Statistics 583, Final Exam Solutions

Wellner; 6/6/2007

1. (40 points) **Define** the following terms.
(a) A *continuous functional* $T(F)$ with respect to a metric d on distribution functions.
(b) A *Gateaux - differentiable* functional $T(F)$.
(c) The *influence function* corresponding to a Gateaux differentiable functional $T(F)$.
(d) The kernel estimator of a density function f on $\mathbb{R}$.

Solution: See Class notes, chapters 6, 7, 8, 14.

2. (30 points). Give a complete *statement* of **two** of the following results or theorems:
(a) An example of a functional $T(F)$ for which the jack-knife *does not* "work."
(b) A limit theorem for the bootstrap of the sample mean $\overline{X}_n$.
(c) The Politis-Romano without replacement bootstrap limit theorem.
(d) Any theorem about asymptotic normality of an estimator via differentiability of the corresponding statistical functional.

Solution: See Class notes, chapters 6, 7, 8, 14.

3. (40 points). (a) Suppose that $Z_1, \ldots, Z_k$ are i.i.d. with distribution function F_0. Let $V_k \equiv \min_{1 \le j \le k} Z_j$, and suppose that K is a random integer with truncated Poisson distribution

$$P(K = k) = \frac{e^{-\theta}}{1 - e^{-\theta}} \frac{\theta^k}{k!}, \qquad k = 1, 2, \ldots, \quad \theta > 0.$$

Show that V_K has survival function given by

$$P(V_K > v) = \frac{\exp(-\theta F_0(v)) - \exp(-\theta)}{1 - \exp(-\theta)},$$

and hence

$$P(V_K \le v) = \frac{1 - e^{-\theta F_0(v)}}{1 - e^{-\theta}} \equiv \psi_\theta(F_0(v)).$$

(Note that this is valid distribution function for $\theta \in \mathbb{R}$, not just $\theta > 0$.)
(b) Now suppose that $X_1, \ldots, X_m$ are i.i.d. F and $Y_1, \ldots, Y_n$ are i.i.d. with

distribution function $G = \psi_\theta(F)$. Find the locally most powerful rank test of $\theta = 0$ (i.e. $F = G$) versus $\theta > 0$.
(c) How would you implement the test you found in (b) to achieve a type one error $\alpha = .01$?

Solution: (a) First note that

$$\begin{aligned} P(V_k > v) &= P(\min_{1 \le j \le k} Z_j > v) = P(Z_1 > v, \ldots, Z_k > v) = P(Z_1 > v)^k \\ &= (1 - F_0(v))^k. \end{aligned}$$

Therefore,

$$\begin{aligned} P(V_K > v) &= \sum_{k=1}^{\infty} P(V_k > v,\ K = k) = \sum_{k=1}^{\infty} P(V_k > v) P(K = k) \\ &= \sum_{k=1}^{\infty} (1 - F_0(v))^k \frac{e^{-\theta}}{1 - e^{-\theta}} \frac{\theta^k}{k!} \\ &= \frac{e^{-\theta}}{1 - e^{-\theta}} \sum_{k=1}^{\infty} \frac{(1 - F_0(v))^k \theta^k}{k!} \\ &= \frac{e^{-\theta}}{1 - e^{-\theta}} \left(e^{\theta(1 - F_0(v))} - 1 \right) \\ &= \frac{\exp(-\theta F_0(v)) - e^{-\theta}}{1 - e^{-\theta}} \\ &= 1 - \frac{1 - \exp(-\theta F_0(v))}{1 - \exp(-\theta)}, \end{aligned}$$

and hence

$$P(V_K \le v) = \frac{1 - \exp(-\theta F_0(v))}{1 - \exp(-\theta)}.$$

(b) If $X_1, \ldots, X_m$ are i.i.d. F and $Y_1, \ldots, Y_n$ are i.i.d. $G = \psi_\theta(F)$, then by Hoeffding's formula

$$P_\theta(Q = q) = \frac{1}{\binom{N}{n}} E \prod_{j=1}^{n} \psi_\theta'(U_{(q_j)})$$

where the expectation is with respect to the distribution of the Uniform order statistics $0 < U_{(1)} < \cdots < U_{(N)} < 1$ and where

