Statistics 583, Problem Set 3 Solutions

Wellner; 4/29/2015

1. (See also van der Vaart (1998), page 303, problem 3.) For distribution functions F on R^+ and $t_0 > 0$, consider the functional $T(F) = \Lambda(t_0) \equiv \int_0^{t_0} \frac{1}{1-F_-} dF$, the *cumulative hazard function* corresponding to F at t_0.
(a) Find the influence function of $T(F)$.
(b) What does this mean about asymptotic normality of the natural estimator $T(\mathbb{F}_n)$ of $T(F)$?
(c) Can you prove asymptotic normality of $T(\mathbb{F}_n)$ directly?

Solution: (a) 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

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

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:

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

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

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

(b) Gateaux differentiability as established in (a) gives useful information about the form of the linear term, but does not yield a proof of asymptotic normality. To prove asymptotic normality we can proceed in several ways:
(i) Establish some stronger form of differentiability such as Fréchet differentiability or Hadamard differentiability with respect to some metric d_* compatible with the empirical distribution function (or empirical measure).
(ii) Recognize some special structure associated with the functional under consideration.
(iii) Proceed directly by showing that the remainder term

$$\begin{aligned} R_n &\equiv \sqrt{n}(T(\mathbb{F}_n) - T(F)) - \dot{T}(F; \sqrt{n}(\mathbb{F}_n - F)) \\ &= \sqrt{n}(T(\mathbb{F}_n) - T(F)) - \int \psi_F(x) d\{\sqrt{n}(\mathbb{F}_n - F)(x)\} \end{aligned}$$

satisfies $R_n = o_p(1)$. We will take route (iii) in (c) below. For (i) in the current case, see van der Vaart (1998), *Asymptotic Statistics*, example 20.15, page 301 (and his Lemma 20.10, page 298). With regard to (ii) in the present case, note that for $X_{(n)} < t_0$ we have

$$\begin{aligned} \sqrt{n}(T(\mathbb{F}_n) - T(F)) &= \int_0^{t_0} \frac{1}{1 - \mathbb{F}_n(t-)} d\left(\mathbb{F}_n(t) - \int_0^t (1 - \mathbb{F}_n(s-)) d\Lambda(s)\right) \\ &= \int_0^{t_0} \frac{1}{1 - \mathbb{F}_n(t-)} d\mathbb{M}_n(t) \end{aligned}$$

where $\mathbb{M}_n(t) \equiv \sqrt{n}(\mathbb{F}_n(t) - \int_0^t (1 - \mathbb{F}_n(s-)) d\Lambda(s))$ is a mean-zero martingale, and hence the integral in the last display is also a martingale (in t_0). Then asymptotic normality (even as a process in t_0) follows from martingale CLTs; see e.g. Shorack & W (1986), chapters 6 and 7 and especially (7.1.4) page 295 and Theorem 7.5.2 page 312.

(c) 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 in the third equality 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 F be a distribution function on $\mathbb{R}^2$ with finite second moments, and let $\rho(F)$ be the correlation coefficient

$$\rho(F)=\frac{Cov_F(X,Y)}{\sqrt{Var_F(X)Var_F(Y)}}.$$

Assume that $|\rho(F)|<1$.
(a) Give an example of a sequence of bivariate distributions $\{F_n\}$ satisfying $F_n \to_d F$, but $\rho(F_n)\to 1 \neq \rho(F)$.
(b) Find a collection $\mathcal{F}$ of distribution functions on $\mathbb{R}^2$ so that ρ is weakly continuous on $\mathcal{F}$.

