Specification of Variance-Covariance Structure in Bivariate Mixed Model for Unequally Time-Spaced Longitudinal Data

Melike Bahçecitapar, Özge Karadağ and Serpil Aktaş*

Department of Statistics, Faculty of Science, Hacettepe University, Ankara, Turkey

Abstract: In medical studies, the longitudinal data sets obtained from more than one response variables and covariates are mostly analyzed to investigate the change in repeated measurements of each subject at different time points. In this study, the usability of multivariate models in the analysis of these kind of data sets is investigated, because it provides the joint analysis of multiple response variables over time and enables researchers to examine both the correlations of response variables and autocorrelation between measurements from each response variable over time. It has been shown that different parameter estimation methods affect the results in the analysis of multivariate unbalanced longitudinal data. We investigated that autocorrelation structure over time between measurements from same response variable should be truly specified. We also illustrated and compared the simpler, more standard models for fixed effects with multivariate models provided by SAS on a real-life data set in the joint analysis of two response variables. Results show that misspecification of autocorrelation structures has a negative impact on the parameter estimates and parameter estimation method should become of interest.

Keywords: Multivariate longitudinal data, mixed models, covariance structures.

1. INTRODUCTION

A longitudinal study is a long-term observational study performed over a period of time to study measurements gathered for the same subjects repeatedly. In many medical and epidemiological trials, longitudinal studies have repeated measurements for more than one response variables and covariates over time. As an example, longitudinal studies of Alzheimer's disease investigate different predictors and different response variables likely to be affected by these different predictors over time in the decline of disease [1].

Life course researches are interested in the individual development over time by analyzing several longitudinal covariates and responses related to the growth, health and lifestyle of subjects over time [2].

These longitudinal multivariate studies require special statistical methods in order to take into account 1) errors likely to be correlated for each response variable over time, 2) errors correlated among response variables measured at the same time point and 3) variances likely to be different for different response variables.

Multivariate repeated measurements model with a Kronecker product covariance, random coefficient mixed model and structural equation models can provide an opportunity to study this kind of longitudinal data sets under the consideration of the joint evaluation of multiple response variables over time [3-5]. Gao *et al*. (2006) [6] compared three approaches by using balanced longitudinal data with equally spaced time points. However, under the joint analysis of multivariate longitudinal data, some points such as the case that the sequence of time points is no longer common for all subjects, the misspecification of autocorrelation structure for errors within subject over time and parameter estimation methods (maximum likelihood (ML), restricted maximum likelihood (REML), minimum variance quadratic unbiased estimation (MIVQUE0)) should become of interest.

In this study we investigate multivariate repeated measurements model with a Kronecker product covariance for multiple responses measured uncommon set of time points under the consideration of first-order autoregressive (AR1), compound symmetry (CS) and unstructured (UN) autocorrelated errors within subject over time. Because variance components are estimated from unbalanced data, research was directed herein toward estimation methods (ML, REML and MIVQUE0) whose properties do not depend on balanced data in SAS version 9.2. We consider bivariate repeated measurements model with a Kronecker product covariance as the special case of multivariate models for two response variables. We also compared these models with the simpler and more standard fixed effects models in SAS PROC REG and PROC GENMOD procedures [15].

We illustrate a real-life Genetic Analysis Workshop 19 (GAW 19) longitudinal data containing replications taken unequally spaced time points between 1991 and

*Address correspondence to this author at the Department of Statistics, Faculty of Science, Hacettepe University, Ankara, Turkey; Tel: +90 312 297 79 30; Fax: +90 312 297 79 13; E-mail: spxl@hacettepe.edu.tr

E-ISSN: 1929-6029/15

2011. Subjects are measured at four different time points during the study. Our primary interest is to assess the interrelationships among two responses, when unequally spaced time points are coded in the model as equally spaced time points (i.e., 1, 2, 3, 4) or unequally spaced time points (i.e. 1,2,…, 16).

