

EE 3070 Statistics

Homework #1 Solution

Due at 23:59, March 31, 2023
online submission to eclass systems

1. At the beginning of a study of individuals, 15% were classified as heavy smokers, 30% were classified as light smokers, and 55% were classified as nonsmokers. In the five-year study, it was determined that the death rates of the heavy and light smokers were five and three times that of the nonsmokers, respectively. A randomly selected participant died over the five-year period: calculate the probability that the participant was a nonsmoker.

Solution:

Let the total number of participant be S and the death rate of nonsmokers be x .

$$\begin{aligned} P(\text{nonsmokers} \mid \text{total death}) &= \frac{P(\text{the nonsmokers who actually died})}{P(\text{total death})} \\ &= \frac{(0.55 \times S) \times x}{(0.55 \times S) \times x + (0.30 \times S) \times 3x + (0.15 \times S) \times 5x} \\ &= \frac{0.55}{0.55 + 0.9 + 0.75} = \frac{0.55}{2.2} = 0.25 \\ &= \frac{1}{4} \end{aligned}$$

■

2. Let the space of the random variable X be $\mathcal{D} = \{x : 0 < x < 1\}$. If $D_1 = \{x : 0 < x < \frac{1}{2}\}$ and $D_2 = \{x : \frac{1}{2} \leq x < 1\}$, find $P_X(D_2)$ if $P_X(D_1) = \frac{1}{4}$.

Solution:

We have $P_X(D_1) = \frac{1}{4}$ and we want to find $P_X(D_2)$. Note that \mathcal{D} is the space of the random variable X and we have partitioned it into two sets D_1 and D_2 .

By the law of total probability, we have:

$$P_X(D_2) = P_X(\mathcal{D}) - P_X(D_1) = 1 - \frac{1}{4} = \frac{3}{4}$$

Therefore

$$P_X(D_2) = \frac{3}{4}$$

which means that the probability of X taking on a value in the interval $[\frac{1}{2}, 1]$ is $\frac{3}{4}$.

■

3. Let X have the pmf $p(x) = \frac{1}{3}$, $x = -1, 0, 1$. Find the pmf of $Y = X^2$.

Solution:

Since $Y = X^2$, $X = \{-1, 0, 1\}$, we have $Y = \{0, 1\}$. Therefore

$$P(Y = 1) = P(X = 1) + P(X = -1) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

$$P(Y = 0) = P(X = 0) = \frac{1}{3}$$

■

4. Let X have the pdf $f(x) = \frac{1}{9}x^2$, $0 < x < 3$, zero elsewhere. Find the pdf of $Y = X^3$.

Solution:

For $0 < y < 27$,

$$x = y^{\frac{1}{3}}, \quad \frac{dx}{dy} = \frac{1}{3}y^{-\frac{2}{3}}, \quad f_Y(y) = \frac{1}{3}y^{-\frac{2}{3}} \cdot \frac{1}{9}y^{\frac{2}{3}} = \frac{1}{27} \quad \text{or simply} \quad f_Y(y) = \frac{f_X(x)}{\left| \frac{dy}{dx} \right|} = \frac{\frac{1}{9}x^2}{3x^2} = \frac{1}{27}$$

■

5. Let X have the pdf $f(x) = 3x^2$, $0 < x < 1$, zero elsewhere.

- Compute $E(X^3)$.
- Show that $Y = X^3$ has a uniform $(0, 1)$ distribution.
- Compute $E(Y)$ and compare this result with the answer obtained in part.

Solution:

(a) We can use the definition of the expected value to find $E(X^3)$:

$$E(X^3) = \int_{-\infty}^{\infty} x^3 f(x) dx = \int_0^1 x^3 (3x^2) dx = 3 \int_0^1 x^5 dx$$

$$= 3 \cdot \frac{1}{6} \cdot x^6 \Big|_0^1 = \frac{1}{2}$$

(b) To show that $Y = X^3$ has a uniform $(0, 1)$ distribution, we need to find the CDF of Y and show that it is equal to the CDF of a uniform distribution on $(0, 1)$. The CDF of Y is:

$$F_Y(y) = P(Y \leq y) = P(X^3 \leq y) = P(X \leq y^{1/3}) = \int_0^{y^{1/3}} 3x^2 dx = y$$

where we can clearly see that $F_Y(y) = y$ for $0 < y < 1$ is a uniform distribution on $(0, 1)$. Therefore, we have shown that $Y = X^3$ has a uniform $(0, 1)$ distribution.