$$\begin{aligned}
\psi_\theta(u) &= \frac{1-\exp(-\theta u)}{1-\exp(-\theta)}, \qquad \text{so} \\
\psi'_\theta(u) &= \frac{\theta\exp(-\theta u)}{1-\exp(-\theta)}, \qquad \text{and, moreover} \\
\frac{\partial}{\partial\theta}\psi'_\theta(u) &= \frac{e^{-\theta u}}{1-\exp^{-\theta}} - \frac{u\theta e^{-\theta}}{1-\exp(-\theta)} - \frac{\theta\exp(-\theta u)\exp(-\theta)}{(1-\exp(-\theta)^2} \\
&\to \frac{1}{2} - u \qquad \text{as} \;\; \theta \to 0.
\end{aligned}$$

To find the locally most powerful rank test of $\theta > 0$ versus $\theta = 0$, we want to find a test function ϕ which maximizes the slope at $\theta = 0$ of the power $\beta_\phi(\theta) = E_\theta\phi(Q)$. But since this power is given by

$$\beta_\phi(\theta) = \sum \phi(\underline{q}) P_\theta(\underline{Q} = \underline{q}),$$

it suffices to find those q's for which

$$\frac{\partial}{\partial\theta} P_\theta(Q = q)\Big|_{\theta=0} \qquad \text{is large.}$$

Thus we calculate

$$\begin{aligned}
\frac{\partial}{\partial\theta} P_\theta(Q = q)\Big|_{\theta=0} &= \frac{\partial}{\partial\theta}\frac{1}{\binom{N}{n}} E \prod_{j=1}^{n} \psi'_\theta(U_{(q_j)})\Big|_{\theta=0} \\
&= \frac{1}{\binom{N}{n}} E \sum_{j=1}^{n} \frac{\partial}{\partial\theta}\psi'_\theta(U_{(q_j)})\Big|_{\theta=0} \\
&= \frac{1}{\binom{N}{n}} E \sum_{j=1}^{n} \left(\frac{1}{2} - U_{(q_j)}\right) \\
&= \frac{1}{\binom{N}{n}} \sum_{j=1}^{n} \left(\frac{1}{2} - \frac{q_j}{n+1}\right).
\end{aligned}$$

Thus the most powerful rank test of $\theta = 0$ versus $\theta > 0$ rejects for small values of $\sum_{j=1}^{n} Q_j$. Note that this makes good sense since for $\theta > 0$, the alternatives $G(x) = \psi_\theta(F(x))$ are stochastically smaller than F.

(c) To implement this test we could use the WWNH finite-sampling central limit theorem: Under the null hypothesis $\overline{Y}_n \equiv n^{-1}\sum_{j=1}^{n} Q_j$ has the same distribution as the sample mean of a sample of size n drawn without replacement from the finite population $\{1, \ldots, N\}$, and this population has

$$\overline{z}_N = (N+1)/2,$$
$$\sigma_z^2 = N^{-1}\sum_{i=1}^{N} z_i^2 - \overline{z}_N^2 = (N+1)(2N+1)/6 - (N+1)^2/4 = \frac{(N+1)(N-1)}{12}$$

so that

$$Var(\overline{Y}_n) = \frac{\sigma_z^2}{n}\left(1 - \frac{n-1}{N-1}\right) = \frac{(N+1)(N-1)}{12n}\frac{N-n}{N-1}.$$

Thus the WWNH finite -sampling CLT says that

$$\frac{\overline{Y}_n - (N+1)/2}{\sqrt{\frac{(N+1)(N-1)}{12n}\frac{N-n}{N-1}}} \rightarrow_d N(0,1),$$

and hence a reasonable way to implement the test found in (b) for moderately large m, n is to reject $\theta > 0$ if

$$\sum_{j=1}^{n} Q_j < n(N+1)/2 - z_{.01}\sqrt{n}\sqrt{\frac{(N+1)(N-1)}{12}\frac{N-n}{N-1}}.$$

Note that $P(Z > z_{.01}) = .01$ so $z_{.01} = ...$ is positive!

