Statistics 581, Problem Set 2 Solutions
Wellner; 10/12/2018

1. Suppose that $X_1, X_2, \ldots$ is a sequence of random variables such that $X_1 \sim \text{Uniform}(0,1)$, and for $n = 1, 2, \ldots$ the conditional distribution of X_{n+1} given $X_1, \ldots, X_n$ is uniform on $[0, cX_n]$ for a number $c \in (\sqrt{3}, 2)$.
   (a) Compute $E(X_n^r)$ for $r > 0$.
   (b) Show that X_n converges to 0 in mean, but X_n does not converge to 0 in quadratic mean.
   (c) Does $X_n \rightarrow_{a.s.} 0$?

   Solution: (a) We compute

$$
\begin{aligned}
E(X_{n+1}^r) &= E\{E(X_{n+1}^r | X_n)\} \\
&= E\left(\int_0^{cX_n} y^r \frac{1}{cX_n} dy \right) \\
&= E\left(\frac{1}{cX_n(r+1)} (cX_n)^{r+1} \right) = \frac{c^r}{(r+1)} E(X_n^r) \\
&= \frac{c^{2r}}{(r+1)^2} E(X_{n-1}^r) = \cdots = \left(\frac{c^r}{(r+1)} \right)^n E(X_1) \\
&= \left(\frac{c^r}{(r+1)} \right)^n \frac{1}{r+1}.
\end{aligned}
$$

   (b) When $r = 1$ the expression on the right side in the last display become $(c/2)^n \cdot 2^{-1} \rightarrow 0$ since $c/2 < 1$. When $r = 2$ it reduces to $(c^2/3)^n \cdot (1/3) \rightarrow \infty$ since $c^2/3 > 1$.
   (c) Note that for any $\epsilon > 0$ and $r = 1$ we have

$$
P(X_n \geq \epsilon) \leq \epsilon^{-1} E(X_n) \leq \epsilon^{-1} (c/2)^{(n-1)}/2
$$

   where $(c/2) < 1$ and hence

$$
\sum_{n=1}^{\infty} P(X_n \geq \epsilon) \leq \frac{1}{2\epsilon} \sum_{n=1}^{\infty} (c/2)^{(n-1)} < \infty.
$$

   Thus $X_n \rightarrow_{a.s.} 0$ by the Borel-Cantelli lemma.

2. Wellner 581 Course Notes, Chapter 1, Exercise 4.1, page 19. (Show just the first equality in each case; we will do the second equalities later.)

   Solution: To see that the first equality in (11) holds, we use Fubini's theorem as follows:

$$
\begin{aligned}
E(X) &= \int_0^{\infty} x dF(x) = \int_0^{\infty} \int_0^x dt dF(x) = \int_0^{\infty} \int_0^{\infty} 1_{[0,x)}(t) dt dF(x) \\
&= \int_0^{\infty} \int_0^{\infty} 1_{[0,x)}(t) dF(x) dt = \int_0^{\infty} \int_{(t,\infty)} dF(x) dt = \int_0^{\infty} (1 - F(t)) dt.
\end{aligned}
$$

To see that equality holds in (12), we proceed much as the proof of (11) after separating the expectation into two terms:

$$
\begin{aligned}
EX &= \int_0^\infty x dF(x) + \int_{-\infty}^0 x dF(x) \\
&= \int_0^\infty \left(\int_0^x dt \right) dF(x) - \int_{-\infty}^0 \left(\int_x^0 dt \right) dF(x) \\
&= \int_0^\infty \int_0^\infty 1\{t < x\} dF(x) dt - \int_{-\infty}^0 \int_{-\infty}^0 1\{x \le t\} dF(x) dt \\
&= \int_0^\infty (1 - F(t)) dt - \int_{-\infty}^0 F(t) dt.
\end{aligned}
$$

The proof of (13) goes much as the proof of (11), but using the substitution $x^r = r \int_0^x t^{r-1} dt$:

$$
\begin{aligned}
E(X^r) &= \int_0^\infty x^r dF(x) = \int_0^\infty \int_0^x r t^{r-1} dt dF(x) = \int_0^\infty \int_0^\infty 1_{[0,x)}(t) r t^{r-1} dt dF(x) \\
&= \int_0^\infty \left(\int_0^\infty 1_{[0,x)}(t) dF(x) \right) r t^{r-1} dt \\
&= \int_0^\infty \int_{(t,\infty)} dF(x) r t^{r-1} dt = \int_0^\infty r t^{r-1} (1 - F(t)) dt.
\end{aligned}
$$

3. Ferguson, ACILST, #4, page 6:
   (a) Give an example of random variables X_n such that $E|X_n| \to 0$ and $E|X_n|^2 \to 1$.
   (b) Give an example of a sequence of random variables X_n such that $X_n \to_p 0$ and $EX_n \to 0$, but $X_n \to_{a.s.} 0$ fails.
   (c) Suppose that Y has a standard Cauchy distribution with density $f(y) = (\pi(1+y^2))^{-1}$. Find a sequence of random variables Y_n such that $Y_n \to_2 Y$, but Y_n does not converge to Y almost surely.

