Question 1

Given the joint PMF, answer the following questions.

$X \setminus Y$	0	1	2
1	0.08	0.2	0.12
2	0.06	0.15	0.09
3	0.04	0.12	0.04
4	0.02	0.03	0.05

1. Compute the marginal PMF's of X and Y.
2. Compute $E(X), E(Y), E(X+Y), E(XY), \mathrm{Var}(X), \mathrm{Var}(Y)$.
3. Find the conditional distributions of $X \mid \{Y=0\}$, $Y \mid \{X=1\}$, $Y \mid \{X=2\}$, $Y \mid \{X=3\}$.
4. Calculate $P(X=1, Y=2 \mid X+Y<5)$.

Solution

1. Using the law of total probability, the marginal probabilities $\Pr[X=x]$ for $x \in \{1,2,3,4\}$ can be obtained by summing the appropriate rows the joint PMF. The results are as follows:

x	1	2	3	4
$\Pr[X=x]$	0.4	0.3	0.2	0.1

   Similarly, the marginal probabilities $\Pr[Y=y]$ for $y \in \{0,1,2\}$ can be obtained by summing the appropriate columns of the joint PMF:

y	0	1	2
$\Pr[Y=y]$	0.2	0.5	0.3

2. Using the definition of expectation, we have that

$$E(X) = \sum_x x \Pr[X=x] \quad \text{and} \quad E(Y) = \sum_y y \Pr[Y=y].$$

   Using these formulas, we find that

$$E(X) = 1 \cdot 0.4 + 2 \cdot 0.3 + 3 \cdot 0.2 + 4 \cdot 0.1 = 2$$

   and

$$E(Y) = 0 \cdot 0.2 + 1 \cdot 0.5 + 2 \cdot 0.3 = 1.1.$$

   The expectation of $X+Y$ can be computed using the linearity of expectation:

$$E(X+Y) = E(X) + E(Y) = 2.0 + 1.1 = 3.1.$$

   For $E[XY]$ we can use the definition of expectation:

$$E[XY] = \sum_x \sum_y xy \Pr[X=x, Y=y].$$

   Summing each entry in the table multiplied by the product of the corresponding x and y values, we find that

$$E[XY] = 1 \cdot (1 \cdot 0.2 + 2 \cdot 0.15 + 3 \cdot 0.12 + 4 \cdot 0.03) + 2 \cdot (1 \cdot 0.12 + 2 \cdot 0.09 + 3 \cdot 0.04 + 4 \cdot 0.05) = 2.22.$$

To compute the variances, we compute $E[X^2]$ and $E[Y^2]$:

$$E(X^2) = \sum_x x^2 \Pr[X = x] \quad \text{and} \quad E(Y^2) = \sum_y y^2 \Pr[Y = y].$$

Using these formulas, we find that

$$E(X^2) = 1^2 \cdot 0.4 + 2^2 \cdot 0.3 + 3^2 \cdot 0.2 + 4^2 \cdot 0.1 = 5$$

and

$$E(Y^2) = 0^2 \cdot 0.2 + 1^2 \cdot 0.5 + 2^2 \cdot 0.3 = 1.7.$$

The variance of X is then given by

$$\mathrm{Var}[X] = E[X^2] - (E(X))^2 = 5 - 2^2 = 1$$

and the variance of Y is given by

$$\mathrm{Var}[Y] = E[Y^2] - (E[Y])^2 = 1.7 - 1.1^2 = 0.49.$$

3. Recall that $\Pr[A \mid B] = \frac{\Pr[A,B]}{\Pr[B]}$. Therefore, to find the required conditional distributions, we take the relevant row or column of the joint PMF and divide by the marginal probability (which is the sum of that row or column). The conditional distributions are as follows:

x	1	2	3	4
$\Pr[X = x \mid Y = 0]$	$\frac{0.08}{0.2} = 0.4$	$\frac{0.06}{0.2} = 0.3$	$\frac{0.04}{0.2} = 0.2$	$\frac{0.02}{0.2} = 0.1$

y	0	1	2
$\Pr[Y = y \mid X = 1]$	$\frac{0.08}{0.4} = 0.2$	$\frac{0.2}{0.4} = 0.5$	$\frac{0.12}{0.4} = 0.3$

y	0	1	2
$\Pr[Y = y \mid X = 2]$	$\frac{0.06}{0.3} = 0.2$	$\frac{0.15}{0.3} = 0.5$	$\frac{0.09}{0.3} = 0.3$

y	0	1	2
$\Pr[Y = y \mid X = 3]$	$\frac{0.04}{0.2} = 0.2$	$\frac{0.12}{0.2} = 0.6$	$\frac{0.04}{0.2} = 0.2$

4. To compute the probability $P(X = 1, Y = 2 \mid X + Y < 5)$, we first compute the joint probability $P(X = 1, Y = 2, X + Y < 5)$. This is given by the sum of the joint probabilities of the relevant outcomes:

$$P(X = 1, Y = 2, X + Y < 5) = P(X = 1, Y = 2) = 0.12.$$

We then compute the probability $P(X + Y < 5)$ by summing the joint probabilities of the relevant outcomes:

$$P(X + Y < 5) = 0.08 + 0.2 + 0.12 + 0.06 + 0.15 + 0.09 + 0.04 + 0.12 + 0.02 = 0.88.$$

Finally, we compute the conditional probability:

$$P(X = 1, Y = 2 \mid X + Y < 5) = \frac{P(X = 1, Y = 2, X + Y < 5)}{P(X + Y < 5)} = \frac{0.12}{0.88} \approx 0.136.$$

Question 2

Let $X \sim U(-1, 1)$ and $Y = X^2$.

