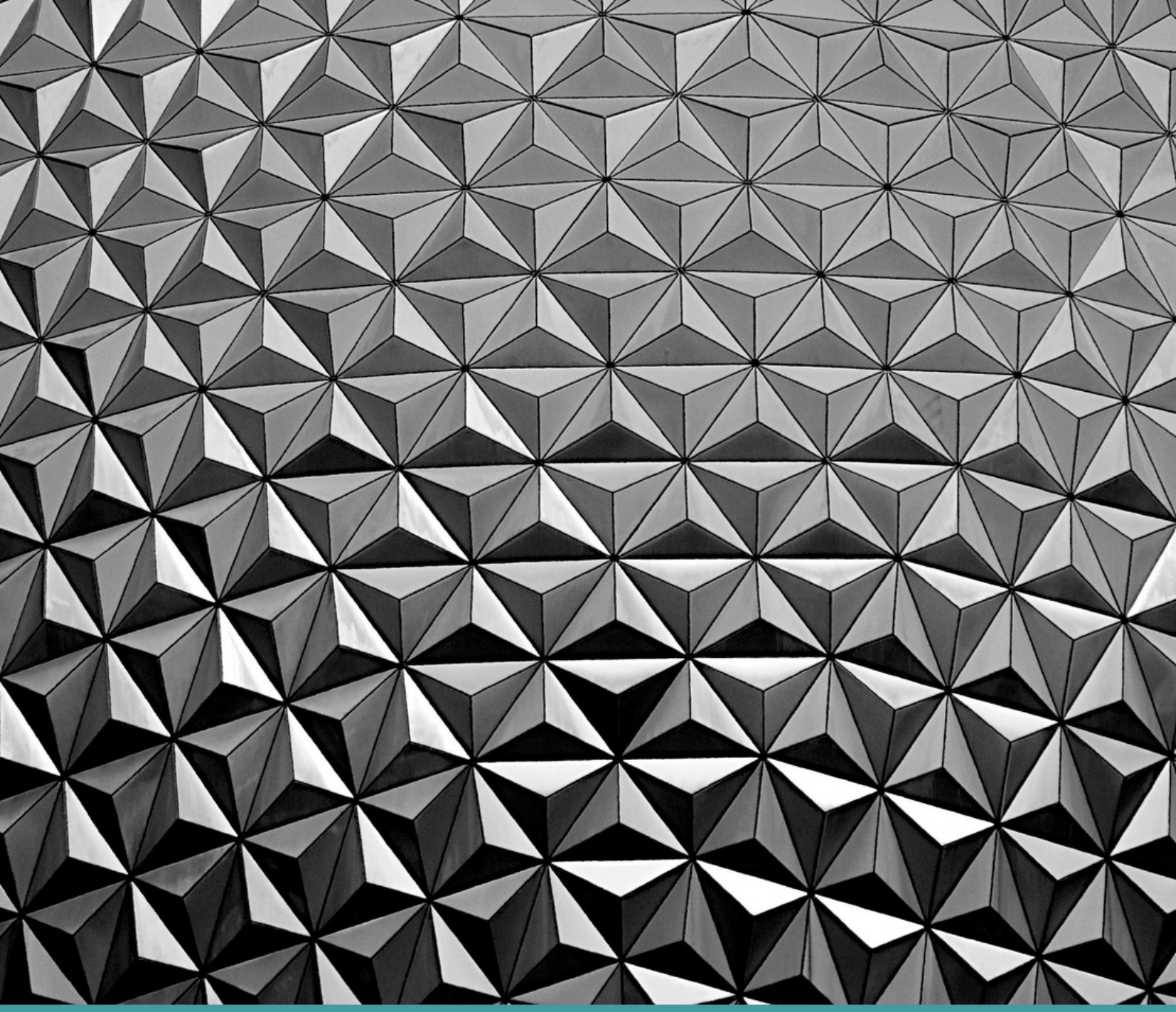


STATISTICAL LEARNING FOR RELIABILITY ANALYSIS

Prof. Monalisa Sarma
Computer Science and Engineering
IIT Kharagpur

INDEX

S.NO	TOPICS	PAGE.NO
	Week 1	
1	Lecture 01: Introduction to Reliability Engineering	4
2	Lecture 02: Introduction to Statistical Methods in Reliability	28
3	Lecture 03: Concept of Probability and Probability Theory	56
4	Lecture 04: Tutorial on Introduction to RE, SL and Probability Theory (Part-I)	77
5	Lecture 05: Conditional, Total and Reverse Probability	103
	Week 2	
6	Lecture 06 : Tutorial on Conditional Probability and Total Probability	122
7	Lecture 07 : Introduction to Probability Distributions	144
8	Lecture 08 : Introduction to Probability Distributions(Contd.)	161
9	Lecture 09 : Discrete Probability Distribution (Part 1)	173
10	Lecture 10 : Discrete Probability Distribution (Part 2)	199
	Week 3	
11	Lecture 11 : Tutorial on Discrete Probability Distributions	216
12	Lecture 12 : Continuous Probability Distributions (Part 1)	242
13	Lecture 13 : Continuous Probability Distributions (Part 2)	263
14	Lecture 14 : Tutorial on Continuous Probability Distribution Functions (Part-1)	283
15	Lecture 15 : Tutorial on Continuous Probability Distribution Functions (Part 2)	304
	Week 4	
16	Lecture 16 : Sampling Distributions (Part 1)	322
17	Lecture 17 : Sampling Distributions (Part 2)	341
18	Lecture 18 : Sampling Distribution (Part 3)	359
19	Lecture 19 : Sampling Distribution (Part 4)	378

20	Lecture 20 : Sampling Distribution (Part 5)	395
	Week 5	
21	Lecture 21 : Tutorial on Sampling Distributions	414
22	Lecture 22 : Statistical Inference (Part 1)	437
23	Lecture 23 : Statistical Inference (Part 2)	465
24	Lecture 24 : Statistical Inference (Part 3)	485
25	Lecture 25 : Tutorial on Statistical Inference	505
	Week 6	
26	Lecture 26 : Statistical Inference (Part 4)	526
27	Lecture 27 : Statistical Inference (Part 5)	550
28	Lecture 28 : Tutorial on Confidence Interval	568
29	Lecture 29 : Statistical Inference (Part 6)	587
30	Lecture 30 : Statistical Inference (Part 7)	615
	Week 7	
31	Lecture 31 : Statistical Inference (Part 8)	631
32	Lecture 32 : ANOVA-I	654
33	Lecture 33 : ANOVA-II	677
34	Lecture 34 : ANOVA – III	698
35	Lecture 35 : ANOVA – IV	715
	Week 8	
36	Lecture 36 : ANOVA – V	731
37	Lecture 37 : ANOVA – VI	745
38	Lecture 38 : Correlation Analysis (Part-I)	759
39	Lecture 39 : Correlation Analysis (Part-II)	787
40	Lecture 40 : Regression Analysis (Part-I)	806
	Week 9	
41	Lecture 41 : Regression Analysis (Part-II)	822
42	Lecture 42 : Regression Analysis (Part-III)	842
43	Lecture 43 : Tutorial on Relation Analysis	857

44	Lecture 44 : Auto-Regression Analysis	880
45	Lecture 45 : Logistic Regression (Part-I)	892
	Week 10	
46	Lecture 46: Logistic Regression (Part-II)	905
47	Lecture 47: Logistic Regression (Part-III)	916
48	Lecture 48: Tutorial on Logistic Regression	929
49	Lecture 49: Introduction	939
50	Lecture 50: Bayes Classification (Part-I)	948
	Week 11	
51	Lecture 51 : Bayesian Classification (Part—II)	959
52	Lecture 52 : k-Nearest Neighbor Classification	960
53	Lecture 53 : Tutorial on Classification Techniques	969
54	Lecture 54 : Support Vector Machine (Part-I)	981
55	Lecture 55 : Support Vector Machine (Part-II)	992
	Week 12	
56	Lecture 56: SVM (Part-III)	1004
57	Lecture 57: SVM (Part-IV)	1016
58	Lecture 58: SVM (Part-V)	1027
59	Lecture 59	1039
60	Lecture 60: Tutorial on SVM	1055

Statistical Learning for Reliability Analysis
Dr. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology – Kharagpur

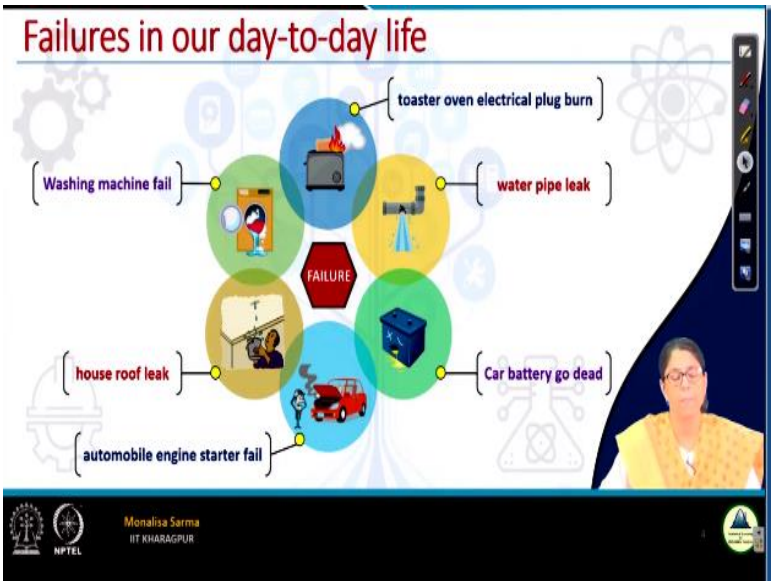
Lecture – 01
Introduction to Reliability Engineering

Warm greetings, welcome to the course on statistical learning for reliability analysis. As this is the first course on this topic, first I will introduce you to the world of reliability engineering. However, I will not go to the details of it. Because this is not a course on reliability engineering, I will mainly focus on statistical methods and which I will to explain the different concepts. I will take the examples from the level of engineering however of course I will also consider some example of other application of statistical methods as well.

(Refer Slide Time: 01:01)

So, in this class basically first I introduced reliability engineering, as I already mentioned, then I will discuss why reliability engineering is necessary then I will talk of failures in the parlance of reliability engineering, causes and types of failure. And then I will talk about a concern and objective of a reliability program. So, now reliability it is something which we see in our day to day life, like a human being has reached to such a stage where we are dependent on almost all our activities we are dependent on some of the other machineries.

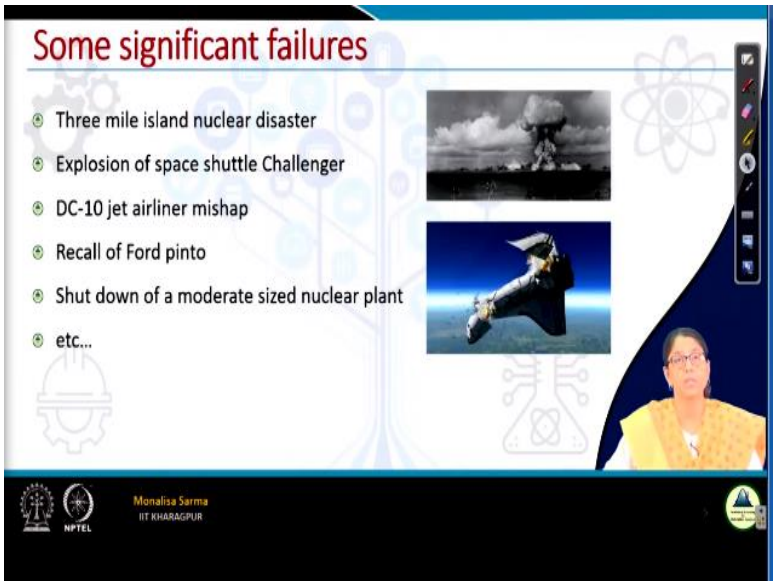
(Refer Slide Time: 01:41)



So, if we just look around, we will see many of the things almost failed daily, we will see some of the things which fail one or the other things fail, like if we consider say, washing machine. So, washing machine fail, like washing machine failed, a failure washing machine maybe due to the wear out period like it has done for the normal life and his wear out has crept in that so, maybe there was washing machine and failed. Then the car battery go there that also may be the after walking for the normal life it has gone it has worn out and so it is battery has gone dead.

But again there are something like toaster oven electric plugs burn, this may be the defect in their design, the design of the electric plug, then this house roof leak, house roof leak is maybe due to the faulty construction, then the water heater leakage, water heater leakage is again maybe this may be due to corrosion. The corrosion clips in a video to not proper maintenance which could have easily avoided by doing preventive maintenance. So, these are some simple failures which we experience in our day to day life.

(Refer Slide Time: 02:37)

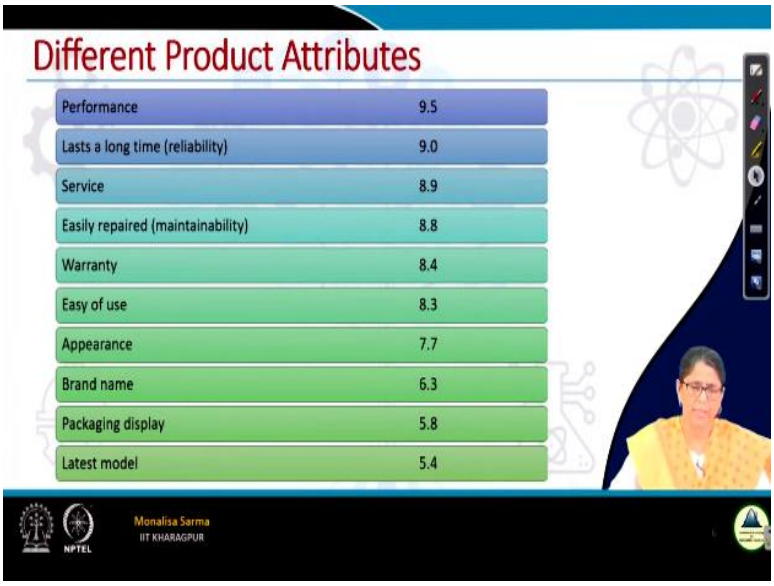


Similarly, there are some failures which are quite significant and whose effects are much more than this, what we have seen are in our last slide like to talk about the significant failure first, let me talk of one failure which has occurred one nuclear failure like which we nuclear disaster is which is happened in Ukraine because I am talking about it because Ukraine is very much in news nowadays. So, in Ukraine there was a wars nuclear disaster is Chernobyl nuclear disaster.

It is most in nuclear disaster in terms of cost in terms of causality that has many billion dollars has been needed to clean up the place. And moreover, the causality during that time is also quite high and the effect was seen for quite a long time affected because of the radiation, we could see many people are suffering from cancer different sort of cancer after many years of disaster as well. Then we are going for back home we can see we can talk about the Bhopal gas tragedy that is one of the worst industrial disaster again in terms of causality.

So, these are some significant failures there are many significant failure like I can talk about the 3 mile Island nuclear disaster which is in US and the explosion of the Challenger Space Shuttle, then we call a Ford Pinto that has happened long back that was in 1970s due to many defects of Ford Pinto like nowadays in news there are so many electric 2 wheelers, there were some electric 2 wheelers we could see it as burning here and there. So, recall of that 2 wheelers is very much in demand now. So, these are some of the significant failure.

(Refer Slide Time: 04:12)



Now from these 2 slides what we can see is that like, because of this failure, that there may be some may be minor irritants or it from starting from minor it may to and it may lead to some catastrophic failures like economy in terms of economic as well as in terms of casualties. So, way back in 1970s, a social science survey was done around they have interviewed around 1000 people that was it was like this survey was mainly done to find out people give what people give importance to which parameters while they are buying a product.

So, this products, maybe any sort of product, and they were like different parameters when you talk a product attributes the different product attributes to product attributes may be performance, my warranty, ease of use, brand name that is more than many things, many parameters, all functional parameters as well as non functional parameters. And it was noticed that people give me quite a bit of importance to reliability and maintainability.

Of course performance is something which people really definitely they will give importance to it, but along with that, people gives quite a bit of importance to reliability and maintainability other things like latest model brand name, appearance and all those things, they are not very much important while purchasing a product.

(Refer Slide Time: 05:37)

So, from this, we could conclude that engineers responsible for product design much therefore, include both reliability and maintainability as design criteria meaning like, when we developed a product, it is not that we should only focus on the functionality of the product like, if I am developing a washing machine, what is the functionality of the washing machine? Functionality of the washing machine maybe just washes clothes. So, it is not that we should not only focus on the functionality, but at the same time we should also focus on the reliability.

How well this washing machine will perform over a period of time, maintainability if something goes back, is it repairable is it how much time it takes to bring back the machines to its original condition like a working condition. So, this criteria reliability criteria, maintainability criteria this has to be incorporated in the design criteria as well. So, that was the conclusion from the survey, but it was done from the around 1000 customers.

(Refer Slide Time: 06:36)

So, now coming to what is the compulsion for a reliable product? It is only economy compulsion or is it only safety compulsion what we have seen in the last few slides, no it is not only that, there are 3 different types of compulsion for reliable products that is one is economy compulsion, environmental compulsion and performance compulsion. Now, these 3 compulsion are not independent to each other, it is to compulsion are very much inter linked.

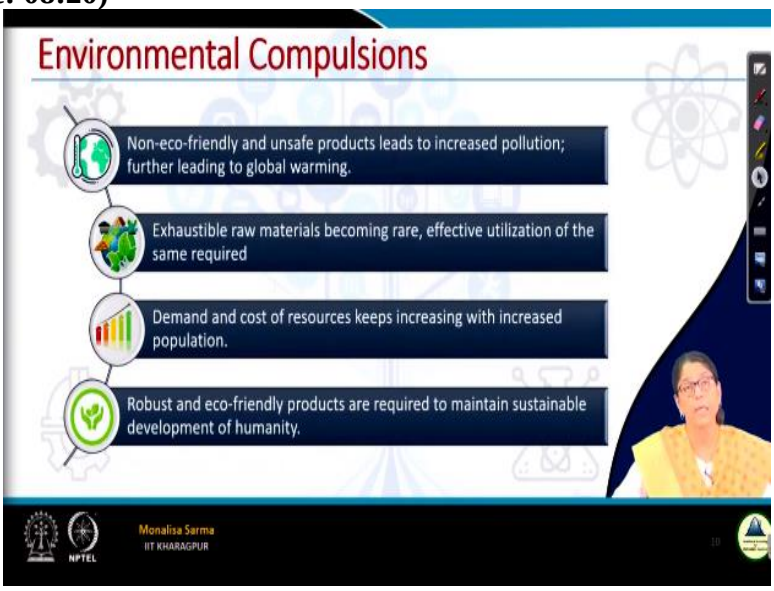
(Refer Slide Time: 07:06)

Like if we talk about economy compulsion, first like the old concept of using throw, it is no longer applicable in today's world wide because the cost of harnessing material and energy use while making product it goes very high. So, the product has to be long lasting from the economy consideration provider has to be long lasting, as well as it needs to be recyclable and moreover, for a company to survive in this market.

The one important criteria is that today's market as you know it is a very competitive market for similar type of product, there are many customers, many producers, they are producing similar type of product. So, to survive in this competitive market, a product must provide trouble free service, it must provide better customer satisfaction, along with the functionality of course, functionality is definitely very important, along with the proper functionality.

It should provide trouble free service, better customer satisfaction and all this at the very cheaper price. So, this is the economy compulsion of reliability product, unless the product is reliable you cannot achieve this.

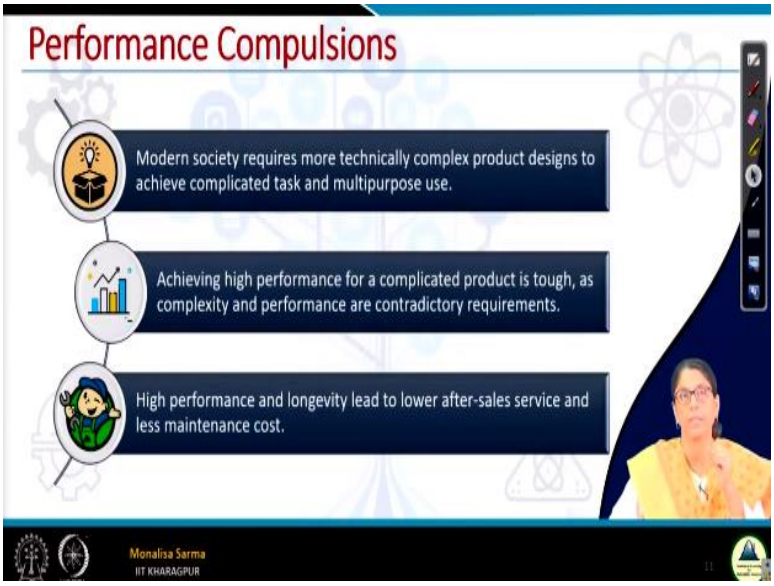
(Refer Slide Time: 08:20)



Now, coming to the environmental compulsion, so, environmental compulsion when we talk about the environmental compulsion, we see that if we use non eco friendly or unsafe products, what happened? It leads to increase pollution and increase pollutions what happened? It leads to depletion of ozone layer global warming, these are the consequences basically and the main thing is that now the resource are depleting. So, because of this exhaustible raw material, it is becoming rare, effective utilization of same are very much required.

So, for the sustainability and sustainability it is to be the motto for a company for considering the health of an environmental health of heart. So, these are the different environmental compulsion.

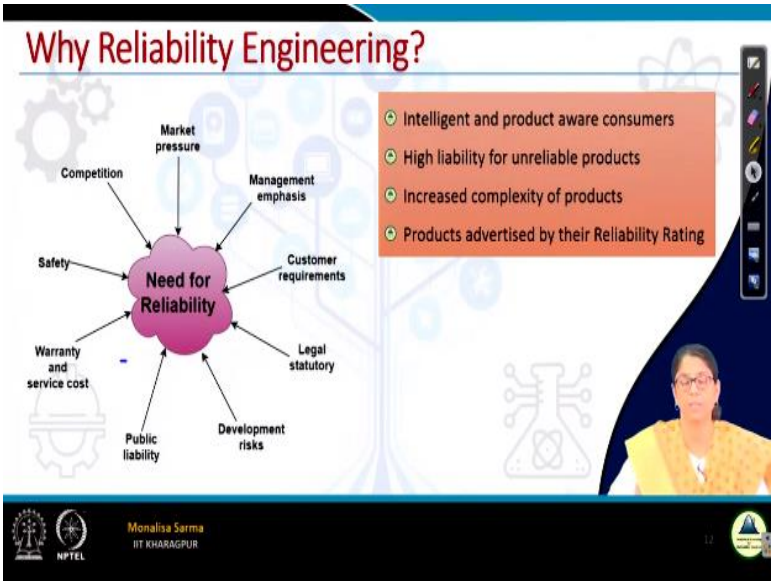
(Refer Slide Time: 09:06)



Now coming to the performance compulsions. Now, the main thing is that nowadays the products that we use, it has to be like multipurpose and like it teach to cater to different kinds of needs. Let me talk about a washing machine only, washing machine and wait like around 20 30 years back washing machine, it says wash clothes. Now, we have many requirements, we need different type of functionality for different types of clothes.

So, our demands are more, demands has become more and more complex and the demands has become more and more complex, definitely the design becomes more complex and the design becomes more complex, achieving high performance has in reliability, it becomes a challenge. So that is what high performance and longevity leads to aft er lower after sale service and lower after sale service means it leads to lower maintenance costs. And does it reduce the lifecycle costs?

(Refer Slide Time: 10:04)

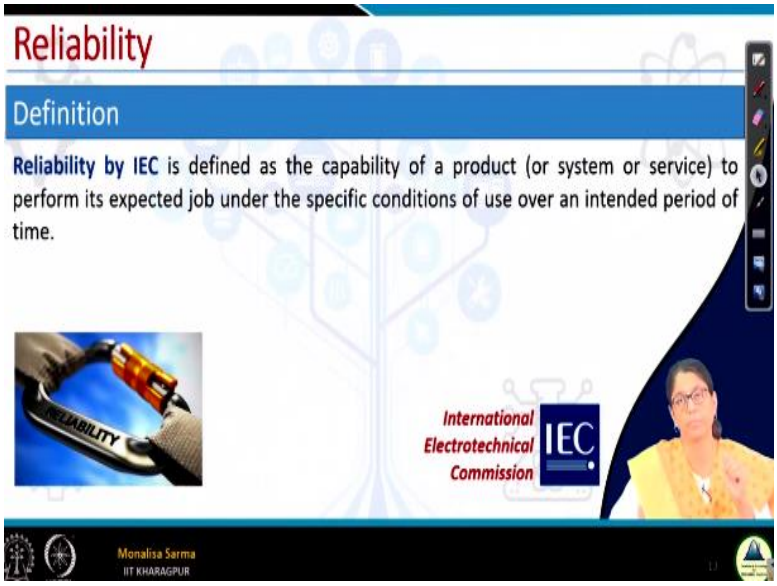


So, now having the knowledge of why the different compulsion of the reliability engineer, now we are in a position to answer why reliability engineering. So, why reliability engineering basically the first point is now, our customers are intelligent, we cannot just fool them, they are very intelligent customers and they are very product our customers. So, when the customers are intelligence and product our customers, you just cannot give any product you have to be very, very careful about all aspects of the product.

And the main thing is that we have to face high liability for an unreliable product high liability maybe in terms of a legal suit, maybe in terms of warranty costs and with the increased cost increased complexity of the product achieving this becomes very big challenge. And moreover, now as I told you for the same type of similar type of product, there are many players in the market. So, now, the products are advertised by the reliability rating.

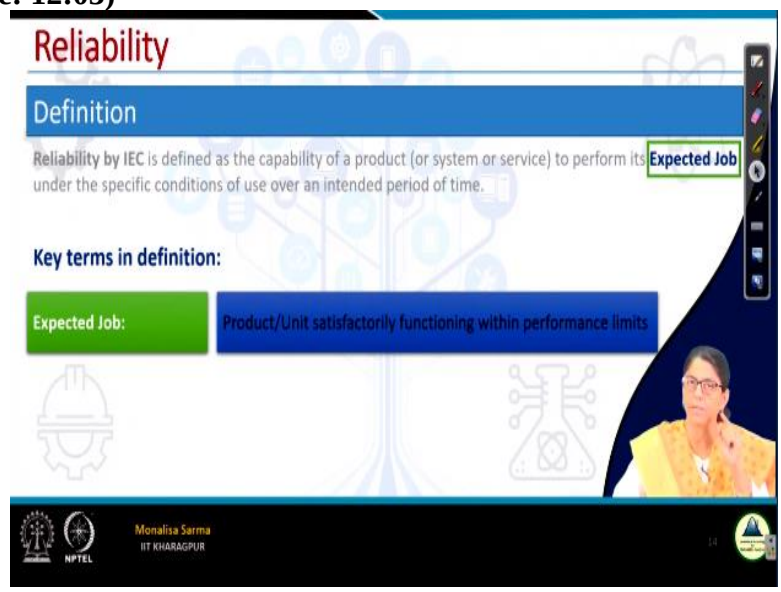
Now promotion, promotion is very much necessary for the sale of a product. So, until and unless you do not have a proper reliability rating, you cannot advertise your product. So, these are some of the reasons why we need to go for reliability engineering.

(Refer Slide Time: 11:18)



Now, we have talked enough about reliability engineering, we have basically understood what is. Now let us formally, define what is reliability engineering? So, reliability as defined by IEC What does IEC means IEC: International Electrotechnical Commission, it is an industry in international standard organizations that publishes standards for electronic, electrical, electronics and related technologies. So, reliability as defined by IEC as the capability of a product, a system of service to perform its expected job under specific conditions of use over an intended period of time.

(Refer Slide Time: 12:03)



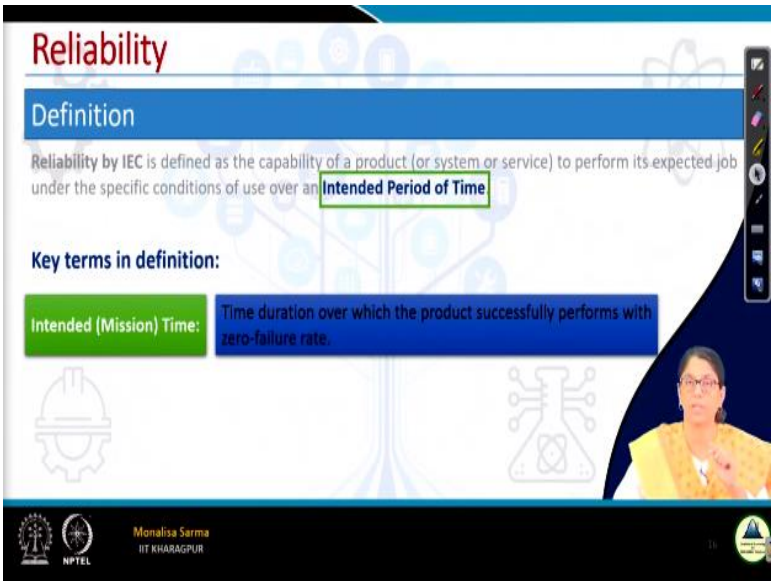
Now, if you see the definitions, there are 3 different elements to it, the first element is the performance expected job. So, what is the expected job like if we talk about a missile, what is the expected job of the missile? Expected job of the missile is to hit a target with very high accuracy

and high speed that is the expected job of the missile. So, that is the expected job after the job is product or unit should function satisfactorily within performance limit. So, now, what it may be the performance limit of the missile? Performance limits of the missile maybe the speed, maybe the accuracy. So, this is one element when we consider reliability that is the expected job.

(Refer Slide Time: 12:39)

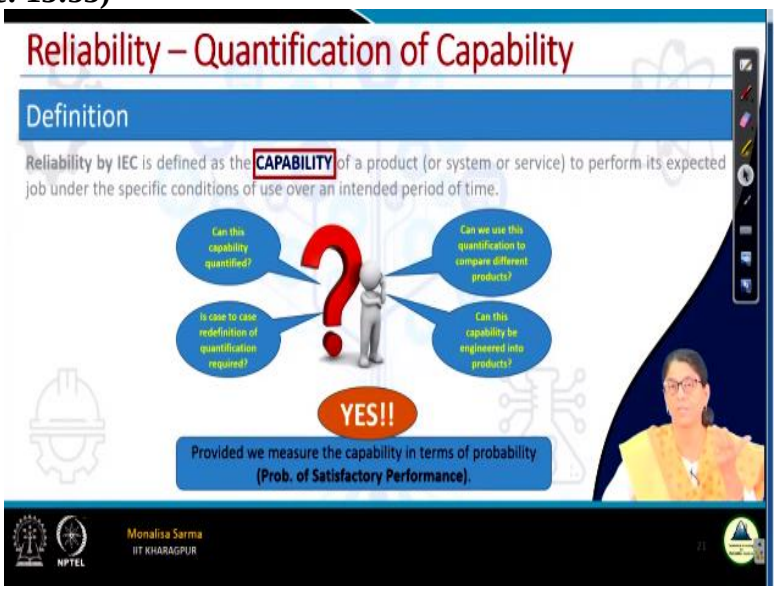
Next element is a condition of use. So, condition of use now, again when we try to find out the reliability of the product of course, the condition of use is also very important, like a product is built to operate in a particular condition of it, let us talk about luxury car, a luxury car is not expected to run in a mountain terrain when so, if you are trying to run a luxury car in a mountain terrain and then if you try to find out the reliability of the luxury car then it is not done basically. So, condition of use is the environmental condition under which the product or system has to operate.

