All of Statistics - Chapter 2 Solutions

May 1, 2020

Acknowledgements: Thanks to Ben S. for correcting some mistakes.

1.

By Lemma 2.15, $\mathbb{P}(X = x) = F(x) - F(x-)$. Since F is right-continuous, $F(x) = F(x+)$.

2.

By Lemma 2.15,

$$\mathbb{P}(2 < X \leq 4.8) = F(4.8) - F(2) = 1/10$$

and

$$\mathbb{P}(2 \leq X \leq 4.8) = \mathbb{P}(X = 2) + \mathbb{P}(2 < X \leq 4.8) = F(4.8) - F(2-) = 2/10.$$

3.

1)

Since F is monotone, we can write $F(x-) = \lim_n F(x_n)$ where (x_n) is some strictly increasing sequence converging to x. Let $A_n = \{X \leq x_n\}$ so that $\{X < x\} = \cup_n A_n$. By continuity of probability, $\mathbb{P}(X < x) = \lim_k \mathbb{P}(A_n) = \lim_n F(x_n)$.

2)

By additivity, $\mathbb{P}(X \leq x) + \mathbb{P}(x < X \leq y) = \mathbb{P}(X \leq y)$. The desired result follows by moving some terms around.

3)

Taking complements, $\mathbb{P}(X > x) = 1 - \mathbb{P}(X \leq x) = 1 - F(x)$.

4)

If X is continuous, $\mathbb{P}(X = x) = 0$ for all x by Part 1. The desired result follows from combining this fact with the findings from Part 2.

4.

a)

We can express the CDF using indicator functions:

$$F_X(x) = \frac{x}{4} I_{[0,1)}(x) + \frac{1}{4} I_{[1,\infty)}(x) + \frac{3}{8}(x-3)\, I_{[3,5)}(x) + \frac{3}{4} I_{[5,\infty)}(x).$$

b)

Since $Y = 1/X$ and $F_X(0) = 0$, it follows that $F_Y(0) = 0$. For $y > 0$,

$$F_Y(y) = \mathbb{P}(X \geq 1/y) = 1 - \mathbb{P}(X < 1/y) = 1 - F_X(1/y).$$

5.

Suppose X and Y are independent. Then,

$$f_{X,Y}(x,y) = \mathbb{P}(X \in \{x\}, Y \in \{y\}) = \mathbb{P}(X \in \{x\})\mathbb{P}(Y \in \{Y\}) = f_X(x) f_Y(y).$$

To establish the converse, suppose that $f_{X,Y} = f_X f_Y$. For a subset A of the support of X and a subset B of the support of Y,

$$\begin{aligned} \mathbb{P}(X \in A, Y \in B) &= \sum_{(x,y) \in A \times B} f_{X,Y}(x,y) = \sum_{x \in A} f_X(x) \sum_{y \in B} f_Y(y) \\ &= \mathbb{P}(X \in A)\mathbb{P}(Y \in B). \end{aligned}$$

6.

Note that

$$F_Y(y) = \begin{cases} 0 & \text{if } y < 0 \\ \mathbb{P}(X \notin A) & \text{if } 0 \leq y < 1 \\ 1 & \text{if } y \geq 1. \end{cases}$$

7.

Since

$$\mathbb{P}(Z > z) = \mathbb{P}(\min\{X, Y\} > z) = \mathbb{P}(X > z)\mathbb{P}(Y > z) = (1 - F_X(z))\,(1 - F_Y(z))\,,$$

it follows that

$$F_Z(z) = 1 - (1 - F_X(z))(1 - F_Y(z)) = F_X(z) + F_Y(z) - F_X(z)F_Y(z).$$

When X and Y have the same distribution F, $F_Z(z) = 2F(z) - F(z)^2$ and hence $f_Z(z) = 2f(z) - 2F(z)f(z)$. In particular, when F is a uniform distribution on $(0,1)$,

$$f_Z(z) = 2\,(1 - z)\, I_{(0,1)}(z).$$

8.

Let $Y = X^+$. First, note that $F_Y(0-) = 0$ and $F_Y(0) = F_X(0)$. Moreover, $F_Y(x) = F_X(x)$ for $x > 0$.

