Biostatistics (2009), **1**, 1, *pp.* i–ix
doi:10.1093/biostatistics/bio000

Supplementary materials for "Evaluating Prognostic Accuracy of Biomarkers under Nested Case-control Studies"

Tianxi Cai

Department of Biostatistics,
Harvard School of Public Health, Boston, MA 02115

YINGYE ZHENG

Department of Biostatistics and Biomathematics,
Fred Hutchinson Cancer Research Center, Seattle, WA 98109
yzheng@fhcrc.org

APPENDIX

Throughout, we drop the superscript w in the manuscript for ease of presentation. Define $\mathcal{N}_i(t) = I(X_i \leqslant t)\delta_i$, $\bar{\mathcal{N}}(t) = N^{-1}\sum_{i=1}^N \mathcal{N}_i(t)$, $\bar{\pi}(t) = N^{-1}\sum_{i=1}^N I(X_i \geqslant t)$, $\widehat{\mathcal{N}}(t) = N^{-1}\sum_{i=1}^N \widehat{w}_i\mathcal{N}_i(t)$, $\pi(t) = P(X_i \geqslant t)$, $A_{\rm NCC}(t) = E\{\mathcal{N}_i(t)\}$,

$$\Lambda_{\rm NCC}(t) = \int_0^t \frac{dA_{\rm NCC}(u)}{\pi(u)}, \quad \text{and} \quad G(t) = \exp\{-m\Lambda_{\rm NCC}(t)\}$$

We assume that the censoring time C has a finite support $[0, \tau]$, which is shorter than that of the event time T with $P(T > \tau) > 0$. The marker Y is assumed to be continuous and bounded with $|Y| \leqslant \mathcal{Y}_0 < \infty$ and the true parameter value β_0 is assumed to be an interior point of a compact parameter space Ω. Without loss of generality, we assume that $Y \geqslant 0$. Throughout, unless noted otherwise, the sup over time t is taken over $[0, \tau]$ and the sup over β is taken over Ω. We use the notation $\lesssim$ to denote bounded up to a constant and $\simeq$ to denote equivalence up to $o_p(1)$. In addition, we assume the joint density of Y, T and C has continuous derivatives.

For sampling probabilities, we note that from similar arguments for the consistency of the product limit Kaplan-Meier estimator that

$$\sup_t |\widehat{G}(t) - G(t)| = O_p(N^{-\frac{1}{2}}) \tag{A.1}$$

and thus $\max_i |\widehat{p}_i - p_i| = O_p(N^{\frac{1}{2}})$, where $p_i = \delta_i + (1 - \delta_i)\{1 - G(X_i)\}$.

A. ASYMPTOTIC VARIANCE WITH FINITE POPULATION SAMPLING

Under a finite population sampling of a NCC design, let $U(\cdot)$ be any given function of $\mathbf{D}_i$ such that $U_i = U(\mathbf{D}_i)$ has mean 0 and finite variance, and the total variation of $U(\mathbf{D}_i)$ is bounded by a constant. We first derive a generic form of the asymptotic variance of $N^{-\frac{1}{2}} \sum_{i=1}^{N} \widehat{w}_i U_i$, where $\widehat{w}_i = V_i/p_i$. We note that the NCC sampling variables $\{V_1, \cdots, V_n\}$ are weakly correlated and thus the asymptotic variance of weighted estimators needs to account for such a correlation. We have

$$\begin{aligned}
&\text{var}\left\{N^{-\frac{1}{2}} \sum_{i=1}^{N} \widehat{w}_i U_i\right\} \\
&= E\left[\text{var}\left\{N^{-\frac{1}{2}} \sum_{i=1}^{N} w_i U_i \mid \mathscr{D}\right\}\right] + \text{var}\left[E\left\{N^{-\frac{1}{2}} \sum_{i=1}^{N} \widehat{w}_i U_i \mid \mathscr{D}\right\}\right] \\
&= E\left[N^{-1} \sum_{i=1}^{N} \frac{1 - \widetilde{p}_i}{\widetilde{p}_i} U_i^2 + N^{-1} \sum_{i \neq j} \text{cov}(\widehat{w}_i, \widehat{w}_j \mid \mathscr{D}) U_i U_j\right] + \sigma_{0U}^2
\end{aligned}$$

where $\sigma_{0U}^2 = E(U_i^2)$. Following from the arguments given in Samuelsen (1997), it can be shown that

