

Discrete Probability Distributions

Learning Objectives

You will learn:

- The properties of a probability distribution
- To compute the expected value and variance of a probability distribution
- To compute probabilities from the binomial, Poisson
- How to use these distributions to solve business problems

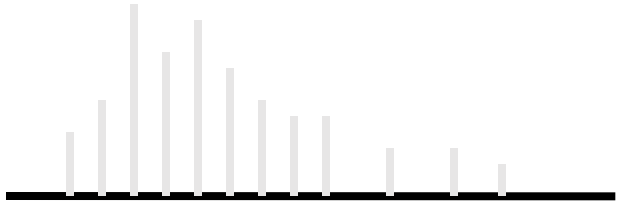
Definitions
Random Variables

- A **random variable** represents a possible numerical value from an uncertain event.
- **Discrete** random variables produce outcomes that come from a counting process (i.e. number of classes you are taking).
- **Continuous** random variables produce outcomes that come from a measurement (i.e. your annual salary, or your weight).

Definitions
Random Variables

Random Variables

- **Discrete Random Variable**
- **Continuous Random Variable**





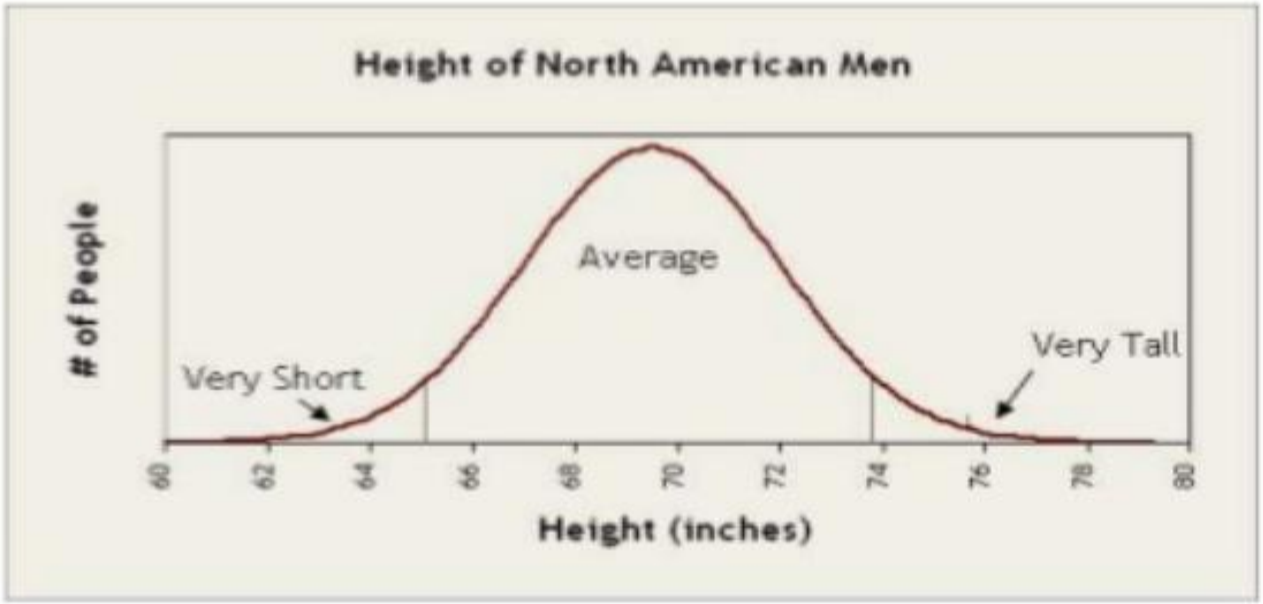
- A continuous random variable is one which takes an **infinite** number of possible values.
- Continuous random variables are usually measurements.

Discrete Random Variables
Examples

- **Examples:**
  - number of children in a family
  - the Friday night attendance at a cinema
  - the number of patients a doctor sees in one day
  - the number of defective light bulbs in a box of ten
  - the number of "heads" flipped in 3 trials

Continuous Random Variables

A **continuous random variable** *X* takes on all values in an interval of numbers. The probability distribution of *X* is described by a **density curve**. The probability of any event is the area under the density curve and above the values of *X* that make up the event.

