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

Applications and Applied Mathematics:
An International Journal
(AAM)

Vol. 13, Issue 1 (June 2018), pp. 19–29

Approximation to Multivariate Normal Higher Dimensional Probabilities in Mixed - Data Model with Missing Responses

E. Bahrami Samani

Department of Statistics
Shahid Beheshti University
Tehran, Iran
Ehsan_bahrami_samani@yahoo.com

Received: June 25, 2017; Accepted: December 23, 2018

Abstract

Multivariate data with mixed ordinal and continuous responses with the possibility of non-ignorable missingness are often common in follow up social studies and their analysis need to be promoted. One of the standard methods of analysis is based on joint modelling. In this method, we use that simultaneously allow modelling non-ignorable mechanism and a full likelihood-based approach is used to obtain maximum likelihood estimates of parameters of joint modelling. In this approach, when the dimension of the vector responses are increased, includes somehow troubling computations which are often time-consuming. Another alternative is an approximation for multivariate normal probabilities for higher dimensional regions, based conditional expectations in joint modelling. A comparison between approximation method and full likelihood-based approach is used to obtain maximum likelihood estimates of parameters of joint modelling. To illustrate the utility of the proposed model, a large data set excerpted from the British Household Panel Survey (BHPS) is analyzed. For these data, the simultaneous effects of some covariates on life satisfaction, income and the amount of money spent on leisure activities per month as three mixed correlated responses are explored.

Keywords: Approximation; Joint modelling; Latent variable; Mixed data; Ordinal and Continuous responses; Higher Dimensional probabilities; Missing Responses

MSC 2010 No.: 62F03, 62J05

1. Introduction

Higher dimensional probabilities from the multivariate normal distribution have many applications in statistics. These include the multivariate probit model (Ochi and Prentice (1984)), the multivariate ordinal responses model (Anderson and Pemberton (1985)), the multivariate mixed ordinal and continuous responses model. In general, these probabilities require multidimensional integrals, which can be evaluated by multidimensional numerical quadrature or Monte Carlo simulation, alternatively, approximations can be used if they are good enough. If estimation of parameters in one of the foregoing models is obtained using a quasi-Newton routine applied to the log-likelihood, then the use of Monte Carlo simulation to three or four decimal place accuracy of evaluation of integrals works poorly, because numerical derivatives of the log-likelihood with respect to parameters are not smooth. Numerical quadrature has computational time (and memory requirements) exponentially increasing in the dimension of the integral, so faster approximations are useful for parameter estimation. For example, a faster approximation method can be used to get preliminary estimates of the parameters, then, if necessary, one used a better approximation method or numerical quadrature, taking the preliminary estimate as a starting point, when computing the likelihood more accurately. Recent approximation for special cases have been given by Solow (1990) and Joe (1995). Earlier approximations are mentioned in the book by Tong (1990), including one by Mendell and Elston (1974) that is mentioned later in this article.

For joint modelling 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 of the categorical variables and a conditional multivariate normal distribution for the continuous variables, given the categorical variables (for a mixed Poisson and continuous responses where Olkin and Tate's method is used see Yang et al. (2007) and for joint modelling of mixed outcomes using latent variables see McCulloch (2007)). A second method for joint modelling is to decompose the joint distribution as a multivariate marginal distribution of 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 simultaneous modelling 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.

Rubin (1976), Little and Rubin (2002), and Diggle and Kenward (1994) made important distinctions between the various types of missing mechanisms for each of the above-mentioned patterns. They define the missing mechanism as missing completely at random (MCAR) if missingness is dependent neither on the observed responses nor on the missing responses, and missing at random (MAR) if, given the observed responses, it is not dependent on the missing responses. Missingness is defined as non-random if it depends on the unobserved responses. From a likelihood point of view MCAR and MAR are ignorable but not missing at random (NMAR) is non-ignorable. For mixed data with missing outcomes, Little and Schuchter (1987) and Fitzmaurice and Laird (1997) used the general location model of Olkin and Tate (1961) with the assumption of missingness at random (MAR) to justify ignoring the missing data mechanism.

The aim of this paper is to use and extend an approach similar to that of Joe (1995), for an approximation of the conditional probabilities in the joint model for mixed correlated continuous and ordinal responses with missing data. The model is described in terms of a correlated multivariate normal distribution for the underlying latent variables of ordinal responses and continuous responses. We compare an approximation method and full likelihood-based approach is used to obtain maximum likelihood estimates of parameters of our model.

In Section 2, we introduce briefly the model and the likelihood. In Section 3, an approximation conditional probabilities based on the moment-generating function of a truncated multivariate normal distribution is given. In Section 4, the proposed methodology is applied on the BHPS data. Finally, concluding remarks are given.

