Statistics 583, Problem Set 8 Solutions
Wellner; 5/25/2016

1. Van der Vaart (1998), problem 23.8, page 340: Suppose that $\sqrt{n}(\hat{\theta}_n - \theta) \rightarrow_d T$ and $\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n) \rightarrow_d T$ in probability given the original observations. Show that, unconditionally, $\sqrt{n}(\hat{\theta}_n - \theta, \hat{\theta}_n^* - \hat{\theta}_n) \rightarrow (S, T)$ for independent copies S and T of T. Use this to find the unconditional limit distribution of $\sqrt{n}(\hat{\theta}_n^* - \theta)$.

   Solution: Let $S_n \equiv \sqrt{n}(\hat{\theta}_n - \theta)$ and let $T_n \equiv \sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n)$. For a fixed bounded and continuous functions ψ and φ we have

$$\begin{aligned} &E\{\psi(S_n)\} \rightarrow E\{\psi(S)\}, \quad \text{and} \\ &E\{\varphi(T_n)|\underline{X}_n\} \rightarrow_p E\{\varphi(T)\} = E\{\varphi(S)\}. \end{aligned}$$

   Since S_n is a function only of the original observations $\underline{X}_n$, it follows that

$$\begin{aligned} E\{\psi(S_n)\varphi(T_n)\} &= E\{E\{\psi(S_n)\varphi(T_n)|\underline{X}_n\}\} = E\{\psi(S_n)E\{\varphi(T_n)|\underline{X}_n\}\} \\ &\rightarrow E\left\{\psi(S)E\{\varphi(T)\}\right\} = E\psi(S) \cdot E\varphi(T) \end{aligned}$$

   This implies that $(S_n, T_n) \rightarrow_d (S, T)$ where S and T are independent and $T \stackrel{d}{=} S$. Thus by the continuous mapping (or Mann-Wald) theorem

$$\begin{aligned} \sqrt{n}(\hat{\theta}_n^* - \theta) &= \sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n) + \sqrt{n}(\hat{\theta}_n - \theta) \\ &= T_n + S_n \rightarrow_d T + S. \end{aligned}$$

2. The expression for the jackknife variance estimator for the median, in the display (1) on page 11 (3rd line from the bottom) in chapter 8 was derived under the assumption $n = 2m$ and that $T(\mathbb{F}_n) = X_{(m)}$ if $n = 2m - 1$, $T(\mathbb{F}_n) = (X_{(m)} + X_{(m+1)})/2$ if $n = 2m$.
   (a) Derive the first equality in (1), page 11, using this definition of the sample median.
   (b) Derive versions of the development in (1), page 11, using $T(F) = F^{-1}(1/2)$ (strictly). Does the asymptotic result in (1) still hold? Here is some further explanation of what I mean by "strictly" here: let $T_1(\mathbb{F}_n) = X_m$ if $n = 2m - 1$, $T_1(\mathbb{F}_n) = (X_{(m)} + X_{(m+1)})/2$ if $n = 2m$. This is one common definition of the median, and this is the definition used in (a). Let $T_2(\mathbb{F}_n) = \mathbb{F}_n^{-1}(1/2)$. This is my favorite definition of the median. Note that $T_2(\mathbb{F}_n) = T_1(\mathbb{F}_n)$ if $n = 2m - 1$, but $T_2(\mathbb{F}_n) \neq T_1(\mathbb{F}_n)$ if $n = 2m$. (What is the value of $T_2(\mathbb{F}_n)$ in this case?) T_2 is the definition of the median to be considered in 2(b)!

Solution: (a). For $n = 2m$,

$$T_{n,i} = \begin{cases} X_{(m+1)} & \text{if} \quad i \le m \\ X_{(m)} & \text{if} \quad i > m \end{cases}$$

and $T_{n,\cdot} = (X_{(m)} + X_{(m+1)})/2$. Hence

$$\begin{aligned} n\widehat{\mathrm{Var}}_n &= (n-1)\left\{ m(X_{(m+1)} - \frac{1}{2}(X_{(m)} + X_{(m+1)}))^2 \right. \\ &\quad + \left. m(X_{(m)} - \frac{1}{2}(X_{(m)} + X_{(m+1)}))^2 \right\} \\ &= n(n-1)\{\frac{X_{(m+1)} - X_{(m)}}{2}\}^2 . \end{aligned} \tag{1}$$

