Empirical Likelihood for Difference of Two Linear Model with Mixing dependent Samples

Naiyi Li [1, 2, a,*], Yuan Li [1, b], Yongming Li [3, c] and Chengdong Wei [4, d]

[1]School of Economics and Statistics, Guangzhou University, Guangzhou, China.

[2]School of science, Guangdong Ocean University, Zhanjiang, China.

[3]School of Mathematics and Computer Science, Shang Rao Normal University, Shangrao, China.

[4]College of Mathematics and Statistics, Guangxi Teachers Education University, Nanning, China.

[a]linaiyi1979@163.com, [b]mathly@gzhu.edu.cn, [c]lym1019@163.com, [d]wcdbbh@163.com

Abstract

Applying empirical likelihood method , we discuss difference of two linear model with strongly stationary $\phi-$ mixing sample. We obtain the large samples property.

Keywords

Mixing sample; two linear model; empirical likelihood.

1. Introduction

Owen ([3-4]) firstly proposed empirical likelihood. It has good characteristics. For example, the empirical likelihood method is not necessary to estimate the variance and its shape is determined by the data itself and so on. So, it has been widely applied in many fields.

The empirical likelihood method is applied to the two population problem ([6-8]). Using empirical likelihood method with $\phi-$mixing dependent samples, In this paper, we obtain the large samples property for difference of two linear model.

Definition[1] The random variable sequences are ϕ-mixed. If there exists a non increasing positive sequence $\{\phi(n)\mid n\in N\}$, $\lim\limits_{n\to\infty}\varphi(n)=0$,for $n\in N, i\geq 1$,we obtain

$$|P(B)-P(B\mid \mathrm{A})|\leq \phi(n)$$

where $A\in F_1^i, B\in F_{i+n}^{\infty}, F_j^m=\sigma\{X_i\mid j\leq i\leq m\}$.

Consider the two linear model:

$$Y_i=X_i\beta_1+\varepsilon_i,\ i=1,\cdots,n$$

$$Y_i^{'}=X_i\beta_2+\zeta_i,\ i=1,\cdots,n$$

Where $E\varepsilon_i=0$, $E\zeta_i=0$, $D\varepsilon_i=\sigma_1^2>0,$, $D\zeta_i=\sigma_2^2>0$, ε_i and ζ_i are independent.

Let $(X_1,Y_1),(X_2,Y_2),\cdots,(X_n,Y_n)$ ($n\geq 2$) be ϕ -mixed. random sample. Let $(X_1,Y_1^{'}),(X_2,Y_2^{'}),\cdots,(X_n,Y_n^{'})$ $(m\geq 2)$ be ϕ-mixed random sample .

Denote that $\nabla\beta=\beta_2-\beta_1$.

Put

$$\Theta_i=X_i(Y_i^{'}-Y_i-X_i^{'}\nabla\beta),\quad i=1,2\cdots,n$$

Empirical likelihood is obtained by

$$R(\nabla\beta)=\sup\left\{\prod_{i=1}^n nw_i, w_i\geq 0,\sum_{i=1}^n w_i=1,\sum_{i=1}^n w_i\Theta_i=0\right\}$$

empirical likelihood ratio is obtained by

$$l(\nabla\beta) = -2\log R(\nabla\beta) = 2\sum_{i=1}^{n}\log(1+s\Theta_i).$$

where $s \in R^1$, s is determined by $\Pi(s) = \frac{1}{n}\sum_{i=1}^{n}\frac{\Theta_i}{1+s\Theta_i} = 0$.

2. The main conclusions and proofs Condition

Let $(X_1, Y_1, Y_1^{'}), (X_2, Y_2, Y_2^{'}), \cdots, (X_n, Y_n, Y_n^{'})$ $(n \geq 2)$ be ϕ-mixed. random sample;

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

$E|X|^r < \infty$, where $r > 2$;

$E\varepsilon_i = 0$, $E\zeta_i = 0$, $E|\varepsilon_i|^r < \infty$, $E|\zeta_i|^r < \infty$, ε_i and ζ_i are independent, where $r > 2$.

Theorem 1 If the above conditions are established, we obtained

$$l(\nabla\beta) \rightarrow_d \frac{\sigma_3^2}{\sigma_4^2}\chi_{(1)}^2, n \rightarrow \infty.$$

where

$$\sigma_3^2 = Var\{\Theta_i\} + 2\sum_{i=1}^{\infty}Cov(\Theta_i, \Theta_{i+1}),$$

$$\sigma_4^2 = Var\{\Theta_i\}.$$

σ_3^2 and σ_4^2 aren't known. We apply the following method of empirical likelihood to obtain new limit distribution which don't contain unknown parameters.

