

# Bayesian Test of Homogeneity in Transition Model for Longitudinal Ordinal Response Data with Missing Values

S. Noorian<sup>1</sup> and M. Ganjali<sup>2</sup> \*

<sup>1</sup> Department of Statistics, Faculty of Mathematical Sciences, Tarbiat Modares University, Tehran, Iran

<sup>2</sup> Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran

sajad.noorian@gmail.com, m-ganjali@sbu.ac.ir

Received 13 January 2012

Accepted 8 May 2013

## Abstract

For analyzing longitudinal ordinal response data, many methods are available to take into account the correlations between responses of the same individual where different methods use different approaches. One of the methods to consider this correlation is the transition (Markov) model. In this paper, we present a Bayesian test of homogeneity of a transition model for analyzing longitudinal ordinal response data. In other words we test whether the transition probabilities  $\pi_{mabt}$ 's, which are probabilities of moving to state  $b$  from state  $a$  of  $m^{\text{th}}$  individual at time  $t$ , are the same for any given time  $t$ , i.e.,  $\pi_{mabt} = \pi_{mab}$ . A cumulative logistic regression model and the Bayesian method, using MCMC, are implemented for testing the null hypothesis of homogeneity. We also define what a specific homogeneous covariate is. Our approach is applied on three-period longitudinal Fluvoxamine (a treatment for deregulation of serotonin in brain) data.

*Keywords:* Bayesian Analysis; Fluvoxamine Data; Homogeneity; Logistic Regression Model; Non-homogeneity; Transition (Markov) Models.

*AMS (2000) Subject Classifications:* 62H17, 62F15

## 1. Introduction

In many areas of medical research, longitudinal ordinal response data are often collected where the response of interest for each individual is recorded on an ordinal scale and is observed on several regular or irregular time points. Since in longitudinal study there is a sequence of responses measured for the same individual, we should take into account the correlation between these responses.

For handling such correlation between responses, some approach are proposed such as marginal modelling, random effects modelling and transition (Markov) modelling. Marginal modelling allows for inferences about parameters averaged over the whole population or trend over time (Ten Have et al., 1996; Kim, 1995; Liang et al., 1992; Molenberghs and Lesaffre, 1994). Random effects modelling deliberately provides inferences about variability between respondents. In this approach, individual behavior is often of scientific interest (Harville and

---

\*corresponding author; address: Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran.  
Email: m-ganjali@sbu.ac.ir

Mee, 1984; Verbeke and Lesaffre, 1996; Tutz and Hennevogl, 1996; Verbeke and Molenberghs, 1997; Diggle et al., 2002; Tutz, 2005). However, both of these approaches are generally appropriate for long sequences of measurements. Another appropriate approach to investigate the reasons for the change of the responses is the use of Markov (transition) models (Garber, 1989; Francom et al., 1989; Chung et al., 2005; Rezaee and Ganjali, 2009; Rezaee et al., 2009) where we can consider the effect of previous response on current response.

In longitudinal studies, it is not unusual in practice that complete follow-up data were not available for all subjects. In other words, some subject's responses are missing due to dropout, death or any reason outside the control of investigator. Any subject's response so affected is often called a missing response.

Using the terminology of Diggle and Kenward (1994), the dropout mechanism is called completely random dropout if the probability of dropping out is independent of both responses (observed and unobserved). It is called random dropout if the above probability depends on observed response. Finally a dropout mechanism that depends on unobserved responses is called non-random dropout mechanism or non-ignorable. For a wide discussion about the dropout mechanisms you can see Little and Rubin (2002).

In analysing longitudinal data using Markov models some classical tests of homogeneity of transition probabilities are used, (Weißbach and Dette, 2007; Schweinberger, 2012). In this paper we propose a Bayesian test of homogeneity of transition probabilities in longitudinal ordinal response data with random dropout. We use the cumulative logit regression model to describe the ordinal structure of the data.

By setting some prior distributions on coefficients and cut-points, the Bayesian analysis of the data are implemented and the Bayesian estimates of parameters are presented. The proposed model is applied to Fluvoxamine data (a treatment for deregulation of serotonin in the brain).

The structure of this paper is as follows: In Section 2, first the Fluvoxamine data are described briefly. Then, the transition model and the Bayesian computation are presented. Also to test the homogeneity assumption of the transition model, Bayes factor and some other criteria are presented. The proposed model and test are then applied to the Fluvoxamine data. Finally in Section 3 conclusion is given.

