ESTIMATION METHOD 9: Cumulative Distribution Function and Variance for Proportion of a Finite Population; Parametric Jackknife Estimator

1 Scope and Application

An important aspect of environmental statistics is to measure specific indicators in order to monitor the status of the environment. Frequently these indicators are subject to measurement error. When sample units are measured with error, the naive estimator of the population cumulative distribution obtained when the measurement error is ignored are biased and may be misleading. The purpose of this report is to describe a bias-adjusted estimator proposed by Stefanski and Bay (1996) for the cumulative distribution of a finite population in the presence of measurement error. This estimator, called the parametric jackknife, reduces much of the bias induced by the measurement error. A variance estimator for the parametric jackknife estimator is obtained using Horvitz-Thompson estimation.

2 Statistical Estimation Overview

A sampling model is assumed in which a sample of size n is selected from a population $\mathbf{U} = \{U_1, U_2, \cdots, U_N\}$ with inclusion probabilities $\{\pi_i\}_{i=1}^{N}$ and joint inclusion probabilities $\{\pi_{ij}\}_{1 \le j < i \le N}$. The observed data consists of $\{X_i\}_{i=1}^{n}$, where X_i is a measured value of U_i and is subject to measurement error. It is assumed that $X_i = U_i + \sigma Z_i$ for $i = 1, \cdots, n$, where $\{Z_i\}_{i=1}^{n}$ are mutually independent, independent of random sampling, and identically distributed standard normal random variables. Thus, the measurement errors in the observed sample are normally distributed with mean zero and variance σ^2.

The estimation procedure involves adding additional measurement error in known increments to the observed data, computing cumulative distribution estimates from these contaminated data, establishing a relationship between these estimates and the measurement error variance, and extrapolating this relationship back to the case of no measurement error.

Computing $X_i^*(\lambda) = X_i + \sigma\sqrt{\lambda}\, Z_i^*$, for $i = 1, \cdots, n$, where Z_i^* is a standard normal pseudo-random variable and $\lambda > 0$ is a constant, increases the variability of the measurement error. The total measurement error variance of the resulting data is $\sigma^2(1 + \lambda)$. Cumulative distribution estimates are calculated at a fixed argument from $\{X_i^*(\lambda)\}_{i=1}^{n}$ over a range of values of λ. The expectation of these estimates is approximated by a quadratic function in λ. Least squares regression of the cumulative distribution estimates on λ estimates the parameters of this quadratic model. Extrapolation to the case of no measurement error, i.e., $\lambda = -1$, gives the parametric jackknife estimator.

Refer to Section 8.3. for a more detailed explanation of this estimation procedure and for details on calculating a variance estimate of a parametric jackknife estimate.

3 Conditions Under Which This Method Applies

- Probability sample with known inclusion and joint inclusion probabilities.
- Finite population.
- Data observed with error.
- Additive independent and normally distributed measurement errors with mean zero and common variance σ^2.

4 Required Elements

4.1 Input Data

N = population size.
$\{X_1, \cdots, X_n\}$ = probability sample of size n, where X_i is a measured value of U_i, the i^{th} element in population **U**.
$\{\pi_1, \cdots, \pi_n\}$ = vector of inclusion probabilities, where π_i is the probability of selecting element U_i from population **U**.
$[\pi_{ij}]$ = matrix of joint inclusion probabilities, where π_{ij} for $i, j = 1, \cdots, n$ is the probability of selecting elements U_i and U_j from population **U**; $\pi_{ii} = \pi_i$, and $\pi_{ij} = \pi_{ji}$ for $i, j = 1, \cdots, n$.
σ^2 = measurement error variance.
$V\hat{a}r(\hat{\sigma}^2)$ = variance estimate of $\hat{\sigma}^2$, when σ^2 is estimated; zero otherwise.

5 Formulas and Definitions

Let t denote a fixed argument in the following definitions and formulas. Define:

- $F_{U,N}(t)$ = estimand (i.e., the population cumulative distribution (CD)).

- $\hat{F}_{U,n}(t)$ = estimator of the population CD in the absence of measurement error.

