RECEIVER OPERATING CHARACTERISTIC CURVE FOR COMPLEX SURVEY DATA

TAMY H. M. TSUJIMOTO, JIANWEN CAI

University of North Carolina at Chapel Hill

Supplementary Material

S1 Assumptions and Notations

Following Han and Wellner (2021), we impose the following assumptions for the asymptotic convergence:

(A1.1) $\min_{1\leq i\leq N} \pi_i \geq \pi_0 > 0$

(A1.2) $\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{\xi_i}{\pi_i} - 1\right) = O_{\mathbb{P}_{d,m}}(1)$

(A1.3) There exist constant $K > 0$ such that

$$\sup_{N\in\mathbb{N}} \sup_{1\leq i\neq j\leq N} N|\pi_{ij} - \pi_i\pi_j| \leq K$$

(A1.4) The sample size n increases as the population size N increases, with

$$\lim_{N\to\infty} \frac{n}{N} = \lambda, \quad 0 < \lambda < 1$$

Let $\{V_i\}$ be a sequence of bounded i.i.d random variables defined on $(\Omega, \mathfrak{F}, \mathbb{P}_m)$. Let S_N^2 be the design-based variance of the Horvitz-Thompson estimator of the population mean, that is,

$$S_N^2 = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\pi_{ij} - \pi_i\pi_j}{\pi_i\pi_j} V_i V_j$$

(A2.1) Suppose that for N sufficiently large

$$\frac{1}{S_N}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\xi_i}{\pi_i}V_i - \frac{1}{N}\sum_{i=1}^{N}V_i\right) \to N(0,1)$$

holds under $\mathbb{P}_d$ ω-almost surely.

(A2.2) There exist constants $\mu_{\pi 1}, \mu_{\pi 2} \in \mathbb{R}$ such that

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\pi_{ii} - \pi_i^2}{\pi_i^2} \to \mu_{\pi 1} \quad \text{in } \mathbb{P}_m$$

$$\frac{1}{N}\sum_{i\neq j}^{N}\frac{\pi_{ij} - \pi_i\pi_j}{\pi_i\pi_j} \to \mu_{\pi 2} \quad \text{in } \mathbb{P}_m$$

Following Han and Wellner (2021), for $\{\pi_i\}_{i=1}^N$, $\{\xi_i\}_{i=1}^N$, $\{Y_i\}_{i=1}^N$, and a class $\mathcal{F}$ of real functions f we define the Horvitz-Thompson empirical measure as

$$\mathbb{P}_N^{\pi}(f) = \frac{1}{N}\sum_{i=1}^{N}\frac{\xi_i}{\pi_i}f(Y_i), \quad f \in \mathcal{F}, \tag{S1.1}$$

and the associated Horvitz-Thompson empirical process as

$$\mathbb{G}_N^{\pi}(f) = \sqrt{N}(\mathbb{P}_N^{\pi} - P)(f), \quad f \in \mathcal{F}. \tag{S1.2}$$

The usual empirical measure and empirical process ($\xi_i/\pi_i = 1$ for all $i = 1, \cdots, N$) will be denoted as $\mathbb{P}_N$ and $\mathbb{G}_N$.

S2 Proofs

Proof of Theorem 2.1 (a). Let $\mathcal{F} = \{f_{s,l} \equiv f_{s,l}(x,d) = I(x \leq s, d = l) : s \in \mathbb{R}, l \in \{0,1\}\}$, and $f_{s,d}, f_{u,d'} \in \mathcal{F}$, with $s, u \in \mathbb{R}$, and $d, d' \in \{0,1\}$. From Corollary 3.13 in Han and Wellner (2021), it follows that

$$\sqrt{n}(\mathbb{P}_N^{\pi} - \mathbb{P}_N) \rightsquigarrow \mathbb{G}^{\pi} \quad \text{in} \quad \ell^{\infty}(\mathcal{F}),$$

where $\mathbb{G}^\pi$ is a tight Gaussian process with covariance function

$$\mathrm{Cov}(\mathbb{G}^\pi(f_{s,d}), \mathbb{G}^\pi(f_{u,d'})) = \lambda\left(\mu_{\pi_1} P(f_{s,d} f_{u,d'}) + \mu_{\pi_2}(P f_{s,d})(P f_{u,d'})\right)$$

$$= \begin{cases} \lambda(1-p)\left[\mu_{\pi_1} G(s \wedge u) + \mu_{\pi_2}(1-p)G(s)G(u)\right], \quad d = d' = 0 \\ \lambda p\left[\mu_{\pi_1} F(s \wedge u) + \mu_{\pi_2} p F(s) F(u)\right], \quad d = d' = 1 \\ \lambda \mu_{\pi_2} p(1-p) G(s) F(u), \quad 0 = d \neq d' = 1 \\ \lambda \mu_{\pi_2} p(1-p) G(u) F(s), \quad 1 = d \neq d' = 0 \end{cases}$$

