应用概率统计 第 37 卷
第 3 期 2021 年 6 月

Chinese Journal of Applied Probability and Statistics
Jun., 2021, Vol. 37, No. 3, pp. 259-273
doi: 10.3969/j.issn.1001-4268.2021.03.003

Empirical Likelihood Estimators of the Error Variances in Linear Models*

QIN Yongsong* ZHANG Ping

(*Department of Statistics, Guangxi Normal University, Guilin, 541004, China*)

Abstract: We apply the empirical likelihood technique to propose a new class of estimators of the error variance in linear models. It is shown that the proposed estimators are asymptotically normally distributed with asymptotic variances not greater than that of the usual estimators of the error variance. And the closed forms of the asymptotic variances of the estimators are presented.

Keywords: linear model; empirical likelihood; estimator of error variance; relative efficiency

2020 Mathematics Subject Classification: 62G05; 62E20

Citation: QIN Y S, ZHANG P. Empirical likelihood estimators of the error variances in linear models [J]. Chinese J Appl Probab Statist, 2021, 37(3): 259–273.

§1. Introduction

Consider the following linear regression model

$$Y = x^{\mathsf{T}}\beta + \epsilon, \tag{1}$$

where Y is a scalar response variable, $x \in \mathbb{R}^r$ is a vector of fixed design variable, $\beta \in \mathbb{R}^r$ is a vector of regression parameters, and error $\epsilon \in \mathbb{R}$ is a random variable. Let $x_1, x_2, \cdots, x_n$ be the design vectors, $Y_1, Y_2, \cdots, Y_n$ be the corresponding observations, and $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ be the random errors. We assume that $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ are i.i.d. random errors with mean 0 and unknown variance σ^2, where $0 < \sigma^2 < \infty$.

Let

$$X_{(n)} = \begin{pmatrix} x_1^{\mathsf{T}} \\ x_2^{\mathsf{T}} \\ \vdots \\ x_n^{\mathsf{T}} \end{pmatrix}, \qquad Y_{(n)} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

*The project was supported by the National Natural Science Foundation of China (Grant Nos. 12061017; 11671102).

*Corresponding author, E-mail: ysqin@gxnu.edu.cn.

Received June 18, 2019. Revised June 2, 2020.

and suppose that $\mathrm{Rank}(X_{(n)}) = r$. Then the least square estimator of β is

$$\widehat{\beta}_n = (X_{(n)}^{\mathsf{T}} X_{(n)})^{-1} X_{(n)}^{\mathsf{T}} Y_{(n)}.$$

A commonly used estimator of σ^2 is

$$\widehat{\sigma}_n^2 = \frac{1}{n-r}(Y_{(n)} - X_{(n)}\widehat{\beta}_n)^{\mathsf{T}}(Y_{(n)} - X_{(n)}\widehat{\beta}_n) = \frac{1}{n-r}\sum_{i=1}^{n} e_{ni}^2, \tag{2}$$

where $e_{ni} = y_i - x_i^{\mathsf{T}}\widehat{\beta}_n$, $1 \leqslant i \leqslant n$.

The estimation of error variances plays an important role in regression predictions, diagnostics and model selection procedures. For example, based on (15) below, to construct the confidence region of β in the linear model, we need to obtain a good estimator of the variance. The asymptotic properties of $\widehat{\sigma}_n^2$ are studied extensively. For instance, Zhao[1] investigated the strong convergence of $\widehat{\sigma}_n^2$. Chen[2] obtained the Berry-Esseen bound of $(\widehat{\sigma}_n^2 - \sigma^2)/\sqrt{\mathsf{Var}\,\widehat{\sigma}_n^2}$. Fu et al.[3] obtained the precise asymptotics in the law of the logarithm for the first moment of the error variance estimator.

Note that $\mathsf{E}(e) = 0$ and $\mathsf{E}(e^2) = \sigma^2$. Using $\mathsf{E}(e^2) = \sigma^2$ and the Quasi observation values e_{ni} of e, the empirical likelihood (EL) estimator of σ^2 would be $n^{-1}\sum_{i=1}^{n} e_{ni}^2$, which has the same asymptotic efficiency as $\widehat{\sigma}_n^2$ while $\widehat{\sigma}_n^2$ is unbiased. Our aim in this article is to construct an empirical likelihood (EL) estimator of σ^2 by using the auxiliary information $\mathsf{E}(e) = 0$. Under some regularity conditions, it is shown that the asymptotic distribution of the estimator is a normal distribution with asymptotic variance not larger than that of $\widehat{\sigma}_n^2$. We note that the closed form of the asymptotic variance of $\widehat{\sigma}_n^2$ is not given in the existing literatures. In this article, the closed forms of the asymptotic variances of $\widehat{\sigma}_n^2$ and its improved version are all given, which are especially useful to construct the confidence intervals for the error variance. When the design points form a random sequence, Shi[4] studied the EL estimator of error variance in a linear model. We note that the statistical linear models with non-random design points is also an important research topic theoretically and practically and usually the techniques used to treat non-random designs are more involved and the non-random case occurs extensively in application fields.

The EL method as a nonparametric technique for statistical inference in the nonparametric setting has been introduced by Owen[5,6] and has many advantages over its counterparts like the normal-approximation-based method and the bootstrap method; see, for instance, [7] and [8]. Three striking properties of the empirical likelihood are the Wilks' theorem, Bartlett correction and ability to use auxiliary information. Chen and Qin[9] have shown that the empirical likelihood method can be naturally applied to make more accurate statistical inference in finite population estimation problems by employing auxiliary information efficiently. Zhang[10] applied the empirical likelihood technique to propose a new class of M-functional estimators as well as quantile estimators in the presence of some auxiliary information. More results of the empirical likelihood inference with auxiliary information can be seen in [11] and [12], among others.

The rest of this paper is organized as follows. The main results of this paper are presented in Section 2. Results of a simulation study on the finite sample performance of the new variance estimator are reported in Section 3. Some lemmas to prove the main results are presented in Section 4. The proof of the main results is presented in Section 5.

