Journal of Sciences, Islamic Republic of Iran 30(4): 363 - 369 (2019) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

Modelling of Correlated Ordinal Responses, by Using Multivariate Skew Probit with Different Types of Variance Covariance Structures

M. R. Ghalani and M. R. Zadkarami*

Department of Statistics, Faculty of Mathematical Sciences and Computer, Shadid Chamran University of Ahvaz, Ahvaz, Islamic Republic of Iran

Received: 13 August 2018 / Revised: 11 November 2018 / Accepted: 14 September 2019

Abstract

In this paper, a multivariate fundamental skew probit (MFSP) model is used to model correlated ordinal responses which are constructed from the multivariate fundamental skew normal (MFSN) distribution originate to the greater flexibility of MFSN. To achieve an appropriate VC structure for reaching reliable statistical inferences, many types of variance covariance (VC) structures are considered to model MFSN. Simulation methods are used to find the properties of the parameters estimate. The Schizophrenia Collaborative Study data invokes the proposed MFSN model. The results confirm that the first-order autoregressive (AR(1)) structure substantially enhances the estimation of the parameters. Furthermore, over time the drugs effect the schizophrenia treatment, noticeably.

Keywords: Ordinal response; Multivariate fundamental skew probit; Variance covariance structures.

Introduction

Correlated categorical ordinal data arise in many applications related to medical, behavioral and social survey researches [1-4]. The multivariate probit (MP) model has been a popular method to model this type of data and many people have studied various aspects of the MP method [5-9]. In the MP model, the measured observations for different individuals are independent, but the measured observations at different occasions for a given individual are assumed to be correlated where the ordinal categorical responses are obtained by using discrete threshold values from a multivariate normal variable [1]. However, we can use the more flexible skew normal distribution to obtain the correlated ordinal responses. This model is called a multivariate skew probit (MSP) model.

In the literature, various forms of the multivariate skew normal distributions are defined [10-13]. However [14], discusses a class of a fundamental skew normal distribution that is more flexible to model the ordinal data. This is because the marginal and conditional distributions of the MSFN sub-vectors are also fundamental skew normal.

Usually, special VC structures have been applied in modeling the MP to avoid using expensive computational methods to analysis correlated data [15-17, 2]. However, we need to determine an appropriate VC structure to make a valid inference about the model parameters. This is achieved by finding a correct standard error such that the model parameter estimates are been consistent and asymptotically unbiased [8].

* Corresponding author: Tel: +986133331043; Fax: +986133331041; Email: zadkarami_m@scu.ac.ir

Furthermore, the procedure of the estimation of the parameters for the proposed model is involved in complicated numerical computations for the MP model. However, considering the MFSN instead of the multivariate normal yields a closed form likelihood function of the MFSP model. Therefore, we do not need to employ approximate methods that require an extensive amount of computation.

Therefore, the main point of this paper is to use the MFSP model with different VC structures to model ordinal correlated response variables to obtain more efficient parameter estimates, using the maximum likelihood method.

The paper is organized as follows. The MFSP model is illustrated in Section 2. In Section 3, the analysis of the real data is presented in conjunction with evaluating the maximum likelihood estimators properties of the MFSP model. Finally, some conclusions are demonstrated in Section 4.

Materials and Methods

Consider m individuals in a longitudinal study such that there are n_j measurements for jth individual. Now, consider a continuous latent variable $y_j^* = \left(y_{1j}^*, \dots, y_{n_j j}^*\right)$ which contains fixed covariates as:

$$y_j^* = x_j\beta + \varepsilon_j, \quad j = 1, \dots, m \qquad (2.1)$$

where β and x_j denote a p-dimensional vector of unknown regression coefficients and a $(n_j \times p)$ matrix of the fixed covariates associated with β respectively, and ε_j is the $(n_j \times 1)$ vector of the model errors. The ordinal longitudinal response variable Y_{ij} is obtained by employing the latent variable y_j^* as follows. Assume the threshold levels $-\infty = a_0 < a_1 < \cdots < a_{k+1} = \infty$, the response variable $Y_{ij} = l_{ij}$ whenever $a_{l_{ij}-1} < y_{ij}^* \leq a_{l_{ij}}$, $l_{ij} = 1,2, \dots, k$. Thus, for the model error ε_{ij} distribution function F_ε, we have:

$$P\left(Y_{ij} = l_{ij}\right) = P\left(a_{l_{ij}-1} < y_{ij}^* \leq a_{l_{ij}}\right) = F_\varepsilon\left(a_{l_{ij}} - x_{ij}\beta\right) - F_\varepsilon\left(a_{l_{ij}-1} - x_{ij}\beta\right).$$

