Statistics 521, Midterm Exam Solutions

Wellner; 11/16/2016

1. (24 points). **Define** *three* of the following five terms:
   (a) $\limsup A_n$ for a sequence of events $\{A_n\}$, and give the common (intuitive) abbreviation for this set.
   (b) A measurable function $X : \Omega \to \Omega'$ where $(\Omega, \mathcal{A})$ and $(\Omega', \mathcal{A}')$ are measurable spaces.
   (c) A *simple function* defined on a measurable space $(\Omega, \mathcal{A})$.
   (d) The Lebesgue integral $\int X d\mu$ of a (real-valued) measurable function X defined on a measure space $(\Omega, \mathcal{A}, \mu)$.
   (e) A Lebesgue - Stieltjes measure on the real line $\mathbb{R}$.

   Solution: See PfS, chapters 1 - 3.

2. (20 points). Give careful **statements** of *three* of the following five theorems or results:
   (a) Fatou's lemma
   (b) A theorem relating Lebesgue-Stieltjes measures to (generalized) distribution functions.
   (c) A theorem relating convergence in measure ($\to_\mu$) to convergence almost everywhere ($\to_{a.e.}$).
   (d) Liapunov's inequality.
   (e) The Caratheodory extension theorem.

   Solution: See PfS, chapters 1 - 3.

3. (36 points).
   (a) State Hölder's inequality for random variables X and Y.
   (b) State Jensen's inequality, and prove that $\left(\prod_{i=1}^n x_i\right)^{1/n} \le \{x_1 + \cdots + x_n\}/n$ for any real numbers $x_i \ge 0$ for $i \in \{1, \ldots, n\}$. (Be careful about the case with some $x_j = 0$.)
   (c) State Markov's inequality.
   (d) Use Markov's inequality or a special case to give a bound for $P(|\overline{X}_n - \mu| \ge t)$ 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. Does the resulting bound show that $\overline{X}_n \to_p \mu$ under these assumptions?

   Solution: (a) Let $r, s > 1$ satisfy $r^{-1} + s^{-1} = 1$. Then $E|XY| \le \{E|X|^r\}^{1/r}\{E|Y|^s\}^{1/s}$.
   (b) If $E|X| < \infty$, $g : I \to \mathbb{R}$ is convex on I, and $E(X) \in I^0$, the interior of I, then $g(EX) \le Eg(X)$. For the inequality $\left(\prod_{i=1}^n x_i\right)^{1/n} \le \{x_1 + \cdots + x_n\}/n$, note that the left side is 0 if any $x_i = 0$, and then the inequality holds trivially. Thus we may

assume that $x_i > 0$ for all i. But then with $P(X = x_i) = 1/n$ for $i = 1, \ldots, n$,

$$\begin{aligned} \log\left(\prod_{i=1}^n x_i\right)^{1/n} &= \frac{1}{n}\log\prod_{i=1}^n x_i = \frac{1}{n}\sum_{i=1}^n \log x_i = E\{\log X\} \\ &\le \log E(X) \quad \text{since} \quad y \mapsto \log(y) \quad \text{is concave} \\ &= \log\left(n^{-1}(x_1 + \cdots + x_n)\right). \end{aligned}$$

Taking the exponential of both sides yields the claimed inequality.
(c) $P(|Y| \ge \lambda) \le \lambda^{-r} E|Y|^r$ for all $\lambda > 0$.
(d) Markov's inequality applied with $r = 2$ and $Y \equiv \overline{X}_n - \mu$ yields

$$\begin{aligned} P(|\overline{X}_n - \mu| \ge \lambda) &\le \frac{E(\overline{X}_n - \mu)^2}{\lambda^2} = \frac{Var(\overline{X}_n)}{\lambda^2} = \frac{\sigma^2/n}{\lambda^2} \\ &= \frac{\sigma^2}{n\lambda^2} \to 0 \end{aligned}$$

for every $\lambda > 0$. This implies that $\overline{X}_n - \mu \to_p 0$; equivalently $\overline{X}_n \to_p \mu$.