§2. Main Results

To use the auxiliary information $\mathsf{E}(e)=0$, based on the Quasi observation values $\{e_{ni},\ 1\leqslant i\leqslant n\}$ of e, we now define the following empirical likelihood function

$$R=\sup_{p_1,p_2,\cdots,p_n}\Big\{\prod_{i=1}^n p_i\,\Big|\,\sum_{i=1}^n p_i=1,\ p_i\geqslant 0,\ \sum_{i=1}^n p_ie_{ni}=0\Big\}.$$

It is easy to obtain $R=\prod\limits_{i=1}^n p_i$, where

$$p_i=\frac{1}{n}\frac{1}{1+\lambda e_{ni}},\qquad 1\leqslant i\leqslant n,\tag{3}$$

where $\lambda\in\mathbb{R}$ is determined by

$$\sum_{i=1}^n\frac{e_{ni}}{1+\lambda e_{ni}}=0.\tag{4}$$

Thus, with the use of the auxiliary information $\mathsf{E}(e)=0$, the new estimator of σ^2 is

$$\widehat{\sigma}^2_{n,el}=\sum_{i=1}^n p_ie^2_{ni},\tag{5}$$

Use $\|x\|$ to denote the L_2 norm of x, and $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ to denote the least and largest eigenvalues of a matrix A respectively. To obtain the asymptotical distribution of $\widehat{\sigma}^2_{n,el}$, we need the following assumptions:

(A1) The $e_1,e_2,\cdots,e_n$ are i.i.d. random variables with $\mathsf{E}(e_1)=0$, $0<\mathsf{Var}\,(e_1)=\sigma^2<\infty$ and there exists $\delta>0$ such that $\mathsf{E}|e_1|^{4+\delta}<\infty$.

(A2) $\max\limits_{1\leqslant i\leqslant n}\|x_i\|=O(1)$ and there are constants $C_i>0$, $i=1,2$ such that $0<C_1\leqslant\lambda_{\min}(n^{-1}X^{\mathsf T}_{(n)}X_{(n)})\leqslant\lambda_{\max}(n^{-1}X^{\mathsf T}_{(n)}X_{(n)})\leqslant C_2<\infty$.

(A3) $ns_n^6 \to \infty$ and $\mu_4 - \sigma^4 - \mu_3^2\sigma^{-2}s_n^2 \geqslant c$ for some constant $c > 0$, where

$$s_n^2 = 1 - \Big(n^{-1}\sum_{i=1}^n x_i\Big)^{\mathsf{T}}(n^{-1}X_{(n)}^{\mathsf{T}}X_{(n)})^{-1}\Big(n^{-1}\sum_{i=1}^n x_i\Big).$$

Remark 1 Note that $n^{-1}X_{(n)}^{\mathsf{T}}X_{(n)} = n^{-1}\sum_{i=1}^n x_i x_i^{\mathsf{T}}$ and in many published articles $n^{-1}X_{(n)}^{\mathsf{T}}X_{(n)} = \Sigma + o_p(1)$ for some positive matrix Σ is supposed, which implies

$$0 < C_1 \leqslant \lambda_{\min}(n^{-1}X_{(n)}^{\mathsf{T}}X_{(n)}) \leqslant \lambda_{\max}(n^{-1}X_{(n)}^{\mathsf{T}}X_{(n)}) \leqslant C_2 < \infty.$$

Further, $s_n^2 = n^{-1}1_n^{\mathsf{T}}[I_n - X_{(n)}(X_{(n)}^{\mathsf{T}}X_{(n)})^{-1}X_{(n)}^{\mathsf{T}}]1_n$, where $1_n = (1, 1, \cdots, 1)^{\mathsf{T}}$. It can be seen that $0 \leqslant s_n^2 \leqslant 1$. From (6) below, one can see that $\mu_4 - \sigma^4 - \mu_3^2\sigma^{-2}s_n^2$ is the asymptotic variance of $\sqrt{n}\widehat{\sigma}_{n,el}^2$. We need the assumption that the asymptotic variance is bounded away from 0.

We now state the main results, which establish the asymptotic normality of the estimators of error variances.

Theorem 2 Suppose that conditions (A1) to (A3) are satisfied. Then as $n \to \infty$,

$$\frac{\sqrt{n}(\widehat{\sigma}_{n,el}^2 - \sigma^2)}{\sqrt{\mu_4 - \sigma^4 - \mu_3^2\sigma^{-2}s_n^2}} \xrightarrow{\mathrm{d}} \mathrm{N}(0,1), \tag{6}$$

and

$$\frac{\sqrt{n}(\widehat{\sigma}_n^2 - \sigma^2)}{\sqrt{\mu_4 - \sigma^4}} \xrightarrow{\mathrm{d}} \mathrm{N}(0,1), \tag{7}$$

where $\mu_j = \mathsf{E}(e_1^j)$, $j = 3, 4$.

Remark 3 Theorem 2 shows that $\widehat{\sigma}_{n,el}^2$ has smaller asymptotic variance than $\widehat{\sigma}_n^2$ or the same asymptotic variance as $\widehat{\sigma}_n^2$. To the best of our knowledge, this is the first time the explicit form of the asymptotic variance of $\widehat{\sigma}_n^2$ is given. We also note that $\widehat{\sigma}_{n,el}^2$ and $\widehat{\sigma}_n^2$ have the same asymptotic variance when the skewness of e_1 is 0, especially as $e_1 \sim \mathrm{N}(0, \sigma^2)$. In addition, it is also noted that $s_n^2 = 0$ when the covariate x in (1) includes 1 (i.e. intercept term), which indicates that the new estimator $\widehat{\sigma}_{n,el}^2$ has the same asymptotic variance as $\widehat{\sigma}_n^2$. In this case, $\widehat{\sigma}_{n,el}^2$ may not be recommended. However, in many application cases, centralization or standardization of $\{x_i\}$ and $\{y_i\}$ is used, where the intercept term vanishes and $\widehat{\sigma}_{n,el}^2$ may be used.

Remark 4 Let u_α satisfy $\mathsf{P}(|\mathrm{N}(0,1)| \leqslant u_\alpha) = \alpha$ for $0 < \alpha < 1$, where $\mathrm{N}(0,1)$ is the standard normal random variable. It follows from Theorem 2 that the EL based confidence interval for σ^2 with asymptotically correct coverage probability α can be constructed as

$$[\widehat{\sigma}_{n,el}^2 - u_\alpha\sqrt{a},\ \widehat{\sigma}_{n,el}^2 + u_\alpha\sqrt{a}],$$

where $a = (\widehat{\mu}_4 - \widehat{\sigma}^4 - \widehat{\mu}_3^2\widehat{\sigma}^{-2}s_n^2)/n$, $\widehat{\mu}_j = n^{-1}\sum_{i=1}^{n} e_{ni}^j$, $j = 2, 3, 4$ with $\widehat{\sigma}^2 = \widehat{\mu}_2$. On the other hand, the confidence interval for σ^2 based on the estimator $\widehat{\sigma}_n^2$ with asymptotically correct coverage probability α can be constructed as

$$[\widehat{\sigma}_n^2 - u_\alpha\sqrt{b},\ \widehat{\sigma}_n^2 + u_\alpha\sqrt{b}],$$

where $b = (\widehat{\mu}_4 - \widehat{\sigma}^4)/n$.

