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

Applications and Applied Mathematics:
An International Journal
(AAM)

Vol. 10, Issue 1 (June 2015), pp. 81 - 105

Gaussian Copula Mixed Models with Non-Ignorable Missing Outcomes

N. Jafari, E. Tabrizi and E. Bahrami Samani*
Department of Statistics
Faculty of Mathematical Science
Shahid Beheshti University
Tehran, 1983963113, Iran

* Correspondence author, e-mail: ehsan_bahrami_samani@yahoo.com

Received: September 24, 2014; Accepted: November 21, 2014

Abstract

This paper is concerned with the analysis of mixed data with ordinal and continuous outcomes with the possibility of non-ignorable missing outcomes. A copula-based regression model is proposed that accounts for associations between ordinal and continuous outcomes. Our approach entails specifying underlying latent variables for the mixed outcomes to indicate the latent mechanisms which generate the ordinal and continuous variables. Maximum likelihood estimation of our model parameters is implemented using standard software such as function nlminb in R. Results of simulations concern the relative biases of parameter estimates of joint and marginal models using data with non-ignorable outcomes. The proposed methodology is illustrated using a medical data obtained from an observational study on women with three correlated responses, an ordinal response of osteoporosis of the spine and two continuous responses of body mass index and waistline. The effect of the amount of total body calcium (Ca), job status (Job), type of dwelling (Ta) and age on all responses are investigated simultaneously.

Keywords: Nonignorable missing outcomes; Mixed outcomes; Latent Variables; Likelihood-based; Gaussian copula.

AMS (2010) No.: 62F03

1. Introduction

Many statistical applications involve joint analysis of multivariate data including mixed ordinal and continuous outcomes with non-ignorable missing values. For example, in health studies pertaining to the maternal smoking effect on respiratory illness in children, we have a continuous measure of the pulmonary function and an ordinal measure of chronic symptoms in children. In medical data set of osteoporosis of the spine, correlated outcomes are the ordinal outcome of osteoporosis of the spine and continuous outcomes of body mass index and waistline and covariates that might be due to this type of job and dwelling.

For the first example, separate analysis cannot assess the effect of maternal smoking on both outcomes. Also, separate analysis give biased estimates for the parameters and we need to consider a method in which these variables can be modelled jointly. So, we need to model responses simultaneously. In the second example the simultaneous effect of the type of job and accommodation on body mass index, waistline and osteoporosis of the spine should be modelled jointly considering missing mechanisms for each outcomes. Multivariate joint modelling of such missing data often leads to complications in computation due to a relative lack of standard models.

A number of joint modelling strategies for mixed outcomes have been studied in the literature. The first formulation that has received much attention in mixed data literature was introduced by Olkin and Tate's (1961) which is called general location model. This model assumes multivariate normal distribution for continuous outcomes given values of discrete outcomes. The second formulation includes the Cox and Wermuth (1992) approach who suggest a logistic or probit conditional distribution for the binary variable given continuous outcomes. The third formulation was presented by Heckman (1978) in which a general model for simultaneously analysing two mixed correlated responses is introduced and Catalano and Ryan (1997) extended and used the model for a cluster of discrete and continuous outcomes (vide also, Fitzmaurice and Laird, (1995) and Fitzmaurice and Laird, (1997)). All the above references consider correlated nominal and continuous responses. Poon and Lee (1987) and Moustaki and Knott (2000) used a model for ordinal and continuous responses without considering any covariate effect. De Leon an Carriere (2007) extended an approach similar to that of Heckman (1978) and Sammel et al. (1997) for jointly modelling of a nominal and a continuous variable to joint modelling of bivariate ordinal and continuous outcomes. All the above references discuss identifiability with imposing some restrictions on the correlation structure. Pinto and Normand (2009) proposed a new parametric constrained latent variable model to have identifiability without restrictions on the correlation structure.

In such medical studies, often some of the subjects do not respond in some occasions which cause for missing outcomes. Much has been written about statistical methods for handling incomplete data. Rubin (1976) and Little and Rubin (2002) define the missing mechanism as: (1) Missing Completely At Random (MCAR): if missingness is dependent neither on the observed responses nor on the missing responses, (2) Missing At Random (MAR): if it is not dependent on the missing responses (given the observed responses), (3) Not Missing At Random (NMAR): if it depends on the unobserved responses. MCAR and MAR are ignorable but NMAR is non-ignorable.

A number of joint modelling strategies for mixed outcomes with possibility of missing values have been studied in the literature. Little and Schluchter (1987) proposed the general location model with the assumption of ignoring the missing data mechanism. Ganjali (2003), used A model for mixed continuous and discrete binary responses with possibility of missing responses. Bahrami Samani et al. (2008, 2010) extended the model of Ganjali. Also, Bahrami Samani et al. (2011) proposed a multivariate latent random effect model for mixed continuous and ordinal longitudinal responses with missing responses. Yang et al. (2007) investigate an inferential method for mixed Poisson and continuous longitudinal data with non-ignorable missing values. The challenge is that models for joint distributions of mixed outcomes with non-ignorable missing values are uncommon.

A recent alternative strategy involves the use of copulas, as discussed in Sklar (1959), Song et al. (2000), Niewiadomska-Bugaj and Kowalczyk (2005), Zimmer and Trivedi (2006), Kolev et al. (2006) and Song et al. (2009). A number of transition regression models for non-Gaussian responses have been proposed in literature, vide Benjamin et al. (2003) for a review. Several authors have recently adopted copulas to indirectly construct mixed-outcome joint models. Copulas have been proved to be useful in practice when the joint distribution of interest is either not available or difficult to specify but marginal distributions can be specified with confidence like in mixed-outcome settings. Song et al. (2000) investigate some copula-based regression models for bivariate continuous outcomes, Zimmer and Trivedi (2006) proposed trivariate copulas to model sample selection and treatment effects. De Leon and Wu (2011) proposed copula-based regression models for bivariate mixed discrete and continuous outcomes.

Our paper is concerned with joint regression models for correlated mixed ordinal and continuous outcomes with possibility of non- ignorable missing outcomes constructed by using copulas. We will also extend De Leon and Wu' (2011)'s approach and consider missing data of the outcomes, so our models are copula-based joint modelling of mixed data for bivariate and multivariate mixed ordinal and continuous outcomes with non-ignorable missing outcomes.

This paper is organized as follows. We introduce a class of copula-based regression models and the full likelihood of the model for bivariate mixed outcomes with non-ignorable outcomes in Section 2. Simulation results on the sample properties of estimates are reported in Section 3. Section 4 illustrates the application of the model to the medical data. Finally, the paper concludes in Section 5.

2. Model and Likelihood

2.1. Bivariate Outcomes with non-ignorable missing values

Let X_i be an ordinal outcome with D level and Y_i be a continuous outcome. These outcomes are recorded for N individuals, correlated and modeled simultaneously. Some outcome values may be missing due to some reasons. Let X_i^*, $R_{X_i}^*$ and $R_{Y_i}^*$ denote the underlying latent variables for ordinal outcome X_i, the non-response mechanism for the ordinal variable and non-response

mechanism for the continuous variable, respectively. The ordinal variable of the i^{th} individual with D levels is defined as

$$X_i = \begin{cases} l_1, & \text{if } \mathrm{X}_i^* \in (\theta_0, \theta_1), \\ \vdots & \vdots \\ l_{k+1}, & \text{if } \mathrm{X}_i^* \in [\theta_k, \theta_{k+1}), \\ \vdots & \vdots \\ l_D, & \text{if } \mathrm{X}_i^* \in [\theta_{D-1}, \theta_D), \end{cases}$$

where $\theta_1 < \cdots < \theta_{D-1}$ are the cut-point parameters with $\theta_0 = -\infty$ and $\theta_D = \infty$. Also, for the response variables for responding to X and Y are defined, respectively, as

$$R_{X_i} = \begin{cases} 1, & R_{X_i}^* > 0, \\ 0, & o.w, \end{cases}$$

and

$$R_{Y_i} = \begin{cases} 1, & R_{Y_i}^* > 0, \\ 0, & o.w, \end{cases}$$

$R_{X_i}^*$ and $R_{Y_i}^*$ may be interpreted as propensity of individual i as a latent variable to respond to X_i and Y_i, respectively.

The joint model is assumed to take the form:

$$\begin{aligned} X_i^* &= \mu_{1i}(z_{1i}, \alpha) + \varepsilon_{1i}, \\ Y_i &= \mu_{2i}(z_{2i}, \beta) + \varepsilon_{2i}, \\ R_{X_i}^* &= \mu_{3i}(z_{3i}, \gamma) + \varepsilon_{3i} \\ R_{Y_i}^* &= \mu_{4i}(z_{4i}, \eta) + \varepsilon_{4i}, \end{aligned} \tag{1}$$

where $E(\varepsilon_{ki}) = 0.$ for $k = 1, 2, 3, 4,$ and the covariance matrix of the vector of errors $(\varepsilon_{1i}, \varepsilon_{2i}, \varepsilon_{3i}, \varepsilon_{4i})'$ is

$$\Sigma_{1234} = \begin{pmatrix} 1 & \sigma\rho_{12} & \rho_{13} & \rho_{14} \\ \sigma\rho_{12} & \sigma^2 & \sigma\rho_{23} & \sigma\rho_{24} \\ \rho_{13} & \sigma\rho_{23} & 1 & \sigma\rho_{34} \\ \rho_{14} & \sigma\rho_{24} & \sigma\rho_{34} & 1 \end{pmatrix},$$

where α, β, γ and η are vectors of regression coefficients, also β includes an intercept parameter but α does not include any intercept. Also z_{1i}, z_{2i}, z_{3i} and z_{4i} are outcome specific covariate vectors, and $\mu_{1i}, \mu_{2i}, \mu_{3i}$ and μ_{4i} are link functions specifying how the covariates are incorporated in the marginal means. For example, in the linear models

$$\mu_{1i} = z_{1i}^T\alpha,\ \mu_{2i} = z_{2i}^T\beta,\ \mu_{3i} = z_{3i}^T\gamma \text{ and } \mu_{4i} = z_{4i}^T\eta.$$

