Pooled Autoregressive Models for Categorical Data

Zhenqiu (Laura) Lu and Zhiyong (Johnny) Zhang

Abstract Time series capture time dependent intra-individual variation within a single participant. When data are collected from more than one subject, methods developed for single subject intra-individual relationship may not fully work and laws governing inter-individual relationship may not apply to intra-individual relationship, especially when outcomes are categorical or ordinal data. These data are usually collected by the Likert table. This article aims to investigate the performance of four estimation methods for pooling time series data focusing on categorical outcomes and to address related issues through an an autoregressive model, AR(1). In this article, models for pooling time series were formulated, estimation methods were derived, simulation studies were conducted, results were summarized and compared.

Keywords Pooling time series · Autoregressive model · Categorical data · Conditional likelihood · Exact likelihood · Maximum likelihood estimation

1 Introduction

The variation analysis in psychological, social, and behavioral researches has many ramifications. Among them two main branches are inter-individual variation and intra-individual variation. Inter-individual variation is the variation between individuals, and also widely known as the analysis of cross-sectional data in many researches. Intra-individual variation is the time dependent variation within a single participant's time series. It is also known as the analysis of time series data or P-technique in Cattell's (1952) data-box Cattell (1952). In this type of study, usually

Z. (Laura) Lu (✉)
University of Georgia, Athens, GA, USA
e-mail: zlu@uga.edu

Z. (Johnny) Zhang
University of Notre Dame, Notre Dame, IN, USA

M. Wiberg et al. (eds.), *Quantitative Psychology*, Springer Proceedings in Mathematics & Statistics 393, https://doi.org/10.1007/978-3-031-04572-1_14

one subject is measured and the variables of interests are collected from each of a large number of occasions. Many methods are available for single time series analysis (e.g., Cattell et al., 1947; Molenaar, 1985; Nesselroade & Molenaar, 2003).

However, data collected in this way do not have inter-individual differences since there is only one subject involved, but they can reflect changes across occasions. Intra-individual analysis has become popular advanced by Nesselroade, Molenaar, and colleagues. So many researches on intra-individual relationship, data are collected from more than one subject. When multiple subjects are involved, methods developed for single subject intra-individual relationship may not fully work. Also, laws governing inter-individual relationship may not apply to intra-individual relationship (Molenaar, 2004; Nesselroade & Ram, 2004, e.g.,). There are few methods in literature dealing with the analysis of pooling multiple time series (Cattell & Scheier, 1961; Daly et al., 1974; Molenaar et al., 2003; Nesselroade & Molenaar, 1999, e.g.,). The attention of this article will be drawn to multiple subjects intra-individual variation analysis. Also, the data in educational and social areas are usually collected by Likert tables. But the research on multiple subjects time series for categorical outcomes is very few. So we fill the gap by focusing our research in this area.

This article aims to investigate the performance of different estimation methods for pooling time series data focusing on categorical outcomes and to address related issues through an AR(1) model. We focus on four estimation methods for multiple time series: pooling conditional likelihood estimation, pooling exact likelihood estimation, connecting data conditional likelihood, and connecting data exact likelihood.

This article is organized as follows. In the next section some introductory remarks about time series are given. First single series and multiple series focusing on the AR(1) model are described and formulated. And then different estimation methods for multiple time series are introduced and derived. Then follows a section of simulation studies in which the performance of four estimating methods are investigated under various conditions. Simulation results are provided after simulation design and implementation. The closing part of this article summarizes the simulation results, compares different estimation methods of aggregating time series, and provides practical implication.

2 Autoregressive Model and Categorical Data

2.1 First-Order Autoregressive Model, AR(1)

We first consider a model for a single subject (or individual). And then we extend it to the model for multiple subjects. Suppose we are interested in a first-order autoregressive model, AR(1), as follows.

$$y_1 : \text{ the initial value}$$

$$y_t = \mu + \alpha\, y_{t-1} + z_t \quad (t > 1) \quad with \quad z_t \sim i.i.d.\ N(0, \phi) \tag{1}$$

where y_t is the observed value at time point t, α is the model autoregressive coefficient, μ is a parameter correlated with the mean of y, z is a shock variable, or a white noise sequence, satisfying a normal distribution with mean 0 and variance ϕ. In this case, the vector of population parameters to be estimated consists of $\boldsymbol{\theta} = (\mu, \alpha, \phi)'$. When $|\alpha| < 1$, there is a covariance stationary process for y_t satisfying Eq (1). Thus, the remainder of this discussion of AR(1) assumes that $|\alpha| < 1$. By algebra and Taylor expansion, we have the mean, the variance, and the j^{th} autocovariance of y_t.

$$E(y_t) = \frac{\mu}{1-\alpha}, \tag{2}$$

$$\text{Var}(y_t) = \frac{\phi}{1-\alpha^2}, \tag{3}$$

$$\text{Cov}(y_t, y_{t-j}) = \alpha^j \frac{\phi}{1-\alpha^2} \tag{4}$$

So we have the following distribution of y_t

$$\begin{cases} y_1 \sim N(\frac{\mu}{1-\alpha}, \frac{\phi}{1-\alpha^2}), \\ y_t | y_{t-1} \sim N(\mu + \alpha\, y_{t-1}, \phi),\ (t > 1). \end{cases}$$

For multiple subjects, suppose there are N individuals, we can express the constant coefficient AR(1) model as follows:

$$y_{it} = \mu + \alpha\, y_{i(t-1)} + z_{it},\ (i = 1, ..., N;\ t = 2, ..., T)$$

where $z_{it}\ i.i.d. \sim N(0, \phi)$ and the parameters μ, α and ϕ are constants which keep the same values across all individuals. This model is very useful when the sample size (or number of participants) is small but with fairly large measurement occasions.

