Resampling Inference with Complex Survey Data

Stanislav Kolenikov

skolenik@unc.edu

117 New West, Cameron Ave,

University of North Carolina,

Chapel Hill, NC 27599-3260, US

Outline

- Main concepts of survey sampling statistics (Cochran 1977)
  - Simple random sample
  - Stratified random sample
  - Cluster/multistage random sample
  - Estimates of the mean, total, ratio
- Estimation of variance as the main problem
  - Linearization estimate (Huber 1967)
  - Jackknife estimate (Krewski and Rao 1981)
  - Balanced repeated replication (McCarthy 1969)
  - The bootstrap (Rao and Wu 1988)
- Summary and comparison of performance (Shao 1996)

Simple random sample - I

- Finite population, size N
- Observed characteristic of interest: y, auxiliary x — **nonrandom**
- Might be interested in the estimates of the total $Y = \sum_{i=1}^{N} Y_i$, or in the average $\bar{Y} = Y/N$, or in the ratio $R = \bar{Y}/\bar{X}$.
- Simple random sample: size $n = fN$, f is the sampling fraction
- Estimator of the average: $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$;
  its variance: $\mathbb{V}[\bar{y}] = \frac{S^2}{n}(1-\lambda f)$.
  The last factor is the finite sample correction, $\lambda = 1$ for sampling without replacement, 0 for sampling with replacement
  $S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$
- Unbiased estimator of S^2 is $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$.
- The only randomness is in the sampling process

Simple random sample - II

Estimation of the total:

$$\hat{Y} = \frac{N}{n}\bar{y}, \qquad \mathbb{V}[\hat{Y}] = \left(\frac{N}{n}\right)^2 \mathbb{V}[\bar{y}]$$

Estimation of the ratio:

$$\hat{R} = \bar{y}/\bar{x}$$

Not unbiased, although $\mathbb{B}^2[\hat{R}] = O(\frac{1}{n})$.

$$MSE(\hat{R}) \approx \mathbb{V}[\hat{R}] \approx \frac{1-f}{n\bar{X}^2}\frac{\sum_{i=1}^{N}(y_i - Rx_i)^2}{N-1}$$

Stratified Random Sample - I

The population of N units is first divided into non-overlapping subpopulations, called *strata*, of $N_1, N_2, \dots, N_L$ units, with strata weights $W_1 = N_1/N, \dots$, number of sampled units $n_1, \dots$, sampling fractions $f_1 = n_1/N_1, \dots$, strata averages $\bar{Y}_1, \dots$ and variances $S_1^2, \dots$

Reasons to stratify:

- Efficiency gains if within strata variances are small
- Need data on those subpopulations
- Administrative division
- Markedly different sampling problems (different types of objects)

$\forall\, h\; f_h = f$: *proportional* allocation of n_h's; *self-weighting* sample.

Resampling with complex survey data

Stratified Random Sample - II

Estimate of the mean:

$$\bar{y}_{st} = \sum_{h=1}^{L} W_h \bar{y}_h$$

Unbiased, with variance

$$\mathbb{V}[\bar{y}_{st}] = \sum_{i=1}^{L} W_i^2\, \mathbb{V}[\bar{y}_h] = \sum_{h=1}^{L} W_h^2 \frac{S_h^2}{n_h}(1 - \lambda_h f_h)$$

so the variance of $\bar{y}_{st}$ can be unbiasedly estimated by

$$s^2[\bar{y}_{st}] = \frac{1}{N^2} \sum_{h=1}^{L} N_h (N_h - \lambda_h n_h) \frac{s_h^2}{n_h} = \sum_{h=1}^{L} \frac{W_h^2 s_h^2}{n_h} - \sum_{h=1}^{L} \frac{\lambda_h W_h s_h^2}{N}$$

Stratified Random Sample - III

The variance of the mean is minimized, for a fixed total sample size n, when (*Neyman-Tschuprow allocation*)

$$n_h = n \frac{N_h S_h}{\sum N_h S_h}$$

Stratification provides gains in precision: if terms of the order $1/N_h$ are ignored,

$$\mathbb{V}_{opt}[\bar{y}_{st}] \leq \mathbb{V}_{prop}[\bar{y}_{st}] \leq \mathbb{V}_{SRS}[\bar{y}]$$

S_h's unknown $\Longrightarrow$ optimal variance unattainable; N_h's (W_h's) are wrong $\Longrightarrow \bar{y}_{st}$ is biased.