## 2. Fluvoxamine Data Study and Transition Model

In this section the Fluvoxamine data are introduced. Then we define the transition model, the homogeneous model, the non-homogeneous model and the homogeneity of a specific covariate (or a specific set of covariates). The cumulative logistic regression model for ordinal response is also defined. By setting up some priors on cut-points and regression coefficients, the Bayesian test of homogeneity is presented.

### 2.1. Fluvoxamine Data

The data are collected from a study in which the Fluvoxamine drug was used to treat psychiatric symptoms. In initial visits, some covariates such as age, gender, initial severity (on an ordinal scale with seven levels), and duration of actual mental illness were collected from each individual. After recruitment, the individuals were asked to return for four visits at weeks 2, 4, 8 and 12 and also the Fluvoxamine drug is prescribed to them. On each follow-up visit, the side effect is measured on an ordinal scale with 4 possible levels as 1 (none), 2 (not interfering significantly with functionality of patient), 3 (significant interference with functionality of patient) and 4 (side effect surpass therapeutic effect). Since level 4 of side effect occurred only 2 percent of time, in this paper level 4 is combined with level 3 for simplicity in computations. Following Hines and Hines (2005) and for having the random dropout mechanism, we use the data on weeks 2, 4 and 12 and denote these responses by  $Y_1$ ,  $Y_2$  and  $Y_3$ , respectively. Hines and Hines (2005) and Rezaee and Ganjali (2009) and our Bayesian results show that the coefficients of the covariate 'gender' are not significant at the 5% level in any time and transition from any state. So we don't consider this covariate in the data analysis. Rezaee et al. (2009) show that for analysing

this data set one can use the first order transition model and the use of second order is not necessary. For more details on Fluvoxamine data, see Molenberghs and Lesaffre (1994).

## 2.2. Transition Models

Suppose that  $\{Y_{m1}, Y_{m2}, \dots\}$  be a set of random variables indexed by time associated with  $m^{\text{th}}$  subject where each element of this set can take finite values in  $\mathcal{Y} = \{1, 2, \dots, J\}$ . In a transition (Markov) model, a random variable  $Y_{mt}$  is described as a function of previous outcomes  $(Y_{m1}, \dots, Y_{m,t-1})$ . In other words one can model the  $Y_{mt}$  based on  $(Y_{m1}, \dots, Y_{m,t-1})$ . The order of a transition model is the number of previous measurements that is still considered to influence the current one (Molenberghs and Verbeke, 2005, page 236).

For the Fluvoxamine data, suppose that  $\{Y_{m1}, Y_{m2}, \dots\}$  forms a first-order Markov chain, [i.e., the conditional probability of  $Y_{mt}$  given  $(Y_{m1}, \dots, Y_{m,t-1})$  is equal to the conditional probability of  $Y_{mt}$  given  $Y_{m,t-1}$ ]. Here  $Y_{mt}$  represents the state of subject  $m$  at time  $t$ . The transition probability matrix  $\Pi_{mt}$  for  $m^{\text{th}}$  individual is

$$\Pi_{mt} = \begin{bmatrix} \pi_{m11t} & \cdots & \pi_{m1Jt} \\ \vdots & \ddots & \vdots \\ \pi_{mJ1t} & \cdots & \pi_{mJJt} \end{bmatrix}$$

where the  $(a, b)^{\text{th}}$  element of  $\Pi_{mt}$  is  $\pi_{mabt} = P(Y_{mt} = b | Y_{m,t-1} = a)$  and represents the  $m^{\text{th}}$  subject's probability of making transition from state  $a$  at time  $t - 1$  to state  $b$  at time  $t$ , that called transition probability from state  $a$  to state  $b$  at time  $t$ .

If these transition probabilities are independent of time, i.e.,  $\Pi_{mt} = \Pi_m$  and hence  $\pi_{mabt} = \pi_{mab}$ , for all  $t = 1, 2, \dots, T$ , then the Markov chain is called time homogeneous Markov chain. If the probabilities depend on time then we have a non-homogeneous Markov chain. When  $\pi_{mabt}$ 's are modelled based on some cut-points and some covariates, the homogeneity of a specific covariate (or a specific set of covariates) is occur when the cause of non-homogeneity is just because of different cut-points or other covariates effects and not that specific covariate (or a specific set of covariates). In other words, that specific covariate (or a specific set of covariates), has the same effect(s) and also has the same coefficient(s) over the time.