For $s \in \mathbb{R}$, the estimators for cdfs $G(s)$ and $F(s)$ are

$$G_n(s) = \frac{\mathbb{P}_N^\pi(f_{s,0})}{\mathbb{P}_N^\pi(f_{\infty,0})} \qquad F_n(s) = \frac{\mathbb{P}_N^\pi(f_{s,1})}{\mathbb{P}_N^\pi(f_{\infty,1})}$$

The ratio map $\phi(A, B) = A/B$ is Hadamard-differentiable with derivative $\phi'(\alpha, \beta) = \alpha/B - (A\beta)/B^2$. It follows from Functional Delta Method (Vaart and Wellner, 1996) that

$$\sqrt{n}\begin{bmatrix} G_n(s) - G_N(s) \\ F_n(s) - F_N(s) \end{bmatrix} \rightsquigarrow \begin{bmatrix} \mathbb{G}_0^\pi(s) \\ \mathbb{G}_1^\pi(s) \end{bmatrix} = \begin{bmatrix} (1-p)^{-1}(\mathbb{G}^\pi(f_{s,0}) - G(s)\mathbb{G}^\pi(f_{\infty,0})) \\ p^{-1}(\mathbb{G}^\pi(f_{s,1}) - F(s)\mathbb{G}^\pi(f_{\infty,1})) \end{bmatrix}.$$

The covariance structure of $\mathbb{G}_0^\pi$ and $\mathbb{G}_1^\pi$ can be computed as follows:

$$\begin{aligned}
\mathrm{Cov}(\mathbb{G}_0^\pi(s), \mathbb{G}_0^\pi(u)) =& (1-p)^{-2}\Big[\mathrm{Cov}(\mathbb{G}^\pi f_{s,0}, \mathbb{G}^\pi f_{u,0}) \\
& + G(s)G(u)\mathrm{Cov}(\mathbb{G}^\pi f_{\infty,0}, \mathbb{G}^\pi f_{\infty,0}) \\
& - G(u)\mathrm{Cov}(\mathbb{G}^\pi f_{s,0}, \mathbb{G}^\pi f_{\infty,0}) \\
& - G(s)\mathrm{Cov}(\mathbb{G}^\pi f_{\infty,0}, \mathbb{G}^\pi f_{u,0})\Big] \\
=& \lambda(1-p)^{-1}\Big[\mu_{\pi_1} G(s\wedge u) + \mu_{\pi_2}(1-p)G(s)G(u) + \\
& + \mu_{\pi_1} G(s)G(u) + \mu_{\pi_2}(1-p)G(s)G(u) \\
& - \mu_{\pi_1} G(u)G(s) - \mu_{\pi_2}(1-p)G(u)G(s) \\
& - \mu_{\pi_1} G(s)G(u) - \mu_{\pi_2}(1-p)G(s)G(u)\Big] \\
=& \lambda(1-p)^{-1}\mu_{\pi_1}\Big[G(s\wedge u) - G(s)G(u)\Big]
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(\mathbb{G}_1^\pi(s), \mathbb{G}_1^\pi(u)) =& p^{-2}\Big[\mathrm{Cov}(\mathbb{G}^\pi f_{s,1}, \mathbb{G}^\pi f_{u,1}) \\
& + F(s)F(u)\mathrm{Cov}(\mathbb{G}^\pi f_{\infty,1}, \mathbb{G}^\pi f_{\infty,1}) \\
& - F(u)\mathrm{Cov}(\mathbb{G}^\pi f_{s,1}, \mathbb{G}^\pi f_{\infty,1}) \\
& - F(s)\mathrm{Cov}(\mathbb{G}^\pi f_{\infty,1}, \mathbb{G}^\pi f_{u,1})\Big] \\
=& \lambda p^{-1}\Big[\mu_{\pi_1} F(s\wedge u) + \mu_{\pi_2} pF(s)F(u) \\
& + \mu_{\pi_1} F(s)F(u) + \mu_{\pi_2} pF(s)F(u) \\
& - \mu_{\pi_1} F(u)F(s) - \mu_{\pi_2} pF(u)F(s) \\
& - \mu_{\pi_1} F(s)F(u) + \mu_{\pi_2} pF(s)F(u)\Big] \\
=& \lambda p^{-1}\mu_{\pi_1}\Big[F(s\wedge u) - F(s)F(u)\Big]
\end{aligned}$$

