POSTERIOR PROPRIETY IN SOME HIERARCHICAL EXPONENTIAL FAMILY MODELS

EDWARD I. GEORGE and ZUOSHUN ZHANG [1]

The University of Texas at Austin and Quintiles Inc.

June 2000

SUMMARY

For Bayesian analysis of hierarchical models, it may be of interest to try and use improper priors to represent ignorance. In this article, we consider hierarchical exponential family models and obtain conditions on prior tail behavior that guarantee posterior propriety. In particular, we study three special cases, the Poisson-Gamma, the Binomial-Beta, and the Multinomial-Dirichlet models. Interestingly, in all three cases, flat priors produce improper posteriors. Although some data sets can yield proper posteriors under improper priors with the Poisson-Gamma model, we show that this cannot happen with the Binomial-Beta and Multinomial-Dirichlet models.

Keywords: Hierarchical Bayes Models, Improper Priors, Poisson-Gamma, Binomial-Beta, Multinomial-Dirichlet.

1 INTRODUCTION

Hierarchical models provide a useful approach to combining data from several similar populations. A typical setup for such models consists of pairs $(x_1, \theta_1), \ldots, (x_p, \theta_p)$ where x_i is the observed data from population i, and θ_i is a parameter which identifies $F(x_i \mid \theta_i)$, the distribution of x_i given

[1] Edward I. George holds the Ed and Molly Smith Chair and is Professor of Statistics, Department of MSIS, University of Texas, Austin, TX 78712-1175, egeorge@mail.utexas.edu. Zuoshun Zhang is a Biostatistician at Quintiles, Inc., 1801 Rockville Pike, Suite 300, Rockville, MD 20852-1366, zzhang@qrrg.quintiles.com.

θ_i. A hierarchical model is then obtained by using another model to describe the variation across the parameters $\theta_1, \ldots, \theta_p$. Such hierarchical models enable a statistical analysis that simultaneously uses all the data $x_1, \ldots, x_p$ to make inference about all the parameters $\theta_1, \ldots, \theta_p$.

In this paper, we focus on the following conditionally independent, three stage hierarchical model (Lindley and Smith 1972, Kass and Steffey 1989, George, Markov and Smith 1993, 1994):

1. Conditionally on $\theta_1, \ldots, \theta_p$, each x_i has distribution $F(x_i \mid \theta_i)$, and the data $x_1, \ldots, x_p$ are independent of each other and are independent of λ.

2. Conditionally on λ, the parameters $\theta_1, \ldots, \theta_p$ are iid with density $\pi(\theta \mid \lambda)$.

3. The hyperparameter λ has a density $\pi(\lambda)$, which may or may not be proper.

In particular, we consider setups for which $F(x \mid \theta)$ belongs to an exponential family (Brown 1986)

$$dF(x \mid \theta) = \exp[x \cdot \theta - \psi(\theta)]\, d\nu(x), \text{ for } \theta \in \Theta,$$

where $d\nu(x)$ is a fixed σ-finite measure on the Borel sets of R^k, Θ is a subset of the natural parameter space $N = \{\theta \in R^k : \int \exp(x \cdot \theta)\, d\nu(x) < \infty\}$, and $\psi(\theta) = \ln \int \exp(x \cdot \theta)\, d\nu(x)$ is the cumulant generating function. For such $F(x \mid \theta)$, we consider conjugate priors on θ (Diaconis and Ylvisaker 1979), namely

$$\pi(\theta \mid x_0, n_0) = \exp[x_0 \cdot \theta - n_0 \psi(\theta) - \varphi(x_0, n_0)] I_\Theta(\theta),$$

where $I_\Theta(\theta)$ is the indicator function and $\varphi(x_0, n_0) = \ln \int_\Theta \exp[x_0 \cdot \theta - n_0 \psi(\theta)]\, d\theta$, with the range of (x_0, n_0) properly chosen to make $\pi(\theta \mid x_0, n_0)$ a proper density. Finally, the hyperparameters (x_0, n_0) are given some prior $\pi(x_0, n_0)$, which may be proper or improper.