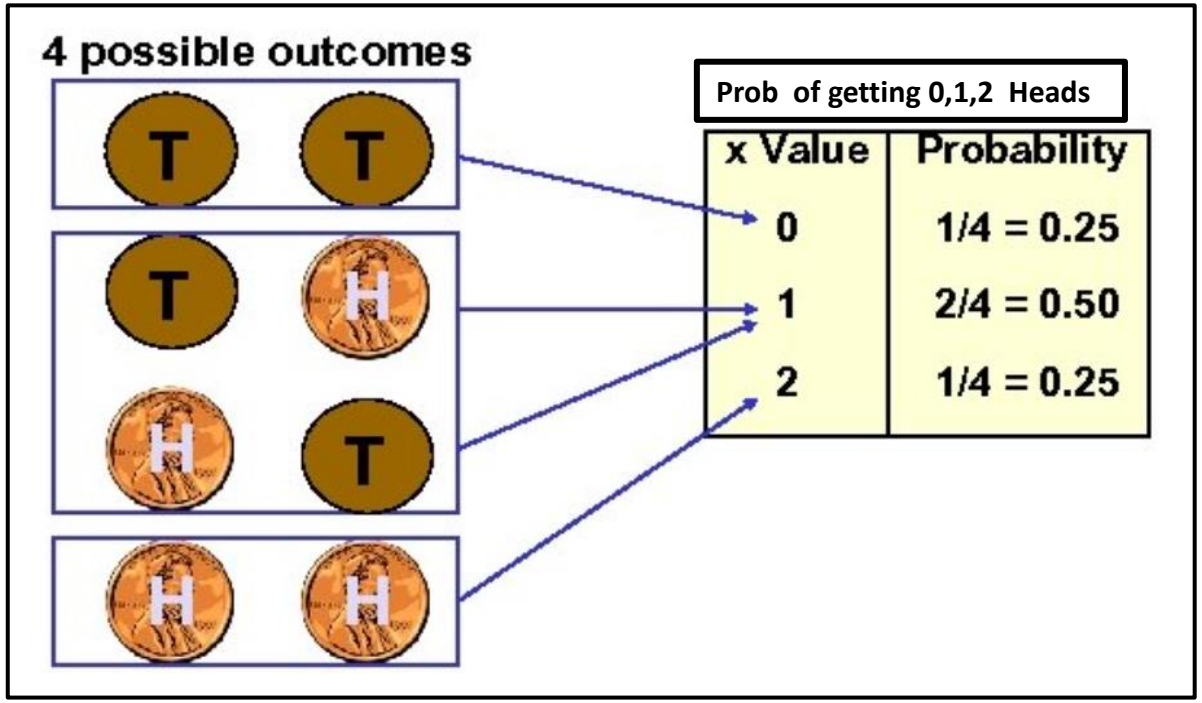


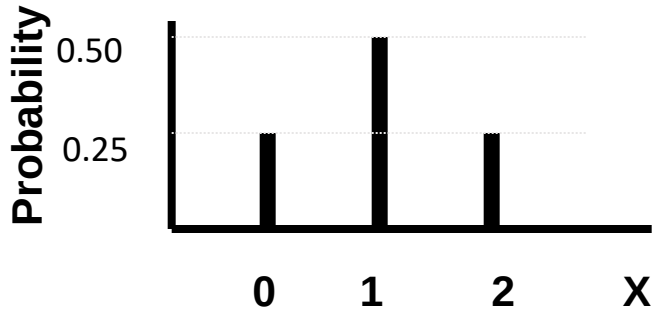
Definitions Probability Distribution

The **probability distribution** of a discrete random variable is a list of probabilities associated with each of its possible values

Experiment: Toss 2 Coins. Let X = # heads.

Probability Distribution





Definitions
Probability Distribution

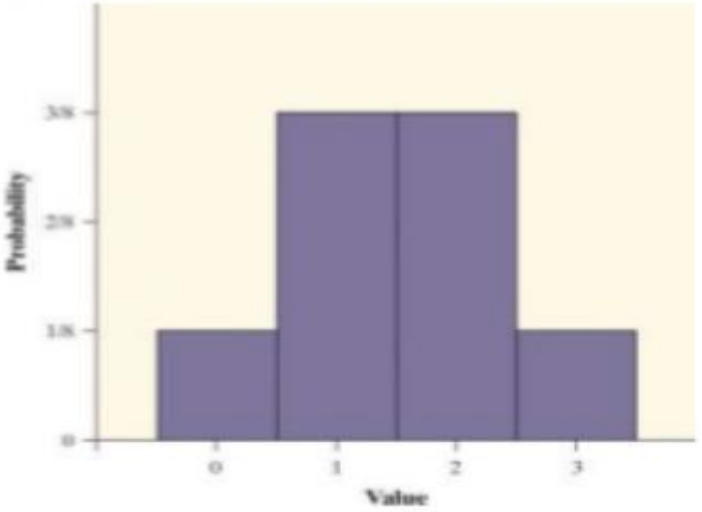
The **probability distribution** of a discrete random variable is a list of probabilities associated with each of its possible values

Consider tossing a fair coin 3 times.
Define X = the number of heads obtained

	Prob
HHH	1/8
HHT	
HTH	
HTT	
THH	
THT	
TTH	
TTT	1/8

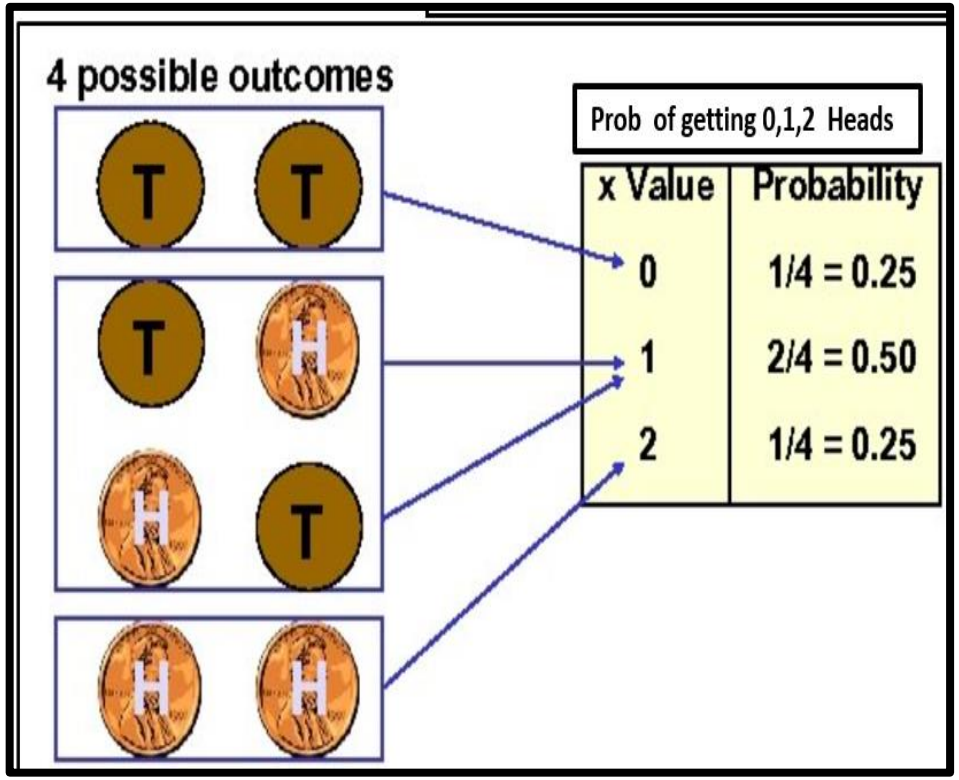
X = 0: TTT
X = 1: HTT THT TTH
X = 2: HHT HTH THH
X = 3: HHH

Value	0	1	2	3
Probability	1/8	3/8	3/8	1/8



Cumulative Probability Distribution

Number of heads: x	Probability: P(X = x)	Cumulative Probability: P(X ≤ x)
0	0.25	
1	0.50	
2	0.25	



P(x≤ 1) = P(x=0) + P(x=1) = 0.75

- A **cumulative probability** refers to the probability that the value of a random variable falls within a specified range.
  - Probability for at most (less than equal to: ≤) one head? → 0.25+0.5=0.75

Discrete Random Variables Expected Value

Meaning of expected value: *The **return you can expect for some kind of action**, like how many questions you might get right if you guess on a multiple choice test.*

