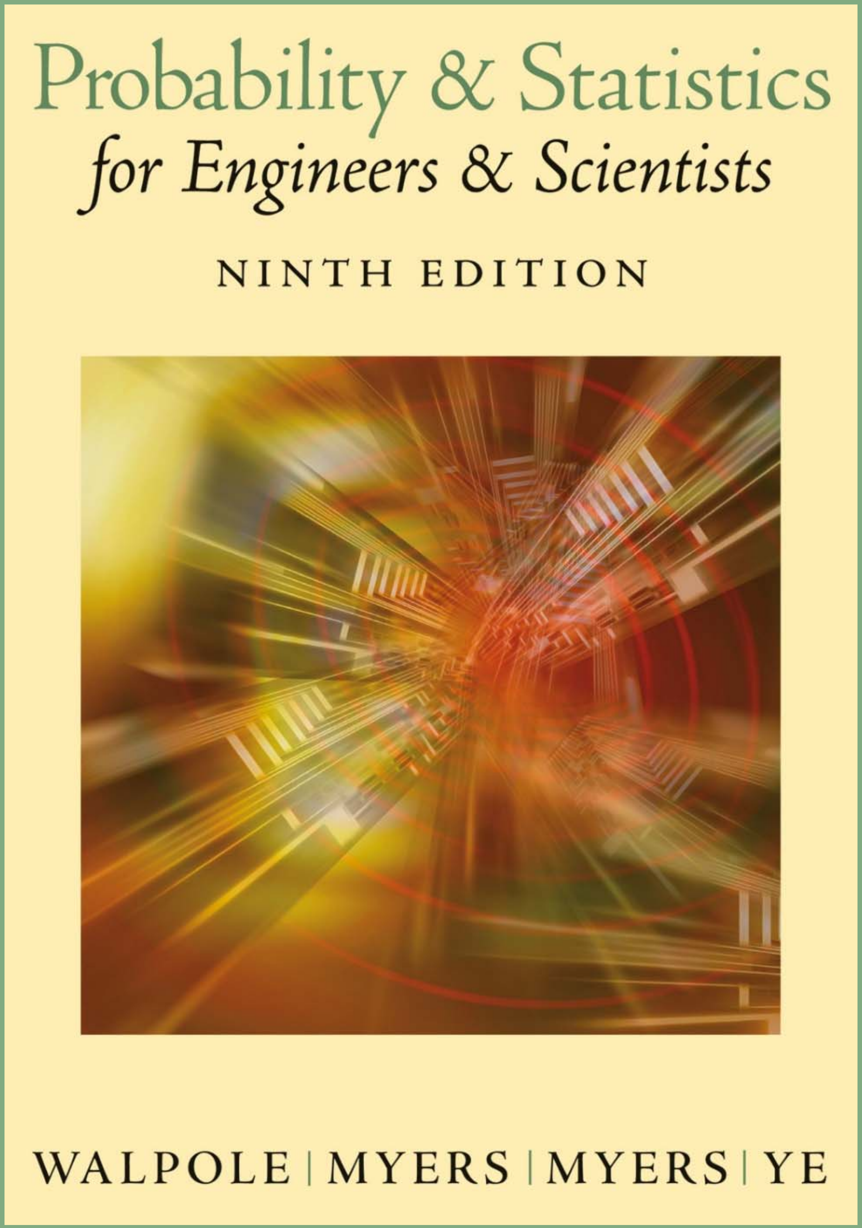
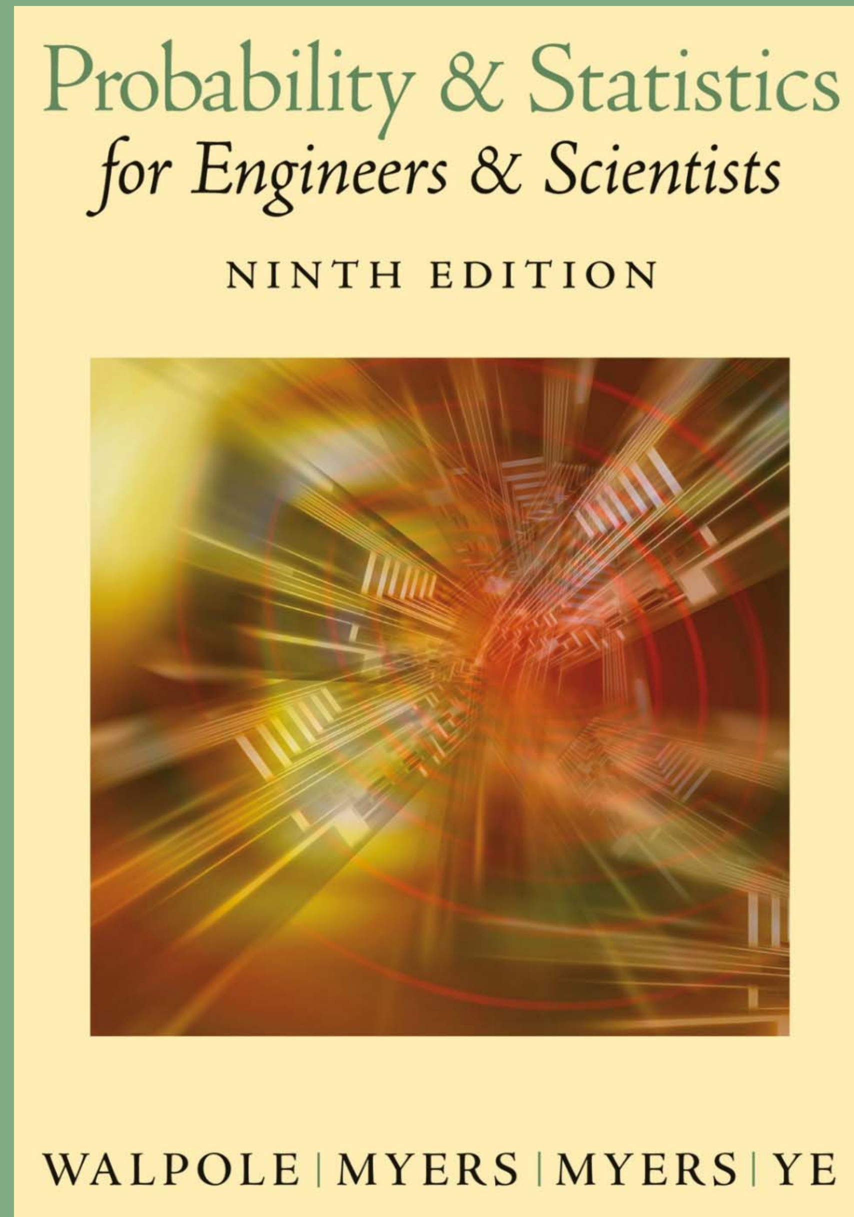
Chapter 4