For any of these hierarchical exponential family setups, the only remaining specification issue is the choice of the hyperparameter prior $\pi(x_0, n_0)$. In many problems, there will be little available prior information about the hyperparameters, and it will be desirable to use a prior which is noninformative in some sense. Because improper priors are often considered for this purpose, it is of interest to know whether the resulting posterior $\pi(\theta_1, \ldots, \theta_p, x_0, n_0 \mid x_1, \ldots, x_p)$ is proper. This issue has increased in importance with the advent of Markov chain Monte Carlo posterior computation, which relies on such propriety to guarantee overall convergence of the simulated Markov chain, see George and Casella (1992), Hobert and Casella (1996, 1997) and Casella (1996). Berger and Strawderman (1996) and Hobert and Casella (1996) contain results for proper posteriors related to normal models. Natarajan and McCulloch (1995) contains posterior propriety results for a class of mixed models for Binomial data.

In this article, we obtain posterior propriety results for some special cases of the above hierarchical exponential family model, namely the Poisson-Gamma, the Binomial-Beta and Multinomial-Dirichlet hierarchical models. These results are conditions on the hyperparameter prior tail behavior that guarantee posterior propriety. For the Poisson-Gamma hierarchical model, these conditions allow for some improper priors to guarantee posterior propriety. However, for the Binomial-Beta and Multinomial-Dirichlet hierarchical models, it turns out that no improper prior can guarantee a proper posterior. In particular, under the traditional parameterizations, flat priors can never guarantee posterior propriety in any of our three hierarchical setups.

2 THE POISSON-GAMMA MODEL

We first consider the Poisson-Gamma model, where the data have Poisson distributions, and the conjugate prior on the parameters is the Gamma distribution. The three stages for the conditionally independent hierarchical model are then the following. For $i = 1, \ldots, p$,

$$\begin{aligned}
f(x_i \mid \theta_i) &= \frac{e^{-\theta_i}\theta_i^{x_i}}{x_i!},\ x_i = 0, 1, \ldots, \\
\pi(\theta_i \mid \alpha, \beta) &= \frac{\beta^\alpha}{\Gamma(\alpha)}\theta_i^{\alpha-1}e^{-\beta\theta_i},\ \theta_i > 0, \\
(\alpha, \beta) &\sim \pi(\alpha, \beta),\ \alpha > 0,\ \beta > 0
\end{aligned}$$

where $\pi(\alpha, \beta)$ is a (possibly improper) prior on the hyperparameters α and β.

The joint posterior for all the parameters is

$$\begin{aligned}
&\pi(\theta_1, \ldots, \theta_p, \alpha, \beta \mid x_1, \ldots, x_p) \\
&\sim \left[\left(\frac{\beta^\alpha}{\Gamma(\alpha)}\right)^p \prod_{i=1}^p \theta_i^{x_i+\alpha-1} \exp\left(-(\beta+1)\sum_{i=1}^p \theta_i\right)\right]\pi(\alpha, \beta).
\end{aligned}$$

Integrating out $\theta_1, \ldots, \theta_p$ yields

$$\begin{aligned}
&\pi(\alpha, \beta \mid x_1, \ldots, x_p) \\
&\sim \left[\left(\frac{\beta^\alpha}{\Gamma(\alpha)}\right)^p \prod_{i=1}^p \frac{\Gamma(x_i+\alpha)}{(\beta+1)^{x_i+\alpha}}\right]\pi(\alpha, \beta) \\
&= \left[\frac{\beta^{\alpha p}}{(\beta+1)^{\alpha p+s}} \prod_{i=1}^p (\alpha + x_i - 1)\cdots(\alpha+1)\alpha\right]\pi(\alpha, \beta) \\
&= \left[\frac{1}{(\beta+1)^s}\sum_{i=k}^s c(i)\alpha^i e^{-p\alpha\log(1+\frac{1}{\beta})}\right]\pi(\alpha, \beta),
\end{aligned}$$

where $s = \sum_{i=1}^p x_i$, k is the number of nonzero entries in the data vector $(x_1, \ldots, x_p)$, and $\sum_{i=k}^s c(i)\alpha^i := \prod_{i=1}^p (\alpha+x_i-1)\cdots(\alpha+1)\alpha$. We use the convention that the above polynomial is identically 1 when $k = s = 0$. Hence we have $0 \le k \le p,\ k \le s$, and $c(i) > 0, i = k, \ldots, s$.

