Statistics 522, Final Exam Solutions

Wellner; 3/20/2008

1. (24 points). **Define** *three* of the following six terms:
   (a) An asymptotically tight sequence $\{X_n\}$ in a metric space (M, d).
   (b) Weak convergence of probability measures $\{P_n\}$ to a probability measure P on $(M, \mathcal{M})$ for a metric space (M, d).
   (c) The characteristic function of a real-valued random variable X and of a random vector $\underline{X}$ with values in $\mathbb{R}^d$.
   (d) A Brownian motion process $\mathbb{S}$ on $[0, \infty)$.
   (e) A Brownian bridge process $\mathbb{U}$ on $[0, 1]$.
   (f) The "Uniformly Asymptotically Negligible" (or UAN) property of a triangular array of random variables $\{X_{nk} : 1 \le k \le n,\ n \ge 1\}$ (in the context of the Lindeberg-Feller central limit theorem).

   Solution: See the course notes.

2. (30 points). Give careful **statements** of *any three* of the following six theorems or results:
   (a) A result connecting convergence of characteristic functions of a sequence of random variables X_n to tightness of the sequence X_n.
   (b) Donsker's theorem for the uniform empirical process $\mathbb{U}_n(t) = \sqrt{n}(\mathbb{G}_n(t) - t)$ where $\mathbb{G}_n(t) = n^{-1} \sum_{i=1}^n 1_{[0,t]}(\xi_i)$ and $\xi_1, \ldots, \xi_n$ are i.i.d. Uniform$(0, 1)$ random variables.
   (c) Any result connecting a Brownian motion process $\mathbb{S}$ on $[0, 1]$ or $[0, \infty)$ to a Brownian bridge process $\mathbb{U}$ on $[0, 1]$.
   (d) The Cramér - Lévy continuity theorem for characteristic functions.
   (e) The law of the iterated logarithm for $S_n = X_1 + \cdots + X_n$ where $X_1, \ldots, X_n$ are i.i.d. with $EX_i = 0$ and $Var(X_i) = \sigma^2$.
   (f) The Berry-esseen theorem for $S_n = X_1 + \cdots + X_n$ where $X_1, \ldots, X_n$ are i.i.d. with $E(X_1) = 0$, $Var(X_1) = \sigma^2$, and $E|X_1|^3 < \infty$.

   Solution: See the course notes.

3. (36 points) Consider the following sequences of distribution functions F_n on $\mathbb{R}$ with densities f_n with respect to Lebesgue measure:
   (a) $f_n(x) = n^{-1} 1_{[0,n]}(x)$.
   (b) $f_n(x) = n e^{-nx} 1_{[0,\infty)}(x)$.
   (c) $f_n(x) = n^{-1} e^{-x/n} 1_{[0,\infty)}(x)$.
   (i) In each of (a) - (c), describe these distributions in terms of random variables of

the form $X_n = a_n Y$ where Y has a fixed distribution F_0 with density f_0; that is, find random variables of this form so that X_n has distribution function F_n with density f_n.
(ii) Which of (a) - (c), if any, are uniformly (or asymptotically) tight sequences? In the case of a tight sequence, identify the collection of limiting distributions.

Solution: (i) (a) Let $Y \sim$ Uniform$[0,1]$ and $a_n = n$. Then $X_n \equiv nY \sim$ Uniform$[0,n]$ with density $f_n(x) = n^{-1}1_{[0,n]}(x)$.
Proof:

$$F_n(x) = P(X_n \le x) = P(Y \le x/n) = \{(x/n) \vee 0\} \wedge 1.$$

Thus $f_n(x) = n^{-1}1_{[0,n]}(x)$.
(b) Let $Y \sim$ Exponential(1) and let $a_n = n^{-1}$. Then $X_n = n^{-1}Y \sim$ Exponential(n) with density $f_n(x) = ne^{-nx}1_{[0,\infty)}(x)$.
Proof:

$$F_n(x) = P(X_n \le x) = P(Y \le nx) = (1 - e^{-nx}) \vee 0$$

and hence $f_n(x) = ne^{-nx}1_{[0,\infty)}(x)$.
(c) Let $Y \sim$ Exponential(1) and let $a_n = n$. Then $X_n = nY \sim$ Exponential(n^{-1}) with density $f_n(x) = n^{-1}e^{-x/n}1_{[0,\infty)}(x)$.
Proof:

$$F_n(x) = P(X_n \le x) = P(Y \le x/n) = (1 - e^{-x/n}) \vee 0$$

and hence $f_n(x) = n^{-1}e^{-x/n}1_{[0,\infty)}(x)$.
(ii) From the representations in (i) it is clear that:
(a) $X_n = nY$ is *not tight*; $X_n \to \infty$ a.s.
(b) $X_n = n^{-1}Y$ *is tight*; $X_n = n^{-1}Y \to 0$ a.s.
(c) $X_n = nY$ is *not tight*; $X_n = nY \to \infty$ a.s.
In the tight case (b), $X_n \to_d 0$ and hance the only limit point is the distribution δ_0 (point mass at 0) corresponding to the distribution function $F_0(x) = 1_{[0,\infty)}(x)$.

4. (25 points). Suppose that X and Y are random variables on the probability space $(\Omega, \mathcal{A}, P)$ with $X \in L_2(P)$ and $Y \in L_2(P)$ (so that $XY \in L_1(P)$), and suppose that $\mathcal{D}$ is a sub sigma-field of $\mathcal{A}$. Show that

$$Cov(X,Y) = E\left[Cov(X,Y|\mathcal{D})\right] + Cov(E(X|\mathcal{D}), E(Y|\mathcal{D}))$$

where

$$Cov(X,Y|\mathcal{D}) = E\left[(X - E(X|\mathcal{D})(Y - E(Y|\mathcal{D}))|\mathcal{D}\right].$$

(This generalizes our formula for the variance of a random variable X obtained in the midterm exam.)

Solution: By adding and subtracting $E(X|\mathcal{D})$ and $E(Y|\mathcal{D})$ we find

$$\begin{aligned} Cov(X,Y) &= E\{(X - EX)(Y - EY)\} \\ &= E\{(X - E(X|\mathcal{D}) + E(X|\mathcal{D}) - EX)(Y - E(Y|\mathcal{D}) + E(Y|\mathcal{D}) - EY)\} \end{aligned}$$

$$\begin{aligned}
&= E\{(X - E(X|\mathcal{D}))(Y - E(Y|\mathcal{D}))\} + E\{(E(X|\mathcal{D}) - EX)(E(Y|\mathcal{D}) - EY)\} \\
&\qquad + E\{(X - E(X|\mathcal{D})(E(Y|\mathcal{D}) - EY)\} + E\{(E(X|\mathcal{D}) - EX)(Y - E(Y|\mathcal{D}))\} \\
&= E\left[E\{(X - E(X|\mathcal{D}))(Y - E(Y|\mathcal{D}))|\mathcal{D}\}\right] \\
&\qquad + Cov\left[E(X|\mathcal{D}), E(Y|\mathcal{D})\right] + 0 + 0
\end{aligned}$$

since

$$\begin{aligned}
E\{(X - E(X|\mathcal{D})(E(Y|\mathcal{D}) - EY)\} &= E\left[E\{(X - E(X|\mathcal{D})(E(Y|\mathcal{D}) - EY)|\mathcal{D}\}\right] \\
&= E\left[(E(Y|\mathcal{D}) - EY)E\{(X - E(X|\mathcal{D})|\mathcal{D}\}\right] \\
&\qquad \text{since } (E(Y|\mathcal{D}) - EY) \text{ is } \mathcal{D}-\text{measurable} \\
&= E\left[(E(Y|\mathcal{D}) - EY)\{E(X|\mathcal{D}) - E(X|\mathcal{D})\}\right] \\
&\qquad \text{by linearity of conditional expectation} \\
&= E\left[(E(Y|\mathcal{D}) - EY) \cdot 0\right] \\
&= 0,
\end{aligned}$$

and similarly for the second "cross term" .

