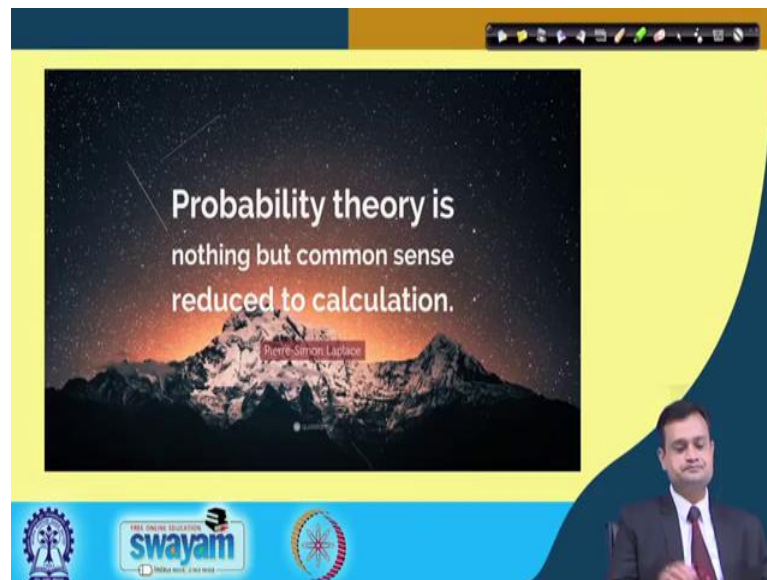


Six Sigma
Prof. Jitesh J Thakkar
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture – 23
Probability Theory

Hello friends. I welcome you to the lecture 24 of our Six Sigma journey. And we will discuss probability theory as a part of this lecture, so I would like to remind you that we are discussing the major phase of DMAIC, Six sigma cycle and we are discussing the various topics specific to this.

(Refer Slide Time: 00:44)



So, let us begin this lecture with a beautiful quote given by a great mathematician and researcher Laplace. So, he says probability theory is nothing, but common sense reduced to calculation.

So, if it is a common sense, is it really required. I would say yes. If you are exposed with a phenomena, which is occurring randomly and it is difficult for you to converge to a particular decision then definitely you need a tool or a technique like statistics to make better decisions. Suppose you make the forecast about the rain season you make forecast about the demand and you try to say control the quality of your process through certain statistical measures. So, in all the cases statistics has a vital role to play and it enables the

intuitive judgmental capability of the managers, executives and they can take better informed decisions.

(Refer Slide Time: 02:07)



A presentation slide titled "Recap" with a yellow background and a dark blue sidebar on the left. The slide lists seven topics in a bulleted format, each preceded by a red square icon. At the bottom, there are logos for Swayam and other educational institutions, along with a small video inset of a man in a suit.

Recap

- ❑ Basics of Statistics
- ❑ Descriptive v/s Inferential Statistics
- ❑ Measures of Central Tendency and Dispersion
- ❑ Shape of a Distribution
- ❑ Covariance and Coefficient of Correlation
- ❑ Numerical Descriptive Measures for a Population
- ❑ The Central Limit Theorem

So, we can see that as a part of recap, we have discussed in the previous lecture basics of statistic, descriptive versus inferential, measures of central tendency, relationship, covariance, correlation and numerical descriptive measures and the central limit theorem. I have also introduced the concept of random number.

(Refer Slide Time: 02:31)



A presentation slide titled "CONCEPTS COVERED" with a yellow background and a dark blue sidebar on the left. The slide lists six concepts in a bulleted format, each preceded by a red square icon. At the bottom, there are logos for Swayam and other educational institutions, along with a small video inset of a man in a suit.

CONCEPTS COVERED

Concepts Covered:

- ❑ Basic concepts of Probability
- ❑ Complementary Rule of Probability
- ❑ Conditional Probability
- ❑ Mutually Exclusive Events
- ❑ Probability Distributions

Now, in this particular lecture we will talk about basic concepts of the probability, complementary rule of probability, conditional probability, mutually exclusive event and probability distributions.

(Refer Slide Time: 02:46)

What is Probability Theory?

It is a branch of mathematics concerned with the analysis of random phenomena.

The outcome of a random event cannot be determined before it occurs, but it may be any one of several possible outcomes.

The actual outcome is considered to be determined by chance.

So, the question is that what is probability theory. It is important, but what exactly we mean by probability theory. So, many phenomena they are random in nature and they occur by chance. So, typically statistics is a branch of mathematics which is concerned with the analysis of random phenomena and that can help us to draw some conclusion about the problem under investigation or that phenomena itself.

(Refer Slide Time: 03:22)

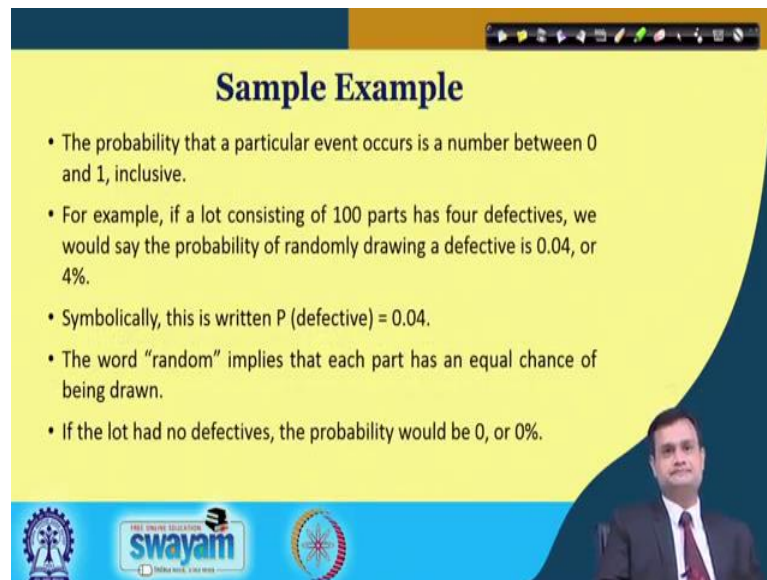
Basic Concepts

- **Classic definition:** If an event, say A, can occur in m ways out of a possible n equally likely ways, then
$$P(A) = \frac{m}{n}$$
- **Relative-frequency definition:** The proportion that an event, say A, occurs in a series of repeated experiments as the number of those experiments approaches infinity. According to this definition, probability is a limiting value that may or may not be possible to quantify.

Now, just see the basic concepts that what is the classic definition of a probability. So, when I say probability of some event A then $P(A)$ is m/n and this is typically m ways out of n equally likely ways. So, there are n equal possible opportunities, out of that m times let us say something is happening then $P(A)=m/n$. In the context of our six sigma let us say I am getting say 10 defectives out of may be 1000, then $10/1000$ is the probability of getting a defective when I select a particular sample.

Now, there is a relative frequency definition also. So, it says that the proportion that an event, say A, occurs in a series of repeated experiments as the number of those experiments approaches infinity and according to this definition probability is a limiting value that may or may not be possible to quantify. So, it is little bit the extension of the classical definition, and typically it considers the repeated experiment, number of times and we keep the window open up to infinity.

(Refer Slide Time: 04:51)



Sample Example

- The probability that a particular event occurs is a number between 0 and 1, inclusive.
- For example, if a lot consisting of 100 parts has four defectives, we would say the probability of randomly drawing a defective is 0.04, or 4%.
- Symbolically, this is written $P(\text{defective}) = 0.04$.
- The word "random" implies that each part has an equal chance of being drawn.
- If the lot had no defectives, the probability would be 0, or 0%.

swayam
All India Council for Technical Education

Now, just see the example, sample example, that the probability that a particular event occur is obviously, between 0 to and 1 and your total probability will be 1.

Now, let us say for example, if a lot consisting of 100 parts has 4 defectives we would say that probability of randomly drawing a defective is 0.04. So, $4/100$ will give me the probability of getting or drawing a defective part. Now, symbolically I would write P in bracket defective is equal to 0.04, and the word random implies that each part has a equal opportunity for say chance of being drawn. If the lot had no defective the probability of getting a defective will be 0 or 0 %.

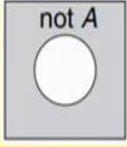
(Refer Slide Time: 05:53)

Example

- If the lot had 100 defectives, the probability would be 1, or 100%.
- The probability that event A will occur can be illustrated with a **Venn diagram**, as shown in Figure.



Venn diagram illustrating the probability of event A



Venn diagram illustrating the complementary rule of probability

swayam

Then, if you see some basic concepts on probability then here if a lot of 100 defective the probability would be 1, it means entire lot is defective. So, probability is 100 percent and the probability that event A will occur you can illustrate by a Venn diagram. You can see the first Venn diagram here.

(Refer Slide Time: 06:18)

Basic Concepts

- An experiment is the process by which an observation (or measurement) is obtained.
- An event is an outcome of an experiment, usually denoted by a capital letter.
 - The basic element to which probability is applied
 - When an experiment is performed, a particular event either happens, or it doesn't!
- Experiment: Record an age
 - A: person is 30 years old
 - B: person is older than 65
- Experiment: Toss a die
 - A: observe an odd number
 - B: observe a number greater than 2

swayam

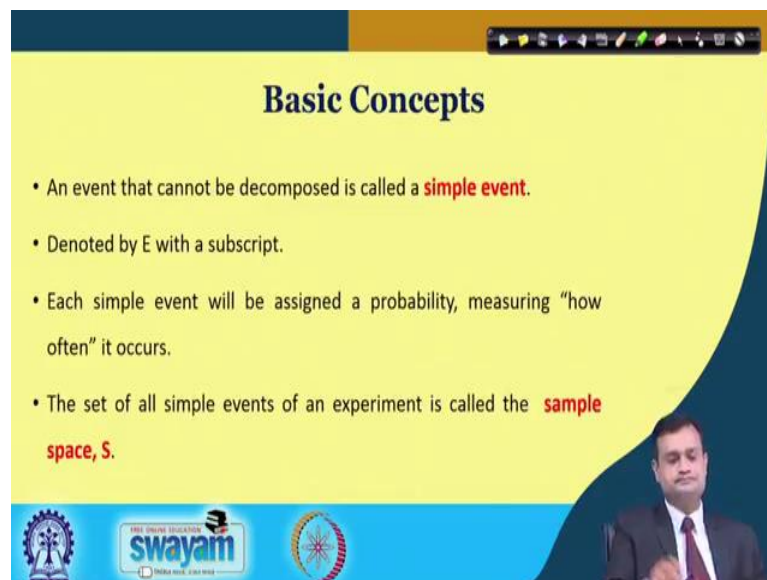
You can see the first Venn diagram here, and this illustrates the probability of event A. Now, if you see the second Venn diagram here it typically illustrates the complementary rule of probability. So, the area shaded in grey, grey color says that not A, it means that

this is the probability that A will not occur and the circle if I consider that as A than that is the probability that A will occur.

Now, suppose you are conducting an experiment and typically you are analyzing a particular manufacturing process which an observation or measurement you are trying to take. And an event is an outcome of this particular experiment, so the event may be the defective parts or defects produced, so the basic element to which probability is applied. So, this particular event is typically denoted by a capital letter. And when an experiment is performed a particular event either happens or it does not happen. So, it is a complementary. It happens or it does not happen.

