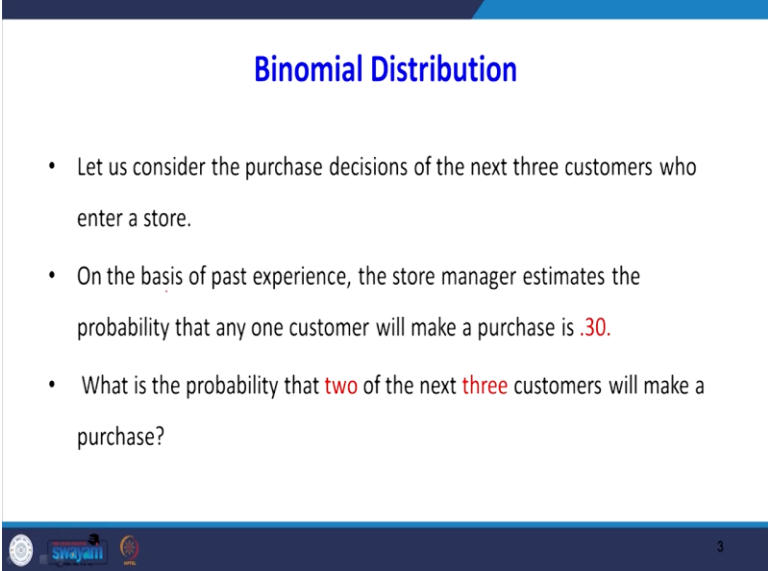


Data Analytics with Python
Prof. Ramesh Anbanadham
Department of Management Studies
Indian Institute of Technology – Roorkee

Lecture - 09
Probability Distribution – II

Dear students in this lecture, we are going to see some of the discrete distributions and continuous distributions, the discrete we are going to study the binomial poisson hypergeometric in continuous that where we going to study uniform exponential and normal, first you will see, what is the application of this Binomial Distribution?

(Refer Slide Time: 00:48)



Binomial Distribution

- Let us consider the purchase decisions of the next three customers who enter a store.
- On the basis of past experience, the store manager estimates the probability that any one customer will make a purchase is .30.
- What is the probability that two of the next three customers will make a purchase?

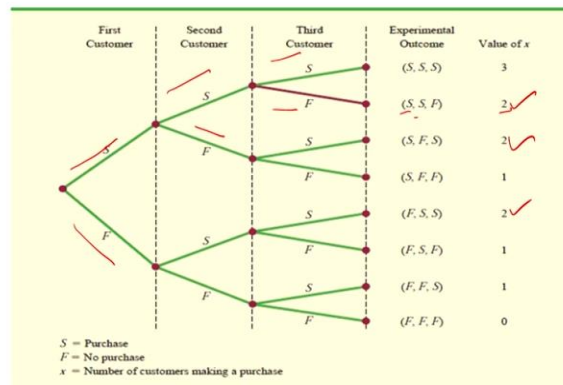
The slide is a presentation slide with a blue header and footer. The title 'Binomial Distribution' is in blue. The content consists of three bullet points. The footer contains logos for IIT Roorkee, Swayam, and a small circular logo, along with the number 3.

Otherwise what is the need of this binomial distribution. We will take one practical example that I will solve manually with the help of our concept of probability then I will tell you how the concept of binomial distribution will help us to solve this problem quick way, but just consider the purchase decision of the next 3 customers who enter your store and the basis of past experience.

The store manager estimates the probability that any one customer will make your purchase is 0.30. Now, what is the probability that 2 of the next 3 customers will make a purchase?

(Refer Slide Time: 01:29)

Tree diagram for the Martin clothing store problem



That problem is drawn in the form of a tree diagram the first customer, second customer, third customer when the first customer comes, there are 2 possibility he can go for purchase or not purchase, purchase mention as say S success failure. The second customer also may go for purchase, not purchase third customer purchase, not purchase. Look at the first possibility the First customer is purchasing, so, success, success, success that is S, S, S.

If you take X look at the bottom of this slide, if X equals the number of customers making a purchase, so, in this category, all 3 customers all 3 are going to buy look at the second possibility success, success failure. So, in this out of 3 customers 2 customers S S going to purchase like this we have displayed all possibilities. Now, the problem is X what is asked is 2 out of 3 customers what they have purchased.

So, number of customers making a purchases how many possibilities are there 2 customers 1, 2, 3 how we got this 3 is nothing but your 3C_2 the 3C_2 is out of 3, what is the probability? What is the chance that 2 customers will buy there will be a possibility of 3C_2 , so the 3C_2 is 3.

(Refer Slide Time: 02:57)

Trial Outcomes

Trial Outcomes				Probability of Experimental Outcome
1st Customer	2nd Customer	3rd Customer	Experimental Outcome	
Purchase	Purchase	No purchase	(S, S, F)	$pp(1-p) = p^2(1-p)$ $= (.30)^2(.70) = .063$
Purchase	No purchase	Purchase	(S, F, S)	$p(1-p)p = p^2(1-p)$ $= (.30)^2(.70) = .063$
No purchase	Purchase	Purchase	(F, S, S)	$(1-p)pp = p^2(1-p)$ $= (.30)^2(.70) = .063$

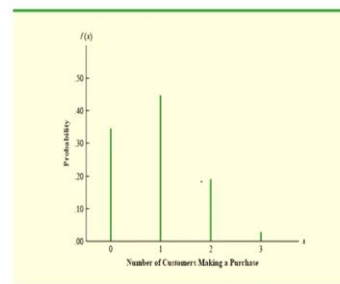
$P(x=2) = {}^nC_x p^x q^{n-x}$

Now, we will see the first customer all possibilities purchase purchase no purchase success success, so, success we know, we are calling it is a p failure is 1 - p. So, it is a p.p into 1 - p, p is $(0.30)^2$ multiply by 0.70, it is a 0.63. For second chance also first customer may buy second customer did not buy third customer also buy, so, it is success failure success. So, p 1 - p into p again $p^2 \cdot (1 - p) = 0.063$, third choice, no purchase purchase purchase FSS, so 1 - p, p,p so it is a $p^2 \cdot (1 - p) = 0.063$.