5. (30 points). Suppose that $\xi_1, \xi_2, \ldots$ are i.i.d. Uniform$(0,1)$ random variables, and let $\mathbb{G}_n(t) = n^{-1}\sum_{i=1}^n 1_{[0,t]}(\xi_i)$ be the empirical distribution function of the ξ_i's. Define $\mathcal{F}_n(t) \equiv \sigma\{\mathbb{G}_n(r) : \ t \le r \le 1\}$ for $0 < t \le 1$.
(a) Let $0 < s < t < 1$. Show that

$$(n\mathbb{G}(s), n(\mathbb{G}_n(t) - \mathbb{G}_n(s)), n(1 - \mathbb{G}_n(t))) \sim \text{Mult}_3(n, (s, t-s, 1-t)).$$

(b) Find the conditional distribution of $n\mathbb{G}_n(s)$ conditional on $n(1 - \mathbb{G}_n(t))$.
(c) Use the result of (b) to show that

$$E\left\{\frac{\mathbb{G}_n(s)}{s}\Big|\mathbb{G}_n(t)\right\} = \frac{\mathbb{G}_n(t)}{t} \qquad \text{almost surely.}$$

Note: It is true that the conditional distribution of $\mathbb{G}_n(s)$ given the $\sigma-$field $\mathcal{F}_n(t)$ is the same as the conditional distribution of $\mathbb{G}_n(s)$ given the random variable $\mathbb{G}_n(t)$; this is the Markov property of $\mathbb{G}_n$. Thus (c) shows that

$$E\left\{\frac{\mathbb{G}_n(s)}{s}\Big|\mathcal{F}_n(t)\right\} = \frac{\mathbb{G}_n(t)}{t} \qquad \text{almost surely;}$$

that is, $\{(\mathbb{G}_n(t)/t, \mathcal{F}_n(t)) : \ 0 < t \le 1\}$ is a reverse martingale.

Solution: (a) Note that

$$(n\mathbb{G}_n(s), n(\mathbb{G}_n(t) - \mathbb{G}_n(s)), n(1 - \mathbb{G}_n(t))) = \sum_{i=1}^n (1_{[0,s]}(\xi_i), 1_{(s,t]}(\xi_i), 1_{(t,1]}(\xi_i))$$

where $(1_{[0,s]}(\xi_i), 1_{(s,t]}(\xi_i), 1_{(t,1]}(\xi_i))$ are i.i.d. Multinomial$_3(1,(s,t-s,1-t))$ for $i = 1,\ldots,n$. Thus the sum has a Mult$_3(n,(s,t-s,1-t))$ distribution by definition.
(b) For any integers $k_1, k_2 \in \{0,\ldots,n\}$ with $k_1 + k_2 \le n$ we have

$$\begin{aligned}
&P(n\mathbb{G}_n(s) = k_1 \Big| n(1-\mathbb{G}_n(t)) = k_2) \\
&= \frac{P(n\mathbb{G}_n(s) = k_1, n(1-\mathbb{G}_n(t)) = k_2)}{P(n(1-GG_n(t)) = k_2)} \\
&= \frac{P(n\mathbb{G}_n(s) = k_1, n(\mathbb{G}_n(t)-\mathbb{G}_n(s)) = n-k_1-k_2, n(1-\mathbb{G}_n(t)) = k_2)}{P(n(1-GG_n(t)) = k_2)} \\
&= \frac{\frac{n!}{k_1!(n-k_1-k_2)!k_2!} s^{k_1}(t-s)^{n-k_1-k_2}(1-t)^{k_2}}{\frac{n!}{k_2!(n-k_2)!}(1-t)^{k_2}t^{k_2}} \\
&= \frac{(n-k_2)!}{k_1!(n-k_2-k_1)!}\left(\frac{s}{t}\right)^{k_1}\left(1-\frac{s}{t}\right)^{n-k_2-k_1}, \qquad 0 \le k_1 \le n-k_2.
\end{aligned}$$

