Applications and Applied Mathematics: An International Journal (AAM)

Volume 8 | Issue 2 | Article 4

12-2013

Local Influence in Bayesian Elliptically Contoured-Ordinal Model for Mixed Data

Ehsan B. Samani
Shahid Beheshti University

Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466

Vol. 8, Issue 2 (December 2013), pp. 391 – 403

Applications and Applied Mathematics:
An International Journal
(AAM)

Local Influence in Bayesian Elliptically Contoured-Ordinal Model for Mixed Data

Ehsan Bahrami Samani
Department of Statistics
Shahid Beheshti University
Tehran, Iran
ehsan_bahrami_samani@yahoo.com

Received: December 29, 2012; Accepted: October 29, 2013

Abstract

This paper develops a new class of joint modeling of mixed correlated ordinal and continuous responses with elliptically contoured errors. This joint model includes the latent variable approach of using an elliptically contoured distribution for mixed ordinal and continuous responses. A Markov Chain Monte Carlo sampling algorithm is described for estimating the posterior distribution of the parameters. For sensitivity analysis to investigate the perturbation from associate responses, it is demonstrated how one can use some elements of covariance structure. Influence of small perturbation of these elements on the posterior normal curvature is also studied. To illustrate the application of such modeling the data (medical) is analyzed.

Keywords: Bayesian inference; Joint modeling; Gibbs sampler; Ordinal and continuous responses; Markov Chain; Monte Carlo; Sensitivity Analysis; Posterior Normal Curvature

AMS-MSC (2010) No.: 62F03

1. Introduction

Some biomedical data include some correlated discrete and continuous outcomes. Outcomes related to mixed ordinal and continuous responses data are pervasive and research on analyzing them needs to be encouraged. For joint modeling of responses, one method is to use the general location model of Olkin and Tate (1961), where the joint distribution of the continuous and categorical variables is decomposed into a marginal multinomial distribution for the categorical variables and a conditional multivariate normal distribution for the continuous variables. The categorical variables (for a mixed Poisson and continuous responses) were used in the method by Olkin and Tate's (1961), see Yang et al. (2007). For joint modeling of mixed outcomes using latent variables see McCulloch (2007). A second method for joint modeling is to decompose the joint distribution as a multivariate marginal distribution for the continuous responses and a conditional distribution for categorical variables given the continuous variables. Cox and Wermuth (1992) empirically examined the choice between these two methods. The third method uses the simultaneous modeling of categorical and continuous variables to take into account the association between the responses by the correlation between errors in the model for responses. For more details of this approach see, for example, Heckman (1978) in which a general model for simultaneously analyzing two mixed correlated responses is introduced and Catalano and Ryan (1992) who extended and used the model for a cluster of discrete and continuous outcomes. Poon and Lee (1987) presented a model for the ordinal and continuous responses without considering any covariate effect.

The class of elliptical distribution includes a vast set of known symmetric distributions, for example, the normal and Student *t* distributions [see Kelker (1970) and Fang et al. (1990)]. In the case of normal distribution, see Azzalini and Dalla Valle (1996), Azzalini and Capitanio (1999), and Arnold and Beaver (2000). They obtained the multivariate distribution by conditioning on one suitable random variable being greater than Zero. Sahu et al. (2003) conditioned on as many random variables as the dimension of the multivariate distribution (multivariate skew-normal and *t* distributions). They obtained analytical forms of densities and studied distributional properties.

Bayesian analysis of regression problems under heavy-tailed error distributions were researched by Zellner (1976), Geweke (1993) and Fernandez and Steel (1998). Extensions of those results for elliptical distributions are considered in Chib et al. (1998). The Bayesian approach has several important advantages. First, the exact posterior distribution of the parameters can be estimated by using MCMC methods. Means and quantities based on the estimated posterior are appropriate regardless of the sample size. In contrast, standard errors and confidence limits for the maximum likelihood estimates are typically based on strong asymptotic normality assumptions. Second, the Bayesian approach allows for the direct incorporation of prior knowledge. This is a major advantage in structural equation modeling. Classical methods often require that a subset of the parameters is known to ensure identifiability. Although constrain on the threshold parameters and the variance of the latent variables are often reasonable, additional less justifiable constraints can be avoided by using a prior distribution to allow for prior uncertainty in the parameters. In addition, by assigning parameters about which there is previous information, more precise estimates of the parameters of interest can be obtained.

In this paper, the author proposed a new class of latent variable models for mixed correlated ordinal and continuous responses with elliptically contoured errors. Markov chain Monte Carlo (MCMC) algorithms (Tierney, 1994) are developed for estimating the posterior distribution of the parameters. The aim of this paper is to adapt and extend an approach similar to that of Heckman (1978), for a joint modeling of multivariate ordinal and continuous outcomes with elliptically contoured errors.