2. Model and Likelihood

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

$$Y_{ij} = \begin{cases} 1, & Y^*_{ij} < \theta_{1,j}, \\ k+1, & \theta_{k,j} \leq Y^*_{ij} < \theta_{k+1,j}, \quad k = 1, ..., c_j - 2, \\ c_j, & Y^*_{ij} \geq \theta_{c_j-1,j}, \end{cases}$$

where $i = 1, ..., n$, $j = 1, ..., M_1$. $\theta_{1j}, ..., \theta_{c_j-1,j}$ are the cut-point parameters and Y^*_{ij} denotes the underlying latent variable for Y_{ij}.

The ordinal response vector for the ith individual and the continuous response vector for the ith individual are denoted by $Y_i = (Y_{i1}, ..., Y_{iM_1})'$ and $Z_i = (Z_{i(M_1+1)}, ..., Z_{iM})'$. Typically, when missing data occur in an outcome, assume $R_{y_i} = (R_{y_{i1}}, ..., R_{y_{iM_1}})'$ as the indicator vector of corresponding to Y_i and $R_{y_{ij}}$ is defined as

$$R_{y_{ij}} = \begin{cases} 1, & R^*_{y_{ij}} > 0, \\ 0, & otherwise, \end{cases}$$

$R_{z_i} = (R_{z_{i(M_1+1)}}, ..., R_{z_{iM}})'$ is the indicator vector for corresponding to Z_i and $R_{z_{ij}}$ is defined as,

$$R_{z_{ij}} = \begin{cases} 1, & R^*_{z_{ij}} > 0, \\ 0, & otherwise, \end{cases}$$

where $R^*_{y_{ij}}$ and $R^*_{z_{ij}}$ denote the underlying latent variable of the non-response mechanism respectively, for the ordinal and continuous variables.

The joint model takes the form:

$$\begin{aligned} Y^*_{ij} &= \beta'_j X_i + \varepsilon^{(1)}_{ij}, & j &= 1, ..., M_1, \\ Z_{ij} &= \beta'_j X_i + \varepsilon^{(2)}_{ij}, & j &= M_1 + 1, ..., M, \\ R^*_{y_{ij}} &= \alpha'_j X_i + \varepsilon^{(3)}_{ij}, & j &= 1, ..., M_1, \\ R^*_{z_{ij}} &= \alpha'_j X_i + \varepsilon^{(4)}_{ij}, & j &= M_1 + 1, ..., M, \end{aligned}$$

where X_i is the design matrix for the ith individual. In the above presented model, the vector of parameters β_j for $j = 1, ..., M$, the parameters $\theta_{1j}, ..., \theta_{c_j-1,j}$ for $j = 1, ..., M$, should be estimated. The vector, β_j for $j = M_1 + 1, ..., M$, includes an intercept parameter but β_j, for $j = 1, ..., M_1$, due to having cutpoint parameters, are assumed not to include any intercept.

Let

$$\varepsilon_i = (\varepsilon_i^{(1)'}, \varepsilon_i^{(2)'}, \varepsilon_i^{(3)'}, \varepsilon_i^{(4)'})' \stackrel{iid}{\sim} MVN(0, \Sigma_\varepsilon),$$

where $\varepsilon_i^{(u)} = (\varepsilon_{i1}^{(u)}, ..., \varepsilon_{iM_1}^{(u)})'$, for $u = 1, 3$, $\varepsilon_i^{(u)} = (\varepsilon_{i(M_1+1)}^{(u)}, ..., \varepsilon_{iM}^{(u)})'$, for $u = 2, 4$ and

$$\Sigma_\varepsilon = \begin{pmatrix} \Sigma_{11}^\varepsilon & \Sigma_{12}^\varepsilon & \Sigma_{13}^\varepsilon & \Sigma_{14}^\varepsilon \\ \Sigma_{21}^\varepsilon & \Sigma_{22}^\varepsilon & \Sigma_{23}^\varepsilon & \Sigma_{24}^\varepsilon \\ \Sigma_{31}^\varepsilon & \Sigma_{32}^\varepsilon & \Sigma_{33}^\varepsilon & \Sigma_{34}^\varepsilon \\ \Sigma_{41}^\varepsilon & \Sigma_{42}^\varepsilon & \Sigma_{43}^\varepsilon & \Sigma_{44}^\varepsilon \end{pmatrix},$$

where $\Sigma_{uu}^\varepsilon = Var(\varepsilon_{\mathbf{i}}^{(\mathbf{u})})$, for $u = 1, 2, 3, 4$ and $\Sigma_{uv}^\varepsilon = Cov(\varepsilon_{\mathbf{i}}^{(\mathbf{u})}, \varepsilon_{\mathbf{i}}^{(\mathbf{v})})$, $u < v$, $u, v = 1, 2, 3, 4$ and $\Sigma_{uv}^\varepsilon = {\Sigma_{vu}^\varepsilon}'$. Because of identifiability problem we have to assume $Var(Y_{ij}^*) = Var(R_{y_{ij}}^*) = Var(R_{y_{ij}}^*) = 1$, so $\Sigma_{jj}^\varepsilon = I$ for $j = 1, 3, 4$.

