IIT ROORKEE

NPTEL ONLINE CERTIFICATION COURSE

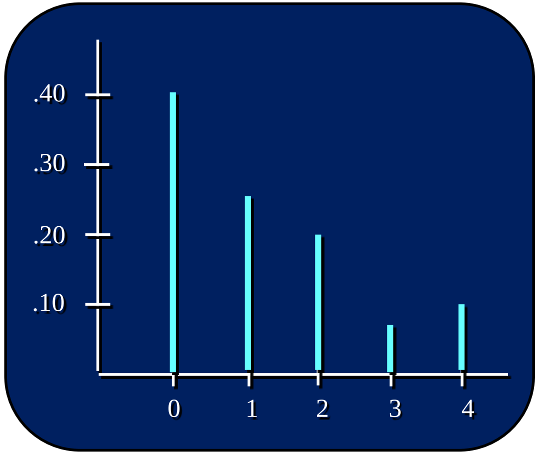
Business Statistics

Introduction

Dr M K BARUA
DEPARTMENT OF MANAGEMENT STUDIES

Discrete Probability Distributions

- Random Variables
- Discrete Probability Distributions
- Expected Value and Variance
- Binomial Distribution
- Poisson Distribution
- Hypergeometric Distribution



Random Variables

- A random variable is a numerical description of the outcome of an experiment.
- A discrete random variable may assume either a **finite number** of values or an **infinite sequence** of values.
- A continuous random variable may assume any numerical value in an interval or collection of intervals.

Example: JSL Appliances

- Discrete random variable with a finite number of values

Let x = number of TVs sold at the store in one day, where x can take on 5 values (0, 1, 2, 3, 4)

Example: JSL Appliances

- Discrete random variable with an <u>infinite</u> sequence of values

Let x = number of customers arriving in one day,
where x can take on the values 0, 1, 2, . . .

We can count the customers arriving, but there is **<u>no finite upper limit</u>** on the number that might arrive.

Random Variables

Question	Random Variable x	Type
Family size	x = Number of dependents reported on tax return	
Distance from home to store	x = Distance in miles from home to the store site	
Own dog or cat	x = 1 if own no pet; = 2 if own dog(s) only; = 3 if own cat(s) only; = 4 if own dog(s) and cat(s)	

Random Variables

Question	Random Variable x	Type
Family size	x = Number of dependents reported on tax return	Discrete
Distance from home to store	x = Distance in miles from home to the store site	Continuous
Own dog or cat	x = 1 if own no pet; = 2 if own dog(s) only; = 3 if own cat(s) only; = 4 if own dog(s) and cat(s)	Discrete

Discrete Probability Distributions

- The probability distribution for a random variable describes how **probabilities are distributed** over the values of the **random** variable.

- We can describe a discrete probability distribution with **a table, graph, or equation**.

Discrete Probability Distributions

- The probability distribution is defined by a <u>probability function</u>, denoted by $f(x)$, which provides the probability **for each value of the random variable.**

- The required conditions for a discrete probability function are:

$$f(x) \geq 0$$

$$\Sigma f(x) = 1$$

Discrete Probability Distributions

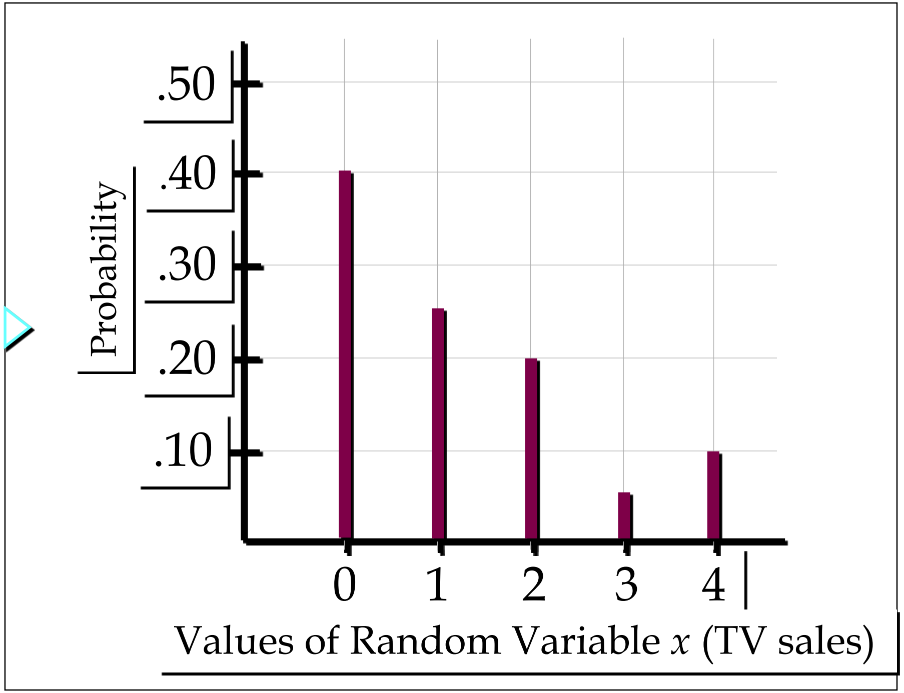
- Using past data on TV sales, …
- a tabular representation of the probability distribution for TV sales was developed.

