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. 609 - 619

Sensitivity Analysis in Correlated Bivariate Continuous and Binary Responses

Zh. Thmasebinejad[1] and E. Tabrizi[2]

[1]Institute for Endocrine Sciences
Shahid Beheshti University of Medical
Tehran, Iran
Email: jtahmasebinejad@yahoo.com

[2]Department of Statistics
Shahid Beheshti University
Tehran, Iran

Received: April 26, 2013; Accepted: Mach 24, 2015

Abstract

Factorization models for correlated binary and continuous responses are proposed. Full likelihood-based approach that yields maximum likelihood estimates of the model parameters is used. A common way to investigate if perturbations of model components influence key results of the analysis is to compare the results derived from the original and perturbed models using an influence graph. So small perturbation influence of the correlation parameters of the models on likelihood displacement and a general index of sensitivity (ISNI) are also studied. The model is illustrated using data from arthritis and body mass index data. The effect of systolic blood pressure, gender and age on arthritis and body mass index are investigated.

Keywords: Factorization models; Likelihood Displacement; Continuous and binary Outcomes; Medical Data

AMS (2010) No.: 62F03

1. Introduction

Some biomedical and health sciences data include both categorical (ordinal or nominal) and continuous outcomes. The analysis of biomedical data set with variables Arthritis and BMI as response variables and systolic blood pressure (SBP), gender and age as explanatory variables for 61 diabetic patients is a good example for such studies. Notice that body mass index (BMI) and arthritis are continuous and nominal variables, respectively (section 4). In

this example, separate analyses cannot assess the effect of SBP, age and gender on body mass index and Arthritis. Furthermore, separate analyses give biased estimates for the parameters and misleading inference. Consequently, we need to consider a method in which the mentioned variables can be modelled jointly.

For mixed correlated outcomes, one method is to use the general location model of Olkin and Tate (1961), where the joint distribution of the continuous and nominal variables is decomposed into a marginal multinomial distribution for the nominal variables and a conditional multivariate normal distribution for the continuous variables given the categorical variables. A second method for mixed correlated outcomes is to decompose the joint distribution as a multivariate marginal distribution for the continuous outcomes and a conditional distribution for nominal variables given the continuous variables. Heckman (1978) presented simultaneous models to analyze two mixed correlated outcomes. Catalano and Ryan (1992) extended the mentioned models for data sets of a clustered structure contained discrete and continuous outcomes. Consequently, the interested case effects cannot be characterized by a single outcome, but instead of multiple outcomes need to be measured on each individual under study. Cox and Wermuth (1992) discussed two possible factorizations for modeling a continuous and a binary outcome as functions of covariates.

Some researchers have investigated and proposed the mixed correlated outcomes, for example, Lin et al.(2000) proposed a scaled linear mixed model for multiple outcomes, Gueorguieva and Agresti (2001) and Gueorguieva and Sancora (2006) investigated correlated probit Model for joint modeling of clustered binary and continuous response and repeatedly observed continuous and ordinal measures of disease Severity , McClluch (2007), Deleon and Carrier (2007), Bahrami Samani et al. (2008), Bahrami samani et al. (2010), Deleon and Wu (2011) and Bahrami Samani and Ganjali (2011) proposed joint modelling of mixed outcome type using latent variables.

Some authors have investigated sensitivity analysis for example Cook (1977), Belsey et al. (1980), Cook and Weisberg (1982), and Chatterjee and Hadi (1988), among others. In this paper, we used sensitivity analysis in correlated bivariate continuous and binary responses. The main idea of the factorization method is to write the likelihood as the pro duct of the marginal distribution of one outcome and the conditional distribution of the second outcome given the previous outcome.

In the next Section, the model and the likelihood are presented and then in the subsequent Section, the used methodology will be applied on the data from the medical study medical. As a sensitivity analysis for these data, small perturbation influence of the correlation parameters of the model on likelihood displacement will also be investigated. Finally, concluding remarks are given.

