**Statistics 581**
**Problem Set 5 Solutions**
Wellner; 10/29/2014

1. Suppose that $X_1, \ldots, X_n$ are i.i.d. Cauchy$(0,1)$; so the density of each $X_i$ with respect to Lebesgue measure on $R$ is $f(x) = \pi^{-1}(1+x^2)^{-1}$, $x \in R$.
(a) Compute the distribution function $F$ of the $X_i$'s.
(b) Compute and plot the inverse distribution function $F^{-1}$ corresponding to $F$.
(c) For what values of $r > 0$ is $E|X_1|^r < \infty$?
(d) Find the distribution function of $M_n \equiv \max_{1 \le i \le n} X_i$.
(e) For what values of $r$ is $E|M_n|^r < \infty$?
(f) Find a sequence of constants $b_n$ so that $M_n/b_n \rightarrow_d$ and find the limiting distribution. [Hint: see Ferguson, ACLST, Theorem 14, page 95.]

**Solution:** (a) $F(x) = (1/\pi)\int_{-\infty}^x (1+t^2)^{-1} dt = (1/\pi)\{\arctan(x) + \pi/2\}$.
(b) Setting $F(x) = u$ and solving for $x = F^{-1}(u)$ yields $F^{-1}(u) = \tan(\pi(u - 1/2))$. Note that $F^{-1}(1/2) = \tan(0) = 0$; $F^{-1}(1) = \tan(\pi/2) = \infty$, and $F^{-1}(0) = \tan(-\pi/2) = -\infty$.
(c) We compute

$$\begin{aligned} E|X_1|^r &= \frac{1}{\pi}\int_{-\infty}^{\infty} |x|^r \frac{1}{1+x^2} dx \\ &= \frac{2}{\pi}\left\{\int_0^1 \frac{x^r}{1+x^2} dx + \int_1^\infty \frac{x^r}{1+x^2} dx\right\} \\ &\le \frac{2}{\pi}\left\{\int_0^1 \frac{x^r}{1+x^2} dx + \int_1^\infty \frac{x^r}{x^2} dx\right\} \\ &= \frac{2}{\pi}\left\{\int_0^1 \frac{x^r}{1+x^2} dx + \frac{1}{1-r}\right\} < \infty \end{aligned}$$

if $r < 1$. Since

$$E|X_1| = \frac{2}{\pi}\int_0^\infty \frac{x}{1+x^2} dx = \infty\,,$$

$E|X_1|^r < \infty$ if and only if $r < 1$.
(d) Since the $X_i$'s are i.i.d. with distribution function $F$,

$$F_{M_n}(x) = P(M_n \le x) = P(X_1 \le x, \ldots, X_n \le x) = F(x)^n.$$

(e) First, note that

$$1 - F_{|M_n|}(x) = P(|M_n| > x) = P(\cup_{i=1}^n [|X_i| > x]) \le \sum_{i=1}^n P(|X_i| > x) = n(1 - F_{|X_1|}(x))$$

where $F_{|X_1|}(x) = P(|X_1| \le x) = F(x) - F(-x)$. Hence

$$\begin{aligned} E|M_n|^r &= \int_0^\infty r t^{r-1}(1 - F_{|M_n|}(t)) dt \\ &\le \int_0^\infty r t^{r-1} n(1 - F_{|X|}(t)) dt \\ &= nE|X_1|^r < \infty \end{aligned}$$

if $r < 1$ by part (d). But since $E|M_n|^r \ge E|X_1|^r = \infty$ if $r \ge 1$, we conclude that $E|M_n|^r < \infty$ if and only if $r < 1$.
(f) Note that $1-F(x) = \pi^{-1}\int_x^\infty (1+t^2)^{-1}dt \sim 1/(\pi x)$ in the sense that $x(1-F(x)) \to 1/\pi$ as $x \to \infty$. [This follows easily by writing the left side as $(1-F(x))/(x^{-1})$ and using L'Hopital's rule.] Hence for $b_n \to \infty$ and $x > 0$

$$F_{M_n/b_n}(x) \;=\; P(M_n \le x b_n) = F(xb_n)^n \qquad \text{by part (d)}$$

and, with $c_n \equiv x b_n(1-F(xb_n)) \to 1/\pi$,

$$\begin{aligned} F_{M_n/b_n}(x) &= F(xb_n)^n = (1-(1-F(xb_n)))^n \\ &= (1-[xb_n(1-F(xb_n))]/(xb_n))^n \\ &= (1-c_n/xb_n)^n. \end{aligned}$$

