Chapter 4:

Mathematical Expectation

Mean of a Random Variable:

Definition:

Let X be a random variable with a probability distribution $f(x)$. The mean (or expected value) of X is denoted by μ_X (or E(X)) and is defined by:

$$E(X) = \mu_X = \begin{cases} \sum_{all\ x} x\, f(x)\ ; & if\ X\ is\ discrete \\ \int_{-\infty}^{\infty} x\, f(x)\, dx\ ; & if\ X\ is\ continuous \end{cases}$$

Example:

A shipment of 8 similar microcomputers to a retail outlet contains 3 that are defective and 5 are non-defective. If a school makes a random purchase of 2 of these computers, find the expected number of defective computers purchased

Solution:

Let X = the number of defective computers purchased. we found that the probability distribution of X is:

x	0	1	2
f(x)= P(X=x)	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

or:

$$f(x) = P(X = x) = \begin{cases} \dfrac{\binom{3}{x} \times \binom{5}{2-x}}{\binom{8}{2}}; & x = 0, 1, 2 \\ 0; \; otherwise & \end{cases}$$

The expected value of the number of defective computers purchased is the mean (or the expected value) of X, which is:

$$E(X) = \mu_X = \sum_{x=0}^{2} x.f(x)$$

$$= 0.f(0) + 1.f(1) + 2.f(2)$$

$$= (0)\,\frac{10}{28} + (1)\,\frac{15}{28} + (2)\,\frac{3}{28}$$

$$= \frac{15}{28} + \frac{6}{28} = \frac{21}{28} = 0.75 \text{ (computers)}$$

Example:

A box containing 7 components is sampled by a quality inspector; the box contains 4 good components and 3 defective components. A sample of 3 is taken by the inspector. Find the expected value of the number of good components in this sample.

Solution:

Let X represent the number of good components in the sample. The probability distribution of X is

$$f(x) = \frac{\binom{4}{x}\binom{3}{3-x}}{\binom{7}{3}}, \qquad x = 0, 1, 2, 3.$$

Simple calculations yield $f(0) = 1/35$, $f(1) = 12/35$, $f(2) = 18/35$, and $f(3) = 4/35$. Therefore,

$$\mu = E(X) = (0)\left(\frac{1}{35}\right) + (1)\left(\frac{12}{35}\right) + (2)\left(\frac{18}{35}\right) + (3)\left(\frac{4}{35}\right) = \frac{12}{7} = 1.7.$$

Thus, if a sample of size 3 is selected at random over and over again from a box of 4 good components and 3 defective components, it will contain, on average, 1.7 good components.

Example

Let X be a continuous random variable that represents the life (in hours) of a certain electronic device. The pdf of X is given by:

$$f(x) = \begin{cases} \dfrac{20{,}000}{x^3} \; ; \; x > 100 \\ 0 \; ; \; elsewhere \end{cases}$$

Find the expected life of this type of devices.

Solution:

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

$$= \int_{100}^{\infty} x \frac{20000}{x^3}\, dx$$

$$= 20000 \int_{100}^{\infty} \frac{1}{x^2}\, dx$$

$$= 20000 \left[-\frac{1}{x} \Big|_{x=100}^{x=\infty} \right]$$

$$= -20000 \left[0 - \frac{1}{100} \right] = 200 \text{ hours}$$

Therefore, we can expect this type of device to last, on *average*, 200 hours.

Theorem

Let X be a random variable with a probability distribution $f(x)$, and let $g(X)$ be a function of the random variable X. The mean (or expected value) of the random variable $g(X)$ is denoted by $\mu_{g(X)}$ (or $E[g(X)]$) and is defined by:

$$E[g(X)] = \mu_{g(X)} = \begin{cases} \sum_{all\ x} g(x) f(x) ; & if\ X\ is\ discrete \\ \int_{-\infty}^{\infty} g(x) f(x) dx ; & if\ X\ is\ continuous \end{cases}$$

Example:

Let X be a discrete random variable with the following probability distribution

x	0	1	2
f(x)	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

Find $E[g(X)]$, where $g(X) = (X - 1)^2$.

