BULLETIN *of the*
MALAYSIAN MATHEMATICAL
SCIENCES SOCIETY
http://math.usm.my/bulletin

Bull. Malays. Math. Sci. Soc. (2) **37**(1) (2014), 129–137

Dependence in Binary Outcomes: A Quadratic Exponential Model Approach

[1]MAHBOOBEH ZANGENEH SIRDARI, [2]M. ATAHARUL ISLAM
AND [3]NORHASHIDAH AWANG

[1,3]School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Pulau Pinang, Malaysia
[2]Department of Applied Statistics, East West University, Aftabnagar, Dhaka 1212, Bangladesh
[1]zangeneh_m@yahoo.com, [2]mataharul@yahoo.com, [3]shidah@usm.my

Abstract. Repeated measurements data appear in many applications of study subjects such as correlated binary data. Most of studies often focus on the dependence of marginal response probabilities. There is a lack of study based on joint probability distributions that yield estimation and test procedure using conditional probabilities, marginal means and correlated binary data. In this paper, the quadratic exponential form model has been extended for a Markov chain framework. This study extends the quadratic exponential model for displaying the estimation procedure for the nature and extent of dependence among the binary outcomes. In addition, a test procedure is extended to test for the goodness of fit of the model as well as for testing the order of the underlying Markov chain. The proposed model and the test procedures have been examined thoroughly with an application to elderly population data from the Health and Retirement Study (HRS) data.

2010 Mathematics Subject Classification: 62H20, 62F03, 62F10, 62N03

Keywords and phrases: Conditional model, marginal model, repeated observations, quadratic exponential form, Markov model, and dependence in outcomes.

1. Introduction

Binary Markov chain arise in various application areas such as studies of disease occurrence among family members, longitudinal studies and studies involving repeated measurements on the study subjects. The models for correlated binary data often focus on the dependence of marginal response probabilities on covariates and experimental conditions, although, there is a lack of study based on joint probability distributions that contain convenient estimation of marginal means and correlations for correlated binary data. The models based on generalized estimating equations (GEE) provide very attractive and useful results but the estimates are often inefficient [3, 8]. The classical marginal models (such as GEE) are constrained with the selection of underlying correlation structure, which may or may not be functions of the marginal expectations [12]. As an alternative, the subject-specific models are proposed taking into consideration the random effects by allowing random effect

Communicated by Anton Abdulbasah Kamil.
Received: July 19, 2012; *Revised:* March 5, 2013.

terms in the linear predictor [2]. It has been demonstrated by Lee and Nelder [12] that conditional models are fundamental and the advantages of conditional models over marginal models are obvious because marginal predictions can be made from conditional models. At this backdrop, logistic representations has been suggested by Cox [5] which noted a probability distribution of this quadratic exponential form can be reparameterized in terms of marginal parameter of ready interpretation. This has been further discussed by Cox and Wermuth [6, 7]. Zhao and Prentice [18] provided a comprehensive estimation procedure for quadratic exponential quadratic form models and provided measures based on covariances [16]. The quadratic exponential form models have been employed by Hudson *et al.* [10] for analyzing familial aggregation of two disorders. In this paper a test procedure for association of order is proposed by extending the Tsiatis [15] procedure test based on quadratic exponential model which was defined by Bahadur [1].

2. Quadratic exponential form model

According to Bahadur [1], consider Y as a specified set of all points $y = (y_1, \ldots, y_n)$ with $y_i = 0$ or 1; with 2^n points. Let $P(y)$ be a given probability distribution on Y, such that $P(y) \geq 0$ for all y and $\sum_{(y \in Y)} P(y) = 1$.