The cumulative logit model for ordinal longitudinal responses is specified in terms of cumulative transition probabilities, i.e.

$$C_{m0b1} = P(Y_{m1} \leq b) = \sum_{k=1}^b \pi_{m0k1},$$

$$C_{mabt} = P(Y_{mt} \leq b | Y_{m,t-1} = a) = \sum_{k=1}^b \pi_{makt}, \text{ for } t > 1$$

where

$$\text{logit}(C_{m0b1}) = \log\left(\frac{C_{m0b1}}{1 - C_{m0b1}}\right) = \log\left(\frac{P(Y_{m1} \leq b)}{P(Y_{m1} > b)}\right),$$

$$\text{logit}(C_{mabt}) = \log\left(\frac{C_{mabt}}{1 - C_{mabt}}\right) = \log\left(\frac{P(Y_{mt} \leq b | Y_{m,t-1} = a)}{P(Y_{mt} > b | Y_{m,t-1} = a)}\right), \text{ for } t > 1$$

for  $m = 1, 2, \dots, M$ ;  $a = 1, 2, \dots, J$ ;  $b = 1, 2, \dots, J - 1$  and  $t = 1, 2, \dots, T$ . In ordinal data, we have a term called latent variable,  $Z_{mat}$ , where given  $Y_{m,t-1} = a$ , we will have  $Y_{mt} = j$  if  $\alpha_{at,j-1} < Z_{mat} < \alpha_{at,j}$ . The index  $a$  in  $Z_{mat}$  is used to show the first order Markov structure of data.

We write the cumulative logit model as

$$\text{logit}(C_{mabt}) = \alpha_{atb} - X_{mt} \beta_t^a,$$

In other words, If we have an ordinal variable with  $J$  levels, then the form of the transition model for  $T$  response variables, (in the Fluvoxamine data,  $J = 3$  and  $T = 3$ ), is:

$$\begin{aligned} \text{logit}[P(Y_{m1} \leq b; \alpha_{1b}, \beta_1)] &= \alpha_{1b} - X_{m1}\beta_1 = \alpha_{1b} - \sum_{k=1}^K \beta_{1k}X_{m1k}, & \text{for } t = 1, \\ \text{logit}[P(Y_{mt} \leq b|Y_{m,t-1} = a, \alpha_{atb}, \beta_t^a)] &= \alpha_{atb} - X_{mt}\beta_t^a = \alpha_{atb} - \sum_{k=1}^K \beta_{tk}^a X_{mtk}, & \text{for } t > 1, \end{aligned} \tag{1}$$

where  $\alpha_1 := (\alpha_{11}, \dots, \alpha_{1,J-1})$  and  $\alpha_{at} := (\alpha_{at1}, \dots, \alpha_{at,J-1})$  are cut-point parameters in which  $\alpha_{11} \leq \alpha_{12} \leq \dots \leq \alpha_{1,J-1}$  and for any  $a$  and  $t$ ,  $\alpha_{at1} \leq \alpha_{at2} \leq \dots \leq \alpha_{at,J-1}$ , respectively,  $X_{mt} := (X_{mt1}, \dots, X_{mtK})$  is a  $1 \times K$  vector of covariates for the  $m^{\text{th}}$  individual, the parameters  $\beta_1 := (\beta_{11}, \dots, \beta_{1K})'$  and  $\beta_t^a = (\beta_{t1}^a, \dots, \beta_{tK}^a)'$  are  $K \times 1$  vectors of regression parameters. In model (1) if  $\beta_t^a = \beta^a$ ,  $\alpha_{atb} = \alpha_{ab}$  and covariates be time-invariant, then we have a homogeneous model.

Using the transition model the likelihood function is given in Rezaee et al. (2009). With considering the presence of dropout we can write the likelihood as below;

$$L = \prod_{m=1}^M P(Y_{m1} = y_{m1}) \times \prod_{m=1}^{M_1} P(Y_{m2} = y_{m2}|Y_{m1} = y_{m1}) \times \dots \times \prod_{m=1}^{M_{T-1}} P(Y_{mT} = y_{mT}|Y_{m,T-1} = y_{m,T-1}),$$

where  $M$  is the number of all individuals,  $M_1$  is the number of subjects for whom both  $Y_{i1}$  and  $Y_{i2}$  are observed,  $M_2$  is the number of subjects for whom all  $Y_{i1}$ ,  $Y_{i2}$  and  $Y_{i3}$  are observed and so on.