(c) To compute $E(Y)$, we can use the definition of the expected value:

$$E(Y) = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_0^1 y \cdot 1 dy = \frac{1}{2} \cdot y^2 \Big|_0^1 = \frac{1}{2}$$

■

6. Suppose X_1 and X_2 have the joint pdf $f_{X_1, X_2}(x_1, x_2) = e^{-(x_1+x_2)}$, $0 < x_i < \infty, i = 1, 2$, zero elsewhere.

(a) Use formula (2.2.2) to find the pdf of $Y_1 = X_1 + X_2$.

(b) Find the mgf of Y_1 .

Note. formula (2.2.2) $f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{X_1, X_2}(y_1 - y_2, y_2) dy_2$

Solution:

(a) We have $f_{X_1, X_2}(x_1, x_2) = e^{-(x_1+x_2)}$, for $0 < x_i < \infty, i = 1, 2$, and zero elsewhere. To find the pdf of $Y_1 = X_1 + X_2$, we use the transformation method as follows:

$$\begin{aligned}
 F_{Y_1}(y_1) &= P(Y_1 \leq y_1) = P(X_1 + X_2 \leq y_1) \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{y_1-x_2} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 \quad (\text{by definition of joint pdf}) \\
 &= \int_0^{y_1} \int_0^{y_1-x_2} e^{-(x_1+x_2)} dx_1 dx_2 \quad (\text{since the pdf is zero elsewhere}) \\
 &= \int_0^{y_1} e^{-x_2} \int_0^{y_1-x_2} e^{-x_1} dx_1 dx_2 \\
 &= \int_0^{y_1} e^{-x_2} (1 - e^{-(y_1-x_2)}) dx_2 \\
 &= \int_0^{y_1} e^{-y_1} dx_2 \\
 &= y_1 e^{-y_1}, \quad \text{for } y_1 > 0,
 \end{aligned}$$

Therefore, the pdf of Y_1 is given by

$$f_{Y_1}(y_1) = \frac{d}{dy_1} F_{Y_1}(y_1) = y_1 e^{-y_1}, \quad \text{for } y_1 > 0$$

(b) To find the mgf of Y_1 , we use the definition:

$$\begin{aligned}
 M_{Y_1}(t_1) &= E(e^{t_1 Y_1}) = \int_{-\infty}^{\infty} e^{t_1 y_1} f_{Y_1}(y_1) dy_1 = \int_0^{\infty} e^{t_1 y_1} y_1 e^{-y_1} dy_1 \\
 &= \int_0^{\infty} y_1 \cdot e^{-(1-t_1)y_1} dy_1 \quad (\text{by using integration by parts } \int u dv = u \cdot v - \int v du) \\
 &= \int_0^{\infty} y_1 \cdot d \left(\frac{1}{-(1-t_1)} e^{-(1-t_1)y_1} \right) \quad (\because \int x e^{ax} dx = x \cdot \frac{1}{a} e^{ax} - \int \frac{1}{a} \cdot e^{ax} dx) \\
 &= \left(y_1 \cdot \frac{1}{-(1-t_1)} e^{-(1-t_1)y_1} \right) \Big|_0^{\infty} - \int_0^{\infty} \frac{1}{-(1-t_1)} e^{-(1-t_1)y_1} dy_1 \\
 &= (\infty \cdot e^{-\infty} - 0 \cdot e^0) - \frac{1}{-(1-t_1)} \cdot \left(\frac{1}{-(1-t_1)} \cdot e^{-(1-t_1)y_1} \Big|_0^{\infty} \right) \quad (\because \frac{1}{a} \int e^{ax} dx = \frac{1}{a} \cdot \frac{1}{a} e^{ax}) \\
 &= (0 - 0) - \frac{1}{(1-t_1)^2} \cdot (e^{-\infty} - e^0) = 0 - \frac{1}{(1-t_1)^2} \cdot (0 - 1) \\
 &= \frac{1}{(1-t_1)^2}, \quad \text{for } t_1 < 1.
 \end{aligned}$$

Therefore, the moment generating function of Y_1 is given by $M_{Y_1}(t_1) = \frac{1}{(1-t_1)^2}$, for $t_1 < 1$.