Let

$$p_i = P(Y_i = 1), \quad i = 1, \ldots, n,$$

so

$$\mu_i = \mathrm{E}_P(Y_i) = p_i, \quad 0 < p_i < 1,\ i = 1, \ldots, n;$$

where E_P is expected value. Then define

$$\begin{aligned} Z_i &= \frac{(Y_i - p_i)}{\sqrt{p_i(1-p_i)}}, \quad i = 1, \ldots, n \\ \eta_{ij} &= \mathrm{E}_P(Z_i Z_j), \quad i < j \\ \eta_{ijk} &= \mathrm{E}_P(Z_i Z_j Z_k), \quad i < j < k \\ &\cdots \\ \eta_{12\ldots n} &= \mathrm{E}_P(Z_1 Z_2 \cdots Z_n). \end{aligned}$$

Let $P_{[1]}(y_1, \ldots, y_n)$ denote the joint probability distribution of Y_i 's when Y_i 's are independent identify distribution. i.e.,

$$P_{[1]}(y_1, \ldots, y_n) = \prod_{i=1}^{n} p_i^{y_i} (1-p_i)^{1-y_i}.$$

A representation of the distribution defined [1] as

$$P(y) = P_{[1]}(y).f(y) \quad \text{for every } y = (y_1, \ldots, y_n)$$

where

$$f(y) = 1 + \sum_{i<j} \eta_{ij} z_i z_j + \sum_{i<j<k} \eta_{ijk} z_i z_j z_k + \cdots + \eta_{12\ldots n} z_1 z_2 \cdots z_n.$$

Now, let $y_k^T = (y_{k1}, \ldots, y_{kn_k}), k = 1, \ldots, K$ be a sample of k-independent multivariate binary observation. The distribution of Y_k is defined by Zhao and Prentice [18]

$$Pr(y_k) = \Delta_k^{-1} \exp\left\{ y_k^T \theta_k + w_k^T \lambda_k + c_k(y_k) \right\}, \tag{2.1}$$

where

$$\begin{aligned} w_k^T &= (y_{k1}y_{k2}, y_{k1}y_{k3}, \ldots, y_{k2}y_{k3}, \ldots), \\ \theta_k^T &= (\theta_{k1}, \ldots, \theta_{kn_k}), \\ \lambda_k^T &= (\lambda_{k12}, \lambda_{k13}, \ldots, \lambda_{k23}, \ldots) \end{aligned}$$

are canonical parameters, and $\Delta_k = \Delta_k(\theta_k, \lambda_k)$ is a normalization constant defined by

$$\Delta_k = \sum \exp\left\{y_k^T \theta_k + w_k^T \lambda_k + c_k(y_k)\right\},$$

with summation over all 2^{n_k} possible values of Y_k. If $\lambda_k \equiv 0$ and $c_k(.) \equiv 0$, the elements of Y_k will be statistically independent.

By fixing the functions, $c_k(.)$, for example $c_k(.) \equiv 0$, and considering a fixed number of canonical parameters (θ_k, λ_k), $k = 1, \ldots, K$, parametric inference based on Equation (2.1) could proceed. Alternatively the response probabilities, $\mu_k = \mathrm{E}(Y_k)$, and pairwise correlations can be model by $\mu_k = \mu_k(\beta)$ and covariances $\sigma_k = (\sigma_{k12}, \sigma_{k13}, \ldots, \sigma_{k23}, \ldots) = \sigma_k(\beta, \alpha)$ in terms of parameter vectors β and α. Then can show that the transformation from (θ_k, λ_k) to (μ_k, σ_k) has Jacobian the covariance matrix for (Y_k^T, W_k^T). Although the parameters (θ_k, λ_k) are complicated functions of the corresponding (μ_k, σ_k) values, the score estimating equations for β and α turn out to have a particularly simple estimation procedures for mean and covariance parameters.

Zhao and Prentice [18] defined the score estimating equation for β and α from Equation (2.1), with specified the response means μ_k and covariances σ_k, written as

$$K^{-1/2} \sum_{k=1}^{K} D_k^T V_k^{-1} f_k = 0, \tag{2.2}$$

where

$$D_k = \begin{bmatrix} \partial\mu_k/\partial\beta & 0 \\ \partial\sigma_k/\partial\beta & \partial\sigma_k/\partial\alpha \end{bmatrix}, \quad V_k = \begin{bmatrix} \mathrm{var}(Y_k) & \mathrm{cov}(Y_k, S_k) \\ \mathrm{cov}(S_k, Y_k) & \mathrm{var}(S_k) \end{bmatrix}, \quad f_k = \begin{bmatrix} Y_k - \mu_k \\ S_k - \sigma_k \end{bmatrix},$$

and where

$$S_k^T = (S_{k12}, S_{k13}, \ldots, S_{k23}, \ldots), \quad S_{kij} = (Y_{ki} - \mu_{ki})(Y_{kj} - \mu_{kj})$$

