Statistics & Probability Letters 70 (2004) 25–35

STATISTICS & PROBABILITY LETTERS

Generating correlated ordinal categorical random samples

Atanu Biswas*

Applied Statistics Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata 700 108, India

Received 18 November 2002; received in revised form 30 June 2004

Abstract

Ordinal categorical random variables are common in many studies. In different context it is important to appropriately define and simulate from such ordinal categorical random variables with a desired pattern of the correlation structure. This is an important problem in longitudinal studies as well as analyzing clustered data involving ordinal categorical responses. The present paper deals with the theoretical presentation and the construction of multivarite ordinal categorical random variables with some desired patterns of correlation structure. Algorithms for generating samples for the AR-type correlation with particular illustration of AR(1) and AR(2), and equicorrelation are discussed using some urn models.

Keywords: Longitudinal autocorrelation; Equicorrelation model; Ordinal categorical random variable; Urn model

1. Introduction

Correlated random variables with some known pattern of correlation structure are often important in statistical study. In the recent years, with the revolution of longitudinal studies and clustered analysis having mixed effects, theoreticians as well as practitioners are to deal with different types of correlated ordinal categorical data. In several datasets involving pain, post-operative conditions, etc., correlated ordinal random variables (classified as nil/mild/moderate/severe, for example) comes under the purview of study. Some examples of correlated categorical random variables in the literature are due to Dale (1986), Klein et al. (1984), Koch et al. (1989) and Molenberghs and Lesaffre (1994), among others.

Zeger et al. (1985) discussed the construction of AR(1) autocorrelation structure for repeated binary data. Prentice (1988) dealt with correlated binary regression with common value of pairwise correlations. Correlation structure for the multivariate binary data can now be easily defined and represented by using the Bahadur representation (Bahadur, 1961; Prentice, 1988; Lipsitz et al., 1991). But relatively little attention has been paid to polytomous categorical variables. In the recent years, among the explosion of papers on repeated measurement problems, models have been developed for modeling repeated observations of some ordinal categorical response obtained over time on the same individual. One of the first approaches to the analysis of repeated categorical responses is due to Koch et al. (1977). In such modeling, both the transition models describing the probability distribution of subject's future events given the subject's prior history and the marginal models utilizing various methodological strategies to account for the correlation between repeated measurements can be employed (see Ware et al., 1988). Consequently, there has been some attempts to model correlated ordinal responses. In practice, we need a flexible model for such multivariate categorical responses. Based on the work of Dale (1986), Molenberghs and Lesaffre (1994) used the multivariate Plackett distribution to explain multivariate ordinal data. Note that none of the existing models for categorical responses incorporate simultaneously a simple model for the conditional and marginal approach (see Ashby et al., 1992). The present paper is motivated to fulfil that gap.

To study the performance of several concerned theory, one may need to simulate random samples from a properly correlated setup. The latent variable approach is not suitable as correlation between the derived categorical random variables are not of simple form or of simple interpretation. The present paper provides some simple algorithms to generate such random samples for some specific correlation structures. Consequently, one can write down the joint probability mass function of such correlated categorical random variables which, of course, may not have a simple form. But the sample generation technique is quite easy and elegant. In Section 2, we propose our technique with the AR-type correlation with illustration with AR(1) and AR(2) models. In Section 3 we deal with the equicorrelation structure. Section 4 provides the psuedocodes of the algorithms of Sections 2 and 3. Section 5 ends with some concluding remarks.

2. Autocorrelation models

In the present paper, we discuss the AR(1) and AR(2)-type autocorrelation models. A general AR(p)-type autocorrelation model can similarly be described.

2.1. AR(1) model

Suppose we need to generate $Y_1, Y_2, \ldots, Y_T$, which are T identically distributed ordinal random variables, longitudinally obtained at T successive time points and we are interested to introduce a desired correlation structure within them. Suppose each Y_t can take the possible values $0,1,\dots,k$. Again, for known $a_j(>0)$, $j=0,1,\dots,k$, we want to have

$$P(Y_i=j)\propto a_j, \tag{1}$$

where ρ_{il} be the correlation coefficient between Y_i and Y_l. To achieve $\rho_{il}=\rho^{|i-l|}$, for some $\rho=b/(\sum_{j=0}^{k}a_j+b)$, we employ the following algorithm. The pseudocode of the algorithm is presented in Section 4.

Algorithm A1.