So, for experiment like record an age. So, A, person is 30 years old; B, person is older than 65, I want to find the probability. Toss a die, so this is a coin also or toss a die then observe an odd number observe a number greater than 2. So, I would be interested in finding the probability of occurrences that how many times I can observe the odd number or I can observe the number greater than 2.

(Refer Slide Time: 08:29)



Basic Concepts

- An event that cannot be decomposed is called a **simple event**.
- Denoted by E with a subscript.
- Each simple event will be assigned a probability, measuring “how often” it occurs.
- The set of all simple events of an experiment is called the **sample space, S**.

At the bottom of the slide, there are logos for Swamyam and other educational institutions, and a small video inset of a man in a suit.

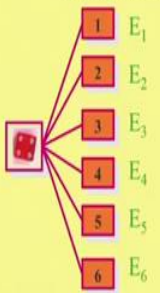
So, an event that cannot be decomposed is called simple event. It is typically denoted by E with a subscript. Each simple event will be assigned a probability measuring how often it occurs. And the set of all simple event of an experiment is called the sample space denoted as capital S.

(Refer Slide Time: 08:56)

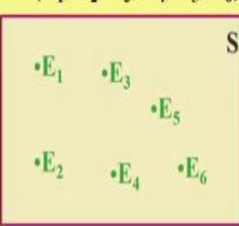
Example

• The die toss:

Simple events:



Sample space:

$$S = \{E_1, E_2, E_3, E_4, E_5, E_6\}$$


The slide features a yellow background with a blue header and footer. The footer includes the Swayam logo and the text 'FREE ONLINE EDUCATION swayam'.

Just see the example that would be make the concept better clear. Suppose I am tossing the die, rolling the die then my simple event I may get say 1 as a number that is E_1 event 1; E_2 , 2 ; E_3 number 3 and so on. So, my sample space would include all the possible outcomes E_1, E_2, E_3, E_4, E_5 and E_6 . And I am just representing this in the box so this becomes my sample space. So, all the events comprised in a space typical space is my sample space.

(Refer Slide Time: 09:41)

Basic Concepts

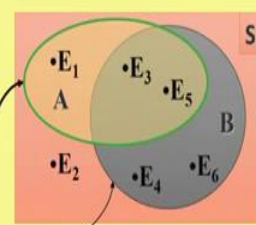
• An **event** is a collection of one or more **simple events**.

• The die toss:

- A: an odd number
- B: a number > 2

$A = \{E_1, E_3, E_5\}$

$B = \{E_3, E_4, E_5, E_6\}$



The slide features a yellow background with a blue header and footer. The footer includes the Swayam logo and the text 'FREE ONLINE EDUCATION swayam'.

Now, just see the little bit extended version of the example. Suppose I consider event A as odd number and event B a number greater than 2. Now, in your sample space which is typically shown as pink color you have all the possible events $E_1, E_2, E_3, E_4, E_5, E_6$, but when I put a boundary with a condition that A would have only odd number than E_1, E_3 and E_5 . So, my subset of the sample space A would be E_1, E_3, E_5 . Similar way B would comprise all the number greater than 2, so it would be E_3, E_4, E_5 , and E_6 .

(Refer Slide Time: 10:28)

Example

Toss a fair coin twice. What is the probability of observing at least one head?

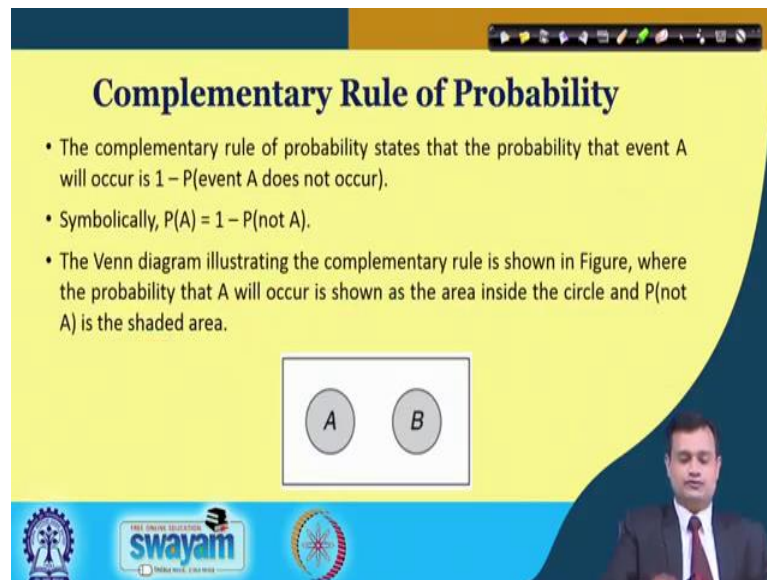
1st Coin	2nd Coin	E_i	$P(E_i)$
H	H	HH	$1/4$
	T	HT	$1/4$
T	H	TH	$1/4$
	T	TT	$1/4$

$P(\text{at least 1 head})$
 $= P(E_1) + P(E_2) + P(E_3)$
 $= 1/4 + 1/4 + 1/4 = 3/4$

Now, I will just explain the other example that is tossing a coin I may get head or tail. So, if you toss the coin, first coin then I will get head or tail. So, if there is a head and if I coin second time then I will get either head or tail. Similar way first time if I got tail and I coin it second time I may get head or tail. So, the probability I would be interested to find that how many times or two times I will get head 1 time head 1 time tail, 1 time tail 1 time head and 2 times tail.

So, you can easily find it that probability of at least one head is $P(E_1) + P(E_2) + P(E_3)$ and each individual will have equal probability $1/4, 1/4, 1/4, 1/4$ and when I sum it up this $P(E_1), P(E_2), P(E_3)$ I cannot consider the 4th one that is TT because there is no head, so my probability of getting at least one head would be $3/4$.

(Refer Slide Time: 11:51)



Complementary Rule of Probability

- The complementary rule of probability states that the probability that event A will occur is $1 - P(\text{event A does not occur})$.
- Symbolically, $P(A) = 1 - P(\text{not } A)$.
- The Venn diagram illustrating the complementary rule is shown in Figure, where the probability that A will occur is shown as the area inside the circle and $P(\text{not } A)$ is the shaded area.

The slide features a yellow background with a dark blue header and footer. A Venn diagram with two circles, A and B, is shown in the center. The footer includes logos for Swayam and other educational institutions, along with a small video inset of a man in a suit.

Now, there is a complementary rule of probability which is very important and this says that the probability that event A will occur is 1 minus probability that event A does not occur. So, simply if I ask what is the probability that today it will rain, you will say if I know the probability that today it will not rain then total probability of any event cannot be more than 1. So, 1 minus the probability of not happening, the event will be your probability of happening the event.

So, typically it is written as $P(A) = 1 - P(\text{not } A)$ and if you see the Venn diagram then complementary rule is shown in this figure where the probability that A will occur is shown as the area inside the circle and probability not A is shaded area. So, this can help us to appreciate the complementary rule.

(Refer Slide Time: 13:01)

Contingency Tables

- Contingency tables provide an effective way of arranging attribute data while allowing us to readily determine relevant probabilities.
- This statement is best illustrated by the following scenario: Suppose each part in a lot is one of four colors (red, yellow, green, or blue) and one of three sizes (small, medium, or large).

Size	Red	Yellow	Green	Blue	Totals
Small	16	21	14	19	70
Medium	12	11	19	15	57
Large	18	12	21	14	65
Totals	46	44	54	48	192

Now, there is another important representation of the data in the form of contingency table, and contingency table will basically represent the number or the frequency of some phenomena occurring.

So, for example, here I have a table and I say that I am trying to create the different scenario and suppose I have a lot which contains the 4 color and it is a 3 size. So, you may consider that it is a lot of ball and there are 3 different colored sizes of ball and 4 different colors. I have small, medium and large, red, yellow, green, blue and I can find the total in the last column as well as in the last row.

Now, once you have represented the data in the form of contingency you can very well compute the probabilities of various events.

(Refer Slide Time: 14:03)

Examples	
Example	Solutions
Find the probability that a part selected at random is red:	$P(\text{red}) = \frac{46}{192} = 0.24$
Find the probability that a part selected at random is small:	$P(\text{small}) = \frac{70}{192} = 0.365$
Find the probability that a part selected at random is both red and small:	$P(\text{red and small}) = P(\text{red} \cap \text{small})$ $= \frac{16}{192} = 0.083$
Find the probability that a part selected at random is red or small:	$P(\text{red or small}) = P(\text{red}) + P(\text{small}) - P(\text{red and small})$ $P(\text{red or small}) = \frac{46}{192} + \frac{70}{192} - \frac{16}{192}$ $P(\text{red or small}) = 0.521$

So, let us see that suppose I am interested to find the probability of red. It means find the probability that a part selected at random is red, everything is mixed. So, probability of red is equal to I can go back you just see this particular number that is 46. So, this my total red. So, I want to find what is the probability that red can occur. So, this is 46 and if you see the computation I have 46/192, so 192 is the total number. So, what is the probability that I will get a red is 46/192.

Similar way what is the probability that a part selected at random is small respective of the color. So, I will say small is basically 70 here and my total is 192. So, I can take the ratio of this two and it would be 0.365. You can further analyze that find the probability that part selected at random is both red and small, fine, this is my interest particular phenomena of interest and I would say that probability red intersection small means both should occur red and small, so 16/192 and same way you can find the probability that part selected at random may be red or may be small. So, you will say probability of red or small basically expressed as $P(\text{red}) + P(\text{small}) - P(\text{red and small})$. So, you can just say plug in the values and find the probability of red or small.

(Refer Slide Time: 16:02)

Conditional Probability

- Concern the odds of one event occurring, given that another event has occurred

$$P(A|B) = \frac{P(A \cap B)}{P(A)}$$

Example on Conditional Probability	Solutions
$P(\text{large} \text{green})$	$P(\text{green}) = \frac{54}{192} = 0.281$ $P(\text{green and large}) = P(\text{green} \cap \text{large}) = \frac{21}{192} = 0.109$ Therefore $P(\text{large} \text{green}) = \frac{0.109}{0.281} = 0.389$
$P(\text{small} \text{red})$	$P(\text{small} \text{red}) = \frac{P(\text{small} \cap \text{red})}{P(\text{red})} = \frac{16/192}{46/192} = 0.348$

Now, many a times say events are not occurring independently and we may have some conditions. So, under this conditions when we are analyzing and we want to find the probability we basically take the help of conditional probability rules.

So, let us say I say that probability that event A will occur given that event B will occur, so this can be expressed as $\frac{P(A \cap B)}{P(A)}$. So, I can just continue with my example of say 3 different colored balls and 4 may be 3 different sizes and 4 different colored balls. So, probability that I get a larger ball given that it is green, then I can plug in the values in this particular expression. So, I know what is the probability of green. Total green is 54, total number of say balls or parts are 192. So, this probability of green is 0.281.

I can find probability of green and say large. So, probability green intersection large that would be 21/192, so it would be 0.109 when I put it in the expression so my numerator will be 0.109, denominator is 0.281, I will get 0.389. Similar way you can find the probability of small given red and you will get the conditional probability.