(Refer Slide Time: 13:14)



That the third element is the mission time and then that order mission time, mission time is the time as a time which we do not want any failure for the system to occur. Like for the missile, missile the mission time is a time when we deployed a missile till it hits the target that is the mission time. So, we do not want any failure in this mission time. So, what we defined time duration over which the product successfully performed with 0 failures. So now, when we talk about reliability? We should talk about these 3 terms that is the expected job, condition of use, as well as the intended period of time.

(Refer Slide Time: 13:53)

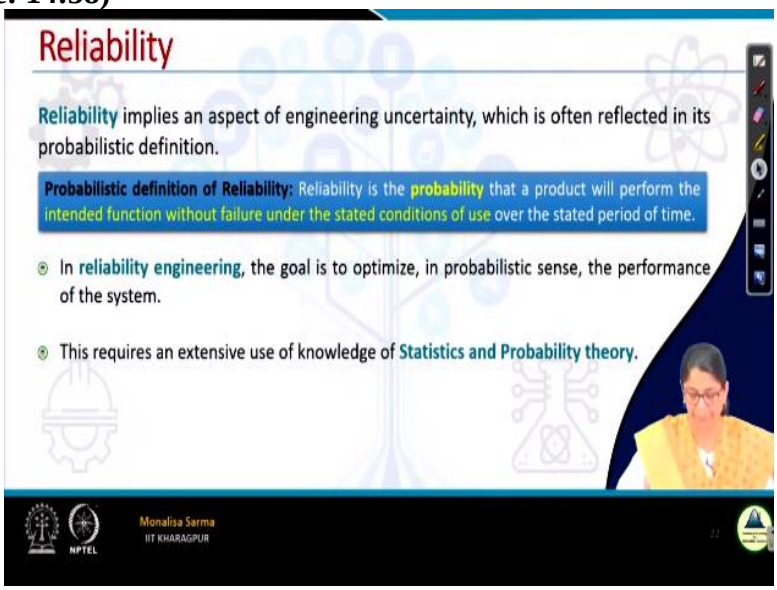


Now, there is something else also that is the capability it is defined as the capability of a product. Now what is this capability? Can this capability be quantified that is the first question. Second question, this capability do we have to resort to re-definition from case to case like for different

products? This capability we will have to define it separately as it like that or is the same definition is applicable for all types of product.

Next, can we use the quantification to compare different products? Now; similar kinds of products produced by different players using maybe different technologies. So, can we use this quantification to compare these different products? And also, can this capability be engineered into the products? So yes, the answer to all this question is a big YES. But provided we measured a capability in terms of probability that is the probability of satisfactory performance. So, when this probability is the science of basically uncertainty, right?

(Refer Slide Time: 14:58)

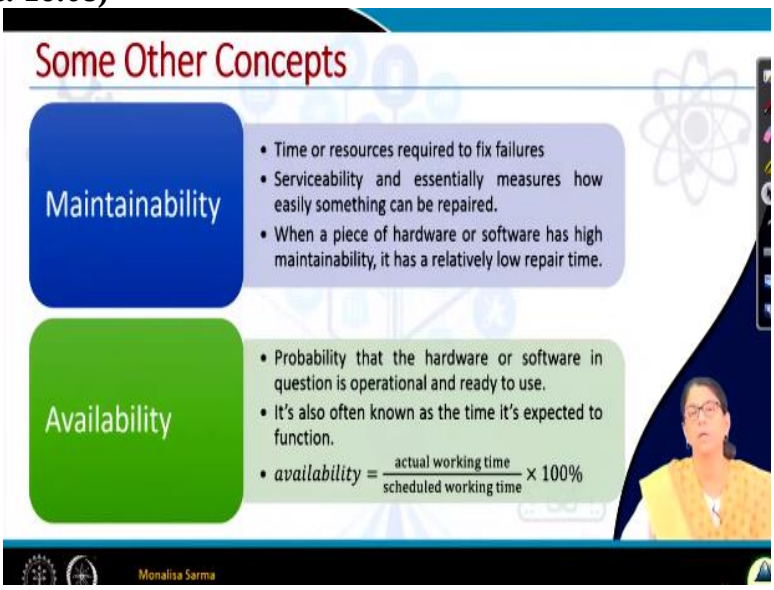


So, reliability that means it implies an aspect of engineering uncertainty. So, hence, we will come to a probabilistic definition of reliability. So, what is that? Reliability is the probability that a product will perform the intended function without failure under the stated condition of use over the stated period of time, everything remains same just instead of capability we have changed it to the Reliability that is the probability that a product will perform, the intended function without failure under a stated condition of use over a stated period of time.

So, in reliability engineering, the main goal is to basically to optimize in probabilistic sense the performance of the system. So, now to do this, it requires an extensive use of the knowledge of statistics and probability theory. So, this is basically the goal of this course. So, we will be learning the statistics and probability theory. Again, I am mentioning here which I have already

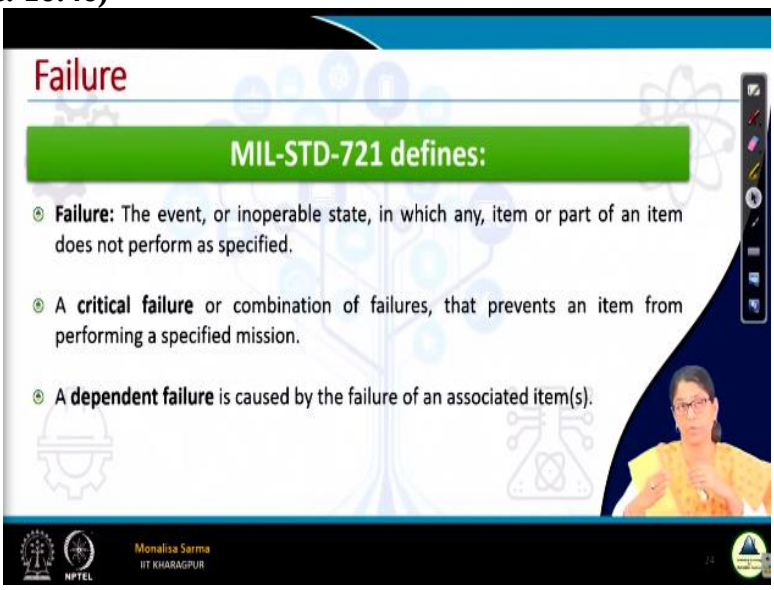
mentioned in my introductory class, that the statistics and probability theory do I will be focusing mainly from the reliability analysis application, but this is statistical as you know, it is applicable to many other fields as well.

(Refer Slide Time: 16:08)



So, there are some other related concepts of reliability which you need to know first is the maintainability. Maintainability it is the extent of time or resource required to fix failure. It is also known as serviceability and essentially it measures how easily something can be repaired. Now, what is availability? Availability is the probability that the hardware or software or any product in question is operational and ready to you. How you define mathematically availability is actually working time and schedule working time in to 100%.

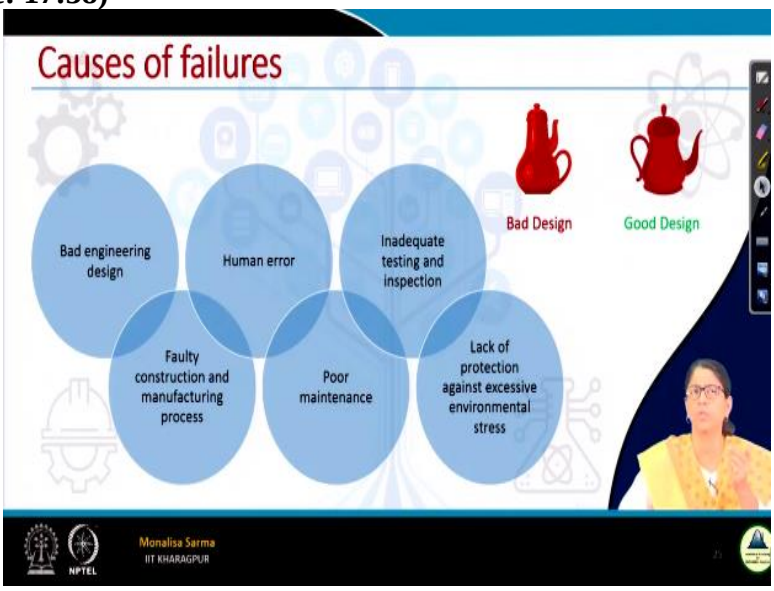
(Refer Slide Time: 16:46)



So, now, having talked of reliability say, in the reliability definition we have encountered what different failure now, we will define failure in the parlance of reliability engineering. So, what is a failure in the parlance of reliability engineering? Failure is the event or inoperable state in which any item or part of an item does not perform as specified. Again, there is some order definition, order type of failures which we call is a critical failure, critical failure is that prevents an item from performing a specified mission.

First is a failure which may not stop you from performing a specified mission, but if it is a critical failure when submission will be aborted it will not perform, then there is something called a dependent failure, dependent failure is caused by the failure of an associated item one failure may lead to another failure that is dependent failure.

(Refer Slide Time: 17:38)

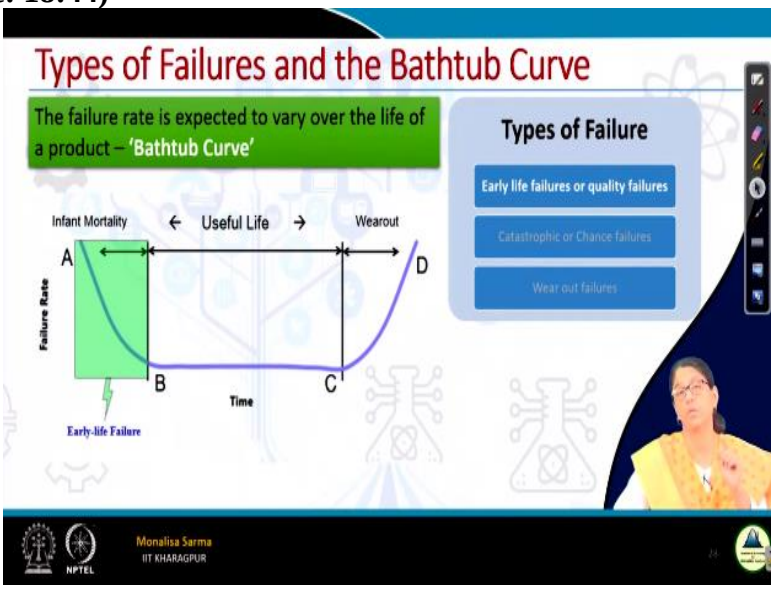


So, on different causes of failure like see, you can see the figure the first I can say different causes of failure maybe a bad design you can see in the figure the 2 designs of teapots take at least one we can see very easily we can find out why this first one is called a bad design and the second picture why we call it a good design. Now, all of you can easily see that. So, the different causes failure means maybe the bad engineering design human error, human error was one of the main reasons for failure.

Then the inadequate testing and inspections maybe faulty construction like what we have seen the house roof failure, in the first light house roof failure may be due to a faulty construction,

then again oven plug and burning that may be the faulty construction of the plug, then poor maintenance, poor maintenance what we have seen that water heater leakage is due to the corrosion has crept in why corrosion could creep in because of poor maintenance, then the lack of protection against excessive environmental stress. So, these are may be the different causes of failures there may be n number of causes for failure.

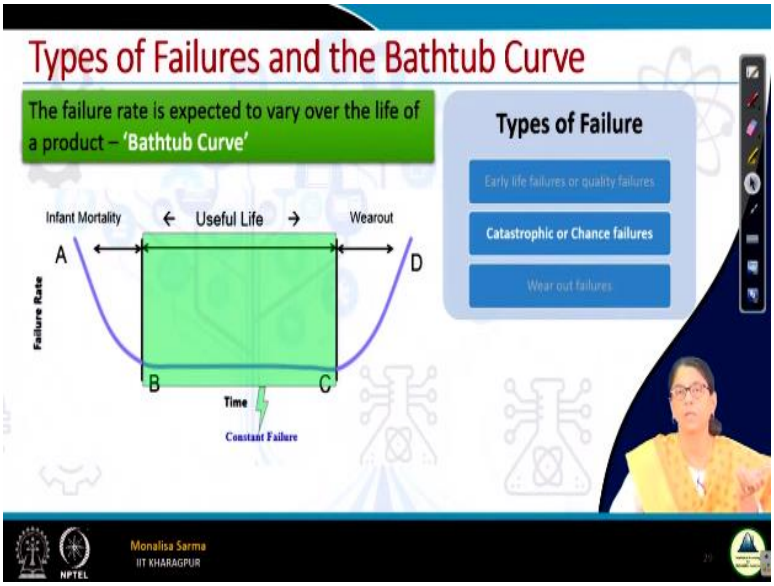
(Refer Slide Time: 18:44)



Now, what are the types of failures? Talking about the types of failures, there are basically 3 types of failures that we can see in the lifetime of a product. So, the 3 different the lifetime of a product if we consider, this we call it as a, it basically follows a curve which is very similar to a bathtub. So, it is also called a Bathtub Curve. So, the 3 different failures, the first is the early failures or which you call it a quality failures, It happens in the early life that is also called infant mortality.

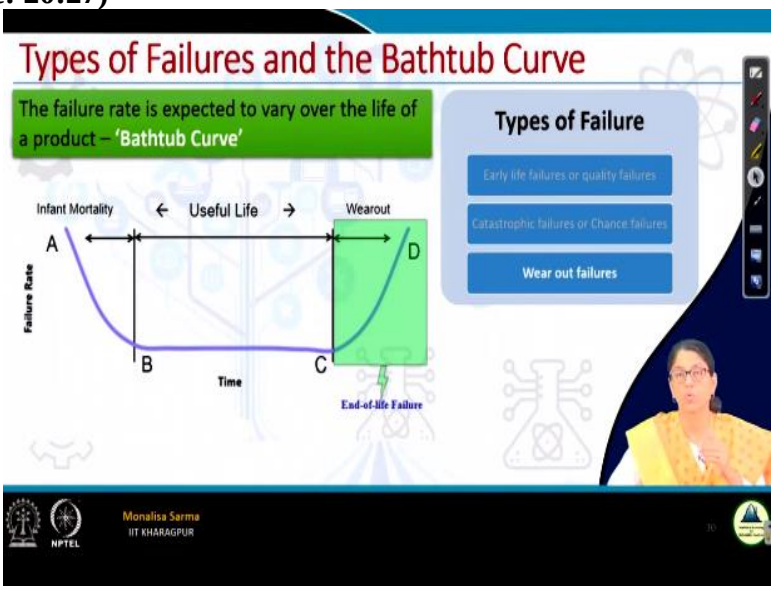
This early life this failure happens, why the failure happens in early life? The reasons maybe quality, then the design is defective, maybe the proper quality was not used, quality of material and maybe the design process was not proper. These are the many issues because of which we get failure in the initial period, it is also called the teaching problem. So, that is called the early life failure or the quality failure.

(Refer Slide Time: 19:48)



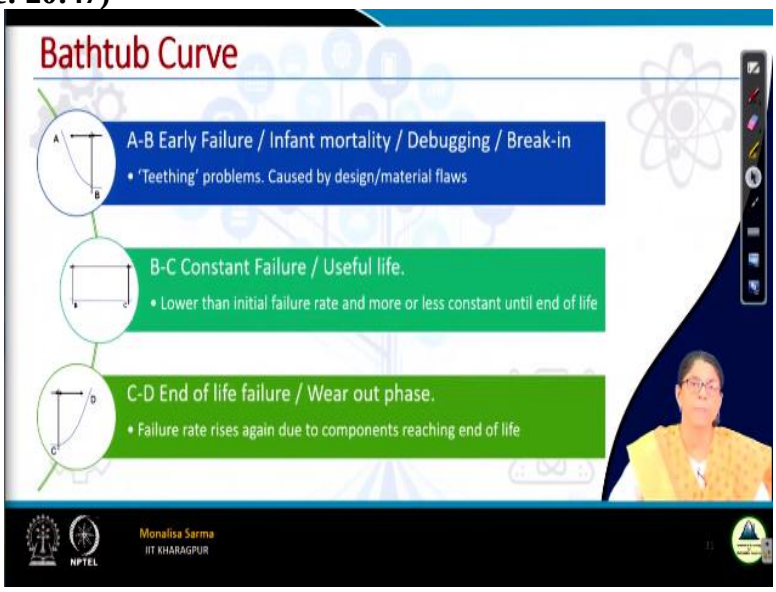
After that, then comes the normal life once this after the system has become stabilized then the normal life of the product starts. In the normal life the failure rate is more or less constant here we are do not encounter any of the quality failures here mainly the failures are chance failure, maybe due to natural environment natural causes like maybe due to earthquake, maybe due to tsunami. So, like when the stress accumulation of the products becomes more and the strength of the product than the failure occurs in the useful life time. So, that is called a catastrophic or chance failure, the failure rate at this time is constant.

(Refer Slide Time: 20:27)



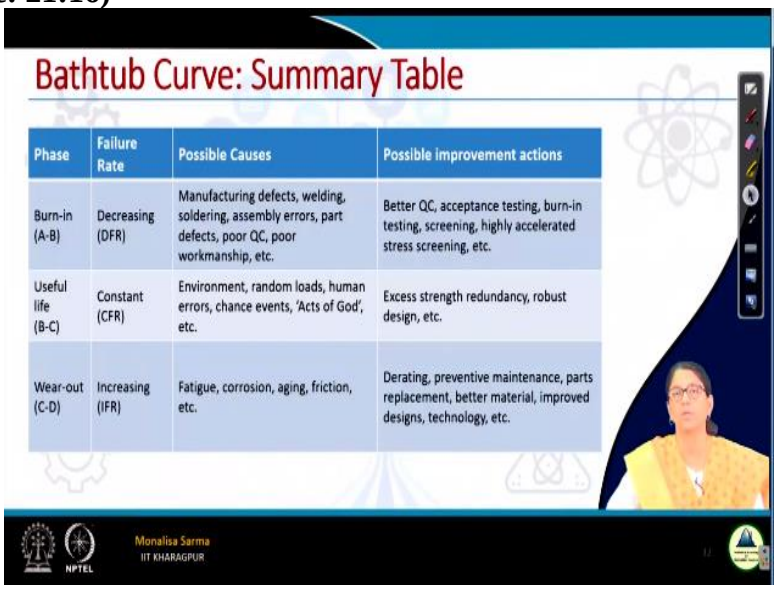
Then the last is that is wearout failures when the component or the system ages. So, it is the end of the life failure. So, aging happens, so corrosion creeps in, wearout happens, so, because of this again our failure rate increases. So this is the wearout failure.

(Refer Slide Time: 20:47)



So, here we see the first initial the early failure, early failure what happens, gradually our failure rate decreases. So, you see the first figure top figure, top failure rate visually decreases, then it is the constant failure or that is the useful life and the useful life or failure rate remains more or less constant than the wearout life wearout again failure it increases. So, these are the 3 different types of failure that we encounter usually in a component.

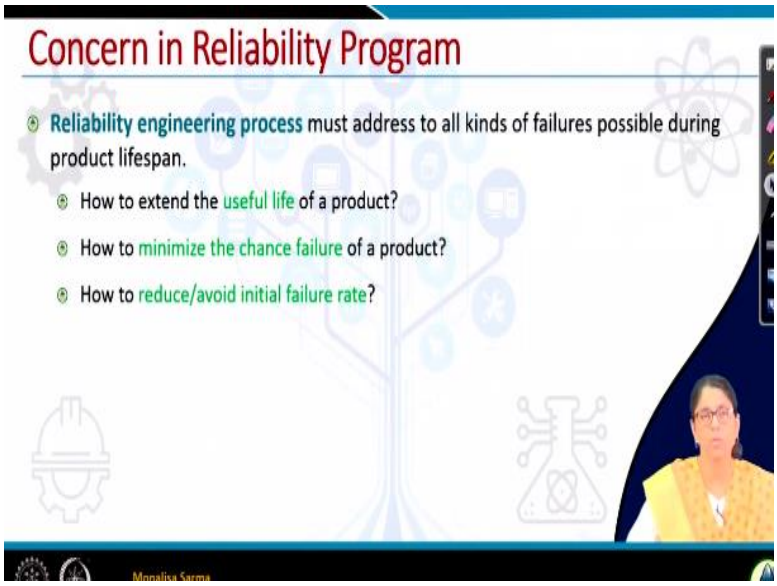
(Refer Slide Time: 21:16)



Phase	Failure Rate	Possible Causes	Possible improvement actions
Burn-in (A-B)	Decreasing (DFR)	Manufacturing defects, welding, soldering, assembly errors, part defects, poor QC, poor workmanship, etc.	Better QC, acceptance testing, burn-in testing, screening, highly accelerated stress screening, etc.
Useful life (B-C)	Constant (CFR)	Environment, random loads, human errors, chance events, 'Acts of God', etc.	Excess strength redundancy, robust design, etc.
Wear-out (C-D)	Increasing (IFR)	Fatigue, corrosion, aging, friction, etc.	Derating, preventive maintenance, parts replacement, better material, improved designs, technology, etc.

So, now, again so this is a basically summary table what may be the possible causes, why this failure has happened in 3 different phases and what are possible improvement action which I will not be explaining and it was can just go through it as scattered as a summary table here.

(Refer Slide Time: 21:34)

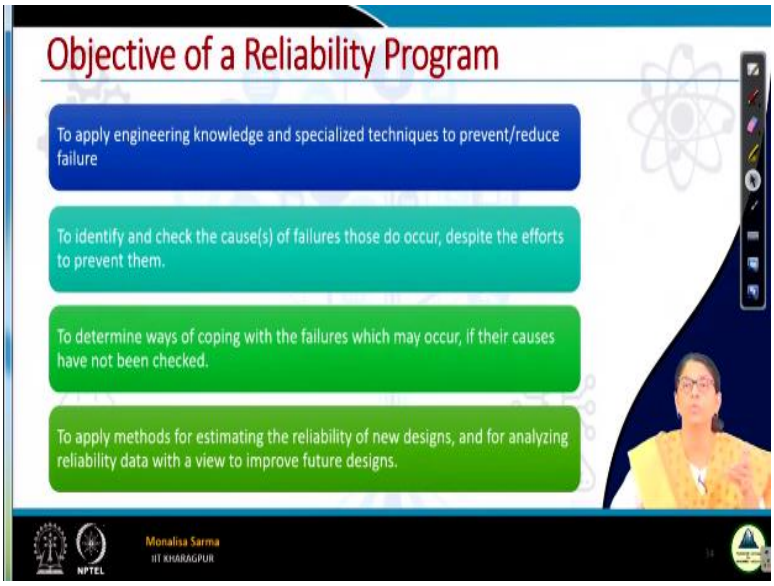


Now, we have seen the bathtub curve of a system so we know at what stage what type of failures occur? Now with this knowledge now we know what exactly to do in a reliability program. So, what are the main concerns in a reliability program? So, the main concern in a reliability program is basically, first thing is the how to extend the useful life of the product? We need our whole life to be more and more before the aging creeps in.

So, how we can extend the useful life of the product? This may be one of the concerns, second concern may how to minimize the chance failure of a product and or use chance failure that is not a catastrophic failure that against happens in normal life failure. So, how to minimize this? And then thirdly, then again how to reduce or avoid the initial failure rate , Initial failure Rate that is the teaching problem, how we can reduce the failure rate?

Then again, there may be some other concern as well how to what to say in reduce the failure rate and wearout period how to reduce the what to say reduce the increase the normal lifetime, so, that wearout happens, wearout does not happen quite fast, how we can do that? So, maybe these are the some of the concerns of the reliability program.

(Refer Slide Time: 22:48)

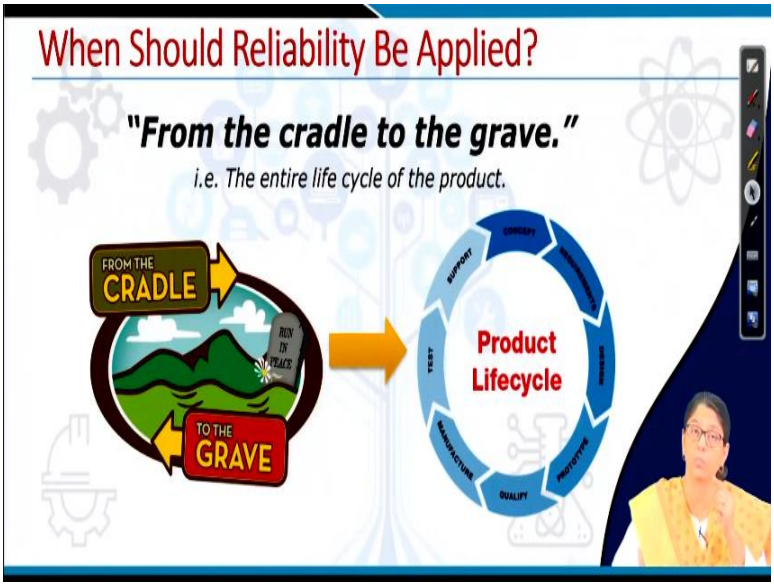


Now, we know what is reliability engineering? Why reliability is necessary? What are the different concerns of the reliability? With this now, we can basically stay down our objective of reliability program. So, what are the objectives of reliability program, there may be other objective as well as just noted down here to 4 objectives. So, first one maybe to apply engineering knowledge and specialized technique to prevent or reduce failure reduce failure needs, in all of the stages, in the initial state, the normal life and the wearout stage.

So, next to identify and check the causes of failures those 2 occur despite the effort to prevent them, how we can check the causes of the failure? We have taken steps to prevent them, but still those failures are happening, how we can check those failure? Then next to determine ways of coping with a failure which may occur if their causes have not been checked, some causes which goes, we just overlook it, then if such type of failure occurs, how we can cope with it?

This is also another objective or reliability program. Again, we can apply methods for estimating the reliability of a new design with our gain knowledge, how we can estimate the reliability of a new design? And then we can analyze the reliability data with a view to improve future designs, this may be the some of the objective of a reliability program.

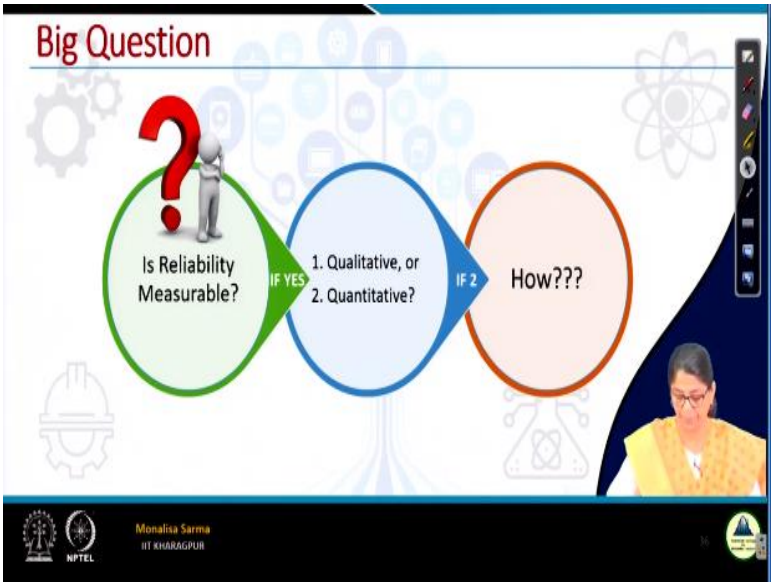
(Refer Slide Time: 24:17)



Now, when should be apply reliability? There is not a fixed timing when we should apply reliability in the whole lifetime of the product. It is basically the answer is if simple answer from the cradle to the grave. That means when I tell is from the cradle to the grave that means it is the entire lifecycle of the product. Lifecycle of the product starts from the concept when we are developing a new product, we are first we conceptualize the product, once we conceptualize the product, then we think of the different requirements of the product.

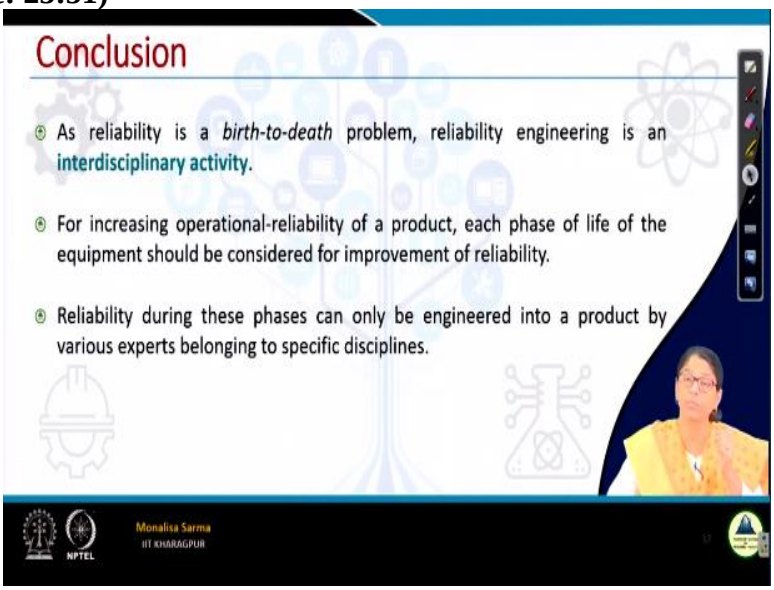
Then we design the product, we develop the prototype, qualify a prototype, then we manufacture and once we manufacture, after we manufacture we again test the product, then they are just sent to the market, and then our support. So, these are the different stages of the product, in the different stages the product goes product lifecycle, reliability should be incorporated in all these stage that is why it is called from the cradle to the grave.

(Refer Slide Time: 25:09)



So, now the big question, is reliability measurable? If yes, then it is qualitative or quantitative? Already we have seen that reliability is measurable; already we have seen it is the capability is measurable. Therefore, I think around 6 or 7 slides we have talked about that. So, now, it is definitely it is quantitative, it is measurable, it is quantitative and how we will measure it, there are different ways how we can measure reliability. One of the ways of measuring reliability, one of the most used methods of measurability is by using statistical method which is basically the genesis of this course.