- *For example, if you take a 20 question multiple-choice test with A,B,C,D as the answers, and you guess all "A", then you can expect to get 25% right (5 out of 20).*
- *The **math** behind this kind of expected value is: The probability (P) of getting a question right if you guess: 0.25 The number of questions on the test (n)*: 20*
  P x n = .25 x 20 = 5

- Expected Value (or mean) of a discrete distribution (Weighted Average)

Example: Toss 2 coins, X = # of heads

$$\mu = \mathrm{E}(X) = \sum_{i=1}^{N} X_i P(X_i)$$

Compute expected value of X

X Value	Probability
0	1/4 = .25
1	2/4 = .50
2	1/4 = .25

E(X) = (0)(.25) + (1)(.50) + (2)(.25) = 1.0

Discrete Random Variables
Expected Value

- Compute the expected value for the given distribution:

Number of Classes Taken	Probability
2	0.2
3	0.4
4	0.24
5	0.16

$E(X) = 2(.2) + 3(.4) + 4(.24) + 5(.16) = 3.36$

So, the average number of classes taken is 3.36.

Discrete Random Variables Dispersion

- Variance of a discrete random variable

$$\sigma^2 = \sum_{i=1}^{N} [X_i - E(X)]^2 P(X_i)$$

- Standard Deviation of a discrete random variable

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^{N} [X_i - E(X)]^2 P(X_i)}$$

where:

E(X) = Expected value of the discrete random variable X

X_i = the i^{th} outcome of X

$P(X_i)$ = Probability of the i^{th} occurrence of X

Discrete Random Variables Dispersion

- Example: Toss 2 coins, X = # heads, compute standard deviation (recall that E(X) = 1)

Possible number of heads = 0, 1, or 2

X Value	Probability
0	1/4 = .25
1	2/4 = .50
2	1/4 = .25

$E(X) = (0)(.25) + (1)(.50) + (2)(.25) = 1.0$

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^{N} [X_i - E(X)]^2 P(X_i)}$$

$$\sigma = \sqrt{(0-1)^2(.25) + (1-1)^2(.50) + (2-1)^2(.25)} = \sqrt{.50} = .707$$

Using Excel

	A	B	C	D	E
1					
2			=A4*B4		
3	x	P(x)	x* P(x)	(x-mean)^2 * P(x)	
4	0	0.25	0	0.25	=(A4-C$7)^2 *B4
5	1	0.5	0.5	0	
6	2	0.25	0.5	0.25	
7		E(x)= Mean	1		=SUM(C4:C6)
8			Variance	0.5	=SUM(D4:D6)
9			Standard Deviation	0.707106781	=SQRT(D8)

Problem

- The distribution of the number of mortgages approved per week at the local branchoffice of a bank is given below

(1) Calculate the expected value $\mu = E(X) = \sum_{i=1}^{N} X_i P(X_i)$

(2) Variance $\sigma^2 = \sum_{i=1}^{N} [X_i - E(X)]^2 P(X_i)$

(3) Standard Deviation $\sigma = \sqrt{\sigma^2}$

X	P(X)
0	0.1
1	0.1
2	0.2
3	0.3
4	0.15
5	0.1
6	0.05

Using Excel worksheet

DISCRETE RANDOM VARIABLE PROBABILITY DISTRIBUTION

ENTER THE DATA

X	P(X)	X* P(x)	(X- E(X))^2	(X-E(X))^2 * P(X)
0	0.1	0	7.84	0.784
1	0.1	0.1	3.24	0.324
2	0.2	0.4	0.64	0.128
3	0.3	0.9	0.04	0.012
4	0.15	0.6	1.44	0.216
5	0.1	0.5	4.84	0.484
6	0.05	0.3	10.24	0.512

STATISTICS	
EXPECTED VALUE	
VARIANCE	
STANDARD DEVIATION	

Expected value = E(X) **2.80**

Variance = **2.46**

Discrete Probability Distributions:

1.BINOMIAL

2.POISSON

Definitions
Random Variables

- A **random variable** represents a possible numerical value from an uncertain event.
- **Discrete** random variables produce outcomes that come from a counting process (i.e. number of classes you are taking).
- **Continuous** random variables produce outcomes that come from a measurement (i.e. your annual salary, or your weight).

BINOMIAL DISTRIBUTION

Binomial distribution can be thought of as simply the probability of a SUCCESS or

FAILURE outcome in an experiment

BINOMIAL PROBABILITY - EXAMPLES

EXPERIMENT	Event
Repayment by borrowers	EMI paid regularly,not paid regularly
Auditing a bill or invoice	Containing an error or no error
Inspection of an item	Good or Bad
Obseving a machine	Machine working or not working
Tossing a coin	Head or Tail
Birth of a Baby	Boy or Girl

Applications of the Binomial Distribution

- A manufacturing plant labels items as either defective or acceptable
- A firm bidding for contracts will either get a contract or not
- A marketing research firm receives survey responses of “yes I will buy” or “no I will not”
- New job applicants either accept the offer or reject it
- Your team either wins or loses the football game at the company picnic
- Control Chart – For Defective items

The Binomial Distribution: Properties

- **A fixed number of observations, n**
  - ex. 15 tosses of a coin; ten light bulbs taken from a warehouse
- **Two mutually exclusive and collectively exhaustive categories**
  - ex. head or tail in each toss of a coin; defective or not defective light bulb; having a boy or girl
  - Generally called “success” and “failure”
  - Probability of success is p, probability of failure is 1 – p
- **Constant probability for each observation**
  - ex. Probability of getting a tail is the same each time we toss the coin

The Binomial Distribution: Properties

- **Observations are independent**
  - The outcome of one observation does not affect the outcome of the other
- **Two sampling methods**
  - Infinite population without replacement
  - Finite population with replacement

Binomial Distribution

- The binomial distribution is a discrete probability distribution which describes the binomial experiment.
- An experiment is characterized as binomial if the following four conditions are satisfied
    1. There are fixed number of trials. Denote it by n.
    2. Each trial is independent. For example, while tossing a coin, 3rd toss is independent of any of the previous two tosses.

    4. The probability of success denoted as p remains constant from trial to trial. The probability of failure is denoted by q which equal to 1-p.

- An experiment is said to be Bernoulli process if 2, 3 and 4 conditions are satisfied.

