ISSN: 2688-9323

Empirical Likelihood for Line Regression Model with Ranked Set Sampling under Dependent Sample

Naiyi Li[1], Yongming Li[2], Juan Huang[1,*]

[1]School of Mathematics and Computer, Guangdong Ocean University, Zhanjiang 524088, China

[2]School of Mathematics and Computer Science, Shang Rao Normal University Shangrao, 334001, China

Abstract

Applying block empirical likelihood method which are many obvious advantages, we discuss the line regression model with dependent sample and ranked set sampling. We can obtain the good large samples property under the regular condition. It has important theoretical value and extensive academic value to popularize the research results of predecessors.

Keywords

Dependent Sample; Ranked Set Sampling; Line Regression Model; Empirical Likelihood.

1. Introduction

Owen proposed empirical likelihood method was which has good characteristics. Because it is not necessary to estimate the variance and its shape is determined by the data, which has been used in some aspects[1-11].The method is used to the problem. In this paper, by empirical likelihood inference under dependent and ranked set sampling, we establish the large samples property for quantile of the population.

Assume that:

$X_{1,1}, Y_{1,1}, X_{1,2}, Y_{1,2} \cdots, X_{1,m}, Y_{1,m}, X_{2,1}, Y_{2,1} X_{2,2}, Y_{2,2} \cdots, X_{2,m}, Y_{2,m}, \cdots, X_{k,1}, Y_{k,1} X_{k,2}, Y_{k,2} \cdots, X_{k,m}, Y_{k,m}$ be α -mixed random sample, which satisfied the linear models:

$Y_{i,j} = X_{i,j}\beta + \xi_{i,j}$ Where $E(\xi_{i,j})=0, D(\xi_{i,j}) = \sigma^2$

$$\text{Write } Z_{ij} = X_{i,j}^T (Y_{i,j} - X_{i,j}\beta), \quad i = 1, 2, \cdots, k. j = 1, 2, \cdots, m.$$

empirical likelihood is established by

$$R(\beta) = \sup\left\{ \prod_{i=1}^{k}\prod_{j=1}^{m} kmp, p_{ij} \ge 0, \sum_{i=1}^{k}\sum_{j=1}^{m} p_{ij} = 1, \sum_{i=1}^{k} km \sum_{j=1}^{m} p_{ij} Z_{ij} = 0 \right\}.$$

empirical likelihood ratio is established by

$$l(\beta) = -2\log R(\beta) = 2\sum_{i=1}^{k}\sum_{j=1}^{m} \log(1 + sZ_{ij}).$$

where $s \in R$, s is determined by

$$\Pi(s) = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{m} \frac{Z_{ij}}{1 + sZ_{ij}} = 0$$

2. Theorem

2.1. Condition

2.1.1. Sub-section Headings

(a) Let $X_{1,1}, Y_{1,1}, X_{1,2}, Y_{1,2} \cdots, X_{1,m}, Y_{1,m}, X_{2,1}, Y_{2,1} X_{2,2}, Y_{2,2} \cdots, X_{2,m}, Y_{2,m}$, be α -mixed random sample and ranked set sampling;

(b) mixing coefficient satisfies: $\sum_{i=1}^{\infty} \alpha^{\frac{1}{2}}(i) < \infty$;

Theorem If the above conditions are established, we get

$$l(\beta) \to_d \frac{\sigma_1^2}{\sigma_2^2} \chi_{(1)}^2, \quad m \to \infty .$$

where

$$\sigma_1^2 = \sum_{i=1}^{k} Var\{Z_i\} + 2\sum_{i=1}^{k}\sum_{j=1}^{m} Cov(Z_{i,j}, Z_{i,j+1}),$$

$$\sigma_3^2 = \sum_{i=1}^{k} Var\{Z_i\}.$$

2.2. Proof of Theorem

Because

$$P\{Z_{11} < 0\} \geq c > 0, P\{Z_{11} > 0\} \geq c > 0, \tag{1}$$

we show that

0 is the set of convex hull $\{Z_{11}, \cdots, Z_{km}\}$, and

$$R(F(x)) = \sup\left\{R(F) \mid \int Z_{ij} dF = 0, F \ll F_n\right\} \text{ exist as a positive.} \tag{2}$$

It established that

$$R(F(x)) = \sup \prod_{i=1}^{k} \prod_{j=1}^{m} kmp_{ij} \tag{3}$$

where $p_{ij} \geq 0, \sum_{i=1}^{k}\sum_{j=1}^{m} p_{ij} = 1, \sum_{i=1}^{k}\sum_{j=1}^{m} p_{ij} Z_{ij} = 0$.

Applying Lagrange multiplier method, we have

$$p_{ij} = \frac{1}{n(1 + sZ_{ij})}, 1 \leq i \leq k, 1 \leq j \leq m. \tag{4}$$

where s$\in R^1$, s is determined by

$$\Pi(s) = \frac{1}{km} \sum_{i=1}^{k}\sum_{j=1}^{m} \frac{Z_{ij}}{1 + sZ_{ij}} = 0.$$

$$0 = |\Pi(s)|$$

$$\geq \frac{|s| \frac{1}{km} \sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij}^{2}}{1+|s| \Delta_n} - \left| \frac{1}{km} \sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij} \right| \tag{5}$$

$$\frac{|s|}{1+|s| \Delta_n} = \mathrm{O}_p\left(\frac{1}{\sqrt{n}}\right).$$

we get

$$s = \mathrm{O}_p\left(\frac{1}{\sqrt{n}}\right). \tag{6}$$

Put $\gamma_{ij} = sZ_{ij}$, and s is determined by $\Lambda(s) = 0$.
By (6) , we have

$$\max_{\substack{1 \le i \le k \\ 1 \le j \le m}} |\gamma_{ij}| = o_p(1). \tag{7}$$

$$\begin{aligned} 0 &= \Pi(s) \\ &= \frac{1}{km} \sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij} - \frac{1}{km} \sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij}^2 \\ &+ \frac{1}{km} \sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij} \frac{\gamma_{ij}^2}{1+\gamma_{ij}}. \end{aligned}$$