(Refer Slide Time: 25:51)

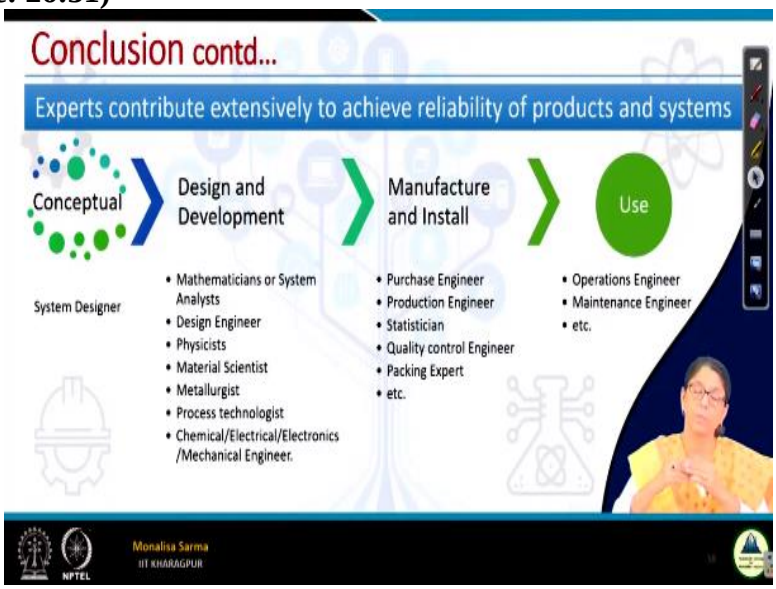


So, now, to conclude this lecture for I would like to bring in front of you this reliability when we talk with a birth to death problem, it is this reliability when starting from birth to death, it is very much an interdisciplinary activity, it is not that a simple when you are developing a mechanical

product is not a mechanical engineer is involved in so, many types of different people are involved in the development of a product. So, in the different phases, so, people from different disciplines are involved.

So, when we talk about reliability during this phase, it can be engineered into a product by various experts belonging to different disciplines.

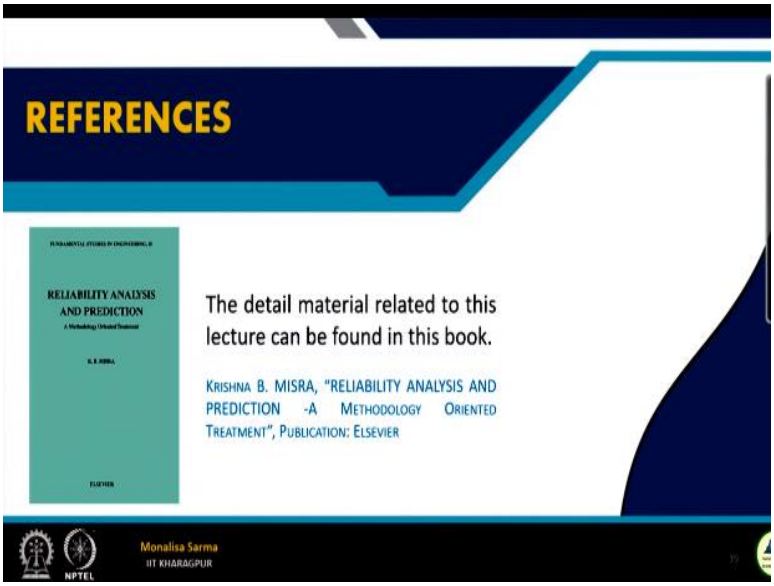
(Refer Slide Time: 26:31)



Like we can say and when we talks about the conceptualizing a product conceptualizing maybe a system designer, then the design and development and mathematician maybe there and the design engineer, physicist, physicists maybe there metrologists, process technologists, chemical engineer, electrical engineer, mechanical engineer, design and development, many different types of engineering disciplines from there.

And then again in manufacture and install them a purchase engineer, production engineer, then packing expert, then in use, use means during maintenance time for the support there can be operation engineer, maintenance engineer. So, we see different, it is a totally an interdisciplinary activity, people from different fields of disciplines are involved and incorporating reliability from the life to death process.

(Refer Slide Time: 27:20)



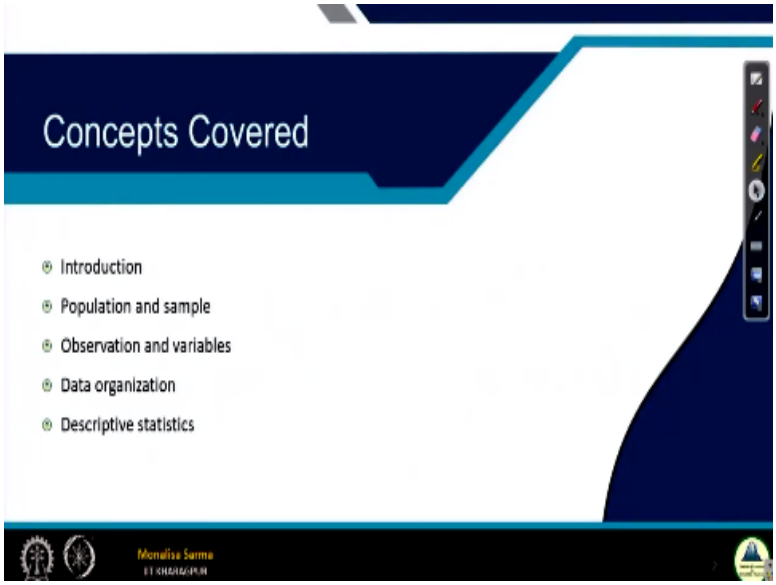
With this, I end this presentation this lecture and the detail material which I have covered in this lecture it will be found in this book, reliability analysis and prediction by Professor Krishna B Misra and thank you.

Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology - Kharagpur

Lecture – 02
Introduction to Statistical Methods in Reliability

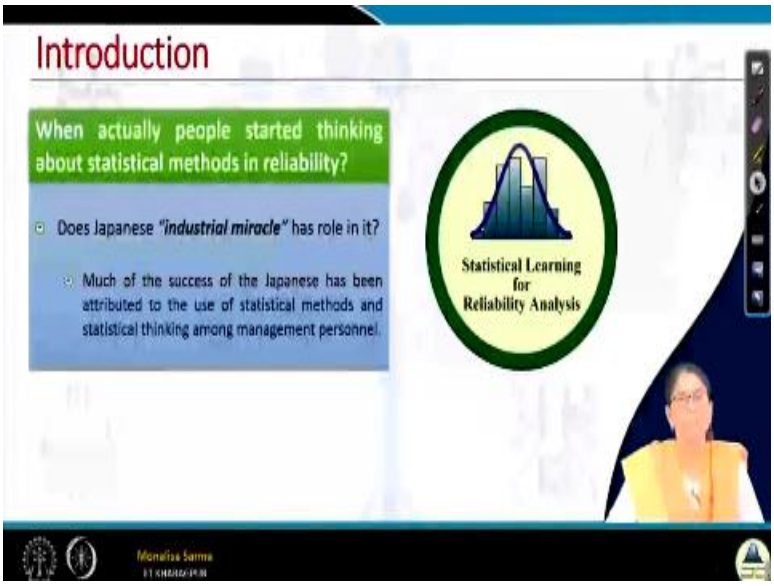
Hello everyone. So, this is the second lecture on the course statistical learning for reliability analysis. And the last class we got introduced to reliability engineering and this is the introductory class on statistical methods.

(Refer Slide Time: 00:39)



So, in this class basically we will try to learn about statistical methods and we will learn various terms related to statistical method that is like population sample, what is population, what is sample what is observation, what is variable and then we will learn also some data organization tool and then we will go to discuss some descriptive statistics.

(Refer Slide Time: 01:04)



So, statistics when we talk it is not a new concept it was there from long back like, but long back it was there, but then it was never used in industry long back like when actually people started thinking of statistical methods in reliability and from the reliability concept and asking. So, does Japanese industrial miracle has a role in it? So, what is Japanese industrial miracle? If you just Google it will be able to say at as in the middle of the 20th century, Japanese were very much successful, where we India or many other countries has failed miserably, where we have failed it is like developing such an environment where we can develop some good quality and reliable product. So, it was in the middle of 20th century Japanese were very much successful very developing very quality product, very reliable product but other countries that involvement was not there. So, as you can develop a reliable and good performance product.

So, much of the success of the Japanese has been attributed to the use of statistical methods and statistical thinking among management personnel. So, now, when we talk about use of statistical methods, like this data has when we talk about statistics definitely we talk of data like so, this using data some collecting data, summarizing data and then reporting data all this was quite prevalent in many other industries from long back, but then collection of data and statistical treatment of this data, these are 2 different issues.

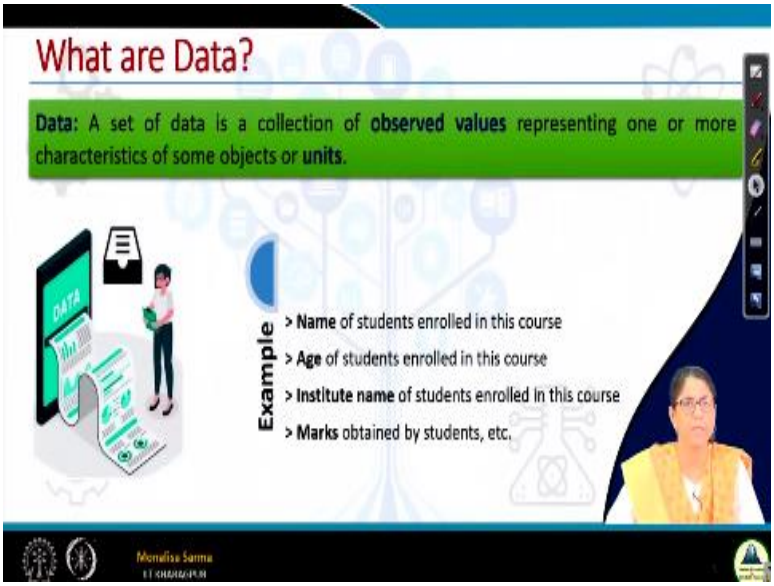
So, before this Japanese industrial miracle people were just collecting data they were collecting, they were summarizing the data and some other format, they were reporting the data and storing

the data for the purpose of result but they were not doing any statistical thinking on those data. So, it has started this Japanese successive, Japanese company this has made people think and then people started visually started using these statistical methods in their industries.

(Refer Slide Time: 03:13)

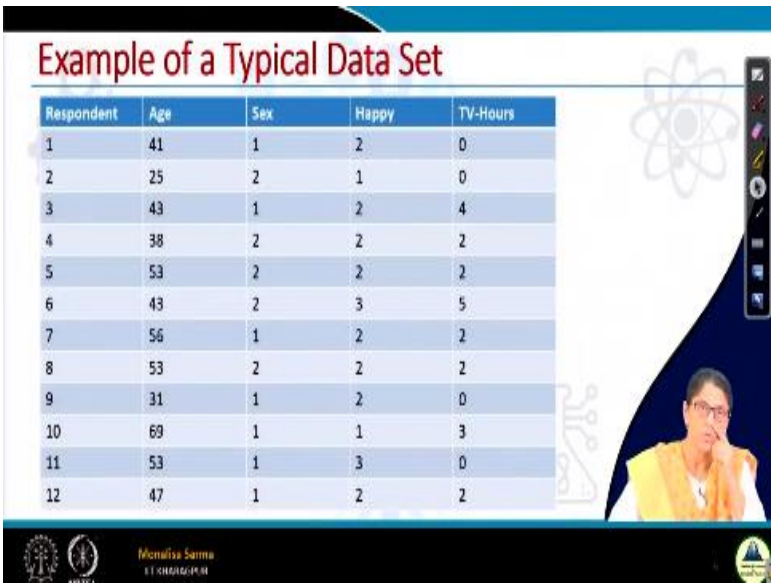
Now, when we talk of statistics as I told you, the first thing that comes to mind is what does it comes the comes to mind it is like a big, big tables of many numbers and volumes and volumes of figures, figures maybe pertaining to birth, death, taxes. So, that is why because statistics is very much used synonymously with data, this is actually in fact not very wrong also, because when we talk of statistics, Statistics basically it largely deals with principles and procedures of collecting data, collecting and summarizing and what to say taking information out of this data. So basically, when we talk of statistical method, so a good idea of what is data is very important. So, first we will go to that what is the data?

(Refer Slide Time: 04:12)



So, data is how we define data, a set of data is a collection of observed values representing 1 or more characteristics of some units, objects or units. Now, like let us take an example of a student when we if the student is a unit in particular, one student is a unit so the different data pertaining to that unit is the name of the students maybe the age of the students, institute name marks obtained by the students. These are the different data pertaining to the unit students. So, all this together, we can tell it is a set of data.

(Refer Slide Time: 04:45)



Respondent	Age	Sex	Happy	TV-Hours
1	41	1	2	0
2	25	2	1	0
3	43	1	2	4
4	38	2	2	2
5	53	2	2	2
6	43	2	3	5
7	56	1	2	2
8	53	2	2	2
9	31	1	2	0
10	69	1	1	3
11	53	1	3	0
12	47	1	2	2

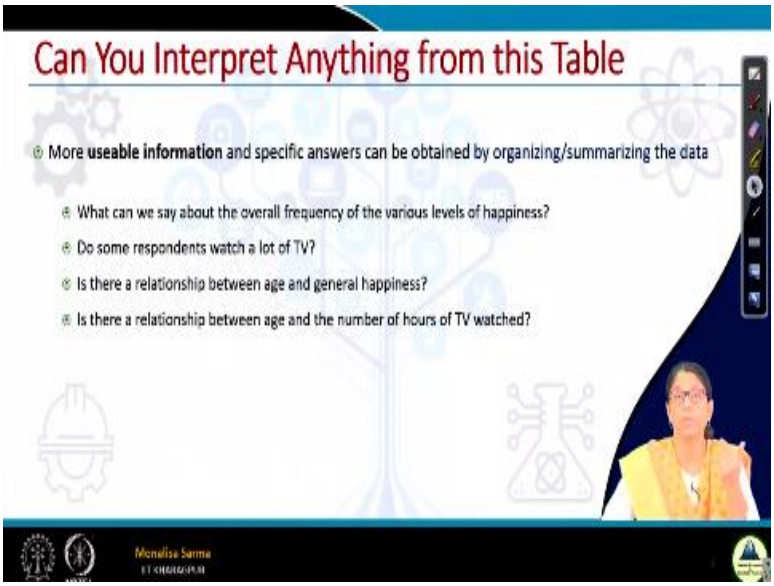
So now, it is a typical data set. So, it has not surveyed done by for some robot to say for some research purpose when people were interviewed. Today is mainly social science survey where many people were interviewed around 2000 people were interviewed, and they were asked

different questions there were around 70 questions, there is based on their political beliefs, their lifestyle, their standard of living and differences for social things, they have different questions are asked.

And I have just taken here out of this different 70 questions, I have this taken around 4 questions to bring give you an example of a typical dataset. So, here if we might have seen the different questions maybe the age of the respondent then the gender of the respondent and how happy is the respondent happy maybe I have given him the 3 index maybe he is not happy, pretty happy, very happy. So, in this index are may not happy maybe 1, pretty happy maybe 2, very happy maybe 3.

Then TV hours means how many hours the respondent watch TV, so basis. So, if I consider this data, this table of data is has many numbers like now, from this number, what we can conclude? It is very difficult to conclude anything from just if we look at this table.

(Refer Slide Time: 06:04)

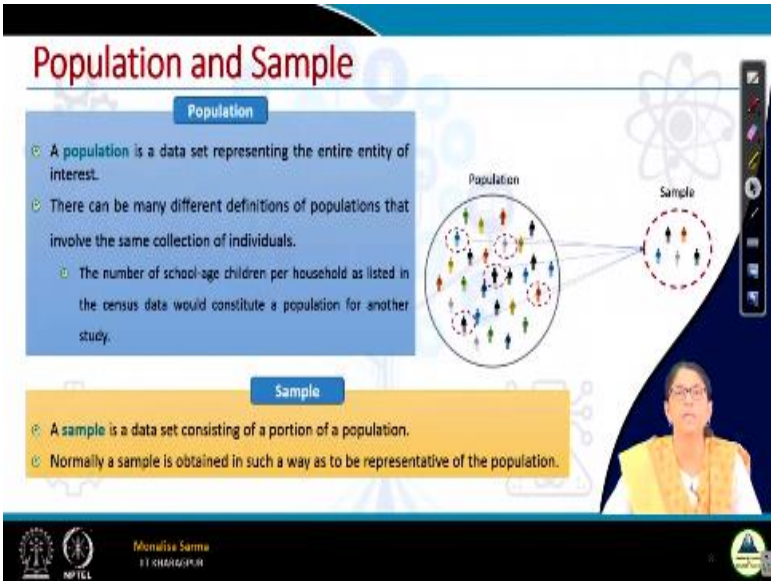


So, however, if we can somehow summarize this data, summarized this data from this table, then we can get many useful information like we can, what can we say that the overall frequency of the various level of happiness, Do some respondents watch a lot of TV these are some information you might get from the data if you can, able to summarize this data. Like is there a

relationship between age and general happiness is more or lesser age happiness is more is this such a relation is there or is there the relation between age and the number of TV hours?

So, for this, we need a data organization tools. So, we will be discussing some data organization tools, how we can summarize the data from a table of data.

(Refer Slide Time: 06:49)



Now, before summarizing, before going to the data organization tool first we need to know few other terms also like what is a population? Population, is like all of us we know what is the population it is something that is school children also know but still for completeness, let us go through it what is the population a population is a data set representing the entire entity of interest, like we know census data. So, however, there can be different definition of populations that involve the same collection of individuals.

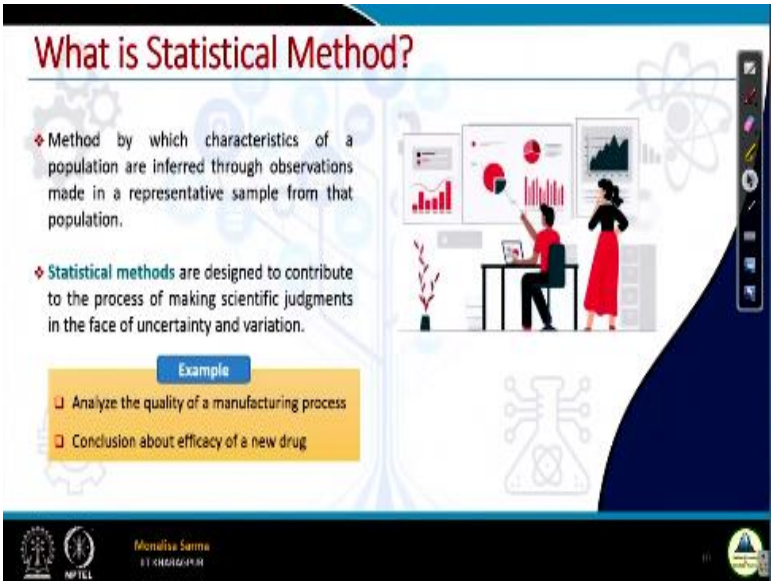
Like from the census data day one population may be the number of school aged children per household, whatever we got in the census data, it constitutes a population for another study again the number of say unemployed youth per household that again may constitute a population for another study from the same census data we might get different definition of populations. Now, then coming to sample what is a sample?

Sample is a data set consisting of a portion of a population, but portion of a population means not any portion will just extract some portion and we get that is a sample no, that is not a sample. Normally, a sample is obtained in such a way so as to be representative of the population. Right, when I talk about the representative of the population, like let me give an example of exit poll what we do in exit poll?

So, suppose an election in an election if we were interested which party will be winning this election how we do? We definitely do not try to find out from each and every person who has come to cast their vote, we do not go and ask each and every one of them by. So, be the person who whoever some news media, whoever does the exit poll day just select a particular some person and from there try to get the information from them and that is the result of the exit poll from that day trying to infer on the whole election.

So, now, when they select these people, how do they select? The selected representative of the population when they select a representative meaning, representative meaning means they might select people from different age category they might seem select people from different profession, they might select people from different agendas mayors, and then they might select people from different area different locality this way, so, there may be so as they are the represents the whole population. Now, what is this statistical method?

(Refer Slide Time: 09:14)



So, when we talk of statistical method, method by which characteristics of a population are inferred through observation made in a representative sample from that population. So, the exit poll that was a very good example, like what we have done, we have taken the views of a representation of some people, from their views, we try to infer for the populations. So, method by which characteristics of a population are inferred so, method by which characteristics of a populations are inferred through observation made in a representative samples.

So, whatever observation we have made in the people which we have interviewed from there we are trying to infer for the whole populations. So, that is statistical method. Again and let me first give you some other example, like analyzing the quality of a manufacturing process if you are considering the manufacturing process, suppose, this manufacturing process does it has come out it manufacture in some batches.

So, when you consider the batches the material density is producing some material maybe the density may be different vary from batch to batch even while the some it produces or in continuous then also there may be differences in the material density from different items from each item. So, we need to analyze how, how much it is different from the target density. And then we have engineer should be able to realize where the process needs improvement so, as to bring the material to close to the target density.

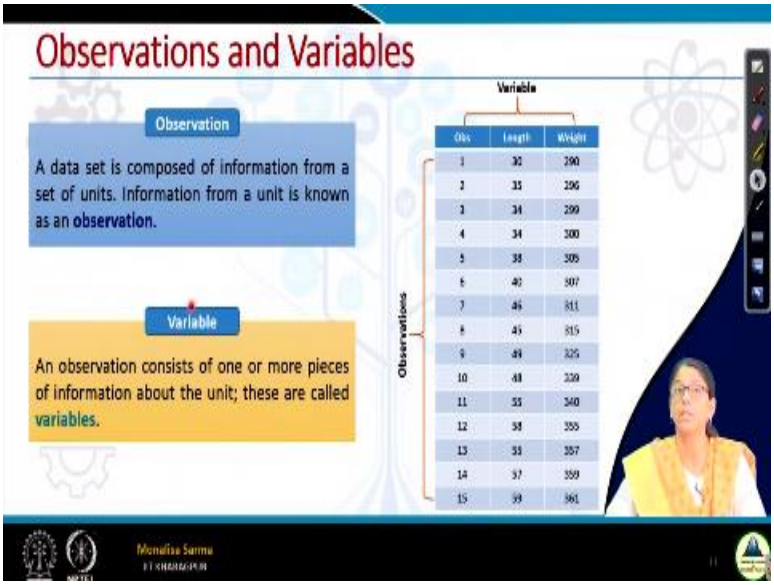
One another good example, conclusion about the efficacy of a new drug, let us take an example of a like a drug for curing headache suppose there is an old drug. So, which resists which gives 80% of the patients gets cured headache gets cured with this old drug and maybe this old drug we get at a cheaper price now, a new drug has come to the market. So, which it is a claim that it gives a success rate is 90%. However, the price is a bit higher.

So, what before bringing into market we need to check whether it really gives 90% of the success whether it really can achieve so, what it does, so, it has to test on few individually, it has to test on few subjects to find out whether it is really 90% efficient. Now, it has taken it has tested on a few subjects and it has found that it is 90% effective. Again suppose it has done the test again on

different set of subjects and it has come there suppose it is found, it is only 75% effective. So, there is some variance in that we got it the difference in effectiveness from study to study.

So, that is what come to the second point statistical methods are designed to contribute to the process of making scientific judgment in the face of uncertainty and variation. So, different study there may be uncertainty there may be variation. So, statistical methods are design to contribute to the process of making scientific judgment in the face of this uncertainty and variation. So, that is how we define statistical methods.

(Refer Slide Time: 12:27)



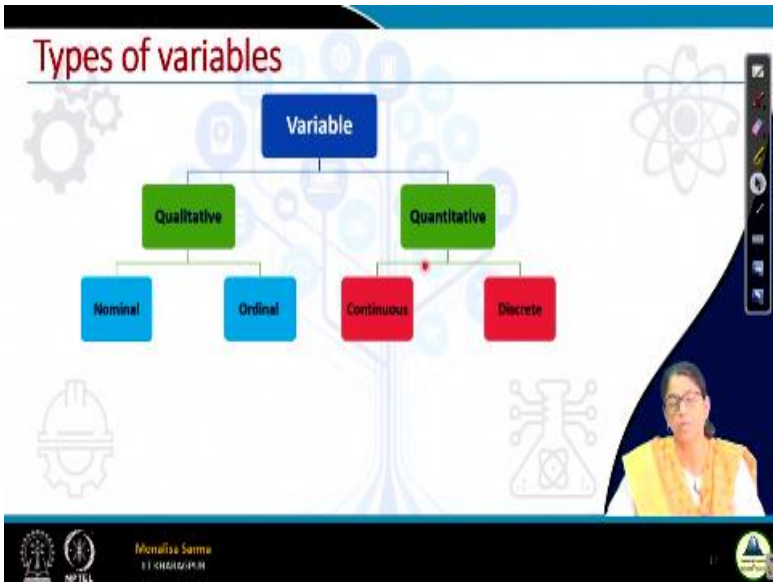
Now, some more definitions, like first is what is an observation? So, a data set is composed of information what we have already seen right from a set of units, now information from a unit is known an observation. So, now, let us take the example of these drugs that cures headache. So, what are the units in suppose we have interviewed 20 persons, so 20 total there are 20 units. So, 20 units, so 20 units means to how many observations? So, there are 20 observations each unit one observation.

So, now, next is that variable and observation consists of one or more pieces of information about a unit these are called variables. So, like this example, the drug for headache, what are the variables here variables may be the 2 variables, one is the drug use and another is the maybe the

whether the headache is cured or not, but a headache. So, this may be the second what to say variable. So, here there are total 20 which means we have interviewed 20 persons.

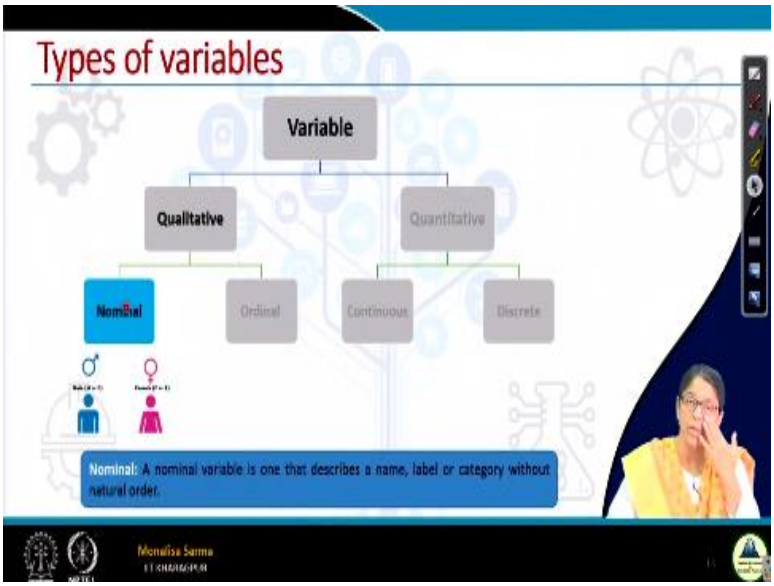
So, we got 20 observation and each observation there how many variables, there are 2 variables variable one is the drug use and another is the whether the drug is successfully using the patient determines whether a patient is cured or not cured, maybe the answer may be yes or no. So, there are 2 variables. So, similar here the example is given there are 2 variables this length is one variable weight is one variable. So, there are total say 15 observations here. So, each unit leads to one observation.

(Refer Slide Time: 14:02)



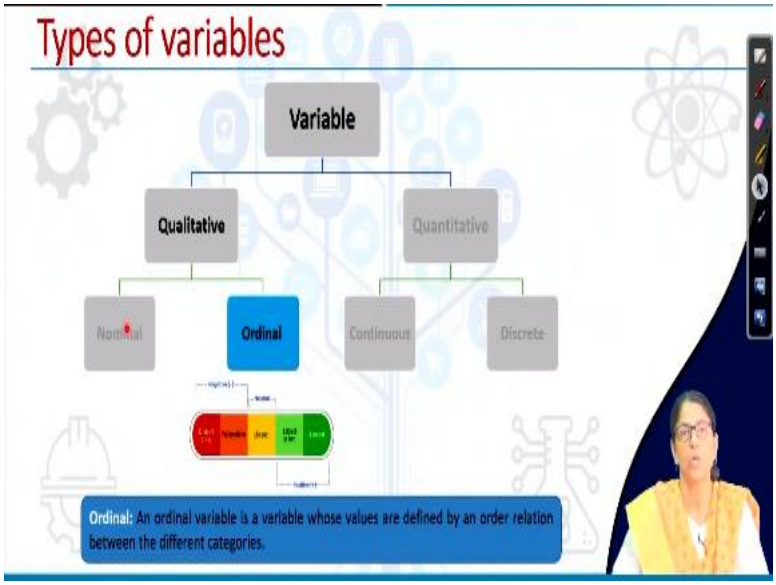
Now, this variables when you talk of variables there are different types of variables, there will be 2 types of variables one is qualitative variables it is also called categorical variable and another is quantitative variables. So, qualitative variable that is the categorical variables again there are 2 types of qualitative variables that is nominal variable and ordinal variable.

(Refer Slide Time: 14:27)



So, what is nominal variable? So, a nominal variable is one that describes a name, label or category without natural ordering like breed of animals, colours, different brands these are nominal variables.

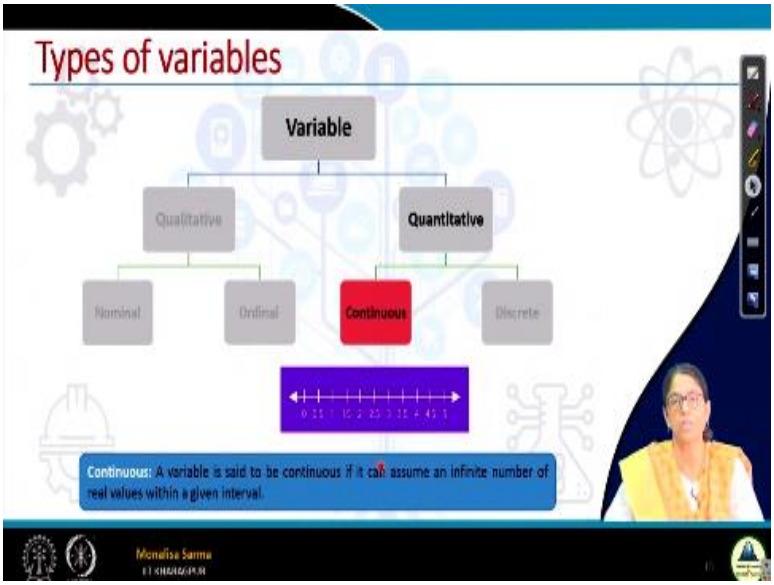
(Refer Slide Time: 14:43)



So, next is the ordinal variable, ordinal variable is a variable whose values are defined by an order relation between the different categories. Now, different categories like let me give you an example. Suppose I have asked one person, I have given him 5 different chocolate pies and I have asked him to do it in a scale of how much he has liked the chocolate pies I have given in a scale of 5 maybe he did not like it at all, he had like it bit like it he liked it very much.