2. BIVARIATE REPEATED MEASUREMENTS MODEL WITH A KRONECKER PRODUCT COVARIANCE

We assume that the multivariate distribution of longitudinal responses follow the multivariate normal distribution. The general model definition is as follows

$$\boldsymbol{Y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{\varepsilon}_i \quad (1)$$

where $\boldsymbol{Y}_i$ is the qx1 vector of outcomes for i^{th} subject, $\boldsymbol{X}_i$ is the $q \times (p+1)$ matrix of p predictors for i^{th} subject, $\boldsymbol{\beta}$ is the $(p+1)\times 1$ vector of fixed effects and $\boldsymbol{\varepsilon}_i$ is the error vector for i^{th} subject.

Let a bivariate longitudinal data set arise when a set of response variables $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ and the covariates $\boldsymbol{X}_1,\ldots \boldsymbol{X}_p$ are measured repeatedly over time for multiple subjects. Y_{ijk} is the response of the $i^{th}(i=1,2,\ldots,N)$ subject on the $j^{th}(j=1,2,\ldots,n_i)$ time point for $k^{th}(k=1,2)$ response variable [7,8].

We used bivariate repeated measurements model using a Kronecker product covariance which allows researchers to analyze bivariate longitudinal data sets under three alternative variance-covariance structures of the measurement error within each subject by the REPEATED statement in the PROC MIXED procedure in SAS: UN@AR(1), UN@CS and UN@UN [6]. The Kronecker products specifying the covariance of UN@AR(1), UN@CS and UN@UN for the relationship between $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are given in Eq.(2):

$$\boldsymbol{V} \otimes \boldsymbol{\Sigma_i} = \begin{pmatrix} \sigma_{Y_1}^2 & \sigma_{Y_1Y_2} \\ \sigma_{Y_1Y_2} & \sigma_{Y_2}^2 \end{pmatrix} \otimes \boldsymbol{\Sigma_i} = \begin{pmatrix} \sigma_{Y_1}^2 \boldsymbol{\Sigma_i} & \sigma_{Y_1Y_2} \boldsymbol{\Sigma_i} \\ \sigma_{Y_1Y_2} \boldsymbol{\Sigma_i} & \sigma_{Y_2}^2 \boldsymbol{\Sigma_i} \end{pmatrix} \quad (2)$$

where $\boldsymbol{\Sigma_i} = \begin{pmatrix} 1 & \rho & \cdots & \rho^{n_i-1} \\ \rho & 1 & \cdots & \rho^{n_i-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n_i-1} & \rho^{n_i-2} & \cdots & 1 \end{pmatrix}$ for AR(1) autocorrelated errors, $\boldsymbol{\Sigma_i} = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}$ for CS autocorrelated errors and $\boldsymbol{\Sigma_i} = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{ni-1} \\ \rho_1 & 1 & \cdots & \rho_{ni-1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{ni-1} & \rho_{ni-2} & \cdots & 1 \end{pmatrix}$ for UN autocorrelated errors [9].

This approach using the covariance structures for bivariate repeated measurements obtained by Kronecker product covariance allows to specify inter-response correlation, i.e. the correlated errors for the same response variable measured at different time points (eg. $\boldsymbol{\Sigma_i} = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}$ in UN@CS structure), intra-response correlation, i.e. the errors likely to be correlated among response variables measured at the same time point $\left(\rho_{Y_1Y_2} = \frac{\sigma_{Y_1Y_2}}{\sqrt{\sigma_{Y1}^2 \sigma_{Y2}^2}} \right)$ and the variances of errors likely to be different for two response variables. The inter-responses correlation $\rho_{Y_1Y_2}$ of the bivariate process is assumed to be same for two responses measured at the same time point. If $\rho_{Y_1Y_2}$ is 0, then the responses are independent. Bivariate model also assumes a common intra-response correlation [4].

