Statistics 521, Final Exam Solutions

Wellner; Wednesday 12/12/2012

1. (30 points). **Define** five of the following eight terms:
   (a) *Absolute continuity* of a signed measure ϕ with respect to a measure μ, **and** *singularity* of ϕ with respect to μ.
   (b) The *product* $\sigma-$*field* $\mathcal{A} \times \mathcal{A}'$ for two measurable spaces $(\Omega, \mathcal{A})$ and $(\Omega', \mathcal{A}')$.
   (c) *Almost sure convergence* of a sequence of random variables $\{X_n\}$.
   (d) *Independent random variables* $X_1, \ldots, X_n$ and *independent events* $A_1, \ldots, A_n$.
   (e) The *tail* $\sigma-$*field* of a sequence of random variables $X_1, X_2, \ldots$.
   (f) A $\overline{\pi}-$system $\mathcal{C}$.
   (g) A $\lambda-$system $\mathcal{D}$.
   (h) *Khintchine - equivalent* sequences of random variables.

   Solution: See PfS, pages 1-174.

2. (30 points). Give careful **statements** of three of the following six theorems or results:

   (a) The first Borel-Cantelli lemma.
   (b) The Kolmogorov zero-one law.
   (c) Feller's weak law of large numbers.
   (d) The strong law of large numbers.
   (e) The $\pi - \lambda$ theorem.
   (f) Fatou's lemma.

   Solution: See PfS, pages 1-174.

Do **either** 3 **or** 4:

3. (30 points). Let $X : \Omega \to \mathbb{R}$ be a random variable defined on the probability space $(\Omega, \mathcal{A}, P)$, let P_X denote the induced distribution of X on $(\mathbb{R}, \mathcal{B})$, and let g be a measurable function from $\mathbb{R}$ to $\mathbb{R}$.
(a) State the *theorem of the unconscious statistician* in this context.
(b) Sketch a proof of the theorem you stated in (a).

Solution: (a) In this context the theorem says

$$E\left\{g(X)1_B(X)\right\} = \int_B g(x)dP_X(x)$$

for all $B \in \mathcal{B}$.
(b) We can replace take $h(x) = g(x)1_B(x)$ without loss of generality. Thus it suffices to show that

$$E\{h(X)\} = \int h(x)dP_X(x) \tag{1}$$

for all measurable functions h from $(\mathbb{R}, \mathcal{B})$ to $(\mathbb{R}, \mathcal{B})$.
Step 1. Let $h = 1_A$ for $A \in \mathcal{B}$. Then

$$Eh(X) = \int 1_X(X)dP = P(X \in A) = P(X^{-1}(A) = P_X(A) = \int 1_A(x)dP_X(x),$$

so (**??**) holds.
Step 2. Let h be a simple function, $h(x) = \sum_1^n c_i 1_{A_i}(x)$ for $c_i \geq 0$ and $A_{\in}\mathcal{B}$ with $\sum_1^n A_i = \mathbb{R}$. Then by the linearity of the integral and step 1 we have

$$\begin{aligned} Eh(X) &= \sum_1^n c_i E 1_{A_i}(X) = \sum_1^n c_i \int 1_{A_i}(x)dP_X(x) \\ &= \int \sum_1^n c_i 1_{A_i}(x)dP_X(x) = \int h(x)dP_X(x), \end{aligned}$$

so (**??**) holds.
Step 3. Suppose $h \geq 0$. Then there exist simple functions $h_m \nearrow h$. Then by the Monotone Convergence Theorem, Step 2, and the Monotone Convergence theorem again, we have

$$Eh(X) = E(\lim_m h_m(X)) = \lim_m Eh_m(X) = \lim_m \int h_m(x)dP_X(x) = \int h(x)dP_X(x),$$

so (**??**) holds.
Step 4. h measurable and one side or the other in (**??**) well-defined. Then let $h = h^+ - h^-$. By Step 3 and linearity of the integral we find that (**??**) holds for h. □

4. (30 points) Suppose that X and Y are independent random variables and that f and g are real-valued measurable functions from $(\mathbb{R}, \mathcal{B})$ to $(\mathbb{R}, \mathcal{B})$ such that $f(X)$ and $g(Y)$ are measurable. Suppose that $E|f(X)| < \infty$ and $E|g(Y)| < \infty$. Show that

$$E[f(X)g(Y)] = E[f(X)]E[g(Y)]. \tag{2}$$

Solution: See PfS page 1-124.