Write $u = [n^{\gamma}]$, $g = \left[\frac{n}{2u}\right]$, where $[\cdot]$ is the integral function, $0 < \gamma < \frac{1}{2}$, $n = 2ug$.

Write

$$\rho_i = \sum_{j=1}^{u} v_{2(i-1)u+j} \qquad \xi_i = \sum_{j=1}^{u} v_{(2i-1)u+j}$$

$$\mathrm{H}_{2i-1} = \frac{\rho_i}{u}, Y_{2i} = \frac{\xi_i}{u} \qquad (\text{for } i = 1, 2, \cdots, g)$$

The empirical likelihood ratio is obtained by

$$R^{'}(\nabla\beta) = \sup\left\{\prod_{i=1}^{2g}2gP_i^{'} \mid \sum_{i=1}^{2g}P_i^{'} = 1, P_i^{'} \geq 0, \sum_{i=1}^{2g}P_i^{'}\sqrt{u}\mathrm{H}_i = 0\right\}.$$

Log empirical likelihood ratio is obtained by

$$l^{'}(\nabla\beta) = -2\log R^{'}(\nabla\beta) = 2\sum_{i=1}^{2g}\log(1+\lambda\sqrt{u}\mathrm{H}_i).$$

where $\lambda \in R^1$, λ is determined by $\Pi^{'}(\lambda) = \frac{1}{2g}\sum_{i=1}^{2g}\frac{\sqrt{u}\mathrm{H}_i}{1+\lambda\sqrt{u}\mathrm{H}_i} = 0$.

Theorem 2 Under the conditions of Theorem 1, we obtain

$$l^{'}(\nabla\beta) \rightarrow_d \chi_{(1)}^2, \ n \rightarrow \infty.$$

Lemma 1 [5] Write $\Delta_n = \max_{1\le i\le n} |\Theta_i|$, we obtain $\Delta_n = O_p(n^{1/2})$.

Lemma2 [2] Assume that $\{X_i \mid i\ge 1\}$are a strongly stationary $\phi-$mixing sequence, $\sum_{i=1}^{\infty}\phi^{\frac{1}{2}}(i)<\infty$, Assume that $EX_1 = 0, E|X_1|^r < \infty$, where $r\ge 2$,we obtain

$$E|\sum_{i=1}^{n} X_i|^r \le cn^{\frac{r}{2}-1}\sum_{i=1}^{n} E|X_i|^r$$

Lemma 3 [9] Assume that $\{X_i \mid i\ge 1\}$are strong stationary $\phi-$ mixing sequence, $\sum_{i=1}^{\infty}\phi^{\frac{1}{2}}(i)<\infty$, Assumed $EX_1 = 0, E|X_1|^2 < \infty$, we obtain $A_0^2 = EX_1^2 + 2\sum_{i=1}^{\infty} E(X_1 X_{1+i})$ is convergent,

$$\sup_{-\infty<x<\infty} |P(\frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i}{A_0} < x) - \Phi(x)| \to 0,$$

where $\Phi(x)$ is the standard normal distribution.

Lemma 4[5] Let $(X_1,Y_1,Y_1^{'}),(X_2,Y_2,Y_2^{'}),\cdots,(X_n,Y_n,Y_n^{'})$ be ϕ -mixed. random sample. If the conditions of theorem are established, we obtain

$$\frac{1}{2g}\sum_{i=1}^{2g} u(\mathrm{H}_i)^2 = \sigma_3^2 + o_p(1).$$

Proof of theorem 1

Since

$$P\{\Theta_1 < 0\} \ge c > 0, P\{\Theta_1 > 0\} \ge c > 0, \tag{1}$$

we obtain that 0 is the set of convex hull $\{\Theta_1,\cdots,\Theta_n\}$, and

$$R(F(x)) = \sup\{R(F) \mid \int \Theta_i dF = 0, F 《F_n\} \text{ exist as a positive.} \tag{2}$$

$$\text{It sees that } R(F(x)) = \sup\prod_{i=1}^{n} nw_i, \tag{3}$$

where $w_i \ge 0, \sum_{i=1}^{n} w_i = 1, \sum_{i=1}^{n} w_i\Theta_i = 0$.