is the vector of empirical pairwise covariances. And the corresponding Fisher information matrix defines as

$$W = K^{-1} \sum_{k=1}^{K} D_k^T V_k^{-1} D_k. \tag{2.3}$$

Also, expectation of W_k from (2.1) is written by

$$\eta_k^T = \mathrm{E}\left(W_k^T\right), \quad \eta_{kij} = \sigma_{kij} + \mu_{ki}\mu_{kj} \quad \text{that} \quad \sigma_{kij} = \mathrm{E}[(Y_{ki} - \mu_{ki})(Y_{kj} - \mu_{kj})].$$

Whenever $(V_k^{-1})_{12} \equiv 0$, or equivalently $V_{k12} \equiv 0$, and $D_{k21} = \partial\sigma_k/\partial\alpha$ is replaced by $D_{k21} \equiv 0$. These estimating equations will be unbiased for mean parameters in spite of whether $\mathrm{E}(S_k) = \sigma_k$.

Application of the chain rule to the log likelihood contribution

$$l_k = y_k^T \theta_k + w_k^T \lambda_k + c_k(y_k) - \log \Delta_k \tag{2.4}$$

gives

$$\begin{bmatrix} \partial l_k/\partial\beta \\ \partial l_k/\partial\alpha \end{bmatrix} = \tilde{D}_k^T \tilde{V}_k^{-1} \tilde{f}_k,$$

where

$$\tilde{D}_k = \begin{bmatrix} \partial\mu_k/\partial\beta & 0 \\ \partial\eta_k/\partial\beta & \partial\eta_k/\partial\alpha \end{bmatrix}, \quad \tilde{f}_k = \begin{bmatrix} Y_k - \mu_k \\ W_k - \eta_k \end{bmatrix}.$$

Since $\tilde{V}_k = \mathrm{E}(\tilde{f}_k \tilde{f}_k^T)$ the corresponding information matrix contribution is $\tilde{D}_k^T \tilde{V}_k^{-1} \tilde{D}_k$.

3. Quadratic exponential model for transition probabilities

Consider a single stationary process $Y_k = (Y_1, \ldots, Y_n)$ generated by a binary Markov chain taking values 0 and 1. The transition matrix defines by

$$P = \begin{bmatrix} 1 - p_{0t} & p_{0t} \\ 1 - p_{1t} & p_{1t} \end{bmatrix}$$

where, $p_{jt} = \Pr(Y_t = 1 | Y_{t-1} = j);\ j = 0, 1,\ t = 1, \ldots, T$.

For expanding Equation (2.1) for transition probabilities of first order Markov chain, let $n = 2$. Thus, the equations can be written as [18]

$$y_k^T = (y_{k1},\ y_{k2}), \quad \text{and} \quad Pr(y_k) = \Delta_k^{-1} \exp\left\{ y_k^T \theta_k + w_k^T \lambda_k + c_k(y_k) \right\},$$

where

$$w_k^T = (y_{k1} y_{k2}), \quad \theta_k^T = (\theta_{k1}, \theta_{k2}), \quad \lambda_k^T = (\lambda_{k12}).$$

Here, λ_{k12} is the association parameter.

In this section, a more natural way to represent the association parameter, in terms of the underling dependence between outcome variables as well as between the outcome and explanatory variables, is proposed. The proposed model for the expected value of the outcome variables, Y_{k1} and Y_{k2}, are shown in Equations (3.1) and (3.2).

Let $\mu_k = \mathrm{E}(Y_k) = (\mu_{k1}, \mu_{k2})^T$, $\mu_k = \mu_k(\beta, \alpha)$ and covariance matrix

$$\sigma_k = \begin{bmatrix} \sigma_{k11} & \sigma_{k12} \\ \sigma_{k21} & \sigma_{k22} \end{bmatrix}, \quad \sigma_k = \sigma_k(\beta,\ \alpha).$$

β and α are vector parameters $\beta = (\beta_0, \beta_1, \ldots, \beta_p)$, $\alpha = (\alpha_{12})$ and

$$\mu_{k1} = P(Y_{k1} = 1 | x_k) = \frac{\exp(\beta_1 x_k)}{1 + \exp(\beta_1 x_k)}, \tag{3.1}$$

$$\mu_{k2} = P(Y_{k2} = 1 | y_{k1},\ x_k) = \frac{\exp(\beta_2 x_k + \alpha_{12} y_{k1})}{1 + \exp(\beta_2 x_k + \alpha_{12} y_{k1})}, \tag{3.2}$$

where, the vector x_k contains covariates and for the kth individual is equal to $x_k = (1, x_{k1}, \ldots, x_{kp})$ and $\beta_j = (\beta_{j0}, \beta_{j1}, \ldots, \beta_{jp});\ j = 0, 1$.