Solution:

$$g(X)=(X-1)^2$$

$$E[g(X)]=\mu_{g(X)}=\sum_{x=0}^{2} g(x)\,f(x)=\sum_{x=0}^{2}(x-1)^2\,f(x)$$

$$=(0-1)^2\,f(0)+(1-1)^2\,f(1)+(2-1)^2\,f(2)$$

$$=(-1)^2\,\frac{10}{28}+(0)^2\,\frac{15}{28}+(1)^2\,\frac{3}{28}$$

$$=\frac{10}{28}+0+\frac{3}{28}=\frac{13}{28}$$

Example:

Suppose that the number of cars X that pass through a car wash between 4:00 P.M. and 5:00 P.M. on any sunny Friday has the following probability distribution:

x	4	5	6	7	8	9
$P(X = x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Let $g(X) = 2X - 1$ represent the amount of money, in dollars, paid to the attendant by the manager. Find the attendant's expected earnings for this particular time period.

Solution:

$$E[g(X)] = E(2X - 1) = \sum_{x=4}^{9} (2x - 1) f(x)$$

$$= (7)\left(\frac{1}{12}\right) + (9)\left(\frac{1}{12}\right) + (11)\left(\frac{1}{4}\right) + (13)\left(\frac{1}{4}\right)$$

$$+ (15)\left(\frac{1}{6}\right) + (17)\left(\frac{1}{6}\right) = \$12.67.$$

Example

Let X be a continuous random variable that represents the life (in hours) of a certain electronic device. The pdf of X is given by:

$$f(x)=\begin{cases}\dfrac{20{,}000}{x^3} & ;\ x>100\\ 0 ;\ elsewhere\end{cases}$$

Find $E\left(\frac{1}{x}\right)$ $\left\{\text{note: } g(X)=\dfrac{1}{X}\right\}$

Solution:

$$f(x)=\begin{cases}\dfrac{20{,}000}{x^3} \;;\; x>100\\ 0 \;;\; elsewhere\end{cases}$$

$$g(X)=\frac{1}{X}$$

$$E\left(\frac{1}{X}\right) = E[g(X)]=\mu_{g(X)} = \int_{-\infty}^{\infty} g(x)\,f(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{x} f(x)\,dx$$

$$= \int_{100}^{\infty} \frac{1}{x}\frac{20000}{x^3}dx = 20000\int_{100}^{\infty} \frac{1}{x^4}dx = \frac{20000}{-3}\left[\frac{1}{x^3}\middle|\begin{matrix}x=\infty\\x=100\end{matrix}\right]$$

$$= \frac{-20000}{3}\left[0-\frac{1}{1000000}\right] = 0.0067$$

Definition

Let X and Y be random variables with joint probability distribution $f(x, y)$. The mean, or expected value, of the random variable $g(X, Y)$ is

$$\mu_{g(X,Y)} = E[g(X, Y)] = \sum_{x} \sum_{y} g(x, y) f(x, y)$$

if X and Y are discrete, and

$$\mu_{g(X,Y)} = E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) \, dx \, dy$$

if X and Y are continuous.

Example

Let X and Y be the random variables with joint probability distribution

	$f(x,y)$	x = 0	1	2	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Find the expected value of $g(X,Y) = XY$.

Solution:

$$E(XY) = \sum_{x=0}^{2}\sum_{y=0}^{2} xyf(x,y)$$

$$= (0)(0)f(0,0) + (0)(1)f(0,1)$$

$$+ (1)(0)f(1,0) + (1)(1)f(1,1) + (2)(0)f(2,0)$$

$$= f(1,1) = \frac{3}{14}.$$

Exercises

4.1 – 4.2 – 4.10 – 4.13 and 4.17 on page 117

Variance (of a Random Variable)

The most important measure of variability of a random variable X is called the variance of X and is denoted by $Var(X)$ or σ^2

Definition

Let X be a random variable with a probability distribution $f(x)$ and mean μ. The variance of X is defined by:

$$\mathrm{Var}(X) = \sigma_X^2 = E[(X-\mu)^2] = \begin{cases} \sum_{all\ x} (x-\mu)^2 f(x)\ ; & if\ X\ is\ discrete \\ \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx\ ; & if\ X\ is\ continuous \end{cases}$$

Definition: (standard deviation)

The positive square root of the variance of X, $\sigma_X = \sqrt{\sigma_X^2}$, is called the standard deviation of X.