That is, $(n\mathbb{G}_n(s)|n(1-\mathbb{G}_n(t))) \sim$ Binomial$(n\mathbb{G}_n(t),(s/t))$.
(c) It follow from the calculation in (b) that

$$E\{n\mathbb{G}_n(s)|\mathbb{G}_n(t)\} = n\mathbb{G}_n(t)\left(\frac{s}{t}\right), \qquad \text{almost surely}$$

or, equivalently, that

$$E\left\{\frac{\mathbb{G}_n(s)}{s}\Big|\mathbb{G}_n(t)\right\} = \frac{\mathbb{G}_n(t)}{t} \qquad \text{almost surely.}$$

6. (36 points)
(a) Suppose that $\epsilon_1,\ldots,\epsilon_n,\ldots$ are i.i.d. Rademacher random variables (i.e. $P(\epsilon_i = \pm 1) = 1/2$ for $i = 1,2,\ldots$. Let $X_{ni} = (i/n)\epsilon_i$ for $1 \le i \le n$. If $S_n \equiv \sum_{i=1}^n X_{ni}$, find constants σ_n such that $S_n/\sigma_n \to_d Z \sim N(0,1)$.
(b) Suppose we replace $\epsilon_1,\ldots,\epsilon_n$ in (a) by $Y_1,\ldots,Y_n$ i.i.d. with $E(Y_i) = 0$, $Var(Y_i) = \sigma^2$, and we replace $\{(i/n) : 1 \le i \le n\}$ in (a) with $\{a_{ni} : 1 \le i \le n\}$ with $\max_{i\le n} a_{ni}^2/\sum_1^n a_{nj}^2 \to 0$. Let $X_{ni} \equiv a_{ni}Y_i$ for $1 \le i \le n$ and define $S_n \equiv \sum_{i=1}^n X_{ni}$. Again find constants σ_n^2 such that $S_n/\sigma_n \to_d Z \sim N(0,1)$.

Solution: (a) Now $E(X_{ni}) = (i/n)E\epsilon_i = (i/n)\cdot 0 = 0$ and

$$\sigma_{ni}^2 = Var(X_{ni}) = (i/n)^2 \cdot Var(\epsilon_i) = (i/n)^2 \cdot 1 = (i/n)^2$$

for $1 \le i \le n$. Thus

$$\sigma_n^2 \equiv \sum_{i=1}^n \sigma_{ni}^2 = n^{-2}\sum_{i=1}^n i^2 = \frac{1}{n^2}\frac{n(n+1)(2n+1)}{6} \sim \frac{2n+1}{6}.$$

Then

$$\begin{aligned}
LF_n^\epsilon &= \frac{1}{\sigma_n^2}\sum_{i=1}^n E\{X_{ni}^2 1_{[|X_{ni}|\ge\epsilon\sigma_n]}\} \\
&= \frac{1}{\sigma_n^2}\sum_{i=1}^n (i/n)^2 E\{\epsilon_i^2 1_{[|\epsilon_i|\ge\epsilon\sigma_n/(i/n)]}\} \\
&\le \frac{1}{\sigma_n^2}\sum_{i=1}^n (i/n)^2 E\{\epsilon_i^2 1_{[|\epsilon_i|\ge\epsilon\sigma_n/\max_{1\le i\le n}(i/n)]}\} \\
&= \frac{1}{\sigma_n^2}\sum_{i=1}^n (i/n)^2 E\{\epsilon_1^2 1_{[|\epsilon_1|\ge\epsilon\sigma_n/1]}\} \\
&\qquad \text{since the } \epsilon_i' s \text{ are i.i.d.} \\
&\to 0 \qquad \text{as } n\to\infty
\end{aligned}$$

since $\epsilon\sigma_n \to \infty$ for every $\epsilon > 0$ while $|\epsilon_1| \le 1$ a.s. Thus the expectation is 0 exactly as soon as $\epsilon\sigma_n > 1$.

