theta_{r+1}^{-1}|\theta_2|Q^{-1}(\theta_1). \quad (3.7)$$

Both $\pi^J(\boldsymbol{\theta})$ and $\pi^S(\boldsymbol{\theta})$ belong to the general class of priors of the form $\pi_a^{(2)}(\boldsymbol{\theta}) \propto \theta_{r+1}^{-a}|\theta_2|Q^{-1}(\theta_1)$. The former has $a=r+1$ while the latter has $a=1$. Back to the original parameterization, this class of priors transforms to

$$\pi_a^{(2)}(\beta_1,\ldots,\beta_r,\sigma) \propto \sigma^{-a}. \quad (3.8)$$

The reference priors are first order but not second order matching priors. We now find conditions under which the joint posterior $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ is proper under the two proposed classes of priors.

Theorem 2. *Under the class of priors $\pi_a^{(1)}(\boldsymbol{\theta}) \propto \theta_{r+1}^{-a}Q^{-1}(\theta_1)$, the joint posterior $\pi_a^{(1)}(\boldsymbol{\theta}|\boldsymbol{y})$ is proper if and only if $n+a>r$. When $n+a>r$, the marginal posterior of θ_1 is*

$$\pi_a^{(1)}(\theta_1|\boldsymbol{y}) \propto Q^{-1}(\theta_1)\Big\{SSE+\frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1-\widehat{\beta}_1)^2}{Q(\theta_1)}\Big\}^{-\frac{n+a-r}{2}}.$$

Theorem 3. *Under the class of priors* $\pi_a^{(2)}(\boldsymbol{\theta}) \propto \frac{|\theta_2|}{\theta_{r+1}^a} Q^{-1}(\theta_1)$, *the joint posterior* $\pi_a^{(2)}(\boldsymbol{\theta}|\boldsymbol{y})$ *is proper if and only if* $n+a > r+1$. *When* $n+a > r+1$, *the marginal posterior of* θ_1 *is*

$$\frac{1}{Q^{\frac{1}{2}}(\theta_1)}\Big\{SSE + \frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1 - \widehat{\beta}_1)^2}{Q(\theta_1)} + nQ(\theta_1)\omega^2(\theta_1)\Big\}^{-\frac{n+a-r-1}{2}}$$
$$+2\omega(\theta_1)\Big\{SSE + \frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1 - \widehat{\beta}_1)^2}{Q(\theta_1)}\Big\}^{-\frac{n+a-r}{2}} \int_0^A \frac{dz}{(1+z^2)^{\frac{n+a-r+1}{2}}},$$

where $\omega(\theta_1) = \frac{\theta_1(c_{11}\widehat{\beta}_1 + c_{12}\widehat{\beta}_2) + c_{12}\widehat{\beta}_1 + c_{22}\widehat{\beta}_2}{Q(\theta_1)}$ *and* $A = Q^{\frac{1}{2}}(\theta_1)\,\omega(\theta_1)\,[SSE + \frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1 - \widehat{\beta}_1)^2}{Q(\theta_1)}]^{-1/2}$.

The proof of Theorem 3 is given in the appendix. The proof of Theorem 2 is simpler, and is omitted. It may be noted that for all these posteriors $\pi(\theta_1|\boldsymbol{y}) \to 0$ as $|\theta_1| \to \infty$. Thus, unlike the usual frequentist approach, credible intervals for θ_1 based on these posteriors avoid the problem of potentially being the entire real line.

4. Examples and Numerical Illustration

We first demonstrate how the slope-ratio and parallel-line assays are special cases of the general problem considered in this paper. Other examples mentioned in the introduction can be handled similarly. Later, we provide a numerical illustration of the performance of the different priors for a parallel-line assay example.