§3. Simulation Results

We conducted a small simulation study on the finite sample performances of $\widehat{\sigma}_{n,el}^2$ and $\widehat{\sigma}_n^2$. For this purpose, we used the model

$$Y_i = x_i\beta + e_i \tag{8}$$

with $x_i = i/(n+1) + u_i$, $i = 1, 2, \cdots, n$, the u_i's were generated from the uniform distribution $\mathrm{U}[-1, 1]$ with a given seed so that they remained fixed in the simulations, and the $\{e_i, i = 1, 2, \cdots, n\}$ were generated respectively from $e_1 \sim \mathrm{N}(0, 1)$, $e_1 \sim \chi_1^2 - 1$, $e_1 \sim \exp(1) - 1$ and $e_1 \sim \mathrm{U}(-0.5, 0.5)$, where $\beta = 2$.

For each of the four cases, we generated 3,000 random samples of incomplete data $\{(x_i, Y_i), i = 1, 2, \cdots, n\}$ for $n = 60$, 100, 150, 500 and 1,000 from model (8). Based on the simulation samples, we evaluated the biases, sample variances, estimated variance and relative efficiencies for $\widehat{\sigma}_{n,el}^2$ and $\widehat{\sigma}_n^2$. In more details, $\mathrm{Bias}(\widehat{\sigma}_{n,el}^2)$ and $\mathrm{Var}(\widehat{\sigma}_{n,el}^2)$ denote, respectively, the average of 3,000 biases of $\widehat{\sigma}_{n,el}^2$ and the sample variance of 3,000 estimators $\widehat{\sigma}_{n,el}^2$. In addition, $\widehat{\mathrm{Var}}(\widehat{\sigma}_{n,el}^2)$ denotes the average of 3,000 estimated variances, i.e. the estimators $(\widehat{\mu}_4 - \widehat{\sigma}^4 - \widehat{\mu}_3^2\widehat{\sigma}^{-2}s_n^2)/n$, where $\widehat{\mu}_j = n^{-1}\sum_{i=1}^{n} e_{ni}^j$, $j = 2, 3, 4$ with $\widehat{\sigma}^2 = \widehat{\mu}_2$. $\mathrm{Bias}(\widehat{\sigma}_n^2)$, $\mathrm{Var}(\widehat{\sigma}_n^2)$ and $\widehat{\mathrm{Var}}(\widehat{\sigma}_n^2)$ are defined similarly. Finally, $e(\widehat{\sigma}_n^2, \widehat{\sigma}_{n,el}^2) = \mathrm{Var}(\widehat{\sigma}_{n,el}^2)/\mathrm{Var}(\widehat{\sigma}_n^2)$ denotes the relative efficiency. Simulation results for $e_1 \sim \mathrm{N}(0, 1)$, $e_1 \sim \chi_1^2 - 1$, $e_1 \sim \exp(1) - 1$ and $e_1 \sim \mathrm{U}(-0.5, 0.5)$ were reported in Tables 1, 2, 3 and 4, respectively. In additional simulations, we considered the comparison of the mean squared errors (MSE) of the estimators and founded that the results were similar to the comparison of the $\mathrm{Var}(\widehat{\sigma}_{n,el}^2)$ and $\mathrm{Var}(\widehat{\sigma}_n^2)$. To save the space, we do not present these results here.

We observe from the simulation results that the proposed variance estimator has significantly improved the usual variance estimator in terms of the asymptotic efficiency when the underline error has a skewed distribution.

Table 1 Biases, variances and efficiencies when $e_1 \sim \mathrm{N}(0,1)$

n	$\mathrm{Bias}(\widehat{\sigma}^2_{n,el})$	$\mathrm{Bias}(\widehat{\sigma}^2_n)$	$\mathrm{Var}(\widehat{\sigma}^2_{n,el})$	$\mathrm{Var}(\widehat{\sigma}^2_n)$	$\widehat{\mathrm{Var}}(\widehat{\sigma}^2_{n,el})$	$\widehat{\mathrm{Var}}(\widehat{\sigma}^2_n)$	$e(\widehat{\sigma}^2_n, \widehat{\sigma}^2_{n,el})$
60	-0.0081	0.0100	0.0346	0.0350	0.0308	0.0327	0.9892
100	-0.0048	0.0057	0.0200	0.0199	0.0193	0.0201	1.0077
150	-0.0131	-0.0066	0.0131	0.0132	0.0127	0.0130	0.9912
500	0.0013	0.0030	0.0041	0.0041	0.0040	0.0040	1.0001
1,000	-0.0007	0.0001	0.0020	0.0020	0.0020	0.0020	0.9983

Table 2 Biases, variances and efficiencies when $e_1 \sim \chi^2 - 1$

n	$\mathrm{Bias}(\widehat{\sigma}^2_{n,el})$	$\mathrm{Bias}(\widehat{\sigma}^2_n)$	$\mathrm{Var}(\widehat{\sigma}^2_{n,el})$	$\mathrm{Var}(\widehat{\sigma}^2_n)$	$\widehat{\mathrm{Var}}(\widehat{\sigma}^2_{n,el})$	$\widehat{\mathrm{Var}}(\widehat{\sigma}^2_n)$	$e(\widehat{\sigma}^2_n, \widehat{\sigma}^2_{n,el})$
60	-0.0978	-0.0321	0.5597	0.8797	0.4608	0.8361	0.6363
100	-0.0410	0.0354	0.3771	0.5639	0.3078	0.5455	0.6687
150	-0.0144	0.0025	0.2206	0.3365	0.2122	0.3621	0.6558
500	-0.0093	0.0030	0.0744	0.1142	0.0691	0.1118	0.6517
1,000	0.0076	0.0084	0.0396	0.0568	0.0364	0.0575	0.6974