(Refer Slide Time: 03:48)

Graphical representation of the probability distribution for the number of customers making a purchase

x	P(x)
0	$0.7 \times 0.7 \times 0.7 = 0.343$
1	$0.3 \times 0.7 \times 0.7 + 0.7 \times 0.3 \times 0.7 + 0.7 \times 0.7 \times 0.3 = 0.441$
2	0.189
3	0.027



If I plot this one see x possibility, x possibility this possibility is see this possibility, when $x = 0$, that means, what is the chance that all 3 customers will not buy? That is this case 0.7, 0.7, 0.7. If I say $x = 1$, what is the probability that 1 customer will buy, if $x = 2$? What is the probability 2

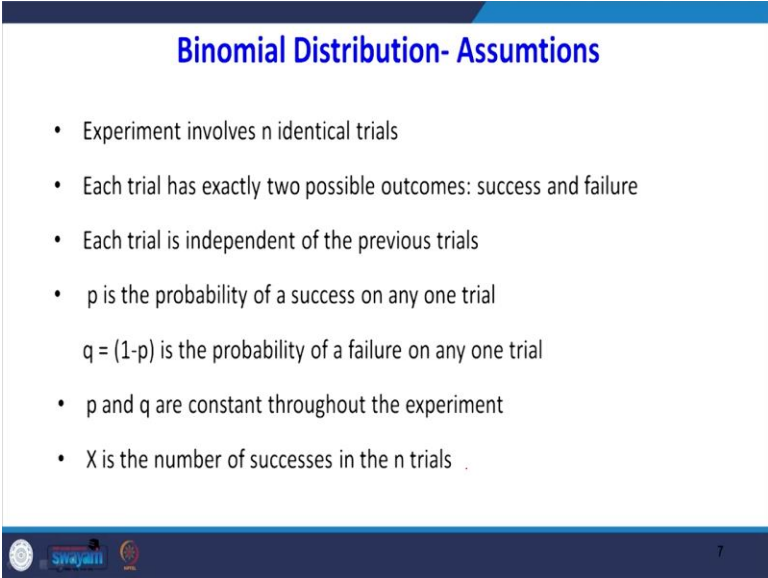
customers will come, 2 customers will buy with $x = 3$? What is the probability that 3 customers will buy? So, we can do with the concept of probability manually like this.

So, the question is asked is this one? What is the probability that 2 customers will buy out of 3. So, this is your probability distributions. So, this we can do manually but it will take a lot of time. So, here with the help of your binomial distributions, you can get right this all possibilities going back this way nC_x So,

$P(X=2) = {}^nC_x \cdot p^x \cdot q^{n-x}$, where $n = 3$ where $x = 2$ p is the probability success q = probability of failure.

So, this nC_x will give you different combinations of that event to happen. So, this p^x , see that always here $p^2 q^{n-2}$. So, this will give you the probability. So, that is the purpose of this binomial distributions.

(Refer Slide Time: 05:35)



Binomial Distribution- Assumptions

- Experiment involves n identical trials
- Each trial has exactly two possible outcomes: success and failure
- Each trial is independent of the previous trials
- p is the probability of a success on any one trial
 $q = (1-p)$ is the probability of a failure on any one trial
- p and q are constant throughout the experiment
- X is the number of successes in the n trials

What is the property of this binomial distribution experiment involves in identical trials, each trial has exactly 2 possible outcomes success, failures. Each trial is independent of the previous trial for example, the previous case customer one may buy or may not buy. That is not depending upon the for example for the second customer is intension of purchase is not based on the previous customer.

So, p is the probability of success on any one trial, the q is probability of failure on any one trial. So, p and q are constant throughout the experiment, this is important assumption. So, x is the number of success in the end trials, there the previous case we are seeing $x = 2$.

(Refer Slide Time: 06:23)

Binomial Distribution

- Probability function $P(X) = \frac{n!}{X!(n-X)!} p^X \cdot q^{n-X}$ for $0 \leq X \leq n$
- Mean value $\mu = n \cdot p$ ✓
- Variance and standard deviation $\sigma^2 = n \cdot p \cdot q$
 $\sigma = \sqrt{\sigma^2} = \sqrt{n \cdot p \cdot q}$

So the probability function as I told you can written as

$${}^nC_x = \frac{n!}{X!(n-X)!} \cdot p^X \cdot q^{n-X}$$

the mean of this distribution is $\mu = n \cdot p$ directly we are going to answered. So, you can find out I think but using the concept of expected value, when you multiply this $P(X)$ with the x , $\sum x \cdot P(x)$, when you simplify will get this formula. Similarly, variance directly we are going to use only the result variance = nPq

standard deviation = root of nPq .

(Refer Slide Time: 07:03)

Binomial Table

SELECTED VALUES FROM THE BINOMIAL PROBABILITY TABLE

EXAMPLE: $n = 10, x = 3, p = .40; f(3) = .2150$

n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
9	0	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020
	1	.2985	.1874	.1079	.0600	.0328	.0175	.0090	.0046	.0023	.0010
	2	.0629	.0290	.0131	.0054	.0025	.0011	.0005	.0002	.0001	.0000
	3	.0077	.0046	.0025	.0013	.0006	.0003	.0001	.0000	.0000	.0000
	4	.0006	.0004	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000
10	0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010
	1	.3151	.1874	.1079	.0600	.0328	.0175	.0090	.0046	.0023	.0010
	2	.0746	.0290	.0131	.0054	.0025	.0011	.0005	.0002	.0001	.0000
	3	.0105	.0046	.0025	.0013	.0006	.0003	.0001	.0000	.0000	.0000
	4	.0010	.0004	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	5	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	6	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

Sometime there is ready made table is there to find out binomial probability values for example, $n = 10$ and n is number of trials, $x = 3$, $p = 0.40$, $n = 10$, $x = 3$ that value is this one 0.2150. So, the tables are provided you do not use your calculator it may take more time.