   Solution: (a) If $X_n = a_n$ with probability p_n and $X_n = 0$ with probability $1 - p_n$, then $E(X_n) = a_n p_n$ and $E(X_n^2) = a_n^2 p_n = 1$ if $p_n = 1/a_n^2$. Then $E(X_n) = a_n / a_n^2 = 1/a_n \to 0$ if $a_n \to \infty$. Ferguson's solution on page 173 takes $a_n = n$; the same holds for any sequence $a_n \to \infty$.
   (b) Let $U \sim \text{Uniform}(0,1)$. The "dancing functions" are defined by $X_{n,k} = 1_{[(k-1)/2^n, k/2^n)}(U)$, $k = 1, \ldots, 2^n$, $n = 1, 2, \ldots$. Let $\{Y_m\}_{m \ge 1}$

be defined by $Y_m = X_{n,k}$ if $m = (\sum_{j=1}^n 2^j) + k = 2^{n+1} - 2 + k$ with $1 \le k \le 2^n$. Then for $\epsilon \in (0,1)$,

$$P(|Y_m| > \epsilon) = P(|X_{n,k}| > \epsilon) = 2^{-n} \to 0$$

so $Y_m \to_p 0$, but for every $U(\omega) \in (0,1)$ we have $Y_m(\omega) = 1$ for infinitely many m's and also $Y_m(\omega) = 0$ for infinitely many m's. Hence

$$0 = \liminf Y_m < \limsup Y_m = 1 \qquad a.s.$$

and it follows that Y_m does not converge to 0 almost surely. To see that $EY_m \to 0$, note that Y_m takes on only the values 0 and 1, and hence $Y_m \sim \text{Bernoulli}(p_m)$ where $p_m = p_{n,k} = 2^{-n}$ for $m = 2^{n+1} + k$ with $1 \le k \le 2^n$. Thus $E(Y_m^2) \le 1 \cdot E(Y_m) = p_m \to 0$ as $m \to \infty$.
(c) Let V_n be a sequence of random variables as in (b) satisfying $V_n \to_p 0$, and $EV_n^2 \to 0$, but such that V_n does not converge to 0 a.s., and define $Y_n \equiv Y + V_n$. Then $E(Y_n - Y)^2 = EV_n^2 \to 0$ but $Y_n = Y + V_n$ does not converge almost surely to Y since V_n does not converge a.s. to 0.

4. vdV, *Asymp. Statist.*, problem 5, page 24: Find an example of a sequence (X_n, Y_n) such that $X_n \to_d X$, $Y_n \to_d Y$, but (X_n, Y_n) does not converge in distribution.

   Solution: Suppose that $X_n = U \sim \text{Uniform}(0,1)$ for every n and let $Y_{2n} = U$, $Y_{2n-1} = 1 - U$ for $n = 1, 2, \ldots$. Then $X_n \stackrel{d}{=} U$ for every n and $Y_n \sim U$ for every n since $1 - U \stackrel{d}{=} U$. But (X_n, Y_n) does not converge in distribution: for every even integer n the random vector has a uniform distribution on $\{(x,x) : \ 0 \le x \le 1\}$ while for every odd integer n the random vector (X_n, Y_n) has a uniform distribution on $\{(x, 1-x) : \ 0 \le x \le 1\}$. (Note that since $\{X_n\}$ is tight and $\{Y_n\}$ is tight it follows that $\{(X_n, Y_n)\}$ is tight, and by Prohorov's theorem the exist subsequences $(X_{n'}, Y_{n'})$ which do converge in distribution. In the present example there are exactly two such subsequences.

5. (See vdV, *Asymp. Stat.*, section 11.1, pages 153 - 156.)
   Suppose that Y is a random variable with $E(Y^2) < \infty$, let X be another random variable on the same probability space as Y, and consider finding a (measurable) function g of X with $Eg^2(X) < \infty$ so that

$E(Y - g(X))^2$ is "small".
(a) Show that

$$\inf_{g:\mathbb{R}\to\mathbb{R}, Eg^2(X)<\infty} E(Y - g(X))^2 = E(Y - E(Y|X))^2$$

so that the minimizer is exactly $g_0(X) \equiv E(Y|X)$.
(b) Show that $E\{(Y - E(Y|X))g(X)\} = 0$ for all $g(X) \in L_2(P)$.
(c) Interpret the results in (a) and (b) geometrically (i.e. in the Hilbert space $L_2(P)$ of square integrable random variables with the inner product $\langle X, Y\rangle \equiv E(XY)$.

Solution: (a) Note that

$$\begin{aligned} E(Y - g(X))^2 &= E(Y - E(Y|X) + E(Y|X) - g(X))^2 \\ &= E(Y - E(Y|X))^2 + E(E(Y|X) - g(X))^2 \\ &\quad + 2E\{(Y - E(Y|X))(E(Y|X) - g(X))\} \\ &= E(Y - E(Y|X))^2 + E(E(Y|X) - g(X))^2 \qquad (1) \end{aligned}$$

since, by computing conditionally on X,

$$\begin{aligned} &E\{(Y - E(Y|X))(E(Y|X) - g(X))\} \\ &= EE\{(Y - E(Y|X))(E(Y|X) - g(X))|X\} \\ &= E\left\{(E(Y|X) - g(X))E\{(Y - E(Y|X))|X\}\right\} \\ &= E\left\{(E(Y|X) - g(X)) \cdot 0\right\} = 0. \end{aligned}$$

From (1) we conclude that

$$E(Y - g(X))^2 \geq E(Y - E(Y|X))^2$$

with equality if and only if $g(X) = E(Y|X)$ almost surely.
(b) By a computation similar to that in (a) we have, for any $g(X) \in L_2(P)$,

$$\begin{aligned} E\{(Y - E(Y|X))g(X)\} &= EE\{(Y - E(Y|X))g(X)|X\} \\ &= E\{g(X)E\{(Y - E(Y|X))|X\}\} \\ &= E\{g(X) \cdot 0\} = 0. \end{aligned}$$

(c) The result in (a) shows that $E(Y|X)$ is the "projection" of Y onto the sub-space of all $L_2(P)$ random variables which are measurable functions of X. The result in (b) shows that the "residual" $Y - E(Y|X)$ resulting from projecting Y onto $L_2(P_X)$ is orthogonal to the subspace $L_2(P_X)$ of all square integrable functions of X.