Our aim is to find conditions on the hyperparameter prior $\pi(\alpha, \beta)$ such that the posterior

$$\pi(\theta_1, \ldots, \theta_p, \alpha, \beta \mid x_1, \ldots, x_p)$$

is proper, which is equivalent to

$$m(x_1, \ldots, x_p) = \int_0^\infty \int_0^\infty \pi(\alpha, \beta \mid x_1, \ldots, x_p)\, d\alpha\, d\beta < \infty,$$

for any nonnegative integers $x_1, , \ldots, x_p$. When $\pi(\alpha, \beta)$ is proper, the posterior is always proper. Hence we concentrate on improper priors.

We impose the following boundary conditions:

$$
\begin{aligned}
&(C1) \quad \pi(\alpha,\beta) = O(\alpha^{-a_0}\beta^{-b_0}) \quad as\ \alpha \to 0,\ \beta \to 0 \\
&(C2) \quad \pi(\alpha,\beta) = O(\alpha^{-a_0}\beta^{-b_\infty}) \quad as\ \alpha \to 0,\ \beta \to \infty \\
&(C3) \quad \pi(\alpha,\beta) = O(\alpha^{-a_\infty}\beta^{-b_0}) \quad as\ \alpha \to \infty,\ \beta \to 0 \\
&(C4) \quad \pi(\alpha,\beta) = O(\alpha^{-a_\infty}\beta^{-b_\infty}) \ as\ \alpha \to \infty,\ \beta \to \infty
\end{aligned}
$$

Our problem is now to determine the values of $a_0, b_0, a_\infty, b_\infty$ such that the posterior is proper. Let us state our results and then carry out the proofs.

Theorem 2.1 *Let the prior $\pi(\alpha, \beta)$ satisfy the conditions (C1)–(C4) and let the data $x_1, \ldots, x_p$ be fixed. Then the posterior*

$$\pi(\theta_1, \ldots, \theta_p, \alpha, \beta \mid x_1, \ldots, x_p)$$

is proper if and only if one of the following two groups of conditions hold:

(I) $a_0 < k+1, b_0 < 1, b_\infty + s > 1, a_\infty < k+1, a_\infty + b_\infty > 2;$

(II) $a_0 < k+1, b_0 < 1, b_\infty + s > 1, a_\infty \geq k+1.$

Corollary 2.2 *The posterior $\pi(\theta_1, \ldots, \theta_p, \alpha, \beta \mid x_1, \ldots, x_p)$ is proper for all values of $x_1, \ldots, x_p$ if and only if $a_0 < 1, b_0 < 1, b_\infty > 1, a_\infty + b_\infty > 2$.*

Corollary 2.3 *Let $\pi_1(\alpha) = \pi_2(\beta) = 1$ be the flat priors. Then the posterior $\pi(\theta_1, \ldots, \theta_p, \alpha, \beta|x_1, \ldots, x_p)$ is improper for all values of $x_1, \ldots, x_p$.*

Proof: First we partition the integral over $(0,\infty) \times (0,\infty)$ four parts,

$$
\begin{aligned}
&\left[\int_0^1\int_0^1 + \int_0^1\int_1^\infty + \int_1^\infty\int_0^1 + \int_1^\infty\int_1^\infty\right] \pi(\alpha,\beta \mid x_1,\ldots,x_p)\, d\alpha\, d\beta \\
&= A_1 + A_2 + A_3 + A_4.
\end{aligned}
$$

We proceed to derive necessary and sufficient conditions for each of A_1, A_2, A_3 and A_4 to be finite.

$$
\begin{aligned}
A_1 &\sim \int_0^1\int_0^1 \frac{\beta^{-b_0}}{(\beta+1)^s}\sum_{i=k}^{s} c(i)\alpha^{i-a_0}e^{-p\alpha\log(1+\frac{1}{\beta})}\, d\alpha\, d\beta \\
&\sim \int_0^1\int_0^{p\log(1+\frac{1}{\beta})}\sum_{i=k}^{s}\frac{\beta^{-b_0}}{(\beta+1)^s[p\log(1+\frac{1}{\beta})]^{i-a_0+1}}c(i)u^{i-a_0}e^{-u}\, du\, d\beta
\end{aligned}
$$

so that $A_1 < \infty$ if and only if $b_0 < 1$ and $a_0 < k+1$.