Resampling with complex survey data

Multistage Random Sample - I

The process of sampling can be repeated:

1. Get a random sample (SRS, PPS, etc.) of the primary sampling units (PSU) at the first level;
2. Get a sample of second stage units (SSU) from each chosen primary unit;
3. . . .
4. Get a sample of individual observations.

Sometimes, unfortunately, also called *subsampling* . . .

Stratified multistage samples: PSUs are stratified; small # of PSUs per strata.

Multistage Random Sample - II

Denote by $\mathbb{E}_1$ expectation over the first stage (over all possible first-stage selections), $\mathbb{E}_2$, the expectation over the second stage selections, etc. Then for an estimate $\hat{\theta}$ of some parameter θ,

$$\mathbb{E}[\hat{\theta}] = \mathbb{E}_1[\mathbb{E}_2[\hat{\theta}]],$$
$$\mathbb{V}[\hat{\theta}] = \mathbb{V}_1[\mathbb{E}_2[\hat{\theta}]] + \mathbb{E}_1[\mathbb{V}_2[\hat{\theta}]],$$
$$\mathbb{V}[\hat{\theta}] = \mathbb{V}_1[\mathbb{E}_2[\mathbb{E}_3[\hat{\theta}]]] + \mathbb{E}_1[\mathbb{V}_2[\mathbb{E}_3[\hat{\theta}]]] + \mathbb{E}_1[\mathbb{E}_2[\mathbb{V}_3[\hat{\theta}]]],$$

etc.

Resampling with complex survey data

Multistage Random Sample - III

In particular, in two stage sampling with equal number of subunits m in each of n units, the mean, its variance, and the estimate of the variance are

$$\bar{\bar{y}} = \sum_{i=1}^{n} \bar{y}_i / n,$$
$$\mathbb{V}[\bar{\bar{y}}] = \frac{N-n}{n}\frac{S_1^2}{n} + \frac{M-m}{M}\frac{S_2^2}{mn},$$
$$s^2[\bar{\bar{y}}] = \frac{1-f_1}{n} s_1^2 + \frac{f_1(1-f_2)}{mn} s_2^2$$

Notation

The survey sampling world lives on its own... Below is some notation for Shao (1996).

The estimate of the distribution function: $\hat{F}(x)$

The parameter of interest: $\theta = g(Z)$ for some vector of the population characteristics / means / totals Z.

Variance estimates: $v_{\text{something}}$ for different methods.

Linearization Estimate - I

General case: $\theta = g(\mathbf{Y})$, $\mathbf{Y} \in \mathbb{R}^{\mathbf{p}}$ is the p-variate population total (or mean $\bar{\mathbf{Y}}$; note that sometimes the number of units in the population may be unknown). Estimate θ by $\hat{\theta} = g(\bar{\mathbf{y}})$.

Linearization estimate of $\mathbb{V}[g(\bar{\mathbf{y}})]$ is available whenever there is an unbiased (consistent?) estimator of $\mathbb{V}[\bar{\mathbf{y}}]$. By the delta method / first order Taylor series expansion,

$$\hat{\theta} - \theta = \sum_{k=1}^{p} \nabla_k g(\bar{\mathbf{y}})(\bar{y}_k - \bar{Y}_k) + o(\|\bar{y}_k - \bar{Y}_k\|) \Longrightarrow$$

$$\mathbb{V}[\hat{\theta}] \approx \nabla g(\bar{\mathbf{y}})' \; \mathbb{V}[\bar{\mathbf{y}}] \; \nabla g(\bar{\mathbf{y}})$$

Derivatives might be tedious (do numerical approximations instead?). The above estimates of variances for the mean and the ratio are the linearization estimates.

Linearization Estimate - II

Plug-in estimate ($\lambda_h = 1$ for w/o replacement sample, 0 for with replacement):

$$\widehat{\mathbb{V}}_L[\hat{\theta}] = \sum_{h=1}^{L} \frac{1-\lambda_h f_h}{n_h} \nabla g(\bar{\mathbf{y}})' \, s_h^2 \, \nabla g(\bar{\mathbf{y}})$$

Also known / can be viewed as Huber (1967) robust variance estimate similar to (information) sandwich estimate.

Idea works for many sorts of second moment assumption violations. In econometrics, heteroskedasticity — White estimator; autocorrelation — Newey-West estimator.

Jackknife Estimate - I

Pretty clear for SRS; what to do with clustered / stratified samples? Except for stratified SRS, need original sampling of PSUs to be with replacement (not what is done in practice) to justify the jackknife.