Mathematical Expectation



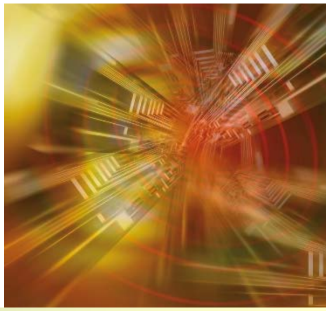
Section 4.1

Mean of a Random Variable



PEARSON
Addison Wesley

Mean / Expected value
Definition



Definition:
Let X be a random variable with probability / Density function f(x). The mean or expected value of X is give by

$$\mu = E(X) = \sum_{x} x\, f(x)$$

if X is discrete, and

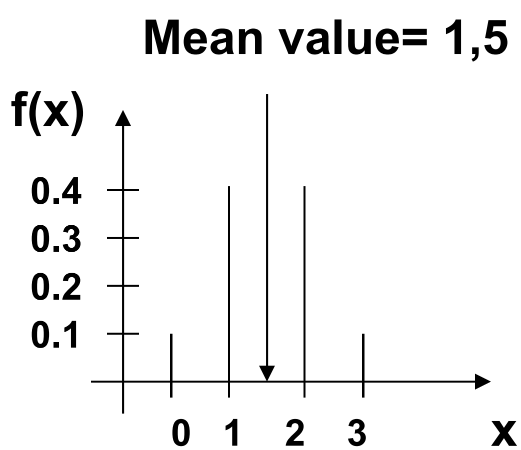
$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

if X is continuous.

Mean / Expected value
Interpretation

Interpretation:
The total contribution of a value multiplied by the probability of the value – a weighted average.

Example:



Example

Example: **(Example 3.3)**

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

Solution:

Let X = the number of defective computers purchased.

In Example 3.3, 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:

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

Solution

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)

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

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

Example

Example 4.3:

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:

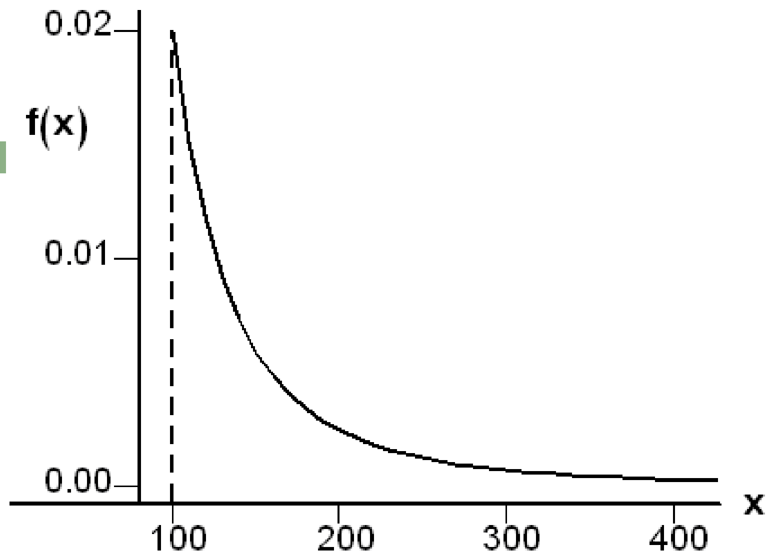
$$E(X)=\mu_X=\int_{-\infty}^{\infty} x\, f(x)\, dx$$

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

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

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

$$=-20000\left[0-\frac{1}{100}\right]=200 \quad \textbf{(hours)}$$



Therefore, we expect that this type of electronic devices to last, on average, 200 hours.