2.2 Categorical Data

Let c be the number of categories and $\boldsymbol{\tau} = (\tau_1, ..., \tau_{c-1})$ be thresholds. Assume y_{it} is a continuous normality distributed variable following AR(1) model $y_{it} = \mu + \alpha y_{i(t-1)} + z_{it}$. With the assumption of the normality distribution of y_t, the thresholds $\boldsymbol{\tau}$ can be created from the standardized thresholds $\boldsymbol{\tau}_z$ as

$$\boldsymbol{\tau} = \mu_y + \boldsymbol{\tau}_z \sigma_y$$

where $\mu_y = \frac{\mu}{1-\alpha}$ and $\sigma_y = \sqrt{\frac{\phi}{1-\alpha^2}}$. With thresholds $\boldsymbol{\tau}$, categorical data y_{it}^* can be created by

$$\begin{cases} y_{it}^* = 1, \text{ when } y_{it} \le \tau_1; \\ y_{it}^* = k, \text{ when } \tau_{k-1} < y_{it} \le \tau_k; \\ y_{it}^* = c, \text{ when } y_{it} > \tau_{c-1}. \end{cases}$$

The scale of y_{it}^* is from 1 to c. Let $\boldsymbol{\pi}$ be a c-dimensional vector $\boldsymbol{\pi} = (\pi_1, ..., \pi_c)$ which is defined as

$$\begin{aligned} \pi_1 &= \Phi(\tau_1) \\ \pi_k &= \Phi(\tau_k) - \Phi(\tau_{k-1}) \quad (2 \le k \le c-1) \\ \pi_c &= 1 - \Phi(\tau_{c-1}), \end{aligned}$$

then each π_k $(1 \le k \le c)$ is defined to be the probability of corresponding kth category. With $\boldsymbol{\pi}$, the mean of c categories is

$$M_c = \sum_{k=1}^{c} \pi_k \, k$$

Therefore, the true μ and ϕ of c categories are

$$\begin{aligned} \mu_c &= M_c \, (1-\alpha), \\ \phi_c &= \left[\sum_{k=1}^{c} \pi_k \, (k - M_c)^2 \right] (1-\alpha^2), \end{aligned}$$

3 Estimation Methods and Likelihoods

This study investigates two MLE estimation methods: (1) exact MLE estimation method: the parameters are estimated by maximizing the exact log-likelihood function including the distribution of deterministic y_1 which requires stationarity assumption, and (2) conditional MLE estimation method: the parameters are estimated by maximizing the conditional log-likelihood function without y_1. For multiple subjects we pool likelihood functions for all individuals. In practice, there is another method to deal with time series data by connecting all similar time series from multiple subjects together as from a single subject (reference here). It assumes there is some relationship between y_{iT} and $y_{(i+1)1}$. We have pooled data exact MLE and pooled data conditional MLE.

3.1 Exact MLE for Pooled Likelihood Function

The exact likelihood function of the stationary AR(1) model described in Eq. (1) and its corresponding log likelihood function are

$$L_i(\alpha, \mu, \phi|\mathbf{y}_i) = \frac{1}{\sqrt{2\pi(\frac{\phi}{1-\alpha^2})}} \exp\left[-\frac{(y_{i1} - \frac{\mu}{1-\alpha})^2}{2(\frac{\phi}{1-\alpha^2})}\right] \times \left\{\prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\phi}} \exp\left[-\frac{(y_{it} - \mu - \alpha\, y_{i(t-1)})^2}{2\phi}\right]\right\},$$

$$log(L) = \frac{N}{2} log(1-\alpha^2) - \frac{1-\alpha^2}{2\phi}\sum_{i=1}^{N}(y_{i1} - \frac{\mu}{1-\alpha})^2 - \frac{NT}{2} log(2\pi\phi) - \frac{1}{2\phi}\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{it} - \mu - \alpha\, y_{i(t-1)})^2.$$

In order to obtain the maximum likelihood estimates (MLE) of parameters μ, α and ϕ, we make all of their first order derivatives with respective to these parameters 0 and their corresponding second order derivatives negative. The MLE obtained through solving the exact likelihood function is called the exact MLE. Unfortunately, there is no simple solution for $\boldsymbol{\theta}$ in terms of $(\{y_{it}\}, 1 \le i \le N, 1 \le t \le T)$. But with the help of computers, we can use iterative or numerical procedures to solve the equation.

3.2 Exact MLE for Pooled Data

The exact likelihood function of the connected stationary AR(1) model and its corresponding log likelihood function are

$$L(\alpha, \mu, \phi|\mathbf{y}) = \frac{1}{\sqrt{2\pi(\frac{\phi}{1-\alpha^2})}} \exp\left[-\frac{(y_1 - \frac{\mu}{1-\alpha})^2}{2(\frac{\phi}{1-\alpha^2})}\right] \times \left\{\prod_{t=2}^{NT} \frac{1}{\sqrt{2\pi\phi}} \exp\left[-\frac{(y_t - \mu - \alpha\, y_{t-1})^2}{2\phi}\right]\right\},$$

$$log(L) = \frac{1}{2} log(1-\alpha^2) - \frac{1-\alpha^2}{2\phi}(y_1 - \frac{\mu}{1-\alpha})^2 - \frac{NT}{2} log(2\pi\phi) - \frac{1}{2\phi}\sum_{t=2}^{NT}(y_t - \mu - \alpha\, y_{t-1})^2.$$

By making the first derivatives zero equal to 0 to obtain the solution. Again, unfortunately, there is no simple solution for $\boldsymbol{\theta}$ in terms of $(\{y_{it}\}, 1 \leq i \leq N, 1 \leq t \leq T)$.

3.3 Conditional MLE for Pooled Likelihood Function

The conditional likelihood function of the stationary AR(1) model does not take the distribution of y_1 into consideration, so the likelihood and its log likelihood function are

$$L_i(\alpha, \mu, \phi|\mathbf{y}_i) = \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\phi}} \exp\left[-\frac{(y_{it} - \mu - \alpha\, y_{i(t-1)})^2}{2\phi}\right], \tag{5}$$

$$log(L(\alpha, \mu, \phi|\mathbf{y})) = -\frac{N(T-1)}{2} log(2\pi\phi) - \frac{1}{2\phi} \sum_{i=1}^{N} \sum_{t=2}^{T} (y_{it} - \mu - \alpha\, y_{i(t-1)})^2. \tag{6}$$

