STATISTICAL ANALYSIS - LECTURE 5

Dr. Mahmoud Mounir

mahmoud.mounir@cis.asu.edu.eg

Discrete Probability Distributions

Note: This PowerPoint is only a summary and your main source should be the book.

Introduction

5-1 Probability Distributions

5-2 Mean , Variance, Standard Deviation ,and Expectation

5-3 The Binomial Distribution

Note: This PowerPoint is only a summary and your main source should be the book.

Probability Distributions

- A **random variable**
  - It is a function that associates a real number with each element in the sample space.
  - It is a variable whose values are determined by chance.
- Classify variables as **discrete** or **continuous**.
- A **discrete probability distribution** consists of the values a random variable can assume and the corresponding probabilities of the values.

Note: This PowerPoint is only a summary and your main source should be the book.

For example 1: $S = \{T, H\}$

X= number of heads

X= 0 , 1

$P(x=0) = P(T) = \frac{1}{2}$

$\frac{no.\,of\ events}{no.\,of\ sample\ space}$

$P(x=1) = P(H) = \frac{1}{2}$

Probability Distribution Table

X	0	1
P(X)	$\frac{1}{2}$	$\frac{1}{2}$

Note: This PowerPoint is only a summary and your main source should be the book.

For example 2: S={TT , HT , TH , HH}

X= number of heads

X= 0 , 1 , 2

$$P(x=0) = P(TT) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$P(x=1) = P(HT) + P(TH) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{2}{4}$$

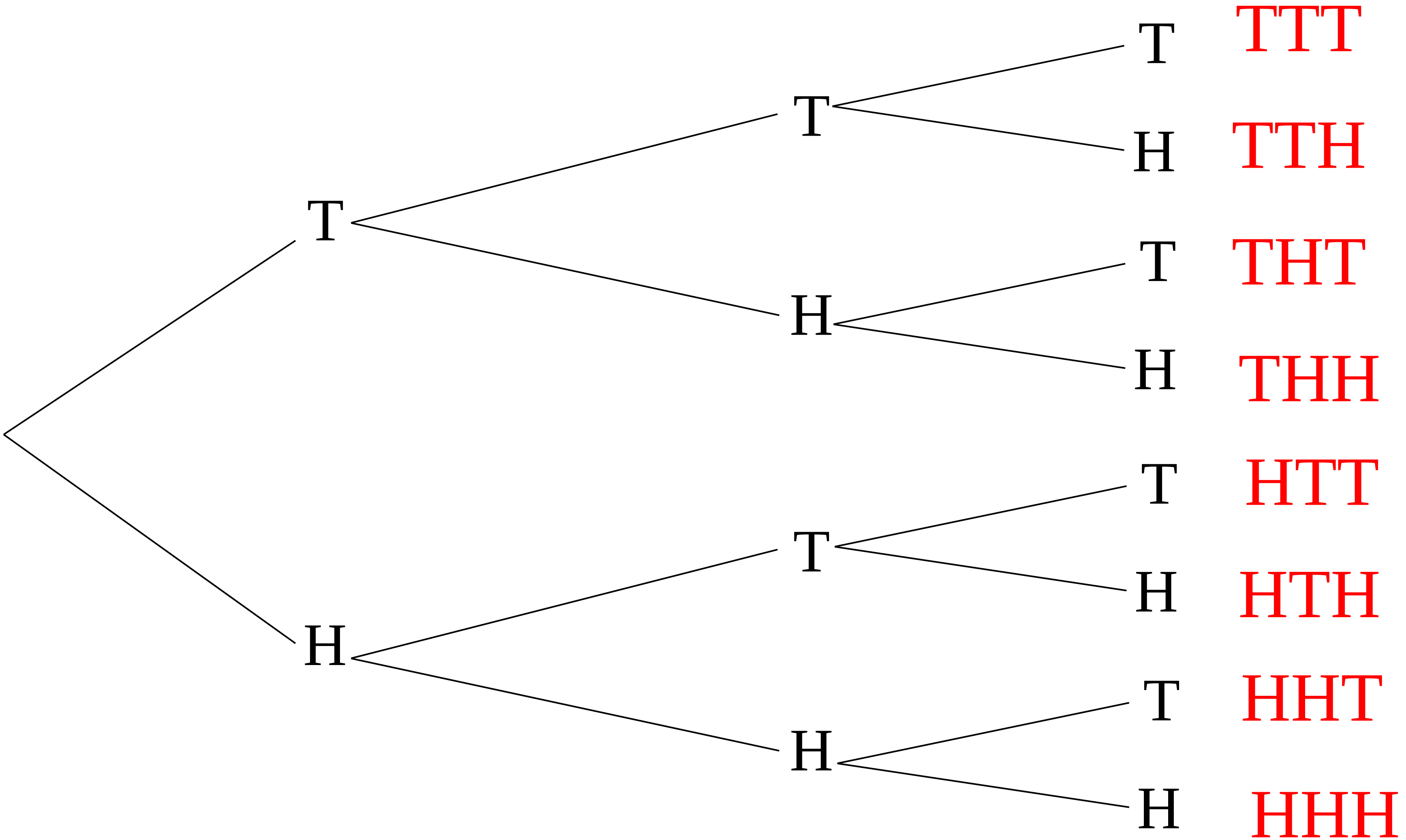
$$P(x=2) = P(HH) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Probability Distribution Table

