

# **Improper Priors for Models with Grouped or Partially Observed Data**

by

**Ranjini Natarajan**

Statistics Center  
Caldwell Hall and School of Operations Research  
Engineering and Theory Center  
Cornell University  
Ithaca, NY 14853

**Charles E. McCulloch**

Biometrics Unit  
and Statistics Center  
Warren Hall  
Cornell University  
Ithaca, NY 14853

BU-1237-M

May 1994

# Improper Priors for Models with Grouped or Partially Observed Data

Ranjini Natarajan <sup>1</sup> and Charles E McCulloch <sup>2</sup>

Cornell University

Ithaca, NY 14853

## Abstract

In multiparameter situations, the elicitation of prior information and subsequent formulation into a distribution can be a difficult task. In such situations, it is not unusual to consider analyses with conventionally chosen improper priors to reflect vague prior information. However, such prior distributions do not always lead to finite posterior distributions. In this work, we give conditions which guarantee the propriety of the resulting posterior distributions of the variance components for a class of improper priors for models with partially observed or grouped data. We study the normal linear mixed model where some or all of the dependent variables are unobserved. We illustrate the implications of our results for some popular incomplete data models: censored regression, grouped and truncated data.

*Keywords:* Gibbs sampler, latent data, mixed models, variance components.

---

<sup>1</sup>Statistics Center, Caldwell Hall and School of Operations Research, Engineering and Theory Center.

<sup>2</sup>Biometrics Unit and Statistics Center, Warren Hall.

# 1 Introduction

An important issue in any Bayesian analysis is the specification of a prior distribution. Formulation of prior knowledge into a distribution is difficult in multiparameter situations. In such situations, one may consider analyses with improper priors to reflect vague information. These priors can serve as reference priors for comparison with contextually appropriate *informative* priors [5]. Improper priors may also be used in a frequentist context due to the equivalence of flat prior Bayes and maximum likelihood estimation.

Some work has been done on studying the conditions under which certain classes of improper priors guarantee the existence of posterior distributions for specific models. Ibrahim and Laud [5] studied the propriety of the posterior distribution for the family of generalized linear models using Jeffrey's prior on the fixed effects parameters. They developed sufficient and necessary conditions for the existence of the posterior moment generating function and integrability of the posterior distribution. However, their conditions involve checking the existence of certain integrals, which can be quite a formidable task, especially in a multiparameter context. Hobert and Casella ([4]) studied some improper priors for the one parameter exponential family and hierarchical linear mixed models and developed conditions under which the posterior distributions of the variance components were guaranteed to exist.

We are interested in investigating the conditions under which a class of improper priors on the variance components leads to proper posterior distributions for models with incomplete data. We are specifically interested in situations where the data might be grouped (e.g., categorical variables), or only partially observed (e.g., censored or truncated data). We assume that the dependent variable of interest follows a linear model. We do not focus on the analytic tractability of the posterior distribution, but rather on its existence. We give very specific conditions under which the posterior distributions are guaranteed to exist for these models. The results developed

in this work also have implications for the use of Monte Carlo Markov chain methods (e.g. the Gibbs sampler) for such models. We show that the existence of the full conditional distributions does not guarantee the existence of the posterior distributions; thus, use of the Gibbs sampler in such situations can give seriously misleading results.

In section 2, we formulate a class of mixed models for partially observed data and state the lemmas that guarantee the existence of the posterior distribution of the variance components for a specific family of prior distributions. We illustrate the implications of these lemmas on censored regression data, Tobit regression and truncated normal data. In section 3, we formulate a class of random effects models for grouped data and study the integrability of the posterior distributions. We discuss the implications of the results developed in section 3 on binary data. Finally, in section 4 we provide a summary discussion.

## 2 Hierarchical Linear Mixed Models for Partially Observed Data

### 2.1 Known Fixed Effects:

We start with a hierarchical model:

$$\begin{aligned}
 Y | \beta, u, \theta_0 &\sim N_n(X\beta + Zu, \theta_0 I) \\
 u | D &\sim N_q(0, D)
 \end{aligned}
 \tag{2.1.1}$$

where  $Y \in \mathfrak{R}^{n \times 1}$  is the data vector which is only partially observed,  $X \in \mathfrak{R}^{n \times p}$  is the design matrix and  $Z \in \mathfrak{R}^{n \times q}$  is the incidence matrix corresponding to the random effects vector  $u \in \mathfrak{R}^{q \times 1}$ .

We first consider the case where the fixed effects parameters  $\beta$  are known. We assume that  $u$  can be partitioned in the following way:  $u = (u_1, u_2, \dots, u_r)'$  where  $u_k \in \mathfrak{R}^{q_k \times 1}$  and we make the simplifying assumption that  $D = \text{diag}(\theta_1 I_{q_1}, \dots, \theta_r I_{q_r})$  where  $q = \sum_{k=1}^r q_k$ . Denote  $q_0 = n$ .

We say a component of  $Y$ ,  $Y_i$  is unobserved if the only data information available about

$Y_i$  is that it lies in some interval  $(a_i^*, b_i^*)$ , where  $-\infty \leq a_i^* < b_i^* \leq \infty$ , and at least one of  $a_i^*$  or  $b_i^*$  is finite. Such applications arise when “experimental conditions or measuring devices permit sample points to be trapped only within specified limits” ([9]) as in censored or truncated data. Define  $m (> 0)$  to be the number of  $Y$  that are observed exactly. We reorder the  $Y$  such that  $Y = (Y^{(1)'}, Y^{(2)'})'$ , where  $Y^{(1)} \in \Re^{m \times 1}$  is observed exactly and  $Y^{(2)} \in \Re^{(n-m) \times 1}$  denotes the components of  $Y$  that are unobserved. Let  $X^{(1)} \in \Re^{m \times p}$ ,  $Z^{(1)} \in \Re^{m \times q}$  denote the partition of the design and incidence matrix that correspond to  $Y^{(1)}$ , and  $X^{(2)} \in \Re^{(n-m) \times p}$ ,  $Z^{(2)} \in \Re^{(n-m) \times q}$  denote the partition corresponding to  $Y^{(2)}$ . We assume that  $(Z^{(1)' } Z^{(1)})^{-1}$  and  $(X^{(1)' } X^{(1)})^{-1}$  exist. The probability distribution of the observed data conditional on  $u$ , is given by:

$$[Y^{(1)}, Y^{(2)} \in \mathcal{C} | u] = \prod_{i=1}^m \phi\left(\frac{y_i - x_i' \beta - z_i' u}{\sqrt{\theta_0}}\right) \prod_{i=m+1}^n \left(\Phi\left(\frac{b_i^* - x_i' \beta - z_i' u}{\sqrt{\theta_0}}\right) - \Phi\left(\frac{a_i^* - x_i' \beta - z_i' u}{\sqrt{\theta_0}}\right)\right)$$

where  $[\cdot]$  denote densities,  $\mathcal{C} = (a_{m+1}^*, b_{m+1}^*) \times \dots \times (a_n^*, b_n^*)$ ,  $x_i'$  and  $z_i'$  are the  $i^{th}$  rows of  $X$  and  $Z$  respectively and  $\phi(\cdot)$  and  $\Phi(\cdot)$  are the standard normal probability density function and cumulative distribution function respectively.