It is assumed $a_1 = 0$ for the identifiability of the model [4, 19]. We assume that the response variables for different individuals are independent, but the response variables on different occasions for a given individual are correlated. Therefore, we have a multivariate probit (MP) model for ordinal data if we consider the normal distribution for the model error ε_{ij} [18, 3]. The likelihood function is given by

$$l(\theta; y) = \prod_{j=1}^{m} P\left(Y_{1j} = l_1, \dots, Y_{n_j j} = l_{n_j} \middle| x_j; \theta\right)$$

$$= \prod_{j=1}^{m} \int_{A_{l_1}} \dots \int_{A_{l_{n_j}}} f\left(y_{1j}^*, \dots, y_{n_j j}^* \middle| x_j; \theta\right) dy_{ij}^* \dots dy_{n_j j}^*$$

where $x_j = \left(x_{1j}, \dots, x_{n_j j}\right)$ $A_{l_i} = (a_{i-1}, a_i)$ for $l_i = 1,2, \dots, k$, $i = 1, \dots, n_j$ and, θ is the model unknown parameters. Stingo et al., [19] have illustrated that in the presence of selectivity bias, the probit model is not convenient. Therefore, they applied multivariate skew normal distribution which is defined in [20] for the model error $\varepsilon_j = \left(\varepsilon_{1j}, \dots, \varepsilon_{n_j j}\right)$ which is called a multivariate skew probit (MSP) model. In this paper, a multivariate fundamental skew normal distribution defined by [14] was employed for the model error ε_j which is more flexible than the previous versions of the multivariate skew normal distribution. This model is called a multivariate fundamental skew probit (MFSP) model

Multivariate Fundamental Skew Normal

A multivariate skew normal [4], which are used by authors to model various types of correlated longitudinal responses is given by

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = 2\phi_k(\boldsymbol{z})\Phi_1(\boldsymbol{\lambda}^T\boldsymbol{z}), \quad \boldsymbol{z} \in \boldsymbol{R}^k, \quad \boldsymbol{\lambda} \in \boldsymbol{R}^k, \qquad (2.2)$$

where $\phi_k(\boldsymbol{z})$ and $\Phi_k(\boldsymbol{z})$ represents the density and cumulative distribution functions of the k-dimensional normal vector, $N_k(\boldsymbol{0}, \boldsymbol{I}_k)$, and $\boldsymbol{\lambda}$ is the vector of skewness parameters. For $\boldsymbol{\lambda} = \boldsymbol{0}$, the density function $f_{\boldsymbol{Z}}(\boldsymbol{z})$ declines to $N_k(\boldsymbol{0}, \boldsymbol{I}_k)$. Diverse versions of the multivariate skew normal distribution have been introduced by [5, 17, 9]. The MFSN distribution [14] is:

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = 2^m\phi_k(\boldsymbol{z}|\mu, \Omega + \Lambda\Lambda^T)\Phi_m(\Lambda^T(\Omega + \Lambda\Lambda^T)^{-1}(\boldsymbol{z} - \boldsymbol{\mu})|I_m - \Lambda^T(\Omega + \Lambda\Lambda^T)^{-1}\Lambda)$$

which is denoted by $\boldsymbol{Z} \sim SN_{k,m}(\mu, \Omega, \Lambda)$ and by $\boldsymbol{Z} \sim SN_m(\mu, \Omega, \Lambda)$ for $k = m$. For $\Lambda = 0$ $SN_{k,m}(\mu, \Omega, \Lambda)$ mitigates to the multivariate normal $N_k(\mu, \Omega)$ distribution. Sahu et al, [17] defined the multivariate fundamental skew normal which is a generalized version of the multivariate skew normal. Arellano-Valle and Genton [14] illustrated that the MFSN which has

many interesting properties which are not available in (2.2). Furthermore, a random variable $\boldsymbol{Z}$ is distributed as a multivariate canonical fundamental skew normal, $CSN_{k,m}(\Delta)$, when its density function is:

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = 2^m \phi_k(\boldsymbol{z})\Phi_m(\Delta^T \boldsymbol{z}|I_m - \Delta\Delta^T), \qquad \boldsymbol{z}\epsilon R^k. \quad (2.3)$$

Lemma 1 Let $\boldsymbol{Z}$ distribute as the multivariate canonical fundamental skew normal $CSN_{k,m}(\Delta)$ in the form of (2.3). Then, the cumulative distribution function $\boldsymbol{Z}$ is given by

$$F_{\boldsymbol{Z}}(\boldsymbol{z}) = 2^m \Phi_{k+m}\big((\boldsymbol{z}^T \quad 0^T)\big|\Omega\big), \qquad \boldsymbol{z}\epsilon R^k$$

where $\Omega = \begin{pmatrix} I_k & -\Delta \\ -\Delta^T & I_m \end{pmatrix}$

Proof: [1]

Likelihood Function for MFSP Model

Let the model error ε_j in the latent variable Y_j^* defined in (2.1) be distributed as a multivariate fundamental skew normal as:

$$\varepsilon_j \overset{ind}{\sim} SN_{n_j}(0, D, \Delta_e) \quad (2.4)$$

where D and Δ_e represent respectively the variance covariance and skewness matrices of the model errors, ε_j, $(j = 1,2,\dots,m)$. Therefore, for the latent variable Y_j^*, the density function is:

$$f_{Y_j^*}\big(y_j^*\big|\theta\big) = 2^{n_j}\phi_{n_j}\big(y_j^*\big|X_j\beta, D + \Delta_e\Delta_e^T\big)\Phi_{n_j}$$
$$\Big(\Delta_e^T(D + \Delta_e\Delta_e^T)^{-1}\big(y_j^* - x_j\beta\big)\Big|I_{n_j} - \Delta_e^T(D + \Delta_e\Delta_e^T)^{-1}\Delta_e\Big) \quad (2.5)$$

$j = 1,2,\dots,m$ and the cumulative distribution function of Y^* using the density function (2.5) and lemma 1 is as follows.

$$F_{Y_j^*}\big(y_j^*\big) = 2^{2n_j}\Phi_{2n_j}\left(\begin{pmatrix} y_j^* - x_j\beta \\ 0 \end{pmatrix}\middle|\begin{pmatrix} D + \Delta_e\Delta_e^T & \Delta_e^T \\ \Delta_e & \mathrm{I}_{n_j} \end{pmatrix}\right) \quad (2.6)$$

Moreover, the ordinal response variable Y is obtained by employing the latent variable Y* in a way that for threshold levels $-\infty = a_0 < a_1 < \cdots < a_{k+1} = \infty$, the response variable $Y_{ij} = l_{ij}$, when $a_{l-1} < y_{ij}^* \leq a_l$, $l_{ij} = 1,2,\dots,k, i = 1,2,\dots,n_j,\ j = 1,2,\dots,m$. The jth likelihood function for the ordinal response Y_j is:

$$L_j(\theta) = P\big(Y_j = \boldsymbol{l_j}\big) = P\left(\boldsymbol{a_{l_j-1}} < Y_j^* \leq \boldsymbol{a_{l_j}}\right) =$$
$$F_{Y_j^*}\left(\boldsymbol{a_{l_j}}\right) - F_{Y_j^*}\left(\boldsymbol{a_{l_j-1}}\right)$$

where $Y_j = \left(Y_{1j},\dots,Y_{n_jj}\right)$, $Y_j^* = \left(Y_{1j}^*,\dots,Y_{n_jj}^*\right)^T$ $\boldsymbol{a_l} = \left(a_{1l},\dots,a_{n_jl}\right)^T$ and $\boldsymbol{l_j} = \left(l_{1j},\dots,l_{n_jj}\right)^T$.