In Section 2, the general modeling framework is described. A general MCMC sampling algorithm for posterior estimation is outlined in Section 3. In Section 4, we have a simulation study. In Section 5, a common way to investigate if perturbations of model components influence key results of the analysis is to compare the results derived from the original and perturbed models using posterior normal curvature. In Section, the proposed methodology is applied on the medical data. In this section, as sensitivity analysis for these data the influence of a small perturbation of the associate responses on posterior normal curvature will be also investigated. Finally, concluding remarks are given.

2. Latent Variable Model

Consider a $m-$dimensional random vector W having elliptical distribution with probability density function of the form

$$f(w|\mu, \Sigma, g^{(M)}) = |\Sigma|^{-\frac{1}{2}} g^{(M)}[(w-\mu)'\Sigma^{-1}(w-\mu)],$$

where Σ is an $m \times m$ positive definite matrix with the covariance matrix of the random vector given by W $g^{(M)}$ is a function R^+ to R^+ called the density generator of the random vector W, defined by

$$g^{(M)} = \frac{\Gamma(\frac{M}{2})}{\pi^{\frac{M}{2}}} . \frac{g(u;M)}{\int_0^\infty u^{\frac{M}{2}-1} g(u;M) du},$$

where $g(u;M)$ is a function R^+ to R^+ such that the $\int_0^\infty u^{\frac{M}{2}-1} g(u;M)du < \infty$. We shall denote an elliptically distributed $M-$ dimensional vector with location μ, scale Σ and characteristic generator $g^{(M)}$, by $EC_M(\mu, \Sigma; g^{(M)})$.

We use Y_{ij} to denote jth ordinal response for the ith individual with c_j levels defined as,

$$Y_{ij} = \begin{cases} 1 & Y_{ij}^* < \theta_{1,j}, \\ k+1 & \theta_{k,j} \le Y_{ij}^* < \theta_{k+1,j}, \quad k = 1, \dots, c_j - 2, \\ c_j & Y_{ij}^* \ge \theta_{c_j,j}, \end{cases}$$

where $i = 1, \dots, n$, $j = 1, \dots, M_1$. $\theta_{1,j}, \dots, \theta_{c_j-1,j}$ are the cut-point parameters and Y_{ij}^* denotes the underlying latent variable for Y_{ij}. The joint model takes the form:

$$\begin{aligned} Y_{ij}^* &= \beta_j{}' X_i + \varepsilon_{ij}^{(1)} \qquad j = 1, \dots, M_1, \\ Z_{ij} &= \beta_j{}' X_i + \varepsilon_{ij}^{(2)} \qquad j = M_1 + 1, \dots, M, \quad (M_{\mathbf{1}} < M). \end{aligned} \tag{1}$$

Let

$$\varepsilon_i = (\varepsilon_i^{(1)}, \varepsilon_i^{(2)})' \overset{iid}{\sim} E\,C_M(0, \Sigma; g),$$

where $\varepsilon_i^{(1)} = (\varepsilon_{i1}^{(1)}, \dots, \varepsilon_{iM_1}^{(1)})', \varepsilon_i^{(2)} = (\varepsilon_{i(M_1+1)}^{(2)}, \dots, \varepsilon_{iM}^{(2)})'$, $\theta_j = (\theta_{1,j}, \dots, \theta_{c_j-1,j})', j = 1, \dots, M_1$, is the vector of cutpoint parameters for the jth ordinal response and X_i is the vector of explanatory variables for the ith individual and Σ is the $M \times M$ covariance matrix which for illustration, when $M_1 = 2$ and $M = 3$ has the following structure,

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \sigma\rho_{13} \\ \rho_{12} & 1 & \sigma\rho_{23} \\ \sigma\rho_{13} & \sigma\rho_{23} & \sigma^2 \end{pmatrix},$$

where σ^2 is the variance of the continuous response, and $\rho_{jj'}$ for $j \neq j'$, $j = j' = 1,2,3$ is the correlation between jth and j'th responses. The vector of coefficients β_j, cut points parameters θ_j for $j = 1, \dots, M_1$ and Σ should be estimated. The parameter vector, β_j for $j = M_1 + 1, \dots, M$, includes an intercept parameter but β_j, for $j = 1, \dots, M_1$, due to having cut point parameters, are assumed not to include any intercept. In this model any elliptical distribution can be assumed for the errors in the model.

3. Bayesian Estimation

In this section, prior distributions are chosen for the parameters, and the general MCMC algorithm is outlined for estimating the posterior distributions of the parameters and the latent variables via Dunson (2000). The prior distributions are conjugate if the underlying variables are normal.