$$\begin{aligned}\mathrm{Cov}(\mathbb{G}_0^\pi(s), \mathbb{G}_1^\pi(u)) =& p^{-1}(1-p)^{-1}\Big[\mathrm{Cov}(\mathbb{G}^\pi f_{s,0}, \mathbb{G}^\pi f_{u,1})\\ &+ G(s)F(u)\mathrm{Cov}(\mathbb{G}^\pi f_{\infty,0}, \mathbb{G}^\pi f_{\infty,1})\\ &- F(u)\mathrm{Cov}(\mathbb{G}^\pi f_{s,0}, \mathbb{G}^\pi f_{\infty,1})\\ &- G(s)\mathrm{Cov}(\mathbb{G}^\pi f_{\infty,0}, \mathbb{G}^\pi f_{u,1})\Big]\\ =&\lambda\mu_{\pi_2}G(s)F(u) + \lambda\mu_{\pi_2}G(s)F(u)\\ &- \lambda\mu_{\pi_2}G(s)F(u) - \lambda\mu_{\pi_2}G(s)F(u) = 0\end{aligned}$$

So, the covariance function is given by

$$\mathrm{Cov}(\mathbb{G}_d^\pi(s), \mathbb{G}_{d'}^\pi(u)) = \begin{cases}\lambda(1-p)^{-1}\mu_{\pi_1}(G(u\wedge s) - G(u)G(s)), \quad d = d' = 0\\ \lambda p^{-1}\mu_{\pi_1}(F(u\wedge s) - F(u)F(s)), \quad d = d' = 1\\ 0, \quad d \neq d'\end{cases} \tag{S2.3}$$

which implies that

$$\sqrt{n}\begin{bmatrix}G_n(s) - G_N(s)\\ F_n(s) - F_N(s)\end{bmatrix} \rightsquigarrow \begin{bmatrix}\mathbb{G}_0^\pi(s)\\ \mathbb{G}_1^\pi(s)\end{bmatrix} = \begin{bmatrix}\{\lambda(1-p)^{-1}\mu_{\pi_1}\}^{1/2}B_1(G(s))\\ \{\lambda p^{-1}\mu_{\pi_1}\}^{1/2}B_2(F(s))\end{bmatrix}, \tag{S2.4}$$

where $B_1(.)$ and $B_2(.)$ denote two independent Brownian bridges and $p = P(D = 1)$.

□

Proof of Theorem 2.1 (b). The ROC estimator depends on the pair (G_n, F_n) through the map $\psi(A, B) = B(A^{-1})$, where A^{-1} is the inverse map of A. The map $\psi(G, F)$ is Hadamard-differentiable (Lemma 12.2 and Lemma 12.7 from Kosorok (2008)) with derivative

$$\psi'(\alpha, \beta) = \beta(G^{-1}) - \frac{f(G^{-1})}{g(G^{-1})}\alpha(G^{-1})$$

It follows from (S2.4) and Functional Delta Method that

$$\begin{aligned}&\sqrt{n}(F_n \circ G_n^{-1}(s) - F_N \circ G_N^{-1}(s)) \\ &\rightsquigarrow \sqrt{\lambda\mu_{\pi_1}}\left\{p^{-1/2}B_1(F \circ G^{-1}(s)) - (1-p)^{-1/2}\frac{f(G^{-1})}{g(G^{-1})}B_2(s)\right\}\end{aligned}$$

where $B_1(.)$ and $B_2(.)$ denote two independent Brownian bridges and $p = P(D = 1)$.