We consider the following class of improper priors on the variance components  $(\theta_0, \theta_1, \dots, \theta_r)$ :

$$\pi(\theta_k | a_k) \propto \frac{1}{\theta_k^{a_k+1}}, \quad (k = 0, 1, \dots, r) \quad (2.1.2)$$

where  $a_k$ ,  $(k = 0, 1, \dots, r)$  are prespecified constants characterizing the prior distribution. Note that when  $a_k = 0$ ,  $(k = 0, 1, \dots, r)$  we have the usual non-informative prior on a variance component ([1]).

While using a data augmentation approach such as the Gibbs sampler, it is typical to generate the unobserved  $Y$  from its full conditional specification ([2], [10]). For the case when  $a_k > -\frac{q_k}{2}$ ,  $(k = 0, 1, \dots, r)$  Hobert and Casella ([4]) derive the full conditionals for  $\theta_0, \theta_1, \dots, \theta_r$  and  $u$ . The full conditional for  $Y^{(2)}$  can easily be shown to be that of a normal distribution constrained to lie within the set  $\mathcal{C}$ . Thus, implementation of the Gibbs sampler to compute posterior estimates

for this model is straightforward and certainly attractive! However, the improper priors described earlier do not always lead to finite posterior distributions and we now state the lemmas that guarantee the propriety of the posterior distribution of the variance components for this class of models. Proofs are given in appendix 1.

*Lemma 1:* Given the prior specifications in (2.1.2) and the model in (2.1.1), the joint posterior distribution of  $\theta = (\theta_0, \theta_1, \dots, \theta_r)$  will exist *iff* the following integral converges:

$$\int_{\theta} \int_{\mathcal{C}} \frac{\exp(-\frac{1}{2}(Y - X\beta)'(\theta_0 I_n + ZDZ')^{-1}(Y - X\beta)) dY^{(2)} \prod_{k=0}^r d\theta_k}{|\theta_0 I_n + ZDZ'|^{1/2} \prod_{k=0}^r \theta_k^{a_k+1}} \quad (2.1.3)$$

*Lemma 2:* Given the priors in (2.1.2) and the model in (2.1.1), the following conditions are necessary and sufficient for the posterior distribution of  $\theta$  to exist.

- (i)  $a_k < 0$ , ( $k = 1, \dots, r$ )
- (ii)  $a_k + \frac{q_k}{2} > 0$ , ( $k = 1, \dots, r$ )
- (iii)  $m + 2 \sum_{k=0}^r a_k > 0$

These results are very similar to those obtained by Hobert and Casella ([4]), the only difference being that the dimension of the  $Y$  vector in their case, is replaced by the number of  $Y$  that is exactly observed in our case. Also, it is interesting to note that the classic non-informative prior on the variance components  $\theta_1, \theta_2, \dots, \theta_r$ , i.e., ( $a_k = 0$ ,  $k = 1, \dots, r$ ) does not lead to proper posterior distributions for this model even if the fixed effects parameters are assumed known.

## 2.2 Unknown Fixed Effects:

We now consider the case where the fixed effects parameters  $\beta$  are unknown. We postulate the following hierarchical model:

$$Y | \beta, u, \theta_0 \sim N_n(X\beta + Zu, \theta_0 I)$$

$$u | D \sim N_q(0, D)$$

$$\pi(\beta) = 1 \quad -\infty < \beta < \infty \quad (2.2.4)$$

where  $\beta \in \mathfrak{R}^{n \times p}$  is the vector of unknown fixed effects and  $\pi(\beta)$  corresponds to the usual non-informative prior for a location parameter. We assume that  $p + q < m$ . We consider the prior specifications in (2.1.2). Again, while implementing the Gibbs sampler for this situation, we simply generate the fixed effects parameters  $\beta$  from its full conditional specification, which can be shown to be a  $p$ -variate normal distribution.

Before we state the lemmas that guarantee proper posterior distributions, we define some quantities of interest. Let  $P = (I_n - X(X'X)^{-1}X')$  and  $t = \text{rank}(Z'PZ)$ . Also define  $P^{(1)} = (I_m - X^{(1)}(X^{(1)'}X^{(1)})^{-1}X^{(1)'})$  and  $t^{(1)} = \text{rank}(Z^{(1)'}P^{(1)}Z^{(1)})$ . Proofs of the lemmas are given in appendix 2.

*Lemma 3: Necessity* Given the prior specifications in (2.1.2) and the model in (2.2.4), there are two cases:

Case 1: If  $t = q$  or  $r = 1$ , then conditions (iv), (v) and (vi) below are necessary for the propriety of the posterior distribution.

Case 2: If  $t < q$  and  $r > 1$ , then conditions (iv), (v') and (vi) are necessary for the existence of the posterior distribution.

$$(iv) \ a_k < 0, \ k = 1, \dots, r$$

$$(v) \ q_k > q - t - 2a_k, \ k = 1, \dots, r$$

$$(v') \ q_k > -2a_k, \ k = 1, \dots, r$$

$$(vi) \ m + 2 \sum_{k=0}^r a_k - p > 0$$

*Lemma 4: Sufficiency* Given the prior specifications in (2.1.2) and the model in (2.2.4), conditions (iv), (v) and (vi) are sufficient, with  $t$  replaced by  $t^{(1)}$ .

Again, it is interesting to note that the classic non-informative prior on the variance components does not lead to proper posterior distributions, thus reiterating the message that while imposing improper priors, one must first verify if the resulting posterior distribution exists! We will now give examples of some popular incomplete data models that fall within the framework developed in this paper.

### 2.3 Examples

**Censored Regression Data with a Known Censoring Time** We consider models in which some of the observations are right censored. This can occur when the response is a waiting time and a typical member of the population of physical or biological units is observed till an event of interest (or censoring) occurs. Such data arise in medical applications (time till the first tumor), reliability (repairable systems and software reliability) or labor economics (period of successive layoffs). The observed data is the pair  $(\min(Y_i, a_i^*), I(Y_i \leq a_i^*), i = 1, \dots, n)$  where  $I$  is the indicator function and  $a_i^*$  ( $i = 1, \dots, n$ ), are known constants. The censored observations are considered to be latent variables or unobserved data. The response vector is assumed to satisfy the mixed model defined in (2.1.1) or (2.2.4). In this context,  $m (> 0)$ , is the number of exact/uncensored observations. The lemmas stated earlier provide conditions on the number of uncensored observations and the hyperparameters of the improper prior distributions that need to be satisfied in order to ensure proper posterior distributions.