### 2.3. Bayesian Analysis

Due to the ordering property of the cut-points parameters the multivariate normal distributions may not be appropriate priors for these parameters, Johnson and Albert (1999). Therefore, the independent truncated normal distributions are used for each element of the cut-point parameter vector. For implementation, the non-informative proper priors are used. The first order Markov structure is set on  $\beta_t^a$  by

$$\begin{aligned} \beta_1 | \Sigma &\sim N(\mathbf{0}, \Sigma), & \text{for } t = 1, \\ \beta_t^a | \beta_{t-1}^a, \Sigma &\sim N(\beta_{t-1}^a, \Sigma), & \text{for } t > 2, \end{aligned}$$

where  $\Sigma$  is a  $K$ -dimensional diagonal covariance matrix defined as  $\Sigma = \sigma_\beta^2 \mathbf{I}$  and  $\sigma_\beta^2 = (\sigma_{\beta_{i1}^a}^2, \dots, \sigma_{\beta_{iK}^a}^2)'$  is a known vector. The following prior distributions are set for cut point parameters:

$$\begin{aligned} \alpha_{11} &\sim TN(\mu, \sigma^2), & -\infty \leq \alpha_{11} \leq \alpha_{12}, \\ \alpha_{1b} &\sim TN(\mu, \sigma^2), & \alpha_{1,b-1} \leq \alpha_{1,b} \leq \alpha_{1,b+1} \\ \alpha_{1,J-1} &\sim TN(\mu, \sigma^2), & \alpha_{1,J-2} \leq \alpha_{1,J-1} \leq +\infty, \\ \alpha_{at1} &\sim TN(\mu, \sigma^2), & -\infty \leq \alpha_{at1} \leq \alpha_{at2} \\ \alpha_{atb} &\sim TN(\mu, \sigma^2), & \alpha_{at,b-1} \leq \alpha_{at,b} \leq \alpha_{at,b+1} \\ \alpha_{at,J-1} &\sim TN(\mu, \sigma^2), & \alpha_{at,J-2} \leq \alpha_{at,J-1} \leq +\infty, \end{aligned}$$

where  $b = 2, 3, \dots, J - 2$ ,  $t = 2, 3, \dots, T$  and also  $TN(\mu, \sigma^2)$  is a truncated normal distribution with parameters  $\mu$  and  $\sigma^2$ .

Given the transition data on  $M$  individuals, the joint posterior distribution is proportional to

$$\begin{aligned} &\left( \prod_{m=1}^M P(Y_{m1} = y_{m1}) \times \prod_{m=1}^{M_1} P(Y_{m2} = y_{m2}|Y_{m1} = y_{m1}) \times \dots \times \prod_{m=1}^{M_{T-1}} P(Y_{mT} = y_{mT}|Y_{m,T-1} = y_{m,T-1}) \right) \times \\ &\left( \prod_{t=1}^T \pi(\beta_t^a | \beta_{t-1}^a, \Sigma) \right) \times \pi(\alpha), \end{aligned}$$

where  $\pi(\alpha)$  is the prior distribution for  $\alpha$  which is the vector including all cut-points parameters.

We set independent truncated normal distributions,  $N(0, 10000)$ , for cut point parameters  $\alpha_{1b}$  and  $\alpha_{atb}$ . For the regression coefficients  $\beta_1$  and  $\beta_t^a$  we use  $N(0, 10000)$  and  $N(\beta_{t-1}^a, 10000)$  priors, respectively. We use Metropolis-Hastings algorithm to generate samples from posterior distribution. We use prior distributions with large variances ( $\sigma^2 = 10000$  and  $\Sigma$  has diagonal components equal to 10000) to have low-informative priors. We run the program in the WinBUGS 1.4.3 for 100000 iterations with a burn-in time 70000 and the thin 10.

In transition modelling, because of conditioning on the previous responses, the number of cut-points parameters to estimate may be reduced. For example, conditioning on  $Y_1 = 1$ , there is no observation of  $Y_2 = 3$ . In this case, one of the transition probabilities is estimated zero. Therefore one of the cut-points parameters should be omitted, thus we can use a binary logit model instead of a cumulative logit model. For Fluvoxamine data in models for  $Y_2|Y_1 = 1$  and  $Y_3|Y_2 = 3$  we have only one cut-point where they are  $\alpha_{121}$  and  $\alpha_{332}$ . In other states and times, two cut-points exist.

For using the model, given in Eq. (1), to analyze the Fluvoxamine data we don't consider all covariates, because some of them aren't statistically significant. With  $K$  covariates, and no interactions, the number of possible models that may be used including these covariates is  $2^K$ . To determine the significant covariates coefficients, we use the Occam's window approximation (Madigan and Raftery, 1994) which starts with calculating the posterior probabilities of all models. Then this method identifies the 'best' model  $M_b$ , that is, the one with the highest posterior probability. We will use this method in transitions from each state and also in any time to determine the best parsimonious model.

