应用概率统计 第二十二卷
第四期 2006 年 11 月

Chinese Journal of Applied Probability
and Statistics Vol. 22 No. 4 Nov. 2006

Empirical Likelihood Method under Stratified Random Sampling Using Auxiliary Information and the Information in the Strata Population Size*

Wu Changchun[1,2] Zhang Runchu[2]

([1] *School of Mathematics and Information Sciences, Jiaxing University, Jiaxing, 314001*)

([2] *LPMC and Department of Statistics, School of Mathematical Sciences, Nankai University, Tianjin, 300071*)

Abstract

The empirical likelihood method under stratified random sampling is used for getting estimators of finite population parameters, we show in this paper that the empirical likelihood approach is well-suited to incorporate auxiliary information and can accommodate this information contained in the population size for each stratum quite naturally. Our results show that it can lead to efficient estimators.

Keywords: Empirical likelihood, stratified random sampling, auxiliary information, strata population size, jackknife variance estimator.

AMS Subject Classification: 62D05, 62F12, 62E20.

§1. Introduction

In sample surveys, auxiliary information on the finite population is regularly used to increase the precision of estimators. The empirical likelihood method, introduced by Owen (1988, 1990) in the context of independent identically distributed random variables, provides a systematic nonparametric way of utilizing auxiliary information in making inference about the parameters of interest. Hartley and Rao (1968) presented a similar method in the sample survey context, using their "Scale-load approach". Under simple random sampling, they obtain the empirical maximum likelihood estimator of $\overline{Y}$ when only $\overline{X}$ is known. Chen and Qin (1993) propose an empirical likelihood method to make effective use of auxiliary information in simple random sampling without replacement (srswor). Zhong and Rao (2000) used empirical likelihood to deal with the case of stratified simple random sampling when only $\overline{X}$ is known, they, in essence, obtain estimators of the finite population mean by estimating the strata population distributions.

In this paper, we are devoted to develop an empirical likelihood method to estimate the whole finite population distribution directly, which is suitable for both stratified sampling and post stratification. In section 2, we give a brief review of empirical likelihood in finite population. In section 3, first, we naively apply the empirical likelihood method of Chen and Qin (1993) to stratified

*Supported by National Natural Science Foundation of China (No. 10571093)

Received 2003. 10. 28.

random sampling, the resulting estimator cannot be the best one, sometimes it would be bad. As in Chen and Sitter (1999), considering the fact that we have the additional information contained in the strata population sizes, we develop an empirical likelihood method which effectively incorporate auxiliary information and accommodate this information contained in the population size for each stratum. In section 4, we derive the empirical maximum likelihood estimator (EMLE) for the parameters of interest when auxiliary information in the form $\mathsf{E}_N\{u(x)\}=0$ is known. In section 5, we establish the Large-sample properties of EMLE.

§ 2. Empirical Likelihood Estimation in Finite Population

Suppose a finite population U, consists of N distinct units with values $Z_i=(y_i,x_i)^T$, $i=1,\cdots,N$. Denote the population means and distribution function of (y_i,x_i) by $(\overline{Y}_N,\overline{X}_N)$, $F_N=(1/N)\sum\limits_{i=1}^{N}\delta z_i$, respectively. Now suppose the sample, S, is, a simple random sample without replacement (srswor). The log-empirical likelihood is defined as

$$l=\sum_{i\in S}\log p_i, \tag{2.1}$$

where $p_i=\mathsf{P}(Z=Z_i)$. For auxiliary information of the form $\mathsf{E}_N\{u(x)\}=(1/N)\sum\limits_{i=1}^{N}u(x_i)=0$, where $u(\cdot)$ is known, the problem then reduces to maximizing (2.1) subject to

$$\sum_{i\in S}p_i=1,\qquad \sum_{i\in S}p_iu(x_i)=0,\qquad (0\le p_i\le 1). \tag{2.2}$$