Let us consider

$$\theta_{k1} = \mathrm{logit}(\mu_{k1}) = \log \frac{\mu_{k1}}{1 - \mu_{k1}} = \beta_1 x_k,$$

$$\theta_{k2} = \mathrm{logit}(\mu_{k2}) = \log \frac{\mu_{k2}}{1 - \mu_{k2}} = \beta_2 x_k + \alpha_{12} y_{k1},$$

and

$$\lambda_{k12} = \log\left(\frac{\mu_{k2}/(1-\mu_{k2})}{\mu_{k1}/(1-\mu_{k1})}\right) = (\beta_2 x_k + \alpha_{12} y_{k1}) - \beta_1 x_k = (\beta_2 - \beta_1) x_k + \alpha_{12} y_{k1}. \tag{3.3}$$

Hence, the proposed association parameter, λ_{k12}, appears to depend on the values of $(\beta_2 - \beta_1)$ and α_{12}. In other words, if $(\beta_2 - \beta_1) = 0$ and $\alpha_{12} = 0$, then can say that there is no dependence.

The log-likelihood function is defined

$$\begin{aligned} l &= \sum_{k=1}^{K} l_k = \sum_{k=1}^{K} \left\{ y_k^T \theta_k + w_k^T \lambda_k + c_k(y_k) - \log \Delta_k \right\} \\ &= \sum_{k=1}^{K} \left\{ y_{k1} \operatorname{logit}(\mu_{k1}) + y_{k2} \operatorname{logit}(\mu_{k2}) + y_{k1} y_{k2} \log\left(\frac{\mu_{k2}/(1-\mu_{k2})}{\mu_{k1}/(1-\mu_{k1})}\right) - \log \Delta_k \right\} \\ &= \sum_{k=1}^{K} \left\{ y_{k1}(\beta_1 x_k) + y_{k2}(\beta_2 x_k + \alpha_{12} y_{k1}) + y_{k1} y_{k2}(\beta_2 x_k + \alpha_{12} y_{k1} - \beta_1 x_k) - \log \Delta_k \right\}. \end{aligned}$$

Here, $c_k(y_k) \equiv 0$ and $\Delta_k = \sum \exp\{y_k^T \theta_k + w_k^T \lambda_k + c_k(y_k)\}$ is a normalization constant.

For estimating means and covariance parameters let $V_{k11} = \operatorname{var}(Y_k)$, $V_{k12} = V_{k21}^T$ and V_{k22} denote the block submatrices of V_k. In this paper is considered: Independence among the elements of Y_k implies $V_{k12} \equiv 0$, and a diagonal V_{k22} with entries $\operatorname{var}(S_{kij}) = \sigma_{kii}\sigma_{kjj}$, where $\sigma_{kii} = \mu_k i(1-\mu_k i)$ [13].

The first derivative can be written as

$$\begin{bmatrix} \partial l_k/\partial \beta \\ \partial l_k/\partial \alpha \end{bmatrix} = \tilde{D}_k^T \tilde{V}_k^{-1} \tilde{f}_k,$$

where

$$\tilde{D}_k = \begin{bmatrix} \partial \mu_k/\partial \beta & \partial \mu_k/\partial \alpha \\ \partial \eta_k/\partial \beta & \partial \eta_k/\partial \alpha \end{bmatrix}, \quad \tilde{f}_k = \begin{bmatrix} Y_k - \mu_k \\ W_k - \eta_k \end{bmatrix}.$$