$$\text{cov}\left\{\widehat{w}_i U_i, \widehat{w}_j U_j \mid \mathscr{D}\right\} = -\frac{m}{n} \int \eta_u(t; X_i, \delta_i) \eta_u(t; X_j, \delta_j) \frac{d\Lambda_{\text{NCC}}(t)}{\pi(t)} + O_p(n^{-3/2}),$$

where $\eta_u(t; X_i, \delta_i) = E\{U_i I(X_i > t)(1 - p_i)/p_i\}$. Therefore, $N^{-\frac{1}{2}} \sum_{i=1}^{N} \widehat{w}_i U_i$ has asymptotic variance

$$\sigma_U^2 = E\left\{\frac{U_i^2}{p_i}\right\} - m \int \frac{\eta_U(t)^2 d\Lambda_{\text{NCC}}(t)}{\pi(t)} \tag{A.1}$$

When the controls are selected based on an additional matching variable Z with L strata, we may obtain $\widehat{\rho}_{ij}$ accordingly. Following some algebraic manipulation, it can be shown that adjustment for variance would be

$$
\begin{aligned}
&- m \sum_{l=1}^{L} \int \frac{E\{I(X_i \geqslant t, Z_i = \mathfrak{z}_l) U_i\}^2}{\pi^{(l)}(t) \wp_l} d\Lambda_{\text{NCC}}^{(l)}(t) \\
= &- m \sum_{l=1}^{L} \wp_l \int \frac{E\{I(X_i \geqslant t) U_i \mid Z_i = \mathfrak{z}_l\}^2}{\pi^{(l)}(t)} d\Lambda_{\text{NCC}}^{(l)}(t),
\end{aligned}
$$

where $A_{\text{NCC}}^{(l)}(t)s = E\{I(X_k \leqslant t)\delta_k \mid Z_k = \mathfrak{z}_l\}$, $\Lambda_{\text{NCC}}^{(l)}(t) = \int_0^t dA_{\text{NCC}}^{(l)}(u)/\pi^{(l)}(u)$, $\pi^{(l)}(t) = P(X_k \geqslant t \mid Z_k = \mathfrak{z}_l)$ and $\{\mathfrak{z}_1, \cdots \mathfrak{z}_L\}$ are the unique values of Z among the cases and $\wp_l = P(Z_i = \mathfrak{z}_l)$.

B. Consistency and Asymptotic Normality for a Generic IPW estimator

To derive the asymptotic properties of our IPW estimators, we first note that since the NCC sampling variables $\{V_1, \cdots, V_n\}$ are weakly correlated, and thus the standard convergence theory derived for independent identically distributed (i.i.d.) case does not apply. Here we use the results on the strong and weak convergence of weighted sums of negative associated (NegA) dependent variables (Liang *and others*, 2004; Liang and Baek, 2006) , to establish the consistency and asympotitic normality of the IPW process for NCC design with finite population sampling. The key is to show that $\widehat{U} = N^{-1}\sum_{i=1}^{N} \widehat{w}_i U_i$ can be viewed as weighted sums of NegA dependent variables, and it satisfies the conditions required for tightness and weak convergence of the NegA process.

If we let V_{0ji} denote whether the ith subject was selected as a control for the jth failure time in the NCC sample, then $\{V_{0ji}\}$ are NegA random variables and thus $\{V_{0i} = I(\sum_{j=1}^{n} V_{0ji} > 0), i = 1, \cdots n\}$ and $\{a_i V_{0i}, i = 1, ..., n\}$ are also NegA for $a_i \geqslant 0$ (Joag-Dev and Proschan, 1983). This indicates that conditional on $\mathscr{D}$, $\{\widehat{w}_i - 1, i = 1, ..., n\}$ are negatively associated random variables with mean 0. For simplicity, we focus the setting with no additional matching variables but note that the same arguments can be used to justify the case with additional matching on Z.

We next provide justifications for the convergence of $\widehat{U} = N^{-1}\sum_{i=1}^{N} \widehat{w}_i U_i \to 0$ in probability and $N^{\frac{1}{2}}\widehat{U} \to N(0, \sigma_U^2)$. To this end, we write $\widehat{U} = \widetilde{U} + \widehat{\mathcal{U}}_w$, where $\widetilde{U} = N^{-1}\sum_{i=1}^{N} U_i$ and $\widehat{\mathcal{U}}_w = N^{-1}\sum_{i=1}^{N}(\widehat{w}_i - 1)U_i$. Obviously, $\widetilde{U}$ and $\widehat{\mathcal{U}}_w$ are independent given $\mathscr{D}$. By the standard law of large numbers and central limit theorem, $\widetilde{U} \to 0$ almost surely and $N^{\frac{1}{2}}\widetilde{U} \to N(0, \sigma_{0U}^2)$. On the other hand, conditional on $\mathscr{D}$, $\widehat{\mathcal{U}}_w$ is a weighted sum of negatively associated random variables. Since U_i's are bounded and $\widehat{p}_i$ are bounded away from 0, it is straightforward to see that $\limsup_{N\to\infty} N^{-1}\sum_{i=1}^{N} U_i^2 <$