Mean / Expected value
Function of a random variable

Theorem:
Let X be a random variable with probability / density function f(x). The expected value of g(X) is

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

if X is discrete, and

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

if X is continuous.

Example

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:

In Example 4.3, find $E\left(\frac{1}{X}\right)$ **.**

Solution:

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

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

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

$$= \int_{100}^{\infty} \frac{1}{x} \frac{20000}{x^3} dx = 20000 \int_{100}^{\infty} \frac{1}{x^4} dx = \frac{20000}{-3} \left[\frac{1}{x^3} \Bigg|_{x=100}^{x=\infty} \right]$$

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

Expected value
Linear combination

Theorem: Linear combination
Let X be a random variable (discrete or continuous), and let a and b be constants. For the random variable aX + b we have

$$E(aX+b) = aE(X)+b$$

Mean / Expected value
Joint Function of two random variables

Definition:
Let X and Y be random variables with joint probability / density function f(x,y). The expected value of 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.

Mean / Expected value
Function of two random variables

Problem:

Burger King sells both via "drive-in" and "walk-in".
Let X and Y be the fractions of the opening hours that "drive-in" and "walk-in" are busy.
Assume that the joint density for X and Y are given by

$$f(x,y) = \begin{cases} 4xy & 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

The turn over g(X,Y) on a single day is given by

$$g(X,Y) = 6X + 9Y$$

What is the expected turn over on a single day?

