Introduction to Empirical Processes and Semiparametric Inference Lecture 16: The Delta Method, Continued

Michael R. Kosorok, Ph.D.

Professor and Chair of Biostatistics

Professor of Statistics and Operations Research

University of North Carolina-Chapel Hill

Integration, Continued

Recall, for given functions $A \in D[a,b]$ and $B \in BV_M[a,b]$ and domain $\mathbb{D}_M \equiv D[a,b] \times BV_M[a,b]$, the maps $\phi : \mathbb{D}_M \mapsto \mathbb{R}$ and $\psi : \mathbb{D}_M \mapsto D[a,b]$ defined by

$$\phi(A,B) = \int_{(a,b]} A(s)dB(s) \quad \text{and} \quad \psi(A,B)(t) = \int_{(a,t]} A(s)dB(s). \tag{1}$$

Also recall the following lemma which we introduced last time:

LEMMA 1. *For each fixed* $M < \infty$*, the maps* $\phi : \mathbb{D}_M \mapsto \mathbb{R}$ *and* $\psi : \mathbb{D}_M \mapsto D[a, b]$ *defined in (1)are Hadamard differentiable at each* $(A, B) \in \mathbb{D}_M$ *with* $\int_{(a,b]} |dA| < \infty$.

The derivatives are given by

$$\begin{aligned} \phi'_{A,B}(\alpha, \beta) &= \int_{(a,b]} A d\beta + \int_{(a,b]} \alpha dB, \quad \text{and} \\ \psi'_{A,B}(\alpha, \beta)(t) &= \int_{(a,t]} A d\beta + \int_{(a,t]} \alpha dB. \end{aligned}$$

Note that in the above lemma we define

$$\int_{(a,t]} A d\beta = A(t)\beta(t) - A(a)\beta(a) - \int_{(a,t]} \beta(s-) dA(s)$$

so that the integral is well defined even when β does not have bounded variation.

We now look at a statistical applications of Lemma 1 to the two-sample Wilcoxon rank sum statistic.

Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be independent samples from distributions F and G on the reals.

If $\mathbb{F}_m$ and $\mathbb{G}_n$ are the respective empirical distribution functions, the Wilcoxon rank sum statistic for comparing F and G has the form

$$T_1 = m \int_{\mathbb{R}} (m\mathbb{F}_m(x) + n\mathbb{G}_n(x)) d\mathbb{F}_m(x).$$

If we temporarily assume that F and G are continuous, then

$$\begin{aligned} T_1 &= mn \int_{\mathbb{R}} \mathbb{G}_n(x) d\mathbb{F}_m(x) + m^2 \int_{\mathbb{R}} \mathbb{F}_m(x) d\mathbb{F}_m(x) \\ &= mn \int_{\mathbb{R}} \mathbb{G}_n(x) d\mathbb{F}_m(x) + \frac{m^2 + m}{2} \\ &\equiv mnT_2 + \frac{m^2 + m}{2}, \end{aligned}$$

where T_2 is the Mann-Whitney statistic.

When F or G have atoms, the relationship between the Wilcoxon and Mann-Whitney statistics is more complex.

We will now study the asymptotic properties of the Mann-Whitney version of the rank sum statistic, T_2.

For arbitrary F and G, $T_2 = \phi(\mathbb{G}_n, \mathbb{F}_m)$, where ϕ is as defined in Lemma 1.

Note that F, G, $\mathbb{F}_m$ and $\mathbb{G}_n$ all have total variation ≤ 1.

Thus Lemma 1 applies, and we obtain that the Hadamard derivative of ϕ at $(A, B) = (G, F)$ is the map

$$\phi'_{G,F}(\alpha, \beta) = \int_{\mathbb{R}} G d\beta + \int_{\mathbb{R}} \alpha dF,$$

which is continuous and linear over $\alpha, \beta \in D[-\infty, \infty]$.

If we assume that $m/(m+n) \to \lambda \in [0, 1]$, as $m \wedge n \to \infty$, then

$$\sqrt{\frac{mn}{m+n}} \begin{pmatrix} \mathbb{G}_n - G \\ \mathbb{F}_m - F \end{pmatrix} \rightsquigarrow \begin{pmatrix} \sqrt{\lambda}\,\mathbb{B}_1(G) \\ \sqrt{1-\lambda}\,\mathbb{B}_2(F) \end{pmatrix},$$

where $\mathbb{B}_1$ and $\mathbb{B}_2$ are independent standard Brownian bridges.