Note:

Var(X)=E[g(X)], where $g(X)=(X-\mu)^2$

Theorem

The variance of the random variable X is given by:

$$\text{Var}(X) = \sigma_X^2 = E(X^2) - \mu^2$$

$$\text{where } E(X^2) = \begin{cases} \sum_{all\ x} x^2 f(x)\,; & if\ X\ is\ discrete \\ \int_{-\infty}^{\infty} x^2 f(x)\,dx\,; & if\ X\ is\ continuous \end{cases}$$

Proof : See page 121

Example

Let X be a discrete random variable with the following probability distribution

x	0	1	2	3
f(x)	0.51	0.38	0.10	0.01

Find Var(X)=σ_X^2 .

Solution:

$$\mu = \sum_{x=0}^{3} x\, f(x) = (0)\, f(0) + (1)\, f(1) + (2)\, f(2) + (3)\, f(3)$$

$$= (0)\,(0.51) + (1)\,(0.38) + (2)\,(0.10) + (3)\,(0.01)$$

$$= 0.61$$

1. First method:

$$\mathrm{Var}(X)=\sigma_X^2=\sum_{x=0}^{3}(x-\mu)^2 f(x)$$

$$=\sum_{x=0}^{3}(x-0.61)^2 f(x)$$

$$=(0-0.61)^2\,f(0)+(1-0.61)^2\,f(1)+(2-0.61)^2\,f(2)+(3-0.61)^2\,f(3)$$

$$=(-0.61)^2\,(0.51)+(0.39)^2\,(0.38)+(1.39)^2\,(0.10)+(2.39)^2\,(0.01)$$

$$=0.4979$$

2. Second method:

$$Var(X) = \sigma_X^2 = E(X^2) - \mu^2$$

$$E(X^2) = \sum_{x=0}^{3} x^2 f(x) = (0^2)\, f(0) + (1^2)\, f(1) + (2^2)\, f(2) + (3^2)\, f(3)$$

$$= (0)\,(0.51) + (1)\,(0.38) + (4)\,(0.10) + (9)\,(0.01)$$

$$= 0.87$$

$$Var(X) = \sigma_X^2 = E(X^2) - \mu^2 = 0.87 - (0.61)^2 = 0.4979$$

Example

Let X be a continuous random variable with the following pdf:

$$f(x)=\begin{cases}2(x-1) \; ; \; 1<x<2 \\ 0 \; ; \; elsewhere\end{cases}$$

Find the mean and the variance of X.

Solution:

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{1}^{2} x[2(x-1)]\, dx = 2\int_{1}^{2} x(x-1)\, dx = 5/3$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \int_{1}^{2} x^2[2(x-1)]\, dx = 2\int_{1}^{2} x^2(x-1)\, dx = 17/6$$

$$Var(X) = \sigma_X^2 = E(X^2) - \mu^2 = 17/6 - (5/3)^2 = 1/18$$

Means and Variances of Linear Combinations of Random Variables

If $X_1, X_2, \ldots, X_n$ are n random variables and $a_1, a_2, \ldots, a_n$ are constants, then the random variable :

$$Y = \sum_{i=1}^{n} a_i X_i = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$$

is called a linear combination of the random variables $X_1, X_2, \ldots, X_n$.

Theorem

If X is a random variable with mean $\mu = E(X)$, and if a and b are constants, then:

$$E(aX \pm b) = a\,E(X) \pm b \Leftrightarrow \mu_{aX \pm b} = a\mu_X \pm b$$

Corollary 1: E(b) = b (a=0 in Theorem)

Corollary 2: E(aX) = a E(X) (b=0 in Theorem)

Example

Let X be a random variable with the following probability density function:

$$f(x)=\begin{cases}\frac{1}{3}x^2 \ ; -1<x<2\\ 0 \ ; \ elsewhere\end{cases}$$

Find E(4X+3).