After conducting the Occam's window approximation, it is shown that covariate 'severity' is not significant in all states and all times. But covariate 'age' is significant in the response at week 2 (model for  $Y_1$ ) and in the follow-up responses of week 4 when the previous response is 2 (models for  $Y_2|Y_1 = 2$ ). Covariate 'duration' is also significant in the response at week 2 (model for  $Y_1$ ) and in the follow-up responses when the previous response is 2, (i.e., models for  $Y_2|Y_1 = 2$  and  $Y_3|Y_2 = 2$ ). We call this parsimonious model as Model (I). In fact in this parsimonious model we have these five effect coefficients parameters:  $\beta_{11}$  and  $\beta_{21}^2$  for covariate 'age' in models for  $Y_1$  and  $Y_2|Y_1 = 2$ ;  $\beta_{12}$ ,  $\beta_{22}^2$  and  $\beta_{32}^2$  for covariate 'duration' in models for  $Y_1$ ,  $Y_2|Y_1 = 2$  and  $Y_3|Y_2 = 2$  (see the results given in Table 1). To obtain these results for cut point parameters  $\alpha_{1b}$  and  $\alpha_{atb}$ , the independent truncated normal distributions,  $N(0, 10000)$  are considered. For regression coefficients we use the following priors;

$$\begin{aligned} \beta_{11} , \beta_{12} &\sim N(0, 10000), \\ \beta_{21}^2 &\sim N(\beta_{11}, 10000), \\ \beta_{22}^2 &\sim N(\beta_{12}, 10000), \\ \beta_{32}^2 &\sim N(\beta_{21}^2, 10000). \end{aligned}$$

We also present two other versions of model (I) for Fluvoxamine data. One of them is Model (II) which is the time-invariant effects version of Model (I) for all covariates (partial homogeneous Markov model), i.e.,  $\beta_{2k}^a = \beta_{3k}^a$  where  $k = 2$  and  $a = 2$  since the second covariate (duration) is significant when the previous response is 2. Another model is the homogeneous Markov model that is model (I) which is restricted to have  $\alpha_{a2b} = \alpha_{a3b}$  and  $\beta_{22}^2 = \beta_{32}^2$ , (Model (III)). With these constraints we have  $\pi_{mab2} = \pi_{mab3}$  and hence the homogeneity property of Markov model is attained. In Model (II), for cut point parameters  $\alpha_{1b}$  and  $\alpha_{atb}$ , the independent truncated normal distributions,  $N(0, 10000)$  are considered. For regression coefficients we use the following priors;

$$\begin{aligned} \beta_{11} , \beta_{12} &\sim N(0, 10000), \\ \beta_{21}^2 &\sim N(\beta_{11}, 10000), \\ \beta_{22}^2 &\sim N(\beta_{12}, 10000). \end{aligned}$$

In Model (III), we suppose that  $\alpha_{a2b} = \alpha_{a3b}$  and set the independent truncated normal distributions as prior distribution for cut point parameters  $\alpha_{1b}$  and  $\alpha_{a2b}$ . For the regression coefficients the prior distributions are the same as priors in Model (II). The posterior results for the parameters of marginal modelling for the response at week 2 ( $Y_1$ ) and the results of the transition modelling of  $Y_2$  given  $Y_1$  and of  $Y_3$  given  $Y_2$  based upon three models are given in Tables 1 to 3.

In Bayesian paradigm, one approach to model selection is the use of Bayes factor. The comparison between two models  $H_i$  versus  $H_j$ ,  $i, j \in \{I, II, III\}$  and  $i \neq j$ , is possible by computing the Bayes factor. If  $\pi(H_i)$  and  $\pi(H_j)$  are the prior probabilities of these models then the Bayes factor is obtained by

$$B_{ij} = \frac{f(Y|H_i)}{f(Y|H_j)} = \frac{\frac{\pi(H_i|Y)}{\pi(H_j|Y)}}{\frac{\pi(H_i)}{\pi(H_j)}}$$

The Bayes factor for each couple of models, is given in Table 4. According to this table, we see that between two models (II) and (III) we approve the model (II) because of the value of  $B_{23}$ ,  $B_{23} = 13.755$ . According to Kass and Raftery (1995) and Jeffreys (1961), since  $10 < B_{23} < 100$  we have strong evidence in favor of the model (II) against the model (III). Between two models (I) and (II), the Bayes factor  $B_{12} = 1.580$ . this value of  $B_{12}$  didn't say that model (I) is better than model (II). In fact model (I) is a bit better than model (II), but we have no strong evidence that (I) dominates (II). Hence, we can accept both of models (I) and (II), but the more parsimonious one is model (II).

We also use the deviance information criterion (DIC, Spiegelhalter et al., 2002) as a measures of goodness of fit when we compare models (I) and (II). The DIC is in favour of model (II) as implied by the lower DIC value, (for model (I), DIC=1160.570 and for model (II), DIC= 1158.460).

