Improvement on Estimating Quantiles in Finite Population Using Indirect Methods of Estimation

M. Rueda García[1], A. Arcos Cebrián[1], E. Artés Rodríguez[2]

[1] Department of Statistics and Operations Research
University of Granada, 18071 Granada, Spain
[2] Department of Statistics and Operations Research
University of Almería, 04120 Almería, Spain

Abstract. New methods for estimating confidence limits for quantiles in a finite population are proposed. These methods use auxiliary information through the ratio, difference and regression estimator of the population distribution function. They may be applied to any type of sampling. Simulation studies based of two real populations show that the methods proposed in this paper can be considerably more efficient than the customary classic method.
Key words: auxiliary information, finite population quantiles, ratio, difference and regression type estimator, confidence intervals.

1 Introduction

In survey practice, it is often of interest to study variables with a highly skewed distribution. In such situations, it is useful to make inferences about finite population quantiles. Sample medians have long been recognized as simple robust alternatives to sample means, for estimating location of heavy-tailed or markedly skewed populations from simple random samples. A large class of robust estimates of location, including the sample median, was investigated in the Princeton simulation study (D.F. Andrew et al (1972)). Although the sample median did not emerge as best estimate in many nonstandard populations simulated in the study, its robustness in small samples for medium and large deviations from normality was clearly demonstrated. Its simplicity relative to other robust estimates, indicated its choice. Unfortunately, while there is an extensive literature on the estimation of means and totals, relatively less research has been done to development of efficient methods for estimating finite population quantiles. Moreover, most of these methods in simple random sampling (*Gross* 1980, *Sedransk* and *Meyer* 1978, *Smith* and *Sedransk* 1983) does not make explicit use of auxiliary variable is available, it is natural to expect that the auxiliary information can be incorporated to construct an estimator more efficient than the direct estimator (sample quantile).

X. Liu, P. Cohen, M. Berthold (Eds.): "Advances in Intelligent Data Analysis" (IDA-97)
LNCS 1280, pp. 491–499, 1997.

The use of indirect methods for estimating a finite population mean has been widely studied (see *Cochran* 1977), however, it is not immediately clear how these well–established techniques, such as the regression estimator, can be extended to the case of estimating the quantiles.

Increasingly, this need is being recognized, so point estimation of finite population quantiles that uses auxiliary information has received considerable attention (*Chambers* and *Dunstan*, 1986, and *Rao*, *Kovar* and *Mantel*, 1990), both suggested estimating quantiles by in inverting improved estimates of the distribution functions in presence of auxiliary information. Other references are *Kuk* and *Mak* (1989,94), *Mak* and *Kuk* (1993).

In this paper, we suggest alternative procedures for determining confidence intervals for a finite population median and other quantiles, under simple random sampling and using an auxiliary variable.

Let $y_1, y_2, \ldots, y_N$ denote the values of the population elements U_1, U_2, ..., U_N, for the variable of interest y. For any y $(-\infty < y < \infty)$, as the population distribution $F_Y(y)$ is defined as the proportion of elements in the population that are less than or equal to y.

The finite population β quantile is defined as

$$Q_Y(\beta) = F_Y^{-1}(\beta),$$

where F_Y^{-1} is the inverse function of F_Y.

The general procedure to estimate the population quantile $Q_Y(\beta)$, using data y_k for $k \in s$, where s is a simple random sample can be summarized as follows: we first produce an estimated distribution function, $\widehat{F}_Y(y)$, and then, estimate $Q_Y(\beta) = F_Y^{-1}(\beta)$ as $\widehat{Q}_Y(\beta) = \widehat{F}_Y^{-1}(\beta)$, where the inverse $\widehat{F}_Y^{-1}$ is to be understood in the same way as F_Y^{-1} above. This method has been in use for a long time; the first published account is probably *Woodruff* (1952).

Woodruff (1952) describes a general method of obtaining confidence intervals for medians and others positions measures using a principle that has been applied to sample random sampling and extending it to any type of sampling. These confidence limits can be approximated for any sampling design where the variance of the percentage of items less than a stated value can be acceptably estimated (in general, where large samples are involved).

We present new methods to derive the confidence interval finite population quantiles in Section 2 and 3. Ratio, difference and regression estimators of the population distribution function based on an auxiliary variable is the key to these methods. We also compare the methods that we propose and Woodruff's method using simulation studies, in Section 4.

