IIT ROORKEE

NPTEL ONLINE CERTIFICATION COURSE

Business Statistics

Introduction

Dr M K BARUA

DEPARTMENT OF MANAGEMENT STUDIES

• 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

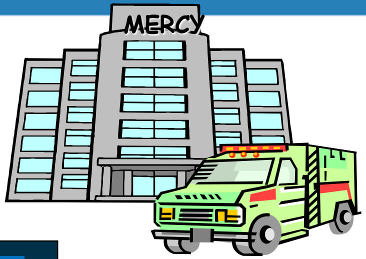
Poisson Distribution

Example: Mercy Hospital

- Patients arrive at the emergency room of Mercy Hospital at the average rate **of 6 per hour** on weekend evenings.

  What is the probability **of 4 arrivals** in **30 minutes** on a weekend evening?

Poisson Distribution

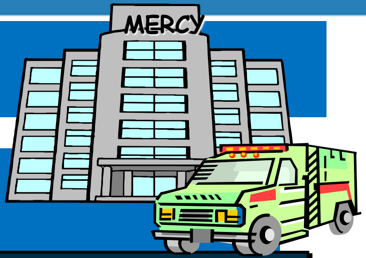


- Using the Poisson Probability Function

$$\mu = 6/\text{hour} = 3/\text{half-hour}, \;\; x = 4$$

$$f(4) = \frac{3^4 (2.71828)^{-3}}{4!} = .1680$$

Poisson Distribution

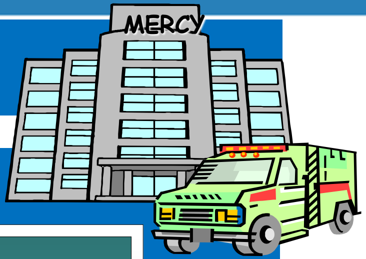


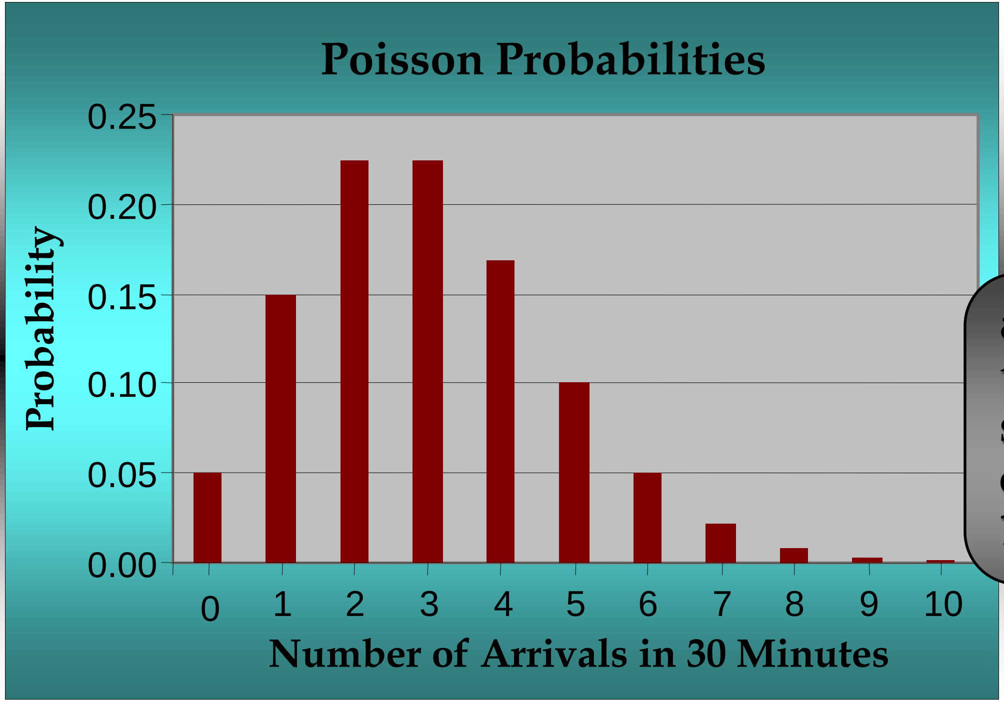
- Using Poisson Probability Tables

					μ					
x	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	.1225	.1108	.1003	.0907	.0821	.0743	.0672	.0608	.0550	.0498
1	.2572	.2438	.2306	.2177	.2052	.1931	.1815	.1703	.1596	.1494
2	.2700	.2681	.2652	.2613	.2565	.2510	.2450	.2384	.2314	.2240
3	.1890	.1966	.2033	.2090	.2138	.2176	.2205	.2225	.2237	.2240
4	.0992	.1082	.1169	.1254	.1336	.1414	.1488	.1557	.1622	.1680
5	.0417	.0476	.0538	.0602	..0668	.0735	.0804	.0872	.0940	.1008
6	.0146	.0174	.0206	.0241	.0278	.0319	.0362	.0407	.0455	.0504
7	.0044	.0055	.0068	.0083	.0099	.0118	.0139	.0163	.0188	.0216
8	.0011	.0015	.0019	.0025	.0031	.0038	.0047	.0057	.0068	.0081