(Refer Slide Time: 07:34)

Mean and Variance

- Suppose that for the next month the Clothing Store forecasts 1000 customers will enter the store.
- What is the expected number of customers who will make a purchase?
- The answer is $\mu = np = (1000)(.3) = 300$.
- For the next 1000 customers entering the store, the variance and standard deviation for the number of customers who will make a purchase are

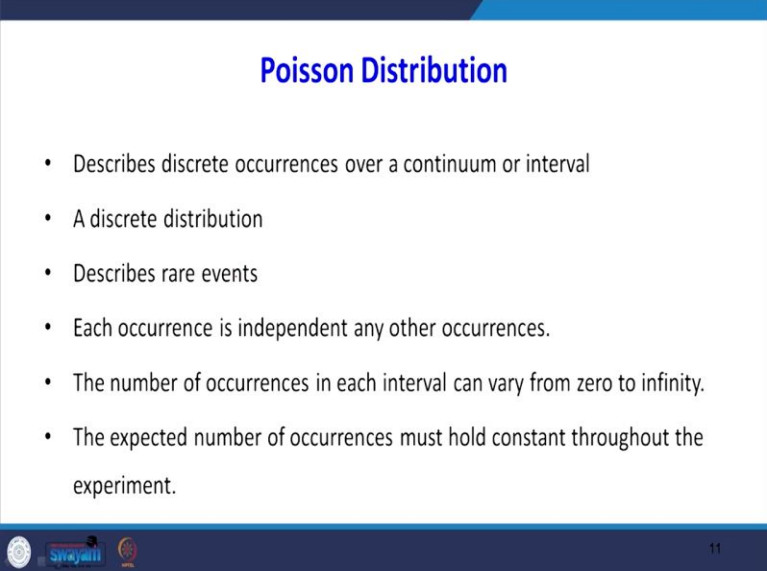
$$\sigma^2 = np(1 - p) = 1000(.3)(.7) = 210$$

$$\sigma = \sqrt{210} = 14.49$$

Now we will find out mean and variance suppose, for the next month the clothing store forecast 1000 customers will enter the store the previous problem their 1000 customers. So, what is the expected number of customers who will make the purchase out of 1000. So, the answer is $\mu = np$ so n this 1000 probability of success point 3, so it is equal to 300. So, that is out of 1000 customers, there is a chance that 300 customers will make the purchase.

For the next 1000 customers entering the store the variance and standard deviation of the number of customers who will make the purchase can you written as npq . So, n is given p is given $1 - p$ 70 210 the standard deviation is 14.49.

(Refer Slide Time: 08:24)



Poisson Distribution

- Describes discrete occurrences over a continuum or interval
- A discrete distribution
- Describes rare events
- Each occurrence is independent any other occurrences.
- The number of occurrences in each interval can vary from zero to infinity.
- The expected number of occurrences must hold constant throughout the experiment.

11

Then will go to next distribution Poisson distribution describes discrete occurrences over continuous or interval. Generally Poisson distribution is for rare events, it is an discrete distribution. X can have only few discrete values. Yes, it describes the rare events, each occurrence is independent of any other occurrences. The number of occurrences in each interval can vary from 0 to infinity. The expected number of occurrences must hold constant throughout the experiment. These are the assumption of Poisson distribution.

(Refer Slide Time: 09:09)

Poisson Distribution: Applications

- Arrivals at queuing systems
 - airports -- people, airplanes, automobiles, baggage
 - banks -- people, automobiles, loan applications
 - computer file servers -- read and write operations
- Defects in manufactured goods
 - number of defects per 1,000 feet of extruded copper wire
 - number of blemishes per square foot of painted surface
 - number of errors per typed page

Some of the examples are application of Poisson distribution. So, arrival at the queuing system follow Poisson distribution, any airport people may arrive, airplane may arrive automobile may arrive, baggage may arrive that follow Poisson distribution. The banks people automobiles loan applications this arrival pattern follow Poisson distribution in computer file services, read and write operations will follow Poisson distribution

The defects in manufacturing goods can be considered an example of Poisson distribution. For example, number of defects per 1000 feet of extruded copper wire see that the n is very large. The probability of success is very low. Another example of Poisson distribution is number of blemishes per square foot of painted surface blemishes kind of a defect in paint number of errors per typed page, these are the example of your Poisson distributions.

(Refer Slide Time: 10:15)

Poisson Distribution

- Probability function

$$P(X) = \frac{\lambda^X e^{-\lambda}}{X!} \text{ for } X = 0, 1, 2, 3, \dots$$

where:
 λ = long-run average
 $e = 2.71828\dots$ (the base of natural logarithms)

Mean value

$$\lambda$$

Variance

$$\lambda$$

Standard deviation

$$\sqrt{\lambda}$$

So, the probability function for Poisson distribution = $(\lambda^X e^{-\lambda})/X!$

Where X can take only discrete value, the lambda is mean that is a long run average the value of e is 2.71. The mean of this Poisson distribution is lambda variance of the Poisson distribution lambda standard deviation of Poisson distribution is root of lambda. So, this distribution called the univariate distribution.

Because it has unique parameter that means, it is a special property of your Poisson distribution where mean and variance is same. So, that is mean and variances. Same that is lambda.

