Statistics 583, Problem Set 6 Solutions

Wellner; 5/12/2016

1. Let $T(F) \equiv \Lambda_F(x_0) = \int_0^{x_0} (1-F_-)^{-1} dF$ where $1 - F_-(x) = 1 - F(x-) = \lim_{y \nearrow x}(1 - F(y)) = P_F(X \ge x)$ and where x_0 is fixed. This is the cumulative hazard function of F at x_0.
(a) Is T a weakly continuous function of F (at continuity points of F)? Is it continuous with respect to the Kolmogorov (i.e. the uniform metric) on distribution functions?
(b) Find the influence function of $T(F)$. **Hint:** see van der Vaart, Lemma 20.10, page 298, and Lemma 20.14, page 300.

Solution: (a) No, T is not weakly continuous (at all continuity points of F). Suppose that $F_p(x) = (1-p)1_{[0,\infty)}(x) + p1_{[1,\infty)}(x)$ where $0 \le p \le 1$; this is the family of Bernoulli(p) distributions. Then we compute

$$\Lambda_p(x) = \int_0^x \frac{1}{1 - F_{p-}} dF_p = \begin{cases} 0, & x < 0 \\ 1 - p, & 0 \le x < 1 \\ (1-p) + \frac{p}{p} = 2 - p, & 1 \le x < \infty \end{cases}$$

Then with $F_n(x) \equiv F_{p_n}(x)$ where $p_n \to 0$ we have $F_n \to F_0 \equiv 1_{[0,\infty)}$, the distribution which has mass 1 at 0. But for any $x_0 \ge 1$, x_0 is a continuity point of F_0 and

$$\begin{aligned} T(F_n) = \Lambda_n(x_0) &\equiv \Lambda_{p_n}(x_0) = \begin{cases} 0, & x_0 < 0 \\ 1 - p_n, & 0 \le x_0 < 1 \\ 2 - p_n, & 1 \le x_0 < \infty \end{cases} \\ &\to \begin{cases} 0, & x_0 < 0 \\ 1, & 0 \le x_0 < 1 \\ 2, & 1 \le x_0 < \infty \end{cases} \\ &\ne 1_{[0,\infty)}(x_0) = \int_0^{x_0} \frac{1}{1 - F_{0-}} dF_0 = T(F_0) \end{aligned}$$

I have not yet found a counterexample if $x_0 \in \text{int}(\text{supp}(F_0))$.
(a), part 2 (Kolmogorov metric): Suppose that $\|F_n - F\|_\infty \to 0$ and assume that $1 - F(x_0-) > 0$ where x_0 is a continuity point of F. Then, via integration by

parts for the second term,

$$
\begin{aligned}
|T(F_n) - T(F)| &= \left| \int_0^{x_0} \frac{dF_n(y)}{1 - F_n(y-)} - \int_0^{x_0} \frac{dF(y)}{1 - F(y-)} \right| \\
&\le \left| \int_0^{x_0} \left(\frac{1}{1 - F_n(y-)} - \frac{1}{1 - F(y-)} \right) dF_n(y) \right| \\
&\quad + \left| \int_0^{x_0} \frac{1}{1 - F(y-)} d(F_n - F)(y) \right| \\
&= \int_0^{x_0} \left| \frac{(1 - F(y-) - (1 - F_n(y-))}{(1 - F(y-))(1 - F_n(y-))} \right| dF_n(y) \\
&\quad + \left| \int_0^{x_0} \frac{1}{1 - F(y-)} d(F_n - F)(y) \right| \\
&\le \int_0^{x_0} \frac{|F_n(y-) - F(y-)|}{(1 - F(y-))(1 - F_n(y-))} dF_n(y) \\
s &\quad + \left| \frac{(F_n(x_0) - F(x_0))}{1 - F(x_0-)} - \int_{[0,x_0]} (F_n - F)(y) d\frac{1}{1 - F(y-)} \right| \\
&\le \|F_n - F\|_\infty \left\{ \int_0^{x_0} \frac{1}{(1 - F(y-))(1 - F_n(y-))} dF_n(y) \right. \\
&\qquad \left. + \frac{2}{1 - F(x_0-)} \right\} \\
&\to 0
\end{aligned}
$$