4. (40 points). Let $X_1, \ldots, X_n$ be i.i.d. with unknown density function f. Consider the kernel estimator

$$\widehat{f}_n(x) = \int \frac{1}{h_n} K\left(\frac{x-y}{h_n}\right) d\mathbb{F}_n(y)$$

of an unknown density f at a point x with the "box-car" or uniform kernel $K(x) = 2^{-1}1_{[-1,1]}(x)$. Assume that f' exists at x and is continuous in a neighborhood of x.

(a) Compute $E\widehat{f}_n(x)$ at x explicitly in terms of F and h after taking advantage of the given kernel K.

(b) Use the result of (a) and Taylor expansion to give an expression for $\text{bias}_n(x) = E\widehat{f}_n(x) - f(x)$ in terms of $f'(x)$ assuming that $h_n \rightarrow 0$.

(c) Give a formula for $Var(\widehat{f}_n(x))$ in terms of K, F, and h_n, and then use it to give an asymptotic expression for the variance assuming that $h_n \rightarrow 0$ and $nh_n \rightarrow \infty$.

(d) What is the optimal choice of $h_n = Cn^{-r}$ in this case?

(e) For the optimal choice of r in (d), sketch how you would prove that

$$\sqrt{nh_n}(\widehat{f}_n(x) - f(x)) \rightarrow_d N(b(x), \sigma^2(x))$$

and identify the functions $b(x)$ and $\sigma^2(x)$ that will appear here, including their dependence on C.
(f) For the limit distribution identified in (e), find the optimal choice of C (as a function of $f(x)$ and $f'(x)$ and other constants). Explain briefly why this optimal C makes some intuitive sense (as a function of x).

Solution: I had this problem goofed up, as you will see below. I will first try to solve the problem in the form stated, and then will solve a modified version of the problem which has as its solution something more along the lines I had intended. (My apologies on this one!)

(a) Since

$$\widehat{f}_n(x) = \int \frac{1}{h_n} K\left(\frac{x-y}{h_n}\right) d\mathbb{F}_n(y),$$

we compute

$$\begin{aligned}
E(\widehat{f}_n(x)) &= \int \frac{1}{h} K\left(\frac{x-y}{h_n}\right) dF(y) \\
&= \int K(z) f(x-hz)dz = \frac{1}{2}\int_{-1}^{1} f(x-hz)dz = -\frac{1}{2h} F(x-hz)\Big|_{-1}^{1} \\
&= \frac{1}{2h}(F(x+h) - F(x-h)) \\
&= \frac{1}{2h}\left(F(x) + f(x)h + \frac{1}{2}f'(x^*)h^2 - (F(x) - f(x)h + \frac{1}{2}f'(x^{**})h^2)\right) \\
&\qquad \text{for some } \ x^* \in [x, x+h], \quad x^{**} \in [x-h, x] \\
&= f(x) + \frac{h}{4}(f'(x^*) - f'(x^{**})).
\end{aligned}$$

(b) It follows from the expectation computed in (a) that the bias of $\widehat{f}_n(x)$ is given by

$$\text{bias}(\widehat{f}_n(x)) = \frac{h}{4}(f'(x^*) - f'(x^{**}) = \frac{h}{4}o(1)$$

since $f'(x^*) \to f'(x)$ and $f'(x^{**}) \to f'(x)$. (I made a sign error here in my initial calculations, and had thought I was getting $(h/2)f'(x) + o(h)$ for the bias.)

(c) Now for the variance:

$$
\begin{aligned}
Var(\widehat{f}_n(x)) &= Var\left(\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)\right) \\
&= \frac{1}{nh^2} Var\left(K\left(\frac{x-X_i}{h}\right)\right) \\
&= \frac{1}{nh^2}\left\{\int K^2\left(\frac{x-y}{h}\right) dF(y) - \left(\int K\left(\frac{x-y}{h}\right) dF(y)\right)^2\right\} \\
&= \frac{1}{nh}\int K^2(z) f(x-hz)dz - \left(\frac{1}{n}(f(x)+o(1)h)^2\right) \\
&= \frac{f(x)}{nh}\int K^2(z)dz(1+o(1)) + O(n^{-1}).
\end{aligned}
$$