(Refer Slide Time: 17:41)

Independent and Dependent Events

- We can redefine independence in terms of conditional probabilities:

Two events A and B are **independent** if and only if

$$P(A|B) = P(A) \quad \text{or} \quad P(B|A) = P(B)$$

Otherwise, they are **dependent**.



Now, in probability you cannot always have dependent event, independent event both the possibilities exist. So, how to express it? So, suppose event A and B are independent and if I have to express it then I would say they are only independent if the conditional probability, $P(A|B) = P(A)$, means there is no effect of B when I am finding the probability of A, so this is my say case of independent events. Similar way you can say $P(B|A) = P(B)$, otherwise if this is not true, they are dependent.

(Refer Slide Time: 18:31)

Example

- From the table:

$$P(\text{small}|\text{red}) = \frac{P(\text{small} \cap \text{red})}{P(\text{red})} = \frac{16/192}{46/192} = 0.348$$
$$P(\text{small}) = \frac{70}{192} = 0.365$$

Size	Red	Yellow	Green	Blue	Totals
Small	16	21	14	19	70
Medium	12	11	19	15	57
Large	18	12	21	14	65
Totals	46	44	54	48	192

- So

$$P(\text{small}|\text{red}) \neq P(\text{small})$$

- Therefore, we can conclude that $P(\text{small})$ and $P(\text{red})$ are dependent events.



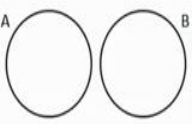
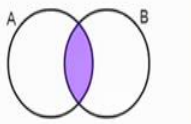
We can just see the example and that would be say interesting to clarify our understanding and strengthen it. So, let us say probability of small ball as I am considering a component or item as a ball given it is red I can find and the probability is 0.348. Suppose, I find probability of simply small I am not considering any given red as I did in the first part, so it would be 0.365.

Now, what is the conclusion that $P(\text{small}|\text{red})$ is not equal to probability of small and we can say that they are dependent events. So, we can conclude that probability of small and probability of red are dependent events and we cannot just consider them independent of each other.

(Refer Slide Time: 19:34)

Mutually Exclusive Events

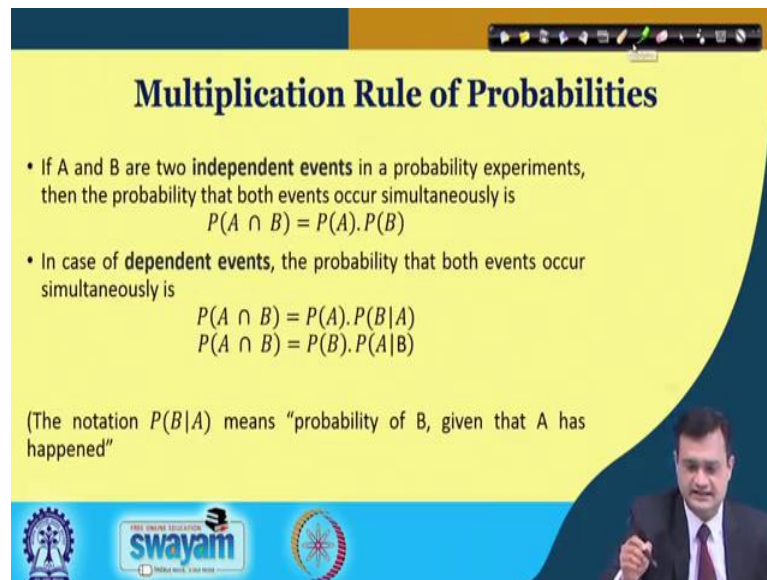
- Two events, A and B , are said to be mutually exclusive if both events cannot occur at the same time.

Mutually Exclusive Events	Non-Mutually Exclusive Events
	
$P(A \text{ or } B) = P(A) + P(B)$	$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$



Now, another dimension to express my event is to exclusiveness. So, you can have mutually exclusive event. You can see it very well in the first particular say figure that mutually exclusive event there is no common area and they are exclusive if both events cannot occur at the same time. So, $P(A \text{ or } B) = P(A) + P(B)$, and if you see the second one then in this case these are non-mutually exclusive event. So, $P(A \text{ or } B)$ is given by $P(A) + P(B) - P(A \text{ and } B)$.

(Refer Slide Time: 20:34)



Multiplication Rule of Probabilities

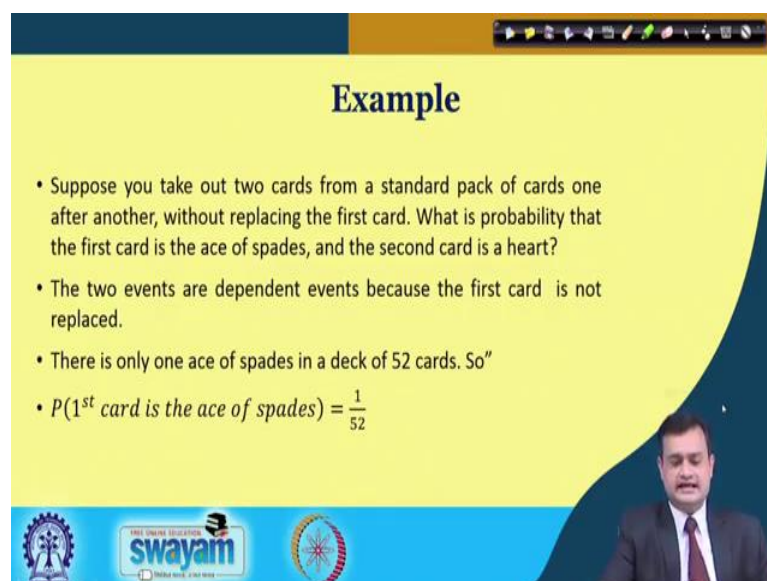
- If A and B are two **independent events** in a probability experiments, then the probability that both events occur simultaneously is
$$P(A \cap B) = P(A).P(B)$$
- In case of **dependent events**, the probability that both events occur simultaneously is
$$P(A \cap B) = P(A).P(B|A)$$
$$P(A \cap B) = P(B).P(A|B)$$

(The notation $P(B|A)$ means “probability of B, given that A has happened”)

swayam

Now, there is a multiplication rule of probabilities and if I say probability of two independent events it means probability if I want to find A intersection B then it would be $P(A).P(B)$ and here my consideration is that two events are independent of each other. When I say dependent event then the probability it depends on occurrence of both the events and $P(A \cap B) = P(A).P(B|A)$, and same way it can be expressed as $P(B).P(A|B)$. So, B by A or B given A means probability of B given that A has happened because they are dependent events.

(Refer Slide Time: 21:40)



Example

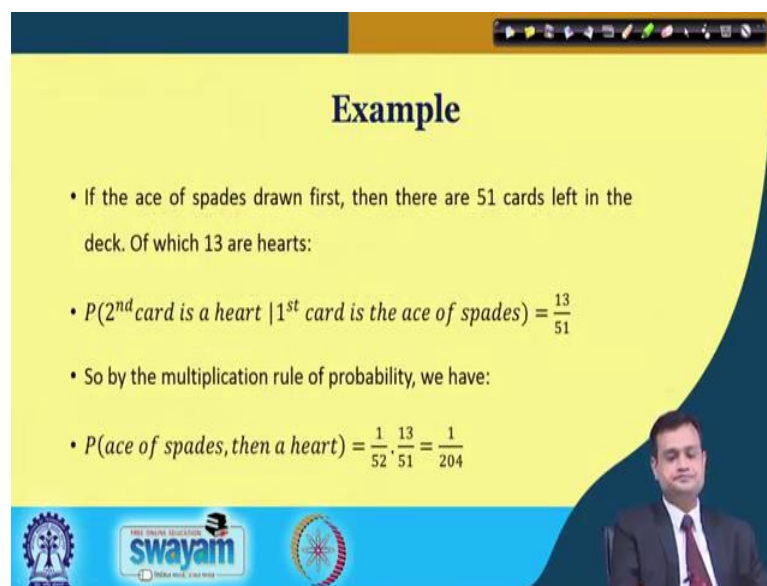
- Suppose you take out two cards from a standard pack of cards one after another, without replacing the first card. What is probability that the first card is the ace of spades, and the second card is a heart?
- The two events are dependent events because the first card is not replaced.
- There is only one ace of spades in a deck of 52 cards. So”
- $P(1^{st} \text{ card is the ace of spades}) = \frac{1}{52}$

swayam

We can see the example to appreciate the concept better. Suppose you take out two cards from a standard pack of cards one after another without replacing first card means I am not putting it back. So, what is the probability that the first card is ace of spades and the second card is heart?

So, your two events are dependent events because the first card is not replaced, please remember. So, there is only one ace of spades in a deck of 52 cards, so probability of first card is ace of spades is equal to $1/52$.

(Refer Slide Time: 22:25)



Example

- If the ace of spades drawn first, then there are 51 cards left in the deck. Of which 13 are hearts:
- $P(2^{nd} \text{ card is a heart} | 1^{st} \text{ card is the ace of spades}) = \frac{13}{51}$
- So by the multiplication rule of probability, we have:
- $P(\text{ace of spades, then a heart}) = \frac{1}{52} \cdot \frac{13}{51} = \frac{1}{204}$

Now, if I say if the ace of spades drawn first then there are 51 cards left because I am not putting it back. Now, of which 13 are hearts, so probability that second card is a heart given that first card ace of spades $13/51$, and I can say that the multiplication rule probability of ace of spade then heart is $1/52$ into $13/51$, so it would be $1/204$.

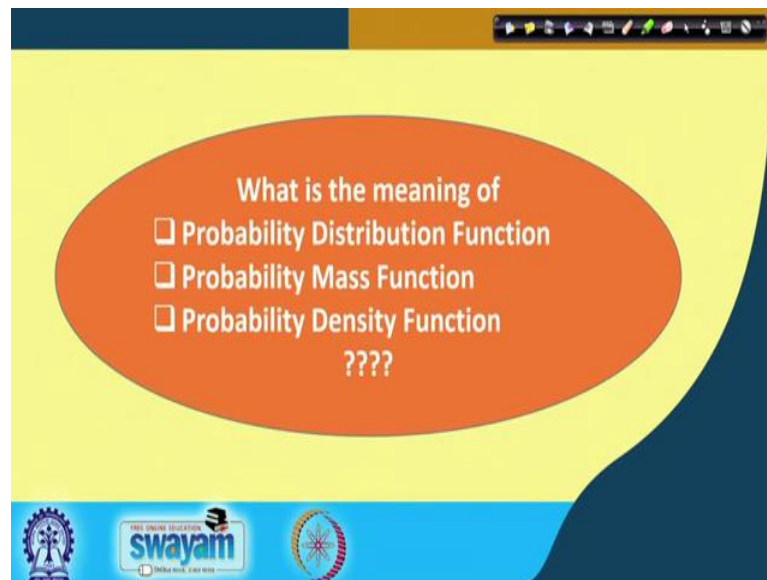
(Refer Slide Time: 23:02)

	Independent	Dependent
Events	$P(A B) = P(A)$ $P(B A) = P(B)$	$P(A B) \neq P(A)$ $P(B A) \neq P(B)$
Addition rule	$P(A \cup B) = P(A) + P(B)$	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Multiplication rule	$P(A \cap B) = P(A) \times P(B)$	$P(A \cap B) = P(A) \times P(B A)$ $= P(B) \times P(A B)$
	Mutually exclusive	Not mutually exclusive
	$P(A \cap B) = 0$	$P(A \cap B) \neq 0$

So, I would just like to present some summary of the various rules. So, if you have independent event and if you have dependent event and suppose you are talking about the event $P(A|B) = P(A)$, $P(B|A)$ when it is independent it is $P(B)$ but it is just reverse not equal to when it is dependent.

