Finite Population Sampling and Inference: A Prediction Approach

Errata

p.12 . Line 9 should read

$$\pi_i = \binom{N-1}{n-1} \Big/ \binom{N}{n}$$

p. 23, problem 1.11. $\hat{T}_2$ should be

$$\hat{T}_2 = \sum_{i\in s} Y_i + (N-n) \sum_{i\in s} Y_i x_i^{-1} \Big/ \sum_{i\in s} x_i^{-1}$$

[not $i \in r$ in second two summations]

p. 24, problem 1.12. Line 3 should read

expansion estimator $\hat{T}_0$ of the total of the *y*'s corresponding to each of the 10 samples. Is there
[inserting "of the total of the *y*'s"]

p. 26, line 10 from bottom. The formula should read

$$\mathbf{g}_s'\mathbf{Y}_s - \boldsymbol{\gamma}'\mathbf{Y} = (\mathbf{g}_s' - \boldsymbol{\gamma}_s')\mathbf{Y}_s - \boldsymbol{\gamma}_r'\mathbf{Y}_r .$$

p. 28, line 8 from bottom. The line should read

BLU estimator of β is $\hat{\beta} = \bar{Y}_s / \bar{x}_s$, which is the least squares estimator introduced

[putting subscript on the $\bar{x}$]

p. 29, line 9 from top. The line should read

θ under model (2.1.1) is [dropping the hat on theta]

p. 31, line 5 from bottom. The line should read

$\mathbf{X} = \mathbf{1}_N$, $\mathbf{V} = \sigma^2 \mathbf{I}_N$, and $\hat{\boldsymbol{\beta}} = \bar{Y}_s \equiv \sum_s Y_i / n$. The *BLU* predictor is $\hat{T}_0 = \sum_s Y_i +$

[replacing $\mathbf{X} = \mathbf{I}_N$ by $\mathbf{X} = \mathbf{1}_N$]

p. 32, line 19 from top. The line should read

values 1 or 0 only, with probabilities p and $q = 1 - p$, respectively, and that

p. 34, line 2 from bottom. The line should read

***Theorem* 2.5.1**. Under conditions (i)-(iii), as $N, n \to \infty$, if $f \to 0$, then

[not Theorem 2.5.2]

p. 38, line 2 from top. The line should read

noninformative. Ignorable and nonignorable sampling has received much atten-

[putting a period after *noninformative*]

p. 53, line 16 from bottom. The line should read

the data better, perhaps considerably better, if we had all the data in the

p. 55, section 3.2.3, line 1 should read

We noted in Section 2.1 that the ratio estimator $\hat{T}_R = \hat{T}(0,1:x) = N\overline{Y}_s\, \overline{x}/\overline{x}_s$

[putting the bar on x_s]

p. 57, line 3 from top. The formula should read

$\hat{T}_{LR} = \hat{T}(1,1:1) = N\left[\overline{Y}_s + b_1(\overline{x} - \overline{x}_s)\right]$

[putting the bar on x_s]

p. 94, problem 3.18. The x density should be $f(x) = 0.04x\exp(-x/5)$.

p. 95, line 5 from bottom. The line should read

Consider the minimal model estimator $\hat{T}\left(x^{1/2}, x : x\right)$ or, more simply, the

[putting the hat on T]

p. 96, line 15 from top. The line should read

more general minimal estimator $\hat{T}\left(x^{1/2}, x : x\right)$. One can certainly make the ratio

[putting the hat on T]

p. 97, line 12 from top. The line should read

p-vector **c**. It follows that $\mathbf{c}'\mathbf{X}_s'\mathbf{V}_{ss}^{-1} = \mathbf{1}_s'$ or $\mathbf{c}'\mathbf{X}_s'\mathbf{V}_{ss}^{-1}\mathbf{Y}_s = \mathbf{1}_s'\mathbf{X}_s\hat{\boldsymbol{\beta}}$. Substituting