The results for Model (II) are given in Table 2. According to this table, 'age' and 'duration' have a significant effect on side effects in week 2 (model for  $Y_1$ ), there is no effect of severity'. In fact the covariate 'severity' hasn't significant effects in all states and times. Results of marginal modelling of response in week 2 (model for  $Y_1$ ) show that for older patients and patients with longer durations of illness, the probability of having more severe side effects is higher.

The results of the transition modelling of  $Y_2$  given  $Y_1$  show that when there is no side effects in week 2 ( $Y_1 = 1$ ) or when there is significant interference with functionality in week 2 ( $Y_1 = 3$ ), none of the covariates 'age' and 'duration' have the significant effects on side effect. But, when ( $Y_1 = 2$ ), there are significant effects of covariates 'age' and 'duration' on side effects in week 4. When ( $Y_1 = 2$ ), the odds of observing more severe side effects increase with 'age' and 'duration'.

For different values of  $Y_1$ , assuming there is no effect of any other covariates, cut-point parameters indicate the log odds of less severe, rather than more severe, side effects in week 4 ( $Y_2$ ). For example, when ( $Y_1 = 1$ ), the log odd of having no side effects in week 4 ( $Y_2 = 1$ ) is 2.008, which it means that if an individual has no side effects in week 2, then he/she will has a high probability (0.882) of having no side effects in week 4, i.e.,  $P(Y_2 = 1|Y_1 = 1) = \frac{\exp(2.008)}{1+\exp(2.008)} = 0.882$ . When ( $Y_1 = 3$ ), the log odd of having no side effects in week 4 when ( $Y_2 = 1$ ) is reduced to  $-1.567$ , which it means that if an individual has significant interference with functionality in week 2, then the probability that he/she has no side effects in week 4 is a low probability of 0.173. Also, when ( $Y_1 = 3$ ), the log odd of ( $Y_2 \leq 2$ ), i.e., having either no side effects or no significant interference with functionality in week 4 are  $-0.146$  which it means that  $P(Y_2 \leq 2|Y_1 = 3) = \frac{\exp(-0.146)}{1+\exp(-0.146)} = 0.464$ . In other words, if an individual has significant interference with functionality in week 2, then the probability that he/she has no side effects or no significant interference with functionality in week 4 is equal to 0.464. According to Table 2, in the transition modelling of  $Y_2$  given  $Y_1$ , the value of cut-points reduce when the value of  $Y_1$  increase. It shows that the probability of having less severe side effects in week 4 ( $Y_2$ ) decreases when the previous response ( $Y_1$ ) increases.

In Table 2 the results of the transition modelling of  $Y_3$  given  $Y_2$  are also given. These results show that the covariate ‘age’ wasn’t significant in the transition model of  $Y_3$  given  $Y_2$ . The covariate ‘duration’ has a significant effect on side effects, only when an individual has no interfering significant with functionality ( $Y_2 = 2$ ).

For different values of  $Y_2$ , cut-point parameters indicate the log odds of less severe, rather than more severe, side effects in week 12 ( $Y_3$ ). For example,  $P(Y_3 = 1|Y_2 = 1) = 0.896 = \frac{\exp(2.158)}{1+\exp(2.158)}$  which means that if an individual has no side effects in week 4, then he/she will has a high probability (0.896) of having no side effects in week 12. When ( $Y_2 = 1$ ), the probability of  $Y_3 \leq 2$  is equal to 0.996 which means that if an individual has no side effects in week 4, then with probability 0.996 he/she would have no side effects at all or no significant interference with functionality.  $P(Y_3 \leq 2|Y_2 = 3) = 0.812$  means that if the Fluvoxamine drug has significantly interfered with functionality in week 4, then with high chance of (0.812), Fluvoxamine drug will produce no side effects at all or will not significantly interfere with functionality in week 12. When ( $Y_2 = 2$ ), the odds of observing more severe side effects increase with duration of illness.

Table 1. Posterior mean of cut-points and regression coefficient parameters for the parsimonious non-homogeneous model (Model (I)). .