Solution: (a) Without loss of generality we may suppose that F is a bivariate distribution function with zero means, $E_F(X)=E_F(Y)=0$. Let $F_n=(1-$

$n^{-1})F + n^{-1}\delta_{(a_n,b_n)}$ with $(a_n, b_n) \in \mathbb{R}^2$. Note that F_n has marginal distribution functions $F_{n,X} = (1-n^{-1})F_X + n^{-1}\delta_{a_n}$, $F_{n,Y} = (1-n^{-1})F_Y + n^{-1}\delta_{b_n}$ respectively where F_X and F_Y are the marginal df's of F. Thus we compute

$$\begin{aligned}
Cov_{F_n}(X,Y) &= E_{F_n}(XY) - E_{F_n}(X)E_{F_n}Y \\
&= (1-n^{-1})E(XY) + n^{-1}a_nb_n - ((1-n^{-1})E_FX + n^{-1}a_n)((1-n^{-1})E_FY + n^{-1}b_n) \\
&= (1-n^{-1})\{E_F(XY) - E_FX \cdot E_FY\} \\
&\quad + (1-n^{-1})E_FX \cdot E_FY - (1-n^{-1})^2 E_FX \cdot E_FY \\
&\quad - (1-n^{-1})E_FX \cdot n^{-1}b_n - (1-n^{-1})E_FY \cdot n^{-1}a_n - n^{-2}a_nb_n \\
&= (1-n^{-1})Cov_F(X,Y) + n^{-1}(1-n^{-1})a_nb_n \\
&\quad - n^{-1}(1-n^{-1})\{E_FX \cdot b_n + E_FY \cdot a_n\} + n^{-1}(1-n^{-1})E_FX \cdot E_FY \\
&= (1-n^{-1})Cov_F(X,Y) + n^{-1}(1-n^{-1})a_nb_n
\end{aligned}$$

since $E_FX = E_FY = 0$. Similarly,

$$\begin{aligned}
Var_{F_n}(X) &= E_{F_n}(X^2) - (E_{F_n}(X))^2 \\
&= (1-n^{-1})E_FX^2 + n^{-1}a_n^2 - ((1-n^{-1})E_FX + n^{-1}a_n)^2 \\
&= (1-n^{-1})\{E_F(X^2) - (E_FX)^2\} + n^{-1}(1-n^{-1})(E_FX)^2 \\
&\quad + n^{-1}(1-n^{-1})a_n^2 - 2n^{-1}(1-n^{-1})E_FX \cdot a_n \\
&= (1-n^{-1})Var_F(X) + n^{-1}(1-n^{-1})a_n^2 + n^{-1}(1-n^{-1})\{(E_FX)^2 - 2E_FX \cdot a_n\} \\
&= (1-n^{-1})Var_F(X) + n^{-1}(1-n^{-1})a_n^2, \qquad \text{and} \\
Var_{F_n}(Y) &= E_{F_n}(Y^2) - (E_{F_n}(Y))^2 = (1-n^{-1})Var_F(Y) + n^{-1}(1-n^{-1})b_n^2.
\end{aligned}$$

Choosing $a_n = b_n = n$ yields

$$\begin{aligned}
Cov_{F_n}(X,Y) &= n + o(n) = n(1+o(1)), \\
Var_{F_n}(X) &= n + o(n) = n(1+o(1)), \\
Var_{F_n}(Y) &= n + o(n) = n(1+o(1)).
\end{aligned}$$

Thus we find that

$$\rho(F_n) = \frac{Cov_{F_n}(X,Y)}{\sqrt{Var_{F_n}(X)Var_{F_n}(Y)}} = \frac{n(1+o(1))}{n(1+o(1))} \to 1$$

as $n \to \infty$. Thus ρ is weakly discontinuous at every F.

(b) Consider the following collection of distributions on R^2: for some $r > 2$ and $M < \infty$

$$\mathcal{F}_{r,M} \equiv \{F : \quad E_F|X|^r \le M, \; E_F|Y|^r \le M\}.$$

Then ρ is weakly-continuous on $\mathcal{F}_{r,M}$ at any F with $Var_F(X) > 0$ and $Var_F(Y) > 0$. Here is a proof: let $\{F_n\} \subset \mathcal{F}_{r,M}$ satisfy $F_n \to_d F$. Then with $(X_n, Y_n) \sim F_n$