Mean / Expected value
Sums and products

Theorem: Sum/Product
Let X and Y be random variables then

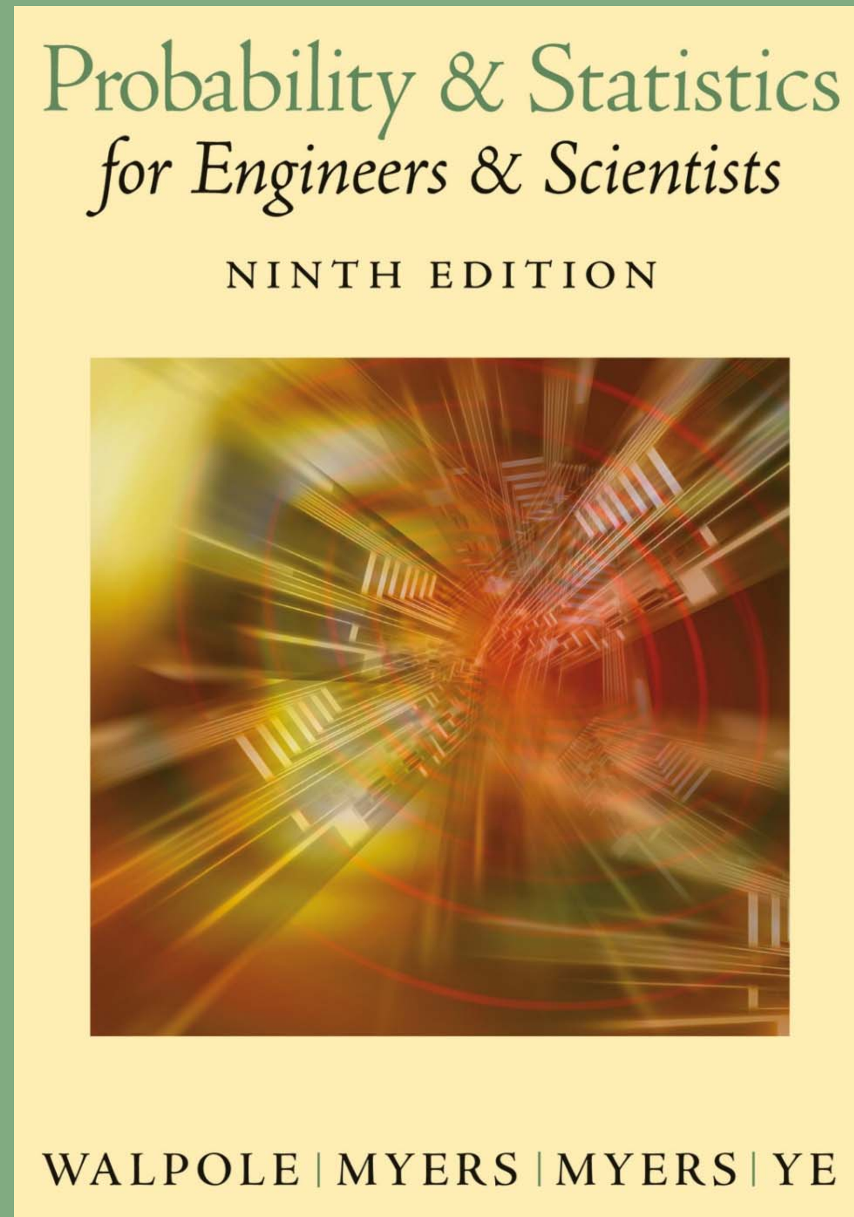
$$E[X+Y] = E[X] + E[Y]$$

If X and Y are independent then

$$E[X \cdot Y] = E[X] \cdot E[Y]$$

Section 4.2

Variance and Covariance of Random Variables



Variance

4.2 Variance (of a Random Variable): σ_X^2

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

Definition 4.3:

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

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

Definition:

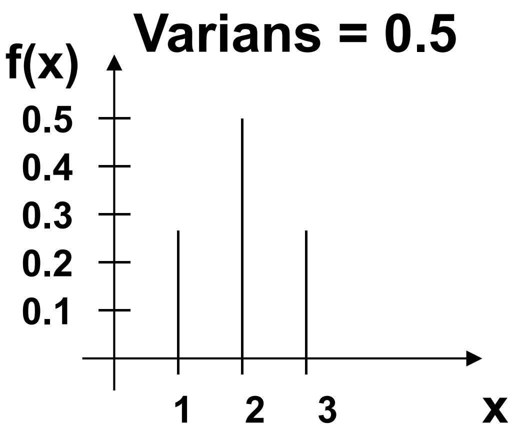
The positive square root of the variance of X, $\sigma_X=\sqrt{\sigma_X^2}$,is called the **standard deviation** of X.

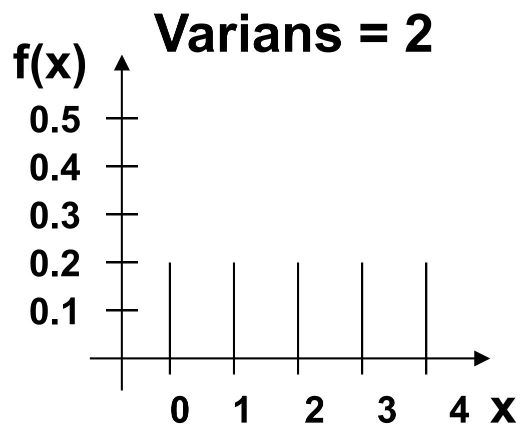
Note:

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

Variance Interpretation

The variance expresses, how dispersed the density / probability function is around the mean.





Rewrite of the variance: $\boxed{\sigma^2 = Var(X) = E[X^2] - \mu^2}$

Theorem 4.2:

The variance of the random variable X is given by:

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

where

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

Example 4.9:

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

x	0	1	2	3
f(x)	0.15	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:

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

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

$$= (0)\,(0.51) + (1)\,(0.38) + (4)\,(0.10) + (9)\,(0.01) = 0.87$$

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

Example

Example 4.10:

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:

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

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

$$Var(X) = \sigma_X^2 = E(X^2) - \mu^2 = \mathbf{17/6 - (5/3)^2 = 1/8}$$

Variance
Linear combinations

Theorem: Linear combination
Let X be a random variable, and let a be b constants.
For the random variable aX + b the variance is

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

Examples:

Var (X + 7) = Var (X)
Var (-X) = Var (X)
Var (2X) = 4 Var (X)

Covariance
Definition

Definition:
Let X and Y be to random variables with joint probability / density function f(x,y). The covariance between X and Y is

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

if X and Y are discrete, and

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

if X and Y are continuous.

Covariance
Interpretation

Covariance between X and Y expresses how X and Y influence each other.

Examples: Covariance between

- **X = sale of bicycle and Y = bicycle pumps is positive.**
- **X = Trips booked to Spain and Y = outdoor temperature is negative.**
- **X = # eyes on red dice and Y = # eyes on the green dice is zero.**

Covariance
Properties

Theorem:
The covariance between two random variables X and Y with means μ_X and μ_Y, respectively, is

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

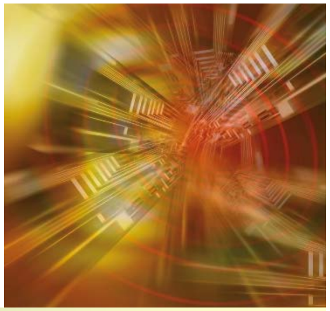
Notice! **Cov (X,X) = Var (X)**

If X and Y are independent random variables, then

Cov (X,Y) = 0

Notice! Cov(X,Y) = 0 does not imply independence!

Variance/Covariace
Linear combinations



Theorem: **Linear combination**

Let X and Y be random variables, and let a and b be constants.