Hence $\mathbb{G}_G(\cdot) \equiv \mathbb{B}_1(G(\cdot))$ and $\mathbb{G}_F(\cdot) \equiv \mathbb{B}_2(F(\cdot))$ both live in $D[-\infty, \infty]$.

Now Theorem 2.8 yields

$$\sqrt{\frac{mn}{m+n}} T_2 \rightsquigarrow \sqrt{\lambda} \int_{\mathbb{R}} G d\mathbb{G}_F + \sqrt{1-\lambda} \int_{\mathbb{R}} \mathbb{G}_F dG,$$

as $m \wedge n \to \infty$.

When $F = G$ and F is continuous, this limiting distribution is mean zero normal with variance $1/12$.

The delta method bootstrap, Theorem 12.1, is also applicable and can be used to obtain an estimate of the limiting distribution under more general hypotheses on F and G.

We now consider a second integration example which involves the Nelson-Aalen estimator under right censoring.

In the right censored survival data setting, an observation consists of the pair (X, δ), where $X = T \wedge C$ is the minimum of a failure time T and censoring time C, and $\delta = \mathbf{1}\{T \leq C\}$.

T and C are assumed to be independent.

Let F be the distribution function for T, and define the integrated baseline hazard for F to be

$$\Lambda(t) = \int_0^t \frac{dF(s)}{S(s-)},$$

where $S \equiv 1 - F$ is the survival function.

The Nelson-Aalen estimator for Λ, based on the i.i.d. sample $(X_1, \delta_1), \ldots, (X_n, \delta_n)$, is

$$\hat{\Lambda}_n(t) \equiv \int_{[0,t]} \frac{d\hat{N}_n(s)}{\hat{Y}_n(s)},$$

where

$$\hat{N}_n(t) \equiv n^{-1} \sum_{i=1}^{n} \delta_i \mathbf{1}\{X_i \leq t\}$$

and

$$\hat{Y}_n(t) \equiv n^{-1} \sum_{i=1}^{n} \mathbf{1}\{X_i \geq t\}.$$

It is easy to verify that the classes $\{\delta \mathbf{1}\{X \leq t\}, t \geq 0\}$ and $\{\mathbf{1}\{X \geq t\} : t \geq 0\}$ are both Donsker and hence that

$$\sqrt{n} \left(\begin{array}{c} \hat{N}_n - N_0 \\ \hat{Y}_n - Y_0 \end{array} \right) \rightsquigarrow \left(\begin{array}{c} \mathbb{G}_1 \\ \mathbb{G}_2 \end{array} \right), \qquad (2)$$

where

- $N_0(t) \equiv P(T \leq t, C \geq T)$,
- $Y_0(t) \equiv P(X \geq t)$,
- and $\mathbb{G}_1$ and $\mathbb{G}_2$ are tight Gaussian processes
- with respective covariances $N_0(s \wedge t) - N_0(s)N_0(t)$ and $Y_0(s \vee t) - Y_0(s)Y_0(t)$
- and with cross-covariance $\mathbf{1}\{s \geq t\} \left[N_0(s) - N_0(t-)\right] - N_0(s)Y_0(t)$.

Note that while we have already seen this survival set-up several times (eg., Sections 2.2.5 and 4.2.2), we are choosing to use slightly different notation than previously used to emphasize certain features of the underlying empirical processes.

The Nelson-Aalen estimator depends on the pair $(\hat{N}_n, \hat{Y}_n)$ through the two maps

$$(A, B) \mapsto \left(A, \frac{1}{B}\right) \mapsto \int_{[0,t]} \frac{1}{B} dA.$$

From Section 12.1.1, Lemma 1, and the chain rule (Lemma 6.19), it is easy to see that this composition map is Hadamard differentiable on a domain of the type

$$\{(A, B) : \int_{[0,\tau]} |dA(t)| \leq M, \inf_{t\in[0,\tau]} |B(t)| \geq \epsilon\}$$

for a given $M < \infty$ and $\epsilon > 0$, at every point (A, B) such that $1/B$ has bounded variation.

Note that the interval of integration we are using, $[0, \tau]$, is left-closed rather than left-open as in the definition of ψ given in (1).