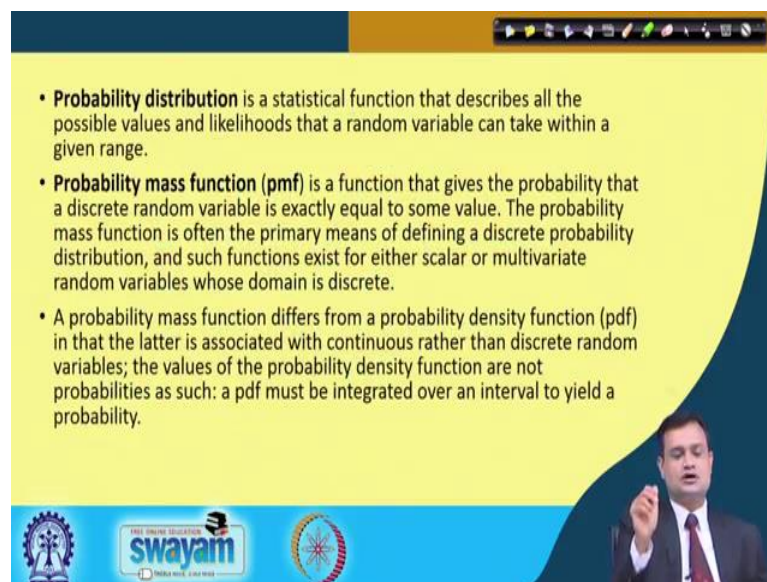
Addition rule, $P(A \cup B) = P(A) + P(B)$ when you look at the dependent event you have to subtract the $P(A \cap B)$ and there are multiplication rules for independent event as well as dependent event we have seen. Mutually exclusive event $P(A \cap B) = 0$ there is no commonality. $P(A \cap B) \neq 0$ they are not mutually exclusive.

(Refer Slide Time: 23:59)



Now, let us try to appreciate the couple of things. What is the meaning of probability distribution function? Probability mass function? Probability density function? And then subsequently we can have a quick overview of couple of distributions.

(Refer Slide Time: 24:22)



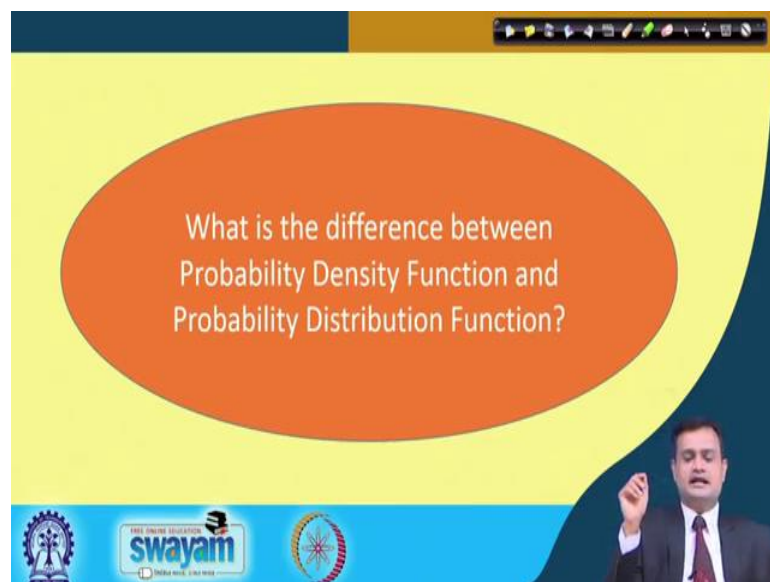
So, before that let us try to appreciate these 3 terms. So, probability distribution typically it is a statistical function that describes all the possible events and likelihood that a random variable can take within a given range. For example, suppose you want to see that what is the distribution of waiting time at particular bank ATM? For a specific

period of time you collect the data, plot the data with respect to time and then you will get an idea that how my data is distributed; in what way typically entire range of data is represented so, we have already seen that you can also get an idea when you put a histogram and skewness, kurtosis, those things can also help you. So, typically it is a mathematical function, statistical function that describes how a particular phenomena is occurring.

We have probability mass function, this gives a probability that a discrete. So, you have two types of variables one is discrete, other is continuous. If I have discrete random variable then it gives the probability that a discrete random variable is exactly equal to some value, and the probability mass function is often the primary means of defining the discrete probability distribution.

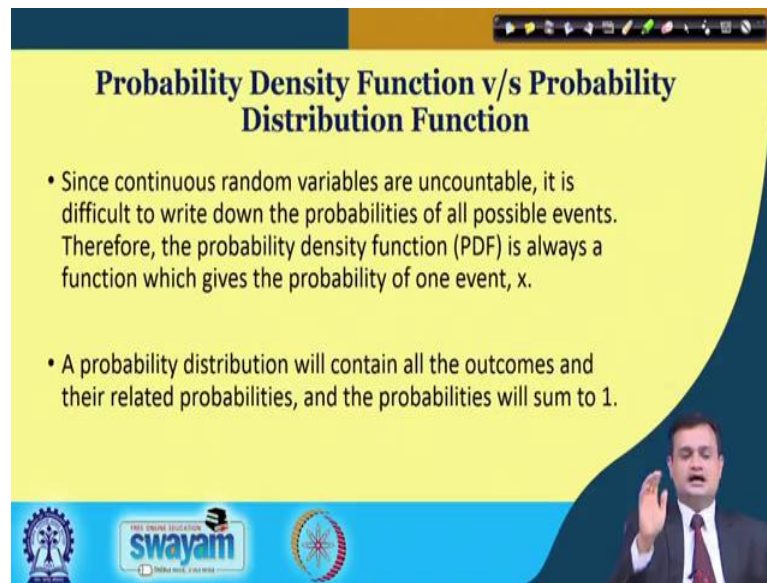
So, such function exist either scalar or multivariate random variable whose domain is discrete and typically probability mass function differs from probability density function. When I say a probability density function again it is for a specific value, but density function pertains to the continuous random variable.

(Refer Slide Time: 26:20)



So, typically what is the difference between probability distribution function and probability density function if I talk about a continuous random variable?

(Refer Slide Time: 26:32)



Probability Density Function v/s Probability Distribution Function

- Since continuous random variables are uncountable, it is difficult to write down the probabilities of all possible events. Therefore, the probability density function (PDF) is always a function which gives the probability of one event, x .
- A probability distribution will contain all the outcomes and their related probabilities, and the probabilities will sum to 1.

The slide features a yellow background with a dark blue curved border on the right. At the bottom, there is a blue banner with logos for 'swayam' and 'All India Open University'. A presenter in a suit is visible in the bottom right corner of the slide frame.

Now, the difference is like this, that since you have the continuous random variable and they are uncountable it is difficult to write down the probabilities of all the possible events and therefore, you take the help of probability density function, typically expressed in the form of mathematical expression and it is a function which gives the probability of a particular event, x . So, a probability distribution function will contain all the outcomes and their related probabilities and the probability will sum up to 2.

There are commonly used distribution I would quickly like to give you the idea. Broadly because we have two types of random variables, continuous and discrete, you have two types of distributions. Distribution again I am saying it is a way mathematical way of describing a particular phenomena which is occurring in a random fashion, and we have a family of distribution both under continuous and discrete. There is a huge say number of distributions available. We cannot discuss all, we are just trying to appreciate some of the widely used distributions here as continuous distribution and discrete.

(Refer Slide Time: 27:59)

Normal Distribution

- The normal distribution is the most important probability distribution in statistics because it fits many natural phenomena.
- For example, heights, blood pressure, measurement error, and IQ scores follow the normal distribution. It is also known as the Gaussian distribution and the bell curve.
- If X is a normal random variable with expected value μ (that is, $E(X) = \mu$) and finite variance σ^2 , then the random variable Z is defined as

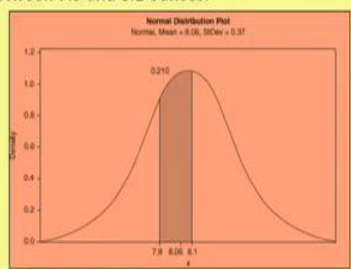
$$Z = \frac{x - \mu}{\sigma}$$


So, you have continuous distribution the most widely used one is the bell shaped distribution which is a normal distribution, and typically a standard normal distribution the critical value Z is expressed as $\frac{x - \mu}{\sigma}$. So, x is the individual value, μ is your population mean, σ is your population standard deviation.

(Refer Slide Time: 28:27)

Example

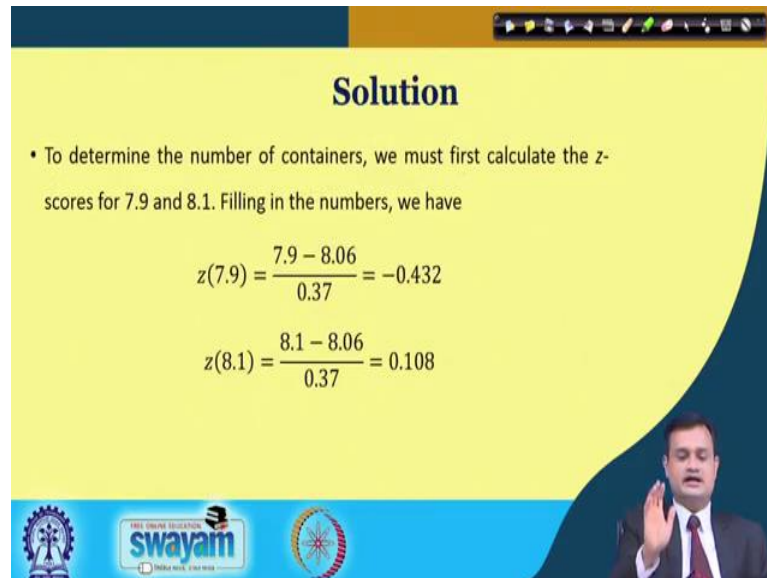
- A product fill operation produces net weights that are normally distributed. A random sample has mean 8.06 ounces and standard deviation 0.37 ounces, as illustrated in the following Figure. Estimate the percentage of the containers that have a net weight between 7.9 and 8.1 ounces.



Just see the example to appreciate the normal distribution. You have a product fill operation produces net weight that are normally distributed in a random sample has mean

8.06 ounces and the standard deviation 0.37 ounces as illustrated in the given figure. And this bell shaped curve also shows that your data has some variability.