- $\hat{F}_{X,\lambda,n}(t)$ = the CD estimator based on the data $\{X_i + \sigma\sqrt{\lambda}Z_i^*\}_{1=1}^{n}$, where $\{X_i\}_{i=1}^{n}$ is the observed sample, $\{Z_i^*\}_{i=1}^{n}$ are standard normal pseudo-random variables, and $\lambda > 0$ is a constant.

- $\hat{F}_{X,\lambda,JK}(t)$ = parametric jackknife estimator.

- $V\hat{a}r\{\hat{F}_{X,\lambda,JK}(t)\}$ = variance estimator of the parametric jackknife estimator.

- $\hat{L}(t)$ = lower $100(1-\alpha)\%$ confidence limit for $\hat{F}_{X,\lambda,JK}(t)$.

- $\hat{U}(t)$ = upper $100(1-\alpha)\%$ confidence limit for $\hat{F}_{X,\lambda,JK}(t)$.

The formulas for the above definitions are as follows:

- $$F_{U,N}(t) = \frac{1}{N}\sum_{i=1}^{N} I(U_i \le t)$$

- $$\hat{F}_{U,n}(t) = \frac{1}{N}\sum_{i=1}^{n} \frac{1}{\pi_i} I(U_i \le t)$$

If σ^2 is known,

- $$\hat{F}_{X,\lambda,n}(t) = \frac{1}{N}\sum_{i=1}^{n} \frac{1}{\pi_i}\,\Phi\left(\frac{X_i - t}{\sqrt{\sigma^2 \lambda}}\right)$$

If σ^2 is estimated by $\hat{\sigma}^2$,

- $$\hat{F}_{X,\lambda,n}(t) = \frac{1}{N}\sum_{i=1}^{n} \frac{1}{\pi_i}\,\Phi\left(\frac{X_i - t}{\sqrt{\hat{\sigma}^2 \lambda}}\right)$$

If σ^2 is known,

- $$\hat{F}_{X,n,JK}(t) = \frac{1}{N}\sum_{i=1}^{n} \frac{1}{\pi_i}\,G(t; X_i, \sigma^2)$$

If σ^2 is estimated by $\hat{\sigma}^2$,

- $$\hat{F}_{X,n,JK}(t) = \frac{1}{N}\sum_{i=1}^{n} \frac{1}{\pi_i}\,G(t; X_i, \hat{\sigma}^2)$$

If σ^2 is known,

- $$\hat{Var}\{\hat{F}_{X,n,JK}(t)\} = \frac{1}{N^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(1 - \frac{\pi_i \pi_j}{\pi_{ij}}\right)\frac{G(t; X_i, \sigma^2)}{\pi_i}\cdot\frac{G(t; X_j, \sigma^2)}{\pi_j}$$

If σ^2 is estimated by $\hat{\sigma}^2$,

- $$\hat{Var}\{\hat{F}_{X,n,JK}(t)\} = \frac{1}{N^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(1 - \frac{\pi_i \pi_j}{\pi_{ij}}\right)\frac{G(t; X_i, \hat{\sigma}^2)}{\pi_i}\cdot\frac{G(t; X_j, \hat{\sigma}^2)}{\pi_j} + \frac{1}{N^2}\hat{Var}(\hat{\sigma}^2)\left\{\sum_{1=1}^{n}\frac{1}{\pi_i}\,g(t; X_i, \hat{\sigma}^2)\right\}^2$$

- $$\hat{L}(t) = \hat{F}_{X,n,JK}(t) - z_{1-\alpha/2}\cdot\hat{Var}\{\hat{F}_{X,n,JK}(t)\}$$

- $$\hat{U}(t) = \hat{F}_{X,n,JK}(t) + z_{1-\alpha/2}\cdot\hat{Var}\{\hat{F}_{X,n,JK}(t)\}$$

where

$I(\cdot)$ = indicator function,

N = population size,

$\mathbf{U}$ = population of interest,

n = number of elements sampled from U,

$\mathbf{X} = \{X_i\}_{i=1}^{n}$ = sample observed with measurement error,

π_i = probability of selecting U_i from $\mathbf{U}$, where U_i is the true (unobserved) value of X_i,

π_{ij} = probability of selecting U_i and U_j from $\mathbf{U}$, where U_k is the true value of X_k, k = i, j,