and $(X,Y) \sim F$ we have $(X_n,Y_n) \to_d (X,Y)$, and by a Skorokhod construction there exist $(X_n^*,Y_n^*) =_d (X_n,Y_n)$ and $(X^*,Y^*) =_d (X,Y)$ defined on a common probability space and satisfying $(X_n^*,Y_n^*) \to_{a.s.} (X^*,Y^*)$. But because $\{F_n \subset \mathcal{F}_{r,M}$, X_n^2, Y_n^2, and $|X_nY_n|$ are all uniformly integrable: since $r > 2$,

$$EX_n^2 1_{[X_n^2 \geq \lambda]} \;\; \leq \;\; \frac{1}{\lambda^{r-2}} E|X_n|^r \leq \frac{M}{\lambda^{r-2}}$$

so

$$\lim_{\lambda\to\infty} \limsup_{n\to\infty} EX_n^2 1_{[X_n^2 \geq \lambda]} \;\; \leq \;\; \lim_{\lambda\to\infty} \frac{M}{\lambda^{r-2}} = 0$$

and similarly for $\{Y_n^2\}$, so the uniform integrability of $|X_nY_n|$ follows by Cauchy-Schwarz. The same holds true for the (X_n^*,Y_n^*) pairs since the uniform integrability only depends on the (marginal) distributions. Thus by Vitali's theorem it follows that

$$EX_n^s = EX_n^{*s} \to EX^{*s} = EX^s$$

and

$$EY_n^s = EY_n^{*s} \to EY^{*s} = EY^s$$

for $s = 1,2$, while Vitali also yields

$$EX_nY_n = EX_n^*Y_n^* \to EX^*Y^* = EXY\,.$$

Therefore

$$Var_{F_n}(X_n) \to Var_F(X), Var_{F_n}(Y_n) \to Var_F(Y), \tag{1}$$

and

$$Cov_{F_n}(X_n,Y_n) \to Cov_F(X,Y). \tag{2}$$

Since we have assumed that $Var_F(X) > 0$ and $Var_F(Y) > 0$, (1) and (2) yield

$$\rho(F_n) = \frac{Cov_{F_n}(X_n,Y_n)}{\sqrt{Var_{F_n}(X_n)\cdot Var_{F_n}(Y_n)}} \to \frac{Cov_F(X,Y)}{\sqrt{Var_F(X)\cdot Var_F(Y)}} = \rho(F);$$

i.e. ρ is continuous on $\mathcal{F}_{r,M}$ at any F with positive variances.

It is interesting to note that the hypothesis $\{F_n\} \subset \mathcal{F}_{r,M}$ *cannot be weakened to* $\{F_n\} \subset \mathcal{F}_{2,M}$ (and hence it can also not be weakened to the still larger class $\mathcal{F}_{2,\infty}$). Here is a counterexample. Let F be a d.f. on R^2 with $EX = 0 = EY$ and $EX^2 = 1 = EY^2$, and $\rho(F) < 1$ where $(X,Y) \sim F$. Let $M > 1$ be a big number, and consider the class

$$\mathcal{F}_{2,M} = \{F \text{ on } R^2 : \;\; E_FX^2 \leq M, E_FY^2 \leq M\}\,.$$

Let $a_n, b_n > 0$; we will specify them in terms of M shortly. Consider the sequence of d.f.'s $\{F_n\} \subset \mathcal{F}_{2,M}$ defined by

$$F_n(x,y) = (1-\frac{1}{n})F(x,y) + \frac{1}{2n}\delta_{(a_n,b_n)} + \frac{1}{2n}\delta_{(-a_n,-b_n)}.$$