The likelihood function is given by

$$l(\theta; y) = \prod_{j=1}^{m} L_j(\theta) = \prod_{j=1}^{m}\left[F_{Y_j^*}(\boldsymbol{a_l}) - F_{Y_j^*}(\boldsymbol{a_{l-1}})\right]$$

where $F_{Y_j^*}$ denotes the cumulative distribution function of Y_j^*. Therefore, utilizing the fundamental multivariate skew normal $SN_m(0, D, \Delta_e)$ for the model error (2.1) has this benefit that the distribution function of the latent variable, Y_j^*, and the response variable, Y_j, have closed forms and it results in closed forms for the maximum likelihood function of the desired model. Thus, there is no need to employ extensive numerical methods to achieve the parameters estimate. After employing the Newton-Raphson method, Maximization the likelihood function is done applying appropriate functions in R software.

The Structure of VC Matrix

To verify an appropriate VC structure for the response variables over time, several VC matrix structures have been considered by different authors [18]. Let $\varepsilon_j = \left(\varepsilon_{1j}, \varepsilon_{2j}, \varepsilon_{3j}, \varepsilon_{4j}\right)$ be the model error. The structures which are usually used for the VC matrix are simple, equal correlations, Toeplitz 2 bands, Toeplitz, first-order autoregressive and unstructured independence.

In the simple (SIM) structure of VC, the model error components are uncorrelated $(Cov\left(\varepsilon_{ij}, \varepsilon_{kj}\right) = 0, i \neq k = 1,2,3,4)$ with identical variance ${\sigma_e}^2$ $(Var\left(\varepsilon_{ij}\right) = {\sigma_e}^2, i = 1,2,3,4)$. Thus, the VC matrix will be diagonal, i.e., a multiple of identity matrix (${\sigma_e}^2\mathrm{I}$). As a result, the latent variables components Y_j^* and the components of the response variable Y_j are independent over time. If the VC matrix has an equal correlation (EC) structure, all correlations between model error components will be the same and equal ρ_e $(Var\left(\varepsilon_{ij}\right) = {\sigma_e}^2, i = 1,2,3,4\ and\ Cov\left(\varepsilon_{ij}, \varepsilon_{kj}\right) = {\sigma_e}^2\rho_e, i \neq k = 1,2,3,4)$. Therefore, all correlations between response variables Y_j are the same. Thus the VC structure reduces to SIM's whenever $\rho_e = 0$. In the Toeplitz 2 bands (TOEP2) structure of VC, the components of model error are correlated with adjacent components, however independent of other elements $(Var\left(\varepsilon_{ij}\right) = {\sigma_e}^2, i = 1,2,3,4\ and\ Cov\left(\varepsilon_{ij}, \varepsilon_{kj}\right) = {\sigma_e}^2\rho_e\ where\ abs(i -$

$k) = 1$). The Toeplitz (TOEP) type of VC structure generalizes TOEP2 to k-th adjacent component ($Var(\varepsilon_{ij}) = {\sigma_e}^2, i = 1,2,3,4\ and\ Cov(\varepsilon_{ij}, \varepsilon_{kj}) = {\sigma_e}^2 \rho_{de}\ where\ abs(i-k) = d$). As a result, the TOEP structure can be perceived as a moving average (MA) structure of order equal to the size of the matrix. Moreover, a TOEP2 matrix corresponds to MA structures of lower order. Therefore, model error components are correlated to each other with different values. In first-order autoregressive (AR(1)) structure of the VC matrix, the correlation coefficient decreases over time ($Var(\varepsilon_{ij}) = {\sigma_e}^2, i = 1,2,3,4\ and\ Cov(\varepsilon_{ij}, \varepsilon_{kj}) = {\sigma_e}^2 \rho_e^d\ where\ abs(i-k) = d$). Finally, in an unstructured independent (UN1) structure, the VC matrix is diagonal ($Var(\varepsilon_{ij}) = {\sigma_{ie}}^2, i = 1,2,3,4\ \ and\ Cov(\varepsilon_{ij}, \varepsilon_{kj}) = 0, i \neq k = 1,2,3,4$). The following relationship is found between various types of VC structure matrices. SIM is a sub-model of UN1 when in UN1 considered $\sigma_{1e}^2 = \sigma_{2e}^2 = \sigma_{3e}^2 = \sigma_{4e}^2 = {\sigma_e}^2$ and EC is a sub-model of TOEP when in TOEP structure it is supposed $\rho_{1e} = \rho_{2e} = \rho_{3e} = \rho_e$. Structures AR(1) and TOEP2 are the sub-models of TOEP when in TOEP we assume $\rho_{he} = \rho_e^h$, $h = 1,2,3$ and $\rho_{1e} = \rho_e$ and $\rho_{2e} = \rho_{3e} = 0$, respectively.