2. Models and Likelihoods

Suppose two responses, Y_b as a binary response and Y_C as a continuous response are measured on the i^{th} individual. Fitzmaurice and Laird (1992) proposed a model for a correlated binary and a continuous outcome based on the factorization of the joint distribution of the outcomes. The factorization Model is assumed to take the form:

$$probit\,(E(Y_{b_i} \mid (X_{b_i})) = probit(\mu_{b_i}) = X'_{b_i}\beta_b\,, \tag{2.1}$$

$$(Y_{c_i} \mid Y_{b_i}, X_{b_i}, X_{c_i}) = X'_{c_i}\beta_c + \tau(Y_{b_i} - \mu_{b_i}) + \varepsilon_{c_i}$$

where $\varepsilon_{c_i} \sim N(0, \sigma_c^2), \tau$ is the parameter for the regression of Y_c on Y_b. X_b and X_c are vectors of explanatory variables which may be different because some variables have effect on Y_b, but not on Y_c and vice versa. The correlation that results from this model is

$$\rho = \begin{cases} \dfrac{sign(\tau)}{\sqrt{1 + \dfrac{\sigma_c^2}{\tau^2 Var(Y_{b_i} \mid X_{b_i})}}}, & \tau \neq 0, \\ 0, & \tau = 0. \end{cases}$$

The vector of parameters β_b and β_c and the scale parameters ρ, σ_c and τ should be estimated. The log-likelihood function under the factorization model (2.1) is

$$l(y_b, y_c) = \log\left[\prod_{i=1}^{n} f(y_{b_i}, y_{c_i} \mid x_{b_i}, x_{c_i})\right] = \log\left[\prod_{i=1}^{n} f(y_{c_i} \mid y_{b_i}, x_{b_i}, x_{c_i}) f(y_{b_i} \mid x_{b_i})\right]$$

$$= \sum_{i=1}^{n} \left[-\frac{1}{2}\log(2\pi\sigma_c^2) - \frac{1}{2\sigma_c^2}(y_{c_i} - X'_{c_i}\beta_c - \tau(y_{b_i} - \mu_{b_i}))^2\right]$$

$$+ \sum_{i=1}^{n} \left[y_{b_i}\log(\mu_{b_i}) + (1 - y_{b_i})\log(1 - \mu_{b_i})\right].$$

where $\mu_{b_i} = \Phi(X'_{b_i}\beta_b)$ and $\Phi(0)$ represents the cdf of the standard normal distribution. The factorization of the joint distribution of y_{b_i} and y_{c_i} can also be considered in the reverse order:

$$f(Y_b, Y_c) = f(Y_b \mid Y_c) f(Y_c).$$

The vector of parameters β_b and β_c and the scale parameters ρ', σ_c and τ' should be estimated. The model for the two outcomes is written as

$$probit\left(E(Y_{b_i} \mid Y_{c_i}, X_{b_i})\right) = probit(\mu_{b_i}) = X'_{b_i}\beta_b + \tau'\left(Y_{c_i} - X'_{c_i}\beta_c\right),$$

$$\left(Y_{c_i} \mid X_{c_i}\right) = X'_{c_i}\beta_c + \varepsilon_{c_i},$$

where $\varepsilon_{c_i} \sim N(0, \sigma_c^2), \tau'$ is the parameter for the regression of Y_b on Y_c. The correlation that results from this model is

$$\rho' = \frac{E\left[Y_{c_i}\Phi\left(X'_{b_i}\beta_b + \tau'(Y_{c_i} - X'_{c_i}\beta_c)\right)\right] - E\left[\Phi\left(X'_{b_i}\beta_b + \tau'(Y_{c_i} - X'_{c_i}\beta_c)\right)\right]\left(X'_{c_i}\beta_c\right)}{\sqrt{Var(Y_{b_i})Var(Y_{c_i})}}$$

The log-likelihood function under factorization model (2.2) is

$$\begin{aligned} l(y_b, y_c) &= \log\left[\prod_{i=1}^{n} f(y_{b_i}, y_{c_i} \mid x_{b_i}, x_{c_i})\right] \\ &= \log\left[\prod_{i=1}^{n} f(y_{c_i} \mid x_{b_i}) f(y_{b_i} \mid y_{c_i}, x_{b_i}, x_{c_i})\right] \\ &= \sum_{i=1}^{n}\left[-\frac{1}{2}\log(2\pi\sigma_c^2) - \frac{1}{2\sigma_c^2}(y_{c_i} - X'_{c_i}\beta_c)^2\right] \\ &\quad + \sum_{i=1}^{n}\left[y_{b_i}\log(\mu_{b_i}) + (1 - y_{b_i})\log(1 - \mu_{b_i})\right], \end{aligned}$$

where

$$\mu_{b_i} = \Phi\left(X'_{b_i}\beta_b + \tau'(Y_{c_i} - X'_{c_i}\beta_c)\right).$$