- If Bernoulli process also satisfies the condition of fixed numbers of trails, then it is called binomial experiment.

Binomial Experiment

SUMMARY

- The experiment involves **repeated trials**.
- Each trial has **only two possible outcomes** - a success or a failure (i.e., head/tail, goal/no goal).
- The probability that a particular outcome will occur on any given trial is **constant**.
  - 0.5 every trial
- All of the trials in the experiment are **independent**.
  - The outcome on one trial does not affect the outcome on other trials.

- # What is Factorial?
  - In many problems product of consecutive integers are needed. For instance the product of 6, 5, 4, 3, 2, 1 is called 6 factorial and denoted by *6!*.
  - In general, *n!* is called n-factorial which denotes the product of first *n* natural numbers.
  - Few Examples
    - 0! = 1
    - 1! = 1
    - 2! = 2x1 = 2
    - 3! = 3x2x1 =6
    - 4! = 4x3x2x1= 24 and so on.

Combinations

- It refers to the number of arrangements which can be made from a group of items irrespective of the order.

- The combination of 'n' different objects taken r at a time is given by the following expression:

$$^{n}C_{r} = \frac{n!}{r!\,(n-r)!}$$

(a)

$$^{12}C_{5} = \frac{n!}{r!\,(n-r)!} = \frac{12!}{5!\,7!} = \frac{12\times11\times10\times9\times8\times7!}{5\times4\times3\times2\times1\times7!} = \frac{95040}{120} = 792$$

(b)

$$^{4}C_{2} = \frac{n!}{r!\,(n-r)!} = \frac{4!}{2!\,2!} = \frac{4\times3\times2!}{2\times1\times2!} = \frac{12}{2} = 6$$

Counting Techniques - Rule of Combinations

- The number of combinations of selecting X objects out of n objects is:

$$ {}_{n}C_{X} = \binom{n}{X} = \frac{n!}{X!(n-X)!} $$

where:

$n! = n(n-1)(n-2)\ldots(2)(1)$

$X! = X(X-1)(X-2)\ldots(2)(1)$

$0! = 1$ (by definition)

Counting Techniques - Rule of Combinations

- How many possible 3 scoop combinations could you create at an ice cream parlor if you have 31 flavors to select from?
- The total choices is n = 31, and we select X = 3.

$$_{31}C_3 = \binom{31}{3} = \frac{31!}{3!(31-3)!} = \frac{31!}{3!28!} = \frac{31 \bullet 30 \bullet 29 \bullet 28!}{3 \bullet 2 \bullet 1 \bullet 28!} = 31 \bullet 5 \bullet 29 = 4495$$

The Binomial Distribution Formula

$$P(X) = \frac{n!}{X!(n-X)!} p^X (1-p)^{n-X}$$

Where x can take the value 0, 1,2 …. n

P(X) = probability of **X** successes in **n** trials,
with probability of success **p** on each trial

X = number of 'successes' in sample,
(X = 0, 1, 2, ..., n)

n = sample size (number of trials or observations)

p = probability of "success"

Using Excel

=BINOM.DIST(J7,4,0.5,FALSE)

x	P(x)
0	0.0625
1	0.25
2	0.375
3	0.25
4	0.0625

Example: Flip a coin four times, let x = # heads:

n = 4

p = 0.5

1 - p = (1 - .5) = .5

X = 0, 1, 2, 3, 4

$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ Given that x=18, n = 20, p = 0.8

$P(X = 18) = \binom{20}{18} (0.8)^{18} (0.2)^2 = 0.1369$

Using Excel

= BINOM.DIST(18,20,0.8,FALSE)

P(x=18) = 0.136909

The Binomial Distribution Characteristics

- Mean

$$\mu = E(x) = np$$

- Variance and Standard Deviation

$$\sigma^2 = np(1-p) \qquad \sigma = \sqrt{np(1-p)}$$

Where n = sample size

p = probability of success

(1 – p) = probability of failure

Binomial Distribution Notation

P(x=2) =Probability of getting exactly 2 defectives

P(x<2) =Probability of getting less than 2 defectives

= P(x=0) + P(x =1)

P(x>=2) =Probability of getting atleast 2 defectives

= 1-(P(x=0) + P(x =1))

P(x> 2) = Prob of getting more than 2 defectives

= 1- (P (x=0) + P (x=1) + P (x=2))

P (x<=2) = Prob of getting at the most 2 defectives

= P(x=0) + P(x =1) + P(x=2)

Using Excel – Find the Probability

=BINOM.DIST(I9,10,0.6,1) - CUM PROB

Find the prob for x=0 1,…10

=BINOM.DIST(C5,10,0.6,0)

	Data Given : n- 10 p= 0.6 x=2		
Notation	Out of 10 cars selected from the assy line What is the Prob of getting	Excel Command	Value
P(X=2)	Probability of getting exactly 2 defectives car	BINOM.DIST(2,10,0.6,FALSE)	0.0106
P(x<2)	Probability of getting less than 2 defectives car= P(x=0) +P(x-1)	BINOM.DIST(1,10,0.6,TRUE)	0.002
P(x≥2)	Probability of getting atleast 2 defectives car = 1-(P(x=0) + P(x=1))	1-BINOMDIST(1,10,0.6,TRUE)	0.998
P(X>2)	Prob of getting more than 2 defective cars= 1-(P(x=0)+P(x-1)+P(x=2))	1-BINOM.DIST(2,10,0.6,TRUE)	0.988
P(X≤2)	Prob of getting at the most 2 defectives car = P(x=0)+P(x=1)+P(x=2)	BINOM.DIST(2,10,0.6,True)	0.012

x	P(X)	Cum ≤
0	0.0001	0.000
1	0.0016	0.002
2	0.0106	0.012
3	0.0425	0.055
4	0.1115	0.166
5	0.2007	0.367
6	0.2508	0.618
7	0.215	0.833
8	0.1209	0.954
9	0.0403	0.994
10	0.006	1.000

The Binomial Distribution
Example-1

What is the probability of one success in five observations if the probability of success is .1?

$$X = 1,\ n = 5,\ \text{and}\ p = .1$$

$$P(X=1) = \frac{n!}{X!(n-X)!}\,p^{X}(1-p)^{n-X}$$

$$= \frac{5!}{1!(5-1)!}(.1)^{1}(1-.1)^{5-1}$$

$$= (5)(.1)(.9)^{4}$$

$$= .32805$$

Probability of 1 success in 5 observations
P(x= 1) = 0.32805

Using Excel : x=1, n=5 & p=0.1

= BINOM.DIST(1,5, 0.1,FALSE)

P(x=1) = 0.32805

Probability of getting less than 1 success in 5 observations

P(x<1) = P(x =0)

= BINOM.DIST(0,5, 0.1,TRUE)

P(x<1) = P(x=0) = 0. 59049

The Binomial Distribution
Example-2

Suppose the probability of purchasing a defective computer is 0.02. What is the probability of purchasing 2 defective computers is a lot of 10?