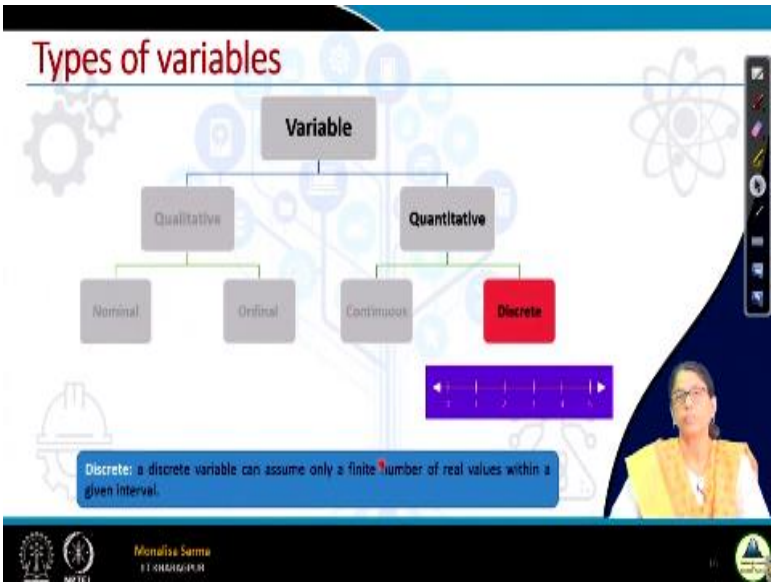
And it is excellent that maybe I have given 5 different scale and I have given him a 5 chocolate pies and accordingly rate this 5 chocolate pies in these 5 scales. So, what is this? There is a relation among the order it among the different categories. So, this is ordinal variable. In the first example, where does I have talked about a social science survey where we have armed 70 questions have asked one is the one question is the happiness index. So, how happy he is? Is not happy, pretty happy, quite happy. So that is again ordinal variable.

(Refer Slide Time: 15:47)



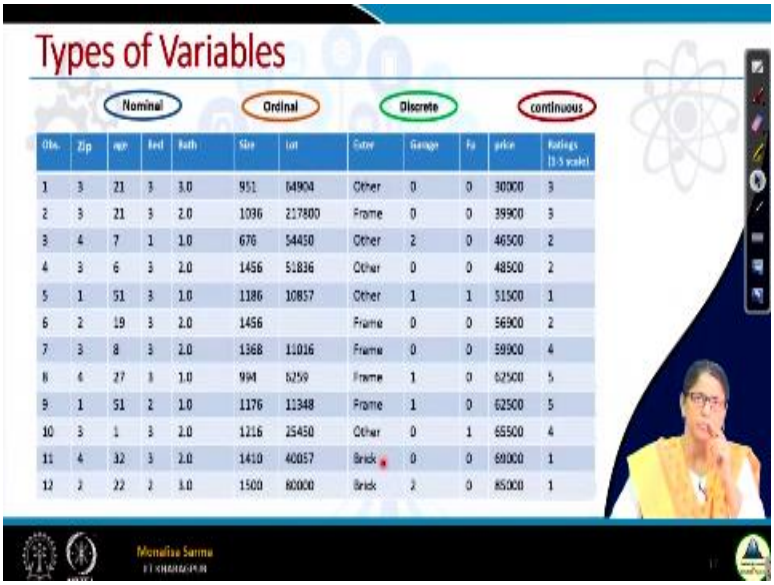
So, what is a continuous variable, continuous variable all of us of course, as we know it, so, a variable is said to be continuous if it can assume an infinite number of real values within a given interval, some example of continuous variable height of a person is a continuous variable, this car mileage is continuous variables.

(Refer Slide Time: 16:08)



Discrete variable, a discrete variable can assume only a finite number real values within a given interval. So, discrete variables like a number of persons in a family in our census data number of person in a family that is a discrete variable.

(Refer Slide Time: 16:22)

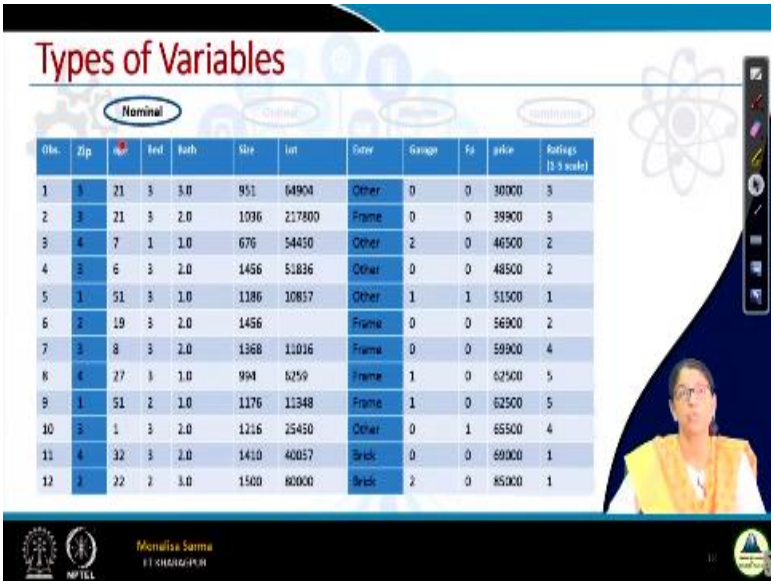


Obs.	Zip	age	Bed	Bath	Size	Lot	Outer	Garage	Fp	price	Ratings (1-5 scale)
1	3	21	3	3.0	951	64904	Other	0	0	30000	3
2	3	21	3	2.0	1036	217800	Frame	0	0	39900	3
3	4	7	1	1.0	676	54450	Other	2	0	46500	2
4	3	6	3	2.0	1456	51836	Other	0	0	48500	2
5	1	51	3	1.0	1186	10857	Other	1	1	51500	1
6	2	19	3	2.0	1456		Frame	0	0	56900	2
7	3	8	3	2.0	1368	11016	Frame	0	0	59900	4
8	4	27	3	1.0	994	6259	Frame	1	0	62500	5
9	1	51	2	1.0	1176	11348	Frame	1	0	62500	5
10	3	1	3	2.0	1216	25450	Other	0	1	65500	4
11	4	32	3	2.0	1410	40057	Brick	0	0	69000	1
12	2	22	2	3.0	1500	80000	Brick	2	0	85000	1

Here I have again explained the type of variables I have taken another one example, suppose this is a data of some houses which are sold in an area. In an area that is on the last set and yes, there are different houses that are sold totally you can see there are 12 observations. So, different houses are basically the same belongs to different zip code. So, we have different zip code 1 2 3 4 to maybe whatever they are, we have some zip codes and what is the age of the house? How old is the house?

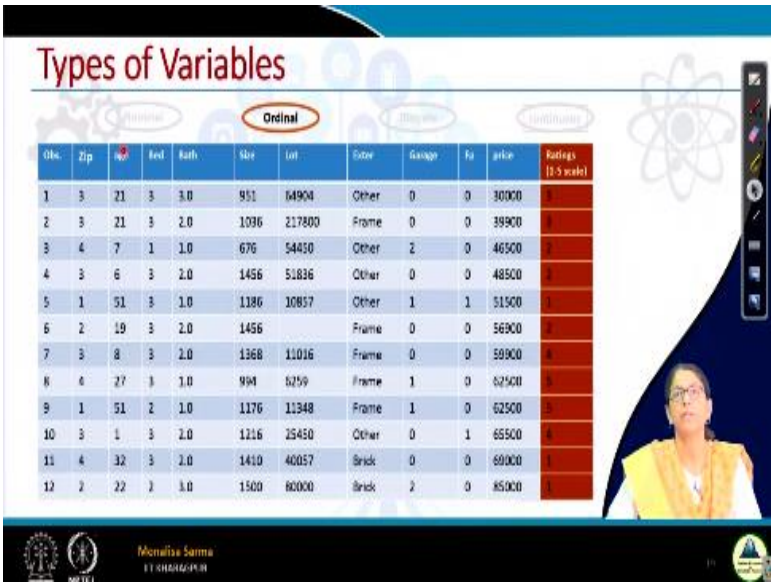
And how many bedrooms are there, how many bathrooms are there, what is the size of the house? what is the lot means what is the overall area of the property, then what is how the exterior is designed, then how many garages are there fireplace, price and ratings, when people have sell, give people has given some rating to the houses. So, this may be the different these are the different variables of these observations. To total 12 observation and how many variables zip is 1 2 3 4 5 6 7 8 9 10 11 total data 11 variables now we will see which is which variable.

(Refer Slide Time: 17:30)



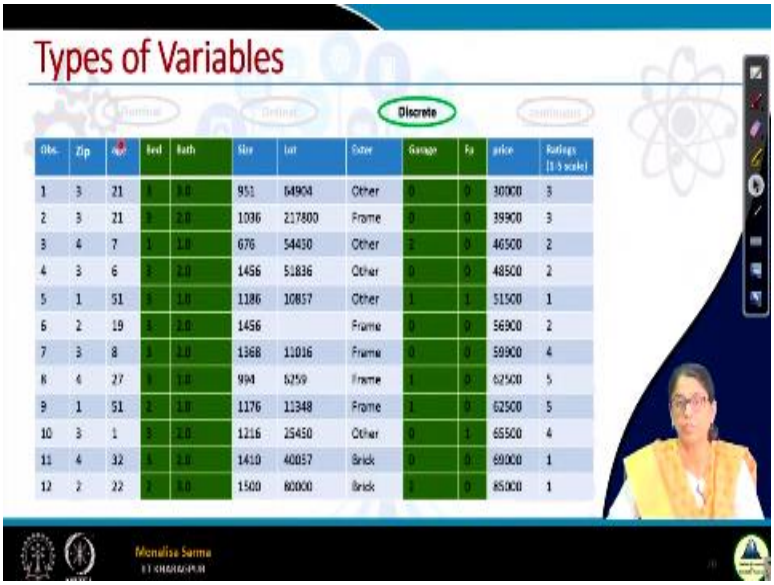
So, here zip is a nominal variable, right? There is a it is defines a different given different classes which are not it has it does not have any relation as such. So, then similarly external, it is also a nominal variable.

(Refer Slide Time: 17:45)



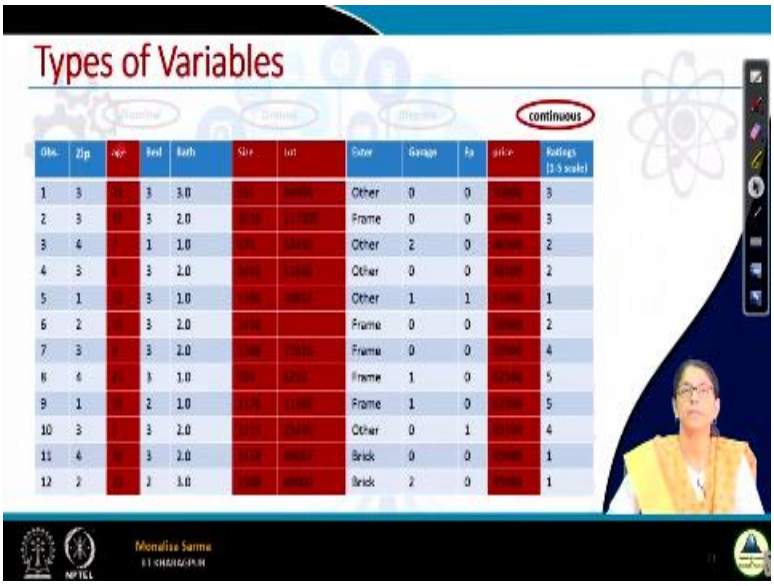
Then ratings that is an ordinal variable, there is some relative ratings is not it. So, that is an ordinal variable.

(Refer Slide Time: 17:53)



Then the number of bedrooms, number of bathrooms, garage, fireplaces are discrete variable this connect the continuous variables.

(Refer Slide Time: 18:00)



Than the size, lot size, price all this is these are all these are continuous variables. Now, if we see this type of table, as I have already told before this type of table we hardly can get any information like at the most the information that we can get is the price of the house may range from particularly 30000 to say 85000 or the houses maybe in the zip code of say 1 to 5 or something like that, not much information we can gather if this data is are given in this sort of big tables. So, we need to organize this data in some form or the other for this have we needed data organizations tool.

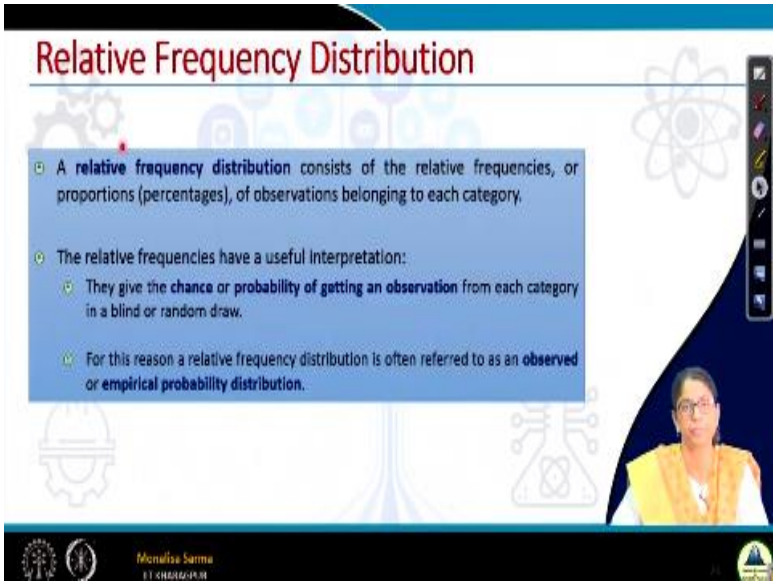
(Refer Slide Time: 18:40)

So, one of the data organization tool is the frequency distribution. So, what is frequency distribution? Frequency distribution is constructed by grouping the data into categories, we can

group the data into categories like how what are the what may be the categories, like we have seen the different zip codes where houses belonging to different area different areas, different zip codes. So, we are all different, we can if we can group the houses into different zip codes.

That is we are grouping the data into different categories based on the zip code. So, a frequency distribution is a listing of frequencies of all categories of the observed values of a variable.

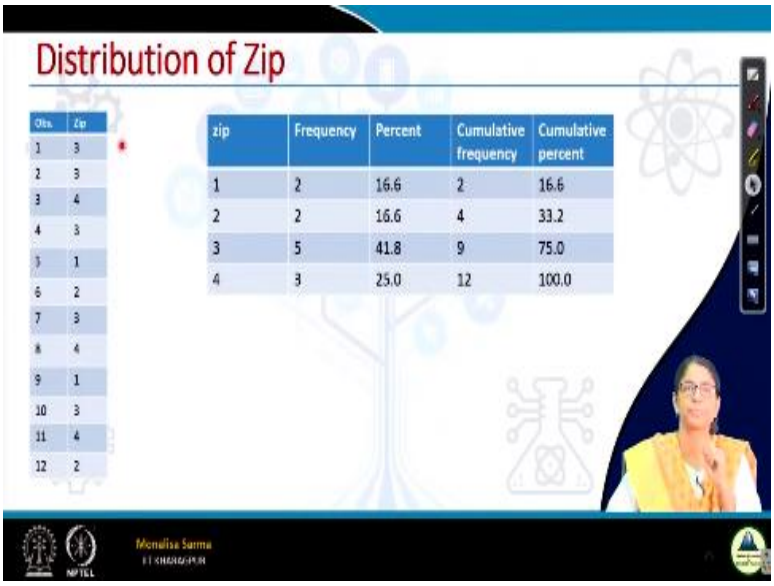
(Refer Slide Time: 19:19)



Then we have something already relative frequency distribution, relative frequency distribution is like it is nothing but it consists of relative frequency or the proportion, proportion of observation belonging to a each category like relative frequency distribution again, it can we can also call it as in terms of percentage, a relative frequency distribution again, it has some useful interpretation also. So see, they give the chance or probability of getting an observation from each category in a blind or random draw.

So, for this reason, a relative frequency distribution is often referred to as an observed or empirical probability distribution.

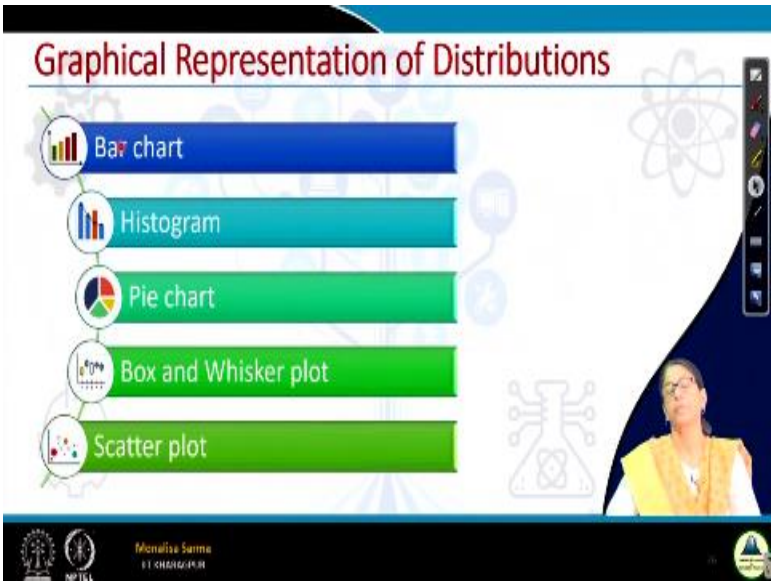
(Refer Slide Time: 20:03)



So, there is an example of how we can categorize based on this zip code what I have just told. So, these are the different observations. So, based on different zip code, we have found out total zip in zip code 1 suppose there are 2 houses, zip code 2 there are 2 houses zip code 3 there 5 similarly 4 there are 3 houses, so, this is one set of category and this next is the percent that means that the relative frequency that is the proportion 16.6% of the houses total houses that are sold 16.6% of the houses belong to zip code 1 then 41.8% belongs to the zip code 3.

Then we have called cumulative frequency, cumulative frequency the first column you can see this, so, 2 so, what this second let us see the second column cumulative frequencies 4 that means, number of houses that belongs to zip code 2 and 1 together that is 4 that is the cumulative frequency. Similarly, cumulative percentage that is cumulative of relative frequency.

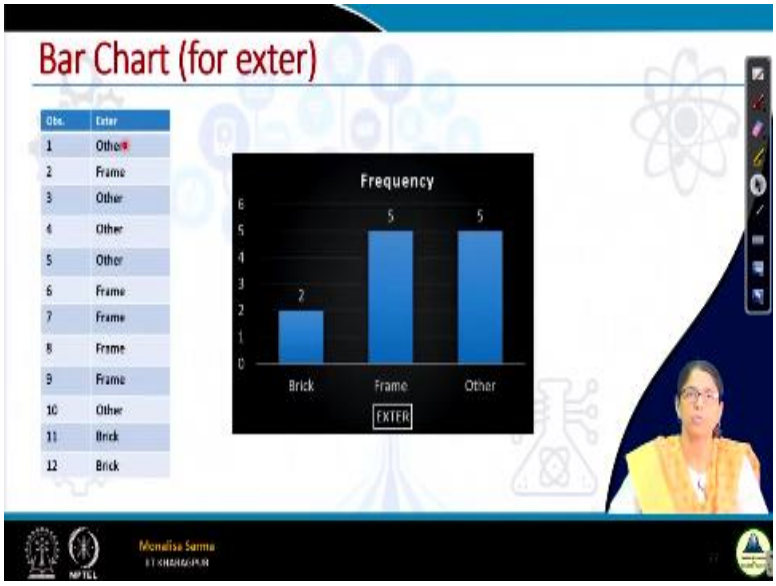
(Refer Slide Time: 21:04)



So fine, now a picture is equal to 1000 words and if we in fact instead of writing something, if a figure is given to us, it is more easier to grasp it is more easier to understand the things so, like this data organizations, if instead of just using it in some tables, if we put it in some sort of figure, some sort of graph, it is more easy to understand. So, we have some graphical representation of this distribution as well as distribution I mean the frequency distribution.

So, what the different graphical representation may be bar chart histogram pie chart, box and whisker plot, scatter plot there are some other graphical representation as well. So, we will be just discussing this.

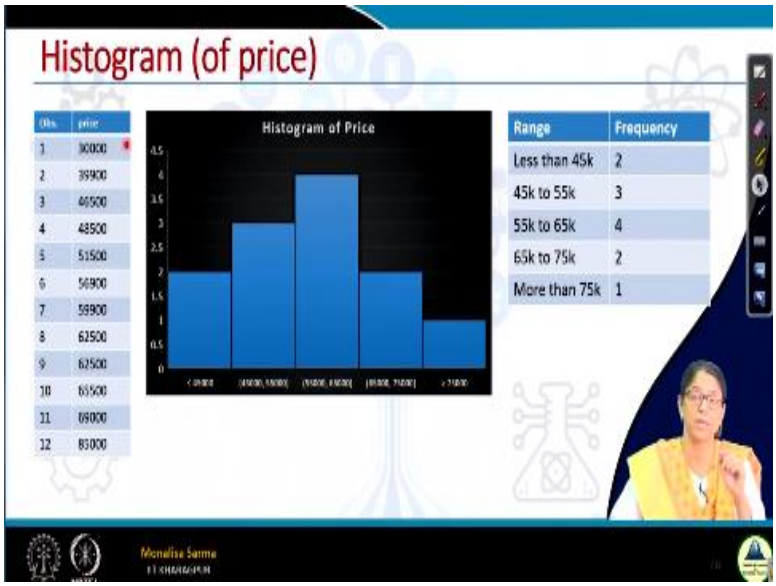
(Refer Slide Time: 21:47)



So, when we talk a bar chart bar chart is applicable only for categorical variables what I mean by categorical variables means qualitative variables like for nominal variables for ordinal variables we use bar chart. So, here suppose for the exterior we have seen when seeing the figure we have drawn the figure for exterior so, what are the different types of exterior we had that is brick we have frame and one others. So, this is how we draw the bar chart in the y axis we get the frequency and the x axis we define the class.

So, your brick one class frame is another class others is another class and in a y axis we give the frequency. Now here the width in the bar chart width has no meaning as such, but it is just for to look it beautiful. So, as the picture does not look cluttered, we usually give the weight of each class we keep it same. So, this is the bar chart.

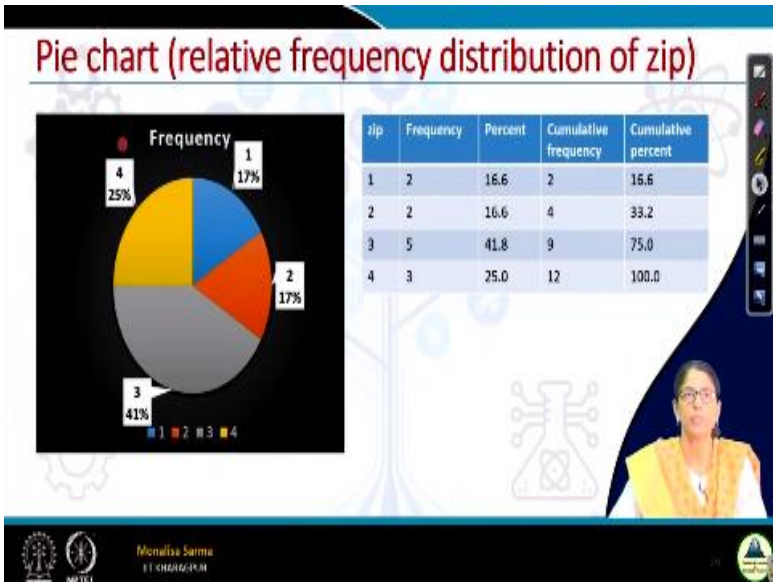
(Refer Slide Time: 22:45)



Now, if you are interested in putting in a graphical representation of this continuous data or discrete data we use histogram. Histogram is same as bar chart just that same in the sense y axis we have frequency x axis we have the different basically here x axis we have the range of data instead of class we have the range of data. So, here the width represents the range here the width has meaning. And you see here between each block there is no space because it is continuous.

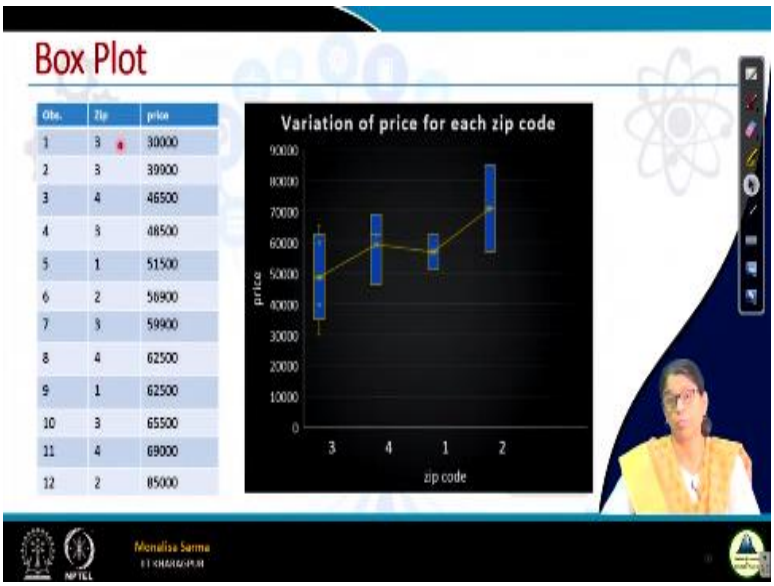
So, this is the histogram. So, we have done the histogram basically for the price histogram for the price that is for less than 45k then 45k to 55k then 55 to 65k and so on and so forth.

(Refer Slide Time: 23:36)



Then pie chart, pie chart is nothing but a circle we divide into few slices, each slice represents a proportion of the each slice represented basically category. So, category is basically the proportion, proportion of occurrence like here we have seen the for different zip code, what is the percentage of it houses are in different zip code, that is the relative proportion so pie chart, we can use pie chart to visualize.

(Refer Slide Time: 24:03)

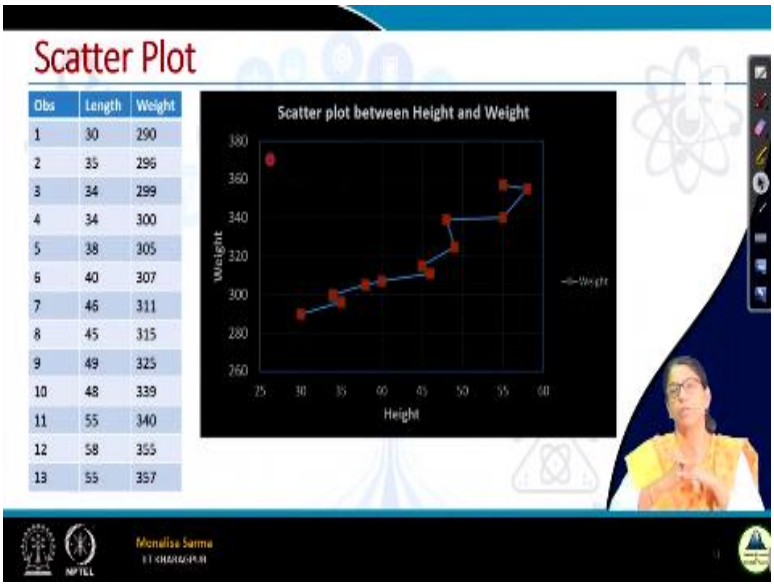


So, then we have box plot, box node is also called 5 point summary and a box plot how we draw a box plot and there are basically 3 horizontal lines, the down horizontal line is the 25th percentile, and the upper horizontal line is the 75 percentile and the middle line is the median and

then we have a whisker this is coming down, you can see a vertical line which is coming down. So, this vertical line this is coming down and here it plots the lowest value and then we have the highest value is the again a vertical line which is going up which plots the highest value.

So, here we have tried box, box plot we may use to compare different set of data and different distributions, so here we have tried to compare variation of price for each zip code for each zip code what are the variation or prices like for zip code 3 we have seen this is the variation of prices, we have drawn the box plot similarly, this is the variation of prices for box plot 4. So, when we have this sort of figure we can easily compare how the price is different from different zip code. So, this is the box plot.

(Refer Slide Time: 25:18)



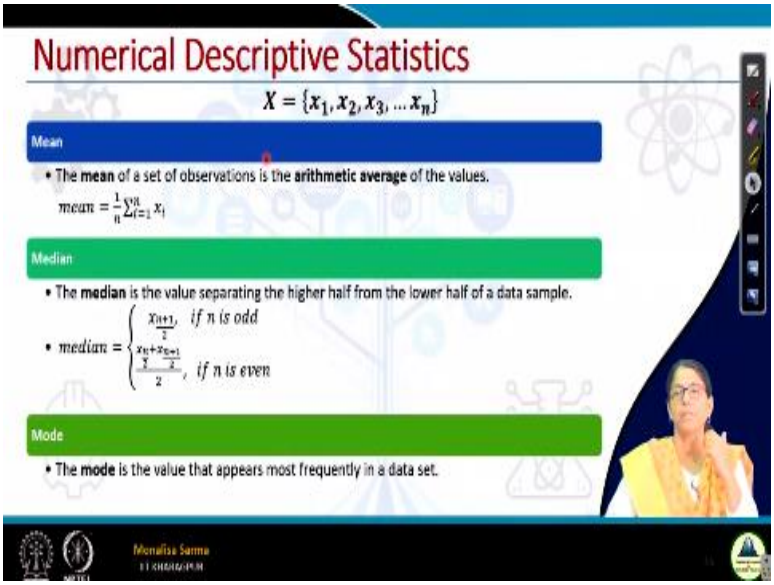
Then we have scatter plot, scatter plot is like when we try to find out how one variable varies compared to other variables like how height varies compared to weight or weight varies with in accordance to height. So, this is how we plot scatter plot. So, now, okay, graphical description is okay, so again graphical description we get lots of information however some sometimes we do not need such we do not need lots of information.

We may need a simple one just one data, we may need some summary data of some certain population, we just need a summary data then we do not go for a graphical description of this

data then what we will do? We will take help of descriptive statistics to describe the characteristics of our populations okay.

This graphical description also does that it describes the characteristics of the population, but we have lots of information there, but if we want just a one value to describe the characteristics of the population this one value can be anything.

(Refer Slide Time: 26:23)



So, that the different descriptive statistics may be mean all of us so, we know what is a mean? Mean is a set of observation of a set of observation is the arithmetic average of the values all of us we know what is arithmetic mean. So, then median; median is the middle value. Now, if given a set of data, given a set of data, whether we will be considering the mean or median that is also that also it is very much subjective like if there are many outliers like let me take consider the example of a rainfall.

