Statistics 522, Problem Set 4 Solutions

Wellner; 2/5/99

1. Exercise 8.4.2, PfS page 141. Repeat Example 8.4.1 for sampling without replacement.

   Solution: (a) When the sampling is done without replacement the joint probability distribution for (X_1, X_2) is as follows:

			X_1		
		1	2	3	
	1	0	2/30	3/30	5/30
X_2	2	2/30	2/30	6/30	10/30
	3	3/30	6/30	6/30	15/30
		5/30	10/30	15/30	1

   Hence the marginal distribution of $S = X_1 + X_2$ is given by

k	2	3	4	5	
$P(S = k)$	4/30	8/30	12/30	6/30	1

   It is easy to compute the conditional distribution of $Y = X_2$ given S (or given $\mathcal{D} = S^{-1}(\mathcal{B})$): letting $D_j = [S = j]$,

		Y		
	1	2	3	$E(Y\vert\mathcal{D})$
D_3	1/2	1/2	0	3/2
D_4	3/8	2/8	3/8	2
D_5	0	1/2	1/2	5/2
D_6	0	0	1	3
$P(Y = i)$	5/30	10/30	15/30	

Note that

$$P(Y = i|\mathcal{D}) = \sum_{j=3}^{6} \frac{P([Y = i] \cap D_j)}{P(D_j)} 1_{D_j}$$

satisfies $P(Y = i) = E\{P(Y = i|\mathcal{D})\}$. Also note that $E(Y) = 7/3$, and

$$E(E(Y|\mathcal{D})) = (3/2)(4/30) + 2(8/30) + (5/2)(12/30) + 3(6/30) = 7/3\,.$$

2. Exercise 8.4.3, PfS page 141. Let Y denote a rv on some $(\Omega, \mathcal{A}, P)$ that takes on the eight values $1, \ldots, 8$ with probabilities $1/32, 2/32, 3/32, 4/32, 15/32, 4/32, 1/32, 1/32$. Let $\mathcal{C} = \mathcal{F}(Y)$, and let $C_i \equiv [Y = i]$ and $p_i = P(C_i)$ for $i = 1, \ldots, 8$. Let $\mathcal{D} \equiv \sigma[\{C_1 + C_5, C_2 + C_6, C_3 + C_7, C_4 + C_8\}]$, $\mathcal{E} \equiv \sigma[\{C_1 + C_5 + C_2 + C_6, C_3 + C_7 + C_4 + C_8\}]$, and $\mathcal{F} \equiv \{\Omega, \}$.
(a) Represent Ω as a 2×4 rectangle having eight cells representing $C_1, \ldots, C_4$ in the first row and $C_5, \ldots, C_8$ in the second row. Enter the appropriate values of $Y(\omega)$ and p_i in each cell, forming a table. Evaluate $E(Y)$.
(b) Evaluate $E(Y|\mathcal{D})$ and $E(E(Y|\mathcal{D}))$.
(c) Evaluate $E(Y|\mathcal{E})$ and $E(E(Y|\mathcal{E}))$.
(d) Evaluate $E(Y|\mathcal{F})$ and $E(E(Y|\mathcal{F}))$.

Solution: (a) With the events C_i as given, we compute the following probability distribution:

i	1	2	3	4
$P(Y = i)$	1/32	2/32	3/32	4/32
$P(Y = i + 4)$	15/32	4/32	1/32	2/32
$P(Y = i \text{ or } i + 4)$	16/32	6/32	4/32	6/32

This results in

$$E(Y) = (1 \cdot 1 + 2 \cdot 2 + 3 \cdot 3 + 4 \cdot 4 + 5 \cdot 15 + 6 \cdot 4 + 7 \cdot 1 + 8 \cdot 2)/32 = 152/32\,.$$

(b) With $\mathcal{D}$ as given we compute the conditional probability distribution and conditional expectations to be:

D_i	1	2	3	4
$P(Y = i\|\mathcal{D})$	1/16	2/6	3/4	4/6
$P(Y = i+4\|\mathcal{D})$	15/16	4/6	1/4	2/6
$E(Y\|\mathcal{D})$	76/16	28/6	16/4	32/6