Then for any bounded and continuous function $\psi : R^2 \to R$,

$$\begin{aligned}\int \psi dF_n &= (1-\frac{1}{n})\int \psi dF + \frac{1}{2n}\psi(a_n,b_n) + \frac{1}{2n}\psi(-a_n,-b_n) \\ &\to \int \psi dF,\end{aligned}$$

so $F_n \to_d F$. Furthermore, with $(X_n, Y_n) \sim F_n$,

$$EX_n = (1-1/n)EX = 0, EY_n = 0,$$

$$EX_n^2 \ = \ = (1-1/n)EX^2 + \frac{a_n^2}{n} = (1-1/n) + \frac{a_n^2}{n} = M$$

if $a_n^2 = n\{M-(1-1/n)\}$. Similarly,

$$EY_n^2 = (1-1/n) + \frac{b_n^2}{n} = M$$

if $b_n^2 = n\{M-(1-1/n)\}$. With these choices of a_n and b_n,

$$Cov(X_n,Y_n) = (1-1/n)Cov(X,Y) + \frac{a_n b_n}{n},$$

$$\begin{aligned}\rho(F_n) &= \frac{Cov(X_n,Y_n)}{Var(X_n)Var(Y_n)} \\ &= \frac{(1-1/n)Cov(X,Y) + M - (1-1/n)}{\sqrt{M^2}} \\ &\to \frac{\rho(F) + M - 1}{M} \neq \rho(F).\end{aligned}$$

Thus $\rho(F)$ is not continuous on $\mathcal{F}_{2,M}$.

3. Exercise 3.8.1, Wasserman, page 39. [Hint: the formula given by Wasserman, page 29, is not correct.] Under what additional hypotheses can we establish $\sqrt{n}(T(\mathbb{F}_n) - T(F)) \to N(0, E_F\psi_F^2(X))$? (Here my ψ_F equals Wasserman's L_F.)

Solution: (a) 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 $T(F)$:

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

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 (3). 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 $1.76 \pm 2(.163) = (1.434, 2.086)$.

4. (a) Exercise 2.7.9, Wasserman, page 24.
(b) What additional hypotheses are needed to show that $\sqrt{n}(T(\mathbb{F}_n) - T(F))$ is asymptotically normal for this particular functional $T(F)$?
Reminder: This exercise gives the same result as we derived last Fall in Stat 581.

Solution: (a) I will take a slightly different tack than Wasserman: write

$$T(F) = \frac{E_F(X - \mu_X)(Y - \mu_Y)}{\sqrt{Var_F(X) Var_F(Y)}} \equiv \rho(F) = g(T_1(F), T_2(F), T_3(F))$$

where

$$\begin{aligned} g(u, v, w) &\equiv \frac{u}{\sqrt{vw}}, \\ T_1(F) &= E_F(X - \mu_X)(Y - \mu_Y), \\ T_2(F) &= Var_{F_X}(X), \\ T_3(F) &= Var_{F_Y}(Y). \end{aligned}$$

From example 4.3 we know that

$$\dot{T}_2(F; G - F) = \int \psi_{2,F}(x) dG_X(x), \qquad \psi_{2,F}(x) = (x - \mu_X)^2 - \sigma_X^2,$$
$$\dot{T}_3(F; G - F) = \int \psi_{3,F}(y) dG_Y(y), \qquad \psi_{3,F}(y) = (y - \mu_Y)^2 - \sigma_Y^2.$$

Now

$$\begin{aligned}
\frac{d}{dt}T_1(F_t)\Big|_{t=0} &= \frac{d}{dt}\int xy dF_t(x,y)\Big|_{t=0} - \frac{d}{dt}\mu_X(F_t)\mu_Y(F_t)\Big|_{t=0} \\
&= \int\int xy d(G-F)(x,y) - \mu_Y \int (x-\mu_X) dG_X(x) - \mu_X \int (y-\mu_Y) dG_Y(y) \\
&= \int\int \psi_{1,F}(x,y) dG(x,y)
\end{aligned}$$

where

$$\psi_{1,F}(x,y) = (x-\mu_X)(y-\mu_Y) - E_F(X-\mu_X)(Y-\mu_Y).$$

Since $\nabla g(u,v,w) = (1, -u/(2v), -u/(2w))/\sqrt{vw}$, it follows by the chain rule that