■

7. Let the joint pdf of X and Y be given by

$$f(x, y) = \begin{cases} \frac{2}{(1+x+y)^3} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{elsewhere} \end{cases}$$

- (a) Compute the marginal pdf of X and the conditional pdf of Y , given $X = x$.
(b) For a fixed $X = x$, compute $E(1+x+Y|x)$ and use the result to compute $E(Y|x)$.

Solution:

(a) To find the marginal pdf of X , we integrate the joint pdf over all possible values of Y :

$$\begin{aligned} f_X(x) &= \int_0^\infty f(x, y) dy \\ &= \int_0^\infty \frac{2}{(1+x+y)^3} dy \\ &= \frac{1}{(1+x)^2} \end{aligned}$$

To find the conditional pdf of Y given $X = x$, we use the conditional probability formula:

$$\begin{aligned} f_{Y|X}(y|x) &= \frac{f(x, y)}{f_X(x)} \\ &= \frac{2}{(1+x+y)^3} \cdot \frac{(1+x)^2}{1} \\ &= \frac{2(1+x)^2}{(1+x+y)^3} \end{aligned}$$

(b) For a fixed $X = x$, the conditional expectation of $1+x+Y$ is:

$$\begin{aligned} E(1+x+Y|X=x) &= E(1+x|X=x) + E(Y|X=x) \\ &= 1+x + E(Y|X=x) \end{aligned}$$

To find $E(Y|X=x)$, we use the definition of conditional expectation:

$$\begin{aligned} E(Y|X=x) &= \int_0^\infty y f_{Y|X}(y|x) dy \\ &= \int_0^\infty y \cdot \frac{2(1+x)^2}{(1+x+y)^3} dy \\ &= 1+x \end{aligned}$$

Therefore, we have:

$$\begin{aligned} E(1+x+Y|X=x) &= 1+x + E(Y|X=x) \\ &= 1+x + 1+x \\ &= 2+2x \end{aligned}$$

■

8. Let X and Y have the joint pmf $p(x, y) = \frac{1}{7}, (0, 0), (1, 0), (0, 1), (1, 1), (2, 1), (1, 2), (2, 2)$, zero elsewhere. Find the correlation coefficient ρ .

Solution:

We first need to compute the mean and variance of X and Y as follows:

$$E(X) = \sum_{x=0}^2 x \sum_{y=0}^2 p(x, y) = (0 + 1 + 0 + 1 + 2 + 1 + 2) \cdot \frac{1}{7} = \frac{7}{7} = 1$$

$$E(Y) = \sum_{y=0}^2 y \sum_{x=0}^2 p(x, y) = (0 + 0 + 1 + 1 + 1 + 2 + 2) \cdot \frac{1}{7} = \frac{7}{7} = 1$$

$$E(XY) = \sum_{x=0, y=0}^2 xy \sum_{x=0}^2 p(x, y) = (0 + 0 + 0 + 1 + 2 + 2 + 4) \cdot \frac{1}{7} = \frac{9}{7}$$

$$E(X^2) = \sum_{x=0}^2 x^2 \sum_{y=0}^2 p(x, y) = (0^2 + 1^2 + 0^2 + 1^2 + 2^2 + 1^2 + 2^2) \cdot \frac{1}{7} = \frac{11}{7}$$

$$E(Y^2) = \sum_{y=0}^2 y^2 \sum_{x=0}^2 p(x, y) = (0^2 + 0^2 + 1^2 + 1^2 + 1^2 + 2^2 + 2^2) \cdot \frac{1}{7} = \frac{11}{7}$$

Now, we can compute the covariance:

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{9}{7} - 1 \cdot 1 = \frac{2}{7}$$

Finally, we can compute the standard deviations:

$$\begin{aligned} \sigma_x &= \sqrt{E(X^2) - (E(X))^2} = \sqrt{\text{Var}(X)} \\ &= \sqrt{\left(\frac{11}{7}\right) - (1)^2} = \sqrt{\frac{11}{7} - \frac{7}{7}} \\ &= \sqrt{\frac{4}{7}} \\ \sigma_y &= \sqrt{E(Y^2) - (E(Y))^2} = \sqrt{\text{Var}(Y)} \\ &= \sqrt{\left(\frac{11}{7}\right) - (1)^2} = \sqrt{\frac{11}{7} - \frac{7}{7}} \\ &= \sqrt{\frac{4}{7}} \end{aligned}$$