For example, $u(x_i)=x_i-\overline{X}_N$. Using the lagrange multiplier method it is easily show that, for any finite population parameters that can be written as $\theta_N=\theta(F_N)$, the resulting empirical maximum likelihood estimator (EMLM) is $\widehat{\theta}_n=\theta(\widehat{F}_n)$, $\widehat{F}_n=\sum\limits_{i\in S}\widehat{p}_i\delta z_i$, where δz_i is the point measure at z_i, $\widehat{p}_i=1/[n\{1+\lambda u(x_i)\}]$, for $i\in S$, and the lagrange multiplier λ is the solution to

$$\sum_{i\in S}\frac{u(x_i)}{\{1+\lambda u(x_i)\}}=0. \tag{2.3}$$

Chen and Qin show that the empirical likelihood method indeed has some desirable properties in this context.

Theorem 1 Suppose that as $\nu\longrightarrow\infty$, the population size N, sample size n, and $N-n$ go to infinity, and

$$\frac{1}{N}\{|u(x_i)|^3+\cdots+|u(x_N)|^3\},\qquad \frac{1}{N}\{|g(y_i)|^3+\cdots+|g(y_N)|^3)\},$$

have an upper bound independent of ν. Then

$$\frac{n^{1/2}(\widehat{\theta}_n-\theta)}{\sigma_\nu}\xrightarrow{d}N(0,1),$$

where $\sigma_\nu^2 = (1-n/N)(\sigma_g^2 - \sigma_{gu}^2/\sigma_u^2)$, $\sigma_g^2 = [1/(N-1)]\sum_{i=1}^{N}\{g(y_i)-\overline{g}\}^2$, $\sigma_u^2 = [1/(N-1)]\sum_{i=1}^{N}u^2(x_i)$, $\sigma_{gu} = [1/(N-1)]\sum_{i=1}^{N}\{g(y_i)-\overline{g}\}u(x_i)$, and $\overline{g} = (1/N)\sum_{i=1}^{N}g(y_i)$.

§ 3. EMLE in Stratified Sampling

In stratified sampling, the population of N units is partitioned into nonoverlapping subpopulations called strata, of size $N_1, \cdots, N_H$ units, respectively. For the purpose of illustration, suppose we have obtained srswor $\{Z_{hi}, i \in S_h\}$ with $Z_{hi} = (y_{hi}, x_{hi})^T$ from stratum h, and that the samples are selected independently across the strata. As in Chen and Qin (1993), the log-empirical likelihood for the above sampling scheme is defined as

$$l = \sum_{i=1}^{H}\sum_{i\in S_h}\log p_{hi}, \tag{3.1}$$

where $p_{hi} = \mathsf{P}(Z_h = Z_{hi})$. We consider the problem of estimating $\theta = \theta(F_N) = \mathsf{E}_N(g(y))$ with auxiliary information $\mathsf{E}_N\{u(x)\} = 0$. suppose we naively apply the method in section 2 without considering the fact that we have the additional information contained in the knowledge of $N_1, \cdots, N_H$. Let $\overline{g} = N^{-1}\sum_{h=1}^{H}\sum_{i=1}^{N_h}g(y_{hi})$, and $\sigma_g^2 = [1/(N-1)]\sum_{h=1}^{H}\sum_{i=1}^{N_h}\{g(y_{hi})-\overline{g}\}^2$, $\sigma_u^2 = [1/(N-1)]\sum_{h=1}^{H}\sum_{i=1}^{N_h}u^2(x_{hi})$, $\sigma_{gu} = [1/(N-1)]\sum_{h=1}^{H}\sum_{i=1}^{N_h}\{g(y_{hi})-\overline{g}\}u(x_{hi})$. Then from Theorem 1 we have

Corollary 1 Suppose that as $\nu \longrightarrow \infty$, N, $n = \sum_{h=1}^{H} n_h$, $N-n$ go to infinity, and both

$$\frac{1}{N}\sum_{h=1}^{H}\sum_{i=1}^{N_h}|u(x_{hi})|^3, \qquad \frac{1}{N}\sum_{h=1}^{H}\sum_{i=1}^{N_h}|g(y_{hi})|^3,$$

have an upper bound independent of ν. Then

$$\sigma_\nu^{-1}(\widehat{\theta}-\theta) \xrightarrow{d} N(0,1),$$

where $\sigma_\nu^2 = (1/n)(1-n/N)(\sigma_g^2 - \sigma_{gu}^2/\sigma_u^2)$, $N = \sum_{h=1}^{H} N_h$, $\widehat{\theta}$ is the EMLE of θ.