1. We discuss the construction of $Y_1,Y_2,\dots,Y_T$ successively with the help of T urns, labeled $1,2,\dots,T$, each having $\sum_{j=0}^{k}a_j$ balls at the outset, a_j balls of kind A_j, $j=0,1,\dots,k$. A ball of kind A_j represents the value of corresponding Y as j.
2. At first we take the urn labeled '1', draw a ball from it and notice the kind of the drawn ball. If the drawn ball is of kind A_{j_1}, then the value of Y_1 will be j_1. Then we add an additional b balls of kind j_1 to the urn labeled '2'. This urn will now have a total of $(\sum_{j=0}^{k}a_j+b)$ balls of which $(a_{j_1}+b)$ balls of kind A_{j_1} and a_j balls of every other kind A_j. This urn will now reflect the conditional probability distribution of Y_2 given Y_1. We now draw a ball from this urn to find Y_2. Let the realized value of Y_2 be j_2.
3. We now take the urn labeled '3', add new b balls of kind A_{j_2} in it which makes the total number of balls in that urn to be $(\sum_{j=0}^{k}a_j+b)$, of which $(a_{j_2}+b)$ balls of kind A_{j_2} and a_j balls of every other kind A_j. We draw a ball from the urn to find Y_3.
4. We continue this procedure up to the Tth urn.

Note that all the positive values of ρ are covered by this approach and an interval of the negative values. From the urn model formulation (2), in order the right-hand side of (2) to be nonnegative, we need $b\geqslant -\min\{a_j\}$, and hence

$$\rho\in\left[-\frac{\min(a_j)}{\sum a_u-\min(a_j)},1\right].$$

Result 1. The observations $\{Y_1,Y_2,\dots,Y_T\}$ obtained using the Algorithm A1 is such that

(a) The marginal distribution of any Y_i is given by (1).
(b) Here $\rho_{il}=\rho^{|i-l|}$ for $\rho=b/(\sum a_u+b)$.

Proof. (a) From the urn model formulation it is easy to note that from the composition of the urn '1' we have

$$P(Y_1=j)=\frac{a_j}{\sum_{u=0}^{k}a_u},\quad j=0,1,\dots,k.$$

Again, the conditional probability distribution of any Y_i, $i=2,3,\dots,T$, given Y_{i-1} is

$$P(Y_i=j|Y_{i-1})=\frac{a_j+bI(Y_{i-1},j)}{\sum a_u+b},\quad j=0,1,\dots,k, \tag{2}$$

where $I(Y,j)$ is the indicator variable which takes the value 1 if $Y=j$ and 0 elsewhere. Taking expectation in both sides of (2) with respect to Y_{i-1}, noting that $E(I(Y_{i-1},j)) = P(Y_{i-1}=j) = a_j/\sum a_u$, the unconditional probability distribution of Y_i is given by (1).

(b) Let M and V be the expectation and variance of any Y_i, respectively, where

$$M = \frac{\sum j a_j}{\sum_j},$$

$$V = \frac{(\sum a_j)(\sum j^2 a_j) - (\sum j a_j)^2}{(\sum a_j)^2}.$$

Now, the conditional probability distribution of Y_{i+2} given Y_i can be obtained by taking expectation over the distribution of Y_{i+1} given Y_i. Thus

$$P(Y_{i+2}=j|Y_i) = \frac{a_j + bP(Y_{i+1}=j|Y_i)}{\sum a_u + b} = \frac{a_j + b\left(\frac{a_j+bI(Y_i,j)}{\sum a_u+b}\right)}{\sum a_u + b}$$
$$= \frac{a_j(\sum a_u + b) + b a_j + b^2 I(Y_i,j)}{(\sum a_u + b)^2}.$$

Proceeding in this way and taking expectations recursively, we have for $t = 1,2,\ldots,$

$$P(Y_{i+t}=j|Y_i) = \frac{a_j\sum_{w=0}^{t-1}(\sum a_u + b)^w b^{t-1-w} + b^t I(Y_i,j)}{(\sum a_u + b)^t}.$$

Consequently, the conditional expectation of Y_{i+t} given Y_i is

$$E(Y_{i+t}|Y_i) = \frac{(\sum j a_j)(\sum_{w=0}^{t-1}(\sum a_j + b)^w b^{t-1-w}) + b^t \sum j I(Y_i,j)}{(\sum a_j + b)^t}.$$