Also the correlation parameters $\rho_{jj'}$ for $j < j'$, $j = 1,2,3$, and $j' = 2,3,4$ should be estimated. If one of the correlation parameters $\rho_{jj'}$ for $j < j'$, $j = 1,2$, and $j' = 3,4$ is found to be significant, then we have a NMAR mechanism and missing mechanism cannot be ignored. On the other hand, if $\rho_{jj'}$ for $j < j'$, $j = 1,2$, and $j' = 3,4$ are found to be 0s, the missing data is MCAR and can be ignored. In this model any multivariate distribution can be assumed for the errors in the model. Here, a multivariate Gaussian copula is assumed. We have to restrict at least one parameter to obtain an identifiable model. For identifiability reasons, we assume that

$$Var(X_i^*) = Var(R_{X_i}^*) = Var(R_{Y_i}^*) = 1.$$

To obtain the likelihood function, we used the multivariate Gaussian copula. we can specify the joint CDF of

$$F_{X_i^*,Y_i},\ F_{X_i^*,Y_i,R_{Y_i}^*,R_{X_i}^*},\ F_{X_i^*,Y_i,R_{X_i}^*},\ F_{X_i^*,R_{Y_i}^*,R_{X_i}^*} \text{ and } F_{Y_i,R_{Y_i}^*,R_{X_i}^*}$$

as

$$\begin{aligned}
&F_{X_i^*,Y_i}(x_i^*, y_i) = \Phi_2(\Phi^{-1}\{F_{X_i^*}(x_i^*)\}, \Phi^{-1}\{F_{Y_i}(y_i)\} \big| \rho_{12}),\\
&F_{X_i^*,Y_i,R_{Y_i}^*,R_{X_i}^*}(x_i^*, y_i, r_{y_i}^*, r_{x_i}^*) = \Phi_4(\Phi^{-1}\{F_{X_i^*}(x_i^*)\}, \Phi^{-1}\{F_{Y_i}(y_i)\}, \Phi^{-1}\{F_{R_{X_i}^*}(r_{x_i}^*)\}, \Phi^{-1}\{F_{R_{Y_i}^*}(r_{y_i}^*)\} \big| \Sigma_{1234}),\\
&F_{X_i^*,Y_i,R_{Y_i}^*}(x_i^*, y_i, r_{x_i}^*) = \Phi_3(\Phi^{-1}\{F_{X_i^*}(x_i^*)\}, \Phi^{-1}\{F_{Y_i}(y)\}, \Phi^{-1}\{F_{R_{Y_i}^*}(r_{y_i}^*)\} \big| \Sigma_{123}),\\
&F_{X_i^*,R_{Y_i}^*,R_{X_i}^*} = \Phi_3(\Phi^{-1}\{F_{X_i^*}(x_i^*)\}, \Phi^{-1}\{F_{R_{y_i}^*}(r_{y_i}^*)\}, \Phi^{-1}\{F_{R_{x_i}^*}(r_{x_i}^*)\} \big| \Sigma_{134}),\\
&F_{Y_i,R_{Y_i}^*,R_{X_i}^*} = \Phi_3(\Phi^{-1}\{F_{Y_i}(y_i)\}, \Phi^{-1}\{F_{R_{y_i}^*}(r_{y_i}^*)\}, \Phi^{-1}\{F_{R_{x_i}^*}(r_{x_i}^*)\} \big| \Sigma_{134}),
\end{aligned}$$

where $\Phi(\cdot)$ is the standard normal distribution function, $\Phi_3(\cdot,\cdot,\cdot,\cdot)$ is the cumulative standard multivariate normal distribution with covariance matrix

$$\Sigma_{123} = \begin{pmatrix} 1 & \sigma\rho_{12} & \rho_{13} \\ \sigma\rho_{12} & \sigma^2 & \sigma\rho_{23} \\ \rho_{13} & \sigma\rho_{23} & 1 \end{pmatrix},\quad \Sigma_{134} = \begin{pmatrix} 1 & \rho_{13} & \rho_{14} \\ \rho_{13} & 1 & \rho_{34} \\ \rho_{14} & \rho_{34} & 1 \end{pmatrix},\quad \Sigma_{234} = \begin{pmatrix} \sigma^2 & \sigma\rho_{23} & \sigma\rho_{24} \\ \sigma\rho_{23} & 1 & \rho_{34} \\ \sigma\rho_{24} & \rho_{34} & 1 \end{pmatrix},$$

and $\Phi_4(\cdot,\cdot,\cdot,\cdot;\cdot)$ is the cumulative standard multivariate normal distribution with matrix covariance

$$\Sigma_{1234}=\begin{pmatrix} 1 & \sigma\rho_{12} & \rho_{13} & \rho_{14} \\ \sigma\rho_{12} & \sigma^2 & \sigma\rho_{23} & \sigma\rho_{24} \\ \rho_{13} & \sigma\rho_{23} & 1 & \sigma\rho_{34} \\ \rho_{14} & \sigma\rho_{24} & \sigma\rho_{34} & 1 \end{pmatrix},$$

where the marginal distributions $F_{X_i^*}, F_{Y_i^*}, F_{R_{X_i}^*}$ and $F_{R_{Y_i}^*}$ are absolutely continuous distributions.

To obtain joint the distribution of X_i and Y_i and missing mechanism, we consider the following four cases:

Case 1

For the ith individual with both X_i and Y_i observed the joint distribution of X_i, Y_i and missing mechanisms is

$$P(X_i = x_i, Y_i \le y_i, R_{Y_i}=1, R_{X_i}=1),$$

so, we have

$$\begin{cases} \left[F_{X_i^*,Y_i}(\theta_1,y_i)-F_{X_i^*,Y_i,R_{Y_i}^*,R_{X_i}^*}(\theta_1,y_i,0,0)\right], & x_i=l_1, \\ \vdots \qquad\qquad\qquad \vdots & \\ \left[F_{X_i^*,Y_i}(\theta_{k+1},y_i)-F_{X_i^*,Y_i,R_{Y_i}^*,R_{X_i}^*}(\theta_k+1,y_i,0,0)\right] & \\ \qquad -\left[F_{X_i^*,Y_i}(\theta_k,y_i)-F_{X_i^*,Y_i,R_{Y_i}^*,R_{X_i}^*}(\theta_k,y_i,0,0)\right], & x_i=l_{k+1}, \\ \vdots \qquad\qquad\qquad \vdots & \\ \left[F_{Y_i}(y_i)-F_{Y_i,R_{Y_i}^*,R_{X_i}^*}(y_i,0,0)\right] & \\ \qquad -\left[F_{X_i^*,Y_i}(\theta_D,y_i)-F_{X_i^*,Y_i,R_{Y_i}^*,R_{X_i}^*}(\theta_D,y_i,0,0)\right], & x_i=l_D, \end{cases}$$

where

$$F_{X_i^*,Y_i}(\theta_k,y_i)=\Phi_2(\Phi^{-1}\{F_{X_i^*}(\theta_k)\},\Phi^{-1}\{F_{Y_i}(y_i)\}\big|\rho_{12}),$$
$$F_{X_i^*,Y_i,R_{Y_i}^*}(\theta_k,y_i,0)=\Phi_3(\Phi^{-1}\{F_{X_i^*}(\theta_k)\},\Phi^{-1}\{F_{Y_i}(y_i)\},\Phi^{-1}\{F_{R_{Y_i}^*}(0)\}\big|\Sigma_{123}),$$
$$F_{X_i^*,Y_i,R_{Y_i}^*,R_{X_i}^*}(\theta_k,y_i,0,0)=\Phi_3(\Phi^{-1}\{F_{X_i^*}(\theta_k)\},\Phi^{-1}\{F_{Y_i}(y_i)\},\Phi^{-1}\{F_{R_{Y_i}^*}(0),F_{R_{X_i}^*}(0)\}\big|\Sigma_{1234}).$$

Case 2

For the i^{th} individual neither whose X_i nor Y_i is observed the joint distribution of R_{X_i} and R_{Y_i} is

$$P(R_{Y_i}=0,R_{X_i}=0)=P(R_{Y_i}^{*}\le 0,R_{X_i}^{*}\le 0)=F_{R_{Y_i}^{*},R_{X_i}^{*}}(0,0).$$

In other words, the joint distribution of R_{X_i} and R_{Y_i} with possibility of missing for both outcomes is specified using bivariate Gaussian copula, as follow:

$$P(R_{Y_i}=0,R_{X_i}=0)=\Phi_2(\Phi^{-1}\{F_{R_{Y_i}^{*}}(0)\},\Phi^{-1}\{F_{R_{X_i}^{*}}(0)\}\,|\rho_{34}).$$