3. PARAMETER ESTIMATION METHODS

Two frequently used approaches for estimating all variance and covariance parameters are maximum likelihood (ML) and residual/restricted maximum likelihood (REML) estimation methods [10]. Both methods simplifies many common statistical analyses involving repeated measures [4]. In particular, REML is used as a method for fitting linear models. In contrast to ML estimation, REML can produce unbiased estimates of variance and covariance parameters.

Minimum variance quadratic unbiased estimator (***MIVQUE***) [11] available in SAS PROC MIXED for estimating variance components is an alternative method which does not require a normality assumption as in ML and REML and may be used when ML does not converge. Instead of ML or REML, the non-iterative MIVQUE method can be used to estimate variance components [12]

Newton Raphson (NR) or Expectation Maximization (EM) can be carried out as an optimization technique in PROC MIXED for the maximization of the likelihood functions [13, 14]. Lindstrom and Bates (1988) [14] showed that NR algorithm is preferred to the EM algorithm. PROC MIXED uses NR algorithm to optimize either a full (ML) or residual (REML) likelihood function.

4. NUMERICAL EXAMPLE

The data is taken from 19th Genetic Analysis Workshop (GAW 19) (http://www.gaworkshop.org/). Longitudinal real phenotype data (systolic and diastolic blood pressure, hypertension diagnosis, anti-hypertensive medication usage, age, sex, cigarette smoking) collected at up to four time points taken from 959 subjects in 20 large families. The systolic blood pressure (SBP) and diastolic blood pressure (DBP) are measured in four different time periods for Hispanic samples. The measurement periods are unequally time-spaced as Visit1: between 1991-1996, Visit2: between 1997-2000, Visit3: between 1998-2006 and Visit4: between 2009-2011.

In this study, we first analyzed this multivariate longitudinal data in SAS by both simple linear regression in PROC REG procedure and longitudinal data analysis in PROC GENMOD procedure. In all analysis, two different presentations for specifying visits of subjects, fist one is VISIT which indicates that all assessments are equally spaced and the other is UNEQVISIT which indicates real visits of subjects. Because subjects are measured four times during the study, the variable VISIT used in the MODEL statement takes 1, 2, 3 and 4 values. However, the variable UNEQVISIT takes values from 1 to 16. The illustration of some raw data from GAW 19 is given in Table **1**.

The following commands invoke the PROC REG procedure. We fit a linear regression model, with SBP or DBP as the Y (outcome) variable and VISIT or UNEQVISIT as the X (independent or predictor) variable. We analyzed four linear regression models to predict the relationship between time and SBP/DBP outcomes. Information concerning model fits were obtained. The *F-statistics* for the overall models for SBP and DBP were found to be highly significant. Under VISIT, the results were as for SBP, F=41.63, p<.0001; for DBP, F=24.44, p<.0001 and under UNEQVISIT, for SBP, F=46.41, p<.0001; for DBP, F=29.26, p<.0001. From Table **2**, each parameter was found to be significantly different from zero (p<.0001). Table **2** shows that there is a positive relationship between these two variables. When VISIT increases by one, SBP is predicted to increase by 3.56 units, and this is a significant relationship (t-value = 6.45, p<.0001). In the regression analysis of unequally time points, it is found that when UNEQVISIT increases by one, SBP is predicted to increase by 0.82 units, and this is a significant relationship (t-value = 6.81, p<.0001). The R-square values for each model which are the square of the correlation between the two variables are given in Table **3**.

```
proc reg data=gen;
model SBP DBP = VISIT;
run;

proc reg data=gen;
model SBP DBP = UNEQVISIT;
run;
```

In order to take into account errors likely to be autocorrelated over time, we also analyzed this data set with SAS PROC GENMOD [15]. The main advantage of PROC GENMOD is that it can handle general linear models as well as more complex ones. It can analyze correlated data in repeated measures

Table 1: Raw Data from GAW 19 for PROC REG Procedure

İd	SEX	YEAR	VISIT	UNEQVISIT	SBP	DBP	Bp	VAL
200031	1	1992	1	1	117	77	SBP	117
200031	1	1997	2	6	129	81	SBP	129
200031	1	2002	3	10	127	82	SBP	127
200031	1	2009	4	15	133.5	86	SBP	133.5
300138	1	1992	1	1	132	79	DBP	79
300138	1	1998	2	7	129	75	DBP	75
300138	1	2003	3	11	125	71	DBP	71
300138	1	2009	4	15	126	68.5	DBP	68.5
: .	: .	: .	: .	: .	: .	: .	: .	: .

