Math 431, Homework 11 Solutions

1. **Exercise 7.1**

   The marginal probability mass functions of X and Y are

   $$\begin{aligned} p_X(k) &= e^{-2}\tfrac{2^k}{k!} \quad \text{if } k \geq 0 \\ p_Y(k) &= \left(\tfrac{1}{3}\right)^{k-1}\tfrac{2}{3} \quad \text{if } k \geq 1. \end{aligned}$$

   Thus,

   $$\begin{aligned} P(Z=3) &= P(X+Y=3) = P(X=0, Y=3) + P(X=1, Y=2) + P(X=2, Y=1) \\ &= P(X=0)P(Y=3) + P(X=1)P(Y=2) + P(X=2)P(Y=1) \\ &= e^{-2}\left(\tfrac{1}{3}\right)^2\tfrac{2}{3} + e^{-2}2\cdot\tfrac{1}{3}\cdot\tfrac{2}{3} + e^{-2}\tfrac{2^2}{2}\cdot\tfrac{2}{3} \\ &= \tfrac{50}{27}e^{-2} \\ &\approx 0.2506. \end{aligned}$$

2. **Exercise 7.18**

   Using the convolution formula

   $$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx = \int_{-\infty}^{\infty} f_Y(y)f_X(z-y)dy$$

   Since $X, Y > 0$, the same will be true for $X+Y$. Thus $f_{X+Y}(z) = 0$ for $z \leq 0$.

   If $z > 0$ then $f_Y(y) > 0$ and $f_X(z-y) > 0$ if and only if $y > 0$ and $z - y > 0$, which gives $0 < y < z$. Thus for $z > 0$ we only need to integrate between 0 and z:

   $$\begin{aligned} f_{X+Y}(z) &= \int_{-\infty}^{\infty} f_Y(y)f_X(z-y)\,dy = \int_0^z f_Y(y)f_X(z-y)\,dy \\ &= \int_0^z 4xe^{-2x}\cdot 2e^{-2(z-x)}dx = \int_0^z 8xe^{-4z}\,dx \\ &= [4x^2e^{-4z}]_{x=0}^{z} \\ &= 4z^2e^{-4z} \end{aligned}$$

   Thus

   $$f_Z(z) = \begin{cases} 4z^2e^{-4z} & z > 0 \\ 0 & z \leq 0. \end{cases}$$

3. **Exercise 8.1**

(a)

$$E[X+Y] = EX + EY = \frac{1}{p} + nr$$

(b) This is impossible, unless we know something about the joint distribution of X and Y. We are not given such information.

(c)

$$\begin{aligned} E[X^2+Y^2] &= EX^2 + EY^2 \\ &= (\text{Var}(X) + (EX)^2) + (\text{Var}(Y) + (EY)^2) \\ &= (\frac{1-p}{p^2} + \frac{1}{p^2}) + (nr(1-r) + n^2r^2) \\ &= \frac{2-p}{p^2} + nr - nr^2 + n^2r^2 \end{aligned}$$

(d) Expanding the square gives

$$E[(X+Y)^2] = E[X^2 + 2XY + Y^2] = EX^2 + 2EXY + EY^2.$$

As explained in the solution to (b), we cannot evaluate EXY without some information about the joint distribution of X and Y.

4. **Exercise 8.6**

(a) No change to this solution.

$$E[X+Y] = EX + EY = \frac{1}{p} + nr$$

(b) With the independence of X and Y we have

$$E[XY] = E[X]E[Y] = \frac{1}{p} \cdot nr = \frac{nr}{p}.$$

(c) No change to this solution.

$$\begin{aligned} E[X^2+Y^2] &= EX^2 + EY^2 \\ &= (\text{Var}(X) + (EX)^2) + (\text{Var}(Y) + (EY)^2) \\ &= (\frac{1-p}{p^2} + \frac{1}{p^2}) + (nr(1-r) + n^2r^2) \\ &= \frac{2-p}{p^2} + nr - nr^2 + n^2r^2 \\ &= \frac{2-p}{p^2} + nr + n(n-1)r^2 \end{aligned}$$

(d) Expanding the square gives

$$E[(X+Y)^2] = E[X^2 + 2XY + Y^2] = EX^2 + 2EXY + EY^2.$$

We can use our solutions to (b) and (c) to get

$$E[(X+Y)^2] = \frac{2-p}{p^2} + nr + n(n-1)r^2 + \frac{2nr}{p}.$$

5. **Exercise 8.9**

   (a)
   $$E[3X - 2Y + 7] = 3E[X] - 2E[Y] + 7 = 3 \cdot 3 - 2 \cdot 5 + 7 = 6.$$

   (b) Because X and Y are independent,
   $$\mathrm{Var}(3X - 2Y + 7) = \mathrm{Var}(3X - 2Y) = 9\mathrm{Var}(X) + 4\mathrm{Var}(Y) = 9 \cdot 2 + 4 \cdot 3 = 30.$$