$$\begin{aligned}
\dot{T}(F; G-F) &= \int\int \frac{\psi_{1,F}(x,y)}{\sigma_X\sigma_Y} dG(x,y) \\
&\quad - \frac{T(F)}{2}\int \frac{\psi_{2,F}(x)}{\sigma_X^2} dG_X(x) - \frac{T(F)}{2}\int \frac{\psi_{3,F}(x)}{\sigma_Y^2} dG_Y(y) \\
&= \int\int \psi_F(x,y) dG(x,y)
\end{aligned}$$

where

$$\begin{aligned}
\psi_F(x,y) &= \frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} - T(F) - \frac{T(F)}{2}\left\{\frac{(x-\mu_X)^2}{\sigma_X^2} - 1 + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - 1\right\} \\
&\equiv \tilde{x}\tilde{y} - \rho(F) - \frac{\rho(F)}{2}\left\{\tilde{x}^2 - 1 + \tilde{y}^2 - 1\right\}.
\end{aligned}$$

Thus I get a "centered" version of Wasserman's formula (page 21, line 2).
(b) To show that $\sqrt{n}(T(\mathbb{F}_n) - T(F)) \to_d N(0, E_F\psi_F^2(X,Y))$, it is clear from the influence calculation in (a) that we will need to assume that $E(X^4) < \infty$ and $E(Y^4) < \infty$.

Here is a neater organization of the the calculation due to Silas Bergen:

$$
\begin{aligned}
\frac{d}{dt}T(F_t)_{t=0} &= \frac{d}{dt}\int \frac{(x-\mu_{X,t})(y-\mu_{Y,t})}{\sqrt{Var_{F,t}(X)Var_{F,t}(Y)}} d\left((1-t)F+tG\right) \\
= & \int \frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} d(G-F) - \left(\int \frac{y-\mu_Y}{\sigma_X\sigma_Y} dF(x,y)\right)\dot{\mu}_X(F,G-F) \\
& - \left(\int \frac{x-\mu_X}{\sigma_X\sigma_Y} dF(x,y)\right)\dot{\mu}_Y(F,G-F) \\
& - \frac{1}{2}\left(\int \frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X^3\sigma_Y} dF(x,y)\right)\dot{\sigma}_X^2(F,G-F) \\
& - \frac{1}{2}\left(\int \frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y^3} dF(x,y)\right)\dot{\sigma}_Y^2(F,G-F) \\
= & \int \left\{\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} - T(F)\right\} dG(x,y) \\
& - 0 - 0 - \frac{1}{2\sigma_X^2}T(F)\int \{(x-\mu_X)^2-\sigma_X^2\} dG(x,y) \\
& - - \frac{1}{2\sigma_Y^2}T(F)\int \{(y-\mu_Y)^2-\sigma_Y^2\} dG(x,y)
\end{aligned}
$$

since $\int(y-\mu_Y)dF(x,y) = 0 = \int(x-\mu_X)dF(x,y)$ and using

$$
\begin{aligned}
\dot{\sigma}_Y^2(F;G-F) &= \int \{(y-\mu_Y)^2-\sigma_Y^2\} dG(x,y), \\
\dot{\sigma}_X^2(F;G-F) &= \int \{(x-\mu_X)^2-\sigma_X^2\} dG(x,y),
\end{aligned}
$$

as we calculated in class.

5. **Optional bonus problem 1:** Consider the collection $\mathcal{F}_0$ of distribution functions F on R^+ with $0 < E_F X < \infty$ and $E_F X^2 < \infty$. Let $T(F) \equiv \sigma(F)/\mu(F)$ for $F \in \mathcal{F}_0$ where $\sigma^2(F) = Var_F(X)$ and $\mu(F) = E_F(X)$. This is the *coefficient of variation of* F. Find the influence function of $T(F)$.