Data given : **X= 2 , n= 10 and p = 0.02**

$$P(X=2)=\frac{n!}{X!(n-X)!}p^{X}(1-p)^{n-X}$$

$$=\frac{10!}{2!(10-2)!}(.02)^{2}(1-.02)^{10-2}$$

$$=(45)(.0004)(.8508)$$

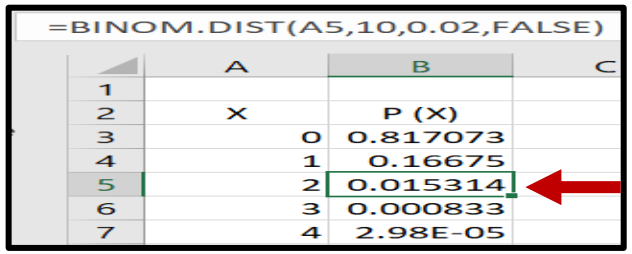
$$=.01531$$

P(x=2) = 0.015314

Using Excel : x=2, n=10 & p=0.02

= BINOM.DIST(2,10, 0.02,FALSE)

Probability of purchasing 2 defective computer P(x =2) = 0.015314



Example-3

- Assume that on an average one telephone no out of fifteen is busy. What is the probability that if six randomly selected telephone numbers are called
- (1) Not more than three will be busy $P(x \le 3)$ $P(x \le 3) = P(x=0)+P(x=1)+P(x=2)+P(x=3)$ **0.9997**
- (2) At least 3 of them will be busy $P(x \ge 3)$ $P(x \ge 3) = 1-((P(x=0)+P(x=1)+P(x=2))$ **0.0049**

1- (0.6639+0.2815+0.0497) = 1-0.9951 = 0.0049

Data Given :

p = probability that a telephone no is busy = 1/15 = 0.066

n = no of selected telephone nos = 6

Data Given : p= 0.066 , n = 6 , 1-p = 1-0.066= 0.934					
X	**nCx**	**value**	**p^x**	**(1-p)^ (n-x)**	**P(X)**
0	6C0	1	1	0.66386727	0.6639
1	6C1	6	0.066	0.710778661	0.2815
2	6C2	15	0.0044	0.761004991	0.0497
3	6C3	20	0.0003	0.814780504	0.0047
				P(X <=3)	0.9997

Using Excel :

Data Given : x = 3 , n=6, p= 0.066, To find $P(x \le 3)$

$P(x \le 3) = P(x=0)+P(x=1)+P(x=2)+P(x=3)$

=BINOM.DIST(3,6,0.066,TRUE) P(x≤3)= 0.999745

Using Excel :

Data Given : x = 3 , n=6 ,p= 0.066, To find $P(x \ge 3)$

$P(x \ge 3) = 1-((P(x=0)+P(x=1)+P(x=2))$

= 1-BINOM.DIST(2,6,0.066,TRUE) = 0.00494

Example -4

- Suppose that the half of the population of town are consumers of rice. One hundred investigators are appointed to find out its truth . Each investigator interviewed 10 individuals. How many investigators do you expect to report that 3 or less of the people interviewed are consumers of rice

Data Given :
p = Probability of a person being consumer of rice = 0.5
n = no of individuals = 10 N= 100

To find :
P (x ≤3) & No of investigators to report that 3 or less people are consumers of rice = N * P(x ≤ 3)

P (x ≤3) = P(x=0) +P(x=1) +)(x=2) + P(x=3)

$$P(X) = \frac{n!}{X!(n-X)!} p^{x}(1-p)^{n-x}$$

Using Excel :
x=3 , n= 10 , p= 0,5

To find : P (x ≤3)

= BINOM.DIST(3,10,0.5,TRUE)

P(x≤ 3) = 0.171875

Data : n=10 p=0.5					
P(x) = nCx p^x q^n-x					
x	nCx	nCx	p^x	(1-p)^ (10 - c6)	P(X)
0	10Co	1	1	0.000976563	0.000976563
1	10C1	10	0.5	0.001953125	0.009765625
2	10C2	45	0.25	0.00390625	0.043945313
3	10C3	120	0.125	0.0078125	0.1171875
				P(x <= 3)	0.171875

No of investigators to report that 3 or less people are consumers of rice is given by
= N * P(x ≤ 3) = 0.17 * 100 = 17(approx.)

Example -5

- The incidence of occupational disease in an industry is such that the workers have have a 20 % chance of suffering from it .What is the probability that out of six workers 4 or more will contract disease

Data Given :
p = probability of a worker who is suffering from the disease = 0.2
n = no of workers = 6

To find :
What is the probability out of 6 workers 4 or more will contract the disease P (x ≥ 4)

$P(x \geq 4) = 1 - (P(x=0)+P(x=1)+P(x=2)+P(X=3))$ = 1- 0.983 = 0.017

Data Given : p= 0.2 , n = 6 , 1-p = 1-0.2= 0.8

X	nCx	value	p^x	(1-p)^ (n-x)	P(X)
0	6C0	1	1	0.262144	0.2621
1	6C1	6	0.2	0.32768	0.3932
2	6C2	15	0.04	0.4096	0.2458
3	6C3	20	0.008	0.512	0.0819
				P(X <=3)	0.983

USING EXCEL

= 1-BINOM.DIST(3,6,0.2,TRUE)

P(x ≥4)= 0.01696

The Binomial Distribution
Using Binomial Tables

n = 10									
x	…	p=.20	p=.25	p=.30	**p=.35**	p=.40	p=.45	p=.50	
0	…	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010	10
1	…	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098	9
2	…	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439	8
3 ←	…	0.2013	0.2503	0.2668	**0.2522**	0.2150	0.1665	0.1172	7
4	…	0.0881	0.1460	0.2001	0.2377	0.2508	0.2384	0.2051	6
5	…	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461	5
6	…	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051	4
7	…	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172	3
8	…	0.0001	**0.0004**	0.0014	0.0043	0.0106	0.0229	0.0439	**2**
9	…	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098	1
10	…	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0
	…	p=.80	**p=.75**	p=.70	p=.65	p=.60	p=.55	p=.50	**x**