**Tobit Regression** Data available to economists are often incomplete in one way or another. As an example, sometimes dependent variables can be observed only in a limited range, as in the case of Tobin's model of the demand for consumer durables ([7]). In this model, the utility maximizing amount of expenditures on a durable good ( $Y$ ) satisfies a linear regression. However, the problem for a household is that there is some minimum level of expenditure required to purchase a durable

good, say  $c$ . Thus, for any household, the amount of expenditure actually observed is:

$$\begin{aligned} Y^* &= Y \text{ if } Y > c \\ &= 0 \text{ otherwise} \end{aligned}$$

i.e., the desired level or 0 which simply reflects that  $Y \leq c$ . This model is known in the econometrics literature as the Tobit model. To put this model in the framework developed in section 2, we can define  $Y^{(1)} = \{Y : Y > c\}$  and  $Y^{(2)} = \{Y : Y \leq c\}$ . In this context,  $m (> 0)$  is the number of  $Y$  that exceed  $c$ . We postulate a mixed model as in (2.2.4) for the dependent variable  $Y$ . Again, the lemmas stated earlier provide a word of caution on the use of improper priors for this model. In particular, they specify restrictions on the number of  $Y$  that are exactly observed and the parameters of the improper prior distribution.

**Truncated Bivariate Normal Data** We consider a bivariate normal process  $(X_i, Y_i), i = 1, \dots, n$ , where some of the  $Y_i$  are not observed. Such data might arise in paired survival time studies ([10]) where observation  $Y_i$  of the second of the paired patients is terminated when the first of the pair dies; so that  $Y_i$  is observed only if  $Y_i \leq X_i$ . Thus, the observed data is the pair  $(X_i, W_i)$  where  $W_i = Y_i$  if  $Y_i \leq X_i$ ; otherwise we observe  $(X_i, *)$  where  $W_i = *$  indicates that  $Y_i > X_i$ . We reorder the observations so that  $(X_i, W_i), i = 1, \dots, l$ , are observed exactly and the remaining  $(n-l)$  pairs are  $(X_i, *)$ ,  $i = l+1, \dots, n$ . We can define  $Y^{(1)} = (X_1, Y_1, \dots, X_l, Y_l, X_{l+1}, X_{l+2}, \dots, X_n)$  and  $Y^{(2)} = (Y_{l+1}, \dots, Y_n)$ . Thus, in this context,  $m (> 0)$  is equal to  $n + l$ . Denote  $Y^* = (X_1, Y_1, X_2, Y_2, \dots, X_n, Y_n) \in \mathfrak{R}^{2n \times 1}$ . We postulate the following model:

$$\begin{aligned} Y^* &= \mu + (I_n \otimes \mathbf{1}_2) u + \epsilon \\ u &\sim N_n(0, \theta_1 I) \\ \epsilon &\sim N_n(0, \theta_0 I) \end{aligned} \tag{2.3.5}$$

where  $\mathbf{1}_2$  is the  $2 \times 1$  vector of ones and  $\otimes$  is the direct product. (2.3.5) simply states that the pair  $(X_i, Y_i)$  are independently distributed as a bivariate normal distribution with unknown mean  $\mu = (\mu_1, \mu_2)$  and variance-covariance matrix given by:

$$\begin{pmatrix} \theta_0 + \theta_1 & \theta_1 \\ \theta_1 & \theta_0 + \theta_1 \end{pmatrix}$$

This is the mixed model described in (2.2.4), with design matrix  $X = \mathbf{1}_n \otimes I_2$ , incidence matrix  $Z = I_n \otimes \mathbf{1}_2$ ,  $q = n$  and  $r = 1$ . Kuo and Smith ([10]) outlined a Bayesian analysis of this model using an inverse Wishart prior on the variance-covariance matrix. By lemmas 3 and 4, we claim that the posterior distribution of the variance components exists for the class of priors in (2.1.1) iff  $a_1 < 0$ ,  $n - 1 + 2a_1 > 0$  and  $n + l + 2 \sum_{k=0}^1 a_k - 2 > 0$ .

We now study the existence of proper posterior distributions of the variance components for a class of models for grouped data.

### 3 Random Effects Model for Grouped Data

Our model is a random effects threshold model where  $Y$  represents an unobserved, continuous variable and we observe only  $W_i = I_{(Y_i > 0)}$ . We postulate the following model for the underlying variable  $Y$ :

$$\begin{aligned} Y &= \sum_{k=1}^r Z_k u_k + \epsilon \\ u_k &\sim N_{q_k}(0, \theta_k I), \quad k = 1, \dots, r \\ \epsilon &\sim N_n(0, I) \end{aligned} \tag{3.0.6}$$

We can assume that the error variance in (3.0.6) is one without loss of generality ([3]). This is a normal regression problem where the data is grouped. McCulloch ([12]) studied a mixed model

version of the above and computed maximum likelihood estimators of the variance components using the E M algorithm.

We consider the class of priors defined in (2.1.2) on the variance components. These priors do not necessarily lead to proper posterior distributions for this model and we state the lemmas that guarantee the existence of the posterior distribution of  $\theta = (\theta_1, \dots, \theta_r)$ . Proofs are given in appendix 3. Re-order the  $Y$  such that  $(Y_1 < 0, Y_2 < 0, \dots, Y_l < 0, Y_{l+1} > 0, \dots, Y_n > 0)$  where  $l \in (0, 1, \dots, n)$ . Define  $Z_k^* = P Z_k, \forall k = 1, \dots, r$ , where  $P = \text{diag}(-I_l, I_{n-l})$ .

*Lemma 5:* Given the priors in (2.1.2) and the model in (3.0.6), the posterior distribution of  $\theta$  exists iff the following integral converges.

$$\int_{\theta} \int_{\mathfrak{R}^q} \frac{P(T > \theta_1^{1/2} Z_1^* v_1 + \dots + \theta_r^{1/2} Z_r^* v_r) e^{-\frac{v'v}{2}} dv \prod_{k=1}^r d\theta_k}{\prod_{k=1}^r \theta_k^{a_k+1}} \quad (3.0.7)$$

where  $T \sim N_n(0, I)$ ,  $v_k \in \mathfrak{R}^{q_k \times 1}$ , ( $k = 1, \dots, r$ ) and  $q = \sum_{k=1}^r q_k$ .

*Lemma 6: (Necessity)* Given the priors in (2.1.2) and the model in (3.0.6), the following conditions are necessary for the propriety of the posterior distribution.