X	0	1	2
P(X)	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Note: This PowerPoint is only a summary and your main source should be the book.

For example 3: Tossing three coins



Note: This PowerPoint is only a summary and your main source should be the book.

S={TTT , TTH , THT , HTT , HHT , HTH , THH , HHH}

X= number of heads

X= 0 , 1, 2 , 3

P(x=0) = P(TTT) = $\frac{1}{2} \bullet \frac{1}{2} \bullet \frac{1}{2} = \frac{1}{8}$

P(x=1) = P(TTH) + P(THT) + P(HTT)

$$= \frac{1}{2} \bullet \frac{1}{2} \bullet \frac{1}{2} + \frac{1}{2} \bullet \frac{1}{2} \bullet \frac{1}{2} + \frac{1}{2} \bullet \frac{1}{2} \bullet \frac{1}{2} = \frac{3}{8}$$

P(x=2) = P(HHT) + P(HTH) + P(THH)

$$= \frac{1}{2} \bullet \frac{1}{2} \bullet \frac{1}{2} + \frac{1}{2} \bullet \frac{1}{2} \bullet \frac{1}{2} + \frac{1}{2} \bullet \frac{1}{2} \bullet \frac{1}{2} = \frac{3}{8}$$

P(x=3) = P(HHH)= $\frac{1}{2} \bullet \frac{1}{2} \bullet \frac{1}{2} = \frac{1}{8}$

Note: This PowerPoint is only a summary and your main source should be the book.

Probability Distribution Table

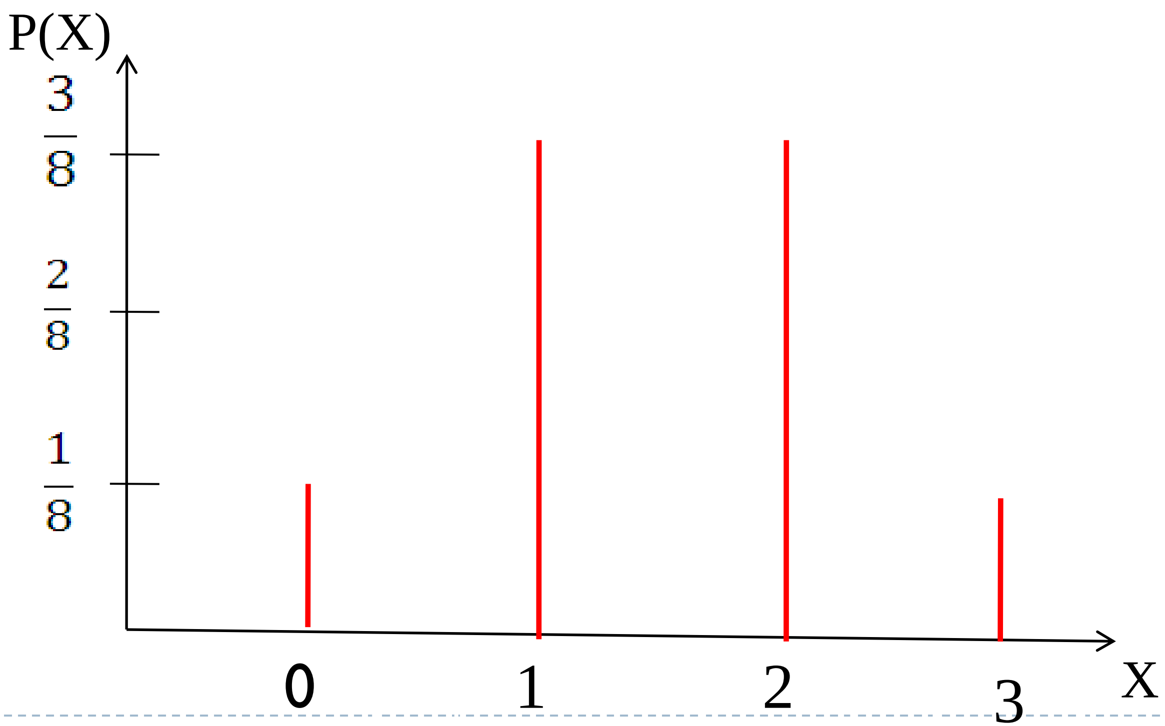
Number of heads (X)	0	1	2	3
Probability P(X)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Example 5-2: Tossing Coins

