Statistics 521, Midterm Exam Solutions

Wellner; 11/04/2019

1. (24 points). **Define** *three* of the following five terms:
   (a) Convergence in probability of a sequence of random variables $\{X_n\}$ defined on a probability space $(\Omega, \mathcal{A}, P)$.
   (b) A measurable function $X : \Omega \to \Omega'$ where $(\Omega, \mathcal{A})$ and $(\Omega', \mathcal{A}')$ are measurable spaces.
   (c) The induced distribution P_X of a random variable $X : \ (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B})$.
   (d) A *simple function* defined on a measurable space $(\Omega, \mathcal{A})$.
   (e) Given a measurable space $(\Omega, \mathcal{A})$ define a $\lambda-$system of subsets of Ω and a $\pi-$system of subsets of Ω.

   Solution: See course notes.

2. (24 points). Give careful **statements** of *three* of the following five theorems or results:
   (a) The monotone convergence theorem.
   (b) A theorem relating Lebesgue-Stieltjes measures μ to (generalized) distribution functions on $\mathbb{R}$
   (c) A result connecting a $\pi-$system $\mathcal{C}$ and a $\lambda-$system $\mathcal{D}$ where $\mathcal{C}$ and $\mathcal{D}$ are collections of subsets of some set Ω.
   (d) Hölder's inequality.
   (e) The Caratheodory extension theorem.

   Solution: See course notes.

3. (30 points).
   Show that if $X_n \to_p X$ then $X_n \to_d X$.

   Solution 1: This is Proposition 4.1 on pages 33-34 of the course notes, but here is the proof again just for the record. Let $\epsilon > 0$. Then

$$\begin{aligned} F_n(t) &= P(X_n \le t) \\ &= P([X_n \le t] \cap [|X_n - X| \le \epsilon]) + P([X_n \le t] \cap [|X_n - X| > \epsilon]) \\ &\le P(X \le t + \epsilon) + P(|X_n - X| > \epsilon) \le F(t + \epsilon) + \epsilon \end{aligned}$$

   for all $n \ge$ some N_ϵ. Also

$$\begin{aligned} F(t - \epsilon) &= P(X \le t - \epsilon) \\ &= P([X \le t - \epsilon] \cap [|X_n - X| \le \epsilon]) + P([X \le t - \epsilon] \cap [|X_n - X| > \epsilon]) \\ &\le P(X_n \le t) + P(|X_n - X| > \epsilon). \end{aligned}$$

   Rearranging the inequality in the last display yields

$$F_n(t) \ge F(t - \epsilon) - P(|X_n - X| > \epsilon) \ge F(t - \epsilon) - \epsilon$$

for all $n \geq$ some N'_ϵ. Thus for $n \geq \max\{N_\epsilon, N'_\epsilon\}$ we have

$$F(t-\epsilon) - \epsilon \leq \underline{\lim} F_n(t) \leq \overline{\lim} F_n(t) \leq F(t+\epsilon) + \epsilon.$$

If t is a continuity point of F, then letting $\epsilon \to 0$ in the last display yields $F_n(t) \to F(t)$. Thus $F_n \to_d F$ (or $X_n \to_d X$).

Solution 2: Here is a second solution of this problem which might be instructive. Suppose that $X_n \to_p X$. Then if g is continuous it follows that $g(X_n) \to_p g(X)$. (This is easily proved by arguing along subsequences n' and n'' for which almost sure convergence holds.) Thus $Y_n \equiv g(X_n) - g(X) \to_p 0$. Moreover,

$$|Y_n| = |g(X_n) - g(X)| \leq 2\|g\|_\infty \equiv 2 \sup_{x \in \mathbb{R}} |g(x)| < \infty$$

since g is bounded. Thus Y_n is uniformly integrable and by Vitali's theorem $E|Y_n| \to 0$. But then

$$|Eg(X_n) - Eg(X)| = |E(g(X_n) - g(X))| \leq E|g(X_n) - g(X)| = E|Y_n| \to 0.$$

This implies that $Eg(X_n) \to Eg(X)$ for all $g \in C_B(\mathbb{R})$. Thus $X_n \to_d X$.