Noting that $E[Y_i I(Y_i,j)] = jP(Y_i=j)$, we obtain

$$E(Y_iY_{i+t}) = E\{Y_i E(Y_{i+t}|Y_i)\}$$
$$= \frac{(\sum j a_j)^2(\sum_{w=0}^{t-1}(\sum a_j + b)^w b^{t-1-w}) + b^t(\sum j^2 a_j)}{(\sum a_j)(\sum a_j + b)^t}.$$

Then the covariance of Y_i and Y_{i+t} is

$$\mathrm{cov}(Y_i, Y_{i+t}) = \frac{(\sum j a_j)^2(\sum_{w=0}^{t-1}(\sum a_j + b)^w b^{t-1-w}) + b^t(\sum j^2 a_j)}{(\sum a_j)(\sum a_j + b)^t} - \left(\frac{\sum j a_j}{\sum a_j}\right)^2$$
$$= \frac{b^t(\sum j^2 a_j)(\sum a_j) - (\sum j a_j)^2[(\sum a_j + b)^t - (\sum a_j)(\sum_{w=0}^{t-1}(a_j + b)^w b^{t-1-w})]}{(\sum a_j)^2(\sum a_j + b)^t}. \quad (3)$$

Noting that

$$(C+D)^n - C\sum_{i=0}^{n-1}(C+D)^i D^{n-1-i} = D^n,$$

the expression under (3) reduces to

$$\left(\frac{b}{\sum a_j + b}\right)^t V.$$

Thus the correlation between Y_t and Y_s becomes

$$\mathrm{corr}(Y_t, Y_s) = \rho_{ts} = \rho^{|t-s|},$$

with $\rho = b/(\sum a_j + b)$. □

The joint probability mass function of $Y_1, \ldots, Y_T$ is

$$P(Y_1 = y_1, \ldots, Y_T = y_T) = \left(\frac{a_{y_1}}{\sum_{u=0}^{k} a_u}\right) \prod_{t=2}^{T} \left(\frac{a_{y_t} + bI(y_t, y_{t-1})}{\sum a_u + b}\right).$$

2.2. AR(2) model

Here, we want to achieve $\rho_{t,t+l} = (b\rho_{t,t+l-1} + c\rho_{t,t+l-2})/(\sum a_u + b + c)$ with $\rho_{12} = b/(\sum a_u + b + c)$.

Algorithm A2.

1. Here, as earlier, we start with T urns, initially each having $\sum a_u$ balls, a_j balls of kind A_j.
2. We draw a ball from urn '1' to find Y_1. Suppose the observed value of Y_1 is j_1. We add $(b + c)$ balls to the second urn, $(b + ca_{j_1}/\sum a_u)$ balls of kind A_{j_1} and $ca_j/\sum a_u$ balls of every remaining kind A_j. Now this urn '2' will have a total of $(\sum a_u + b + c)$ balls of which $(a_{j_1} + b + ca_{j_1}/\sum a_u)$ balls are of kind A_{j_1} and all the remaining kind A_j have $(a_j + ca_j/\sum a_u)$ balls.
   Here, at the tth time point, b balls reflect the influence of Y_{t-1} and c balls reflect the influence of Y_{t-2}. At $t = 2$, there is no Y_{t-2} to add 'c' balls to the urn model. Hence, by convention, we distribute these c balls according to the weights of the possible $(k + 1)$ values.
3. We draw a ball from this urn to get Y_2, let it be j_2. From the third urn onwards, for any urn labeled 'i', we add $(b + c)$ balls to the urn, b balls of kind j_{i-1}, the realized value of Y_{i-1}, and also add c balls of kind j_{i-2}, the realized value of Y_{i-2}.
4. We continue this procedure up to the Tth urn.

Result 2. The observations generated using the Algorithm A2 are such that

(a) The marginal distribution of any Y_t is given by (1).
(b) Here

$$\rho_{12} = \frac{b}{\sum a_u + b + c} \tag{4}$$

and all other correlations satisfy the recursive relation

$$\rho_{i,i+t} = \frac{b\rho_{i,i+t-1} + c\rho_{i,i+t-2}}{\sum a_u + b + c}. \tag{5}$$