4.1. Slope-ratio assay

Consider an experiment where p doses $(x_{11}, \ldots, x_{1p})$ of a standard drug S are assayed m times and q doses $(x_{21}, \ldots, x_{2q})$ of a test drug T are assayed u times so that a set $\{Z_{1ik},\ i = 1, \ldots, p;\ k = 1, \ldots, m;\ Z_{2jk},\ j = 1, \ldots, q;\ k = 1, \ldots, u\}$ of $pm + qu$ observations is obtained.

The assumed model for slope-ratio assay is

$$Z_{1ik} = \alpha + \beta x_{1i} + \epsilon_{1ik}, \quad k = 1, \ldots, m;\ i = 1, \ldots, p,$$
$$Z_{2jk} = \alpha + \beta\rho x_{2j} + \epsilon_{2jk}, \quad k = 1, \ldots, u;\ j = 1, \ldots, q, \tag{4.1}$$

where the ϵ_{1ik} and ϵ_{2jk} are i.i.d $\mathcal{N}(0, \sigma^2)$. To see how this problem arises as a special case of the general regression model given in (2.1), let

$$\begin{array}{cccc}
Y_1 = Z_{111}, & Y_2 = Z_{112}, & \ldots, & Y_m = Z_{11m} \\
Y_{m+1} = Z_{121}, & Y_{m+2} = Z_{122}, & \ldots, & Y_{2m} = Z_{12m} \\
\vdots & \vdots & \vdots & \\
Y_{(p-1)m+1} = Z_{1p1}, & Y_{(p-1)m+2} = Z_{1p2}, & \ldots, & Y_{pm} = Z_{1pm} \\
Y_{pm+1} = Z_{211}, & Y_{pm+2} = Z_{212}, & \ldots, & Y_{pm+u} = Z_{21u} \\
\vdots & \vdots & \vdots & \\
Y_{pm+(q-1)u+1} = Z_{2q1}, & Y_{(p+q-1)n+2} = Z_{2q2}, & \ldots, & Y_{pm+qu} = Z_{2qu}.
\end{array}$$

We can then identify this model as a special case of the general regression model with $n = pm + qu$, $r = 3$, $\beta_1 = \beta\rho$, $\beta_2 = \beta$, and $\beta_3 = \alpha$. Also, the design matrix $\boldsymbol{X}$ is

$$\boldsymbol{X}^T = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & x_{21} & \cdots & x_{21} & \cdots & x_{2q} & \cdots & x_{2q} \\ x_{11} & \cdots & x_{11} & \cdots & x_{1p} & \cdots & x_{1p} & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 1 & \cdots & 1 & \cdots & 1 & \cdots & 1 & 1 & \cdots & 1 & \cdots & 1 & \cdots & 1 \end{bmatrix}. \tag{4.2}$$

In this example, $c_{11} = u\sum_{j=1}^q (x_{2j} - \bar{x}_2)^2 + \frac{pmqu}{pm+qu}\bar{x}_2^2$, $c_{12} = -\frac{pmqu}{pm+qu}\bar{x}_1\bar{x}_2$ and $c_{22} = m\sum_{i=1}^p (x_{1i} - \bar{x}_1)^2 + \frac{pmqu}{pm+qu}\bar{x}_1^2$.

4.2. Parallel-line assay

The set up is the same as in the previous subsection, but the assumed model here is

$$\begin{aligned} Z_{1ik} &= \alpha + \beta x_{1i} + \epsilon_{1ik}, \quad k = 1, \ldots, m;\ i = 1, \ldots, p, \\ Z_{2jk} &= \alpha + \beta(x_{2j} + \rho) + \epsilon_{2jk}, \quad k = 1, \ldots, u;\ j = 1, \ldots, q, \end{aligned} \tag{4.3}$$

where the ϵ_{1ik} and ϵ_{2jk} are i.i.d. $\mathcal{N}(0, \sigma^2)$. Once again, ρ is the parameter of interest.