if the limsup of the first term inside the brackets is finite since the second and third terms are clearly finite and $\|F_n - F\|_\infty \to 0$. But by monotonicity,

$$
\begin{aligned}
\int_0^{x_0} \frac{1}{(1 - F(y-))(1 - F_n(y-))} dF_n(y) &\le \frac{1}{(1 - F(x_0-))(1 - F_n(x_0-))} F_n(x_0) \\
&\to \frac{1}{(1 - F(x_0-))^2} < \infty,
\end{aligned}
$$

so we have $T(F_n) \to T(F)$ if x_0 is a continuity point of F. If x_0 is not a continuity point of F the same proof apparently works if $1 - F(x_0) > 0$.
(b) **Solution:** To find the influence function of $T(F)$, let $F_t = (1-t)F + t\delta_x$. The distribution function corresponding to $G \equiv \delta_x$ is $1_{(-\infty,y]}(x)$, $y \in R$, so the left limit is $G_-(y) = 1_{[x<y]}$, and the corresponding "at risk" function $1 - G_-(y) =$

$1_{[x \ge y]} = 1_{[y,\infty)}(x)$. We need to compute

$$\begin{aligned}
&\lim_{t \to 0} \frac{T(F_t) - T(F)}{t} = \frac{d}{dt} T(F_t)|_{t=0} \equiv IC(x; T, F) \equiv \psi_F(x) \\
&= \frac{d}{dt} \left\{ \int_0^{t_0} \frac{1}{1 - (F_t)_-} dF_t \right\} \Big|_{t=0} \\
&= \int_0^{t_0} \frac{1}{1 - F_-} d(\delta_x - F) + \int_0^{t_0} \frac{(\delta_x - F_-)}{(1 - F_-)^2} dF \\
&= \frac{1_{[0,t_0]}(x)}{1 - F_-(x)} - \int_0^{t_0} \frac{1}{1 - F_-} dF + \int_0^{t_0} \frac{1}{1 - F_-} dF - \int_0^{t_0} \frac{(1 - \delta_{x-})}{(1 - F_-)^2} dF \\
&= \frac{1_{[0,t_0]}(x)}{1 - F_-(x)} - \int_0^{t_0} \frac{1_{[x \ge y]}}{(1 - F_-(y))^2} dF(y) \\
&= \frac{1_{[0,t_0]}(x)}{1 - F_-(x)} - \int_0^{x} \frac{1_{[0,t_0]}(y)}{1 - F_-(y)} d\Lambda(y) \\
&= \begin{cases} \frac{1}{1 - F_-(x)} - \int_0^x \frac{1}{(1 - F_-)^2} dF & \text{if} \quad x \le t_0 \\ - \int_0^{t_0} \frac{1}{(1 - F_-)^2} dF & \text{if} \quad x > t_0 \,. \end{cases}
\end{aligned}$$

The next to last formula for the influence function of $\Lambda(t_0)$ is natural from a martingale perspective. When F is continuous $F_- = F$, and the influence function computed above reduces to:

$$IC(x; T, F) \quad = \quad 1_{[x \le t_0]} - \frac{F(t_0)}{1 - F(t_0)} 1_{[x > t_0]} = \frac{1_{[x \le t_0]} - F(t_0)}{1 - F(t_0)} \,.$$

Note that $E_F \psi_F(X) = 0$ and (in the case of a continuous d.f. F)

$$E_F \psi_F^2(X) = \frac{F(t_0)}{1 - F(t_0)} \,.$$

To prove asymptotic normality of $T(\mathbb{F}_n)$ (assuming that F satisfies $F(t_0) < 1$), write

