MTH102 Solution to Tutorial 06 Jointly distributed random variables & limit theorems

Question 1

Let the joint pmf of X and Y be

$$f(x,y)=\frac{1}{4},\ (x,y)\in S=\{(0,0),\ (1,1),\ (1,-1),\ (2,0)\}.$$

(a) Are X and Y independent?

(b) Calculate the covariance and correlation coefficient of X and Y.

Answer:

(a) The marginal pmfs of X and Y are the following:

$$f_X(x)=\begin{cases}1/4, & x=0,\\ 1/2, & x=1,\\ 1/4, & x=2,\end{cases}\quad f_Y(y)=\begin{cases}1/4, & y=-1,\\ 1/2, & y=0,\\ 1/4, & y=1.\end{cases}$$

Note that

$$f(0,0)=\frac{1}{4}\neq f_X(0)f_Y(0)=\frac{1}{8},$$

therefore X and Y are not independent.

(b) We compute the following

$$E(X)=1,\ E(Y)=0,\ \text{and } E(XY)=\frac{1}{4}\times(0+1-1+0)=0.$$

Hence

$$\operatorname{Cov}(X,Y)=E(XY)-E(X)E(Y)=0,\ \rho=0.$$

Question 2

The joint probability density function X and Y is given by

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & x > 0,\ y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the marginal pdf of X and Y.

(b) Find the mean and variance of X and Y.

(c) Find the correlation coefficient of X and Y.

(d) Find $P(X > 1, Y < X)$.

Answer:

(a) The marginal pdf of X is

$$f_X(x) = \begin{cases} \int_0^\infty 2e^{-x}e^{-2y}dy = e^{-x} & x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

The marginal pdf of Y is

$$f_Y(y) = \begin{cases} \int_0^\infty 2e^{-x}e^{-2y}dx = 2e^{-2y} & y > 0, \\ 0 & \text{otherwise.} \end{cases}$$

(b) X and Y both have an exponential distribution with parameters 1 and 2 respectively. Therefore,

$$E(X) = 1,\ Var(X) = 1,\ E(Y) = \frac{1}{2},\ Var(Y) = \frac{1}{4}.$$

(c) Note that $f(x,y) = f_X(x)f_Y(y)$, we have thus X and Y are independent. Therefore $\rho = 0$.

(d)

$$P(X > 1, Y < X) = \int_1^\infty \int_0^x 2e^{-x}e^{-2y}dydx = e^{-1} - \frac{1}{3}e^{-3}.$$

Question 3

Let X and Y have the joint pdf

$$f(x,y) = 2,\ 0 \le x \le y, 0 \le y \le 1.$$

Find the covariance of X and Y.

Answer:

$$E(X) = \int_0^1 \int_x^1 2x dy dx = \frac{1}{3},$$
$$E(Y) = \int_0^1 \int_x^1 2y dy dx = \frac{2}{3},$$
$$E(XY) = \int_0^1 \int_x^1 2xy dy dx = \frac{1}{4},$$
$$Cov(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{36}.$$

Question 4

Let X be a random variable with mean -2 and variance 1, and Y be a random variable with mean 2 and variance 4. It is known that the correlation coefficient ρ of X and Y is -0.5. Use Chebyshev's inequality to find an upper bound of $P(|X+Y| \geq 6)$.

Answer:

Let $Z = X + Y$. Then

$$E(Z) = E(X+Y) = E(X) + E(Y) = 0,$$

and

$$\begin{aligned} Var(Z) &= Var(X+Y) \\ &= Var(X) + Var(Y) + 2Cov(X, Y) \\ &= Var(X) + Var(Y) + 2\rho\sqrt{Var(X)Var(Y)} \\ &= 3. \end{aligned}$$

Then by Chebyshev's inequality

$$P(|X+Y| \geq 6) = P(|Z - E(Z)| \geq 6) \leq \frac{Var(Z)}{6^2} = \frac{1}{12}.$$

Question 5

Let $X_1, \ldots, X_{25}$ be independent Poisson random variables with mean 1.