   Solution: Let $F_t \equiv (1-t)F + t\delta_x$. From our previous calculations we know that

$$
\begin{aligned}
& \frac{d}{dt}\sigma^2(F_t)\Big|_{t=0} = \psi_{2,F}(x) = (x-\mu_F)^2 - \sigma_F^2, \quad \text{and} \\
& \frac{d}{dt}\mu(F_t)\Big|_{t=0} = \psi_{1,F}(x) = x - \mu_F.
\end{aligned}
$$

Thus, by the chain rule,

$$
\begin{aligned}
\frac{d}{dt}T(F_t)\Big|_{t=0} &= \frac{d}{dt}\frac{[\sigma^2(F_t)]^{1/2}}{\mu(F_t)}\Big|_{t=0} \\
&= \frac{1}{2}\frac{[\sigma^2(F_t)]^{1/2}}{\mu(F_t)}\frac{d}{dt}\sigma^2(F_t)\Big|_{t=0} - \frac{[\sigma^2(F_t)]^{1/2}}{\mu(F_t)^2}\frac{d}{dt}\mu(F_t)\Big|_{t=0} \\
&= \frac{1}{2\sigma_F\mu_F}\psi_{2,F}(x) - \frac{\sigma_F}{\mu_F^2}\psi_{1,F}(x) \\
&= \frac{\sigma_F}{2\mu_F}\frac{1}{\sigma_F^2}\psi_{2,F}(x) - \frac{\sigma_F^2}{\mu_F^2}\frac{1}{\sigma_F}\psi_{1,F}(x) \\
&= T(F)\left\{\frac{1}{2\sigma_F^2}\psi_{2,F}(x) - T(F)\frac{1}{\sigma_F}\psi_{1,F}(x)\right\} \\
&\equiv T(F)\eta_F(x)
\end{aligned}
$$

where $\psi_{1,F}$ and $\psi_{2,F}$ are as given above. Thus we expect to have, if $E_F X^4 < \infty$,

$$\sqrt{n}(T(\mathbb{F}_n) - T(F)) \to_d N(0, T^2(F)E_F\eta_F^2(X))$$

where η_F involves the standardized variable $(X - \mu_F)/\sigma_F$.

6. **Optional bonus problem 2:** 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 $T(F)$.
(c) Can you prove asymptotic normality of $T(\mathbb{F}_n)$ directly?

Solution: (a) Let $\mathcal{F}_{r,M}$ denote the collection of all distribution functions F on $\mathbb{R}$ such that $E_F|X|^r \le M$ for some fixed $M < \infty$ and some $r > 1$. Then I claim that $T(F)$ is weakly continuous on $\mathcal{F}_{r,M}$ at any F with $F(x_0) < 1$ and $x_0 \in C_F$. To see this, suppose that $\{F_n\} \subset \mathcal{F}_{r,M}$ satisfies $F_n \to_d F$ where $F(x_0) < 1$ and $x_0 \in C_F$. Then clearly $1 - F_n(x_0) \to 1 - F(x_0) > 0$. Furthermore,

$$
\begin{aligned}
\int_{x_0}^\infty (1 - F_n(x))dx &= \int_{x_0}^\infty \int_{(x,\infty)} dF_n(y)dx = \int_0^\infty \int_0^\infty 1_{(x<y)}1_{(x_0,\infty)}(x)dxdF_n(y) \\
&= \int_0^\infty (y - x_0)1_{[y>x_0]}dF_n(y) = E\{(X_n - x_0)1_{[X_n>x_0]}
\end{aligned}
$$

where $X_n \sim F_n$ and $X_n \to_d X \sim F$. Now $\{X_n\}$ is uniformly integrable since $\limsup_n E|X_n|^r \le M < \infty$ with $r > 1$. Thus

$$\begin{aligned}\int_{x_0}^{\infty}(1-F_n(x))dx &= E\{(X_n - x_0)1_{[X_n > x_0]}\} \\ &\to E\{(X - x_0)1_{[X > x_0]}\} = \int_{x_0}^{\infty}(1-F(x))dx\end{aligned}$$

by Vitali's theorem. Thus

$$T(F_n) = \frac{\int_{x_0}^{\infty}(1-F_n(x))dx}{1-F_n(x_0)} \to \frac{\int_{x_0}^{\infty}(1-F(x))dx}{1-F(x_0)} = T(F).$$

Thus the claimed weak continuity of T on $\mathcal{F}_{r,M}$ at F with $x_0 \in C_F$ and $F(x_0) < 1$ holds.