- (i)  $a_k < 0, k = 1, \dots, r$
- (ii)  $a_k + \frac{q_k}{2} > 0, k = 1, \dots, r$
- (iii) The set  $(v_k \in \mathfrak{R}^{q_k \times 1} : Z_k^* v_k \leq 0)$  must be empty with probability 1,  $\forall (k = 1, \dots, r)$ .

*Lemma 7: (Sufficiency)* Given the priors in (2.1.2) and the model in (3.0.6), the following conditions are sufficient for the propriety of the posterior distribution.

- (i)  $-\frac{1}{2} < a_k < 0, k = 1, \dots, r$
- (ii)  $\forall v \in \mathfrak{R}^q \times 1 = (v_1 \in \mathfrak{R}^{q_1 \times 1}, \dots, v_r \in \mathfrak{R}^{q_r \times 1}) \neq 0, \exists j_v \in (1, \dots, n)$  such that:

$$z_{k,j_v}^{*'} v_k > 0, \forall k = 1, \dots, r$$

where  $z_{k,j_v}^{*'}$  is the  $j_v^{\text{th}}$  row of  $Z_k^*$ .

Again, we note that the non-informative prior on the variance components does not lead to finite posterior distributions. These results have implications for the logit link as well, since the logit link can be very closely approximated by the probit link ([6]). Zeger and Karim ([8]) used a logit normal regression to analyze the salamander mating data set published in McCullagh and Nelder ([11]). They considered the following non-informative prior on the variance-covariance matrix  $D \in \mathfrak{R}^{m \times m}$  of the random effects:  $\pi(D) = |D|^{-\frac{m+1}{2}}$ . They advocate use of the Gibbs sampler to obtain Bayesian estimates of the parameters. For the special case  $m = 1$ , the above prior reduces to the usual non-informative prior on a single variance component ([1]), i.e.,  $a_k = 0$  in (2.1.2). Lemmas 6 and 7 suggest that the posterior distributions do not exist for this choice of prior distribution. We now consider a simple example to illustrate the implications of the conditions developed in lemmas 6 and 7.

### 3.1 Binary Data

Let  $W_1, W_2, \dots, W_n$  be  $n$  correlated Bernoulli variates with probability of success  $p_i$ ,  $i = 1, \dots, n$ . We introduce  $n$  latent/unobserved variables  $Y_1, Y_2, \dots, Y_n$  where the  $Y$  satisfy the random effects model (3.0.6). The observed data  $W_i$  arises in the following manner:

$$\begin{aligned} W_i &= 1 \text{ if } Y_i > 0, \quad i = 1, \dots, n \\ &= 0 \text{ otherwise} \end{aligned} \tag{3.1.8}$$

We consider a simple, special structure of the incidence matrix, namely,  $Z = (I_k \otimes \mathbf{1}_n)$  and  $r = 1$ . This corresponds to  $n$  repeat observations being taken on each of the  $k$  levels of the random effect and  $Z$  is simply a group indicator matrix. Condition (ii) of lemma 7 and condition (iii) of lemma 6 have a very interesting interpretation for this special structure of the incidence matrix. Since  $z_j^{*'} = (0, 0, \dots, \underbrace{1}_j, \dots, 0) \times (2W_i - 1)$ , condition (ii) of lemma 7 translates into “if there is at least one success and failure in each group” and condition (i) holds, the posterior

distributions exist; while condition (iii) from lemma 6 translates into “if there is any group for which there are only successes or only failures, the posterior distribution does not exist”. This interpretation coincides with the likelihoodist claim for the purely fixed effects probit model, which states that if there is a group with only successes or failures, the maximum likelihood estimators of the fixed effects do not exist.

## 4 Summary

This paper is meant to provide a word of caution while using improper priors for incomplete data models. Such priors do not always lead to proper posterior distributions and it is imperative that one must check the existence of the resulting posterior distributions while using them. We derive conditions under which the posterior distribution of the variance components are guaranteed to exist for a class of conditionally independent hierarchical models with latent data. The results developed in this paper have important implications for the use of a Monte Carlo Markov Chain method such as the Gibbs sampler for a Bayesian analysis of incomplete data models. It is quite common, while using the Gibbs sampler, to impose non-informative priors on the parameters. However, we have shown that for the class of incomplete data models considered in this paper, these priors do not lead to proper posterior distributions and it is unclear as to what the sampler converges to.

**Acknowledgements** I wish to thank Prof Samorodnitsky for his invaluable help and boundless patience.

## 5 Appendix 1:

### 5.1 Lemma 1:

The posterior distribution of  $\theta$  is given by:

$$[\theta | Y^{(1)}, Y^{(2)} \in \mathcal{C}] = \frac{L(\theta) \prod_{k=0}^r \pi(\theta_k | a_k)}{\underbrace{\int_{\mathcal{C}} L(\theta) \prod_{k=0}^r \pi(\theta_k | a_k) d\theta_k}_{m(Y^{(1)}, Y^{(2)} \in \mathcal{C})}} \quad (5.1.9)$$

where  $m(\cdot)$  is the marginal distribution of the observed data and  $L(\cdot)$  is the observed data likelihood.

It is clear that the posterior distribution of the variance components will exist *iff*  $m(\cdot)$  converges, i.e., the integral in (5.1.9) converges. The observed data likelihood is given by:

$$\begin{aligned} L(\theta | Y^{(1)}, Y^{(2)} \in \mathcal{C}) &= \int_{\mathcal{C}} \int_u [Y, u | \theta] du dY^{(2)} \\ &\propto \int_{\mathcal{C}} \int_u \frac{\exp(-\frac{1}{2\theta_0}(Y - X\beta - Zu)'(Y - X\beta - Zu)) \exp(-\frac{1}{2}u'D^{-1}u) du dY^{(2)}}{\theta_0^{\frac{n}{2}} |D|^{1/2}} \end{aligned}$$

Completing a multivariate square and integrating over  $u$ , we have:

$$L(\theta | Y^{(1)}, Y^{(2)} \in \mathcal{C}) \propto \int_{\mathcal{C}} \frac{\exp(-\frac{1}{2\theta_0}(Y - X\beta)'(I_n - Z(Z'Z + \theta_0 D^{-1})^{-1}Z')(Y - X\beta)) dY^{(2)}}{\theta_0^{\frac{(n-q)}{2}} |D|^{1/2} |Z'Z + \theta_0 D^{-1}|^{1/2}}$$