Finally, we can plug all the values into the formula for the correlation coefficient:

$$\rho = \frac{\text{Cov}(X, Y)}{\sigma_x \cdot \sigma_y} = \frac{\frac{2}{7}}{\sqrt{\frac{4}{7}} \cdot \sqrt{\frac{4}{7}}} = \frac{1}{2}$$

Therefore, the correlation coefficient between X and Y is given by

$$\rho = \frac{1}{2}$$

■

9. Let $f(x_1, x_2, x_3) = e^{-(x_1+x_2+x_3)}, 0 < x_1 < \infty, 0 < x_2 < \infty, 0 < x_3 < \infty$, zero elsewhere, be the joint pdf of X_1, X_2, X_3 .

(a) Compute $P(X_1 < X_2 < X_3)$ and $P(X_1 = X_2 < X_3)$.

(b) Determine the joint mgf of X_1, X_2 and X_3 . Are these random variables independent?

Solution:

(a) To compute $P(X_1 < X_2 < X_3)$, we use the definition:

$$\begin{aligned}
 P(X_1 < X_2 < X_3) &= \iiint_{x_1 < x_2 < x_3} f(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\
 &= \int_0^\infty \int_0^{x_3} \int_0^{x_2} e^{-(x_1+x_2+x_3)} dx_1 dx_2 dx_3 \\
 &= \int_0^\infty \int_0^{x_3} \int_0^{x_2} e^{-x_1} \cdot e^{-(x_2+x_3)} dx_1 dx_2 dx_3 \\
 &= \int_0^\infty \int_0^{x_3} e^{-(x_2+x_3)} \cdot \left(\int_0^{x_2} e^{-x_1} dx_1 \right) dx_2 dx_3 \\
 &= \int_0^\infty \int_0^{x_3} e^{-(x_2+x_3)} \cdot \left(-e^{-x_1} \Big|_0^{x_2} \right) dx_2 dx_3 \\
 &= \int_0^\infty \int_0^{x_3} e^{-(x_2+x_3)} \cdot (1 - e^{-x_2}) dx_2 dx_3 \\
 &= \int_0^\infty e^{-x_3} \cdot \left(\int_0^{x_3} e^{-x_2} (1 - e^{-x_2}) dx_2 \right) dx_3 \\
 &= \int_0^\infty e^{-x_3} \cdot \left(\int_0^{x_3} (e^{-x_2} - e^{-2x_2}) dx_2 \right) dx_3 \\
 &= \int_0^\infty e^{-x_3} \cdot \left(-e^{-x_2} \Big|_0^{x_3} - \left(-\frac{1}{2} e^{-2x_2} \Big|_0^{x_3} \right) \right) dx_3 \\
 &= \int_0^\infty e^{-x_3} \cdot \left(1 - e^{-x_3} + \frac{1}{2} e^{-2x_3} - \frac{1}{2} \right) dx_3 \\
 &= \int_0^\infty \frac{1}{2} e^{-x_3} - e^{-2x_3} + \frac{1}{2} e^{-3x_3} dx_3 \\
 &= -\frac{1}{2} e^{-x_3} \Big|_0^\infty - \left(-\frac{1}{2} e^{-2x_3} \Big|_0^\infty \right) + \frac{1}{2} \cdot \left(-\frac{1}{3} e^{-3x_3} \Big|_0^\infty \right) \\
 &= -\frac{1}{2} \cdot (0 - 1) + \frac{1}{2} \cdot (0 - 1) - \frac{1}{6} \cdot (0 - 1) \\
 &= \frac{1}{2} - \frac{1}{2} + \frac{1}{6} \\
 &= \frac{1}{6}
 \end{aligned}$$