To obtain the MLE of parameters μ, α and ϕ, we make their first derivatives zero and the second order derivatives negative, and we have

$$\hat{\mu} = \frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=2}^{T} (y_{it} - \hat{\alpha}\, y_{i(t-1)}) \tag{7}$$

$$\hat{\alpha} = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} \left[(y_{it} - \hat{\mu}) y_{i(t-1)}\right]}{\sum_{i=1}^{N} \sum_{t=2}^{T} y_{i(t-1)}^2} \tag{8}$$

$$\hat{\phi} = \frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=2}^{T} (y_{it} - \hat{\mu} - \hat{\alpha}\, y_{i(t-1)})^2 \tag{9}$$

We can also use the ordinal least square (OLS) estimation method to obtain μ and α,

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\mu} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} N(T-1) & \sum_{i=1}^{N} \sum_{t=2}^{T} y_{i(t-1)} \\ \sum_{i=1}^{N} \sum_{t=2}^{T} y_{i(t-1)} & \sum_{i=1}^{N} \sum_{t=2}^{T} y_{i(t-1)}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{N} \sum_{t=2}^{T} y_{it} \\ \sum_{i=1}^{N} \sum_{t=2}^{T} y_{i(t-1)} y_{it} \end{bmatrix}.$$

And ϕ can be obtained by inserting the estimates of (μ, α) into Eq. (9). The OLS solution is exactly the same as the MLE solution.

3.4 Conditional MLE for Pooled Data

The conditional likelihood function and its corresponding log likelihood function of the connected stationary AR(1) model are

$$L(\alpha, \mu, \phi|\mathbf{y}) = \prod_{t=2}^{NT} \frac{1}{\sqrt{2\pi\phi}} \exp\left[-\frac{(y_t - \mu - \alpha\, y_{t-1})^2}{2\phi}\right],$$

$$log(L) \quad = -\frac{NT-1}{2} log(2\pi\phi) - \frac{1}{2\phi}\sum_{t=2}^{NT}(y_t - \mu - \alpha\, y_{t-1})^2.$$

4 Simulation Study

We conduct a simulation study to investigate the performance of the exact MLE estimation method and conditional MLE estimation method fitting different models fitting categorical data. We use iterative or numerical procedures to solve the equations which have no explicit solutions.

4.1 Data Generation

The true values in this simulation are set as $\mu = \mu_c$, $\alpha = 0.5$, $\phi = \phi_c$. The replication number is 1000. We use the following 3 steps to generate the categorical data $y{*}_{it}$.

Step 1: Generate the continuous data according to the constant coefficients AR(1) model $y_{it} = \mu + \alpha y_{i(t-1)} + z_{it}$.
Step 2: Generate thresholds $\boldsymbol{\tau} = (\tau_1, ..., \tau_{c-1})$. With the assumption of the normality distribution of y_t, the thresholds $\boldsymbol{\tau}$ are created by (1) obtaining the standardized thresholds $\boldsymbol{\tau}_z$ by dividing the segment $[-2, 2]$ into $c - 2$ parts evenly, and then (2) transform $\boldsymbol{\tau}_z$ to $\boldsymbol{\tau}$ according to the original data scale. For example, if $c = 5$, the standardized thresholds are $\boldsymbol{\tau}_z = (\tau_{z1}, \tau_{z2}, \tau_{z3}, \tau_{z4}) = (-2, -2/3, 2/3, 2)$, then $\boldsymbol{\tau} = \mu_y + \boldsymbol{\tau}_z\sigma_y$.
Step 3: Generate the categorical data y^*_{it} by

$$\begin{cases} y^*_{it} = 1, \text{ when } y_{it} \le \tau_1; \\ y^*_{it} = k, \text{ when } \tau_{k-1} < y_{it} \le \tau_k; \\ y^*_{it} = c, \text{ when} y_{it} > \tau_{c-1}. \end{cases}$$

Simulation condition factors in this study include the initial value, the number of categories, the lengths of series, and the number of subjects. (1) The initial value

y_1 has 3 cases: (i) a fixed yc_1 based on a fixed $y_1 = 0$; (ii) a random yc_1 based on a random y_1 from $N(0, \phi)$; and (iii) a random yc_1 based on a random y_1 from $N(\frac{\mu}{1-\alpha}, \frac{\phi}{1-\alpha^2})$. (2) The number of categories is $c = (5, 7, 9)$. (3) The lengths of series is set as $T = (5, 10, 15, 20, 30, 40, 50)$ to catch the change patterns. (4) The number of subjects is $N = (50, 100, 150, 200)$. In total, there are $3 * 3 * 7 * 4 = 252$ conditions, with each condition having 1000 replications.

4.2 Model Estimation and Evaluation

When the categorical data are ready, we use four estimation methods: pooled likelihood exact MLE, pooled likelihood conditional MLE, pooled data exact MLE, and pooled data conditional MLE. We use MSE, the mean square error of the estimate, to compare accuracy of estimates.

$$MSE = Bias.abs^2 + SE.emp^2$$

where $Bias.abs$ is the absolute bias of the estimate, and $SE.emp$ is the empirical standard error across 1000 replications.

R language was used to generate data, estimate parameters, and summarize results. The main R functions for data generating and model estimation are attached in Appendix 1.

5 Results, Conclusions and Discussion

5.1 Results

In total, there are 252 (= 3 initial values $\times$ 3 numbers of categories $\times$ 7 lengths of series $\times$ 4 number of subjects) simulation conditions. For each condition, there are 4 estimation methods. Part of simulation results are summarized and shown in Tables 1, 2, and 3. For example, Table 1 summarized part of the estimation results from 1000 replications with $c = 5$ categories, including sample size $N = 50$ or $N = 200$ individuals, time series length $T = 5$ or $T = 50$ observations per individual, and the initial value of y from $y_1 = 0$, $y_1 \sim N(0, \psi)$ or $y_1 \sim N(\frac{\mu}{1-\alpha}, \frac{\psi}{1-\alpha^2})$.