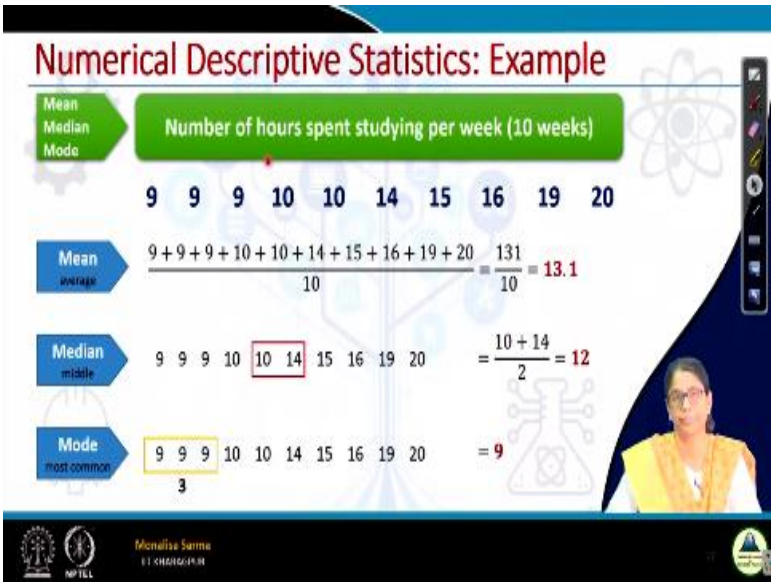
Rainfall in an area, rain fall in an area if we consider say suppose in a 30 day period 30 day period if we want to find out what is the rainfall in the area, average rainfall in an area, suppose there are lots of outlayers like sometimes there is no for somebody and then there is no rainfall at all and sometimes some days they are very huge rain fall say around to the extent of say around 120 Mm, or 200 Mm, some like something like that.

So, then if we try to find out if we try to take the mean, the mean will not be a proper reflective of the picture, because there are some very small data, some very high data. In that case, maybe we may go to the to find out the median and what is the middle data means in a 30 days period we will see in the 15th day what is the data, but then median the disadvantage with a median is that median does not take care of all the values when we consider mean, we take the average of all the values.

But when we consider median, we just take the middle value, middle value, if there is a odd number, we take the middle value, if it is even number, we take this and by 2th value and n plus 1 by 2th values, and we take the average of these 2 values. So that is the median. So that basically in which situation we will use mean in which situation you will use median that is totally very subject and dependent on the situations.

Then what is mode? Mode is the value that appears most frequently in a data set is a value that appears most frequently the data set, that appears most frequently the data set.

(Refer Slide Time: 28:24)

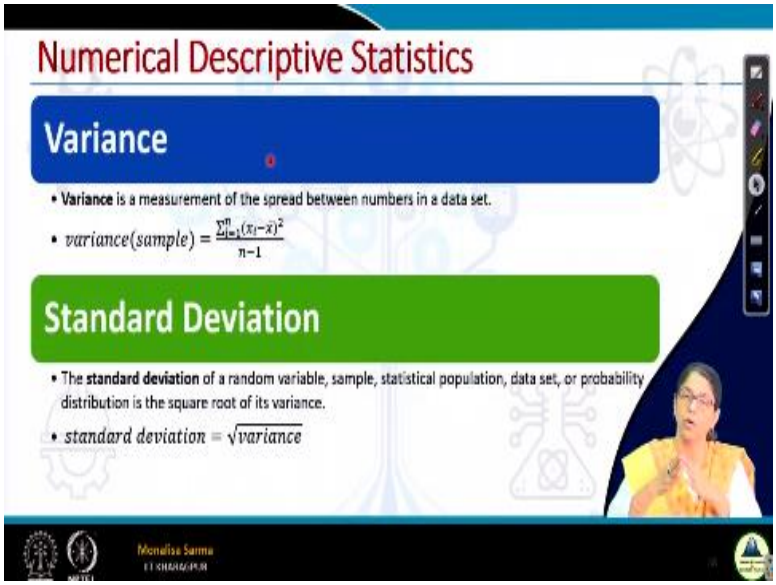


I will show you some example. Like suppose here is the data for number of hours spent studying per weeks and is number of hours in the given for certain weeks. That is in the number of hours a student has spent in studying. So, if I am interested in finding the mean, how do, I find the mean, I will add up all this data since it is a data for 10 weeks, then will be divided by 10. That is how I

got the mean, if I am interested in the median, so there are total 10 data sources there is an even number I will have to consider 2, middle data.

So, to the middle data that is the 10 and the 14 as a red dot the data in the red box. So that is 10 + 14 / by 2 that is my median is 12. Now if I am interested in finding the mode, mode see, your 9 has appeared more number of times so my mode is 9.

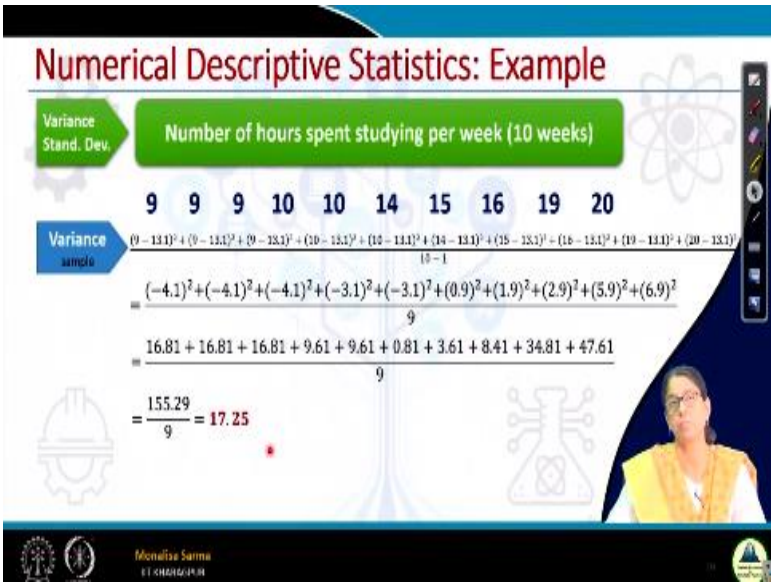
(Refer Slide Time: 29:14)



Now, there is one more descriptive statistics that is variance, variance basically it gives us spread how spread out my data is how my data is different, how much spread out my data is from the mean compared to the mean how spread out the my data is compared to the mean. So, how do I calculate variance? This is the formula for calculating variance. All of you of course, that variance standard deviation mean, median, mode, you have studied in class level only.

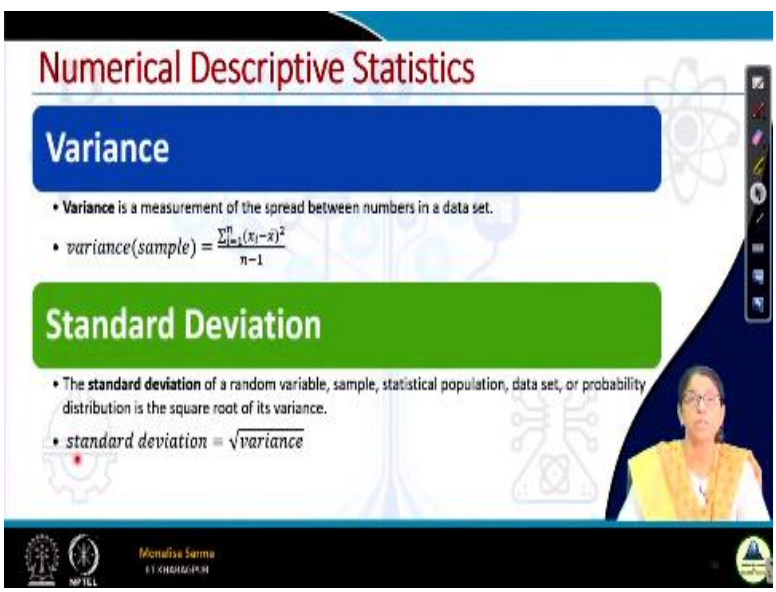
This just a quick recap basically. So, this is variance we have calculated a variance. So, then from the variance, how we calculate the standard deviation? Standard deviation is nothing but the square root of variance.

(Refer Slide Time: 29:58)



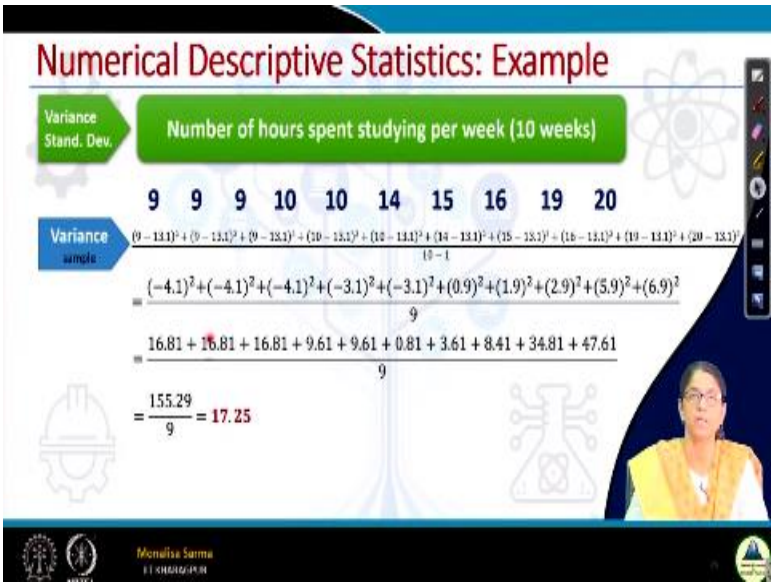
So now, we can again the same example, but what we have considered for number of hours spent studying per week that is for 10 weeks. Let us take the same example for same example how do I calculate the variance?

(Refer Slide Time: 30:11)



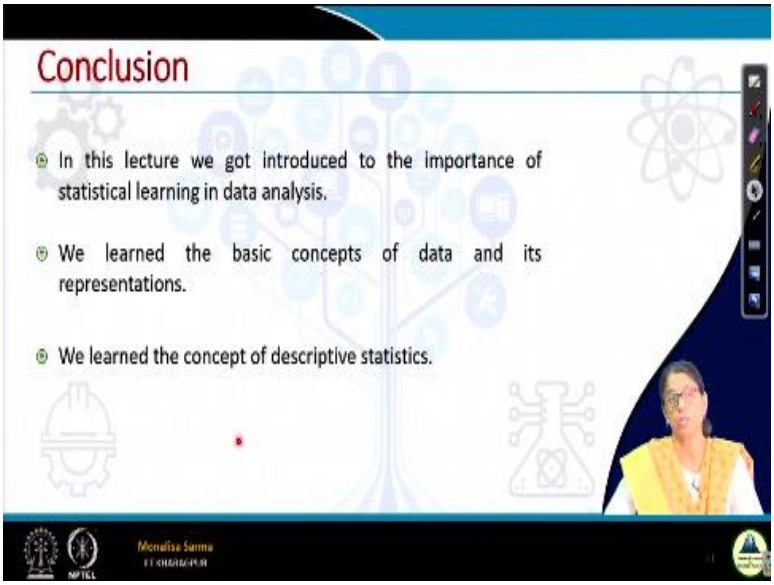
You see in the variance what is the definition for variance we for each data, I try to find out how much each data is different from the mean that is the x bar, x bar is mean. And since I may get negative values as well as positive values, so, negative values may bring out to be 0 that is why I have squared it that is you are see this a square and this is the what is the degree of freedom degree of freedom is n - 1. So, it is I have divided by n - 1 see I have not divided by n. So, since the degree of freedom is not n degrees of freedom is n - 1. So, this is my variance.

(Refer Slide Time: 30:46)



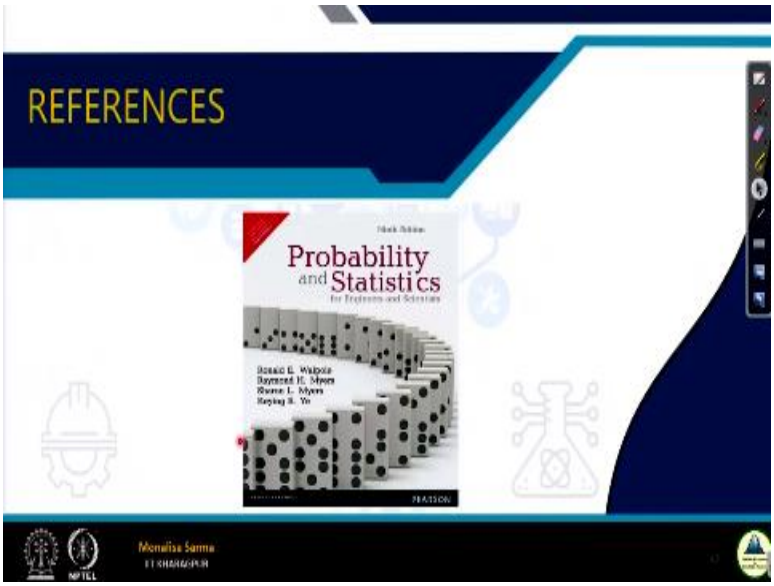
So, here so, accordingly what is my mean I have calculated mean earlier it is 13.1. So, I have substituted each data from 13.1 and divided by 10 - 1 total this 10 data so, and I got this variance now, from the variance I could calculate the standard deviation.

(Refer Slide Time: 31:04)



So, in this lecture we got introduced to the importance of statistical learning and data analysis and reliability analysis whatever it is, let me tell you the general terms data analyzes this because reliability means we will be analyzing the reliability data only. That is where we will be using statistics. So, we learned the basic concept of data and its representation. We also learned the concept of descriptive statistics.

(Refer Slide Time: 31:32)



In the reference for this lecture, as I have taken the help from this book probability and statistics for engineers and scientists by Walpole, Myers, Myers and Ye is a very good book. Thank you.

Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology – Kharagpur

Lecture – 03
Concept of Probability and Probability Theory

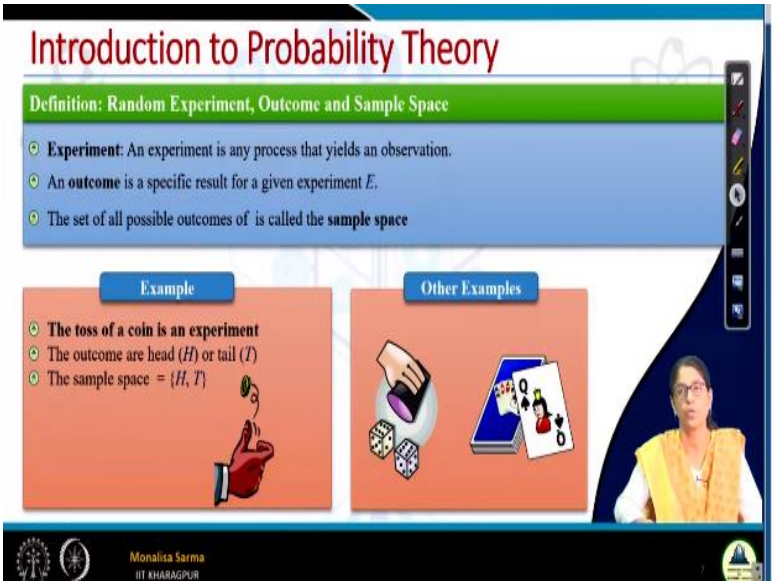
Hello everyone. So, in today's lecture we will be learning probability theory. Now the question is why we will learn probability theory in this course on statistical learning for reliability analysis. Nothing is that in this while we are learning statistical methods that are in this course on statistical learning, we will be doing inferences on many applications. After that, we will have to evaluate the reliability of these inferences.

To evaluate the reliability of these inferences, we will be reading the concepts and principles of probability. And that is the reason we will be discussing some concepts and theories of probability in the coming 2 lectures.

(Refer Slide Time: 01:04)

So in this class, I will cover what is an event, then I will cover different types of events like complement, union, intersection and also mutually exclusive event, exhaustive events. Then after that when of course, I will also cover when we will be discussing probably definitely we will have to talk about sample space also, now sample space how to calculate the sample space for that also we will be discussing few points that is counting sample points and then we will be discussing few of course 2, 3 problems also in this lecture.

(Refer Slide Time: 01:39)

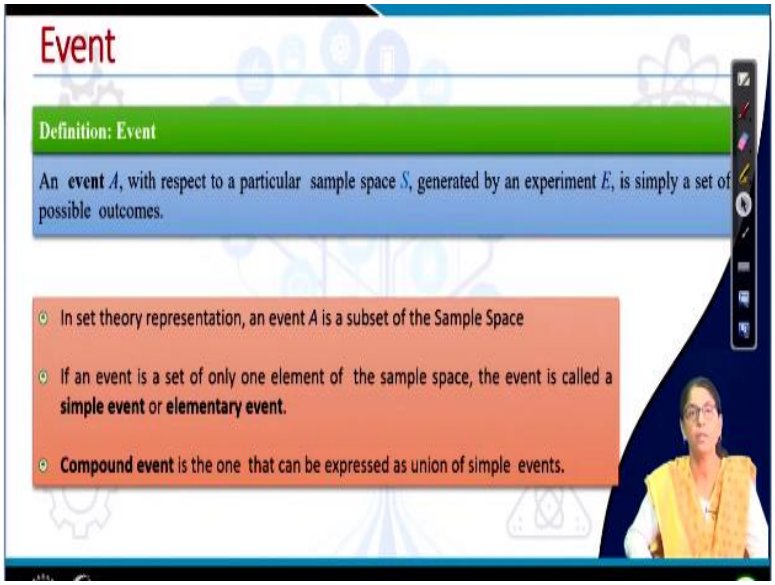


So to start with, we will go through a few definitions before learning what is the probability theory, the first definition that we need to know is and what is an experiment. So, an experiment is any process that yields observation or I can say an outcome. So, what is an outcome then? An outcome is a specific result for a given experiment E. Now, what is sample space? The set of all possible outcomes is called a sample space.

So, first is, we need to know what is an experiment. So, once we know what is an experiment, so experiment is output outcome, then outcome then an experiment they can be different possible outcomes, all the outcomes put together we call it as a sample space. So, there is an example suppose tossing of a coin simple tossing of a coin is an experiment. So, what if we toss a coin what output we can get one observation that can bring come head or tail?

So, these are the 2 outcomes so then what is the sample space is? Sample space is a set H and T. Similarly, another example of a tossing of a dice, that is an experiment, so what are the different outcomes? Outcome is getting different values from 1 to 6 and so what is the sample space? Sample space is 1, 2, 3, 4, 5 and 6 this is the sample space.

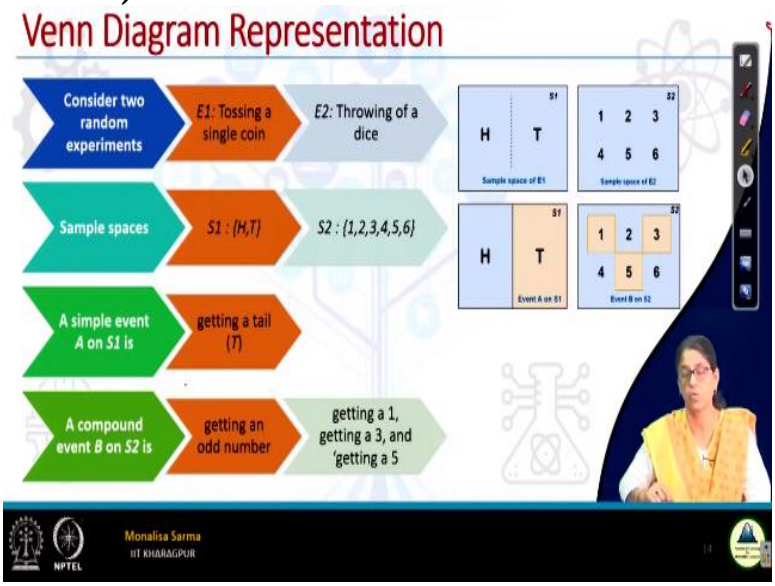
(Refer Slide Time: 03:04)



Now, let us see what is an event? So, an event A with respect to a particular sample space S generated by an experiment E is simply a set of possible outcome. So, an experiment for an experiment there can be different events, so for any event A, it will be basically a subset of the sample space. So, an experiment can have different outcomes or different events. And all events have some order other value in a sample space. So, for an event, what we get? It is simply a subset of the sample space.

So, set theory representation and event A is nothing but a subset of a sample space. Now, if the event is a set of only one element, then we call it a simple event. And if it is a union of many simple events, then we call it a compound event.

(Refer Slide Time: 03:58)



So, we can see the Venn diagram representation of this, consider 2 random experiment like tossing a coin or throwing a single dice, if we toss a coin so what is the sample space, we

have already seen sample spaces H and T. A sample space contains 2 elements and similarly tossing a dice let us I should say throwing a dice, throwing a dice there is a sample space sample space consists of 6 elements, so 1, 2, 3, 4, 5 and 6.

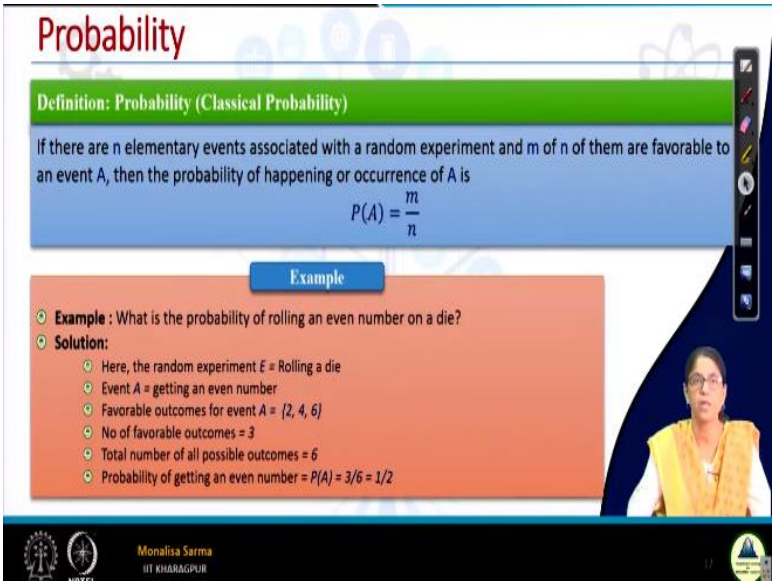
So now, what is an event? Now we will see that event is a simple event A1 S1. Suppose I am talking about simple event on the sample space S1. So, events say my one event maybe getting a tail. And of course, an experiment that is tossing a single coin. They can be just 2 events, either getting a tail or getting a head there can be just 2 events. So, if I talk about getting a tail, so my event is my sample spaces this subset of this, this is this T.

Let me take the pen so I will be able to point it properly so this is T ok, so once again this is not writing properly, so this is the sample events for getting a tail now, if I considered a second sample space this is S2 there may be different event for this the event maybe getting a particular number getting 1 getting 2 getting 3 like that event maybe getting a number or even maybe getting an even number, even maybe getting an odd number, even maybe getting as number combination of these numbers like what is ?

If I am interested in finding event of getting a total of 7 if I total of 7 that will be discussing, what is the event of getting a total of 7 in a sample space S2. I will come to that what is that event? So, 7 is there in the sample space it is not there. So, we will see what is that. Now, we will talk about compound event. So, in a compound event on B I will talk of compound event on B so getting an odd number, what is getting an odd number in sample space B that is a how many odd numbers are there that is 1, 5 and 3.

So that is the set of events that is propagating an odd number, this is getting 1, 5 and 3, that is the subset of that set as to which corresponds to this event.

(Refer Slide Time: 06:38)



Now we will define probability. So, if there are n elementary events associated with a random experiment, like as I told you in a first sample spaces S1 which we have discussed in this previous slide, sample space S1, how many probable events are there probably event 1 event of getting a tail another event of getting a head. So similar, that is what I am telling if there are n elementary or events associated with n random experiment and m of the n of them are favourable to an event A.

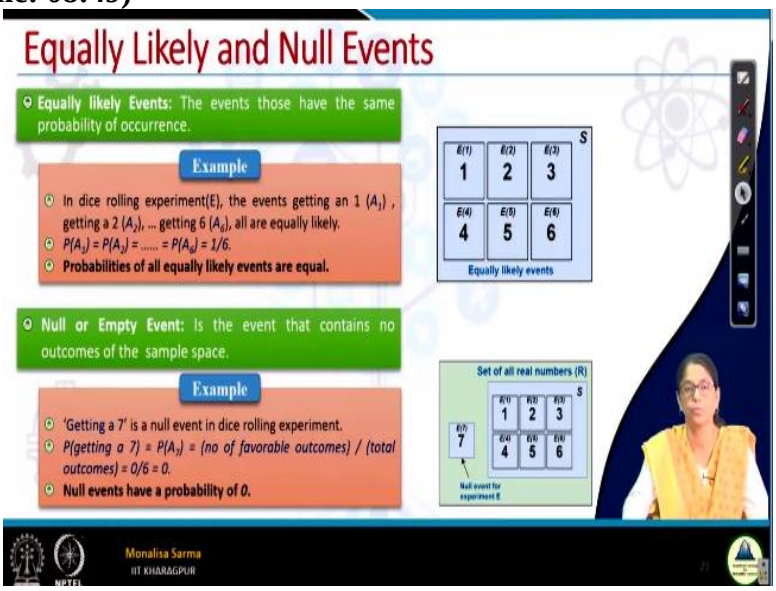
Any event out of this n events, we are just interested in even a so if we are interested in finding out the probability of A so to find out the probability of A, then we have to find out the out of this sample space which favours the event A, so then if there is n elementary event associated with random experiment and m of n of them are favourable to an event A. Then the probability of happening or occurrence of A is nothing but m by n.

Where n is the size of the sample space and m is the size corresponds to an event A, m is basically the number of elements that corresponds to the event A number of elements in the sample space. So, like, what is the probability of rolling an even number on a die? This is an event. So, what is the event? Event is finding an even number. So, how many elements correspond to this in a sample space of rolling a die there it is 6 sample space size of the sample space is 6.

So, now what are the; how many elements correspond to this even rolling an even number that is 3. So, probability will be 3 / 6, see so your random experiment E is rolling a die, event A is getting an even number, favourable outcomes for even A is 2, 4, 6, number of favourable

items total 2, 4, 6 total number is 3. So, total number of possible outcomes sample space that is the size of the sample size is 6, so probably getting an even number is 3 by 6.

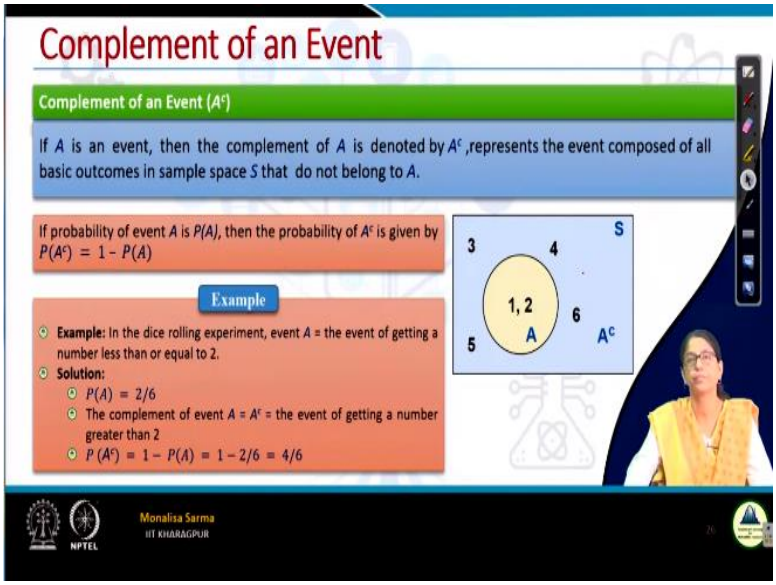
(Refer Slide Time: 08:49)



Now what is an equally likely event? The events, those are the same probability of occurrence then we call it equally likely events. Like when we toss a coin, the probability of getting a head and probability getting a tail it is the same probability. So, it is called equally likely events. Similarly, tossing a die probability of getting any number is equally likely. So, in a die rolling the events of getting 1, 2 or 3 anything 1, 2 up to 6 are all equally likely.

So, what is this probability? Probability is equal to 1 by 6 so this is the probability of all like equally likely events are equal. Now null or empty event remember, when I asked in a sample space S2 what is the probability of getting 7, so that was a I was telling that is an event when probability of getting a 7, so 7 available in the sample space it is not there. So that is an empty event or null event It is the event that contains no outcomes of the sample space So, getting a 7 is a null event in a dice rolling experiment. So, null event have a probability of 0.

(Refer Slide Time: 10:01)

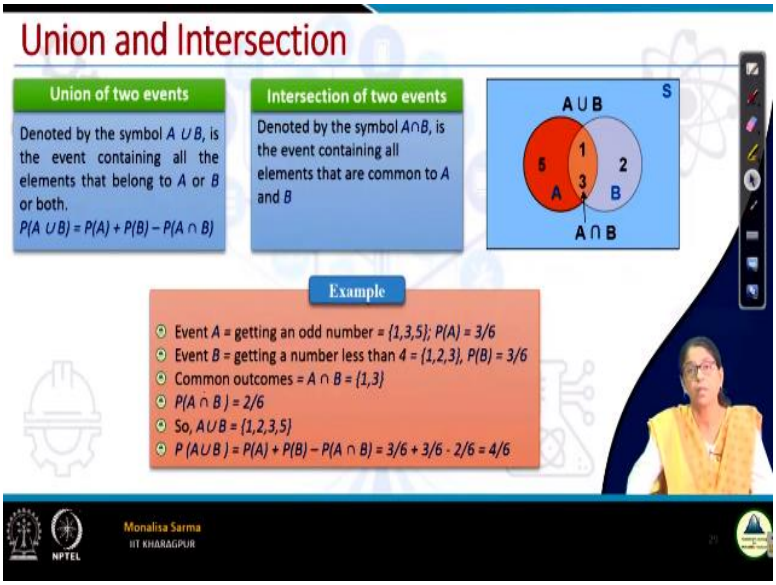


So, compliment of the event, it is easier if A is an event compliment of A is just 1 minus because probability will always range from 0 to 1. So, if one we know the probability of a certain event happening, and then what is the probability of that event not happening is just 1 minus of that event that is the complement. So, if A is an event then a complement of A is denoted by A complement represent the events composed of all basic outcomes in a sample space that do not belong to A.

So, 1 - probability of A, probability of A compliment is 1 - probability of A so example, in the dice rolling of when event A is the event of getting a number less than or equal to 2. So, if I am interested in the event of getting a number less than or equals to 2, how many numbers corresponds to this less than or equal to 2 that is 1 and 2. So, size of this event is 2 and what is the size of the sample space? Size of the sample space is 6.

So, what is the probability of that probability that is 2 / 6. So, this is, this light yellow coloured circle, what you can see this is the this corresponds to the even that event of getting a number less than or equal to 6, then in this the whole sample space which is shown in a light blue colour, this whole sample space and this is the whole sample space, this is S and this circle, this small circle corresponds to event A then what is A complement? A complement is this 1 minus of this A that means the out of portion of this circle. The complement of event A is the event of getting a number greater than 2. So, what is 1 - P (A).

(Refer Slide Time: 11:55)



Now union of 2 events denoted by the symbol A union B is the event containing all the elements that belongs to A or B are both. Union of 2 events means all the elements containing to event A event B or both. So, A union B is P(A) + P(B) and why we are doing -P(A intersection B), because there are some elements which belongs to both A and B, those elements will be counted twice. So that is why we are just subtracting it -P(A intersection B).

So, intersection of 2 events, intersection of 2 events is the event containing all the elements that are common to A and B. So, if you see the Venn diagram, so here say red colour corresponds to event A and this violet colour it corresponds to event B, then this orange colour, this portion it belongs to both A and B. So, when we are trying to find a probability A or B, so this orange portion we counting it again, you are counting it twice, we have counted it for when we have counted for A.