(d) Thus the bias term is of order $ho(1) = o(h)$, so the squared bias is of order $o(h^2)$, while the variance term is of order $O((nh_n)^{-1})$. If $h = h_n = Cn^{-1/3}$, then the squared bias is of order $o(n^{-2/3})$ while the variance is of order $O(n^{-2/3})$.
(e) We can write

$$
\begin{aligned}
\sqrt{nh_n}(\widehat{f}_n(x) - f(x)) &= \sqrt{nh_n}(\widehat{f}_n(x) - E\widehat{f}_n(x)) + \sqrt{nh_n}(E\widehat{f}_n(x) - f(x)) \\
&= \text{stochastic term} \quad + \quad \text{bias term},
\end{aligned}
$$

where

$$
\begin{aligned}
& Var\left(\sqrt{nh_n}(\widehat{f}_n(x) - E\widehat{f}_n(x))\right) \\
&= f(x)\int K^2(z)dz + o(1) \equiv \sigma^2(x) + o(1) \qquad (1)
\end{aligned}
$$

by the variance calculation in (c), and

$$
\sqrt{nh_n}(E\widehat{f}_n(x) - f(x)) = \sqrt{n}h_n^{3/2}o(1) = C^{3/2}o(1) \to 0
$$

if $h_n = Cn^{-1/3}$. Moreover, we can easily show (for example by use of the Lindeberg-Feller CLT), that

$$
\sqrt{nh_n}(\widehat{f}_n(x) - E\widehat{f}_n(x)) \to_d N(0, \sigma^2(x))
$$

where $\sigma^2(x)$ is as given in (1). Thus it follows that if $h_n = Cn^{-1/3}$, then

$$
\sqrt{nh_n}(\widehat{f}_n(x) - f(x)) \to_d N(0, \sigma^2(x)),
$$

or, equivalently,

$$
n^{1/3}(\widehat{f}_n(x) - f(x)) \to_d N(0, \sigma^2(x)/C).
$$

(f) For the current choice of the kernel $b(x) = 0$ and there is no trade-off in the choice of C. Apparently we improve the estimator $\widehat{f}_n$ by choosing C to be very large. However, the following analysis for a different kernel does give a non-trivial tradeoff at this point and clarifies the situation to some degree.

Modified problem and solution: Suppose that the kernel is taken to be $K(x) = 1_{[-1,0]}(x)$.
(a) Now the expectation of $\widehat{f}_n(x)$ is given by

$$
\begin{aligned}
E(\widehat{f}_n(x)) &= \int \frac{1}{h} K\left(\frac{x-y}{h_n}\right) dF(y) \\
&= \int K(z) f(x-hz)dz = \int_{-1}^{0} f(x-hz)dz = -\frac{1}{h} F(x-hz)\Big|_{-1}^{0} \\
&= \frac{1}{h}(F(x+h) - F(x)) \\
&= \frac{1}{h}\left(f(x)h + \frac{1}{2} f'(x^*)h^2\right) \\
&\qquad \text{for some} \quad x^* \in [x-h, x], \\
&= f(x) + \frac{h}{2} f'(x^*).
\end{aligned}
$$

(b) It follows from the expectation computed in (a) that the bias of $\widehat{f}_n(x)$ is given by

$$
\text{bias}(\widehat{f}_n(x)) = \frac{h}{2} f'(x^*) = \frac{h}{2} f'(x) + o(h)
$$

since $f'(x^*) \to f'(x)$.
(c) The variance calculation is unchanged in this case.
(d) Now the variance is of the order $(1/nh)$ while the squared bias is of the order h^2, and these are both exactly of the order $n^{-2/3}$ if $h = h_n = Cn^{-1/3}$. Thus $r = 1/3$ gives the optimal rate.
(e) Now for $h_n = Cn^{-1/3}$ we have

$$
\sqrt{nh_n}(\widehat{f}_n(x) - f(x)) \to_d N(b(x), \sigma^2(x))
$$

where $b(x) \equiv C^{3/2} f'(x)/2$. Equivalently,

$$
n^{1/3}(\widehat{f}_n(x) - f(x)) \to_d N(Cf'(x)/2, \sigma^2(x)/C),
$$

and hence the asymptotic mean squared error is

$$
C^2 \frac{(f'(x))^2}{4} + \frac{\sigma^2(x)}{C}.
$$

This is minimized as a function of C by

$$C_{opt} = \left(\frac{\sigma^2(x)}{2(f'(x)/2)^2} \right)^{1/3} = \left(\frac{2f(x) \int K^2(z)dz}{(f'(x))^2} \right)^{1/3}.$$

Note that this is very sensible: if $(f'(x))^2$ is small, then the bandwidth can be large (more smoothing), while if $(f'(x))^2$ is big, then the bandwidth (at x) should be narrower so that we do less smoothing at x.