Obviously, this can not be a good method. To see why we call this application of Theorem 1 naive, note that $F_N = \sum_h W_h F_{N_h}$, where $F_{N_h} = (1/N_h)\sum_{i=1}^{N_h}\delta_{Z_{hi}}$, $W_h = N_h/N$ for $h = 1, \cdots, H$. So $\overline{Z}_N = \int Z\mathrm{d}F_N = \sum_{h=1}^{H} W_h \int Z\mathrm{d}F_{N_h}$ and $\int u(x)\mathrm{d}F_N = 0$ can be written $\sum_h W_h \int u(x)\mathrm{d}F_{N_H} = 0$. This knowledge of the form of F_N contained in W_h should be incorporated in constructing the EMLE. The empirical likelihood method is well-suited to incorporate auxiliary information and can accommodate this information contained in the population size for each stratum quite naturally. To see this, let $Z_i = (y_i, v_i^T)$ for $i = 1, \cdots, N$, where $v_i = (u(x_i),\ v_{1i}, \cdots, v_{Hi})^T$ and $v_{hi} = 1$ if $i \in S_h$ and 0 otherwise. Then $\overline{v}_N = (0, W_1, \cdots, W_H)^T$ is known. the problem then reduces to maximizing (3.1) subject to

$$\sum_{i\in S_h} p_{hi} = W_h \quad \text{for} \quad h=1,\cdots,H, \qquad \text{and} \qquad \sum_{h=1}^{H}\sum_{i\in S_h} p_{hi}u(x_{hi}) = 0, \tag{3.2}$$

and using the resulting $\widehat{p}_{hi}$ to get $\widehat{F}_n = \sum\limits_{h=1}^{H}\sum\limits_{i\in S_h} p_{hi}\delta_{Z_{hi}}$, and thus $\widehat{\theta} = \theta(\widehat{F}_n)$ a EMLE of $\theta = \theta(F_N)$. When no auxiliary information beyond the stratum sizes is available, the resulting EMLE of the population mean is the usual unbiased estimator of the mean $\overline{y}_{st}$. Suppose, instead, that $\mathsf{E}_N u(x) = (1/N)\sum\limits_{h=1}^{H}\sum\limits_{i=1}^{N_h} u(x_{hi}) = 0$ is also known. In this case, the empirical likelihood (3.1) should be maximized with restriction

$$\sum_{i\in S_h} p_{hi} = W_h \quad \text{for} \quad h=1,\cdots,H, \qquad \text{and} \qquad \sum_{h=1}^{H}\sum_{i\in S_h} p_{hi}u(x_{hi}) = \sum_{h=1}^{H} W_h\widetilde{u}_h = 0, \tag{3.3}$$

with $W_h\widetilde{u}_h = \sum\limits_{i\in S_h} p_{hi}u(x_{hi})$. Then, how do we solve it numerically? We have the following simple numerical method.

§4. Numerical Method for Stratified Random Sampling

Let $\widetilde{u}_h$ be a group of numbers such that $\sum\limits_h W_h\widetilde{u}_h = 0$. Using the lagrange multiplier method, we maximize (3.1) subject to

$$\sum_{i\in S_h} p_{hi} = W_h, \qquad \sum_{i\in S_h} p_{hi}u(x_{hi}) = W_h\widetilde{u}_h, \qquad h=1,\cdots,H.$$