□

Proof of Theorem 2.1 (c). It follows from Theorem 2.1 (b) and Continuous Mapping Theorem that

$$\begin{aligned}\sqrt{n}(A_n - A_N) &= \int_0^1 \sqrt{n}(R_n(s) - R_N(s))\ ds \\ &\rightsquigarrow \int_0^1 \mathbb{W}\{G^{-1}(1-s)\}ds \sim N(0, \sigma^2)\end{aligned}$$

where

$$\begin{aligned}\sigma^2 &= \int_0^1 \int_0^1 \sigma^2\{G^{-1}(1-s), G^{-1}(1-t)\}\ ds\ dt \\ &= \int_0^1 \int_0^1 \sigma^2\{G^{-1}(s), G^{-1}(t)\}\ ds\ dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sigma^2(s,t)\ dG(s)\ dG(t)\end{aligned}$$

□

S3 Supplemental Simulation Tables

Table 1: Relative Bias (in %) of the SVY, WT, and UN, estimators for the super-population ROC curve with finite population size N, disease proportion p, and sampling fraction λ under SSRS.

FPR	Method	N = 50,000				N = 100,000			
		$p = 5\%$		$p = 25\%$		$p = 5\%$		$p = 25\%$	
		$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$
0.1	SVY	0.8	0.2	0.1	0.2	0.8	0.6	0.0	0.2
	WT	0.8	0.2	0.1	0.2	0.8	0.6	0.0	0.2
	UN	32.5	32.0	31.7	31.8	32.4	32.4	31.8	31.8
0.2	SVY	0.7	0.4	-0.2	-0.4	0.6	0.6	-0.4	-0.2
	WT	0.7	0.4	-0.2	-0.4	0.6	0.6	-0.4	-0.2
	UN	20.4	20.2	19.4	19.3	20.4	20.4	19.4	19.4
0.3	SVY	0.2	0.1	0.0	0.0	0.1	0.1	0.0	0.1
	WT	0.2	0.1	0.0	0.0	0.1	0.1	0.0	0.1
	UN	13.6	13.5	13.5	13.5	13.6	13.6	13.5	13.5
0.4	SVY	0.1	0.2	0.0	0.0	0.1	0.1	0.0	0.1
	WT	0.1	0.2	0.0	0.0	0.1	0.1	0.0	0.1
	UN	9.6	9.6	9.4	9.4	9.6	9.6	9.4	9.5
0.5	SVY	0.0	0.0	0.1	0.0	-0.1	0.0	0.0	0.1
	WT	0.0	0.0	0.1	0.0	-0.1	0.0	0.0	0.1
	UN	6.5	6.5	6.6	6.6	6.5	6.6	6.6	6.6
0.6	SVY	0.0	0.0	0.1	0.1	-0.1	0.0	0.0	0.1
	WT	0.0	0.0	0.1	0.1	-0.1	0.0	0.0	0.1
	UN	4.4	4.4	4.5	4.5	4.4	4.4	4.5	4.5
0.7	SVY	-0.1	-0.1	0.0	0.0	-0.1	-0.1	-0.1	0.0
	WT	-0.1	-0.1	0.0	0.0	-0.1	-0.1	-0.1	0.0
	UN	2.7	2.7	2.8	2.8	2.6	2.7	2.7	2.7
0.8	SVY	-0.2	-0.1	0.0	0.0	-0.2	-0.1	0.0	0.0
	WT	-0.2	-0.1	0.0	0.0	-0.2	-0.1	0.0	0.0
	UN	1.3	1.4	1.5	1.5	1.4	1.4	1.5	1.5
0.9	SVY	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	WT	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	UN	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6

Table 2: Relative Bias (in %) of the SVY, WT, and UN estimators for the super-population ROC curve with finite population size N, disease proportion p, and sampling fraction λ under STSCS.