Proof. (a) From the above urn model we observe that the unconditional probability distribution of Y_1 is

$$P(Y_1 = j) = \frac{a_j}{\sum a_u}, \quad j = 0, 1, \ldots, k,$$

the conditional probability distribution of Y_2 given Y_1 is

$$P(Y_2 = j | Y_1) = \frac{a_j + bI(Y_1, j) + ca_j / \sum a_u}{\sum a_u + b + c}, \quad j = 0, 1, \ldots, k \tag{6}$$

and, for $i = 3, 4, \ldots, T$, the conditional probability distribution of Y_i given Y_{i-1} and Y_{i-2} is

$$P(Y_i = j | Y_{i-1}, Y_{i-2}) = \frac{a_j + bI(Y_{i-1}, j) + cI(Y_{i-2}, j)}{\sum a_u + b + c}, \quad j = 0, 1, \ldots, k. \tag{7}$$

Taking expectations in both sides of (6) with respect to Y_1 and in both sides of (7) with respect to Y_{i-1} and Y_{i-2}, we find that Y_i's, $i = 1, 2, \ldots, T$, are identically distributed as (1).

(b) Clearly, the expectation and variance of any Y_i will be M and V, respectively. From (7), for any $i + t \geqslant 3$, taking expectations on both sides, we get the conditional probability distribution of Y_{i+t} given Y_i as

$$P(Y_{i+t} = j | Y_i) = \frac{a_j + bP(Y_{i+t-1} = j | Y_i) + cP(Y_{i+t-2} = j | Y_i)}{\sum a_u + b + c}, \quad j = 0, 1, \ldots, k.$$

Consequently, the conditional expectation of Y_{i+t} given Y_i is

$$E(Y_{i+t} | Y_i) = \frac{\sum j a_j + bE(Y_{i+t-1} | Y_i) + cE(Y_{i+t-2} | Y_i)}{\sum a_u + b + c},$$

yielding

$$E(Y_i Y_{i+t}) = \frac{\frac{(\sum j a_j)^2}{\sum a_j} + bE(Y_i Y_{i+t-1}) + cE(Y_i Y_{i+t-2})}{\sum a_u + b + c}.$$

Consequently, we find

$$\operatorname{cov}(Y_i, Y_{i+t}) = \frac{b \operatorname{cov}(Y_i, Y_{i+t-1}) + c \operatorname{cov}(Y_i, Y_{i+t-2})}{\sum a_u + b + c}$$

and hence (5) follows. Using the same technique, one can easily obtain (4). □

We can use the recursion relation (5) to find several correlations provided we know ρ_{12}. It can be observed from (5) that for any other pair $(i, i+1)$ except $(1, 2)$, we have

$$\rho_{i,i+1} = \frac{b + c\rho_{i-1,i}}{\sum a_u + b + c},$$

which also holds for ρ_{12} if we define $\rho_{0,1}=0$. In that case (5) holds for any $(i,t): i=1,2,\dots,T-1; t=1,2,\dots,T-i$. Again, from (4), we can argue that $(b+c)$ should be at least as large as $-\min\{a_j\}$.

The joint probability distribution of $Y_1,\dots,Y_T$ can be written as

$$P(Y_1=y_1,\dots,Y_T=y_T)=\left(\frac{a_{y_1}}{\sum_{u=0}^{k}a_u}\right)\left(\frac{a_{y_2}+bI(y_2=y_1)+ca_{y_1}/\sum a_u}{\sum a_u+b+c}\right)\times\prod_{i=3}^{T}\left(\frac{a_{y_i}+bI(y_i=y_{i-1})+cI(y_i=y_{i-2})}{\sum a_u+b+c}\right)$$

3. Equicorrelation model

Equicorrelation structures are important in cluster analysis, where the random variables have equal correlation among them due to some random effect. To obtain equal correlation, $(b/\sum a_u+b)^2$, for correlated ordinal categorical random variables we proceed as follows. Suppose, in a clustered analysis, a random effect is affecting each of $Y_1,Y_2,\dots,Y_T$ in the same way. Our object is to model that effect of the random component and obtain the correlations between any two Y_i's.

Algorithm A3.

1. Suppose the random effect is denoted by Y_0 which is also ordered categorical, taking values $0,1,\dots,k$, with $P(Y_0=j)=a_j/\sum a_u$.
2. We start with T urns for generating $Y_1,\dots,Y_T$, each with a_j balls of kind A_j at the outset.
3. If the realized value of Y_0 is j_0, we add b balls of kind A_{j_0} in each of the T urns. Each of the urns have now $(\sum a_j+b)$ balls in total, $(a_{j_0}+b)$ balls of kind A_{j_0} and a_j balls of all the remaining kind A_j.
4. Then generate $Y_1,\dots,Y_T$ by drawing one ball from each of the urns.

