Statistics 582, Problem Set 2 Solutions

Wellner; 10/19/2001

1. Ferguson, ACILST, #2, page 6:
(a) Suppose that $X_n \sim \text{Uniform}\{1/n, 2/n, \ldots, n/n\}$. Show that $X_n \rightarrow_d X \sim \text{Uniform}(0,1)$. Does $X_n \rightarrow_p X$?
(b) Suppose that $Y_n = \sum_{k=1}^n (k/n) 1_{[(k-1)/n, k/n)}(U)$ where $U \sim \text{Uniform}[0,1]$. Show that $Y_n \sim \text{Uniform}\{1/n, 2/n, \ldots, n/n\}$, and $Y_n \rightarrow_p U$.
Solution: (a) For $0 \le x \le 1$,

$$P(X_n \le x) = \frac{1}{n}\sum_{i=1}^n 1_{[i/n,1]}(x) = \frac{1}{n}\sum_{i=1}^n 1_{[i/n \le x]} = [nx]/n \rightarrow x\,;$$

here $[x]$ = greatest integer less than or equal to x. Thus $X_n \rightarrow X \sim \text{Uniform}(0,1)$. X_n does not necessarily converge in probability to X because all the different random variables involved could be defined on different probability spaces.
(b) Now $P(Y_n = k/n) = P(U \in [(k-1)/n, k/n)) = 1/n$ for $k = 1, \ldots, n$, so $Y_n \sim \text{Uniform}\{1/n, \ldots, n/n\}$. Furthermore,

$$P(|Y_n - U| \ge \epsilon) = \begin{cases} n(1/n - \epsilon), & \text{if } 0 \le \epsilon \le 1/n \\ 0, & \text{if } \epsilon > 1/n\,, \end{cases}$$

and this clearly converges to 0 as $n \rightarrow \infty$.

2. Ferguson, ACILST, #4, page 6:
Give an example of random variables X_n such that $E|X_n| \rightarrow 0$ and $E|X_n|^2 \rightarrow 1$.
Solution: If $X_n = a_n$ with probability p_n and $X_n = 0$ otherwise, then $E|X_n| = a_n p_n$ while $E|X_n|^2 = a_n^2 p_n$. In particular we choose $p_n = 1/a_n^2$, then $E|X_n| = 1/a_n$ and $E|X_n|^2 = 1$. Thus for any choice $a_n \rightarrow 0$ we have $E|X_n| \rightarrow 0$ and $E|X_n^2| = 1$ for all n.

3. Suppose that $U \sim \text{Uniform}(0,1)$, $\alpha > 0$, and

$$X_n \equiv (n^\alpha / \log(n+1)) 1_{[0,1/n^\alpha]}(U)\,.$$

(a) Show that $X_n \to_{a.s.} 0$ and $E(X_n) \to E(0) = 0$.

(b) Can you find a random variable Y with $|X_n| \le Y$ for all n with $E(Y) < \infty$ for any α?
(c) For what values of α does the uniform integrability condition

$$\limsup_{n\to\infty} E\{|X_n| 1_{[|X_n|\ge M]}\} \to 0 \qquad \text{as} \quad M \to \infty$$

hold?
Solution: (a) $X_n \to_{a.s.} 0$ since $X_n(\omega) = 0$ for $1/n^\alpha < U(\omega)$, or equivalently $n > (1/U(\omega))^{1/\alpha}$ and since $P(0 < U \le 1) = 1$. Moreover,

$$E(X_n) = \frac{n^\alpha}{\log(n+1)} \frac{1}{n^\alpha} = \frac{1}{\log(n+1)} \to 0 = E(0)\,.$$

(b) Now the smallest possible random variable Y satisfying $|X_n| \le Y$ for all n is Y defined by

$$Y = \sum_{k=1}^{\infty} \frac{k^\alpha}{\log(k+1)} 1_{(1/(k+1)^\alpha, 1/k^\alpha]}(U)\,.$$

But we compute

$$\begin{aligned} E(Y) &= \sum_{k=1}^{\infty} \frac{k^\alpha}{\log(k+1)} \left\{ \frac{1}{k^\alpha} - \frac{1}{(k+1)^\alpha} \right\} \\ &= \sum_{k=1}^{\infty} \frac{1}{\log(k+1)} \left\{ 1 - \left(\frac{k}{k+1} \right)^\alpha \right\} \\ &= \sum_{k=1}^{\infty} \frac{1}{\log(k+1)} \left\{ 1 - \left(1 - \frac{1}{k+1} \right)^\alpha \right\} \\ &\ge \sum_{k=1}^{k(\alpha)} \frac{1}{\log(k+1)} \left\{ 1 - \left(1 - \frac{1}{k+1} \right)^\alpha \right\} \\ &\qquad + \sum_{k=k(\alpha)}^{\infty} \frac{1}{\log(k+1)} \frac{\alpha/2}{k+1} \\ &= \infty\,. \end{aligned}$$

since $(1-x)^\alpha \le 1 - \alpha x/2$ for $x \le x(\alpha)$. Thus there is no integrable dominating function Y for any value of α.