By Lagrange multiplier method, we obtain

$$w_i = \frac{1}{n(1+s\Theta_i)}, 1\le i\le n. \tag{4}$$

where $s\in R^1$, s is determined by $\Pi(s) = \frac{1}{n}\sum_{i=1}^{n}\frac{\Theta_i}{1+s\Theta_i} = 0$.

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

$$\ge \frac{|s|\frac{1}{n}\sum_{i=1}^{n}\Theta_i^2}{1+|s|\Delta_n} - |\frac{1}{n}\sum_{i=1}^{n}\Theta_i| \tag{5}$$

By lemma 4, we obtain $\frac{|s|}{1+|s|\Delta_n} = O_p(\frac{1}{\sqrt{n}})$.

By lemma 2, we obtain

$$s = O_p(\frac{1}{\sqrt{n}}) . \tag{6}$$

Write $\gamma_i = s\Theta_i, s$ is determined by $\Pi(s) = 0$.

By (6) and lemma 2 , we obtain

$$\max_{1\le i\le n} |\gamma_i| = O_p(\frac{1}{\sqrt{n}}) o_p(\sqrt{n}) = o_p(1) . \tag{7}$$

and

$$\begin{aligned} 0 &= \Pi(s) \\ &= \frac{1}{n}\sum_{i=1}^{n}\Theta_i - s\frac{1}{n}\sum_{i=1}^{n}\Theta_i^{\ 2} \\ &+ \frac{1}{n}\sum_{i=1}^{n}\Theta_i \frac{\gamma_i^2}{1+\gamma_i} \end{aligned}$$

Write $s = \frac{\frac{1}{n}\sum_{i=1}^{n}\Theta_i}{\frac{1}{n}\sum_{i=1}^{n}\Theta_i^{\ 2}} + \beta,$

By (6), (7) and lemma 2, we obtain

$$\beta = \frac{\frac{1}{n}\sum_{i=1}^{n}\Theta_i^{\ 3} s^2}{\frac{1}{n}\sum_{i=1}^{n}\Theta_i^{\ 2}(1+\gamma_i)} = o_p(\sqrt{n}) O_p(\frac{1}{n}) = o_p(\frac{1}{\sqrt{n}}) .$$

By Taylor expansion, we obtain

$\log(1+\gamma_i) = \gamma_i - \frac{\gamma_i^2}{2} + \eta_i$, where A is a positive

$P\{|\eta_i| \le A|\gamma_i|^3, 1\le i\le n\} \to 1, n\to\infty$.

We obtain $l(\nabla\beta) = -2\log R(\nabla\beta) = 2\sum_{i=1}^{n}\log(1+\gamma_i) = 2\sum_{i=1}^{n}\gamma_i - \sum_{i=1}^{n}\gamma_i^2 + 2\sum_{i=1}^{n}\eta_i$

$$= \frac{n(\frac{1}{n}\sum_{i=1}^{n}\Theta_i)^2}{\frac{1}{n}\sum_{i=1}^{n}\Theta_i^{\ 2}} - n\beta^2\frac{1}{n}\sum_{i=1}^{n}\Theta_i^2 + 2\sum_{i=1}^{n}\eta_i$$

$$\hat{=} T_1 + T_2 + T_3 \tag{8}$$

By lemma 4 and lemma 3, we obtain

$$T_1 \to_d \frac{\sigma_3^2}{\sigma_4^2}\chi_{(1)}^2 \text{ , } n\to\infty . \tag{9}$$

$$T_2 = o_p(1) . \tag{10}$$

By (6) and lemma 2 , we obtain

$$|2\sum_{i=1}^{n}\eta_i| \le 2A|s|^3|r_i|^3 = o_p(1)$$

We show that

$$T_3 = o_p(1). \tag{11}$$

By(9)-(11), we obtain

$$l(\nabla\beta) \to_d \frac{\sigma_3^2}{\sigma_4^2}\chi^2_{(1)},\ , \ n \to \infty. \ \square \tag{12}$$

Proof of theorem 2

We obtain $R^{'}(\nabla\beta) = \sup\prod_{i=1}^{2g} 2gP_i^{'}$,

where $P_i^{'} \ge 0, \sum_{i=1}^{2g} P_i^{'} = 1, \sum_{i=1}^{2g} P_i^{'}\sqrt{u}H_i = 0$.

