Estimation of the Serially Correlated Oneway Random Effect Model

Byoung Cheol Jung
Ewha Womans University, Department of Statistics,
Daehyun-Dong Seodamun-Ku, Seoul, Korea
bcjung@heitel.net

Juneyoung Lee
Korea University, Department of Preventive Medicine
Anam-Dong, Sungbuk-Ku, Seoul, Korea
jyleeuf@mail.korea.ac.kr

Seuck Heun Song
Korea University, Department of Statistics
Anam-Dong, Sungbuk-Ku, Seoul, Korea
ssong@korea.ac.kr

1. Introduction

The analysis of longitudinal data or repeated measurement data by using the variance components model is widely used in statistics and biometrics. In some instant, it is unavoidable to have an unbalanced data. For example, subjects are observed at different time points. Moreover, in certain situation, we may need to allow the remainder disturbances to be serially correlated. In this paper, three estimation methods for the between group variance component as well as the serial correlation coefficient for serially correlated oneway random model are investigated. To compare their estimation capability, three designs having different degrees of unbalancedness are considered. The so-called empirical quantile dispersion graphs(EQDGs) proposed by Lee and Khuri (1999) are used to compare estimation methods as well as designs.

2. Model

Consider a serially correlated unbalanced one-way random model,

$$\begin{aligned} y_{it} &= \mu + u_{it} \\ &= \mu + \alpha_i + \varepsilon_{it}, \quad i = 1, \cdots, N, \text{ and } t = 1, \cdots, T_i, \end{aligned} \tag{1}$$

where μ is a fixed unknown parameter, u_{it} is an overall disturbance, α_i is an effect of level i of a treatment factor which is distributed as $i.i.d\ N(0, \sigma_\alpha^2)$, and ε_{it} is a remainder random error which follows a stationary AR(1) process such that $\varepsilon_{it} = \rho\varepsilon_{i,t-1} + e_{it}$ with $|\rho| < 1$. In here, ρ represents a serial correlation coefficient and e_{it} has $i.i.d$ $N(0, \sigma_e^2)$. The α_i and e_{it} are independently distributed.

3. Estimation Methods

(1) Conditional ANOVA estimator
(2) ML estimator
(3) REML estimator.

4. Comparison of Estimators with Monte Carlo Study

In this section we compare estimation methods described in previous section with three selected designs. In model (1), without loss of generality, we fix $\mu = 5$. By letting $\phi = \sigma_\alpha^2/\sigma_e^2$, model (1) can then be written as

$$y_{it} = 5 + \alpha_i + \varepsilon_{it}, \qquad i = 1, \cdots, N, \quad t = 1, \cdots, T_i, \tag{2}$$

where $\alpha_i \sim iid\ N(0, \phi\sigma_e^2)$ and ε_{it} follows a stationary AR(1) process, such that $\varepsilon_{it} = \rho\varepsilon_{i,t-1} + e_{it}$ with $|\rho| < 1$ where $e_{it} \sim iid\ N(0, \sigma_e^2)$. In order to compare estimation capabilities of methods, three designs, namely, a balanced design and two unbalanced designs, one being moderately and the other being severely unbalanced, are considered.

5. Conclusion

The comparison is made based on the empirical quantile dispersion graphs(EQDGs) using quantiles of estimates. Results show that the proposed conditional ANOVA estimation is robust for design imbalance. However, ML estimate is preferred to other estimates regardless of design imbalance in the sense that ML estimation provides more stable and less variable estimates of a variance component than both an ANOVA type and REML estimation. Moreover, similar stability of ML and REML estimates is achieved unless a design is severely unbalanced.

REFERENCES

Lee, J. and Khuri, A.I., (1999), Graphical technique for comparing designs for random models, *Journal of Applied Statistics* 26, 933-947.