		$N = 50{,}000$				$N = 100{,}000$			
		$p = 5\%$		$p = 25\%$		$p = 5\%$		$p = 25\%$	
FPR	Method	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$
0.1	SVY	-0.1	0.1	0.2	0.0	-0.2	-0.2	0.2	0.0
	WT	-0.1	0.1	0.2	0.0	-0.2	-0.2	0.2	0.0
	UN	30.6	30.8	30.9	30.8	30.8	30.6	31.1	30.9
0.2	SVY	0.2	0.5	0.0	-0.1	0.3	0.0	0.1	-0.1
	WT	0.2	0.5	0.0	-0.1	0.3	0.0	0.1	-0.1
	UN	20.6	20.9	20.3	20.4	20.7	20.4	20.5	20.3
0.3	SVY	0.0	0.3	-0.4	-0.4	0.2	0.1	-0.2	-0.4
	WT	0.0	0.3	-0.4	-0.4	0.2	0.1	-0.2	-0.4
	UN	14.6	15.0	14.2	14.2	14.8	14.7	14.3	14.2
0.4	SVY	-0.3	0.0	-0.3	-0.3	-0.4	-0.3	-0.3	-0.3
	WT	-0.3	0.0	-0.3	-0.3	-0.4	-0.3	-0.3	-0.3
	UN	10.3	10.5	10.3	10.3	10.3	10.3	10.4	10.3
0.5	SVY	-0.4	-0.2	-0.4	-0.3	-0.5	-0.4	-0.2	-0.3
	WT	-0.4	-0.2	-0.4	-0.3	-0.5	-0.4	-0.2	-0.3
	UN	7.2	7.4	7.4	7.4	7.2	7.3	7.5	7.4
0.6	SVY	-0.3	-0.1	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2
	WT	-0.3	-0.1	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2
	UN	5.0	5.3	5.2	5.2	5.2	5.2	5.3	5.2
0.7	SVY	-0.1	-0.1	-0.1	-0.1	-0.1	-0.2	-0.1	-0.1
	WT	-0.1	-0.1	-0.1	-0.1	-0.1	-0.2	-0.1	-0.1
	UN	3.4	3.4	3.4	3.5	3.4	3.4	3.5	3.5
0.8	SVY	-0.1	0.0	0.0	0.0	-0.1	-0.1	0.0	0.0
	WT	-0.1	0.0	0.0	0.0	-0.1	-0.1	0.0	0.0
	UN	2.0	2.0	2.1	2.1	2.0	2.0	2.1	2.0
0.9	SVY	0.1	0.1	0.0	0.0	0.1	0.1	0.0	0.0
	WT	0.1	0.1	0.0	0.0	0.1	0.1	0.0	0.0
	UN	0.9	1.0	0.9	0.9	1.0	1.0	0.9	0.9

Table 3: Estimates of empirical (EMP) and asymptotic standard error of the SVY, WT, and UN estimators for the super-population ROC curve with finite population size N, disease proportion p, and sampling fraction λ under SSRS.

FPR	Method	$N = 50{,}000$				$N = 100{,}000$			
		$p = 5\%$		$p = 25\%$		$p = 5\%$		$p = 25\%$	
		$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$
0.1	EMP	0.049	0.035	0.026	0.019	0.035	0.025	0.018	0.013
	SVY	0.051	0.036	0.026	0.019	0.036	0.025	0.019	0.013
	WT	0.045	0.032	0.023	0.016	0.032	0.023	0.016	0.012
	UN	0.046	0.032	0.023	0.017	0.033	0.023	0.017	0.012
0.2	EMP	0.051	0.036	0.026	0.018	0.036	0.026	0.018	0.013
	SVY	0.051	0.036	0.026	0.018	0.036	0.026	0.018	0.013
	WT	0.046	0.032	0.023	0.016	0.032	0.023	0.016	0.012
	UN	0.043	0.031	0.021	0.015	0.031	0.022	0.015	0.011
0.3	EMP	0.050	0.035	0.025	0.017	0.036	0.025	0.017	0.012
	SVY	0.049	0.035	0.024	0.017	0.035	0.025	0.017	0.012
	WT	0.043	0.031	0.022	0.015	0.031	0.022	0.015	0.011
	UN	0.040	0.028	0.019	0.014	0.028	0.020	0.014	0.010
0.4	EMP	0.047	0.033	0.023	0.016	0.034	0.023	0.016	0.011
	SVY	0.045	0.032	0.022	0.016	0.032	0.023	0.016	0.011
	WT	0.040	0.028	0.020	0.014	0.028	0.020	0.014	0.010
	UN	0.035	0.025	0.017	0.012	0.025	0.018	0.012	0.009
0.5	EMP	0.042	0.030	0.020	0.014	0.031	0.021	0.015	0.010
	SVY	0.040	0.028	0.020	0.014	0.028	0.020	0.014	0.010
	WT	0.035	0.025	0.017	0.012	0.025	0.018	0.012	0.009
	UN	0.031	0.022	0.015	0.010	0.022	0.016	0.010	0.007
0.6	EMP	0.037	0.026	0.018	0.013	0.027	0.018	0.013	0.008
	SVY	0.034	0.024	0.017	0.012	0.024	0.017	0.012	0.008
	WT	0.031	0.022	0.015	0.010	0.022	0.015	0.010	0.007
	UN	0.026	0.019	0.012	0.009	0.019	0.013	0.009	0.006
0.7	EMP	0.031	0.021	0.015	0.010	0.022	0.015	0.010	0.007
	SVY	0.028	0.020	0.014	0.010	0.020	0.014	0.010	0.007
	WT	0.025	0.018	0.012	0.009	0.018	0.013	0.009	0.006
	UN	0.021	0.015	0.010	0.007	0.015	0.011	0.007	0.005
0.8	EMP	0.023	0.016	0.011	0.008	0.016	0.011	0.008	0.005
	SVY	0.021	0.015	0.010	0.007	0.015	0.011	0.007	0.005
	WT	0.019	0.013	0.009	0.006	0.013	0.010	0.006	0.005
	UN	0.016	0.011	0.008	0.005	0.011	0.008	0.005	0.004
0.9	EMP	0.015	0.010	0.007	0.005	0.011	0.007	0.005	0.003
	SVY	0.015	0.010	0.006	0.005	0.010	0.007	0.005	0.003
	WT	0.013	0.008	0.006	0.004	0.009	0.006	0.004	0.003
	UN	0.011	0.007	0.005	0.003	0.007	0.005	0.003	0.002