(Refer Slide Time: 29:02)



Solution

- To determine the number of containers, we must first calculate the z-scores for 7.9 and 8.1. Filling in the numbers, we have

$$z(7.9) = \frac{7.9 - 8.06}{0.37} = -0.432$$
$$z(8.1) = \frac{8.1 - 8.06}{0.37} = 0.108$$

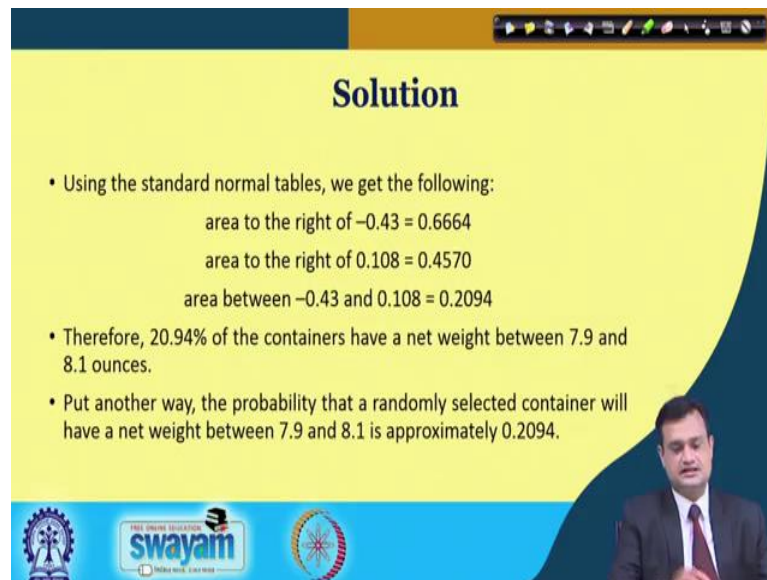
The slide also features a small video inset of a man in a suit in the bottom right corner and logos for Swayam and other institutions at the bottom.

So, now estimate the percentage of containers that have a net weight between 7.9 to 8.1 ounces.

So, I am using the expression of standard normal distribution $z(7.9)$, so $\frac{7.9-8.06}{0.37}$, this will give me the value minus 0.432. Similar way $z(8.1)$, I want to find the score between these two. So, I will get Z value 0.108. We have the tables for all the distributions typically called statistical tables, these tables are derived using the standard mathematical expression of the distribution typically called as probability distribution function. And this table documents the values of various parameters as well as the critical values, and it gives us the critical values as well the probability values for decision making.

So, these tables are available in any suggested book by me and you can easily find the probability values or critical values which can help you to make the inferences or the decisions.

(Refer Slide Time: 30:17)



Solution

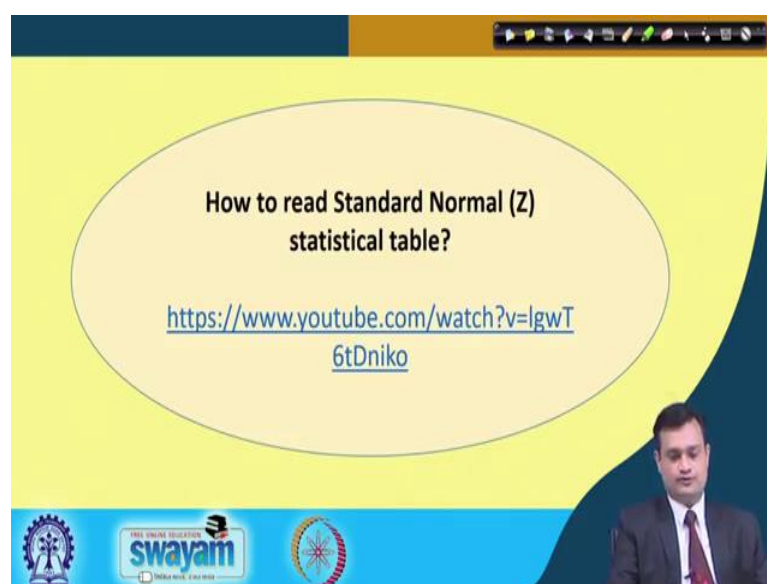
- Using the standard normal tables, we get the following:
area to the right of $-0.43 = 0.6664$
area to the right of $0.108 = 0.4570$
area between -0.43 and $0.108 = 0.2094$
- Therefore, 20.94% of the containers have a net weight between 7.9 and 8.1 ounces.
- Put another way, the probability that a randomly selected container will have a net weight between 7.9 and 8.1 is approximately 0.2094.

The slide features a yellow background with a blue header and footer. The footer includes logos for Swamyam and other educational institutions, along with a small video inset of a man in a suit.

So, now in case of my typical normal distribution I would like to find the area to the right of minus 0.43 and area to the right of 0.108, and from the table, I get this area 0.6664 with respect to the critical value and 0.4570. So, I will say that 20.94 percent of the containers have a net weight between 7.9 and 8.1 ounces.

So, another way I can say that the probability randomly selected container will have a net weight 7.9, 8.1 is approximately 20.94 percent or 0.2094.

(Refer Slide Time: 31:07)



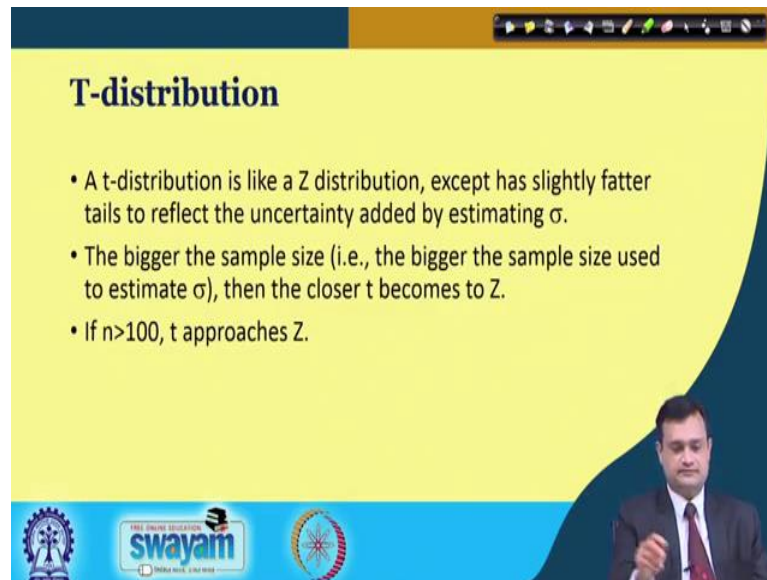
How to read Standard Normal (Z) statistical table?

<https://www.youtube.com/watch?v=lgwT6tDniko>

The slide features a yellow background with a blue header and footer. The footer includes logos for Swamyam and other educational institutions, along with a small video inset of a man in a suit.

You only need little bit practice to use the statistical table. I am suggesting here a very small video link. You can refer and just try to get acquainted with the use of statistical table.

(Refer Slide Time: 31:21)



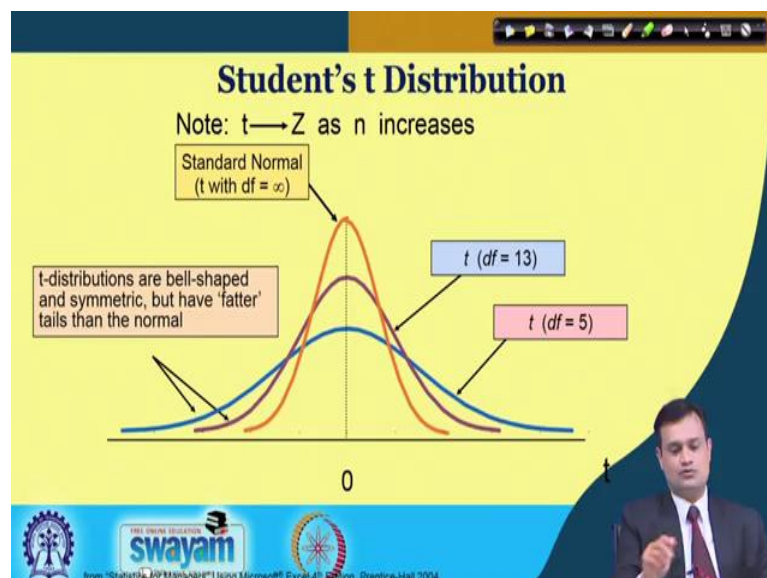
T-distribution

- A t-distribution is like a Z distribution, except has slightly fatter tails to reflect the uncertainty added by estimating σ .
- The bigger the sample size (i.e., the bigger the sample size used to estimate σ), then the closer t becomes to Z.
- If $n > 100$, t approaches Z.

The slide features a yellow background with a dark blue header and footer. The footer includes the Swayam logo and a small circular emblem. A presenter is visible in the bottom right corner.

Another important distribution T distribution or also called as the student T distribution, very much similar to Z distribution, but typically when the sample size is small you make use of T distribution when your n sample size is greater than 100 it almost approaches to the normal distribution.

(Refer Slide Time: 31:42)



So, you can see it very well that you have t with degree of freedom infinity, t with degree of freedom 13, t with degree of freedom 5. So, when you will look at the t distribution table you need to find the critical value with the help of degree of freedom, and already I explained what is degree of freedom. So, how much opportunity? How many opportunity you have to maneuver in the data set? Typically that is your degree of freedom.

(Refer Slide Time: 32:15)

How to read t statistical table?

<https://www.youtube.com/watch?v=lccjSiVTx7c>

The slide features a yellow background with a large white oval in the center containing the title and the YouTube link. At the bottom, there are logos for 'swayam' and 'All India Open University' on the left, and a small video feed of a man in a suit on the right.

So, again I am suggesting a link to refer the T statistical table and this you can do very easily just by going this particular reference.

(Refer Slide Time: 32:26)

F-Distribution

- The F distribution is the probability distribution associated with the f statistic.
- The **f statistic**, also known as an **f value**, is a random variable that has an F distribution.

Compute f statistic:

- Select a random sample of size n_1 from a normal population, having a standard deviation equal to σ_1 .
- Select an independent random sample of size n_2 from a normal population, having a standard deviation equal to σ_2 .
- The f statistic is the ratio of s_1^2/σ_1^2 and s_2^2/σ_2^2 .

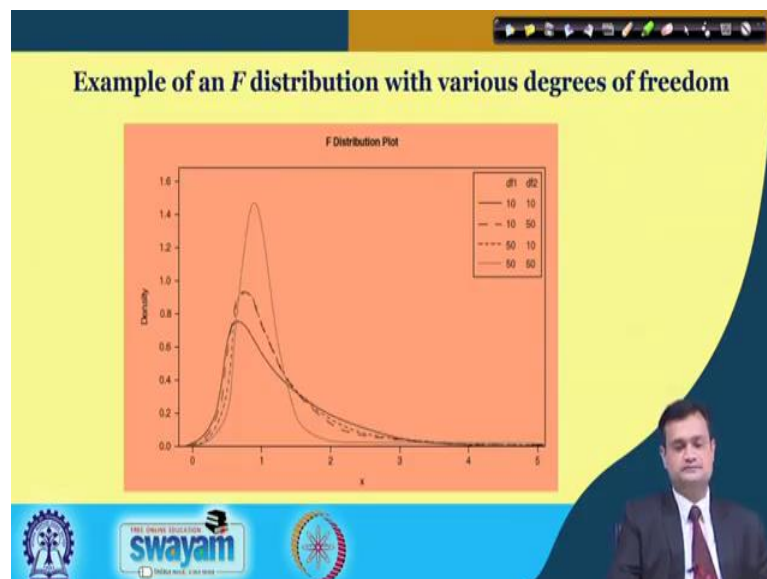
The distribution of all possible values of the f statistic is called an F distribution, with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ degrees of freedom.