Since $\tilde{V}_k = \mathrm{E}(\tilde{f}_k \tilde{f}_k^T)$ the corresponding information matrix contribution can be written as $\tilde{D}_k^T \tilde{V}_k^{-1} \tilde{D}_k$ and parameters can be estimated by solving

$$\begin{bmatrix} \partial l_k/\partial \beta \\ \partial l_k/\partial \alpha \end{bmatrix} = \tilde{D}_k^T \tilde{V}_k^{-1} \tilde{f}_k = 0.$$

The null hypothesis for testing order of Markov chain model in first order is $H_0 : \alpha = 0$; and under true null hypothesis, the log-likelihood function is written as

$$l = \sum_{k=1}^{K} \{ y_{k1}(\beta_1 x_k) + y_{k2}(\beta_2 x_k) + y_{k1} y_{k2}(\beta_2 x_k - \beta_1 x_k) - \log \Delta_k \}.$$

By using efficient score test the test statistic is defined by

$$T = Z' V^{-1} Z,$$

where Z is the $(\partial l/\partial \alpha)$, and the matrix V is

$$V = A - BC^{-1}B',$$

where

$$A = \frac{-\partial^2 l}{\partial \alpha \partial \alpha}, \quad B = \frac{-\partial^2 l}{\partial \alpha \partial \beta}, \quad C = \frac{-\partial^2 l}{\partial \beta \partial \beta}.$$

All above terms were evaluated at $\alpha = 0$ and $\beta = \hat{\beta}$. Where, $\hat{\beta}$ is the maximum likelihood estimate of the parameters when H_0 is true [14].

The Wald test statistic is written as

$$W = K^{-1}\left(\tilde{D}_k^T \tilde{V}_k^{-1} \tilde{f}_k\right)^T \left(\tilde{D}_k^T \tilde{V}_k^{-1} \tilde{D}_k\right)^{-1} \left(\tilde{D}_k^T \tilde{V}_k^{-1} \tilde{f}_k\right).$$

where under true null hypothesis and $\beta = \hat{\beta}$ is distributed asymptotically chi-square with $2p$ degree of freedom.

For testing null hypothesizes $H_0 : \beta_1 = \beta_2 = \beta$ and $\alpha_{12} = 0$ likelihood ratio test statistics can be used

$$LRT = -2(\ln L_0 - \ln L_1),$$

where L_0 is likelihood function under true null hypothesis and L_1 is likelihood function based on alternative hypothesis, which is distributed asymptotically chi-square with $p+1$ degree of freedom. Also, following test statistic with asymptotic chi square distribution with p degree of freedom,

$$\hat{X}^2 = \left(\hat{\beta}_1 - \hat{\beta}_2\right)^T \left(\operatorname{var}\left(\hat{\beta}_1 - \hat{\beta}_2\right)\right)^{-1} \left(\hat{\beta}_1 - \hat{\beta}_2\right).$$

can be used.

It is evident from Equation (3.3) that the odds ratio can be expressed as

$$OR_k = \exp(\lambda_{k12}) = \frac{\mu_{k2}/(1-\mu_{k2})}{\mu_{k1}/(1-\mu_{k1})} = \exp\{(\beta_2 - \beta_1)x_k + \alpha_{k12} y_{k1}\}. \tag{3.4}$$

From the above expression, it can be stated that if β_1 and β_2 as well as α_{k12} are zero then the odds ratio, OR_k, is 1 indicating no dependence between the outcome variables. In Equation (3.3), if $\alpha_{k12} = 0$, then association will depend on the difference between β_1 and β_2. In other words, if there is significant difference between β_1 and β_2, it will still indicate the significant association between outcomes, Y_1 and Y_2.

4. Application

To demonstrate an application of the proposed method, we have employed the Health and Retirement Study (HRS) data [9]. The data were collected from 1992 to 2006 by the RAND Centre for 30,405 people in 8 waves, for considering repeated measures. In this case, only individuals who attended to the program in 1992 and the follow up until 2006 have been considered. The study takes into account the factors affecting depression during the elderly. Depression (0 for no depression and 1 for depression) is considered as dependent variable, and age (in year), gender (0 for male and 1 for female), body mass index (BMI), and drink (0 for no drink and 1 for drink) as covariate variables. Some of variables contained missing values because reference person did not respond to the all waves. Thus, these individuals are dropped completely from studying if there were missing value in the covariate variables, but were kept if the value of dependent variable (depression) was missing. For estimating the parameters of model, S-Plus program which has been developed by Chowdhury *et al.* [4], is modified and used.