(Refer Slide Time: 10:55)

Poisson Distribution: Example

$\lambda = 3.2$ customers/4 minutes
 $X = 10$ customers/8 minutes
 Adjusted λ
 $\lambda = 6.4$ customers/8 minutes

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

$$P(X=10) = \frac{6.4^{10} e^{-6.4}}{10!} = 0.0528$$

$\lambda = 3.2$ customers/4 minutes
 $X = 6$ customers/8 minutes
 Adjusted λ
 $\lambda = 6.4$ customers/8 minutes

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

$$P(X=6) = \frac{6.4^6 e^{-6.4}}{6!} = 0.1586$$

And another caution while using Poisson distribution is the unit of this lambda and X should be same, for example, lambda equal to 3.2 customers for 4 minutes probability of around 10 customers in 8 minutes. So, you have to adjust the lambda, how will you adjust, you multiply by 2, so it will be 6.4 into 8. Now, the unit of X and lambda is same, then you can use your probability function, $P(X) = (\lambda^X e^{-\lambda})/X!$ and substitute $X = 10$.

We are getting 0.025. See the second one, lambda = 3.2 customers by 4 minutes, $X = 6$ customers for 8 minutes. So, you have to multiply lambda. So, it will be 6.4 in 8 then P of X will get the answer. So, what is the point here is the unit of lambda and X should be same.

(Refer Slide Time: 11:49)

Poisson Probability Table
Example: $\mu = 10, x = 5; f(5) = .0378$

x	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10
0	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0000
1	.0010	.0009	.0009	.0008	.0007	.0007	.0006	.0005	.0005	.0005
2	.0046	.0043	.0040	.0037	.0034	.0031	.0029	.0027	.0025	.0023
3	.0140	.0131	.0123	.0115	.0107	.0100	.0093	.0087	.0081	.0076
4	.0319	.0302	.0285	.0269	.0254	.0240	.0226	.0213	.0201	.0189
5	.0581	.0555	.0530	.0506	.0483	.0460	.0439	.0418	.0398	.0378
6	.0881	.0851	.0822	.0793	.0764	.0736	.0709	.0682	.0656	.0631
7	.1145	.1118	.1091	.1064	.1037	.1010	.0982	.0955	.0928	.0901
8	.1302	.1266	.1239	.1211	.1182	.1152	.1121	.1090	.1058	.1026
9	.1317	.1285	.1251	.1216	.1180	.1143	.1104	.1064	.1023	.0981
10	.1198	.1160	.1121	.1080	.1038	.0994	.0949	.0903	.0856	.0808
11	.0991	.0952	.0911	.0868	.0823	.0777	.0730	.0682	.0633	.0583
12	.0752	.0716	.0678	.0638	.0596	.0553	.0509	.0464	.0418	.0371
13	.0526	.0492	.0457	.0421	.0384	.0346	.0307	.0268	.0228	.0188
14	.0342	.0311	.0279	.0245	.0210	.0174	.0137	.0100	.0062	.0024
15	.0208	.0179	.0149	.0117	.0084	.0050	.0016	.0004	.0001	.0000
16	.0118	.0097	.0075	.0052	.0030	.0017	.0008	.0003	.0001	.0000
17	.0063	.0049	.0035	.0022	.0013	.0007	.0003	.0001	.0000	.0000
18	.0032	.0023	.0015	.0009	.0005	.0002	.0001	.0000	.0000	.0000
19	.0015	.0010	.0006	.0003	.0001	.0000	.0000	.0000	.0000	.0000
20	.0007	.0004	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000
21	.0003	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
22	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
23	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
24	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

Here also there is a probability Poisson probability table is available. So, when $\mu = 10$ $X = 5$, you see in the column shows mu. So, $\mu = 10$ here, $X = 5$ is this value. So, this value is 0.378. So, you cannot, you need not to use your calculator directly you can read it from the table. But we are going to use these things in your practical class we are going to find out probabilities with the help of Python.

(Refer Slide Time: 12:23)

The Hypergeometric Distribution

- The **binomial distribution** is applicable when selecting from a finite population with replacement or from an infinite population without replacement.
- The hypergeometric distribution is applicable when selecting from a **finite population without replacement**.

We will go for hypergeometric distribution, see the binomial distribution is applicable, when selecting from finite population with replacement or for an infinite population without replacement. So, whenever the concept of without replacement comes, then we have to think of using hypergeometric distribution. The hypergeometric distribution is applicable when selecting from your finite population without replacement.

(Refer Slide Time: 12:52)

Hyper Geometric Distribution

- Sampling without replacement from a finite population
- The number of objects in the population is denoted N .
- Each trial has exactly two possible outcomes, success and failure.
- Trials are not independent
- X is the number of successes in the n trials
- The binomial is an acceptable approximation, if $N/10 > n$ Otherwise it is not.

The properties of hypergeometric distribution, so, sampling without replacement from your finite population then you should go for hypergeometric distribution. The number of objects in the population is denoted as N , each trial has exactly 2 possible outcomes, success or failure. Similar to binomial distribution trials are not independent, this is one different property, when compared

to binomial. In the binomial distributions, since we are going with replacement, the trials are independent; here we are going without replacement.

So, the trials are not independent, it is dependent. Then, that means, the P will not be fixed, the probability of P will not be fixed, every time we will get different answer for the P. X is the number of success in the n trials. The binomial is acceptable approximation $N/10 \geq n$, otherwise, it is not.

(Refer Slide Time: 13:53)

Hypergeometric Distribution

- Probability function
 - N is population size ✓
 - n is sample size ✓
 - A is number of successes in population ✓
 - x is number of successes in sample ✓
- Mean Value
- Variance and standard deviation

$$P(x) = \frac{{}^A C_x ({}^{N-A} C_{n-x})}{{}^N C_n}$$

$$\mu = \frac{A \cdot n}{N}$$

$$\sigma^2 = \frac{A(N-A)n(N-n)}{N^2(N-1)}$$

$$\sigma = \sqrt{\sigma^2}$$

That we will see the probability function with discrete values use probability mass function, if it is a continuous will use PDF probability density function. So, the probability mass function is $P(x) = ({}^A C_x) ({}^{N-A} C_{n-x}) / {}^N C_n$. So, here see the capital letters represent for the population. So N is for population size, A is the number of success in the population, small letter, n for the sample size, the x is number of success in the sample.