(b) In this slightly more general case $E(X_{ni}) = a_{ni}E(Y_i) = a_{ni}\cdot 0 = 0$ while

$$\sigma_{ni}^2 = Var(X_{ni}) = a_{ni}^2\sigma^2, \qquad \sigma_n^2 = \sum_1^n \sigma_{ni}^2 = \sigma^2\sum_1^n a_{ni}^2,$$

and hence we compute

$$\begin{aligned}
LF_n^\epsilon &= \frac{1}{\sigma_n^2}\sum_{i=1}^n E\{X_{ni}^2 1_{[|X_{ni}|\ge\epsilon\sigma_n]}\} \\
&= \frac{1}{\sigma_n^2}\sum_{i=1}^n a_{ni}^2 E\{Y_i^2 1_{[|Y_i|\ge\epsilon\sigma_n/|a_{ni}|]}\} \\
&\le \frac{1}{\sigma_n^2}\sum_{i=1}^n a_{ni}^2 E\{Y_i^2 1_{[|Y_i|\ge\epsilon\sigma_n/\max_{1\le i\le n}|a_{ni}|]}\} \\
&= \frac{1}{\sigma_n^2}\sum_{i=1}^n a_{ni}^2 E\{Y_1^2 1_{[|Y_1|\ge\epsilon\sigma_n/\max_{1\le i\le n}|a_{ni}|]}\} \\
&\qquad \text{since the } Y_i' s \text{ are i.i.d.} \\
&= \frac{1}{\sigma^2} E\{Y_1^2 1_{[|Y_1|\ge\epsilon\sigma_n/\max_{1\le i\le n}|a_{ni}|]}\} \\
&\to 0 \qquad \text{as } n\to\infty
\end{aligned}$$

by the dominated convergence theorem since $E(Y_1^2) < \infty$ and $\max_{1\le i\le n}|a_{ni}|/\sqrt{\sum_1^n a_{ni}^2} \to 0$. It follows from the Lindeberg - Feller CLT that $S_n/\sigma_n \to Z \sim N(0,1)$ with $\sigma_n^2 = \sigma^2\sum_1^n a_{ni}^2$.

Do either problem 7 or problem 8.

7. (30 points)
Suppose that $X_1, X_2, \ldots$ are i.i.d. Cauchy random variables with density function $f(x) = \pi^{-1}(1+x^2)^{-1}$. We showed in class that the characteristic function of each of these X_i's is $\phi(t) = e^{-|t|}$.
(a) If $a_i \in \mathbb{R}$ for $i = 1, \ldots, n$, find the characteristic function of $\sum_{i=1}^n a_i X_i$.
(b) Give conditions on the numbers $\{a_i : 1 \le i \le n\}$ which imply that $\sum_1^n a_i X_i \to_d$ something.
(c) What is the limiting distribution in (b)?

Solution: (a) Since the c.f. of a sum of independent random variables is the product of the c.f.'s of the individual variables we compute

$$\begin{aligned}
\phi_{\sum_1^n a_i X_i}(t) &= E\exp(it\sum_1^n a_i X_i) = \prod_{i=1}^n E\exp(ita_i X_i) \\
&= \prod_{i=1}^n \exp(-|a_i t|) = \prod_{i=1}^n \exp(-|a_i||t|) \\
&= \exp(-\sum_{i=1}^n |a_i||t|).
\end{aligned}$$

(b) If $\sum_{i=1}^n |a_i| \to A \in [0, \infty)$, then it follows from (a) that

$$\phi_{\sum_1^n a_i X_i}(t) \to \exp(-A|t|) \equiv \phi(t).$$

Since ϕ is continuous at 0, it follows from the Cramér - Lévy continuity theorem that ϕ corresponds to a bona-fide distribution function F and $\sum_1^n a_i X_i \to_d Y$ where Y has c.f. ϕ.
(c) Of course ϕ is just the c.f. of $Y \equiv AY_0$ where Y_0 is standard Cauchy (and Y is equal in distribution to each of the X_i's).