However, if we pick some $\eta > 0$, then in fact integrals over $[0, t]$, for any $t > 0$, of functions which have zero variation over $(-\infty, 0)$ are unchanged if we replace the interval of integration with $(-\eta, t]$.

Thus we will still be able to utilize Lemma 1 in our current set-up.

In this case, the point (A, B) of interest is $A = N_0$ and $B = Y_0$.

Thus if t is restricted to the interval $[0, \tau]$, where τ satisfied $Y_0(\tau) > 0$, then it is easy to see that the pair $(\hat{N}_n, \hat{Y}_n)$ is contained in the given domain with probability tending to 1 as $n \to \infty$.

The derivative of the composition map is given by

$$(\alpha, \beta) \mapsto \left(\alpha, \frac{-\beta}{Y_0^2}\right) \mapsto \int_{[0,t]} \frac{d\alpha}{Y_0} - \int_{[0,t]} \frac{\beta dN_0}{Y_0^2}.$$

Thus from (2), we obtain via Theorem 2.8 that

$$\sqrt{n}(\hat{\Lambda}_n - \Lambda) \quad \rightsquigarrow \quad \int_{[0,(\cdot)]} \frac{d\mathbb{G}_1}{Y_0} - \int_{[0,(\cdot)]} \frac{\mathbb{G}_2 dN_0}{Y_0^2}. \tag{3}$$

The Gaussian process on the right side of (3) is equal to

$$\int_{[0,(\cdot)]} \frac{d\mathbb{M}}{Y_0},$$

where

$$\mathbb{M}(t) \equiv \mathbb{G}_1(t) - \int_{[0,t]} \mathbb{G}_2 d\Lambda$$

can be shown to be a Gaussian martingale with independent increments and covariance

$$\int_{[0,s\wedge t]} (1 - \Delta\Lambda) d\Lambda,$$

where $\Delta A(t) \equiv A(t) - A(t-)$ is the mass at t of a signed-measure A.

This means that the Gaussian process on the right side of (3) is also a Gaussian martingale with independent increments but with covariance

$$\int_{[0,s\wedge t]} (1-\Delta\Lambda)\frac{d\Lambda}{Y_0}.$$

A useful discussion of continuous time martingales arising in right censored survival data can be found in Fleming and Harrington (1991).

The delta method bootstrap, Theorem 12.1, is also applicable here and can be used to obtain an estimate of the limiting distribution.

However, when Λ is continuous over $[0, \tau]$, the independent increments structure implies that the limiting distribution is time-transformed Brownian motion.

More precisely, the limiting process can be expressed as $\mathbb{W}(v(t))$, where $\mathbb{W}$ is standard Brownian motion on $[0, \infty)$ and

$$v(t) \equiv \int_{(0,t]} \frac{d\Lambda}{Y_0}.$$

As discussed in Chapter 7 of Fleming and Harrington (1991), this fact can be used to compute asymptotically exact simultaneous confidence bands for Λ.

Proof of Lemma 1. For sequences $t_n \to 0$, $\alpha_n \to \alpha$, and $\beta_n \to \beta$, define $A_n \equiv A + t_n\alpha_n$ and $B_n \equiv B + t_n\beta_n$.

Since we require that $(A_n, B_n) \in \mathbb{D}_M$, we know that the total variation of B_n is bounded by M.

Consider first the derivative of ψ, and note that

$$\frac{\int_{(a,t]} A_n dB_n - \int_{(a,t]} AdB}{t_n} - \psi'_{A,B}(\alpha_n, \beta_n) \quad = \quad \int_{(a,t]} \alpha d(B_n - B) + \int_{(a,t]} (\alpha_n - \alpha) d(B_n - B). \tag{4}$$

Since it is easy to verify that the map

$$(\alpha, \beta) \mapsto \psi'_{A,B}(\alpha, \beta)$$

is continuous and linear, the desired Hadamard differentiability of ψ will follow provided the right side of (4) goes to zero.

To begin with, the second term on the right side goes to zero uniformly over $t \in (a, b]$, since both B_n and B have total variation bounded by M.

Now, for the first term on the right side of (4), fix $\epsilon > 0$.

Since α is cadlag, there exists a partition $a = t_0 < t_1 < \cdots < t_m = b$ such that α varies less than ϵ over each interval $[t_{i-1}, t_i)$, $1 \leq i \leq m$, and $m < \infty$.