[putting a prime on 1_s]

p. 97, lines 17-20 from top. The lines should read

We have $\text{var}_M\left(\mathbf{1}_N'\mathbf{X}\hat{\boldsymbol{\beta}}\right) = \mathbf{1}_N'\mathbf{X}\mathbf{A}_s^{-1}\mathbf{X}'\mathbf{1}_N\,\sigma^2$,
$\text{var}_M\left(\mathbf{1}_s'\mathbf{Y}_s\right) + \text{var}_M\left(\mathbf{1}_r'\mathbf{Y}_r\right) = \mathbf{1}_N'\mathbf{V}\mathbf{1}_N\,\sigma^2$, and
$\text{cov}_M\left(\mathbf{1}_N'\mathbf{X}\hat{\boldsymbol{\beta}}, \mathbf{1}_s'\mathbf{Y}_s\right) = \mathbf{1}_N'\mathbf{X}\mathbf{A}_s^{-1}\mathbf{X}_s'\mathbf{1}_s\sigma^2$. Because $\mathbf{1}_s = \mathbf{V}_{ss}^{-1}\mathbf{X}_s\mathbf{c}$, we have
$\text{cov}_M\left(\mathbf{1}_N'\mathbf{X}\hat{\boldsymbol{\beta}}, \mathbf{1}_s'\mathbf{Y}_s\right) = \mathbf{1}_N'\mathbf{X}\mathbf{c}\sigma^2 = \mathbf{1}_N'\mathbf{V}\mathbf{1}_N\sigma^2$. Substituting the various
components produces (4.2.3) above.■

[adding σ^2's where appropriate]

p. 98, line 7 from bottom. The line should read

Proof: Lemma 4.2.1 applies so that the variance in (4.2.3) is minimized when

[not (4.2.5)]

p. 99, lines 9-11 from top. The lines should read

reduced form of $\hat{T}$, note that $\hat{T}(\mathbf{X}:\mathbf{V}) = \mathbf{1}_N'\mathbf{X}\hat{\boldsymbol{\beta}}$ from Lemma 4.2.1,
$\mathbf{1}_N'\mathbf{X} = \left(\mathbf{1}_N'\mathbf{V}^{1/2}\mathbf{1}_N\right)\left(\mathbf{1}_s'\mathbf{V}_{ss}^{-1/2}\mathbf{X}_s\right)\big/n$ since $s \in B(\mathbf{X}:\mathbf{V})$, $\mathbf{1}_s'\mathbf{V}_{ss}^{-1/2} = \mathbf{c}_1'\mathbf{X}_s'\mathbf{V}_{ss}^{-1}$ and
$\mathbf{c}_1'\mathbf{X}_s' = \mathbf{1}_s\mathbf{V}_{ss}^{1/2}$ since $\mathbf{V}^{1/2}\mathbf{1}_N = \mathbf{X}\mathbf{c}_1$ for some *p*-vector $\mathbf{c}_1$. Making these substitutions

[Lemma 4.2.1 not 3.2.1; $\mathbf{c}_1$ not $\mathbf{c}_2$]

p. 100, line 7 from top. The line should read

proof is left to exercise 4.6.

[not 4.5]

p. 103, line 5 from bottom. The line should read

$\mathbf{W} = 0.04\,diag\left(x^2\right)$. It may be verified that the median coefficient of variation

p. 126, line 12 from top. The line should read

linear function of the Y's, we can write $\hat{\Theta} = \mathbf{a}'\mathbf{Y}_s = \sum_s a_i Y_i$, where the a's are

[putting subscript s on $\mathbf{Y}$]

p. 130, line 1 from bottom. The formula should read

$$\hat{T} = \sum_{i \in s} Y_i + \sum_{i \in r} \hat{Y}_i = \sum_{i \in s} Y_i + \sum_{i \in r} \hat{\beta}\, x_i = \frac{\sum_{i=1}^{N} x_i \sum_{i \in s} Y_i}{\sum_{i \in s} x_i} = N\bar{x}\,\frac{\overline{Y}_s}{\overline{x}_s}$$

[putting bar on Y_s]

p. 131, line 4 from top. The formula should read

$$\mathrm{var}_M\left(\hat{T} - T\right) = \mathrm{var}_M\left(\sum_{i \in r} \hat{Y}_i - \sum_{i \in r} Y_i\right) = \mathrm{var}_M\left(\sum_{i \in r} \hat{Y}_i\right) + \mathrm{var}_M\left(\sum_{i \in r} Y_i\right), \qquad (5.1.9)$$