This yields

$$E(E(Y|\mathcal{D})) = \frac{76}{16}\cdot\frac{16}{32} + \frac{28}{6}\cdot\frac{6}{32} + \frac{16}{4}\cdot\frac{4}{32} + \frac{32}{6}\cdot\frac{6}{32} = \frac{152}{32} = E(Y)\,.$$

(c) With $\mathcal{E}$ as given, we compute the conditional probability distribution and conditional expectations to be:

i	1	2	3	4
$P(Y = i\|\mathcal{E})$	1/22	2/22	3/10	4/10
$P(Y = i+4\|\mathcal{E})$	15/22	4/22	1/10	2/10
$P(E_j)$	22/32		10/32	

This yields

$$E(Y|\mathcal{E})(\omega) = \frac{1+4+75+24}{22} = \frac{104}{22} \qquad \omega\in E_1\,,$$

and

$$E(Y|\mathcal{E})(\omega) = \frac{9+16+7+16}{10} = \frac{48}{10} \qquad \omega\in E_2\,,$$

and hence

$$E(E(Y|\mathcal{E})) = \frac{104}{22}\frac{22}{32} + \frac{48}{10}\frac{10}{32} = \frac{152}{32} = E(Y)\,.$$

3. Exercise 8.4.4, PfS page 145: i.e. prove (24) and (26) on page 142:
(24) The C_r-, Hölder, Liapounov, Minkowski, and Jensen inequalities hold for $E(\cdot|\mathcal{D})$.
(26) h is a determination of $E(Y|\mathcal{D})$ if and only if $E(XY) = E(Xh)$ for all $\mathcal{D}-$measurable rv's X.

Solution: (24): C_r: For $r \ge 1$, $|x|^r$ is a convex function of x, so $|(x+y)/2|^r \le (1/2)(|x|^r + |y|^r)$. Thus $|X+Y|^r \le 2^{r-1}\{|X|^r + |Y|^r\}$.

Taking condition expectations across this inequality and using (16) yields $E(|X+Y|^r|\mathcal{D}) \leq 2^{r-1}\{E(|X|^r|\mathcal{D}) + E(|Y|^r|\mathcal{D})\}$. For $0 < r \leq 1$, $|X+Y|^r \leq |X|^r + |Y|^r$, so taking conditional expectations across this inequality yields $E(|X+Y|^r|\mathcal{D}) \leq E(|X|^r|\mathcal{D}) + E(|Y|^r|\mathcal{D})$.
Hölder's inequality: for arbitrary $a, b \in R$ and r, s satisfying $(1/r) + (1/s) = 1$, we have

$$|ab| \leq \frac{|a|^r}{r} + \frac{|b|^s}{s}$$

with equality only if $|b| = |a|^{1/(s-1)}$. Taking $a = |X|/E^{1/r}(|X|^r|\mathcal{D})$ and $b = |Y|/E^{1/s}(|Y|^s|\mathcal{D})$, we find that

$$\frac{|X||Y|}{E^{1/r}(|X|^r|\mathcal{D})E^{1/s}(|Y|^s|\mathcal{D})} \leq \frac{|X|^r}{rE(|X|^r|\mathcal{D})} + \frac{|Y|^s}{sE(|X|^s|\mathcal{D})},$$

and taking conditional expectations across this inequality and using (16) gives

$$\frac{E\{|X||Y||\mathcal{D}\}}{E^{1/r}(|X|^r|\mathcal{D})E^{1/s}(|Y|^s|\mathcal{D})} \leq \frac{1}{r} + \frac{1}{s} = 1\,.$$

This yields $E\{|X||Y||\mathcal{D}\} \leq E^{1/r}(|X|^r|\mathcal{D})E^{1/s}(|Y|^s|\mathcal{D})$ with equality if and only if

$$\frac{|Y|}{E^{1/s}(|X|^s|\mathcal{D})} = \left(\frac{|X|}{E^{1/r}(|X|^r|\mathcal{D})}\right)^{1/(s-1)} \qquad \text{a.s.}$$

Liapunov's inequality: Suppose that $E|X|^q < \infty$. and let $0 < p \leq q$. Then by the conditional Hölder inequality with $1/r = p/q$, $1/s = 1 - p/q$, we find that

$$E(|X|^p|\mathcal{D}) \leq E(|X|^q|\mathcal{D})^{p/q}E(1^{1/(1-p/q)}|\mathcal{D})^{1-p/q} = E(|X|^q|\mathcal{D})^{p/q} \qquad \text{a.s.}$$

This implies that $E(|X|^p|\mathcal{D})^{1/p} \leq E(|X|^q|\mathcal{D})^{1/q}$ a.s.
Minkowski's inequality: This follows from the conditional Hölder inequality in the same way that Minkowski's inequality follows from the unconditional Hölder inequality.
Jensen's inequality: see the nice proof in Williams, page 89, and note the "important corollary" to Williams' (h).