2 Confidence intervals for the quantiles using ratio, difference and regression estimators

Consider the simple random sampling design. Suppose that the population under study consists of N units, and attached to each of these units are the values of the survey variable y of interest and an auxiliary variable x. It's assumed that only the population βth quantile $Q_X(\beta)$ of x is known and that $Q_Y(\beta)$ is to be estimated on the basis of a simple random sample of size n. Let $(x_1, y_1), \ldots, (x_n, y_n)$ be the associated values of the variables x and y for the units in the sample.

Consider the ratio, difference and regression estimators

$$\widehat{F}_R(Q_Y(\beta)) = \frac{\widehat{F}_Y(Q_Y(\beta))}{\widehat{F}_X(Q_X(\beta))}\beta$$

$$\widehat{F}_D(Q_Y(\beta)) = \widehat{F}_Y(Q_Y(\beta)) + \left(\beta - \widehat{F}_X(Q_X(\beta))\right)$$

$$\widehat{F}_{Reg}(Q_Y(\beta)) = \widehat{F}_Y(Q_Y(\beta)) + b\left(\beta - \widehat{F}_X(Q_X(\beta))\right)$$

(b is a known constant) and the constants c_1^i and c_2^i, $i = 1, 2, 3$ such that

$$P\left\{c_1^1 \le \widehat{F}_R(Q_Y(\beta)) \le c_2^1\right\} = 1 - \alpha.$$

$$P\left\{c_1^2 \le \widehat{F}_D(Q_Y(\beta)) \le c_2^2\right\} = 1 - \alpha.$$

$$P\left\{c_1^3 \le \widehat{F}_{Reg}(Q_Y(\beta)) \le c_2^3\right\} = 1 - \alpha.$$

Thus, the approximated $100(1-\alpha)\%$ confidence intervals for $Q_Y(\beta)$ will be

$$\left[\widehat{F}_Y^{-1}\left(c_1^1 \frac{\widehat{F}_X(Q_X(\beta))}{\beta}\right), \widehat{F}_Y^{-1}\left(c_2^1 \frac{\widehat{F}_X(Q_X(\beta))}{\beta}\right)\right]$$

$$\left[\widehat{F}_Y^{-1}\left(c_1^2 - \left(\beta - \widehat{F}_X(Q_X(\beta))\right)\right), \widehat{F}_Y^{-1}\left(c_2^2 - \left(\beta - \widehat{F}_X(Q_X(\beta))\right)\right)\right]$$

$$\left[\widehat{F}_Y^{-1}\left(c_1^3 - b\left(\beta - \widehat{F}_X(Q_X(\beta))\right)\right), \widehat{F}_Y^{-1}\left(c_2^3 - b\left(\beta - \widehat{F}_X(Q_X(\beta))\right)\right)\right].$$

For large samples, $\widehat{F}_Y(Q_Y(\beta))$, $\widehat{F}_X(Q_X(\beta))$ and $\widehat{F}_R(Q_Y(\beta))$ are approximately normally distributed (see *Kuk* and *Mak*, 1989). Then, the asymptotic distributions of the estimation $\widehat{F}_D(Q_Y(\beta))$ and $\widehat{F}_{Reg}(Q_Y(\beta))$ approach a normal distribution, and we would choose the smallest confidence interval as

$$c_1^1 = \beta - z_{\frac{\alpha}{2}}\left\{V\left(\widehat{F}_R(Q_Y(\beta))\right)\right\}^{\frac{1}{2}} \quad , \quad c_2^1 = \beta + z_{\frac{\alpha}{2}}\left\{V\left(\widehat{F}_R(Q_Y(\beta))\right)\right\}^{\frac{1}{2}},$$

$$c_1^2 = \beta - z_{\frac{\alpha}{2}}\left\{V\left(\widehat{F}_D(Q_Y(\beta))\right)\right\}^{\frac{1}{2}} \quad , \quad c_2^2 = \beta + z_{\frac{\alpha}{2}}\left\{V\left(\widehat{F}_D(Q_Y(\beta))\right)\right\}^{\frac{1}{2}},$$

and

$$c_1^3 = \beta - z_{\frac{\alpha}{2}} \left\{ V\left(\widehat{F}_{Reg}(Q_Y(\beta)) \right) \right\}^{\frac{1}{2}} \quad , \quad c_2^3 = \beta + z_{\frac{\alpha}{2}} \left\{ V\left(\widehat{F}_{Reg}(Q_Y(\beta)) \right) \right\}^{\frac{1}{2}} .$$