(b). When $n = 2m$ and $T(F) = F^{-1}(1/2)$, we have $T(\mathbb{F}_n) = X_{(m)}$ and $T_{n,i}$ are exactly as in (a) above. Hence (1) continues to hold.
When $n = 2m - 1$, then $T(\mathbb{F}_n) = X_{(m)}$,

$$T_{n,i} = \begin{cases} X_{(m)} & \text{if} \quad i \le m-1 \\ X_{(m-1)} & \text{if} \quad i \ge m \end{cases} ,$$

and $T_{n,\cdot} = \{(m-1)X_{(m)} + mX_{(m-1)}\}/(2m-1)$. Therefore

$$\begin{aligned} n\widehat{\mathrm{Var}}_n &= (n-1)\left\{ (m-1)\{X_{(m)} - \frac{1}{2m-1}[(m-1)X_{(m)} + mX_{(m-1)}]\}^2 \right. \\ &\quad + \left. m\{X_{(m-1)} - \frac{1}{2m-1}[(m-1)X_{(m)} + mX_{(m-1)}]\}^2 \right\} \\ &= \frac{(n-1)^2(n+1)}{n}\left\{\frac{X_{(m)} - X_{(m-1)}}{2}\right\}^2 \\ &\to_d \frac{1}{4f^2(F^{-1}(1/2))}\left(\frac{\chi_2^2}{2}\right)^2 \end{aligned}$$

just as before.
Remark: The only case left out in (a) and (b) is that of an odd sample size, $n = 2m - 1$ in part (a). In this case,

$$T_{n,i} = \begin{cases} (X_{(m)} + X_{(m+1)})/2 & \text{if} \quad i \le m-1 \\ (X_{(m-1)} + X_{(m+1)})/2 & \text{if} \quad i = m \\ (X_{(m-1)} + X_{(m)})/2 & \text{if} \quad i \ge m+1 \end{cases} .$$

Thus

$$\begin{aligned} T_{n,\cdot} &= \frac{1}{n}\left\{ \frac{(m-1)}{2}(X_{(m)} + X_{(m+1)}) \right. \\ &\quad + \left. \frac{1}{2}(X_{(m-1)} + X_{(m+1)}) + \frac{(m-1)}{2}(X_{(m-1)} + X_{(m)}) \right\} . \end{aligned}$$

The analysis from this point proceeds not just by algebra, but by careful grouping of terms and observing which terms are negligible. I will not present a full analysis here, but will record the result:

$$\begin{aligned} n\widehat{\mathrm{Var}}_n &= \frac{(m-1)m^2}{2n^3}\left\{n(X_{(m+1)}-X_{(m-1)})\right\}^2 + o_p(1) \\ &\to_d \frac{1}{4f^2(F^{-1}(1/2))}\left(\frac{\chi_4^2}{4}\right)^2 \end{aligned}$$

since, with $g \equiv F^{-1}$,

$$n(X_{(m+1)}-X_{(m-1)}) \to_d g'(1/2)W$$

where $W =_d Y_1 + Y_2 \sim \mathrm{Gamma}(2,1)$ for independent exponential rv's Y_1, Y_2, so that $2W \sim \chi_4^2$. Thus for this definition of the sample median, it is true that $n\widehat{\mathrm{Var}}_n = O_p(1)$ for the full sequence of nonnegative integers n but it converges in distribution to one limit as $n = 2m \to \infty$ and a different limit as $n = 2m-1 \to \infty$.

3. (a) Wasserman, problem 3.8.3, page 39, modified. Show that the claimed expression for v_{boot} given in the display for this problem is incorrect and find the correct expression. Here $v_{boot} = Var_{\mathbb{F}_n}(T_n)$ where $T_n = \overline{X}_n^2$. [Hint: see Dodd and Korn, *The American Statistician* **61** (2007), 127 - 131, and especially their appendix B, pages 130-131. Apparently the formula given by Wasserman in his problem is from Shao and Tu (1995), page 10; as noted by Dodd and Korn, the expression in Shao and Tu is incorrect.]
(b) Explain how the resulting formulas relate to how you would estimate the variance of $\overline{X}_n^2$ via the delta method.

