Additional Practice Questions 1 (Partial) Solutions

2.6 (a) Solution not provided.

2.16

x^2 is a convex function and, therefore, Jensen's inequality implies that

$$\mathrm{E}[X^2] \geq \mathrm{E}[X]^2 = 1$$

To show that the inequality is strict, notice that, if X is not degenerate, it implies that $\mathrm{var}(X) > 0$ and, therefore, that

$$\mathrm{E}[X^2] - \mathrm{E}[X]^2 > 0$$

which implies that $\mathrm{E}[X^2] > \mathrm{E}[X]^2 = 1$.

2.21 [This is just a sketch of the solution]

a) For the categorical wage data in the problem, there are five possible values that wage takes: 0, 10, 20, 30, 40, 50.

   Mass on right:

$$\begin{aligned} \mathrm{F}(0) &= 0 & \pi(0) &= 0 \\ \mathrm{F}(10) &= 0.1 & \pi(10) &= 0.1 \\ \mathrm{F}(20) &= 0.5 & \pi(20) &= 0.4 \\ \mathrm{F}(30) &= 0.8 & \pi(30) &= 0.3 \\ \mathrm{F}(40) &= 1 & \pi(40) &= 0.2 \end{aligned}$$

   the cdf is flat between each value of wage, and the pmf is 0 for other values of wages.

   Mass on left:

$$\begin{aligned} \mathrm{F}(0) &= 0.1 & \pi(0) &= 0.1 \\ \mathrm{F}(10) &= 0.5 & \pi(10) &= 0.4 \\ \mathrm{F}(20) &= 0.8 & \pi(20) &= 0.3 \\ \mathrm{F}(30) &= 1 & \pi(30) &= 0.2 \\ \mathrm{F}(40) &= 1 & \pi(40) &= 0 \end{aligned}$$

   the cdf is flat between each value of wage, and the pmf is 0 for other values of wages.

   For the true cdf of wages, we can say that it sits in between the two cdfs above.

b) For the first discrete distribution,

$$\mathrm{E}[X] = 0(0) + 10(0.1) + 20(0.4) + 30(0.3) + 40(0.2) = 26$$

   For the second discrete distribution,

$$\mathrm{E}[X] = 0(0.1) + 10(0.4) + 20(0.3) + 30(0.2) + 40(0) = 16$$

c) For this distribution, the cdf just connects the "mass on the right" cdf with straight lines between the points. For example, $\mathrm{F}(5) = 0.05$ and $\mathrm{F}(25) = 0.65$. The pdf is constant in between each cutoff and takes the value of $(\mathrm{F}(\text{upper cutoff}) - \mathrm{F}(\text{lower cutoff}))/10$. For example, $f(5) = (0.1 - 0)/10 = 0.01$, and $f(25) = (0.8 - 0.5)/10 = 0.03$. And we can calculate $\mathrm{E}[X]$ as

$$
\begin{aligned}
\mathrm{E}[X] &= \int_0^{40} x f(x)\, dx \\
&= \int_0^{10} x\, 0.01\, dx + \int_{10}^{20} x\, 0.04\, dx + \int_{20}^{30} x\, 0.03\, dx + \int_{30}^{40} x\, 0.02\, dx \\
&= \frac{0.01}{2} x^2 \Big|_0^{10} + \frac{0.04}{2} x^2 \Big|_{10}^{20} + \frac{0.03}{2} x^2 \Big|_{20}^{30} + \frac{0.02}{2} x^2 \Big|_{30}^{40} \\
&= \frac{0.01}{2}\Big(100 - 0\Big) + \frac{0.04}{2}\Big(400 - 100\Big) + \frac{0.03}{2}\Big(900 - 400\Big) + \frac{0.02}{2}\Big(1600 - 900\Big) \\
&= 0.5 + 6 + 7.5 + 7 = 21
\end{aligned}
$$

This is sort of a tedious calculation, but notice that this is exactly halfway between the $\mathrm{E}[X]$ under "mass on the right" and "mass on the left", which, given that we are supposing that the distribution is uniform between the cutoffs, is exactly what you would expect.