So, $P(x) = ({}^A C_x) ({}^{N-A} C_{n-x}) / {}^N C_n$. The mean value = $A \cdot n / N$. The variance and of hypergeometric distribution is $A(N-A)n(N-n) / N^2(N-1)$, root of variance is the formula for standard deviation.

(Refer Slide Time: 14:57)

The Hypergeometric Distribution Example

- Different computers are checked 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?
- So, N = 10, n = 3, A = 4, X = 2

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

$N = 10$
 $A = 4$
 $n = 3$
 $X = ?$

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

Will see one problem here, different computers are checked from 10 in the department, 4 out of 10 computers have illegal software loaded. What is the probability that 2 of the 3 selected computers will have illegal software loaded. By looking at the problem we are feeling that it is a finite population. Whenever there is a finite population, we should think of hypergeometric distribution.

So, N is given that N is your population size, A is 4, because we know 4 is out of 10, 4 computers are illegal software. Because we know the how much illegal software is installed in the population A = 4. Then, what is the probability that x is 2 of the 3 selected the computers will have illegal software loaded. So, x = 2, here n is 3 there are 2 things is there, one is for the population and population N = 10, A = 4.

For the sample n = 3, then what is the probability of that 2 out of 3 selected computers will have illegal software loaded. So now substituted, everything is given, so we will get 0.3. So what is the meaning of this 0.3, the probability that 2 of the 3 selected computers will have illegal software loaded is 30%.

(Refer Slide Time: 16:38)

Continuous Probability Distributions

- A continuous random variable is a variable that can assume any value on a continuum (can assume an uncountable number of values)
 - thickness of an item
 - time required to complete a task
 - temperature of a solution
 - height
- These can potentially take on any value, depending only on the ability to measure precisely and accurately.

Then we will go to the continuous probability distributions. A continuous random variable is a variable that can assume any value in the continuum. That is, can assume an uncountable number of values. Thickness of an item, time required to complete a task, temperature of a solution, height, these are example of continuous random variable. This can potentially take any value depending only on the ability to measure precisely and accurately.

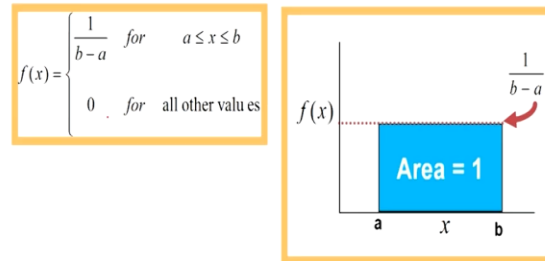
The Uniform Distribution

- The uniform distribution is a probability distribution that has equal probabilities for all possible outcomes of the random variable
- Because of its shape it is also called a rectangular distribution

First we will see uniform distribution. The uniform distribution is the probability distribution that has equal probabilities for all possible outcomes of a random variable. That is the probability of random numbers also. When we say random numbers, the probability of choosing any number is same, because of its shape, it is also called a rectangular distribution.

(Refer Slide Time: 17:31)

Uniform Distribution



Look at this, Uniform Distribution. The uniform distribution is defined on interval, $a \leq x \leq b$. So the probability function is $1/(b-a)$, where 'b' is upper limit, 'a' is the lower limit. In other interval, the value of $f(x)$ is 0, the area = 1.

(Refer Slide Time: 17:51)

Uniform Distribution: Mean and Standard Deviation

Mean

$$\mu = \frac{a + b}{2} \quad \checkmark$$

Standard Deviation

$$\sigma = \frac{b-a}{\sqrt{12}} \quad \checkmark$$

The mean of a uniform distribution is $\mu = (a + b)/2$,

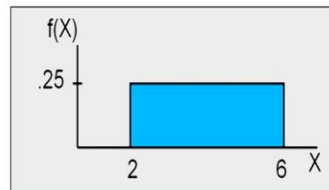
Standard deviation = $(b - a) / \sqrt{12}$. These are very standard results, the derivation of this one, you can refer some textbook.

(Refer Slide Time: 18:07)

The Uniform Distribution

Example: Uniform probability distribution over the range $2 \leq X \leq 6$: 345

$$f(X) = \frac{1}{6-2} = .25 \text{ for } 2 \leq X \leq 6$$



$$\mu = \frac{a+b}{2} = \frac{2+6}{2} = 4$$

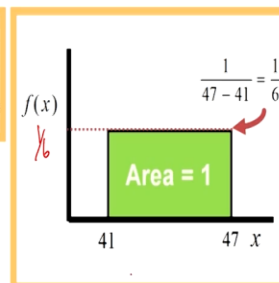
$$\sigma = \sqrt{\frac{(b-a)^2}{12}} = \sqrt{\frac{(6-2)^2}{12}} = 1.1547$$

Now, we will see some problem using uniform distribution, suppose uniform probability distribution over the range is defined this way. $2 \leq x \leq 6$. So if you want to know $f(x) = 1/(b-a)$, where b is 6, a is 2, so, $(1/(6-2)) = 0.25$. What is the meaning of this 0.25 means, in between 2 and 6, you can select any random variable, maybe 3, you may select 3 or 4 or 5, the probability of choosing 3 or 4 5 is, 0.25. So, the mean $= (a+b)/2$, where a is 2 b is 6 $= (8/2) = 4$. Similarly, standard deviation $= (b-a) / \sqrt{12}$, that is 1.17.

(Refer Slide Time: 18:56)

Uniform Distribution Example

$$f(x) = \begin{cases} \frac{1}{47-41} & \text{for } 41 \leq x \leq 47 \\ 0 & \text{for all other values} \end{cases}$$



We will see another example, suppose a random variable is defined between 41 to 47. First, we have to find out probability density function 1 divided by b - a, so 47 - 41, so that is 1 / 6. So 1 / 6 the height of this rectangle of distribution is 1 / x this value is your 1 / 6. The lower limit 41 upper limit 47.