Case 3

For the i^{th} individual whose X_i is only observed the joint distribution of X_i, R_{X_i} and R_{Y_i} is

$$P(X_i=x,R_{Y_i}=0,R_{X_i}=1).$$

So, we have:

$$\begin{cases} F_{X_i^{*},R_{Y_i}^{*}}(\theta_1,0)-F_{X_i^{*},R_{Y_i}^{*},R_{X_i}^{*}}(\theta_1,0,0), x_i=\mathrm{l}_1, & x_i=l_1,\\ \vdots \qquad\qquad \vdots & \\ \left[F_{X_i^{*},R_{Y_i}^{*}}(\theta_{k+1},0)-F_{X_i^{*},R_{Y_i}^{*}}(\theta_k,0)\right] & \\ \qquad -\left[F_{X_i^{*},R_{Y_i}^{*},R_{X_i}^{*}}(\theta_{k+1},0,0)-F_{X_i^{*},R_{Y_i}^{*},R_{X_i}^{*}}(\theta_k,0,0)\right], & x_i=l_{k+1},\\ \vdots \qquad\qquad \vdots & \\ \left[F_{R_{Y_i}^{*}}(0)-F_{X_i^{*},R_{Y_i}^{*}}(\theta_D,0)\right] & \\ \qquad -\left[F_{R_{Y_i}^{*},R_{X_i}^{*}}(0,0)-F_{X_i^{*},R_{Y_i}^{*},R_{X_i}^{*}}(\theta_D,0,0)\right], & x_i=l_D, \end{cases}$$

where

$$F_{X_i^{*},R_{yi}^{*}}(\theta_1,0)=\Phi_2(\Phi^{-1}\{F_{X_i^{*}}(\theta_1)\},\Phi^{-1}\{F_{R_{yi}^{*}}(0)\}\,|\rho_{24}),$$

and

$$F_{X_i^{*},R_{Y_i}^{*},R_{X_i}^{*}}(\theta_1,0,0)=\Phi_3(\Phi^{-1}\{F_{X_i^{*}}(\theta_1)\},\Phi^{-1}\{F_{R_{yi}^{*}}(0)\},\Phi^{-1}\{F_{R_{x_i}^{*}}(0)\}|\Sigma_{134}).$$

Case 4

For the i^{th} individual whose Y_i is observed the joint distribution of R_{X_i} and R_{Y_i} is

$$P(Y_i \le y_i, R_{Y_i} = 1, R_{X_i} = 0) = P(Y_i \le y_i, R^*_{X_i} \le 0, R^*_{Y_i} > 0),$$
$$= P(Y_i \le y_i, R^*_{X_i} \le 0) - P(Y_i \le y_i, R^*_{X_i} \le 0, R^*_{Y_i} \le 0),$$
$$= F_{Y_i, R^*_{X_i}}(y_i, 0) - F_{Y_i, R^*_{X_i}, R^*_{Y_i}}(y_i, 0, 0),$$

in other words, the joint distribution of Y_i and missing mechanisms is specified using Gaussian copula, as follow

$$P(Y_i \le y_i, R_{Y_i} = 1, R_{X_i} = 0) = \Phi_2(\Phi^{-1}\{F_{Y_i}(y_i)\}, \Phi^{-1}\{F_{R^*_{X_i}}(0)\} \big| \rho_{23}),$$
$$- \Phi_3(\Phi^{-1}\{F_{Y_i}(y_i)\}, \Phi^{-1}\{F_{R^*_{X_i}}(0)\}, \Phi^{-1}\{F_{R^*_{Y_i}}(0)\} \big| \Sigma_{243}).$$

Also, we have:

$$f_{X_i, Y_i}(x_i, y_i) = \frac{\partial}{\partial y_i} P(X_i = x_i, Y_i \le y_i).$$

Likelihood function is the product of the joint distribution of X_i and Y_i and missing mechanism, for four cases and shows the simplification obtained by using the assumption of multivariate Gaussian copula for errors in the model.

2.2. Multivariate Outcomes with non-ignorable missing values

Suppose the vector of response for the i^{th} individual is

$$W_i = (X_{i1}, ..., X_{ip}, Y_{i(p+1)}, ..., Y_{iq})',$$

where X_{is} for $s = 1, ..., p,$ are ordinal responses each with D_s levels and Y_{is} for $s = p+1, ..., q$, are continuous responses. All responses are correlated.

Let X^*_{is} for $s = 1, ..., p,$ denote the underlying random variable of the ordinal response for the i^{th} individual and $s-th$ outcomes with D_s levels. Define

$$X_{is} = \begin{cases} l_{1s}, & \theta_{0s} < X^*_{is} < \theta_{1s}, \\ l_{j+1}, & \theta_{sj} \le X^*_{is} \le \theta_{s(j+1)}, \quad j = 1, ..., D_s - 2, \\ l_{D_s}, & \theta_{s(D_s-1)} < X^*_{is} < \theta_{sD_s}, \end{cases}$$

where $\theta_{0s} = -\infty$, $\theta_{sD_s} = \infty$ and $\theta_s = (\theta_{s_1}, ..., \theta_{s(D_s-1)})'$ is the vector of cut-points parameters for $s = 1, ..., p$. Typically, when missing data occur in an outcome, we assume $R_{X_i} = (R_{X_{i1}}, ..., R_{X_{ip}})'$ as the indicator vector of responding to X_i and $R_{X_{is}}$. It is defined as

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

and $R_{Y_i} = (R_{Y_{i(p+1)}}, ..., R_{Y_{iq}})'$ the indicator vector for responding to Y_i and $R_{Y_{is}}$ is defined as

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

where $R^*_{X_{is}}$ and $R^*_{Y_{is}}$ denote the underlying latent variables of the non-response mechanism, respectively, for the ordinal and continuous variables.

The joint model takes the form:

$$\begin{aligned} X^*_{is} &= \mu_{1i}(z_{i1}, \alpha_s) + \varepsilon^{(1)}_{is}, \quad s = 1, ..., p, \\ Y_{is} &= \mu_{2i}(z_{i2}, \beta_s) + \varepsilon^{(2)}_{is}, \quad s = p+1, ..., q, \\ R^*_{X_{is}} &= \mu_{3i}(z_{i3}, \gamma_s) + \varepsilon^{(3)}_{is}, \quad s = 1, ..., p, \\ R^*_{Y_{is}} &= \mu_{4i}(z_{i4}, \eta_s) + \varepsilon^{(4)}_{is}, \quad s = p+1, ..., q, \end{aligned} \tag{2}$$

where

$$\varepsilon_i = (\varepsilon_i^{(1)'}, \varepsilon_i^{(2)'}, \varepsilon_i^{(3)'}, \varepsilon_i^{(4)'})' \overset{iid}{\sim} MVN(0, \Sigma\varepsilon),$$
$$\varepsilon_i^{(u)} = (\varepsilon_{i1}^{(u)}, ..., \varepsilon_{ip}^{(u)})', \text{ for } u = 1, 3,$$
$$\varepsilon_i^{(u)} = (\varepsilon_{i(p+1)}^{(u)}, ..., \varepsilon_{iq}^{(u)})', \text{ for } u = 2, 4,$$

and

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

where

$$\Sigma^\varepsilon_{uu} = Var(\varepsilon_i^{(u)}), \text{ for } u = 1, 2, 3, 4, \ \Sigma^\varepsilon_{uv} = Cov(\varepsilon_i^{(u)}, \varepsilon_i^{(v)}), \ u < v, \ u, v = 1, 2, 3, 4, \text{ and } \Sigma^\varepsilon_{uv} = \Sigma^{\varepsilon\,\prime}_{vu}.$$

Because of identifiability problem we have to assume

$$Var(Y^*_{is}) = Var(R^*_{y_{is}}) = Var(R^*_{y_{is}}) = 1,$$

so

$$\Sigma^\varepsilon_{ss} = j, \text{ for } j = 1, 3, 4.$$

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. The vector β_s for $s = p+1, ..., q$, includes an intercept parameter but α_s and γ_s, for $s = 1, ..., p$ and η_s, for $s = p+1, ..., q$, due to having cut-point parameters are assumed not to include any intercept.

Let

$$J_{obs} = (J_{obs}^X, J_{obs}^Y),\ J_{Mis} = J_{obs}^C,\ J_{obs}^X = \{s : X_{is}\ is\ observed\},\ J_{Mis}^X = (J_{obs}^X)^C,$$
$$\text{and } J_{obs}^{\ Y} = \{s : Y_{is}\ is\ observed\},\ J_{Mis}^Y = (J_{obs}^Y)^C.$$

Also, let

$$Y_{i,obs} = \{Y_{is}, \forall s \in J_{obs}^Y\}$$

and

$$X_{i,obs} = \{X_{is}, \forall s \in J_{obs}^X\}$$

denote the set of continuous and ordinal variables observed.

Further, let $X_{i,obs}^*$ denote the set of underlying random variable of the ordinal response of the i^{th} individual defined as:

$$X_{i,obs}^* = \{\theta_{X_{is}-1} \le X_{is}^* \le \theta_{X_{is}}, \forall s \in J_{obs}^X\}.$$

Also, the set of non-response mechanism for the continuous and ordinal random variables which is defined, respectively, as:

$$R_{Y_{i,obs}} = \{R_{Y_{is}} = 1, \forall s \in J_{obs}^Y\} = R^*_{Y_{i,obs}} = \{R_{Y_{is}}^* > 0, \forall s \in J_{obs}^Y\},$$

$$R_{X_{i,obs}} = \{R_{X_{is}} = 1, \forall s \in J_{obs}^X\} = R^*_{X_{i,obs}} = \{R_{X_{ij}}^* > 0, \forall s \in J_{obs}^X\}.$$