Since ([13])

$$(\theta_0^{-1} I_n - \theta_0^{-2} Z(D^{-1} + \theta_0^{-1} Z'Z)^{-1}Z') = (\theta_0 I_n + Z D Z')^{-1} \text{ and}$$

$$\theta_0^{n-q} |D| |\theta_0 D^{-1} + Z'Z| = |\theta_0 I_n + Z D Z'|$$

we have:

$$L(\theta | Y^{(1)}, Y^{(2)} \in \mathcal{C}) \propto \int_{\mathcal{C}} \frac{\exp(-\frac{1}{2}(Y - X\beta)'(\theta_0 I_n + Z D Z')^{-1}(Y - X\beta)) dY^{(2)}}{|\theta_0 I_n + Z D Z'|^{1/2}}$$

From (5.1.9),

$$m(Y^{(1)}, Y^{(2)} \in \mathcal{C}) \propto \int_{\theta} \int_{\mathcal{C}} \frac{e^{-\frac{1}{2}(Y - X\beta)'(\theta_0 I_n + Z D Z')^{-1}(Y - X\beta)} dY^{(2)} \prod_{k=0}^r d\theta_k}{|\theta_0 I_n + Z D Z'|^{1/2} \prod_{k=0}^r \theta_k^{\alpha_k+1}} \quad (5.1.10)$$

Hence the posterior distribution exists *iff* (2.1.3) converges.

□

## 5.2 Lemma 2:

*Necessity:* Denote  $\theta_{-(0)} = (\theta_1, \dots, \theta_r)$ . Let

$$f(\theta_{-(0)}) = \exp(-\frac{1}{2}(Y - X\beta)'(\theta_0 I_n + Z D(\theta_{-(0)}) Z')^{-1}(Y - X\beta)).$$

Hobert and Casella ([4]) proved that  $f(\cdot)$  is non-decreasing in  $(\theta_1, \dots, \theta_r)$ . Thus,

$$\begin{aligned} f(\theta_{-(0)}) &\geq \exp(-\frac{1}{2\theta_0}(Y - X\beta)'(Y - X\beta)), \quad \forall \theta_{-(0)} \\ \Rightarrow \int_{\mathcal{C}} f(\theta_{-(0)}) dY^{(2)} &\geq \int_{\mathcal{C}} \exp(-\frac{1}{2\theta_0}(Y - X\beta)'(Y - X\beta)) dY^{(2)} \\ &= \theta_0^{-\frac{(n-m)}{2}} \exp(-\frac{1}{2\theta_0}(Y^{(1)} - X^{(1)}\beta)'(Y^{(1)} - X^{(1)}\beta)) \\ &\quad \prod_{i=m+1}^n \left( \Phi\left(\frac{b_i^* - x_i'\beta}{\sqrt{\theta_0}}\right) - \Phi\left(\frac{a_i^* - x_i'\beta}{\sqrt{\theta_0}}\right) \right) \end{aligned}$$

where  $x_i'$  is the  $i^{\text{th}}$  row of  $X$ . Thus from (5.1.10),

$$\begin{aligned} m(Y^{(1)}, Y^{(2)} \in \mathcal{C}) &\geq \int_{\theta} \frac{e^{-\frac{1}{2\theta_0}(Y^{(1)} - X^{(1)}\beta)'(Y^{(1)} - X^{(1)}\beta)} \prod_{i=m+1}^n \left( \Phi\left(\frac{b_i^* - x_i'\beta}{\sqrt{\theta_0}}\right) - \Phi\left(\frac{a_i^* - x_i'\beta}{\sqrt{\theta_0}}\right) \right) \prod_{k=0}^r d\theta_k}{\theta_0^{-\frac{(n-m)}{2}} |\theta_0 I_n + Z D Z'|^{1/2} \prod_{k=0}^r \theta_k^{\alpha_k+1}} \\ &= \int_{\theta_0} \frac{e^{-\frac{1}{2\theta_0}(Y^{(1)} - X^{(1)}\beta)'(Y^{(1)} - X^{(1)}\beta)} \prod_{i=m+1}^n \left( \Phi\left(\frac{b_i^* - x_i'\beta}{\sqrt{\theta_0}}\right) - \Phi\left(\frac{a_i^* - x_i'\beta}{\sqrt{\theta_0}}\right) \right)}{\theta_0^{\frac{m-q+2(\alpha_0+1)}{2}}} d\theta_0 \\ &\quad \times \int_{\theta_1, \dots, \theta_r} \frac{\prod_{k=1}^r d\theta_k}{|D|^{1/2} |Z' Z + \theta_0 D^{-1}|^{1/2} \prod_{k=1}^r \theta_k^{\alpha_k+1}} \end{aligned}$$

Hobert and Casella ([4]) proved that the iterated integral over  $(\theta_1, \dots, \theta_r)$  converges *iff* conditions

(i) and (ii) hold, in which case it is equal to:  $c \theta_0^{-(\frac{q}{2} + \sum_{k=1}^r \alpha_k)}$  where  $c$  is a constant independent of

$\theta_0$ . Thus,

$$\begin{aligned} m(Y^{(1)}, Y^{(2)} \in \mathcal{C}) &\geq c \int_{\theta_0} \theta_0^{-\frac{w}{2}} \exp(-\frac{1}{2\theta_0}(Y^{(1)} - X^{(1)}\beta)'(Y^{(1)} - X^{(1)}\beta)) \\ &\quad \prod_{i=m+1}^n \left( \Phi\left(\frac{b_i^* - x_i'\beta}{\sqrt{\theta_0}}\right) - \Phi\left(\frac{a_i^* - x_i'\beta}{\sqrt{\theta_0}}\right) \right) d\theta_0 \end{aligned} \quad (5.2.11)$$

where  $w = m + 2 \sum_{k=0}^r \alpha_k + 2$ . (5.2.11) converges *iff* condition (iii) holds, since otherwise

$\lim_{\theta_0 \rightarrow \infty} \theta_0^{-\frac{w}{2}} = \infty$ . Thus, conditions (i), (ii) and (iii) of lemma 2 are necessary.