Results

In this study, the schizophrenic data [21] is used to investigate risk factors associated with the intensity of the schizophrenic disease. The data included information on schizophrenic patients. The disease intensity of these has been measured for four consecutive weeks of the illness. The information from 312 complete cases of observations is selected. The data is analyzed, via MP and MSFP models such that the response variable Y demonstrates the disease intensity. The response variable was 85 items of the Inpatient Multidimensional Psychiatric Scale, scored as follows: (a) normal or borderline mentally ill, (b) mildly or moderately ill, (c) markedly ill, and (d) severely or the most extremely ill which are denoted by 1, 2, 3 and 4 respectively so that number one demonstrates the lowest intensity and number four indicates the highest. The patients were randomly allocated to receive one of four medications: placebo, chlorpromazine, fluphenazine, or thioridazine. Because previous studies illustrated similar effects for the three antipsychotic drug groups, they were assigned as one group (drug) in the analysis [21]. The covariates used in the previous analyses were: TX (=0,1) the type of drug administered to the patient (i.e. real drug=1, placebo=0), SW is the square root of the patient's week to referral. The measurement taken place at weeks 0, 1, 3, and 6 and finally interaction term, i.e., $SW.TX$.

The following latent variable is used to analyze the data.

$$y_{ij}^* = \beta_0 + \beta_1 TX_{ij} + \beta_2 SW_{ij} + \beta_3 SW_{ij}.TX_{ij} + \varepsilon_{ij} \quad i = 1, \dots, n_j \qquad (3.1)$$

where $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$ denotes the covariate coefficients, $n_j = 4$ and, the distribution of the model error is the MFSN. The skewness matrix Δ_b of the model error is assumed to be diagonal with values $\delta_{1e} = \delta_{2e} = \delta_{3e} = \delta_{4e} = \delta_e$. VC matrix D can take different structures. Furthermore, it is assumed that $\sigma_e^2 = 1$ for scale normalization in using EC, TOEP, TOEP2, and AR(1) structures. Many VC structures, e.g. TOEP, TOEP2, and AR(1) have merely been employed when the measurements are considered at equal time intervals. The VC structures TOEP, TOEP2 and AR(1) are altered to be suitable for non-equal time intervals by omitting the rows and columns of the corresponding time intervals which have not observations in the data.

Finally, two MP ($\varepsilon_j \overset{ind}{\sim} N_q(0, D)$) and MFSP ($\varepsilon_j \overset{ind}{\sim} SN_q(0, D, \Delta_e)$) models are fitted to the data with different structures for the VC matrix D. The values of the Akaike Information Criterion (AIC) which is a criterion of the good modeling measure regarding different types of the VC structures for models MP and MFSP are presented in Table 1. Table 1 shows that MFSP produces a better fit because it has lower AIC for all VC structures. Thus, the MFSP model is more reliable than the MP model to analyze this data.

As can see in Table 1, the VC structure AR(1) produces the lowest AIC (7620.404) among the entire VC structures in the MFSP model. Therefore, it is the

Table 1. The values of AIC goodness for multivariate probit (MP) and multivariate fundamental skew probit (MFSP) models

The structure of the VC matrix	Multivariate Probit	Multivariate Fundamental Skew Probit
UN1	9779,810	9320,742
SIM	9086,822	8648,004
TEOP	7801,650	7724,948
TEOP2	8447,410	8123,196
AR(1)	7975,716	7620.404
EC	8102,992	7983,062

best model among 8 models.