Markov chain Monte Carlo (MCMC) methods use computer simulation of Markov chains in the parameter space. The Markov chains are defined in such a way that the posterior distribution, in a given statistical inference problem, is the asymptotic distribution. One of the standard approaches to define such Markov chains is Gibbs sampling. We will use MCMC techniques for posterior computation in the models proposed in section 2. In the special case where all the underlying and latent variables have normal distribution the MCMC algorithm is Gibbs sampler that follows a simple form.

3.1. Prior Distributions

The parameters $\eta = (\beta_1, \dots, \beta_M, \theta_1, \dots, \theta_{M_1})'$ are assigned a normal prior,

$$\eta \sim EC_q(\mu_0, \Sigma_0; g),$$

where μ_0 is a vector of location parameters, Σ_0 is a covariance matrix and $q = M + M_1$. To choose a vague prior distribution for θ, set $\mu_0 = 0$ and $\Sigma_0 = diag\{\sigma_1^2, \dots, \sigma_q^2\}$. Wishart prior are specified for the precision matrix Σ in expressions (1):

$$\Sigma \sim Wishart(\nu, \Lambda),$$

with degrees of freedom ν and precision Λ . A prior can be assigned by choosing $\nu \geq M$, where M is the the dimension of Σ.

3.2. The Gibbs Sampler

Consider the model with all equations given in (1). We assume that η and Σ are a priori independent with $p(\eta) = EC_q(\mu_0, \Sigma_0; g)$ and $p(\Sigma) = Wishart(\nu, \Lambda)$, where $p(.)$ is the prior distribution. The conditional posterior distribution of $p(\eta|y, z, \Sigma)$ and $p(\Sigma|y, z, \eta)$ are computed in subsection 3.3.

To estimate the posterior distributions of the parameters, we define a Markov chain in $\xi = (\eta, \Sigma)$. Denote with $\xi^{(t)} = (\eta^{(t)}, \Sigma^{(t)})$ the state parameter of the Markov chain after t iterations. Given the nature of a Markov chain, all we need to define is the transition probability, i.e., given a current value for $\xi^{(t)}$, we need to generate a new value $\xi^{(t+1)}$. We do so by sampling from the complete conditional posterior distributions for η and Σ

$$\eta^{(t+1)} \sim p\left(\eta^{(t)} \middle| y, z, \Sigma^{(t)}\right),$$
$$\Sigma^{(t+1)} \sim p(\Sigma^{(t)}|yz, \eta^{(t)}).$$

Step 1 and Step 2 define a Markov chain $\xi^{(t)}$ which converges to $p(\eta, \Sigma|y, z)$, as desired. The described Markov chain Monte Carlo simulation is a special case of a Gibbs sampler. In general, let $\xi^* = (\xi_1, \ldots, \xi_p)$ denote the parameter vector. The Gibbs sampler proceeds iteratively, for $j = 1, \ldots, p$, generating from the conditional posterior distributions

$$\xi_j^{(t+1)} \sim p(\xi|\xi_1^{(t+1)}, \ldots, \xi_{j-1}^{(t+1)}, \xi_{j+1}^{(t)}, \ldots, \xi_p^{(t)}, y, z).$$

3.3. Posterior Computations

We now obtain the form of the joint posterior distribution. Let $\mathrm{y} = (\mathrm{y}_1', \ldots, \mathrm{y}_\mathrm{n}')'$, $\mathrm{z} = (\mathrm{z}_1', \ldots, \mathrm{z}_\mathrm{n}')'$ and $\mathrm{x} = (\mathrm{x}_1', \ldots, \mathrm{x}_\mathrm{n}')'$ where $\mathrm{y_i} = (\mathrm{y_{i1}}, \ldots, \mathrm{y_{iM_1}})'$, $\mathrm{z_i} = (\mathrm{z_{i(M_1+1)}}, \ldots, \mathrm{z_{iM}})'$ and $\mathrm{x_i} = (\mathrm{x_{i1}}, \ldots, \mathrm{x_{ip}})'$, and p is the number of explanatory variables for the ith individual (the number of components in this vector may also be dependent on the chosen variable, i.e., $\mathrm{x_i}$ be $\mathrm{x_{im}}$ and p be $\mathrm{p_m}$, here, we ignore this for simplicity).