Result 3. For the observations $\{Y_1,Y_2,\dots,Y_T\}$ generated using the Algorithm A3, we have

(a) The marginal distribution of any Y_i is given by (1).
(b) The correlation coefficient between any Y_i and Y_s is given by

$$\rho_{is}=\left(\frac{b}{\sum a_u+b}\right)^2.$$

Proof. (a) From the urn model, we have for $i=1,2,\dots,T$, the conditional distribution of Y_i given Y_0 is

$$P(Y_i=j|Y_0)=\frac{a_j+bI(Y_0,j)}{\sum a_u+b},$$

whence taking expectation we get the unconditional distribution as given by (1).

(b) Taking conditional expectation of the above, we get

$$E(Y_i|Y_0)=\frac{\sum ja_j+b\sum jI(Y_0,j)}{\sum a_j+b}$$

and consequently,

$$E(Y_0Y_i)=\frac{(\sum ja_j)^2+b\sum j^2a_j}{(\sum a_j)(\sum a_j+b)},$$

yielding

$$\rho_{0i}=\frac{b}{\sum a_j+b}$$

with ρ_{0i} being the correlation between Y_0 and Y_i. Exactly in the same way we find

$$P(Y_i=j|Y_s)=\frac{a_j+bP(Y_0=j|Y_s)}{\sum a_j+b}$$

and, consequently,

$$\rho_{it}=\left(\frac{b}{\sum a_j+b}\right)\rho_{0t}=\left(\frac{b}{\sum a_j+b}\right)^2 \qquad \square$$

Note that the random effect Y_0 is affecting all the Y_i's in the same way, and as the correlation between Y_i's are through this random effect, we get a positive correlation in this case, which is the case, in general, in any random effect model. The joint probability distribution of $Y_1,\ldots,Y_T$ in this setup can be written as

$$P(Y_1=y_1,\ldots,Y_T=y_T)=\sum_{y_0=0}^{k}\left\{\prod_{i=1}^{T}\left(\frac{a_{y_i}+bI(y_i=y_0)}{\sum a_u+b}\right)\right\}\frac{a_{y_0}}{\sum_{u=0}^{k}a_u}$$

4. Implementation

In this section, we provide the pseudocodes of the Algorithms A1–A3 in the spirit of Paatero (1999).

Pseudocode of the Algorithm A1:

1. Initialize the probability distribution of Y_1.
   1.1. Set $p_s^{(1)}=P(Y_1=s)=a_s/\sum_{j=0}^{k}a_j$, $s=0,1,\ldots,k$.
2. Drawing random sample from the probability distribution of Y_1.
   2.1. Find the cumulative probability distribution of Y_1 as $Q_s^{(1)}=P(Y_1\leqslant s)=\sum_{j=0}^{s}p_j^{(1)}$, $s=0,1,\ldots,k$.
   2.2. Set $Q_{-1}^{(1)}=0$.
   2.3. Draw a random number r_1 between [0, 1].
   2.4. For $j=0,1,\ldots,k$, if $Q_{j-1}^{(1)}<r_1\leqslant Q_j^{(1)}$, then $Y_1=j$.

3. Drawing random sample from the probability distribution of Y_t, $t = 2, 3, \ldots, T$.
   3.1. Find the probability distribution of Y_t, $t = 2, 3, \ldots, T$, as follows. If $Y_{t-1} = Z$, then $p_Z^{(t)} = P(Y_t = Z) = (a_Z + b)/(\sum_{j=0}^{k} a_j + b)$; and $p_s^{(t)} = P(Y_t = s) = a_s/(\sum_{j=0}^{k} a_j + b)$ for $s = 0, 1, \ldots, k$, but $s \neq Z$.
   3.2. Set the cumulative probability distribution of Y_t as $Q_s^{(t)} = P(Y_t \leqslant s) = \sum_{j=0}^{s} p_j^{(t)}$, $s = 0, 1, \ldots, k$.
   3.3. Set $Q_{-1}^{(t)} = 0$.
   3.4. Draw a random number r_t between [0, 1].
   3.5. For $j = 0, 1, \ldots, k$, if $Q_{j-1}^{(t)} < r_t \leqslant Q_j^{(t)}$, then $Y_t = j$.