(a) Use the Markov's inequality to obtain a bound on

$$P\left(\sum_{i=1}^{25} X_i > 30\right).$$

(b) Use the central limit theorem to approximate

$$P\left(\sum_{i=1}^{25} X_i > 30\right).$$

Answer:

1.
$$P\left(\sum_{i=1}^{25} X_i > 30\right) \leq \frac{E\left(\sum_{i=1}^{25} X_i\right)}{30} = \frac{5}{6}.$$

2. We have $n = 25$, $\mu = E(X_i) = 1$ and $\sigma^2 = Var(X_i) = 1$.

$$P\left(\sum_{i=1}^{25} X_i > 30\right) = P\left(\frac{\sum_{i=1}^{25} X_i - n\mu}{\sigma\sqrt{n}} > \frac{30-25}{5}\right) \approx 1 - \Phi(1) = 0.1587.$$

Question 6

A worker goes to work by bus and the waiting time for a bus on every working day follows an exponential distribution with mean 5 (in minutes). Find the approximate probability that the worker has spent more than 24 hours on waiting the bus in total during a period of 225 working days.

Answer:

Let $n = 225$, $\mu = 5$ and $\sigma^2 = 5^2$. For $i = 1, \ldots, n$, let X_i be the waiting time on the $i-$th day. Then $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$. By the central limit theorem,

$$\begin{aligned} P\left(\sum_{i=1}^{n} X_i \geq 24 \times 60\right) &= P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \geq \frac{1440 - 1125}{5\sqrt{225}}\right) \\ &\approx 1 - \Phi(4.2). \end{aligned}$$

Question 7

Suppose each of 300 real numbers are rounded to the nearest integer and then added. Assume the individual roundoff errors are independent and uniformly distributed over the interval $[-0.5, 0.5]$. Using the central limit theorem to find the approximate probability that the absolute value of the sum of the errors is greater than 5.

Answer:

For $i = 1, 2, \ldots, 300$, let X_i be the rounded error of the i-th number. Then

$$n = 300,\ \mu = E(X_i) = 0,\ \sigma^2 = Var(X_i) = \frac{1}{12}.$$

Hence,

$$\begin{aligned}
P\left(\left|\sum_{i=1}^{n} X_i\right| > 5\right) =& P\left(\left\{\sum_{i=1}^{n} X_i > 5\right\} \cup \left\{\sum_{i=1}^{n} X_i < -5\right\}\right) \\
=& P\left(\sum_{i=1}^{n} X_i > 5\right) + P\left(\sum_{i=1}^{n} X_i < -5\right) \\
=& P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} > \frac{5-0}{\sqrt{300/12}}\right) + P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} < \frac{-5-0}{\sqrt{300/12}}\right) \\
\approx& 1 - \Phi(1) + \Phi(-1) = 2(1-\Phi(1)) = 2 \times 0.1587 = 0.3174.
\end{aligned}$$

Question 8

Let $X_1, X_2, \ldots, X_{48}$ be a random sample of size 48 from the distribution with pdf $f(x) = 1/x^2,\ 1 < x < \infty$. Approximate the probability that at most 10 of these random variables have values greater than 4.

Answer:

Let the ith trial be a success if $X_i > 4$, $i = 1, 2, \ldots, 48$, and let Y be the number of successes.

$$P(X_i > 4) = \int_4^{\infty} \frac{1}{x^2} dx = \frac{1}{4}.$$

And thus Y has a binomial distribution $b(48, p)$ with $p = 1/4$. Hence,

$$P(Y \leq 10) \approx \Phi\left(\frac{10 + 0.5 - 48p}{\sqrt{48p(1-p)}}\right) = \Phi(-0.5) = 0.3085.$$

Question 9

Let X equal the forced vital capacity (the volume of air a person can expel from his or her lungs) of an athlete. 17 observations of X, which have been ordered, are

3.4 3.6 4.1 4.3 4.5 4.9 5.2 5.4 5.5 5.7 5.8 6.0 6.1 6.1 6.9 6.9 7.5.

Find the mean, the median, the first quartile and the third quartile.

Answer:

- The mean is
$$\frac{1}{17}\sum_{i=1}^{17} x_i \approx 5.4059.$$
- Then median is $x_9 = 5.5$.
- The first quartile is the median of $x_1, \ldots, x_8$, i.e. $\frac{1}{2}(x_4 + x_5) = 4.4$.
- The third quartile is the median of $x_{10}, \ldots, x_{17}$, i.e. $\frac{1}{2}(x_{13} + x_{14}) = 6.1$.