$$
\begin{aligned}
A_2 &\sim \int_0^1\int_1^\infty \frac{\beta^{-b_0}}{(\beta+1)^s}\sum_{i=k}^{s} c(i)\alpha^{i-a_\infty}e^{-p\alpha\log(1+\frac{1}{\beta})}\, d\alpha\, d\beta \\
&\sim \int_0^1\int_{p\log(1+\frac{1}{\beta})}^\infty\sum_{i=k}^{s}\frac{\beta^{-b_0}}{(\beta+1)^s[p\log(1+\frac{1}{\beta})]^{i-a_\infty+1}}c(i)u^{i-a_\infty}e^{-u}\, du\, d\beta
\end{aligned}
$$

so that $A_2 < \infty$ if and only if $b_0 < 1$.

$$\begin{aligned}
A_3 &\sim \int_1^\infty \int_0^1 \frac{\beta^{-b_\infty}}{(\beta+1)^s} \sum_{i=k}^{s} c(i)\alpha^{i-a_0} e^{-p\alpha \log(1+\frac{1}{\beta})}\, d\alpha\, d\beta \\
&\sim \int_1^\infty \int_0^1 \frac{1}{\beta^{b_\infty+s}} \sum_{i=k}^{s} c(i)\alpha^{i-a_0}\, d\alpha\, d\beta
\end{aligned}$$

so that $A_3 < \infty$ if and only if $b_\infty + s > 1$, and $a_0 < k+1$.

$$\begin{aligned}
A_4 &\sim \int_1^\infty \int_1^\infty \frac{\beta^{-b_\infty}}{(\beta+1)^s} \sum_{i=k}^{s} c(i)\alpha^{i-a_\infty} e^{-p\alpha \log(1+\frac{1}{\beta})}\, d\alpha\, d\beta \\
&\sim \int_1^\infty \int_{p\log(1+\frac{1}{\beta})}^\infty \sum_{i=k}^{s} \frac{c(i)u^{i-a_\infty}e^{-u}}{\beta^{s+b_\infty}[p\log(1+\frac{1}{\beta})]^{i-a_\infty+1}}\, du\, d\beta \\
&\sim \int_1^\infty \int_{\frac{p}{\beta}}^\infty \sum_{i=k}^{s} \frac{c(i)u^{i-a_\infty}e^{-u}}{\beta^{s+b_\infty+a_\infty-i-1}}\, du\, d\beta.
\end{aligned}$$

To establish conditions for A_4 to be finite, we consider three subcases.

Subcase 1: When $a_\infty < k+1$,

$$A_4 \sim \int_1^\infty \sum_{i=k}^{s} \frac{\Gamma(i-a_\infty+1)}{\beta^{s+b_\infty+a_\infty-i-1}}\, d\beta < \infty,$$

if and only if $b_\infty + a_\infty > 2$.

Subcase 2: When $a_\infty = k+1$,

$$\begin{aligned}
A_4 &\sim \int_1^\infty \int_{\frac{p}{\beta}}^\infty \sum_{i=k}^{s} c(i)\frac{u^{i-k-1}e^{-u}}{\beta^{s+b_\infty+k-i}}\, du\, d\beta \\
&\sim \int_1^\infty \frac{c(k)\log(\beta)}{\beta^{s+b_\infty}}\, d\beta < \infty,
\end{aligned}$$

if and only if $s + b_\infty > 1$.

Subcase 3: When $a_\infty > k+1$,

$$\begin{aligned}
A_4 &\sim \int_1^\infty \int_{\frac{p}{\beta}}^\infty \sum_{i=k}^{s} c(i)\frac{u^{i-a_\infty}e^{-u}}{\beta^{s+b_\infty+a_\infty+-i-1}}\, du\, d\beta \\
&\sim \int_1^\infty \int_{\frac{p}{\beta}}^\infty \frac{c(k)u^{k-a_\infty}e^{-u}}{\beta^{s+b_\infty+a_\infty-k-1}}\, d\beta\, du \\
&\sim \int_1^\infty \frac{1}{\beta^{s+b_\infty}}\, d\beta < \infty,
\end{aligned}$$

if and only if $s + b_\infty > 1$.