Do **either** 5 **or** 6:

5. (30 points). Suppose that $X, X_1, X_2, \ldots$ are independent and identically distributed random variables.
(a) Show that the following identities holds: for all $\lambda > 0$

$$P(\max_{1 \le k \le n} |X_k| > \lambda) = P(|X| > \lambda) \sum_{k=1}^{n} P(|X| \le \lambda)^{k-1} = 1 - P(|X| \le \lambda)^n.$$

[Hint: For the first identity use the same type of decomposition of the event on the left side as we used in the proof of Kolmogorov's inequality.]
(b) Use the identities in (a) to show that for $\epsilon > 0$

$$P(\max_{1 \le k \le n} |X_k| > n\epsilon) \left\{ \begin{array}{l} \le nP(|X| > n\epsilon) \\ \ge 1 - \exp(-nP(|X| > n\epsilon)). \end{array} \right.$$

(c) Use the results of (b) to show that $M_n \equiv n^{-1} \max_{1 \le k \le n} |X_k| \to_p 0$ if and only if $xP(|X| > x) \to 0$ as $x \to \infty$ (i.e. X is weak-L_1).

Solution: (a) Let $A_k \equiv \{|X_1| \le \lambda, \ldots, |X_{k-1}| \le \lambda, |X_k| > \lambda\}$ for $k = 1, \ldots, n$. Then the A_k's are disjoint, $\sum_{k=1}^n A_k = \{\max_{1 \le k \le n} |X_k| > \lambda\}$, and

$$\begin{aligned}
P(\max_{1\le k\le n} |X_k| > \lambda) &= P(\sum_1^n A_k) = \sum_1^n P(A_k) \\
&= \sum_1^n P(|X_1| \le \lambda, \ldots, |X_{k-1}| \le \lambda, |X_k| > \lambda) \\
&= P(|X_1| \le \lambda) \cdots P(|X_{k-1}| \le \lambda) \cdot P(|X_k| > \lambda) \\
&= P(|X| > \lambda) \sum_{k=1}^n P(|X| \le \lambda)^{k-1}
\end{aligned}$$

by using independent of the X_j's in the third equality and the fact that the X_j's are identically distributed to get the fourth equality. This proves the first identity in (a). On the other hand

$$\begin{aligned}
P(\max_{k\le n} |X_k| > \lambda) &= 1 - P(\max_{k\le n} |X_k| \le \lambda) \\
&= 1 - P(|X_1| \le \lambda, \ldots, |X_n| \le \lambda) = 1 - \prod_{k=1}^n P(|X_k| \le \lambda) \\
&= 1 - P(|X| \le \lambda)^n,
\end{aligned}$$

and hence the second identity in (a) holds.
(b) The first inequality holds by replacing λ by $n\epsilon$ and then noting that $\sum_{k=1}^n P(|X| \le n\epsilon)^{k-1} \le n$. The second inequality holds by noting that $P(|X| \le n\epsilon)^n = (1 - P(|X| > n\epsilon)^n \le \exp(-nP(|X| > n\epsilon))$ since $1 - x \le e^{-x}$.
(c) Suppose that $xP(|X| > x) \to 0$ as $x \to \infty$. Then the first inequality in (b) yields

$$P(\max_{k\le n} |X_k| > n\epsilon) \le \frac{1}{\epsilon} n\epsilon P(|X| > n\epsilon) \to \frac{1}{\epsilon} \cdot 0 = 0$$

as $n \to \infty$ for every $\epsilon > 0$, so $n^{-1} \max_{k\le n} |X_k| \to_p 0$. On the other hand, if $n^{-1} \max_{k\le n} |X_k| \to_p 0$, then by the second inequality in (b) we have

$$P(\max_{k\le n} |X_k| > n\epsilon) \ge 1 - \exp(-nP(|X| > n\epsilon))$$

where the left side converges to 0 for every $\epsilon > 0$. But this implies that $n\epsilon P(|X| > n\epsilon) \to 0$ for every $\epsilon > 0$, and this holds if and only if $xP(|X| > x) \to 0$ as $x \to \infty$; i.e. X is weak-L_1.

6. (30 points). Give an example of a distribution function F with density function f with respect to Lebesgue measure λ such that $E|X| = \infty$ but $\tau(x) \equiv xP(|X| > x) \to 0$ as $x \to \infty$. Thus if $X_1, \ldots, X_n$ are i.i.d. F, the WLLN holds: $\overline{X}_n - \mu_n \to_p 0$ for some sequence μ_n (where $\mu_n = E(X_1 1_{[|X_1| \le n]})$ works), but the strong law of large numbers fails: $\limsup_n |\overline{X}_n| = +\infty$ a.s.