σ^2 = measurement error variance,

$\hat{Var}(\hat{\sigma}^2)$ = variance estimate of $\hat{\sigma}^2$, where $\hat{\sigma}^2$ is the estimate of σ^2,

$z_{1-\alpha/2}$ = $100(1-\alpha/2)th$ percentile in the standard normal distribution,

Φ = standard normal cumulative distribution function,

ϕ = standard normal density function,

$G(t; X_i, \tau) = (1, -1, 1)(D^T D)^{-1} D^T Y$ (i.e., least squares solution)

$$\text{with } D = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 \\ \vdots & \vdots & \vdots \\ 1 & \lambda_m & \lambda_m^2 \end{bmatrix} \text{ and } Y = \left[\Phi\left(\frac{X_i - t}{\sqrt{\tau\lambda_1}}\right), \Phi\left(\frac{X_i - t}{\sqrt{\tau\lambda_2}}\right), \ldots, \Phi\left(\frac{X_i - t}{\sqrt{\tau\lambda_m}}\right)\right]^T,$$

$g(t; X_i, \tau) = (1, -1, 1)(D^T D)^{-1} D^T y$

with D as defined above and

$$y = \left[\frac{X_i - t}{\sqrt{\tau^3\lambda_1}}\phi\left(\frac{X_i - t}{\sqrt{\tau\lambda_1}}\right), \frac{X_i - t}{\sqrt{\tau^3\lambda_2}}\phi\left(\frac{X_i - t}{\sqrt{\tau\lambda_2}}\right), \ldots, \frac{X_i - t}{\sqrt{\tau^3\lambda_m}}\phi\left(\frac{X_i - t}{\sqrt{\tau\lambda_m}}\right)\right]^T$$

6 Procedure

6.1 Generate a sequence of k grid points, $t_1 < t_2 < \cdots < t_k$, spanning the range of the observed data $\{X_i\}_{i=1}^{n}$.

For example, suppose $\min\{X_1, X_2, \cdots, X_n\} = 0$ and $\max\{X_1, X_2, \cdots, X_n\} = 25$. We could let $k = 51$ and define $t_1 = 0$, $t_2 = 0.5$, $t_3 = 1.0$, $t_4 = 1.5$, $\cdots$, $t_{50} = 24.5$, $t_{51} = 25.0$.

6.2 Generate a sequence of values $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_m$.

See Section 8.1 for more information.

6.3 For each grid point t_h, $h = 1, \cdots, k$,

6.3.1 For each data value X_i, $i = 1, \cdots, n$,

6.3.1.1 Calculate $G(t_h; X_i, \sigma^2)$ or $G(t_h; X_i, \hat{\sigma}^2)$, when σ^2 is estimated).

$$G(t_h; X_i, \sigma^2) = (1, -1, 1)(D^T D)^{-1} D^T Y$$

where

$$D = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 \\ \vdots & \vdots & \vdots \\ 1 & \lambda_m & \lambda_m^2 \end{bmatrix}$$

$$Y = \left[\Phi\left(\frac{X_i - t}{\sqrt{\tau \lambda_1}}\right), \Phi\left(\frac{X_i - t}{\sqrt{\tau \lambda_2}}\right), \cdots, \Phi\left(\frac{X_i - t}{\sqrt{\tau \lambda_m}}\right) \right]^T,$$

(Note, Φ is the standard normal cumulative distribution function).

6.3.1.2 If σ^2 is estimated, calculate $g(t_h; X_i, \hat{\sigma}^2)$.

$$g(t_h; X_i, \hat{\sigma}^2) = (1, -1, 1)(D^T D)^{-1} D^T y$$

where D is defined in Section 6.3.1.1 above, and

$$y = \left[\frac{X_i - t}{\sqrt{\hat{\sigma}^6 \lambda_1}} \phi\left(\frac{X_i - t}{\sqrt{\hat{\sigma}^2 \lambda_1}}\right), \frac{X_i - t}{\sqrt{\hat{\sigma}^6 \lambda_2}} \phi\left(\frac{X_i - t}{\sqrt{\hat{\sigma}^2 \lambda_2}}\right), \cdots, \frac{X_i - t}{\sqrt{\hat{\sigma}^6 \lambda_m}} \phi\left(\frac{X_i - t}{\sqrt{\hat{\sigma}^2 \lambda_m}}\right) \right]^T$$

(Note, ϕ is the standard normal density function).