We don't know $Q_Y(\beta)$, then the problem of evaluating the last unknown variances is not so simple. For example, to evaluate the variance $V\left(\widehat{F}_R(Q_Y(\beta))\right)$, we make the variables

$$e_0 = \frac{\widehat{F}_Y(Q_Y(\beta)) - F_Y(Q_Y(\beta))}{F_Y(Q_Y(\beta))}, \quad e_1 = \frac{\widehat{F}_X(Q_X(\beta)) - F_X(Q_X(\beta))}{F_X(Q_X(\beta))}.$$

Then, Taylor's series expansion yields

$$V\left(\widehat{F}_R(Q_Y(\beta))\right) \simeq F_Y(Q_Y(\beta))^2 \left(E\left(e_0^2\right) + E\left(e_1^2\right) - 2E\left(e_1 e_0\right)\right) =$$

$$= \left(V\left(\widehat{F}_X(Q_X(\beta))\right) + V\left(\widehat{F}_Y(Q_Y(\beta))\right) - 2\mathrm{Cov}\left(\widehat{F}_X(Q_X(\beta)), \widehat{F}_Y(Q_Y(\beta))\right)\right) =$$

$$= 2\frac{1-f}{n}\beta(1-\beta) - 2\mathrm{Cov}\left(\widehat{F}_X(Q_X(\beta)), \widehat{F}_Y(Q_Y(\beta))\right). \tag{1}$$

We have to calculate the value of $\mathrm{Cov}\left(\widehat{F}_X(Q_X(\beta)), \widehat{F}_Y(Q_Y(\beta))\right)$, therefore we consider the two–way classification

	$x_k \le Q_X(\beta)$	$x_k > Q_X(\beta)$	
$y_k \le Q_Y(\beta)$	$n_{11} \setminus N_{11}$	$n_{12} \setminus N_{12}$	$N_{1\cdot}$
$y_k > Q_Y(\beta)$	$n_{21} \setminus N_{21}$	$n_{22} \setminus N_{22}$	$N_{2\cdot}$,
	$N_{\cdot 1}$	$N_{\cdot 2}$	

where n_{11} denotes the number of units in the sample with $x \le Q_X(\beta)$ and $y \le Q_Y(\beta)$; and N_{11} is the number of units in the population with $x \le Q_X(\beta)$ and $y \le Q_Y(\beta)$. Thereby,

$$(n_{11}, n_{12}, n_{21}, n_{22}) \simeq HG(N, n, N_{11}, N_{12}, N_{21}),$$

$n\widehat{F}_Y(Q_Y(\beta)) = n_{11} + n_{12}$ and similarly $n\widehat{F}_X(Q_X(\beta)) = n_{11} + n_{21}$.

Besides, we can verify that

$$\mathrm{Cov}\left(n\widehat{F}_Y(Q_Y(\beta)), n\widehat{F}_X(Q_X(\beta))\right) = \frac{N-n}{N-1}\frac{n}{N^2}\left(N_{11}N_{22} - N_{12}N_{21}\right).$$

Substituting the last expression in (1) we have

$$V\left(\widehat{F}_R(Q_Y(\beta))\right) = \frac{1-f}{n}2\left(\beta(1-\beta) - \frac{N_{11}N_{22} - N_{12}N_{21}}{N^2}\right). \tag{2}$$

Denoting Cramer's V coefficient as

$$\phi_\beta = \frac{N_{11}N_{22} - N_{12}N_{21}}{\sqrt{N_{1\cdot}N_{2\cdot}N_{\cdot 1}N_{\cdot 2}}},$$

we can rewrite the following expression

$$V\left(\widehat{F}_R(Q_Y(\beta))\right) = \frac{1-f}{n} 2\beta(1-\beta)(1-\phi_\beta). \tag{3}$$

Analogously, we evaluate the variances of $\widehat{F}_D(Q_Y(\beta))$ and $\widehat{F}_{Reg}(Q_Y(\beta))$ which are given by (3) and

$$V\left(\widehat{F}_{Reg}\left(Q_Y(\beta)\right)\right) = \frac{1-f}{n}\beta(1-\beta)(1+b^2-2b\phi_\beta). \tag{4}$$

