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 $\psi$ and $\varphi$ we have

$$E\{\psi(S_n)\} \rightarrow E\{\psi(S)\}, \quad \text{and} \quad E\{\varphi(T_n) | X_n\} \rightarrow_p E\{\varphi(T)\} = E\{\varphi(S)\}.$$ 

Since $S_n$ is a function only of the original observations $X_n$, it follows that

$$E\{\psi(S_n) \varphi(T_n)\} = E\{E\{\psi(S_n) \varphi(T_n) | X_n\}\} = E\{\psi(S_n) E\{\varphi(T_n) | X_n\}\} \rightarrow E\{\psi(S) E\{\varphi(T)\}\} = E\psi(S) \cdot E\varphi(T)$$

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

$$\sqrt{n} (\hat{\theta}_n^* - \theta) = \sqrt{n} (\hat{\theta}_n^* - \hat{\theta}_n) + \sqrt{n} (\hat{\theta}_n - \theta)$$

$$= T_n + S_n \rightarrow_d T + S.$$ 

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 \( n = 2m \),

\[
T_{n,i} = \begin{cases} 
X_{(m+1)} & \text{if } i \leq m \\
X_{(m)} & \text{if } i > m 
\end{cases}
\]

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

\[
n\hat{\text{Var}}_n = (n - 1) \left\{ m(X_{(m+1)} - \frac{1}{2}(X_{(m)} + X_{(m+1)}))^2 \\
+ m(X_{(m)} - \frac{1}{2}(X_{(m)} + X_{(m+1)}))^2 \right\} 
\]

\[
= n(n - 1) \left( \frac{X_{(m+1)} - X_{(m)}}{2} \right)^2. \tag{1}
\]

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

When \( n = 2m - 1 \), then \( T(F_n) = X_{(m)} \),

\[
T_{n,i} = \begin{cases} 
X_{(m)} & \text{if } i \leq m - 1 \\
X_{(m-1)} & \text{if } i \geq m 
\end{cases}
\]

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

\[
n\hat{\text{Var}}_n = (n - 1) \left\{ (m - 1)(X_{(m)} - \frac{1}{2m - 1}[(m - 1)X_{(m)} + mX_{(m-1)}])^2 \\
+ (m - 1)(X_{(m-1)} - \frac{1}{2m - 1}[(m - 1)X_{(m)} + mX_{(m-1)}])^2 \right\} 
\]

\[
= \frac{(n - 1)^2(n + 1)}{n} \left\{ \frac{X_{(m)} - X_{(m-1)}}{2} \right\}^2 
\rightarrow_d \frac{1}{4f^2(F^{-1}(1/2))} \left( \frac{\chi^2_2}{2} \right)^2.
\]

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,

\[
T_{n,i} = \begin{cases} 
(X_{(m)} + X_{(m+1)})/2 & \text{if } i \leq m - 1 \\
(X_{(m-1)} + X_{(m+1)})/2 & \text{if } i = m \\
(X_{(m-1)} + X_{(m)})/2 & \text{if } i \geq m + 1 
\end{cases}
\]

Thus

\[
T_n,\cdot = \frac{1}{n} \left\{ \frac{(m - 1)}{2}(X_{(m)} + X_{(m+1)}) \\
+ \frac{1}{2}(X_{(m-1)} + X_{(m+1)}) + \frac{(m - 1)}{2}(X_{(m-1)} + X_{(m)}) \right\}.
\]
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:

\[
\hat{\text{Var}}_n = \frac{(m - 1)m^2}{2n^3} \{n(X_{(m+1)} - X_{(m-1)})\}^2 + o_p(1)
\]

\[
\rightarrow_d \frac{1}{4f^2(F^{-1}(1/2))} \left( \frac{\chi^2_1}{4} \right)^2
\]

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

\[n(X_{(m+1)} - X_{(m-1)}) \rightarrow_d g'(1/2)W\]

where \( W =_d Y_1 + Y_2 \sim \text{Gamma}(2, 1) \) for independent exponential rv’s \( Y_1, Y_2 \), so that \( 2W \sim \chi^2_2 \). Thus for this definition of the sample median, it is true that \( n\hat{\text{Var}}_n = O_p(1) \) 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 \( v_{\text{boot}} \) given in the display for this problem is incorrect and find the correct expression. Here \( v_{\text{boot}} = \text{Var}_{\text{f}}(T_n) \) where \( T_n = \bar{X}^2_n \). [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 \( \bar{X}^2_n \) 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

\[
\text{Var}(\bar{X}^2) = \frac{4\mu^2\sigma^2}{n} + \frac{2\sigma^4}{n^2} + \frac{4\mu\mu_3}{n^2} + \frac{\mu_4 - 3\sigma^4}{n^3}
\]