Note, if one of the matrices $\Sigma_{13}^\varepsilon, \Sigma_{14}^\varepsilon, \Sigma_{23}^\varepsilon, \Sigma_{24}^\varepsilon$ is not zero, then the missing mechanism of response is not at random.

We start with some notation. Let

$$\begin{aligned}
J_{obs}^y &= \{j:\ y_{ij} \text{ is observed}\}, & J_{Mis}^y &= (J_{obs}^y)^C, \\
J_{obs}^z &= \{j:\ z_{ij} \text{ is observed}\}, & J_{Mis}^z &= (J_{obs}^z)^C, \\
Z_{i,obs} &= \{Z_{ij}, \forall j \in J_{obs}^z\}, & Y_{i,obs} &= \{Y_{ij}, \forall j \in J_{obs}^y\}, \\
Y_{i,obs}^* &= \{\theta_{y_{ij}-1} \le Y_{ij}^* \le \theta_{y_{ij}}, \forall j \in J_{obs}^y\}, & R_{z_{i,obs}} &= \{R_{z_{ij}} = 1, \forall j \in J_{obs}^z\}, \\
R_{y_{i,obs}} &= \{R_{y_{ij}} = 1, \forall j \in J_{obs}^y\}, & R_{z_{i,obs}}^* &= \{R_{z_{ij}}^* > 0, \forall j \in J_{obs}^z\}, \\
R_{z_{i,Mis}}^* &= (R_{z_{i,obs}}^*)^C, & R_{y_{i,obs}}^* &= \{R_{y_{ij}}^* > 0, \forall j \in J_{obs}^y\}, \\
R_{y_{i,Mis}}^* &= (R_{y_{i,obs}}^*)^C, & R_{y_i} &= (R_{y_{i1}}, ..., R_{y_{iM_1}})', \\
R_{z_i} &= (R_{z_{i(M_1+1)}}, ..., R_{z_{iM}})', & C_{Mis}^* &= \{R_{y_{ij}} = 0;\ \forall j \in J_{Mis}^y, R_{z_{ij}} = 0;\ \forall j \in J_{Mis}^z\}.
\end{aligned}$$

The likelihood for this model is

$$L = \prod_{i=1}^{n} P(Y_{i,obs}^*, R_{y_{i,obs}}^*, R_{z_{i,obs}}^*, C_{Mis}^* | Z_{i,obs}, X_i)\, f(Z_{i,obs}|X_i),$$

where

$$P(Y_{i,obs}^*, R_{y_{i,obs}}^*, R_{z_{i,obs}}^*, C_{Mis}^* | Z_{i,obs}, X_i) = \gamma_i^* - \Gamma_{y_i}^* - \Gamma_{z_i}^* + \Gamma_{y_i z_i}^*,$$

and

$$\begin{aligned}\gamma_i^* &= P(Y_{i,obs}^*, C_{Mis}^* | Z_{i,obs}, X_i),\\ \Gamma_{z_i}^* &= P(Y_{i,obs}^*, C_{Mis}^*, R_{z_{i,Mis}}^* | Z_{i,obs}, X_i),\\ \Gamma_{y_i}^* &= P(Y_{i,obs}^*, C_{Mis}^*, R_{y_{i,Mis}}^* | Z_{i,obs}, X_i),\\ \Gamma_{y_i z_i}^* &= P(Y_{i,obs}^*, C_{Mis}^*, R_{y_{i,Mis}}^*, R_{z_{i,Mis}}^* | Z_{i,obs}, X_i).\end{aligned}$$

We use the following definition and theorem for multivariate distributions to obtain the form of the likelihood.

Definition 2.1.

If $F(w_1, \cdots, w_{M_1}) = P(W_1^* \leq w_1, ..., W_{M_1}^* \leq w_{M_1})$ is a distribution function, operator $\Delta_{b_j a_j} F(w_1, \cdots, w_{M_1})$ is defined as, $(a_j \leq b_j)$,

$$F(w_1, \cdots, w_{(j-1)}, b_j, w_{(j+1)}, ..., w_{M_1}) - F(w_1, \cdots, w_{(j-1)}, a_j, w_{(j+1)}, ..., w_{M_1}).$$

Theorem 2.2.

If for $j = 1, ..., M_1$, $a_j \leq b_j$, then

$$P(a_1 < W_1^* \leq b_1, ..., a_{M_1} < W_{M_1}^* \leq b_{M_1}) = \Delta_{b_1 a_1} \cdots \Delta_{b_{M_1} a_{M_1}} F(w_1, \cdots, w_{M_1}),$$

where

$$\Delta_{b_1 a_1} \cdots \Delta_{b_{M_1} a_{M_1}} F(w_1, \cdots, w_{M_1}) = F_0 - F_1 + F_2 - ... + (-1)^{M_1} F_{M_1},$$

and F_j is the sum of all $\binom{M_1}{j}$ terms of the from $F(g_1, ..., g_{M_1})$ with $g_k = a_k$ for exactly j integers in $\{1, ..., M_1\}$, and $g_k = b_k$ for the remaining $M_1 - j$ integers.