∞. On the other hand, $|\widehat{w}_i - 1| \prec \widehat{c}_w = 2 + 1/\{1 - \widehat{G}(\tau)\}$ and $P\{\exp(h\widehat{c}_w) < \exp(2hc_w) < \infty\} \to 1$ as $n \to \infty$ for any positive constant $h < \infty$, where $c_w = 2 + 1/\{1 - \widehat{G}(\tau)\}$. This implies that conditions (2) and (4) of Liang and Baek (2006) are satisfied. Thus, by Corollary 2.2 of Liang and Baek (2006), $\widehat{\mathcal{U}}_w \to 0$ almost surely and thus $\widehat{U} \to 0$ almost surely. On the other hand, by Theorem 3.1 of Liang *and others* (2004), a central limit theorem for NegA random variables (CLT_{NegA}), along with the variance calculation given in Appendix A, $N^{\frac{1}{2}}\widehat{\mathcal{U}}_w$, conditional on $\mathscr{D}$, converges in distribution to $N(0, \sigma_U^2 - \sigma_{0U}^2)$. It follows that $N^{\frac{1}{2}}\widehat{U} \to N(0, \sigma_U^2)$.

C. Asymptotic Properties of Proposed Estimators

C·1 Asymptotic Properties of $\widehat{\beta}$

We first establish the consistency of $\widehat{\beta}$. To this end, we note that $N^{-1}\widehat{\mathcal{L}}(\beta)$ is a concave function and the limiting partial likelihood function, $\mathcal{L}_0(\beta)$, is uniquely maximized at β_0. Thus, from Theorem 2.7 of Newey and McFadden (1994), the consistency of $\widehat{\beta}$ can be established if $N^{-1}\widehat{\mathcal{L}}(\beta) \to \mathcal{L}_0(\beta)$ in probability for any given β. To show this convergence, we write

$$N^{-1}\widehat{\mathcal{L}}(\beta) = N^{-1}\sum_{i=1}^{N} \widehat{w}_i \left\{\beta Y_i - \log \widehat{\Pi}^{(0)}(X_i, \beta)\right\}$$

where $\widehat{\Pi}^{(k)}(t, \beta) = N^{-1}\sum_{i=1}^{N} \widehat{w}_i I(X_i \geqslant t) e^{\beta Y_i} Y_i^k$. We next show that for any given $\beta \in \Omega$,

$$\sup_t |\widehat{\Pi}^{(k)}(t, \beta) - \Pi^{(k)}(t, \beta)| \to 0, \quad \text{in probability.} \tag{C.1}$$

where $\Pi^{(k)}(t, \beta) = E\{I(X_i \geqslant t)e^{\beta Y_i} Y_i^k\}$. To this end, we note that $|\widehat{\Pi}^{(k)}(t, \beta) - \Pi^{(k)}(t, \beta)| \leqslant |\widehat{\Pi}^{(k)}(t, \beta) - \widetilde{\Pi}^{(k)}(t, \beta)| + |\widetilde{\Pi}^{(k)}(t, \beta) - \Pi^{(k)}(t, \beta)|$, where $\widetilde{\Pi}^{(k)}(t, \beta) = N^{-1}\sum_{i=1}^{N} I(X_i \geqslant t)e^{\beta Y_i} Y_i^k$. From a uniform law of large numbers (Pollard, 1990), $\sup_t |\widetilde{\Pi}^{(k)}(t, \beta) - \Pi^{(k)}(t, \beta)| \to 0$ in probability. Thus it remains to show that $\widehat{\Pi}^{(k)}(t, \beta) - \widetilde{\Pi}^{(k)}(t, \beta) = N^{-1}\sum_{i=1}^{N}(\widehat{w}_i - 1)I(X_i \geqslant t)e^{\beta Y_i} Y_i^k \to 0$ in probability, uniformly in t. Noticing that conditional on $\mathscr{D}$, $\{(\widehat{w}_i - 1)e^{\beta Y_i} Y_i^k, i = 1, ..., N\}$ is a sequence of NegA random variables and $E(\widehat{w}_i - 1 \mid \mathscr{D}) = 0$, we may sort the sequence based on the order X_i such that $\{(\widehat{w}_{(i)} - 1)e^{\beta Y_{(i)}} Y_{(i)}^k, X_{(1)} \leqslant X_{(2)} \leqslant ... \leqslant X_{(n)}\}$ remains a sequence of mean zero NegA random variables. This sorting will allow us to use the maximum inequality for partial sums of NegA random variables given in Theorem 2 of Shao (2000) and obtain