Concerning continuity with respect to $d_F(F,G) \equiv \|F - G\|_\infty$, we also need to restrict so some subclass $\mathcal{F} \subset= \mathcal{F}_{r,M}$. Again, we need $F(x_0) < 1$, but now $x_0 \in C_F$ is not essential. Note that if $d_K(F_n, F) = \|F_n - F\|_\infty \to 0$, then $d_L(F_n, F) \to 0$, and hence $F_n \to_d F$. Thus it follows from the previous argument that $T(F)$ is continuous at F with respect to $d_K = \|\cdot\|_\infty$ on $\mathcal{F}_{r,M}$ for any $r > 1$ if $F(x_0) < 1$.

(b) To calculate the influence function of $T(F) = e_F(x_0)$, let F be a distribution function with $F(x_0) < 1$, and let $F_t = (1-t)F + tG$ for $0 \le t \le 1$. We need to compute

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

Taking $G = \delta_x = 1_{[x,\infty)}$, so that $1 - G(y) = 1_{[0,x)}(y)$ and $1 - G(x_0) = 1_{[0,x)}(x_0) = 1_{[x>x_0]}$, yields

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

(c) Proceeding directly we find that

$$
\begin{aligned}
&\sqrt{n}(T(\mathbb{F}_n) - T(F)) \\
&= \sqrt{n}\left\{ \frac{\int_{x_0}^{\infty}(1-\mathbb{F}_n(x))dx}{1-\mathbb{F}_n(x_0)} - \frac{\int_{x_0}^{\infty}(1-F(x))dx}{1-F(x_0)} \right\} \\
&= -\frac{\int_{x_0}^{\infty}\sqrt{n}(\mathbb{F}_n(x)-F(x))dx}{1-\mathbb{F}_n(x_0)} + \sqrt{n}\left\{ \frac{\int_{x_0}^{\infty}(1-F(x))dx}{1-\mathbb{F}_n(x_0)} - \frac{\int_{x_0}^{\infty}(1-F(x))dx}{1-F(x_0)} \right\} \\
&= -\frac{\int_{x_0}^{\infty}\sqrt{n}(\mathbb{F}_n(x)-F(x))dx}{1-F(x_0))} \cdot \frac{1-F(x_0)}{1-\mathbb{F}_n(x_0)} \\
&\qquad + \frac{\int_{x_0}^{\infty}(1-F(x))dx}{1-F(x_0)}\sqrt{n}\left\{ \frac{(1-F(x_0)) - (1-\mathbb{F}_n(x_0))}{1-\mathbb{F}_n(x_0)} \right\} \\
&= -\frac{\int_{x_0}^{\infty}\sqrt{n}(\mathbb{F}_n(x)-F(x))dx}{1-F(x_0))} \cdot \frac{1-F(x_0)}{1-\mathbb{F}_n(x_0)} \\
&\qquad + e_F(x_0)\frac{\sqrt{n}(\mathbb{F}_n(x_0)-F(x_0))}{1-F(x_0)} \cdot \frac{1-F(x_0)}{1-\mathbb{F}_n(x_0)} \\
&= \int_0^{\infty}\left\{ \frac{(x-x_0)1_{[x>x_0]}}{1-F(x_0)} - e_F(x_0)1_{[x>x_0]} \right\} d\sqrt{n}(\mathbb{F}_n(x)-F(x)) + o_p(1) \\
&\to_d N(0, E_F\psi_F^2(X))
\end{aligned}
$$

where

$$
E_F\psi_F^2(X) = \frac{Var(X-x_0|X>x_0)}{1-F(x_0)}.
$$

This convergence has been extended to hold uniformly in x_0 by Yang (1978), *Ann. Statist.* **6**, 112-116; Hall and Wellner (1979), and Csörgő, Csörgő, and Horváth (1986, 2009); see also Shorack and W (1986), *E P*, page 778.