Proof:

See Ash and Dolens (2000). ■

Suppose the g_1- elements of Y_i and g_2- elements of Z_i are observed, so $J_{obs}^y = \{o_1, ..., o_{g_1}\}$ and $J_{obs}^z = \{o_1, ..., o_{g_2}\}$. Using the Theorem 2.2,

$$\begin{aligned}\gamma_i^* &= \Delta_{b_{o_1} a_{o_1}} ... \Delta_{b_{og_1} a_{og_1}} P(Y_{io_1}^* \leq \omega_{io_1}, ..., Y_{iog_1}^* \leq \omega_{iog_1}, C_{Mis}^* | Z_{i,obs}, X_i),\\ &= F_{i0}^{(1)} - F_{i1}^{(1)} + F_{i2}^{(1)} - ... + (-1)^{g_1} F_{ig_1}^{(1)},\end{aligned}$$

$$\begin{aligned}\Gamma_{z_i}^* &= \Delta_{b_{o_1} a_{o_1}} ... \Delta_{b_{og_1} a_{og_1}} P(Y_{io_1}^* \leq \omega_{io_1}, ..., Y_{iog_1}^* \leq \omega_{iog_1}, C_{Mis}^*, R_{z_{i,Mis}}^* | Z_{i,obs}, X_i),\\ &= F_{i0}^{(2)} - F_{i1}^{(2)} + F_{i2}^{(2)} - ... + (-1)^{g_1} F_{ig_1}^{(2)},\end{aligned}$$

$$\begin{aligned}\Gamma_{y_i}^* &= \Delta_{b_{o_1} a_{o_1}} ... \Delta_{b_{og_1} a_{og_1}} P(Y_{io_1}^* \leq \omega_{io_1}, ..., Y_{iog_1}^* \leq \omega_{iog_1}, C_{Mis}^*, R_{y_{i,Mis}}^* | Z_{i,obs}, X_i),\\ &= F_{i0}^{(3)} - F_{i1}^{(3)} + F_{i2}^{(3)} - ... + (-1)^{g_1} F_{ig_1}^{(3)},\end{aligned}$$

$$\begin{aligned}\Gamma_{y_i z_i}^* &= \Delta_{b_{o_1} a_{o_1}} ... \Delta_{b_{og_1} a_{og_1}} P(Y_{io_1}^* \leq \omega_{io_1}, ..., Y_{iog_1}^* \leq \omega_{iog_1}, C_{Mis}^*, R_{y_{i,Mis}}^*, R_{z_{i,Mis}}^* | Z_{i,obs}, X_i),\\ &= F_{i0}^{(4)} - F_{i1}^{(4)} + F_{i2}^{(4)} - ... + (-1)^{g_1} F_{ig_1}^{(4)},\end{aligned}$$

where $b_{o_j} = \theta_{o_j y_{io_j}}$ and $a_{o_j} = \theta_{o_j(y_{io_j}-1)}$ for $j = 1, ..., g_1$ and $F_{ij}^{(1)}, F_{ij}^{(2)}, F_{ij}^{(3)}$ and $F_{ij}^{(4)}$ are the sum of all $\begin{pmatrix} g_1 \\ j \end{pmatrix}$ terms of the form

$$\begin{aligned}
&P(Y_{ic_1}^* \le c_1, ..., Y_{ic_{g_1}}^* \le c_{g_1}, C_{Mis}^* | Z_{i,obs}, X_i),\\
&P(Y_{ic_1}^* \le c_1, ..., Y_{ic_{g_1}}^* \le c_{g_1}, C_{Mis}^*, R_{z_{i,Mis}}^* | Z_{i,obs}, X_i),\\
&P(Y_{ic_1}^* \le c_1, ..., Y_{ic_{g_1}}^* \le c_{g_1}, C_{Mis}^*, R_{y_{i,Mis}}^* | Z_{i,obs}, X_i),\\
&P(Y_{ic_1}^* \le c_1, ..., Y_{ic_{g_1}}^* \le c_{g_1}, C_{Mis}^*, R_{y_{i,Mis}}^*, R_{z_{i,Mis}}^* | Z_{i,obs}, X_i),
\end{aligned}$$

with $c_k = a_k$ for exactly j integers in $\{1, ..., g_1\}$, and $c_k = b_k$ for the remaining $g_1 - j$ integers.