Table 3 Biases, variances and efficiencies when $e_1 \sim \exp(1) - 1$

n	$\mathrm{Bias}(\widehat{\sigma}^2_{n,el})$	$\mathrm{Bias}(\widehat{\sigma}^2_n)$	$\mathrm{Var}(\widehat{\sigma}^2_{n,el})$	$\mathrm{Var}(\widehat{\sigma}^2_n)$	$\widehat{\mathrm{Var}}(\widehat{\sigma}^2_{n,el})$	$\widehat{\mathrm{Var}}(\widehat{\sigma}^2_n)$	$e(\widehat{\sigma}^2_n, \widehat{\sigma}^2_{n,el})$
60	-0.0278	-0.0100	0.0953	0.1217	0.0735	0.1207	0.7835
100	-0.0165	-0.0005	0.0464	0.0757	0.0455	0.0724	0.6133
150	-0.0187	-0.0097	0.0360	0.0508	0.0316	0.0493	0.7087
500	0.0048	0.0104	0.0121	0.0167	0.0113	0.0169	0.7229
1,000	-0.0061	-0.0034	0.0053	0.0078	0.0053	0.0078	0.6812

Table 4 Biases, variances and efficiencies when $e_1 \sim \mathrm{U}(-0.5, 0.5)$

n	$\mathrm{Bias}(\widehat{\sigma}^2_{n,el})$	$\mathrm{Bias}(\widehat{\sigma}^2_n)$	$\mathrm{Var}(\widehat{\sigma}^2_{n,el})$	$\mathrm{Var}(\widehat{\sigma}^2_n)$	$\widehat{\mathrm{Var}}(\widehat{\sigma}^2_{n,el})$	$\widehat{\mathrm{Var}}(\widehat{\sigma}^2_n)$	$e(\widehat{\sigma}^2_n, \widehat{\sigma}^2_{n,el})$
60	-0.00170	-0.00028	0.00009	0.00009	0.00009	0.00009	0.9810
100	-0.00079	0.00007	0.00005	0.00006	0.00005	0.00006	0.9818
150	-0.00064	-0.00009	0.00003	0.00003	0.00004	0.00004	0.9926
500	-0.00006	0.00010	0.00001	0.00001	0.00001	0.00001	0.9974
1,000	-0.00001	0.00008	0.00001	0.00001	0.00001	0.00001	0.9967

In addition, using the same models and the simulated samples as above and taking the nominal level $\alpha = 0.95$, we conducted a small simulation study to compare the finite sample performances of the confidence intervals for the error variance σ^2 based on $\widehat{\sigma}^2_{n,el}$ and $\widehat{\sigma}^2_n$, respectively. The coverage probabilities (CP) and the average lengths (AL) of the confidence intervals for σ^2 in 3,000 simulations were shown in Tables 5 – 8. From the simulation results, we can see that the CP of the confidence intervals based on $\widehat{\sigma}^2_{n,el}$ and $\widehat{\sigma}^2_n$ all converge to the nominal level $\alpha = 0.95$ as n is large enough. In addition, the AL of the confidence intervals based on $\widehat{\sigma}^2_{n,el}$ are uniformly shorter than that of the confidence intervals based on $\widehat{\sigma}^2_n$, which agrees with the results of Theorem 2.

Table 5 Coverage probabilities (CP) and average lengthes (AL) of the $\widehat{\sigma}^2_{n,el}$ and $\widehat{\sigma}^2_n$ based confidence intervals for σ^2 with $e_1 \sim \mathrm{N}(0,1)$

	CP		AL	
n	$\widehat{\sigma}^2_{n,el}$	$\widehat{\sigma}^2_n$	$\widehat{\sigma}^2_{n,el}$	$\widehat{\sigma}^2_n$
60	0.909	0.915	0.6694	0.6894
100	0.917	0.926	0.5352	0.5451
150	0.921	0.929	0.4360	0.4413
500	0.941	0.945	0.2466	0.2474
1,000	0.950	0.948	0.1748	0.1751

Table 6 Coverage probabilities (CP) and average lengthes (AL) of the $\widehat{\sigma}^2_{n,el}$ and $\widehat{\sigma}^2_n$ based confidence intervals for σ^2 with $e_1 \sim \chi^2 - 1$

	CP		AL	
n	$\widehat{\sigma}^2_{n,el}$	$\widehat{\sigma}^2_n$	$\widehat{\sigma}^2_{n,el}$	$\widehat{\sigma}^2_n$
60	0.729	0.754	2.1563	2.8599
100	0.799	0.827	1.8781	2.4710
150	0.858	0.872	1.6068	2.0895
500	0.893	0.908	0.9717	1.2389
1,000	0.915	0.926	0.7207	0.9090

Table 7 Coverage probabilities (CP) and average lengthes (AL) of the $\widehat{\sigma}^2_{n,el}$ and $\widehat{\sigma}^2_n$ based confidence intervals for σ^2 with $e_1 \sim \exp(1) - 1$

	CP		AL	
n	$\widehat{\sigma}^2_{n,el}$	$\widehat{\sigma}^2_n$	$\widehat{\sigma}^2_{n,el}$	$\widehat{\sigma}^2_n$
60	0.793	0.811	0.9156	1.1476
100	0.866	0.862	0.7585	0.9446
150	0.874	0.874	0.6488	0.8033
500	0.906	0.917	0.4005	0.4892
1,000	0.926	0.920	0.2800	0.3401

Table 8 **Coverage probabilities (CP) and average lengthes (AL) of the $\widehat{\sigma}^2_{n,el}$ and $\widehat{\sigma}^2_n$ based confidence intervals for σ^2 with $e_1 \sim \mathrm{U}(-0.5, 0.5)$**

	CP		AL	
n	$\widehat{\sigma}^2_{n,el}$	$\widehat{\sigma}^2_n$	$\widehat{\sigma}^2_{n,el}$	$\widehat{\sigma}^2_n$
60	0.925	0.934	0.0366	0.0373
100	0.932	0.937	0.0291	0.0287
150	0.952	0.954	0.0236	0.0238
500	0.957	0.955	0.0130	0.0130
1,000	0.952	0.950	0.0092	0.0092

§4. Lemmas

To prove our main results, we need some lemmas. The leading term of the proposed variance estimator is presented in a linear-quadratic form of independent random variables in the first lemma. A central limit theorem for linear-quadratic forms of independent random variables is given in the second lemma.