Table 2: Parameter Estimates for Linear Regression Models in PROC REG Procedure with SBP/DBP and VISIT/UNEQVISIT Variables

Model 1: $SBP = \beta_0 + \beta_1$ ***VISIT***					
Variable	**df**	**Parameter Estimates**	**St.Error**	**t-value**	**P > t**
β_0	1	112.730	1.512	74.54	<.0001
β_1	1	3.560	0.552	6.45	<.0001
Model 2: $SBP = \beta_0 + \beta_1$ ***UNEQVISIT***					
Variable	**df**	**Parameter Estimates**	**St.Error**	**t-value**	**P > t**
β_0	1	114.464	1.219	93.85	<.0001
β_1	1	0.824	0.120	6.81	<.0001
Model 3: $DBP = \beta_0 + \beta_1$ ***VISIT***					
Variable	**df**	**Parameter Estimates**	**St.Error**	**t-value**	**P > t**
β_0	1	68.547	1.041	65.82	<.0001
β_1	1	1.880	0.380	4.94	<.0001
Model 4: $DBP = \beta_0 + \beta_1$ **UNEQVISIT**					
Variable	**df**	**Parameter Estimates**	**St.Error**	**t-value**	**P > t**
β_0	1	69.325	0.839	82.54	<.0001
β_1	1	0.450	0.083	5.41	<.0001

study when measures are assumed to be multivariate normal.

Table 3: The R-Square Values of PROC REG Analysis

	R^2
Model1	0.0599
Model2	0.0663
Model3	0.0346
Model4	0.0428

We could perform this analysis in PROC GENMOD under unequally time points with the following syntax:

```
proc genmod data=gen;
class id Bp;
model VAL=UNEQVISIT*Bp / dist=normal;
repeated subject=id / corrw type=UN;
run;
```

```
proc genmod data=gen;
class id Bp;
model VAL= UNEQVISIT *Bp / dist=normal;
repeated subject=id / corrw type=CS;
run;
```

```
proc genmod data=gen;
class id Bp;
model VAL= UNEQVISIT *Bp / dist=normal;
repeated subject=id / corrw type=AR(1);
run;
```

Tables **4** and **5** display model-based parameter estimates for each autocorrelation structure (AR(1), UN, CS) in the model with VISIT and UNEQVISIT variables, respectively.

From Tables **2**, **4** and **5**, it can be seen that in PROC GENMOD procedure, the change in both SBP and DBP over time can be analyzed separately. In PROC GENMOD, the examination of time points in the format of unequally time points change the results. When time points are assigned as equally time points, for all autocorrelation structures the change in both DBP and SBP over time was found to be significant ($p<.0001$). PROC GENMOD under consideration of unbalanced data and UN autocorrelation structure found the change over time in DBP responses insignificant ($p=0.5703$).

PROG REG and PROC GENMOD give merely simpler solution and these approaches do not correspond to joint analysis. For the joint analysis of this multivariate longitudinal data set, the following SAS

Table 4: Parameter Estimates Obtained by PROC GENMOD Procedure with VISIT Variable

	UN			CS			AR(1)		
Parameter	Estimate (St.Error)	Z	p-value	Estimate (St.Error)	Z	p-value	Estimate (St.Error)	Z	p-value
Intercept	93.0239 (1.3743)	67.69	<.0001	90.6387 (1.1946)	75.88	<.0001	89.7423 (1.1480)	78.17	<.0001
VISIT×DBP	-7.7879 (0.5738)	-13.57	<.0001	-5.4836 (0.3989)	-13.75	<.0001	-3.4780	-9.60	<.0001
VISIT×SBP	9.3727 (0.4897)	19.14	<.0001	10.9269 (0.3771)	28.98	<.0001	11.8903	32.38	<.0001