3. Approximation

Suppose the g_1- elements of Y_i and g_2- elements of Z_i are observed, so $J_{obs}^y = \{y_1, ..., y_{g_1}\}$, $J_{Mis}^y = \{q_1, ..., q_{M_1-g_1}\}$, $J_{obs}^z = \{z_1, ..., z_{g_2}\}$, $J_{Mis}^z = \{p_1, ..., p_{M-g_2-M_1-1}\}$ and $I_{ijl} = I_{\{\theta_{j,y_{ij}-1} \le Y_{ij}^* \le \theta_{j,y_{ij}}, R_{y_{ij}}=1, R_{z_{il}}=1 | z_{i,obs}\}}$, $i = 1, ..., n$, $j \in J_{obs}^y$ and $l \in J_{obs}^z$, and $J_{ijl} = J_{\{R_{y_{ij}}=0, R_{z_{il}}=0 | z_{i,obs}\}}$,$i = 1, ..., n$, $j \in J_{Mis}^y$ and $l \in J_{Mis}^z$, where $I(A)$ and $J(B)$ denotes the indicator of the even A and B.

For an approximation of the likelihood function, we are interested in the probability

$$P(Y_{i,obs}^*, R_{y_{i,obs}}^*, R_{z_{i,obs}}^*, C_{Mis}^* | Z_{i,obs}, X_i),$$

which can be decomposed as the product of conditional probabilities,

$$\begin{aligned}
&E(I_{iy_1z1}) E(J_{iq_1p_1} | I_{iy_1z_1} = 1,..., I_{i,y_{g_1},z_{g_2}} = 1)\\
&\times \prod_{j=y_2}^{y_{g_1}} \prod_{l=z_2}^{z_{g_2}} E[I_{ijl} \, | I_{iy_1z_1} = 1, ..., I_{i,j-1,l-1} = 1]\\
&\times \prod_{j'=q_2}^{Q} \prod_{l'=p_2}^{P} E[J_{ijl} \left| I_{iy_1z_1} = 1, ..., I_{i,y_{g_1},z_{g_2}} = 1 \right.,\\
&\qquad J_{iq_1p_1} = 1, ..., J_{i,j'-1,l'-1} = 1],
\end{aligned}$$

where $Q = q_{M_1-g_1}$ and $P = p_{M-g_1-M_1-1}$. Now, the expectations in the right hand can be, approximated as (vide Joe (1995)), we have

$$\begin{aligned}
&E(J_{iq_1p_1} | I_{iy_1z_1} = 1,..., I_{i,y_{g_1},z_{g_2}} = 1) \simeq\\
&\qquad E(J_{iq_1p_1}) + \Omega_{21}\Omega_{11}^{-1}(1 - E(I_{iy_1z_1}), ..., 1 - E(I_{i,y_{g_1},z_{g_2}}))',\\
&E[I_{ijl} \, | I_{iy_1z_1} = 1,..., I_{i,j-1,l-1} = 1] ,\\
&\simeq E(I_{ijl}) + \Delta_{21}\Delta_{11}^{-1}(1 - E(I_{iy_1z_1}), ..., 1 - E(I_{i,j-1,l-1}))'\\
&\times E[J_{ijl} \left| I_{iy_1z_1} = 1,..., I_{i,y_{g_1},z_{g_2}} = 1, J_{iq_1p_1} = 1, ..., J_{i,j'-1,l'-1} = 1 \right.] ,\\
&\simeq E(I_{ijl}) + \Gamma_{21}\Gamma_{11}^{-1}(1 - E(I_{iy_1z_1}), ..., 1 - E(I_{i,j-1,l-1}), 1 - E(J_{iq_1p_1}),\\
&, ..., 1 - E(J_{i,j'-1,l'-1}))',
\end{aligned}$$

where Ω_{21} is a row vector consisting of the entries $cov(I_{iy_1zl}, I_{ihk})$, $h = y_1, ..., y_{g_1}$ and $k = z_1, ..., z_{g_2}$, Δ_{21} is a row vector consisting of the entries $cov(I_{iy_1zl}, I_{ihk})$, $h = y_1, ..., j-1$ and $k = z_1, ..., l-1$ and Γ_{12} is a row vector consisting of the entries $(cov(I_{iy_1zl}, I_{ihk}), cov(J_{iq_1pl}, J_{ih'k'}))'$, $h = y_1, ..., j-1$, $k = z_1, ..., l-1$ and $h' = q_1, ..., j'-1$, $k' = p_1, ..., l'-1$.

Ω_{11} is a $(y_{g_1}+z_{g_2})\times(y_{g_1}+z_{g_2})$ matrix with (j,l) element $cov(I_{ijl}, I_{ihk})$, $h = y_1, ..., y_{g_1}$, $k = z_1, ..., z_{g_2}$, Δ_{11} is a $a \times a$, $a = j+l-2$, matrix with (j,l) element $cov(I_{ijl}, I_{ihk})$, $h = y_1, ..., j-1$, $k = z_1, ..., l-1$ and Γ_{11} is a $b \times b$, $b = j+l+j'+l'-4$, matrix with (j,l) element $(cov(I_{ijl}, I_{ihk}), cov(J_{iq_1pl}, J_{ih'k'}))'$, $h = y_1, ..., j-1$, $k = z_1, ..., l-1$ and $h' = q_1, ..., j'-1$, $k' = p_1, ..., l'-1$.