6.3.2. Calculate $\hat{F}_{X,n,JK}(t_h)$.

If σ^2 is known, then

$$\hat{F}_{X,n,JK}(t_h) = \frac{1}{N}\sum_{i=1}^{n}\frac{1}{\pi_i}G(t_h; X_i, \sigma^2)$$

If σ^2 is estimated by $\hat{\sigma}^2$,

$$\hat{F}_{X,n,JK}(t) = \frac{1}{N}\sum_{i=1}^{n}\frac{1}{\pi_i}G(t; X_i, \hat{\sigma}^2)$$

6.3.3 Calculate $\hat{\text{Var}}\{\hat{F}_{X,n,JK}(t_h)\}$.

If σ^2 is known,

$$\hat{\text{Var}}\{\hat{F}_{X,n,JK}(t_h)\} = \frac{1}{N^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(1 - \frac{\pi_i\pi_j}{\pi_{ij}}\right)\frac{G(t_h; X_i, \sigma^2)}{\pi_i}\cdot\frac{G(t_h; X_j, \sigma^2)}{\pi_j}$$

If σ^2 is estimated by $\hat{\sigma}^2$,

$$\hat{\text{Var}}\{\hat{F}_{X,n,JK}(t_h)\} = \frac{1}{N^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(1 - \frac{\pi_i\pi_j}{\pi_{ij}}\right)\frac{G(t_h; X_i, \hat{\sigma}^2)}{\pi_i}\cdot\frac{G(t_h; X_j, \hat{\sigma}^2)}{\pi_j} + \frac{1}{N^2}\hat{\text{Var}}(\hat{\sigma}^2)\left\{\sum_{1=1}^{n}\frac{1}{\pi_i}g(t_h; X_i, \hat{\sigma}^2)\right\}^2$$

6.3.4 Calculate approximate $100(1-\alpha)\%$ confidence limits, $\hat{L}(t_h)$ and $\hat{U}(t_h)$.

$$\hat{L}(t_h) = \hat{F}_{X,n,JK}(t_h) - z_{1-\alpha/2}\cdot\hat{\text{Var}}\{\hat{F}_{X,n,JK}(t_h)\}$$

$$\hat{U}(t_h) = \hat{F}_{X,n,JK}(t_h) + z_{1-\alpha/2}\cdot\hat{\text{Var}}\{\hat{F}_{X,n,JK}(t_h)\}$$

where $z_{1-\alpha/2}$ is the $100(1-\alpha/2)$*th* percentile of the standard normal distribution.

6.4 Apply isotonic regression to $\{\hat{F}_{X,n,JK}(t_1), \cdots, \hat{F}_{X,n,JK}(t_k)\}$ on $\{t_1, \cdots, t_k\}$.

While$_1$ $\left\{\hat{F}_{X,n,JK}(t_h)\right\}_{h=1}^{k}$ is NOT a non-decreasing sequence

Let $i = 1$ and $j = 2$

While$_2$ $j < n$

While$_3$ $\hat{F}_{X,n,JK}(t_{j-1}) > \hat{F}_{X,n,JK}(t_j)$

Increase j by 1 (i.e., $j = j + 1$)

End of While$_3$

For $h = 1, \cdots, j - 1$,

$$\hat{F}_{X,n,JK}(t_h) = \frac{1}{j-i} \sum_{q=i}^{j-1} \hat{F}_{X,n,JK}(t_q)$$

End of For

Let $i = j$ and $j = j+1$

End of While$_2$

End of While$_1$

(Note, isotonic regression simply ensures that the function $\hat{F}_{X,n,JK}(t)$ is non-decreasing on $[t_1, t_k]$. This is accomplished by averaging the function values on intervals where the function is decreasing. The procedure is repeated until $\hat{F}_{X,n,JK}(t)$ is non-decreasing. For more information on isotonic regression see Barlow et al. (1972).)