The solutions are

$$\widehat{p}_{hi} = \frac{W_h}{n_h\{1+\lambda_h(u_{hi}-\widetilde{u}_h)\}},$$

with λ_h satisfying

$$\sum_{i\in S_h} \frac{W_h(u_{hi}-\widetilde{u}_h)}{n_h\{1+\lambda_h(u_{hi}-\widetilde{u}_h)\}} = 0, \tag{4.1}$$

where $u_{hi} = u(x_{hi})$. Clearly, the maximum of the original likelihood (3.1) equals

$$-\sum_{h=1}^{H}\sum_{i\in S_h} \log\{1+\lambda_h(u_{hi}-\widetilde{u}_h)\} - \sum_{h=1}^{H}\sum_{i\in S_h} \log n_h + \sum_{h=1}^{H}\sum_{i\in S_h} \log W_h,$$

with the best choice of feasible values of $\widetilde{u}_h$. Hence, the problem reduces to maximizing

$$-\sum_{h=1}^{H}\sum_{i\in S_h} \log\{1+\lambda_h(u_{hi}-\widetilde{u}_h)\},$$

with respect to the $\widetilde{u}_h$'s under the restriction $\sum\limits_{h=1}^{H} w_h\widetilde{u}_h = 0$. Note that, in this problem, λ_h is a function of $\widetilde{u}_h$ defined by (4.1). Using the lagrange multiplier method, we get the function

$$l(\widetilde{u}_1,\cdots,\widetilde{u}_H,t) = -\sum_{h=1}^{H}\sum_{i\in S_h} \log\{1+\lambda_h(u_{hi}-\widetilde{u}_h)\} - t\Big(\sum_{h=1}^{H} W_h\widetilde{u}_h\Big).$$

Taking derivatives with respect to $\widetilde{u}_h$ and setting to zero, we get

$$-\sum_{i\in S_h}\frac{\lambda_(u_{hi}-\widetilde{u}_h)}{\{1+\lambda_h(u_{hi}-\widetilde{u}_h)\}}-tW_h=n_h\lambda_h-tW_h=0,$$

where $\lambda'_h=\partial\lambda_h/\partial\widetilde{u}_h$. Hence, we obtain $\lambda_h=(W_h/n_h)t$ and

$$\sum_{i\in S_h}\frac{\frac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)}{\left\{1+t\frac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\right\}}=0,\tag{4.2}$$

for $h=1,\cdots,H$. The other equation is

$$\sum_{h=1}^{H}W_h\widetilde{u}_h=0.\tag{4.3}$$

For each given t, $\widetilde{u}_h$ takes a different value and thus we denote it as $\widetilde{u}_h(t)$. It can be obtained from (4.2) easily. In addition, it is simple to show that $\sum\limits_{h=1}^{H}W_h\widetilde{u}_h(t)$ is a monotone function of t. Hence, numerically, we need only increase or decrease the size of t to determine the existence of the solution and the uniqueness is a simple consequence of the monotonicity.

§ 5. Asymptotic Properties

We now study the asymptotic properties of the empirical maximum likelihood estimator $\widehat{\theta}$. Let

$$\sigma_{hgu}=\frac{1}{N_h-1}\sum_{i=1}^{N_h}(g(y_{hi})-\overline{G}_h)(u_{hi}-\overline{U}_h),$$
$$\sigma_{hg}^2=\frac{1}{N_h-1}\sum_{i=1}^{N_h}(g(y_{hi})-\overline{G}_h)^2,$$
$$\sigma_{hu}^2=\frac{1}{N_h-1}\sum_{i=1}^{N_h}(u_{hi}-\overline{U}_h)^2,$$

where $\overline{G}_h=(1/N_h)\sum\limits_{i=1}^{N_h}g(y_{hi})$, and $\overline{U}_h=(1/N_h)\sum\limits_{i=1}^{N_h}u_{hi}$. Let $\overline{g}_h$, $\overline{u}_h$, S_{hg}^2, S_{hu}^2 and S_{hgu} denote the corresponding sample versions. Suppose that there is a sequence of finite populations indexed by ν such that when $\nu\longrightarrow\infty$:

(i) N_h, n_h go to infinity, n_h/N_h goes to zero, $\max\{n_h^{-1}W_h\}=O(n^{-1})$;

(ii) $0<c_1\le\sum\limits_{h=1}^{H}W_h\sigma_{hu}^2\le c_2<\infty$;

(iii) $N^{-1}\sum\limits_{h=1}^{H}\sum\limits_{i=1}^{N_h}|u_{hi}|^3=O(1)$;

(iv) $N^{-1}\sum\limits_{h=1}^{H}\sum\limits_{hi=1}^{N_h}|g(y_{hi})|^3=O(1)$.