$$E \sup_t \left| N^{-1} \sum_{i=1}^{N} I(X_i \leqslant t)(\widehat{w}_i - 1)e^{\beta Y_i} Y_i^k \right| \leqslant E \max_{1 \leqslant k \leqslant n} \left| N^{-1} \sum_{i=1}^{N} (\widehat{w}_{(i)} - 1) e^{\beta Y_{(i)}} Y_{(i)}^k \right|$$

$$\leqslant 2E \left| N^{-1} \sum_{i=1}^{N} (\widehat{w}^*_{(i)} - 1) e^{\beta Y_{(i)}} Y_{(i)}^k \right| = 2E \left| N^{-1} \sum_{i=1}^{N} (\widehat{w}^*_i - 1) e^{\beta Y_i} Y_i^k \right|$$

where $\{\widehat{w}^*_{(i)}, i = 1, ..., N\}$ is a sequence of independent random variables conditional on $\mathscr{D}$ such that $\widehat{w}^*_{(i)}$ and $\widehat{w}_i$ have the same distribution given $\mathscr{D}$. The independence among $\{\widehat{w}^*_{(i)}, i = 1, ..., N\}$ allows us to invoke the standard law of large numbers and the dominated convergence theorem to show that $E \left| N^{-1} \sum_{i=1}^{N} (\widehat{w}^*_i - 1) e^{\beta Y_i} Y_i^k \right| \to 0$. Therefore, $\sup_t |\widehat{\Pi}^{(k)}(t, \beta) - \widetilde{\Pi}^{(k)}(t, \beta)| \to 0$ in probability and hence (C.1) holds.

To derive the large sample distribution for $\widehat{\beta}$, by Taylor series expansion we have

$$N^{\frac{1}{2}}(\widehat{\beta} - \beta_0) = N^{-\frac{1}{2}} \sum_{i=1}^{N} \widehat{w}_i U_{\beta i} + o_p(1), \tag{C.2}$$

where

$$U_{\beta i} = \mathbb{A}(\beta_0)^{-1} \int \left\{ Y_i - \frac{\Pi^{(1)}(t)}{\Pi^{(0)}(t)} \right\} dM_i(t), \tag{C.3}$$

$\widehat{\mathbb{A}}(\beta) = -\frac{\partial \widehat{U}(\beta)}{\partial \beta} = N^{-1} \sum_{i=1}^{N} \widehat{w}_i \delta_i \left\{ \frac{\widehat{\Pi}^{(2)}(X_i,\beta)\widehat{\Pi}^{(0)}(X_i,\beta) - \widehat{\Pi}^{(1)}(X_i,\beta)^2}{\widehat{\Pi}^{(0)}(X_i,\beta)^2} \right\}$, $M_i(t) = \mathcal{N}_i(t) - A_i(t)$ and $A_i(t) = \int_0^t I(X_i \geqslant u) e^{\beta_0 Y_i} d\Lambda_0(u)$. Here and in the sequel, for the ease of notation, we let $\Pi^{(k)}(t) = \Pi^{(k)}(t, \beta_0)$ and $\widehat{\Pi}^{(k)}(t) = \widehat{\Pi}^{(k)}(t, \beta_0)$. We next show that $\widehat{\mathcal{W}}(t) = N^{-\frac{1}{2}} \sum_{i=1}^{N} \widehat{w}_i M_i(t)$ converges weakly to a zero-mean Gaussian process. To this end, we first note that the finite dimensional weak convergence follows directly from CLT_{NegA}. To establish the tightness of the process $\widehat{\mathcal{W}}(t)$, we write $\widehat{\mathcal{W}}(t) = \widetilde{\mathcal{W}}(t) + \widehat{\mathcal{W}}_{w1}(t) - \widehat{\mathcal{W}}_{w2}(t)$, where $\widetilde{\mathcal{W}}(t) = N^{-\frac{1}{2}} \sum_{i=1}^{N} M_i(t)$, $\widehat{\mathcal{W}}_{w1}(t) = N^{-\frac{1}{2}} \sum_{i=1}^{N} (\widehat{w}_i - 1)\mathcal{N}_i(t)$, and $\widehat{\mathcal{W}}_{w2}(t) = N^{-\frac{1}{2}} \sum_{i=1}^{N} (\widehat{w}_i - 1) A_i(t)$. The functional central limit theorem (Pollard, 1990) ensures that the process $\widetilde{\mathcal{W}}(t)$ is tight. By Theorem 8.4 of Billingsley (1962) along with the fact that $\{(\widehat{w}_{(i)} - 1)\delta_{(i)}, X_{(1)} \leqslant X_{(2)} \leqslant ... \leqslant X_{(n)}\}$ is a sequence of mean zero NegA random variables, the tightness of $\widehat{\mathcal{W}}_{w1}(t)$ holds if for any $\epsilon > 0$, there exists a constant c_0 and an integer N_0 such that for every $N \geqslant N_0$,