Now define the function $\tilde{\alpha}$ to be equal to $\alpha(t_{i-1})$ over the interval $[t_{i-1}, t_i)$, $1 \leq i \leq m$, with $\tilde{\alpha}(b) = \alpha(b)$.

Thus

$$
\begin{aligned}
&\left\|\int_{(a,t]} \alpha d(B_n - B)\right\|_\infty \\
&\quad \le \left\|\int_{(a,t]} (\alpha - \tilde{\alpha}) d(B_n - B)\right\|_\infty + \left\|\int_{(a,t]} \tilde{\alpha} d(B_n - B)\right\|_\infty \\
&\quad \le \|\alpha - \tilde{\alpha}\|_\infty 2M + \sum_{i=1}^m |\alpha(t_{i-1})| \times |(B_n - B)(t_i) - (B_n - B)(t_{i-1})| \\
&\qquad + |\alpha(b)| \times |(B_n - B)(b)| \\
&\quad \le \epsilon 2M + (2m+1)\|B_n - B\|_\infty \|\alpha\|_\infty \\
&\quad \to \epsilon 2M,
\end{aligned}
$$

as $n \to \infty$.

Since ϵ was arbitrary, we have that the first term on the right side of (4) goes to zero, as $n \to \infty$, and the desired Hadamard differentiability of ψ follows.

Now the desired Hadamard differentiability of ϕ follows

- from the trivial but useful Lemma 2 below,
- by taking the extraction map $f : D[a, b] \mapsto \mathbb{R}$ defined by $f(x) = x(b)$,
- noting that $\phi = f(\psi)$,
- and then applying the chain rule for Hadamard derivatives (Lemma 6.19).$\Box$

LEMMA 2. *Let* T *be a set and fix* $T_0 \subset T$.

Define the extraction map $f : \ell^\infty(T) \mapsto \ell^\infty(T_0)$ *as* $f(x) = \{x(t) : t \in T_0\}$.

Then f *is Hadamard differentiable at all* $x \in \ell^\infty(T)$ *with derivative*

$$f'_x(h) = \{h(t) : t \in T_0\}.$$

Proof. Let t_n be any real sequence with $t_n \to 0$, and let $\{h_n\} \in \ell^\infty(T)$ be any sequence converging to $h \in \ell^\infty(T)$.

The desired conclusion follows after noting that

$$\begin{aligned} t_n^{-1}[f(x + t_n h_n) - f(x)] &= \{h_n(t) : t \in T_0\} \\ &\to \{h(t) : t \in T_0\}, \end{aligned}$$

as $n \to \infty$.$\square$

Product Integration

For a function $A \in D(0,b]$, let $A^c(t) \equiv A(t) - \sum_{0<s\leq t} \Delta A(s)$, where ΔA is as defined in the previous section, be the continuous part of A.

We define the product integral to be the map $A \mapsto \phi(A)$, where

$$\phi(A)(t) \equiv \prod_{0<s\leq t} (1 + dA(s)) = \exp(A^c(t)) \prod_{0<s\leq t} (1 + \Delta A(s)).$$

The first product is merely notation, but it is motivated by the mathematical definition of the product integral:

$$\phi(A)(t) = \lim_{\max_i |t_i - t_{i-1}| \to 0} \prod_i \{1 + [A(t_i) - A(t_{i-1})]\},$$

where the limit is over partitions $0 = t_0 < t_1 < \cdots t_m = t$ with maximum separation decreasing to zero.

We will also use the notation

$$\phi(A)(s,t] = \prod_{s<u\leq t} (1 + dA(u)) \equiv \frac{\phi(A)(t)}{\phi(A)(s)},$$

for all $0 \leq s < t$.

The two terms on the left are defined by the far right term.

Three alternative definitions, as solutions of two different Volterra integral equations and as a "Peano series," are given in Exercise 12.3.2.

The following lemma verifies that product integration is Hadamard differentiable:

LEMMA 3. *For fixed constants* $0 < b, M < \infty$, *the product integral map* $\phi : BV_M[0, b] \subset D[0, b] \mapsto D[0, b]$ *is Hadamard differentiable with derivative*

$$\phi'_A(\alpha)(t) = \int_{(0,t]} \phi(A)(0, u)\phi(A)(u, t]d\alpha(u).$$

When $\alpha \in D[0, b]$ *has unbounded variation, the above quantity is well-defined by integration by parts.*

From the discussion of the Nelson-Aalen estimator $\hat{\Lambda}_n$ in Section 12.2.2, it is not hard to verify that in the right-censored survival analysis setting $S(t) = \phi(-\Lambda)(t)$, where ϕ is the product integration map.