Here $U(-1, 1)$ denotes the discrete uniform distribution, supported on the set $\{-1, 0, 1\}$.

1. Compute the PMF of Y.
2. Compute the mean and variance of Y.
3. Compute the covariance of X and Y.
4. Are X and Y independent? Explain.

Solution

1. $\Pr[Y = 0] = \Pr[X = 0] = \frac{1}{3}$, and $\Pr[Y = 1] = \Pr[X \in \{-1, 1\}] = \frac{2}{3}$.
2. $E[Y] = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$. To compute the variance of Y, we first compute $E[Y^2] = 0^2 \cdot \frac{1}{3} + 1^2 \cdot \frac{2}{3} = \frac{2}{3}$. Then, $\mathrm{Var}[Y] = E[Y^2] - (E[Y])^2 = \frac{2}{3} - \left(\frac{2}{3}\right)^2 = \frac{2}{9}$.
3. The covariance of X and Y is given by $\mathrm{Cov}[X, Y] = E[XY] - E[X]E[Y]$. Note that $XY = X^3 = X$, and therefore $E[XY] = E[X] = 0$ since X is symmetric around 0. Thus, $\mathrm{Cov}[X, Y] = 0 - 0 \cdot \frac{2}{3} = 0$.
4. Y is a function of X and therefore X and Y are clearly not independent, though they are uncorrelated. To prove that they are not independent, it suffices to show that $\Pr[Y = y \mid X = x] \neq \Pr[Y = y]$ for some x and y. For example, $\Pr[Y = 0 \mid X = 0] = 1 \neq \frac{2}{3} = \Pr[Y = 0]$.

Question 3

Let X_1, X_2, X_3 be i.i.d. (independent and identically distributed) random variables, each with mean μ and variance σ^2. Denote their average by $\bar{X} = \frac{X_1+X_2+X_3}{3}$. Calculate the mean and variance of $\bar{X}$ in terms of μ and σ^2.

Solution

Recall the linearity of expectation: $E[aX + bY] = aE[X] + bE[Y]$. Therefore, $E[\bar{X}] = E\left[\frac{X_1+X_2+X_3}{3}\right] = \frac{E[X_1]+E[X_2]+E[X_3]}{3} = \frac{3\mu}{3} = \mu$.

To compute the variance of $\bar{X}$, we use the fact that the variance of a sum of **independent** random variables is the sum of their variances, and that $\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$. Therefore, $\mathrm{Var}[\bar{X}] = \mathrm{Var}\left[\frac{X_1+X_2+X_3}{3}\right] = \frac{\mathrm{Var}[X_1]+\mathrm{Var}[X_2]+\mathrm{Var}[X_3]}{3^2} = \frac{3\sigma^2}{9} = \frac{\sigma^2}{3}$.

Question 4

Say that you are analyzing the performance of cloud computing infrastructure, and are interested in the time until the failure of a machine. The time until failure is modeled as an exponential random variable, and the mean time until failure is 5 years.

1. What is the probability that a machine fails in less than 5 years?
2. What is the median time until failure, and what are the 0.025, 0.05, 0.95 and 0.975 quantiles?
3. Say that you had modeled the time until failure as a normal random variable with the same mean and variance. What would then be your answers to the previous questions? HINT: You can use standard normal distribution tables. Does it make sense to do that?
4. Using the exponential distribution model, given that a machine did not fail within the first five years, what is the probability that it will fail within the next five years?

Solution

1. Denote the time until failure by $X \sim \mathrm{Exp}(\lambda)$. Recall that the CDF of an exponential distribution is given by $F(x) = 1 - e^{-\lambda x}$. Since the mean time until failure is 5 years, we have that $\lambda = \frac{1}{5}$. Therefore, the probability that a machine fails in less than 5 years is given by

$$P(X < 5) = 1 - e^{-\frac{1}{5} \cdot 5} = 1 - e^{-1} \approx 0.632.$$

2. To compute the q quantile of X, we solve for x in the equation $q = 1 - e^{-\frac{1}{5}x}$. In other words, $x = -5\ln(1 - q)$. Therefore, the median time until failure is $-5\ln(1 - 0.5) = 5$ years. The 0.025

quantile is $-5\ln(1-0.025) \approx 0.13$ years, the 0.05 quantile is $-5\ln(1-0.05) \approx 0.26$ years, the 0.95 quantile is $-5\ln(1-0.95) \approx 14.98$ years, and the 0.975 quantile is $-5\ln(1-0.975) \approx 18.44$ years.

3. Note that $E[X] = \frac{1}{\lambda} = 5$ and $\text{Var}[X] = \frac{1}{\lambda^2} = 25$. Therefore, in this case we would model the time until failure as $X \sim N(5, 25)$. The median time until failure is equal to the mean, which is 5 years. The quantiles can be computed using the standard normal distribution tables. The 0.025 quantile is $5+5\cdot\Phi^{-1}(0.025) \approx 5-1.96\cdot 5 = -4.8$, the 0.05 quantile is $5+5\cdot\Phi^{-1}(0.05) \approx 5-1.645\cdot 5 = -3.225$, the 0.95 quantile is $5+5\cdot\Phi^{-1}(0.95) \approx 5+1.645\cdot 5 = 11.225$, and the 0.975 quantile is $5+5\cdot\Phi^{-1}(0.975) \approx 5+1.96\cdot 5 = 14.8$. As we can see, the normal distribution model predicts negative times until failure, which is nonsensical.

4. Using the memoryless property of the exponential distribution, we have that

$$P(X > 5 \mid X > 0) = P(X > 5) = 1 - P(X < 5) \approx 0.368.$$