3.1

a)

$$\sum_{x=0}^{1} \pi(x|p) = (1-p) + p = 1$$

b)

$$
\begin{aligned}
\mathrm{E}[X] &= \sum_{x=0}^{1} x \pi(x|p) \\
&= 0(1-p) + 1(p) \\
&= p
\end{aligned}
$$

c)

$$
\begin{aligned}
\mathrm{var}(X) &= \sum_{x=0}^{1} (x - \mathrm{E}[X])^2 \pi(x|p) \\
&= (0-p)^2(1-p) + (1-p)^2 p \\
&= p\Big(p(1-p)\Big) + (1-p)\Big(p(1-p)\Big) \\
&= p(1-p)
\end{aligned}
$$

where the third equality factors out p from the first term and $1-p$ from the second term, and the fourth equality adds these up. [Side-comment: I calculated this from the definition of $\mathrm{var}(X)$, but it is perhaps somewhat easier to calculate $\mathrm{E}[X^2]$ and then use $\mathrm{var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2$ – this will give you the same answer.]

3.4

a)

$$\begin{aligned}\int_a^b \frac{1}{b-a}\,dx &= \frac{1}{b-a}\int_a^b 1\,dx\\ &= \frac{1}{b-a}x\Big|_a^b\\ &= \frac{b-a}{b-a} = 1\end{aligned}$$

b)

$$\begin{aligned}\mathrm{E}[X] &= \int_a^b x\frac{1}{b-a}\,dx\\ &= \frac{1}{b-a}\frac{1}{2}x^2\Big|_a^b\\ &= \frac{b^2-a^2}{2(b-a)}\\ &= \frac{(b-a)(b+a)}{2(b-a)}\\ &= \frac{a+b}{2}\end{aligned}$$

c) First, let's calculate $\mathrm{E}[X^2]$.

$$\begin{aligned}\mathrm{E}[X^2] &= \int_a^b x^2\frac{1}{b-a}\,dx\\ &= \frac{1}{b-a}\frac{1}{3}x^3\Big|_a^b\\ &= \frac{b^3-a^3}{3(b-a)}\\ &= \frac{(b-a)(a^2+2ab+b^2)}{3(b-a)}\\ &= \frac{a^2+ab+b^2}{3}\end{aligned}$$

Thus,

$$\begin{aligned}\mathrm{var}(X) &= \mathrm{E}[X^2]-\mathrm{E}[X]^2\\ &= \frac{a^2+ab+b^2}{3}-\left(\frac{a+b}{2}\right)^2\\ &= \frac{4(a^2+ab+b^2)-3(a^2+2ab+b^2)}{12}\\ &= \frac{a^2-2ab+b^2}{12}\\ &= \frac{(a-b)^2}{12}\end{aligned}$$

4.2 Solution not provided.

4.5

[As a side-comment, to me this problem seems pretty straightforward to understand by drawing a graph, but somewhat harder to prove (though perhaps you can figure out a shorter solution than what I provide below).]

To start with, we will show that $\mathrm{P}(a < X \leq b, Y \leq c) = \mathrm{F}(b,c) - \mathrm{F}(a,c)$. To see this, notice that,

$$\mathrm{P}(Y \leq c) = \mathrm{P}(X > b, Y \leq c) + \mathrm{P}(a < X \leq b, Y \leq c) + \mathrm{P}(X \leq a, Y \leq c)$$

which covers all possible values that X can take. Moreover, using a similar sort of argument

$$\mathrm{P}(Y \leq c) = \mathrm{P}(X > b, Y \leq c) + \mathrm{P}(X \leq b, Y \leq c)$$

Plugging this expression in above and canceling terms implies that

$$\begin{aligned} \mathrm{P}(X \leq b, Y \leq c) &= \mathrm{P}(a < X \leq b, Y \leq c) + \mathrm{P}(X \leq a, Y \leq c) \\ \implies \mathrm{P}(a < X \leq b, Y \leq c) &= \mathrm{F}(b,c) - \mathrm{F}(a,c) \end{aligned}$$

which holds by re-arranging terms and the definition of joint cdf. This is the result that we were trying to show.