9.

For $x > 0$, $F_X(x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$. Therefore, $F^{-1}(q) = -\ln(1-q)/\lambda$.

10.

If X and Y are independent, then

$$\begin{aligned}\mathbb{P}(g(X) \in A, h(Y) \in B) &= \mathbb{P}(X \in g^{-1}(A), Y \in h^{-1}(B)) \\ &= \mathbb{P}(X \in g^{-1}(A))\mathbb{P}(Y \in h^{-1}(B)) = \mathbb{P}(g(X) \in A)\mathbb{P}(h(Y) \in B)\end{aligned}$$

under some lax conditions on g and h (Borel measurable).

11.

a)

The two variables are dependent because

$$\mathbb{P}(X = 1, Y = 0) = 0 \neq p(1-p) = \mathbb{P}(X = 1)(Y = 0).$$

b)

The two variables are independent because

$$\mathbb{P}(X = i, Y = j) = \frac{\lambda^{i+j} e^{-\lambda}}{(i+j)!} \binom{i+j}{i} p^i (1-p)^j = e^{-\lambda} \frac{\lambda^i p^i}{i!} \frac{\lambda^j (1-p)^j}{j!}$$

is decomposable into the form $g(i)h(j)$.

12.

If X and Y admit a joint density satisfying $f(x, y) = g(x)h(y)$, then

$$\mathbb{P}(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s,t) dt ds = \int_{-\infty}^{x} g(s) ds \int_{-\infty}^{y} h(t) dt.$$

The marginal distribution for X is $\mathbb{P}(X \leq x) = c_h \int_{-\infty}^{x} g(s) ds$ where $c_h = \int_{-\infty}^{\infty} h(t) dt$. It follows that $f_X = h c_h$. We can similarly define c_g to find that $f_Y = g c_g$. Moreover, $c_h c_g = 1$ and hence $c_g = 1/c_h$. It follows that $f_{X,Y} = f_X f_Y$, as desired.

13.

a)

Note that

$$F_Y(y) = \mathbb{P}(e^X \le y) = \mathbb{P}(X \le \ln y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\ln y} \exp\left(-\frac{x^2}{2}\right) dy.$$

Taking derivatives,

$$f_Y(y) = \frac{1}{y\sqrt{2\pi}} \exp\left(-\frac{(\ln y)^2}{2}\right).$$

b)

TODO (Computer Experiment)

14.

Let $0 < r < 1$. Then, $F_R(r) = \pi r^2/\pi = r^2$ and hence $f_R(r) = 2r$.

15.

For $0 \le y \le 1$,

$$F_Y(y) = \mathbb{P}(F(X) \le y) = \mathbb{P}(X \le F^{-1}(y)) = F(F^{-1}(y)) = y.$$

For all x,

$$F_X(x) = \mathbb{P}(F^{-1}(U) \le x) = \mathbb{P}(U \le F(x)) = F(x).$$

16.

Note that

$$\mathbb{P}(X = x \mid X + Y = n) = \frac{\mathbb{P}(X = x, Y = n - x)}{\mathbb{P}(X + Y = n)}.$$

Moreover,

$$\mathbb{P}(X = x, Y = n - x) = \frac{e^{-\lambda}\lambda^x}{x!} \frac{e^{-\mu}\mu^{n-x}}{(n-x)!}.$$

As per the hint,

$$\mathbb{P}(X + Y = n) = e^{-\lambda-\mu} \frac{(\lambda + \mu)^n}{n!}.$$

Letting $\pi = \lambda/(\lambda+\mu)$, combining these facts yields

$$\mathbb{P}(X = x \mid X + Y = n) = \binom{n}{x} \pi^x (1-\pi)^{n-x}.$$

17.

First, note that

$$f_Y(1/2) = \int_0^1 f(x, 1/2) dx = c \int_0^1 \left(x + \frac{1}{4}\right) dx = \frac{3}{4} c.$$

Therefore,

$$f_{X|Y}(x \mid 1/2) = \frac{f_{X,Y}(x, 1/2)}{f_Y(1/2)} = \frac{4}{3}\left(x + \frac{1}{4}\right) I_{(0,1)}(x).$$

It follows that

$$\mathbb{P}(X < 1/2 \mid Y = 1/2) = \frac{4}{3} \int_0^{1/2} \left(x + \frac{1}{4}\right) dx = \frac{1}{3}.$$

18.