In practice ϕ_β is unobservable since $Q_Y(\beta)$ is unknown and therefore has to be estimated from the sample. Substituting n_{ij} for $\tilde{n}_{ij}$, based on a similar cross–classification

$$\begin{array}{lll} & x_k \le \widehat{Q}_X(\beta) & x_k > \widehat{Q}_X(\beta) \\ y_k \le \widehat{Q}_Y(\beta) & \tilde{n}_{11} \setminus \tilde{N}_{11} & \tilde{n}_{12} \setminus \tilde{N}_{12} \\ y_k > \widehat{Q}_Y(\beta) & \tilde{n}_{21} \setminus \tilde{N}_{21} & \tilde{n}_{22} \setminus \tilde{N}_{22} \end{array},$$

and then, we would consider the following estimator for ϕ_β

$$\tilde{\phi}_\beta = \frac{\tilde{n}_{11}\tilde{n}_{22} - \tilde{n}_{12}\tilde{n}_{21}}{\sqrt{\tilde{n}_{1\cdot}\tilde{n}_{2\cdot}\tilde{n}_{\cdot 1}\tilde{n}_{\cdot 2}}}.$$

So, the intervals
$\left[\widehat{F}_Y^{-1}\left(\tilde{c}_1^1 \frac{\widehat{F}_X(Q_X(\beta))}{\beta}\right), \widehat{F}_Y^{-1}\left(\tilde{c}_2^1 \frac{\widehat{F}_X(Q_X(\beta))}{\beta}\right)\right]$,
$\left[\widehat{F}_Y^{-1}\left(\tilde{c}_1^2 - \left(\beta - \widehat{F}_X(Q_X(\beta))\right)\right), \widehat{F}_Y^{-1}\left(\tilde{c}_2^2 - \left(\beta - \widehat{F}_X(Q_X(\beta))\right)\right)\right]$ and
$\left[\widehat{F}_Y^{-1}\left(\tilde{c}_1^3 - b\left(\beta - \widehat{F}_X(Q_X(\beta))\right)\right), \widehat{F}_Y^{-1}\left(\tilde{c}_2^3 - b\left(\beta - \widehat{F}_X(Q_X(\beta))\right)\right)\right]$ where

$$\tilde{c}_i^1(\beta) = \tilde{c}_i^2(\beta) = \beta + (-1)^i z_{\frac{\alpha}{2}} \left\{\frac{1-f}{n} 2\beta(1-\beta)(1-\tilde{\phi}_\beta)\right\}^{\frac{1}{2}} \quad i = 1, 2,$$

$$\tilde{c}_i^3(\beta) = \beta + (-1)^i z_{\frac{\alpha}{2}} \left\{\frac{1-f}{n}\beta(1-\beta)(1+b^2-2b\tilde{\phi}_\beta)\right\}^{\frac{1}{2}} \quad i = 1, 2,$$

are $100(1-\alpha)\%$ confidence intervals for $Q_Y(\beta)$. When the sample size is small, the method should be applied with caution, as this method relies on several approximations.

3 Confidence intervals for quantiles using the optimum regression estimator

The optimum regression estimator, $\widehat{F}_{Reg}^{opt}(Q_Y(\beta))$, that is, the regression type estimator with the smallest variance, is obtained in this section. The variance (4) is minimum to $b = \phi_\beta$, and then we consider the regression type estimator for the population distribution function as follows:

$$\widehat{F}_{Reg}^{opt}(Q_Y(\beta)) = \widehat{F}_Y(Q_Y(\beta)) + \phi_\beta\left(\beta - \widehat{F}_X(Q_X(\beta))\right),$$

and its variance is given by

$$V\left(\widehat{F}_{Reg}^{opt}(Q_Y(\beta))\right) = \frac{1-f}{n}\beta(1-\beta)(1-\phi_\beta^2).$$

This regression type estimator is always more precise than the simple estimator $\widehat{F}_Y(Q_Y(\beta))$ (except to $\phi_\beta = 0$), although it has the same difficulty that the previous estimators since ϕ_β is unobservable, because $Q_Y(\beta)$ is unknown. To resolve this difficulty, we take the coefficient estimation corresponding to ϕ_β, but with $Q_Y(\beta)$ and $Q_X(\beta)$ replaced by $\widehat{Q}_Y(\beta)$ and $\widehat{Q}_X(\beta)$, respectively, which gives

$$\widehat{F}_{Reg}^{*}(Q_Y(\beta)) = \widehat{F}_Y(Q_Y(\beta)) + \tilde{\phi}_\beta\left(\beta - \widehat{F}_X(Q_X(\beta))\right).$$