In order to recognize this model as a special case of (2.1), we first represent the Z_{1ik}'s and Z_{2jk}'s as the $\boldsymbol{Y}$ vector as in the previous example. Also, as before, $n = pm + qu$, $r = 3$, $\beta_1 = \beta\rho$, $\beta_2 = \beta$ and $\beta_3 = \alpha$. The design matrix $\boldsymbol{X}$ in this case is

$$\boldsymbol{X}^T = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 1 & \cdots & 1 & \cdots & 1 & \cdots & 1 \\ x_{11} & \cdots & x_{11} & \cdots & x_{1p} & \cdots & x_{1p} & x_{21} & \cdots & x_{21} & \cdots & x_{2q} & \cdots & x_{2q} \\ 1 & \cdots & 1 & \cdots & 1 & \cdots & 1 & 1 & \cdots & 1 & \cdots & 1 & \cdots & 1 \end{bmatrix}. \tag{4.4}$$

Also, $c_{11} = \frac{pmqu}{pm+qu}$, $c_{12} = \frac{pmqu}{pm+qu}(\bar{x}_2 - \bar{x}_1)$ and $c_{22} = m\sum_{i=1}^p (x_{1i} - \bar{x}_1)^2 + u\sum_{j=1}^q (x_{2j} - \bar{x}_2)^2 + \frac{pmqu}{pm+qu}(\bar{x}_2 - \bar{x}_1)^2$.

In order to see the performance of the noninformative priors, we analyze one parallel-line assay dataset using Jeffreys' prior, the reference priors and the second order probability matching prior. The data given in Table 1, originally analyzed in Finney (1978), pertain to turbidimetric measurements on the growth response of *Lactobacillus leichmannii* to vitamin B_{12}.

Table 1. Responses in an assay of vitamin B_{12}.

Stimulus	Standard			Test		
Dose	-1.0	0.0	1.0	-1.0	0.0	1.0
	0.96	1.06	1.17	0.91	1.09	1.15
	0.91	1.07	1.14	0.93	1.04	1.15
	0.92	0.99	1.14	0.98	0.97	1.14
	0.76	0.86	1.13	0.96	1.06	1.16
	1.03	1.06	1.13	0.89	1.04	1.10
	0.93	1.02	1.15	1.01	1.02	1.15

In this case, $n = 36$, $r = 3$, $p = 3$, $q = 3$, $SSE = 0.0961$ and

$$\widehat{\boldsymbol{\beta}} = \begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \widehat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \widehat{\beta\rho} \\ \widehat{\beta} \\ \widehat{\alpha} \end{bmatrix} = \begin{bmatrix} 0.0178 \\ 0.105 \\ 1.0238 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 0 & 24 \end{bmatrix}.$$

Finney obtained Fieller's 95% confidence interval for this data as $(-0.193, 0.551)$. Table 2 provides the different equal-tailed 95% credible intervals. It appears from this table that the noninformatives priors all performed well and produced intervals shorter than Finney's. Although Jeffreys' and the two-group reference priors seem to perform best lengthwise, our simulation results indicated that the second-order matching prior and the one-at-a-time reference prior produced intervals with probability content closer to the target 95% frequentist coverage probability than the other priors.

Table 2. Posterior quantiles.

	$P_{0.025}$	$P_{0.975}$	Length
π^J	-0.165	0.523	0.688
π^{3R}	-0.181	0.529	0.710
π^{2R}	-0.170	0.517	0.687
π^{3R}_*	-0.176	0.523	0.699
π^S	-0.182	0.541	0.723

5. Concluding Remarks

One of the common evaluation procedures of objective Bayesian priors is to judge their frequentist performance. For instance, Mendoza (1990) compared the Bayesian credible interval based on the two-group reference prior with the classical interval given by Finney. Similar comparisons are provided in Philippe and Robert (1998), and Yin and Ghosh (2000). One of the objectives of this paper is to consider several objective priors, and we have considered an enlarged class comprising the reference priors as well as the matching priors. Based on our numerical findings for the slope-ratio and parallel-line assays, it appears that the posterior quantiles for the one-at-a-time reference prior are closest to those for the second order probability matching prior. Thus, our general recommendation is the use of one-at-a-time reference prior. It meets both the criteria of maximization of entropy, as well as matching the coverage probability of Bayesian credible intervals with the corresponding frequentist coverage probability. Also, this prior is also computationally attractive since evaluation of percentiles based on this prior requires only one-dimensional numerical integration.