Lemma 5 Suppose that conditions (A1) to (A3) are satisfied. Then

$$\sqrt{n}(n^{-1}X^{\mathsf{T}}_{(n)}X_{(n)})^{1/2}(\widehat{\beta}_n - \beta) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \sigma^2 I), \tag{9}$$

$$\omega_n = \max_{1\leqslant i\leqslant n} |e_{ni}| = o_p(n^{1/(4+\delta)}), \tag{10}$$

$$\frac{1}{\sqrt{ns_n^2}}\sum_{i=1}^{n} e_{ni} \xrightarrow{\mathrm{d}} \mathrm{N}(0, \sigma^2) \text{ for the } s_n^2 \text{ appearing in (A3)}, \tag{11}$$

$$\frac{1}{n}\sum_{i=1}^{n} e_{ni}^j = \mu_j + o_p(1), \qquad j = 2, 3, 4, \tag{12}$$

$$\lambda = \sigma^{-2}\cdot\frac{1}{n}\sum_{i=1}^{n} e_{ni} + n^{-1/2}o_p(1), \tag{13}$$

$$\widehat{\sigma}^2_{n,el} = \frac{1}{n}\sum_{i=1}^{n}(e_{ni}^2 - \sigma^{-2}\mu_3 e_{ni}) + o_p(n^{-1/2}), \tag{14}$$

where $\mu_j = \mathrm{E}(e_1^j)$, $j = 2, 3, 4$.

Proof Proof of (9): To prove (9), we only need to show, for any given $l \in \mathbb{R}^r$ with $\|l\| = 1$, that

$$S_n = l^{\mathsf{T}}\sqrt{n}(n^{-1}X^{\mathsf{T}}_{(n)}X_{(n)})^{1/2}(\widehat{\beta}_n - \beta) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \sigma^2). \tag{15}$$

Note that

$$S_n = n^{-1/2}l^{\mathsf{T}}(n^{-1}X^{\mathsf{T}}_{(n)}X_{(n)})^{-1/2}\sum_{i=1}^{n} x_i e_i.$$

It can be shown that

$$\mathsf{Var}\,(S_n) = \sigma^2 l^{\mathsf{T}} l = \sigma^2$$

and

$$\sum_{i=1}^{n} \mathsf{E}\big|n^{-1/2} l^{\mathsf{T}} (n^{-1} X_{(n)}^{\mathsf{T}} X_{(n)})^{-1/2} x_i e_i\big|^3 \leqslant C n^{-1/2} \max_{1\leqslant i\leqslant n} \|x_i\|^3 = o(1),$$

where the conditions (A1) and (A2) are used. We thus have (15) by the Liapounov's central limit theorem.

Proof of (10): (9) and the condition (A2) imply that $\widehat{\beta}_n - \beta = O_p(n^{-1/2})$. Further, $\mathsf{E}|e_1|^{4+\delta} < \infty$ leads to $\max\limits_{1\leqslant i\leqslant n} |e_i| = o_p(n^{1/(4+\delta)})$. It follows that

$$\omega_n = \max_{1\leqslant i\leqslant n} |x_i^{\mathsf{T}}(\beta - \widehat{\beta}_n) + e_i| = O_p(n^{-1/2}) \max_{1\leqslant i\leqslant n} \|x_i\| + o_p(n^{1/(4+\delta)}) = o_p(n^{1/(4+\delta)}),$$

which implies (10).

Proof of (11): Put

$$A_n = I_n - X_{(n)}(X_{(n)}^{\mathsf{T}} X_{(n)})^{-1} X_{(n)}^{\mathsf{T}}, \quad E_n = (e_1, e_2, \cdots, e_n)^{\mathsf{T}}, \quad 1_n = (1, 1, \cdots, 1)^{\mathsf{T}},$$
$$\delta_{ij} = 1 \text{ if } i = j \quad \text{and} \quad \delta_{ij} = 0 \text{ if } i \neq j,$$

where I_n is the identity matrix. Then $\sum\limits_{i=1}^{n} e_{ni} = 1_n^{\mathsf{T}} A_n E_n$. Noting that $X_{(n)}^{\mathsf{T}} 1_n = \sum\limits_{i=1}^{n} x_i$, We have

$$\mathsf{Var}\Big(\sum_{i=1}^{n} e_{ni}\Big) = \mathsf{Var}\,(1_n^{\mathsf{T}} A_n E_n) = \sigma^2 1_n^{\mathsf{T}} A_n 1_n = n\sigma^2 s_n^2,$$

for the s_n^2 appearing in (A3), and by the moment inequalities for the sums of independent random variables,

$$\begin{aligned}
\sum_{i=1}^{n} \mathsf{E}|e_{ni}|^3 &= \sum_{i=1}^{n} \mathsf{E}\Big|\sum_{j=1}^{n} [\delta_{ij} - n^{-1} x_i^{\mathsf{T}} (n^{-1} X_{(n)}^{\mathsf{T}} X_{(n)})^{-1} x_j] e_j\Big|^3 \\
&\leqslant C \sum_{i=1}^{n} \Big\{ \sum_{j=1}^{n} \big|\delta_{ij} - n^{-1} x_i^{\mathsf{T}} (n^{-1} X_{(n)}^{\mathsf{T}} X_{(n)})^{-1} x_j\big|^3 \mathsf{E}|e_j|^3 \\
&\quad + \Big[\sum_{j=1}^{n} \big|\delta_{ij} - n^{-1} x_i^{\mathsf{T}} (n^{-1} X_{(n)}^{\mathsf{T}} X_{(n)})^{-1} x_j\big|^2 \mathsf{E}(e_j^2)\Big]^{3/2} \Big\} \\
&\leqslant C \sum_{i=1}^{n} \Big(1 + C n^{-2} \max_{1\leqslant i\leqslant n} \|x_i\|^6 + C n^{-3/2} \max_{1\leqslant i\leqslant n} \|x_i\|^6\Big) \\
&= C n \Big(1 + C n^{-3/2} \max_{1\leqslant i\leqslant n} \|x_i\|^6\Big) \leqslant C n.
\end{aligned}$$

Thus (11) follows by $n/(ns_n^2)^{3/2} = 1/(ns_n^6)^{1/2} \to 0$ and the Liapounov's central limit theorem.