By Lagrange method, we obtain

$$P_i^{'} = \frac{1}{2g(1+\lambda\sqrt{u}H_i)}, 1 \le i \le 2g, \tag{13}$$

where $\lambda \in R^1, \lambda$ is determined by $\Pi^{'}(\lambda) = \frac{1}{2g}\sum_{i=1}^{2g}\frac{\sqrt{\mathrm{u}}H_i}{1+\lambda\sqrt{u}H_i} = 0$.

we obtain

$$\max_{1\le i\le 2g}|Y_i| = o_p(g^{\frac{1-2\gamma}{4}}), \tag{14}$$

and $\max_{1\le i\le 2g}|\sqrt{\mathrm{u}}H_i| = \sqrt{u}\max_{1\le i\le 2g}|H_i| = \sqrt{m}o_p(g^{\frac{1-2\gamma}{4}})$

$$= o_p(n^{\frac{1}{4}}) = o_p(g^{\frac{1}{2}}). \tag{15}$$

We can obtain

$$\lambda = \mathrm{O}_p(g^{-\frac{1}{2}}). \tag{16}$$

And

$$0 = |\Pi^{'}(\lambda)| = \frac{1}{2g}\left|\sum_{i=1}^{2g}\frac{\sqrt{\mathrm{u}}H_i}{1+\lambda\sqrt{u}H_i}\right|$$

$$\ge \frac{|\lambda|S_2}{1+Z_{2g}|\lambda|} - \frac{1}{2g}\left|\sum_{i=1}^{2g}\sqrt{\mathrm{u}}H_i\right|, \tag{17}$$

where $\Delta_{2g} = \max_{1\le i\le 2g}|\sqrt{u}\mathrm{H}_i|$.

By Lemma 3 , we get $\frac{1}{2g}\sum_{i=1}^{2g}\sqrt{u}\mathrm{H}_i = \mathrm{O}_p(\frac{\sqrt{u}}{\sqrt{n}}) = \mathrm{O}_p(g^{-\frac{1}{2}})$. (18)

By (17) and (18), we obtain

$$\frac{|\lambda| S_2}{1+\Delta_{2g}|\lambda|} = O_p(g^{-\frac{1}{2}}). \tag{19}$$

By lemma 4 and 6, (17) and (19), we obtain

$|\lambda| = O_p(g^{-\frac{1}{2}})$. Write $\gamma_i = \lambda\sqrt{u}\mathrm{H}_i$

By (16) and (17), we obtain

$$\max_{1\le i\le 2g}|\gamma_i| = O_p(g^{-\frac{1}{2}})o_p(g^{\frac{1}{2}}) = o_P(1). \tag{20}$$

It is shown that

$$0 = \Pi'(\lambda) = \frac{1}{2g}\sum_{i=1}^{2g}\frac{\sqrt{u}\mathrm{H}_i}{1+\lambda\sqrt{u}\mathrm{H}_i} = \frac{1}{2g}\sum_{i=1}^{2g}\frac{\sqrt{u}\mathrm{H}_i}{1+\gamma_i}$$

$$= \sqrt{u}\overline{\mathrm{H}} - S_2\lambda + \frac{1}{2g}\sum_{i=1}^{2g}\sqrt{u}\mathrm{H}_i\frac{\gamma_i^2}{1+\gamma_i}.$$

Write $\beta = \frac{1}{2g}\sum_{i=1}^{2g}\sqrt{u}\mathrm{H}_i\frac{\gamma_i^2}{1+\gamma_i}$, we obtain

$$\lambda = S_2^{-1}\sqrt{u}\overline{\mathrm{H}} + S_2^{-1}\beta. \tag{21}$$

By lemma 4 and (17), we obtain

$$\frac{1}{2g}\sum_{i=1}^{2g}(\sqrt{u}|\mathrm{H}_i|)^3 \le \frac{1}{2g}\sum_{i=1}^{2g}(\sqrt{u}|\mathrm{H}_i|)^2\max_{1\le i\le 2g}\sqrt{u}|\mathrm{H}_i| = o_p(g^{\frac{1}{2}}). \tag{22}$$

By (20)- (22), we obtain

$$|\beta| \le \frac{1}{2g}\sum_{i=1}^{2g}(\sqrt{u}|\mathrm{H}_i|)^3|\lambda|^2(1+\gamma_i)^{-1} = o_p(g^{\frac{1}{2}})O_p(g^{-1}) = o_p(g^{-\frac{1}{2}}). \tag{23}$$