Proof. Let F be defined by

$$1 - F(x) = \begin{cases} \frac{e}{x \log(x)}, & x \ge e, \\ 1, & x < e. \end{cases}$$

Then F has density

$$f(x) = \frac{e}{x^2 \log x} \left\{ 1 + \frac{1}{\log x} \right\} 1_{[e,\infty)}(x)$$

and

$$xP(|X| > x) = x(1 - F(x)) = \frac{e}{\log(x)} \to 0$$

as $x \to \infty$, so X is weak-L_1. But

$$\begin{aligned} E(X) &= \int_0^\infty (1 - F(x))dx = \int_0^e 1 \cdot dx + \int_e^\infty \frac{e}{x \log(x)} dx \\ &= e\left(1 + \int_e^\infty \frac{1}{x \log(x)} dx\right) \\ &= e\left(1 + \int_1^\infty \frac{1}{y} dy\right) \quad \text{by the change of variables } y = \log(x) \\ &= e\,(1 + \infty) = \infty. \end{aligned}$$

Thus the weak law of large numbers holds, but the strong law of large numbers fails: $\overline{X}_n - \mu_n \to_p 0$ while $\limsup_{n\to\infty} |\overline{X}_n| = \infty$ a.s.

Do **either** 7 **or** 8:

7. (30 points). Suppose that X_1 and X_2 are independent Rademacher random variables, and set $X_3 = X_1 X_2$. (Thus $P(X_j = \pm 1) = 1/2$ for $j = 1, 2$.)
(a) Show that X_3 is a Rademacher random variable: $P(X_3 = \pm) = 1/2$.

(b) Show that each pair of X_1, X_2, X_3 are independent random variables.
(c) Show that X_1, X_2, X_3 are *not* independent random variables.

Solution: (a) Now $X_3 \in \{-1, +1\}$ a.s.. Furthermore,

$$\begin{aligned} P(X_3 = 1) &= P([X_1 = 1, X_2 = 1] \cup [X_1 = -1, X_2 = -1]) \\ &= P(X_1 = 1)P(X_2 = 1) + P(X_1 = -1)P(X_2 = -1) = \frac{1}{2}\frac{1}{2} + \frac{1}{2}\frac{1}{2} \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \end{aligned}$$

and hence $P(X_3 = -1) = 1/2$.
(b) Now

$$\begin{aligned} P(X_1 = 1, X_3 = 1) &= P(X_1 = 1, X_2 = 1) = \frac{1}{2}\frac{1}{2} = \frac{1}{4} \\ &= P(X_1 = 1) \cdot P(X_3 = 1), \end{aligned}$$

so X_1 and X_3 are independent (using the fact that both X_1 and X_3 take on just two values). Similarly, X_2 and X_3 are independent.
(c) Now

$$\begin{aligned} P(X_1 = 1, X_2 = 1, X_3 = 1) &= P(X_1 = 1, X_2 = 1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \\ &\neq \frac{1}{8} = P(X_1 = 1)P(X_2 = 1)P(X_3 = 1), \end{aligned}$$

so X_1, X_2, X_3 are *not* independent.

8. (30 points). Let $(\Omega, \mathcal{A}, P)$ denote the probability space $([0,1], \mathcal{B} \cap [0,1], \lambda)$ where λ is Lebesgue measure. For $n = 1, 2, \ldots$ define

$$X_n(\omega) = \begin{cases} 1, & \text{if } 0 \le \omega < 1/3, \\ 2, & \text{if } 1/3 \le \omega < 1/3 + 2/3^n, \\ 3, & \text{if } 1/3 + 2/3^n \le \omega < 1. \end{cases}$$

(a) Are the X_n's independent?
(b) What is the tail $\sigma-$field of the X_n's?

Solution: (a) Note that

$$P(X_n = j) = \begin{cases} 1/3, & \text{if } j = 1, \\ 2/3^n, & \text{if } j = 2, \\ 2/3 - 2/3^n, & \text{if } j = 3. \end{cases}$$

On the other hand, for $n \geq 1$ and $m > n$ we have

$$P(X_n = 1,\ X_m = 1) = \frac{1}{3} \neq \frac{1}{9} = P(X_n = 1) \cdot P(X_m = 1).$$

Thus the X_n's are *not independent*.
(b) Note that for each $n \geq 1$,

$$X_n^{-1}(\{1\}) = [0, 1/3), \qquad X_n^{-1}(\{2\}) = [1/3, 1/3+2/3^n), \qquad X_n^{-1}(\{3\}) = [1/3+2/3^n, 1).$$

Hence

$$\cap_{n=1}^{\infty} \mathcal{F}(X_n, X_{n+1}, \ldots) = \{\emptyset, [0, 1/3), [1/3, 1], [0, 1]\}.$$

Do **either** 9 **or** 10:

9. (30 points). Let $X_1, X_2, \ldots$ be i.i.d. with d.f. $F(x) = 1 - \exp(-x^\alpha)$ for $x \geq 0$ where $\alpha > 0$.
(a) Find a sequence b_n so that $\limsup_{n\to\infty}(X_n/b_n) = 1$ almost surely.
(b) Let $M_n \equiv \max_{1\leq k\leq n} X_k$. In the case $\alpha = 1$, find a sequence of numbers c_n so that $M_n - c_n \to_d$ "something" and find the distribution of "something".