Examples:

n = 10, p = .35, x = 3

P(x = 3|n =10, p = .35) = .2522

The Binomial Distributio- MEAN & VARIANCE Example

- If the Probability of defective bolt is 1/10 , find the tracing for the Binomial distribution of defctive bolts in a sum of 400
- (a) The Mean (b) The variance

Data Given : n= 400 , p = 1/10 = 0.10 , q = 1-p = 1-0.1 = 0.9

Mean = np = 400 * 0.1 = 40

Variance = npq = 400*0.1*0.9 = 36

Binomial Distribution

- The screws produced by a certain machine were found out by examining sample of 12.The following table shows the distribution of 128 samples according to the ordinal number of defective items they carried.
- **1. Fit a Binomial distribution & find the Expected frequencies if th chance of m/c being defective is 0.5**
- **2.Find the mean & S.d.of the fitted distribution**

No of defectives in a sample of 12	0	1	2	3	4	5	6	7	Total
No of samples	7	6	19	35	30	23	7	1	128

Data given

n= 7 , p =0.5 N=128

FIND THE PROBABILITY WHEN,X=0,1,2,3,4,5,6 &7

x	P(x)	Expected frequency = N*P(x)	E(X) = x*P(x)	(X-E(x))^2	((X-E(X)^2)*P(X)
0	0.0078125	1	0	12.25	0.095703125
1	0.0546875	7	0.055	6.25	0.341796875
2	0.1640625	21	0.328	2.25	0.369140625
3	0.2734375	35	0.82	0.25	0.068359375
4	0.2734375	35	1.094	0.25	0.068359375
5	0.1640625	21	0.82	2.25	0.369140625
6	0.0546875	7	0.328	6.25	0.341796875
7	0.0078125	1	0.055	12.25	0.095703125
		MEAN	3.5	Variance	1.75

S.D	1.32287566

FIND THE PROBABILITY , USING EXCEL

=BINOM.DIST(D5,7,0.5,FALSE)

Data Given : p= 0.5 , n = 7 , 1-p = 1-0.5= 0.5

X	nCx	value	p^x	(1-p)^ (n-x)	P(X)
0	7C0	1	1	0.0078125	0.0078
1	7C1	7	0.5	0.015625	0.0547
2	7C2	21	0.25	0.03125	0.1641
3	7C3	35	0.125	0.0625	0.2734
4	7C4	35	0.0625	0.125	0.2734
5	7C5	21	0.0313	0.25	0.1641
6	7C6	7	0.0156	0.5	0.0547
7	7C7	1	0.0078	1	0.0078

Practice Problem 1– Binomial Distribution

- The mean Binomial distribution is 40 and standard deviation is 6

  Calculate n ,p, and q

Data given :

Mean = np = 40 and standard deviation = 6 (i.e) $\sigma = \sqrt{np(1-p)}$

Variance = 36 = npq

$q = \frac{36}{np} = \frac{36}{40} = 0.9,$

p = (1-q) =(1-0.9) = 0.1

` Mean = np=40 $n = \frac{40}{p} = \frac{40}{0.1}$ = 400

Solution :
n = 400 p= 0.1 & q = 0.9

Practice Problem 2– Binomial Distribution

The overall percentage of failures in a certain examination is 30 . What is the probability that out of a group of 6 candidates at least 4 passed the examination

Data Given :

Probability of a student failing in the examination is 0.30 , n = No of candidates = 6

We know that P (passing) + P(Failing) = 1

Therefore P(passing) = 1-P(failing) = 1-0.3 =0.7

To find the probability that at least 4 students passing the examination = P ($x \geq 4$) here p= 0.7 & q= 0.3

$P(x \geq 4) = P(x=4) + P(x=5) + P(x=6)$ **or** $1 - (P(x=0) + P(x=1) + P(x=2) + P(x=3))$

Data Given : p= 0.7 , n = 6 , 1-p = 1-0.7= 0.3					
X	**nCx**	**value**	**p^x**	**(1-p)^ (n-x)**	**P(X)**
4	6C4	15	0.2401	0.09	0.3241
5	6C5	6	0.1681	0.3	0.3025
6	6C6	1	0.1176	1	0.1176
				P(X>=4)	0.744

Probability that out of a group of 6 candidates at least 4 passed the examination = P($X \geq 4$) is 0.744

Practice Problem 3– Binomial Distribution

Suppose A and B are two equally strong table tennis players ,which of the following two events is more probable

(i) A beats B exactly in 3 games out of 4 or (ii) A beats B exactly in 5 games out of 8

(i) A beats B in exactly 3 games out of 4

Here n= 4 , x= 3 , p =0.5

Data Given : p= 0.5 , n = 4 , 1-p = 1-0.5= 0.5					
X	nCx	value	p^x	(1-p)^ (n-x)	P(X)
3	4C3	4	0.125	0.5	0.25

P(X=3) = 0.25

(ii) A beats B in exactly 5 games out of 8

Here n= 8 , x=5 , p =0.5

Data Given : p= 0.5 , n = 4 , 1-p = 1-0.5= 0.5					
X	nCx	value	p^x	(1-p)^ (n-x)	P(X)
5	8C5	56	0.0313	0.125	0.2188

P(X=5) = 0.218

The First Event is more probable

USING EXCEL

(i) A beats B in exactly 3 games out of 4

Here n= 4 , x= 3 , p =0.5

P(X=3) =BINOM.DIST(3,4,0.5,FALSE) **P(X=3) = 0.25**

(ii) A beats B in exactly 5 games out of 8

Here n= 8 , x=5 , p =0.5

P(X=5) =BINOM.DIST(5,8,0.5,FALSE) **P(X=5) = 0.218**

QUIZ

1. In a Binomial Distribution, if ‘n’ is the number of trials and ‘p’ is the probability of success, then the mean value is given by _____

a) np
b) n
c) p
d) np(1-p)

$\mu = np.$

2. In a Binomial Distribution, if p, q and n are probability of success, failure and number of trials respectively then variance is given by _____

a) np
b) npq
c) np^2q
d) npq^2

Variance = npq.