This likelihood can be maximized by function “nlminb” in R software. The function “nlminb” uses a sequential quadratic programming (SQP) method to minimize the requested function. The details of this method can be found in Fletcher (2000). The observed Hessian matrix may be obtained by “nlminb” function or may be provided by “fdHess” function. Standard deviations are obtained by square-root of the inverse of the diagonal elements of the observed Hessian matrix. All parameters are found to be identifiable in this section.

3. Sensitivity Analysis

Sensitivity analysis methods may be broadly classified as statistical methods. Sensitivity analysis in statistical models, under normality assumptions has been studied by Co ok (1977), Co ok and Weisberg (1982), Co ok (1986) and Chatterjee and Hadi (1988). Frey and Patil (2002) used statistical methods for sensitivity analysis including linear regression analysis, analysis of variance, response surface method, Fourier Amplitude Sensitivity Test, and Mutual Information Index. Based on a correlated bivariate continuous and binary model (model (2.1)), we used Likelihood displacement and derivation of ISNI.

3.1. Likelihood Displacement

Likelihood displacement is a useful measure of influence. We would like to have a complete influence graph of the likelihood displacement. Cook (1896) presented some general methods for assessing the local influence of minor perturbations of a statistical model.

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

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

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

$$C=\frac{\left|\dot{\alpha}_1\ddot{\alpha}_2-\dot{\alpha}_2\ddot{\alpha}_1\right|}{(\dot{\alpha}_1^2-\dot{\alpha}_2^2)^{3/2}}, \tag{3.1}$$

where the first and second derivations $\dot{\alpha}_i$ and $\ddot{\alpha}_i$ are evaluated at ω_0. Since $\dot{\alpha}_1=1$ and $\dot{\alpha}_2=\ddot{\alpha}_1=0$, C reduces to

$$C=\ddot{\alpha}_2=\ddot{L}D(\omega 0).$$

When $q>1$, an influence graph is a surface in R^{q+1}. The normal curvature C_l of the lifted line in the direction l can now be obtained by applying (3.1) to the plan curve $(\alpha, LD(\omega(\alpha)))$, where $\omega(\alpha)=\omega_0+\alpha l$, $\alpha\in R$, and l is a fixed nonzero vector of unit length in R^q. Cook (1986) proposed to look at local influences, i.e., he proposed to look at the normal curvature C_l of $\alpha(\omega)$ in ω_0, in the direction of some *q*-dimensional vector l of unit length. Let Δ_i be the *p*-dimensional vector defined by

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

and define Δ as the $p\times n$ matrix with Δ_i as its i^{th} column. Further, let $\ddot{L}$ denote the $p\times p$ matrix of second derivatives of $L(\eta\,|\,\omega_0)$ with respect to η, also evaluated at $\eta=\hat{\eta}$.

Cook (1986) has then shown that C_l can be easily calculated by

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

Obviously, C_l can be calculated for any direction l. One evident choice is the vector l_i containing one in the i^{th} position and zero elsewhere, corresponding to the perturbation of the i^{th} weight only. The corresponding local influence measure, denoted by C_i, then becomes

$$C_i = 2\left|\Delta_i^T \ddot{L}^{-1}\Delta_i\right|.$$

Another important direction is the direction $l_{\max}$ of maximal normal curvature $C_{\max}$.