(26): Suppose that $E(XY) = E(Xh)$ for all $\mathcal{D}-$measurable rv's X. Then, in particular with $h = 1_D$ for $D \in \mathcal{D}$, we have $E(1_D Y) = E(1_D h)$

for $D \in \mathcal{D}$, and hence h is a version (or "determination") of $E(Y|\mathcal{D})$.
On the other hand, suppose that h is a version of $E(Y|\mathcal{D})$; i.e. $E(1_D Y) = E(1_D h)$ for all $D \in \mathcal{D}$. Note that this implies $E(1_D Y^+) = E(1_D h^+)$ and $E(1_D Y^-) = E(1_D h^-)$ for all $D \in \mathcal{D}$.

Suppose first that $X \geq 0$. Then there is a sequence of $\mathcal{D}-$measurable simple functions $X_n = \sum_{j=1}^n d_j 1_{D_j} \nearrow X$. Then by the monotone convergence theorem

$$\begin{aligned}
E(XY) &= E(X(Y^+ - Y^-)) = E(XY^+) - E(XY^-) \\
&= \lim_n E(X_n Y^+) - \lim_n E(X_n Y^-) \qquad \text{by the MCT} \\
&= \lim_n E(\sum_1^n d_j 1_{D_j} Y^+) - \lim_n E(\sum_1^n d_j 1_{D_j} Y^-) \\
&= \lim_n \sum_1^n d_j E(1_{D_j} Y^+) - \lim_n \sum_1^n d_j E(1_{D_j} Y^-) \\
&= \lim_n \sum_1^n d_j E(1_{D_j} h^+) - \lim_n \sum_1^n d_j E(1_{D_j} h^-) \text{by the equality for sets} \\
&= \lim_n E(X_n h^+) - \lim_n E(X_n h^-) \qquad \text{by reversing the above steps} \\
&= E(Xh^+) - E(Xh^-) \qquad \text{by the MCT} \\
&= E(X(h^+ - h^-)) = E(Xh).
\end{aligned}$$

Now suppose that X is arbitrary with $E|XY| < \infty$. Then

$$\begin{aligned}
E(XY) &= E((X^+ - X^-)Y) = E(X^+Y) - E(X^-Y) \\
&= E(X^+h) - E(X^-h) \qquad \text{by the result for } X \geq 0 \\
&= E((X^+ - X^-)h) = E(Xh).
\end{aligned}$$

4. Exercise E9.2, PwMG, page 231. Suppose that $X, Y \in L_1(\Omega, \mathcal{F}, P)$ and that $E(Y|X) = X$ a.s. and $E(X|Y) = Y$ a.s. Prove that $P(X = Y) = 1$.

Solution: Suppose first that $X, Y \in L_2(P)$. Then, by Pythagoras,

$$EX^2 = E(E(X|Y)^2) + E((X - E(X|Y))^2),$$

and since $E(X|Y) = Y$ a.s. this yields

$$E(X^2) = E(Y^2) + E((X-Y)^2)\,. \tag{1}$$

Reversing the roles of X and Y, we also obtain, upon using $E(Y|X) = X$ a.s.,

$$E(Y^2) = E(X^2) + E((Y-X)^2)\,. \tag{2}$$

Adding (1) and (2) gives

$$E(X^2) + E(Y^2) = E(X^2) + E(Y^2) + 2E((X-Y)^2)\,,$$

and this implies that $E(X-Y)^2 = 0$, which in turn implies $P(X = Y) = 1$.