	$Y_1$	$Y_2 Y_1 = 1$	$Y_2 Y_1 = 2$	$Y_2 Y_1 = 3$
Par.	Est. (S.E.)	Est. (S.E.)	Est. (S.E.)	Est. (S.E.)
$\alpha_{at1}$	0.650 (0.388)	1.999 (0.284)	0.870 (0.735)	-1.551 (0.523)
$\alpha_{at2}$	2.827 (0.426)	—	5.097 (0.971)	-0.120 (0.399)
$\beta_{t1}^a$ (Age)	0.018 (0.009)	—	0.038 (0.017)	—
$\beta_{t2}^a$ (Duration)	0.015 (0.005)	—	0.028 (0.009)	—
		$Y_3 Y_2 = 1$	$Y_3 Y_2 = 2$	$Y_3 Y_2 = 3$
Par.		Est. (S.E.)	Est. (S.E.)	Est. (S.E.)
$\alpha_{at1}$		2.168 (0.299)	-0.319 (0.252)	—
$\alpha_{at2}$		5.411 (1.276)	3.655 (0.649)	1.453 (0.896)
$\beta_{t1}^a$ (Age)		—	—	—
$\beta_{t2}^a$ (Duration)		—	0.030 (0.015)	—

Table 2. Posterior mean of cut-points and regression coefficient parameters for the partial homogeneous model (Model (II)). .

	$Y_1$	$Y_2 Y_1 = 1$	$Y_2 Y_1 = 2$	$Y_2 Y_1 = 3$
Par.	Est. (S.E.)	Est. (S.E.)	Est. (S.E.)	Est. (S.E.)
$\alpha_{at1}$	0.675 (0.381)	2.008 (0.290)	0.838 (0.678)	-1.567 (0.529)
$\alpha_{at2}$	2.856 (0.424)	—	5.057 (0.890)	-0.146 (0.411)
$\beta_{t1}^a$ (Age)	0.019 (0.009)	—	0.037 (0.015)	—
$\beta_{t2}^a$ (Duration)	0.015 (0.005)	—	0.028 (0.008)	—
		$Y_3 Y_2 = 1$	$Y_3 Y_2 = 2$	$Y_3 Y_2 = 3$
Par.		Est. (S.E.)	Est. (S.E.)	Est. (S.E.)
$\alpha_{at1}$		2.158 (0.295)	-0.326 (0.238)	—
$\alpha_{at2}$		5.471 (1.282)	3.583 (0.588)	1.464 (0.901)
$\beta_{t1}^a$ (Age)		—	—	—
$\beta_{t2}^a$ (Duration)		—	0.028 (0.008)	—

Table 3. Posterior mean of cut-points and regression coefficient parameters for the homogeneous model (Model (III)).

	$Y_1$	$Y_2 Y_1 = 1$	$Y_2 Y_1 = 2$	$Y_2 Y_1 = 3$
Par.	Est. (S.E.)	Est. (S.E.)	Est. (S.E.)	Est. (S.E.)
$\alpha_{at1}$	0.680 (0.395)	2.068 (0.210)	-0.237 (0.224)	-1.356 (0.499)
$\alpha_{at2}$	2.862 (0.432)	—	3.753 (0.417)	0.194 (0.338)
$\beta_{t1}^a$ (Age)	0.019 (0.009)	—	0.014 (0.006)	—
$\beta_{t2}^a$ (Duration)	0.015 (0.005)	—	0.027 (0.008)	—

  

	$Y_3 Y_2 = 1$	$Y_3 Y_2 = 2$	$Y_3 Y_2 = 3$
Par.	Est. (S.E.)	Est. (S.E.)	Est. (S.E.)
$\alpha_{at1}$	2.068 (0.210)	-0.237 (0.224)	—
$\alpha_{at2}$	5.329 (1.225)	3.753 (0.417)	0.194 (0.338)
$\beta_{t1}^a$ (Age)	—	—	—
$\beta_{t2}^a$ (Duration)	—	0.027 (0.008)	—

Table 4. Bayes factors for comparing models (I) to (III). In each cell the value is the Bayes factor of column model against row model .

Hypothesis	model (I)	model (II)	model (III)
model (I)	1	0.633	0.046
model (II)	1.580	1	0.073
model (III)	21.732	13.755	1

### 3. Conclusion

In this paper, we presented Bayesian tests of homogeneity in cumulative logit transition model for modelling and analysis of longitudinal ordinal response data with possibility of drop-out. Since the posterior distributions of parameters could not be obtained in analytically tractable forms, we used simulation (MCMC)-based approaches. Three versions of the cumulative logit regression models were implemented using real data from a psychiatric treatment program (Fluvoxamine) and the Bayesian estimates of the parameters of each model are presented. Bayes Factor and DIC are used to compare these models to find the model that fits the data the best. Model (II) is found to be good enough for fitting the data. Model (II) considers the cumulative logit transition to describe the ordinal structure of the data which has time-invariant covariate ‘Duration’ and hence means that the coefficients of this covariate does not change with time. So in view of hypothesis testing, the homogeneity assumption (model(III)) is rejected against the partial homogeneity assumption (model(II)) and also against the non-homogeneity assumption of transition model (model (I)).