Poisson Distribution

- Poisson Distribution of Arrivals





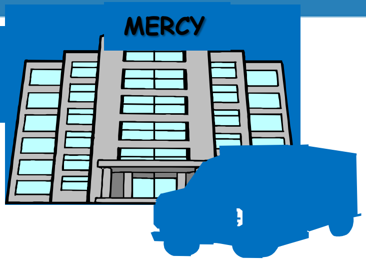
actually, the sequence continues: 11, 12, …

Poisson Distribution

A property of the Poisson distribution is that the mean and variance are equal.

$$\mu = \sigma^2$$

Poisson Distribution



- Variance for Number of Arrivals During 30-Minute Periods

$$\mu = \sigma^2 = 3$$

PD as an approximation of BD:

To avoid tedious calculation of BD, PD can be a reasonable approximation of the BD, but under certain conditions. When “n” is large **(>= to 20)** and “p” is small **(less than or equal to 0.05).**

We can substitute mean of BD that is **“np”** in place of mean of PD that is **λ**.

Ex: We have 20 m/c in a hospital, chance that one will malfunction on any day is 0.02.
What is the prob of 3 m/c malfunction on a day.

PD

$$P(x) = (20*0.02)^{3} * e^{-((20*0.02)} / 3!$$

$$=0.00715$$

BD

$$P(3) = \frac{20! * (.02)^{3} *(0.98)^{17}}{3!\,(20-3)!}$$

$$= 0.0065$$

PD

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

BD

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

$P(x) = (20*0.02)^{3} * e^{-((20*0.02)} / 3!$

$=0.00715$

$$P(3) = \frac{20! * (.02)^3 * (0.98)^{17}}{3!\ (20-3)!}$$

$= 0.0065$

Ex. Given a binomial distribution with $n = 30$ trials and $p = 0.04$, use the Poisson approximation to the binomial to find

a) $P(r = 25)$

b) $P(r = 3)$

c) $P(r = 5)$

Binomial, n =30, p = 0.04; λ= np = 1.2; $e^{-1.2}$ = 0.30119

a) $P(r = 25) = \dfrac{(1.2)^{25}\, e^{-1.2}}{25!} = 0.0000$

b) $P(r = 3) = \dfrac{(1.2)^{3}\, e^{-1.2}}{3!} = 0.0867$

c) $P(r = 5) = \dfrac{(1.2)^{5}\, e^{-1.2}}{5!} = 0.0062$

Hypergeometric Distribution

Statisticians often use the **hypergeometric distribution** to ***complement*** the types of analyses that can be made by using **the binomial distribution**.

Recall that the binomial distribution applies, in theory, only to experiments in which the **trials are done with replacement (independent events)**.

The hypergeometric distribution applies only to experiments in which the **trials are done without replacement.**

The hypergeometric distribution has the following **characteristics**:

- It is **discrete** distribution.
- Each outcome consists of either a success or a failure.
- Sampling is done **without** replacement.
- The population, N, is **finite and known**.
- The number of **successes** in the population, r, **is known**.
- The probability of success **changes from trial to trial**.

Hypergeometric Distribution

- Hypergeometric Probability Function

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \qquad \text{for } 0 \leq x \leq r$$

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

n = number of trials

N = number of elements in the population

r = number of elements in the population labeled success

Hypergeometric Distribution

- Hypergeometric Probability Function

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \quad \text{for } 0 \le x \le r$$

number of ways
x successes can be selected
from a **total** of r successes
in the population

number of ways
$n - x$ failures can be selected
from a **total** of $N - r$ failures
in the population

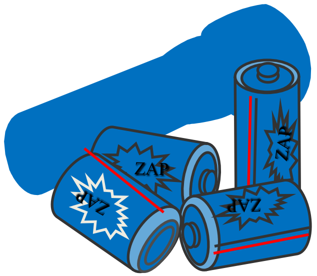
number of ways
a **sample of size** n can be selected
from a population of size N