To obtain the likelihood function, we used the multivariate Gaussian copula. The likelihood of the model (2) is

$$\begin{aligned} L &= \prod_{\{i \in J_{obs}\}} f(X_{i,obs}, Y_{i,obs}, R_{X_{i,obs}}, R_{Y_{i,obs}} \mid z_{i1}, ..., z_{i4}), \\ &= \prod_{\{i \in J_{obs}\}} P(X_{i,obs}^*, R_{X_{i,obs}}^*, R_{Y_{i,obs}}^*, Y_{i,obs}, C_{i,Mis}^* \mid z_{i1}, ..., z_{i4}), \\ &= \prod_{\{i \in J_{obs}\}} (\Gamma_i^* - \Gamma_{X_i}^* - \Gamma_{Y_i}^* + \Gamma_{X_iY_i}^*), \end{aligned}$$

where

$$C_{i,Mis}^* = \{R_{X_{is}} = 0; \forall s \in J_{Mis}^X, R_{Y_{ij}} = 0; \forall s \in J_{Mis}^Y\}$$

and

$$\begin{aligned}\Gamma_i^* &= P(X_{i,obs}^*, C_{Mis}^*, Y_{i,obs} \mid z_{i1},...,z_{i4}) \\ &= \Phi_q(\Phi^{-1}\{F_{X_{is}^*}(\theta_{(j+1)s}); \forall s \in J_{obs}^X\}, \Phi^{-1}\{F_{Y_{is}}(y_{is}); \forall s \in J_{obs}^Y\}, \\ &\qquad \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \forall s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \forall s \in J_{Mis}^Y\}) \\ &\quad - \Phi_q(\Phi^{-1}\{F_{X_{is}^*}(\theta_{js}); \forall s \in J_{obs}^X\}, \Phi^{-1}\{F_{Y_{is}}(y_{is}); \forall s \in J_{obs}^Y\}, \\ &\qquad \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \forall s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \forall s \in J_{Mis}^Y\}),\end{aligned}$$

$$\begin{aligned}\Gamma_{X_i}^* &= P(X_{i,obs}^*, C_{Mis}^*, R_{Y_{i,Mis}}^*, Y_{i,obs} \mid z_{i1},...,z_{i4}) \\ &= \Phi_q(\Phi^{-1}\{F_{X_{is}^*}(\theta_{(j+1)s}); \forall s \in J_{obs}^X\}, \Phi^{-1}\{F_{Y_{is}}(y_{is}); \forall s \in J_{obs}^Y\}, \\ &\qquad \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \exists s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \forall s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \forall s \in J_{Mis}^Y\}) \\ &\quad - \Phi_q(\Phi^{-1}\{F_{X_{ij}^*}(\theta_{js}); \forall s \in J_{obs}^X\}, \Phi^{-1}\{F_{Y_{is}}(y_{is}); \forall s \in J_{obs}^Y\}, \\ &\qquad \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \exists s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \forall s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \forall s \in J_{Mis}^Y\}),\end{aligned}$$

$$\begin{aligned}\Gamma_{Y_i}^* &= P(X_{i,obs}^*, C_{Mis}^*, R_{X_{i,Mis}}^* Y_{i,obs} \mid z_{i1},...,z_{i4}) \\ &= \Phi_q(\Phi^{-1}\{F_{X_{ij}^*}(\theta_{(j+1)s}); \forall s \in J_{obs}^X\}, \Phi^{-1}\{F_{Y_{is}}(y_{is}); \forall s \in J_{obs}^Y\}, \\ &\qquad \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \exists s \in J_{Mis}^Y\}, \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \forall s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \forall s \in J_{Mis}^Y\}) \\ &\quad - \Phi_q(\Phi^{-1}\{F_{X_{is}^*}(\theta_{js}); \forall s \in J_{obs}^X\}, \Phi^{-1}\{F_{Y_{is}}(y_{is}); \forall s \in J_{obs}^Y\}, \\ &\qquad \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \exists s \in J_{Mis}^y\}, \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \forall s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \forall s \in J_{Mis}^Y\}),\end{aligned}$$

and

$$\begin{aligned}\Gamma_{X_iY_i}^* &= P(X_{i,obs}^*, C_{Mis}^*, R_{X_{i,Mis}}{}^*, R_{Y_{i,Mis}}^*, Y_{i,obs} \mid z_{i1},...,z_{i4}) \\ &= \Phi_q(\Phi^{-1}\{F_{X_{ij}^*}(\theta_{(j+1)s}); \forall s \in J_{obs}^X\}, \Phi^{-1}\{F_{Y_{is}}(y_{is}); \forall s \in J_{obs}^Y\}, \\ &\qquad \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \exists s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \exists s \in J_{Mis}^Y\}, \\ &\qquad \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \forall s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \forall s \in J_{Mis}^Y\}) \\ &\quad - \Phi_q(\Phi^{-1}\{F_{X_{ij}^*}(\theta_{sj}); \forall s \in J_{obs}^X\}, \Phi^{-1}\{F_{Y_{is}}(y_{is}); \forall s \in J_{obs}^Y\}, \\ &\qquad \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \exists s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \exists s \in J_{Mis}^Y\}, \\ &\qquad \Phi^{-1}\{F_{R_{X_{is}}^*}(0); \forall s \in J_{Mis}^X\}, \Phi^{-1}\{F_{R_{Y_{is}}^*}(0); \forall s \in J_{Mis}^Y\}).\end{aligned}$$

2.3. Multivariate Outcomes with ignorable missing values

We consider model (2) for finding the condition for MAR. Let

$$W=(X,Y)=(W_{obs},W_{mis}),\ W^*=(X^*,Y)\ \ (W^*_{obs},W^*_{mis}) \quad \text{and}\ R^*=(R^*_X,R^*_Y),$$

where

$$X=(X_1,...,X_p)',\ Y=(Y_{p+1},...,Y_q)',\ X^*=(X^*_1,...,X^*_p)',$$

W^*_{obs} is the vector of latent variables related to the observed part of $W=(X,Y)$, and W^*_{mis} is the vector of latent variables related to the missing part of $W=(X,Y)$. According to our joint model, the vector of responses along with the missing indicators $(W^*,R^*)=(W^*_{obs},W^*_{mis},R^*)$ has a multivariate normal distribution with the following covariance structure,

$$\Sigma=\begin{pmatrix}\Sigma_{o,o} & \Sigma_{o,m} & \Sigma_{o,R^*}\\ \Sigma_{m,o} & \Sigma_{m,m} & \Sigma_{m,R^*}\\ \Sigma_{R^*,o} & \Sigma_{R^*,m} & \Sigma_{R^*,R^*}\end{pmatrix},$$

where

$$\begin{aligned}\Sigma_{o,o}&=\operatorname{cov}(W^*_{obs},W^*_{obs}),\\ \Sigma_{m,m}&=\operatorname{cov}(W^*_{mis},W^*_{mis}),\\ \Sigma_{o,m}&=\operatorname{cov}(W^*_{obs},W^*_{mis}),\\ \Sigma_{o,R^*}&=\operatorname{cov}(W^*_{obs},R^*),\\ \Sigma_{R^*,R^*}&=\operatorname{cov}(R^*,R^*).\end{aligned}$$

The joint density function of W^* and R^* can also be partitioned as

$$f(W^*,R^*)=f(W^*_{mis},R^*\,|W^*_{obs})\,f(W_{obs}{}^*),$$

where $f(W^*_{mis},R^*\,|W^*_{obs})$ and $f(W^*_{obs})$ have, respectively, a conditional and a marginal normal distribution. According to the missing mechanism definitions, to have a MAR mechanism the covariance matrix of the above mentioned conditional normal distribution,

$$\begin{aligned}\Sigma_{m,R^*|o}&=\operatorname{cov}(W_{mis}{}^*,R^*\,|W^*_{obs})\\ &=\begin{pmatrix}\Sigma_{m,m} & \Sigma_{m,R^*}\\ \Sigma_{R^*,m} & \Sigma_{R^*,R^*}\end{pmatrix}-\begin{pmatrix}\Sigma_{m,o} & \Sigma_{R^*,o}\end{pmatrix}'\Sigma^{-1}_{o,o}\begin{pmatrix}\Sigma_{m,o} & \Sigma_{R^*,o}\end{pmatrix}\\ &=\begin{pmatrix}\Sigma_{m,m}-\Sigma'_{m,o}\Sigma^{-1}_{o,o}\Sigma_{m,o} & \Sigma_{m,R^*}-\Sigma'_{m,o}\Sigma^{-1}_{o,o}\Sigma_{R^*,o}\\ \Sigma_{R^*,m}-\Sigma_{R^*,o'}\Sigma^{-1}_{o,o}\Sigma_{m,o} & \Sigma_{R^*,R^*}-\Sigma'_{R^*,o}\Sigma^{-1}_{o,o}\Sigma_{R^*,o}\end{pmatrix}\end{aligned}$$

should satisfy the following constraint,

$$\Sigma_{m,R^*} - \Sigma'_{m,o}\Sigma^{-1}_{o,o}\Sigma_{R^*,o} = \mathbf{0}. \quad (3)$$

So, for obtaining the likelihood function, we used the multivariate Gaussian copula with constraint (3).

2.4. Estimation

2.4.1. Joint Estimation

Putting $l(\theta)$ as the log-likelihood function, then let $\breve{S}(\theta) = \partial l(\theta)/\partial\theta$ be the score function and $\hbar(\theta) = \partial^2 l(\theta)/\partial\theta\partial\theta^\top$ the Hessian matrix, for obtaining the maximum likelihood estimate (MLE) $\hat{\theta}$, we must solve the $\breve{S}(\theta) = 0$, (joint estimation). We know that the Fisher information matrix is $I(\theta) = E\{-\hbar(\theta)\} = E\{\breve{S}(\theta)\breve{S}^\top(\theta)\}$. It can be shown that $\hat{\theta}$ is consistent and it has asymptotically multivariate normal distribution with mean θ and covariance matrix given by the inverse of the $I(\theta)$. So the standard errors (SE) for $\hat{\theta}$ are calculated from diagonals of $E\left\{\breve{S}(\hat{\theta})\breve{S}^\top(\hat{\theta})\right\}^{-1}$. We used the function pnorm for likelihood evaluation and the function nlminb, which do not require the score function for optimization in R. One may choose different starting values over multiple runs of the iteration procedure and then examine the results to see whether the same solution is obtained repeatedly. When that happens, one can conclude with some confidence that a global maximum has been found. For good initial values for our application we suggest the use of the results of separately analyzing continuous and ordinal variables.