Proof of (12): Let $\Delta_n = \sqrt{n}(\widehat{\beta}_n - \beta) = O_p(1)$. We then have

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n e_{ni}^2 &= \frac{1}{n}\sum_{i=1}^n (e_i - n^{-1/2}x_i^{\mathsf{T}}\Delta_n)^2 \\ &= n^{-1}\sum_{i=1}^n e_i^2 - 2n^{-3/2}\sum_{i=1}^n e_i x_i^{\mathsf{T}}\Delta_n + n^{-2}\sum_{i=1}^n x_i^{\mathsf{T}}\Delta_n^2 x_i.\end{aligned}$$

Note that

$$\begin{aligned}n^{-3/2}\sum_{i=1}^n e_i x_i^{\mathsf{T}}\Delta_n &= n^{-1/2}\max_{1\leqslant i\leqslant n}\|x_i\|O_p(1)\cdot n^{-1}\sum_{i=1}^n |e_i| \\ &= n^{-1/2}\max_{1\leqslant i\leqslant n}\|x_i\|O_p(1)\cdot O_p(1) = o_p(1)\end{aligned}$$

and that

$$n^{-2}\sum_{i=1}^n x_i^{\mathsf{T}}\Delta_n^2 x_i = n^{-1}\max_{1\leqslant i\leqslant n}\|x_i\|^2 O_p(1) = o_p(1).$$

Therefore, by the law of large numbers,

$$\frac{1}{n}\sum_{i=1}^n e_{ni}^2 = n^{-1}\sum_{i=1}^n e_i^2 + o_p(1) = \mu_2 + o_p(1).$$

Similarly, one can show that

$$\frac{1}{n}\sum_{i=1}^n e_{ni}^j = \mu_j + o_p(1), \qquad j = 3, 4.$$

We thus have (12).

Proof of (13): From (4), we have

$$0 = \frac{1}{n}\sum_{i=1}^n e_{ni} - \frac{\lambda}{n}\sum_{i=1}^n \frac{e_{ni}^2}{1+\lambda e_{ni}}.$$

It follows that

$$\Big|\frac{1}{n}\sum_{i=1}^n e_{ni}\Big| \geqslant \frac{|\lambda|}{1+|\lambda|\omega_n}\frac{1}{n}\sum_{i=1}^n e_{ni}^2,$$

where ω_n is defined in (10). Then from (11) and (12), we have

$$\frac{|\lambda|}{1+|\lambda|\omega_n} = n^{-1/2}s_n O_p(1).$$

Therefore,

$$\lambda = n^{-1/2}s_n O_p(1) \tag{16}$$

provided $n^{1/(4+\delta)}\cdot n^{-1/2}s_n \to 0$. Let $\gamma_i = \lambda e_{ni}$. Then

$$\max_{1\leqslant i\leqslant n}|\gamma_i| = o_p(1). \tag{17}$$

Using (4) again, we have

$$0 = \frac{1}{n}\sum_{i=1}^{n} e_{ni} - \frac{1}{n}\sum_{i=1}^{n} \frac{\lambda e_{ni}^2}{1+\lambda e_{ni}}$$
$$= \frac{1}{n}\sum_{i=1}^{n} e_{ni} - \lambda \cdot \frac{1}{n}\sum_{i=1}^{n} e_{ni}^2 + \lambda^2 \cdot \frac{1}{n}\sum_{i=1}^{n} \frac{e_{ni}^3}{1+\gamma_i}.$$

Therefore, combining with (12), (16) and (17), we may write

$$\lambda = [\sigma^{-2} + o_p(1)] \cdot \frac{1}{n}\sum_{i=1}^{n} e_{ni} + n^{-1}s_n^2 O_p(1) \cdot \frac{1}{n}\sum_{i=1}^{n} |e_{ni}|^3.$$

Using $n^{-1}\sum_{i=1}^{n} |e_{ni}|^3 = O_p(1)$, $0 \leqslant s_n^2 \leqslant 1$ and (11), we have

$$\lambda = \sigma^{-2} \cdot \frac{1}{n}\sum_{i=1}^{n} e_{ni} + [n^{-1/2}s_n o_p(1) + n^{-1}s_n^2 O_p(1)] = \sigma^{-2} \cdot \frac{1}{n}\sum_{i=1}^{n} e_{ni} + n^{-1/2} o_p(1),$$

which implies (13).

Proof of (14): From (5), (3) and (4), we have

$$\widehat{\sigma}_{n,el}^2 = \frac{1}{n}\sum_{i=1}^{n} \frac{e_{ni}^2}{1+\lambda e_{ni}} = \frac{1}{n}\sum_{i=1}^{n} e_{ni}^2 - \frac{\lambda}{n}\sum_{i=1}^{n} \frac{e_{ni}^3}{1+\lambda e_{ni}}$$
$$= \frac{1}{n}\sum_{i=1}^{n} e_{ni}^2 - \frac{\lambda}{n}\sum_{i=1}^{n} e_{ni}^3 + \frac{\lambda^2}{n}\sum_{i=1}^{n} \frac{e_{ni}^4}{1+\lambda e_{ni}}.$$

Using (11), (12), (13), (16) and $\max_{1\leqslant i\leqslant n} |\lambda e_{ni}| = o_p(1)$, we have

$$\widehat{\sigma}_{n,el}^2 = \frac{1}{n}\sum_{i=1}^{n} (e_{ni}^2 - \sigma^{-2}\mu_3 e_{ni}) + o_p(1)n^{-1/2}s_n + O_p(1)n^{-1}s_n^2$$
$$= \frac{1}{n}\sum_{i=1}^{n} (e_{ni}^2 - \sigma^{-2}\mu_3 e_{ni}) + o_p(n^{-1/2}s_n).$$

Thus (14) holds true by using $0 \leqslant s_n^2 \leqslant 1$. □

Let

$$Q_n = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{nij}\epsilon_{ni}\epsilon_{nj} + \sum_{i=1}^{n} b_{ni}\epsilon_{ni},$$

where ϵ_{ni} are real valued random variables, and the a_{nij} and b_{ni} denote the real valued coefficients of the linear-quadratic form. We need the following assumptions in Lemma 6.

(C1) $\{\epsilon_{ni},\ 1 \leqslant i \leqslant n\}$ are independent random variables with mean 0 and

$$\sup_{1\leqslant i\leqslant n, n\geqslant 1} \mathsf{E}|\epsilon_{ni}|^{4+\eta_1} < \infty$$

for some $\eta_1 > 0$;

(C2) For all $1 \leqslant i, j \leqslant n$, $n \geqslant 1$, $a_{nij} = a_{nji}$,

$$\sup_{1\leqslant j\leqslant n, n\geqslant 1} \sum_{i=1}^{n} |a_{nij}| < \infty,$$

and

$$\sup_{n\geqslant 1} n^{-1} \sum_{i=1}^{n} |b_{ni}|^{2+\eta_2} < \infty$$

for some $\eta_2 > 0$.