Combining the conditions of all the cases above, shows that $m(x_1, \ldots, x_p) = A_1 + A_2 + A_3 + A_4 < \infty$ if and only if either of the following two groups of conditions hold:

(I) $a_0 < k+1, b_0 < 1, b_\infty + s > 1, a_\infty < k+1, a_\infty + b_\infty > 2$;

(II) $a_0 < k+1, b_0 < 1, b_\infty + s > 1, a_\infty \geq k+1$.

which are exactly the conditions of the theorem. Combining them and considering all possible values of k and s from the data yields the two corollaries.

3 THE BETA-BINOMIAL MODEL

Next, we turn to the Binomial-Beta model, where the data have Binomial distributions, and the conjugate prior on the parameters is the Beta distribution. The three stages for the conditionally independent hierarchical model are then the following. For $i = 1, \ldots, p$,

$$
\begin{aligned}
&f(x_i \mid \theta_i) = \begin{pmatrix} n_i \\ x_i \end{pmatrix} \theta_i^{x_i}(1-\theta_i)^{n_i - x_i},\ x_i = 0, 1, \ldots, n_i, \\
&\pi(\theta_i \mid \alpha, \beta) = \frac{1}{B(\alpha,\beta)}\theta_i^{\alpha-1}(1-\theta_i)^{\beta-1}, \\
&(\alpha,\beta) \sim \pi(\alpha,\beta),\ \alpha > 0,\ \beta > 0,
\end{aligned}
$$

where $\pi(\alpha,\beta)$ is a (possibly improper) prior on the hyperparameters α and β.

The joint posterior for all the parameters is

$$
\begin{aligned}
&\pi(\theta_1, \ldots, \theta_p, \alpha, \beta \mid x_1, \ldots, x_p) \\
&\propto \left[\frac{1}{B(\alpha,\beta)^p} \prod_{i=1}^p \begin{pmatrix} n_i \\ x_i \end{pmatrix} \theta_i^{\alpha + x_i - 1}(1-\theta_i)^{\beta + n_i - x_i - 1} \right] \pi(\alpha,\beta).
\end{aligned}
$$

Integrating out $\theta_1, \ldots, \theta_p$ yields

$$
\begin{aligned}
&\pi(\alpha, \beta \mid x_1, \ldots, x_p) \\
&\propto \left[\prod_{i=1}^p \frac{B(\alpha + x_i, \beta + n_i - x_i)}{B(\alpha,\beta)} \right] \pi(\alpha,\beta) \\
&= \left[\prod_{i=1}^p \frac{\Gamma(\alpha + x_i)\Gamma(\beta + n_i - x_i)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(n_i + \alpha + \beta)} \right] \pi(\alpha,\beta) \\
&= \left[\prod_{i=1}^p \frac{(\alpha + x_i - 1)\cdots(\alpha+1)\alpha(\beta + n_i - x_i - 1)\cdots(\beta+1)\beta}{(n_i + \alpha + \beta - 1)\cdots(\alpha+\beta+1)(\alpha+\beta)} \right] \pi(\alpha,\beta) \\
&:= g(\alpha,\beta)\pi(\alpha,\beta).
\end{aligned}
$$

Our aim is to find conditions on the hyperparameter prior $\pi(\alpha,\beta)$ such that the posterior $\pi(\theta_1, \ldots, \theta_p, \alpha, \beta \mid x_1, \ldots, x_p)$ is a proper distribution. For any number c such that $0 < c < 1$, let $S_c = \{(\alpha,\beta) : 0 < c\beta < \alpha < \beta/c\}$ be a cone-like region in the first quadrant of the (α,β) plane. Let us state our results and then carry out the proofs.

Theorem 3.1 *(A) If the posterior $\pi(\theta_1, \ldots, \theta_p, \alpha, \beta \mid x_1, \ldots, x_p)$ is proper then*

$$\int_{S_c} \pi(\alpha,\beta)\, d\alpha\, d\beta < \infty$$

for any S_c with $0 < c < 1$.