(Refer Slide Time: 19:21)

Uniform Distribution: Mean and Standard Deviation

<p>Mean</p> $\mu = \frac{a + b}{2}$	<p>Mean</p> $\mu = \frac{41 + 47}{2} = \frac{88}{2} = 44$
<p>Standard Deviation</p> $\sigma = \frac{b - a}{\sqrt{12}}$	<p>Standard Deviation</p> $\sigma = \frac{47 - 41}{\sqrt{12}} = \frac{6}{3.464} = 1.732$

So, the mean of the distribution is (41 + 47) divided by 2

= 88 divided by 2 = 44

standard deviation (b - a) / $\sqrt{12}$. So, (47 - 41) / $\sqrt{12}$

= (47-41)/3.464

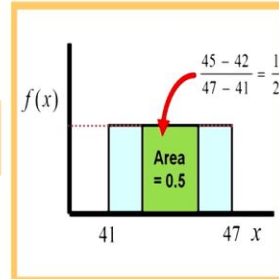
=1.732.

(Refer Slide Time: 19:37)

Uniform Distribution Probability

$$P(x_1 \leq X \leq x_2) = \frac{x_2 - x_1}{b - a}$$

$$P(42 \leq X \leq 45) = \frac{45 - 42}{47 - 41} = \frac{1}{2}$$



Now, see the uniform distribution is defined in this interval 41 to 47 right. Suppose, we want to know the probability between 42 and 45, is it 42 and 45. So, that had not been done by 45 x, that is:

$$x_2 - x_1 / (b - a)$$

x_2 is 45, x_1 is 42 So,

$45 - 42 = 3$, 3 divided by 6 = $1/2$. So, this area is 0.5.

(Refer Slide Time: 20:04)

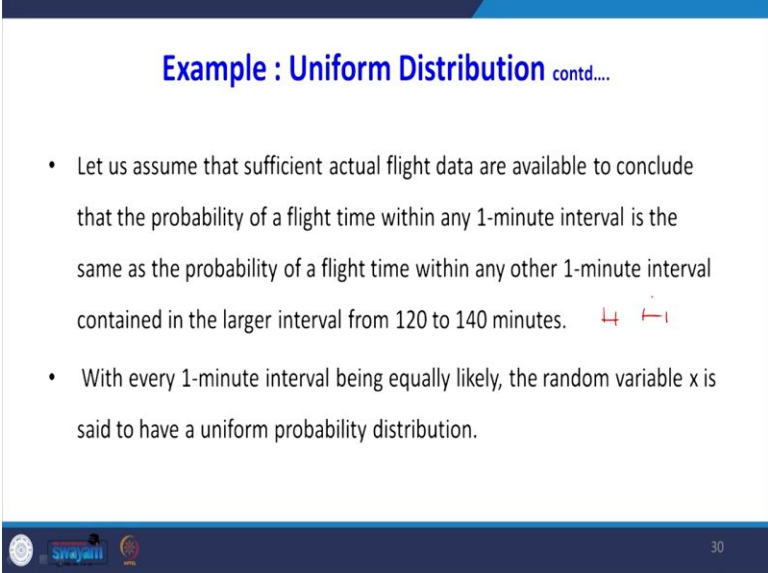
Example : Uniform Distribution

- Consider the random variable x representing the flight time of an airplane traveling from Delhi to Mumbai.
- Suppose the flight time can be any value in the interval from 120 minutes to 140 minutes.
- Because the random variable x can assume any value in that interval, x is a continuous rather than a discrete random variable

You will see another example of this uniform distribution. Consider a random variable x representing the flight time of an airplane travelling from Delhi to Mumbai. Suppose the flight

time can any value in the interval between 122 to 140 minutes. Because of the random variable x can assume any value in the interval, x is a continuous rather than a discrete random variable.

(Refer Slide Time: 20:30)



Example : Uniform Distribution contd....

- Let us assume that sufficient actual flight data are available to conclude that the probability of a flight time within any 1-minute interval is the same as the probability of a flight time within any other 1-minute interval contained in the larger interval from 120 to 140 minutes. $\frac{1}{20}$ $\frac{1}{20}$
- With every 1-minute interval being equally likely, the random variable x is said to have a uniform probability distribution.

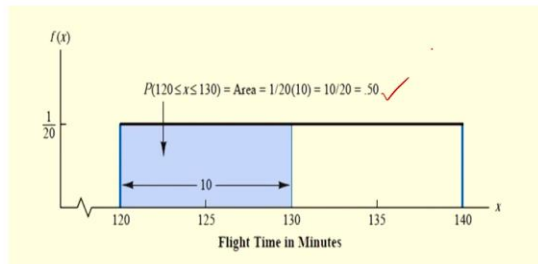
30

Let us assume that sufficient actual flight data are available to conclude that the probability of your flight time within any one minute interval is the same as the probability of a flight at time within any other one minute interval, contained in the large interval from 120 to 140. So, that is the properties of your uniform distribution. That means, in any one, any small interval you should have the same value.

If any one interval, if any between this interval the probability will say next, another one minute interval, the probability same. With every one minute interval being equally likely the random variable x is set to have a uniform probability distribution. So, the upper limit is 140, the lower limit is 120, $1/(140 - 120)$.

(Refer Slide Time: 21:20)

Probability of a flight time between 120 and 130 minutes



So, the height of this rectangle distribution is 1 divided by 20, it is starting a value is 120, b value is 140. Suppose, you are asked to find out, probability of your flight time between 120 and 130 minutes. So, somebody is asking what is the probability of that, flight arriving between 120 and 130 minutes. So, what you have to do, this is $(130 - 120)$ divided by $(140 - 120)$. When you simplify you will get 0.50. So the probability of that flight arrive between 120, 130 minutes is 50%.