$$P \left\{ \max_{1 \leqslant k \leqslant n} \left| \sum_{i=1}^{k} (\widehat{w}_{(i)} - 1) \delta_{(i)} \right| \geqslant c_0 N^{\frac{1}{2}} \right\} \leqslant \epsilon c_0^{-2}.$$

The above inequality can be established with (1.10) of Theorem 3 of Shao (2000) by setting $x = c_0 N^{\frac{1}{2}}$ and $a = c_0 N^{\frac{1}{2}}/48$. The same argument can be used to show that $N^{-\frac{1}{2}} \sum_{i=1}^{N} (\widehat{w}_i - 1) I(X_i \geqslant t) e^{\beta_0 Y_i}$ is tight, which implies the tightness of $\widehat{\mathcal{W}}_{w2}(t) = \int_0^t \{N^{-\frac{1}{2}} \sum_{i=1}^{N} (\widehat{w}_i - 1) I(X_i \geqslant u) e^{\beta_0 Y_i}\} d\Lambda_0(u)$. It follows that the process $\widehat{\mathcal{W}}(t)$ is tight and thus in view of Theorem 15.1 of Billingsley (1962), we have the weak convergence of $\widehat{\mathcal{W}}(t)$ to a zero-mean Gaussian process. It then follows from a $\mathrm{CLT}_{\mathrm{NegA}}$ that $N^{\frac{1}{2}}(\widehat{\beta} - \beta_0)$ is asymptotically normal with mean 0 and variance

$$\sigma_\beta^2 = E(U_{\beta i}^2/p_i) + E\{N\mathrm{cov}(\widehat{w}_i, \widehat{w}_j \mid \mathscr{D}) U_{\beta i} U_{\beta j}\} = E(U_{\beta i}^2/p_i) - m\mathcal{R}_{U_\beta}^2$$

where for any random variable U,

$$\mathcal{R}_U^2 = \int E\{\eta_U(t, X, \delta)\}^2 \frac{d\Lambda_{\mathrm{NCC}}(t)}{\pi(t)} \tag{C.4}$$

C·2 Asymptotic Properties of $\widehat{S}(t \mid y)$

We first obtain asymptotic properties for the baseline cumulative hazard estimator $\widehat{\Lambda}_0(t) = \int_0^t \frac{d\widehat{\mathcal{N}}(t)}{\widehat{\Pi}^{(0)}(t, \widehat{\beta})}$. It follows from Corollary 2.2 of Liang and Baek (2006) and the monotonicity of $\widehat{\mathcal{N}}(t)$ that $\widehat{\mathcal{N}}(t) \to E\{\mathcal{N}_i(t)\}$ in probability. On the other hand, from the consistency of $\widehat{\beta}$ and (C.1), we have $\sup_t |\widehat{\Pi}^{(0)}(t, \widehat{\beta}) - \Pi^{(0)}(t, \beta_0)| \leqslant |\widehat{\beta} - \beta_0| + \sup_t |\widehat{\Pi}^{(0)}(t, \beta_0) - \Pi^{(0)}(t, \beta_0)| = o_p(1)$. This, together with Lemma A1 of Bilias *and others* (1997) , implies the uniform consistency of $\widehat{\Lambda}_0$.