(B) For any improper prior $\pi(\alpha, \beta)$, there exist values of $x_1, \ldots, x_p$ for which the posterior $\pi(\theta_1, \ldots, \theta_p, \alpha, \beta \mid x_1, \ldots, x_p)$ is improper.

Corollary 3.2 *Let $\pi(\alpha, \beta) = 1$ be the flat prior. Then the posterior $\pi(\theta_1, \ldots, \theta_p, \alpha, \beta \mid x_1, \ldots, x_p)$ is improper for all values of $x_1, \ldots, x_p$.*

PROOF: For any set S_c, it can be easily shown that there exist constants $0 < a < b$ such that $a \leq g(\alpha, \beta) \leq b$, where $(\alpha, \beta) \in S_c$. Hence a proper posterior implies $\int_{S_c} \pi(\alpha, \beta)\, d\alpha\, d\beta < \infty$, the conclusion of (A).

Next we prove (B). First we pick $x_1 = \cdots = x_p = 0$. Let $R_1 = \{(\alpha, \beta) : 0 < \alpha \leq \beta\}$ be the upper half of the first quadrant of the (α, β) plane. The function $g(\alpha, \beta)$ can be expressed as

$$g(\alpha, \beta) = \prod_{i=1}^{p} \frac{(\beta + n_i - 1) \cdots (\beta + 1)\beta}{(\alpha + \beta + n_i - 1) \cdots (\alpha + \beta + 1)(\alpha + \beta)},$$

from which we can deduce that $1/2^s \leq g(\alpha, \beta) \leq 1$, for $(\alpha, \beta) \in R_1$, where $s = \sum_{i=1}^{p} n_i$. It follows that

$$\int \int_{R_1} \pi(\alpha, \beta \mid x_1, \ldots, x_p)\, d\alpha\, d\beta < \infty$$

if and only if $\int \int_{R_1} \pi(\alpha, \beta)\, d\alpha\, d\beta < \infty$.

Similarly, we may pick $x_1 = n_1, \cdots, x_p = n_p$ and $R_2 = \{(\alpha, \beta) : \alpha \geq \beta > 0\}$, in which case,

$$\int \int_{R_2} \pi(\alpha, \beta \mid x_1, \ldots, x_p)\, d\alpha\, d\beta < \infty$$

if and only if $\int \int_{R_2} \pi(\alpha, \beta)\, d\alpha\, d\beta < \infty$.

Because an improper prior cannot yield finite integrals on both R_1 and R_2, it follows that we can always find values of $x_1, \ldots, x_p$ such that the posterior is improper.

4 THE MULTINOMIAL-DIRICHLET MODEL

Similar results as the last section hold for its natural generalization, the Multinomial-Dirichlet model (Good 1983, Leonard 1977). Here, the data have Multinomial distributions, and the conjugate prior on the parameters is a Dirichlet distribution. The three stages for the conditionally independent hierarchical model are then the following. For $i = 1, \ldots, p$,

$$f(x^{(i)} \mid \theta^{(i)}) = \frac{n_i!}{\prod_{j=1}^{t} x_j^{(i)}!} \prod_{j=1}^{t} (\theta_j^{(i)})^{x_j^{(i)}},$$

$$\pi(\theta^{(i)} \mid \beta) = \frac{\Gamma(\sum_{j=1}^{t} \beta_j)}{\prod_{j=1}^{t} \Gamma(\beta_j)} \prod_{j=1}^{t} (\theta_j^{(i)})^{\beta_j - 1},$$

$$\pi(\beta) = \pi(\beta_1, \ldots, \beta_t), \;\; \beta_j > 0, j = 1, \ldots, t$$

where $x_j^{(i)}$ are non-negative integers, $\sum_{j=1}^t x_j^{(i)} = n_i$, $0 < \theta_j^{(i)} < 1$, $\sum_{j=1}^t \theta_j^{(i)} = 1$, $x^{(i)} = (x_1^{(i)}, \ldots, x_t^{(i)})$, $\theta^{(i)} = (\theta_1^{(i)}, \ldots, \theta_t^{(i)})$ and $\beta = (\beta_1, \ldots, \beta_t)$.