QUIZ

3. If ‘X’ is a random variable, taking values ‘x’, probability of success and failure being ‘p’ and ‘q’ respectively and ‘n’ trials being conducted, then what is the probability that ‘X’ takes values ‘x’? Use Binomial Distribution

a) $P(X = x) = {}^{n}C_{x}\, p^{x}\, q^{x}$
b) $P(X = x) = {}^{n}C_{x}\, p^{x}\, q^{(n-x)}$
c) $P(X = x) = {}^{x}C_{n}\, q^{x}\, p^{(n-x)}$
d) $P(x = x) = {}^{x}C_{n}\, p^{n}\, q^{x}$

$P(X = x) = {}^{n}C_{x}\, p^{x}\, q^{(n-x)}$.

4 In a Binomial Distribution, the mean and variance are equal.

a) True
b) False

FALSE

5 Binomial Distribution is a _____

a) Continuous distribution
b) Discrete distribution
c) Irregular distribution
d) Not a Probability distribution

It is applied to a discrete Random variable, hence it is a discrete distribution.

Poisson Distribution

It is used to find the probability value of discrete random variable. This is a distribution of rare events (ie) the events whose probability of occurrence is very small but the number of trials which could lead to the occurrence of the event are very large.

- The examples of poisson distribution are :
  1. The no. of calls received per day
  2. The number of customers arriving at super market/day
  3. No. of defects per item
  4. No. of goals scored per match in a soccer tournament

The Poisson Distribution Definitions

- An **area of opportunity** is a continuous unit or interval of time, volume, or such area in which more than one occurrence of an event can occur.
  - ex. The number of scratches in a car's paint
  - ex. The number of mosquito bites on a person
  - ex. The number of computer crashes in a day

POISSON DISTRIBUTION - EXAMPLES

- The number of cars arriving at a service station in 1 hour
- The number of flaws in a bolt of cloth
- The number of accidents in 1 day on a particular stretch of highway

The Poisson Distribution Properties

Apply the Poisson Distribution when:

- You wish to count the number of times an event occurs in a given area of opportunity
- The probability that an event occurs in one area of opportunity is the same for all areas of opportunity
- The number of events that occur in one area of opportunity is independent of the number of events that occur in the other areas of opportunity
- The probability that two or more events occur in an area of opportunity approaches zero as the area of opportunity becomes smaller
- The average number of events per unit is λ (lambda)

The Poisson Distribution Formula

$$P(X) = \frac{e^{-\lambda}\lambda^{x}}{X!}$$

Parameters of the Poisson Distribution

- The mean is μ. **or Lambda (λ)**

where:

X = the probability of X events in an area of opportunity

λ = expected number of events

e = mathematical constant approximated by 2.71828…

The Poisson Distribution
Example -1

- Suppose that, on average, 5 cars enter a parking lot per minute. What is the probability that in a given minute, 7 cars will enter?

- **Poisson Probabilities**

Data						1
Average/Expected No of Successes		5				
Poisson Probabilties table						
	X	P(X)	P(<=X)	P(<X)	P(>X)	P(>=X)
	0	0.006738	0.006737947	0	0.993262	1
	1	0.03369	0.040427682	0.006738	0.959572	0.993262
	2	0.084224	0.124652019	0.040428	0.875348	0.959572
	3	0.140374	0.265025915	0.124652	0.734974	0.875348
	4	0.175467	0.440493285	0.265026	0.559507	0.734974
	5	0.175467	0.615960655	0.440493	0.384039	0.559507
	6	0.146223	0.762183463	0.615961	0.237817	0.384039
	7	0.10444	0.866628326	0.762183	0.133372	0.237817

P(x=7)	=POISSON.DIST(7,5,FALSE)
P(x ≤7)	=POISSON.DIST(7,5,True)
P(x <7)	=POISSON.DIST(6,5,True)

So, X = 7 and λ = 5

$$P(7) = \frac{e^{-\lambda}\lambda^{x}}{X!} = \frac{e^{-5}5^{7}}{7!} = 0.104$$

USING EXCEL

P(x=7)	=POISSON.DIST(7,5,FALSE)	= 0.10444486

- So, there is a 10.4% chance 7 cars will enter the parking in a given minute.

The Poisson Distribution Example -2

- The average no of customers , who appears at a counter of a certain bank per minute is two. Find the probability that during a given minute

  (1) No customer appears (2) 3 or more customers appear

- Data given λ =2

$$P(X) = \frac{e^{-\lambda}\lambda^{x}}{X!}$$

Find (1) P (x=0)

(2) P (x >= 3) = 1-(P(x=0) +P(x=1)+P(x=2))

x	Lambda	e^{-2}	2^x	x!	P(X)
0	2	0.1353	1	1	0,1353
1			2	1	0.2706
2			4	2	0.2706
				P(x≤ 2)	0.6765

(1) No customer appears P(x=0) = 0.1353

(2) 3 or more customers appear =1- P(x≤ 2)
= 1- 0.6765
= 0.3235

USING EXCEL

To find P(x=0)

=-POISSON.DIST(0,2,FALSE)

P(x=0)=0.1353

- Prob that no customer appears = 0.1353.

Find P(x ≥ 3)

P(x≥ 3) = 1-(P(x=0) +P(x=1)+P(x=2))

P(x≥ 3) =1-POISSON.DIST(2,2,TRUE)

- Prob that 3 or more customer appears
- P(x >= 3) = 0.3233.