From this last expression it becomes clear that the choice $b_n = n$ yields,

$$F_{M_n/b_n}(x) \to \exp(-1/\pi x) \equiv G(x), \qquad \text{for } x > 0,$$

while for $x \le 0$

$$F_{M_n/b_n}(x) \to 0$$

since $F(xb_n) \le 1/2$ for $x \le 0$. Note that $G(0) = \exp(-\infty) = 0$, $G$ is monotone increasing, and $G(\infty) = \exp(0) = 1$. In fact, $G$ is a member of the Weibull family with shape parameter $-1$, and is one of the three different families that can arise as limit distributions of maxima of independent rv's; see e.g. Ferguson (1996), *A Course in Large Sample Theory*, page 95.

2. Suppose that $X_1, \ldots, X_n$ are i.i.d. random vectors with values in $R^k$ with $E(X_1) = \mu$ and $E(X_1^T X_1) < \infty$ so that $\Sigma = E(X_1-\mu)(X_1-\mu)^T$ is well-defined. Thus

$$Z_n \equiv \sqrt{n}(\overline{X}_n - \mu) \to_d Z \sim N_k(0, \Sigma)\,.$$

Suppose that $g : R^k \to R$ is a function, and suppose that $\nabla g = \dot{g}$ exists at $\mu$. Then the delta-method (or $g'$ theorem) tells us that

$$(1) \qquad \sqrt{n}(g(\overline{X}_n) - g(\mu)) \to_d \nabla g(\mu)^T Z \sim N(0, \nabla g(\mu)^T \Sigma \nabla g(\mu))\,.$$

(a) Show that we can strengthen (**??**) as follows: Suppose that $\nabla g = \dot{g}$ is continuous at $\mu$. Then $\sqrt{n}(g(\overline{X}_n) - g(\mu))$ is *asymptotically linear* at $\mu$:

$$\begin{aligned} \sqrt{n}(g(\overline{X}_n) - g(\mu)) &= \nabla g(\mu)^T \sqrt{n}(\overline{X}_n - \mu) + o_p(1) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi(X_i) + o_p(1) \end{aligned}$$

where

$$(2) \qquad \psi(x) = \nabla g(\mu)^T (x - \mu)$$

which is called the *influence function* of $g(\overline{X}_n)$ as an estimator of $g(\mu)$, has mean $E\psi(X_i) = 0$ and $Var(\psi(X_i)) = \nabla g(\mu)^T \Sigma \nabla g(\mu)$.
(b) Does the result of (a) apply to the situation considered in problem 3(a) of

problem set #4? If not, formulate another result of the same type as in (a) which does apply, and use it to find the influence function of $S_n^2/\overline{X}_n$.

**Solution:** By Taylor's theorem, for some $Y_n^*$ satisfying $|Y_n^* - \mu| \le |\overline{X}_n - \mu| \to_p 0$ it follows that

$$\begin{aligned}\sqrt{n}(g(\overline{X}_n) - g(\mu)) &= \nabla g(Y_n^*)\sqrt{n}(\overline{X}_n - \mu) \\ &= \nabla g(\mu)\sqrt{n}(\overline{X}_n - \mu) \\ &\qquad + \{\nabla g(Y_n^*) - \nabla g(\mu)\}\sqrt{n}(\overline{X}_n - \mu) \\ &= \nabla g(\mu)\sqrt{n}(\overline{X}_n - \mu) + o_p(1)\end{aligned}$$

since $\nabla g(Y_n^*) \to_p \nabla g(\mu)$ by continuity of $\nabla g$ at $\mu$ and since $\sqrt{n}(\overline{X}_n - \mu) = O_p(1)$. Now note that

$$\nabla g(\mu)\sqrt{n}(\overline{X} - \mu) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \nabla g(\mu)(X_i - \mu) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi(X_i)$$

with $\psi$ as in (**??**).