From Table 1, we see that the max value of MSE under all conditions of $N = 200$ and $T = 50$ is 0.0516 and the min value is 0.0009. But under the conditions with fewer individual participated $N = 50$ and shorter time series $T = 5$, the max and min values of MSE are 0.3596 and 0.0021, respectively. The smaller MSE value, the more accurate the estimate. So the longer the time series or the more individual participated, the more accurate the estimate. By comparing all three Tables Tables 1, 2, and 3 with difference categories, we can also see

Table 1 Results summarized from 1000 replications for $c = 5$ categories

				$N = 50$ $T = 5$							$N = 200$ $T = 50$					
			True[a]	Est.[b]	Bias.abs[c]	Bias.rel[d]	SE.emp[e]	SE.avg[f]	MSE[g]	Cover[h]	Est.	Bias.abs	Bias.rel	SE.emp	SE.avg	MSE
$y_1 = 0$																
P.L.[i]	Exact[j]	μ	1.5	2.0002	0.5002	0.3335	0.1847	0.1839	0.2844	0.2180	1.7245	0.2245	0.1496	0.0289	0.0279	0.0512
		α	0.5	0.3333	0.1667	0.3334	0.0602	0.0600	0.0314	0.2070	0.4251	0.0749	0.1497	0.0094	0.0090	0.0057
		ψ	0.4811	0.4136	0.0675	0.1404	0.0404	0.0372	0.0062	0.5250	0.5098	0.0287	0.0597	0.0068	0.0072	0.0009
	Cond.[k]	μ	1.5	1.7317	0.2317	0.1544	0.2360	0.2363	0.1094	0.8430	1.6985	0.1985	0.1324	0.0295	0.0286	0.0403
		α	0.5	0.4228	0.0772	0.1544	0.0766	0.0769	0.0118	0.8370	0.4338	0.0662	0.1324	0.0096	0.0092	0.0045
		ψ	0.4811	0.5123	0.0312	0.0648	0.0503	0.0512	0.0035	0.9390	0.5202	0.0391	0.0812	0.0069	0.0074	0.0016
P.D.[l]	Exact	μ	1.5	2.0737	0.5737	0.3824	0.1746	0.1848	0.3596	0.1060	1.7253	0.2253	0.1502	0.0291	0.0281	0.0516
		α	0.5	0.3090	0.1910	0.3820	0.0566	0.0600	0.0397	0.0920	0.4249	0.0751	0.1503	0.0095	0.0091	0.0057
		ψ	0.4811	0.4280	0.0531	0.1105	0.0419	0.0383	0.0046	0.6500	0.5121	0.0310	0.0645	0.0068	0.0072	0.0010
	Cond.	μ	1.5	2.0700	0.5700	0.3800	0.1753	0.1857	0.3556	0.1140	1.7252	0.2252	0.1501	0.0291	0.0281	0.0515
		α	0.5	0.3102	0.1898	0.3795	0.0568	0.0603	0.0392	0.1030	0.4249	0.0751	0.1502	0.0095	0.0091	0.0057
		ψ	0.4811	0.4297	0.0514	0.1069	0.0421	0.0385	0.0044	0.6690	0.5122	0.0311	0.0646	0.0068	0.0072	0.0010
$y_1 \sim N(0, \psi)$																
P.L.	Exact	μ	1.5	1.7948	0.2948	0.1966	0.1938	0.1883	0.1245	0.6750	1.7039	0.2039	0.1359	0.0287	0.0280	0.0424
		α	0.5	0.4027	0.0973	0.1945	0.0624	0.0612	0.0134	0.6660	0.4321	0.0679	0.1358	0.0093	0.0090	0.0047
		ψ	0.4811	0.4969	0.0157	0.0327	0.0435	0.0449	0.0021	0.9580	0.5181	0.0370	0.0769	0.0071	0.0073	0.0014
	Cond.	μ	1.5	1.7468	0.2468	0.1645	0.2089	0.2048	0.1045	0.7900	1.6985	0.1985	0.1323	0.0290	0.0283	0.0402
		α	0.5	0.4192	0.0808	0.1615	0.0669	0.0661	0.0110	0.7770	0.4339	0.0661	0.1322	0.0093	0.0091	0.0045
		ψ	0.4811	0.5153	0.0342	0.0711	0.0495	0.0516	0.0036	0.9430	0.5203	0.0392	0.0814	0.0073	0.0074	0.0016
P.D.	Exact	μ	1.5	2.0071	0.5071	0.3381	0.1864	0.1849	0.2920	0.2050	1.7247	0.2247	0.1498	0.0289	0.0281	0.0513
		α	0.5	0.3321	0.1679	0.3358	0.0593	0.0596	0.0317	0.1830	0.4252	0.0748	0.1496	0.0093	0.0091	0.0057
		ψ	0.4811	0.5306	0.0495	0.1029	0.0476	0.0475	0.0047	0.8620	0.5224	0.0413	0.0858	0.0072	0.0074	0.0018
	Cond.	μ	1.5	2.0068	0.5068	0.3379	0.1867	0.1851	0.2917	0.2070	1.7247	0.2247	0.1498	0.0289	0.0281	0.0513
		α	0.5	0.3322	0.1678	0.3355	0.0593	0.0597	0.0317	0.1880	0.4252	0.0748	0.1496	0.0093	0.0091	0.0057
		ψ	0.4811	0.5309	0.0498	0.1034	0.0477	0.0476	0.0048	0.8630	0.5224	0.0413	0.0858	0.0072	0.0074	0.0018

(continued)

Table 1 (continued)