Put $s = \dfrac{\frac{1}{km} \sum_{i=1}^{k} \sum_{j=1}^{m} \Upsilon_{ij}}{\frac{1}{km} \sum_{i=1}^{k} \sum_{j=1}^{m} \Upsilon_{ij}^{2}} + \psi,$

Applying (6)and(7), we have

$$\psi = \frac{\frac{1}{km} \sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij}^{3} s^{2}}{\frac{1}{km} \sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij}^{2} (1+\gamma_{ij})} = o_p\left(\frac{1}{\sqrt{n}}\right).$$

By Taylor expansion, We obtain

$$l(\beta) = -2\log R(\beta) = 2\sum_{i=1}^{k} \sum_{j=1}^{m} \log(1+\gamma_{ij}) = 2\sum_{i=1}^{k} \sum_{j=1}^{m} \gamma_{ij} - \sum_{i=1}^{k} \sum_{j=1}^{m} \gamma_{ij}^{2} + 2\sum_{i=1}^{k} \sum_{j=1}^{m} \eta_{ij}$$

$$= \frac{\frac{2}{km} \left(\sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij}\right)^2}{\frac{1}{km} \sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij}^{2}} + 2\beta \sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij} - \frac{\frac{1}{km} \left(\sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij}\right)^2}{\sum_{i=1}^{k} \sum_{j=1}^{m} Z_{ij}^{2}}$$

$$=\frac{km(\frac{1}{km}\sum_{i=1}^{k}\sum_{j=1}^{m}Z_{ij})^2}{\frac{1}{km}\sum_{i=1}^{k}\sum_{j=1}^{m}Z_{ij}^2}-km\beta^2\frac{1}{km}\sum_{i=1}^{k}\sum_{j=1}^{m}Z_{ij}^2+2\sum_{i=1}^{k}\sum_{j=1}^{m}\eta_{ij}$$

$$\hat{=}Z_1+Z_2+Z_3 \quad (8)$$

$$Z_1\to_d\frac{\sigma_1^2}{\sigma_3^2}\chi_{(1)}^2,\quad n\to\infty. \quad (9)$$

$$Z_2=o_p(1). \quad (10)$$

Applying (6), we obtain

$$|2\sum_{i=1}^{k}\sum_{j=1}^{m}\eta_{ij}|\le 2A|s|^3|r_{ij}|^3=o_p(1).$$

We show that

$$Z_3=o_p(1). \quad (11)$$

By(9)-(11), we obtain

$$l(\beta)\to_d\frac{\sigma_1^2}{\sigma_3^2}\chi_{(1)}^2,\ m\to\infty. \quad (12)$$

3. Conclusion

Using empirical likelihood method, we discuss quantile of population with under dependent and ranked set sampling. We obtain the good large property.

Acknowledgments

This paper was financially supported by Natural Science Foundation of China (12161074), Natural Science Foundation of Guangdong Province (2022A1515010978) Natural Science Foundation of Jiangxi Province (20122ABC201006) and Natural Science Foundation of Guangdong Ocean University (R17083, C17201, P16091).

References

[1] Wang Q, Jing B. Empirical likelihood for a class of functionals of survival distribution with censored data[J]. Annals of the Institute of Statistical Mathematics, 2001, 53(3):517-527.

[2] Wang Q, Li G. Empirical likelihood semiparametric regression analysis under random censorship[J]. Journal of Multivariate Analysis, 2002, 83(2):469-486.

[3] Li G, Wang Q. Empirical likelihood regression analysis for right censored data[J]. Statistica Sinica, 2003, 13(1):51-68.

[4] Qin G, Min T. Empirical likelihood inference for median regression models for censored survival data[J]. Journal of Multivariate Analysis, 2003, 85(2):416-430.

[5] Jing B, Yuan J, Zhou W. Jackknife empirical likelihood[J]. Journal of American Statistical Association, 2009, 104 (487):1224-1232.

ISSN: 2688-9323

[6] Zhao Y. Semiparametric inference for transformation models via empirical likelihood [J]. Journal of multivariate analysis, 2010, 101(8):1846-1858.

[7] Yu W, Sun Y, Zheng M. Empirical likelihood method for linear transformation models[J]. Annals of the Institute of Statistical Mathematics, 2011, 63(2):331-346.

[8] He S, Liang W, Shen J, Yang G. Empirical likelihood for right censored lifetime data[J]. Statistical Papers, 2012, 55(3):827-839.

[9] George B, Seals S, Aban I. Survival analysis and regression models[J]. Journal of Nuclear Cardiology, 2014, 21(4):686-694.

[10] Bouadoumou M, Zhao Y, Lu Y. Jackknife empirical likelihood for the accelerated failure time model with censored Data[J]. Communications in Statistics-Simulation and Computation, 2015, 44 (7): 1818-1832.

[11] Zhou M. Empirical likelihood method in survival analysis[M]. Taylor & Francis, 2016.