Solution:

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{-1}^{2} x[\frac{1}{3}x^2]\, dx = \frac{1}{3}\int_{-1}^{2} x^3\, dx = \frac{1}{3}\left[\frac{1}{4}x^4 \Big|_{x=-1}^{x=2}\right] = 5/4$$

$$E(4X+3) = 4\,E(X)+3 = 4(5/4) + 3 = 8$$

Another solution:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx \quad ;\ g(X) = 4X+3$$

$$E(4X+3) = \int_{-\infty}^{\infty} (4x+3)\, f(x)\, dx = \int_{-1}^{2} (4x+3)\, [\frac{1}{3}x^2]\, dx = \cdots = 8$$

Theorem:

If X_1, X_2, …, X_n are n random variables and a_1, a_2, …, a_n are constants, then:

$$E(a_1X_1+a_2X_2+\ldots+a_nX_n) = a_1E(X_1)+a_2E(X_2)+\ldots+a_nE(X_n)$$

$$\Leftrightarrow$$

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i)$$

Corollary:

If X, and Y are random variables, then:

$$E(X \pm Y) = E(X) \pm E(Y)$$

Theorem

If X is a random variable with variance

$$Var(x) = \sigma_x^{\ 2}$$

and if a and b are constants, then:

$$\mathrm{Var}(aX \pm b) = a^2\ \mathrm{Var}(X)$$

$$\Leftrightarrow$$

$$\sigma^2_{aX+b} = a^2 \sigma^2_X$$

Theorem:

If $X_1, X_2, \ldots, X_n$ are n independent random variables and $a_1, a_2, \ldots, a_n$ are constants, then:

$$Var(a_1X_1+a_2X_2+\ldots+a_nX_n) = a_1^2 Var(X_1)+ a_2^2 Var(X_2)+\ldots+ a_n^2 Var(X_n)$$

$$\Leftrightarrow$$

$$Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 Var(X_i)$$

$$\Leftrightarrow$$

$$\sigma^2_{a_1X_1+a_2X_2+\ldots+a_nX_n} = a_1^2\sigma^2_{X_1} + a_2^2\sigma^2_{X_2} + \ldots + a_n^2\sigma^2_{X_n}$$

Corollary:

If X, and Y are independent random variables, then:

- $Var(aX+bY) = a^2 Var(X) + b^2 Var(Y)$
- $Var(aX-bY) = a^2 Var(X) + b^2 Var(Y)$
- $Var(X \pm Y) = Var(X) + Var(Y)$

Example:

Let X, and Y be two independent random variables such that $E(X)=2$, $Var(X)=4$, $E(Y)=7$, and $Var(Y)=1$. Find:

1. $E(3X+7)$ and $Var(3X+7)$
2. $E(5X+2Y-2)$ and $Var(5X+2Y-2)$.

Solution:

1. $E(3X+7) = 3E(X)+7 = 3(2)+7 = 13$

$Var(3X+7)= (3)^2 \, Var(X)=(3)^2 \, (4) =36$

2. $E(5X+2Y-2)= 5E(X) + 2E(Y) -2= (5)(2) + (2)(7) - 2= 22$

$Var(5X+2Y-2)= Var(5X+2Y)= 5^2 \; Var(X) + 2^2 \; Var(Y) =$

$(25)(4)+(4)(1) = 104$

Chebyshev's Theorem

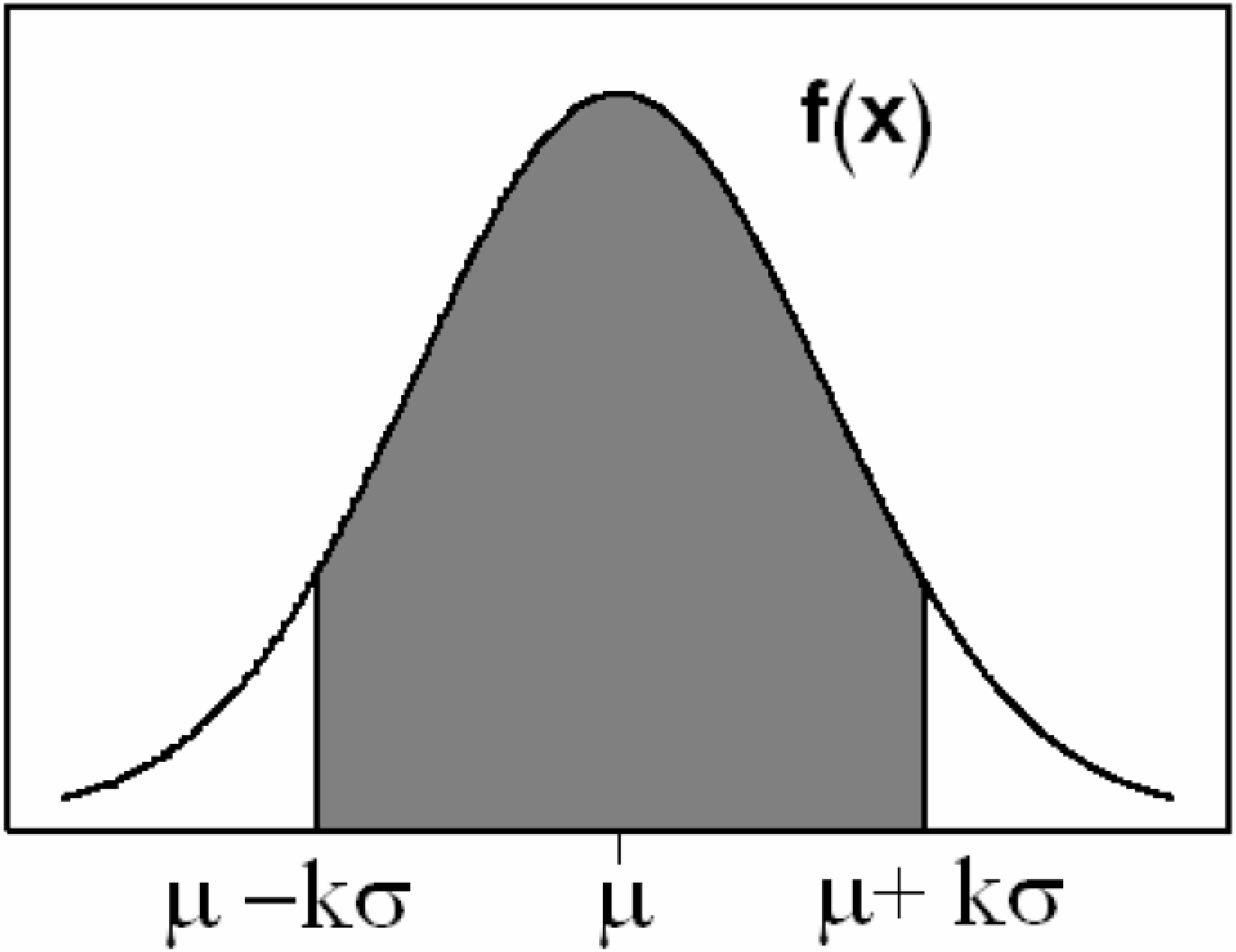
Suppose that X is any random variable with mean $E(X) = \mu$ and variance $Var(X) = \sigma^2$ and standard deviation σ.