(Refer Slide Time: 22:05)

Exponential Probability Distribution

- The exponential probability distribution is useful in describing the time it takes to complete a task.
- The exponential random variables can be used to describe:

Time between vehicle arrivals at a toll booth



Time required to complete a questionnaire



Distance between major defects in a highway



Next we will go to the exponential probability distribution. The exponential probability distribution is useful in describing the time it takes to complete a task. The exponential random variable can be used to describe the time between vehicle arrivals at a toll booth, time required to

complete a questionnaire, the distance between major defects in a highway. For example, whenever the time between arrival and time required to complete the questionnaire, whenever the word the between comes. Then that is the then it is appropriate to use exponential distribution.

(Refer Slide Time: 22:43)

Exponential Probability Distribution

- Density Function

$$f(x) = \frac{1}{\mu} e^{-x/\mu}$$

where: μ = mean
 $e = 2.71828$

The density function for a exponential probability distribution is $f(x) = (1/\mu) \cdot e^{(-x/\mu)}$. Here, μ is the mean.

(Refer Slide Time: 22:52)

Exponential Probability Distribution

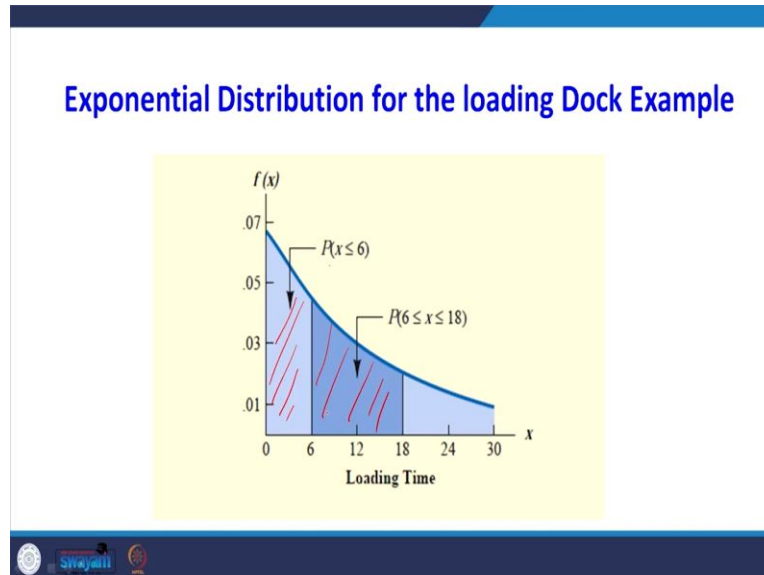
- Suppose that x represents the loading time for a truck at loading dock and follows such a distribution.
- If the mean, or average, loading time is 15 minutes ($\mu = 15$), the appropriate probability density function for x is

$$f(x) = \frac{1}{15} e^{-x/15}$$

We will see what is the how to construct the exponential distribution. Suppose that x represents the loading time for a truck at a loading dock and follow such a distribution. If the mean or

average loading time is 15 minutes, $\mu = 15$, then appropriate probability density function is f of $x = 1$ divided by $15 e$ to the power $-x$ divided by 15 .

(Refer Slide Time: 23:13)



So, what is the meaning of this exponential distribution is supposed, if the loading time, the probability of loading time less than 6 is this area. The probability of loading time between 6 and 18 is this one this time, this is a probability, this shadowed portion represents the probability of that loading time is between 6 and 18.

(Refer Slide Time: 22:39)

**Exponential Probability Distribution
Cumulative Probabilities**

$$\underline{P(x \leq x_0) = 1 - e^{-x_0/\mu}}$$

where:
 x_0 = some specific value of x

Because in many application of exponential distribution will use the cumulative probability density function. So, for an exponential distribution, the formula for finding the cumulative

probability density function = $1 - e^{(-x_0 / \mu)}$ where x_0 is some specific value of x . it is nothing but if you integrate that to distribution, if you integrate the distribution between the intervals when you simplify it, you will get this answer.

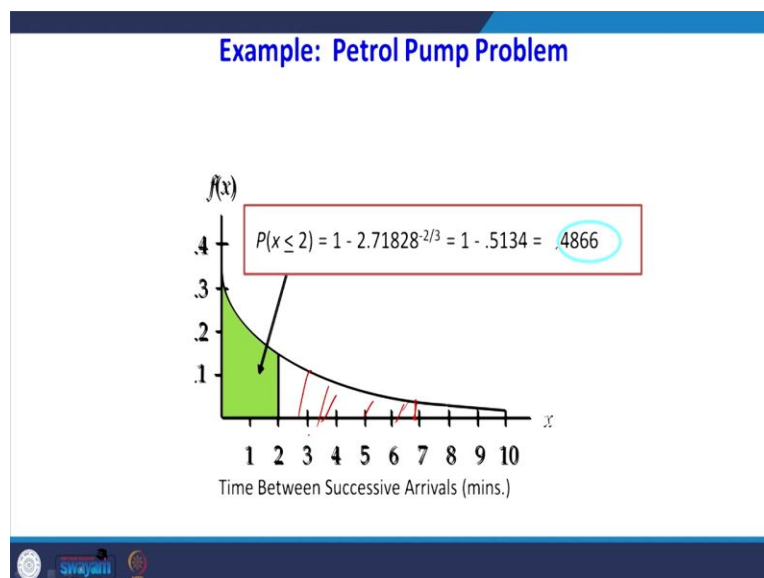
(Refer Slide Time: 24:08)

Example: Exponential Probability Distribution

- The time between arrivals of cars at a Petrol pump follows an exponential probability distribution with a mean time between arrivals of 3 minutes.
- The Petrol pump owner would like to know the probability that the time between two successive arrivals will be 2 minutes or less.