To derive the limiting distribution of $\widehat{\mathcal{U}}_\Lambda(t) = N^{\frac{1}{2}}\{\widehat{\Lambda}_0(t) - \Lambda_0(t)\}$, we note that

$$\widehat{\mathcal{U}}_\Lambda(t) \simeq N^{-\frac{1}{2}} \sum_{i=1}^{N} \widehat{w}_i \int_0^t \frac{dM_i(u)}{\Pi^{(0)}(u)} + N^{\frac{1}{2}} \int_0^t \frac{\widehat{\Pi}^{(0)}(u, \beta_0) - \widehat{\Pi}^{(0)}(u, \widehat{\beta})}{\Pi^{(0)}(u)} d\Lambda_0(u)$$

$$\simeq N^{-\frac{1}{2}} \sum_{i=1}^{N} \widehat{w}_i U_{\Lambda i}(t)$$

where $U_{\Lambda i}(t) = \int_0^t \frac{dM_i(u) - U_{\beta i}\Pi^{(1)}(u) d\Lambda_0(u)}{\Pi^{(0)}(u)}$. It then follows from the same arguments as given in C·1 that $\widehat{\mathcal{U}}_\Lambda(t)$ is asymptotically normal with mean 0 and variance $E\{U_{\Lambda i}^2/p_i\} - m\mathcal{R}_{U_\Lambda}^2$.

We next derive the asymptotic properties of $\widehat{S}(t \mid y) = e^{-\widehat{\Lambda}_0(t) e^{\widehat{\beta} y}}$. The uniform consistency of $\widehat{\Lambda}_0(t)$ and the consistency of $\widehat{\beta}$ implies that $\sup_{t, |y| \leqslant \mathcal{Y}_0} |\log \widehat{S}(t \mid y) -$

$\log S(t \mid y)|$ is bounded by

$$\sup_t \left|\widehat{\Lambda}_0(t) - \Lambda_0(t)\right| e^{|\widehat{\beta}|\mathcal{Y}_0} + \Lambda_0(\tau) \sup_{|y| \leqslant \mathcal{Y}_0} |e^{\widehat{\beta} y} - e^{\beta_0 y}| = o_p(1).$$

Now, by a taylor series expansion and the asymptotic expansions given above, we have

$$N^{\frac{1}{2}} \left\{ \widehat{S}(t \mid y) - S(t \mid y) \right\} \simeq N^{-\frac{1}{2}} \sum_{i=1}^{N} \widehat{w}_i U_{Si}(t \mid y) \tag{C.5}$$

where

$$U_{Si}(t \mid y) = -S(t \mid y) \left\{ e^{\beta_0 y} U_{\Lambda i}(t) + \Lambda_0(t) e^{\beta_0 y} y U_{\beta i} \right\}. \tag{C.6}$$

It follows from the CLT_{NegA} that $N^{\frac{1}{2}}\{\widehat{S}(t \mid y) - S(t \mid y)\}$ is asymptotically normal with mean 0 and variance $E\{U_{Si}(t \mid y)^2/p_i\} - m\mathcal{R}^2_{U_S(t|y)}$.

C·3 Asymptotic Properties of $\widehat{\mathcal{S}}(t, c)$

To derive large sample properties for $\widehat{\mathcal{S}}(t, c)$, we write $\widehat{\mathcal{S}}(t, c) = \int_c^{\mathcal{Y}_0} \widehat{S}(t \mid y) d\widehat{F}_Y(y)$. By Corollary 2.2 of Liang and Baek (2006) and the monotonicity of $\widehat{F}_Y$, we obtain the uniform consistency of $\widehat{F}_Y(y)$ for $F_Y(y)$. This, together with the uniform consistency of $\widehat{S}(t \mid y)$ and Lemma A1 of Bilias *and others* (1997), implies the consistency of $\widehat{\mathcal{S}}(t, c)$.

To derive the asymptotic distribution for $\widehat{\mathcal{S}}(t, c)$, we first note that similar arguments as given in Appendix C·1 for the weak convergence of $\widehat{\mathcal{W}}(t)$ can be used to establish the weak convergence of

$$\widehat{\mathcal{U}}_{\mathcal{F}}(y) = N^{\frac{1}{2}}\{\widehat{F}_Y(y) - F_Y(y)\} = N^{-\frac{1}{2}} \sum_{i=1}^{N} \widehat{w}_i \{I(Y_i \leqslant y) - F_Y(y)\}$$

to a zero-mean Gaussian process. This, together with (C.5), implies that $\widehat{\mathcal{U}}_{\mathcal{S}}(t, c) = N^{\frac{1}{2}}\{\widehat{\mathcal{S}}(t, c) - \mathcal{S}(t, c)\}$ is asymptotically equivalent to