6.5 Restrict range of $\left\{\hat{F}_{X,n,JK}(t_h)\right\}_{h=1}^{k}$ to [0, 1].

Set $h = 1$

While $\hat{F}_{X,n,JK}(t_h) < 0$

$\hat{F}_{X,n,JK}(t_h) = 0$

$h = h + 1$

End of While

Set $h = k$

While $\hat{F}_{X,n,JK}(t_h) > 1$

$\hat{F}_{X,n,JK}(t_h) = 1$

$h = h - 1$

End of While

6.6 Apply isotonic regression to $\{\hat{L}(t_1), \cdots, \hat{L}(t_k)\}$ on $\{t_1, \cdots, t_k\}$ and restrict range of $\{\hat{L}(t_h)\}_{h=1}^{k}$ to [0, 1].

See Sections 6.4 and 6.5 above.

6.7 Apply isotonic regression to $\{\hat{U}(t_1), \cdots, \hat{U}(t_k)\}$ on $\{t_1, \cdots, t_k\}$ and restrict range of $\{\hat{U}(t_h)\}_{h=1}^{k}$ to [0, 1].

See Sections 6.4 and 6.5 above.

7 Associated Methods

A related method for estimating a cumulative distribution in the presence of measurement error is described in Estimation Method 11, the simulation-extrapolation method. This method does not assume a particular sampling model nor does it require a finite population.

8 Notes

The algorithm given in Section 6 requires specification of $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_m$. Stefanski and Bay (1996) propose taking equally spaced values over the interval [0.05, 2.00]. They also suggests $m \geq 5$, although the exact number of values is not critical.

The algorithm in Section 6 calculates estimates of the cumulative proportion. Estimates of the cumulative total may be obtained by multiplying the estimates $\{\hat{F}_{X,n,JK}(t_h)\}_{h=1}^{k}$ by N, the population size. The variance estimator for the cumulative total is equal to the variance estimator for the cumulative proportion times N^2. Confidence limits would need to be recalculated. Additionally, the range of the estimates of the cumulative total and its confidence limits would be [0, N] rather than [0, 1] as specified for the cumulative proportion.

This method of bias-adjustment is closely related to Quenouille's jackknife. The usual jackknife increases sampling variance by decreasing sample size. In this method measurement error variance is increased by adding pseudo-random errors to the observed data, achieving the same “variance-inflation" effect as in the jackknife method. This is done by calculating $X_i^*(\lambda) = X_i + \sigma\sqrt{\lambda}Z_i^*$, for $i = 1, \cdots, n$, where $\{X_i\}_{i=1}^{n}$ is the observed sample with measurement error variance σ^2, $\{Z_i^*\}_{i=1}^{n}$ are standard normal pseudo-random variables, and $\lambda > 0$ is a constant. The variance of the additional error is $\sigma^2\lambda$, and the variance of the total measurement error in $X_i^*(\lambda)$ is $\sigma^2(\lambda + 1)$. The usual CD estimator based on these data $\{X_i^*(\lambda)\}_{i=1}^{n}$ is given by

$$\frac{1}{N}\sum_{i=1}^{n}\frac{1}{\pi_i} I(X_i^*(\lambda) \le t).$$

Taking the expectation with respect to the pseudo-random error distribution, we obtain

$$\hat{F}_{X,\lambda,n}(t) = \frac{1}{N}\sum_{i=1}^{n}\frac{1}{\pi_i}\Phi\left(\frac{X_i - t}{\sigma\sqrt{\lambda}}\right)$$

It can be shown that

$$E\{\hat{F}_{X,\lambda,n}(t)\} = E_Z\{E_\pi\{\hat{F}_{X,\lambda,n}(t)\}\} = \frac{1}{N}\sum_{i=1}^{n}\frac{1}{\pi_i}\Phi\left(\frac{U_i - t}{\sigma\sqrt{\lambda+1}}\right)$$

where $E_Z(\cdot)$ and $E_\pi(\cdot)$ denote the expectation with respect to the error and sampling distributions, and

$$\lim_{\lambda \to -1}\frac{1}{N}\sum_{i=1}^{N}\Phi\left(\frac{U_i - t}{\sigma\sqrt{\lambda+1}}\right) \approx \frac{1}{N}\sum_{i=1}^{N} I(U_i \le t) = F_{U,N}(t)$$