So

$$\begin{aligned}L \simeq \prod_{i=1}^{n} & f(Z_{i,obs}|X_i)E(I_{iy_1z1})[E(J_{iq_1p_1}) \\ & +\Omega_{21}\Omega_{11}^{-1}(1-E(I_{iy_1z_1}), ..., 1-E(I_{i,y_{g_1},z_{g_2}}))')] \\ & \times \prod_{j=y_2}^{y_{g_1}} \prod_{l=z_2}^{z_{g_1}} [E(I_{ijl}) + \Delta_{21}\Delta_{11}^{-1}(1-E(I_{iy_1z_1}), ..., \\ & 1-E(I_{i,j-1,l-1}))'] \times \prod_{j'=q_2}^{Q} \prod_{l'=p_2}^{P} [E(I_{ijl}) + \Gamma_{21}\Gamma_{11}^{-1} \\ & \times(1-E(I_{iy_1z_1}), ..., 1-E(I_{i,j-1,l-1}), 1-E(J_{iq_1p_1}))'].\end{aligned}$$

4. Application

4.1. Data

The data used in this paper is excerpted from the 15th wave (2005) of the British Household Panel Survey (BHPS); a longitudinal survey of adult Britons, being carried out annually since 1991 by the ESRC UK Longitudinal Studies Center with the Institute for Social and Economic Research at the University of Essex. These data are recorded for 11251 individuals. The selected variables which will be used in this application are explained in the following. One of the responses is the life Satisfaction (LS), which is measured by directly asking the level of an individual's satisfaction with life overall, resulting in three categories ordinal variable [1: Not satisfied at all (10.300%). 2: Not satisfied/dissent (45.400%) and 3: Completely satisfied (44.300 %)]. In our application, the percentage of missing values of LS is 5.000%. The amount of money spent on leisure activities per month including money spent on entertainment and hobbies (AM) is also measured as an ordinal response with three categories, [0: Nothing (17.515 %). 1: Under 50 Pound (53.449%) and 2: 50 Pound or over (29.036%.)]. Moreover, the exact amount of individuals annual income (INC) in the past year in thousand pounds, considered here in the logarithmic scale, is also excerpted as a continuous response variable (mean: 4.068). As some values of annual income of thousand pounds is between 0 and 1, some of the logarithms of incomes are less than 0. These three responses, LS, AM and the logarithm of income are endogenous correlated variables and should be modelled as a multivariate vector of responses.

Socio-demographic characteristics, namely: Gender (male: 44.200% and female: 55.800%), Marital Status (MS)[married or living as a couple: 68.500%, widowed: 8.300%, divorced or separated: 8.400% and never married: 14.800%], Age (mean: 49.180) and Highest Educational Qualification

(HEQ) [higher or first degree: 15.100%, other higher QF: 64.600%, other QF: 2.000% and no qualification: 18.300%] are also included in the model as covariates. The vector of explanatory variables is

$$X = (Gender, Age, MS_1, MS_2, MS_3, HEQ_1, HEQ_2, HEQ_3),$$

where MS_1, MS_2 and MS_3 are dummy variables for married or living as a couple, widowed and divorced or separated, respectively, and HEQ_1, HEQ_2, HEQ_3 are dummy variables for higher or first degree, other higher QF, and other QF, respectively.

4.2. Models for BHPS Data

We apply the model described in section 2 to evaluate the effect of Age, Gender, HEQ and MS simultaneously on LS, AM and Income. The model is

$$LS^* = \beta_{11}\ MS_1 + \beta_{12}\ MS_2 + \beta_{13}\ MS_3 + \beta_{14}\ HEQ_1 + \beta_{15}\ HEQ_2 + \beta_{16}\ HEQ_3$$

$$+\beta_{17}Gender + \beta_{18}AGE + \varepsilon_1,$$

$$AM^* = \beta_{21}\ MS_1 + \beta_{22}\ MS_2 + \beta_{23}\ MS_3 + \beta_{24}\ HEQ_1 + \beta_{25}\ HEQ_2 + \beta_{26}\ HEQ_3$$

$$+\beta_{27}Gender + \beta_{28}AGE + \varepsilon_2,$$

$$\log(INC) = \beta_{30} + \beta_{31}\ MS_1 + \beta_{32}\ MS_2 + \beta_{33}\ MS_3 + \beta_{34}\ HEQ_1 + \beta_{35}\ HEQ_2$$

$$+\beta_{36}\ HEQ_3 + \beta_{37}Gender + \beta_{38}AGE + \varepsilon_3,$$

$$R^*_{LS} = \alpha_{11}\ MS_1 + \alpha_{12}\ MS_2 + \alpha_{13}\ MS_3 + \alpha_{14}\ HEQ_1 + \alpha_{15}\ HEQ_2 + \alpha_{16}\ HEQ_3$$

$$+\alpha_{17}Gender + \alpha_{18}AGE + \varepsilon_4.$$

For this model the covariance matrix takes the form,

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \sigma\rho_{13} & \rho_{14} \\ \rho_{21} & 1 & \sigma\rho_{23} & \rho_{24} \\ \sigma\rho_{31} & \sigma\rho_{32} & \sigma^2 & \sigma\rho_{34} \\ \rho_{14} & \rho_{24} & \sigma\rho_{34} & 1 \end{pmatrix}.$$

Here, a multivariate normal distribution with correlation parameters $\rho_{12}, \rho_{13}, \rho_{14}, \rho_{23}, \rho_{24}$ and ρ_{34} is assumed for the errors and these parameters should be also estimated.