Given the above assumptions (C1) and (C2), the mean and variance of Q_n are given as (e.g. [13])

$$\mu_Q = \sum_{i=1}^{n} a_{nii}\sigma_{ni}^2,$$

$$\sigma_Q^2 = 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_{nij}^2\sigma_{ni}^2\sigma_{nj}^2 + \sum_{i=1}^{n} b_{ni}^2\sigma_{ni}^2 + \sum_{i=1}^{n}[a_{nii}^2(\mu_{ni}^{(4)} - 3\sigma_{ni}^4) + 2b_{ni}a_{nii}\mu_{ni}^{(3)}]$$

with $\sigma_{ni}^2 = \mathsf{E}(\epsilon_{ni}^2)$ and $\mu_{ni}^{(s)} = \mathsf{E}(\epsilon_{ni}^s)$ for $s = 3, 4$.

Lemma 6 Suppose that assumptions (C1) and (C2) hold true and $n^{-1}\sigma_Q^2 \geqslant c$ for some constant $c > 0$. Then

$$\frac{Q_n - \mu_Q}{\sigma_Q} \xrightarrow{\text{d}} \mathrm{N}(0, 1).$$

Proof See Theorem 1 and the remark 12 in [13]. □

§5. Proof of Theorem 2

Proof of Theorem 2 Denote $\Sigma_n = n^{-1}X_{(n)}^{\mathsf{T}}X_{(n)}$ and

$$Q_n = E_n^{\mathsf{T}}A_nE_n - \sigma^{-2}\mu_3 \cdot 1_n^{\mathsf{T}}A_nE_n = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{nij}e_ie_j + \sum_{i=1}^{n} b_{ni}e_i,$$

where

$$A_n = I_n - n^{-1}X_{(n)}\Sigma_n^{-1}X_{(n)}^{\mathsf{T}}, \quad E_n = (e_1, e_2, \cdots, e_n)^{\mathsf{T}}, \quad 1_n = (1, 1, \cdots, 1)^{\mathsf{T}}.$$

Our first step is to obtain the asymptotic distribution of Q_n. To this end, we first check the conditions (C1) and (C2). By using the condition (A2), it can be shown that

$$\begin{aligned}\sum_{i=1}^{n} |a_{nij}| &= |1 - n^{-1}x_j^{\mathsf{T}}\Sigma_n^{-1}x_j| + n^{-1}\sum_{i\neq j}^{n} |x_i^{\mathsf{T}}\Sigma_n^{-1}x_j| \\ &\leqslant 1 + n^{-1}\sum_{i=1}^{n} |x_i^{\mathsf{T}}\Sigma_n^{-1}x_j| \leqslant 1 + \max_{1\leqslant i\leqslant n} \|x_i\|^2 \leqslant C,\end{aligned}$$

and that

$$n^{-1}\sum_{i=1}^{n}|b_{ni}|^3 = n^{-1}\sum_{i=1}^{n}\left|\left[1-\left(n^{-1}\sum_{i=1}^{n}x_j\right)^{\mathsf T}\Sigma_n^{-1}x_i\right]\mu_3\sigma^{-2}\right|^3$$
$$\leqslant C + C\max_{1\leqslant i\leqslant n}\|x_i\|^6 \leqslant C.$$

Therefore, the conditions (C1) and (C2) are satisfied.

We now derive the mean and variance of Q_n. Noting that $n^{-1}\sum_{i=1}^{n}x_ix_i^{\mathsf T} = \Sigma_n$ and $n^{-1}\sum_{i=1}^{n}x_i^{\mathsf T}\Sigma_n^{-1}x_i = r$, we have

$$\mu_Q = \sigma^2\sum_{i=1}^{n}a_{nii} = \sigma^2\sum_{i=1}^{n}(1-n^{-1}x_i^{\mathsf T}\Sigma_n^{-1}x_i) = \sigma^2(n-r),$$

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n}a_{nij}^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n}(\delta_{ij}-n^{-1}x_i^{\mathsf T}\Sigma_n^{-1}x_j)^2\\
&= \sum_{i=1}^{n}\left(1-2n^{-1}x_i^{\mathsf T}\Sigma_n^{-1}x_i + n^{-2}\sum_{j=1}^{n}x_i^{\mathsf T}\Sigma_n^{-1}x_jx_j^{\mathsf T}\Sigma_n^{-1}x_i\right)\\
&= \sum_{i=1}^{n}(1-n^{-1}x_i^{\mathsf T}\Sigma_n^{-1}x_i) = n-r,\\
\sum_{i=1}^{n}a_{nii}^2 &= \sum_{i=1}^{n}(1-n^{-1}x_i^{\mathsf T}\Sigma_n^{-1}x_i)^2\\
&= \sum_{i=1}^{n}[1-2n^{-1}x_i^{\mathsf T}\Sigma_n^{-1}x_i + n^{-2}(x_i^{\mathsf T}\Sigma_n^{-1}x_i)^2]\\
&= n-2r+O\left(n^{-1}\max_{1\leqslant i\leqslant n}\|x_i\|^4\right),\\
\sum_{i=1}^{n}b_{ni}^2 &= \mu_3^2\sigma^{-4}\sum_{i=1}^{n}\left[1-\left(n^{-1}\sum_{j=1}^{n}x_j\right)^{\mathsf T}\Sigma_n^{-1}x_i\right]^2\\
&= \mu_3^2\sigma^{-4}\sum_{i=1}^{n}\left[1-2\left(n^{-1}\sum_{j=1}^{n}x_j\right)^{\mathsf T}\Sigma_n^{-1}x_i\right.\\
&\qquad\left.+\left(n^{-1}\sum_{j=1}^{n}x_j\right)^{\mathsf T}\Sigma_n^{-1}x_ix_i^{\mathsf T}\Sigma_n^{-1}\left(n^{-1}\sum_{j=1}^{n}x_j\right)\right]\\
&= \mu_3^2\sigma^{-4}\cdot ns_n^2,\\
\sum_{i=1}^{n}b_{ni}a_{nii} &= -\mu_3\sigma^{-2}\sum_{i=1}^{n}\left[1-\left(n^{-1}\sum_{j=1}^{n}x_j\right)^{\mathsf T}\Sigma_n^{-1}x_i\right](1-n^{-1}x_i^{\mathsf T}\Sigma_n^{-1}x_i)\\
&= -\mu_3\sigma^{-2}\sum_{i=1}^{n}\left[1-\left(n^{-1}\sum_{j=1}^{n}x_j\right)^{\mathsf T}\Sigma_n^{-1}x_i - n^{-1}x_i^{\mathsf T}\Sigma_n^{-1}x_i\right.\\
&\qquad\left.+n^{-1}\left(n^{-1}\sum_{j=1}^{n}x_j\right)^{\mathsf T}\Sigma_n^{-1}x_i\cdot x_i^{\mathsf T}\Sigma_n^{-1}x_i\right]\\
&= -\mu_3\sigma^{-2}\cdot\left[ns_n^2-r+O\left(\max_{1\leqslant i\leqslant n}\|x_i\|^4\right)\right].
\end{aligned}$$