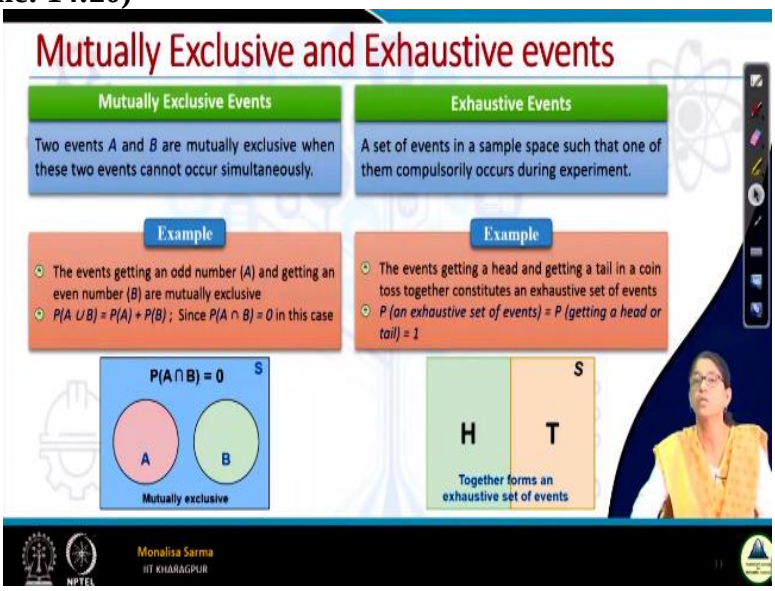
We have counting it again when you have counting for B, so we are counting it twice, but now we need to count it once. That is why we are again subtracting once that is minus of A intersection B that is why we have this minus of A intersection B in probability of A union B. So, simple example, even A maybe getting an odd number, what is the n rolling of a die? So, what is the odd number enrolling of a die is 1, 3, 5. So, what is property of A is 3 by 6.

Similarly event B means getting a number less than 4, what is the number less than 4 is 1, 2, 3. So, probability B is again 3 by 6. Now I am interested in finding out the common outcomes that is probability A intersection B, probability A intersection means that is which is common

to both A and B, if we can see which are the common to both A and B, A and B 1 and 3 is common to both A and B. So, it will be 2 by 6.

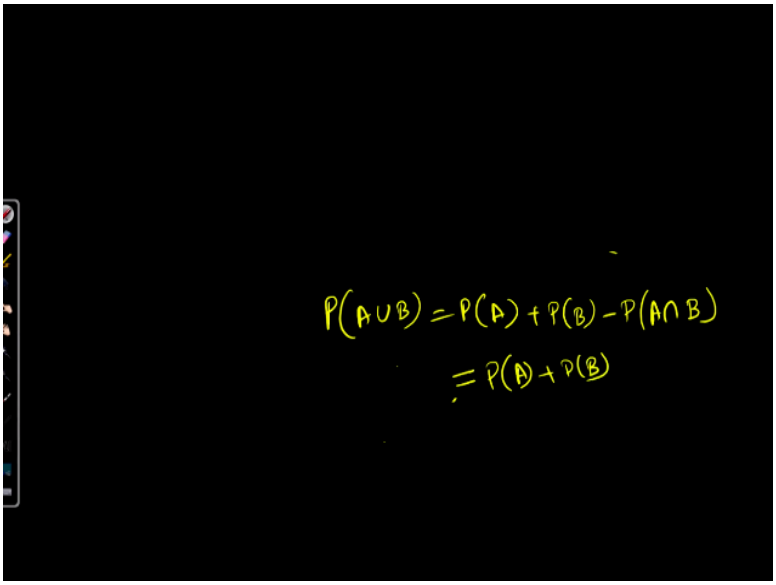
And now if you are interested in finding A union B, so union name is added up if you see 1 and 3 will be added twice so we are again subtracting that, so probability of A union B is probability of A + probability of B – probability of A intersection B that is why we got 4 by 6.

(Refer Slide Time: 14:20)



Mutually exclusive event, mutually exclusive events are when 2 events cannot occur simultaneously when we toss a coin at the same time head and tail both cannot occur. When you toss a tail at the same time 1 and 2 both cannot occur, any 2 numbers cannot occur at the same time. So, those we called as a mutually exclusive event. So, in case of mutually exclusive events like when we have done A union B.

(Refer Slide Time: 14:46)



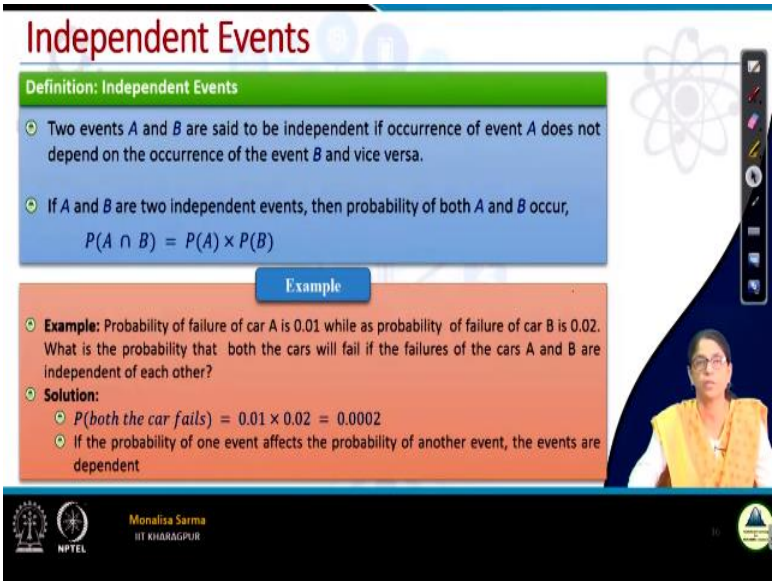
So, what we have got initially probability of A union B = probability of A + probability of B - probability of A intersection B. Now, if then mutually exclusive event, mutually exclusive events means both cannot happen together that means probability A intersection B is 0. So, if A and B are mutually exclusive events, so probability of A union B, we will just get probability A + probability of B.

So, the events getting an odd number and getting an even number, we at the same time we cannot get an odd number as well as you cannot get an even number. So, these are mutually exclusive event. So, probability of A union B is a simple probability of A + probability of B what we have just now discussed. Exhaustive events, exhaustive event is a set of events in a sample space such that one of them compulsorily occurs during experiment.

Actually when I when we talk about exhaustive events means all possible events that may happen in an experiment. In an experiment what events that can happen? So, this is we are considering all the events. So, when in an experiment when we are considering all the events, when we are doing an experiment one of the event has to occur. So, that exhaustive event is a set of events in a sample space such that one of them compulsorily occurs during an experiment.

The events of getting a head and getting a tail in a coin toss together constitutes exhaustive event. So, what is the probability of an exhaustive event is always one because it is the whole sample space basically.

(Refer Slide Time: 16:36)



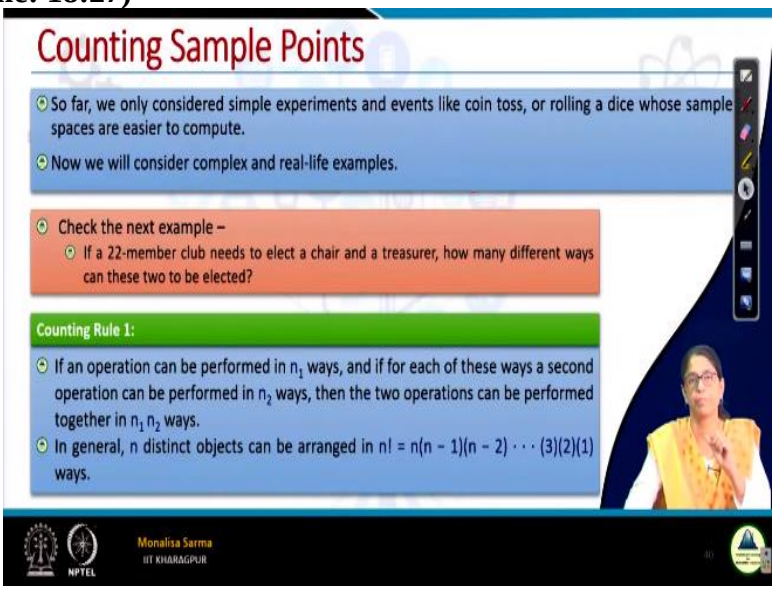
Now, the definition of independent events. So, if 2 events A and B are said to be independent, that is the occurrence of event A does not depend on the occurrence of event B that the event A has occurred, it will no way affect the occurrence of event B then we call as 2 events are independent. If then in that case, if you are interested in probability of occurring of the both the events is simply the multiplication of both the events.

So, probability of independent events independent probability of both events occurring we have seen when we try to find out probability about events occurring, it is intersection right . So, probability of A intersection B is probability of A x probability of B if they are independent events. A very good example is here, probability of failure of car A is 0.01 while the probability of failure of car B is 0.02 these 2 events cannot be in no way dependent.

They are totally independent events failures car A is failing, car B is failing what dependency can be there. So, what is the probability that both the cars will fail if the failures of car A and B are independent of each other. So, just a multiplication of both the probabilities, probability of both the car failures 0.01 and 0.02. If the probability of one event affects the probability of another event, then we call it a dependent event we will see dependent event later.

Now, see till now, how to find the sample space we have taken very small small example tossing a head, tossing a tail. So, we know the size of the sample where, but for bigger problems for real life problems, it is very difficult to find out the total size of the sample space to find the type and size of the sample space basically we have to use some methods so that we can count the sample points.

(Refer Slide Time: 18:27)



So, we will see that so like see this example, if a 22 member club needs to elect a chair and a treasurer how many different ways can these 2 be elected? We will learn, let us not talk about it definitely will be, we are finding this so that we can calculate some sort of probability. Now we are just interested in finding the sample space. Now, if you 22, member clubs needs to elect a chair and a treasurer, how many different ways can these 2 be selected?

Essentially, what are the different possible ways we can select means essentially, this is the sample space. In the sample space, how many elements will be there, so, in how many ways we can do so, there are 22 members, suppose we select a chair, this chair can be selected in 22 ways, from 22 person any person can be chair. So, chair can be selected 22 ways once his chair is selected, the rest remaining is 21.

So, now to select a treasurer we have to select any one from this 21 so basically, since both are again independent events, so it will be 22 x 21. So, event operation can be performed in n 1 ways. And if for each of these ways, a second operation can be performed in n 2 ways. Then these 2 perform operation can be performed together in n 1 x n 2 ways. So, how many if a 22 member clubs need to elect a chair and a treasurer?

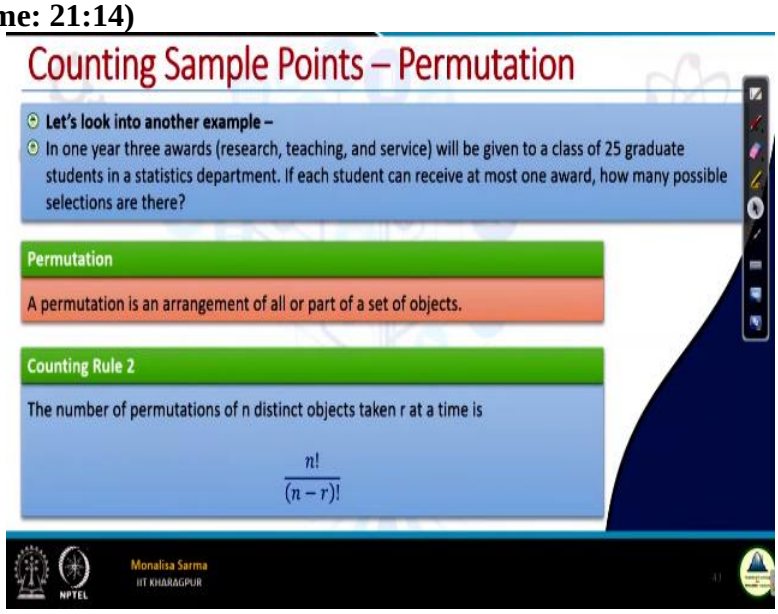
How many different ways we can select these to be selected? That is basically 22 ways. Suppose we select a chair, then we seem to select to treasurer their 21, so 22 x 21. Now, if in general, if you are interested in n distinct object, how does distinct object can be arranged? So, then in general that n distinct object can be arranged in n factorial ways. Let me take a

very small example suppose we have 3 letters a, b, c. So, in how many ways this all these 3 letters are distinct a, b and c, and how many ways we can have how many different elements we can have from this 3 letter?

How many ways we can so all are different? So it is abc, bca you see, you try to find out all the combination, you will totally you will be getting total 6 components and how do we get that. So basically, for the first position, we are considering arrangement of 3 digits for the first position so how many digits are there total 3 digits, for the second position once you have selected the first digit for the second position, how many digits remaining 2 digits remaining.

So, once we have selected the second position, now for the selection the third digit, how many options are remaining? Just 1, so it is 3 cross 2 cross 1. So, 3 cross 2 cross 1 is what basically, it is 3 factorial. So, in general n distinct object can be arranged in n factorial ways.

(Refer Slide Time: 21:14)

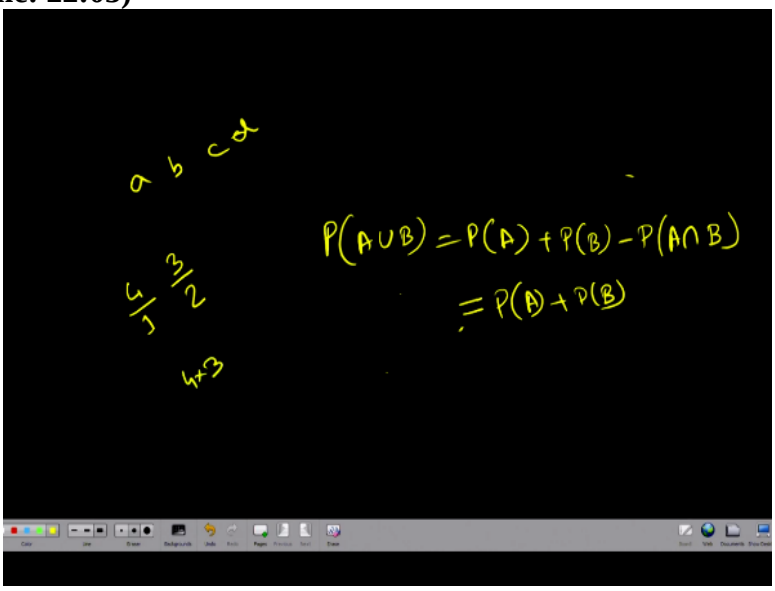


Now, something some different calculation from the one which we have just learned. So, in just see an example, in 1 year 3 awards will be given to a class of 25 graduate students in statistics department. There are 3 graduates, 3 awards will be given to a class of 25 students in a statistic, if each student can receive at most 1 award how many possible selection are there? Like, instead of doing that, we will take a simple example.

Like suppose there are 4 letters from this 4 letters, how many different ways we can have a combination of 2 letters? There are 4 letters a, b, c, d from these 4 letters in how many

different combinations we can have 2 letters. So, basically, what so for 4 letters to have a combination of, we can take the help of the board here.

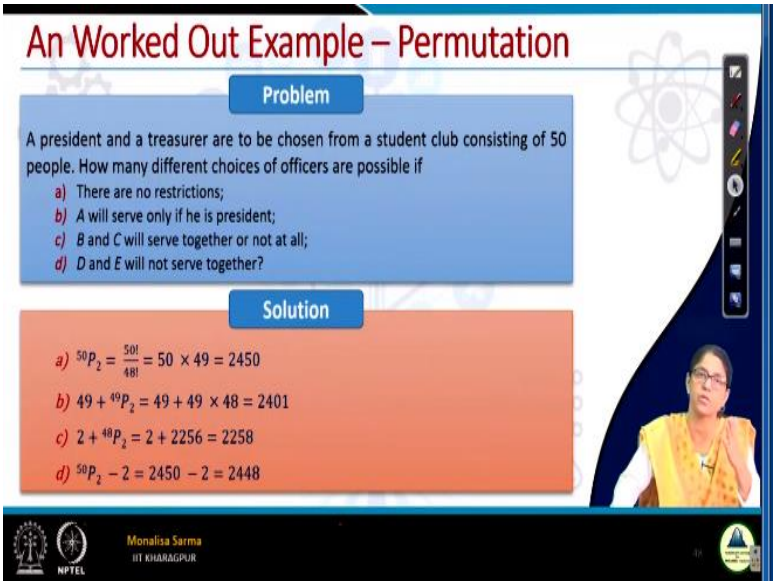
(Refer Slide Time: 22:03)



So first, let me rub this, that is not rub, let us just use it. So, we have 4 letters, a, b, c, d and we need a combination of 2 letters, to 2 letters just to position this is the first position, this is the second position. First position, we can fill it in how many ways we can fill it in 4 ways. Second position, we can fill it in how many ways second position we can fill it in 3 ways their total remaining 3 ways. So, this is 4 x 3. So, what is this 4 x 3?

So, this is 4 x 3 is nothing but n ! / n - r !, you see if you interested in finding out that is 4 x 3. So, from n if you are interested in selecting r objects, so, if you will do basically n into n - 1 into n - 2 this will go to out what level n - r + 1. So, n x n - 1 x n - 2 it will go to till n - r + 1. So, what is that, that is nothing but simple n ! / n - r !. So, now, this question in 1 year 3 awards will be given a class of 25 so, 3 of us will be given a class of 25 basically, so, this is nothing but 25 b 3.

(Refer Slide Time: 23:24)



So, now this problem how will you solve it? A president and treasurer to be chosen from a student club consisting of 50 people. How many different choices of officers are possible if there are no restriction? So, we have to select from 50 people a president and a treasurer. How many different sizes of offices if there is first of all, there is no restriction we can any people can any person can be a president, any person can be a treasurer in how many ways we can do it using the counting rule.

So, first if we select the president, president can be selected in 50 ways. And then if we select the treasurer, treasurer can be selected in 49 ways, because 1 already has become president. So, basically, what is that 50 P 2, is not it? So, it is 50 P 2, that second A will serve only if he is president, for then there are 2 possible one, we have taken one such choice where A is selected as a president or A is not selected because A will not take any other job when A is selected as a president then.

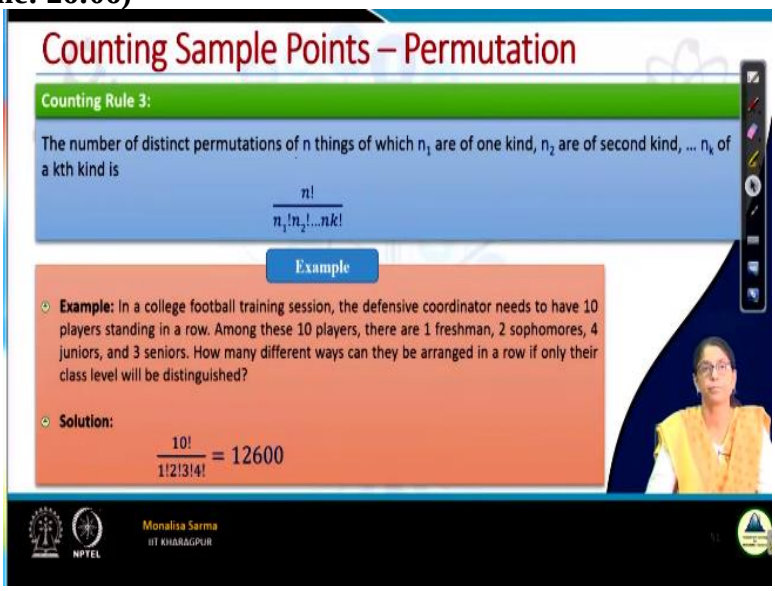
And next task we have to select a treasurer, treasurer will be selected can be selected how many ways treasurer can be selected out of the remaining 49 that is 49 way. So, 1 x 49 that is one option. Another option is A is not selected, A is not selected, then president and treasurer can be selected from any of the remaining 49 and that is from 49 P 2. So, it will be 49 + 49 P 2, 49 x 1 basically 49 into 1 + 49 P 2.

Now, next is B and C will serve together are not at all what does that mean B and C will serve together means B and C will be either B will be president C will be treasurer or B will be treasurer C will be president B and C will serve together means how many ways we can do

that, we can do that in 2 ways or not at all then if they are not serving depends remaining is 48 so, that means 2 + 48 P 2. Cool right? So, next is D and E will not serve together, D and E will not serve to get this here we will be using the concept of complement.

First we will find out when how many ways D and E will serve together, D and E will serve together and how many ways if we can find out then 1 - of that will give us D and E will not serve together and D and E will serve together? How did you find out that we have just seen in C, B and C will serve together? So, in what way D and E will serve together there is only 2 ways and D and E will not serve together the total sample space is 50 P 2 - 2.

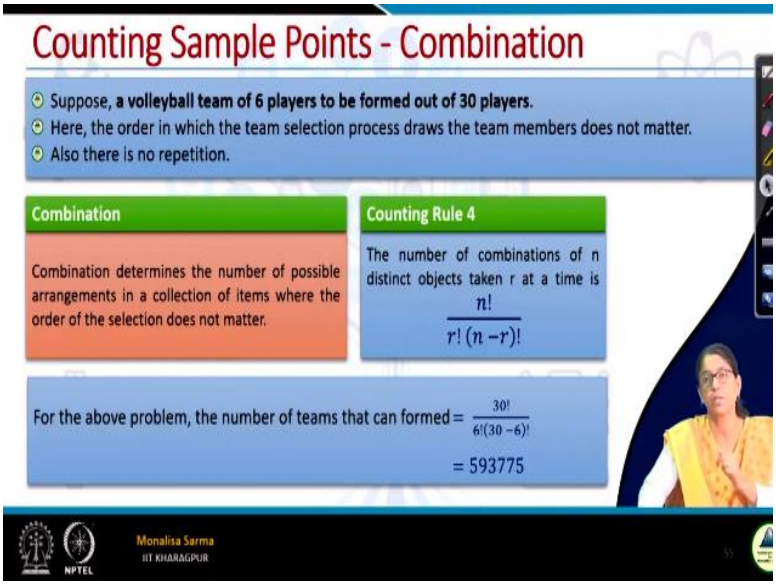
(Refer Slide Time: 26:06)



Now, the number of distinct permutations of n things of which n 1 are of one kind n 2 are of second kind and say n k of the kth kind, then what is the formula for that n ! / n 1 ! x n 2 ! … n k ! we will understand with an example. Suppose, in a college football training session, the defensive coordinator needs to have 10 players standing in a row they need 10 players.

Among these 10 players, there are 1 freshman, 1 first year student, 2 sophomores second year students, 4 juniors for third year students 3 senior students. So, these are not distinguishable that 1 freshman only 1, 2 second year students this 2 second year students are not distributable there we will consider both second year students only. Similarly 4 juniors are not distinguishable. So, basically there are 1 kind, 2 or 1 kind then 4 juniors are of 1 kind 3 seniors are 1 kind and among these I have to select total 10. How will I select, so, it is 10 ! divided by this 2 ! x 4 ! x 3 !.

(Refer Slide Time: 27:20)



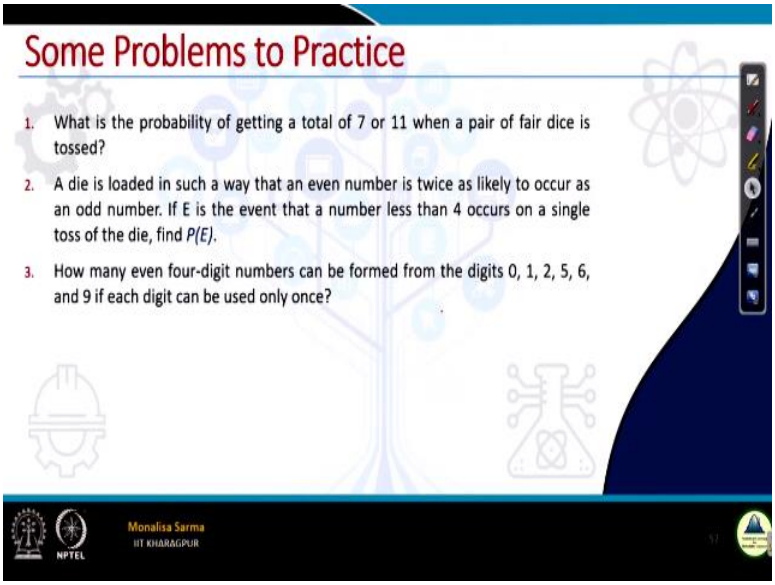
So now, below what we have discussed is permutation where the ordering methods abc and cba are 2 different things where we are distinguishing among the digits. So, now, if suppose the volleyball team of 6 players to be formed out of 30 players and the order will be the team selection process draws the team member does not matter, there are 30 players, we are not distinguishing the players from the 30 we need to the select 6, we are not distinguishing the players somehow.

Like there are 3 digits we are not distinguish that this is a, this is b, this is c we are not distinguish. So, there are 30 players from this 30 player we have just 2, we just need to say select c, then what we will be used combination and also there is no repetition. So, this is a concept of combination, combination determines the number of possible arrangement in a collection of items where the order of selection does not matter.

So, number of combination of n distinct objects taken r at a time. So, it is n distinct and the formula for n distinct n ! / r ! n - r ! all this actually it is like a quick recap for all of you because I am sure all of you can have done this in your school level. And if any of you have forgotten then we can consider this is a brush up basically. So, actual lecture will be start once we complete a probability theory basically.

From for the above question like when from the 30 players, we need to select 60 players. So, it is 30 c 3. So, how do we write the 30 c 3, 30 ! / 6 ! into 30 - 6 !.

(Refer Slide Time: 29:07)



So now, there are a few problems which I have kept here of course I will be doing some tutorial classes also. But these 3 problems I have kept here I want you to solve these problems. So, if you have any problems, then we will be having doubt learning session then of course, we can clear this problem. But still, I want you to practice this problem. So, I will give you a hint on how to solve this problem?

So, first is what is the probability of getting a 7 or 11 when a pair of fair dice is tossed? I am tossing a pair of fair dice. What is the probability of getting a 7 or 11? See here that means 2 events. One is probability of getting a 7 if I consider it is an event A and probability of getting 11 it is considered is an event B basically that means I am interested in probability A union B A or B, but these 2 events are mutually exclusive both 7 and 11 cannot happen. So, it is a probability of A union B is just probability of A + probability of B.

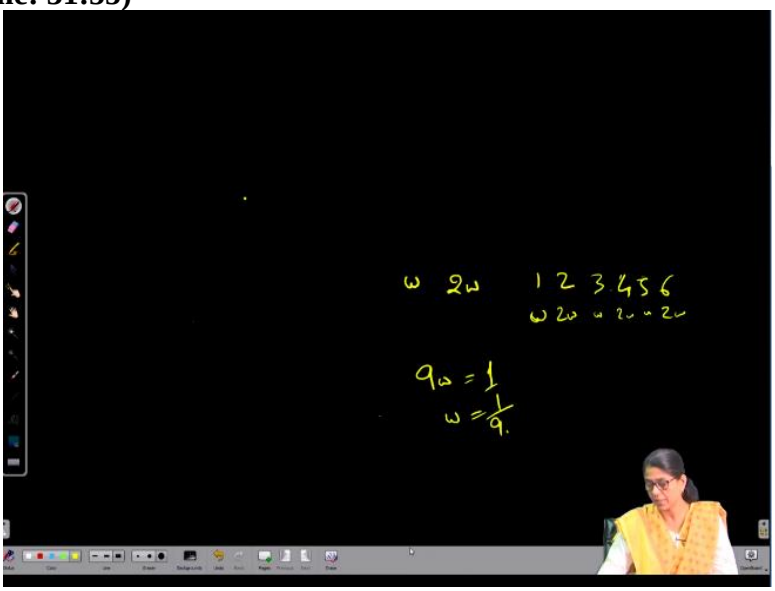
So, now what is the probability of A from the sample space when we are tossing 2 dice? What will be the sample space size of the sample space will be 36 is not it 6 cross 6. Now, out of these 36 how many corresponds to 7 how many of this corresponds to a total of 7 like if I get 3 and 4, 4 and 3, 5 and 2, 2 and 5, 6 and 1, 1 and 6. So, these are the thing which corresponds to 7. So, I think around 6 of them will corresponds to a total of 7. So, what will be the probability of getting event A that is a total of 7 is 6 / 36.

Now, probability of getting 11 when will I get 11 when I get 5, 6 and 6, 5, so that is 2 that is 2 by 11. So, 6 / 36 + 2 by 36 is the answer for the first question. Second a die is loaded in such a way that an even number is twice as likely to occur as an odd number. So, we have seen

when we toss a die we have considered as it is equally likely, equally likely means so, there are total 6 phases. So, what is the probability of occurrence of is as 1 by 6 since this sample size is 6 and it is equally likely so, it is occurrence of any one is 1 by 6.

Now, here it is different here it is not in fair die, if it is a fair die then it is equally likely that it is not in fair die, but it is given is loaded in such a way that an even number is twice as likely to occur as an odd number. So, what I will do? I will assume the probability of occurrence of odd number, let us take the probability of occurrence of odd number is w. So, now, what was the question? The question was probability of even number is a double the probability of getting an odd number.

(Refer Slide Time: 31:55)



So, if the probability of getting an odd number is w, then I can take probability of getting an even number is 2w. So, I will take 1 as w and another as 2w. So, total there how many numbers in a die, so 1, 2, 3, 4, 5, 6 so, if 1 is w, this is 2w, so w 2w, w 2w so how much I got? I got 9w and this probability will be equals to 1 probability of getting this number total is 1. So, what is w? w will be 1 / 9. So, I got w 1 / 9.

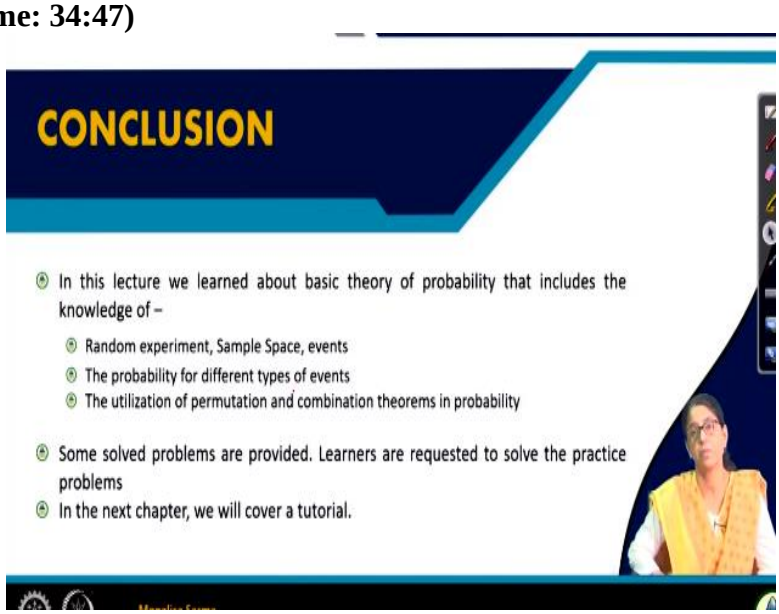
Now let us come to the equation if E is an event number that a number less than 4 occurs on a single toss of a die less than 4 what is less than 4 that means 1, 2 and 3. So, what is the probability of occurrence of 1? Probability of occurrence of 1 is 1 / 9. Probability of occurrence of 2 is 2 / 9. Probability of occurrence of 3 is 1 / 9 so the probability is the sum of all this 3 I am not doing it just but I to explain you in fact you will be able to do it.