4. (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 $E(X_n) \to 0, 1,$ 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 $X_n \not\to_{a.s.} 0$.
(c) Give an example of a sequence of random variables $\{X_n\}$, X which satisfy $X_n \to_d X$ but for which $X_n \not\to_p X$ and $X_n \not\to_{a.s.} X$.

Solution:
(a) Let $U \sim$ 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 $P(Y_n(U) = 1 i.o.) = P(U \in (0,1]) = 1$, so $Y_n \not\to_{a.s.} 0$.
(c) Suppose that $X_1, X_2, \ldots$ are independent random variables with any fixed distribution function F on $\mathbb{R}$. Then $F_{X_n}(x) = P(X_n \le x) = F(x)$ for every n and x, so $X_n \to_d X$ with distribution function F. But $X_n \not\to_p X$ and $X_n \not\to_{a.s.} X$.

5. (42 points).
   (a) State the definition of $X_n \to_d X$ for random variables X_n and X.
   (b) State the Helly-Bray theorem.
   (c) What is the resulting equivalent formulation of $X_n \to_d X$ which results as a corollary of the Helly-Bray theorem.
   (d) State the Mann-Wald theorem.
   (e) Suppose that $X_n \sim$ Uniform on$\{1/n, 2/n, \cdots, n/n = 1\}$; i.e. $P(X_n = k/n) = 1/n$ for $k \in \{1, \ldots, n\}$. Show that $X_n \to_d X$ where $X \sim$ Uniform$(0,1)$.
   (f) Find a sequence of rv's $\{Y_n\}$ and Y, all on a common probability space, such that $Y_n \to_{a.s.} Y$.

Solution:
(a) $X_n \to X$ if $F_n(x) \equiv P(X_n \le x) \to P(X \le x) = F_X(x)$ for all $x \in C_{F_X}$.
(b) Helly-Bray: If $X_n \to_d X$ and g is bounded and continuous a.s. P_X, the distribution of X, then $Eg(X_n) \to Eg(X)$. On the other hand, if $Eg(X_n) \to Eg(X)$ for all bounded and continuous functions g, then $X_n \to_d X$.
(c) By virtue of the two parts of the Helly-Bray theorem we may define $X_n \to X$ is equivalent to $Eg(X_n) \to Eg(X)$ for all bounded and continuous functions g.
(d) Mann-Wald: if $X_n \to_d X$ and g is continuous a.s. P_X, then $g(X_n) \to_d g(X)$.
(e) $F_n(x) = P(X_n \le x) = \lfloor nx \rfloor / n$ for $0 \le x \le 1$; for example, for $x = k/n$, $F_n(k/n) = k/n$ for $k \in \{1, \ldots, n\}$. But then $F_n(x) \to x$ for every $x \in [0,1]$; i.e. $X_n \to_d X$ where $F(x) = P(X \le x) = 0 \vee (x \wedge 1)$.
(f) The inverse distribution functions F_n^{-1} of the F_n's in (e) are given by $F_n^{-1}(u) = k/n = \lceil nx \rceil / n$ for $(k-1)/n < u \le k/n$, while $F^{-1}(u) = u$ for $0 \le u \le 1$. Then with $U \sim$ Uniform$(0,1)$ we have $Y_n \stackrel{d}{=} X_n$ and $Y \equiv U \stackrel{d}{=} X$ where X_n and X are as in (e), but now $Y_n = F_n^{-1}(U) \to F^{-1}(U) \equiv Y$, and in particular $Y_n \to_{a.s.} Y$.

6. (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 $(\Omega, \mathcal{A})$. Suppose that $X(\omega) = -\log(\omega)$ for $\omega \in (0,1)$.
   (c) For the simple functions X_n as given in (a), evaluate
$$\lim_{n\to\infty} \int X_n dP = \lim_{n\to\infty} E(X_n).$$
   (d) 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) and (c): 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

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

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

(d) 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)-(c); this is the standard exponential distribution.