Theorem 2 Under conditions (i)–(iv) above, and the stratum size information is incorporated, we have

$$\sigma_\nu^{-1}(\widehat{\theta}-\theta)\xrightarrow{d}N(0,1),$$

where $\sigma_\nu^2=\sum\limits_{h=1}^{H}W_h^2(n_h^{-1}-N_h^{-1})\sigma_{hg}^2-\Big\{\Big[\sum\limits_{h=1}^{H}W_h^2(n_h^{-1}-N_h^{-1})\sigma_{hgu}\Big]^2\Big/\Big[\sum\limits_{h=1}^{H}W_h^2(n_h^{-1}-N_h^{-1})\sigma_{hu}^2\Big]\Big\}$.

Proof For any $h=1,\cdots,H$, we have

$$\sum_{i\in S_h}\frac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)=t\sum_{i\in S_h}\Big\{\Big[\frac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]^2\Big/\Big[1+t\frac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]\Big\},$$

and hence

$$\sum_{h=1}^{H}\frac{W_h}{n_h}\sum_{i\in S_h}(u_{hi}-\widetilde{u}_h)=t\sum_{h=1}^{H}\sum_{i\in S_h}\Big\{\Big[\frac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]^2\Big/\Big[1+t\frac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]\Big\}.$$

So we have

$$|t|\le\Big(1+\frac{W_h}{n_h}|t|u^*\Big)\cdot\Big\{\Big|\sum_{h=1}^{H}\frac{W_h}{n_h}\sum_{i\in S_h}(u_{hi}-\widetilde{u}_h)\Big|\Big/\Big[\sum_{h=1}^{H}\frac{W_h^2}{n_h^2}\sum_{i\in S_h}(u_{hi}-\widetilde{u}_h)^2\Big]\Big\},$$

where $u^*=\max\{|u_{hi}|:i\in S_h\}$. Since

$$0=\sum_{h=1}^{H}W_h\widetilde{u}_h=\sum_{h=1}^{H}W_h(\widetilde{u}_h-\overline{u}_h)+\sum_{h=1}^{H}W_h\overline{u}_h.$$

Note that the second term has mean zero and variance $\sum\limits_{h=1}^{H}W_h^2(n_h^{-1}-N_h^{-1})\sigma_{hu}^2=O(\max n_h^{-1}W_h)=O(n^{-1})$ by assumption (i), (ii). Thus the second term is of order $n^{-1/2}$, and, consequently, the first term is of the same order, i.e. $\Big|\sum\limits_{h=1}^{H}(W_h/n_h)\sum\limits_{i\in S_h}(u_{hi}-\widetilde{u}_h)\Big|=\Big|\sum\limits_{h=1}^{H}W_h(\widetilde{u}_h-\overline{u}_h)\Big|=O(n^{-1/2})$. Applying this to the inequality for $|t|$, we find $t=O_p(n^{-1/2})$. The third moment condition implies $u^*=O_p(n^{1/2})$. Hence $t(W_h/n_h)(u_{hi}-\widetilde{u}_h)=o_p(1)$ uniformly over sampled units. We therefore get

$$\begin{aligned}t&=\Big[\sum_{h=1}^{H}W_h(\overline{u}_h-\widetilde{u}_h)\Big]\Big/\Big[\sum_{h=1}^{H}W_h^2n_h^{-1}S_{hu}^2\Big]+o_p(n^{-1/2})\\&=\Big(\sum_{h=1}^{H}W_h^2n_h^{-1}S_{hu}^2\Big)^{-1}\overline{u}_{st}+o_p(n^{-1/2}),\end{aligned}$$

where $\overline{u}_{st}=\sum\limits_{h=1}^{H}W_h\overline{u}_h$. With this expansion for t and the relation $\widehat{p}_{hi}=(W_h/n_h)/\{1+t(w_h/n_h)\cdot(u_{hi}-\widetilde{u}_h)\}$, we obtain