4.3. Results

In the first method (full likelihood approach) and the second method (approximation) are used to obtain maximum likelihood estimates of parameters of the model for BHPS data. In this model show a significant effect of age (the older the individual the more the life satisfaction), MS (married people are more satisfied than never-married people and divorced or separated people are less satisfied than never-married people), HEQ (the higher the qualification the higher the life satisfaction) and gender (males are more satisfied than females) on the life satisfaction status. All explanatory variables have significant effect on the ordinal the response of amount of money spent on leisure activities. Never married people spend more on leisure time activities than other people. The higher the education the more the leisure time activities. Females spend more amount of money than males for leisure time and older people spend less money than younger ones. Also, the effect of all explanatory variables are significant on the logarithm of income. Parameter estimates indicate that as the degree of educational qualification increases log(INC) increases. Never married people have less the logarithm of income than married people and divorced or separated people. Females have more logarithm of income than males and the older people earn less money than younger ones. By these results, we can conclude that the two responses are correlated and also the missing indicator for LS is not related to three responses. This leads to having an at random missing mechanism (A missing at random mechanism for LS which means no correlation between error terms of (LS, R^*_{LS}), (AM, R^*_{LS}) and ($\log(Income)$, R^*_{LS}).

In the first method (full likelihood approach) has computational time and memory requirements exponentially increasing in the dimension of the vector of responses (computational time: 7 days), but the second method (approximation)is faster than the first method (full likelihood approach), a faster approximation method can be used to get preliminary estimates of the parameters, then, if necessary, one used a better approximation method, taking the preliminary estimate as a starting point, when computing the likelihood more accurately (computational time: 3 days).

5. Conclusion

In this paper an approximation for multivariate latent variable model is presented for simultaneously modelling of ordinal and continuous correlated responses. We assume a multivariate normal distribution of errors in the model. However, any other multivariate distribution such as t or logistic can also be used. Binary responses are a special case of ordinal responses. So, our the model can also be used for mixed binary and continuous responses. For correlated nominal, ordinal and continuous responses Deleon and Carriere (2007) have developed a model by extending general location model. However, the kind of scientific question they can answer is different with what our model can do (vide Section 1). Generalization of our model for nominal, ordinal and continuous responses is an ongoing research on our part.

REFERENCES

Anderson, J. A., and Pemberton, J. D. (1985). The grouped continuous model for multi-variate orders categorical variables and covariate adjustment, Biometrics, Vol. 41, pp. 875–885.

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

Catalano, P. and Ryan, L. M. (1992). Bivariate latent variable models for clustered discrete and continuous outcomes, Journal of the American Statistical Association, Vol. 50, No. 3, pp. 1078–1095.

Cox, D. R., and Wermuth, N. (1992). Response models for mixed binary and quantitative variable, Biometrika, Vol. 79, No. 3, pp. 441–461.

Deleon, A. R. and Carriere, K. C. (2007). General mixed-data model: Extension of general location and group continuous models, The Canadian Journal of Statistics, Vol. 35, No. 4, pp. 533–548.

Diggle, P.J. and Kenward, M. G. (1994). Informative drop-out in longitudinal data analysis, Journal of Applied Statistics, Vol. 43, pp. 49–93.

Fitzmaurice, G. M. and Laird, N. M. (1997). Regression models for mix discrete and continuous responses with potentially missing values, Biometrics, Vol. 53, pp. 110–122.

Joe, H. (1995). Approximations multivariate normal rectangle probabilities based on conditional expectation, Journal of the American Statistical Association, Vol. 90, pp. 957–967.

Heckman, J. J. (1978). Dummy endogenous variable in a simultaneous equation system. Econometrica, Vol. 46, No. 6, pp. 931–959.

Little, R. J. and Schluchter M. (1987). Maximum likelihood estimation for mixed continuous and categorical data with missing values, Biometrika, Vol. 72, pp. 497–512.

Little, R. J. and Rubin, D. (2002). *Statistical analysis with missing data,* second edition, Wiley, New York.

McCulloch, C. (2007). Joint modelling of mixed outcome type using latend variables, Statistical Methods in Medical Research , pp. 1–27.

Mendell, N. R. and Elston, R. C. (1974). Multifactorial qualitative traits: genetic analysis and prediction of recurrence risk, Biometrics, Vol. 30, pp. 41–57.

Rubin, D. B. (1976). Inference and missing data, Biometrica, Vol. 82, pp. 669–710.