Solution: First note that $P(X_n > x) = 1 - F(x) = \exp(-x^\alpha)$ for $x \geq 0$. For $\epsilon > -1$, set $b_n(\epsilon) = [(1+\epsilon)\log n]^{1/\alpha}$. Then

$$\begin{aligned} P(X_n > b_n(\epsilon)) &= P(X_n > (1+\epsilon)^{1/\alpha}(\log n)^{1/\alpha}) = e^{-(1+\epsilon)\log n} \\ &= n^{-(1+\epsilon)}. \end{aligned}$$

Thus

$$\sum_{n=1}^{\infty} P(X_n > b_n(\epsilon)) = \sum_{n=1}^{\infty} n^{-(1+\epsilon)} \begin{cases} < \infty, & \text{if } \epsilon > 0, \\ = \infty, & \text{if } \epsilon \leq 0. \end{cases}$$

Thus by the Borel-Cantelli lemmas $P(X_n \geq b_n(\epsilon)) = 0$ if $\epsilon > 0$, while $P(X_n \geq b_n(\epsilon)) = 1$ if $\epsilon \leq 0$. Thus we have

$$\limsup_{n\to\infty} \frac{X_n}{(\log n)^{1/\alpha}} = 1 \text{ almost surely.}$$

(b) First note that

$$P(\max_{k\leq n} X_k \leq \lambda) = P(|X_1| \leq \lambda)^n = (1 - P(X_1 > \lambda))^n.$$

Thus when $\alpha = 1$ we have

$$\begin{aligned} P(\max_{k\leq n} X_k - \log n \leq x) &= (1 - P(X_1 > x + \log n))^n = \{1 - \exp(-(x + \log n))\}^n \\ &= \left\{1 - \frac{e^{-x}}{n}\right\}^n \to \exp(-\exp(-x)) \end{aligned}$$

for all $x \in \mathbb{R}$. Thus $\max_{k\leq n} X_k - \log n \to_d M_0$ where M_0 has the double exponential extreme value (or Gumbel) distribution given by $F_0(x) = P(M_0 \leq x) = \exp(-\exp(-x))$.

10. (30 points). Suppose that $X_1, X_2, \ldots$ are uncorrelated and $E(X_j^2) \leq M < \infty$ for all $j \geq 1$.
(a) Show that $\overline{X}_n - E(\overline{X}_n) \to_2 0$.
(b) Show that $\overline{X}_n - E(\overline{X}_n) \to_p 0$.
(c) Show that $n^\alpha(\overline{X}_n - E(\overline{X}_n)) \to_p 0$ for $0 < \alpha < \alpha_0$ for some α_0 (and determine α_0).

Solution: (a) Note that

$$\begin{aligned} E(\overline{X}_n - E(\overline{X}_n))^2 &= Var(\overline{X}_n) = n^{-2}\sum_{j=1}^n Var(X_j) \text{ since the } X_j's \text{ are uncorrelated} \\ &\leq n^{-2}\sum_{j=1}^n E(X_j)^2 \text{ since } Var(X_j) \leq EX_j^2 \\ &\leq n^{-1}M \text{ since } E(X_j^2) \leq M \text{ for all } j \geq 1 \\ &\to 0 \text{ as } n \to \infty. \end{aligned}$$

Therefore $\overline{X}_n - E(\overline{X}_n) \to_2 0$. (b) Since $\to_2$ implies $\to_p$, this follows immediately from (a).
(c) Now by the computation in (a),

$$E\left\{[n^\alpha(\overline{X}_n - E(\overline{X}_n))]^2\right\} = n^{2\alpha} Var(\overline{X}_n) \le M/n^{1-2\alpha} \to 0$$

as $n \to \infty$ if $\alpha < 1/2 \equiv \alpha_0$, so $n^\alpha(\overline{X}_n - E(\overline{X}_n)) \to_2 0$. This implies $n^\alpha(\overline{X}_n - E(\overline{X}_n)) \to_p 0$ just as in (b).