5. (48 points) Consider the functional $T(F) = \int \int |x-y| dF(x) dF(y)$ as a measure of spread or dispersion of the distribution function F. (This functional is sometimes called "Gini's mean difference".)
(a) If $X_1, \ldots, X_n$ are i.i.d. random variables with distribution function F, what is the "principle of substitution" estimator of $T(F)$?
(b) Is the estimator you found in (a) an unbiased estimator of $T(F)$? (Calculate the bias explicitly.)
(c) Use the jackknife to suggest an estimator of $T(F)$ with less bias. Can you find an unbiased estimator of $T(F)$?
(d) Calculate the Gateaux derivative of $T(F)$, and use this to find a formula for the asymptotic variance of $\sqrt{n}(T(\mathbb{F}_n) - T(F))$.
(e) Describe how you would use the bootstrap to estimate $nVar(T(\mathbb{F}_n))$ and

$$H_n(x, F) \equiv P_F(\sqrt{n}(T(\mathbb{F}_n) - T(F)) \le x),$$

distinguishing clearly in your description between the "ideal bootstrap" and the Monte-carlo implementation thereof.

Solution: (a) The principle of substitution estimator is just

$$T(\mathbb{F}_n) = \int \int |x-y| d\mathbb{F}_n(x) d\mathbb{F}_n(y) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} |X_i - X_j|.$$

(b) The principle of substitution estimator is biased: because the diagonal terms (for which $j = i$) in the sum are zero we have

$$\begin{aligned} E_F T(\mathbb{F}_n) &= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} E_F |X_i - X_j| \\ &= \frac{n(n-1)}{n^2} E_F |X_1 - X_2| = \frac{n-1}{n} T(F). \end{aligned}$$

Thus the bias of $T_n \equiv T(\mathbb{F}_n)$ is

$$\text{bias}_n(F) = E_F(T_n) - T(F) = \left(\frac{n-1}{n} - 1 \right) T(F) = -\frac{1}{n} T(F).$$

(c) Here we compute

$$T_{n,i} \equiv T(\mathbb{F}_{n-1,i}) = \frac{1}{(n-1)^2} \sum_{j=1, j\neq i}^{n} \sum_{j'=1, j'\neq i}^{n} |X_j - X_{j'}| \, .$$

Hence the pseudo-values are

$$\begin{aligned} T^*_{n,i} &= nT_n - (n-1)T_{n,i} = \left\{ \frac{1}{n} - \frac{1}{n-1} \right\} \sum_{j=1, j\neq i}^{n} \sum_{j'=1, j'\neq i}^{n} |X_j - X_{j'}| + \frac{2}{n} \sum_{j=1}^{n} |X_i - X_j| \\ &= -\frac{1}{n(n-1)} \sum_{j=1, j\neq i}^{n} \sum_{j'=1, j'\neq i}^{n} |X_j - X_{j'}| + \frac{2}{n} \sum_{j=1}^{n} |X_i - X_j| \, . \end{aligned}$$

Thus we find that

$$\begin{aligned} \overline{T^*_n} &\equiv \frac{1}{n} \sum_{i=1}^{n} T^*_{n,i} \\ &= 2T_n - \frac{1}{n^2(n-1)} \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \sum_{j'=1, j'\neq i}^{n} |X_j - X_{j'}| \\ &= 2T_n - \frac{1}{n^2(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{j'=1}^{n} |X_j - X_{j'}| 1_{[i\neq j]} 1_{[i\neq j']} \\ &= 2T_n - \frac{1}{n^2(n-1)} \sum_{j=1}^{n} \sum_{j'=1}^{n} \sum_{i=1}^{n} |X_j - X_{j'}| 1_{[i\neq j]} 1_{[i\neq j']} \\ &= 2T_n - \frac{n-2}{n^2(n-1)} \sum_{j=1}^{n} \sum_{j'=1}^{n} |X_j - X_{j'}| \\ &= 2T_n - \frac{n-2}{n-1} T_n = \frac{n}{n-1} T_n \\ &= \frac{1}{n(n-1)} \sum_{j=1}^{n} \sum_{j'=1}^{n} |X_j - X_{j'}| \, . \end{aligned}$$

This estimator of $T(F)$ is unbiased: $E_F(\overline{T^*_n}) = T(F)$. In fact $\overline{T^*_n}$ is the usual $U-$statistic form corresponding to the $V-$ statistic $T(\mathbb{F}_n)$.