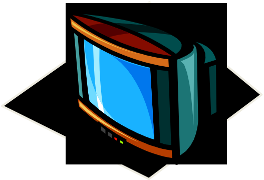
Units Sold	Number of Days	x	$f(x)$
0	80	0	.40
1	50	1	.25
2	40	2	.20
3	10	3	.05
4	20	4	.10
	200		1.00

80/200

Discrete Probability Distributions

- Graphical Representation of Probability Distribution





Discrete Uniform Probability Distribution

- The discrete uniform probability distribution is the simplest example of a discrete probability distribution given by a formula.

- The discrete uniform probability function is

$$\boxed{f(x) = 1/n}$$

the values of the random variable are equally likely

where:

n = the number of values the random variable may assume

No of dots on upside when rolling a die.

$$f(x) = 1/n$$

x	f(x)
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

Expected Value and Variance

- The expected value, or mean, of a random variable is a measure of its central location.

$$E(x) = \mu = \Sigma x f(x)$$

- The variance summarizes the variability in the values of a random variable.

$$\mathrm{Var}(x) = \sigma^2 = \Sigma(x - \mu)^2 f(x)$$

- The standard deviation, σ, is defined as the positive square root of the variance.

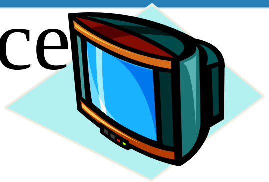
Expected Value and Variance

- Expected Value

x	$f(x)$	$xf(x)$
0	.40	.00
1	.25	.25
2	.20	.40
3	.05	.15
4	.10	.40
	$E(x) =$	1.20

expected number of TVs sold in a day

Expected Value and Variance



Variance and Standard Deviation

x	$x - \mu$	$(x - \mu)^2$	$f(x)$	$(x - \mu)^2 f(x)$
0	-1.2	1.44	.40	.576
1	-0.2	0.04	.25	.010
2	0.8	0.64	.20	.128
3	1.8	3.24	.05	.162
4	2.8	7.84	.10	.784

Variance of daily sales = σ^2 = 1.660

TVs squared

Standard deviation of daily sales = 1.2884 TVs

Binomial Distribution: describes discrete, not continues, data resulting from an experiment know as Bernoulli process.

- Four Properties of a Binomial Experiment

1. The experiment consists of a sequence of n identical trials.
2. Two outcomes, success and failure, are possible on each trial.
3. The probability of a success, denoted by p, does not change (remains fixed) from trial to trial.
4. The trials are independent.

stationarity assumption

Binomial Distribution

- Our interest is in the number of successes occurring in the n trials.
- We let x denote the number of successes occurring in the n trials.

Binomial Distribution

- Binomial Probability Function

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

where:

$f(x)$ = the probability of x successes in n trials

n = the number of trials

p = the probability of success on any one trial

Binomial Distribution

- Binomial Probability Function

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

$$\frac{n!}{x!(n-x)!}$$

Number of experimental outcomes providing exactly **x successes in n trials**

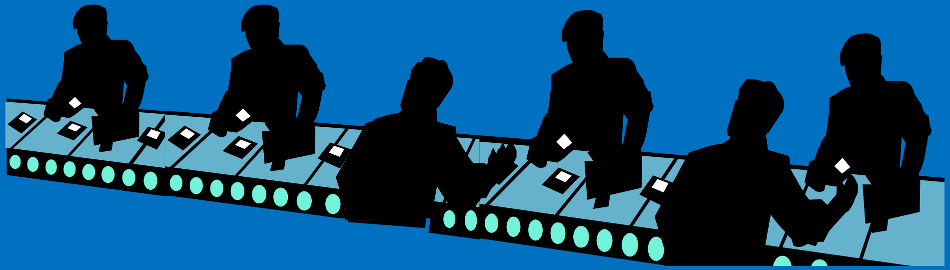
$$p^x (1-p)^{(n-x)}$$

Probability of a particular **sequence** of trial outcomes with x successes in *n* trials

Binomial Distribution

- Example: Evans Electronics

Evans is concerned about a low **retention** rate for employees. In recent years, management has seen a turnover of 10% of the hourly employees annually. Thus, for any hourly employee chosen at random, management estimates a **probability of 0.1** that the person will not be with the company next year.