Calculating the posterior as

$$\pi(\theta^{(1)}, \ldots, \theta^{(p)}, \beta \mid x^{(1)}, \ldots, x^{(p)}) \propto [\prod_{i=1}^{p} f(x^{(i)} \mid \theta^{(i)}) \pi(\theta^{(i)} \mid \beta)] \pi(\beta),$$

and defining

$$S_c = \{(\beta_1, \ldots, \beta_t) : 0 < c\beta_j < \beta_k < \beta_j / c, \;\; j, k = 1, \ldots, t\}$$

for any c such that $0 < c < 1$, the proofs of the following results are very similar to those in the previous section.

Theorem 4.1 *(A) If the posterior $\pi(\theta^{(1)}, \ldots, \theta^{(p)}, \beta \mid x^{(1)}, \ldots, x^{(p)})$ is proper, then*

$$\int_{S_c} \pi(\beta_1, \ldots, \beta_t)\, d\beta_1 \cdots d\beta_t < \infty,$$

for any possible S_c with $0 < c < 1$.
(B) For any improper prior $\pi(\beta_1, \ldots, \beta_t)$, there exist values of $x_1, \ldots, x_p$ for which the posterior $\pi(\beta_1, \ldots, \beta_t)$ is improper.

Corollary 4.2 Let $\pi(\beta) = 1$ be the flat prior. Then the posterior $\pi(\theta^{(1)}, \ldots, \theta^{(p)}, \beta \mid x^{(1)}, \ldots, x^{p})$ is improper for all values of $x_1, \ldots, x_p$.

ACKNOWLEDGEMENT

The authors are grateful to George Casella and James Hobert for helpful suggestions. This work was supported by NSF grants DMS 98.03756 and Texas ARP grants 003658-452,003658-690.

REFERENCES

Berger, J.O. and Strawderman W.E. (1996). Choice of hierarchical priors: admissibility in estimation of normal means. *Ann. Statist.* **24**, 931-951.

Brown, L.D. (1986). *Fundamentals of statistical exponential families.* IMS Lecture Notes Series (ed. S. Gupta). Hayward, California.

Casella, G. (1996). Statistical Inference and Monte Carlo Algorithms (with discussion) *Test.* **5**, 249-344.

Casella, G. and George, E.I. (1992). Explaining the Gibbs sampler. *The American Statistician.* **46**, 167-174.

Diaconis, P. and Ylvisaker D. (1979). Conjugate priors for exponential families. *Ann. Statist.* **7**, 269-281.

George, E.I., Makov, U.E., and Smith, A.F.M. (1993). Conjugate likelihood distributions. *Scand. J. Statist.* **20**, 147-156.

George, E.I., Makov, U.E., and Smith, A.F.M. (1994). Fully Bayesian hierarchical analysis for exponential families via Monte Carlo computation. *Aspects of Uncertainty—A Tribute to D.V.Lindley* (P.R.Freeman, A.F.M.Smith, eds). pp 181-199. Chichester: John Wiley and Sons.

Good, I.J. (1983). The robustness of a hierarchical model for Multinomials and contingency tables. *Statistical Inference, Data Analysis, and Robustness* (G. Box, T. Leonard, C.-F. Wu, eds). pp 191-211. New York: Academic Press.

Hobert, J.P. and Casella, G. (1996). The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *J. Amer. Statist. Assoc.* **91**, 1461-1473.

Hobert, J.P. and Casella, G. (1998). Functional compatibility, Markov chains and Gibbs sampling with improper posteriors. *J. Comp. Graph. Statist.* **7**, 42-60.

Leonard, T. (1977). A Bayesian approach to some Multinomial estimation and pretesting problems. *J. Amer. Statist. Assoc.* **72**, 869-874.

Lindley, D.V. and Smith, A.F.M. (1972). Bayes estimates for the linear model (with discussion). *J. R. Statist. Soc. B.* **34**, 1-41.

Kass, R.E. and Steffey, D. (1989). Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes). *J. Amer. Statist. Assoc.* **84**, 717-726.

Natarajan, R. and McCulloch, C.E. (1995). A note on the existence of the posterior distribution for a class of mixed models for Binomial responses. *Biometrika.* **82**, 639-643.