$$
\begin{aligned}
\sqrt{n}(T(\mathbb{F}_n) - T(F)) &= \sqrt{n}\left\{\int_0^{t_0} \frac{1}{1-\mathbb{F}_n(s-)} d\mathbb{F}_n(s) - \int_0^{t_0} \frac{1}{1-F(s-)} dF(s)\right\} \\
&= \int_0^{t_0} \frac{1}{1-\mathbb{F}_n(s-)} d[\sqrt{n}(\mathbb{F}_n(s) - F(s))] \\
&\quad + \int_0^{t_0} \sqrt{n}\left\{\frac{1}{1-\mathbb{F}_n(s-)} - \frac{1}{1-F(s-)}\right\} dF(s) \\
&= \frac{\sqrt{n}(\mathbb{F}_n(t_0) - F(t_0))}{1-\mathbb{F}_n(t_0-)} - \int_0^{t_0} \sqrt{n}(\mathbb{F}_n(s) - F(s))^2 \frac{1}{(1-\mathbb{F}_n(s-))^2} d\mathbb{F}_n(s) \\
&\quad + \int_0^{t_0} \frac{\sqrt{n}(\mathbb{F}_n(s) - F(s))}{(1-\mathbb{F}_n(s-))(1-F(s-))} dF(s) \\
&= \frac{\sqrt{n}(\mathbb{F}_n(t_0) - F(t_0))}{1-\mathbb{F}_n(t_0-)} + o_p(1) \\
&\to_d \frac{\mathbb{U}(F(t_0))}{1-F(t_0-)} \\
&\sim N(0, \frac{F(t_0)}{1-F(t_0)}) \qquad \text{if } F \text{ is continuous}
\end{aligned}
$$

since the last two terms can be rewritten as

$$
\int_0^{t_0} \frac{\sqrt{n}(\mathbb{F}_n(s) - F(s))}{1-\mathbb{F}_n(s-)} \left\{\frac{d\mathbb{F}_n(s)}{1-\mathbb{F}_n(s-)} - \frac{dF(s)}{1-F(s-)}\right\} = o_p(1)
$$

by arguments similar to those we used to deal with the Mann-Whitney Wilcoxon statistic. Alternatively, martingale methods also work.

2. Let $T(F) \equiv \int (x-\mu(F))^3 dF(y)/\sigma^3(F)$ be the skewness functional where $\mu(F) \equiv \int x dF(x)$ and $\sigma^2(F) = \int (x-\mu(F))^2 dF(x)$.
(a) For what collection of df's F_0 is T weakly continuous at F_0? For what collection of df's F_0 is T continuous at F_0 with respect to the Kolmogorov metric?
(b) Find the influence function of $T(F)$.
Hint: First calculate the influence functions of $\mu(F)$ and $\sigma^2(F)$; then use the chain rule.
Comment: part (b) is problem 1, Wasserman, page 39; the influence function he gives for T on page 29 does not seem to be correct.

Solution: (a) Suppose that $\mathcal{F}_{M,\delta} \equiv \{F : \; E_F|X|^{3+\delta} \le M\}$ for some $0 < M < \infty$ and $\delta > 0$. Then if $\{F_n\} \subset \mathcal{F}_{M,\delta}$ satisfies $F_n \to_d F_0$, it follows from Vitali's theorem that

$$\begin{aligned}(V_3(F_n), V_2(F_n), V_1(F_n)) &\equiv \left(\int x^3 dF_n(x), \int x^2 dF_n(x), \int x dF_n(x)\right) \\ &\to \left(\int x^3 dF(x), \int x^2 dF(x), \int x dF(x)\right).\end{aligned}$$

Slightly more generally this holds for any sequence $\{F_n\}$ for which $F_n \to_d F$ and $|x|^3$ is F_n- uniformly integrable:

$$\limsup_{n\to\infty} \int_{[|x|\geq\lambda]} |x|^3 dF_n(x) \to 0 \quad \text{as} \quad \lambda \to \infty. \tag{1}$$

Since $T(F)$ is a continuous function of $(V_3(F), V_2(F), V_1(F))$ at all F such that $\sigma^2(F) = V_2(F) - V_1^2(F) > 0$, it follows that $T(F_n) \to T(F)$ for all $\{F_n\}$ satisfying $F_n \to F$ and (1) if $\sigma^2(F) > 0$.
Since $\|F_n - F\|_\infty \to 0$ implies that $F_n \to_d F$, if the uniform integrability condition (1) holds and $\sigma^2(F) > 0$, then $T(F_n) \to T(F)$.
(b) To find the influence function of $T(F) = \int (x-\mu)^3 dF(x)/\sigma(F)^3$, let $F_t \equiv (1-t)F + tG$. Then we need to compute $(d/dt)T(F_t)|_{t=0}$. But, by using the