The slide has a yellow background with a white oval for the title. Below the title, there are bullet points and a section for computing the f statistic. At the bottom, there are logos for 'swayam' and 'All India Open University' on the left, and a small video feed of a man in a suit on the right.

Very important distribution F distribution after your normal and T distribution, and then we will discuss the concept of ANOVA, F distribution would be widely used.

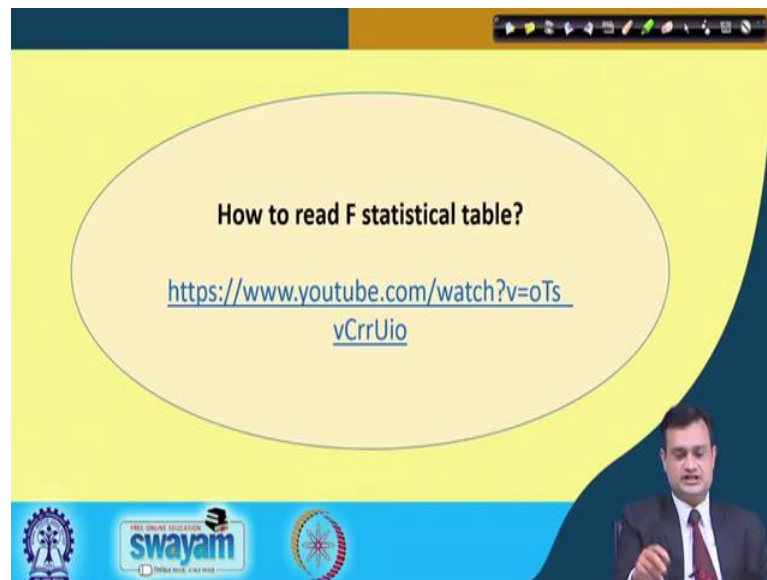
So, typically it is a probability distribution associated with F statistics and F statistics is basically the ratio of your sample variance and the population variance. So, you basically select a random sample of size n_1 from a normal population having a standard deviation σ_1 . Select an independent random sample again of n_2 having a population standard deviation σ_2 , and F statistic basically is a ratio of s_1^2/σ_1^2 and s_2^2/σ_2^2 and this distribution will have two degrees of freedom ν_1 and ν_2 , which is $n_1 - 1$ because you are taking two sample $n_2 - 1$.

(Refer Slide Time: 33:29)



So, you can just appreciate the shape of the F distribution, and it is not like minus infinity to infinity which we have in normal and for different degree of freedom you can see that the shape of F distribution changes.

(Refer Slide Time: 33:46)



How to read F statistical table?

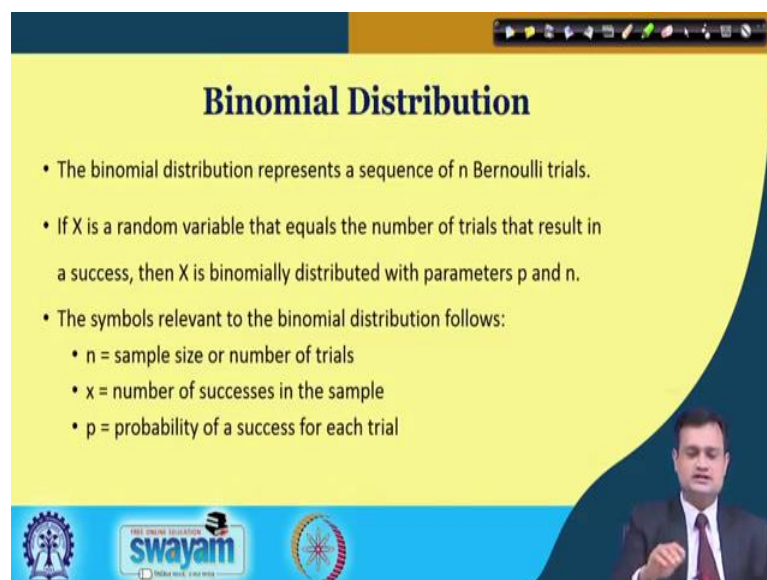
https://www.youtube.com/watch?v=oTs_vCrrUio

The slide features a yellow background with a large white oval in the center containing the title and the YouTube link. At the bottom, there is a blue banner with logos for 'swayam' and other educational institutions. A small inset video of a presenter is visible in the bottom right corner.

Further, you can just refer how to read the F distribution please open the appendix of the suggested book or download the statistical table and use the guidelines to understand how to read the statistical table.

Now, let us try to see quickly couple of discrete distribution.

(Refer Slide Time: 34:06)



Binomial Distribution

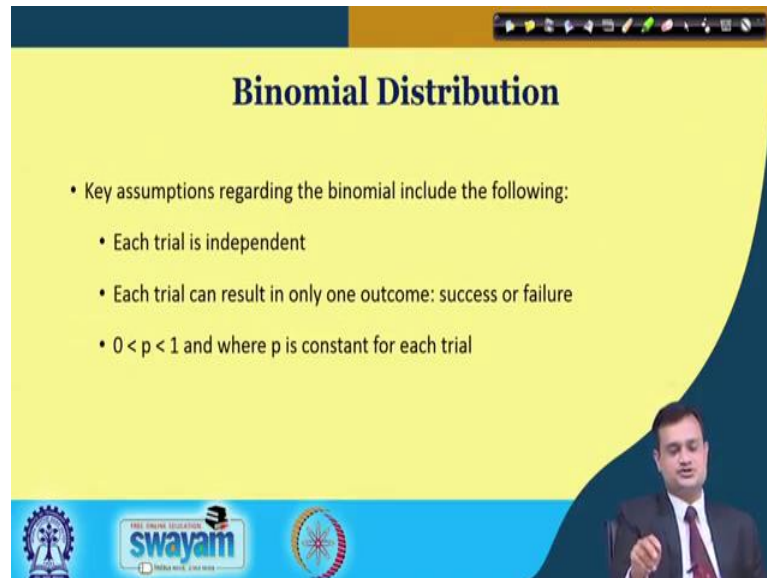
- The binomial distribution represents a sequence of n Bernoulli trials.
- If X is a random variable that equals the number of trials that result in a success, then X is binomially distributed with parameters p and n .
- The symbols relevant to the binomial distribution follows:
 - n = sample size or number of trials
 - x = number of successes in the sample
 - p = probability of a success for each trial

The slide has a yellow background with a blue banner at the bottom containing logos for 'swayam' and other educational institutions. A small inset video of a presenter is visible in the bottom right corner.

So, in quality domain one distribution which is widely used is the binomial distribution and typically it has two outcome success or failure, head or tail, and when you want to capture this kind of phenomena you try to use binomial distribution. The nomenclature it

has n is sample size, x is the number of successes in the sample and p is the probability of a success of each trail.

(Refer Slide Time: 34:34)



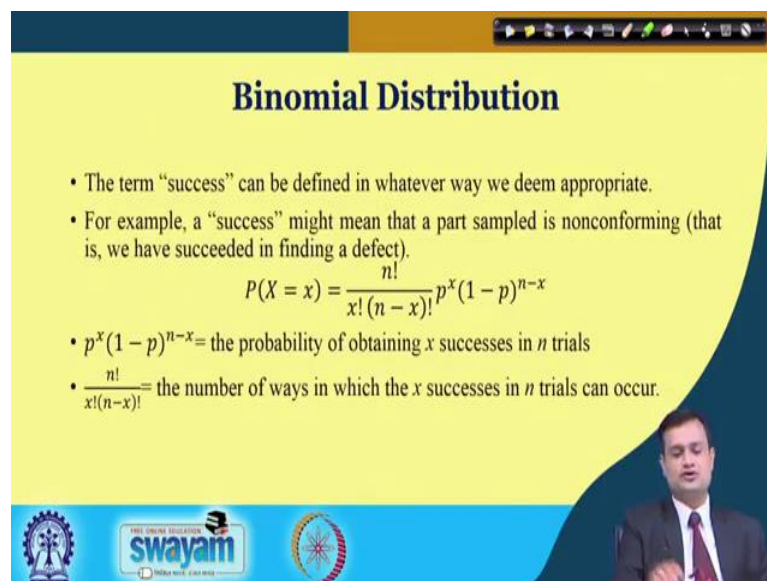
Binomial Distribution

- Key assumptions regarding the binomial include the following:
 - Each trial is independent
 - Each trial can result in only one outcome: success or failure
 - $0 < p < 1$ and where p is constant for each trial

The slide features a yellow background with a dark blue header and footer. The footer contains logos for 'swayam' and 'INDIA RISE, CHINA RISE' along with a small image of a person in a suit.

So, there are certain key assumptions in binomial distribution, each trial is independent, each trial can result in only outcome success or failure and obviously, your probability between 0 to 1.

(Refer Slide Time: 34:48)



Binomial Distribution

- The term “success” can be defined in whatever way we deem appropriate.
- For example, a “success” might mean that a part sampled is nonconforming (that is, we have succeeded in finding a defect).

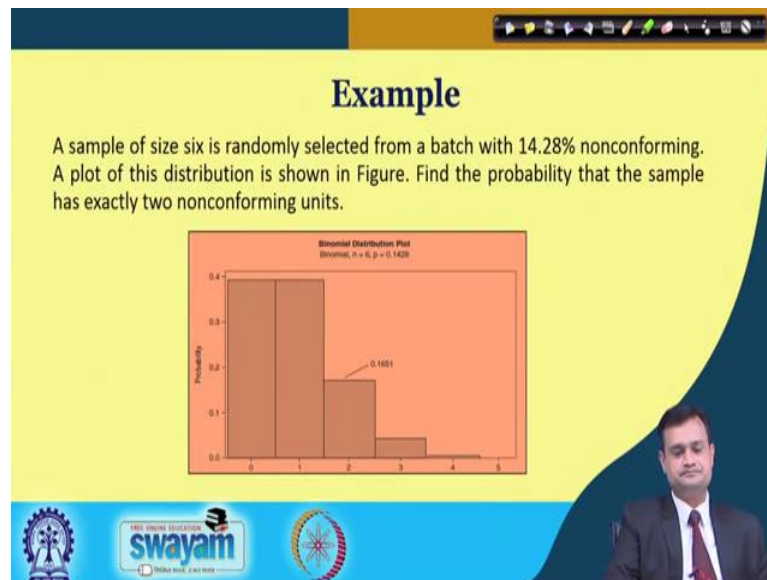
$$P(X = x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

- $p^x (1-p)^{n-x}$ = the probability of obtaining x successes in n trials
- $\frac{n!}{x!(n-x)!}$ = the number of ways in which the x successes in n trials can occur.