Gleser and Hwang ((1987), Theorem 1) have shown that, based on a sample of arbitrary, but fixed size n, any confidence interval for θ_1 of finite expected length has frequentist coverage probability (taking the infimum over all points in the parameter space) zero. Thus, irrespective of which procedure is used, there will be certain regions of the parameter space where the frequentist coverage probability will be far off target. Based on some of our earlier work (e.g., Ghosh, Carlin and Srivastava (1995); Yin and Ghosh (2000)), we have found in the calibration and the Fieller-Creasy problems that the poor frequentist performance of Bayesian procedures occurs in the neighborhood of $\theta_1 = 0$. We want to emphasize, however, that our primary interest is the construction of Bayesian credible sets for θ_1. The coverage probability is then conditional on the data, and is not based on the infimum over all points in the parameter space. Thus, Bayesian credible sets can be constructed to provide any required coverage probability. Moreover, the asymptotic matching of the coverage probability of Bayesian credible sets with the corresponding frequentist coverage probability does not contradict Theorem 1 of Gleser and Hwang (1987). Indeed, these authors have pointed out (p.1361) that large sample $100(1-\alpha)\%$ confidence intervals of any finite length exist for θ_1 for every α in (0,1).

It is possible to generalize the findings of this paper to the more general location-scale families of distributions. To be specific, let $Y_1, \ldots, Y_n$ be independently distributed with pdf's $\sigma^{-1} f((y_i - \boldsymbol{x}_i^T \boldsymbol{\beta})/\sigma)$, $i = 1, \ldots, n$, where $f(z) = f(-z)$ for all real z. Then, after some calculations, one finds the Fisher information matrix

$$\boldsymbol{I}(\boldsymbol{\beta}, \sigma) = nc_1 \sigma^{-2} \begin{bmatrix} ((s_{jl})) & \boldsymbol{0}^t \\ \boldsymbol{0} & c_2/c_1 \end{bmatrix}, \tag{5.1}$$

where $c_1 = \int_{-\infty}^{\infty} (f'(z)/f(z))^2 f(z)dz$ and $c_2 = \int_{-\infty}^{\infty} z^2 (f'(z)/f(z))^2 f(z)dz - 1$. In the $N(0,1)$ case, $c_1 = 1$ and $c_2 = 2$. Thus the same orthogonal transformation derived in Section 2 works in this case, and the objective Bayesian analysis remains the same as before.

6. Appendix

Proof of Theorem 3. Since $\theta_2 = \beta_2 Q^{1/2}(\theta_1)$, the prior $\pi_a^2(\boldsymbol{\theta})$ transforms to $\pi_a^{(2)}(\theta_1, \beta_2, \beta_3, \beta_r, \sigma) \propto |\beta_2|\sigma^{-a}$. Accordingly, the joint posterior of $\theta_1, \beta_2, \ldots, \beta_r, \sigma$ is

$$\begin{aligned}
&\pi_a^{(2)}(\theta_1, \beta_2, \beta_3, \beta_r, \sigma|\boldsymbol{y})\\
&\propto \sigma^{-(n+a-r+2)}|\beta_2| \exp\Big[-\frac{1}{2\sigma^2}\{SSE + n(\beta_2\theta_1 - \hat{\beta}_1, \beta_2 - \hat{\beta}_2)\boldsymbol{C}(\beta_2\theta_1 - \hat{\beta}_1, \beta_2 - \hat{\beta}_2)^T\}\Big]\\
&\quad \times \exp\left(-\frac{n}{2\sigma^2}\mathbf{u}^T \boldsymbol{A}_{22}\mathbf{u}\right),
\end{aligned}$$