To compute $P(X_1 = X_2 < X_3)$, we also use the definition:

$$\begin{aligned}
 P(X_1 = X_2 < X_3) &= \iiint_{x_1 = x_2 < x_3} f(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\
 &= \int_0^\infty \int_0^{x_3} \int_{x_2}^{x_2} e^{-(x_1+x_2+x_3)} dx_1 dx_2 dx_3 \\
 &= \int_0^\infty \int_0^{x_3} e^{-(x_2+x_3)} \cdot \left(\int_{x_2}^{x_2} e^{-x_1} dx_1 \right) dx_2 dx_3 \\
 &= \int_0^\infty \int_0^{x_3} e^{-(x_2+x_3)} \cdot 0 dx_2 dx_3 \\
 &= 0
 \end{aligned}$$

Or, if you think this is hard to understand, here's another way:
Let

$$Z_1 = X_1 + X_2, \quad Z_2 = X_2 - X_3, \quad Z_3 = X_2$$

so that we have

$$X_1 = Z_1 + Z_3, \quad X_2 = Z_3, \quad X_3 = Z_3 - Z_2$$

Then

$$\begin{aligned} f_{Z_1, Z_2, Z_3}(z_1, z_2, z_3) &= f_{X_1, X_2, X_3}(z_1 + z_3, z_3, z_3 - z_2) \cdot \begin{vmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & -1 & 1 \end{vmatrix} \\ &= e^{-(z_1 + z_3 + z_3 + z_3 - z_2)} \cdot (-0 + 0 - (-1)) \\ &= e^{-(z_1 - z_2 + 3z_3)}, \quad -\infty < z_1, z_2 < \infty, \quad 0 < z_3 < \infty \end{aligned}$$

Therefore

$$P(X_1 = X_2 < X_3) = P(Z_1 = 0, Z_2 < 0) = 0$$

(b) To find the joint moment generating function (MGF), we compute:

$$\begin{aligned} M_X(t_1, t_2, t_3) &= E[e^{t_1 X_1 + t_2 X_2 + t_3 X_3}] \\ &= \int_0^\infty \int_0^\infty \int_0^\infty e^{-(x_1 + x_2 + x_3)} \cdot e^{t_1 x_1 + t_2 x_2 + t_3 x_3} dx_1 dx_2 dx_3 \\ &= \int_0^\infty \int_0^\infty \int_0^\infty e^{-(1-t_1)x_1} \cdot e^{-(1-t_2)x_2} \cdot e^{-(1-t_3)x_3} dx_1 dx_2 dx_3 \\ &= \left(\int_0^\infty e^{-(1-t_1)x_1} dx_1 \right) \cdot \left(\int_0^\infty e^{-(1-t_2)x_2} dx_2 \right) \cdot \left(\int_0^\infty e^{-(1-t_3)x_3} dx_3 \right) \\ &= \left(\frac{1}{-(1-t_1)} \cdot (0-1) \right) \cdot \left(\frac{1}{-(1-t_2)} \cdot (0-1) \right) \cdot \left(\frac{1}{-(1-t_3)} \cdot (0-1) \right) \\ &= \frac{1}{(1-t_1)(1-t_2)(1-t_3)} \end{aligned}$$

Note that the MGF of X_1, X_2 , and X_3 factors into the product of their individual MGFs:

$$M_{X_1}(t_1) = \frac{1}{1-t_1}, \quad M_{X_2}(t_2) = \frac{1}{1-t_2}, \quad M_{X_3}(t_3) = \frac{1}{1-t_3}$$

Therefore, X_1, X_2 , and X_3 are independent. ■

10. Let X_1, X_2 and X_3 be iid with common pdf $f(x) = e^{-x}$, $x > 0$, zero elsewhere. Find the joint pdf of $Y_1 = X_1$, $Y_2 = X_1 + X_2$ and $Y_3 = X_1 + X_2 + X_3$.

Solution:

We have X_1, X_2, X_3 are independent and identically distributed random variables with pdf $f(x) = e^{-x}$, $x > 0$, zero elsewhere.