Table 1 shows the results of the test for the null hypothesis $H_0 : \alpha = 0$ evaluated at the maximum likelihood estimates of β_1 and β_2. This indicates apparently that there is no relationship between the outcome variables but as the dependence between the outcome variables depend on α, β_1 and β_2, we need to comment on the basis of tests for all these parameters. It is shown by Islam *et al.* [11] that the dependence between the outcome variables may depend on their association with the explanatory variables as well. In addition, results in Tables 2 and 3 are based on the null hypotheses $H_0 : \beta_1 = \beta_2$ and $H_0 : \beta_1 = \beta_2,\ \alpha_{12} = 0$, respectively. In Table 2, we have used the LRT. In Table 3, the LRT is employed for testing $H_0 : \beta_1 = \beta_2,\ \alpha_{12} = 0$, and the efficient score and the Wald tests are performed for $H_0 : \alpha_{12} = 0$ evaluated at $\beta_1 = \beta_2 = \hat{\beta}$. Both the extended Tsiatis and the Wald test statistics indicate that the models are fitted well. In Table 2 we observe that the null hypothesis of equality of parameters of the two conditional models may be rejected indicating significant differences in the parameters. Hence, following Equation (3.4), we may conclude that there is evidence of dependence between the two outcome variables. It appears that age and drink are negatively associated with the previous outcome while sex is positively associated (females have higher risk). For the subsequent outcome, the role of age and drink remained similar in terms of direction but sex does not show and significant association. The results in Table 3 confirm the results presented in Tables 1 and 2 by all the test statistics.

Table 1. Estimates of parameters of covariate for the quadratic exponential form and testing for the association parameter

$H_0 : \alpha = 0$				
Covariates	Estimated value		s.e.	p_value
$\hat{\beta}_1$	Constant	4.964	0.306	0.000
	Age	-0.098	0.004	0.000
	Sex	0.226	0.060	0.0002
	BMI	-0.008	0.005	0.117
	Drink	-0.136	0.058	0.019
$\hat{\beta}_2$	Constant	-1.029	0.329	0.002
	Age	-0.007	0.004	0.114
	Sex	0.003	0.063	0.230
	BMI	-0.355	0.062	0.000
	Drink	-0.136	0.058	0.019
$\hat{\alpha}$	α_{12}	0.039	0.066	0.553
Tsiatis Test		0.349	p_value= 0.999	
Wald Test		4.80E-06	p_value= 1	

Table 2. Estimates of parameters of covariate for the quadratic exponential form and testing for the equality of parameters for the two outcome variables

$H_0: \beta_1 = \beta_2$				
Covariates	Estimated value		s.e.	p_value
$\hat{\beta}_1$	Constant	4.981	0.307	0.000
	Age	-0.098	0.004	0.000
	Sex	0.226	0.060	0.0002
	BMI	-0.008	0.005	0.118
	Drink	-0.138	0.058	0.018
$\hat{\beta}_2$	Constant	-0.977	0.313	0.002
	Age	-0.007	0.004	0.066
	Sex	0.005	0.063	0.939
	BMI	0.007	0.006	0.234
	Drink	-0.355	0.062	0.000
LRT Test		408.141	p_value=0.000	

Table 3. Estimates of parameters of covariate for the quadratic exponential form and testing for the equality of parameters for two outcome variables and the parameter for association

$H_0: \beta_1 = \beta_2,\ \alpha = 0$				
Covariates	Estimated value		s.e.	p_value
$\hat{\beta}_1$	Constant	4.964	0.306	0.000
	Age	-0.098	0.004	0.000
	Sex	0.226	0.060	0.0002
	BMI	-0.008	0.005	0.117
	Drink	-0.136	0.058	0.019
$\hat{\beta}_2$	Constant	-1.029	0.329	0.002
	Age	-0.007	0.004	0.114
	Sex	0.003	0.063	0.230
	BMI	-0.355	0.062	0.000
	Drink	-0.136	0.058	0.019
$\hat{\alpha}$	α_{12}	0.039	0.066	0.553
LRT Test		401.593	p_value= 0.000	
Tsiatis Test		7.71E-013	p_value= 0.999	
Wald Test		0.025	p_value= 0.874	