For the random variables aX + bY the variance is

$$Var(aX+bY)=a^2\,Var(X)+b^2\,Var(Y)+2ab\,Cov(X,Y)$$

Specielt: Var[X+Y] = Var[X] + Var[Y] +2Cov (X,Y)

If X and Y are independent, the variance is

Var[X+Y] = Var[X] + Var[Y]

Correlation
Definition

Definition:
Let X and Y be two random variables with covariance Cov (X,Y) and standard deviations σ_X and σ_Y, respectively.
The correlation coefficient of X and Y is

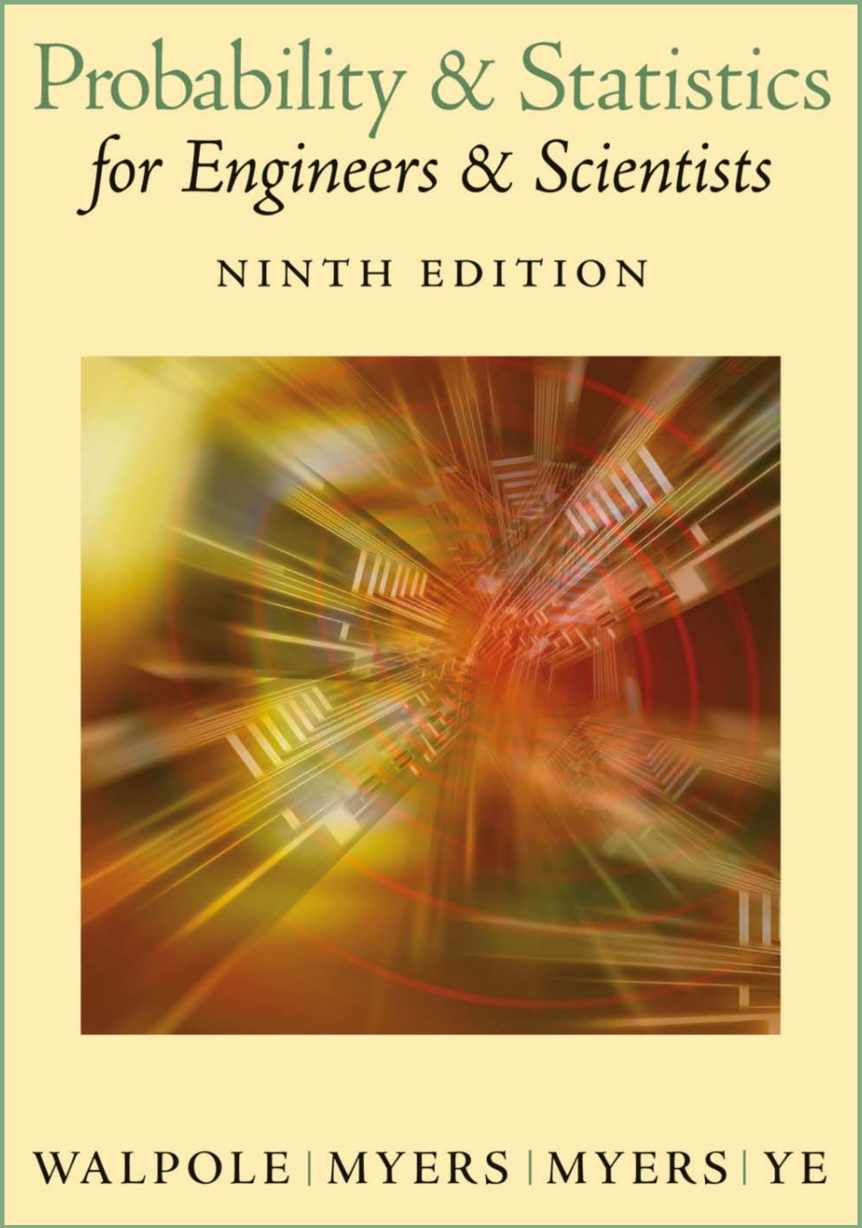
$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

It holds that $-1 \leq \rho_{XY} \leq 1$

If X and Y are independent, then $\rho_{XY} = 0$

Section 4.3

Means and Variances of Linear Combinations of Random Variables



Theorem 4.5:

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 4.5)

Corollary 2: $E(aX) = a\,E(X)$ (b=0 in Theorem 4.5)

Example 4.16:

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:

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

E(4X+3) = 4 E(X)+3 = 4(5/4) + 3=8

Another solution:

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

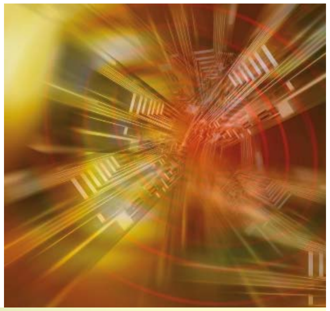
E(4X+3) = $\int_{-\infty}^{\infty} (4x+3)\, f(x)\, dx = \int_{-1}^{2} (4\mathrm{x}+3)\,[\frac{1}{3}\mathrm{x}^2]\ \mathrm{dx} = \cdots = 8$

4.3 Means and Variances of Linear Combinations of Random Variables:

If X_1, X_2, …, X_n are n random variables and a_1, a_2, …, 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,…,X_n.