(1) For $Y_1 = X_1$,

$$f_Y(y) = e^{-y}, \quad y > 0$$

(2) For $Y_2 = X_1 + X_2$,

$$\begin{aligned} M_{Y_2}(y_2) &= E[e^{ty_2}] = E[e^{t(x_1+x_2)}] = E[e^{tx_1}] \cdot E[e^{tx_2}] \\ &= \frac{1}{1-t} \cdot \frac{1}{1-t} = \frac{1}{(1-t)^2} \sim \text{Gamma}(2, 1) \end{aligned}$$

(3) For $Y_3 = X_1 + X_2 + X_3$,

$$\begin{aligned} M_{Y_3}(y_3) &= E[e^{ty_3}] = E[e^{t(x_1+x_2+x_3)}] = E[e^{tx_1}] \cdot E[e^{tx_2}] \cdot E[e^{tx_3}] \\ &= \frac{1}{1-t} \cdot \frac{1}{1-t} \cdot \frac{1}{1-t} = \frac{1}{(1-t)^3} \sim \text{Gamma}(3, 1) \end{aligned}$$

■

11. Find the mean and variance of the sum $Y = \sum_{i=1}^5 X_i$, where X_1, \dots, X_5 are iid, having pdf $f(x) = 6x(1-x), 0 < x < 1$, zero elsewhere.

Solution:

Since X_1, \dots, X_5 are iid, we have $E(X_i) = \int_{-\infty}^{\infty} x \cdot f(x) dx$ and $\text{Var}(X_i) = E(X_i^2) - (E(X_i))^2$. Thus, we have:

$$\begin{aligned} E(X) &= \int_0^1 x \cdot 6x(1-x) dx = 6 \cdot \int_0^1 x^2 - x^3 dx \\ &= 6 \cdot \left(\frac{1}{3} x^3 \Big|_0^1 - \frac{1}{4} x^4 \Big|_0^1 \right) = 6 \cdot \left(\frac{4}{12} - \frac{3}{12} \right) = \frac{6}{12} \\ &= \frac{1}{2} \\ Y &= X_1 + X_2 + X_3 + X_4 + X_5 \\ E(Y) &= E(X_1 + X_2 + X_3 + X_4 + X_5) = 5 \cdot E(X) = 5 \cdot \frac{1}{2} \\ &= \frac{5}{2} \end{aligned}$$

Using independence, we can find the variance of Y as:

$$\begin{aligned} E(X^2) &= \int_0^1 x^2 \cdot 6x(1-x) dx = 6 \cdot \int_0^1 x^3 - x^4 dx \\ &= 6 \cdot \left(\frac{1}{4} x^4 \Big|_0^1 - \frac{1}{5} x^5 \Big|_0^1 \right) = 6 \cdot \left(\frac{5}{20} - \frac{4}{20} \right) \\ &= \frac{3}{10} \\ \text{Var}(X) &= E(X^2) - E(X)^2 = \frac{3}{10} - \left(\frac{1}{2} \right)^2 = \frac{6}{20} - \frac{5}{20} = \frac{1}{20} \\ \text{Var}(Y) &= \text{Var}(X_1 + X_2 + X_3 + X_4 + X_5) = 5 \cdot \text{Var}(X) = 5 \cdot \frac{1}{20} \\ &= \frac{1}{4} \end{aligned}$$

Therefore, the mean and variance of Y are $\frac{5}{2}$ and $\frac{1}{4}$, respectively.

■

12. Let the independent random variables X_1 and X_2 have binomial distribution with parameters $n_1 = 3$, $p = \frac{2}{3}$ and $n_2 = 4$, $p = \frac{1}{2}$, respectively. Compute $P(X_1 = X_2)$.

Solution:

Since X_1 and X_2 are independent, we have

$$\begin{aligned}
 X_1 &\sim B(3, \frac{2}{3}) & P_{X_1}(x_1) &= C_{x_1}^3 \left(\frac{2}{3}\right)^{x_1} \cdot \left(\frac{1}{3}\right)^{3-x_1} \\
 X_2 &\sim B(4, \frac{1}{2}) & P_{X_2}(x_2) &= C_{x_2}^4 \left(\frac{1}{2}\right)^{x_2} \cdot \left(\frac{1}{2}\right)^{4-x_2} = C_{x_2}^4 \left(\frac{1}{2}\right)^4 \\
 P(X_1 = X_2) &= P(X_1 = X_2 = 0) + P(X_1 = X_2 = 1) + P(X_1 = X_2 = 2) + P(X_1 = X_2 = 3) \\
 &= P(X_1 = 0) \cdot P(X_2 = 0) + P(X_1 = 1) \cdot P(X_2 = 1) \\
 &\quad + P(X_1 = 2) \cdot P(X_2 = 2) + P(X_1 = 3) \cdot P(X_2 = 3) \\
 &= \left(1 \cdot \left(\frac{2}{3}\right)^0 \cdot \left(\frac{1}{3}\right)^3 \cdot 1 \cdot \left(\frac{1}{2}\right)^4\right) + \left(3 \cdot \left(\frac{2}{3}\right)^1 \cdot \left(\frac{1}{3}\right)^2 \cdot 4 \cdot \left(\frac{1}{2}\right)^4\right) \\
 &\quad + \left(3 \cdot \left(\frac{2}{3}\right)^2 \cdot \left(\frac{1}{3}\right)^1 \cdot 6 \cdot \left(\frac{1}{2}\right)^4\right) + \left(1 \cdot \left(\frac{2}{3}\right)^3 \cdot \left(\frac{1}{3}\right)^0 \cdot 4 \cdot \left(\frac{1}{2}\right)^4\right) \\
 &= \left(\frac{1}{27} \cdot \frac{1}{16}\right) + \left(\frac{6}{27} \cdot \frac{4}{16}\right) + \left(\frac{12}{27} \cdot \frac{6}{16}\right) + \left(\frac{8}{27} \cdot \frac{4}{16}\right) = \frac{1 + 24 + 72 + 32}{27 \cdot 16} = \frac{129}{432} \\
 &= \frac{43}{144}
 \end{aligned}$$

■

13. Let X have a Poisson distribution. If $P(X = 1) = P(X = 3)$, find the mode of the distribution.

Solution:

$$\begin{aligned}
 \frac{e^{-\mu} \mu}{1!} &= \frac{e^{-\mu} \mu^3}{3!} \text{ requires } \mu^2 = 6 \text{ and } \mu = \sqrt{6}. \\
 \text{Since } \frac{e^{-\sqrt{6}} (\sqrt{6})^2}{2!} &= 3 \cdot e^{-\sqrt{6}} > \frac{e^{-\sqrt{6}} \cdot \sqrt{6}}{1!}, \quad x = 2 \text{ is the mode.}
 \end{aligned}$$

■

14. If X is $N(1, 4)$, compute the probability $P(1 < X^2 < 9)$

Solution:

Since X follows a normal distribution with mean $\mu = 1$ and variance $\sigma^2 = 4$, we know that $Z = \frac{X - \mu}{\sigma}$ follows a standard normal distribution $N(0, 1)$. Therefore, we can transform the inequality $1 < X^2 < 9$ into $-2 < Z < 2$ as follows:

$$1 < X^2 < 9 \Rightarrow -3 < X < -1 \text{ or } 1 < X < 3 \Rightarrow -\frac{3-1}{2} < \frac{X-\mu}{\sigma} < \frac{3-1}{2} \Rightarrow -1 < Z < 1$$

Using a standard normal table, we can find $P(-1 < Z < 1)$ to be approximately 0.477. Therefore, we have:

$$P(1 < X^2 < 9) = P(-1 < Z < 1) \approx 0.477$$

■

15. Let the random variable X have a distribution that is $N(\mu, \sigma^2)$.
- (a) Does the random variable $Y = X^2$ also have a normal distribution?
- (b) Would the random variable $Y = aX + b$, a and b nonzero constants have a normal distribution?
- Hint:* In each case, first determine $P(Y \leq y)$.

Solution:

(a) No, the random variable $Y = X^2$ does not have a normal distribution.

To determine whether the random variable $Y = X^2$ also has a normal distribution, we need to find the distribution of Y .

First, we note that if $X \sim N(\mu, \sigma^2)$, then X has a symmetric bell-shaped distribution. When we square X , we remove the negative values of X since the square of any real number is non-negative. Therefore, the distribution of $Y = X^2$ is not symmetric anymore.

To find the distribution of Y , we can use the cumulative distribution function (CDF) method. Let $y > 0$ be a fixed value. Then we have:

$$\begin{aligned} P(Y \leq y) &= P(X^2 \leq y) \\ &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= \Phi\left(\frac{\sqrt{y} - \mu}{\sigma}\right) - \Phi\left(\frac{-\sqrt{y} - \mu}{\sigma}\right), \end{aligned}$$

where Φ is the standard normal CDF.