Hypergeometric Distribution

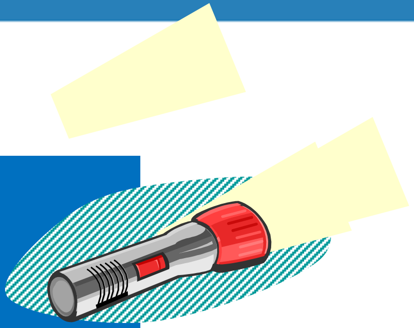
- Example: Neveready

Bob Neveready has removed two dead batteries from a flashlight and inadvertently mingled them with the two good batteries he intended as replacements. The four batteries look identical.

Bob now randomly selects two of the four batteries. What is the probability he selects the two good batteries?



Hypergeometric Distribution



- Using the Hypergeometric Function

$$f(x)=\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}=\frac{\binom{2}{2}\binom{2}{0}}{\binom{4}{2}}=\frac{\left(\frac{2!}{2!0!}\right)\left(\frac{2!}{0!2!}\right)}{\left(\frac{4!}{2!2!}\right)}=\frac{1}{6}=.167$$

where:

x = 2 = number of <u>good</u> batteries selected

n = 2 = number of batteries selected

N = 4 = number of batteries in total

r = 2 = number of <u>good</u> batteries in total

Hypergeometric Distribution

- Mean

$$E(x) = \mu = n\left(\frac{r}{N}\right)$$

- Variance

$$Var(x) = \sigma^2 = n\left(\frac{r}{N}\right)\left(1-\frac{r}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Hypergeometric Distribution

- Mean

$$\mu = n\left(\frac{r}{N}\right) = 2\left(\frac{2}{4}\right) = 1$$

- Variance

$$\sigma^2 = 2\left(\frac{2}{4}\right)\left(1-\frac{2}{4}\right)\left(\frac{4-2}{4-1}\right) = \frac{1}{3} = .33$$

Hypergeometric Distribution

- Example: 3 different computers are checked out from 10 in the department. 4 of the 10 computers have illegal software loaded. What is the probability that 2 of the 3 selected computers have illegal software loaded?

$N = 10$ $\quad$ $n = 3$

$r = A = 4$ $\quad$ $x = 2$

$$P(X=2\,|\,3,10,4)=\frac{\binom{A}{X}\binom{N-A}{n-X}}{\binom{N}{n}}=\frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}}=\frac{(6)(6)}{120}=0.3$$

The probability that 2 of the 3 selected computers have illegal software loaded is 0.30, or 30%.

Hypergeometric Distribution

- Consider a hypergeometric distribution with n trials and let $p = (r/n)$ denote the probability of a success on the first trial.
- If the **population size is large**, the term $(N – n)/(N – 1)$ approaches 1.
- The expected value and variance can be written $E(x) = np$ and $Var(x) = np(1 – p)$.
- Note that these are the expressions for the expected value and variance of a binomial distribution.

Continuous Probability Distributions

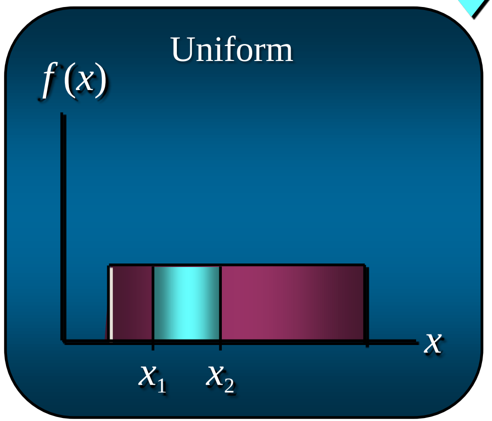
- A continuous random variable can assume any value in **an interval on t**he real line or in a collection of intervals.

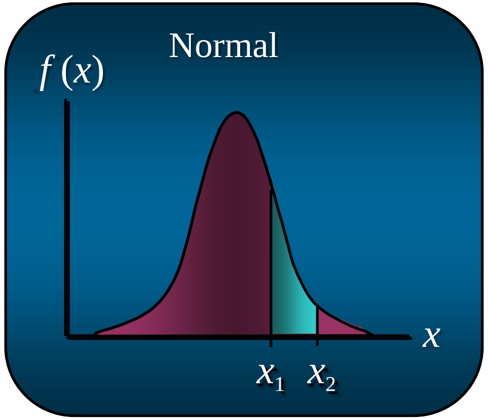
- It is not possible to talk about the probability of the random variable **assuming a particular** value.

