Statistics 522, Midterm Exam Solutions

Wellner; 2/14/2004

1. (24 points). **Define** *three* of the following six terms:
   (a) A tight probability measure on a metric space (M, d).
   (b) Convergence in distribution of a sequence of measures $\{P_n\}$ on a metric space to a measure P.
   (c) The class of bounded Lipschitz functions $BL(M)$ on a metric space (M, d)?
   (d) The Lévy distance $\lambda(F, G)$ between two distribution functions F and G on $\mathbb{R}$.
   (e) The Prohorov distance $\rho(P, Q)$ between two probability measures P and Q on a metric space (M, d).
   (f) A standard Brownian motion process on $[0, 1]$.

   Solution: (a) A measure P is tight if for every $\epsilon > 0$ there is a compact set $K = K_\epsilon$ such that $P(K) > 1 - \epsilon$.
   (b) $P_n \rightarrow_d P$ if and only if $P_n f \rightarrow P f$ for all bounded continuous functions f; i.e. for all $f \in C_b(M)$.
   (c) A function f is bounded Lipschitz if $\sup_x |f(x)| < \infty$ and

   $$\sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)} < \infty .$$

   (d) The Lévy distance $\lambda(F, G)$ between two df's F and G is defined by

   $$\lambda(F, G) = \inf\{\epsilon > 0 : \ F(x - \epsilon) - \epsilon \le G(x) \le F(x + \epsilon) + \epsilon \ \text{ for all } x\} .$$

   (e) The Prohorov distance $\rho(P, Q)$ between two probability measures P and Q is defined by

   $$\rho(P, Q) \equiv \inf\{\epsilon > 0 : \ P(B) \le Q(B^\epsilon) + \epsilon \ \text{ for all } \ B \in \mathcal{M}\} .$$

2. (24 points). Give careful **statements** of *any two* of the following five theorems or results:
   (a) The law of the iterated logarithm for i.i.d. random variables $X_1, X_2, \ldots$ with $E(X_1) = 0$ and $Var(X_1) = \sigma^2$.
   (b) Bennett's inequality for bounded random variables.
   (c) The Lindeberg-Feller central limit theorem for a triangular array of row-wise independent random variables with $E(X_{n,i}) = 0$ and $Var(X_{n,i}) = \sigma_{n,i}^2 < \infty$, $i = 1, \ldots, n$.
   (d) Donsker's theorem for the partial sum process $\mathbb{S}_n$ of i.i.d. mean 0, finite variance

random variables.
(e) The theorem of Prohorov and Le Cam relating tightness to relative compactness.

Solution: See LIL handout and Chapter 11.

3. (24 points). Let $\phi(z) = (2\pi)^{-1/2}\exp(-z^2/2)$ be the standard Normal density function and let $\Phi(z) = \int_{-\infty}^{z}\phi(t)dt$ be the standard Normal distribution function.
(a) Prove the upper half of Mills' ratio inequality: i.e. show that

$$1 - \Phi(z) \le \frac{\phi(z)}{z} \qquad \text{for} \quad z > 0\,.$$

(b) Explain briefly how this is used to prove a part of the law of the iterated logarithm for sums of i.i.d. $N(0,1)$ random variables.

Solution: (a) Note that

$$\begin{aligned} 1 - \Phi(z) &= \int_z^\infty \phi(t)dt \\ &\le \int_z^\infty \frac{t}{z}\phi(t)dt \\ &= \frac{1}{z}\int_z^\infty -\phi'(t)dt \qquad \text{since} \quad \phi'(t) = -t\phi(t) \\ &= \frac{1}{z}(-\phi(t)\Big|_z^\infty = \frac{1}{z}\phi(z)\,. \end{aligned}$$

(b) When $X_1, \ldots, X_n$ are i.i.d. $N(0,1)$, then $S_n = X_1 + \cdots + X_n \sim N(0,n)$ and $S_n/\sqrt{n} \sim N(0,1)$. Then for $n_k \equiv [\alpha^k]$ with $\alpha > 1$ and $\epsilon > 0$,

$$\begin{aligned} P(A_k) &\equiv P(\max_{n_k \le n \le n_{k+1}} \frac{S_n}{\sqrt{2n\log\log n}} \ge 1+\epsilon) \\ &\le P(\sup_{n_k \le n \le n_{k+1}} S_n \ge (1+\epsilon)\sqrt{2n_k \log\log n_k}) \\ &\le 2P(S_{n_{k+1}} \ge (1+\epsilon)\sqrt{2n_k\log\log n_k}) \\ &= 2P(\frac{S_{n_{k+1}}}{\sqrt{n_{k+1}}} \ge (1+\epsilon)\sqrt{\frac{n_k}{n_{k+1}}}\sqrt{2\log\log n_k}) \\ &\le 2\exp\left(-\frac{(1+\epsilon)^2}{\alpha}\log\log\alpha^k\right) \\ &\le 2\frac{1}{k^\tau} \qquad \text{with} \quad \tau > 1 \end{aligned}$$