Table 5: Parameter Estimates Obtained by PROC GENMOD Procedure with UNEQVISIT Variable

	UN			CS			AR(1)		
Parameter	Estimate (St.Error)	Z	p-value	Estimate (St.Error)	Z	p-value	Estimate (St.Error)	Z	p-value
Intercept	115.3872 (16.7080)	6.89	<.0001	91.9565 (1.0072)	91.30	<.0001	91.0556 (0.9675)	94.12	<.0001
UNEQVISIT×DBP	-1.2034 (2.1204)	-0.57	0.5703	-1.4888 (0.0894)	-16.66	<.0001	-0.7614 (0.0786)	-9.68	<.0001
UNEQVISIT×SBP	2.7518 (1.3212)	2.08	0.0373	2.7487 (0.0814)	28.98	<.0001	3.0300 (0.0807)	37.53	<.0001

codes fit a bivariate repeated measurements model using a Kronecker product covariance for two response variables in PROC MIXED procedure. We note that the variable VISIT is used in the model, indicating equally spaced measurements and the variable. UNEQVISIT is used to indicate unequally spaced measurements. The variables VISITC and UNEQVISITC are treated as a class variable in the REPEATED statement in PROC MIXED procedure which have the same values as in VISIT and UNEQVISIT variables.

```
proc mixed data=gen covtest noclprint method=…;
class id Bp VISITC;
model Y=Bp*VISIT/s noint;
repeated Bp VISITC /type=UN@AR1 subject=id r rcorr;
where Y="DBP" or Y="SBP";
run;
```

In SAS, this model under unequally spaced time points for UN@AR1 covariance structure is also implemented with PROC MIXED as follows:

```
proc mixed data=gen covtest noclprint method=…;
class id Bp UNEQVISITC;
model Y=Bp*UNEQVISIT/s noint;
repeated Bp UNEQVISITC /type=UN@AR1 subject=id r rcorr;
where Y="DBP" or Y="SBP";
run;
```

Results are shown in Tables **6-8**. Model 5 and Model 6 represent the multivariate repeated measurement models with a Kronecker product covariance under equally and unequally spaced time intervals, respectively. For Model 6, results cannot be computed for the structure UN@UN because of the inequality of the time intervals under ML and REML estimation methods.

Model 5: Multivariate Repeated Measurement Model with a Kronecker product covariance under equally spaced time intervals

Model 6: Multivariate Repeated Measurement Model with a Kronecker product covariance under unequally spaced time intervals

Smaller values of AIC indicates a better fit. According to the assignment of time points as VISIT variable, Model 5 with UN@UN covariance structure under REML provide a better fit to the data set. On the other hand, in Model 6 (i.e. model based on unequally time points) with UN@UN Kronecker product

Table 6: Parameter Estimates by ML for Multivariate Repeated Measurement Models with a Kronecker Product Covariance Using VISIT or UNEQVISIT Variable

	Model5			Model6		
Fixed Effects (St.Error)	**UN@AR1**	**UN@CS**	**UN@UN**	**UN@AR1**	**UN@CS**	**UN@UN**
Slope DBP	6.072* (0.562)	2.882* (0.333)	19.527* (0.228)	6.385* (0.142)	0.592* (0.073)	NA
Slope SBP	10.337* (0.802)	5.210* (0.470)	32.293* (0.335)	10.622* (0.232)	1.083* (0.103)	NA
Random effects (St.Error)						
$\rho_{DBP \times SBP}$	0.962* (0.006)	0.970* (0.003)	NA	0	0.973* (0.003)	NA
Fit Statistics						
-2 Ln L	11234.8	11146.2	10837.3	12019.1	11156.4	NA
AIC	11246.8	11158.2	10865.3	12031.1	11168.4	NA
BIC	11265.4	11176.8	10908.7	12049.7	11187.0	NA