Jackknife replicate: in the stratum h, omit the whole PSU $\mathbf{y}_{hi}$.

Jackknife estimates:

$$\bar{\mathbf{y}}^{(hi)} = \sum_{h' \neq h} W_{h'} \bar{\mathbf{y}}_{h'} + W_h (n_h \bar{\mathbf{y}}_h - \bar{\mathbf{y}}_{hi})/(n_h - 1)$$

$$\hat{\theta}^{(hi)} = g(\bar{\mathbf{y}}^{(hi)}), \qquad \hat{\theta}^{h} = \sum_{i=1}^{n_h} \hat{\theta}^{(hi)}/n_h$$

Jackknife Estimate - II

Jackknife estimate of variance:

$$\widehat{\mathbb{V}}_J[\hat\theta] = \sum_{h=1}^{L} \frac{(n_h-1)(1-\lambda_h f_h)}{n_h} \sum_{i=1}^{n_h} (\hat\theta^{(hi)} - \hat\theta^{?})^2$$

where $\hat\theta^{?}$ can be any one of $\hat\theta_h$, $\hat\theta$, $\sum_h \sum_i \hat\theta^{(hi)}/n$, or $\sum_h \hat\theta_h/L$.

Might use a random sample of m_h units instead of all n_h in every stratum to reduce the computational effort.

Jackknife Estimate - III

Pseudovalues:

$$\tilde\theta^{(hi)} = n_h\hat\theta - (n_h-1)\hat\theta^{(hi)},$$

$$\tilde\theta_J^{\mathrm{I}} = \sum_{h=1}^{L}\sum_{i=1}^{n_h} \frac{\tilde\theta^{(hi)}}{n}, \qquad \tilde\theta_J^{\mathrm{II}} = \frac{1}{L}\sum_{h=1}^{L} \frac{1}{n_h} \sum_{i=1}^{n_h} \tilde\theta^{(hi)}$$

Couple more variance estimates:

$$\widehat{\mathbb{V}}_J[\hat\theta] = \sum_{h=1}^{L} \frac{(n_h-1)(1-\lambda_h f_h)}{n_h} \sum_{i=1}^{n_h} (\tilde\theta^{(hi)} - \tilde\theta_J^{j})^2, \quad j = \text{I,II}$$

Balanced Repeated Replication - I

Two PSUs per strata ($n_h = 2$) — can use BRR: delete one of two units in each stratum, repeat S times. Can do this efficiently by borrowing from the factorial experiment design literature: a minimal number of BRR resamples (to estimate variance in each stratum) is $L \le S \le L+3$.

$\bar{\mathbf{y}}_{BRR}^{(j)} =$ estimator of $\bar{\mathbf{Y}}$ based on j-th half-sample, $j = 1, \dots, S$;

$\hat{\theta}^{(j)} = g(\bar{\mathbf{y}}_{BRR}^{(j)}),\ \hat{\theta}_c^{(j)}$ is based on the complement of the j-th half-sample,

$$\widehat{\mathbb{V}}_{BRR}^{\mathrm{I}}[\hat{\theta}] = \sum_j \frac{1}{S}(\hat{\theta}^{(j)} - \hat{\theta})^2; \qquad \widehat{\mathbb{V}}_{BRR}^{\mathrm{II}}[\hat{\theta}] = \sum_j \frac{1}{4S}(\hat{\theta}^{(j)} - \hat{\theta}_c^{(j)})^2$$

$$\widehat{\mathbb{V}}_{BRR}^{\mathrm{III}}[\hat{\theta}] = \sum_j \frac{1}{2S}\big((\hat{\theta}^{(j)} - \hat{\theta})^2 + (\hat{\theta}_c^{(j)} - \hat{\theta})^2\big)$$

Asymptotics - I

Typical stratified sample: large number of strata, few PSU in each one.

$$L \to \infty, \quad \max_{1 \le h \le L} n_h = O(1)\ \text{(C2)}\ , \quad \max_{1 \le h \le L} W_h = O(L^{-1})\ \text{(C3)}$$

Also need: finite non-singular limit Γ of $n\,\mathbb{V}[\bar{\mathbf{y}}]$ (C4), something for CLT (e.g. Lyapunov/moment condition (C1)). Those are the conditions (C1)–(C4) of Krewski and Rao (1981).

Under (C1)–(C4), $n^{1/2}(\bar{\mathbf{y}} - \bar{\mathbf{Y}}) \xrightarrow{d} N(\mathbf{0}, \Gamma)$; $n\big(\widehat{\mathbb{V}}[\bar{\mathbf{y}}] - \mathbb{V}[\bar{\mathbf{y}}]\big) \xrightarrow{p} \mathbf{0}$.