We can differentiate this expression with respect to y to find the PDF of Y :

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} P(Y \leq y) \\ &= \frac{1}{2\sqrt{y}\sigma} \left[\phi\left(\frac{\sqrt{y} - \mu}{\sigma}\right) + \phi\left(\frac{-\sqrt{y} - \mu}{\sigma}\right) \right], \end{aligned}$$

where ϕ is the standard normal PDF.

We see that the PDF of Y depends on the PDF of X , which is normal, and on the square root of Y . Therefore, $Y = X^2$ does not have a normal distribution.

Note that if X had a standard normal distribution, then $Y = X^2$ would have a chi-squared distribution with one degree of freedom.

(b)

Let $Y = aX + b$. To determine the distribution of Y , we can use the following formula for the expected value and variance of Y :

$$E(Y) = E(aX + b) = aE(X) + b = a\mu + b$$

$$Var(Y) = Var(aX + b) = a^2 Var(X) = a^2 \sigma^2$$

Therefore, Y is also a normal distribution, since it is a linear transformation of the normal random variable X .

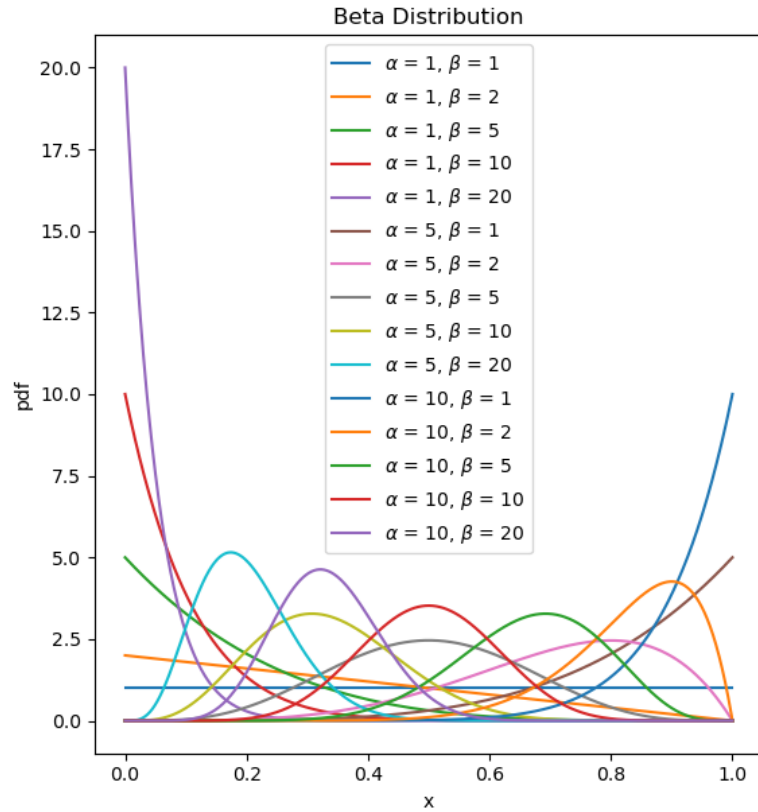
More specifically, Y has a normal distribution with mean $a\mu + b$ and variance $a^2 \sigma^2$.

■

16. (The original question 14.) Using the computer, obtain plots of beta pdfs for $\alpha = 1, 5, 10$ and $\beta = 1, 2, 5, 10, 20$.

Solution:

■



```

1  # -*- coding: utf-8 -*-
2  """
3  Created on Mon Mar 29 17:33:24 2023
4
5  @author: Paul
6  """
7
8  import numpy as np
9  import matplotlib.pyplot as plt
10
11  def plot_beta_dist():
12      alpha = [1, 5, 10]
13      beta_values = [1, 2, 5, 10, 20]
14      x = np.linspace(0, 1, 1000)
15
16      for a in alpha:
17          for b in beta_values:
18              y = x**(a-1) * (1-x)**(b-1)
19              y /= np.trapz(y, x)
20              plt.plot(x, y, label=r'$\alpha$ = {}, $\beta$ = {}'.format(a, b))
21
22      plt.title('Beta Distribution')
23      plt.xlabel('x')
24      plt.ylabel('pdf')
25      plt.legend()
26      plt.show()
27
28
29  if __name__ == "__main__":
30      plot_beta_dist()

```

Figure 1: Beta distribution and code implementation