$$\int_c^{\mathcal{Y}_0} N^{\frac{1}{2}} \left\{ \widehat{S}(t \mid y) - S(t \mid y) \right\} dF_Y(y) + \int_c^{\mathcal{Y}_0} S(t \mid y) d\left[N^{\frac{1}{2}}\{\widehat{F}_Y(y) - F_Y(y)\} \right]$$

and thus $\widehat{\mathcal{U}}_{\mathcal{S}}(t, c) = N^{-\frac{1}{2}} \sum_{i=1}^{N} \widehat{w}_i U_{\mathcal{S}i}(t, c) + o_p(1)$, where

$$U_{\mathcal{S}i}(t, c) = \int_c^{\mathcal{Y}_0} U_{Si}(t \mid y) dF_Y(y) + S(t \mid Y_i) I(Y_i \geqslant c) - \mathcal{S}(t, c) \tag{C.7}$$

It follows from the CLT_{NegA} that $\widehat{\mathcal{U}}_{\mathcal{S}}(t,c)$ is asymptotically normal with mean 0 and variance $E\{U_{\mathcal{S}i}(t,c)^2/p_i\} - m\mathcal{R}^2_{U_{\mathcal{S}}(t,c)}$. The asymptotic properties of the accuracy measure estimates follow directly from the above approximations to $N^{\frac{1}{2}}\{\widehat{F}_Y(y) - F_Y(y)\}$ and $\widehat{\mathcal{U}}_{\mathcal{S}}(t,c)$ as well as applications of delta method.

C·4 Asymptotic Properties of Estimators for ROC Summary Measures

Furthermore, it is not difficult to show that the weak convergence of $\widehat{\mathcal{U}}_{\mathcal{S}}(c,t)$ and $\widehat{\mathcal{U}}_{\mathcal{F}}(c)$ holds jointly. The asymptotic distribution of the accuracy estimators follows directly from the joint distribution of $\widehat{\mathcal{U}}_{\mathcal{S}}(c,t)$ and $\widehat{\mathcal{U}}_{\mathcal{F}}(c)$. This, together with a functional delta theorem, implies the following approximations for $\widehat{\mathcal{U}}_{\text{FPR}_t}(c) = N^{\frac{1}{2}}\{\widehat{\text{FPR}}_t(c) - \text{FPR}_t(c)\}$, $\widehat{\mathcal{U}}_{\text{TPR}_t}(c) = N^{\frac{1}{2}}\{\widehat{\text{TPR}}_t(c) - \text{TPR}_t(c)\}$, $\widehat{\mathcal{U}}_{\text{PPV}_t}(c) = N^{\frac{1}{2}}\{\widehat{\text{NPV}}_t(c) - \text{NPV}_t(c)\}$, and $\widehat{\mathcal{U}}_{\text{NPV}_t}(c) = N^{\frac{1}{2}}\{\widehat{\text{PPV}}_t(c) - \text{PPV}_t(c)\}$,

$$\widehat{\mathcal{U}}_{\text{FPR}_t}(c) \simeq \frac{\widehat{\mathcal{U}}_{\mathcal{S}}(c,t) - \text{FPR}_t(c)\widehat{\mathcal{U}}_{\mathcal{S}}(\mathfrak{c}_l,t)}{\mathcal{S}(t)}, \qquad \widehat{\mathcal{U}}_{\text{TPR}_t}(c) \simeq \frac{\text{TPR}_t(c)\widehat{\mathcal{U}}_{\mathcal{S}}(\mathfrak{c}_l,t) - \widehat{\mathcal{U}}_{\mathcal{F}}(c) - \widehat{\mathcal{U}}_{\mathcal{S}}(c,t)}{1-\mathcal{S}(t)},$$

$$\widehat{\mathcal{U}}_{\text{NPV}_t}(c) \simeq \frac{\{\text{PPV}_t(c)-1\}\widehat{\mathcal{U}}_{\mathcal{F}}(c) - \widehat{\mathcal{U}}_{\mathcal{S}}(c,t)}{1-\mathcal{F}(c)}, \qquad \widehat{\mathcal{U}}_{\text{PPV}_t}(c) \simeq \frac{\widehat{\mathcal{U}}_{\mathcal{S}}(t) - \widehat{\mathcal{U}}_{\mathcal{S}}(c,t) - \text{NPV}_t(c)\widehat{\mathcal{U}}_{\mathcal{F}}(c)}{\mathcal{F}(c)}.$$