2.4.2. Marginal estimation

Often the maximization of $l(\theta)$ computationally is not easy in practice, so we use the method of inference functions for margins (IFM). This method first estimates marginal parameters via margins, then only uses the copula as a basis for estimating the association parameters. In other words the in IFM method the marginal models and the dependence between outcomes are specified separately, (Marginal estimation). The IFM estimate, $\breve{\theta}$, has asymptotically multivariate normal distribution with mean θ and covariance matrix $C = J^{-1}BJ^{-1}$ where J is a block- diagonal matrix with symmetric diagonal blocks and B is a symmetric block matrix. Standard errors (SE) of $\breve{\theta}$ are obtained from the diagonals of $\breve{C} = \breve{J}^{-1}\breve{B}\breve{J}^{-1}$, where $\breve{J}$ and $\breve{B}$ are the respective estimates of J and B obtained from $\breve{\theta}$ [via Joe and Xu (1996) and Harry and James (1998)].

3. Simulation Study

In this section, the first considers a joint model for mixed ordinal and continuous outcomes under the five scenarios, assembled from a marginal normal specification for a marginal normal specification for a marginal normal distribution for the latent variable underlying the ordinal outcome and the continuous outcome, the second a marginal model is based under NMAR and MAR mechanisms for the five scenarios. In both cases, we adopt the Gaussian copula to

construct the joint model. The results indicate that the joint estimation should be preferred to the marginal approach under NMAR and MAR mechanisms, however, the two methods perform generally similarly for mixed ordinal and continuous responses with non-ignorable missing values. The relative biases of the joint and marginal estimates are obtained for the five scenarios with and without non-ignorable outcomes.The formula for the relative bias of θ is as follows:

$$\text{Relative bias}=\frac{\theta-\hat{\theta}}{\theta}\times 100\%.$$

3.1. Ordinal-Normal Model with non-ignorable missing values

Let X_i be an ordinal outcome and Y_i be a continuous outcome. These are obtained for each of N subjects. Some of these values may be missed. Continuous variables, X_i^* and $R_{Y_i}^*$, respectively, represent latent variables for ordinal outcome and latent variable related to missing mechanism of Y_i. We define ordinal variable X_i for the i^{th} subject as follows:

$$X_i=\begin{cases} l_1=1, & \text{if } X_i^*<\theta_1,\\ l_2=2, & \text{if } \theta_1\le X_i^*\le\theta_2,\\ l_3=3, & \text{if } X_i^*>\theta_2. \end{cases}$$

The variables X_i^*, Y_i and $R_{Y_i}^*$ are generated by a multivariate normal distribution with zero mean and covariance matrix

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

We define $R_{Y_i}^*$ as

$$R_{Y_i}=\begin{cases} 1, & R_{Y_i}^*>0,\\ 0. & o.w. \end{cases}$$

We assume the percentage of missing values of Y_i to be 30 %. A total of M=1000 repeated samples $(X_i^*, Y_i, R_{X_i}^*)$ of sizes N=100 and N=200 were generated under five scenarios, where $\alpha_1=1$, $\beta_1=1$, $\beta_2=1$, $\theta_1=-1$, $\theta_2=1$, $\gamma_1=1$ and σ=1, with (A) $\rho_{12}=0.1,\rho_{13}=0.1,\rho_{23}=0.1$ (B) $\rho_{12}=0.25,\rho_{13}=0.25,\rho_{23}=0.25$ (C) $\rho_{12}=0.5,\rho_{13}=0.5,\rho_{23}=0.5$ (D) $\rho_{12}=0.75,\rho_{13}=0.75,\rho_{23}=0.75$ (E) $\rho_{12}=0.9,\rho_{13}=0.9,\rho_{23}=0.9$.

We analyze the following simple model

$$\begin{aligned} E(X_i^*) &= \beta_1 + \beta_2 z_i, \\ E(Y_i) &= \alpha_1 z_i, \\ E(R_{Y_i}^*) &= \gamma_1 z_i, \end{aligned} \tag{4}$$

where the distributions of X_i^*, Y_i and $R_{Y_i}^*$ are, respectively, $N(\beta_1+\beta_2 z_i,1)$, $N(\alpha_1 z_i,\sigma^2)$ and $N(\gamma_1 z_i,1)$. We generate data by the same process as above and in estimating the parameters we assume MAR and NMAR mechanisms.

3.2. Ordinal -Normal Model with ignorable missing values

For our simulation, we have missing values only for our continuous variable and we may have $W_{mis}^* = X^*$ and $W_{obs}^* = Y$. For the missing mechanism we only need to define $R^* = R_X^*$, as we do not have any missing value for our ordinal response and we consider model (4), for finding the condition for MAR, let

$$\Sigma_{m,m} = 1, \Sigma_{m,R^*} = \rho_{23}, \Sigma_{R^*,R^*} = 1, \Sigma_{o,o} = \sigma^2, \Sigma_{m,o} = \sigma\rho_{12}, \Sigma_{R^*,o} = \sigma\rho_{13}.$$

So, the constraint (3) will be reduced to

$$\Sigma_{m,R^*} - \Sigma'_{m,o}\Sigma_{o,o}^{-1}\Sigma_{R^*,o} = \rho_{23} - \rho_{13}\rho_{12} = 0.$$

3.3. Results

Table 1 presents results on the relative biases of joint and marginal estimates obtained under MAR and NMAR mechanisms for the five scenarios. The relative biases for joint and marginal estimates of α_1, α_2, β_1, β_2, θ_1, θ_2, γ_1, ρ_{12}, ρ_{13} and σ under MAR mechanism are generally larger than those for joint and marginal estimates under NMAR mechanism i.e., if data are not missing at random such an assumption on estimating parameters leads to have biased estimates of parameters. So, if the missing mechanism is NMAR, use of model (2) which is assumed to be MAR may lead to biased estimates. A comparision of the relative biases of joint and marginal estimates, relative bias of joint estimates suggest that the were generally smaller than those for marginal estimates.

Figures (1)-(3) show relative Biases of joint and marginal estimates of α_1, β_1, β_2, γ_1, σ, θ_1 and θ_2 under NMAR mechanism versus the values of ρ_{12}, ρ_{23} and ρ_{13}. Solid and dashed plots correspond to relative biases of joint estimates for $N = 100$ and $N = 200$, respectively. Dotted and dashed-dotted plots correspond to those of marginal estimates for $N = 100$ and $N = 200$, respectively. The parameter-specific biases clearly indicates that both full and marginal likelihood approaches yield reasonably unbiased estimates with NMAR mechanism. Comparing parameter-specific estimates show that, relative biases for marginal estimates were generally larger than those for joint estimates. So, according to Figures (1)-(3)

relative biases for marginal estimates of α_1, β_1, β_2, γ_1, θ_1, θ_2 and σ versus the values of ρ_{12}, ρ_{23} and ρ_{13} are generally larger than those for joint estimates.

4. Application

4.1. Osteoporosis of the Spine Data

The osteoporosis of the spine 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 as an ordinal outcome with three levels for 5281 patients.

Albrand et al., (2003) show some epidemiological studies have identified clinical factors that predict the risk of hip fractures in elderly women independently of the level of bone mineral density (BMD), such as low body weight, history of fractures, and clinical risk factors for falls. Also, abdominal obesity needs to be included as a risk factor for osteoporosis and bone loss. Their results showed that having a lot of belly fat is more detrimental to bone health than having more superficial fat or fat around the hips. Excess fat around the belly may increase the risk of women developing the brittle bone disease osteoporosis. So, a bulging waistline puts women at risk of osteoporosis.

We shall also try to find answers for some questions, including: (1) How does the type of dwelling affect the level of osteoporosis, waistline and BMI of the patient? (2) How does the job status effect the level of osteoporosis, waistline and BMI of the patient? (3) How do the amount of total body calcium and age affect the level of osteoporosis, waistline and BMI of the patient?

Also we consider the body mass index (BMI) and waistline as continuous outcomes. Covariates which may affect the osteoporosis of the spine and waistline are amount of total body calcium (Ca), job status (Job), type of the dwelling (Ta) and age.

Table 2: The variable of interest and descriptive statistics for them

Discreet Variables	Type	Levels	Confidence interval
Osteoporosis of the spine	Ordinal		
		None	(26.2,28.2)%
		Mild	(28.1,31.3)%
		Severe	(30.8,35.3)%
		Missing	(7.3,10.11)%
Job status	Binary		
		employee	(39.2,43.4)%
		housekeeper	(56.5,60.9)%
Type of the dwelling	Binary		
		house	(29.11,35.5)%
		apartment	(64.2,68.8)%
Continuous Variables			
Age	Continuous		(45.23,48.34) year
Amount of total body calcium	Continuous		(980.45, 1001.71) mlgr
waistline	Continuous		(76.54 , 83.65) cm
BMI	Continuous		(28.53 , 28.93) kgr/cm^2

Table 2 shows the list, type and descriptive statistics of variables under study. This Table shows that the percentage of severe and mild osteoporosis are more than that of none level. Also 67.4% of women live in apartment and 58.7% of women are housekeeper. A frequency table for the osteoporosis of the spine shows that 39% of values are missing.

The Pearson correlation between osteoporosis of the spine and BMI responses, osteoporosis of the spine and waistline responses and BMI and waistline are $r_{OS,Waistline} = 0.245$, $r_{OS,BMI} = 0.208$ and $r_{BMI,Waistline} = 0.323$. Based on the results our simulation study, we can expect to find a higher value of correlation by our model. These three variables, osteoporosis of the spine, BMI and waistline are endogenous correlated variables, and they have to be modeled simultaneously. Taking into account the correlation, leads us to obtain a more precise estimation of standard errors of estimates and so a better inference.