Do **either** problem 4 **or** problem 5.

4. (36 points).
   (a) State Jensen's inequality.
   (b) Use Jensen's inequality to prove that $\left(\prod_{i=1}^n x_i\right)^{1/n} \leq \{x_1 + \cdots + x_n\}/n$ for any real numbers $x_i \geq 0$ for $i \in \{1, \ldots, n\}$. (Be careful about the case with some $x_j = 0$.) When does equality occur?
   (c) Use Jensen's inequality to prove that $\exp E(\log X) \leq E(X)$ for a non-negative random variable X, assuming that $E(\log X)^+ = E(1_{[X \geq 1]} \log X) < \infty$. When does equality occur?

Solution: (a) If $g : \mathbb{R} \to \mathbb{R}$ is convex and $E|X| < \infty$, then $Eg(X) \geq g(E(X))$.
(b) First note that if any $x_j = 0$, then $\prod_{i=1}^n x_i = 0$ and the inequality is trivially true. Thus we may assume that $x_i > 0$ for all $1 \leq i \leq n$. Then since $x \mapsto \log x$ is concave

$$\begin{aligned} \log\left(\prod_{i=1}^n x_i\right)^{1/n} &= \frac{1}{n}\log\left(\prod_{i=1}^n x_i\right) = \frac{1}{n}\sum_{i=1}^n \log x_i \\ &\leq \log\left(\frac{1}{n}\sum_{i=1}^n x_i\right). \end{aligned}$$

Taking the exponential of both sides yields the claimed inequality. This is exactly the arithmetic mean - geometric mean inequality. Since $\log x$ is strictly concave, equality occurs if and only if all the x_i's are equal (and equal to the arithmetic mean $\overline{x}_n$.
(c) Similarly to the proof in (b), if $P(X = 0) = 0$, then since $\log x$ is concave Jensen's inequality yields

$$E(\log X) \leq \log(E(X)),$$

and exponentiating both side yields the claimed inequality, $\exp(E \log X) \leq E(X)$. Equality occurs if and only if $P(X = E(X)) = 1$; i.e. X is degenerate at some $x_0 = E(X)$.

5. (36 points).
   (a) State Markov's inequality.
   (b) Use Markov's inequality or a special case to give a bound for $P(|\overline{X}_n - \mu| \geq t)$ when $X_1, \ldots, X_n$ are independent and identically distributed with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2 < \infty$ for all i.
   (c) Does the bound in (b) show that $\overline{X}_n \to_p \mu$ under these assumptions?

   Solution: (a) Markov's inequality says that for any $r > 0$ and $t > 0$,

$$P(|X| \geq t) \leq \frac{E|X|^r}{t^r}.$$

   (b) Here we apply Markov's inequality with $r = 2$ and $Y \equiv |\overline{X}_n - \mu|$. Then

$$\begin{aligned} P(|\overline{X}_n - \mu| \geq t) &= P(Y \geq t) \leq \frac{EY^2}{t^2} = \frac{E(\overline{X}_n - \mu)^2}{t^2} \\ &= \frac{Var(\overline{X}_n)}{t^2} = \frac{\sigma^2/n}{t^2} \\ &= \frac{\sigma^2}{nt^2}. \end{aligned}$$

   (c) It follows from (b) that $P(|\overline{X}_n - \mu| \geq t) = \sigma^2/(nt^2) \to 0$ as $n \to \infty$ for every $t > 0$. Thus $\overline{X}_n \to_p \mu$. In fact this argument shows that $\sqrt{n}|\overline{X}_n - \mu| = O_p(1)$; i.e. taking $t = M/\sqrt{n}$ in the above inequalities, we find that

$$P(\sqrt{n}|\overline{X}_n - \mu| \geq M) = P(|\overline{X}_n - \mu| \geq M/\sqrt{n}) \leq \frac{\sigma^2}{M^2},$$

   and therefore

$$\overline{\lim}_n P(\sqrt{n}|\overline{X}_n - \mu| \geq M) \leq \frac{\sigma^2}{M^2} \to 0 \quad \text{as} \quad M \to \infty.$$