where $\boldsymbol{u}^T = (\beta_3 - \hat{\beta}_3, \ldots, \beta_r - \hat{\beta}_r) - (\beta_2\theta_1 - \hat{\beta}_1, \beta_2 - \hat{\beta}_2)\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}$. First, integrating with respect to $\beta_3, \ldots, \beta_r$, one gets

$$\begin{aligned}
&\pi_a^{(2)}(\theta_1, \beta_2, \sigma|\boldsymbol{y})\\
&\propto \sigma^{-(n+a)+r-2}|\beta_2| \exp\Big[-\frac{1}{2\sigma^2}\{SSE + n(\beta_2\theta_1 - \hat{\beta}_1, \beta_2 - \hat{\beta}_2)\boldsymbol{C}(\beta_2\theta_1 - \hat{\beta}_1, \beta_2 - \hat{\beta}_2)^T\}\Big].
\end{aligned}$$

When $n + a > r + 1$, integrating with respect to σ,

$$\pi_a^{(2)}(\theta_1, \beta_2|\boldsymbol{y}) \propto |\beta_2|\Big\{SSE + n(\beta_2\theta_1 - \hat{\beta}_1, \beta_2 - \hat{\beta}_2)\boldsymbol{C}(\beta_2\theta_1 - \hat{\beta}_1, \beta_2 - \hat{\beta}_2)^T\Big\}^{-\frac{n+a-r+1}{2}}.$$

Recalling the definition of $\omega(\theta_1)$ after Theorem 3, one has the identity

$$(\beta_2\theta_1 - \hat{\beta}_1, \beta_2 - \hat{\beta}_2)\boldsymbol{C}(\beta_2\theta_1 - \hat{\beta}_1, \beta_2 - \hat{\beta}_2)^T = Q(\theta_1)[\beta_2 - \omega(\theta_1)]^2 + \frac{|\boldsymbol{C}|(\hat{\beta}_2\theta_1 - \hat{\beta}_1)^2}{Q(\theta_1)}.$$

Hence, $\int_{-\infty}^{\infty} \pi(\theta_1, \beta_2|\boldsymbol{y})d\beta_2$ simplifies to

$$\begin{aligned}
&\int_{-\infty}^{\infty} |\beta_2|\Big\{SSE + nQ(\theta_1)\Big[\beta_2 - \omega(\theta_1)\Big]^2 + \frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1 - \widehat{\beta}_1)^2}{Q(\theta_1)}\Big\}^{-\frac{n+a-r+1}{2}} d\beta_2\\
&= \int_0^{\infty} \beta_2\Big\{SSE + \frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1 - \widehat{\beta}_1)^2}{Q(\theta_1)} + nQ(\theta_1)\Big[\beta_2 - \omega(\theta_1)\Big]^2\Big\}^{-\frac{n+a-r+1}{2}} d\beta_2\\
&\quad - \int_{-\infty}^{0} \beta_2\Big\{SSE + \frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1 - \widehat{\beta}_1)^2}{Q(\theta_1)} + nQ(\theta_1)\Big[\beta_2 - \omega(\theta_1)\Big]^2\Big\}^{-\frac{n+a-r+1}{2}} d\beta_2.
\end{aligned}$$

