EPFL IN, SC | PROBABILITY and STATISTICS
Semester 4 – Spring 2018 | A. C. Davison

TEST - solutions

16 April 2018

Ex. 1 **a)** [*2 marks*] Let c_1 and c_2 denote the outcomes of the first and second coin respectively. The sample space is $\Omega = \{(c_1, c_2) : c_1, c_2 \in \{H, T\}\} = \{(H,H),(T,T),(H,T),(T,H)\}$. The probabilities assigned to the elements of Ω are $\Pr\{(H,H)\} = p^2$, $\Pr\{(T,T)\} = (1-p)^2$ and $\Pr\{(H,T)\} = \Pr\{(T,H)\} = p(1-p)$.

b) [*3 marks*] We have $A_1 = \{(H,H),(H,T)\}$, $A_2 = \{(H,H),(T,H)\}$ and $B = \{(H,T),(T,H)\}$. The corresponding probabilities are

$$\Pr(A_1) = p^2 + p(1-p) = p, \quad \Pr(A_2) = p^2 + p(1-p) = p, \quad \Pr(B) = 2p(1-p).$$

c) [*3 marks*] In this case $\Pr(A_1) = \Pr(A_2) = \Pr(B) = 1/2$, and

$$\Pr(A_1 \cap A_2) = \Pr(A_1 \cap B) = \Pr(A_2 \cap B) = 1/4,$$

so they are pairwise independent. However

$$0 = \Pr(A_1 \cap A_2 \cap B) \neq \Pr(A_1)\Pr(A_2)\Pr(B)$$

so they are not mutually independent.

d) [*2 marks*] We have

$$\Pr(A_1 \mid B) = \frac{\Pr(A_1 \cap B)}{\Pr(B)} = \frac{\Pr\{(H,T)\}}{2p(1-p)} = \frac{1}{2}.$$

Even if the coin is biased, since $\Pr(A_1 \mid B) = \Pr(A_2 \mid B) = 1/2$, one can get the result of a fair coin by performing the experiment until B occurs, and then checking whether the first coin shows a head.

Ex. 2 **a)** [*3 marks*] Let X, G, B denote the negative daily percent returns on a day, on an ordinary day and on a panic day respectively, and let p_G, p_B denote the probability of having an ordinary or panic day respectively. Since

$$\Pr(B > 4) = \int_4^\infty \tfrac{8}{x^3}\mathrm{d}x = \Big[-\tfrac{4}{x^2}\Big]_4^\infty = \tfrac{1}{4},$$

the law of total probability gives

$$\begin{aligned}
\Pr(X > 4) &= p_G \Pr(G > 4) + p_B \Pr(B > 4) \\
&= p_G \Pr\left(\tfrac{G}{\sqrt{4}} > \tfrac{4}{\sqrt{4}}\right) + p_B \Pr(B > 4) \\
&= 0.99\Phi(-4/\sqrt{4}) + 0.01 \times \tfrac{1}{4} \\
&\approx 0.02507.
\end{aligned}$$

b) [*3 marks*] Let H denote the event of a panic day. Our definition in part (a) gives $\Pr(X > 4 \mid H) = \Pr(B > 4)$. We also have $\Pr(H) = p_B = 0.01$, so Bayes' theorem gives

$$\Pr(H \mid X > 4) = \frac{\Pr(X > 4 \mid H)\Pr(H)}{\Pr(X > 4)} = \frac{\Pr(B > 4)\Pr(H)}{\Pr(X > 4)} = \frac{1/4 \times 0.01}{0.02507} \approx 0.09972.$$

c) [*3 marks*] The expectation of B is

$$\mathrm{E}(B) = \int_2^\infty \tfrac{8}{x^2}\mathrm{d}x = \Big[-\tfrac{8}{x}\Big]_2^\infty = 4,$$

but since

$$\mathrm{E}(B^2) = \int_2^\infty \tfrac{8}{x}\mathrm{d}x = \infty,$$

the variance of B is infinite. Because of this, many large losses are possible on a panic day.

d) [*4 marks*] With $y > x > 2$, we have

$$\Pr(B > y \mid B > x) = \frac{\Pr(B > y)}{\Pr(B > x)} = \frac{4/y^2}{4/x^2} = \frac{x^2}{y^2}.$$

Hence, the conditional distribution function of B, given that $B > x$ is

$$\Pr(B \le y \mid B > x) = 1 - \tfrac{x^2}{y^2}, \quad y > x,$$

and the conditional density is

$$f_{B|B>x}(y \mid x) = \frac{\mathrm{d}(1 - x^2/y^2)}{\mathrm{d}y} = \frac{2x^2}{y^3}, \quad y > x.$$

Hence, the conditional expectation is

$$\mathrm{E}(B \mid B > x) = \int_x^\infty y\frac{2x^2}{y^3}\mathrm{d}y = \frac{2x^2}{x} = 2x.$$

Note that this tends to 4 if $x \to 2$, as we should expect from part c).