6. (30 points).
   (a) Give an example of a sequence of measurable functions (or random variables) $\{X_n\}$ defined on a probability space $(\Omega, \mathcal{A}, P)$ (which you should make explicit) for which $X_n \to_{a.s.} 0$ but $E(X_n) \to 0, 1, \text{ or } +\infty$ depending on the value of some number $c > 0$.
   (b) Give an example of a sequence of measurable functions (or random variables) $\{X_n\}$ defined on a probability space $(\Omega, \mathcal{A}, P)$ (which you should make explicit) for which $X_n \to_p 0$, but $X_n \not\to_{a.s.} 0$.

   Solution:
   (a) Let $U \sim \text{Uniform}(0,1)$ on the probability space $([0,1], \mathcal{B}_{[0,1]}, P)$ where P is the uniform distribution (or Lebesgue measure). For $c > 0$, let $X_n \equiv n^c 1_{(1/(n+1),1/n]}(U)$

Then $X_n \to_{a.s.} 0$ for every $c > 0$ (since $1/n < U(\omega) = \omega$ eventually for every $U(\omega) > 0$, and hence $X_n(\omega) = 0$ for all $n > 1/\omega \equiv N_\omega$). But then

$$E(X_n) = n^c(n^{-1} - (n+1)^{-1}) = n^c/(n(n+1)) \to \begin{cases} 0, & \text{if } c < 2, \\ 1, & \text{if } c = 2, \\ \infty, & \text{if } c > 2. \end{cases}$$

(b) Let $U \sim$ Uniform$(0,1)$ on the probability space $((0,1), \mathcal{B}_{[0,1]}, P)$ where P is as in (a). Let $X_{m,k} \equiv 1_{((k-1)/2^m, k/2^m]}(U)$ for $1 \le k \le 2^m$ and $m \ge 1$. Let $Y_n \equiv Y_{2^m+k} \equiv X_{m,k}$. Then $Y_n \to_p 0$ as $n \to \infty$, but $P(Y_n = 1 \text{ i.o.}) = P(U \in (0,1]) = 1$, so $Y_n \not\to_{a.s.} 0$.

7. (30 points).
   (a) Suppose that X is a non-negative measurable function on a measurable space $(\Omega, \mathcal{A})$. Give an explicit sequence of simple functions X_n satisfying $X_n \nearrow X$.
   (b) Now suppose that $(\Omega, \mathcal{A}) = ((0,1), \mathcal{B}_{(0,1)})$, and that we give this measurable space the Lebesgue measure λ, which we call P since it is a probability measure on this $(\Omega, \mathcal{A}, P)$. Suppose that $X(\omega) = -\log(\omega)$ for $\omega \in (0,1)$.

   For the simple functions X_n as given in (a), evaluate

$$\lim_{n\to\infty} \int X_n dP = \lim_{n\to\infty} E(X_n).$$

   (c) Find the (induced) distribution function $F = F_X$ of X on $\mathbb{R}$.

Solution:

(a) A sequence of simple functions converging monotonically to $X \ge 0$ is given by

$$X_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} 1_{[\frac{k-1}{2^n} \le X < \frac{k}{2^n}]} + n 1_{[X \ge n]}.$$

(b) For the simple functions in (a), it follows by the monotone convergence theorem that

$$0 \le \int X_n dP \nearrow \int X dP = \int_0^1 -\log(\omega) d\omega.$$

By the change of variables $v = -\log(\omega)$ we find that

$$\int_0^1 -\log(\omega) d\omega = \int_0^\infty v e^{-v} dv = 1$$

since the mean of the exponential(1) distribution is 1 (or since $\int_0^\infty v e^{-v} dv = \Gamma(2) = 1! = 1$ where $\Gamma(r) \equiv \int_0^\infty v^{r-1} e^{-v} dv$ is the Gamma function).

(c) The induced distribution of X on $\mathbb{R}$ is

$$P(X \le x) = P(-\log\omega \le x) = P(\{\omega \in (0,1) : \ \omega \ge e^{-x}\}) = 1 - e^{-x}$$

in agreement with the computation in (b); this is the standard exponential distribution.