(d) To calculate the Gateaux derivative, we write $F_\epsilon = (1-\epsilon)F + \epsilon G$ and then compute

$$\begin{aligned}
T(F_\epsilon) &= \int\int |x-y| dF_\epsilon(x) dF_\epsilon(y) \\
&= \int\int |x-y| dF(x) dF(y) + \epsilon \int\int |x-y| d(G-F)(x) dF(y) \\
&\quad + \epsilon \int\int |x-y| dF(x) d(G-F)(y) \\
&\quad + \epsilon^2 \int\int |x-y| d(G-F)(x) d(G-F)(y)\,.
\end{aligned}$$

Hence we find that the Gateaux derivative $\dot{T}(F; G-F)$ is given by

$$\begin{aligned}
\dot{T}(F; G-F) &= \frac{d}{d\epsilon} T(F_\epsilon)\Big|_{\epsilon=0} \\
&= \int\int |x-y| d(G-F)(x) dF(y) + \int\int |x-y| dF(x) d(G-F)(y) \\
&= 2\int\int |x-y| d(G-F)(x) dF(y)\,.
\end{aligned}$$

Taking $G = \delta_x$ yields the influence function for T:

$$\begin{aligned}
IC(x; T, F) &= \psi_F(x) \\
&= 2\left(\int |x-y| dF(y) - \int\int |x-y| dF(x) dF(y)\right) \\
&= 2\left(\int |x-y| dF(y) - T(F)\right)\,.
\end{aligned}$$

This leads to the conclusion that the asymptotic variance of $\sqrt{n}(T(\mathbb{F}_n) - T(F))$ will be

$$\begin{aligned}
E_F \psi_F^2(X_1) &= 4\int \left(\int |x-y| dF(y) - T(F)\right)^2 dF(x) \\
&= 4\left\{\int\int |x-y| dF(y) \int |x-y'| dF(y') dF(x) - T^2(F)\right\} \\
&= 4\left\{\int\int\int |x-y||x-y'| dF(x) dF(y) dF(y') - T^2(F)\right\}\,.
\end{aligned}$$

(e) The ideal bootstrap estimator of $nVar_F(T(\mathbb{F}_n))$ is $nVar_{\mathbb{F}_n}(T(\mathbb{F}_n^*))$ where $\mathbb{F}_n^*$ is the empirical distribution function of $X_1^*, \ldots, X_n^*$ i.i.d. with distribution function $\mathbb{F}_n$. Similarly, the ideal bootstrap estimator of

$$H_n(x, F) \equiv P_F(\sqrt{n}(T(\mathbb{F}_n) - T(F)) \le x)\,,$$

is simply

$$H_n(x, \mathbb{F}_n) \equiv P_{\mathbb{F}_n}(\sqrt{n}(T(\mathbb{F}_n^*) - T(\mathbb{F}_n)) \le x)\,.$$

To implement Monte-Carlo approximations of these, we would draw B bootstrap samples of size n from $\mathbb{F}_n$,

$$X_{j,1}^*, \ldots, X_{j,n}^*\,, \qquad j = 1, \ldots, B,$$

form the corresponding empirical d.f. $\mathbb{F}_{j,n}^*$ for each $j = 1, \ldots, B$, and calculate the resulting $T_{n,j}^* \equiv T(\mathbb{F}_{j,n}^*)$. Then

$$n\widehat{Var_{\mathbb{F}_n}(T}(\mathbb{F}_n^*)) = n\frac{1}{B}\sum_{j=1}^{B}(T_{n,j}^* - \overline{T}_n^*)^2$$

while

$$\widehat{H_n(x, \mathbb{F}_n)} = \frac{1}{B}\sum_{j=1}^{B} 1_{[\sqrt{n}(T_{n,j}^* - T_n) \le x]}\,.$$