Chebyshev's Theorem gives a conservative estimate of the probability that the random variable X assumes a value within k standard deviations ($k\sigma$) of its mean μ, which is

$$P(\mu - k\sigma < X < \mu + k\sigma).$$

$$P(\mu - k\sigma < X < \mu + k\sigma) \approx 1 - \frac{1}{k^2}$$

$$\text{area} = P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$



Theorem

Let X be a random variable with mean $E(X) = \mu$ and variance $Var(X) = \sigma^2$, then for $k > 1$, we have:

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

$$\Leftrightarrow \quad P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

Example

Let X be a random variable having an unknown distribution with mean $\mu=8$ and variance $\sigma^2=9$ (standard deviation $\sigma=3$). Find the following probability:

(a) $P(-4 < X < 20)$

(b) $P(|X-8| \geq 6)$

Solution:

(a) $P(-4 < X < 20) = ??$

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

$$(-4 < X < 20) = (\mu - k\sigma < X < \mu + k\sigma)$$

$$-4 = \mu - k\sigma \Leftrightarrow -4 = 8 - k(3)$$
$$\Leftrightarrow -4 = 8 - 3k$$
$$\Leftrightarrow 3k = 12$$
$$\Leftrightarrow \mathbf{k=4}$$

or

$$20 = \mu + k\sigma \Leftrightarrow 20 = 8 + k(3)$$
$$\Leftrightarrow 20 = 8 + 3k$$
$$\Leftrightarrow 3k = 12$$
$$\Leftrightarrow \mathbf{k=4}$$

$$1 - \frac{1}{k^2} = 1 - \frac{1}{16} = \frac{15}{16}$$

Therefore, $P(-4 < X < 20) \geq \frac{15}{16}$, and hence, $P(-4 < X < 20) \approx \frac{15}{16}$

$$(b)\ P(|X - 8| \geq 6) = ??$$

$$P(|X-8| \geq 6) = 1 - P(|X-8| < 6)$$

$$P(|X - 8| < 6) = ??$$

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

$$(|X-8| < 6) = (|X - \mu| < k\sigma)$$

$$6 = k\sigma \Leftrightarrow 6 = 3k \Leftrightarrow \mathbf{k=2}$$

$$1 - \frac{1}{k^2} = 1 - \frac{1}{4} = \frac{3}{4}$$

$$P(|X-8| < 6) \geq \frac{3}{4} \Leftrightarrow 1 - P(|X-8| < 6) \leq 1 - \frac{3}{4}$$

$$\Leftrightarrow 1 - P(|X-8| < 6) \leq \frac{1}{4}$$

$$\Leftrightarrow P(|X-8| \geq 6) \leq \frac{1}{4}$$

Therefore, $P(|X-8| \geq 6) \approx \frac{1}{4}$ (approximately)

Another solution for part (b):

$P(|X-8| < 6) = P(-6 < X-8 < 6)$

$= P(-6 + 8 < X < 6+8)$

$= P(2 < X < 14)$

$(2 < X < 14) = (\mu - k\sigma < X < \mu + k\sigma)$

$2 = \mu - k\sigma \Leftrightarrow 2 = 8 - k(3) \Leftrightarrow 2 = 8 - 3k \Leftrightarrow 3k = 6 \Leftrightarrow \mathbf{k=2}$

$$1-\frac{1}{k^2}=1-\frac{1}{4}=\frac{3}{4}$$

$$P(2<X<14)\geq\frac{3}{4}\iff P(|X-8|<6)\geq\frac{3}{4}$$

$$\iff 1-P(|X-8|<6)\leq 1-\frac{3}{4}$$

$$\iff 1-P(|X-8|<6)\leq\frac{1}{4}$$

$$\iff P(|X-8|\geq 6)\leq\frac{1}{4}$$

Therefore, $P(|X-8|\geq 6)\approx\frac{1}{4}$ (approximately)

Definition:

Let X and Y be random variables with joint probability distribution $f(x, y)$. The covariance of X and Y is

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \sum_x \sum_y (x - \mu_X)(y - \mu_y) f(x, y)$$

if X and Y are discrete, and

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_y) f(x, y)\, dx\, dy$$

if X and Y are continuous.