Solution: (a) This is explained quite well in the appendix of the paper by Dodd and Korn (2007).
(b) The first term of the exact finite sample variance expression

$$Var(\overline{X}^2) = \frac{4\mu^2\sigma^2}{n} + \frac{2\sigma^4}{n^2} + \frac{4\mu\mu_3}{n^2} + \frac{\mu_4 - 3\sigma^4}{n^3}$$

corresponds exactly to what we would get from the delta method: with $g(x) = x^2$ we have $g'(x) = 2x$ and hence

$$\sqrt{n}(\overline{X}_n^2 - \mu^2) \to_d g'(\mu)\sigma Z \sim N(0, 4\mu^2\sigma^2)$$

where $Z \sim N(0,1)$. Thus the delta-method estimator of $Var(\overline{X}_n^2)$ is just $4\overline{X}_n^2 S_n^2$ where S_n is the sample variance. The bootstrap estimator of variance refines this (as shown by Dodd and Korn) by correctly capturing the n^{-2} term when $\mu \neq 0$. When $\mu = 0$, then neither the (first order) delta method nor the (nonparametric) bootstrap tells the complete story.

4. (Continuation of problem 2, problem set #7). As in problem 7.2, let $T(F) = \int (F - F_0)^2 dF_0$.
(a) Find the first Gateaux derivative at $T(F)$ at $F \neq F_0$.
(b) Find the influence function ψ_F of $T(F)$ at $F \neq F_0$ and compute $E_F \psi_F^2(X)$. Is it finite for any distribution function F?
(c) Show that $\sqrt{n}(T(\mathbb{F}_n) - T(F)) \rightarrow_d N(0, A^2)$ for some $A^2 < \infty$ and find A^2.
(d) What does the limit theorem in (c) have to do with approximations of the power of the CvM statistics for testing $H : F = F_0$ versus $K : F \neq F_0$?
(e) How would you use the bootstrap to estimate A^2?

Solution: (a) Let $F_t = (1-t)F + tG$. Then

$$\begin{aligned} T(F_t) &= \int (F_t - F_0)^2 dF_0 = \int \{(F - F_0) + t(G - F)\}^2 dF_0 \\ &= \int \left\{ (F - F_0)^2 + 2t(F - F_0)(G - F) + t^2 (G - F)^2 \right\} dF_0. \end{aligned}$$

Thus the first Gateaux derivative is given by

$$\begin{aligned} \frac{d}{dt} T(F_t) \bigg|_{t=0} &\equiv \dot{T}(F; G - F) = \int 2(F - F_0)(G - F) dF_0 \\ &= 2 \int (F - F_0)(v) \left(\int 1_{[x \leq v]} d(G - F)(x) \right) dF_0(v) \\ &= \int \left(2 \int (F - F_0)(v) 1_{[x \leq v]} dF_0(v) \right) d(G - F)(x). \end{aligned}$$

(b) To find the influence function ψ_F we write

$$\begin{aligned} &\int \left(\int 2(F - F_0)(v) 1_{[x \leq v]} dF_0(v) \right) d(G - F)(x) \\ &\equiv \int \psi(x) d(G - F)(x) = \int \left(\psi(x) - \int \psi dF \right) dG(x) \equiv \int \psi_F(x) dG(x). \end{aligned}$$

where

$$\psi(x) = \int 2(F - F_0)(v) 1_{[x \leq v]} dF_0(v)$$

and where

$$\psi_F(x) \equiv \psi(x) - \int \psi(y) dF(y) = \int 2(F - F_0)(v) \{1_{[x \leq v]} - F(v)\} dF_0(v)$$

is the influence function of T at F and the point x. Before calculating $E_F\psi_F^2(X)$ we note that

$$|\psi_F(x)| \le \int |(F-F_0)(v)||1_{[x\le v]} - F(v)|dF_0(v) \le \int 1\cdot 1\, dF_0(v) = 1,$$

and hence $E_F\psi_F^2(X) < \infty$ for all F and F_0. To calculate $E_F\psi_F^2(X)$ we write

$$\begin{aligned}
&E_F\psi_F^2(X)\\
&= 4\int \left(\int (F-F_0)(u)\{1_{[x\le u]} - F(v)\}dF_0(u)\cdot \int (F-F_0)(v)\{1_{[x\le v]} - F(v)\}dF_0(v)\right) dF(x)\\
&= 4\int \left(\int\int (F-F_0)(u)(F-F_0)(v)\{1_{[x\le u]} - F(u)\}\{1_{[x\le v]} - F(v)\}dF_0(u)dF_0(v)\right) dF(x)\\
&= 4\int\int (F-F_0)(u)(F-F_0)(v)\int \{1_{[x\le u]} - F(u)\}\{1_{[x\le v]} - F(v)\}dF(x)\; dF_0(u)dF_0(v)\\
&= 4\int\int (F-F_0)(u)(F-F_0)(v)(F(u\wedge v) - F(u)F(v))dF_0(u)dF_0(v).
\end{aligned}$$