3.2. Derived ISNI

The index of sensitivity to the correlation parameters of the model (ISNI) measures the extent to which the maximum likelihood estimation (MLE) of Θ for a given vector Γ_1 of the correlation parameters (denoted as $\hat{\Theta}(\Gamma_1)$ depends on Γ_1). Specifically, it measures the sensitivity of $\hat{\Theta}(\Gamma_1)$ to small departures of from its the correlated binary and continuous responses vector of zero. Troxel et al. (1998) defined ISNI as the derivative of $\hat{\Theta}$ with respect to Γ_1 at $\Gamma_1 = 0$, i.e.,

$$ISNI = \frac{\partial \hat{\Theta}(\Gamma_1)}{\partial \Gamma_1^T}\Big|_{\Gamma 1=0}.$$

One obtains $\hat{\Theta}(\Gamma_1)$ from a Taylor- series expansion of the log likelihood around $\Theta = \hat{\Theta}_0$ (the MLE of Θ assuming the correlated binary and continuous responses). A large ISNI implies substantial sensitivity. The difference $\hat{\Theta}(\Gamma_1) - \hat{\Theta}(0)$ is a sensible measure of the sensitivity when Γ_1 is perturbed around the correlated binary and continuous responses. Having a vector of the correlation parameters between binary and continuous responses (Γ_1), we need to adjust ISNI proposed by Troxel et al. (1998):

$$ISNI(\hat{\Theta}) = \frac{\partial \hat{\Theta}(\Gamma 1)}{\partial \Gamma_1^T}\Big|_{\Gamma 1=0} = -\left[\frac{\partial^2 L}{\partial \Theta \partial \Theta^T}\right]^{-1} \frac{\partial^2 L}{\partial \Theta \partial \Gamma_1^T}\Big|_{\hat{\Theta}(0)}. \qquad (4.1)$$

These index vectors measure sensitivity of the MLEs to perturbations in the individual the correlation parameters between binary and continuous responses. Also we can approximate the MLE of a smooth scalar function *f*(.) of Θ using the first order Taylor series expansion around $\Gamma_1 = 0$ as follows:

$$f\left(\hat{\Theta}(\Gamma 1)\right) \approx f\left(\hat{\Theta}(0)\right) + \left[\frac{\partial f}{\partial \Theta^T}\Big|_{\hat{\Theta}(0)} \times \frac{\partial \hat{\Theta}(\Gamma 1)}{\partial \Gamma_1^T}\Big|_{\Gamma 1=0}\right] \times \Gamma_1,$$

where

$$\frac{\partial \hat{\Theta}(\Gamma 1)}{\partial \Gamma_1^T}\Big|_{\Gamma 1=0}$$

is the sensitivity vector defined in (4.1). However, when there are no preferable direction, we would want to take the direction where the sensitivity is greatest among all possible perturbations whose norms is $\sqrt{q}$.

$$ISNI\left(f(\hat{\Theta})\right) = \sqrt{q}\left\|\frac{\partial f}{\partial \Theta^T} \times \frac{\partial \hat{\Theta}(\Gamma 1)}{\partial \Gamma_1^T}\right\|^{1/2}_{\hat{\Theta}(0),\Gamma 1=0}$$

To obtain

$$\frac{\partial^2 L}{\partial \Theta \partial \Theta^T},$$

one can use the Hessian matrix of Θ under the correlated model. For calculating

$$\frac{\partial^2 L}{\partial \Theta \partial \Gamma 1},$$

the Monte Carlo methods of approximating integrals can be utilized to calculate corresponding conditional expectations. Because ISNI depends on the measurement units of Y_{ij}, Troxel et al. (1998) proposed a scale free measure called the sensitivity transformation c defined as

$$c(\hat{\Theta}) = \left|\frac{\text{var}(Y_{ij})^{\frac{1}{2}} SE(\hat{\Theta})}{ISNI(\hat{\Theta})}\right|, \tag{4.2}$$

Where $SE(\hat{\Theta})$ is the standard error (SE) of $\hat{\Theta}$ Large values of c suggest that sensitivity occurs only in cases of extreme the correlation parameters whereas small values suggest that sensitivity may be a problem even when the correlation parameters is modest. Troxel et al. (1998) have suggested to use $c<1$ as a cut off value for important sensitivity.

4. Application

4.1. Data

The medical data set is obtained from an experimental study in the Taleghani hospital in Tehran. The mentioned data record the Arthritis and BMI for 61 diabetic patients. BMI is a statistical measure of the weight of body mass index. A person scaled height body mass index may be accurately calculated using any of the formulas such as

$$BMI = \frac{W}{H^2},$$

where W is weight (kg) and H is height (cm). BMI is indicator of someone's health status. BMI values between 18.5 and 24.9 are considered normal or healthy weight. BMI values between 25 and 29.9 are considered overweight and 30 or over are considered obese.