Binomial Distribution

- Using the Binomial Probability Function

Choosing 3 hourly employees at random, what is the probability that 1 of them will **leave** the company this year?

Let: $p = 0.10$, $n = 3$, $x = 1$

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

$$f(1) = \frac{3!}{1!(3-1)!}(0.1)^1(0.9)^2 = 3(.1)(.81) = .243$$

Binomial Distribution

- Using Tables of Binomial Probabilities

		p									
n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
3	0	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664	.1250
	1	.1354	.2430	.3251	.3840	.4219	.4410	.4436	.4320	.4084	.3750
	2	.0071	.0270	.0574	.0960	.1406	.1890	.2389	.2880	.3341	.3750
	3	.0001	.0010	.0034	.0080	.0156	.0270	.0429	.0640	.0911	.1250

Binomial Distribution

- Expected Value

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

- Variance

$$\mathrm{Var}(x) = \sigma^2 = np(1-p)$$

- Standard Deviation

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

Binomial Distribution

- Expected Value

$$E(x) = \mu = 3(.1) = .3 \text{ employees out of } 3$$

- Variance

$$Var(x) = \sigma^2 = 3(.1)(.9) = .27$$

- Standard Deviation

$$\sigma = \sqrt{3(.1)(.9)} = .52 \text{ employees}$$

Poisson Distribution:

The Poisson distribution and the binomial distribution have some **similarities** but also several **differences**.

The binomial distribution describes a distribution of **two possible outcomes** designated as successes and failures from a given number of trials.

The **Poisson distribution** *focuses only on the number of **discrete occurrences** over some **interval** or continuum.*

For example, whereas a binomial experiment might be used to determine *how many Indian-made cars are in a random sample of 20 cars*, a Poisson experiment might focus on the **number of cars randomly** arriving at an automobile **repair facility** during a 10-minute interval.

The Poisson distribution describes the occurrence of ***rare events***. In fact, the Poisson formula has been referred to as the ***law of improbable events.***

The Poisson Distribution Definitions

- You use the **Poisson distribution** when you are interested in the **number of times** an event occurs in a given **area of opportunity**.
- 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.

  – The number of scratches in a car’s paint
  – The number of mosquito bites on a person
  – The number of computer crashes in a day

Statistics for Managers Using Microsoft Excel® 7e Copyright ©2014 Pearson Education, Inc.

The Poisson Distribution

- 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**)

Poisson Distribution:

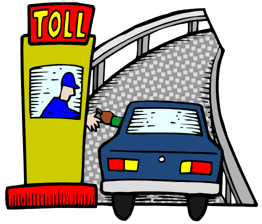
Is used to describe a number of processes, including the distribution of calls going **through a switchboard** system, the demand (needs) of patients for service at a **health institution**, arrival of **trucks and cars at** a toll booth, and the number of **accidents at an intersection**.

- A Poisson distributed random variable is often useful in estimating the number of occurrences over a specified interval of time or space

- It is a discrete random variable that may assume an infinite sequence of values (x = 0, 1, 2, . . .)

Poisson Distribution

- Examples of a Poisson distributed random variable:
- the number of knotholes in 14 linear feet of pine board
- the number of vehicles arriving at a toll booth in **one hour**



Poisson Distribution

- Characteristics of Poisson Prob Dis.

  - 1. The mean no. of vehicles that arrive per rush hour can be **estimated** from **past** traffic data.

  - 2. If we divide rush hour into periods (intervals) of one second each, we will find these statements to be true:

a) The prob that exactly **one vehicle** will arrive at a single booth **per second** is very **small** number and is **constant** for every **one - second interval**.

b) The prob that **two or more vehicles** will arrive **per second** is so **small that we can assign it a zero value.**

c) The no. of arrivals in **any one - second is not dependent on the no. of arrivals in any other one - second** interval.

• Poisson Distribution

- Poisson Probability Function

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

where:

$f(x)$ = probability of x occurrences in an interval

μ = mean number of occurrences in an interval

e = 2.71828