In fact, the hypothesis of continuity of $\nabla g$ can be dropped: consider a new function $h(x) = g(x) - \nabla g(\mu)x$. Then $\nabla h(\mu) = \nabla g(\mu) - \nabla g(\mu) = 0$, and we can write

$$\begin{aligned}\sqrt{n}(g(\overline{X}_n) - g(\mu) - \nabla g(\mu)(\overline{X}_n - \mu)) &= \sqrt{n}(h(\overline{X}_n) - h(\mu)) \\ &\to_d \nabla h(\mu)Z = 0 \cdot Z = 0 \qquad (3)\end{aligned}$$

by the delta-method applied to the function $h$. Since convergence in distribution to a constant implies convergence in probability to the same constant, we conclude from (**??**) that the left side of (**??**) converges in probability to 0. But this is just the claimed asymptotic linearity with $\psi(x) = \nabla g(\mu)(x - \mu)$.

(b) The result in (a) does not quite apply since

$$Z_n \equiv \sqrt{n}(\overline{X}_n - \mu, S_n^2 - \sigma^2)'$$

is not exactly an average of i.i.d. random vectors. But the key features of the proof in (a) do carry through since $n^{-1/2}Z_n \to_p 0$ and $Z_n = n^{-1/2}\sum_{i=1}^n \underline{Y}_i + o_p(1)$ where $\underline{Y}_i = (X_i - \mu, (X_i - \mu)^2 - \sigma^2)'$ are i.i.d. with mean 0 and finite second moment under the assumptions of problem 2(a) of problem set #4. Thus the conclusion continues to hold.

Thus for the setting of problem 4.3(a), with $g(u, v) = v/u$

$$\begin{aligned}\sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - \frac{\sigma^2}{\mu}\right) &= \nabla g(\mu, \sigma^2)\sqrt{n}(\underline{\overline{Y}}_n - \underline{\mu}_Y) + o_p(1) \\ &= \frac{1}{\mu}(-\sigma^2/\mu, 1)\sqrt{n}(\underline{\overline{Y}}_n - \underline{\mu}_Y) + o_p(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{1}{\mu}\left\{(X_i - \mu)^2 - \sigma^2 - \frac{\sigma^2}{\mu}(X_i - \mu)\right\} + o_p(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi(X_i) + o_p(1)\end{aligned}$$

where

$$
\begin{aligned}
\psi(x) &= \frac{1}{\mu}\{(x-\mu)^2 - \sigma^2 - (\sigma^2/\mu)(x-\mu)\} \\
&= \frac{\sigma^2}{\mu}\left\{\left(\frac{x-\mu}{\sigma}\right)^2 - 1 - \frac{\sigma}{\mu}\left(\frac{x-\mu}{\sigma}\right)\right\}
\end{aligned}
$$

has $E\psi(X_i) = 0$ and

$$
Var(\psi(X_i)) = \frac{\sigma^4}{\mu^2}\left\{2 + \gamma_2 - 2\frac{\sigma\gamma_1}{\mu} + \frac{\sigma^2}{\mu^2}\right\} = V^2
$$

as in the solution of problem 4.3(a).

3. (a) Write out a direct proof of (17) on page 27 of the Chapter 2 notes, using the multivariate CLT, and compare with the result of using (10) and (16).
(b) Write out a proof of the corresponding fact concerning the general empirical process $\mathbb{G}_n$: $\mathbb{G}_n \to_{f.d.} \mathbb{G}$ where $\mathbb{G}_n$ and $\mathbb{G}$ are as defined on pages 30-31 of the chapter 2 notes; i.e. for any $f_1, \ldots, f_k \in L_2(P)$, $(\mathbb{G}_n(f_1), \ldots, \mathbb{G}_n(f_k)) \to_d (\mathbb{G}(f_1), \ldots, \mathbb{G}(f_k))$ where $\mathbb{G}$ is a $P-$Brownian bridge process with mean 0 and Covariance $E\{\mathbb{G}(f)\mathbb{G}(g)\} = P(fg) - P(f)P(g)$ as in (39) on page 32.