Table 2 presents the results of the model fitting for VC structures UN1, SIM, TEOP, TEOP2, AR(1) and EC for the MFSP.

The treatment groups (drug users) are substantially different from the baseline group in the various type of VC structures. Significant negative treatment effects emphasizes that the drugs improve patients' health as displayed in Table 2. However, other authors, employed different types of ordinal longitudinal models, have shown that the treatment groups were not considerably different from each other, [21,22]. Furthermore, time covariate is negatively significant which shows that the placebo group also gets better over time but the drug groups have a greater improvement. Table 2 also illustrates that the interaction of drug by time is statistically significant which demonstrates that patients with the highest disease severity show the greatest recovery over time by employing the drug. Furthermore, it can be noticed from Table 2 that the estimation of the threshold levels parameters a_1 and a_2 are 1.582 and 2.714, respectively. This invokes that the drug effects are different for the intensity levels of the schizophrenic disease.

Thus, in schizophrenia treatment, our results support prescribing appropriate drugs; since they decrease the intensity of disease with the passage of time. Also, Table 2 shows in AR(1), a high significant correlation between sequential responses over time ($\rho_{1b} = 0.897$) which decreases as time increases.

According to results in Table 2, $\hat{\delta}_e = -1.081$ which is statistically significant. This demonstrates that the density function of the model error is negatively skewed. Furthermore, in the AR(1) model, the maximum likelihood estimate (MLE) of σ_e is highly statistically significant ($\sigma_e = 1.624$) which shows that the MFSP model works reasonably well in fitting the proposed model to the data

Simulation

The simulation method is used to investigate the properties of the proposed MLE model parameters. We generate $K = 300$ samples with sizes 200, 300 and 500 from the MFSP model as follows. At first, a latent variable Y^* is generated as:

$$y_{ij}^* = x_{1ij}\beta_1 + x_{2ij}\beta_2 + x_{1ij}.x_{2ij}\beta_3 + \varepsilon_{ij} \quad i = 1, \dots, n_j \qquad (4.1)$$

where the covariates X_1 and X_2 are generated independently from the standard normal and binary distribution with success probability 0.6 respectively and $n_j = 4$. The covariates coefficients, $\beta_0 = 1$, $\beta_1 = 1$, $\beta_2 = 0.5$ and $\beta_3 = 0.1$ are considered. The errors $\varepsilon_j = (\varepsilon_{1j}, \varepsilon_{2j}, \varepsilon_{3j}, \varepsilon_{4j})$ are generated from a multivariate fundamental skew normal distribution with mean zero. Also, AR(1) structure is employed for VC matrix D with the variance and correlation coefficient $\sigma_e^2 = 2$ and $\rho_e = 0.5$ respectively and a diagonal skewness matrix with values: $\delta_{1e} = \delta_{2e} = \delta_{3e} = \delta_e = -0.4$. At the end the ordinal response variable Y is achieved from the latent variable Y^*, defend in equation (4.1), by applying the threshold levels: $a_1 = 0$, as a fixed constant value for identifiability, $a_2 = 2$ and $a_3 = 4$. The estimate of parameters are obtained by

Table 2. Results of parameter estimates for various types of VC structures of the MFSP model