The joint posterior distribution for the parameters and latent variables is:

$$P\left(\eta,\boldsymbol{\Sigma}\middle|y,z,x\right)\propto f_{y,z}\left(y,z|\eta,\boldsymbol{\Sigma},x\right).\pi\left(\eta,\boldsymbol{\Sigma}|x\right)$$
$$=\left[\prod_{i=1}^{n} f\left(z_i,\, y_i|x_i,\eta,\boldsymbol{\Sigma}\right)\right]\pi\left(\eta,\boldsymbol{\Sigma}|x\right)$$
$$=[\prod_{i=1}^{n} P\left(\theta_{1,y_{i1-1}}<Y_{i1}^{*}\le\theta_{1,y_{i1}},\ldots,\theta_{M_1,y_{i,M_1-1}}<Y_{iM_1}^{*}\le\theta_{M_1,y_{iM_1}}\left|z_i,x_i\right.\right)$$
$$\pi\left(\eta,\boldsymbol{\Sigma}|x\right)f\left(z_i|x_i\right)],$$

where $\pi(.|x)$ denote the joint prior density and $\eta=(\beta_1,\ldots,\beta_M,\theta_1,\ldots,\theta_{M_1})'$.

Using the above Theorem, the joint posterior distribution could be summarized as,

$$\mathrm{P}\left(\eta,\Sigma\middle|\mathrm{y},\mathrm{z},\mathrm{x}\right)\propto[\prod_{i=1}^{n}\Delta_{\mathrm{b}_{i1}\mathrm{a}_{i1}}\cdots\Delta_{\mathrm{b}_{iM_1}\mathrm{a}_{iM_1}}\mathrm{F}(\mathrm{w}_{i1},\cdots,\mathrm{w}_{iM_1}\mid \mathrm{z}_i,\mathrm{x}_i)$$
$$\mathrm{f}(\mathrm{z}_i\mid \mathrm{x}_i)]\pi(\eta,\Sigma\mid \mathrm{x})]$$
$$\propto\left[\prod_{i=1}^{n}(\mathrm{F}_{i0}-\mathrm{F}_{i1}+\mathrm{F}_{i2}-\ldots+\left(-1\right)^{\mathrm{M}_1}\mathrm{F}_{i\mathrm{M}_1})\right]$$
$$\left|\Sigma\right|^{\frac{\nu-\mathrm{M}-1}{2}}\left|\Sigma_{22}\right|^{\frac{-\mathrm{n}\left(\mathrm{M}-\mathrm{M}_1\right)}{2}}$$
$$\mathrm{g}^{\left(\mathrm{M}-\mathrm{M}_1\right)}\left\{\sum_{i=1}^{n}\left(\mathrm{z}_i-\mu_{\mathrm{z}_i}\right)'\Sigma_{22}^{-1}\left(\mathrm{z}_i-\mu_{\mathrm{z}_i}\right)\right\}$$
$$\mathrm{g}^{\left(\mathrm{M}+\mathrm{M}_1\right)}\{(\eta-\mu_0)'\Sigma_0^{-1}\left(\eta-\mu_0\right)\},$$

where $\mu_{z_i}=(\beta_{M_1+1}{}'X_i,\ldots,\beta_{M_1+1}{}'X_i)'$, $\Sigma_{22}=\mathrm{Var}(Z_i)$, $b_{ij}=\theta_{j,y_{ij}}$ and $a_{ij}=\theta_{j,y_{ij}-1}$ and F_{ij} is the sum of all $\binom{M_1}{j}$ terms of the from $F(g_{i1},\ldots,g_{iM_1}|z_i,x_i)$ with $g_{ik}=a_{ik}$ for exactly j integers in $\{0,1,\ldots,M_1\}$, and $g_{ik}=b_{ik}$ for the remaining M_1-j integers.

We require the full conditional distributions of each unknown parameter. We have

$$P\left(\eta\middle|y,z,\Sigma,x\right)\propto\prod_{i=1}^{n}(F_{i0}-F_{i1}+F_{i2}-\ldots+\left(-1\right)^{M_1}F_{iM_1})]\left|\Sigma_{22}\right|^{\frac{-n\left(M-M_1\right)}{2}}$$
$$g^{\left(M-M_1\right)}\left\{\sum_{i=1}^{n}\left(z_i-\mu_{z_i}\right)'\Sigma_{22}^{-1}\left(z_i-\mu_{z_i}\right)\right\}$$
$$g^{\left(M+M_1\right)}\{(\eta-\mu_0)'\Sigma_0^{-1}\left(\eta-\mu_0\right)\}.$$

The full conditional distribution of the precision matrix Σ is

$$P\left(\Sigma \middle| y,z,\eta,x\right) \propto \prod_{i=1}^{n} (F_{i0} - F_{i1} + F_{i2} - \ldots + \left(-1\right)^{M_1} F_{iM_1})]$$
$$\left|\Sigma\right|^{\frac{\nu-M-1}{2}} \left|\Sigma_{22}\right|^{\frac{-n\left(M-M_1\right)}{2}}$$
$$g^{(M-M_1)}\left\{\sum_{i=1}^{n}\left(z_i-\mu_{z_i}\right)^{'}\Sigma_{22}^{-1}\left(z_i-\mu_{z_i}\right)\right\}.$$