**Solution:** (a) $\sqrt{n}(\mathbb{F}_n - F) \to_{f.d.} \mathbb{U}(F)$. To see this, let $-\infty < x_1 < x_2 < \cdots < x_k < \infty$. Then define random vectors $\underline{Y}_i$ by

$$
\underline{Y}_i = \left(1_{(-\infty,x_1]}(X_i) - F(x_1), \ldots, 1_{(-\infty,x_k]}(X_i) - F(x_k)\right)^T
$$

for $i = 1, \ldots, n$. Note that $E\underline{Y}_1 = 0$ and

$$
\begin{aligned}
E\underline{Y}_1\underline{Y}_1' &= \begin{pmatrix}
F(x_1)(1-F(x_1)) & F(x_1)(1-F(x_2)) & \cdots & F(x_1)(1-F(x_k)) \\
F(x_1)(1-F(x_2)) & F(x_2)(1-F(x_2)) & \cdots & F(x_2)(1-F(x_k)) \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
F(x_1)(1-F(x_k)) & F(x_2)(1-F(x_k)) & \cdots & F(x_k)(1-F(x_k))
\end{pmatrix} \\
&= (F(x_j \wedge x_{j'}) - F(x_j)F(x_{j'}))_{j,j'=1}^k \equiv \Sigma
\end{aligned}
$$

Thus it follows from the multivariate central limit theorem that

$$
\sqrt{n}(\mathbb{F}_n(x_1) - F(x_k), \cdots, \mathbb{F}_n(x_k) - F(x_k))^T = \sqrt{n}\underline{Y}_n \to_d N_k(0, \Sigma).
$$

But for a Brownian bridge process $\mathbb{U}$, $(\mathbb{U}(F(x_1)), \ldots, \mathbb{U}(F(x_k))^T \sim N_k(0, \Sigma)$, so we have shown that

$$
\sqrt{n}(\mathbb{F}_n(x_1) - F(x_k), \cdots, \mathbb{F}_n(x_k) - F(x_k))^T \to_d (\mathbb{U}(F(x_1)), \ldots, \mathbb{U}(F(x_k))^T.
$$

But since this holds for every $k$ and every choice of $x_1 < x_2 < \cdots < x_k$, it follows that $\sqrt{n}(\mathbb{F}_n - F) \to_{f.d.} \mathbb{U}(F)$.
(b) $\mathbb{G}_n \to_{f.d.} \mathbb{G}$. To see this, let $f_1, \ldots, f_k \in L_2(P)$. Then define random vectors $\underline{Y}_i$ by

$$
\underline{Y}_i = (f_1(X_i) - Pf_1, \ldots, f_k(X_i) - Pf_k)^T
$$

for $i = 1, \ldots, n$. Note that $E\underline{Y}_i = 0$ and

$$
\begin{aligned}
E\underline{Y}_1\underline{Y}_1^T &= \begin{pmatrix}
P(f_1^2) - (P(f_1)^2 & P(f_1 f_2) - Pf_1 Pf_2 & \cdots & P(f_1 f_k) - Pf_1 Pf_k \\
P(f_1 f_2) - Pf_1 Pf_2 & P(f_2^2) - (Pf_2)^2 & \cdots & P(f_2 f_k) - Pf_2 Pf_k \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
P(f_1 f_k) - Pf_1 Pf_k & P(f_2 f_k) - Pf_2 Pf_k & \cdots & P(f_k^2) - (Pf_k)^2
\end{pmatrix} \\
&= (P(f_i f_j) - Pf_i Pf_j)_{i,j=1}^k \equiv \Sigma.
\end{aligned}
$$

Thus it follows from the multivariate central limit theorem that

$$(\mathbb{G}_n(f_1), \ldots, \mathbb{G}_n(f_k))^T = \sqrt{n}\underline{Y}_n \to_d N_k(0, \Sigma).$$