### Acknowledgments

We are grateful to the editor and the referees for their valuable comments and helpful insights which lead to improvement of the paper.



## References

1. Chung, H., Park, Y. S. and Lanza, S. T. (2005). Latent transition analysis with covariates: pubertal timing and substance use behaviours in adolescent females. *Statistics in Medicine*, **24**, 2895–2910.
2. Diggle, P. J., Heagerty, P. J., Liang, K. Y. and Zeger, S. L. (2002). *Analysis of longitudinal data*. Oxford University Press, Oxford.
3. Diggle, P. J., and Kenward, M. G. (1994). Informative dropout in longitudinal data analysis (with discussion). *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, **43**, 49–94.
4. Francom, S. F., Chuang-Stein, C. and Landis, J. R. A. (1989). A log-linear model for ordinal data to characterize differential change among treatments. *Statistics in Medicine*, **8**, 571–582.
5. Garber, A. M. (1989). A discrete-time model of the acquisition of antibiotic-resistant infections in hospitalized patients. *Biometrics*, **45**, 797–816.
6. Harville, D. A. and Mee, R. W. (1984). A mixed-model procedure for analyzing ordered categorical data. *Biometrics*, **40**, 393–408.
7. Hines, R. J. O. and Hines, W. G. S. (2005). An appraisal of methods for the analysis of longitudinal categorical data with MAR drop-outs. *Statistics in Medicine*, **24** (23), 3549–3564.
8. Jeffreys, H. (1961). *Theory of probability*. 3rd edition, Oxford university press, Oxford.
9. Johnson, V. E. and Albert, J. H. (1999). *Ordinal data modelling*. Springer, New York.
10. Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, **90** (430), 773–795.
11. Kim, K. (1995). A bivariate cumulative probit Regression model for ordered categorical data. *Statistics in Medicine*, **14**, 1341–1352.
12. Liang, K. Y., Zeger, S. L. and Qaqish B. F. (1992). Multivariate regression analyses for categorical data. *Journal of the Royal Statistical Society, Series B (Methodological)*, **54**, 3–40.
13. Little, R. J. A. and Rubin, D. B. (2002). *Statistical analysis with missing data*. 2nd edition, Wiley, New York.
14. Madigan, D. and Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam’s window. *Journal of the American Statistical Association*, **89** (428), 1535–1546.
15. Molenberghs, G. and Lesaffre, E. (1994). Marginal modelling of correlated ordinal data using a multivariate plackett distribution. *Journal of the American Statistical Association*, **89**, 633–644.
16. Molenberghs, G. and Verbeke, G. (2005). *Models for discrete longitudinal data*, Springer, New York.
17. Rezaei Ghahroodi, Z., Ganjali, M. and Berridge, D. (2009). A transition model for ordinal response data with random dropout: an application to the fluvoxamine data. *Journal of Biopharmaceutical Statistics*, **19** (4), 658–671.
18. Rezaei Ghahroodi, Z. and Ganjali, M. (2009). Testing homogeneity in Markov models for analyzing longitudinal ordinal response data with random dropout. *Journal of Statistical Theory and Applications*, **8** (2), 125–139.
19. Schweinberger, R. (2012). Statistical modelling of network panel data: Goodness of fit. *British Journal of Mathematical and Statistical Psychology*, **65** (2), 263–281.
20. Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Linde, A. V. D. (2002). Bayesian Measures of Model Complexity and Fit (with discussion). *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, **64** (4), 583–639.
21. Ten Have, T. R., Landis, J. R. and Hartzel, J. (1996). Population-averaged and cluster-specific models for clustered ordinal response data. *Statistics in Medicine*, **15**, 2573–2588.
22. Tutz, G. and Hennevogl, W. (1996). Random effects in ordinal regression models. *Computational Statistics & Data Analysis*, **22**, 537–557.
23. Tutz, G. (2005). Modelling of repeated ordered measurements by isotonic sequential regression. *Statistical Modelling*, **5** (4), 269–287.
24. Verbeke, G. and Lesaffre, E. (1996). A linear mixed-effects model with heterogeneity in the random-effects population. *Journal of the American Statistical Association*, **91**, 217–221.
25. Verbeke, G. and Molenberghs, G. (1997). *Linear mixed models in practice: a SAS-oriented approach*. Springer, New York.
26. Weißbach, R. and Dette, H. (2007). Kolmogorov-Smirnov-type testing for the partial homogeneity of Markov processes - with application to credit risk. *Applied Stochastic Models in Business and Industry*, **23** (3), 223–234.