These three outcomes, osteoporosis of the spine, waistline and the indicator variable for missing mechanism of Osteoporosis of the spine are endogenous correlated variables, and they have to be modeled simultaneously. The joint model for these data is

$$\begin{aligned} OS_i^* &= \alpha_1 age_i + \alpha_2 Ca_i + \alpha_3 Ta_i + \alpha_4 Job_i + \varepsilon_{1i}, \\ Waist_i &= \beta_0 + \beta_1 age_i + \beta_2 Ca_i + \beta_3 Ta_i + \beta_4 Job_i + \varepsilon_{2i}, \\ R_{OS_i}^* &= \gamma_1 age_i + \gamma_2 Ca_i + \gamma_3 Ta_i + \gamma_4 Job_i + \varepsilon_{3i}, \\ BMI_i &= \xi_0 + \xi_1 age_i + \xi_2 Ca_i + \xi_3 Ta_i + \xi_4 Job_i + \varepsilon_{4i}. \end{aligned} \tag{5}$$

The covariance matrix of the vector of errors $(\varepsilon_{1i}, \varepsilon_{2i}, \varepsilon_{3i}, \varepsilon_{4i})'$ for model (3) is

$$\Sigma = \begin{pmatrix} 1 & \sigma_1\rho_{12} & \rho_{13} & \sigma_2\rho_{14} \\ \sigma_1\rho_{12} & \sigma_1^2 & \sigma_1\rho_{23} & \sigma_1\sigma_2\rho_{24} \\ \rho_{13} & \sigma_1\rho_{23} & 1 & \sigma_2\rho_{34} \\ \sigma_2\rho_{14} & \sigma_1\sigma_2\rho_{24} & \sigma_2\rho_{34} & \sigma_2^2 \end{pmatrix}.$$

Here,

$$(\alpha_1,...,\alpha_4)', (\beta_0,...,\beta_4)', (\gamma_1,...,\gamma_4)', (\xi_0,...,\xi_4)',$$
$$\sigma_1^2, \sigma_2^2, \theta_1, \theta_2, \rho_{12}, \rho_{13}, \rho_{14}, \rho_{23}, \rho_{24} \text{ and } \rho_{34}$$

are parameters that should be estimated. A joint model to use the marginal and joint estimation for model (3) is specified by a multivariate Gaussian copula.

For our application, we have missing values only for our ordinal variable and we may have $W_{mis}^* = OS^*$ and $W_{obs}^* = (BMI, Waist)'$. For missing mechanism we only need to define $R^* = R_{OS}^*$, as we do not have any missing value for our continuous responses, and

$$\Sigma_{m,m}=1,\Sigma_{m,R^*}=\rho_{13},\Sigma_{R^*,R^*}=1,\Sigma_{o,o}=\begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\rho_{24}\\ \sigma_1\sigma_2\rho_{24} & \sigma_2^2\end{pmatrix},$$
$$\Sigma_{m,o}=(\sigma_1\rho_{12},\sigma_2\rho_{14})',\Sigma_{R^*,o}=(\sigma_1\rho_{23},\sigma_2\rho_{34})',$$

so that the needed constraint will be reduced to

$$\Sigma_{m,R^*}-\Sigma'_{m,o}\Sigma^{-1}_{o,o}\Sigma_{R^*,o}=\rho_{13}-\frac{\rho_{23}\rho_{12}+\rho_{23}\rho_{24}\rho_{12}+\rho_{23}\rho_{24}\rho_{12}-\rho_{34}\rho_{12}}{1-\rho_{24}^2}=0.$$

4.2. Results for Data

Results of analysis the marginal and joint estimation for model (5) with missing mechanism are given in Table 3. For comparative purposes, four models are considered. The first model (model I) and the second model (model II) consider the joint estimation with NMAR and MAR mechanism for model (5) . Also, The third model (model III) and fourth model (model IV) uses the marginal estimation with NMAR and MAR mechanism for model (5).

Model (I) shows a significant effect of Ca, Ta and age on the value of osteoporosis of the spine and significant effect of Ta on waistline and shows a weak significant effect of age on BMI. From these effects we can infer that the older the patient the lower the value of osteoporosis of the spine; people who live in apartment have higher low value of osteoporosis of the spine than that of people who live in a house and the more the amount of calcium of the body of the patient the higher is the low value of osteoporosis of the spine. None of the explanatory variables has any effect on the missing indicator for osteoporosis of the spine.

For model (I) correlation parameters ρ_{12}, ρ_{13}, ρ_{14} and ρ_{24} are strongly significant. They show a positive correlation between waistline and osteoporosis of the spine ($\hat{\rho}_{12}$) and it shows a positive correlation between waistline and BMI ($\hat{\rho}_{24}$) and a positive correlation between osteoporosis and the missing indicator for the spine ($\hat{\rho}_{13}$). This leads to have a NMAR mechanism.

Model (II), model (III) and model (IV) give the same results as model (I). To compare model (I) and model (II) we have deviance =126.011, (p-value < 0.001). So one may prefer model (I). For model (II), the estimated variance of waist and BMI ($\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$) obtained by model (I) are less than those of model (II).

To compare model (I) and model (III) we have deviance =101.08 (P-value < 0.001). Also, for model (I) and model (IV) we have deviance =165.04 (p-value < 0.001). So one may prefer model (I). Also comparing parameter-specific estimates for model (I), model (III), model (II) and model (IV) show that, $-loglikelihoods$ for marginal modelling are generally larger than those of joint modelling.

5. Conclusion

We have extended copula-based regression models for mixed outcomes with non-ignorable missing values. For obtaining joint distribution of discrete and continuous outcomes with possibility of missing values, we consider four cases then using bivariate and multivariate Gaussian copulas we mixed-outcome marginal regression models. Two likelihood estimation strategies are proposed, one method uses full likelihood function to estimate parameters simultaneously, the other applies the IFM method to estimate parameters marginally and shared parameters jointly. A generalization of our model for longitudinal studies is still an ongoing research on our part.

REFERENCES

Albrand. G, Munoz. F, Sornay-Rendu. E, DuBoeuf. F, Delmas. P. D. (2003). Independent predictors of all osteoporosis-related fractures in healthy postmenopausal women: *The OFELY Study. Bone*, **32(1)**:78-85.

Bahrami Samani, E., Ganjali, M. and Khodaddadi, A. (2008). A Latent Variable Model for Mixed Continuous and Ordinal Responses, *Journal of Statistical Theory and Applications*, **7(3)**:337-349.

Bahrami Samani, E., Ganjali, M. and Eftekhari Mahabadi, S. (2010). A Latent Variable Model for Mixed Continuous and Ordinal Responses with Nonignorable Missing Responses, *Sankhya* **72-B**, (1): 38-57.

Bahrami Samani, E., Ganjali, M. and Zahed, H. (2011). A Multivariate Latent Random Effect Model for Mixed Continuous and Ordinal Longitudinal Responses with Missing Responses, *Journal of Statistical Theory and Applications*, **10(3)**:357-378.

Benjamin, M., Rigby, R. Stasinopoulos, D. (2003). Generalized autoregressive moving average models, *Journal of the American Statistical Association*, **98**, 214-223.

Catalano, P. J. (1997). Bivariate modelling of clustered continuous and ordered categorical outcomes, *Statistics in Medicine*, **16**, 883-900.

Chib, S. and Greenberg, E. (1998). Analysis of multivariate probit model, *Biometrika*, **85**, 247-261.

Cox, D. R and Wermuth, N. (1992). Response models for mixed binary and quantitative variables, B*i*ometrika, **79 (3)**, 441-461.

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

De Leon, A. R. and Wu, B. (2011). Copula-based regression models for a bivariate mixed discrete and continuous outcome. *Statistics in Medicine*, **30(2)**, 175-185.

Escarela, G., Mena, R. and Castill-Morales, A. (2006). A flexible class of parametric regression models based on copulas: application to poliomyelitis incidence, *Statistical Methods in Medical Research*, **15**, 593-609.

Fitzmaurice, G. M., and Laird, N. M. (1995). Regression models for a bivariate discrete and continuous outcome with clustering, Journal of the American Statistical Association, **90**, 845-852.

Fitzmaurice, G. M., and Laird, N. M. (1997). Regression models for mixed discrete and continuous responses with potentially missing values, *Biometrics*, **53**, 110-122.

Ganjali, M. (2003). A model for mixed continuous and discrete responses with possibility of missing responses, *Journal of Sciences, Islamic Republic of Iran*, **14 (1)**, 53-60.

Harry, J. and James, J. (1998). The Estimation Method of Inference Functions for Margins for Multivariate Models, *University of British Columbia*.

Heckman, J. J. D. (1978). Endogenous variable in a simultaneous Equation system, *Econometrica*, **46 (6)**, 931-959.

Joe, H. and Xu J. J. (1996). The estimation method of inference functions for margins for multivariate models. *Technical Report, Department of Statistics, University of British Columbia*.

Joe, H. (1993). Parametric families of bivariate distributions with given marginals, Journal of Multivariate Analysis, **46**, 262-282.

Kolev, N., Anjos, U. and Mendes, B. (2006). Copulas: a review and recent developments. *Statistical Models*, **22**, 617-660.

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

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

Moustaki I., Knott M. (2000). Generalized latent trait models. *Psychometrika*, **65**(3):391-411.

Marshall, A. (1996). *Copulas, marginals, and join distributions. In: distributions with Fixed Marginals and Related Topics (L. Ruschendorf, B. Sweitzer and M.D. Taylor (Eds.).* Institute of Mathematical Statistics: Hayward, 213-222.

Niewiadomska-Bugaj, M. and Kowalczyk, T. (2005). On grade transformation and its implications for copulas, *Brazilian Journal of Probability and Statistics*, **19**, 125-137.

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

Oakes, D. and Ritz, J (2000). Regression in a bivariate copula model, *Biometrika*, **87** (2): 345-352.