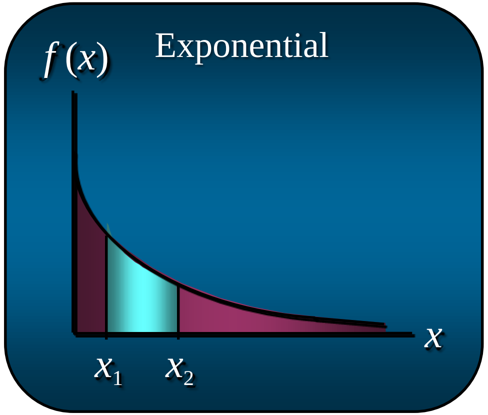
- Instead, we talk about the probability of the random variable assuming a value within a **given interval**.

Continuous Probability Distributions

- The probability of the random variable assuming a value within some given ***interval*** from x_1 to x_2 is **defined** to be the **area under the graph** **of the** **probability density function** between x_1 and x_2.







Uniform Probability Distribution

- A random variable is <u>uniformly distributed</u> whenever the probability is proportional to the *interval's length*.

- The <u>uniform probability density function</u> is:

$$f(x) = 1/(b - a) \quad \text{for } a \le x \le b$$
$$= 0 \quad \text{elsewhere}$$

where: a = smallest value the variable can assume

b = largest value the variable can assume

Uniform Probability Distribution

- Expected Value of x

$$E(x) = (a + b)/2$$

- Variance of x

$$Var(x) = (b - a)^2/12$$

Uniform Probability Distribution

- Example: Slater's Buffet

Slater customers are charged for the amount of salad they take. Sampling suggests that the amount of salad taken is uniformly distributed between 5 ounces and 15 ounces. Find out mean and S.D.

Uniform Probability Distribution

- Uniform Probability Density Function

$$f(x) = 1/10 \quad \text{for } 5 \leq x \leq 15$$
$$= 0 \quad \text{elsewhere}$$

where:

x = salad plate filling weight

Uniform Probability Distribution

- Expected Value of x

$$E(x) = (a + b)/2 = (5 + 15)/2 = 10$$

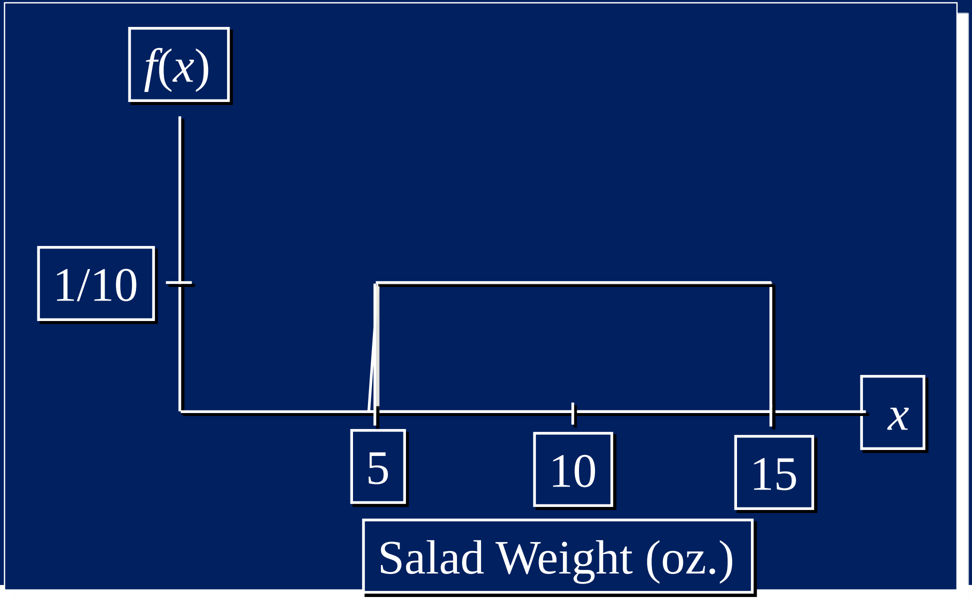
- Variance of x

$$Var(x) = (b - a)^2/12 = (15 - 5)^2/12 = 8.33$$

$$SD = \sqrt{8.33}$$

Uniform Probability Distribution

- Uniform Probability Distribution
  for Salad Plate Filling Weight



Uniform Probability Distribution

What is the probability that a customer will take between 12 and 15 ounces of salad?