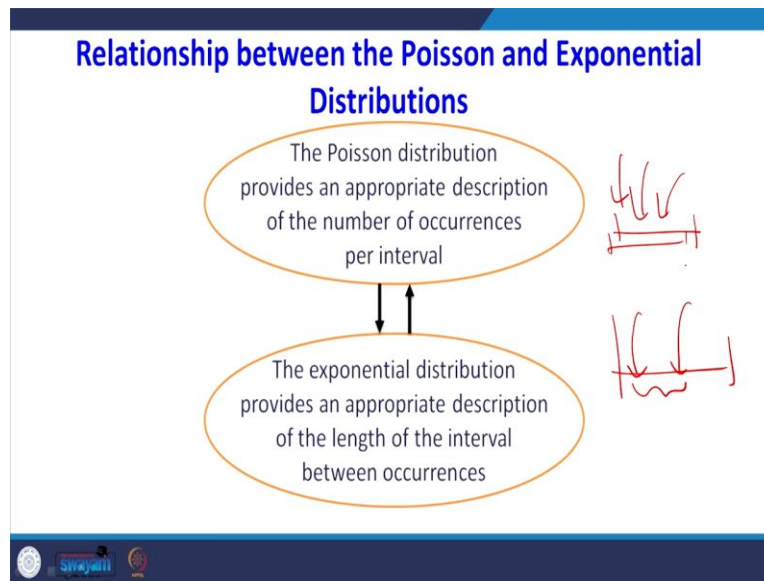
We will see one example of this exponential probability distribution. The time between arrivals of cars at a petrol pump follows an exponential probability distribution with the mean time between arrivals of 3 minutes. See that it is mean time between arrivals. So, the mean is 3 here. The petrol pump owner would like to know, the probability that time between 2 successive arrivals will be 2 minutes or less. So x value is less than or equal to 2.

(Refer Slide Time: 24:43)



So if you want to know P of x less than equal 2, we have to substitute into the, that distribution. That is the cumulative to distribution function, $1 - x$ value, so it is 0.4866. It is a 0.4866 so this one. So this shaded one. Suppose if you want to find out, the time between 2 successive arrivals of vehicle is less than 7 minutes, probability will increase. So, when x increases, but there are more chances that the time between 2 success arrival is this much. So what I am saying this is a way to interpret the exponential distribution.

(Refer Slide Time: 25:23)



Now, there is a very important relationship between poisson and exponential distribution. See, the Poisson distribution provides an appropriate description of number of occurrences per interval. So, one interval is there. In that interval, how many occurrences happened, that is the Poisson distribution. In the exponential distribution provides an appropriate description of length of the interval between the occurrences.

Suppose the same interval between this occurrence to this occurrences, so this phenomena is explained with the help of between 2 occurrences, this phenomena is explained with the help of exponential distribution. Number of occurrences in that interval it is explained with the help of Poisson distribution, that is a relation between poisson and exponential distribution.

(Refer Slide Time: 26:14)

Mean of Poisson and Mean of Exponential Distributions

- Because the average number of arrivals is 10 cars per hour, the average time between cars arriving is

$$\frac{1 \text{ hour}}{10 \text{ cars}} = .1 \text{ hour/car}$$

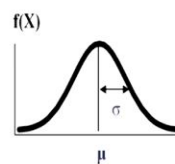
10
 $\mu \rightarrow$ mean for Poisson
 $1/\mu \rightarrow$ mean for Exp
 μ

So, mean of your poisson and mean of your exponential distribution, what is the relationship. Because the average number of arrival is 10 cars per hour, the average. So, if it is a 10 in that interval, there are 10 cars are arriving. So, this mean is taken for the Poisson distribution. But 1 / 10 is taken as the mean for the exponential distribution. That means, time between arrivals. So, that is a link between μ generally we write μ mean for Poisson distribution, $1 / \mu$ mean for exponential distribution that is the relation between poisson and exponential distribution.

(Refer Slide Time: 27:05)

The Normal Distribution: Properties

- 'Bell Shaped'
- Symmetrical
- Mean, Median and Mode are equal
- Location is characterized by the mean, μ
- Spread is characterized by the standard deviation, σ
- The random variable has an infinite theoretical range: $-\infty$ to $+\infty$



Mean = Median = Mode

Next, we are entering into the very important distribution that is a normal distribution, normal distribution is we can say a father of all the distributions. Because suppose some phenomena is happening if you are not aware, you can assume that it follows normal distribution. Normal

distribution is following Bell Shaped, it is symmetrical, mean, median and modes are equal. The location of the normal distribution is characterized by μ , the spread is characterised by σ . The random variable has an infinite theoretical range that is minus infinity to plus infinity.

(Refer Slide Time: 27:50)

The Normal Distribution: Density Function

The formula for the normal probability density function is

$$f(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{(X-\mu)}{\sigma}\right)^2}$$

Where

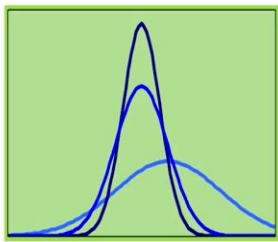
- e = the mathematical constant approximated by 2.71828
- π = the mathematical constant approximated by 3.14159
- μ = the population mean
- σ = the population standard deviation
- X = any value of the continuous variable

Chap 6-43

The density function of a normal distribution is $f(X) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma}} \cdot e^{(-1/2)((x - \mu)/\sigma)^2}$, e is mathematical constant value is 2.71, π is the mathematic constant we know that 3.14, μ is the population mean, σ is the population standard deviation, X is any value of the continuous variable.

(Refer Slide Time: 28:18)

The Normal Distribution: Shape



By varying the parameters μ and σ , we obtain different normal distributions

Chap 6-44

The shape of the normal distribution is by varying the parameter of μ and σ we obtained different normal distributions. Dear students, what we have seen so far is we have seen some of the discrete and continuous distributions in the discrete distribution we have talked about the binomial, Poisson distribution. In the continuous distribution we have seen exponential and uniform distributions. The next class very important distribution that is normal distribution, that will cover in the next class. Thank you very much.