Pinto, A. T. and Normand, S. L. T. (2009). Correlated bivariate continuous and binary outcomes: Issues and applications, *Statistics in Medicine*, **28**, 1753-1773.

Poon, W. Y. and Lee, S. Y. (1987). Maximum likelihood estimation of multivariate polyserial and polychoric correlation coeficients. *Psychometrika*, **52** (3): 409-430.

Rubin, D. B. (1976). Inference and missing data, *Biometrica*, **82**, 669-710.

Sammel, M. D., Ryan, L. M., Legler, J. M. (1997). Latent variable models for mixed discrete and continuous outcomes, *Journal of the Royal Statistical Society, Series B, Methodological*, **59**, 667-678.

Song, P. X-K., Li, M., and Yuan, Y. (2009). Joint regression analysis of correlated data using Gaussian copula, *Biometrics*, **65**, 60-68.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges, Publications de l'Institut de Statistique de l'Université de Paris 8, 229-231.

Song, P. X. (2000). Multivariate dispersion models generated from Gaussian copula, *Scandinavian Journal of Statistics*, **24**, 305-320.

Tiit, E. and Kaarik, E. (1996). *Generation and investigation of multivariate distributions having fixed discrete marginals. In: distributions with Fixed Marginals and Related Topics*

(L. Ruschendorf, B. Sweitzer and M.D. Taylor), Institute of Mathematical Statistics: Hayward, 113-127.

Van Ophem, H. (1999). A general method to estimate correlated discrete random vector. *Econometric Theory*, **15**, 228-237.

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

Zimmer, D. M., and Trivedi, P. K. (2006). Using trivariate copulas to model sample selection and treatment effects: Application to family health care demand, *Journal of Business and Economic Statistics*, **24**, 63-76.

Table 1: Relative bias of joint (J) and marginal (M) estimates of parameters with NMAR and MAR mechanisms under five different scenarios

Scenario (J)		N	Parameter α_1	β_1	β_2	σ	ρ_{12}	ρ_{23}	ρ_{13}	γ_1	θ_1	θ_2
A	NMAR	100	0.06	0.11	-0.13	-1.42	-0.56	0.38	1.73	0.21	-0.02	0.13
		200	0:01	0:04	-0.01	-0.19	-0.043	0.16	0.09	0.14	0.21	0.02
	MAR	100	0.21	0.15	0.31	−0.92	−0.66	−	1.83	0.18	0.31	0.34
		200	0.23	0.09	0.02	−0.04	−0.071	−	0.12	0.23	0.02	0.29
B	NMAR	100	0.12	−0.15	0.23	−1.12	0.23	−0.41	0.12	−0.11	−0.12	0.28
		200	−0.21	0.12	0.34	−0.82	0.04	0.07	0.27	0.25	0.19	−0.01
	MAR	100	0.18	0.20	0.22	−1.01	0.31	−	−0.52	−0.04	0.11	−0.02
		200	0.31	0.14	0.41	−0.71	−0.01	−	0.31	0.24	−0.02	0.1
C	NMAR	100	0.04	0.17	−0.08	−1.51	−0.15	0.33	0.21	0.03	0.13	0.21
		200	−0.01	−0.03	0.06	−0.08	0.22	−0.15	0.33	−0.01	0.04	0.33
	MAR	100	0.14	0.14	0.13	−0.05	0.02	−	0.25	0.12	0.39	0.41
		200	0.22	0.02	0.08	−0.09	0.12	−	−0.09	0.05	−0.01	0.45
D	NMAR	100	0.1	0.02	0.23	−0.98	0.44	−0.08	0.42	0.14	−0.2	0.08
		200	0.24	0.13	0.17	−1.45	0.56	0.52	0.11	0.28	−0.13	0.05
	MAR	100	0.08	−0.01	0.32	−0.72	0.71	−	0.51	0.39	0.01	0.12
		200	0.27	0.25	0.15	−1.13	0.62	−	0.01	0.22	0.51	0.21
E	NMAR	100	0.11	0.17	0.12	−1.09	−0.01	0.42	−0.23	0.12	0.32	0.02
		200	0.35	0.08	−0.02	−0.81	0.10	0.12	−0.11	−0.01	0.43	−0.03
	MAR	100	0.47	0.26	0.28	−0.52	0.02	−	−0.05	0.25	0.58	0.22
		200	0.28	0.14	0.09	−0.96	−0.13	−	0.12	0.05	0.42	−0.02

Scenario (M)	N	Parameter α_1	β_1	β_2	σ	ρ_{12}	ρ_{23}	ρ_{13}	γ_1	θ_1	θ_2

Table 1 (Continues)

A	NMAR	100	0.12	0.41	−0.04	−1.05	0.06	0.49	0.11	0.44	−0.28	0.25
		200	0.11	0.08	0.02	−0.89	−0.11	0.23	0.09	0.31	−0.03	0.12
	MAR	100	0.31	0.18	0.29	−0.81	0.11	−	0.13	0.59	0.48	0.41
		200	0.28	0.30	062	−0.01	0.42	−	0.44	0.48	0.29	0.39
B	NMAR	100	0.14	0.19	0.81	−1.01	0.44	0.52	−0.02	0.19	0.14	0.52
		200	0.12	−0.01	0.25	−0.68	0.14	0.47	0.56	−0.18	0.54	0.11
	MAR	100	0.32	0.64	0.83	−0.82	0.58	−	0.23	0.39	0.19	−0.01
		200	0.43	0.29	0.44	−0.51	0.22	−	0.59	0.81	0.65	−0.12
C	NMAR	100	0.14	0.31	0.11	−1.00	0.16	0.31	0.28	0.05	0.24	0.45
		200	0.13	0.89	0.15	−0.29	0.33	0.52	0.45	0.22	0.02	0.59
	MAR	100	0.32	0.65	0.91	−0.93	−0.11	−	0.23	0.26	−0.20	0.52
		200	0.43	0.85	0.42	−0.09	0.58	−	0.02	0.43	0.38	0.63
D	NMAR	100	−0.12	0.03	0.21	−0.62	0.51	0.18	0.39	−0.08	0.14	0.11
		200	0.48	0.64	0.48	−1.22	0.61	−0.03	0.13	−0.11	0.25	0.24
	MAR	100	0.29	−0.02	0.84	−0.41	0.59	−	0.48	0.18	019	0.17
		200	0.51	0.86	0.66	−0.91	0.85	−	0.56	0.69	0.36	0.32
E	NMAR	100	0.31	0.28	−0.04	−0.86	0.02	0.52	−0.12	0.31	0.61	0.12
		200	0.42	0.17	−0.05	−0.42	−0.05	0.37	−0.10	0.48	0.52	0.16
	MAR	100	0.64	0.39	0.34	−1.03	0.12	−	0.15	0.28	0.79	0.33
		200	0.71	0.42	0.19	−0.54	0.48	−	0.53	0.52	0.49	0.01

Table 3: The marginal and joint estimation for model (3) with NMAR and MAR mechanisms.

		Parameter					
	OS^*	Age(α_1)	Ca(α_2)	Ta (α_3)	Job (α_4)	Cut point (θ_1)	Cut point(θ_2)
Joint	Es.t.	0.09^{**}	0.11^{**}	0.07^{**}	−0.53	0.18	0.47
	S. E.	0.02	0.03	0.01	0.41	0.13	0.17
	Es.t.	0.11^{**}	0.17^{**}	0.13^{**}	−0.03	0.22	0.28
	S. E.	0.05	0.06	0.04	0.65	0.14	0.18
Marginal	Es.t.	0.010^{**}	0.14^{**}	0.05^{**}	−0.43	0.24	0.41
	S. E.	0.03	0.08	0.02	0.48	0.17	0.19
	Es.t.	0.13^{**}	0.21^{**}	0.15^{**}	−0.05	0.19	0.22
	S. E.	0.06	0.07	0.05	0.12	0.22	0.25
	Waist	Constant(β_0)	Age(β_1)	Ca(β_2)	Ta(β_3)	Job(β_4)	σ_1^2
Joint	Es.t.	35.12^{**}	0.08	0.15	0.33^{**}	−0.14	21.9^{**}
	S. E.	6.15	0.09	0.08	0.12	0.11	0.25
	Es.t.	35.76^{**}	0.12	0.20	0.28^{**}	−0.08	22.01^{**}
	S. E.	7.01	0.08	0.14	0.13	0.10	0.26

Marginal	Es.t.	36.33**	0.09	0.13	0.34**	−0.12	24.6**
	S. E.	6.55	0.08	0.07	0.14	0.13	0.23
	Es.t.	36.65**	0.11	0.23	0.24**	−0.06	23.41**
	S. E.	7.44	0.05	0.16	0.12	0.11	0.23

	BMI	Constant(β_0)	Age(β_1)	Ca(β_2)	Ta(β_3)	Job(β_4)	σ_2^2
Joint	Es.t.	33.11**	−1.06**	0.11	0.65	0.45	27.3**
	S. E.	5.05	0.41	0.03	0.12	0.67	0.53
	Es.t.	35.32**	−1.02**	0.12	0.42	0.28	28.01**
	S. E.	5.65	0.58	0.13	0.16	0.29	0.67
Marginal	Es.t.	32.08**	−1.09**	0.15	0.47	0.34	29.87**
	S. E.	6.05	0.09	0.16	0.30	0.23	0.65
	Es.t.	34.11**	−1.11**	0.20	0.41	0.36	29.41**
	S. E.	6.42	0.03	0.21	0.13	0.14	0.63