It follows that

$$\begin{aligned}\sigma_Q^2 &= 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_{nij}^2\sigma^4 + \sum_{i=1}^{n} b_{ni}^2\sigma^2 + \sum_{i=1}^{n}[a_{nii}^2(\mu_4 - 3\sigma^4) + 2b_{ni}a_{nii}\mu_3] \\ &= 2(n-r)\sigma^4 + \mu_3^2\sigma^{-2}ns_n^2 + (\mu_4 - 3\sigma^4)\Big[n - 2r + O\Big(n^{-1}\max_{1\leqslant i\leqslant n}\|x_i\|^4\Big)\Big] \\ &\quad - 2\mu_3^2\sigma^{-2}\Big[ns_n^2 - r + O\Big(\max_{1\leqslant i\leqslant n}\|x_i\|^4\Big)\Big] \\ &= n(\mu_4 - \sigma^4 - \mu_3^2\sigma^{-2}s_n^2) + O(1) + O\Big(\max_{1\leqslant i\leqslant n}\|x_i\|^4\Big).\end{aligned}$$

From $\mu_4 - \sigma^4 - \mu_3^2\sigma^{-2}s_n^2 \geqslant c > 0$ and Lemma 6, we have

$$\frac{Q_n - \sigma^2(n-r)}{\sqrt{n(\mu_4 - \sigma^4 - \mu_3^2\sigma^{-2}s_n^2) + O(1) + O\Big(\max_{1\leqslant i\leqslant n}\|x_i\|^4\Big)}} \xrightarrow{\mathrm{d}} \mathrm{N}(0,1),$$

i.e.

$$\frac{n\widehat{\sigma}_{n,el}^2 + o_p(n^{1/2}) - \sigma^2(n-r)}{\sqrt{n(\mu_4 - \sigma^4 - \mu_3^2\sigma^{-2}s_n^2) + O(1) + O\Big(\max_{1\leqslant i\leqslant n}\|x_i\|^4\Big)}} \xrightarrow{\mathrm{d}} \mathrm{N}(0,1).$$

By Cramer-Wold device, we have

$$\frac{\sqrt{n}(\widehat{\sigma}_{n,el}^2 - \sigma^2)}{\sqrt{\mu_4 - \sigma^4 - \mu_3^2\sigma^{-2}s_n^2}} \xrightarrow{\mathrm{d}} \mathrm{N}(0,1),$$

which implies (6). Letting $b_{ni} = 0$, $1 \leqslant i \leqslant n$ and following the prove of (6), one can see that (7) holds true. The proof of Theorem 2 is thus complete. □

Acknowledgments The authors are thankful to the referees for constructive suggestions. We also thank Professor Qinglei Lei for conducting the simulations.

References

[1] ZHAO L C. A sufficient and necessary condition of strong convergence of the estimator of error variance in a linear model [J]. *Sci Sinica*, 1981, **11(10)**: 1187–1191. (in Chinese)

[2] CHEN X R. Berry-Esseen bounds of the estimator of error variance in a linear model [J]. *Sci Sinica*, 1981, **11(2)**: 129–140. (in Chinese)

[3] FU K A, LIU W D, ZHANG L X. Precise asymptotics for the first moment of the error variance estimator in linear models [J]. *Appl Math Lett*, 2008, **21(6)**: 641–647.

[4] SHI J. Empirical likelihood estimation of the error variance in linear correlation models and its bootstrapping [J]. *Acta Math Sci (Chinese)*, 1997, **17(1)**: 38–46. (in Chinese)

[5] OWEN A B. Empirical likelihood ratio confidence intervals for a single functional [J]. *Biometrika*, 1988, **75(2)**: 237–249.

[6] OWEN A B. Empirical likelihood ratio confidence regions [J]. *Ann Statist*, 1990, **18(1)**: 90–120.

[7] HALL P, LA SCALA B. Methodology and algorithms of empirical likelihood [J]. *Int Stat Rev*, 1990, **58(2)**: 109–127.

[8] HALL P. *The Bootstrap and Edgeworth Expansion* [M]. New York: Springer-Verlag, 1992.

[9] CHEN J H, QIN J. Empirical likelihood estimation for finite populations and the effective usage of auxiliary information [J]. *Biometrika*, 1993, **80(1)**: 107–116.

[10] ZHANG B. M-estimation and quantile estimation in the presence of auxiliary information [J]. *J Statist Plann Inference*, 1995, **44(1)**: 77–94.

[11] ZHONG B, RAO J N K. Empirical likelihood inference under stratified random sampling using auxiliary population information [J]. *Biometrika*, 2000, **87(4)**: 929–938.

[12] WU C B. Weighted empirical likelihood inference [J]. *Statist Probab Lett*, 2004, **66(1)**: 67–79.

[13] KELEJIAN H H, PRUCHA I R. On the asymptotic distribution of the Moran I test statistic with applications [J]. *J Econometrics*, 2001, **104(2)**: 219–257.

线性模型误差方差的经验似然估计

秦永松　　张　萍

(广西师范大学统计系, 桂林, 541004)

摘　要: 本文应用经验似然方法得到了线性模型误差方差的一类新的估计, 证明了估计的渐近分布为正态分布且渐近方差不超过常用的误差方差估计的渐近方差, 同时给出了渐近方差的显式表达.

关键词: 线性模型; 经验似然; 误差方差的估计; 相对效率

中图分类号: O212.7