Statistics 522, Final Exam

Wellner; 3/20/2013

1. (24 points). **Define** *three* of the following six terms:
   (a) An asymptotically tight sequence $\{X_n\}$ in $\mathbb{R}$ (or $\mathbb{R}^d$).
   (b) The Lévy metric (for convergence in distribution of distribution functions on $\mathbb{R}$).
   (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 bridge process $\mathbb{U}$ on $[0,1]$, and Brownian motion process $\mathbb{S}$ on $[0,\infty)$.
   (e) The conditional expectation $E(Y|\mathcal{D})$ of Y given a sub-σ-field $\mathcal{D}$.
   (f) The tail $\sigma-$field of a sequence of random variables $X_1, X_2, \ldots$.

2. (40 points). Give careful **statements** of *any four* of the following seven 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) The Mann-Wald or continuous mapping theorem.
   (c) Helly's selection theorem.
   (d) Any result connecting a Brownian motion process $\mathbb{S}$ on $[0,1]$ or $[0,\infty)$ to a Brownian bridge process $\mathbb{U}$ on $[0,1]$.
   (e) The Cramér - Lévy continuity theorem for characteristic functions.
   (f) A basic triangular array central limit theorem.
   (g) The Lindeberg-Feller central limit theorem.

3. (40 points) Suppose that $\underline{X} = (X_1, \ldots, X_d)$ is a random vector in $\mathbb{R}^d$ with distribution function

$$F_{\underline{X}}(\underline{x}) = P(\underline{X} \le \underline{x}) = P(X_1 \le x_1, \ldots, X_d \le x_d)$$

for all $\underline{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$.
(a) Define the characteristic function $\phi_{\underline{X}}(\underline{t})$ of $\underline{X}$ at $\underline{t} \in \mathbb{R}^d$.
(b) Let $\underline{a} \in \mathbb{R}^d$ be fixed and define the random variable Y by $Y = \underline{a}^T \underline{X} = \sum_{j=1}^d a_j X_j$ Express the characteristic function ϕ_Y in terms of $\phi_{\underline{X}}$, the characteristic function of the random vector $\underline{X}$.
(c) Now suppose that $\underline{X}_1, \underline{X}_2, \ldots$ is a sequence of random vectors in $\mathbb{R}^d$. The Cramér-Wold device says that $\underline{X}_n \to_d \underline{X}$ in $\mathbb{R}^d$ if and only if $\underline{a}^T \underline{X}_n \to_d \underline{a}^T \underline{X}$ in $\mathbb{R}$ for all $\underline{a} \in \mathbb{R}^d$. Sketch a proof of this result using characteristic functions and (b).

4. (40 points)
(a) Suppose that $X_1, X_2, \ldots$ are i.i.d. random variables with $E(X_1) = 0$, $Var(X_1) = 1$, and let $X_{n,i} \equiv a_{n,i} X_i$ for $i = 1, \ldots, n$ where $\{a_{n,i} : 1 \le i \le n\}$ satisfy $\sum_{i=1}^n a_{n,i}^2 \to v^2$ and $\max_{1 \le i \le n} |a_{n,i}| \to 0$. Show that $S_n \equiv \sum_{i=1}^n X_{n,i}$ satisfies $S_n \to_d vZ \sim N(0, 1)$.
(b) Suppose that $X_1, X_2, \ldots$ are i.i.d. as in (a) and let $a_{n,i} = n^{-1/2}(i/n)$ for $i \in \{1, \ldots, n\}$. Show that the hypotheses of (a) are satisfied for some v^2 (find the particular value in this case) and hence that $S_n = n^{-1/2} \sum_{i=1}^n (i/n) X_i \to_d N(0, v^2)$.
(c) Now suppose that the X_i's are as in (a), but that $a_{n,i} = n^{-1/2}(i/n)^\alpha$ for some $\alpha \in \mathbb{R}$. For what values of α does it hold that $S_n \to_d N(0, v_\alpha^2)$ for some $v_\alpha^2 < \infty$? For the values for which this holds, compute v_α^2 as a function of α.

5. (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.)

6. (40 points).
Let $X \sim N(\mu, 1)$ with $\mu > 0$.
(a) Suppose that ϵ is a Rademacher random variable independent of X; i.e. $P(\epsilon = \pm 1) = 1/2$. Compute the distribution function G of the random variable $Y = \epsilon X$ and find its density function g. Give rough plots of g and G for $\mu = 3$.
(b) Now suppose that $\epsilon \equiv 2 \cdot 1_{[X \ge \mu]} - 1$. Find the distribution of ϵ. Is ϵ independent of X?
(c) Find the distribution function H of $W = \epsilon X$ when ϵ is as defined in (b) with $\mu = 0$ and with $X \sim N(0, 1)$. Does it have a density h with respect to Lebesgue measure? If so, compute it. Is the distribution function H the same as the distribution G you found in (a) specialized to have $\mu = 0$?
(d) What is the characteristic function of $X \sim N(\mu, 1)$? What is the characteristic function of $Y \equiv \epsilon X$ with $X \sim N(\mu, 1)$ and ϵ as in (a)? Consider also the characteristic function of W with W as defined in (c). Do these all agree when $\mu = 0$?

7. (40 points)
Suppose that $X_n \sim N(\mu_n, \sigma_n^2)$ where $\mu_n \in \mathbb{R}$, $\sigma_n^2 > 0$.
(a) Suppose that $\{\mu_n\}$ and $\{\sigma_n^2\}$ are bounded. Show that $\{X_n\}$ (or the corresponding sequence of distribution functions $\{F_n\}$ of $\{X_n\}$) are tight.
(b) If $\{\mu_n\}$ is unbounded, show that $\{X_n\}$ is not tight.
(c) If $\{\mu_n\}$ is bounded but $\{\sigma_n^2\}$ is unbounded, show that $\{X_n\}$ is not tight.
(d) Suppose that $\mu_n = (-1)^n a$ with $a > 0$ and $\sigma_n^2 = 1 + \exp(-\pi n(-1)^n)$ for $n \geq 1$. Is there any tight subsequence $\{X_{n'}\}$ of $\{X_n\}$? Identify all the subsequences $\{X_n'\}$ of $\{X_n\}$ which converge in sub-distribution or in distribution.