Now one way to proceed is to reduce the general case of $X, Y \in L_1(P)$ to the $L_2(P)$ case treated above. Instead I will prove it using the hint. Note that

$$\begin{aligned}
&E(X-Y)1_{[X>c,Y\le c]} + E(X-Y)1_{[X\le c,Y\le c]}\\
&\quad= E(X-Y)1_{[Y\le c]} = E(X1_{[Y\le c]}) - E(Y1_{[Y\le c]})\\
&\quad= E(E(X1_{[Y\le c]}|Y)) - E(Y1_{[Y\le c]})\\
&\quad= E(1_{[Y\le c]}E(X|Y)) - E(Y1_{[Y\le c]})\\
&\quad= E(1_{[Y\le c]}Y) - E(Y1_{[Y\le c]}) = 0 \qquad (3)
\end{aligned}$$

using $E(X|Y) = Y$ a.s. in the last line. Similarly, reversing the roles of X and Y,

$$\begin{aligned}
&E(Y-X)1_{[Y>c,X\le c]} + E(Y-X)1_{[Y\le c,X\le c]}\\
&\quad= E(Y-X)1_{[X\le c]} = 0\,. \qquad (4)
\end{aligned}$$

Adding (3) and (4) yields

$$\begin{aligned}
0 &= E(X-Y)1_{[X>c,Y\le c]} + E(X-Y)1_{[X\le c,Y\le c]}\\
&\qquad - E(X-Y)1_{[Y>c,X\le c]} - E(X-Y)1_{[Y\le c,X\le c]}\\
&= E(X-Y)1_{[X>c,Y\le c]} - E(X-Y)1_{[Y>c,X\le c]}\,.
\end{aligned}$$

Since $[X-Y>0] = [X>Y] = \cup_{q\in\mathbf{Q}}[X>q\ge Y]$ and similarly for $[X-Y<0]$, this yields, by summing over rationals q,

$$0 = E(X-Y)1_{[X-Y>0]} - E(X-Y)1_{[X-Y<0]} = E|X-Y|\,.$$

But this implies $P(|X-Y| = 0) = 1$, or $P(X=Y) = 1$.

5. Exercise E9.1, PwMG, page 231: show that if $X \in L_1(\Omega, \mathcal{F}, P)$, $Y \in L_1(\Omega, \mathcal{G}, P)$, and $E(X1_G) = E(Y1_G)$ for all G in a $\overline{\pi}-$system $\mathcal{G}_0$ generating $\mathcal{G}$, then the same equality holds for all $G \in \mathcal{G}$. (And then $Y = E(X|\mathcal{G})$!)

Solution: One way to do this is from the $\pi - \lambda-$theorem: Let the $\overline{\pi}-$system be denoted by $\mathcal{G}_0$. Let $\mathcal{H} \equiv \{G \in \mathcal{F} : E(X1_G) = E(Y1_G)\}$. Then $\mathcal{G}_0 \subset \mathcal{H}$. Suppose that $A, B \in \mathcal{H}$ with $A \supset B$. Then

$$\begin{aligned} E(X1_{A\setminus B}) &= E(X(1_A - 1_B)) = E(X1_A) - E(X1_B) \\ &= E(Y1_A) - E(Y1_B) = E(Y(1_A - 1_B)) = E(Y1_{A\setminus B}) \end{aligned}$$

and hence $A\setminus B \in \mathcal{H}$. Now suppose that $\{A_n\} \subset \mathcal{H}$ with $A_n \nearrow A$. Then $E(X1_{A_n}) = E(Y1_{A_n})$ for each n, and we have both $X1_{A_n} \to X1_A$, and $Y1_{A_n} \to Y1_A$, with $|X1_{A_n}| \le |X|$ and $|Y1_{A_n}| \le |Y|$ where $X, Y \in L_1$. Hence, by the above equality for each n and the dominated convergence theorem twice,

$$E(X1_A) = \lim_n E(X1_{A_n}) = \lim_n E(Y1_{A_n}) = E(Y1_A).$$

Therefore $A \in \mathcal{H}$. Thus $\mathcal{H}$ is a $\lambda-$system. So by the $\pi - \lambda-$ theorem, $\sigma[\mathcal{G}_0] = \mathcal{G} \subset \mathcal{H}$; i.e. the equality holds for all $G \in \mathcal{G}$.