□

*Sufficiency:* Since the integrand in  $m(Y^{(1)}, Y^{(2)} \in \mathcal{C})$  is non-negative:

$$m(\cdot) \leq \int_{\theta} \int_{\mathfrak{R}^{(n-m)}} \frac{e^{-\frac{1}{2}(Y - X \beta)'(\theta_0 I_n + Z D Z')^{-1}(Y - X \beta)} dY^{(2)} \prod_{k=0}^r d\theta_k}{|\theta_0 I_n + Z D Z'|^{1/2} \prod_{k=0}^r \theta_k^{a_k+1}} \quad (5.2.12)$$

$$= \int_{\theta} \frac{e^{-\frac{1}{2}(Y^{(1)} - X^{(1)} \beta)'(\theta_0 I_m + Z^{(1)} D Z^{(1)'})^{-1}(Y^{(1)} - X^{(1)} \beta)} \prod_{k=0}^r d\theta_k}{|\theta_0 I_m + Z^{(1)} D Z^{(1)'}|^{1/2} \prod_{k=0}^r \theta_k^{a_k+1}} \quad (5.2.13)$$

where (5.2.13) follows from (5.2.12) using the result on the marginal distribution of variables that have a joint normal distribution ([13]). Conditions (i), (ii) and (iii) are sufficient for (5.2.13) to converge ([4]) and hence they are sufficient for the existence of the posterior distribution of the variance components. □

## 6 Appendix 2:

### 6.1 Lemma 3:

*Necessity:* In a manner similar to the proof in Lemma 1, we can show that the posterior distribution of  $\theta$  exists *iff* the following integral converges:

$$J \propto \int_{\theta} \int_{\beta} \int_{\mathcal{C}} \frac{\exp(-\frac{1}{2}(Y - X \beta)'(\theta_0 I_n + Z D Z')^{-1}(Y - X \beta)) dY^{(2)} d\beta \prod_{k=0}^r d\theta_k}{|\theta_0 I_n + Z D Z'|^{1/2} \prod_{k=0}^r \theta_k^{a_k+1}} \quad (6.1.14)$$

Denote  $V = (\theta_0 I_n + Z D Z')$ . Integrating (6.1.14) over  $\beta$ , we get ([4]):

$$J \propto \int_{\theta} \int_{\mathcal{C}} \frac{\exp(\frac{1}{2}Y'(V^{-1} X (X' V^{-1} X)^{-1} X' V^{-1} - V^{-1})Y) dY^{(2)} \prod_{k=0}^r d\theta_k}{\theta_0^{\frac{n-q-p}{2}} |D|^{1/2} |X' X|^{1/2} |D^{-1} \theta_0 + Z' P Z|^{1/2} \prod_{k=0}^r \theta_k^{a_k+1}}$$

where  $P = (I - X (X' X)^{-1} X)$ . Now, since

$\exp(\frac{1}{2}Y'(V^{-1} X (X' V^{-1} X)^{-1} X' V^{-1} - V^{-1})Y) \geq \exp(-\frac{1}{2}Y' V^{-1} Y)$ , we have,

$$J \geq \int_{\theta} \int_{\mathcal{C}} \frac{\exp(-\frac{1}{2}(Y' (\theta_0 I_n + Z D Z')^{-1} Y)) dY^{(2)} \prod_{k=0}^r d\theta_k}{\theta_0^{\frac{n-q-p}{2}} |D|^{1/2} |X' X|^{1/2} |D^{-1} \theta_0 + Z' P Z|^{1/2} \prod_{k=0}^r \theta_k^{a_k+1}}$$

In a manner similar to the proof of the *necessary* condition of lemma 1, we can show that:

$$J \geq \int_{\theta_0} \frac{\exp(-\frac{1}{2\theta_0} Y^{(1)'} Y^{(1)}) \prod_{i=m+1}^n (\Phi(\frac{b_i^*}{\sqrt{\theta_0}}) - \Phi(\frac{a_i^*}{\sqrt{\theta_0}})) d\theta_0}{\theta_0^{\frac{m-q-p+2a_0+2}{2}}} \int_{\theta} \frac{\prod_{k=1}^r d\theta_k}{|D|^{1/2} |X' X|^{1/2} |D^{-1} \theta_0 + Z' P Z|^{1/2} \prod_{k=1}^r \theta_k^{\alpha_k+1}}$$

Hobert and Casella ([4]) proved that the conditions specified in Lemma 3 are necessary for the integral on the right hand side to converge and hence they are necessary for the existence of the posterior distribution of the variance components.

□

## 6.2 Lemma 4:

*Sufficiency:* Since the integrand in (6.1.14) is non negative, we have:

$$\begin{aligned} J &\leq \int_{\theta} \int_{\beta} \int_{\mathbb{R}^{n-m}} \frac{\exp(-\frac{1}{2}(Y - X \beta)'(\theta_0 I_n + Z D Z')^{-1}(Y - X \beta)) dY^{(2)} d\beta \prod_{k=0}^r d\theta_k}{|\theta_0 I_n + Z D Z'|^{1/2} \prod_{k=0}^r \theta_k^{\alpha_k+1}} \\ &= \int_{\theta} \int_{\beta} \frac{\exp(-\frac{1}{2}(Y^{(1)} - X^{(1)} \beta)'(\theta_0 I_m + Z^{(1)} D Z^{(1)'})^{-1}(Y^{(1)} - X^{(1)} \beta)) d\beta \prod_{k=0}^r d\theta_k}{|\theta_0 I_m + Z^{(1)} D Z^{(1)'}|^{1/2} \prod_{k=0}^r \theta_k^{\alpha_k+1}} \end{aligned}$$

where the second result follows from the first, due to the result on the marginal distribution of variates that have a joint normal distribution. Hobert and Casella ([4]) proved that the conditions in Lemma 4 are sufficient for the integral on the right hand side to converge and hence they are sufficient for the propriety of the posterior distribution of the variance components.

□

## 7 Appendix 3:

### 7.1 Lemma 5:

Denote the observed data by  $W = (W_1, W_2, \dots, W_n)$ . As in lemma 1, we can show that the marginal distribution of the observed data is:

$$m(W) \propto \int_{\theta} \int_{\mathcal{C}} \frac{\exp(-\frac{1}{2}Y'(I_n + Z D Z')^{-1}Y) dY \prod_{k=1}^r d\theta_k}{|I_n + Z D Z'|^{1/2} \prod_{k=1}^r \theta_k^{\alpha_k+1}} \quad (7.1.15)$$

where  $\mathcal{C} = \underbrace{(-\infty, 0), \dots, (-\infty, 0)}_l \underbrace{(0, \infty), \dots, (0, \infty)}_{n-l}$  and  $D = \text{diag}(\theta_1 I_{q_1}, \dots, \theta_r I_{q_r})$ .

Thus (7.1.15) reduces to:

$$m(W) \propto \int_{\theta} \frac{P(Y_1 < 0, \dots, Y_l < 0, Y_{l+1} > 0, \dots, Y_n > 0) \prod_{k=1}^r d\theta_k}{\prod_{k=1}^r \theta_k^{\alpha_k+1}} \quad (7.1.16)$$

Now, on account of the special variance-covariance structure of the  $Y$ , we can write  $Y = T - Z D^{1/2} V$  where  $T \sim N_n(0, I)$  independently of  $V \sim N_q(0, I)$ . This structure of the variance-covariance matrix of the  $Y$  is rather fortunate, as now, we can reduce a  $n$  dimensional integral into a  $q$  ( $< n$ ) dimensional integration problem. Thus, (7.1.16) reduces to:

$$m(W) \propto \int_{\theta} \int_{\mathfrak{R}^q} \frac{P(T > Z^* D^{1/2} v) e^{-\frac{1}{2}v'v} dv \prod_{k=1}^r d\theta_k}{\prod_{k=1}^r \theta_k^{\alpha_k+1}} \quad (7.1.17)$$

Thus, the posterior distribution of the variance components exists iff (7.1.17) converges.