(c) On the other hand the uniform integrability condition does hold for any $\alpha > 0$:

$$\begin{aligned} E\{|X_n|1_{[|X_n|\geq M]}\} &= E\left\{\frac{n^\alpha}{\log(n+1)}1_{[0,1/n^\alpha]}(U)1\{(n^\alpha/\log(n+1))\geq M, U\leq 1/n^\alpha\}\right\} \\ &= \frac{n^\alpha}{\log(n+1)}E\left\{1_{[0,1/n^\alpha]}(U)\right\}1\{(n^\alpha/\log(n+1))\geq M\} \\ &= \frac{1}{\log(n+1)}1\{(n^\alpha/\log(n+1))\geq M\} \\ &\to 0\cdot 1 = 0 \end{aligned}$$

as $n \to \infty$ for every $\alpha > 0$.

4. Suppose that $X \sim \text{Uniform}(0,1)$ and $Y = X^2$.
(a) Find the joint distribution function $F(x,y) = F_{X,Y}(x,y)$ of (X,Y).
(b) Is F a continuous function?
(c) Is the probability measure P corresponding to F absolutely continuous with respect to Lebesgue measure μ on R^2?
Solution: (a) The joint distribution function $F(x,y)$ of (X,Y) is given by

$$\begin{aligned} F(x,y) &= P(X\leq x,\, Y\leq y) = P(X\leq x, X^2\leq y) \\ &= P(X\leq x,\, X\leq\sqrt{y}) = P(X\leq x\wedge\sqrt{y}) \\ &= \begin{cases} 0 & \text{if } \ x\leq 0 \ \text{ or } \ y\leq 0\,, \\ x\wedge\sqrt{y} & \text{if } \ 0\leq x,y\leq 1\,, \\ 1 & \text{if } \ x>1, y>1\,. \end{cases} \end{aligned}$$

(b) Yes, the joint distribution function F is clearly a continuous function of x and y.
(c) No. The corresponding measure P puts all its mass on the set $A \equiv \{(x,y):\ (x,x^2):\ 0\leq x\leq 1\}$. On the other hand, the Lebesgue measure of this set is zero $\mu(A) = \int_A d\mu(x,y) = 0$. Thus P is not absolutely continuous with respect to μ.

5. (a) Lehmann and Casella, #3.5, page 64.
(b) Lehmann and Casella, #3.6, page 64.
(c) Lehmann and Casella, #3.7, page 64.
Solution: (a) (i) Suppose that S is not closed. Then there exists a

sequence $\{x_n\} \subset S$ such that $x_n \to x_0 \in S^c$. But then, for every $\epsilon > 0$ there is an open ball $B(x_0, \epsilon)$ such that $x_n \in B(x_0, \epsilon)$ for $n \ge N_\epsilon$. Since each x_n is a support point, $P(B(x_0, \epsilon)) > 0$ for each $\epsilon > 0$. But for any open set A with $x_0 \in A$, $B(x_0, \epsilon) \subset A$ for some $\epsilon > 0$, and hence $P(A) \ge P(B(x_0, \epsilon) > 0$. But this implies $x_0 \in S$. Contradiction. Thus S is closed.
(ii) $P(S) = 1$. From (i) S is closed, so S^c is open. Since $x \in S^c$ if and only if $x \in A_x$ with A_x an open rectangle satisfying $P(A) = 0$. Thus $S^c \subset \cup_x A_x$. By the Lindelöf theorem, for any such open covering $\{A_x\}_{x \in S^c}$ of $S^c \subset R^d$, there is a countable subcollection $\{A_{x_n}\}$ which covers S^c: $S^c \subset \cup_n A_{x_n}$. Then we have

$$P(S^c) \le P(\cup_n A_{x_n}) \le \sum_n P(A_{x_n}) = \sum_n 0 = 0\,.$$

Hence $P(S) = 1$.
(iii) We want to show that $S = \cap\{C : \ C \text{ closed}, P(C) = 1\}$. From (i) and (ii) we know that S is in the collection of sets on the right side, so it follows that $S \supset \cap\{C : \ C \text{ closed}, P(C) = 1\}$. Thus it remains to show that $S \subset \cap\{C : \ C \text{ closed}, P(C) = 1\}$. Equivalently, it remains to show that $S^c \supset \cup\{C^c : C^c \text{ open}, \ P(C^c) = 0\}$. But if $x \in \cup\{C^c : C^c \text{ open}, \ P(C^c) = 0\}$, then $x \in C^c$ for some C^c open with $P(C^c) = 0$, and hence also $x \in A \subset C^c$ for some open rectangle A (an open ball centered at x for the metric $\|y\| = \max_{1 \le i \le d} |x_i|$) with $P(A) \le P(C^c) = 0$. Hence $x \in S^c$.
(b) Suppose that P and Q are equivalent: i.e. $Q \prec\prec P$ and $P \prec\prec Q$. Then for any open set A, $P(A) = 0$ if and only if $Q(A) = 0$. This implies that for any closed set A^c,