Table 7: Parameter Estimates by REML for Multivariate Repeated Measurement Models with a Kronecker Product Covariance Using VISIT or UNEQVISIT Variable

	Model5			Model6		
Fixed Effects (St.Error)	**UN@AR1**	**UN@CS**	**UN@UN**	**UN@AR1**	**UN@CS**	**UN@UN**
Slope DBP	6.079* (0.563)	2.885* (0.333)	19.528* (0.229)	6.385* (0.142)	0.593* (0.074)	NA
Slope SBP	10.348* (0.802)	5.215* (0.471)	32.293* (0.336)	10.622* (0.232)	1.084* (0.103)	NA
Random effects (St.Error)						
$\rho_{DBP \times SBP}$	0.962* (0.006)	0.970* (0.003)	NA	0	0.973* (0.003)	NA
Fit Statistics						
-2 Ln L	11233.7	11147.2	10839.5	12025.0	11163.4	NA
AIC	11241.7	11155.2	10863.8	12033.0	11171.4	NA
BIC	11254.1	11167.6	10900.7	12045.4	11183.8	NA

parameter estimates cannot be obtained by using ML and REML methods. Model 6 with UN@UN under MIVQUE0 estimation method gives the smaller AIC value. The slope parameter represents the average change during this longitudinal study for each BP over time and they seem to increase over time. All these changes are significantly different from zero ($p < .0001$). In Tables **6** and **7**, ML and REML give similar results. However, MIVQUE0 method can solve UN@UN Kronecker product covariance. In Table **8**, MIVQUE0 method give bigger parameter estimates under the consideration of equally spaced time points and it is seen that $\rho_{DBP \times SBP}$ value is the smallest for equally spaced time points. For Model 6, although ML, REML and MIVQUE0 give similar results for UN@AR1, MIVQUE0 method give different results for UN@CS in Table **8**. In Tables **6** and **7**, the results under ML and REML methods also reveal a strong positive correlation between DBP and SBP.

PROC GENMOD and PROC MIXED procedures find all parameter estimates significant for equally time

Table 8: Parameter Estimates by MIVQUE0 for Multivariate Repeated Measurement Models with a Kronecker Product Covariance Using VISIT or UNEQVISIT Variable

	Model5			Model6		
Fixed Effects (St.Error)	**UN@AR1**	**UN@CS**	**UN@UN**	**UN@AR1**	**UN@CS**	**UN@UN**
Slope DBP	21.038* (0.565)	25.528* (0.399)	24.729* (0.812)	6.385* (0.142)	6.064* (0.149)	6.175* (0.401)
Slope SBP	34.877* (0.915)	42.454* (0.647)	41.139* (0.978)	10.622* (0.232)	10.094* (0.243)	10.314* (0.477)
Random effects (St.Error)						
$\rho_{DBP \times SBP}$	0.534* (0.042)	-0.053* (0.018)	NA	0	0.054* (0.036)	NA
Fit Statistics						
-2 Ln L	11367.9	11792.4	13566.3	12025.0	11981.5	1.8E308
AIC	11375.9	11800.4	13590.3	12033.0	11989.5	1.8E308
BIC	11388.3	11812.8	13627.5	12045.4	12001.9	1.8E308

points. Besides PROC MIXED, for unequally time points, in the presence of UN autocorrelated errors the change in DBP is insignificant (p=0.5703) in PROC GENMOD.