To go further, also need: $\bar{\mathbf{Y}} \to \mu$ (C5), $\nabla g(\cdot) \in C_{U(\mu)}$ (C6).

Asymptotics - II

Under (C1)–(C6),

$$n^{1/2}(\hat{\theta}-\theta) \xrightarrow{d} N(0,\sigma^2), \quad \sigma^2 = \nabla g(\mu)'\,\Gamma\,\nabla g(\mu)$$

$$n\widehat{\mathbb{V}}_{whatever}[\theta] \xrightarrow{p} \sigma^2,$$

$$T_{whatever} = \frac{\hat{\theta}-\theta}{\left(\widehat{\mathbb{V}}_{whatever}[\hat{\theta}]\right)^{1/2}} \xrightarrow{d} N(0,1)$$

whatever is the estimator we are using (linearization, jackknife, BRR).

Comparison of the Methods - I

Simulation findings reported in Krewski and Rao (1981):

2-sided CI : BRR $\succ$ jackknife $\succ$ linearization

Stability : linearization $\succ$ jackknife $\succ$ BRR

$$\text{Stability index}\left[\widehat{\mathbb{V}}[\hat{\theta}]\right] = \frac{\left(MSE\left[\widehat{\mathbb{V}}[\hat{\theta}]\right]\right)^{\frac{1}{2}}}{MSE[\hat{\theta}]}$$

The smaller the index, the better is the estimator.

The Naive Bootstrap - I

Resamples: $\{y^*_{hi}\}_{i=1}^{n_h}$ i.i.d. with replacement in stratum h from $\hat{F}^h_{n_h}$, independently across strata.

Estimates: $\bar{y}^*_h = n_h^{-1} \sum_i y^*_{hi}$, $\bar{y}^* = \sum_i W_i \bar{y}^*_h$, $\hat{\theta}^* = g(\bar{y}^*)$.

Variance: $\widehat{\mathbb{V}}^*_{NBS}[\hat{\theta}^*] = \mathbb{E}_*\left[(\hat{\theta}^* - \mathbb{E}_*[\hat{\theta}^*])^2\right]$, or its simulation approximation

Linear case: $\widehat{\mathbb{V}}^*_{NBS}[\hat{\theta}^*] = \sum_h \frac{W_h^2}{n_h} \frac{n_h - 1}{n_h} s_h^2$ — biased down, inconsistent! (Think of $n_h = 2$, as it is often the case.)

Resampling with complex survey data

The Rescaled Bootstrap - I

Rao and Wu (1988) propose resamples: $\{y^*_{hi}\}_{i=1}^{m_h}$ i.i.d. sample of size $m_h \geq 1$ with replacement in stratum h from $\hat{F}^h_{n_h}$, independently across strata.

Pseudovalues (?? not called this way in the article):
$\tilde{y}_{hi} = \bar{y}_h + \frac{m_h^{1/2}}{(n_h - 1)^{1/2}} (y^*_{hi} - \bar{y}_h)$

Notable features of the proposed method:

- Captures the dependence structure: the i.i.d. pieces are those within strata
- The bootstrap subsample size $m_h <$ the original size n_h
- Modification of the subsample value

The Rescaled Bootstrap - II

Estimates: $\bar{y}^*_h = m_h^{-1} \sum_i \tilde{y}_{hi}$, $\tilde{y}^* = \sum_i W_i \tilde{y}_h$, $\tilde{\theta} = g(\tilde{y})$.

Variance: $\widehat{\mathbb{V}}^*_{RBS}[\tilde{\theta}] = \mathbb{E}_* \big[(\tilde{\theta} - \mathbb{E}_*[\tilde{\theta}])^2 \big]$, or its simulation approximation.

Linear case: $\widehat{\mathbb{V}}^*_{RBS}[\hat{\theta}^*] = \sum_h \frac{W_h^2}{n_h} s_h^2$ — as desired.

The scale is chosen so as to fit the "right" thing, just as we did for the jackknife estimator of variance of the i.i.d. sample mean.