$$\begin{aligned}\widehat{\theta}&=\sum_{h=1}^{H}\sum_{i\in S_h}\widehat{p}_{hi}g(y_{hi})=\sum_{h=1}^{H}\frac{W_h}{n_h}\sum_{i\in S_h}\Big\{g(y_{hi})\Big/\Big[1+\frac{W_h}{n_h}t(u_{hi}-\widetilde{u}_h)\Big]\Big\}\\&=\sum_{h=1}^{H}\frac{W_h}{n_h}\sum_{i\in S_h}\Big\{1-\Big(\sum_{h=1}^{H}W_h^2n_h^{-1}S_{hu}^2\Big)^{-1}\overline{u}_{st}\frac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big\}g(y_{hi})+o_p(n^{-1/2})\\&=\overline{g}_{st}-\Big(\sum_{h=1}^{H}W_h^2n_h^{-1}S_{hu}^2\Big)^{-1}\Big(\sum_{h=1}^{H}W_h^2n_h^{-1}S_{hgu}\Big)\overline{u}_{st}+o_p(n^{-1/2})\\&=\overline{g}_{st}-\Big[\sum_{h=1}^{H}W_h^2(n_h^{-1}-N_h^{-1})\sigma_{hu}^2\Big]^{-1}\Big[\sum_{h=1}^{H}W_h^2(n_h^{-1}-N_h^{-1})\sigma_{hgu}\Big]\overline{u}_{st}+o_p(n^{-1/2}).\end{aligned}$$

Using the central limit theorem established by Bickel and Freedman (1984), we prove the theorem. #

This result shows that, whenever there is any auxiliary information, the asymptotic variance of $\widehat{\theta}$ with respect to stratified srswor is always smaller than or equal to $\mathsf{Var}(\overline{g}_{st})$. We also note that the reduction of the asymptotic variance depends on the relevance of the auxiliary information. The larger the correlation between $u(X)$ and $g(Y)$, the greater the gain in precision. In particular, when $g(y) = y$, $u(x) = x - \overline{X}_N$, the EMLE of $\overline{Y}_N$ is asymptotically equivalent to the optimal linear estimator given by Rao (1994).

There can be many ways to estimate σ_ν^2. However, one might apply resampling variance estimators such as the jackknife, bootstrap and balanced repeated replications (see Shao and Wu (1989) and (1992), Chen and Qin (1993), Shao (1994)) directly to $\widehat{\theta}$, recalculating the $\widehat{p}_{hi}$ for each resample. These may perform better for finite samples since they are applied directly to $\widehat{\theta}$. As an example, we consider the jackknife estimator of σ_ν^2. A result on its large sample property is given in the following theorem.

Theorem 3 Under the same conditions as Theorem 2, let $\widehat{\theta}_{-kj}$ be the estimator of θ when the jth unit from kth strata removed, and let

$$S_{JE}^2 = \sum_{k=1}^{H}(1-f_k)n_k^{-1}(n_k-1)\sum_{j\in S_k}(\widehat{\theta}-\widehat{\theta}_{-kj})^2$$

be the jackknife variance estimator, where $f_k = n_k/N_k$. Then S_{JE}^2 is consistent.

Proof The lagrange multiplier with jth unit from kth strata removed, t_{-kj} satisfying

$$\sum_{\substack{h=1\\h\neq k}}^{H}\frac{W_h}{n_h}\sum_{i\in S_h}\frac{u_{hi}-\widetilde{u}_h}{1+t_{-kj}\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)} + \frac{W_k}{n_k-1}\sum_{\substack{i\in S_k\\i\neq j}}\frac{u_{ki}-\widetilde{u}_k}{1+t_{-kj}\dfrac{W_k}{n_k}(u_{kj}-\widetilde{u}_k)} = 0.$$