				$N = 50$ $T = 5$								$N = 200$ $T = 50$					
			True[a]	Est.[b]	Bias.abs[c]	Bias.rel[d]	SE.emp[e]	SE.avg[f]	MSE[g]	Cover[h]	Est.	Bias.abs	Bias.rel	SE.emp	SE.avg	MSE
$y_1 \sim N(\frac{\mu}{1-\alpha}, \frac{\psi}{1-\alpha^2})$																
P.L.	Exact	μ	1.5	1.7243	0.2243	0.1496	0.1972	0.1877	0.0892	0.7930	1.6962	0.1962	0.1308	0.0285	0.0281	0.0393
		α	0.5	0.4252	0.0748	0.1497	0.0643	0.0611	0.0097	0.7900	0.4345	0.0655	0.1309	0.0093	0.0091	0.0044
		ψ	0.4811	0.5152	0.0341	0.0708	0.0460	0.0466	0.0033	0.9200	0.5198	0.0387	0.0804	0.0070	0.0074	0.0015
	Cond.	μ	1.5	1.7232	0.2232	0.1488	0.2068	0.1981	0.0926	0.8050	1.6961	0.1961	0.1308	0.0287	0.0282	0.0393
		α	0.5	0.4256	0.0744	0.1489	0.0665	0.0638	0.0100	0.7990	0.4346	0.0654	0.1309	0.0093	0.0091	0.0044
		ψ	0.4811	0.5140	0.0329	0.0684	0.0507	0.0514	0.0036	0.9360	0.5198	0.0387	0.0804	0.0071	0.0074	0.0015
P.D.	Exact	μ	1.5	1.9759	0.4759	0.3173	0.1868	0.1843	0.2614	0.2600	1.7221	0.2221	0.1481	0.0283	0.0281	0.0501
		α	0.5	0.3413	0.1587	0.3173	0.0604	0.0594	0.0288	0.2420	0.4259	0.0741	0.1482	0.0092	0.0090	0.0056
		ψ	0.4811	0.5567	0.0756	0.1572	0.0504	0.0498	0.0083	0.7090	0.5246	0.0435	0.0903	0.0072	0.0074	0.0019
	Cond.	μ	1.5	1.9760	0.4760	0.3173	0.1872	0.1844	0.2616	0.2610	1.7221	0.2221	0.1481	0.0283	0.0281	0.0501
		α	0.5	0.3413	0.1587	0.3174	0.0605	0.0594	0.0288	0.2450	0.4259	0.0741	0.1482	0.0092	0.0090	0.0056
		ψ	0.4811	0.5567	0.0755	0.1570	0.0505	0.0499	0.0083	0.7050	0.5246	0.0435	0.0903	0.0072	0.0074	0.0019

[a] True: True value of the corresponding parameter
[b] Est.: Average of the estimate of the corresponding parameter across 1000 replications
[c] Bias.abs: Absolute bias of the estimate
[d] Bias.rel: Relative bias of the estimate
[e] SE.emp: Empirical s.e. across 1000 replications
[f] SE.avg: Average of the s.e. obtained from the model
[g] MSE: Mean square error of the estimate, $MSE = Bias.abs^2 + SE.emp^2$
[h] Cover: Coverage probability of the estimate
[i] P.L.: Pooling likelihood functions
[j] Exact: Maximizing the exact likelihood function
[k] Cond.: Maximizing the conditional likelihood function
[l] P.D.: Pooling data by connecting data directly

a pattern that the more categories, the more accurate the estimate. Within each table, by comparing the pooled likelihood methods (P.L.) and the pooling data methods (P.D.), the MSE values obtained from P.L. are in general smaller than those obtained from P.D., which indicated that the P.L. methods perform better than the P.D. methods. We further compare the exact-likelihood estimation method and the conditional-likelihood estimation method within each table. For the case of random $y_1 \sim N(\frac{\mu}{1-\alpha}, \frac{\psi}{1-\alpha^2})$, the pooled likelihood exact MLE are the best. For the other two initial values, the pooled likelihood conditional MLE are the best. In other word, the pooled likelihood conditional MLE is not sensitive to initial values.

5.2 Conclusions

Through the simulation, we have the following conclusions: (1) The pooled likelihood methods perform better than the pooling data methods. (2) For the case of random $y_1 \sim N(\frac{\mu}{1-\alpha}, \frac{\psi}{1-\alpha^2})$, the pooled likelihood exact MLE are the best. For the other two initial values, the pooled likelihood conditional MLE are the best. In other word, the pooled likelihood conditional MLE is not sensitive to initial values. (3) The more categories, the more accurate the estimate. (4) The longer the time series, the more accurate the estimate. (5) The more individual participated, the more accurate the estimate.

5.3 Discussion

In this article, the intra-individual variation for multiple subjects with categorical outcomes are examined. In reality, when multiple subjects are involved, methods developed for single subject intra-individual relationship may not fully work. Also, laws governing inter-individual relationship may not apply to intra-individual relationship, especially when outcomes are categorical, which are very common in social and behaviorial fields. The categorical or ordinal data are usually collected by the Likert table. There are few methods in literature dealing with the analysis of pooling multiple time series. Fewer for categorical data. This article fill the gap by investigating the performance of four estimation methods for pooling time series data focusing on categorical outcomes and to address related issues through an AR(1) model.

Appendix 1

```
##-----------------------------------##
##     Data Generation Functions     ##
## Random y1~ N(mu/1-a, sig^2/1-a^2) ##
##-----------------------------------##

ar1.ranI2.sim <- function(T, a, mu, sig, nc){
   ymean <- mu/(1-a)
   ysig  <- sig/(sqrt(1-a^2))
   et    <- rnorm(T, 0, sig)
   y     <- rep(1,T)
   y[1]  <- rnorm(1,ymean,ysig)
   yc    <- rep(1,T)

   ##----------------------------##
   ## Categorical data
   ##----------------------------##
    th    <- seq(-2, 2, length=nc-1)*ysig + ymean  ## thresholds
    categ <- seq(1, nc)                            ## categories

    ## For yc[1] when time=1
        if (y[1] <= th[1])  {yc[1] <- categ[1]}          ## s=1
        if (y[1] > th[nc-1]){yc[1] <- categ[nc]}         ## s=nc
        for (s in 2:(nc-1)){
           if ((th[s-1] < y[1])&(y[1] <= th[s]))
                {yc[1] <- categ[s]} }

    ## For yc[2:T]
        for (i in 2:T){
           temp <- mu+a*y[i-1]+et[i]
           if (temp <= th[1])  {yc[i] <- categ[1]}    ## s=1
           if (temp > th[nc-1]){yc[i] <- categ[nc]}   ## s=nc
           for (s in 2:(nc-1)){
                if ((th[s-1] < temp)&(temp <= th[s]))
                    {yc[i] <- categ[s]} }
                }
        return(yc)}


##-----------------------------------##
##     Model Estimation Functions    ##
##    Exact-likelihood Estimation    ##
##-----------------------------------##
```

```
exactllike <- function(par, y){
        mu  <- par[1]
        a   <- par[2]
        psi <- par[3]
        N   <- nrow(y)
        T   <- ncol(y)
        Y1   <- as.vector(y[,1])
        Yt   <- as.vector(y[,2:T])
        Yt1  <- as.vector(y[,1:(T-1)])
        sum  <- (1-a^2)*t(Y1-mu/(1-a))%*%(Y1-mu/(1-a))
                + t(Yt-mu-a*Yt1)%*%(Yt-mu-a*Yt1)
        llik <- -.5*N*T*log(2*pi*psi) + .5*N*log(1-a^2) - sum/
        (2*psi)-llik}

##-----------------------------------##
##     Model Estimation Functions    ##
## conditional-likelihood Estimation ##
##-----------------------------------##
condllike <- function(par, y){
        mu  <- par[1]
        a   <- par[2]
        psi <- par[3]
        N   <- nrow(y)
        T   <- ncol(y)
        Yt   <- as.vector(y[,2:T])
        Yt1  <- as.vector(y[,1:(T-1)])
        sum  <- t(Yt-mu-a*Yt1)%*%(Yt-mu-a*Yt1)
        llik <- -.5*N*(T-1)*log(2*pi*psi) - sum/(2*psi)
        -llik}

##-----------------------------------------##
## Core Code in Main Program of Estimation ##
##-----------------------------------------##
result <- nlm(exactllike, c(0,1/2,1/2), y, hessian=T)
result <- nlm(condllike, c(0,1/2,1/2), y, hessian=T)
se     <- sqrt(diag(solve(result$hessian)))

y.conn   <- as.vector(t(y))
result   <- arima(y.conn, c(1,0,0), include.mean=T)
se       <- sqrt(diag(result$var.coef))
```