Arthritis (AR) is a form of joint disorder that involves inflammation of one or more joints. There are over 100 different forms of arthritis. The most common form, osteoarthritis (degenerative joint disease), is a result of trauma to the joint, infection of the joint, or age. Other arthritis forms are rheumatoid arthritis, psoriatic arthritis, and related autoimmune diseases. Septic arthritis is caused by joint infection. The major complaint by individuals who have arthritis is joint pain. Pain is often a constant and may be localized to the joint affected. The pain from arthritis is due to inflammation that occurs around the joint, damage to the joint from disease, daily wear and tear of joint, muscle strains caused by forceful movements against stiff painful joints and fatigue. Arthritis is a joint disorder featuring inflammation.

SBP is considered as explanatory variable for both Arthritis and BMI. the systolic blood pressure (SBP) is the peak pressure in the arteries which occurs near the beginning of the cardiac cycle. The normal rate for systolic in adult humans is near but less than 120 mmHg. Two other explanatory variables are age and gender.

4.2. Model

For comparative purposes, two models are considered. The first model (model I) uses only complete cases and does not consider the correlation parameter between two responses. This model is

$$probit[P(AR=1)] = \gamma_0 + \gamma_1 Gender + \gamma_2 Age + \gamma_3 SBP$$
$$BMI = \beta_0 + \beta_1 Gender + \beta_2 Age + \beta_3 SBP + \varepsilon.$$

The second model (model II) uses model I and takes into account the correlation between two responses (τ).

$$probit[P(AR=1)] = \gamma_0 + \gamma_1 Gender + \gamma_2 Age + \gamma_3 SBP$$
$$BMI = \beta_0 + \beta_1 Gender + \beta_2 Age + \beta_3 SBP + \tau(AR - E(AR)) + \varepsilon.$$

The second model shows significant effects of gender, age and SBP on BMI. For males, the average of BMI is more than that for females, and the more the value of SBP the more is the value of BMI and significant effects of age and SBP on AR. For model correlation parameter τ is strongly significant and it shows a positive correlation between BMI and AR ($\hat{\tau}$ =0.503 for model).

Table 1. Estimation of the parameters for medical data (Gender, baseline: Female) and parameter estimates highlighted in **bold** are significant at 5 % level.)

Parameter	Es.t.	S.E.
Response: *BMI*		
Constant	16.21	0.436
Gender	2.534	0.388
Age	0.081	0.007
SBP	0.053	0.002
Response: AR		
Gender	-0.552	0.541
Age	-0.067	0.011
SBP	-0.014	0.006
Variance of BMI	0.154	0.021
τ	0.503	0.003

4.3. Sensitivity for Model

1) Likelihood displacement: Let see how we can use this approach for our purposes. In our application the continuous variable is always observed, and so the condition for independence of $\omega_0 = 0$. For doing sensitivity analysis, we find the likelihood displacement for different values of $\omega = \tau$ in the interval (0:1; 0:5) and these are plotted in Figure 1. This curve shows a high curvature around $\Gamma_1 = 0$ (the response variables are correlated with each other) so that the model results change with changing values of $\omega = \tau$. hence final results of the model is highly sensitive to $\omega = \tau$.

2) Derived ISNI: Let $\Gamma_1 = \tau$ lead to the correlation parameter of binary and continuous responses. The binary and continuous responses are independent if $\Gamma_1 = 0$. To search for sensitivity analysis we find c. This is confirmed by the curvature c = 10:351 computed from (4.2). This curvature indicates extreme local sensitivity.

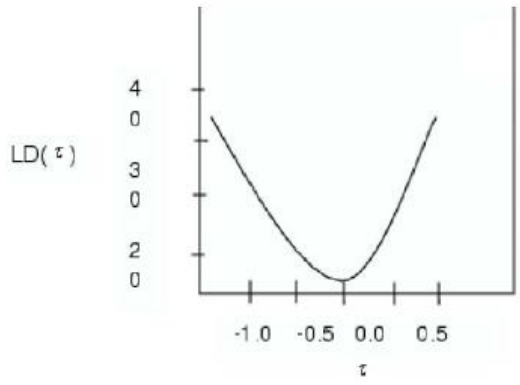

Figure 1: Likelihood displacement against values of τ

5. Conclusion

In this paper, factorization models were presented for simultaneous models with binary and continuous correlated responses. As the binary responses are special cases of ordinal responses, our model can also be used for mixed binary and continuous responses. The results

of using the proposed model for medical data show the ability of the mentioned model to better recognizing the inter dependency between two mixed responses. Generalization of our model for nominal, ordinal and continuous responses is an ongoing research on our part.