Now, the asymptotic distribution of $\widehat{F}_{Reg}^{*}(Q_Y(\beta))$ can be derived through the following reasoning: if p_{11} denotes the proportions of units in the sample with $x \leq \widehat{Q}_X(\beta)$ and $y \leq \widehat{Q}_Y(\beta)$, and P_{11} the proportions of units in the populations with $x \leq Q_X(\beta)$ and $y \leq Q_Y(\beta)$, it can be see that

$$\widehat{F}_{Reg}^{*}(Q_Y(\beta)) = \widehat{F}_Y(Q_Y(\beta)) + \frac{p_{11} - \beta^2}{\beta(1-\beta)}\left(\beta - \widehat{F}_X(Q_X(\beta))\right).$$

Since $\widehat{F}_X(Q_Y(\beta)) \to \beta$ in probability and $p_{11} - P_{11}$ is of order $O_p\left(n^{-\frac{1}{2}}\right)$ (see *Kuk* and *Mak*, 1989), then

$$\widehat{F}_{Reg}^{*}(Q_Y(\beta)) = \widehat{F}_{Reg}^{opt}(Q_Y(\beta)) + O_p\left(n^{-\frac{1}{2}}\right),$$

and $\widehat{F}_{Reg}^{*}(Q_Y(\beta))$ has the same asymptotic distribution of $\widehat{F}_{Reg}^{opt}(Q_Y(\beta))$. Hence $\widehat{F}_{Reg}^{*}(Q_Y(\beta))$ is asymptotically normal with mean β and variance

$$\frac{1-f}{n}\beta(1-\beta)(1-\phi_\beta^2).$$

Considering this new regression estimator we can derive a confidence interval for the βth quantile $Q_Y(\beta)$ as follows:

$$c_j^4 = \beta + (-1)^j z_{\frac{\alpha}{2}}\left\{\widehat{V}\left(\widehat{F}_{Reg}^{*}(Q_Y(\beta))\right)\right\}^{\frac{1}{2}}, \quad j = 1, 2,$$

where

$$\widehat{V}\left(\widehat{F}^{*}_{Reg}\left(Q_Y(\beta)\right)\right) = \frac{1-f}{n}\beta(1-\beta)(1-\tilde{\phi}^2_\beta).$$

Then,

$$\left[\widehat{F}_Y^{-1}\left(c_1^4 - \tilde{\phi}_\beta\left(\beta - \widehat{F}_X(Q_X(\beta))\right)\right), \widehat{F}_Y^{-1}\left(c_2^4 - \tilde{\phi}_\beta\left(\beta - \widehat{F}_X(Q_X(\beta))\right)\right)\right]$$

is a confidence interval with confidence coefficient $1-\alpha$ for $Q_Y(\beta)$.

4 Simulation study

To compare the efficiencies of the proposed methods and Woodfruff's method; we use simulation studies. Choose and fix a $1-\alpha$ level of confidence and a sample size n, consider 1000 samples of size n from the population and for each sample compute the length of the confidence intervals by several methods. The average length of 1000 samples yields information about the precision of each method. Furthermore, their variances yield information about the representatively of the means.

We carry out empirical studies using two finite populations, the first one being the block population (*Kish*, 1965). The data consist of 270 blocks, and Y and X in this example are respectively the number of rented houses and the number of houses in each block, respectively.

Table 1 shows the average length, $\bar{l}$, and the variance length, σ_l^2, of the confidence intervals built using Woodfruff's method (classical) and the ratio, difference and regression methods that we propose, for 1000 samples of size n, for $n = 30, 35, 40, 45, 50$ and 100 selected from the population for $Q_Y(0.5)$ and $100(1-\alpha)\% = 90\%, 95\%$ and 99%.