Theorem:

If $X_1, X_2, \ldots, X_n$ are n random variables and $a_1, a_2, \ldots, a_n$ 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 4.9:

If X is a random variable with variance $Var(X) = \sigma_X^2$ **and if a and b are constants, then:**

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

$$\Leftrightarrow$$

$$\sigma_{aX+b}^2 = a^2 \sigma_X^2$$

Theorem:

If X_1, X_2, …, X_n are n <u>independent</u> random variables and a_1, a_2, …, a_n are constants, then:

$Var(a_1X_1+a_2X_2+\ldots+a_nX_n)$

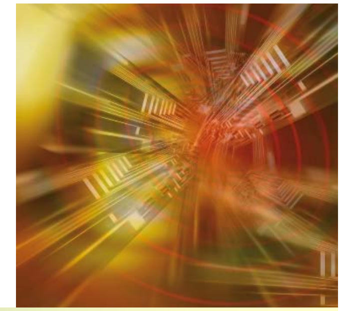
$$= a_1^2 Var(X_1) + a_2^2 Var(X_2) + \ldots + a_n^2 Var(X_n)$$

$$\Leftrightarrow$$

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

$$\Leftrightarrow$$

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



Corollary:

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

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

Let X and Y be two independent random variables. Then $\sigma_{XY} = 0$.

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 $Var(5X+2Y-2)$.

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. $E(5X+2Y-2)= 5E(X) + 2E(Y) -2= (5)(2) + (2)(7) - 2= 22$
   $Var(5X+2Y-2)= Var(5X+2Y)= 5^2\,Var(X) + 2^2\,Var(Y)$
   $= (25)(4)+(4)(1) = 104$

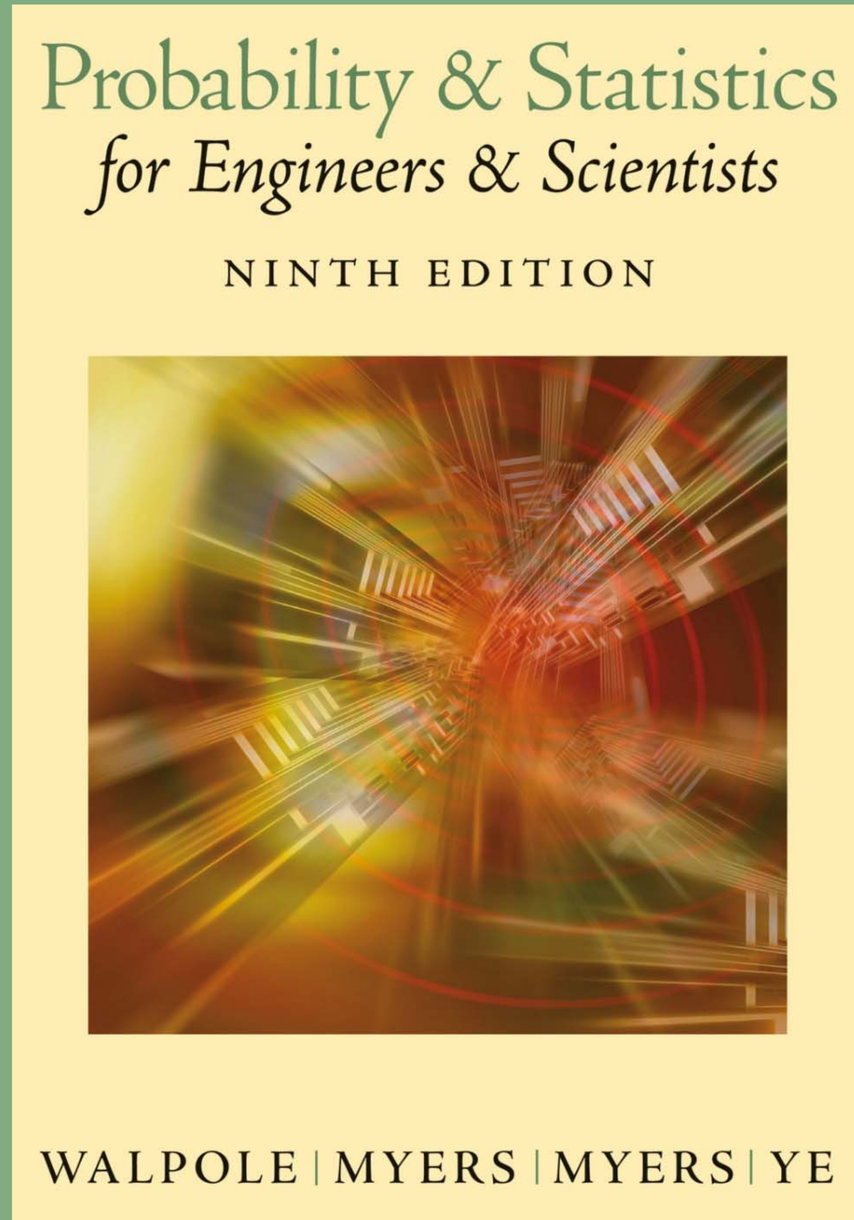
Theorem 4.7

The expected value of the sum or difference of two or more functions of the random variables X and Y is the sum or difference of the expected values of the functions. That is,