TODO (Computer Experiment)

19.

Let r be strictly increasing with differentiable inverse s. Let X be an (absolutely) continuous random variable. Then, for $Y = r(X)$,

$$F_Y(y) = \mathbb{P}(r(X) \leq y) = \mathbb{P}(X \leq s(y)) = F_X(s(y))$$

and hence $f_Y(y) = f_X(s(y))s'(y)$. If r was instead strictly decreasing, then

$$F_Y(y) = \mathbb{P}(X \geq s(y)) = 1 - F_X(s(y))$$

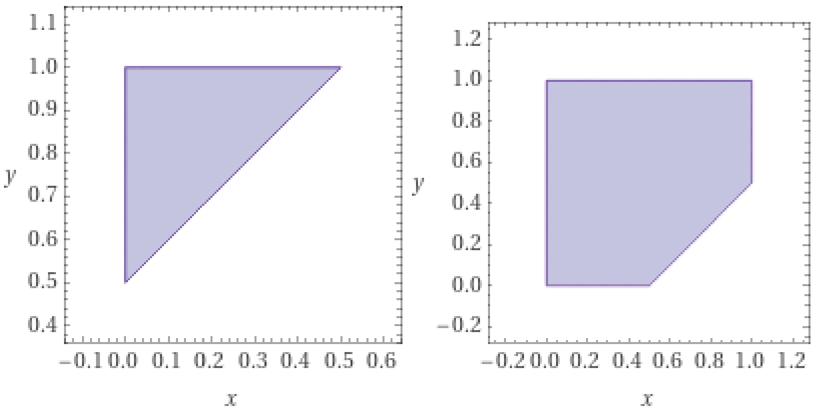
and hence $f_Y(y) = -f_X(s(y))s'(y)$. Since a strictly decreasing function has a strictly decreasing inverse, it follows that $s'(y) < 0$ and hence we can summarize both cases by $f_Y = (f_X \circ s)|s'|$.

20.

Let $W = X - Y$. Then, $F_W(-1) = 0$ and $F_W(1) = 1$. For $-1 < w < 1$, $F_W(w) = \mathbb{P}(Y \geq X - w)$. The region

$$\{(x, y) : y \geq x - w, 0 \leq x, y \leq 1\}$$

is either a triangle or a right trapezoid depending on whether $-1 < w < 0$ or $0 < w < 1$:



By covering these case separately, one can derive $F_W(w) = (1+w)^2/2$ and $F_W(w) = -w^2/2 + w + 1/2$, respectively. It follows that

$$f_W(w) = \begin{cases} 1+w & \text{if } -1 < w < 0 \\ 1-w & \text{if } 0 < w < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Let $V = X/Y$. Then, $F_V(0) = 0$. For $v > 0$, $F_V(v) = \mathbb{P}(Y \geq X/v)$. The region

$$\left\{(x,y): y \geq \frac{x}{v}, 0 \leq x, y \leq 1\right\}$$

is either a triangle or a rectangle plus a right trapezoid depending on whether $0 < v < 1$ or $v > 1$. By covering these cases separately, one can derive $F_V(v) = 2/v$ and $F_V(v) = 1/(2v) + (1 - 1/v)$, respectively. It follows that

$$f_V(v) = \begin{cases} 1/2 & \text{if } 0 < v < 1 \\ 1/(2v^2) & \text{if } v > 1 \\ 0 & \text{otherwise.} \end{cases}$$

21.

Since

$$F_Y(y) = \mathbb{P}(\max\{X_1, \ldots, X_n\} \leq y) = \mathbb{P}(X_1 \leq y)^n = \left(1 - e^{-\beta y}\right)^n,$$

it follows that $f_Y(y) = \beta n e^{-\beta y}(1 - e^{-\beta y})^{n-1}$.