Writing $z = Q^{\frac{1}{2}}(\theta_1)[\beta_2 - \omega(\theta_1)]$, the first of these terms is

$$\frac{1}{\theta^{\frac{1}{2}}(\theta_1)}\frac{1}{n + a - r - 1}\Big\{SSE + \frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1 - \widehat{\beta}_1)^2}{Q(\theta_1)} + nQ(\theta_1)\omega^2(\theta_1)\Big\}^{-\frac{n+a-r-1}{2}}$$

$$+\omega(\theta_1)\Big\{SSE+\frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1-\widehat{\beta}_1)^2}{Q(\theta_1)}\Big\}^{-\frac{n+a-r}{2}}\int_{-A}^{\infty}\frac{dz}{(1+z^2)^{\frac{n+a-r+1}{2}}},$$

where $A=Q^{\frac{1}{2}}(\theta_1)\omega(\theta_1)[SSE+\frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1-\widehat{\beta}_1)^2}{Q(\theta_1)}]^{-1/2}$. Similarly, the second of the terms is

$$\frac{1}{Q^{\frac{1}{2}}(\theta_1)}\frac{1}{n+a-r-1}\Big\{SSE+\frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1-\widehat{\beta}_1)^2}{Q(\theta_1)}+nQ(\theta_1)\omega^2(\theta_1)\Big\}^{-\frac{n+a-r-1}{2}}$$
$$-\omega(\theta_1)\Big\{SSE+\frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1-\widehat{\beta}_1)^2}{Q(\theta_1)}\Big\}^{-\frac{n+a-r}{2}}\int_{A}^{\infty}\frac{dz}{(1+z^2)^{\frac{n+a-r+1}{2}}}.$$

Thus, the marginal posterior pdf of θ_1 is proportional to

$$\frac{2}{(n+a-r-1)Q^{\frac{1}{2}}(\theta_1)}\Big\{SSE+\frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1-\widehat{\beta}_1)^2}{Q(\theta_1)}+nQ(\theta_1)\omega^2(\theta_1)\Big\}^{-\frac{n+a-r-1}{2}}$$
$$+2\omega(\theta_1)\Big\{SSE+\frac{n|\boldsymbol{C}|(\widehat{\beta}_2\theta_1-\widehat{\beta}_1)^2}{Q(\theta_1)}\Big\}^{-\frac{n+a-r}{2}}\int_{0}^{A}\frac{dz}{(1+z^2)^{\frac{n+a-r+1}{2}}}.$$

In order to see the propriety of the posterior, it is more convenient to look at $\pi(\boldsymbol{\beta},\sigma|\boldsymbol{y})$:

$$\pi(\boldsymbol{\beta},\sigma|\boldsymbol{y})\propto\sigma^{-(n+a)}|\beta_2|\exp\Big[-\frac{1}{2\sigma^2}\Big\{SSE+(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})^T\boldsymbol{X}^T\boldsymbol{X}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})\Big\}\Big]. \quad (6.1)$$

Now, integrating out with respect to σ, one gets

$$\pi(\boldsymbol{\beta}|\boldsymbol{y})\propto|\beta_2|\Big\{SSE+(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})^T\boldsymbol{X}^T\boldsymbol{X}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})\Big\}^{-\frac{n+a-1}{2}}.$$

Recognizing the above pdf as proportional to $E\Big(|\beta_2|\Big|\boldsymbol{y}\Big)$, where $\boldsymbol{\beta}$ has a multivariate t-distribution with location parameter $\widehat{\boldsymbol{\beta}}$, scale matrix $\frac{SSE}{r}(\boldsymbol{X}^T\boldsymbol{X})^{-1}$ and degrees of freedom $n+a-r-1$, the propriety of the posterior follows when $n+a>r+1$. On the other hand, if $n+a\le r+1$, $\int\pi(\boldsymbol{\beta}|\boldsymbol{y})d\boldsymbol{\beta}=+\infty$. This proves the theorem.

Acknowledgements

Thanks are due to two reviewers for their very careful reading of the manuscript and making many helpful suggestions. It is a pleasure to acknowledge Gauri Sankar Datta for his meticulous reading of an earlier version of the paper, and catching many typos and errors. This research was partially supported by NSF Grants SBR-9810968 and SBR-9911485.