(c) To show that $\sqrt{n}(T(\mathbb{F}_n) - T(F)) \to_d N(0, A^2)$ for some A^2 we write

$$\begin{aligned}
\sqrt{n}(T(\mathbb{F}_n) - T(F)) &= \sqrt{n}\int \left\{(\mathbb{F}_n - F_0)^2 - (F-F_0)^2\right\} dF_0\\
&= \sqrt{n}\int \{(\mathbb{F}_n - F_0) - (F-F_0)\}\{(FF_n - F_0 + (F-F_0)\}dF_0\\
&= \int \sqrt{n}(\mathbb{F}_n - F)\{\mathbb{F}_n + F - 2F_0\}dF_0\\
&= \int \sqrt{n}(\mathbb{F}_n - F)2\{F - F_0\}dF_0 + \int \sqrt{n}(\mathbb{F}_n - F)2\{\mathbb{F}_n - F\}dF_0\\
&\equiv I_n + II_n.
\end{aligned}$$

Note that

$$\begin{aligned}
|II_n| &= |\int \sqrt{n}(\mathbb{F}_n - F)(\mathbb{F}_n - F)dF_0| \le \|\sqrt{n}(\mathbb{F}_n - F)\|_\infty \cdot \|\mathbb{F}_n - F\|_\infty \cdot 1\\
&= O_p(1)\cdot o_p(1) = o_p(1).
\end{aligned}$$

On the other hand,

$$I_n \overset{d}{=} 2\int \mathbb{U}_n(F)(F-F_0)dF_0 \to_d 2\int \mathbb{U}(F)(F-F_0)dF_0 \sim N(0, A^2)$$

where $A^2 \equiv A^2(F, F_0)$ is given by

$$\begin{aligned}
A^2 &= E\left(\int \mathbb{U}(F)(F - F_0)dF_0\right)^2 \\
&= E\left\{\int \mathbb{U}(F(s))(F - F_0)(s)dF_0(s) \cdot \int \mathbb{U}(F(t))(F - F_0)(t)dF_0(t)\right\} \\
&= E\left\{\int\int (F - F_0)(s)(F - F_0)(t)\mathbb{U}(F(s))\mathbb{U}(F(t))dF_0(s)dF_0(t)\right\} \\
&= \int\int (F - F_0)(s)(F - F_0)(t)E\{\mathbb{U}(F(s))\mathbb{U}(F(t))\}dF_0(s)dF_0(t) \\
&= \int\int (F - F_0)(s)(F - F_0)(t)\{F(s \wedge t) - F(s)F(t)\}dF_0(s)dF_0(t),
\end{aligned}$$

in exact agreement with our influence calculations in (b).
(d) The limit result in (c) gives us a way to approximate the power of the Cramér - von Mises test at a fixed alternative F. The power at F of the Cramér - von Mises test of $H:\ F = F_0$ versus $K:\ F \neq F_0$ is, with $P(\int_0^1 \mathbb{U}(t)^2 dt > t_\alpha) = \alpha$ and with $Z \sim N(0, 1)$,

$$\begin{aligned}
\text{Power}_n(F) &= Pr_F(nT(\mathbb{F}_n) \geq t_\alpha) = Pr_F(T(\mathbb{F}_n) \geq n^{-1}t_\alpha) \\
&= Pr_F(\sqrt{n}(T(\mathbb{F}_n) - T(F)) \geq \sqrt{n}(n^{-1}t_\alpha - T(F))) \\
&\approx P\left(AZ \geq \sqrt{n}(n^{-1}t_\alpha - T(F))\right) \\
&= P\left(Z \geq A^{-1}\sqrt{n}(n^{-1}t_\alpha - T(F))\right).
\end{aligned}$$

(e) The bootstrap estimator of the variance $Var_F(T(\mathbb{F}_n))$ is $Var_{\mathbb{F}_n}(T(\mathbb{F}_n^*))$ where $\mathbb{F}_n^*$ denotes a bootstrap sample of size n from $\mathbb{F}_n$. We would implement this by Monte-Carlo sampling from $\mathbb{F}_n$.