Ex. 3 **a)** [*2 marks*] For $f_{X,Y}$ to be a density function, it must be non-negative (obvious), and must satisfy

$$\int\int f_{X,Y}(x, y)\,\mathrm{d}y\mathrm{d}x = \int_0^3\int_0^3 c(x + y)\mathrm{d}y\mathrm{d}x = 1;$$

integrating yields $c = 1/27$.

b) [*2 marks*] Here

$$\Pr(X + Y \le 3) = \int_0^3\int_0^{3-x} \tfrac{1}{27}(x + y)\mathrm{d}y\mathrm{d}x = \tfrac{1}{3}.$$

c) [*3 marks*] The marginal density of X is

$$f_X(x) = \int_0^3 f_{X,Y}(x, y)\mathrm{d}y = \int_0^3 \tfrac{1}{27}(x + y)\mathrm{d}y = \frac{6x + 9}{54} = \tfrac{x}{9} + \tfrac{1}{6}, \quad 0 < x < 3.$$

By symmetry, $f_Y(y) = \frac{y}{9} + \frac{1}{6}$ $(0 < y < 3)$. Since $f_{X,Y}(x, y) \ne f_X(x)f_Y(y)$, X and Y are not independent (which is obvious anyway, since $x + y$ does not factorise as a function of x and a function of y).

d) [*4 marks*] For any $x \in (0, 3)$, we have

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{(x + y)/27}{(6x + 9)/54} = 2\frac{(x + y)}{6x + 9}, \quad 0 < y < 3,$$

so the expectation of $Y \mid X = x$ is

$$\mathrm{E}(Y \mid X = x) = \tfrac{2}{6x+9}\int_0^3 y(x + y)\mathrm{d}y = \tfrac{2}{6x+9}\Big[\tfrac{xy^2}{2} + \tfrac{y^3}{3}\Big]_0^3 = \frac{9x + 18}{6x + 9} = \frac{3x + 6}{2x + 3}, \quad x \in (0, 3).$$

Setting $x = 2$ gives $\mathrm{E}(Y \mid X = 2) = 12/7$.

Ex. 4 **a)** [*3 marks*] Here is one possible formulation (we accept other equivalent ones). Let D_j denote the document being in box $j \in \{1,2,3\}$. Let S_j denote the event where box j has been searched but the document was not found. From the information given, since $\Pr(D_1) = \Pr(D_2) = \Pr(D_3) = \frac{1}{3}$ and

$$\Pr(S_1) = \Pr(S_1 \mid D_1)\Pr(D_1) + \Pr(S_1 \mid D_2)\Pr(D_2) + \Pr(S_1 \mid D_3)\Pr(D_3) = (1-p_1)\tfrac{1}{3} + 1\times\tfrac{1}{3} + 1\times\tfrac{1}{3},$$

we see that

$$\Pr(D_1 \mid S_1) = \frac{\Pr(D_1 \cap S_1)}{\Pr(S_1)} = \frac{\Pr(S_1 \mid D_1)\Pr(D_1)}{\Pr(S_1)} = \frac{(1-p_1)\times\frac{1}{3}}{(1-p_1)\frac{1}{3} + \frac{1}{3} + \frac{1}{3}} = \frac{1-p_1}{3-p_1}.$$

b) [*3 marks*] Let B denote the event where boxes 1, 2 and 3 have been searched but the document has not been found. Clearly, $B = S_1 \cap S_2 \cap S_3$. Since the S_j's are conditionally independent given D_1,

$$\begin{aligned}
\Pr(D_1 \mid B) &= \frac{\Pr(B \mid D_1)\Pr(D_1)}{\Pr(B)} = \frac{\Pr(S_1 \mid D_1)\Pr(S_2 \mid D_1)\Pr(S_3 \mid D_1)\Pr(D_1)}{\Pr(B)} \\
&= \frac{\frac{1}{3}(1-p_1)\times 1 \times 1}{\frac{1}{3}\{1-p_1+1-p_2+1-p_3\}} = \frac{1-p_1}{3-p_1-p_2-p_3}
\end{aligned}$$

c) [*4 marks*] If the number of searches is N, then we seek

$$\mathrm{E}(N) = \mathrm{E}(N \mid D_1)\Pr(D_1) + \mathrm{E}(N \mid D_2)\Pr(D_2) + \mathrm{E}(N \mid D_3)\Pr(D_3).$$

Now

$$\begin{aligned}
\mathrm{E}(N \mid D_1) &= \sum_{r=0}^{\infty}(1+3r)p_1(1-p_1)^r \\
&= \sum_{r=0}^{\infty} p_1(1-p_1)^r + 3\sum_{r=0}^{\infty} rp_1(1-p_1)^r \\
&= 1 + 3\{\mathrm{E}(G_1) - 1\} \\
&= 1 + 3(1/p_1 - 1) \\
&= 3/p_1 - 2,
\end{aligned}$$

where G_1 is a geometric variable with probability p_1. Equivalently, the hint gives

$$\sum_{r=0}^{\infty} rp_1(1-p_1)^r = p_1(1-p_1)\sum_{r=1}^{\infty} r(1-p_1)^{r-1} = \frac{p_1(1-p_1)}{p_1^2} = 1/p_1 - 1.$$

Similar calculations will clearly yield

$$\mathrm{E}(N \mid D_2) = 1 + (3/p_2 - 2) = 3/p_2 - 1, \quad \mathrm{E}(N \mid D_3) = 2 + (3/p_3 - 2) = 3/p_3.$$

Since $\Pr(D_j = 1/3)$,

$$\begin{aligned}
\mathrm{E}(N) &= (3/p_1 - 2)\times\tfrac{1}{3} + (3/p_2 - 1)\times\tfrac{1}{3} + (3/p_3)\times\tfrac{1}{3} \\
&= 1/p_1 + 1/p_2 + 1/p_3 - 1.
\end{aligned}$$