$$P(A^c) = 1 \qquad \text{if and only if} \qquad Q(A^c) = 1\,.$$

This implies that the minimal closed set S_P with $P(S_P) = 1$ is also the minimal closed set S_Q with $Q(S_Q) = 1$; i.e. $S_P = \text{supp}(P) = \text{supp}(Q) = S_Q$.
(c) Since $P(X = 1/n) = p_n > 0$ for $n = 1, 2, \ldots$ with $\sum_1^\infty p_n = 1$, it follows that $\text{supp}(P) = \{0, \ldots, 1/n, \ldots, 1/2, 1\}$, which is closed. Similarly, Since $Q(X = 1/n) = q_n > 0$ for $n = 1, 2, \ldots$ with $\sum_1^\infty q_n = 1/2$, and $Q(X = 0) = 1/2$, it follows that $\text{supp}(Q) = \{0, \ldots, 1/n, \ldots, 1/2, 1\} =$

supp(P). But $P(\{0\}) = 0$ while $Q(\{0\}) = 1/2$, so $Q \prec\prec P$ fails. Thus Q and P are not equivalent.

6. Suppose that $X \sim F$ on $R^+ \equiv [0, \infty)$, $Y \sim G$ on R^+, and X and Y are independent random variables. Let $Z = \min\{X, Y\} = X \wedge Y$ and $\Delta = 1\{X \le Y\}$. (This is *right-censored data*: if we view X as a survival time, and Y as a censoring time, then $Z = X$ when $X \le Y$, but $Z = Y$ when $X > Y$.)
(a) Find the joint distribution of (Z, Δ).
(b) If $X \sim$ Exponential(λ) and $Y \sim$ Exponential(μ), show that Z and Δ are independent.
[Hint: for (a), compute $P(Z \le z, \Delta = 1)$ and $P(Z \le z, \Delta = 0)$.]
Solution: (a) Since $Z = \min\{X, Y\} = X \wedge Y$ and $\Delta = 1\{X \le Y\}$, it follows that

$$H_{uc}(z) \equiv P(X \le z, X \le Y) = \int_{[0,z]} (1 - G(x-))dF(x)\,,$$

and

$$H_c(z) \equiv P(Y \le z, X > Y) = \int_{[0,z]} (1 - F(y))dG(y)\,.$$

These two sub-distribution functions completely determine the joint distribution function H of (Z, Δ) since

$$P(Z \le z, \Delta \le \delta) = \begin{cases} 0, & \text{if } \delta < 0\,, \\ H_c(z), & \text{if } 0 \le \delta < 1\,, \\ H_c(z) + H_{uc}(z), & \text{if } 1 \le \delta < \infty\,. \end{cases}$$

Note that

$$1 - H_c(z) - H_{uc} = P(Z > z) = (1 - F(z))(1 - G(z))\,,$$

so the marginal d.f. of Z is

$$H(z, 1) = H_c(z) + H_{uc}(z) = 1 - (1 - F(z))(1 - G(z))\,.$$

(b) When $1 - F(x) = \exp(-\lambda x)$ and $1 - G(x) = \exp(-\mu x)$, then

$$1 - H(z, 1) = (1 - F(z))(1 - G(z)) = \exp(-(\lambda + \mu)x)\,,$$

while

$$P(\Delta = 1) = P(X \le Y) = H_{uc}(\infty) = \frac{\lambda}{\lambda + \mu}\,,$$

so $Z \sim \text{Exponential}(\lambda + \mu)$, $\Delta \sim \text{Bernoulli}(\lambda/(\lambda + \mu))$. Furthermore,

$$H_{uc}(z) = \int_0^z e^{-\mu x} \lambda e^{-\lambda x}\, dx = \frac{\lambda}{\lambda + \mu} \left(1 - \exp(-(\lambda + \mu) z)\right)$$

$$H_c(z) = \int_0^z e^{-\lambda x} \lambda e^{-\mu x}\, dx = \frac{\mu}{\lambda + \mu} \left(1 - \exp(-(\lambda + \mu) z)\right)\,,$$

so that Z and Δ are independent in this case.