□

### 7.2 Lemma 6:

Proofs of the conditions (i) and (ii) are similar to the proofs in lemma 2. We now that prove that condition (iii) is necessary. If  $\exists v_k \in \mathfrak{R}^{q_k \times 1} : Z_k^* v_k \leq 0$  for some  $k$ , then in (7.1.17), we have:

$$m(W) \geq \int_{\theta} \int_{\mathfrak{R}^q - q_k} \int_{v_k : Z_k^* v_k \leq 0} \frac{P(T > Z^* D^{1/2} v) e^{-\frac{1}{2}v'v} dv \prod_{k=1}^r d\theta_k}{\prod_{k=1}^r \theta_k^{\alpha_k+1}} \quad (7.2.18)$$

$$\geq \int_{\theta} \int_{\mathbb{R}^q - q_k} \int_{v_k : Z_k^* v_k \leq 0} \frac{P(T > Z_{-(k)}^* D_{-(k)}^{1/2} v_{-(k)}) e^{-\frac{1}{2} v' v} dv \prod_{k=1}^r d\theta_k}{\prod_{k=1}^r \theta_k^{a_k+1}} \quad (7.2.19)$$

where  $Z_{-(k)}^* = (Z_1^*, \dots, Z_{k-1}^*, Z_{k+1}^*, \dots, Z_r^*)$ ,  $D_{-(k)} = \text{diag}(\theta_1 I_{q_1}, \dots, \theta_{k-1} I_{q_{k-1}}, \theta_{k+1} I_{q_{k+1}}, \dots, \theta_r I_{q_r})$

and  $v_{-(k)} = (v'_1, \dots, v'_{k-1}, v'_{k+1}, \dots, v'_r)'$ . (7.2.19) follows from (7.2.18) since

$$Z_{-(k)}^* D_{-(k)}^{1/2} v_{-(k)} \geq Z^* D^{1/2} v \text{ over the set } \{v_k : Z^* v_k \leq 0\}.$$

We can rewrite (7.2.19) as:

$$\int_{\theta_k} \frac{d\theta_k}{\theta_k^{a_k+1}} \int_{\theta_{-(k)}} \int_{\mathbb{R}^q - q_k} \int_{v_k : Z_k^* v_k \leq 0} \frac{P(T > Z_{-(k)}^* D_{-(k)}^{1/2} v_{-(k)}) e^{-\frac{1}{2} v' v} dv \prod_{l \neq k} d\theta_l}{\prod_{l \neq k} \theta_l^{a_l+1}} \quad (7.2.20)$$

where  $\theta_{-(k)} = (\theta_1, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_r)$ . It is clear that (7.2.20) diverges on account of the iterated integral over  $\theta_k$  and hence so does  $m(W)$ . Thus, condition (iii) is necessary for the existence of the posterior distribution of  $\theta$ .

□

### 7.3 Lemma 7:

In (7.1.17) we make a change of variable:  $x_k = \theta_k^{1/2} v_k$ . Then, we can write (7.1.17) as:

$$m(W) \propto \int_{\theta} \int_{\mathbb{R}^q} \frac{P(T > (Z_1^* x_1 + \dots + Z_r^* x_r)) \prod_{k=1}^r e^{-\frac{1}{2\theta_k} x_k' x_k} d\theta_k dx_k}{\prod_{k=1}^r \theta_k^{a_k + \frac{q_k}{2} + 1}} \quad (7.3.21)$$

The iterated integral over  $\theta_k$  is simply the kernel of an inverted gamma distribution and is finite if  $a_k + \frac{q_k}{2} > 0$ ,  $k = 1, \dots, r$ . When this is the case, the integral over  $\theta_k$  is equal to  $c_k (x_k' x_k)^{-(a_k + \frac{q_k}{2})}$ , where  $c_k$  is a constant independent of  $x_k$ . Thus,

$$\begin{aligned} m(W) &\propto \int_{\mathbb{R}^q} P(T > (Z_1^* x_1 + \dots + Z_r^* x_r)) \prod_{k=1}^r (x_k' x_k)^{-(a_k + \frac{q_k}{2})} dx_k \\ &\leq \int_{\mathbb{R}^q} P(T_{j_x} > (z_{1,j_x}^* x_1 + \dots + z_{r,j_x}^* x_r)) \prod_{k=1}^r (x_k' x_k)^{-(a_k + \frac{q_k}{2})} dx_k \end{aligned} \quad (7.3.22)$$

where  $j_x \in (1, \dots, n)$  is an index such that:

$$z_{k,j_x}^{*'} v_k > 0, k = 1, \dots, r$$

and  $z_{k,j_x}^{*'}$  is the  $j_x^{th}$  row of  $Z_k^*$ . It is easy to see that for any  $X \sim N(0, 1)$  we have:

$$P(X > \lambda) \leq \exp(-\frac{1}{2} \lambda^2), \lambda > 0.$$

Using this fact, conditon (ii) of lemma 7 and ignoring the positive cross products in the exponent, we can bound the right hand side of (7.3.22) by:

$$m(W) \leq \int_{\mathbb{R}^q} \exp(-\frac{1}{2}((z_{1,j_x}^{*'} x_1)^2 + \dots + (z_{r,j_x}^{*'} x_r)^2)) \prod_{k=1}^r (x_k' x_k)^{-(a_k + \frac{q_k}{2})} dx_k$$

In order to remove the dependence of the index  $j_x$  on the variable of integration, we can bound it by the sum over all the indices (rows) in the following way:

$$\begin{aligned} &\leq \sum_{j=1}^n \int_{\mathbb{R}^q} \exp(-\frac{1}{2}((z_{1,j}^{*'} x_1)^2 + \dots + (z_{r,j}^{*'} x_r)^2)) \prod_{k=1}^r (x_k' x_k)^{-(a_k + \frac{q_k}{2})} dx_k \\ &= \sum_{j=1}^n \prod_{k=1}^r \int_{\mathbb{R}^{q_k}} \exp(-\frac{1}{2}(z_{k,j}^{*'} x_k)^2) (x_k' x_k)^{-(a_k + \frac{q_k}{2})} dx_k \end{aligned} \quad (7.3.23)$$

We now consider the convergence of one of the iterated integrals in (7.3.23). Thus, we inspect the convergence of the following integral:

$$J = \int_{\mathbb{R}^{q_1}} \exp(-\frac{1}{2}(z_{1,j}^{*'} x_1)^2) (x_1' x_1)^{-(a_1 + \frac{q_1}{2})} dx_1 \quad (7.3.24)$$