Parameters	Parameter Estimates (Standard Deviation)					
	UN1	SIM	TEOP2	TEOP	AR (1)	EC
β_0	1.696(0.048)	2.200(0.007)	2.498(0.022)	2.512(0.017)	2.207(0.019)	2.624(0.007)
β_1	-0.075(0.025)	-0.199(0.013)	-0.167(0.007)	-0.051(0.010)	-0.037(0.003)	-0.105(0.004)
β_2	-0.624(0.010)	-1.231(0.011)	-1.122(0.021)	-1.019(0.006)	-0.954(0.014)	-1.101(0.010)
β_3	0.067(0.050)	0.174(0.013)	0.136(0.007)	0.109(0.007)	0.094(0.010)	0.112(0.019)
a_1	1.140(0.040)	1.873(0.002)	1.737(0.002)	1.671(0.010)	1.582(0.009)	1.783(0.011)
a_2	1.870(0.051)	3.251(0.009)	2.966(0.011)	2.879(0.012)	2.714(0.007)	3.121(0.016)
ρ_{1e}	-	-	0.799(0.006)	0.755(0.008)	0.897(0.019)	0.688(0.019)
ρ_{2e}	-	-	-	0.812(0.004)	-	-
ρ_{3e}	-	-	-	0.675(0.026)	-	-
ρ_{5e}	-	-	-	0.573(0.008)	-	-
ρ_{6e}	-	-	-	0.232(0.014)	-	-
δ_e	-0.521(0.027)	-2.170(0.014)	-1.359(0.012)	-0.961(0.007)	-1.081(0.016)	-1.247(0.012)
Sigma1	0.466(0.053)	1.170(0.012)	1.628(0.007)	1.822(0.007)	1.624(0.014)	1.964(0.011)
Sigma2	0.649(0.020)	-	-	-	-	-
Sigma3	0.762(0.025)	-	-	-	-	-
Sigma4	1.098(0.035)	-	-	-	-	-

Table 3. The MSE and BIAS of the MLE of the model parameters

Samples	**n=200**			**n=300**			**n=500**		
Parameter	Bias	Ratio of bias	MSE	Bias	Ratio of bias	MSE	Bias	Ratio of bias	MSE
β_0	0.017	0.017	0.027	0.002	0.002	0.013	0.002	0.002	0.012
β_1	0.057	0.057	0.034	0.013	0.013	0.014	0.003	0.003	0.010
β_2	0.026	0.051	0.035	0.020	0.039	0.019	-0.003	-0.006	0.007
β_3	0.020	0.195	0.022	-0.012	-0.116	0.019	-0.011	0.114	0.017
a_1	0.005	0.002	0.002	0.004	0.002	0.001	0.002	0.001	0.001
a_2	0.003	0.001	0.002	0.002	0.001	0.001	-0.001	-0.0001	0.001
ρ_e	0.022	0.043	0.023	0.016	0.032	0.013	-0.006	-0.012	0.010
δ_e	0.026	-0.065	0.026	0.007	-0.017	0.019	0.002	-0.006	0.011
σ_e	-0.014	-0.002	0.007	0.007	0.004	0.005	-0.001	-0.001	0.003

utilizing the maximum likelihood method and R software.

We use measures: the mean-squared error, $MSE_j = \frac{1}{K}\sum_k\left(\hat{\beta}_{jk} - \beta_j\right)^2$, the bias, $B_j = \frac{1}{K}\sum_k\left(\hat{\beta}_{jk} - \beta_j\right)$ and the ratio of bias, $R_j = \frac{\hat{\beta}_{jk}}{\beta_j}$ to evaluate parameter estimation where n = 200, 300 or 500 and $\hat{\beta}_{jk}$ represents the estimated value of β_j in the kth simulated data. Therefore, two measures, MSE_j and B_j respectively deal with the precision and accuracy of $\hat{\beta}_j$ respectively.

According to the results in Table 3, the bias and mean-squared error (MSE) of the MLE of entire parameters reduce when the sample size increases from 200 to 500. However, as can be seen that the magnitude of the declining rate is not the same for all parameters estimates. However, the simulation results indicate that the MLE's parameters are efficient and asymptotically unbiased.

Discussion

The aim of this paper is using the multivariate fundamental skew probit (MFSP) model to obtain parameters estimates of the ordinal longitudinal models where different types of VC structures are assumed. The maximum likelihood method is employed to obtain the parameters estimate. The maximum likelihood function was found to have a closed form whenever the MFSP model is used.

We recognized that assuming AR(1) structure for the VC matrix yields the least value of the goodness-of-fit criterion AIC. Therefore, it provides the best results compared to other types of VC structures in the analysis of the real data. We also obtained that using the MFSP model provided the least value of AIC and enhance the model fitting noticeably.

Our analysis of the real data demonstrated that the patients who employed drug substantially improved over time. Finally, the results of simulation show efficiency and asymptotically unbiasedness of the MLE's parameters of the MFSP model.