   (c) From the definition of the variance:
   $$\mathrm{Var}(XY) = E[X^2Y^2] - (E[XY])^2$$
   By the independence of X and Y:
   $$E[XY] = E[X]E[Y] = 3 \cdot 5 = 15.$$
   and
   $$E[X^2Y^2] = E[X^2]E[Y^2].$$
   We can compute the second moments from the first moment and the variance:
   $$E[X^2] = \mathrm{Var}(X) + E[X]^2 = 2 + 3^2 = 11, \quad E[Y^2] = \mathrm{Var}(Y) + E[Y]^2 = 3 + 5^2 = 28.$$
   Putting everything together:
   $$\begin{aligned}\mathrm{Var}(XY) &= E[X^2Y^2] - (E[XY])^2 = E[X^2]E[Y^2] - (E[X]E[Y])^2 \\ &= 11 \cdot 28 - 15^2 = 83.\end{aligned}$$

6. **Exercise 8.14**

   We must compute EX, EY, EX^2, EY^2, and EXY. We can use the joint PMF for all these cases.
   $$\begin{aligned}EX = 1 \cdot \frac{1}{15} + 1 \cdot \frac{1}{15} + 1 \cdot \frac{2}{15} + 1 \cdot \frac{1}{15} \\ 2 \cdot \frac{1}{10} + 2 \cdot \frac{1}{10} + 2 \cdot \frac{2}{5} + 2 \cdot \frac{1}{10} \\ 3 \cdot \frac{1}{30} + 3 \cdot \frac{1}{30} + 3 \cdot 0 + 3 \cdot \frac{1}{10} \\ = \frac{11}{6}\end{aligned}$$

   $$\begin{aligned}EX^2 = 1^2 \cdot \frac{1}{15} + 1^2 \cdot \frac{1}{15} + 1^2 \cdot \frac{2}{15} + 1^2 \cdot \frac{1}{15} \\ 2^2 \cdot \frac{1}{10} + 2^2 \cdot \frac{1}{10} + 2^2 \cdot \frac{2}{5} + 2^2 \cdot \frac{1}{10} \\ 3^2 \cdot \frac{1}{30} + 3^2 \cdot \frac{1}{30} + 3^2 \cdot 0 + 3^2 \cdot \frac{1}{10} \\ = \frac{23}{6}\end{aligned}$$

$$\begin{aligned}
EY &= 0\cdot\frac{1}{15}+1\cdot\frac{1}{15}+2\cdot\frac{2}{15}+3\cdot\frac{1}{15}\\
&\quad 0\cdot\frac{1}{10}+1\cdot\frac{1}{10}+2\cdot\frac{2}{5}+3\cdot\frac{1}{10}\\
&\quad 0\cdot\frac{1}{30}+1\cdot\frac{1}{30}+2\cdot 0+3\cdot\frac{1}{10}\\
&= \frac{5}{3}
\end{aligned}$$

$$\begin{aligned}
EY^2 &= 0^2\cdot\frac{1}{15}+1^2\cdot\frac{1}{15}+2^2\cdot\frac{2}{15}+3^2\cdot\frac{1}{15}\\
&\quad 0^2\cdot\frac{1}{10}+1^2\cdot\frac{1}{10}+2^2\cdot\frac{2}{5}+3^2\cdot\frac{1}{10}\\
&\quad 0^2\cdot\frac{1}{30}+1^2\cdot\frac{1}{30}+2^2\cdot 0+3^2\cdot\frac{1}{10}\\
&= \frac{59}{15}
\end{aligned}$$

$$\begin{aligned}
EXY &= 1\cdot 0\cdot\frac{1}{15}+1\cdot 1\cdot\frac{1}{15}+1\cdot 2\cdot\frac{2}{15}+1\cdot 3\cdot\frac{1}{15}\\
&\quad 2\cdot 0\cdot\frac{1}{10}+2\cdot 1\cdot\frac{1}{10}+2\cdot 2\cdot\frac{2}{5}+2\cdot 3\cdot\frac{1}{10}\\
&\quad 3\cdot 0\cdot\frac{1}{30}+3\cdot 1\cdot\frac{1}{30}+3\cdot 2\cdot 0+3\cdot 3\cdot\frac{1}{10}\\
&= \frac{47}{15}
\end{aligned}$$

Using these pieces we get

$$\mathrm{Cov}(X,Y) = EXY - EXEY = \frac{47}{15}-\frac{11}{6}\cdot\frac{5}{3} = \frac{7}{90}.$$

For the correlation we first get the standard deviations of X and Y.