But for a $P-$Brownian bridge process $\mathbb{G}_P$, $(\mathbb{G}_P(f_1), \ldots, \mathbb{G}_P(f_k))' \sim N_k(0, \Sigma)$, so we have shown that $(\mathbb{G}_n(f_1), \ldots, \mathbb{G}_n(f_k))^T \to_d (\mathbb{G}(f_1), \ldots, \mathbb{G}(f_k))^T$. But since this holds for every $k$ and every choice of $f_1, \ldots, f_k \in L_2(P)$, it follows that $\mathbb{G}_n \to_{f.d.} \mathbb{G}$.

4. Suppose that $X_1, \ldots, X_n$ are i.i.d. with continuous distribution function $F$. Let $F_0$ be a fixed, specified continuous distribution function. Suppose we want to test $H: \ F = F_0$ versus $K: \ F \neq F_0$. Consider the *Cramér - von Mises statistic* given by

$$C_n^2 \equiv \int_{-\infty}^{\infty} n(\mathbb{F}_n(x) - F_0(x))^2 dF_0(x)\,.$$

(a) Show that if the null hypothesis $H$ holds (so that $F_0$ is the true distribution function of the $X_i$'s), then

$$C_n^2 \stackrel{d}{=} \int_0^1 n(\mathbb{G}_n(t) - t)^2 dt\,,$$

where $\mathbb{G}_n$ is the empirical d.f. of $n$ i.i.d. Uniform$(0,1)$ rv's.
(b) Show that when the null hypothesis is true,

$$C_n^2 \to_d \int_0^1 \mathbb{U}(t)^2 dt$$

where $\mathbb{U}$ is a standard Brownian bridge process.
[Hint: Use the fact that $\mathbb{U}_n \Rightarrow \mathbb{U}$ in $(D[0,1], \|\cdot\|_\infty)$ and the continuous mapping theorem.]
(c) Suppose that the null hypothesis fails. Thus $F \neq F_0$. Show that in this case

$$n^{-1} C_n^2 \to_{a.s.} \int_{-\infty}^{\infty} (F(x) - F_0(x))^2 dF_0(x) > 0\,,$$

and hence the test based on $C_n^2$ is consistent for all $F \neq F_0$.

**Solution:** (a) Now $\sqrt{n}(\mathbb{F}_n - F) \stackrel{d}{=} \mathbb{U}_n(F)$ is always true (for any df $F$) when the $X_i$ are i.i.d. $F$, so under the null hypothesis $F = F_0$

$$C_n^2 \equiv \int_{-\infty}^{\infty} n(\mathbb{F}_n(x) - F_0(x))^2 dF_0(x) \stackrel{d}{=} \int_{-\infty}^{\infty} [\mathbb{U}_n(F_0)]^2 dF_0$$

holds. By the change of variable $t = F_0(x)$, the variable $t$ takes on all values in $(0,1)$ when $F_0$ is continuous, and

$$\int_{-\infty}^{\infty} [\mathbb{U}_n(F_0)]^2 dF_0 = \int_0^1 [\mathbb{U}_n(t)]^2 dt.$$

Thus the stated conclusion holds.
(b) Now $\mathbb{U}_n \Rightarrow \mathbb{U}$ and $g(x) = \int_0^1 [x(t)]^2 dt$ is a continuous function from $(D[0,1], \|\cdot\|)$ to $\mathbb{R}$ (since

$$|g(x) - g(y)| = \left| \int_0^1 (x^2(t) - y^2(t)) dt \right| \le \|x - y\|_\infty \|x + y\|_\infty.$$

Thus by the continuous mapping theorem, when the null hypothesis is true,

$$C_n^2 \stackrel{d}{=} g(\mathbb{U}_n) \to_d g(\mathbb{U}) = \int_0^1 \mathbb{U}^2(t) dt.$$