Pseudocode of the Algorithm A2:

1. Initialize the probability distribution of Y_1.
   1.1. Set $p_s^{(1)} = P(Y_1 = s) = a_s/\sum_{j=0}^{k} a_j$, $s = 0, 1, \ldots, k$.
2. Drawing random sample from the probability distribution of Y_1.
   2.1. Find the cumulative probability distribution of Y_1 as $Q_s^{(1)} = P(Y_1 \leqslant s) = \sum_{j=0}^{s} p_j^{(1)}$, $s = 0, 1, \ldots, k$.
   2.2. Set $Q_{-1}^{(1)} = 0$.
   2.3. Draw a random number r_1 between [0, 1].
   2.4. For $j = 0, 1, \ldots, k$, if $Q_{j-1}^{(1)} < r_1 \leqslant Q_j^{(1)}$, then $Y_1 = j$.
3. Drawing random sample from the probability distribution of Y_2.
   3.1. Find the probability distribution of Y_2 as follows. If $Y_1 = Z$, then $p_Z^{(2)} = P(Y_2 = Z) = (a_Z + b + ca_Z/\sum_{u=0}^{k} a_u)/(\sum_{j=0}^{k} a_j + b + c)$; and $p_s^{(2)} = P(Y_2 = s) = (a_s + ca_Z/\sum_{u=0}^{k} a_u)/(\sum_{j=0}^{k} a_j + b + c)$ for $s = 0, 1, \ldots, k$, but $s \neq Z$.
   3.2. Set the cumulative probability distribution of Y_2 as $Q_s^{(2)} = P(Y_2 \leqslant s) = \sum_{j=0}^{s} p_j^{(2)}$, $s = 0, 1, \ldots, k$.
   3.3. Set $Q_{-1}^{(2)} = 0$.
   3.4. Draw a random number r_2 between [0, 1].
   3.5. For $j = 0, 1, \ldots, k$, if $Q_{j-1}^{(2)} < r_2 \leqslant Q_j^{(2)}$, then $Y_2 = j$.
4. Drawing random sample from the probability distribution of Y_t, $t = 3, 4, \ldots, T$.
   4.1. Find the probability distribution of Y_t, $t = 3, 4, \ldots, T$, as follows. Denote $Y_{t-1} = Z$ and $Y_{t-2} = W$. If $Z \neq W$, set $p_Z^{(t)} = P(Y_t = Z) = (a_Z + b)/(\sum_{j=0}^{k} a_j + b + c)$; $p_W^{(t)} = P(Y_t = W) = (a_W + c)/(\sum_{j=0}^{k} a_j + b + c)$; and $p_s^{(t)} = P(Y_t = s) = a_s/(\sum_{j=0}^{k} a_j + b + c)$ for $s = 0, 1, \ldots, k$, but $s \neq Z, W$. If $Z = W$, set $p_Z^{(t)} = P(Y_t = Z) = (a_Z + b + c)/(\sum_{j=0}^{k} a_j + b + c)$; and $p_s^{(t)} = P(Y_t = s) = a_s/(\sum_{j=0}^{k} a_j + b + c)$ for $s = 0, 1, \ldots, k$, but $s \neq Z$.
   4.2. Set the cumulative probability distribution of Y_t as $Q_s^{(t)} = P(Y_t \leqslant s) = \sum_{j=0}^{s} p_j^{(t)}$, $s = 0, 1, \ldots, k$.
   4.3. Set $Q_{-1}^{(t)} = 0$.
   4.4. Draw a random number r_t between [0, 1].
   4.5. For $j = 0, 1, \ldots, k$, if $Q_{j-1}^{(t)} < r_t \leqslant Q_j^{(t)}$, then $Y_t = j$.

Pseudocode of the Algorithm A3:

1. Initialize the probability distribution of Y_0.
   1.1. Set $p_s^{(0)} = P(Y_0 = s) = a_s/\sum_{j=0}^{k} a_j$, $s = 0, 1, \ldots, k$.