if $\alpha > 1$ is sufficiently close to 1. Here we used Mills' ratio inequality in the next to last inequality. Thus by Borel-Cantelli we conclude that $P(A_k\ i.o.) = 0$, and hence that

$$\limsup_n \frac{S_n}{\sqrt{2n\log\log n}} \le 1+\epsilon \qquad \text{a.s.}\,.$$

4. (30 points). Suppose that $X_n \sim \text{Poisson}(n)$; thus $X_n \stackrel{d}{=} \sum_1^n Y_i$ where Y_i are independent Poisson(1) random variables.
(a) Compute

$$E\left\{\left(\frac{X_n - n}{\sqrt{n}}\right)^-\right\}$$

explicitly and thereby show that this expectation equals $n^{n+1}e^{-n}/(\sqrt{n}n!)$.
(Recall that $Y^- = -Y1_{[Y \le 0]}$.)
(b) Show that $(X_n - n)/\sqrt{n} \to_d Z \sim N(0,1)$.
(You may appeal to one of our theorems.)
(c) Use (b) and a uniform integrability argument to show that $E[(X_n - n)^-/\sqrt{n}] \to E[Z^-]$.
(d) Compute $E[Z^-]$.
(e) Combine (a) - (d) to show that $n! \sim \sqrt{2\pi n}(n/e)^n$; i.e. $n!/(\sqrt{2\pi n}(n/e)^n) \to 1$.

Solution: (a) Since $X_n \sim \text{Poisson}(n)$,

$$\begin{aligned} E\left\{\left(\frac{X_n - n}{\sqrt{n}}\right)^-\right\} &= \sum_{k=0}^{n} \frac{n-k}{\sqrt{n}} e^{-n} \frac{n^k}{k!} \\ &= \frac{e^{-n}}{\sqrt{n}}\left\{\frac{n}{0!} + \sum_{k=1}^{n}\left(\frac{n^{k+1}}{k!} - \frac{n^k}{(k-1)!}\right)\right\} \\ &= \frac{e^{-n}n^{n+1}}{\sqrt{n}n!} \end{aligned}$$

upon noting that the sum telescopes to just the last term.
(b) Now $n^{-1/2}(X_n - n) = \sqrt{n}(\overline{Y}_n - 1) \to_d Z \sim N(0,1)$ by the CLT.
(c) Since $Z_n \equiv n^{-1/2}(X_n - n) \to_d Z$, it follows by continuous mapping that $Z_n^- \to_d Z^-$. Since $Z_n^- \le |Z_n|$ where $E(Z_n^2) = n^{-1}Var(X_n) = 1$, it follows that $E(Z_n^-) \to E(Z^-)$.
(d) Now

$$E(Z^-) = E(Z^+) = (2\pi)^{-1/2} \int_0^\infty z\phi(z)dz = (2\pi)^{-1/2}.$$

Combining (a) - (d) yields

$$\frac{n^{n+1}e^{-n}}{\sqrt{n}n!} \to \frac{1}{\sqrt{2\pi}}.$$

This yields

$$\frac{n!}{\sqrt{2\pi n}(n/e)^n} \to 1.$$

5. (32 points). Suppose that $X_1, X_2, \ldots$ are i.i.d. random variables with $E(X_1) = 0$, $Var(X_1) = 1$, and let $S_n = X_1 + \cdots + X_n$ for $n \geq 1$. Let $\mathbb{S}_n$ denote the partial sum process,

$$\mathbb{S}_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} X_i, \qquad 0 \leq t \leq 1 .$$

(a) Rewrite $n^{-3/2} \sum_{i=1}^{n} S_i$ as a (continuous) function g of $\mathbb{S}_n$.
(b) Use the result of (a) and Donsker's theorem to show that $n^{-3/2} \sum_{i=1}^{n} S_i \rightarrow_d$ some Y, and find the distribution of Y.
(c) Use summation by parts and the fact that the X_i's are i.i.d. to show that $n^{-3/2} \sum_{i=1}^{n} S_i \stackrel{d}{=} n^{-3/2} \sum_{j=1}^{n} jX_j$.
(d) Use the result of (c) together with the Lindeberg-Feller CLT to show that $n^{-3/2} \sum_{i=1}^{n} S_i \rightarrow_d Y$ where Y has the same distribution as was found in (b).