5. CONCLUSION

Multivariate repeated models are useful approaches for multiple responses over time and can be computed using standard statistical package like the SAS system. SAS PROC MIXED is easily extendable to multivariate response in longitudinal studies. In this study, we point out that in the analysis of bivariate longitudinal data set, the indication of time points for unbalanced longitudinal study as equally or unequally paced time points differ the parameter estimations. We consider three different autocorrelation structures for errors (UN, CS and AR1) over time in this paper. Data are also analyzed by PROC REG and PROC GENMOD procedures. Bivariate repeated measurements model with a Kronecker product covariance is analyzed by PROC MIXED. Besides PROC REG and PROC GENMOD, multivariate analysis approach in PROC MIXED give the change in responses over time together with the possible correlation of two response variables. We also remark the effect of maximum likelihood, restricted maximum likelihood and minimum variance quadratic unbiased estimation on multivariate joint analysis of unbalanced longitudinal data. In accordance with the coding equally spaced time points, under all autocorrelation structures over time, the model works well. However, ML and REML parameter estimates are non-available for UN@UN Kronecker covariance, when time points were coded as unequally. If we code as unequally space, the parameter estimation method and the specification of errors likely to be autocorrelated should be noticed in the analysis of unbalanced multivariate longitudinal data sets.

REFERENCES

[1] Harvey, DJ, Beckett LA, Mungas DM. Multivariate modelling of two associated cognitive outcomes in a longitudinal study. J Alzheimer's Dis 2003; 5(5): 357-365.

[2] Hoekstra T, Twisk JWR. The analysis of individual health trajectories across the life couse: latent class growth models versus mixed models. The Series Life Course Research and Social Policies 2015; Vol. 4: pp. 179-195.

[3] Galecki AT. General class of covariance structures for two or more repeated factors in longitudinal data analysis. Communications in Statistics - Theory and Methods 1994; 23: 3105-3119. http://dx.doi.org/10.1080/03610929408831436

[4] Littell RC, Milliken GA, Stroup WW, Wolfinger RD. SAS System for Mixed Models, Cary, NC: SAS Institute Inc. 1996.

[5] Hatcher L. A step-by step approach to using the SAS System for factor analysis and structural equation modeling, Cary, NC: SAS Institute Inc. 1998.

[6] Gao F, Xiong C, Thompson P, Miller JP. Analyzing multivariate longitudinal data using SAS, SUGI 31 Proceedings 2006.

[7] Sy JP, Taylor JM,Cumberland WG. A stochastic model for the analysis of bivariate longitudinal AIDS data. Biometrics 1997; 53: 542-55. http://dx.doi.org/10.2307/2533956

[8] Laird NM, Ware JH. Random-effects models for longitudinal data, Biometrics 1982; 38: 963-74. http://dx.doi.org/10.2307/2529876

[9] Hedeker D, Gibbons RD. Longitudinal data analysis. Wiley. 2006.

[10] Harville DA. Maximum Likelihood approaches to variance component estimation and to related problems. J Am Stat Assoc 1977; 72: 320-340. http://dx.doi.org/10.1080/01621459.1977.10480998

[11] Rao CR. Estimation of variance and covariance components-MINQUE theory. Journal of Multivariate Analysis 1971; 1: 257-275.
http://dx.doi.org/10.1016/0047-259X(71)90001-7

[12] Wolfinger RD, Tobias RD, Sall J. Computing Gaussian Likelihoods and Their Derivatives for General Linear Mixed Models. SIAM Journal on Scientific Computing 1994; 15(6): 1294-1310.
http://dx.doi.org/10.1137/0915079

[13] Dempster AB, Laird NM, Rubin DB. Max. Likelihood from incomplete data *via* the EM algorithm.Journal of the Royal Statistical Society B 1977; 39: 1-38.

[14] Lindstrom MJ, Bates DM. Newton-Raphson and EM Algorithm for linear mixed-effects models for repeated-measures data. JASA 1988; 83: 1014-22.

[15] SAS/IML Software: usage and Reference. Version 6., (SAS Institute Inc., Carry, NC), 1990.

Received on 22-09-2015 Accepted on 01-10-2015 Published on 03-11-2015

http://dx.doi.org/10.6000/1929-6029.2015.04.04.6