$$\begin{aligned}
\sigma_X &= \sqrt{EX^2-(EX)^2} = \sqrt{\frac{23}{6}-\left(\frac{11}{6}\right)^2} = \sqrt{\frac{17}{36}}\\
\sigma_Y &= \sqrt{EY^2-(EY)^2} = \sqrt{\frac{59}{15}-\left(\frac{5}{3}\right)^2} = \sqrt{\frac{52}{45}}
\end{aligned}$$

Then the correlation is

$$\mathrm{Corr}(X,Y) = \frac{7/90}{\sqrt{17/36}\sqrt{52/45}} = \frac{7}{2\sqrt{1105}} \approx 0.1053.$$

7. **Exercise 8.53**

   (a) We can use the bilinearity of covariance to get

$$\begin{aligned}
\mathrm{Cov}(3X+2,2Y-3) &= 3\cdot 2\cdot\mathrm{Cov}(X,Y) - 3\mathrm{Cov}(X,3) + 2\cdot 2\cdot\mathrm{Cov}(2,Y) - \mathrm{Cov}(2,3)\\
&= 6\mathrm{Cov}(X,Y)\\
&= 6(E[XY]-E[X]E[Y])\\
&= 6(-1-2) = -18.
\end{aligned}$$

(b) We start with the covariance.

$$\text{Cov}(X,Y) = EXY - EXEY = -1 - 2 = -3$$

Now we need the standard deviations of X and Y.

$$\sigma_X = \sqrt{EX^2 - (EX)^2} = \sqrt{3 - 1^2} = \sqrt{2}$$
$$\sigma_Y = \sqrt{EY^2 - (EY)^2} = \sqrt{13 - 2^2} = 3$$

So the correlation is

$$\text{Corr}(X,Y) = \frac{-3}{\sqrt{2}\cdot 3} = -\frac{1}{\sqrt{2}}.$$

8. **Exercise 8.57**

We start with the covariance.

$$\text{Cov}(X,Y) = EXY - EXEY = -1 - 2 = -3$$

Now we need the standard deviations of X and Y.

$$\sigma_X = \sqrt{EX^2 - (EX)^2} = \sqrt{3 - 1^2} = \sqrt{2}$$
$$\sigma_Y = \sqrt{EY^2 - (EY)^2} = \sqrt{5 - 2^2} = 1$$

So the correlation is

$$\text{Corr}(X,Y) = \frac{-3}{\sqrt{2}\cdot 1} = -\frac{3}{\sqrt{2}} < -1.$$

However, this is impossible. For any random variables $-1 \le \text{Corr}(X,Y) \le 1$. So there cannot be any random variables with the proposed expected values.

9. **Exercise 9.5**

(a) We use the Markov inequality.

$$P(X > 15) \le P(X \ge 15) \le \frac{EX}{15} = \frac{10}{15} = \frac{2}{3}.$$

(b) Now we use the Chebyshev inequality.

$$\begin{aligned} P(X > 15) \le P(X \ge 15) &= P(X - \mu \ge 5) \\ &\le P(|X - \mu| \ge 5) \le \frac{\text{Var}(X)}{5^2} = \frac{3}{25}. \end{aligned}$$

(c) Start by defining

$$S_{300} = Y_1 + Y_2 + \cdots + Y_{300}.$$

This is an IID sum, so the CLT suggests the distribution of S_{300} is approximately normal. To apply the CLT we need

$$\begin{aligned} E[S_{300}] &= 300E[Y_1] = 300 \cdot 10 = 3000 \\ \text{Var}(S_{300}) &= 300\text{Var}(Y_1) = 900 \\ \text{SD}(S_{300}) &= \sqrt{900} = 30. \end{aligned}$$

Now we use these for our CLT approximation.

$$\begin{aligned} P(S_{300} > 3030) &= P\left(\frac{S_{300} - 3000}{30} > \frac{3030 - 3000}{30}\right) \\ &= 1 - P\left(\frac{S_{300} - 3000}{30} \le 1\right) \\ &\approx 1 - \Phi(1) \approx 1 - 0.8413 = 0.1587 \end{aligned}$$

10. **Exercise 9.6**

    We start by defining appropriate random variables.

    $$X_k = \text{Time to eat } k\text{-th hot dog}$$

    $$\begin{aligned} S_{64} &= X_1 + X_2 + \cdots + X_{64} \\ &= \text{Time to eat 64 hot dogs} \end{aligned}$$

    We want to compute $P(S_{64} \le 15 \cdot 60)$. If we are working in units of seconds,

    $$E[X_k] = 15, \quad \mathrm{Var}(X_k) = 16.$$

    So $ES_{64} = 960$, $\mathrm{Var}(S_{64}) = 64 \cdot 16 = 1024$, and $SD(S_{64}) = 32$. Using the central limit theorem we have

    $$\begin{aligned} P(S_{64} \le 900) &= P\left(\frac{S_{64} - 960}{32} \le \frac{900 - 960}{32}\right) \\ &\approx \Phi(-1.875) \\ &= 0.0304. \end{aligned}$$