Thus

$$\sum_{h=1}^{H}\frac{W_h}{n_h}\sum_{i\in S_h}\frac{u_{hi}-\widetilde{u}_h}{1+t_{-kj}\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)} + \frac{W_k}{n_k(n_k-1)}\sum_{i\in S_h}\frac{u_{ki}-\widetilde{u}_k}{1+t_{-kj}\dfrac{W_k}{n_k}(u_{ki}-\widetilde{u}_k)}$$
$$-\Big(\frac{W_k}{n_k-1}\Big)\Big[(u_{kj}-\widetilde{u}_k)\Big/\Big(1+t_{-kj}\frac{W_k}{n_k}(u_{kj}-\widetilde{u}_k)\Big)\Big] = 0,$$

and also,

$$\sum_{h=1}^{H}\frac{W_h}{n_h}\sum_{i\in S_h}\frac{u_{hi}-\widetilde{u}_h}{1+t\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)} = 0.$$

Therefore

$$\sum_{h=1}^{H}\frac{W_h}{n_h}^2\sum_{i\in S_h}\frac{(u_{hi}-\widetilde{u}_h)^2}{\Big[1+t_{-kj}\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]\Big[1+t\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]}(t_{-kj}-t)$$
$$= \frac{W_k}{n_k(n_k-1)}\sum_{i\in S_k}\frac{(u_{ki}-\widetilde{u}_k)}{1+t_{-kj}\dfrac{W_k}{n_k}(u_{ki}-\widetilde{u}_k)} - \Big(\frac{W_k}{n_k-1}\Big)\Big(\frac{u_{kj}-\widetilde{u}_k}{1+t_{-kj}\dfrac{W_k}{n_k}(u_{kj}-\widetilde{u}_k)}\Big). \quad (5.1)$$

Similarly

$$\begin{aligned}
&\widehat{\theta}-\widehat{\theta}_{-kj}\\
=&\ \sum_{h=1}^{H}\frac{W_h}{n_h}\sum_{i\in S_h}\frac{g(y_{hi})}{1+t\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)}\\
&-\Big\{\sum_{\substack{h=1\\h\neq k}}^{H}\frac{W_h}{n_h}\sum_{i\in S_h}\frac{g(y_{hi})}{\Big[1+t_{-kj}\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]}+\frac{W_k}{n_k-1}\sum_{\substack{i\in S_k\\i\neq j}}\frac{g(y_{ki})}{1+\lambda_{-kj}\dfrac{W_k}{m_k}(u_{ki}-\widetilde{u}_k)}\Big\}\\
=&\ \sum_{h=1}^{H}\frac{W_h^2}{n_h}\sum_{i\in S_h}\frac{u_{hi}g(y_{hi})}{\Big[1+t_{-kj}\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]\Big[1+t\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]}(t_{-kj}-t)\\
&-\frac{W_k}{n_k(n_k-1)}\sum_{i\in S_k}\frac{g(y_{ki})}{1+t_{-kj}\dfrac{W_k}{n_k}(u_{ki}-\widetilde{u}_k)}+\Big(\frac{W_k}{n_k-1}\Big)\Big(\frac{g(y_{kj})}{1+t_{-kj}\dfrac{W_k}{n_k}(u_{kj}-\widetilde{u}_k)}\Big).
\end{aligned}\tag{5.2}$$

Note that $t=O_p(n^{-1/2})$, and similarly it can be shown that $t_{-kj}=O_p(n^{-1/2})$ uniformly. letting $S_u^2=\sum\limits_{h=1}^{H}(W_h^2/n_h)S_{hu}^2$, $S_{gu}=\sum\limits_{h=1}^{h}(W_h^2/n_h)S_{hgu}$, we get

$$\sum_{h=1}^{H}\frac{W_h^2}{n_h^2}\sum_{i\in S_h}\frac{(u_{hi}-\widetilde{u}_h)^2}{\Big[1+t_{-kj}\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]\Big[1+t\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]}=S_u^2(1+o_p(1)),\tag{5.3}$$

and

$$\sum_{h=1}^{H}\frac{W_h^2}{n_h^2}\sum_{i\in S_h}\frac{g(y_{hi})(u_{hi}-\widetilde{u}_h)}{\Big[1+t_{-kj}\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]\Big[1+t\dfrac{W_h}{n_h}(u_{hi}-\widetilde{u}_h)\Big]}=S_{gu}(1+o_p(1)),\tag{5.4}$$