Appendix 2

Table 2 Results for $c = 7$ categories

				$N = 50$ $T = 5$							$N = 50$ $T = 5$						
			True	Est.	Bias.abs	Bias.rel	SE.emp	SE.avg	MSE	Cover	Est.	Bias.abs	Bias.rel	SE.emp	SE.avg	MSE	Cover
$y_1 = 0$																	
P.L.	Exact	μ	2	2.5456	0.5456	0.2728	0.2605	0.2427	0.3655	0.3930	2.1454	0.1454	0.0727	0.0379	0.0368	0.0226	0.0250
		α	0.5	0.3637	0.1363	0.2725	0.0624	0.0590	0.0225	0.3750	0.4636	0.0364	0.0727	0.0091	0.0088	0.0014	0.0220
		ψ	1.2053	0.9995	0.2058	0.1708	0.0957	0.0900	0.0515	0.3990	1.2230	0.0177	0.0147	0.0169	0.0173	0.0006	0.8370
	Cond.	μ	2	2.1549	0.1549	0.0775	0.3327	0.3135	0.1347	0.9160	2.1077	0.1077	0.0539	0.0387	0.0378	0.0131	0.1840
		α	0.5	0.4615	0.0385	0.0771	0.0792	0.0759	0.0078	0.9150	0.4731	0.0269	0.0539	0.0093	0.0090	0.0008	0.1440
		ψ	1.2053	1.2360	0.0307	0.0255	0.1190	0.1236	0.0151	0.9660	1.2478	0.0425	0.0353	0.0173	0.0178	0.0021	0.3150
P.D.	Exact	μ	2	2.6562	0.6562	0.3281	0.2461	0.2463	0.4912	0.2390	2.1465	0.1465	0.0733	0.0379	0.0371	0.0229	0.0250
		α	0.5	0.3360	0.1640	0.3280	0.0583	0.0594	0.0303	0.2070	0.4634	0.0366	0.0733	0.0091	0.0089	0.0014	0.0200
		ψ	1.2053	1.0425	0.1628	0.1350	0.1007	0.0933	0.0366	0.5590	1.2299	0.0246	0.0204	0.0170	0.0174	0.0009	0.7220
	Cond.	μ	2	2.6508	0.6508	0.3254	0.2472	0.2475	0.4846	0.2550	2.1464	0.1464	0.0732	0.0379	0.0371	0.0229	0.0250
		α	0.5	0.3373	0.1627	0.3254	0.0586	0.0597	0.0299	0.2190	0.4634	0.0366	0.0732	0.0091	0.0089	0.0014	0.0210
		ψ	1.2053	1.0467	0.1586	0.1316	0.1011	0.0938	0.0354	0.5690	1.2300	0.0247	0.0205	0.0170	0.0174	0.0009	0.7200
$y_1 \sim N(0, \psi)$																	
P.L.	Exact	μ	2	2.2378	0.2378	0.1189	0.2458	0.2459	0.1170	0.8650	2.1161	0.1161	0.0580	0.0370	0.0370	0.0148	0.1150
		α	0.5	0.4404	0.0596	0.1192	0.0586	0.0596	0.0070	0.8610	0.4709	0.0291	0.0582	0.0088	0.0088	0.0009	0.0850
		ψ	1.2053	1.1885	0.0168	0.0140	0.1005	0.1075	0.0104	0.9570	1.2408	0.0355	0.0295	0.0175	0.0176	0.0016	0.4610
	Cond.	μ	2	2.1587	0.1587	0.0794	0.2744	0.2715	0.1005	0.9170	2.1077	0.1077	0.0539	0.0374	0.0374	0.0130	0.1770
		α	0.5	0.4594	0.0406	0.0813	0.0639	0.0650	0.0057	0.9150	0.4730	0.0270	0.0541	0.0089	0.0089	0.0008	0.1480
		ψ	1.2053	1.2349	0.0296	0.0245	0.1183	0.1235	0.0149	0.9650	1.2463	0.0410	0.0340	0.0177	0.0178	0.0020	0.3690