The slide features a yellow background with a dark blue header and footer. The footer contains logos for 'swayam' and 'INDIA RISE, CHINA RISE' along with a small image of a person in a suit.

So, typically the mathematical function describes the binomial distribution is $P(X = x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$

(Refer Slide Time: 35:09)



So, you have a sample size 6 is randomly selected from a batch with 14.28 percent nonconforming.

(Refer Slide Time: 35:23)

- From the problem statement, we know that

$$n = 6$$

$$x = 2$$

$$p = 0.1428$$

Filling in the numbers, we have

$$P(X = 2) = \frac{6!}{2!(6-2)!} 0.1428^2 (0.8572)^{6-2}$$

$$P(X = 2) = \frac{720}{(2)(24)} (0.0204)(0.5399)$$

$$P(X = 2) = (15) \frac{720}{(2)(24)} (0.0204)(0.5399)$$

$$P(X = 2) = 0.1651$$

- Thus, the probability that the sample contains exactly two nonconforming units is 0.1651.

And a plot of this distribution you can see here, and I am just trying to plug in the values in the distribution to give you an idea and what I get as a result for $P(X = 2) = 0.1651$. So, so probability that sample contains exactly two non-conforming units is 16.51.

(Refer Slide Time: 35:42)

Poisson Distribution

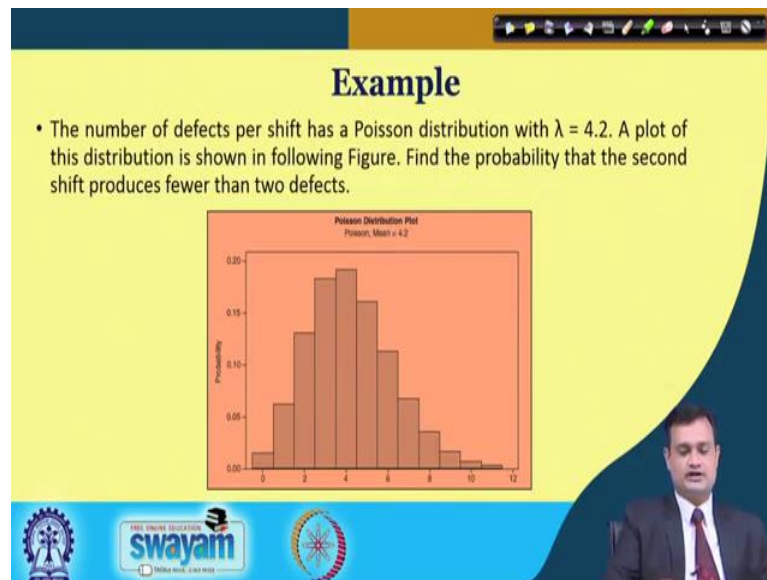
- it is used to model the number of defects per unit of time (or distance) and is the basis for the control limit formulas for the c and u control charts.
- The Poisson has also been used to analyze such diverse phenomena as the number of phone calls received per day and the number of alpha particles emitted per minute. The formula is
- $P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$ for $\lambda > 0$ and $x = 0, 1, 2, 3 \dots$
- The mean and variance of the Poisson distribution are both λ .

THE VEDIC COLLECTION
swayam
Vedika Group, 2014-2015

Another distribution in the family of discrete distribution is your Poisson distribution and typically it is used when you have to model the number of defects per unit of time and the basis for the control limit formulas for the c and u control chart which we will discuss later on.

So, Poisson has been used to analyze such diverse phenomena as the number of phone calls received per day and the number of may be alpha particles emitted per minute or sometimes let us say very simple day to day example that what is the occurrence of mutter in your vegetable pulao. So, then also it follows the Poisson distribution. So, $P(X = x)$ mathematical function which describes the Poisson distribution is $\frac{e^{-\lambda} \lambda^x}{x!}$.

(Refer Slide Time: 36:40)



So, we can see the example here lambda is given 4.2, it is a typical histogram depicting.

(Refer Slide Time: 36:50)

Solution

- Therefore, we will seek

$$P(X < 2) = P(X = 0) + P(X = 1)$$

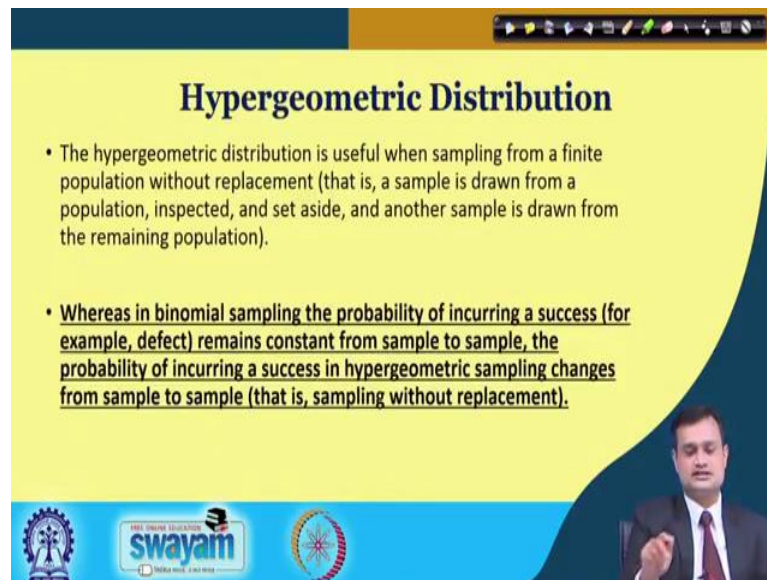
- Filling in the numbers, we have

$$P(X = 0) = \frac{e^{-4.2} 4.2^0}{0!} = 0.015$$
$$P(X = 1) = \frac{e^{-4.2} 4.2^1}{1!} = 0.063$$
$$P(X < 2) = 0.078$$

The slide includes logos for Anna University, Swayam, and the Ministry of Education, India, along with a video player interface at the top and a small inset of the presenter at the bottom right.

The Poisson distribution. I can just plug in the values to find the $P(X < 2)$ defective and I find that it comes out to be; so less than 2 means it could be 0 it could be 1, so when I sum it up 0.015 and 0.063, it is 0.078

(Refer Slide Time: 37:08)



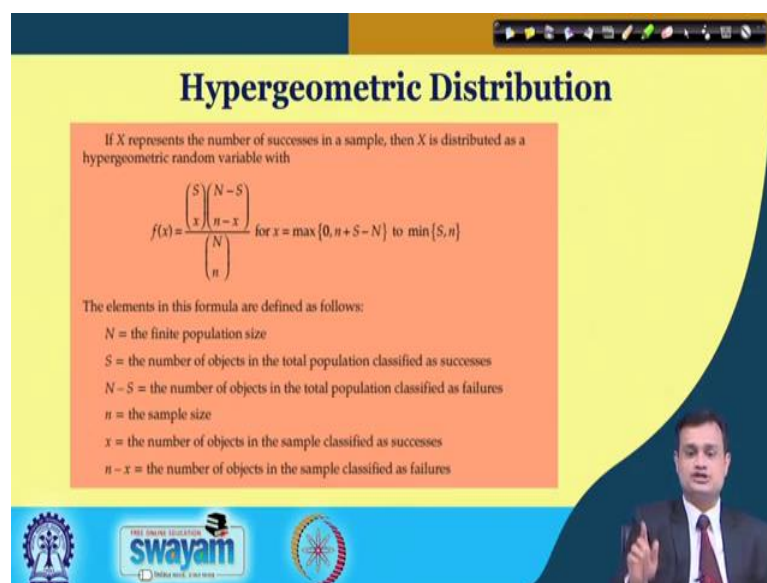
Hypergeometric Distribution

- The hypergeometric distribution is useful when sampling from a finite population without replacement (that is, a sample is drawn from a population, inspected, and set aside, and another sample is drawn from the remaining population).
- Whereas in binomial sampling the probability of incurring a success (for example, defect) remains constant from sample to sample, the probability of incurring a success in hypergeometric sampling changes from sample to sample (that is, sampling without replacement).

The slide features a yellow background with a blue header and footer. The footer includes the logos of the Indian Institute of Space Science and Technology (IIST) and the Swamyam logo, along with the text 'FREE ONLINE EDUCATION' and 'swamyam'.

Hypergeometric distribution, typically is another additional important distribution in the domain of quality and in binomial sampling the probability of incurring a success. For example, defect remains constant from sample to sample. The probability of incurring a success in Hypergeometric changes from sample to sample because this distribution is without replacement. So, I am taking a sample from a lot, I will not replace it and then once again I am taking the sample and checking. So, because of this without replacement phenomena here your probability of getting defective changes every time.

(Refer Slide Time: 37:54)



Hypergeometric Distribution

If X represents the number of successes in a sample, then X is distributed as a hypergeometric random variable with

$$f(x) = \frac{\binom{S}{x} \binom{N-S}{n-x}}{\binom{N}{n}} \text{ for } x = \max\{0, n+S-N\} \text{ to } \min\{S, n\}$$

The elements in this formula are defined as follows:

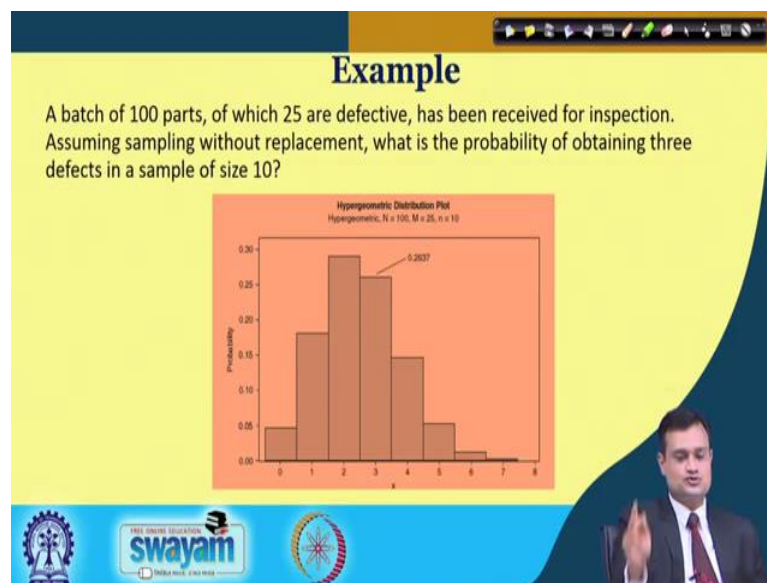
- N = the finite population size
- S = the number of objects in the total population classified as successes
- $N - S$ = the number of objects in the total population classified as failures
- n = the sample size
- x = the number of objects in the sample classified as successes
- $n - x$ = the number of objects in the sample classified as failures

The slide features a yellow background with a blue header and footer. The footer includes the logos of the Indian Institute of Space Science and Technology (IIST) and the Swamyam logo, along with the text 'FREE ONLINE EDUCATION' and 'swamyam'.

So, you have the mathematical function to describe $f(x) = \frac{\binom{s}{x} \binom{N-s}{n-x}}{\binom{N}{n}}$ and you can go through

n capital N is lot size, capital S is number of objects in the total population, N minus S is the number of objects in the total population classified as failures, small n is sample size and so on.