By (23) and Taylor expansion, we obtain

$$l'(\nabla\beta) = -2\log R'(\nabla\beta) = 2\sum_{i=1}^{2g}\log(1+\gamma_i) = 2\sum_{i=1}^{2g}\gamma_i - \sum_{i=1}^{2g}\gamma_i^2 + 2\sum_{i=1}^{2g}\eta_i$$

$$= 2\sum_{i=1}^{2g}\lambda\sqrt{u}\mathrm{H}_i - \sum_{i=1}^{2g}(\lambda\sqrt{u}\mathrm{H}_i)^2 + 2\sum_{i=1}^{2g}\eta_i$$

$$= 2nS_2^{-1}(\overline{\mathrm{H}})^2 + 4\sqrt{u}g\overline{\mathrm{H}}S_2^{-1}\beta - 2ug(S_2^{-1})^2(\overline{\mathrm{H}})^2\frac{1}{2g}\sum_{i=1}^{2g}u\mathrm{H}_i^2$$

$$-2g(S_2^{-1})^2\beta^2\frac{1}{2g}\sum_{i=1}^{2g}u\mathrm{H}_i^2 - 4\sqrt{u}g\overline{\mathrm{H}}\beta\frac{1}{2g}\sum_{i=1}^{2g}u\mathrm{H}_i^2(S_2^{-1})^2 + 2\sum_{i=1}^{2g}\eta_i$$

$$= nS_2^{-1}(\overline{Y})^2 - 2g\beta^2S_2^{-1} + 2\sum_{i=1}^{2g}\eta_i \triangleq F_1 + F_2 + F_3 \tag{24}$$

where $\overline{\mathrm{H}} = \frac{1}{n}\sum_{i=1}^{n}\Theta_i$.

By Lemma 3 and 4, we obtain

$$F_1 \to_d \chi^2_{(1)}, \quad n\to\infty. \tag{25}$$

By (24) , we obtain

$$F_2 = 2g\beta^2 S_2^{-1} = o_p(1)\,. \tag{26}$$

By (18) and (24), we obtain

$$2\left|\sum_{i=1}^{2g}\eta_i\right| \le 2B|\lambda|^3 \sum_{i=1}^{2g}(\sqrt{u}\,|\,\mathrm{H}_i\,|)^3 = \mathrm{O}_p(g^{-1}) = o_p(1)\,,$$

$$F_3 = o_p(1)\,. \tag{27}$$

By (24) - (27), we obtain

$$l'(\nabla\beta)) \to_d \chi^2_{(1)}\ ,\ n\to\infty.\ \square \tag{28}$$

3. Conclusion

In the paper, we obtain the large samples property of difference of two linear model with strongly stationary ϕ – mixing sample.

4. Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of the paper.

Acknowledgements

This work was partially supported by the National Social Science Foundation of China (15BTJ031) and the Natural Science Foundation of Guangdong Province(2016A030313812). The work was also supported by the National Natural Science Foundation of China (11271095; 11461057;11561010), the National Statistics Projects of China(2012LY178). and Project of Enhancing School With Innovation of Guangdong Ocean University (GDOU2014050217).

References

[1] Roussas, G. G., Ioannides, D. A: Moment inequalities for mixing sequences of random variables [J].Stochastic.Anal. Appl, Vol. 5 (1987), p. 61-120.

[2] Shanchao Yang: Moment Bound for strong mixing sequences and their Applications,[J]. Journal of Mathematical Research with Applications, Vol. 20 (2000) No.3, p. 349-359.

[3] Owen, A. B: Empirical likelihood ratio confidence intervals for a single function, [J]. Biometrika, Vol. 75 (1988), p.237-249.

[4] Owen, A, B.: Empirical likelihood confidence regions [J].Ann. Statist., Vol. 18 (1990), p. 90-120.

[5] Naiyi Li, Chengming Wang: Empirical Likelihood Ratio Confidence Regions with A Nuisance Parameter for Dependent Samples [J], Chinese Journal of Engineering Mathematics, Vol. 6 (2006), p. 1039-1047.

[6] Su, H.& Liang H:An empirical likelihood-based method for comparison of treatment effects-test of equality of coefficients in linear models, Comput. Stat. Data. An, Vol. 54 (2010), p. 1079-1088.

[7] Zi, X., Zou, C. & Liu, Y:Two-sample empirical likelihood method for difference between coefficients in linear regression model, Stat. Papers, Vol. 53 (2012), p. 83-93.

[8] Zhou, Y. & Liang, H: Empirical-likelihood-based semi-parametric inference for the treatment effect in the two-sample problem with censoring, Biometrika, Vol. 92 (2005), p. 271-282.

[9] J,J,Samour: Convergence of sums of mixing triangular arrays of random vectors with station arrows[J]. Ann Probability, Vol. 12 (1984), p. 390 -426.