[all sums are over $i \in r$]

p. 135, line 13 from bottom. The line should read

dence intervals for $\Theta = \mathbf{q}'\boldsymbol{\beta}$, where $\boldsymbol{\beta}$ is the unknown parameter implied by M

p. 136, line 9 from top. The formula should read

(a) $h_{ii} = \mathbf{x}_i'(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{x}_i w_i$

p. 136, Lemma 5.3.1 (h). The formula should read

(h) $\sum_i h_{ij}^2 w_i = h_{jj} w_j$

p. 137, line 2 from top. The formula should read

$$\mathbf{a}'\mathbf{a} \le \mathbf{q}'\mathbf{A}^{-1}\mathbf{X}'\mathbf{W}\,diag\left[\max\left(w_i\right)\right]\mathbf{X}\mathbf{A}^{-1}\mathbf{q}$$

p. 142, line 7 from bottom. The formula should read

$$\left|\operatorname{cov}(r_i, r_l)\right| \le B_\psi \left\{ h_{il} + h_{li} + \sum_j \left(h_{ii} h_{jj} h_{ll} h_{jj} w_j^2 w_i^{-1} w_l^{-1} \right)^{1/2} \right\}$$

p. 143, line 6 from top. The line should read

and $\sum a_i^2 \ge c b_w$ by Lemma 5.3.1(j). Continuing from (5.4.5), we have

p. 144, line 9 from bottom. The line should read

Assumption (iv) is probably not entirely necessary, and in 5.5.2 we suggest

[not 5.5.1]

p. 145, lines 11-12 from bottom. The lines should read

For the jackknife, we will use $v_J\left(\hat{T}_r\right)$ as defined in (5.4.6) plus the estimator of $\operatorname{var}_M\left(T_r\right)$ that accompanies v_{J*} so that

p. 145, line 10 from bottom. Equation (5.5.3) should read

$$v_J\left(\hat{T} - T\right) = \frac{n-1}{n} \left\{ \sum_s \frac{a_i^2 r_i^2}{\left(1 - h_{ii}\right)^2} - n^{-1} \left[\sum_s \frac{a_i r_i}{\left(1 - h_{ii}\right)} \right]^2 \right\} \quad + \quad \frac{\sum_r v_j}{\sum_s v_j} \sum_s \hat{\psi}_i$$

[squaring $1 - h_{ii}$ in first summation]

p. 164, problem 5.3(b). Line 2 should read

(5.4.3).

p. 169, line 1 from bottom. The formula should read

$$\frac{n_h}{n} = \frac{N_h \sigma_h \big/ \sqrt{c_h}}{\sum_{h'} N_{h'} \sigma_{h'} \big/ \sqrt{c_{h'}}} \text{ for } h = 1, \ldots, H .$$

[adding prime to c_h in denominator]

p. 177, line 3 from bottom. The line should read

When the optimal allocation (6.2.11) is used and costs are all equal, this

[not 6.2.10]

p. 181, line 9 from top. The line should read

by size is used, then (6.2.11) implies that a stratum h will be a certainty if

[not 6.2.10]

p. 181, line 11 from bottom. The line should read

balanced sample achieves exactly the same error variance as an unstratified, weighted

p. 189, line 6 from top. The line should read

(6.5.5) is $\sigma^2 / n \left[\left(\sum_h N_h \sqrt{\bar{x}_h} \right)^2 - \left(N \bar{x}^{(1/2)} \right)^2 \right]$. This difference is non-negative because

p. 194, Table 6.7 up top. The Title should be

Numbers of Units per Stratum with Four Methods of Strata Formation in the Hospitals and Cancer Populations.

[four methods not three]

p. 194, line 11 from top. The line should read

among the five strata in Table 6.7 giving $n_h = 6$ in each stratum. When strata

[not Table 6.6]

p. 194, line 13 from top. The line should read

an equal allocation is optimal (see exercise 6.17). Thus, there is a logical

[not exercise 6.18]

p. 198, line 7 from bottom. The line should read

error variance under the working model (6.2.1), but are also consistent

[not 6.1.1]

p. 199, line 15 from bottom, The line should read

unit hi is $\hat{\boldsymbol{\beta}}_{h(hi)} = \hat{\boldsymbol{\beta}}_h - \mathbf{A}_h^{-1} \mathbf{x}_{hi} w_{hi} r_{hi} \big/ (1 - h_{hii})$. Substituting this into the expres-