The Rescaled Bootstrap - III

Can we estimate something beyond the variance? What about bias?

$$\mathbb{B}[\hat{\theta}] = \frac{1}{2} \sum_{j,k=1}^{p} \nabla^2_{jk} g(\bar{Y}) \sum_h \frac{W_h^2}{n_h} S_{hjk} + \text{higher order terms}, \qquad (1)$$

$$S_{hjk} = \mathbb{E}[(y_{hij} - \bar{Y}_{hj})(y_{hik} - \bar{Y}_{hk})] \qquad (2)$$

$$\mathbb{B}^*_{RBS}[\hat{\theta}] = \mathbb{E}^*_{RBS}[\tilde{\theta}] - \hat{\theta} = \qquad (3)$$

$$= \frac{1}{2} \sum_{j,k=1}^{p} \nabla^2_{jk} g(\bar{y}) \sum_h \frac{W_h^2}{n_h} \cdot \text{naivety factor} \cdot s_{hjk} + O_p(n^{-3/2}) \qquad (4)$$

where the naivety factor is 1 for the rescaled bootstrap, and $(n_h - 1)/n_h$ for the naive bootstrap. The rescaled bootstrap gives a consistent estimate of bias, while the naive one, does not.

The Rescaled Bootstrap - IV

$m_h = n_h - 1 \Longrightarrow \tilde{y}_{hi} = y^*_{hi}$ — the original bootstrap subsample with the "correct" subsample size.

$n_h = 2$, $m_h = 1 \Longrightarrow$ the (exact) rescaled bootstrap reduces to (complete) BRR, although in simulations may not be as efficient if BRR resamples are chosen wisely.

$n_h \geq 3 \Longrightarrow$ there is an option of choosing m_h so as to match the bootstrap third moment of $\bar{\mathbf{y}}$ to its empirical estimate:

$$m_h = \frac{(n_h - 2)^2}{n_h - 1} \approx n_h - 3$$

Rao and Wu (1988): "The bootstrap histogram of a t statistic captures the second order term of the Edgeworth expansion in the special case of known strata variances" — c.f. the i.i.d. bootstrap histogram properties.

The Rescaled Bootstrap - V

Simulations results: p. 235 of Rao and Wu (1988).

- Not much gain in going beyond $B = 100$
- Substantial biases in one-sided CI coverage for nonlinear statistics (ratio, correlation)
- One sided intervals: the bootstrap with $m_h = n_h - 1$ is clearly superior in terms of coverage
- Two-sided CIs OK for either linearization, jackknife, or the bootstrap
- The choice of $m_h = n_h - 3$ for bootstrapping understates the errors in both tails when the variance of $\bar{y}$ is not known.
- Stability is better for the linearization and jackknife than for the bootstrap

The Rescaled Bootstrap - VI

Rao and Wu (1988) give extensions to:

- sampling without replacement;
- unequal probability sampling;
- two-stage cluster sampling with equal probabilities and without replacement.

Other bootstrap schemes

Shao (1996) gives several other bootstrap schemes used in literature.

With replacement bootstrap, BWR: $m_h = \frac{n_h - 1}{1 - \lambda_h f_h}$, no other modifications / rescaling.

Mirror-match bootstrap, BMM: draw an SRS (w/o replacement) of size $n_h^* < n_h$, repeat k_h times, so $m_h = n_h^* k_h$. BMM keeps higher variability by ruling out the subsamples replicating a single observation; reduces to BWR when $n_h^* = 1$; when the statistic is linear, $\widehat{\mathbb{V}}_{BMM}[\hat{\theta}] = \widehat{\mathbb{V}}_L[\hat{\theta}]$.

Without replacement bootstrap, BWO: mimics the original sampling scheme. Create a pseudopopulation of the size N_h by replication of the strata samples, then take samples of size n_h from this pseudopopulation. May not provide a consistent estimate of the variance even in the linear case.

Shao (1996): " . . . not necessarily all of [the bootstrap procedures] are second-order accurate."

Asymptotics - III

Asymptotics in Shao (1996) is a bit different from that of (Krewski and Rao 1981).

- Array of populations indexed by some k; θ might differ for different k, but $\{\theta_k\}$ is a bounded set
- $n = \sum_h n_h \to \infty$; $\sup f_h < 1$
- $\forall\, k\ \exists\, \mathcal{H}_k \subset \{1, \dots, L\} : \sup_{h \in \mathcal{H}_k, k} n_h < \infty$, $\min_{h \notin \mathcal{H}_k, k} n_h \to \infty$
- Bounded survey weights: $\max_{h,i,j} n_{hi} w_{hij} n/N = O(1)$
- Lyapunov condition on $\bar{Y}$
- $\nabla g \neq 0$ in a neighborhood of $\bar{Y}$.