Ochi, Y. and Prentice, R. L. (1984). Likelihood inference correlated probit regression model, Biomertrika, Vol. 71, pp. 531–543.

Olkin L. and Tate R. F. (1961). Multivariate correlation models with mixed discrete and continuous variables, Annals of Mathematical Statistics, Vol. 32, pp. 448–456.

Solow, A. R. (1990). A method for approximating multivariate normal orthant probabilities, Journal of Statistical Computational and Simulation, Vol. 37, pp. 225–229.

Tong, Y. L. (1990). *The multivariate normal distribution*, Springer-Verlag, New York.

Yang, Y., Kang, J., Mao, K. and Zhang, J. (2007). Regression models for mixed Poisson and continuous longitudinal data, Statistics in Medicine, Vol. 26, pp. 3782–3800.

Table 1. Results of using two models for BHPS data (FL: Full likelihood approach and Appro.: Approximation, parameter estimates highlighted in **bold** are significant at 5 % level.)

method(computational time)	Appro.		FL	
[3ex] parameter	Est.	S.D.	Est.	S.D.
[3ex] Response: LS				
[3ex] MS (baseline: Never married)				
Married or Living as Couple	**0.203**	**0.030**	**0.202**	**0.029**
Widowed	0.031	0.041	0.032	0.039
Divorced or Separated	**-0.334**	**0.046**	**-0.358**	**0.044**
HEQ (baseline: No QF)				
Higher or First Degree	**- 0.060**	**0.097**	**-0.061**	**0.096**
Other higher QF	**-0.206**	**0.373**	**-0.131**	**0.096**
Other QF	**- 0.212**	**0.110**	**-0.210**	**0.100**
Gender (baseline: Female)				
Male	**0.042**	**0.027**	**0.040**	**0.021**
AGE	**0.009**	**0.002**	**0.008**	**0.001**
cutpoint 1	**-1.000**	**0.113**	**-0.987**	**0.102**
cutpoint 2	**0.466**	**0.131**	**0.450**	**0.119**
[3ex] Response: AM				
[3ex] Married or Living as Couple	**-0.190**	**0.030**	**-0.185**	**0.032**
Widowed	**-0.236**	**0.057**	**-0.235**	**0.055**
Divorced or Separated	**-0.262**	**0.042**	**-0.259**	**0.047**
Higher or First Degree	**0.582**	**0.085**	**0.578**	**0.095**
Other higher QF	**0.290**	**0.084**	**0.288**	**0.091**
Other QF	**-0.082**	**0.089**	**-0.079**	**0.088**
Gender (baseline: Female)				
Male	**-0.462**	**0.020**	**-0.449**	**0.022**
AGE	**-0.015**	**0.003**	**-0.014**	**0.001**
cutpoint 1	**-2.353**	**0.125**	**-2.331**	**0.105**
cutpoint 2	**-0.701**	**0.124**	**-0.690**	**0.122**
[3ex] Response: $\log(INC)$				
Constant	**4.246**	**0.038**	**4.245**	**0.039**
Married or Living as Couple	**0.114**	**0.011**	**0.113**	**0.011**
Widowed	**0.217**	**0.019**	**0.216**	**0.019**
Divorced or Separated	**0.215**	**0.077**	**0.216**	**0.016**
Higher or First Degree	**0.391**	**0.037**	**0.393**	**0.036**
Other higher QF	**0.177**	**0.038**	**0.178**	**0.035**
Other QF	0.032	0.037	0.033	0.035
Male	**-0.227**	**0.007**	**-0.225**	**0.007**
AGE	**-0.002**	**0.001**	**-0.003**	**0.001**
σ^2	**0.154**	**0.002**	**0.155**	**0.002**
[3ex] Response: R^*_{LS}				
Cutpoint	**0.958**	**0.028**	**1.458**	**0.010**
Married or Living as Couple	**0.017**	**0.004**	**0.016**	**0.003**
Widowed	-0.003	0.008	-0.004	0.007
Divorced or Separated	**0.0020**	**0.007**	**0.002**	**0.016**
Higher or First Degree	**0.056**	**0.027**	**0.054**	**0.018**
Other higher QF	**0.051**	**0.020**	**0.051**	**0.017**
Other QF	0.024	0.027	0.023	0.018
Male	-0.005	0.003	-0.004	0.018
AGE	**-0.008**	**0.001**	**-0.008**	**0.002**
[3ex] Corr(LS*,AM*)	**0.137**	**0.014**	**0.136**	**0.012**
Corr (LS *,INC)	0.001	0.002	0.001	0.001
Corr(AM *,INC)	**0.139**	**0.012**	**0.138**	**0.010**
Corr(R^*_{LS},LS)	-0.008	0.019	-0.008	0.020
Corr(R^*_{LS},AM)	-0.018	0.011	-0.019	0.015
Corr(R^*_{LS},INC)	-0.048	0.041	-0.048	0.040
-Loglike	34109.030		34107.030	