[r_{hi} is not bold]

p. 200, line 10 from bottom. The line should read

$\hat{T}_{yr} = \hat{T}_{y\pi} + \hat{\mathbf{B}}'\left(\mathbf{T}_x - \hat{\mathbf{T}}_{x\pi}\right)$, defined by (2.7.1). Särndal, Swensson, and Wretman (1992,
[not 2.8.1]

p. 219, line 9 from bottom. The line should read

(7.4.9) and some rearrangement produces the error variance in the statement
[not 7.4.7]

p. 229, Example 7.8.1, line 5 should read

in Table 7.4. We have collapsed together the 35-64 and 65+ age groups, that
[not Table 7.3]

p. 233, line 14 from top. The formula should read

$$\hat{T} = \mathbf{1}_s'\mathbf{Y}_s + \mathbf{1}_r'\mathbf{X}_r\boldsymbol{\beta}^o + \mathbf{1}_r'\mathbf{Z}_r\hat{\boldsymbol{\gamma}} \qquad (7.9.6)$$

[$\mathbf{1}_r'\mathbf{X}_r\boldsymbol{\beta}^o$ not $\mathbf{1}_s'\mathbf{X}_r\boldsymbol{\beta}^o$]

p. 233, line 16 from top. The line should read

***Lemma* 7.9.1** The error variance of $\hat{T} = \mathbf{1}_s'\mathbf{Y}_s + \mathbf{1}_r'\mathbf{X}_r\boldsymbol{\beta}^o + \mathbf{1}_r'\mathbf{Z}_r\hat{\boldsymbol{\gamma}}$ under

[$\mathbf{1}_s'\mathbf{Y}_s$ not $\mathbf{1}_r'\mathbf{Y}_s$]

p. 239, line 15 from bottom. The line should read

As in Chapter 5, $r_i = Y_i\left(1 - h_{ii}\right) - \sum_{j \neq i} h_{ij} Y_j$ and under the working model

p. 240, line 12 from bottom. The line should read

corresponding to v_R, v_D, v_H, and v_{J*}, respectively. Since $h_{ii} = O\left(n^{-1}\right)$, each

p. 264, section 8.5, line 3 should read

used when the correct model requires $E_M\left(Y_{ij}\right)=\mathbf{x}'_{ij}\boldsymbol{\beta}$, the estimator of the total

[$\mathbf{x}'_{ij}$ not $\mathbf{x}'_i$]

p. 264, section 8.5, line 8 should read

$\mathbf{x}'_{ij}\boldsymbol{\beta}$, then

[$\mathbf{x}'_{ij}$ not $\mathbf{x}'_i$]

p. 269, lines 8-9 from top. The lines should read

(8.6.6) with no interaction the vector of nonsample auxiliary totals is $\mathbf{T}'_{xr}$ = (458, 220, 238, 45, 49, 123, 241, 17526) while for model (8.6.6) with

p. 292, problem 8.15, line 1 should read

Generalize Lemma 4.2.1 to the case of **V** a non-diagonal matrix, as

p. 312, line 16 should read

$$\mathbf{X}_{si}=\left(\mathbf{x}_{i1},\ldots,\mathbf{x}_{im_i}\right)',\ i=1,\ldots,n$$

[$i=1,\ldots,n$ not $j=1,\ldots,m_i$]

p. 316, 3 lines from bottom. Replace $\mathbf{W}_{s1}$ by $\mathbf{W}_{si}$ to get

$$\mathbf{A}_s=\begin{bmatrix}\mathbf{X}'_{s1(i)} & \mathbf{X}'_{si1}\\ \mathbf{B}'_{(i)} & \mathbf{B}'_i\end{bmatrix}\begin{bmatrix}\mathbf{W}_{s(i)} & \mathbf{0}\\ \mathbf{0} & \mathbf{W}_{si}\end{bmatrix}\begin{bmatrix}\mathbf{X}_{s1(i)} & \mathbf{B}_{(i)}\\ \mathbf{X}_{si1} & \mathbf{B}_i\end{bmatrix}$$

p. 317, line 13 should read

inverse as in Lemma 9.5.1 is $\boldsymbol{\beta}^o=\mathbf{G}\mathbf{X}'_s\mathbf{W}_s\mathbf{Y}_s$ with **G** defined in (9.5.22). The solution