(Refer Slide Time: 38:22)



So, we can take the example in which we have 25 defectives and I want to find the probability of 3 defect in a sample of size 10.

(Refer Slide Time: 38:33)

N = 100
S = 25
N - S = 75
n = 10
x = 3
n - x = 7

Filling in the numbers, we have

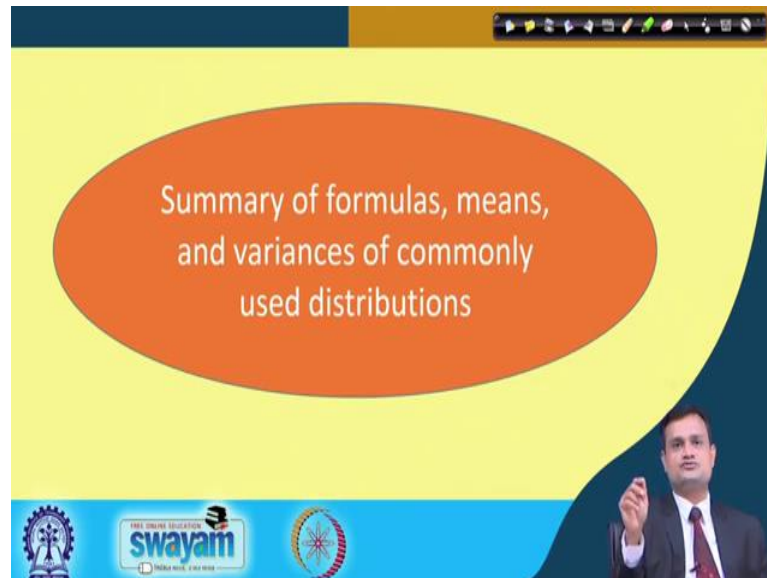
$$P(X = 3) = \frac{\binom{25}{3} \binom{75}{7}}{\binom{100}{10}}$$

$$= \frac{(2300)(1.98 \times 10^9)}{1.73 \times 10^{13}}$$

$$= 0.2637$$

So, you can just plug in the values in the expression, capital N is 100, s is 25, N minus S is 75, n is 10, x is 3, n minus x is 7. So, P (X) is equal to 3 is 0.2637.

(Refer Slide Time: 38:47)



So, summary of formula, mean and variance you can see here.

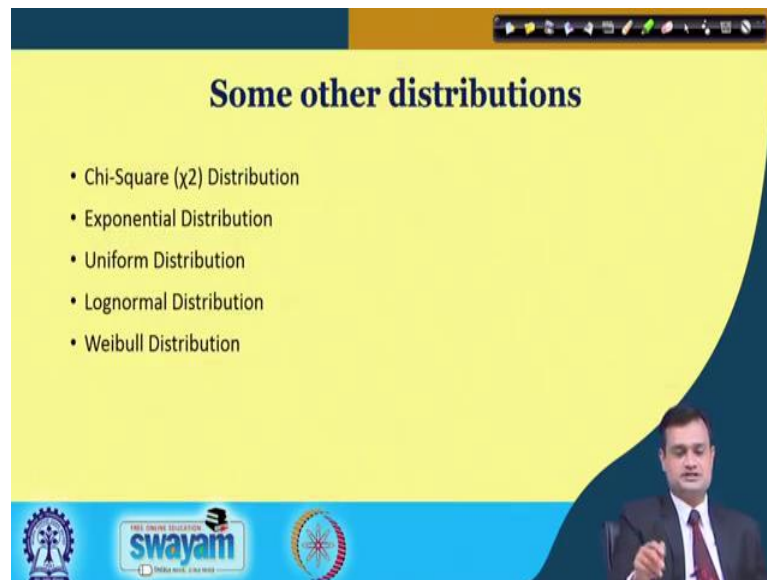
(Refer Slide Time: 38:51)

A presentation slide showing a table of probability distributions. The table has four columns: Distribution, Formula, Mean, and Variance. The rows include Normal, Poisson, Binomial, Chi square, t, and F distributions. At the bottom, there is a blue banner with logos for "swayam" and "INDIA RISE, EDUCATION RISE". A small inset video of a man in a suit is visible in the bottom right corner.

Distribution	Formula	Mean	Variance
Normal	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ for $-\infty < x < \infty$	μ	σ^2
Poisson	$f(x) = \frac{e^{-\lambda} \lambda^x}{x!}$ for $x = 0, 1, 2, \dots$	λ	λ
Binomial	$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0, 1, \dots, n$	np	$np(1-p)$
Chi square	$f(x) = \frac{e^{-x/2} x^{v/2-1}}{2^{v/2} \Gamma(v/2)}$ for $x \geq 0, v = \text{degrees of freedom}$	v	$2v$
t	$f(x) = \frac{1}{\sqrt{\pi v}} \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})} \left(1 + \frac{x^2}{v}\right)^{-\frac{v+1}{2}}$ for $-\infty < x < \infty$	0 for $v \geq 2$	$\frac{v}{v-2}$ for $v \geq 3$
F	$f(x) = \frac{\Gamma(\frac{v_1+v_2}{2}) \Gamma(\frac{v_1}{2}) \Gamma(\frac{v_2}{2})}{\Gamma(\frac{v_1}{2}) \Gamma(\frac{v_2}{2}) \Gamma(\frac{v_1+v_2}{2})} \left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}} \left(\frac{v_2}{v_1}\right)^{\frac{v_2}{2}} \left(1 + \frac{v_1 x}{v_2}\right)^{-\frac{v_1+v_2}{2}}$ for $0 < x < \infty, v_1 > 0, v_2 > 0$	$\frac{v_2}{v_1-2}$ for $v_1 \geq 3$	$\frac{2v_2^2 \{v_1 + v_2 - 2\}}{v_1 (v_1 - 2)^2 (v_1 - 4)}$ for $v_1 \geq 5$

I have put all normal, Poisson, binomial, chi-square, T and F, and their mathematical expressions typically describes the nature of a particular distribution mean and variance that are the parameter of interest and their particular expressions.

(Refer Slide Time: 39:13)

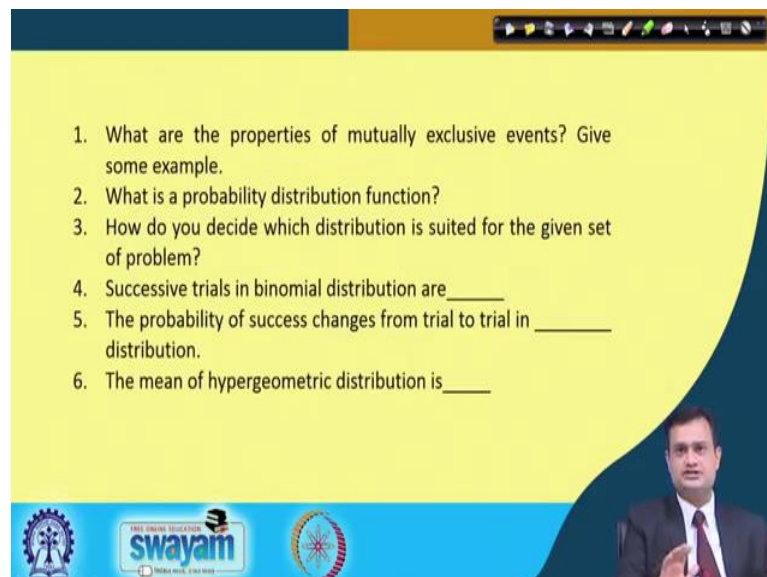


The slide is titled "Some other distributions" in bold black text. It lists five distributions in a bulleted format: Chi-Square (χ^2) Distribution, Exponential Distribution, Uniform Distribution, Lognormal Distribution, and Weibull Distribution. The slide has a yellow background with a dark blue curved border on the right. At the bottom, there are logos for "swayam" and "INDIA RISE, COUNTRY RISE". A small video inset of a man in a suit is visible in the bottom right corner.

- Chi-Square (χ^2) Distribution
- Exponential Distribution
- Uniform Distribution
- Lognormal Distribution
- Weibull Distribution

Some other distributions we are not discussing, but they are also used in many other domain. We have focused on couple of distributions which are very important from six sigma quality point of view, but other distributions may be chi-square, exponential, uniform, lognormal and Weibull distribution.

(Refer Slide Time: 39:35)



The slide contains six numbered questions related to probability distributions. The background is yellow with a dark blue curved border on the right. At the bottom, there are logos for "swayam" and "INDIA RISE, COUNTRY RISE". A small video inset of a man in a suit is visible in the bottom right corner.

1. What are the properties of mutually exclusive events? Give some example.
2. What is a probability distribution function?
3. How do you decide which distribution is suited for the given set of problem?
4. Successive trials in binomial distribution are _____
5. The probability of success changes from trial to trial in _____ distribution.
6. The mean of hypergeometric distribution is _____

Think it. What are the properties of mutually exclusive events? Give some example from your day to day life. What is the probability distribution function? We have seen PMF probability mass function, probability density function, and there is another term which

is called probability distribution function. How do you decide which distribution is most suitable for a given set of problem? Successive trials in binomial distributions are, fill in the blank. The probability of success changes from trial to trial, fill in the blank just now we discussed. And the mean of Hypergeometric distribution, fill in the blank.

(Refer Slide Time: 40:16)

References:

- ❑ Aczel, A., Sounderpandian, J. and Saravanan, P., Complete Business Statistics, McGraw Hill Publication, Seventh Edition.
- ❑ David M. Levine, Timothy C. Krehbiel, Mark L. Berenson and P. K. Vishwanathan, Business Statistics, Fifth Edition, Pearson Publication.
- ❑ T. M. Kubiak and Donald W. Benbow, The Certified Six Sigma Black Belt Handbook, Second Edition, Pearson Publication.
- ❑ Michael H. Kutner, Christopher J. Nachtsheim, John Neter and William Li, Applied Linear Statistical Models, Fifth Edition, McGraw Hill Education.
- ❑ Dimitri Bertsekas and John Tsitsiklis, Introduction to Probability.

The slide features a dark blue background with the word 'References' in a large, yellow, cursive font. A yellow banner on the right contains the reference list. At the bottom right, there is a small video feed of a man in a suit. Logos for 'swayam' and 'MOOCs' are visible at the bottom.

So, with this you refer this couple of references to deepen your understanding.

(Refer Slide Time: 40:21)

Conclusion:

- ❖ **Probability** is the measure of the likelihood that an event will occur
- ❖ The probability that a particular event occurs is a number between 0 and 1, inclusive.
- ❖ **Probability distribution** is a statistical function that describes all the possible values and likelihoods that a random variable can take within a given range.

The slide features a dark blue background with the word 'Conclusion' in a large, yellow, cursive font. A yellow banner on the right contains the conclusion points. At the bottom right, there is a small video feed of a man in a suit. Logos for 'swayam' and 'MOOCs' are visible at the bottom.

And we can conclude that probability is the measure of the likelihood that an event will occur. So, its chance and the probability is always between 0 to 1, and probability

distribution is a statistical mathematical function that describes all possible values and likelihoods that a random variable can take within a given range.

So, thank you very much. Keep revising. Enjoy.