8. (30 points)
(a) Suppose that X is a Rademacher random variable; i.e. $P(X = \pm 1) = 1/2$. Find the characteristic function ϕ_X of X.
(b) Let X_1, X_2 be two independent Rademacher random variables. Find the characteristic function of $Z \equiv X_1 + X_2$.
(c) Use the two independent Rademacher random variables X_1, X_2 in (b) to define two independent Bernoulli random variables Y_1, Y_2. Use the results of (b) to calculate the characteristic function of $Y_1 + Y_2$.
(d) Let $Z_1, \ldots, Z_n$ be i.i.d. each with the same distribution as $Z = X_1 + X_2$ in (b). What is the characteristic function of $S_n \equiv \sum_{i=1}^n Z_i$?
(e) Show that the characteristic function you computed in (c) converges to a function φ that is not a characteristic function.
(f) Find a sequence of numbers a_n such that the characteristic function of S_n/a_n

converges to a proper characteristic function and $S_n/a_n \to_d$ something; identify "something".

Solution: (a) $\phi_X(t) = Ee^{itX} = (e^{it} + e^{-it})/2 = \cos(t)$.
(b) If X_1, X_2 are independent Rademacher rv's, then by (a)

$$\phi_Z(t) = Ee^{it(X_1+X_2)} = \{\cos(t)\}^2.$$

(c) If $X \sim$ Rademacher, then $Y \equiv (X+1)/2 \sim$ Bernoulli(1/2) since

$$\begin{aligned} P(Y = 1) = P((X+1)/2 = 1) = P(X = 2-1 = 1) = 1/2, \quad \text{and} \\ P(Y = 0) = P((X+1)/2 = 0) = P(X = -1) = 1/2. \end{aligned}$$

Thus $Y_1 \equiv (X_1+1)/2$ and $Y_2 \equiv (X_1+1)/2$ are independent Bernoulli (1/2) random variables and

$$\begin{aligned} Ee^{it(Y_1+Y_2)} &= (Ee^{itY_1})^2 = (Ee^{it(X_1+1)/2})^2 \\ &= (e^{it/2}Ee^{itX_1/2})^2 = e^{it}(\cos(t/2))^2 \\ &= \{(e^{it}+1)/2\}^2 \ \text{ by direct calculation.} \end{aligned}$$

(d) The characteristic function of $S_n = \sum_1^n Z_i$ is

$$\phi_{S_n}(t) = \{\phi_{Z_1}(t)\}^n = \{\cos(t)\}^{2n}.$$

(e) Now $\cos(t) = 1$ if t is an even multiple of π, $t = \pm 2k\pi$, $k = 0, 1, \ldots$, while $\cos(t) = -1$ if t is an odd multiple of π, $= \pm(2k+1)\pi$, $k = 0, 1, \ldots$. For all other values of t, $|\cos(t)| < 1$. Thus $\phi_{S_n}(t) \to 1_{\{\pm k\pi : k=0,1,\ldots\}}(t) \equiv \phi(t)$. Since this function ϕ is not a uniformly continuous function on $\mathbb{R}$, it is not the characteristic function of a proper d.f. F (or random variable X with all values finite).
(f) Now $E(Z_i) = E(X_1 + X_2) = 0 + 0 = 0$, while $Var(Z_i) = Var(X_1) + Var(X_2) = 1 + 1 = 2$. Thus by the ordinary CLT

$$\frac{S_n}{\sqrt{2n}} = \frac{\sum_1^n Z_i}{\sqrt{2n}} \to_d Z \sim N(0,1).$$

This implies that

$$\phi_{S_n/\sqrt{2n}}(t) = Ee^{itS_n/\sqrt{2n}} = \{\cos(t/\sqrt{2n})\}^{2n} \to \exp(-t^2/2).$$