P.D.	Exact	μ	2	2.5499	0.5499	0.2749	0.2449	0.2462	0.3624	0.4100	2.1461	0.1461	0.0730	0.0369	0.0372	0.0227	0.0170
		α	0.5	0.3623	0.1377	0.2755	0.0577	0.0589	0.0223	0.3550	0.4634	0.0366	0.0732	0.0088	0.0089	0.0014	0.0070
		ψ	1.2053	1.2892	0.0839	0.0696	0.1151	0.1153	0.0203	0.9200	1.2535	0.0482	0.0400	0.0177	0.0177	0.0026	0.2240
	Cond.	μ	2	2.5487	0.5487	0.2744	0.2452	0.2466	0.3612	0.4100	2.1460	0.1460	0.0730	0.0369	0.0372	0.0227	0.0160
		α	0.5	0.3625	0.1375	0.2750	0.0577	0.0590	0.0222	0.3610	0.4634	0.0366	0.0732	0.0088	0.0089	0.0014	0.0070
		ψ	1.2053	1.2901	0.0848	0.0703	0.1154	0.1156	0.0205	0.9200	1.2535	0.0482	0.0400	0.0177	0.0177	0.0026	0.2240
$y_1 \sim N(\frac{\mu}{1-\alpha}, \frac{\psi}{1-\alpha^2})$																	
P.L.	Exact	μ	2	2.1288	0.1288	0.0644	0.2544	0.2439	0.0813	0.9220	2.1083	0.1083	0.0541	0.0373	0.0371	0.0131	0.1530
		α	0.5	0.4673	0.0327	0.0654	0.0616	0.0591	0.0049	0.9310	0.4729	0.0271	0.0542	0.0090	0.0088	0.0008	0.1330
		ψ	1.2053	1.2343	0.0290	0.0241	0.1095	0.1118	0.0128	0.9630	1.2471	0.0418	0.0347	0.0165	0.0176	0.0020	0.3280
	Cond.	μ	2	2.1308	0.1308	0.0654	0.2714	0.2608	0.0908	0.9160	2.1083	0.1083	0.0542	0.0376	0.0373	0.0131	0.1630
		α	0.5	0.4669	0.0331	0.0662	0.0645	0.0623	0.0053	0.9200	0.4729	0.0271	0.0542	0.0090	0.0089	0.0008	0.1400
		ψ	1.2053	1.2293	0.0240	0.0199	0.1214	0.1230	0.0153	0.9600	1.2472	0.0419	0.0347	0.0166	0.0178	0.0020	0.3310
P.D.	Exact	μ	2	2.5118	0.5118	0.2559	0.2527	0.2456	0.3258	0.4720	2.1462	0.1462	0.0731	0.0375	0.0372	0.0228	0.0230
		α	0.5	0.3714	0.1286	0.2573	0.0603	0.0587	0.0202	0.4360	0.4634	0.0366	0.0732	0.0090	0.0089	0.0014	0.0160
		ψ	1.2053	1.3635	0.1582	0.1312	0.1223	0.1220	0.0400	0.7790	1.2614	0.0561	0.0466	0.0169	0.0178	0.0034	0.0940
	Cond.	μ	2	2.5116	0.5116	0.2558	0.2527	0.2459	0.3255	0.4780	2.1462	0.1462	0.0731	0.0375	0.0372	0.0228	0.0230
		α	0.5	0.3714	0.1286	0.2572	0.0602	0.0587	0.0202	0.4370	0.4634	0.0366	0.0732	0.0090	0.0089	0.0014	0.0170
		ψ	1.2053	1.3636	0.1583	0.1313	0.1225	0.1222	0.0401	0.7830	1.2614	0.0561	0.0466	0.0169	0.0178	0.0034	0.0940

Note: With the same notations as in Table 1

Table 3 Results for $c = 9$ categories

				$N = 50$ $T = 5$							$N = 50$ $T = 5$						
			True	Est.	Bias.abs	Bias.rel	SE.emp	SE.avg	MSE	Cover	Est.	Bias.abs	Bias.rel	SE.emp	SE.avg	MSE	Cover
$y_1 = 0$																	
P.L.	Exact	μ	2.5	3.1544	0.6544	0.2618	0.3066	0.3038	0.5222	0.4300	2.6259	0.1259	0.0503	0.0461	0.0461	0.0180	0.2160
		α	0.5	0.3684	0.1316	0.2633	0.0597	0.0588	0.0209	0.3940	0.4747	0.0253	0.0507	0.0087	0.0087	0.0007	0.1680
		ψ	2.2750	1.8701	0.4049	0.1780	0.1839	0.1684	0.1978	0.3500	2.2738	0.0012	0.0005	0.0305	0.0322	0.0009	0.9580
	Cond.	μ	2.5	2.6606	0.1606	0.0642	0.3922	0.3928	0.1796	0.9370	2.5777	0.0777	0.0311	0.0472	0.0473	0.0083	0.6250
		α	0.5	0.4668	0.0332	0.0664	0.0758	0.0756	0.0068	0.9290	0.4843	0.0157	0.0314	0.0089	0.0090	0.0003	0.5870
		ψ	2.2750	2.3115	0.0365	0.0160	0.2280	0.2312	0.0533	0.9540	2.3199	0.0449	0.0197	0.0311	0.0331	0.0030	0.7460
P.D.	Exact	μ	2.5	3.2921	0.7921	0.3168	0.2903	0.3093	0.7117	0.2620	2.6273	0.1273	0.0509	0.0463	0.0465	0.0184	0.2160
		α	0.5	0.3409	0.1591	0.3183	0.0557	0.0593	0.0284	0.2070	0.4744	0.0256	0.0513	0.0087	0.0088	0.0007	0.1720
		ψ	2.2750	1.9523	0.3227	0.1418	0.1930	0.1747	0.1414	0.5130	2.2874	0.0124	0.0054	0.0307	0.0324	0.0011	0.9510
	Cond.	μ	2.5	3.2853	0.7853	0.3141	0.2915	0.3107	0.7016	0.2750	2.6271	0.1271	0.0508	0.0463	0.0465	0.0183	0.2190
		α	0.5	0.3422	0.1578	0.3155	0.0559	0.0596	0.0280	0.2170	0.4744	0.0256	0.0512	0.0087	0.0088	0.0007	0.1740
		ψ	2.2750	1.9601	0.3149	0.1384	0.1938	0.1757	0.1367	0.5280	2.2876	0.0126	0.0055	0.0307	0.0324	0.0011	0.9500
$y_1 \sim N(0, \psi)$																	
P.L.	Exact	μ	2.5	2.7244	0.2244	0.0898	0.3084	0.3059	0.1455	0.9160	2.5894	0.0894	0.0357	0.0474	0.0463	0.0102	0.5060
		α	0.5	0.4561	0.0439	0.0877	0.0596	0.0589	0.0055	0.9000	0.4822	0.0178	0.0356	0.0090	0.0088	0.0004	0.4690
		ψ	2.2750	2.2075	0.0675	0.0297	0.1886	0.1998	0.0401	0.9330	2.3106	0.0356	0.0157	0.0302	0.0327	0.0022	0.8350
	Cond.	μ	2.5	2.6261	0.1261	0.0504	0.3371	0.3399	0.1295	0.9470	2.5785	0.0785	0.0314	0.0481	0.0469	0.0085	0.6190
		α	0.5	0.4757	0.0243	0.0486	0.0642	0.0644	0.0047	0.9470	0.4843	0.0157	0.0313	0.0091	0.0089	0.0003	0.5810
		ψ	2.2750	2.2932	0.0182	0.0080	0.2134	0.2294	0.0459	0.9640	2.3210	0.0460	0.0202	0.0306	0.0332	0.0030	0.7410
P.D.	Exact	μ	2.5	3.1381	0.6381	0.2552	0.3117	0.3095	0.5044	0.4760	2.6278	0.1278	0.0511	0.0477	0.0466	0.0186	0.2320
		α	0.5	0.3735	0.1265	0.2530	0.0594	0.0586	0.0195	0.4470	0.4745	0.0255	0.0510	0.0091	0.0088	0.0007	0.1860
		ψ	2.2750	2.4139	0.1389	0.0611	0.2081	0.2160	0.0626	0.9340	2.3357	0.0607	0.0267	0.0307	0.0330	0.0046	0.5610
	Cond.	μ	2.5	3.1368	0.6368	0.2547	0.3122	0.3099	0.5031	0.4750	2.6277	0.1277	0.0511	0.0477	0.0466	0.0186	0.2310
		α	0.5	0.3737	0.1263	0.2526	0.0594	0.0587	0.0195	0.4480	0.4745	0.0255	0.0510	0.0091	0.0088	0.0007	0.1870
		ψ	2.2750	2.4151	0.1401	0.0616	0.2081	0.2165	0.0629	0.9370	2.3358	0.0608	0.0267	0.0307	0.0330	0.0046	0.5600