Solution: (a) Note that

$$\begin{aligned} n^{-3/2} \sum_{i=1}^{n} S_i &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{(i-1)/n}^{i/n} S_i dt = \sum_{i=1}^{n} \int_{(i-1)/n}^{i/n} \mathbb{S}_n(t) dt \\ &= \int_0^1 \mathbb{S}_n(t) dt \equiv g(\mathbb{S}_n) \end{aligned}$$

where $g(x(t)) \equiv \int_0^1 x(t) dt$ for $x \in D[0,1]$.
(b) Since g is continuous with respect to the supremum norm, it follows from Donsker's theorem that $g(\mathbb{S}_n) \rightarrow_d g(\mathbb{S}) = \int_0^1 \mathbb{S}(t) dt$.
(c) Writing $S_i = \sum_{j=1}^{i} X_j$ and interchanging the order of summation gives:

$$\begin{aligned} \sum_{i=1}^{n} S_i &= \sum_{i=1}^{n} \sum_{j=1}^{i} X_j = \sum_{j=1}^{n} X_j \sum_{i=1}^{n} 1_{[j \leq i]} \\ &= \sum_{j=1}^{n} (n-j+1) X_j = nX_1 + \cdots + 1 \cdot X_n \\ &\stackrel{d}{=} nX_n + \cdots 1 \cdot X_1 = \sum_{j=1}^{n} jX_j . \end{aligned}$$

Here we used the fact that the X_i's are i.i.d. to get the equality in distribution.
(d) Let $X_{n,j} \equiv jX_j$ for $j = 1, \ldots, n$. Then $\sigma_{n,j} = j^2 Var(X_j) = j^2$, $j = 1, \ldots, n$, and

$$\sigma_n^2 = \sum_1^n \sigma_{n,j}^2 = \sum_1^n j^2 = \frac{n(n+1)(2n+1)}{6},$$

and if the Lindeberg condition

$$L_n(\epsilon) = \sum_1^n E(X_{n,j}^2 1_{[|X_{n,j}|>\epsilon\sigma_n]}) \to 0 \qquad \text{for every} \quad \epsilon > 0 \tag{0.1}$$

holds, then using (c),

$$\frac{\sum_1^n S_i}{\sigma_n} \stackrel{d}{=} \frac{\sum_1^n X_{n,i}}{\sigma_n} \to_d Z \sim N(0,1)\,;$$

that is, since $\sigma_n^2/n^3 \to 1/3$,

$$n^{-3/2}\sum_{i=1}^n S_i \to_d \frac{1}{\sqrt{3}} Z \sim N(0,1/3)\,.$$

This agrees with the result in (b) since $Y \sim \int_0^1 \mathbb{S}(t)dt$ is a Normal random variable (linear combinations of normals are normal) with $E(Y) = \int_0^1 E(\mathbb{S}(t))dt = 0$ and

$$\begin{aligned}
E(Y^2) &= E\left(\int_0^1 \mathbb{S}(s)ds \int_0^1 \mathbb{S}(t)dt\right) \\
&= \int_0^1\int_0^1 E\{\mathbb{S}(s)\mathbb{S}(t)\}dsdt \\
&= \int_0^1\int_0^1 s\wedge t\; dsdt \\
&= 2\int_0^1\int_0^t sdsdt = \int_0^1 t^2 dt \\
&= 1/3\,,
\end{aligned}$$

so $Y \sim N(0,1/3)$. It remains only to check that (0.1) holds. But

$$\begin{aligned}
L_n(\epsilon) &= \frac{1}{\sigma_n^2}\sum_{j=1}^n E\{X_{n,j}^2 1_{[|X_{n,j}|>\epsilon\sigma_n]}\} \\
&= \frac{1}{\sigma_n^2}\sum_{j=1}^n E\{j^2 X_j^2 1_{[|jX_j|>\epsilon\sigma_n]}\} \\
&\le \frac{1}{\sigma_n^2}\sum_{j=1}^n j^2 E\{X_1^2 1_{[|X_1|>\epsilon\sigma_n/\max_{1\le j\le n} j]}\} \\
&= E\{X_1^2 1_{[|X_1|>\epsilon\sigma_n/n]}\} \\
&\to 0 \qquad \text{by the DCT}
\end{aligned}$$

since $\sigma_n/n \to \infty$.