[not 9.2.22]

p. 321, problem 9.5, line 1 should read

9.5 Show that under model (9.2.4),

p. 335, line 7:

and $\bar{Y}_{\ell hs} = \sum_{i \in s_h} Y_{\ell hi} / n_h$ $(\ell = 1, \ldots, p)$. In vector form $\hat{\mathbf{T}} = \sum_h N_h \bar{\mathbf{Y}}_{hs}$ where

p. 336, 6 lines below equation (10.2.11) should read

$\bar{Y}_{\ell hs}^{(\alpha)} = \varsigma_{h1\alpha} Y_{\ell h1} + \varsigma_{h2\alpha} Y_{\ell h2}$. It follows that $\hat{\mathbf{T}}^{(\alpha)} = \sum_h N_h \bar{\mathbf{Y}}_{hs}^{(\alpha)}$, paralleling the full-sample form, $\hat{\mathbf{T}} = \sum_h N_h \bar{\mathbf{Y}}_{hs}$. The *BHS* variance es-

[adding two summation signs]

p. 343, line 9 should read

expansion estimator is used. Rao and Shao (1996) proposed a repeatedly

[not 1993]

p. 359, line 6 from top. The line should read

tion increases. By the time π reaches 20%, the bisquare estimator has a root

p. 384, line 5 from top. The formula should read

$$\frac{\varepsilon_j}{\sqrt{v_j}} \le t_i + \left(\frac{\mathbf{x}_j'}{\sqrt{v_j}} - \frac{\mathbf{x}_i'}{\sqrt{v_i}} \right) \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \cong t_i + \left(\frac{\mathbf{x}_j'}{\sqrt{v_j}} - \frac{\mathbf{x}_i'}{\sqrt{v_i}} \right) \left(\hat{\boldsymbol{\beta}}_{(i)} - \boldsymbol{\beta} \right)$$

[putting a hat on $\boldsymbol{\beta}_{(i)}$]

p. 418, 5th line of proof of Lemma A.10.2 should read

Then (ii) means that $\mathbf{XGX'X} = \mathbf{XFX'X}$. By Lemma A.10.1(ii), $\mathbf{XGX'} =$

[not Then (i)…]

p. 422 The Cancer population is missing units 101 – 200. The full data set can be downloaded from this website.

p. 448, line 22 from top. The line should read

```
A <- t(Xs * Ws) %*% Xs
```

p. 448, line 24 from top. The line should read

```
a <- one.r %*% Xr %*% Ainv %*% t(Xs * Ws)
```

p. 448, line 25 from top. The line should read

```
H <- Xs %*% Ainv %*% t(Xs * Ws)
```

p. 456, line 13 from bottom. The line should read

```
A <- t(Xs * Ws) %*% Xs
```

p. 457, line 5 from top. The line should read

```
h[Zero] <- 0
```

p. 457 lines 1, 17, and 9 from bottom should read

```
#  W = diagonal of weight matrix for entire population
[the input is a vector]
```

p. 458 lines 11 and 15 from bottom should read

```
#  W = diagonal of weight matrix for entire population
[the input is a vector]
```

p. 468, 4 lines from bottom should read

Rao, J.N.K., and Shao, J. (1996), On Balanced Half-sample Variance Estimation in

[not 1993]

p. 475, Answers for Exercise 5.16 should be

(a) $\hat{T}\left(x^{1/2}, x : x\right)$= 12,355.87; $\sqrt{v_R}$ =636.42, $\sqrt{v_D}$ =709.08, $\sqrt{v_H}$ =677.05,

$\sqrt{v_{J*}}$ =832.62, $\sqrt{v_J}$ =812.87.

(b) leverages: 0.104, 0.098, 0.096, 0.094, 0.093, 0.080, 0.078, 0.071, 0.063, 0.055,

0.050, 0.053, 0.054, 0.057, 0.063, 0.071, 0.077, 0.077, 0.104, 0.564

p. 475, Answers for Exercise 5.17 should be

(a) $\hat{T}\left(x, x^2 : x^2\right) = 339{,}232.4$; $\sqrt{v_R} = 13{,}073.63$, $\sqrt{v_D} = 14{,}131.34$,

$\sqrt{v_H} = 14{,}042.81$, $\sqrt{v_{J*}} = 15{,}353.48$, $\sqrt{v_J} = 14{,}977.84$.

(b) leverages: 0.244, 0.187, 0.167, 0.120, 0.082, 0.059, 0.059, 0.050, 0.050, 0.051, 0.055, 0.056, 0.057, 0.057, 0.062, 0.074, 0.112, 0.125, 0.160, 0.171

p. 480, Exercise 8.5. The answer should read

Answer: $E_M(\hat{T}_U - T) = N\mu(\bar{M}_s - \bar{M})$