Represent graphically the probability distribution for the sample space for tossing three coins .

X	0	1	2	3
P(X)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Solution :



Note: This PowerPoint is only a summary and your main source should be the book.

Example 5-3:

The baseball World Series is played by the winner of the National League and the American League. The first team to win four games, wins the world sreies.In other words ,the series will consist of four to seven games, depending on the individual victories. The data shown consist of the number of games played in the world series from 1965 through 2005.The number of games (X) .Find the probability P(X) for each X ,construct a probability distribution, and draw a graph for the data.

x	**Number of games played**
4	8
5	7
6	9
7	16

$$\text{For 4 games} = \frac{8}{40} = 0.200$$

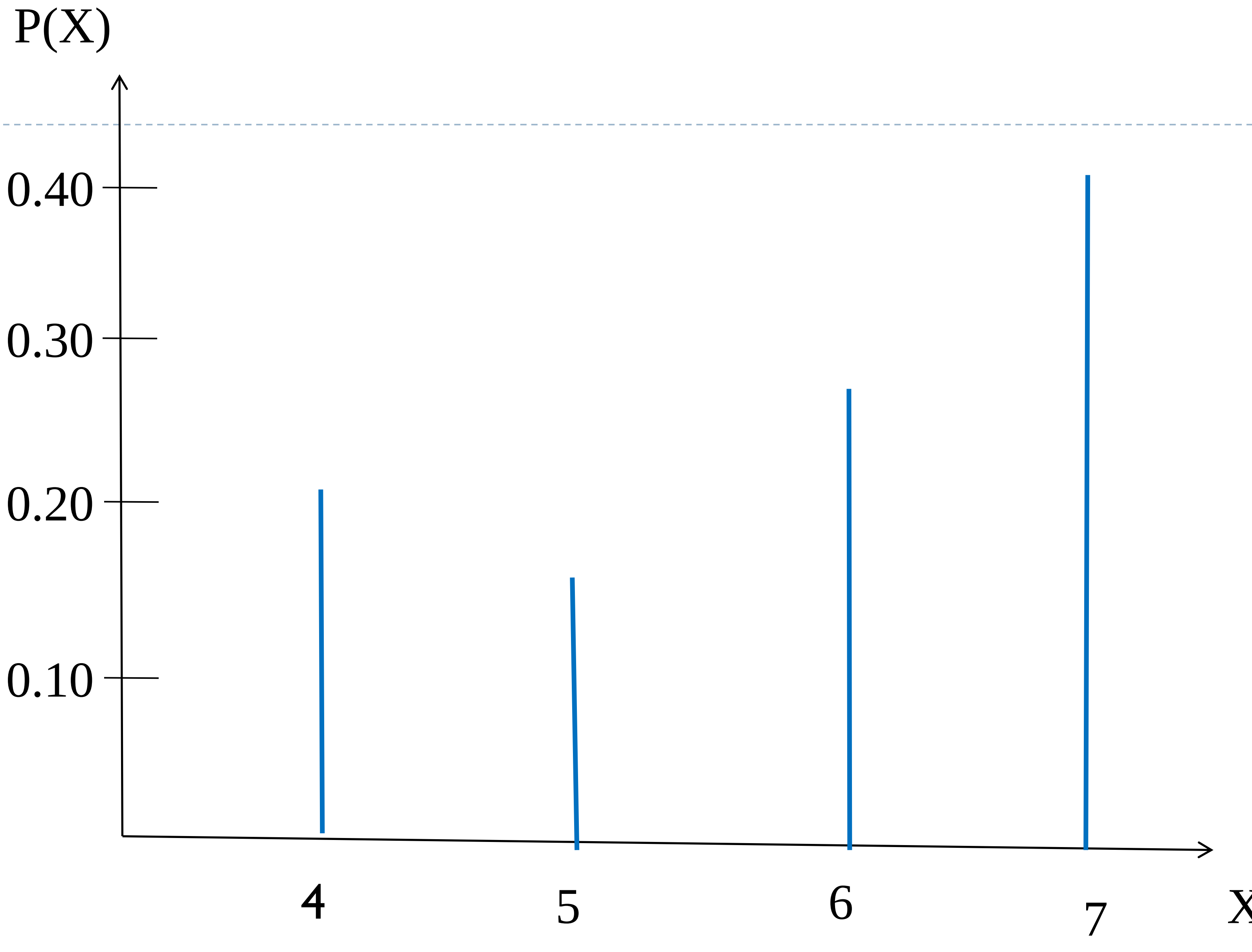
$$\text{For 5 games} = \frac{7}{40} = 0.175$$