		R^*_{OS}	Age(γ_1)	Ca(γ_2)	Ta (γ_3)	Job (γ_4)
Joint	NMAR	Es.t.	0.23	0.37	0.12	0.25
		S. E.	0.19	0.21	0.13	0.23
	MAR	Es.t.	0.16	0.22	0.20	0.19
		S. E.	0.21	0.23	0.19	0.16
Marginal	NMAR	Es.t.	0.31	0.11	0.16	0.23
		S. E.	0.12	0.15	0.10	0.07
	MAR	Es.t.	0.10	0.25	0.21	0.19
		S. E.	0.13	0.34	0.44	0.15

		Correlation	ρ_{12}	ρ_{23}	ρ_{14}	ρ_{23}	ρ_{24}	ρ_{34}
Joint	NMAR	Es.t.	0.49**	0.20**	0.34**	0.16	0.45**	0.10
		S. E.	0.05	0.01	0.09	0.10	0.09	0.11
	MAR	Es.t.	0.47**	–	0.32**	0.15	0.46**	0.12
		S. E.	0.08	–	0.11	0.16	0.11	0.11
Marginal	NMAR	Es. t.	0.48**	0.19**	0.33**	0.14	0.52**	0.12
		S. E.	0.07	0.03	0.08	0.13	0.21	0.15
		Es.t.	0.46**	–	0.38	0.15	0.43**	0.16
		S. E.	0.09	–	0.22	0.14	0.10	0.14

Models			-log-likelihood
Model I	NMAR	Joint	1302.44
Model II		Marginal	1365.45
Model III	MAR	Joint	1352.98
Model IV		Marginal	1384.96

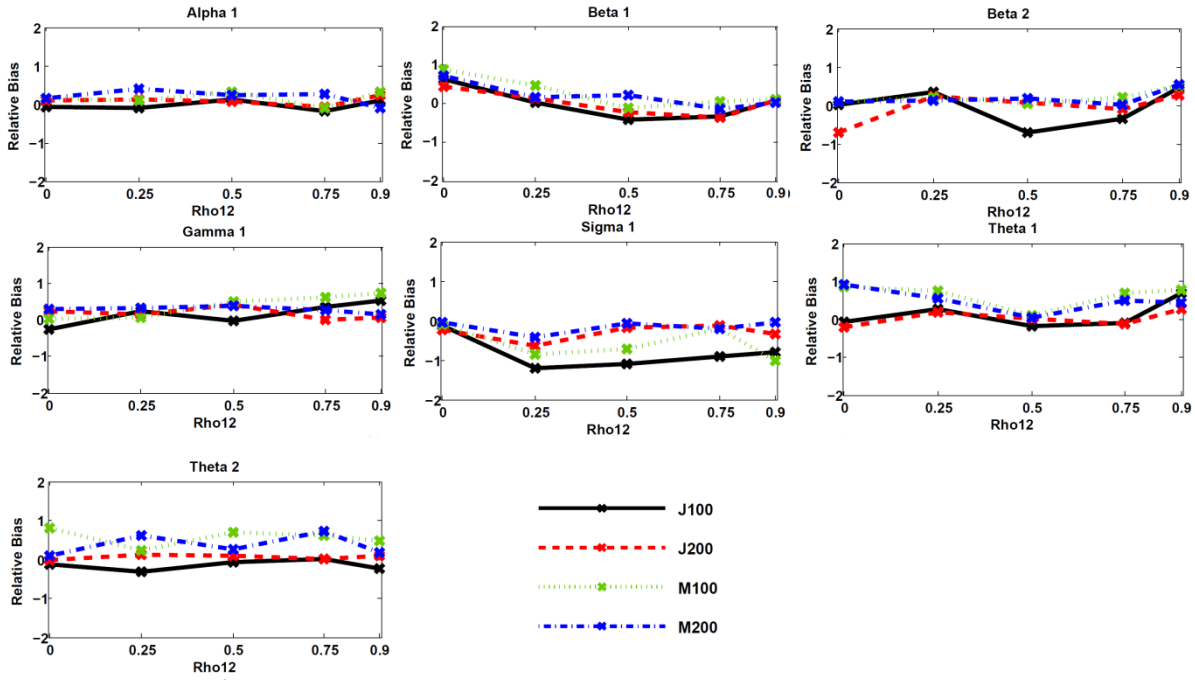

Figure 1: Relative Biases of joint and marginal estimates of α_1, β_1, β_2, γ_1, σ, θ_1 and θ_2 versus $\rho_{12}(Rho12)$ with NMAR mechanism

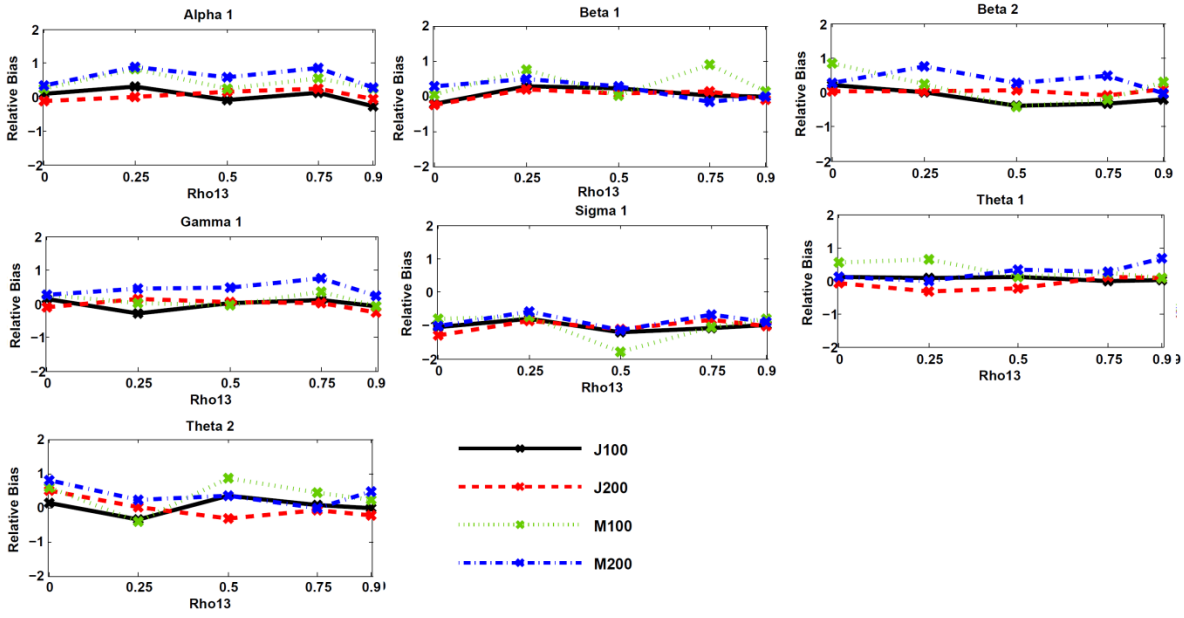

Figure 2: Relative Biases of joint and marginal estimates of α_1, β_1, β_2, γ_1, σ, θ_1 and θ_2 versus $\rho_{13}(Rho13)$ with NMAR mechanism

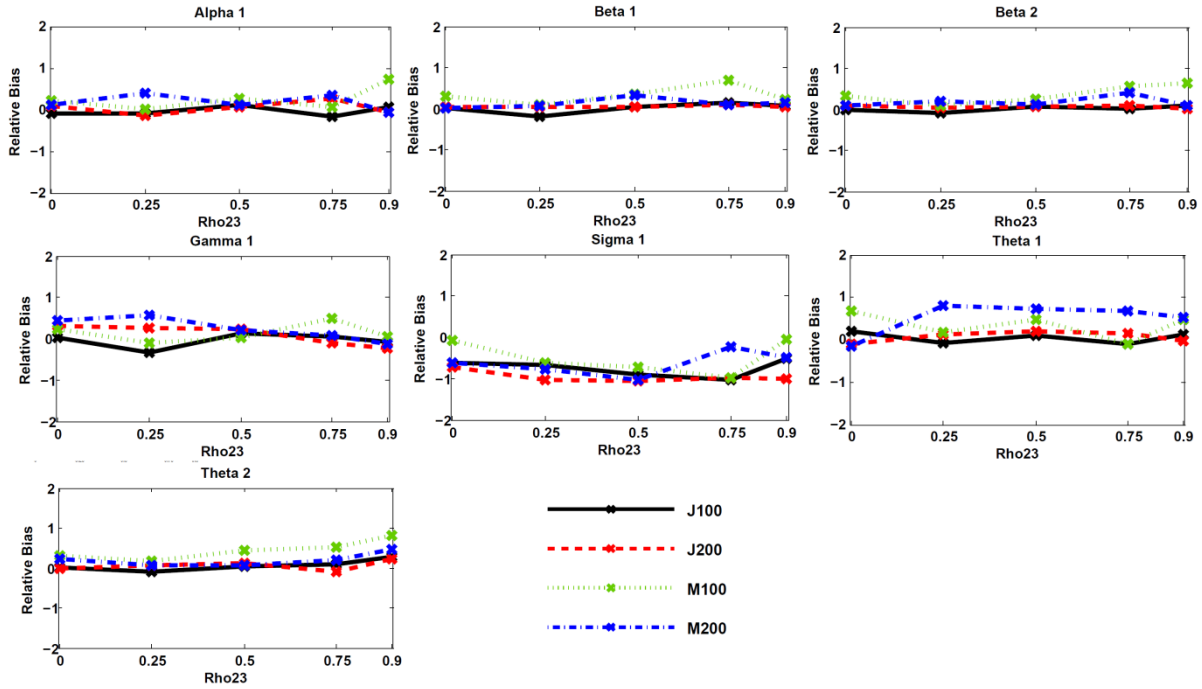

Figure 3: Relative Biases of joint and marginal estimates of α_1, β_1, β_2, γ_1, σ, θ_1 and θ_2 versus $\rho_{23}(Rho23)$ with NMAR mechanism