From table 1, we can see that for this population there is considerable improvement between the average length of the confidence intervals built using the methods proposed in this paper and the classical method, for any quantile and confidence coefficient. For example, if $100(1-\alpha)\%$=95% and $n = 50$, the average length of 1000 confidence intervals determined using the ratio, difference and regression methods are, respectively, 63, 59 and 49 percent of the average length of the respective confidence intervals constructed using the classical method. In this population, the variables Y and X are well correlated, and the concordance is high.

The second population (*Fernández* and *Mayor*, 1994) consist of 1500 households. In this example Y and X are the annual food costs and annual income, respectively. Table 2 shows the results of this second simulation study.

For two population we compute the proportion of intervals that contains the actual population quantile (Cove). We observe this variable doesn't differ a lot from the nominal coverage. Only in the case of the first population this difference is noteworthy for small samples with the regression method. As the sample size increases, the Cove variable is on the increase too and even surpasses the nominal coverage, as it happened in the first population, moreover the average length keeps being lower than the direct method.

In the two populations, we verify that the proposed methods determine more precise confidence intervals for finite population quantiles than Woodruff's method.

References

[1] Andrew, D.F. et al: Robust estimates of location-surveys and advances. Princeton University Press, (1972)

[2] Chambers, R. L., Dunstan, R.: Estimating distribution functions from survey data. Biometrika **73** (1986) 597–604

[3] Chambers, R. L., Dorfman, A. H. and Wehrly, T. E.: Bias robust estimation in finite population using nonparametric calibration. J. Amer. Statist. Assoc. **88** (1991) 268–277

[4] Cochran, A. H.: Sampling Techniques, Third Edition. Wiley, New York (1977)

[5] Fernández García, F. R., Mayor Gallego, J. A.: Muestreo en Poblaciones Finitas: Curso Básico. P.P.U., Barcelona (1994)

[6] Gross, S. T.: Median estimation in sample survey. Proc. Surv. Res. Meth. Sect. Amer. Statist. Ass. (1980) 181–184.

[7] Haskell, J., Sedransk, J.: Confidence interval for quantiles and tolerance intervals of populations. Unpublished Technical Report. SUNY at Albany Dept. of Mathematic Statistics (1980) Albany NY

[8] Kish, L.: Survey Sampling. John Wiley and Sons (1965) New York

[9] Kuk, A. Y. C., Mak, T. K.: Median estimation in presence of auxiliary information. J. R. Statist. Soc. B **51** (1989) 261–269

[10] Kuk, A. Y. C., Mak, T. K.: A functional approach to estimating finite population distribution functions. Commun. Statist. Theor. Meth. **23** (1994) 883–896

[11] Mak, T. K., Kuk, A. Y. C.: A new method for estimating finite–population quantiles using auxiliary information. The Canadian Journal of Statistics **21** (1993) 29–38

[12] Sedransk, J., Meyer, J.: Confidence intervals for the quantiles of a finite population: simple random and stratified simple random sampling. J. Amer. Statist. Ass. **76** (1978) 66–77

[13] Smith, P., Sedransk, J.: Lower bounds for confidence coefficients for confidence intervals for finite population quantiles. Commun. Statist. Theor. Meth. **12** (1983) 1329–1344

[14] Woodruff, R. S.: Confidence intervals for medians and other position measures. J. Amer. Statist. Assoc. **47** (1952) 635–646

Table 1. Block population. $Q_Y(0.5)$.