Now the next question how many even 4 digit number can be formed from the digit 0, 1, 2, 5, 6 and 9 if each digit is used only once? This will not be doing it I will just give you a hint up solve it. So, here we have to find out total even 4 digit number so the digit are 0, 1, 2, 5, 6 and 9 when we can see that number will be even if in the last digit we get either 0, 2 or 6. So, we will take 2 combinations one is when the last digit we got 0 another is when the last digit we did not get 0.

When the; last digit is 0 why we have separated because the first digit cannot be 0 in any of the cases. We are considering 4 digit numbers. So, first digit cannot be 0 in any of the cases. So, when the last digit is 0, then we will get a different way of finding out the probability because then the number of digits remaining will be total how many digits are here 1, 2, 3, 4, 5, 6 to let us 6 digit them for to filling out this other digit this there will be 6 digit.

But in the case of if the last digit is not 0 that means if the last digit is 2 or 6, then for filling the 3 position they will be different for filling the first position we cannot take 0. So, there will be 1 less number of points. So, now I will not explain further in this I want you to try and do this problem. Of course if you cannot do that in doubt clearing session we will be taking up and but only do not restrict to this problem, practice as many problem as you can.

(Refer Slide Time: 34:47)



With this I am concluding this lecture. So, in this lecture what we learn? We learn what is an random experiment, what is a sample space, what are events then probability of different types of events also we have seen, then we have also seen the utilization of permutation combination theorems in probability. Then we have also solved some problems. And as I told

you I request to solve as many problems as possible. In the next lecture, I will be covering the tutorial on the first 3 lectures.

(Refer Slide Time: 35:15)

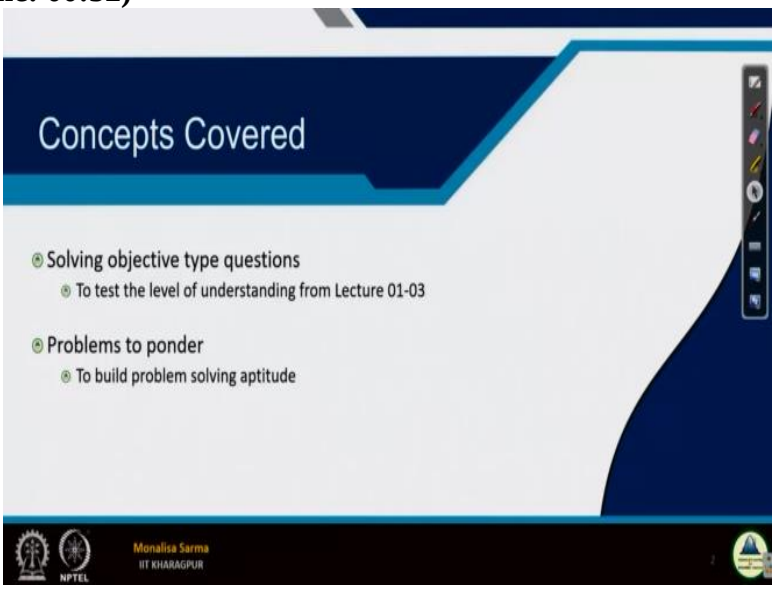
So, these are the basically 3 books which I am referring there do not restrict to this books they can walk from any books on probability and thank you.

Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology - Kharagpur

Lecture - 04
Tutorial on Introduction to RE, SL and Probability Theory (Part - I)

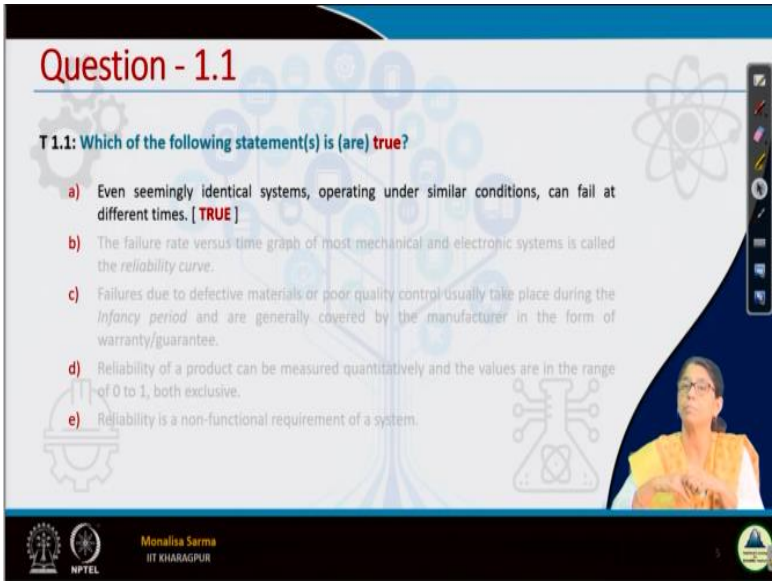
Hello everyone so today is the 4th lecture today basically; we will be doing some tutorial. Till now, we have covered introduction to reliability engineering, we have covered introduction to statistical matter, and we have also discussed, we have also taken a class on probability theory. However there I will take 1 more lecture on probability theory, but before that, let us say we have a tutorial session on the last 3 classes what we have covered.

(Refer Slide Time: 00:52)



So, in this tutorial class I will basically be solving few objective type questions and after the objective type questions, then we will go to some problems which will help your problem solving aptitude.

(Refer Slide Time: 01:12)



So, now, to start with the objective type questions. So, the first question is here I have given few statements, so, you have to identify whether the following statements are true or false. Now, I suggest you please do not go to the answers immediately of course I will give you the answer also here. So, but I says to please do not go to the answers immediately first you try to find out whether you yourself can answer the questions or not.

If you have understood the last 3 lectures, you will definitely be able to answer these questions. So, now, the first statement; what is the first statement? Even seemingly identical systems operating under a similar condition can fail at different times like let me take you an example, let me take an example of a washing machine again suppose 2 friends x and y, x and y both bought the same type of washing machine say Samsung washing machine x and y both bought the same washing machine.

So, same type of washing machine means from the same brand same technology. So, the reliability has to be same, but is it possible that the machine which x is bought it has failed in the first month itself and the machine which y has bought it is running for 5 years and still there is no failure is it possible? So, that is what I want to know. Yes, the answer is yes, it is possible, why it is possible?

Both the components reliability is the same, both are from the same made, so, why it is possible? First to answer this, let me take an analogy from a real life situation. Suppose there are 2 friends both are imperfect human beings, imperfect human means both tell lies, the thing is that suppose person A and person B, person A tells 90% of the time he does not tell

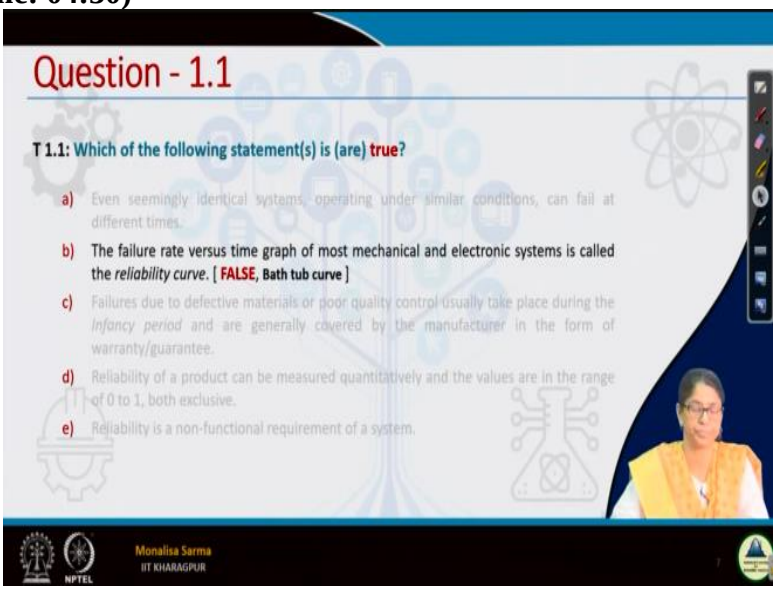
lie and a person B 50% of the time he tells lie that means person A 10% of the times he tells lie.

And person B 50% of the time he tells lie and which person is more dependable, definitely person A is more dependable. Now, I just want to analyse these 2 persons suppose these 2 persons are sitting, I am just observing these 2 persons, these 2 persons are sitting in a restaurant and they are gossiping and I have noticed this. So, it so happened that person A who just tells lie 10% of the time he happens to speak the first lie that day itself.

And a person B who happens to speak 50% of lies he did not he told all the truth. So, does that mean the person B is more reliable more dependable? No, the answer is not it is actually when we talk of reliability, it is basically what we do is that we take on an average assessment based on an average assessment person A is more dependable, but it so happened that he spoke is first truth on that day.

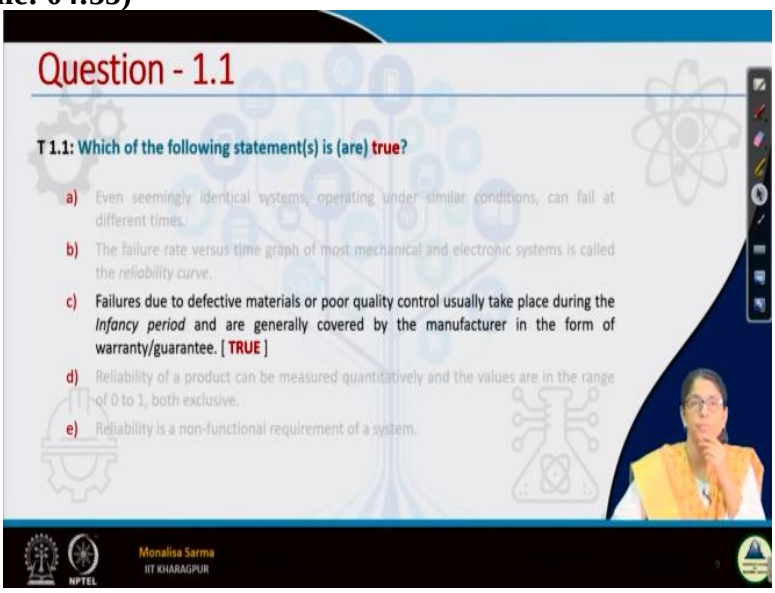
Similarly, so when we are talking of 2 components with the same reliability it may fail at different times the component maybe the component with a lower reliability has not failed for quite a long time, but component with a high reliability he has may have failed in the first stage. So it may happen, because reliability as we have seen it is an engineering uncertainty it is the concept of probabilities. So what is probability? Probability is the science of uncertainty. So, the first question answer is true.

(Refer Slide Time: 04:30)



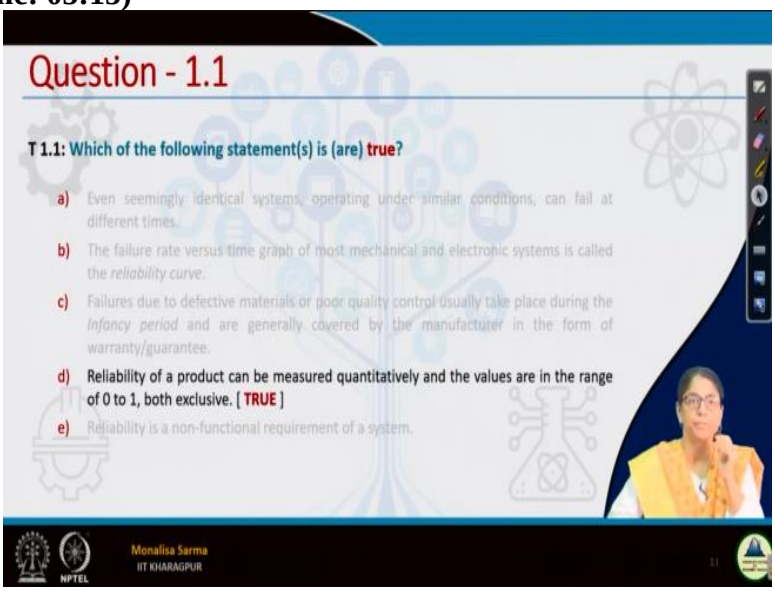
Coming to the second question, the failure rate versus time graph or most mechanical and electronic system is called reliability graph. The failure rate graph what we have seen that is the different types of failure the early failure than the normal life failure wearout failure that is the failure rate curve, right? So failure rate curve, is it called the reliability curve? No, it is not called the reliability curve it is called the bathtub curve.

(Refer Slide Time: 04:55)



And the third question; failures due to defective materials or poor quality control usually takes place during the infancy period and are generally covered by the manufacturer in the form of warranty or guarantee. So, what is the answer? Yes, that is true.

(Refer Slide Time: 05:15)

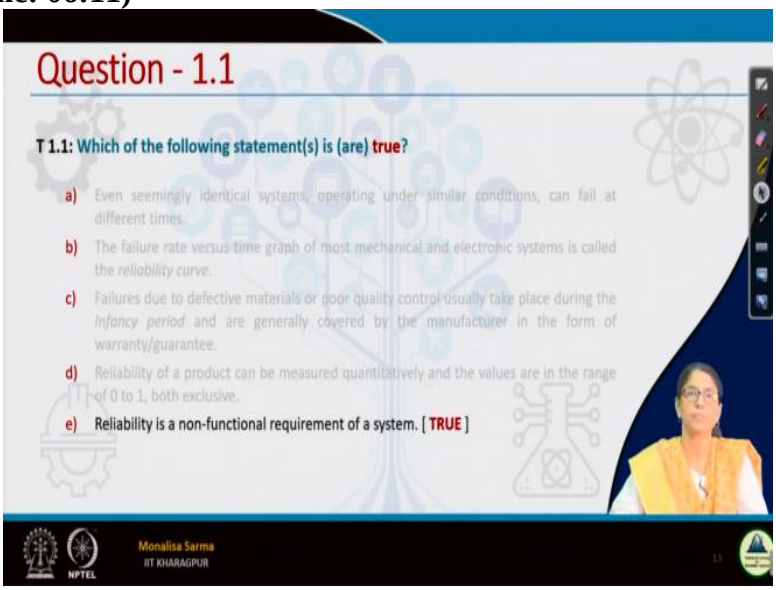


Then the next question next statement reliability of a product can be measured quantitatively and the values are in the range of 0 to 1 both exclusive, this is an interesting question interesting statement, reliability can be measured quantitatively of course, that is true and the

values are in the range of 0 to 1 both exclusive, 0 and 1 are exclusive, but we know probability can take what value? Probability can take value any value from 0 and 1, 0 and 1 are inclusive.

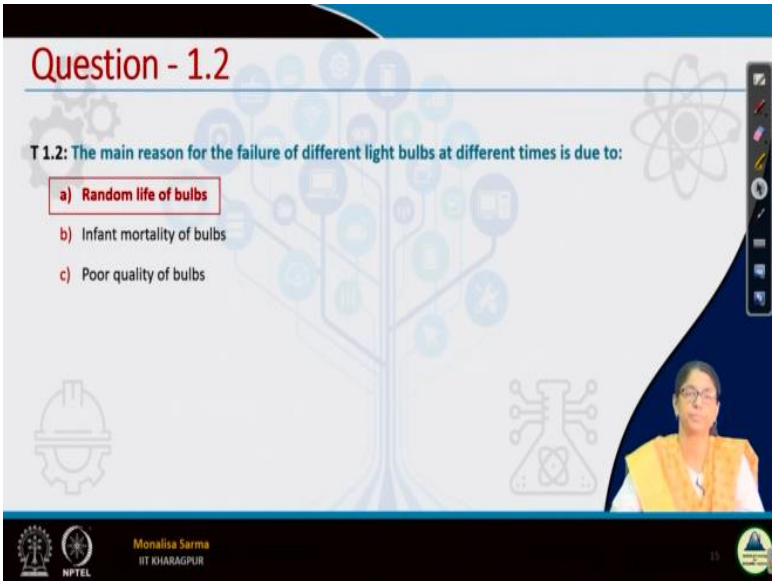
So, but in case of the reliability, we are trying to find out the reliability of a product, definitely we will not try to find out the reliability of the product which is not working which is a failed product that means its reliability is 0, it will not work at all. So, will not try to find out the reliability of the product. Similarly, and nothing is permanent in this universe, stars also it dies and nothing is permanent in this universe. So, when star can die, what about a system a system will definitely have its own death so reliability cannot be 1. So, the statement is true.

(Refer Slide Time: 06:11)



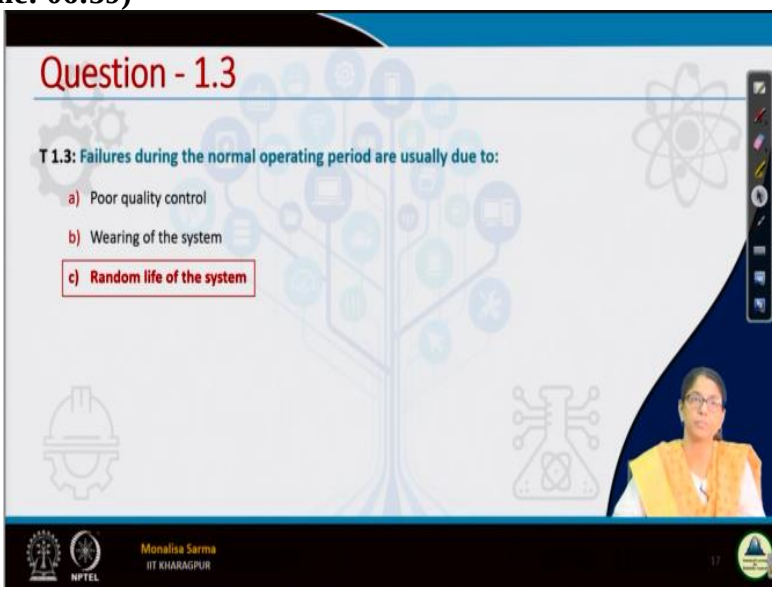
Then reliability is a non-functional requirement of a system. Yes, it is a non-functional requirement of the system, it is not a functional requirement of the system. Execute is a non-functional requirement of the system. User interface, usability. Usability is a non-functional requirement of the system so, true.

(Refer Slide Time: 06:29)



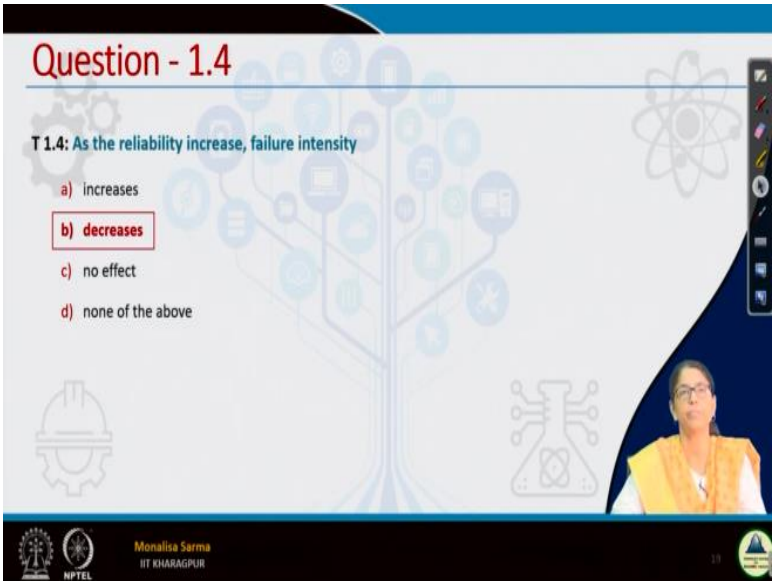
So, the second question the main reason for the failure of different light bulbs at different times what maybe the main reason? Is this random life or the infant mortality or poor quality of bulbs? Definitely it is a random life of failures a different light in the normal lifetime and a normal lifetime why it fails? It may be due to the random reasons. So, it is a random life of the bulbs is the answer.

(Refer Slide Time: 06:59)



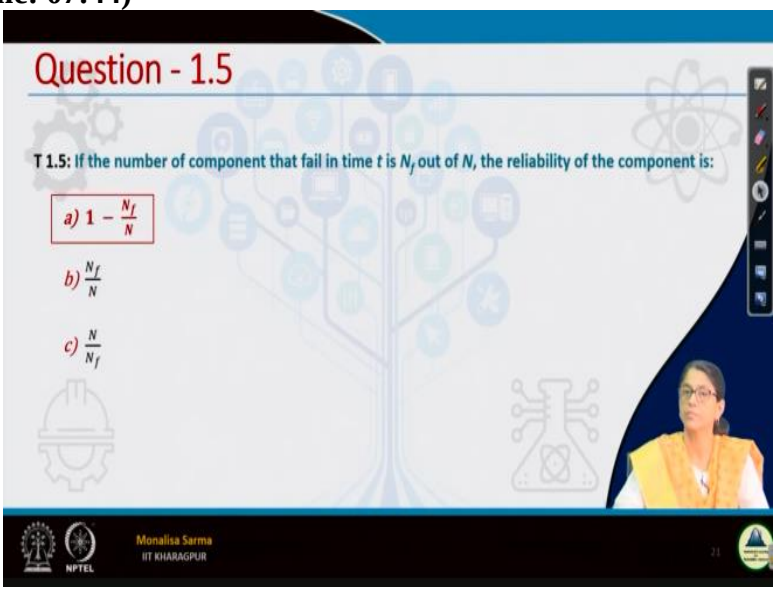
Failures during the normal operating period are usually due to what? Normal operating create is usually due to poor quality control wearing of the system random life of the system, which one is correct? Normal operating period why it fails random life of the system it cannot be poor quality control, poor quality control means it has to be in the infant period, wearing of the system since it is the end part of the light it is the wear out light. So, in a normal operating period, it is due to the normal life of the system.

(Refer Slide Time: 07:32)



As the reliability increases failure intensity, what happens? As the reliability increases failure intensity definitely it is very obvious failure intensity decreases.

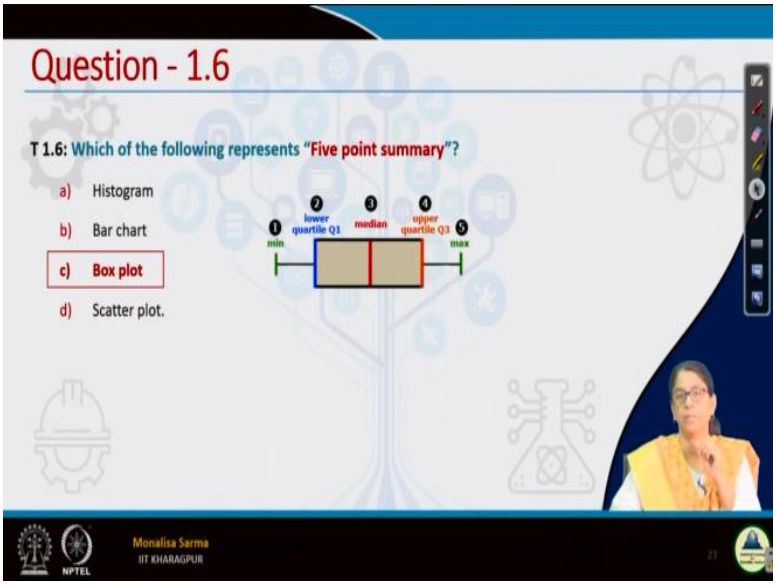
(Refer Slide Time: 07:44)



Now, here, if the number of component that fails in time t is N f out of N then what is the reliability of the component? There are 10 components out of 10 components 6 component had failed at time t. So, what is the reliability always I forgot to mention maybe reliability is always with the perspective of the time when you talk about reliability it is with the perspective of a time. So, what is the reliability of time t?

So, at time t out of 10 components 4 component has failed. So, what is the remaining component? Remaining component is 6 so, what is the reliability at time t? Reliability at time t is 6 / 10. So, here, so, what will be the reliability 1 - N f / N.

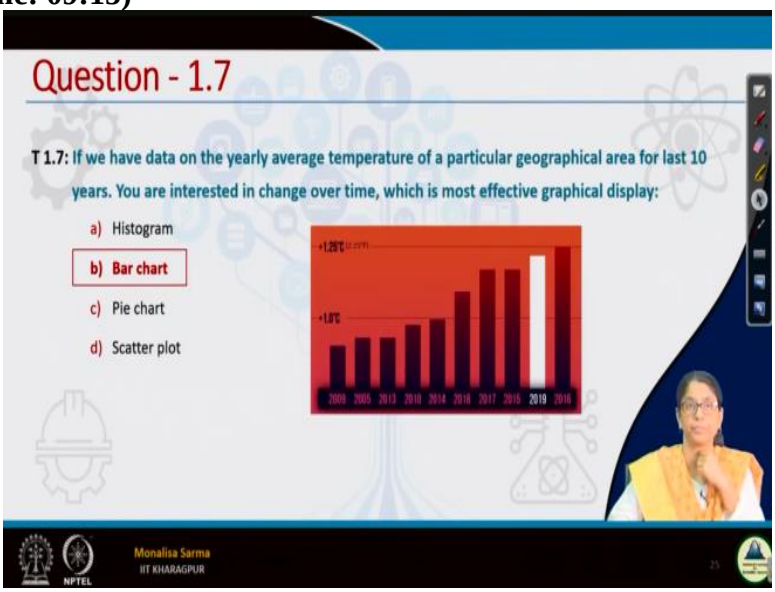
(Refer Slide Time: 08:28)



Which of the following represents 5 points summary? I already told this mentioned this when I was discussing statistical matter, introductory class on statistical matter, that is boxplot. There is 5 points summary why 5 point summary? The first vertical line, that what it gives is the minimum value and the top vertical line it gives the maximum value of a data set, then the blue line that gives the 25th percentile.

That is the lower quartile, than this upper blue upper red line, that is the upper quartile that is the 75th percentile. I am talking percentile not percentage mind it, then the middle red line is the median so, it is a 5 point summary.

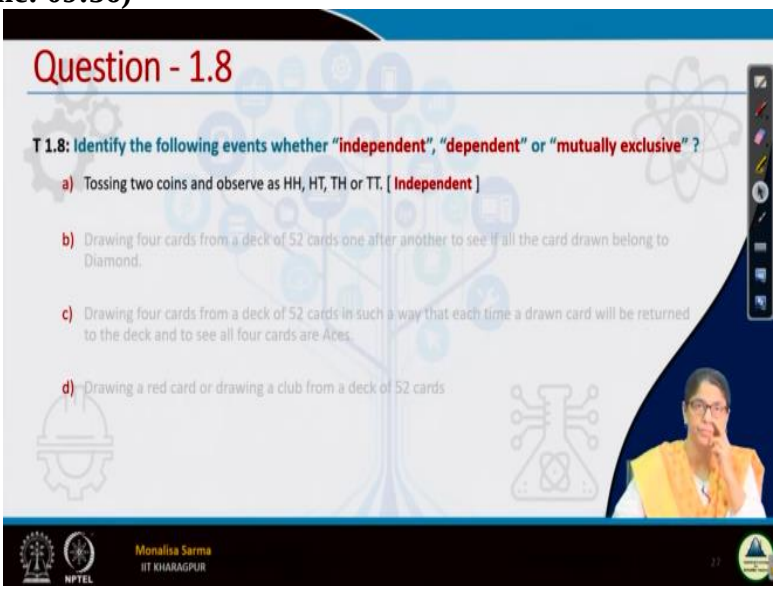
(Refer Slide Time: 09:15)



So if you have data on a yearly average temperature a particular geographical area for last 10 years you are interested in change over time, which is the most effective graphical display? If you are interested in change over time for a year definitely when we talk of year definitely it

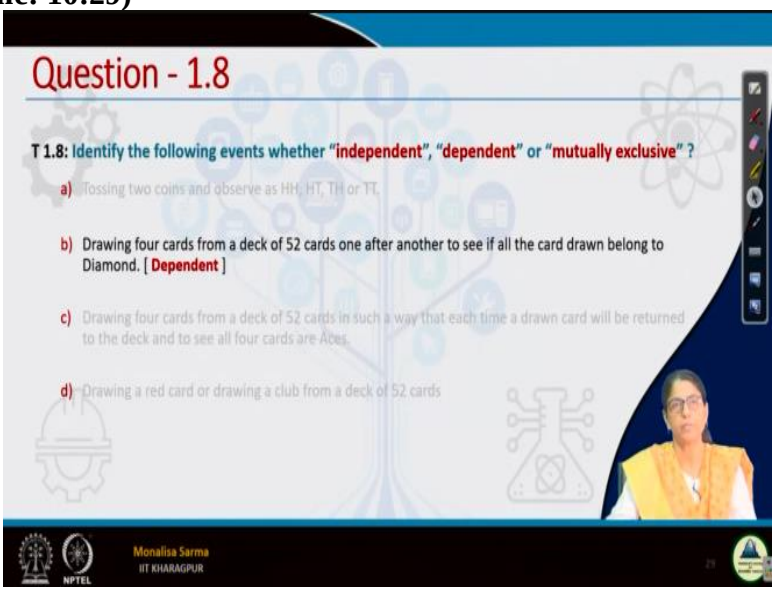
is not an quantitative data that is a categorical data. So per year, that means it is not what type of data? It is a nominal data. So we will be using bar chart here.

(Refer Slide Time: 09:56)



So now this is the question on probability identify the following events whether it is an independent event, dependent event or mutual exclusive, I think you remember what is an independent event? What is a dependent event? What is a mutually exclusive event? Tossing 2 coins we are tossing 2 coins and we have observed either as head head, head, tail, tail, tail or tail, head. So what is it? What is an independent event, dependent or mutually exclusive? You are tossing 2 coins so definitely, this is an very much an independent event.