11. **Exercise 9.7**

    We start by defining appropriate random variables.

    $$X_k = \text{Size of } k\text{-th claim}$$

    Set c = premium cost.

    $$\begin{aligned} S_{2500} &= \sum_{k=1}^{2500} X_k \\ &= \text{Total cost of claims.} \end{aligned}$$

    We want to choose c such that

    $$P(S_{2500} \le 2500c) \ge 0.999.$$

    We will use the CLT to achieve this. First we note that

    $$\begin{aligned} E[S_{2500}] &= 2500 \cdot 1000 = 2500000 \\ \mathrm{Var}(S_{2500}) &= 2500\mathrm{Var}(X_1) = 2500 \cdot 900^2 \\ SD(X_1) &= \sqrt{2500 \cdot 900^2} = 50 \cdot 900 \end{aligned}$$

So we can approximate

$$\begin{aligned}
P\left(\frac{S_{2500}-2500000}{50\cdot 900} \le \frac{2500c-2500000}{50\cdot 900}\right) &\approx \Phi\left(\frac{2500(c-1000)}{50\cdot 900}\right) \\
&= \Phi\left(\frac{c-1000}{18}\right)
\end{aligned}$$

So we want

$$\begin{aligned}
&\Phi\left(\frac{c-1000}{18}\right) \ge 0.999 \\
&c \ge 18\cdot\Phi^{-1}(0.999)+1000 \\
&c \ge 1055.8.
\end{aligned}$$

12. **Exercise 9.17**

   (a)

$$P(X>120) \le \frac{EX}{120} = \frac{100}{120} = \frac{5}{6}.$$

   (b)

$$P(X>120) \le P(|X-100|\ge 20) \le \frac{\mathrm{Var}(X)}{20^2} = \frac{100}{400} = \frac{1}{4}.$$

   (c) Start by defining

$$S_{100} = X_1 + X_2 + \cdots + X_{100}.$$

   where $X_1, X_2, \ldots X_{100} \sim$ Poisson(1) are IID random variables. S_{100} is an IID sum, so the CLT suggests the distribution of S_{100} is approximately normal. To apply the CLT we need

$$\begin{aligned}
E[S_{100}] &= 100E[X_1] = 100 \\
\mathrm{Var}(S_{100}) &= 100\mathrm{Var}(X_1) = 100 \\
\mathrm{SD}(S_{100}) &= \sqrt{100} = 10.
\end{aligned}$$

   Now we use these for our CLT approximation.

$$\begin{aligned}
P(X>120) &= P(S_{100}>120) \\
&= P\left(\frac{S_{100}-100}{10} > \frac{120-100}{10}\right) \\
&= 1 - P\left(\frac{S_{100}-100}{10} \le 2\right) \\
&\approx 1-\Phi(2) \approx 1-0.9772 = 0.0228.
\end{aligned}$$

13. **Exercise 9.21**

   We start by setting

$$Z_k = X_k - Y_k.$$

We can rewrite our probability as

$$
\begin{aligned}
P\left(\sum_{k=1}^{500} X_k > \sum_{k=1}^{500} Y_k + 50\right) &= P\left(\sum_{k=1}^{500}(X_k - Y_k) > 50\right) \\
&= P\left(\sum_{k=1}^{500} Z_k > 50\right) \qquad = 1 - P\left(\sum_{k=1}^{500} Z_k \le 50\right).
\end{aligned}
$$

We use the CLT to approximate the probability of the sum. This first requires the mean and standard deviation.

$$
\begin{aligned}
E\left[\sum_{k=1}^{500} Z_k\right] &= \sum_{k=1}^{500} (E[X_k] - E[Y_k]) = 500(2-2) = 0 \\
\operatorname{Var}\left(\sum_{k=1}^{500} Z_k\right) &= \sum_{k=1}^{500} = 500(\operatorname{Var}(X_k) + (-1)^1 \operatorname{Var}(Y_k)) \\
&= 500(3+2) = 2500 \\
\operatorname{SD}\left(\sum_{k=1}^{500} Z_k\right) &= \sqrt{2500} = 50
\end{aligned}
$$

So we have

$$
\begin{aligned}
P\left(\sum_{k=1}^{500} X_k > \sum_{k=1}^{500} Y_k + 50\right) &= 1 - P\left(\sum_{k=1}^{500} Z_k \le 50\right) \\
&= 1 - P\left(\frac{\left(\sum_{k=1}^{500} Z_k\right) - 0}{50} \le \frac{50-0}{50}\right) \\
&\approx 1 - \Phi(1) \\
&\approx 0.1587.
\end{aligned}
$$