Next, notice that we can write

$$\begin{aligned} \mathrm{F}(b,d) &= \mathrm{P}(X \leq b, Y \leq d) \\ &= \mathrm{P}(a < X \leq b, Y \leq d) + \mathrm{P}(X \leq a, Y \leq d) \\ &= \mathrm{P}(a < X \leq b, c < Y \leq d) + \mathrm{P}(a < X \leq b, Y \leq c) + \mathrm{P}(X \leq a, Y \leq d) \\ &= \mathrm{P}(a < X \leq b, c < Y \leq d) + \mathrm{F}(b,c) - \mathrm{F}(a,c) + \mathrm{F}(a,d) \end{aligned}$$

where the first equality is just the definition of cdf, the second and third equalities hold by repeatedly splitting up the regions of X and Y for which the cdfs are being calculated, and the last equality holds by the definition of cdf. Re-arranging the terms from the last line implies that

$$\mathrm{P}(a < X \leq b, c < Y \leq d) = \mathrm{F}(b,d) - \mathrm{F}(a,d) - \mathrm{F}(b,c) + \mathrm{F}(a,c)$$

which is the result that we were trying to show.

4.6

a) We need to find a c such that $\int_0^1 \int_0^{1-y} cxy \, dx \, dy = 1$

$$\begin{aligned} \int_0^1 \int_0^{1-y} cxy \, dx \, dy &= \int_0^1 \frac{cy}{2} x^2 \Big|_0^{1-y} dy \\ &= \int_0^1 \frac{cy}{2} (1-y)^2 \, dy \\ &= \frac{c}{2} \int_0^1 (y - 2y^2 + y^3) \, dy \\ &= \frac{c}{2} \left(\frac{1}{2} y^2 - \frac{2}{3} y^3 + \frac{1}{4} y^4 \right) \Big|_0^1 \\ &= \frac{c}{2} \left(\frac{1}{12} \right) \\ &= \frac{c}{24} \end{aligned}$$

Thus, c must be equal to 24.

b)

$$\begin{aligned} f_X(x) &= \int_0^{1-x} 24xy\,dy \\ &= \frac{24}{2}xy^2\Big|_0^{1-x} \\ &= 12x(1-x)^2 \\ &= 12(x-2x^2+x^3) \end{aligned}$$

Using the same argument implies that $f_Y(y) = 12(y - 2y^2 + y^3)$.

c) No, they are not independent. Notice that $f(x,y) \neq f_X(x)\, f_Y(y)$ (you can immediately see that the right side will involve higher order terms like $24x^3y^3$). You can also see intuitively that they will not be independent. For example, if you know that $Y = 0.9$ it means that X cannot take a value greater than 0.1, which suggests that they are dependent.

4.11 To calculate the correlation, we'll calculate the covariance of XY and Y and then the variance of XY.

$$\begin{aligned} \mathrm{cov}(XY, Y) &= \mathrm{E}[XY^2] - \mathrm{E}[XY]\mathrm{E}[Y] \\ &= \mathrm{E}[X]\mathrm{E}[Y^2] - \mathrm{E}[X]\mathrm{E}[Y]^2 \\ &= \mathrm{E}[X](\mathrm{E}[Y^2] - \mathrm{E}[Y]^2) \\ &= \mu_X \sigma_Y^2 \end{aligned}$$

where the second equality uses independence. Next,

$$\begin{aligned} \mathrm{var}(XY) &= \mathrm{E}[X^2Y^2] - \mathrm{E}[XY]^2 \\ &= \mathrm{E}[X^2]\mathrm{E}[Y^2] - \mathrm{E}[X]^2\mathrm{E}[Y]^2 \\ &= (\mathrm{var}(X) + \mathrm{E}[X]^2)(\mathrm{var}(Y) + \mathrm{E}[Y]^2) - \mathrm{E}[X]^2\mathrm{E}[Y]^2 \\ &= \Big(\mathrm{var}(X)\mathrm{var}(Y) + \mathrm{var}(X)\mathrm{E}[Y]^2 + \mathrm{var}(Y)\mathrm{E}[X]^2 + \mathrm{E}[X]^2\mathrm{E}[Y]^2\Big) - \mathrm{E}[X]^2\mathrm{E}[Y]^2 \\ &= \sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \sigma_Y^2\mu_X^2 \end{aligned}$$

where the second equality uses independence and third equality uses $\mathrm{var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2$.

Putting this altogether,

$$\mathrm{corr}(X, Y) = \frac{\mu_X \sigma_Y^2}{\sqrt{\sigma_Y^2}\,\sqrt{\sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \sigma_Y^2\mu_X^2}}$$