$$E[g(X,Y) \pm h(X,Y)] = E[g(X,Y)] \pm E[h(X,Y)].$$

Section 4.4

Chebyshev's Theorem

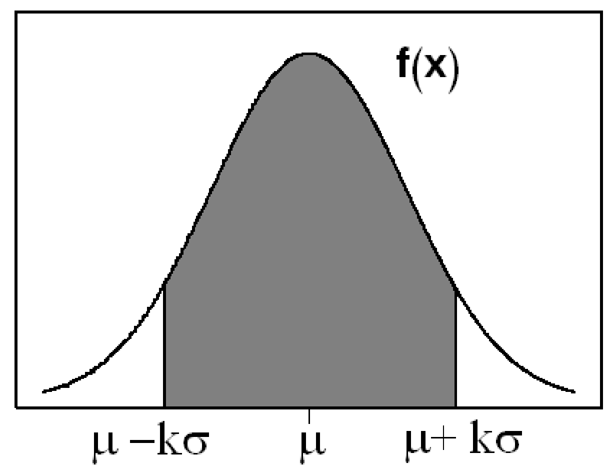


4.4 Chebyshev's Theorem:

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

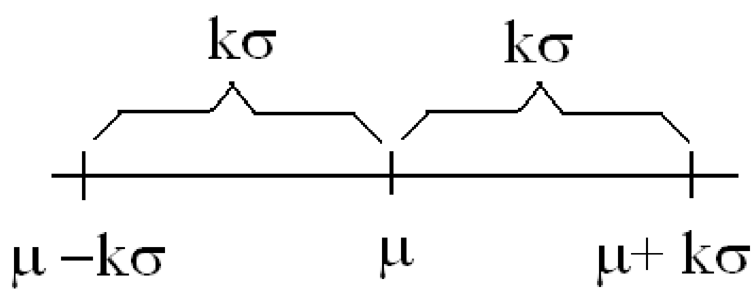
* $P(\mu - k\sigma < X < \mu + k\sigma) \approx 1 - \dfrac{1}{k^2}$

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



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

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



Example 4.22:

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:

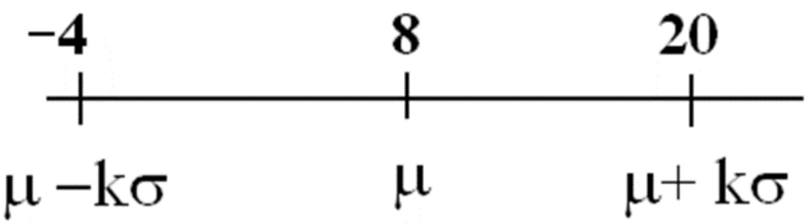
(a) $P(-4 < X < 20)$

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

Solution:

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

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



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

$-4 = \mu - k\sigma \Leftrightarrow -4 = 8 - k(3)$ or $20 = \mu + k\sigma \Leftrightarrow 20 = 8 + k(3)$

$\Leftrightarrow -4 = 8 - 3k$ $\quad\quad \Leftrightarrow 20 = 8 + 3k$

$\Leftrightarrow 3k = 12$ $\quad\quad \Leftrightarrow 3k = 12$

$\Leftrightarrow k = 4$ $\quad\quad \Leftrightarrow k = 4$

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

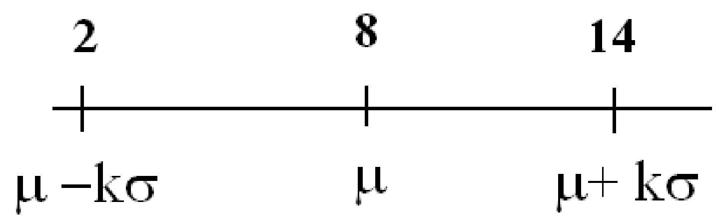
Therefore, $P(-4 < X < 20) \geq \frac{15}{16}$, and hence, $P(-4 < X < 20) \approx \frac{15}{16}$ (approximately)

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

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

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

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



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

$6 = k\sigma \Leftrightarrow 6 = 3k \Leftrightarrow k = 2$

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

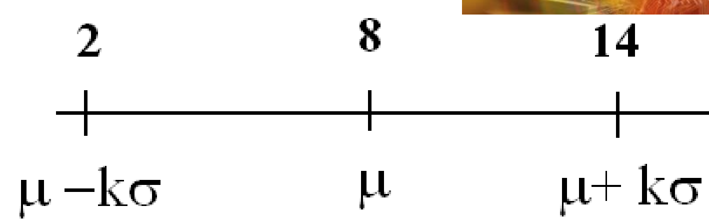
$$P(|X-8| < 6) \geq \frac{3}{4} \Leftrightarrow 1 - P(|X-8| < 6) \leq 1 - \frac{3}{4}$$