4. Simulation

We consider three continuous variables Y_1^*, Y_2^* and Z. The ordinal variables Y_1 and Y_2 with three levels are defined as

$$Y_1 = \begin{cases} 1 & Y^* < \theta_1, \\ 2 & \theta_1 \le Y^* < \theta_2, \\ 3 & Y^* \ge \theta_2, \end{cases}$$

and

$$Y_2 = \begin{cases} 1 & Y^* < \eta_1, \\ 2 & \eta_1 \le Y^* < \eta_2, \\ 3 & Y^* \ge \eta_2, \end{cases}$$

the variables, Y_1^* , Y_2^* and Z are generated by a multivariate normal distribution with zero mean and covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \sigma\rho_{13} \\ \rho_{12} & 1 & \sigma\rho_{23} \\ \sigma\rho_{13} & \sigma\rho_{23} & \sigma^2 \end{pmatrix}.$$

We let $\rho_{12} = \rho_{13} = \rho_{23} = 0.5$ and different values of σ^2, (0.5, 1 and 2). We also let $\theta_1 = \eta_1 = -1$ and $\theta_2 = \eta_2 = 1$. In the set of cut points, not having any covariate in the model for latent variable of Y_1 and Y_2, one expects to have, roughly, 16 percent of Y values to be equal to 1, 16 percent to be equal to 3 and 68 percent to be equal to 2. So, the low and high values have nearly the same frequency but the middle value have the highest frequency. For this we consider 3 values for n (50, 100, and 1000). In this analysis we use 1000 sets of simulation. In each simulation we analyze the following simple model

$$Y_1^* = \varepsilon_1,$$
$$Y_2^* = \varepsilon_2,$$
$$Z = \mu_z + \varepsilon_{3.}$$

Table 1 contains the average estimated values of μ_z, σ^2, ρ_{12} (the correlation between Z and Y_1^*),

ρ_{13} (the correlation between Z and Y_2^*), ρ_{23} (the correlation between Y_1^* and Y_2^*), θ_1, θ_2, η_1 and η_2 for n=50, n=100 and n=1000. The parameter estimates by the model for μ_z, σ^2, ρ_{12}, ρ_{13}, ρ_{23}, θ_1, θ_2, η_1 and η_2 (for $n = 30$, n=100 and n=500) are close to the true values of the parameters. Of course, the more the value of n the better the estimates. We used a Gibbs sampler within winBUGS to estimate from the joint posterior distribution of the parameters. We run three chains with widely varying initial values and used 10000 Gibbs iterates collected after convergence from each chain to compute posterior summaries of the parameters. Posterior summaries of the global parameters for each outcome are shown in Table 1.

Table 1. Results of the simulation study

		n=30		n=100		n=500	
Parameter	True value	Est.	S.E.	Est.	S.E.	Est.	S.E.
μ_z	0.000	0.050	0.151	0.013	0.101	0.001	0.026
σ^2	1.000	1.130	0.136	1.015	0.071	1.003	0.027
	.500	0.441	0.019	0.532	0.011	0 .503	.002
	2.000	2 .128	0.178	2 .013	0.168	2 .013	0.127
ρ_{12}	0.500	0.455	0.145	0.496	0.080	0.502	0.029
ρ_{13}	0.500	0.435	0.139	0.491	0.093	0.506	0.020
ρ_{23}	0.500	0.447	0.178	0.493	0.072	0.509	0.013
θ_1	-1.000	-1.150	0.248	-0.985	0.153	-0.996	0.047
θ_2	1.000	1.085	0.293	1.024	0.164	0.993	0.032
η_1	-1.000	-1.151	0.255	-0.963	0.141	-0.996	0.049
η_2	1.000	1.092	0.226	1.028	0.144	0.984	0.59

5. Sensitivity Analysis

Sensitivity analysis is the study of how model output varies with changes in model inputs. An assessment of the influence of minor perturbations of the model is important, see Cook (1986).

Generally, one introduces perturbations into the model through the $q \times 1$ vector ω which is restricted to some open subset Ω of R^q. Let $P(\eta|\omega)$ denote the log- posterior corresponding to the perturbed model for a given ω in Ω. For a given set of observed data, where η is a $p \times 1$ vector of unknown parameters, the author assume that there is an ω_0 in Ω such that $P(\eta) = P(\eta|\omega_0)$ for all η. Finally, Let $\hat{\eta}$ and $\hat{\eta}_\omega$ denote the maximum posterior estimators under $P(\eta)$ and $P(\eta|\omega)$. To assess the influence of varying ω throughout Ω, the author considers the posterior displacement defined as:

$$PD(\omega) = 2[P(\hat{\eta}) - P(\hat{\eta}_\omega)].$$

A graph $PD(\omega)$ versus ω contains essential information on the influence of the perturbation scheme in questions. It is useful to view this graph as the geometric surface formed by the values of the $(q+1) \times 1$ vector $\alpha(\omega) = (\alpha_1, \alpha_2)' = (\omega, PD(\omega))'$ as ω varies throughout Ω. When $q = 1$, the posterior curvature of such plain curves at ω_0 is,

$$PC = \frac{|\dot{\alpha}_1\ddot{\alpha}_2 - \dot{\alpha}_2\ddot{\alpha}_1|}{(\dot{\alpha}_1^2 - \dot{\alpha}_2^2)^{3/2}}, \quad (2)$$

where the first and second derivations $\dot{\alpha}_i$ and $\ddot{\alpha}_i$ are evaluated at ω_0. Since $\dot{\alpha}_1 = 1$ and $\dot{\alpha}_2 = \ddot{\alpha}_1 = 0$, PC reduces to $PC = \ddot{\alpha}_2 = \ddot{P}D(\omega_0)$. When $q > 1$, an influence graph is a surface in R^{q+1}. The posterior normal curvature PC_l of the lifted line in the direction l can now be obtained by applying (2) to the plan curve $(a, PD(\omega(a)))$, where $\omega(a) = \omega_0 + al$, $a \in R$, and l is a fixed non-zero vector of unit length in R^q. I proposed looking at local influences, i.e., at the posterior normal curvature PC_l of $\alpha(\omega)$ in ω_0, in the direction of some q-dimensional vector l of unit length. Let Δ_i be the p-dimensional vector defined by

$$\Delta_i = \frac{\partial^2 P_i(\eta|\omega_i)}{\partial\omega_i\partial\eta}|\eta = \hat{\eta},\ \omega_i = 0$$

and define Δ as the $p \times n$ matrix with Δ_i as its ith column. Further, let $\ddot{P}$ denote the $p \times p$ matrix of second-order derivatives of $P(\eta|\omega_0)$ with respect to η, also evaluated at $\eta = \hat{\eta}$. Cook (1986) has then shown that C_l can be easily calculated by

$$C_l = 2\left|l^T\Delta^T(\ddot{L})^{-1}\Delta l\right|.$$

I have now shown that PC_l can be easily calculated by

$$PC_l = 2\left|l^T\Delta^T(\ddot{P})^{-1}\Delta l\right|. \quad (3)$$

Obviously, PC_l can be calculated for any direction l. One evident choice is the vector l_i containing one in the $i\,th$ position and zero elsewhere, corresponding to the perturbation of the ith weight only. Another important direction is the direction l_{max} of maximal normal curvature PC_{max}. It shows how to perturb the condition for associate of the responses to obtain the largest local changes in PC_{max}. PC_{max} is the largest eigenvalue of $\Delta_i^T(\ddot{P})^{-1}\Delta_i$ and l_{max} is the corresponding eigenvector.

6. Application

6.1. Medical Data

The medical data set is obtained from an observational study on women in the Taleghani hospital of Tehran, Iran. These data record status of osteoporosis of the spine and Steatosis as ordinal responses and BMI as a continuous response for 163 patients. Osteoporosis of the spine is a disease of bone in which the bone mineral density (BMD) is reduced, bone micro architecture is disrupted and the amount and variety of non-collage nous proteins in bone are altered. BMI is a statistical measure of the weight of body mass index. A person's body mass index may be really accurately calculated using any of the formulas such as BMI $= \frac{W}{H^2}$ (kg/cm^2) where W is weight and H is height. Steatosis is the process describing the abnormal retention of lipids within a cell. It reflects an impairment of the normal process of synthesis and breakdown of triglyceride fat.

Excess lipid accumulates in vesicles that displace the cytoplasm.

These three variables, osteoporosis of the spine, Steatosis and BMI are endogenous correlated variables, and they have to be modeled simultaneously. Explanatory variables which affect these variables are: (1) amount of total body calcium (Ca), (2) job status (Job, employee or housekeeper), (3) type of the accommodation (Ta, house or apartment) and (4) age.

Descriptive statistics (mean and standard deviation for continuous response and frequency or percentage for ordinal responses) are given in Table 2. OS is osteoporosis of the spine of an individual as an ordinal response with 3 levels. These levels defined as 1: individual hasn't osteoporosis of the spine (None), 2: individual has mild osteoporosis of the spine (Mild), 3: individual has severe osteoporosis of the spine (Severe). ST is Steatosis of an individual as an ordinal response with 3 levels. These levels defined as 1: individual hasn't Steatosis (None), 2: individual has mild Steatosis (Mild), 3: individual has severe Steatosis (Severe). BMI is the body mass index of individual

Table 2 . Descriptive statistics for medical data

		No.	Mean	S.E.
BMI		163	29.357	10.806
Steatosis	Levels	No.	Percentage	
	None	41	0.251	
	Mild	58	0.355	
	Severe	64	0.394	
OS				
	None	59	0.362	
	Mild	65	0.399	
	Severe	39	0.239	

Table 2 shows less percentage for severe osteoporosis than those of none and mild levels. The vector of explanatory variable is $X = (Job, Age, Ta, Ca)$.