The model $\beta_0 + \beta_1\lambda + \beta_2\lambda^2$ is used to approximate $E\{\hat{F}_{X,\lambda,n}(t)\}$. Parameters are estimated by least squares regression of $\{\hat{F}_{X,\lambda_j,n}(t)\}_{j=1}^{m}$ on $\{\lambda_j\}_{j=1}^{m}$, where $0 < \lambda_1 < \cdots < \lambda_m$ are fixed constants. Extrapolating to $\lambda = -1$, we obtain

$$\hat{F}_{X,n,JK}(t) = \hat{\beta}_0 - \hat{\beta}_1 + \hat{\beta}_2$$

which may also be expressed as

$$\hat{F}_{X,n,JK}(t) = \frac{1}{N}\sum_{i=1}^{n}\frac{1}{\pi_i} G(t; X_i, \sigma^2)$$

where

π_i = the inclusion probability for selecting the i^{th} element in population **U**,

$g(t; X_i, \sigma^2) = (1, -1, 1)(D^T D)^{-1} D^T Y$

such that

$$D = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 \\ \vdots & \vdots & \vdots \\ 1 & \lambda_m & \lambda_m^2 \end{bmatrix}$$

$$Y = \left[\Phi\left(\frac{X_i - t}{\sqrt{\tau \lambda_1}} \right), \Phi\left(\frac{X_i - t}{\sqrt{\tau \lambda_2}} \right), \cdots, \Phi\left(\frac{X_i - t}{\sqrt{\tau \lambda_m}} \right) \right]^T .$$

When σ^2 is known, the variance of $\hat{F}_{X,n,JK}(t)$ is estimated by the Horvitz-Thompson estimator (Sarndal et al., 1992, p. 43):

$$\hat{Var}\{ \hat{F}_{X,n,JK}(t) \} = \frac{1}{N^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(1 - \frac{\pi_i \pi_j}{\pi_{ij}} \right) \frac{G(t; X_i, \sigma^2)}{\pi_i} \cdot \frac{G(t; X_j, \sigma^2)}{\pi_j}$$

where

π_i and G are given above,

π_{ij} = joint inclusion probability for selecting elements *i* and *j* from population **U**.

When σ^2 is estimated, the Horvitz-Thompson estimator is still used to estimate the variance of the parametric jackknife estimate. However, the additional variation due to estimating   must also be accounted for. Hence, when σ^2 is estimated, the variance estimator is given by:

$$\hat{Var}\{ \hat{F}_{X,n,JK}(t) \} = \frac{1}{N^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(1 - \frac{\pi_i \pi_j}{\pi_{ij}} \right) \frac{G(t; X_i, \hat{\sigma}^2)}{\pi_i} \cdot \frac{G(t; X_j, \hat{\sigma}^2)}{\pi_j} + \frac{1}{N^2} \hat{Var}(\hat{\sigma}^2) \left\{ \sum_{1=1}^{n} \frac{1}{\pi_i} g(t; X_i, \hat{\sigma}^2) \right\}^2$$

where

π_i, π_{ij}, and G are given above,

$$y = \left[\frac{X_i - t}{\sqrt{\hat{\sigma}^6 \lambda_1}} \phi\left(\frac{X_i - t}{\sqrt{\hat{\sigma}^2 \lambda_1}} \right), \frac{X_i - t}{\sqrt{\hat{\sigma}^6 \lambda_2}} \phi\left(\frac{X_i - t}{\sqrt{\hat{\sigma}^2 \lambda_2}} \right), \cdots, \frac{X_i - t}{\sqrt{\hat{\sigma}^6 \lambda_m}} \phi\left(\frac{X_i - t}{\sqrt{\hat{\sigma}^2 \lambda_m}} \right) \right]^T$$

See Stefanski and Bay (1996) for more details.

9 References

Barlow, R. E., Bartholamu, D. J., Brenner, J. M., and Brunk, H. D. 1972. *Statistical Inference under Order Restrictions*. New York: John Wiley & Sons.

Sarndal, C. E., Swensson, B., and Wretman J. 1992. *Model Assisted Survey Sampling*. New York: Springer-Verlag.

Stefanski, L. A. and Bay, J. M. 1996. Simulation extrapolation deconvolution of finite population cumulative distribution function estimators. Biometrika 83: 407-417.