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

\[
\sqrt{n}(\bar{X}^2_n - \mu^2) \rightarrow_d g'(\mu)\sigma Z \sim N(0, 4\mu^2\sigma^2)
\]

where \( Z \sim N(0, 1) \). Thus the delta-method estimator of \( \text{Var}(\bar{X}^2_n) \) is just \( 4\bar{X}^2_n S_n^2 \) 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 $\psi_F$ 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 $\sqrt{n}(T(F_n) - T(F)) \rightarrow_d N(0, A^2)$ 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

$$T(F_t) = \int (F_t - F_0)^2 dF_0 = \int \{(F - F_0) + t(G - F)\}^2 dF_0$$

$$= \int \{(F - F_0)^2 + 2t(F - F_0)(G - F) + t^2(G - F)^2\} dF_0.$$

Thus the first Gateaux derivative is given by

$$\frac{dT(F_t)}{dt}\bigg|_{t=0} = \dot{T}(F; G - F) = \int 2(F - F_0)(G - F) dF_0$$

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

$$= \int \left( 2 \int (F - F_0)(v) 1_{[x \leq v]} dF_0(v) \right) d(G - F)(x).$$

(b) To find the influence function $\psi_F$ we write

$$\int \left( 2(F - F_0)(v) 1_{[x \leq v]} dF_0(v) \right) d(G - F)(x)$$

$$\equiv \int \psi(x) d(G - F)(x) = \int \left( \psi(x) - \int \psi dF \right) dG(x) \equiv \int \psi_F(x) dG(x).$$

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 $E_F\psi^2_F(X)$ we note that

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

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

$$E_F\psi^2_F(X) = 4 \int \left( \int (F - F_0)(u)\{1_{[x \leq u]} - F(u)\}dF_0(u) \cdot \int (F - F_0)(v)\{1_{[x \leq v]} - F(v)\}dF_0(v) \right) dF(x)$$

$$= 4 \int \left( \int \int (F - F_0)(u)(F - F_0)(v)\{1_{[x \leq u]} - F(u)\}\{1_{[x \leq v]} - F(v)\}dF_0(u)dF_0(v) \right) dF(x)$$

$$= 4 \int \int (F - F_0)(u)(F - F_0)(v)\{1_{[x \leq u]} - F(u)\}\{1_{[x \leq v]} - F(v)\}dF(x) dF_0(u)dF_0(v)$$

$$= 4 \int \int (F - F_0)(u)(F - F_0)(v)(F(u \wedge v) - F(u)F(v))dF_0(u)dF_0(v).$$

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

$$\sqrt{n}(T(\mathbb{F}_n) - T(F)) = \sqrt{n} \int \{(\mathbb{F}_n - F_0)^2 - (F - F_0)^2\} dF_0$$

$$= \sqrt{n} \int \{(\mathbb{F}_n - F_0) - (F - F_0)\}\{(FF_n - F_0 + (F - F_0)\}dF_0$$

$$= \int \sqrt{n}(\mathbb{F}_n - F)\{\mathbb{F}_n + F - 2F_0\}dF_0$$

$$= \int \sqrt{n}(\mathbb{F}_n - F)2\{F - F_0\}dF_0 + \int \sqrt{n}(\mathbb{F}_n - F)2\{\mathbb{F}_n - F\}dF_0$$

$$\equiv I_n + II_n.$$

Note that

$$|II_n| = |\int \sqrt{n}(\mathbb{F}_n - F)(\mathbb{F}_n - F)dF_0| \leq \|\sqrt{n}(\mathbb{F}_n - F)\|_\infty \cdot \|\mathbb{F}_n - F\|_\infty \cdot 1$$

$$= o_p(1) \cdot o_p(1) = o_p(1).$$

On the other hand,

$$I_n \overset{d}{=} 2 \int \mathbb{U}_n(F)(F - F_0)dF_0 \xrightarrow{d} 2 \int \mathbb{U}(F)(F - F_0)dF_0 \sim N(0, A^2)$$
where $A^2 \equiv A^2(F, F_0)$ is given by

$$A^2 = E \left( \int \mathbb{U}(F)(F - F_0) dF_0 \right)^2$$

$$= E \left\{ \int \mathbb{U}(s)(F - F_0)(s)dF_0(s) \cdot \int \mathbb{U}(t)(F - F_0)(t)dF_0(t) \right\}$$

$$= E \left\{ \int \int (F - F_0)(s)(F - F_0)(t)\mathbb{U}(s)\mathbb{U}(t)dF_0(s)dF_0(t) \right\}$$

$$= \int \int (F - F_0)(s)(F - F_0)(t)E\{\mathbb{U}(s)\mathbb{U}(t)\}dF_0(s)dF_0(t)$$

$$= \int \int (F - F_0)(s)(F - F_0)(t)\{F(s \wedge t) - F(s)F(t)\}dF_0(s)dF_0(t),$$

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 $Z \sim N(0, 1)$,

$$\text{Power}_n(F) = P_{T,F}(nT(F,n) \geq t_\alpha) = P_{\mathbb{F}_n}(T(F,n) \geq n^{-1}t_\alpha)$$

$$= P_{\mathbb{F}_n}(\sqrt{n}(T(F,n) - T(F)) \geq \sqrt{n}(n^{-1}t_\alpha - T(F)))$$

$$\approx P(AZ \geq \sqrt{n}(n^{-1}t_\alpha - T(F)))$$

$$= P(Z \geq A^{-1}\sqrt{n}(n^{-1}t_\alpha - T(F))).$$

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

6