REFERENCES

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**, Part 1: 38-57.

Bahrami Samani, E. and Ganjali. M. (2011). Bayesian latent variable Model for Mixed Continuous and Ordinal Responses with Possibility of Missing Responses. *Journal of Applied Statistics*. **38**(6): 11031116.

Bahrami Samani, E., Ganjali, M. (2014). Mixed Correlated Bivariate Ordinal and Negative Binomial Longitudinal Responses with Nonignorable Missing Values. *Communications in Statistics - Theory and Methods*. (To Appear).

Besley, D., Kuh, E. and Welsch, R. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* John Wiley Sons, New York.

Catalano, P. and Ryan, L. M. (1992). Bivariate latent variable models for clustered discrete and continuous outcome. *Journal of the American Statistical Association*. **50**(3): 1078-1095.

Chatterjee, S. and Hadi, A. S. (1988). *Sensitivity Analysis in Linear Regression.* John Wiley Sons. New York.

Cook, R. D. (1977). Detection of influential observations in linear regression, *Technometrics*. **19**: 15-18.

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

Cook, R. D. and Weisberg, S. (1982). *Residual and Influence in Regression.* Chapman and Hall, London.

Cook, R. D. (1986). Assessment of Local Influence (with discussion). *Journal Royal Statist. Soc., Ser. B.*,**48**: 133-169.

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

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

Fletcher, R. (2000), *Practical Methods of optimization.* John Wiley. New York.

Frey, H.C., and R. Patil (2002). Identification and Review of Sensitivity Analysis Methods. *Risk Analysis.* **22**(3): 553-577.

Gueorguieva, R.V. and Agresti (2001). A. Correlated Probit Model for Joint Modeling of Clustered Binary and Continuous Response. *Journal of the American Statistical Association.* **96**: 1102-1112.

Gueorguieva, R.V. and Sanacora, G. (2006). Joint Analysis of Repeatedly Observed Continuous and Ordinal Measures of Disease Severity. *Statistics in Medicine.* **25**(8): 1307-1322.

Heckman. (1978). Multivariate regression analyzes for categorical data (with discussion). J. Roy. Statist. Soc. B. 54, 3-40.

Lin X, Ryan L, Sammel M, Zhang D, Padungtod C, Xu X. (2000) A scaled linear mixed model for multiple outcomes. *Biometrics.* **56(**2): 593-601.

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

McCullagh, P. (1980). Regression models for ordinal data (with discussion). *J. Roy. Statist. Soc. B.* **42**: 109-142.

Olkin and R. F. Tate (1961) .Multivariate Correlation Models with Mixed Discrete and Continuous Variables. *Ann. Math. Statist.* **32**: 448-465.

Pinto, T.A., Normand, T. (2009), Correlated bivariate continuous and binary outcomes: Issues and applications. *Statistics in Medicine.* **28**: 13:1753-1773.

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

Sammel MD, Ryan LM, Legler JM. (1997). Latent variable models for mixed discrete and continuous outcomes. *Journal of the Royal Statistical Society, Series B: Methodological*. **59**: 667-678.

Razie, F., Bahrami Samani, E. and Ganjali, M. (2013). Analysis of Mixed Correlated Bivariate Negative Binomial and Continuous Responses. *Applications and Applied Mathematics*. **8**(2), 404-415.

Troxel, A. B., Harrington, D. P., and Lipsitz, S. R. (1998). Analysis of Longitudinal Measurements with Non-ignorable Non-monotone Missing Values. *Applied Statistics*. 47: 425 - 438.

Yang, Y., Kang, J. (2011). Joint analysis of mixed Poisson and continuous longitudinal data with nonignorable missing values. *Computational Statistics Data Analysis*. **54**: 193-207.