calculations in examples 7.4.2 and 7.4.3,

$$\begin{aligned}
\frac{d}{dt}T(F_t)|_{t=0} &= \frac{d}{dt}\frac{\int (x-\mu(F_t))^3 dF_t(x)}{[\sigma^2(F_t)]^{3/2}}\Big|_{t=0} \\
&= \frac{\int (x-\mu(F_t))^3 d(G-F)(x)}{[\sigma^2(F_t)]^{3/2}}\Big|_{t=0} \\
&\quad -\frac{3}{2}\frac{\int (x-\mu(F_t))^3 dF_t(x)}{[\sigma^2(F_t)]^{5/2}}\frac{d}{dt}\sigma^2(F_t)\Big|_{t=0} \\
&\quad -3\frac{\int (x-\mu(F_t))^2 dF_t(x)}{\sigma(F_t)^3}\frac{d}{dt}\mu(F_t)\Big|_{t=0} \\
&= \int \left(\frac{x-\mu(F)}{\sigma(F)}\right)^3 d(G-F)(x) \\
&\quad -\frac{3}{2}T(F)\frac{1}{\sigma^2(F)}\left\{\int (x-\mu(F))^2 dG(x) - \sigma^2(F)\right\} \\
&\quad -3\int \left(\frac{x-\mu(F)}{\sigma(F)}\right) dG(x) \\
&= \int \left(\frac{x-\mu(F)}{\sigma(F)}\right)^3 dG(x) - T(F) \\
&\quad -\frac{3}{2}T(F)\int \left\{\left(\frac{x-\mu(F)}{\sigma(F)}\right)^2 - 1\right\} dG(x) \\
&\quad -3\int \left(\frac{x-\mu(F)}{\sigma(F)}\right) dG(x).
\end{aligned}$$

Hence by taking $G=\delta_x$ we find the influence function of $T(F)$:

$$\begin{aligned}
\dot{T}(F;\delta_x - F) &= \left(\frac{x-\mu(F)}{\sigma(F)}\right)^3 - T(F) - \frac{3}{2}T(F)\left\{\left(\frac{x-\mu(F)}{\sigma(F)}\right)^2 - 1\right\} \\
&\quad -3\left(\frac{x-\mu(F)}{\sigma(F)}\right) \\
&\equiv \psi_F(x). \qquad (2)
\end{aligned}$$

Note that this derivation does not seem to agree with the result stated on page 29 of Wasserman: the third term here does not appear in Wasserman's claimed influence function.

(b) Here is a direct calculation to see the result in (a) another way. Write

$$\begin{aligned}
&\sqrt{n}(T(\mathbb{F}_n) - T(F)) \\
&= \frac{1}{\sigma(\mathbb{F}_n)^3}\sqrt{n}\left\{\int (x-\mu(\mathbb{F}_n))^3 d\mathbb{F}_n(x) - \int (x-\mu(F))^3 dF(x)\right\} \\
&\qquad + \int (x-\mu(F))^3 dF(x)\sqrt{n}\left\{\frac{1}{\sigma(\mathbb{F}_n)^3} - \frac{1}{\sigma(F)^3}\right\} \\
&\equiv A_n + B_n.
\end{aligned}$$

To understand A_n, write

$$\begin{aligned}
(x-\mu(\mathbb{F}_n))^3 &= (x-\mu(F) - (\mu(\mathbb{F}_n) - \mu(F)))^3 \equiv (a-b)^3 \\
&= a^3 - 3a^2 b + 3ab^2 + b^3 \\
&= (x-\mu(F))^3 - 3(x-\mu(F))^2(\mu(\mathbb{F}_n) - \mu(F)) \\
&\qquad + 3(x-\mu(F))(\mu(\mathbb{F}_n)-\mu(F))^2 + (\mu(\mathbb{F}_n) - \mu(F))^3.
\end{aligned}$$