		100(1 − α)%=90%			100(1 − α)%=95%			100(1 − α)%=99%		
n	method	$\bar{l}$	Cove	σ_l^2	$\bar{l}$	Cove	σ_l^2	$\bar{l}$	Cove	σ_l^2
30	classical:	13.08	.941	24.78	15.83	.974	32.21	19.01	.986	36.92
	ratio:	8.75	.911	35.97	10.36	.941	53.06	13.49	.958	76.16
	difference:	8.22	.910	21.57	9.81	.942	26.78	13.05	.958	52.27
	regression:	5.79	.791	15.22	6.98	.832	18.61	9.52	.856	29.72
35	classical:	12.12	.928	17.34	12.06	.932	18.94	17.79	.996	26.34
	ratio:	7.68	.932	23.59	9.41	.956	46.44	11.68	.972	45.41
	difference:	7.35	.927	13.62	8.83	.955	19.79	11.36	.972	25.51
	regression:	5.56	.850	10.51	6.62	.889	13.66	8.92	.909	20.38
40	classical:	9.86	.903	13.83	12.04	.957	15.51	16.48	.990	20.83
	ratio:	7.17	.939	17.02	8.17	.962	28.79	11.42	.979	45.53
	difference:	6.99	.948	10.47	7.89	.962	13.39	10.79	.980	19.77
	regression:	5.36	.885	7.87	6.31	.905	10.67	8.72	.945	13.61
45	classical:	9.47	.937	10.27	11.43	.945	13.66	15.71	.994	18.28
	ratio:	6.20	.950	10.35	7.53	.964	15.05	10.21	.979	30.54
	difference:	5.98	.953	7.16	7.23	.965	9.68	9.66	.981	15.37
	regression:	4.90	.913	5.40	5.95	.929	7.40	8.07	.958	10.68
50	classical:	9.51	.934	11.10	10.99	.965	11.47	14.53	.993	15.99
	ratio:	5.93	.953	8.87	6.94	.967	11.84	9.37	.987	22.68
	difference:	5.63	.947	5.52	6.48	.962	7.56	8.78	.986	11.34
	regression:	4.77	.912	4.38	5.41	.928	5.74	7.44	.967	7.24
100	classical:	5.39	.925	2.49	6.28	.958	2.81	8.97	.996	4.15
	ratio:	3.25	.947	1.32	3.98	.972	1.97	5.03	.994	3.00
	difference:	3.10	.937	1.05	3.79	.969	1.42	4.85	.995	2.02
	regression:	2.88	.940	0.88	3.47	.959	1.06	4.52	.990	1.35

Table 2. Household population. $Q_Y(0.5)$.

		100(1 − α)%=90%			100(1 − α)%=95%			100(1 − α)%=99%		
n	method	$\bar{l}$	Cove	σ_l^2	$\bar{l}$	Cove	σ_l^2	$\bar{l}$	Cove	σ_l^2
30	classical:	902.42	.920	70012.73	1120.36	.961	78944.98	1343.77	.989	87891.63
	ratio:	732.42	.845	102011.39	932.74	.924	163754.75	1269.66	.970	250355.84
	difference:	735.94	.872	72255.67	913.56	.937	96244.71	1244.99	.980	135796.34
	regression:	647.72	.864	57635.63	775.95	.923	68952.73	1043.31	.972	90731.77
35	classical:	869.32	.910	61117.19	1037.03	.967	66873.03	1416.70	.996	81143.97
	ratio:	717.95	.899	74739.91	838.36	.924	106418.91	1122.44	.976	169528.73
	difference:	705.56	.904	60664.25	853.60	.936	76192.88	1132.52	.990	101151.03
	regression:	615.78	.890	47201.46	729.89	.933	56199.54	965.02	.983	71993.33
40	classical:	815.62	.907	48657.10	987.70	.963	53706.22	1316.73	.991	67421.45
	ratio:	651.66	.885	58891.99	765.09	.922	77151.24	1044.07	.974	138418.58
	difference:	637.23	.886	43446.93	771.50	.934	57501.28	1052.02	.987	78183.77
	regression:	554.84	.884	34808.63	670.20	.923	43843.52	901.39	.978	53760.44
45	classical:	659.01	.865	33460.33	801.92	.931	42223.61	1100.41	.986	55879.21
	ratio:	619.80	.892	54675.29	735.40	.946	72341.49	977.01	.979	99075.30
	difference:	603.06	.885	39712.07	737.81	.960	48640.51	984.62	.983	60614.36
	regression:	533.12	.886	29573.68	647.30	.948	35393.71	857.71	.983	47882.70
50	classical:	659.93	.905	29497.58	784.47	.942	37390.02	1033.61	.989	44294.93
	ratio:	569.48	.878	38292.25	678.98	.934	50601.72	906.85	.973	84822.34
	difference:	570.99	.892	30036.28	686.16	.947	39557.14	901.79	.979	52860.35
	regression:	498.01	.895	24763.54	600.00	.937	31051.13	797.53	.984	40655.98
100	classical:	442.95	.895	11055.36	573.00	.963	14610.82	761.55	.998	19171.91
	ratio:	386.65	.877	12057.26	470.89	.944	15444.78	617.80	.985	19659.33
	difference:	385.23	.887	10022.00	473.97	.948	13011.61	612.82	.987	15872.20
	regression:	345.74	.878	8792.42	419.40	.945	10947.29	553.51	.987	14143.49