Acknowledgements

The authors gratefully acknowledge the Associate Editor and referee for their insightful comments and valuable suggestions, which greatly improve this work. This research was supported by the Shahid Chamran University of Ahvaz, Ahvaz, Iran (Grant no. 16670).

References

1. Li Y. and Schafer D.W. Likelihood analysis of the multivariate ordinal probit regression model for repeated ordinal responses. *Comput. Stat. Data Anal.* **52**: 3474-3492 (2008).
2. Xu H. and Craig B.A. Likelihood analysis of multivariate probit models using a parameter expanded MCEM algorithm. *Technometrics* **52** (2): 340-348 (2010).
3. Kim J.S. and Ratchford B.T. A Bayesian multivariate probit for ordinal data with semiparametric random-effects. *Comput. Stat. Data Anal.* **64:** 192-208 (2013).
4. Varin C and Czado C. A mixed autoregressive probit model for ordinal longitudinal data. *Biostatistics* **111** (1): 127-138 (2010).
5. Adugna M., Ketema M., Goshu D. and Debebe S. Market outlet choice decision and its effect on income and productivity of smallholder vegetable producers in lake Tana Basin, Ethiopia. *RAAE*, **22** (1): 83-90 (2019).
6. Chib S. and Greenberg E. Analysis of multivariate probit models. *Biometrika*, **85**: 347-361 (1998).
7. Coppellari L. and Jenkins S.P. Multivariate probit regression using simulated maximum likelihood. *Stata J.* **3** (3): 278-294 (2003).
8. Yano Y., Kato E., Ohe Y. and Blandford D. Examining the opinions of potential consumers about plan-derived cosmetics: an application combining word association, co-occurrence network and multivatiate probit. *J. Sens. Stud*. *e12484*: 1–9 (2018).
9. Li C., Poskitt D.S. and Zhao X. The bivariate probit

model, maximum likelihood estimator, pseudo true parameters and partial identification. *J. Econ*. **209** (1): 94-113 (2019).
10. Azzalini A. and Dalla-Valle A. The multivariate skew-normal distribution, *Biometrika* **83:** 715–726 (1996).
11. Sahu S.K. and Dey D.K. and Branco, M.D. A new class of multivariate skew distributions with applications to Bayesian regression models, *Can. J. Stat.* **31**: 129–150 (2003).
12. Azzalini A. The skew normal distribution and related multivariate families. *Scand. J. Stat. **32:**159-188 (2005).*
13. Genton M.G. and Loperfido N. Generalized skew-elliptical distributions and their quadratic forms. *Ann. Inst. Statist. Math.* **57** (2): 380-401 (2005).
14. Arellano-Valle R.B. and Genton M.G. On fundamental skew distribution. *J. Multivariate Ana. **96***: 93-116 (2005).
15. Mullahy J. Marginal effect in multivariate probit models. *Empir. Econ.* **32 (2)**: 447-461 (2017).
16. Gibbons R. and Wilcox-Gok V. Health service utilization and insurance coverage: A multivariate probit analysis. *J. Am. Stat. Asso.* **93**: 63-72 (1998).
17. Song X. and Lee S. A multivariate probit latent variable model for analyzing dichotomous responses. *Stat. Sinica*, **15**: 645-664 (2005).
18. Fortin M., Daigle G. and Ung C.H. A variance-covariance structure to take into account repeated measurements and heteroscedasticity in growth modeling. *Eur. J. For. Res.* **126**: 573-585 (2007).
19. Stingo F.C., Stanghellini E. and Capobianco R. On the estimation of a binary response model in a selected population. *J. Stat. plan. Infer.* **141** (10): 3293-3393 (2012).
20. Azzalini A. and Capitanio A. Some properties of the multivariate skew-normal distribution. *J. Royal Stat. Soc.: Ser. B* **61:** 579-602 (1999).
21. Hedeker D. Generalized linear mixed models. In *Encyclopedia of Statistics in Behavioral Sciences*, Edited by Everitt D. and Howell, **2**: 729-738 (2005).
22. Hedeker D. and Gibbons R.D. A random-effects ordinal regression model for multilevel analysis, *Biometrics* **50** (4): 933-944 (1994).