$$\text{For 6 games} = \frac{9}{40} = 0.225$$

$$\text{For 7 games} = \frac{16}{40} = 0.400$$

Probability Distribution Table

X	4	5	6	7
P(X)	0.200	0.175	0.225	0.400


P(X)
0.40
0.30
0.20
0.10
4
5
6
7
X

- The sum of the probabilities of all events in a sample space add up to 1.

$$\sum p(x) = 1$$

- Each probability is between 0 and 1, inclusively.

$$0 \leq P(x) \leq 1$$

Example 5-4: Determine whether each distribution is a probability distribution.

X	0	5	10	15	20	
P(X)	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	√

X	0	2	4	6	
P(X)	-1	1.5	0.3	0.2	✖

X	1	2	3	4	
P(X)	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{9}{16}$	√

X	2	3	7	
P(X)	0.5	0.3	0.4	✖

Mean, Variance, Standard Deviation, and Expectation

Mean The mean of a random variable with a discrete probability distribution .

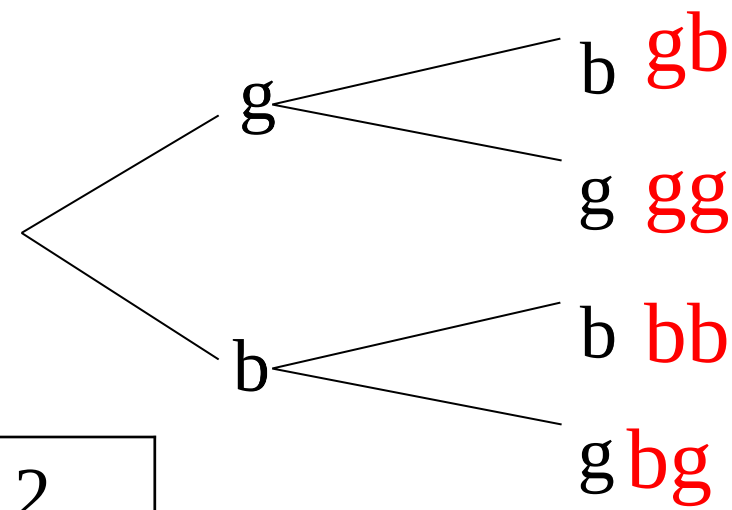
$$\mu = X_1 \,.\, P(X_1) + X_2 \,.\, P(X_2) + X_3 \,.\, P(X_3) + \ldots + X_n \,.\, P(X_n)$$

$$\boxed{\mu = \sum X \,.\, P(X)}$$

Where $X_1, X_2, X_3, \ldots, X_n$ are the outcomes and $P(X_1), P(X_2), P(X_3), \ldots, P(X_n)$ are the corresponding probabilities.

Example 5-6: Children in Family

In a family with two children ,find the mean of the number of children who will be girls.