Moreover, it is easily verified that the Kaplan-Meier estimator $\hat{S}_n$ discussed in Sections 2.2.5 and 4.3 satisfies

$$\hat{S}_n(t) = \phi(-\hat{\Lambda}_n)(t).$$

We can now use Lemma 3 to derive the asymptotic limiting distribution of $\sqrt{n}(\hat{S}_n - S)$.

As in Section 12.2.2, we will restrict our time domain to $[0, \tau]$, where $P(X > \tau) > 0$.

Under these circumstances, there exists an $M < \infty$, such that $\Lambda(\tau) < M$ and $\hat{\Lambda}_n(\tau) < M$ with probability tending to 1 as $n \to \infty$.

Now Lemma 3, combined with (3) and the discussion immediately following, yields

$$\begin{aligned}\sqrt{n}(\hat{S}_n - S) &\rightsquigarrow -\int_{(0,(\cdot)]} \phi(-\Lambda)(0,u)\phi(-\Lambda)(u,t]\frac{d\mathbb{M}}{Y_0} \\ &= -S(t)\int_{(0,(\cdot)]} \frac{d\mathbb{M}}{(1-\Delta\Lambda)Y_0},\end{aligned}$$

where $\mathbb{M}$ is a Gaussian martingale with independent increments and covariance

$$\int_{(0,s\wedge t]} (1-\Delta\Lambda)d\Lambda/Y_0.$$

Thus $\sqrt{n}(\hat{S}_n - S)/S$ is asymptotically time-transformed Brownian motion $\mathbb{W}(w(t))$, where $\mathbb{W}$ is standard Brownian motion on $[0,\infty)$ and where

$$w(t) \equiv \int_{(0,t]} [(1-\Delta\Lambda)Y_0]^{-1} d\Lambda.$$

Along the lines discussed in the Nelson-Aalen example of Section 12.2.2, the form of the limiting distribution can be used to obtain asymptotically exact simultaneous confidence bands for S.

The delta method bootstrap, Theorem 12.1, can also be used for inference on S.

A key element of the proof of Lemma 3, is the following lemma which includes the important *Duhamel equation* for the difference between two product integrals:

LEMMA 4. *For* $A, B \in D(0, b]$, *we have for all* $0 \leq s < t \leq b$ *the following, where* M *is the sum of the total variation of* A *and* B*:*

(i) (the Duhamel equation)

$$(\phi(B) - \phi(A))(s, t] = \int_{(s,t]} \phi(A)(0, u)\phi(B)(u, t](B - A)(du).$$

(ii) $\|\phi(A) - \phi(B)\|_{(s,t]} \leq e^M (1 + M)^2 \|A - B\|_{(s,t]}$.

Inversion

Recall the derivation given in the paragraphs following Theorem 2.8 of the Hadamard derivative of the inverse of a distribution function F.

Note that this derivation did not depend on F being a distribution function per se.

In fact, the derivation will carry through unchanged if we replace the distribution function F with any nondecreasing, cadlag function A satisfying mild regularity conditions.

For a non-decreasing function $B \in D(-\infty, \infty)$, define the left-continuous inverse

$$r \mapsto B^{-1}(r) \equiv \inf\{x : B(x) \geq r\}.$$

We will hereafter use the notation $\tilde{D}[u, v]$ to denote all left-continuous functions with right-hand limits (caglad) on $[u, v]$ and $D_1[u, v]$ to denote the restriction of all non-decreasing functions in $D(-\infty, \infty)$ to the interval $[u, v]$.

Here is a precise statement of the general Hadamard differentiation result for non-decreasing functions:

LEMMA 5. *Let $-\infty < p \leq q < \infty$, and let the non-decreasing function $A \in D(-\infty, \infty)$ be continuously differentiable on the interval*

$$[u, v] \equiv [A^{-1}(p) - \epsilon, A^{-1}(q) + \epsilon],$$

for some $\epsilon > 0$, with derivative A' being strictly positive and bounded over $[u, v]$.

