CHAPTER 4: Mathematical Expectation

Mean of a Random Variable: 4.1

Definition 4.1: Let X be a random variable with probability distribution function $f(x)$. The **mean** or **expected value** of X is

$$\mu = E(X) = \sum_x x f(x)$$

if X is discrete, and

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\ dx$$

if X is continuous.

Example: Among the seven nominees for three vacancies on a city council are 3 men and 4 women. Let X be the number of men that filled these vacancies. Find $E(X)$.

Solution:

Example: Let X be the number of cylinders in the engine of the next car to be tuned up at a certain facility. The cost of a tune-up is related to X by $Y = h(X) = 20 + 3X + .5X^2$. Find $E(Y) = E(h(X))$ if

x	4	6	8
$f(x)$	.5	.3	.2

Solution:

Theorem 4.1: Let X be a random variable with probability distribution function $f(x)$. The **mean** or **expected value** of the random variable $g(X)$ is

$$E(g(X)) = \sum_x g(x) f(x)$$

if X is discrete, and

$$\mu = E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x)\ dx$$

if X is continuous.

Example: Suppose the probability mass function of the length (in millimeters) of computer cables is

$$f(x) = \begin{cases} 0.1 & \text{if } 1200 < x \leq 1210 \\ 0 & \text{otherwise.} \end{cases}$$

Find $E(X)$

Solution:

Example: Assume that

$$f(x) = \begin{cases} \frac{1}{3000} e^{-x/3000} & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Find $E(X)$

Solution:

Example: Assume that

$$f(x) = \begin{cases} x/2 & 0 < x < 2 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $E(X)$
(b) Find $E(X^2)$

Example: Assume that

$$f(x) = \begin{cases} k(1 - x^2) & -1 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the value of k.
(b) What is the probability that the thermometer reads above 0^oC?
(c) What is the probability that the reading is within 0.25^oC of the actual temperature?
(d) What is the mean reading?
(e) What is the median reading?

Definition 4.1 (discrete case only): Let X and Y be discrete random variables with joint probability distribution function $f(x, y)$. The **mean** or **expected value** of the random variable $g(X, Y)$ is

$$E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

Example:
In a randomly chosen lot of 1000 bolts, let X be the number that fail to meet a length specification, and let Y be the number that fail to meet a diameter specification. Assume that the joint probability mass function of X and Y is given in the following table.

x \ y	0	1	2
0	0.40	0.12	0.08
1	0.15	0.08	0.03
2	0.10	0.03	0.01

a. Find $P(X = 0 \text{ and } Y = 2)$.
b. Find $P(X > 0 \text{ and } Y \leq 1)$.
c. Find $P(X \leq 1)$.
d. Find $P(Y > 0)$.
e. Find the probability that all the bolts in the lot meet the length specification.
f. Find the probability that all the bolts in the lot meet the diameter specification.
g. Find the probability that all the bolts in the lot meet both specifications.
h. Find $E(X)$, $E(Y)$, $E(\min(X, Y))$.

Example:
The joint probability mass function for X =automobile policy deductible amount and Y =homeowner policy deductible amount is

x \ y	0	100	200
100	0.2	0.1	0.2
250	0.05	0.15	0.3

(a) Find $E(X))$
(b) Find $E(Y)$
(c) Find $E(XY)$

Solution:

<u>VARIANCE AND COVARIANCE: 4.2</u>

<u>Definition 4.3:</u> Let X be a random variable with probability distribution function $f(x)$ and mean μ. The variance of X is

$$\sigma^2 = E[(X - \mu)^2] = \sum_x (x - \mu)^2 f(x)$$

if X is discrete, and

$$\sigma^2 = E[(X - \mu)^2] == \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx$$

if X is continuous.
The positive square root of the variance σ, is called the **standard deviation** of X.

Formula:

$$\sigma^2 = E(X^2) - \mu^2$$

Example: You and a friend play a game where you each toss a balanced coin. If the upper faces on the coins are both tails, you win \$1; if the faces are both heads, you win \$2; if the coins do not match (one shows a head, the other a tail), you lose \$1 (win −\$1).
(a) Give the probability distribution for your winning, X, on a single play of this game.
(b) Find the mean and variance of X.
(c) How much should you pay to play this game if your net winning , the difference between the payoff and cost of playing, are to have mean 0?
Solution:

Example:
The pH water samples from a specific lake is a random variable X with probability mass function given by

$$f(x) = \begin{cases} (3/8)(7-x)^2, & \text{if } 5 \le x \le 7 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $E(X)$ and $\text{var}(X)$
(b) Would you expect to see a pH measurement below 5.5 very often? Why?