Solution:

Probability Distribution Table

X	0	1	2
P(X)	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

$$\mu = \sum X \,.\, P(X) = 0 \,.\, \frac{1}{4} + 1 \,.\, \frac{2}{4} + 2 \,.\, \frac{1}{4} = 1$$

Note: This PowerPoint is only a summary and your main source should be the book.

Example 5-7: Tossing Coins

If three coins are tossed ,find the mean of the number of heads that occur.

Solution:

Probability Distribution Table

Number of heads (X)	0	1	2	3
Probability P(X)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$\mu = \sum X \,.\, P(X) = 0 \,.\, \frac{1}{8} + 1 \,.\, \frac{3}{8} + 2 \,.\, \frac{3}{8} + 3 \,.\, \frac{1}{8} = 1.5$$

Example 5-8: No. of Trips of 5 Nights or More

The probability distribution shown represents the number of trips of five nights or more that American adults take per year. (That is, 6% do not take any trips lasting five nights or more, 70% take one trip lasting five nights or more per year, etc.) Find the mean.

Number of trips X	0	1	2	3	4
Probability $P(X)$	0.06	0.70	0.20	0.03	0.01

Solution :

$$\mu = \sum X \cdot P(X) = 0(0.06) + 1(0.70) + 2(0.20) + 3(0.03) + 4(0.01)$$

$$= \boxed{1.2} \text{ trips}$$

Note: This PowerPoint is only a summary and your main source should be the book.

Variance and Standard Deviation

- The formula for the variance of a probability distribution is

Variance: $$\sigma^2 = \sum [X^2 . P(X)] - \mu^2$$

Standard Deviation: $$\sigma = \sqrt{\sigma^2}$$

$$\sigma = \sqrt{\sum [X^2 . P(X)] - \mu^2}$$

Example 5-9: Rolling a Die

Compute the variance and standard deviation for the probability distribution in Example 5–5.

Outcome X	1	2	3	4	5	6
Probability $P(X)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Solution :

$$\sigma^2 = \sum\left[X^2 \cdot P(X)\right] - \mu^2$$

$$\sigma^2 = 1^2 \cdot \tfrac{1}{6} + 2^2 \cdot \tfrac{1}{6} + 3^2 \cdot \tfrac{1}{6} + 4^2 \cdot \tfrac{1}{6}$$

$$+ 5^2 \cdot \tfrac{1}{6} + 6^2 \cdot \tfrac{1}{6} - (3.5)^2$$

$$\sigma^2 = \boxed{2.9} \leftarrow \text{Variance}$$

$$, \quad \sigma = \boxed{1.7} \leftarrow \text{standard deviation}$$

Note: This PowerPoint is only a summary and your main source should be the book.

Example 5-10: Selecting Numbered Balls

A box contains 5 balls .Two are numbered 3 , one is numbered 4 ,and two are numbered 5. The balls are mixed and one is selected at random . After a ball is selected , its number is recorded . Then it is replaced . If the experiment is repeated many times , find the variance and standard deviation of the numbers on the balls.

Solution :

Number on each ball (X)	3	4	5
Probability P(X)	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$

Number on each ball (X)	3	4	5
Probability P(X)	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$
X^2	$3^2=9$	$4^2=16$	$5^2=25$

Step 1 :

$$\mu = \sum X \,.\, P(X) = 3 \,.\, \frac{2}{5} + 4 \,.\, \frac{1}{5} + 5 \,.\, \frac{2}{5} = 4$$

Step 2 :

Variance

$$\sigma^2 = \sum [X^2.P(X)] - \mu^2 = [3^2.\frac{2}{5} + 4^2.\frac{1}{5} + 5^2.\frac{2}{5}] - 4^2 = \frac{4}{5}$$

$$\sigma = \sqrt{\frac{4}{5}} = \sqrt{0.8} = 0.894$$

standard deviation

Note: This PowerPoint is only a summary and your main source should be the book.

Days	0	1	2	3	4
probability	0.06	0.7	0.2	0.03	0.01

A) Mean = 1.23, variance= 0.4171
B) Mean = 0.645, variance = 1.23
C) Mean =1.23, variance= 1.93
D) Mean =1.93, variance = 1.23

X	0	1	2	4	6
p(x)	0.2	0.1	K	0.3	0.2

What the value K would be needed to complete the probability distribution?

A) 0.15
B) 0.2
C) -0.25
D) -0.2