Theorem:

The covariance of two random variables X and Y with means μ_X and μ_Y , respectively, is given by

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

Proof: See page 124

Definition

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

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

Notes:

- ρ_{XY} is free of the units of X and Y.
- The correlation coefficient satisfies the inequality $-1 \leq \rho_{XY} \leq 1$.
- It assumes a value of zero when $\sigma_{XY} = 0$

Example

Find the covariance and the correlation coefficient of X and Y for the following joint distribution

	$f(x,y)$	x = 0	1	2	$h(y)$
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
	$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Solution:

$$
\begin{aligned}
E(XY) &= \sum_{x=0}^{2}\sum_{y=0}^{2} xyf(x,y) \\
&= (0)(0)f(0,0) + (0)(1)f(0,1) \\
&\quad + (1)(0)f(1,0) + (1)(1)f(1,1) + (2)(0)f(2,0) \\
&= f(1,1) = \frac{3}{14}.
\end{aligned}
$$

$$\mu_X = \sum_{x=0}^{2} xg(x) = (0)\left(\frac{5}{14}\right) + (1)\left(\frac{15}{28}\right) + (2)\left(\frac{3}{28}\right) = \frac{3}{4},$$

$$\mu_Y = \sum_{y=0}^{2} yh(y) = (0)\left(\frac{15}{28}\right) + (1)\left(\frac{3}{7}\right) + (2)\left(\frac{1}{28}\right) = \frac{1}{2}.$$

Therefore,

$$\sigma_{XY} = E(XY) - \mu_X \mu_Y = \frac{3}{14} - \left(\frac{3}{4}\right)\left(\frac{1}{2}\right) = -\frac{9}{56}.$$

To calculate the correlation coefficient

$$E(X^2) = (0^2)\left(\frac{5}{14}\right) + (1^2)\left(\frac{15}{28}\right) + (2^2)\left(\frac{3}{28}\right) = \frac{27}{28}$$

and

$$E(Y^2) = (0^2)\left(\frac{15}{28}\right) + (1^2)\left(\frac{3}{7}\right) + (2^2)\left(\frac{1}{28}\right) = \frac{4}{7},$$

we obtain

$$\sigma_X^2 = \frac{27}{28} - \left(\frac{3}{4}\right)^2 = \frac{45}{112} \text{ and } \sigma_Y^2 = \frac{4}{7} - \left(\frac{1}{2}\right)^2 = \frac{9}{28}.$$

Therefore, the correlation coefficient between X and Y is

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{-9/56}{\sqrt{(45/112)(9/28)}} = -\frac{1}{\sqrt{5}}.$$

Example:

The fraction *X* of male runners and the fraction *Y* of female runners who compete in marathon races are described by the joint density function

$$f(x,y) = \begin{cases} 8xy, & 0 \le y \le x \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the covariance and the correlation coefficient of X and Y .

Solution:

We first compute the marginal density functions. They are

$$g(x) = \begin{cases} 4x^3, & 0 \le x \le 1, \\ 0, & \text{elsewhere}, \end{cases}$$

and

$$h(y) = \begin{cases} 4y(1-y^2), & 0 \le y \le 1, \\ 0, & \text{elsewhere}. \end{cases}$$

From these marginal density functions, we compute

$$\mu_X = E(X) = \int_0^1 4x^4 \, dx = \frac{4}{5} \text{ and } \mu_Y = \int_0^1 4y^2(1-y^2) \, dy = \frac{8}{15}.$$

From the joint density function given above, we have

$$E(XY) = \int_0^1 \int_y^1 8x^2y^2 \, dx \, dy = \frac{4}{9}.$$

Then

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y = \frac{4}{9} - \left(\frac{4}{5}\right)\left(\frac{8}{15}\right) = \frac{4}{225}.$$

To calculate the correlation coefficient

$$E(X^2) = \int_0^1 4x^5 \, dx = \frac{2}{3} \text{ and } E(Y^2) = \int_0^1 4y^3(1-y^2) \, dy = 1 - \frac{2}{3} = \frac{1}{3},$$

we conclude that

$$\sigma_X^2 = \frac{2}{3} - \left(\frac{4}{5}\right)^2 = \frac{2}{75} \text{ and } \sigma_Y^2 = \frac{1}{3} - \left(\frac{8}{15}\right)^2 = \frac{11}{225}.$$

Hence,

$$\rho_{XY} = \frac{4/225}{\sqrt{(2/75)(11/225)}} = \frac{4}{\sqrt{66}}.$$