The Poisson Distribution
Using Poisson Tables

Example: Find $P(X = 2)$ if $\lambda = .50$

$$P(X=2) = \frac{e^{-\lambda}\lambda^{X}}{X!} = \frac{e^{-0.50}(0.50)^{2}}{2!} = .0758$$

Using Excel

=POISSON.DIST(2,0.5,FALSE)

P(X=2) = 0.07581633

Using Poisson Tables

X	λ								
	0.10	0.20	0.30	0.40	**0.50**	0.60	0.70	0.80	0.90
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659
2	0.0045	0.0164	0.0333	0.0536	**0.0758**	0.0988	0.1217	0.1438	0.1647
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494
4	0.0000	0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111
5	0.0000	0.0000	0.0000	0.0001	0.0002	0.0004	0.0007	0.0012	0.0020
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

EXAMPLE -3

- On the average one in 400 items is defective. If the items are packed in boxes of 100, what is the probability that any given box of items will contain
- (1) No defective
- (2) Less than 2 defectives
- (3) one or more defectives
- (4) More than 3 defetives

P(x=0) =POISSON.DIST(**0,0.25**,FALSE) P(x=0) =0.77880078

P(x< 2) P(x =0)+ P (x= 1) =POISSON.DIST(**1,0.25**,True) P(x< 2)=0.97350098

P(x ≥ 1) 1- P(x =0) = 1-POISSON.DIST(**0,0.25**,True) P(x ≥ 1) = 0.22119922

P(x > 3) 1- (P(x =0) + P(x=1) +P(x=2) + P (x=3) = 1-POISSON.DIST(**3,0.25**,True) P(x > 3) = 0.000133

Data Given :

n = no of items packed in a box = 100

p – Prob of a defective item = 1/400

If n is large & probability is small Binomial can be approximated to Poisson

m = np = 100 * 1/400 = 0.25

Poisson Probabilities

Data		1
Average/Expected No of Successes	0.25	

Poisson Probabilties table

	X	P(X)	P(<=X)	P(<X)	P(>X)	P(>=X)
	0	0.778801	0.778800783	0	0.221199	1
	1	0.1947	0.973500979	0.778801	0.026499	0.221199
	2	0.024338	0.997838503	0.973501	0.002161	0.026499
	3	0.002028	0.99986663	0.997839	0.000133	0.002161

Condition for converting Binomial into Poisson

- The Poisson distribution is actually a limiting case of a Binomial distribution when the number of trials, n, gets very large and p, the probability of success, is small.
- As a rule of thumb, if $n \geq 100$ and $np \leq 10$, the Poisson distribution **(taking $\lambda = np$)** can provide a very good approximation to the binomial distribution.

Example

- A factory produces blades in packets of 10. The Probability of a blade to be defectives is 0.2 %. Find the number of packets having 2 defectives in a consignment of 10,000 packets.

Data Given

P : probability of getting a defective blade =0.002 N= 10,000 Packets
n : no of blades = 10 Mean = n*p =10 X 0.002 =0.02

If np ≤ 10 , Binomial can be approximated to Poisson

Find P (x=2) =Probability of getting 2 defective blades
Find number of packets having 2 defective blades in 10,000 packets = N x P(X = 2)

P(x=2) = =POISSON.DIST(2,0.02,FALSE) = 0.000196

Number of packets having 2 defective blades = N x P (x = 2) = 10,000 x 0.000196 =1.96 say 2 packets

Problem

- The following table gives the no of days in a 50 day period during which automobile accidents occurred in a city. Fit a Poisson distribution to the data

No of accidents	0	1	2	3	4
No of days	21	18	7	3	1

Find Mean

X	No of accidents	0	1	2	3	4	Total
f	No of days	21	18	7	3	1	50
	fx	0	18	14	9	4	45

Mean = m = 45/50 = 0.9

Find the probability of getting 0.1.2.3 & 4 accidents/day
Find the expected frequency = 50 x P(x)

	X	No of accidents	0	1	2	3	4
		P(x)	0.40657	0.36591	0.16466	0.0494	0.01111
Expected Frequency N*P(X)			20.33	18.30	8.23	2.47	0.56

P(X= 0)=POISSON.DIST(K17,0.9,FALSE)

x	O	E
0	21	20.3
1	18	18.3
2	7	8.23
3	3	2.47
4	1	0.56

O- Observed Frequency

E- Expected Frequency

QUIZ

1. In a Poisson Distribution, if ‘n’ is the number of trials and ‘p’ is the probability of success, then the mean value is given by?

a) $m = np$
b) $m = (np)^2$
c) $m = np(1-p)$
d) $m = p$

$\mu = m = np.$

2. . If ‘m’ is the mean of a Poisson Distribution, then variance is given by _____

a) m^2
b) $m^{1/2}$
c) m
d) $m/2$

Variance = Mean

QUIZ

3.In a Poisson Distribution, the mean and variance are equal.

a) True
b) False

True

4 Poisson distribution is applied for _____

a) Continuous Random Variable
b) Discrete Random Variable
c) Irregular Random Variable
d) Uncertain Random Variable

Discrete random variable

5.If ‘m’ is the mean of Poisson Distribution, the P(0) is given by _____

a) e^{-m}
b) e^{m}
c) e
d) m^{-e}

e^{-m}

Binomial or Poisson?

Binomial distribution describes the distribution of binary data from a finite sample. Thus, it gives the **probability of getting r events out of n trials**.
Poisson distribution describes the distribution of binary data from an infinite sample. Thus, it gives the probability of getting r events in a population.

If a *mean* or *average* probability of an event happening per unit time/per page/per mile cycled etc., is given, and you are asked to calculate a probability of *n* events happening in a given time/number of pages/number of miles cycled, then the Poisson Distribution is used.

- If, on the other hand, an *exact probability* of an event happening is given, or implied, in the question, and you are asked to calculate the probability of this event happening *k* times out of *n,* then the Binomial Distribution must be used.

Binomial or Poisson - Quiz

1	A typist makes on average 2 mistakes per page. What is the probability of a particular page having no errors on it?	POISSON	m = 2	P(X=0)
2	A computer crashes once every 2 days on average. What is the probability of there being 2 crashes in one week?	POISSON	m = 3	P(X=2)
3	Components are packed in boxes of 20. The probability of a component being defective is 0.1. What is the probability of a box containing 2 defective components?	BINOMIAL	P=0, 1, n=20	P(X=3)
4	ICs are packaged in boxes of 10. The probability of an IC being faulty is 2%. What is the probability of a box containing 2 faulty ics?	BINOMIAL	P=0.02, n=10	P(X=2)
5	The mean number of faults in a new house is 8. What is the probability of buying a new house with exactly 1 fault?	POISSON	m = 8	P(X=1)
6	A box contains a large number of washers; there are twice as many steel washers as brass ones. Four washers are selected at random from the box. What is the probability that 0, 1, 2, 3, 4 are brass?	BINOMIAL	P=0.33, n=4	P(X=0,1,2,3,4)

Chapter Summary

We have

- Addressed the probability of a discrete random variable
- Discussed the Binomial distribution
- Reviewed the Poisson distribution