$y_1 \sim N(\frac{\mu}{1-\alpha}, \frac{\psi}{1-\alpha^2})$																	
P.L.	Exact	μ	2.5	2.6150	0.1150	0.0460	0.3038	0.3045	0.1055	0.9470	2.5760	0.0760	0.0304	0.0461	0.0464	0.0079	0.6230
		α	0.5	0.4773	0.0227	0.0454	0.0580	0.0586	0.0039	0.9520	0.4848	0.0152	0.0305	0.0087	0.0088	0.0003	0.5950
		ψ	2.2750	2.3029	0.0279	0.0123	0.1964	0.2088	0.0394	0.9560	2.3183	0.0433	0.0191	0.0301	0.0328	0.0028	0.7660
	Cond.	μ	2.5	2.6147	0.1147	0.0459	0.3191	0.3280	0.1150	0.9460	2.5759	0.0759	0.0304	0.0466	0.0468	0.0079	0.6340
		α	0.5	0.4773	0.0227	0.0454	0.0598	0.0620	0.0041	0.9530	0.4848	0.0152	0.0304	0.0088	0.0088	0.0003	0.5970
		ψ	2.2750	2.2954	0.0204	0.0089	0.2152	0.2296	0.0467	0.9570	2.3184	0.0434	0.0191	0.0304	0.0331	0.0028	0.7720
P.D.	Exact	μ	2.5	3.0991	0.5991	0.2396	0.3099	0.3092	0.4550	0.5180	2.6247	0.1247	0.0499	0.0461	0.0466	0.0177	0.2370
		α	0.5	0.3804	0.1196	0.2391	0.0582	0.0584	0.0177	0.4780	0.4750	0.0250	0.0499	0.0087	0.0088	0.0007	0.1910
		ψ	2.2750	2.5538	0.2788	0.1226	0.2207	0.2285	0.1265	0.8320	2.3467	0.0717	0.0315	0.0304	0.0332	0.0061	0.4010
	Cond.	μ	2.5	3.0995	0.5995	0.2398	0.3093	0.3096	0.4551	0.5190	2.6247	0.1247	0.0499	0.0461	0.0466	0.0177	0.2370
		α	0.5	0.3804	0.1196	0.2391	0.0582	0.0585	0.0177	0.4780	0.4750	0.0250	0.0499	0.0087	0.0088	0.0007	0.1910
		ψ	2.2750	2.5540	0.2790	0.1226	0.2208	0.2289	0.1266	0.8300	2.3467	0.0717	0.0315	0.0304	0.0332	0.0061	0.3970

Note: With the same notations as in Table 1

References

Cattell, R. B. (1952). The three basic factor-analytic research designs-their interrelations and derivatives. *Psychological Bulletin, 49*, 499–520.

Cattell, R. B., Cattell, A. K. S., & Rhymer, R. M. (1947). P-technique demonstrated in determining psychophysical source traits in a normal individual. *Psychometrika, 12*, 267–288.

Cattell, R. B., & Scheier, I. H. (1961). *The meaning and measurement of neuroticism and anxiety*. New York: Ronald Press.

Daly, D. L., Bath, K. E., & Nesselroade, J. R. (1974). On the confounding of interand intraindividual variability in examining change patterns. *Journal of Clinical Psychology, 30*, 33–36.

Molenaar, P. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika, 50*, 181–202.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology – this time forever. *Measurement: Interdisciplinary Research and Perspectives, 2*, 201–218.

Molenaar, P. C. M., Huizenga, H. M., & Nesselroade, J. R. (2003). The relationship between the structure of interindividual and intraindividual variability: A theoretical and empirical vindication of developmental systems theory. In U. M. Staudinger & U. Lindenberger (Eds.), *Understanding human development: Dialogues with lifespan psychology* (pp. 339–360). Norwell: Kluwer.

Nesselroade, J. R., & Molenaar, P. (2003). Quantitative models for developmental processes. In J. Valsiner & K. Connolly (Eds.), *Handbook of developmental psychology* (pp. 622–639). London: Sage.

Nesselroade, J. R., & Molenaar, P. C. M. (1999). Pooling lagged covariance structures based on short, multivariate time-series for dynamic factor analysis. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 224–250). Newbury Park: Sage.

Nesselroade, J. R., & Ram, N. (2004). Studying intraindividual variability: What we have learned that will help us understand lives in context. *Research in Human Development, 1*, 9–29.