Table 4: Estimates of empirical (EMP) and asymptotic standard error of the SVY, WT, and UN estimators for the super-population ROC curve with finite population size N, disease proportion p, and sampling fraction λ under STSCS.

FPR	Method	$N = 50{,}000$				$N = 100{,}000$			
		$p = 5\%$		$p = 25\%$		$p = 5\%$		$p = 25\%$	
		$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$
0.1	EMP	0.047	0.034	0.025	0.018	0.033	0.023	0.018	0.012
	SVY	0.047	0.033	0.025	0.017	0.033	0.024	0.017	0.012
	WT	0.042	0.030	0.022	0.015	0.030	0.021	0.015	0.011
	UN	0.045	0.032	0.023	0.016	0.032	0.022	0.016	0.012
0.2	EMP	0.051	0.037	0.027	0.019	0.037	0.026	0.019	0.013
	SVY	0.052	0.036	0.026	0.019	0.036	0.026	0.019	0.013
	WT	0.046	0.032	0.023	0.016	0.032	0.023	0.016	0.012
	UN	0.046	0.032	0.023	0.016	0.032	0.023	0.016	0.012
0.3	EMP	0.052	0.038	0.027	0.019	0.037	0.026	0.019	0.013
	SVY	0.051	0.036	0.026	0.018	0.036	0.026	0.018	0.013
	WT	0.046	0.032	0.023	0.016	0.032	0.023	0.016	0.011
	UN	0.044	0.031	0.022	0.015	0.031	0.022	0.015	0.011
0.4	EMP	0.050	0.036	0.026	0.019	0.035	0.025	0.018	0.012
	SVY	0.049	0.034	0.024	0.017	0.035	0.024	0.017	0.012
	WT	0.043	0.031	0.022	0.015	0.031	0.022	0.015	0.011
	UN	0.041	0.029	0.020	0.014	0.029	0.020	0.014	0.010
0.5	EMP	0.047	0.033	0.024	0.017	0.033	0.024	0.017	0.011
	SVY	0.045	0.032	0.022	0.016	0.032	0.022	0.016	0.011
	WT	0.040	0.028	0.020	0.014	0.028	0.020	0.014	0.010
	UN	0.036	0.026	0.018	0.012	0.026	0.018	0.012	0.009
0.6	EMP	0.042	0.030	0.021	0.015	0.029	0.021	0.015	0.010
	SVY	0.040	0.028	0.020	0.014	0.028	0.020	0.014	0.010
	WT	0.035	0.025	0.017	0.012	0.025	0.018	0.012	0.009
	UN	0.032	0.022	0.015	0.011	0.022	0.016	0.011	0.008
0.7	EMP	0.036	0.025	0.018	0.012	0.025	0.017	0.013	0.009
	SVY	0.034	0.024	0.017	0.012	0.024	0.017	0.012	0.008
	WT	0.030	0.021	0.015	0.010	0.021	0.015	0.010	0.007
	UN	0.026	0.019	0.013	0.009	0.019	0.013	0.009	0.006
0.8	EMP	0.027	0.020	0.014	0.010	0.020	0.014	0.010	0.007
	SVY	0.026	0.019	0.013	0.009	0.019	0.013	0.009	0.006
	WT	0.023	0.016	0.011	0.008	0.016	0.012	0.008	0.006
	UN	0.020	0.014	0.010	0.007	0.014	0.010	0.007	0.005
0.9	EMP	0.018	0.013	0.009	0.006	0.013	0.009	0.006	0.005
	SVY	0.018	0.012	0.008	0.006	0.012	0.009	0.006	0.004
	WT	0.016	0.011	0.007	0.005	0.011	0.008	0.005	0.004
	UN	0.013	0.009	0.006	0.004	0.009	0.006	0.004	0.003

Table 5: Coverage Probabilities (in %) for 95% confidence intervals of the UN, WT, and SVY estimators for the super-population ROC curve with finite population size N, disease proportion p, and sampling fraction λ under SSRS.