6.2. Models

For comparative purposes, two models are considered. The first model (model (I)) does not consider the correlation among responses. This model is

$$\begin{aligned} OS^* &= \beta_{11} Job + \beta_{12} Age + \beta_{13} Ta + \beta_{14} Ca + \varepsilon_1, \\ ST^* &= \beta_{21} Job + \beta_{22} Age + \beta_{23} Ta + \beta_{24} Ca + \varepsilon_2, \\ BMI &= \beta_{30} + \beta_{31} Job + \beta_{32} Age + \beta_{33} Ta + \beta_{34} Ca + \varepsilon_3. \end{aligned}$$

The covariance matrix of the vector of errors $(\varepsilon_1, \varepsilon_2, \varepsilon_3)'$ for this model is $\Sigma_{\text{Ind}} = \text{diag}\{1, 1, \sigma^2\}$. The second model (model (II)) uses model (I) and takes into account the correlation among three errors. For this model covariance matrix is

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \sigma\rho_{13} \\ \rho_{12} & 1 & \sigma\rho_{23} \\ \sigma\rho_{13} & \sigma\rho_{23} & \sigma^2 \end{pmatrix}.$$

Here, a multivariate normal distribution with correlation ρ_{12}, ρ_{13} and ρ_{23} are assumed for the errors and these parameters should be also estimated.

6.3. Results

Results of using two models are given in Table 3. Model (I) shows a weak significant effect of age on ST, a weak significant effect of Ta on BMI and a weak significant effect of Ca on OS. From these effects the author can inferred that the older the patient the less the BMI, people who live in apartment have more BMI than of that people who live in a house and the more the amount of calcium of the body of the patient the higher is the probability of low value of osteoporosis of the spine. To compare model (II) and model (I) the author have deviance =123.318 with two d.f. (P-value$<$0.001). So one may preffered model (II). For model (II) correlation parameters ρ_{23} and ρ_{12} are strongly significant. They show a positive correlation between BMI and Steatosis ($\hat{\rho}_{23}$=0.714) and it shows a negative correlation between BMI and osteoporosis of the spine ($\hat{\rho}_{12}$=-0.210). The estimated variance of BMI ($\hat{\sigma}^2$) obtained by model (II) are less than those of model (I). A consequence of estimating the correlation parameters by model (II) is that the estimated standard errors of constant parameters in models for continuous response are reduced in comparing them with the results obtained by model (I).

Table 3. Results using two models for medical data (**: Significant at %5 level, *: Significant at %10 level and dashed (-): Neither of the correlation among responses was considered in model (I))

	Model (I)		Model (II)	
Parameter	Est.	S.E.	Est.	S.E.
OS				
Job	-0.530	0.545	-0.533	0.501
Age	0.006	0.014	0.005	0.032
Ta	0.005	0.139	0.006	0.154
Ca	0.201*	0.123	0.211*	0.126
θ_1	1.207	1.472	1.237	1.889
θ_2	2.284	1.477	2.315	1.894
ST				
Job	1.633	2.133	1.622	2.110
Age	-0.103 *	0.045	-0.102 *	0.044
Ta	0.951	0.740	0.940	0.744
Ca	-0.183	0.417	-0.180	0.317
BMI				
$Const$	86.151**	14.153	86.151**	9.664
Job	1.653	4.918	1.652	6.498
Age	0.127	0.125	0.126	0.127
Ta	3.001 *	1.744	3.000*	1.746
Ca	0.010	0.173	0.009	0.179
σ^2	116.524**	0.588	112.949**	0.588
ρ_{12}	-	-	-0.210**	0.084
ρ_{13}	-	-	-0.101	0.086
ρ_{23}	-	-	0.715 **	0.038
-loglike	1273.044		1211.385	

6.4. Sensitivity Analysis for Correlated Responses

In our application, suppose $\mathrm{M}_1 = 2$ and $\mathrm{M} = 3$, so Σ is the 3×3 covariance matrix which for illustration has the following structure,