Make the change of variable  $u_1 = z_{1j_1}^* x_{11} + \dots + z_{1j_{q_1}}^* x_{1q_1}$ ,  $u_2 = x_{12}, \dots, u_{q_1} = x_{1q_1}$ . The Jacobian of the transformation is  $\frac{1}{z_{1j_1}^*}$ . We can thus express  $J$  as:

$$J \propto \int_{\mathbb{R}^{q_1}} \exp(-\frac{1}{2}u_1^2) \frac{du_1 \dots du_{q_1}}{\left(\frac{(u_1 - z_{1j_2}^* u_2 - \dots - z_{1j_{q_1}}^* u_{q_1})^2}{z_{1j_1}^{*2}} + u_2^2 + \dots + u_{q_1}^2\right)^{a_1 + \frac{q_1}{2}}} \quad (7.3.25)$$

Make another change of variable  $u_i = u_1 w_i$ ,  $i = 2, \dots, q_1$ . Then (7.3.25) reduces to:

$$\begin{aligned} J \propto &\int_{-\infty}^{\infty} \exp(-\frac{1}{2}u_1^2) u_1^{-2(a_1 + \frac{q_1}{2}) + q_1 - 1} du_1 \\ &\int_{\mathbb{R}^{q_1-1}} \frac{dw_2 \dots dw_{q_1}}{\left((1 - z_{1j_2}^* w_2 - \dots - z_{1j_{q_1}}^* w_{q_1})^2 + z_{1j_1}^{*2} w_2^2 + \dots + z_{1j_1}^{*2} w_{q_1}^2\right)^{a_1 + \frac{q_1}{2}}} \end{aligned}$$

Since  $(1 - z_{1j_2}^* w_2 - \dots - z_{1j_{q_1}}^* w_{q_1})^2 \geq (1 - 2(z_{1j_2}^* w_2 + \dots + z_{1j_{q_1}}^* w_{q_1}))$ , we have:

$$J \leq \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}u_1^2\right) u_1^{-2(a_1 + \frac{q_1}{2}) + q_1 - 1} du_1 \int_{\mathfrak{R}^{q_1-1}} \frac{dw_2 \dots dw_{q_1}}{\left((1 - 2(z_{1j_2}^* w_2 + \dots + z_{1j_{q_1}}^* w_{q_1})) + z_{1j_1}^{*2} w_2^2 + \dots + z_{1j_1}^{*2} w_{q_1}^2\right)^{a_1 + \frac{q_1}{2}}}$$

Collecting terms and completing squares we can rewrite the above equation as:

$$J \leq \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}u_1^2\right) u_1^{-2(a_1 + \frac{q_1}{2}) + q_1 - 1} du_1 \int_{\mathfrak{R}^{q_1-1}} \frac{dw_2 \dots dw_{q_1}}{\left(\frac{z_{1j_1}^{*2} - z_{1j_2}^{*2} - \dots - z_{1j_{q_1}}^{*2}}{z_{1j_1}^{*2}} + z_{1j_1}^{*2} \sum_{l=2}^{q_1} \left(w_l - \frac{z_{1j_l}^*}{z_{1j_1}^{*2}}\right)^2\right)^{a_1 + \frac{q_1}{2}}}$$

Let  $y_l = z_{1j_1}^* \left(w_l - \frac{z_{1j_l}^*}{z_{1j_1}^{*2}}\right)$ ,  $l = 2, \dots, q_1$  and denote the constant  $\frac{z_{1j_1}^{*2} - z_{1j_2}^{*2} - \dots - z_{1j_{q_1}}^{*2}}{z_{1j_1}^{*2}}$  by  $c$ . Then

$$J \leq \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}u_1^2\right) u_1^{-2a_1 - 1} du_1 \int_{\mathfrak{R}^{q_1-1}} \frac{dy_2 \dots dy_{q_1}}{\left(c + (y_2^2 + \dots + y_{q_1}^2)\right)^{a_1 + \frac{q_1}{2}}} \quad (7.3.26)$$

The integral over  $u_1$  converges if  $a_1 < 0$ . We now consider the convergence of the integral over  $y_2, \dots, y_{q_1}$ . Using the idea of spherical co-ordinates, we can reduce this integral to:

$$\int_{\mathfrak{R}^{q_1-1}} \frac{dy_2 \dots dy_{q_1}}{\left(c + (y_2^2 + \dots + y_{q_1}^2)\right)^{a_1 + \frac{q_1}{2}}} = \int_0^{\infty} \frac{r^{q_1-2} dr}{\left(c + r^2\right)^{a_1 + \frac{q_1}{2}}} \quad (7.3.27)$$

and the above converges if  $a_1 > -\frac{1}{2}$ . Thus, conditions (i) and (ii) are sufficient. □

## References

- [1] Box and Tiao (1973) *Bayesian Inference in Statistical Analysis*. Wiley Classics Library.
- [2] Gelfand A. E., Smith A. F. M. and Lee T (1992) *Bayesian Analysis of Constrained Parameter and Truncated Data Problems Using Gibbs Sampling* J A S A 87: No 418, 523-531.
- [3] Harville and Mee R. W. (1984) *Mixed Models for Ordered Categorical Data*. Biometrics (40): 393-408.

- [4] Hobert J. and Casella G. (1993) *Using the Gibbs Sampler with Improper Prior Distributions*. Biometrics Unit Technical Report (BU-1221-M), Cornell University.
- [5] Ibrahim J. G. and Laud P. W (1991) *A Bayesian Analysis of Generalized Linear Models using Jeffrey's Prior*. J A S A 86: No 416, 981-986
- [6] Johnson and Kotz (1970) *Distributions in Statistics: Continuous Univariate Distributions 2* Wiley.
- [7] Judge et. al, (1985) *The Theory and Practice of Econometrics* Wiley.
- [8] Karim M. R: and Zeger S. L. (1992) *Generalized Linear Models with Random Effects; Salamander Mating Revisited*. Biometrics 48: 631-644.
- [9] Dempster A. P., Laird N. M. and Rubin D. B. (1977) *Maximum Likelihood from Incomplete Data via the E M Algorithm*. J R S S (B)
- [10] Kuo Lynn and Smith A. F. M. *Survival Analysis: State of the Art*. 11-24. Kluwer Academic Publishers.
- [11] McCullagh P. and Nelder J. A.(1984) *Generalized Linear Models*. Second Edition
- [12] McCulloch C. E. (1994) *Maximum Likelihood Variance Components Estimations for Binary Data*. JASA 89: No 425, 330-335.
- [13] Searle S. R., Casella G, McCulloch C. E. (1992) *Variance Components*. (Wiley-Interscience Publication)