		$N = 50{,}000$				$N = 100{,}000$			
		$p = 5\%$		$p = 25\%$		$p = 5\%$		$p = 25\%$	
FPR	Method	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$
0.1	SVY	95.9	95.3	95.2	94.8	95.4	95.2	95.9	95.6
	WT	92.5	92.1	92.1	91.8	92.6	92.3	92.4	92.6
	UN	22.9	4.1	0.0	0.0	2.6	0.1	0.0	0.0
0.2	SVY	95.3	94.6	94.7	95.0	95.1	95.3	95.5	95.2
	WT	92.2	91.7	92.1	91.2	91.3	91.8	91.9	92.1
	UN	29.0	6.6	0.2	0.0	5.7	0.1	0.0	0.0
0.3	SVY	94.3	94.1	94.4	95.3	93.7	94.7	94.7	95.3
	WT	90.9	90.7	91.1	92.1	90.8	90.8	91.4	92.3
	UN	40.1	12.5	0.6	0.0	13.7	0.8	0.0	0.0
0.4	SVY	94.1	94.2	94.2	94.4	92.7	93.9	94.2	95.3
	WT	90.4	90.6	90.7	91.4	89.4	90.5	91.0	91.7
	UN	48.8	22.2	2.5	0.0	20.8	2.6	0.0	0.0
0.5	SVY	92.6	93.5	94.1	94.2	93.2	94.1	93.3	94.4
	WT	89.1	89.7	91.2	90.8	88.4	91.1	89.3	91.4
	UN	57.4	34.1	5.9	0.2	34.5	10.5	0.0	0.0
0.6	SVY	92.1	92.8	93.8	93.7	92.2	94.3	92.7	94.5
	WT	88.9	89.5	90.3	90.6	88.9	91.0	89.6	91.1
	UN	64.2	45.9	15.4	1.2	47.2	19.9	1.6	0.0
0.7	SVY	90.4	93.0	92.9	93.7	92.8	94.0	93.2	93.4
	WT	86.9	89.7	88.6	90.1	88.5	90.5	89.6	89.9
	UN	70.7	59.2	33.6	7.0	59.2	39.2	7.3	0.3
0.8	SVY	91.9	91.2	92.5	94.2	92.2	93.3	93.8	93.4
	WT	89.3	87.9	89.0	90.9	88.5	89.6	89.8	89.6
	UN	78.3	71.5	50.2	25.2	72.2	59.2	25.5	4.3
0.9	SVY	98.6	86.9	91.1	92.3	88.0	89.1	92.2	93.6
	WT	97.8	84.0	87.1	88.1	84.5	85.1	88.2	89.7
	UN	99.8	81.1	67.7	54.4	82.0	71.6	53.9	29.9

Table 6: Coverage Probabilities (in %) for 95% confidence intervals of the UN, WT, and SVY estimators for the super-population ROC curve with finite population size N, disease proportion p, and sampling fraction λ under STSCS.