Then the inverse map $B \mapsto B^{-1}$ as a map

$$D_1[u, v] \subset D[u, v] \mapsto \tilde{D}[p, q]$$

is Hadamard differentiable at A tangentially to $C[u, v]$, with derivative

$$\alpha \mapsto -(\alpha/A') \circ A^{-1}.$$

We now restrict ourselves to the setting where A is a distribution function which we will now denote by F.

The following lemma provides two results similar to Lemma 5 but which utilize knowledge about the support of the distribution function F.

Let $D_2[u,v]$ be the subset of distribution functions in $D_1[u,v]$ with support only on $[u,\infty)$, and let $D_3[u,v]$ be the subset of distribution functions in $D_2[u,v]$ which have support only on $[u,v]$.

LEMMA 6. *Let F be a distribution function; we have the following:*

(i) Let $F \in D_2[u, \infty)$, for finite u, and let $q \in (0, 1)$. Assume F is continuously differentiable on the interval

$$[u, v] = [u, F^{-1}(q) + \epsilon],$$

for some $\epsilon > 0$, with derivative f being strictly positive and bounded over $[u, v]$.

Then the inverse map $G \mapsto G^{-1}$ as a map

$$D_2[u, v] \subset D[u, v] \mapsto \tilde{D}(0, q]$$

is Hadamard differentiable at F tangentially to $C[u, v]$.

(ii) Let $F \in D_3[u,v]$, for $[u,v]$ compact, and assume that F is continuously differentiable on $[u,v]$ with derivative f strictly positive and bounded over $[u,v]$.

Then the inverse map $G \mapsto G^{-1}$ as a map

$$D_3[u,v] \subset D[u,v] \mapsto \tilde{D}(0,1)$$

is Hadamard differentiable at F tangentially to $C[u,v]$.

In either case, the derivative is the map $\alpha \mapsto -(\alpha/f) \circ F^{-1}$.

As discussed in Section 2.2.4, an important application of these results is to estimation and inference for the quantile function $p \mapsto F^{-1}(p)$ based on the usual empirical distribution function for i.i.d. data.

Lemma 6 is useful when some information is available on the support of F, since it allows the range of p to extend as far as possible.

These results are applicable to other estimators of the distribution function F besides the usual empirical distribution, provided the standardized estimators converge to a tight limiting process over the necessary intervals.

Several examples of such estimators including

- the Kaplan-Meier estimator,
- the self-consistent estimator of Chang (1990) for doubly-censored data,
- and certain estimators from dependent data

are mentioned in Kosorok (1999).

We now apply Lemma 6 to the construction of quantile processes based on the Kaplan-Meier estimator discussed in Section 12.2.3 above.

Since it is known that the support of a survival function is on $[0, \infty)$, we can utilize Part (i) of this lemma.

Define the Kaplan-Meier quantile process

$$\{\hat{\xi}(p) \equiv \hat{F}_n^{-1}(p), 0 < p \leq q\},$$

where $\hat{F}_n = 1 - \hat{S}_n$, $\hat{S}_n$ is the Kaplan-Meier estimator, and where $0 < q < F(\tau)$ for τ as defined in the previous section.

Assume that F is continuously differentiable on $[0, \tau]$ with density f bounded below by zero and finite.

Combining the results of the previous section with Part (i) of Lemma 6 and Theorem 2.8, we obtain

$$\sqrt{n}(\hat{\xi} - \xi)(\cdot) \rightsquigarrow \frac{S(\xi(\cdot))}{f(\xi(\cdot))} \int_{(0,\xi(\cdot)]} \frac{d\mathbb{M}}{(1 - \Delta\Lambda)Y_0},$$

in $\tilde{D}(0, q]$, where $\xi(p) \equiv \xi_p$ and $\mathbb{M}$ is the Gaussian martingale described in the previous section.

Thus $\sqrt{n}(\hat{\xi} - \xi)f(\xi)/S(\xi)$ is asymptotically time-transformed Brownian motion with time-transform $w(\xi)$, where w is as defined in the previous section, over the interval $(0, q]$.

As described in Kosorok (1999), one can construct kernel estimators for f—which can be shown to be uniformly consistent—to facilitate inference.

An alternative approach is the bootstrap which can be shown to be valid in this setting based on Theorem 12.1.