Thus we see that

$$\begin{aligned}
A_n &= \frac{1}{\sigma(F)^3}\left\{\sqrt{n}\int (x-\mu(F))^3 d(\mathbb{F}_n(x) - F(x))\right. \\
&\qquad - 3\int (x-\mu(F))^2 d\mathbb{F}_n(x)\sqrt{n}(\mu(\mathbb{F}_n) - \mu(F)) \\
&\qquad \left. + 3\int (x-\mu(F)) d\mathbb{F}_n(x)\sqrt{n}(\mu(\mathbb{F}_n)-\mu(F)))^2 + \sqrt{n}(\mu(\mathbb{F}_n)-\mu(F))^3\right\} + o_p(1) \\
&= \frac{1}{\sigma(F)^3}\left\{\sqrt{n}\int (x-\mu(F))^3 d(\mathbb{F}_n(x) - F(x))\right. \\
&\qquad \left. - 3\sigma^2(F)\sqrt{n}\int (x-\mu(F)) d(\mathbb{F}_n(x) - F(x))\right\} \\
&\qquad + o_p(1).
\end{aligned}$$

For B_n we can write, with $m_3(F) \equiv \int (x-\mu(F))^3 dF(x)$

$$\begin{aligned}
B_n &= m_3(F)\sqrt{n}\left\{\frac{1}{\sigma(\mathbb{F}_n)^3} - \frac{1}{\sigma(F)^3}\right\} \\
&= -\frac{m_3(F)}{\sigma(F)^3\sigma(\mathbb{F}_n)^3}\sqrt{n}\left\{\sigma(\mathbb{F}_n)^3 - \sigma(F)^3\right\} \\
&= -\frac{m_3(F)}{\sigma^2(F)^3}\sqrt{n}\left\{\sigma^2(\mathbb{F}_n)^{3/2} - \sigma^2(F)^{3/2}\right\} + o_p(1) \\
&= -\frac{m_3(F)}{\sigma^2(F)^3}\frac{3}{2}\sigma(F)\sqrt{n}(\sigma^2(\mathbb{F}_n) - \sigma^2(F)) + o_p(1) \\
&= -\frac{m_3(F)}{\sigma^3(F)}\frac{3}{2\sigma^2(F)}\sqrt{n}\int \{(x-\mu(F))^2 - \sigma^2(F)\} d\mathbb{F}_n(x).
\end{aligned}$$

Putting the A_n and B_n pieces together we see that we have complete agreement with the result of the influence function calculation:

$$\sqrt{n}(T(\mathbb{F}_n) - T(F)) = \sqrt{n} \int \psi_F(x) d\mathbb{F}_n(x) + o_p(1)$$

where $\psi_F(x)$ is as given in (2). It is clear (from the Central Limit Theorem) that this is asymptotically normal if $E_F X^6 < \infty$.

When I use the influence function derived here to obtain an estimator of the Standard Error of the skewness estimator for the nerve data treated in Wasserman's example 3.10, page 29, I get $\hat{se} = .163$ rather than Wasserman's estimate of .18, a slight reduction. The resulting confidence interval for the population skewness is $1.76 \pm 2(.163) = (1.434, 2.086)$.

3. Suppose that $\mathcal{F}_+$ is the class of distribution functions F on $\mathbb{R}^+$ with mean $\mu_F = E_F X < \infty$, and consider the functional $T(F)$ defined for a fixed $x_0 \in R^+$ by

$$T(F) \equiv e_F(x_0) \equiv E_F(X - x_0 | X > x_0) = \frac{\int_{x_0}^{\infty}(1 - F(t))dt}{1 - F(x_0)} .$$

This functional is the *mean residual life* functional.
(a) For what collection of df's F_0 is T weakly continuous at F_0? For what collection of df's F_0 is T continuous at F_0 with respect to the Kolmogorov metric?
(b) Find the influence function of $T(F)$. (Consider expressing $T(F)$ in terms of two simpler functionals $U(F)$ and $V(F)$ and using the chain rule.)