uniformly in (k,j). It is not difficult to show that replacing $\sum\limits_{i\in S_k}\{g(y_{ki})/[1+t_{-kj}(W_k/n_k)(u_{ki}-\widetilde{u}_k)]\}$ by $\sum\limits_{i\in S_k}g(y_{ki})$ and $g(y_{kj})/[1+t_{-kj}(W_k/n_k)u_{kj}]$ by $g(y_{kj})$ in the expression of $\widehat{\theta}-\widehat{\theta}_{-kj}$ has negligible effect on the jackknife variance estimator. By ignoring these higher order term and using (5.1) and (5.2) we get

$$S_u^2(t_{-kj}-t)\doteq-(n_k-1)^{-1}W_k(u_{kj}-\overline{u}_k),$$

where $\overline{u}_k=(1/n_k)\sum\limits_{i\in S_k}u_{ki}$, and thus by (5.2) and (5.4),

$$\widehat{\theta}-\widehat{\theta}_{-kj}\doteq(n_k-1)^{-1}W_k[(g(y_{kj})-\overline{g}_k)-S_{gu}S_u^{-2}(u_{kj}-\overline{u}_k)],$$

where $\overline{g}_k=(1/n_k)\sum\limits_{i\in S_k}g(y_{ki})$. Therefore

$$S_{JE}^2\doteq\sum_{k=1}^{H}(1-f_k)\frac{W_k^2}{n_k(n_k-1)}\sum_{j\in S_k}[(g(y_{kj})-\overline{g}_k)-S_{gu}S_u^{-2}(u_{kj}-\overline{u_k})]^2.$$

Which implies the desired result.

References

[1] Bickel, P.J. and Freedman, D.A., Asymptotic normality and the bootstrap in stratified sampling, *Ann. Statist.*, **12**(1984), 470–482.

[2] Chen, J. and Qin, J., Empirical likelihood estimation for finite populations and the effective usage of auxiliary information, *Biometrika*, **80**(1993), 107–116.

[3] Chen, J. and Sitter, R.R., A pseudo empirical likelihood approach to the effective use of auxiliary information in complex surveys, *Statistica Sinica*, **9**(1999), 384–406.

[4] Hartley, H.O. and Rao, J.N.K., A new estimation theory for sample surveys, *Biometrika*, **55**(1968), 547–557.

[5] Owen, A.B., Empirical likelihood confidence regions, *Ann. Statist.*, **18**(1990), 90–120.

[6] Rao, J.N.K., Variance estimation in sample surveys, *Handbook of Statistics* (Edited by Krishnaiah, P.K. and Rao, C.R.) 6, 427–447, Elsevier, North-Holland, Amsterdam, 1988.

[7] Shao, J., L-statistics in complex survey problems, *Ann. statist.*, **22**(1994), 946–967.

[8] Shao, J. and Wu, C.F.J., Asymptotic properties of the balanced repeated replication method for sample quartiles, *Ann. statist.*, **20**(1992), 1571–1593.

[9] Zhong, B., Chen, J. and Rao, J.N.K., Empirical likelihood inference in the presence of measurement error, *The Canadian Journal of Statistics*, **28**(2000), 841–852.

在分层抽样下有效利用辅助信息及含于层总体大小中的信息的经验似然方法

伍长春 [1,2]　　张润楚 [2]

([1] 嘉兴学院数学与信息科学学院，嘉兴， 314001)

([2] 南开大学数学科学学院统计系和教育部重点实验室，天津， 300071)

本文考虑用在分层随机抽样下的经验似然方法来获得有限总体参数的估计量．我们指出，经验似然方法非常自然地结合辅助信息和含于层总体大小中的信息．我们的结果显示，由经验似然方法可获得有效估计．

关 键 词：经验似然，分层随机抽样，辅助信息，层总体大小，刀切法估计.

学 科 分 类 号：O212.2, O212.7.