		$N = 50{,}000$				$N = 100{,}000$			
		$p = 5\%$		$p = 25\%$		$p = 5\%$		$p = 25\%$	
FPR	Method	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 5\%$	$\lambda = 10\%$
0.1	SVY	94.4	94.3	94.2	93.6	94.8	95.0	94.0	94.5
	WT	91.3	90.8	91.6	89.7	91.9	92.2	91.2	91.1
	UN	50.2	20.1	3.8	0.0	20.2	2.8	0.1	0.0
0.2	SVY	94.8	95.1	94.3	93.6	95.0	95.3	94.0	94.5
	WT	91.4	91.3	91.1	90.1	91.4	92.0	90.6	90.5
	UN	49.9	19.0	4.4	0.1	20.4	2.6	0.0	0.0
0.3	SVY	94.3	94.2	93.9	93.3	95.0	95.3	93.8	94.4
	WT	90.8	90.1	90.5	89.0	90.9	91.6	90.3	90.8
	UN	53.5	23.6	5.8	0.1	25.4	4.1	0.2	0.0
0.4	SVY	93.9	94.8	94.3	92.0	94.4	94.7	93.5	94.6
	WT	90.8	91.0	90.3	88.7	90.6	91.2	89.9	91.1
	UN	59.1	33.1	10.1	0.6	34.6	9.0	0.8	0.0
0.5	SVY	93.2	93.5	93.5	92.9	93.9	93.8	93.5	94.0
	WT	89.4	89.7	90.0	88.7	90.6	89.4	90.0	90.5
	UN	65.1	42.0	16.7	1.2	44.5	17.2	1.5	0.1
0.6	SVY	92.8	93.6	94.0	93.7	94.2	93.2	92.6	94.1
	WT	89.6	90.5	90.6	90.2	90.7	89.5	88.6	90.1
	UN	69.9	50.1	24.1	3.4	51.1	24.9	4.0	0.1
0.7	SVY	92.2	93.0	93.5	93.3	92.8	94.3	92.4	94.1
	WT	88.8	89.0	89.7	90.2	89.0	90.4	88.8	90.0
	UN	72.5	59.3	36.4	11.1	59.7	38.6	10.9	0.6
0.8	SVY	91.6	92.0	91.8	92.8	92.7	92.8	93.1	93.0
	WT	88.5	88.7	87.5	89.4	89.0	89.7	89.2	88.6
	UN	77.4	68.2	48.7	22.4	68.0	54.0	23.7	4.8
0.9	SVY	95.6	88.5	91.3	92.6	88.1	89.4	91.1	91.7
	WT	91.5	85.8	87.7	88.4	85.9	86.5	87.3	87.6
	UN	81.1	73.6	65.2	47.4	72.1	63.3	46.7	24.3

Table 7: Relative Bias (in %) of the SVY and BIN estimators for the super-population ROC curve with finite population size N, disease proportion p, and sampling fraction λ under SSRS.

FPR	Method	N = 50,000				N = 100,000			
		p = 5%		p = 25%		p = 5%		p = 25%	
		λ = 5%	λ = 10%	λ = 5%	λ = 10%	λ = 5%	λ = 10%	λ = 5%	λ = 10%
0.1	SVY	0.8	0.2	0.1	0.2	0.8	0.6	0.0	0.2
	BIN	2.7	2.6	2.1	2.0	2.6	2.6	2.0	2.0
0.2	SVY	0.7	0.4	-0.2	-0.4	0.6	0.6	-0.4	-0.2
	BIN	1.5	1.4	1.4	1.4	1.4	1.4	1.4	1.4
0.3	SVY	0.2	0.1	0.0	0.0	0.1	0.1	0.0	0.1
	BIN	1.3	1.2	0.8	0.8	1.1	1.2	0.8	0.8
0.4	SVY	0.1	0.2	0.0	0.0	0.1	0.1	0.0	0.1
	BIN	0.7	0.6	0.3	0.3	0.5	0.6	0.3	0.3
0.5	SVY	0.0	0.0	0.1	0.0	-0.1	0.0	0.0	0.1
	BIN	0.3	0.2	-0.1	0.0	0.1	0.2	0.0	0.0
0.6	SVY	0.0	0.0	0.1	0.1	-0.1	0.0	0.0	0.1
	BIN	-0.1	-0.1	-0.2	-0.2	-0.2	-0.1	-0.2	-0.2
0.7	SVY	-0.1	-0.1	0.0	0.0	-0.1	-0.1	-0.1	0.0
	BIN	-0.3	-0.4	-0.4	-0.4	-0.4	-0.3	-0.3	-0.3
0.8	SVY	-0.2	-0.1	0.0	0.0	-0.2	-0.1	0.0	0.0
	BIN	-0.4	-0.4	-0.4	-0.4	-0.5	-0.4	-0.4	-0.4
0.9	SVY	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	BIN	-0.4	-0.3	-0.3	-0.3	-0.4	-0.3	-0.3	-0.3

Bibliography

Han, Q. and J. A. Wellner (2021). Complex sampling designs: Uniform limit theorems and applications. *The Annals of Statistics 49*(1), 459–485.

Kosorok, M. R. (2008). *Introduction to empirical processes and semiparametric inference.* Springer.

Vaart, A. W. and J. A. Wellner (1996). *Weak convergence and empirical processes: with applications to statistics.* Springer.