References

Berger, J. O. and Bernardo, J. M. (1992). On the development of reference priors (with discussion). In *Bayesian Statistics* **4** (Edited by J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith), 35-60. Oxford University Press, London.

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

Buonaccorsi, J. P. and Gatsonis, C. A. (1988). Bayesian inference for ratios of coefficients in a linear model. *Biometrics* **44**, 87-101.

Cox, D. R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *J. Roy. Statist. Soc. Ser. B* **49**, 1-39.

Darby, S. C. (1980). A Bayesian approach to parallel line bioassay. *Biometrika* **67**, 607-612.

Datta, G. S. (1996). On priors providing frequentist validity of Bayesian inference for multiple parameter functions. *Biometrika* **83**, 287-298.

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

Datta, G. S. and Ghosh, J. K. (1995b). Noninformative priors for maximal invariant parameter in group models. *Test* **4**, 95-114.

Datta, G. S. and Ghosh, M. (1995). Some remarks on noninformative priors. *J. Amer. Statist. Assoc.* **90**, 1357-1363.

Fieller, E. C. (1954). Some problems in interval estimation. *J. Roy. Statist. Soc. Ser. B* **16**, 175-185.

Finney, D. J. (1978). *Statistical Methods in Biological Assay.* Charles Griffin and Company, London.

Ghosh, M., Carlin, B. P. and Srivastava, M. S. (1995). Probability matching priors for linear calibration. *Test* **4**, 333-357.

Gleser, L. J. and Hwang, J. T. (1987). The non-existence of $100(1-\alpha)\%$ confidence sets of finite expected diameters in error-in-variable and related models. *Ann. Statist.* **15**, 1351-1362.

Hoadley, B. (1970). A Bayesian look at inverse linear regression. *J. Amer. Statist. Assoc.* **65**, 356-369.

Hunter, W. G. and Lamboy, W. F. (1981). A Bayesian analysis of the linear calibration problem. *Technometrics* **23**, 323-350.

Kim, P. T., Carter, E. M. and Hubert, J. J. (1991). Estimating relative potency using prior information. *Biometrics* **47**, 295-301.

Kim, P. T., Carter, E. M., Hubert, J. J. and Hand, K. J. (1993). Shrinkage estimators of relative potency. *J. Amer. Statist. Assoc.* **88**, 615-621.

Liseo, B. (1993). Elimination of nuisance parameters with reference priors. *Biometrika* **80**, 295-304.

Mendoza, M. (1988). Inference about the ratio of linear combinations of the coefficients in a multiple regression model. In *Bayesian Statistics* **3** (Edited by J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith), 705-711. Oxford University Press, London.

Mendoza, M. (1990). A Bayesian analysis of the slope ratio bioassay. *Biometrics* **46**, 1059-1069.

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.

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

Philippe, A. and Robert, C. (1998). A note on the confidence properties of reference priors for the calibration model. *Test* **7**, 147-160.

Stein, C. (1985). On coverage probability of confidence sets based on a prior distribution. In *Sequential Methods in Statistics* **16**, 485-514. Banach Center Publications, Warsaw.

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.

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

Yin, M. and Ghosh, M. (2000). Bayesian and Likelihood Inference for the Generalized Fieller-Creasy problem. In *Empirical Bayes and Likelihood Inference* (Edited by A. E. Ahmed and N. Reid), 121-139. Springer Verlag, New York.

University of Florida, 103 Griffin-Floyd Hall, P.O. Box 118545, Gainesville, FL 32611-8545, U.S.A.

E-mail: ghoshm@stat.ufl.edu

Analytical Sciences Inc., 2605 Meridian Parkway, Durharn, NC 27713, U.S.A.

E-mail: myin@asciences.com

Chung-Ang University, Dongjak-Gu Heukseol-Dong, Seoul, Korea.

E-mail: ykim@gogators.cau.ac.kr

(Received October 2001; accepted August 2002)