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \sigma\rho_{13} \\ \rho_{12} & 1 & \sigma\rho_{23} \\ \sigma\rho_{13} & \sigma\rho_{23} & \sigma^2 \end{pmatrix},$$

which gives the following condition for independent responses:

$$\omega = (\rho_{12}, \rho_{13}, \rho_{23})' = (0,0,0)'.$$

The perturbation from independent responses to correlated responses is what the author likes to consider for our ordinal response. Posterior normal curvature may be used for the effect of perturbation from correlated responses to independent responses. Here, $\omega = (\rho_{12}, \rho_{13}, \rho_{23})'$, $\omega_0 = (0,0,0)'$ and $q = 3$. Denote the log-posterior function by

$$P(\eta|\omega) = \sum_{i=1}^{n} P_i\,(\eta|\omega),$$

where $P_i(\eta|\omega)$ is the contribution of the ith individual to the log-posterior and η is the parameter vector. Here, $P(\eta|\omega_0)$ is the log-posterior function which corresponds to a normal distribution. Suppose ω perturbed around ω_0. Let $\hat{\eta}$ be the Bayesian estimator for η obtained by maximizing $P(\eta) = P(\eta|\omega_0)$ and let $\hat{\eta}_\omega$ denote the Bayesian estimator for η under $P(\eta|\omega)$. Now one compares $\hat{\eta}_\omega$ and $\hat{\eta}$ as local influences. Strongly different estimates show that the estimation procedure is highly sensitive to such modification. To search for sensitivity analysis, the author finds PC_{max}. This is confirmed by the curvature $PC_{max} = 10.76$ computed from (3). This curvature indicates an extreme local sensitivity.

7. Conclusion

In this paper the Bayesian elliptically contoured-ordinal model for mixed data and assessment of local influence via covariance is presented for simultaneously modeling of ordinal and continuous correlated responses. An elliptically contoured distribution is assumed for errors in the model. However, any other multivariate distribution such as *t* or logistic can be also used with and without missing responses.

REFERENCES

Arnold, B.C. and Beaver, R.J. (2000). Hidden Truncation Models. *Sankhya, Series A.* (1), 23-35.

Ash, R. B. and Dolens-Dade, C. A. (2000). *Probability and measure theory*, Academic Press.

Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distribution , *J. Roy. Statist. Soc.*, B 61, 579-602.

Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution, *Biometrika* 83, 715-726.

Catalano, P. and Ryan, L. M. (1992). Bivariate latent variable models for clustered discrete and continuous outcomes, *Journal of the American Statistical Association*, 3, 1078-109

Chip, R. C. and Jammalamadaka, S. R. (1988). Bayes prediction in regressions with elliptically errors, *Journal of Ecinometrics*, 349-360.

Cox, D. R. and Wermuth, N. (1992). Response models for mixed binary and quantitative variables. *Biometrika*, (3): 441-461.

Cook, R. D. (1986). Assessment of local influence, *Royal Statistical Society*, 48, 133-169.

De Leon, A. R. and Carrire, K. C. (2007). General mixed-data model: Extension of general location and grouped continuous models. *Canadian Journal of Statistics*, 35(4), 533-548.

Dunson D. B. (2000). Bayesian latent variable models for clustered mixed outcomes, *J R Stat Soc B*, 355-366.

Fang, K. T., Kotz, S. and Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions,* Chapman Hall, London.

Fernandez, C. and Steel, M. F. J. (1998). On Bayesian modeling of fat tails and skewness, *Journal of the American Statistical Association*, 359-371.

Geweke, J. (1993). Bayesian Treatment of the Independent Student- *t* Linear Model, *Journal of Applied Econometrics*, 19-40.

Heckman, J. J. D. (1978). Endogenous variable in a simultaneous Equation system, *Econometrica*, (6), 931-959.

Kelker, D. (1970). Distribution theory of spherical distributions and location –scale parameter generalization, *Sankhya*, 419-430.

McCulloch, C. (2007). Joint modeling of mixed outcome type using latent variables, *statistical methods in Medical Research*, 1-27.

Tierney, L. (1994). Markov chains for exploring posterior distributions, *Annals of Mathematical Statistics*, 1701-1786 .

Olkin L. and Tate R. F. (1961). Multivariate correlation models with mixed discrete and continuous variables, *Annals of Mathematical Statistics*, 448-456.

Sahu, S. K., Dey, D. K. and Branco, M. D. (2003). A New Class of Multivariate Skew Distributions with Applications to Bayesian Regression Models. *The Canadian Journal of Statistics,* 129-150.

Poon, W. Y. and LEE, S. Y. (1987). Maximum likelihood estimation of multivariate polyserial and polychoric correlation coefficients, *Psychometrika.*
(3): 409-430.

Yang, Y., Kang, J., Mao, K. and Zhang, J. (2007). Regression models for mixed Poisson and continuous longitudinal data. *Statistics in Medicine*, 3782-3800.

Zellner, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate student-t error terms, *J. Am. Stat. Assoc*., 400-405.