The same arguments as given above can then be used to establish the weak convergence for these processes and obtain the asymptotic variance. For example, since $\widehat{\mathcal{U}}_{\text{FPR}_t}(c) \simeq N^{-\frac{1}{2}}\sum_{i=1}^{N}\widehat{w}_i U_{\text{FPR}_t i}(c)$ with $U_{\text{FPR}_t i}(c) = \mathcal{S}(t)^{-1}\{U_{\mathcal{S}i}(c,t) - \text{FPR}_t(c)U_{\mathcal{S}i}(\mathfrak{c}_l,t)\}$, $\widehat{\mathcal{U}}_{\text{FPR}_t}(c)$ converges in distribution to $N(0,\sigma^2_{\text{FPR}_t}(c))$, where

$$\sigma^2_{\text{FPR}_t}(c) = E\{U^2_{\text{FPR}_t i}(c)/p_i\} - m\int \eta^2_{U_{\text{FPR}_t}}(c,u)d\Lambda_{\text{NCC}}(u)/\pi(u),$$

and $\eta_{U_{\text{FPR}_t}}(c,u) = E\{U_{\text{FPR}_t i}(c)I(X_i > u)(1-p_i)/p_i\}$.

To establish the weak convergence of the ROC curve estimator, we first note that the arguments above can be extended to show that the weak convergences of the two processes, $\widehat{\mathcal{U}}_{\text{FPR}_t}(c)$ and $\widehat{\mathcal{U}}_{\text{TPR}_t}(c)$, hold jointly. This, together with the stochastic equicontinuity of these processes, implies that for $u \in [u_l, u_r] \subset (0,1)$,

$$N^{\frac{1}{2}}\{\widehat{\text{ROC}}_t(u) - \text{ROC}_t(u)\} \simeq \widehat{\mathcal{U}}_{\text{TPR}_t}\{\text{FPR}_t^{-1}(u)\} - \dot{\text{ROC}}_t(u)\widehat{\mathcal{U}}_{\text{FPR}_t}\{\text{FPR}_t^{-1}(u)\},$$

$$\simeq N^{-\frac{1}{2}}\sum_{i=1}^{N}\widehat{w}_i U_{\text{ROC}_t i}\{\text{FPR}_t^{-1}(u)\},$$

where $U_{\text{ROC}_t i} = U_{\text{TPR}_t i}\{\text{FPR}_t^{-1}(u)\} - \dot{\text{ROC}}_t(u)U_{\text{FPR}_t i}\{\text{FPR}_t^{-1}(u)\}$ and $\dot{\text{ROC}}_t(u) = \partial\text{ROC}_t(u)/\partial u$. It follows that $N^{\frac{1}{2}}\{\widehat{\text{ROC}}_t(u) - \text{ROC}_t(u)\}$ converges weakly to a zero-mean Gaussian process, $N(0, \sigma^2_{\text{ROC}_t})$ in distribution, where $\sigma^2_{\text{ROC}_t}(u) = E\{U^2_{\text{ROC}_t i}/p_i\} - m \int \eta^2_{U_{\text{ROC}_t}}(u) d\Lambda_{\text{NCC}}(u)/\pi(u)$.

REFERENCES

BILIAS, YANNIS, GU, MINGGAO AND YING, ZHILIANG. (1997). Towards a general asymptotic theory for Cox model with staggered entry. *The Annals of Statistics* **25**, 662–682.

BILLINGSLEY, P. (1962). *Convergence of probability measures*. New York, NY: Wiley-Interscience.

JOAG-DEV, K. AND PROSCHAN, F. (1983). Negative association of random variables with applications. *The Annals of Statistics* **11**(1), 286–295.

LIANG, H. AND BAEK, J. (2006). Weighted sums of negatively associated random variables. *Aust. N. Z. J. Stat.* **48**(1), 21–31.

LIANG, H., ZHANG, D. AND BAEK, J. (2004). Convergence of weighted sums for dependent random variables. *J. Korean Math. Soc.* **41**(5), 883–94.

NEWEY, W. AND MCFADDEN, D. (1994). Large sample estimation and hypothesis testing. In: *Handbook of Econometrics,* Ed. D. McFadden and R. Engler, Volume IV, pp. 2113–245. Amsterdam: North Holland.

POLLARD, DAVID. (1990). *Empirical processes: theory and applications*. Institute of Mathematical Statistics.

SAMUELSEN, SVEN OVE. (1997). A pseudolikelihood approach to analysis of nested case-control studies. *Biometrika* **84**, 379–394.

SHAO, Q. (2000). A comparison theorem on moment inequalities between negatively associated and independent random variables. *J. Theo. Prob.* **13**(2), 343–55.