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

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

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

Another solution for part (b):



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

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

$= P(2 < X < 14)$

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

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

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

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

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

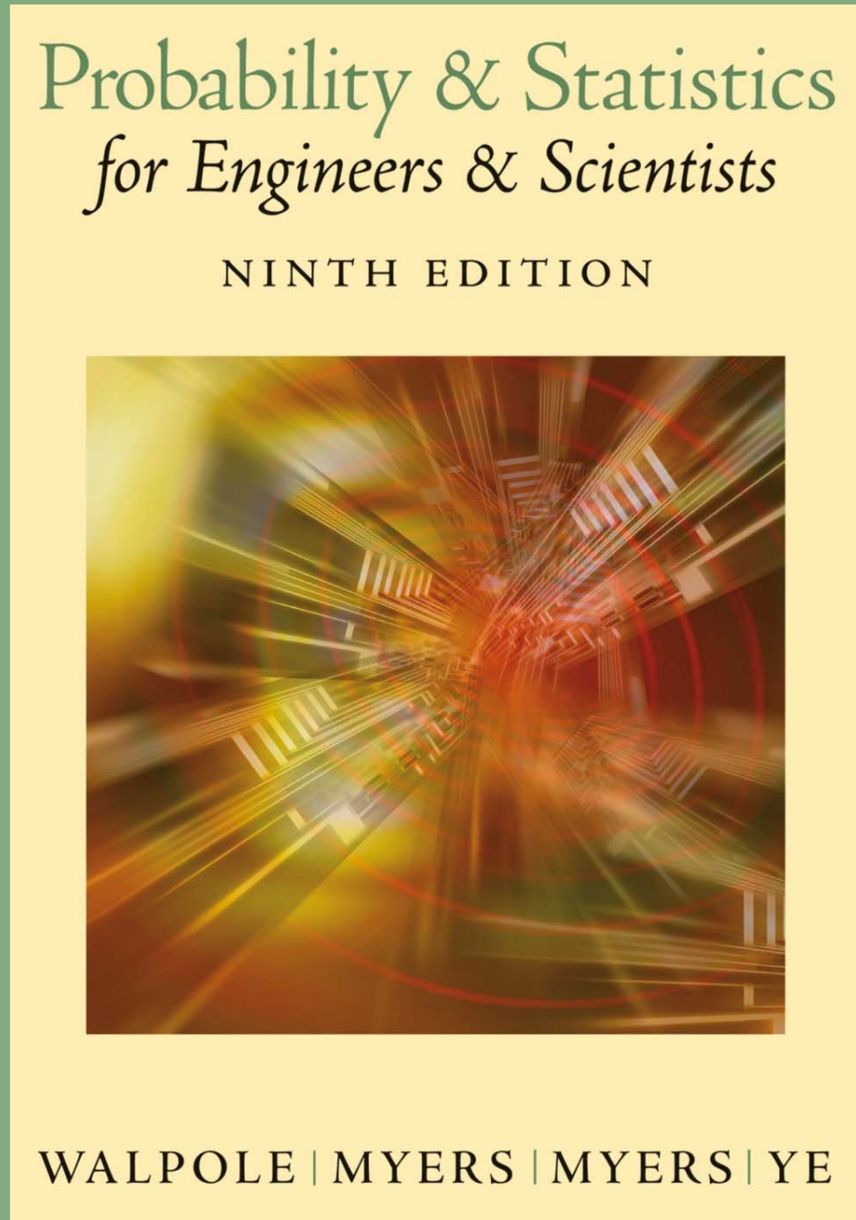
$\Leftrightarrow 1 - P(|X-8| < 6) \leq \frac{1}{4}$

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

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

Section 4.5

Review



Mean / Expected value
Definition

Definition:
Let X be a random variable with probability / Density function f(x). The mean or expected value of X is give by

$$\mu = E(X) = \sum_{x} x f(x)$$

if X is discrete, and

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

if X is continuous.

Definition 4.1

Let X be a random variable with probability distribution $f(x)$. 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.

Theorem 4.2

The variance of a random variable X is

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

Definition 4.3

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

$$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x), \qquad \text{if } X \text{ is discrete, and}$$

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx, \qquad \text{if } X \text{ is continuous.}$$

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

Theorem 4.3

Let X be a random variable with probability distribution $f(x)$. The variance of the random variable $g(X)$ is

$$\sigma^2_{g(X)} = E\{[g(X) - \mu_{g(X)}]^2\} = \sum_x [g(x) - \mu_{g(X)}]^2 f(x)$$

if X is discrete, and

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

if X is continuous.

Definition 4.4

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 4.6

The expected value of the sum or difference of two or more functions of a random variable X is the sum or difference of the expected values of the functions. That is,

$$E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)].$$