2. Drawing random sample from the probability distribution of Y_0.
   2.1. Find the cumulative probability distribution of Y_0 as $Q_s^{(0)} = P(Y_0 \leqslant s) = \sum_{j=0}^{s} p_j^{(0)}$, $s = 0, 1, \ldots, k$.
   2.2. Set $Q_{-1}^{(0)} = 0$.
   2.3. Draw a random number r_0 between [0, 1].
   2.4. For $j = 0, 1, \ldots, k$, if $Q_{j-1}^{(0)} < r_0 \leqslant Q_j^{(0)}$, then $Y_0 = j$.
3. Drawing random sample from the probability distribution of Y_t, $t = 1, 2, \ldots, T$.
   3.1. Find the probability distribution of Y_t, $t = 1, 2, \ldots, T$, as follows. If $Y_0 = Z$, then $p_Z^{(t)} = P(Y_t = Z) = (a_Z + b)/(\sum_{j=0}^{k} a_j + b)$; and $p_s^{(t)} = P(Y_t = s) = a_s/(\sum_{j=0}^{k} a_j + b)$ for $s = 0, 1, \ldots, k$, but $s \neq Z$.
   3.2. Set the cumulative probability distribution of Y_t as $Q_s^{(t)} = P(Y_t \leqslant s) = \sum_{j=0}^{s} p_j^{(t)}$, $s = 0, 1, \ldots, k$.
   3.3. Set $Q_{-1}^{(t)} = 0$.
   3.4. Draw a random number r_t between [0, 1].
   3.5. For $j = 0, 1, \ldots, k$, if $Q_{j-1}^{(t)} < r_t \leqslant Q_j^{(t)}$, then $Y_t = j$.

5. Concluding remarks

The proposed models have quite a large number of potential application in problems regarding multivariate ordinal data. The immediate applicability of the present model is to the analysis of longitudinal data where covariates are not time dependent and also to the analysis of clustered data. As our intention is to provide a theoretical model only, in the present paper we are not going for any real data analysis. Also the present algorithm can be used to study the properties of different inferential approaches concerning correlated categorical random variables.

One obvious but nontrivial generalization could be where the marginal distribution of Y_i's are different and also the case where the number of categories can vary for different Y_i's. The present method cannot be directly applied in that situation. The situation is under study and we hope to pursue some results in a future communication.

Acknowledgements

The author wishes to thank the referee for the constructive suggestions which led to some improvement over two earlier versions of the paper.

References

Ashby, M., Neuhaus, J.M., Hauck, W.W., Bacchetti, P., Heilbron, D.C., Jewell, N.P., Segal, M.R., Fusaro, R.E., 1992. An annotated bibliography of methods for analyzing correlated categorical data. Statist. Med. 11, 67–99.

Bahadur, R.R., 1961. A representation of the joint distribution of responses to n dichotomous items. In: H. Solomon (Ed.), Studies in Item Analysis and Prediction, Stanford Mathematical Studies in the Social Sciences, vol. VI. Stanford University Press, Palo Alto, pp. 158–168

Dale, J.R., 1986. Global cross-ratio models for bivariate, discrete, ordered responses. Biometrics 42, 909–917.
Klein, B.E.K., Davis, M.D., Segal, P., Long, J.A., Harris, W.A., Haug, G.A., Magli, Y., Syrjala, S., 1984. Diabetic retinopathy: assessment of severity and progression. Ophthalmology 91, 10–17.
Koch, G.G., Landis, J.R., Freeman Jr., D.H., Lehnen, R.G., 1977. A general methodology for the analysis of experiments with repeated measurement of categorical data. Biometrics 33, 133–158.
Koch, G.G., Carr, G.J., Amara, I.A., Stokes, M.E., Uryniak, T.J., 1989. Categorical data analysis. In: Berry, D.A. (Ed.), Statistical Methodology in the Pharmaceutical Sciences. Marcel Dekker, New York, pp. 391–475.
Lipsitz, S.R., Laird, N.M., Harrington, D.P., 1991. Generalized estimating equations for correlated binary data: using the odds ratio as a measure of association. Biometrika 78, 153–160.
Molenberghs, G., Lesaffre, E., 1994. Marginal modeling of correlated ordinal data using a multivariate Plackett distribution. J. Amer. Statist. Assoc. 89, 633–644.
Paatero, P., 1999. The multilinear engine—a table-driven least squares program for solving multilinear problems, including the n way parallel factor analysis model. J. Comput. Graphical Statist. 8, 854–888.
Prentice, R.L., 1988. Correlated binary regression with covariates specific to each binary observation. Biometrics 44, 1033–1048.
Ware, J.H., Lipsitz, S., Speizer, F.E., 1988. Issues in the analysis of repeated categorical outcomes. Statist Med. 7, 95–107.
Zeger, S.L., Liang, K.-Y., Self, S.G., 1985. The analysis of binary longitudinal data with time independent covariates. Biometrika 72, 31–38.