(c) When $F \neq F_0$, $\|\mathbb{F}_n - F\|_\infty \to_{a.s.} 0$ by Glivenko-Cantelli, so we define $c^2 \equiv c^2(F, F_0) \equiv \int_{-\infty}^{\infty} (F(x) - F_0(x))^2 dF_0(x)$. Then

$$\begin{aligned}
& |n^{-1} C_n^2 - c^2| \\
= & \left| \int_{-\infty}^{\infty} \{(\mathbb{F}_n(x) - F_0(x))^2 - (F(x) - F_0(x))^2\} dF_0(x) \right| \\
\le & \int_{-\infty}^{\infty} |(\mathbb{F}_n(x) - F_0(x) - (F(x) - F_0(x))(\mathbb{F}_n(x) - F_0(x) + F(x) - F_0(x))| dF_0(x) \\
\le & \int_{-\infty}^{\infty} |\mathbb{F}_n - F| \{|\mathbb{F}_n - F_0| + |F - F_0|\} dF_0 \\
\le & 2\|\mathbb{F}_n - F\|_\infty \to_{a.s.} 0.
\end{aligned}$$

Thus we conclude that $n^{-1} C_n^2 \to_{a.s.} c^2$.

5. **Optional bonus problem:** This is a continuation of the previous problem, and should be thought of in analogy with our development for the Pearson chi-square statistic.
(a) Suppose that $F = F_n$ satisfies $\sqrt{n}(F_n(x) - F_0(x)) \to g(x)$ in $L_2(F_0)$; i.e.

$$\int [\sqrt{n}(F_n(x) - F_0(x)) - g(x)]^2 dF_0(x) \to 0\,.$$

Describe the limiting distribution of $C_n^2$ under the local alternatives $F_n$ in terms of a Brownian bridge process $\mathbb{U}$ and $g$.
(b) Let $c^2$ denote the constant on the right side in Problem 5(c) above. In the set-up of that problem, show that when $F \neq F_0$ it follows that

$$\sqrt{n}(n^{-1} C_n^2 - c^2) \to_d N(0, V^2)$$

and find $V^2$.
[Hint: Use $\sqrt{n}(\mathbb{F}_n - F) =_d \mathbb{U}_n(F)$, $\mathbb{U}_n \Rightarrow \mathbb{U}$, and the continuous mapping theorem.]

**Solution:** (a) If $F = F_n$ satisfies $\sqrt{n}(F_n - F_0) \to g$, then

$$\begin{aligned}
C_n^2 &\stackrel{d}{=} \int [\sqrt{(}\mathbb{F}_n^*(x) - F_n(x)) + \sqrt{n}(F_n(x) - F_0(x))]^2 dF_0(x) \\
&= \int [\mathbb{U}_n(F_n) + \sqrt{n}(F_n - F_0)]^2 dF_0 \\
&\to_d \int [\mathbb{U}(F_0) + g]^2 dF_0
\end{aligned}$$

where only a few more lines are needed to completely justify the convergence in the last line.
(b) For a fixed alternative $F$ we can write

$$\begin{aligned}
\sqrt{n}(C_n - c^2) &= \sqrt{n} \int \left\{(\mathbb{F}_n - F_0)^2 - (F - F_0)^2\right\} dF_0 \\
&= \sqrt{n} \int \{(\mathbb{F}_n - F_0 - (F - F_0))(\mathbb{F}_n - F_0 + F - F_0)\} dF_0 \\
&= \int \sqrt{n}(\mathbb{F}_n - F) \{\mathbb{F}_n + F - 2F_0\} dF_0 \\
&\stackrel{d}{=} \int \mathbb{U}_n(F) \{\mathbb{G}_n(F) + F - 2F_0\} dF_0 \\
&\to_d \int \mathbb{U}(F) 2(F - F_0) dF_0 \sim N(0, V^2)
\end{aligned}$$

where

$$\begin{aligned}
V^2 &\equiv V^2(F_0, F) \\
&= 4 \int\int (F(x) - F_0(x))(F(y) - F_0(y)) (F(x \wedge y) - F(x)F(y)) dF_0(x) dF_0(y).
\end{aligned}$$