Comparison of the Methods - II

Baseline: linearization method

Jackknife: different jackknife estimators are $O_p(n^{-2})$ equivalent;
$\widehat{\mathbb{V}}_J[\hat\theta]/\widehat{\mathbb{V}}_L[\hat\theta] = 1 + O_p(n^{-1})$ in general case;
$\widehat{\mathbb{V}}_J[\hat\theta]/\widehat{\mathbb{V}}_L[\hat\theta] = 1 + O_p(n^{-2})$ when $\forall h\ n_h = 2$;
$\widehat{\mathbb{V}}_J[\hat\theta] = \widehat{\mathbb{V}}_L[\hat\theta]$ for linear statistics or ($n_h = 2$ & $g(\cdot)$ is quadratic).

BRR: $\widehat{\mathbb{V}}^I_{BRR}[\hat\theta]/\widehat{\mathbb{V}}_L[\hat\theta] = 1 + O_p(n^{-1/2})$;
$\widehat{\mathbb{V}}^{II}_{BRR}[\hat\theta]/\widehat{\mathbb{V}}_L[\hat\theta] = 1 + O_p(n^{-1})$.
Unlike other methods, can be applied to estimation of non-smooth functionals (quantiles).

Rescale Bootstrap: $\widehat{\mathbb{V}}_{RBS}[\hat\theta]/\widehat{\mathbb{V}}_L[\hat\theta] = 1 + O_p(n^{-1})$.
Empirical evidence: variance estimate is not quite stable.
The bootstrap histogram $\xrightarrow{p}$ target distribution in sup norm.

Some minor regularity conditions might be omitted.

Comparison of the Methods - III

Shao (1996): “ . . . the choice of the method may depend more on nonstatistical considerations, such as the feasibility of their implementation . . . Blindly applying the resampling methods may yield incorrect results” — that’s often the case!

Other topics - I

Some other topics covered in Shao (1996):

- Confidence intervals for quantiles
- Jackknife estimates of bias
- Resampling with missing / imputed data
- BRR-type methods when the # PSU per cluster varies (not all of them behave nice)

Other topics - II

Discussion part of Shao (1996) provides a number of important comments.

- Jackknife for non-normal data (Binder 1996)
- Design efficiency (Binder 1996)
- Estimating functions (Binder 1996)
- Jackknife with non-smooth functions (Rao and Sitter 1996)
- Post-stratification weights (Rao and Sitter 1996)
- Linearized jackknife estimator (Rao and Sitter 1996)

Other topics - III

- Strata collapsing (Valliant 1996)
- In small / moderate samples, $\mathbb{V}_J[\hat{T}_R] \succ \mathbb{V}_L[\hat{T}_R]$ where $\hat{T}_R = N\bar{y}\bar{X}/\bar{x}$ is the ratio estimator of the total (Valliant 1996)
- Post-stratification: $\widehat{\mathbb{V}}_J$, $\widehat{\mathbb{V}}_J$ might be $\succ \widehat{\mathbb{V}}_L$ (Valliant 1996)
- Parallel computing (Valliant 1996)
- Variance stabilizing transformations help achieving better performance (Valliant 1996)
- Performance of certain bootstrap schemes may deteriorate as complexity of the sampling design increases (Valliant 1996).

References

Cochran, W. G. (1977). *Sampling Techniques*. 3rd edition, Wiley.

Huber, P. J. (1967). The behavior of the maximum likelihood estimates under non-standard conditions. In *Proc. of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, CA, Univ. of California Press, v. 1, pp. 221–233.

Krewski, D., Rao, J. N. K. (1981). Inference From Stratified Samples: Properties of the Linearization, Jackknife and Balanced Repeated Replication Methods. *Annals of Statistics*, **9** (5), pp. 1010–1019.

McCarthy, P. J. (1969). Pseudoreplication: half-samples. *Rev. of Internat. Statist. Inst.*, **37**, pp. 239–264.

Rao, J. N. K., Wu, C. F. J. (1988). Resampling Inference With Complex Survey Data. *JASA*, **83**, pp. 231–241.

Shao, J. (1996). Resampling Methods in Sample Surveys. *Invited paper, Statistics*, **27**, pp. 203–237, with discussion, 237–254.

Binder, D. A. (1996). Discussion of Shao's paper. *Statistics*, **27**, pp. 237–241.

Rao, J. N. K., and Sitter, R. R. (1996). Discussion of Shao's paper. *Statistics*, **27**, pp. 246–247.

Valliant, R. (1996). Discussion of Shao's paper. *Statistics*, **27**, pp. 247–251.