5. Conclusion

In analyzing repeated measures data, we have to encounter the dependence in outcome variables at different times. This has been addressed mainly by the Generalized Estimating Equations which are marginal models and assume the dependence in outcomes. The quadratic exponential form employs different covariance structures and this paper extends the form for transitions arising from repeated measures data. The extended form takes account of previous outcome as a covariate and thus reveals the order of dependence. This procedure can be extended further using higher order transitions and the test for higher order dependence can also be developed. The proposed extension of the quadratic exponential form model is illustrated with a set of depression data from elderly population and reveals first order dependence in the depression outcome.

Acknowledgement. We are grateful to the anonymous reviewers of the paper for their useful comments and to Rafiqul Islam Chowdhury, Western University, Canada for permission of using the program ”kernopt markov.gen” for estimating parameters of the models proposed in this study. We would like to thank the Health and Retirement Study (HRS) center for permission to use RAND data for application of the models.

References

[1] R. R. Bahadur, A representation of the joint distribution of responses to n dichotomous items, in *Studies in Item Analysis and Prediction*, 158–168, Stanford Univ. Press, Stanford, CA, 1961.
[2] N. E. Breslow and D. G. Clayton, Approximate inference in generalized linear mixed models, *J. Amer. Statist. Assoc*. **88** (1993), 9–25.
[3] B. Chen, G. Y. Yi and R. J. Cook, Likelihood analysis of joint marginal and conditional models for longitudinal categorical data, *Canad. J. Statist.* **37** (2009), no. 2, 182–205.
[4] R. I. Chowdhury, M. A. Islam, M. Shah and N. Al-Enezi, A computer program to estimate the parameters of covariate dependence higher order Markov model, *Computer Methods and Program in Biomedicine*, **77** (2005), 175–181.
[5] D. R. Cox, The analysis of multivariate binary data, *Appl. Statist.* **21** (1972), 113–120.
[6] D. R. Cox and N. Wermuth, A note on the quadratic exponential binary distribution, *Biometrika* **81** (1994), no. 2, 403–408.
[7] D. R. Cox and N. Wermuth, On some models for multivariate binary variables parallel in complexity with the multivariate Gaussian distribution, *Biometrika* **89** (2002), no. 2, 462–469.
[8] G. M. Fitzmaurice, N. M. Laird and A. G. Rotnitzky, Regression models for discrete longitudinal responses, *Statist. Sci.* **8** (1993), no. 3, 284–309.
[9] Health and Retirement Study (HRS), (Wave (1-8)/Year (1992-2006)). Public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG09740). Ann Arbor, MI, 2009.
[10] J. I. Hudson, N. M. Laird and R. A. Betensky, Multivariate logistic regression for familial aggregation of two disorders, I. Development of models and method, *American Journal of Epidemiology.* **135** (2001), 500–505.
[11] M. A. Islam, R. I. Chowdhury and L. Briollais, A bivariate binary model for testing dependence in outcomes, *Bull. Malays. Math. Sci. Soc. (2)* **35** (2012), no. 4, 845–858.
[12] Y. Lee and J. A. Nelder, Conditional and marginal models: another view, *Statist. Sci.* **19** (2004), no. 2, 219–238.
[13] R. L. Prentice, Correlated binary regression with covariates specific to each binary observation, *Biometrics* **44** (1988), no. 4, 1033–1048.
[14] C. R. Rao, *Linear Statistical Inference and Its Applications*, second edition, John Wiley & Sons, New York, 1973.
[15] A. A. Tsiatis, A note on a goodness-of-fit test for the logistic regression model, *Biometrika* **67** (1980), 250–251.
[16] G. Y. Yi, W. He and H. Liang, Analysis of correlated binary data under partially linear single-index logistic models, *J. Multivariate Anal.* **100** (2009), no. 2, 278–290.
[17] F. Yu, E. H. Morgenstern and T. R. Berlin, Use of a Markov transition model to analyse longitudinal low-back pain data, *Stat. Methods Med. Res.* **12** (2003), no. 4, 321–331.
[18] L. P. Zhao and R. L. Prentice, Correlated binary regression using a quadratic exponential model, *Biometrika* **77** (1990), no. 3, 642–648.