Solution:

The covariance of X and Y is defined as

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)]$$

- The sign of the covariance indicates whether the relationship between two dependent random variables is positive or negative.
- When X and Y are statistically independent, it can be shown that the covariance is 0.

Formula: $\sigma_{XY} = E(XY) - \mu_X \mu_Y$.

Example:
Contracts for two construction jobs are randomly assigned to one or more of the three firms, A, B, and C. Let X denote the number of contracts assigned to firm A, and Y the number of contracts assigned to firm B. Recall that each firm can receive 0, 1, or 2 contracts.
(a) Find the joint probability function for X and Y.
(b) Find σ_{XY}. Does it surprise you that σ_{XY} is negative? Why?

Solution

Although the covariance between two random variables does provide information regarding the nature of the relationship, the magnitude of σ_{XY} does not indicate anything regarding the strength of the relationship, since σ_{XY} is not scale free. Its magnitude will depend on the units measured for both X and Y. There is a scale-free version of the covariance called the **correlation coefficient**.

Definition 4.5: Let X and Y be random variables with covariance σ_{XY} and standard deviations σ_X and σ_Y, respectively. The correlation coefficient X and Y is

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Example:
The joint probability mass function for X =automobile policy deductible amount and Y =homeowner policy deductible amount is

x \ y	0	100	200
100	0.2	0.1	0.2
250	0.05	0.15	0.3

Find ρ_{XY}.

Proposition

1. For any two random variables X and Y, $-1 \leq \rho_{XY} \leq 1$.
2. If X and Y are independent, then $\rho = 0$, but $\rho = 0$ does not imply independence.
3. $\rho = 1$ or -1 iff $Y = aX + b$ for some numbers a and b with $a \neq 0$.

- The proposition says that ρ is a measure of the degree of **linear** relationship between X and Y.
- A ρ less than 1 in absolute value indicates only that the relationship is not completely linear, but there may still be a very strong nonlinear relation.
- $\rho = 0$ does not imply that X and Y are independent, but only that there is complete absence of a linear relationship. When $\rho = 0$, X and Y are said to be **uncorrelated**.
- Two variables could be uncorrelated yet highly dependent because there is a strong nonlinear relationship, so be careful not to conclude too much from knowing that $\rho = 0$.

Example:
Let X and Y be discrete random variables with joint probability mass function

$$f(x,y) = \begin{cases} \frac{1}{4} & \text{(x,y)=(-4,1),(4,-1),(2,2),(-2,-2)} \\ 0 & \text{otherwise.} \end{cases}$$

Review

1. Let X be a random variable with density function

$$f(x) = \begin{cases} 24x^2 & 0 < x < a \\ 0 & \text{otherwise.} \end{cases}$$

Find a.

2. Suppose that $f(x, y)$, the joint probability mass function of X and Y is given by

$$f(0,0) = 0.5, \quad f(0,1) = 0.1, \quad f(1,0) = 0.1, \quad f(1,1) = 0.3$$

Compute the correlation coefficient of X and Y.

3. A special roulette wheel is made that only has the ten numbers $1, 2, 3, \ldots, 10$ as possible outcomes. A player bets \$2 on a number. If that number is spun, the player gains \$15; otherwise, the player loses the \$2. Find the expected value is:

4. Let

$$F(x) = \begin{cases} 0 & x < -1 \\ \frac{x^3}{2} + \frac{1}{2} & -1 \leq x \leq 1 \\ 1 & x > 1. \end{cases}$$

be a cumulative distribution function. Find $p(-2 \leq X \leq -0.5)$

SOLUTIONS

1. $1 = \int_0^a 24x^2 \; dx = 8a^3$.

2.
$p(X = 0) = 0.6 \quad p(X = 1) = 0.4$
$E(X) = 0.4 \quad E(X^2) = 0.4$ and so $V(X) = 0.4 - (0.4)(0.4) = 0.4 - 0.16 = 0.24$
Also $E(Y) = 0.4$ and $V(Y) = 0.24$.
$E(XY) = 0.3$ and so
$\mathrm{COV}(X,\ Y) = \sigma_{XY} = 0.3 - (0.4)(0.4) = 0.14$
Thus

$$\rho(X, Y) = \frac{0.14}{\sqrt{(0.24)(0.24)}} = \frac{0.14}{0.24} = 0.58$$

3. $-\$0.30$

4. $p(-2 \le X \le -0.5) = F(-0.5) - F(-2) = \frac{(-0.5)^3}{2} + \frac{1}{2} - 0 = 7/16$