Solution: (a) Much as in the previous problem we can regard $T(F)$ as the ratio of two simpler functionals: $T(F) = U(F)/V(F)$ where

$$U(F) \equiv E_F\{(X - x_0)1\{X > x_0\}\} = \int_{x > x_0} (x - x_0) dF(x)$$

and $V(F) \equiv 1 - F(x_0) = \int_{x>x_0} dF(x)$. The second of these is continuous with respect to weak convergence at continuity points x_0 of F_0: $V(F_n) = 1 - F_n(x_0) \to 1 - F_0(x_0)$ if x_0 is a continuity point of F_0. On the other hand, $U(F)$ is weakly discontinuous at every F_0 if we allow arbitrary sequences F_n: taking $F_n(x) = (1 - n^{-1})F_0(x) + n^{-1}1_{[a_n,\infty)}(x)$, then

$$U(F_n) = (1 - n^{-1}V(F_0) + n^{-1}(a_n - x_0)1\{a_n > x_0\} \to V(F_0) + \infty = \infty$$

if a_n satisfies $n^{-1}a_n \to \infty$. But if we restrict to sequences $\{F_n\}$ such that $|x|-$ is F_n uniformly integrable, then we have

$$U(F_n) = \int_{x>x_0} (x - x_0) dF_n(x) \to \int_{x>x_0} (x - x_0) dF_0(x) = U(F_0),$$

and we conclude that $T(F) = U(F)/V(F)$ is weakly continuous at F_0 with $V(F_0) = 1 - F_0(x_0) > 0$ with respect to all sequences $\{F_n\}$ that satisfy $F_n \to F_0$ and for which $|x|$ is F_n- uniformly integrable.

(b) First note that with $F_t \equiv (1-t)F + tG$ we have both

$$\frac{d}{dt}(1 - F_t(x_0))\big|_{t=0} = -(G-F)(x_0)$$

and

$$\frac{d}{dt}\int_{x_0}^{\infty}(1 - F_t(y))dy\big|_{t=0} = -\int_{x_0}^{\infty}(G-F)(y)dy\,.$$

Thus by the product rule we calculate

$$\begin{aligned}\frac{d}{dt}T(F_t)\big|_{t=0} &= -\frac{\int_{x_0}^{\infty}(G-F)(y)dy}{1-F(x_0)} + \frac{\int_{x_0}^{\infty}(1-F(y))dy}{(1-F(x_0))^2}(G-F)(x_0)\\ &= e_F(x_0)\frac{(G-F)(x_0)}{1-F(x_0)} - \frac{\int_{x_0}^{\infty}(G-F)(y)dy}{1-F(x_0)}\,.\end{aligned}$$

Taking $G = \delta_x = 1_{[x,\infty)}$ yields the influence function for T at F:

$$\begin{aligned}IC(x;T,F) &= e_F(x_0)\frac{(1_{[x,\infty)}(x_0) - F(x_0))}{1-F(x_0)} - \frac{\int_{x_0}^{\infty}(1_{[x,\infty)}(y) - F(y))dy}{1-F(x_0)}\\ &= e_F(x_0)\frac{(1_{[0,x_0]}(x) - F(x_0))}{1-F(x_0)} - \frac{\int_{x_0}^{\infty}(1_{[0,y]}(x) - F(y))dy}{1-F(x_0)}\\ &= \begin{cases} e_F(x_0) - \frac{\int_{x_0}^{\infty}1-F(y)dy}{1-F(x_0)} & x \le x_0\\ \frac{-F(x_0)}{1-F(x_0)}e_F(x_0) - \frac{\int_{x_0}^{\infty}1_{[0,y]}(x_0)-F(y)dy}{1-F(x_0)} & x > x_0\end{cases}\\ &= \begin{cases} 0 & x \le x_0\\ \frac{-F(x_0)}{1-F(x_0)}e_F(x_0) - \frac{\int_{x_0}^{x}-F(y)dy}{1-F(x_0)} - \frac{\int_{x}^{\infty}1_{[0,y]}(x_0)-F(y)dy}{1-F(x_0)} & x > x_0\end{cases}\\ &= \begin{cases} 0 & x \le x_0\\ \frac{(x-x_0)-e_F(x_0)}{1-F(x_0)} & x > x_0\end{cases}\\ &= \frac{[(x-x_0) - e_F(x_0)]1_{(x_0,\infty)(x)}}{1-F(x_0)}\,.\end{aligned}$$

Note that

$$E_F[IC^2(X;T,F)] = \frac{Var(X - x_0|X > x_0)}{1-F(x_0)}\,.$$