(Refer Slide Time: 10:29)

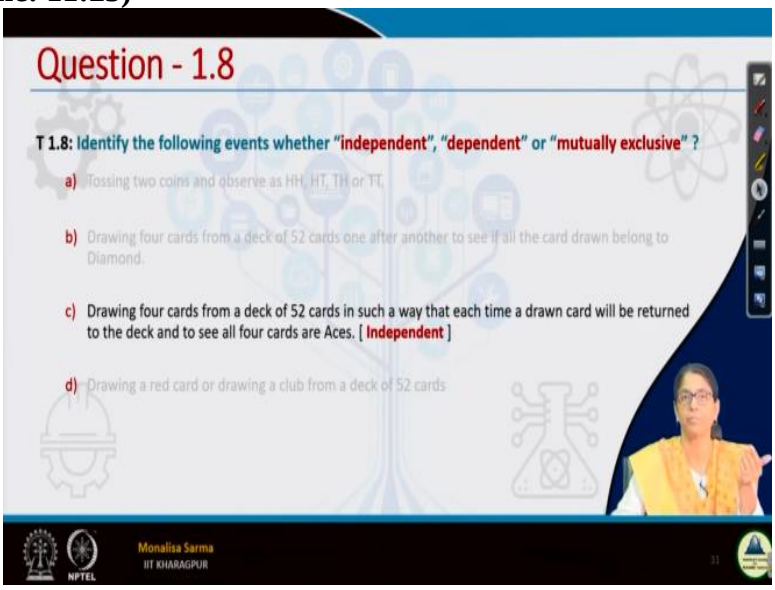


Second, drawing 4 cards from a deck of 52 cards one after another to see if all the cards drawn belong to diamond, what I have done? Deck of 52 cards, I have drawn the first card. I have seen whether this is diamond or not, I have kept the card aside, then I can pick the

second card, I have seen whatever it is diamond or not, I kept it aside, I did not mix it with this pack of cards. Similarly, I have drawn total 4 cards.

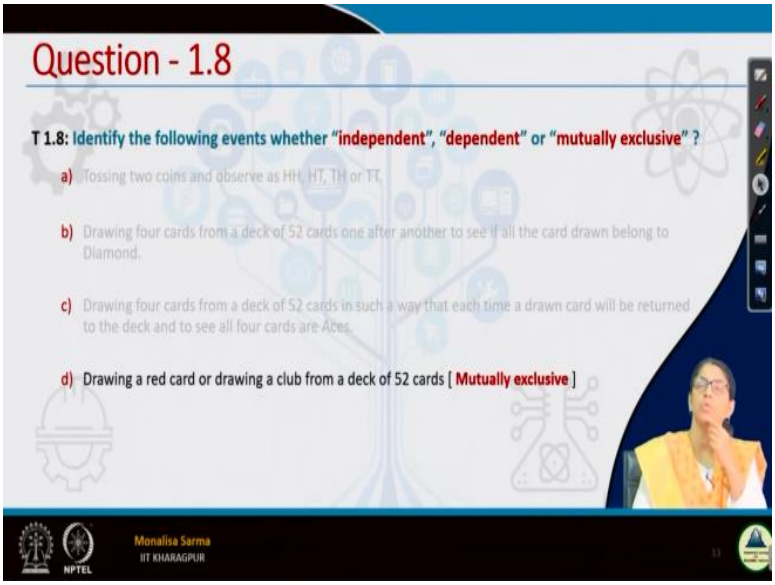
And if all the cards do not belongs to diamond or not, this is an independent event or dependent event? Definitely, it is not a mutually exclusive event. This is very much a dependent event. Because when we are keeping it aside, we are not mixing our number of diamonds changes number of cards in total number of cards changes if and moreover, if you have already picked a diamond and then the number of diamond as well changes in dependent population, is not it? So it is very much dependent event.

(Refer Slide Time: 11:25)



Third, drawing 4 cards from a deck of 52 cards in such a way that each time a drawn card will be returned to the deck and see all the 4 cards are aces each time I am returning it. So this is an independent event.

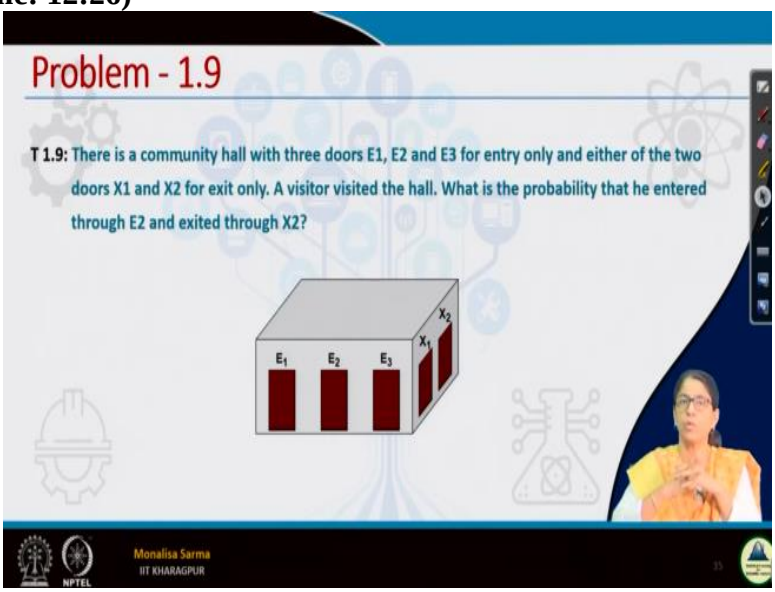
(Refer Slide Time: 11:42)



4th, drawing a red card or drawing a club from a deck of 52 cards club, can club be red it has to be black, right? So red card and a club that means a red card and black card where if I am drawing 1 red card, I cannot draw a black card at the same time. I am drawing 1 card is not it? So these 2 are mutually exclusive event like drawing getting a head or tail that is a mutually exclusive and so that is why our 4th is the mutually exclusive event.

Now, we will be doing some solving some problems. I suggest again immediately do not see the solution. First, see the equations, try to answer it yourself then go to the solution if you do not if you cannot answer it.

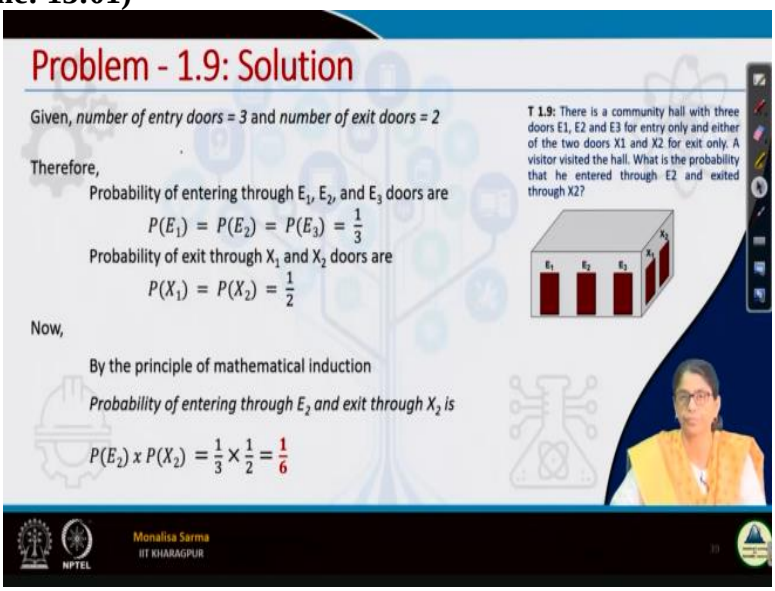
(Refer Slide Time: 12:26)



The first question so there is a community hall with 3 doors. The 3 doors are E1, E2, E3, for entry these 3 doors are for entry. And for exit we have 2 doors X1 and X2 people can enter from any of the door and people can exit from any of the door X1 and X2. A visitor visited

the hall what is the probability that he entered through E2 and exited through X2, so now he can enter through any of the doors E1, E2, E3. And he can exit through any of the doors X1 and X2. What is the probability? Probability of entering to any one of the doors?

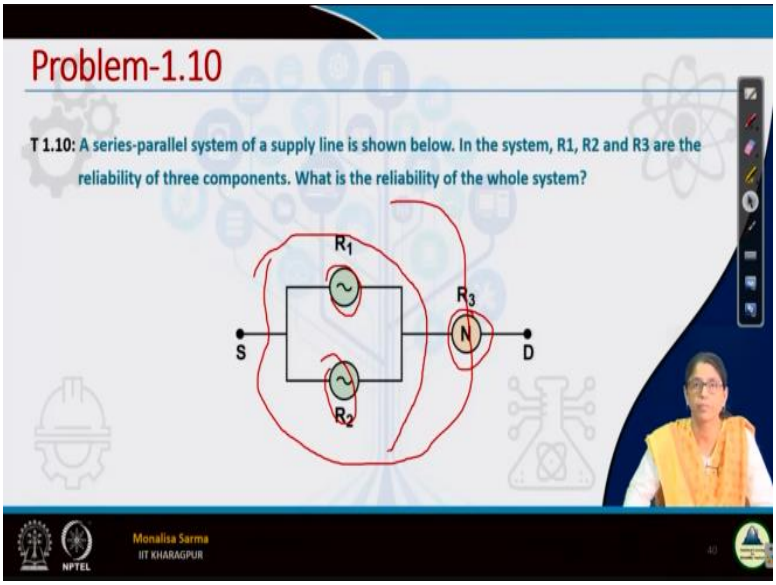
(Refer Slide Time: 13:01)



First answer that was the probability of entering to any one of the doors the equally probable so probably of entering any one E 1, E 2, E 3 is 1 / 3. Similarly, the probability of exiting any of the doors that is also equally probable. So it is 1 / 2 there are 2 doors for exiting and both the events are independent events. It is not if I am entering through 1 I cannot I have to enter exit rather there is no sort of dependency is there, it is totally independent events.

So since this is independent event. If 2 events are independent X and Y are 2 independent events. What is the probability of occurring of X and Y? Probability of occurring of X and Y is probability of X x Y. So, by the principle of mathematical induction I can write that probability of entering through E 2 and exit through X 2 is 1 / 3 x 1 / 2.

(Refer Slide Time: 13:52)



So, now coming to the next question, the next question is also a question on probability, but here we will be seeing the application of reliability as well. So, like this question, what we have seen is there is given a series parallel system of a supply line. So, in the system R 1, R 2, R 3 are reliable, they are the 3 components we can see there are 3 components this is 1 component and this is the second component.

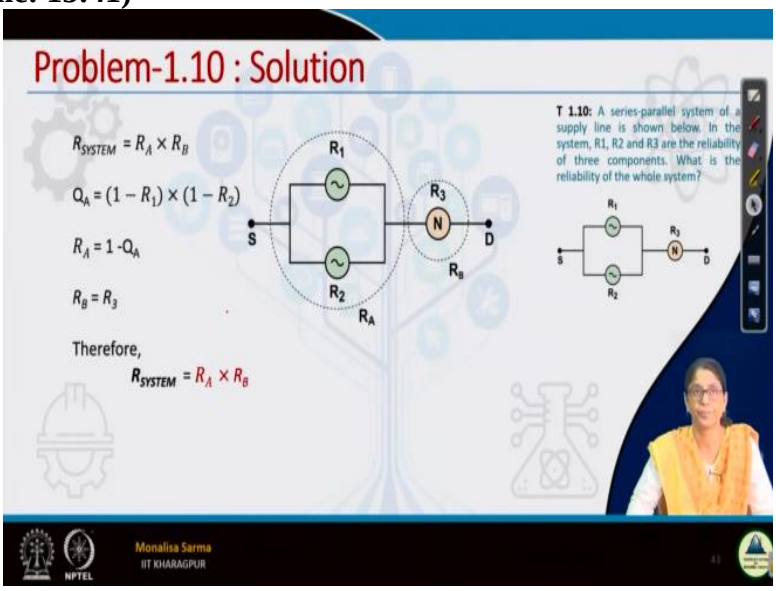
So, then this is the first component and then this is the second component and this is the third component. So, reliability of the first component is R 1, reliability of the second component is R 2, reliability of the third component R 3. Now, what is this reliability? Basically, reliability is the probability of failure for your operation. So, that means, this is the probability that this component will not fail that means, this is the probability of this non failure of the component.

So, essentially R 1 is the probability that the component one will not fail. Similarly, R 2 and R 3. So, now, this just comes to a simple probability problem. So, now what is the here just we need to use the concept of series parallel structure actually, we have learned in physics in class 8, 9 I suppose and even first year engineering also they learn it. So, I will not go to those details, but I am sure all of you know when the 2 components are in parallel that means, for the system to work if the 2 components are parallel.

If one of the components works the system will work however, if both the components failed and the system will not work. So, and for a series if 2 components are in series, both the component has to work then the system will work. So, like for here, let me consider this as

this whole thing, this whole thing as one or sorry, this one, this is one component and the other one is another component.

(Refer Slide Time: 15:41)



So, I have a figure here so, here this is the 1 component, I am just considering this as 1 component, this sub component basically I can say, this component is a combination of 2 parallel components. And so, again, this is 1 component of course, now, if I am interested in finding the reliability of this system reliability means that the system will not fail. So, I will have to find out the reliability of each sub component.

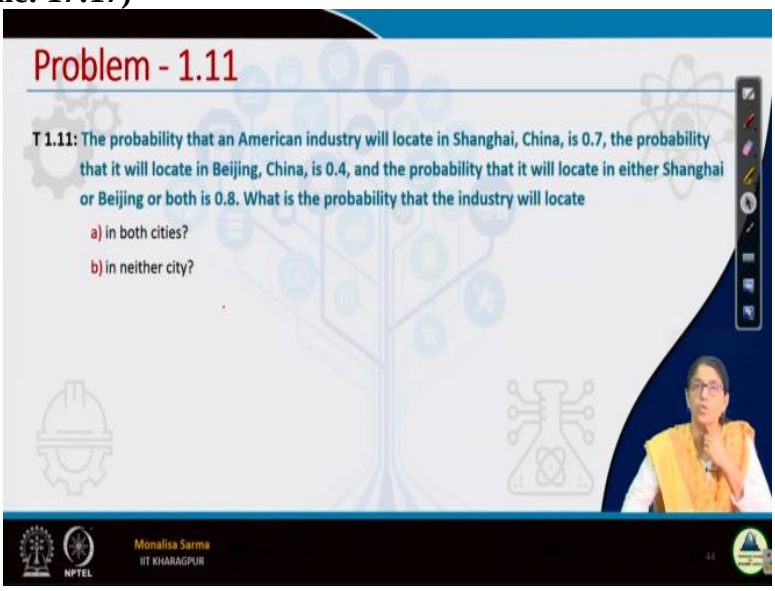
So, first let me find out the reliability of first component that is R A I am naming it as R A. So, what is R A? For finding an R A what we will do is first we will see when the system will not work that is let us, calculate the unreliable the probability of failure. So, when the system will not work, since 2 components are in parallel, so, this system will not work and both the components will not work.

So, if the probability of working is R 1 the probability of not working is 1 - R 1. So, probability reliability of the system is definitely R A cross R B is not it? Because it is a series system this is series and this first component is in series with the second component. So, that is R A cross R B then we are interested in finding out the reliability of the R A to find out R A first time finding out Q A because that will be easier.

So, what is Q A? Q A is that both the components are not working but the components are not working means 1 - R 1 into 1 - R 2. So, what will be my R A? R A is nothing but 1 - Q of

A reliability is one minus unreliability and R B we know it is relevant is given R 3. So, what is my system reliability is R A x R B and so, simple.

(Refer Slide Time: 17:17)



Now, coming to the third problem, what is given first let us go through the problem, the probability that an American industry will locate in Shanghai as 0.7. The probability that it will look at in Beijing is 0.4. And the probability that it will locate in either Shanghai or Beijing or both means it may locate in either Sangha or in Beijing or in both the cities that it has given 0.8. So, what is the probability that the industry will locate in both cities and industry will locate in neither cities?

(Refer Slide Time: 17:51)

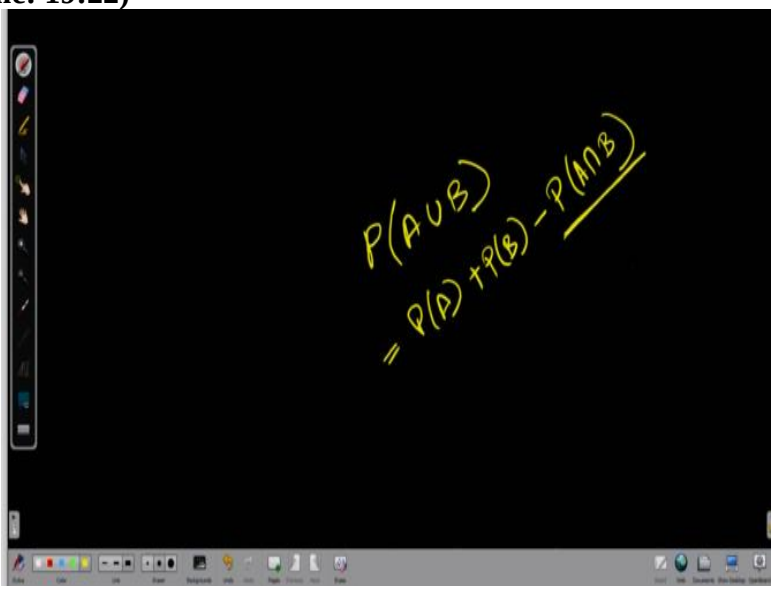
So, first what we will do is that we will consider the events the concept of events we have discussed in our last lecture. So, first we will consider the events what are the 2 events? One event I am calling it as an S, that is the industry we will locate in Shanghai the second event,

let me call it as B that the industry will locate in Beijing. So now since I am interested in finding out the probability that the industry will locate in both the cities.

So, industry will locate in border citizens it has to be located in Shanghai as well as it has to locate in Beijing. So, basically, what I need to find out? I need to find a probability of S intersection B means it needs to basically probability of S into probability of B probability S into probability of B is as told earlier S intersection B. So, I need to find out probability of S intersection B.

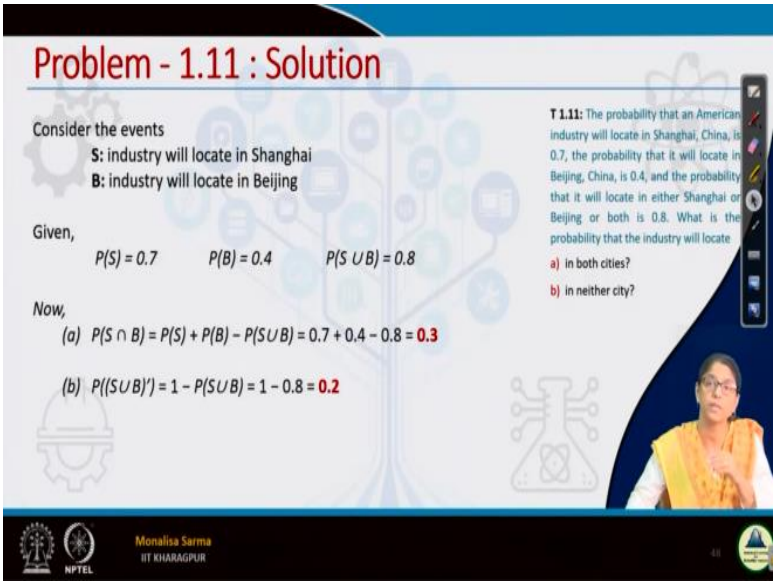
So how do I find probability of S intersection B I have this information PS is 0.7, PB is 0.4 and P S or B it is given that it is either located in Shanghai or Beijing or both. So it is S union D so, this is we have 0.8. Now, we know the formula what is the formula for probability of S union B, so we can go to the let me take a piece of paper and then explain it to you.

(Refer Slide Time: 19:22)



So, like what is probability of A union B? It is probability of A + probability of B - probability of A intersection B is it not? We have learned in the last class. So, now, what we are interested we need to find out this probability of A intersection B. So, what is the property of A intersection B? Just simple this probability of A intersection B will be probability of A + probability of B - probability of A union B.

(Refer Slide Time: 19:54)

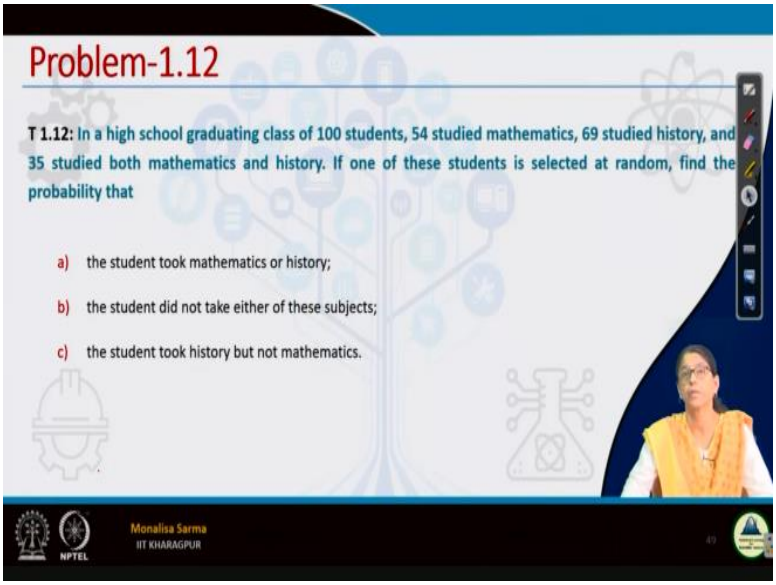


So, this is what all the probability of S intersection B; as this put it in a formula and we just get it. Now, the second is what is the probability that industry will locate in neither city means it will neither to locate in Beijing nor to locate in Shanghai it will not locate in any of the cities. So, what is the probability of that? Basically, we know what is the probability of locating in either Shanghai or Beijing or both.

So, if we know the probability and just finding up in either city is just the complement of that. So, just a complement of that 1 - probability of S union B, so, any problem when you try solving any probability problem first what you do is it first always try to write the events when you write the events, then things becomes very easier to calculate. If you do not write the events if you immediately go on try to calculate the probability then chances are there that you will go wrong.

First try to write down the events then write down what or what probabilities are given which probability is a given.

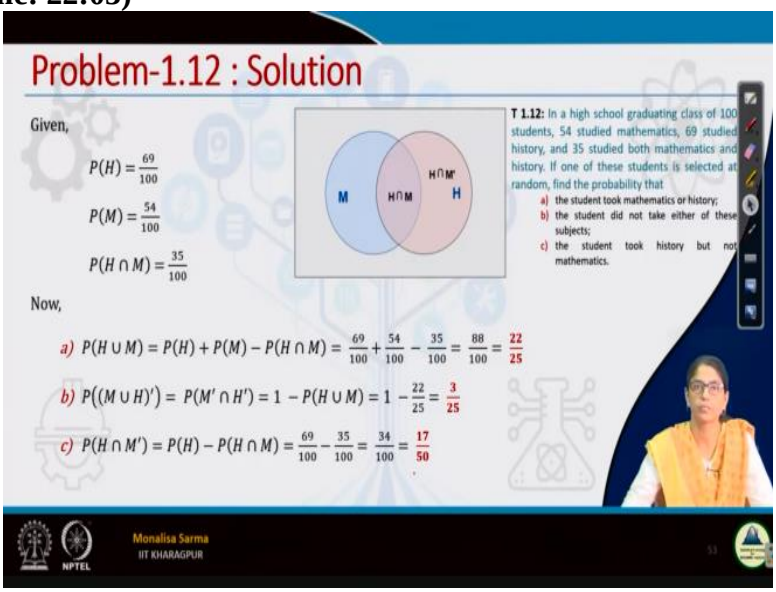
(Refer Slide Time: 21:10)



Let us try another 1 problem so, in a high school graduating class of 100 students, so, there are total 100 students. So, 54 studied mathematics, 69 studied history and 35 studied both mathematics and history. So, it is that means, if I consider 1 event as M another event as H. So, probability of M intersection is or I can get probability of M x probability of H. So, what will be that probability of M x probability of H will be 34 / 100 our total sample space is 100 is not it?

So, now, what is the question there is not a question. Question is if one of the students is selected at random find a probability that the students took mathematics or history.

(Refer Slide Time: 22:03)



So, what is the question if one of these students is selected at random find the probability that the student took mathematic or history what is a probability student of mathematics or history? How will you find that out? Is not it? It is H union M mathematics or history. So,

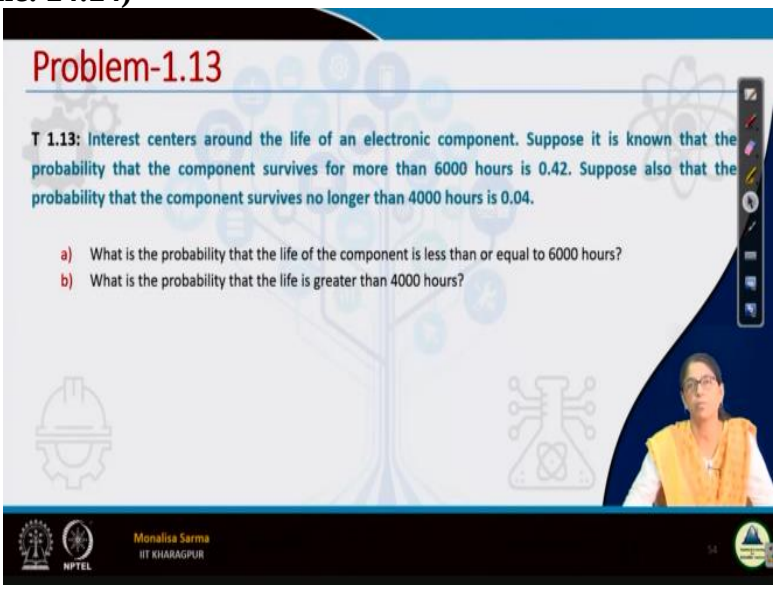
then this is what H union M the formula a probability of H is given M is given, H intersection M is given. So, we could find this out, next is the student did not take either of these subjects what is that this is just the complement of the previous one is not it?

It is just like the previous question what we have done the company located in neither of the cities so the student did not take either of these subjects. So, it is just the complement of the first case. Then second case, second case you have to notice it the student took history, but not mathematics. See, here the student took history but not mathematics. So, if you see the tests what to say this diagram what is this?

This is whole circle is for history and this whole circle this green circle is for math and this between the slight pink we can say this is H intersection M means the students who took both history and mathematics. So, now what I am interested in is the students took history, but the student did not take mathematics that means student history is this circle and did not take mathematics is this portion, this light pink portion, I will have to delete this portion because this portion student is taking both history and mathematics.

So, but I am interested only where the student took only history. So, if I find that out then this means history if I subtract this then I will get this value. So, from probability of history, I am subtracting this part. So, now of all putting the values I get the result.

(Refer Slide Time: 24:24)



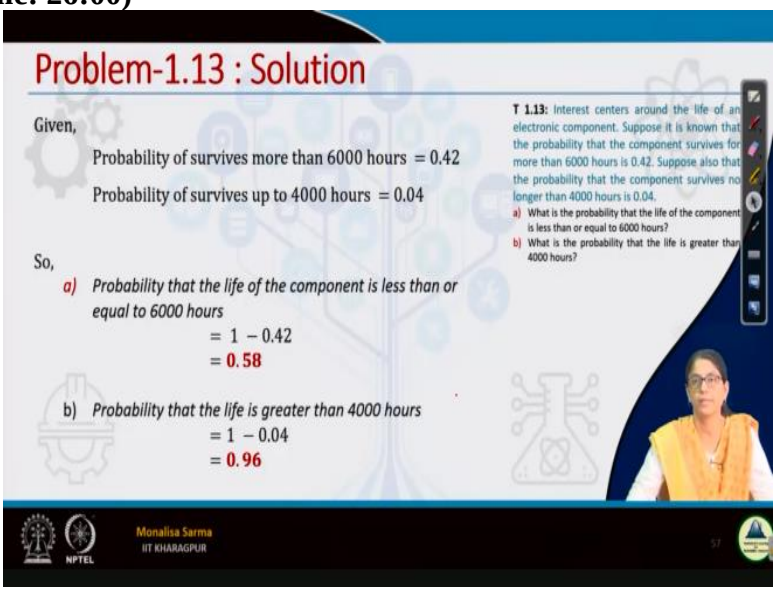
Now, the next question interest centres around the life of an electronic component. Suppose it is known that the probability that the components arise for more than 6000 hours is 0.42 to

that the component that survives more than 6000 hours that probability 0.4 again it is given that suppose also that the probability that a component survives no longer than 4000 hours is 0.04. And sometimes I will also want to mention here something sometimes in some question you see.

But if you see the question, the answer may be very simple but some extra data are given. So, do not get confused with those extra data like here you may get confused with components having no longer than 4000 hours is 0.4 and more than 6000 hours is 0.42 might be do we need to find out some intersection of 6000 and 4000 just for that, you will have to read the question properly. So, do not get confused with the information. So, here see what we have to find out?

What is the probability that the life of the component is less than or equal to 6000 hours? So you might get confused, it is given that no longer than 4000 hours is 0.4 then again for 6000 hours is some below. So do I need to consider these 4000 this value also somewhere? No, and we do not have just it is very simple beyond 6000 is the 0.04 less than or equal to 6000 is just minus of this, this is the complement of that.

(Refer Slide Time: 26:00)



It is just a complement of more than 6000 hours. Similarly, what is the next question? What is the probability that a life is greater than 4000 hours? Similarly, again, in the question only is given the component survives not longer than 4000 hours means the component has survived till 4000 hours is 0.4. Now, what I need to find out that the life is greater than 4000,

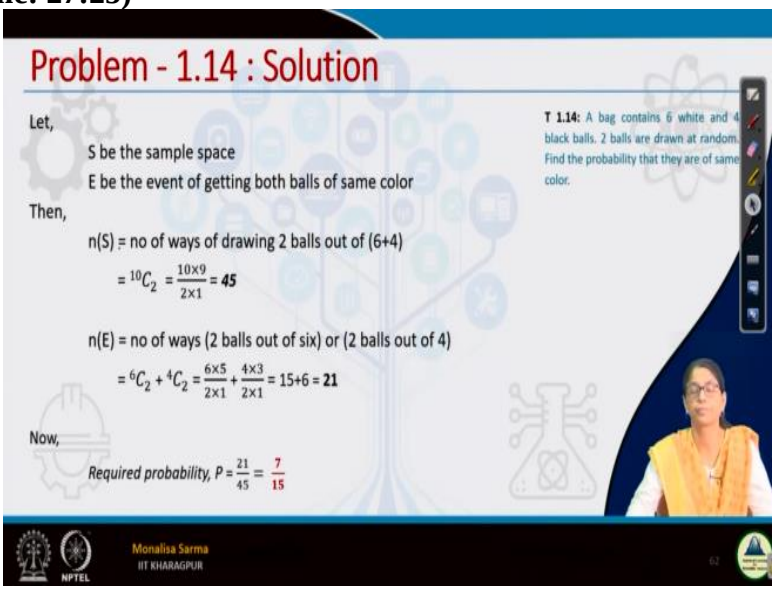
greater than 4000 the component of the death that life is better than 4000 hours is just the complement of that 1 - 0.04 so it is 0.96.

(Refer Slide Time: 26:41)

The bag contains 6 white and 4 black balls, 2 balls are drawn at random find a policy that they are of the same colour say the bag contains how many balls? 6 white and 4 total 10 balls, 2 balls are drawn at random here the position is not important. So, when we want to find out what is the total sample space, we have discussed in our last lecture probability lecture. So, how to find out a sample space for the sample space will be just from 10 we have to select 2 total sample space will be 10 C 2, is not it?

So, that is the total sample space now, what is the probability that they are of the same colour? That is the event we need to find out that event.

(Refer Slide Time: 27:25)

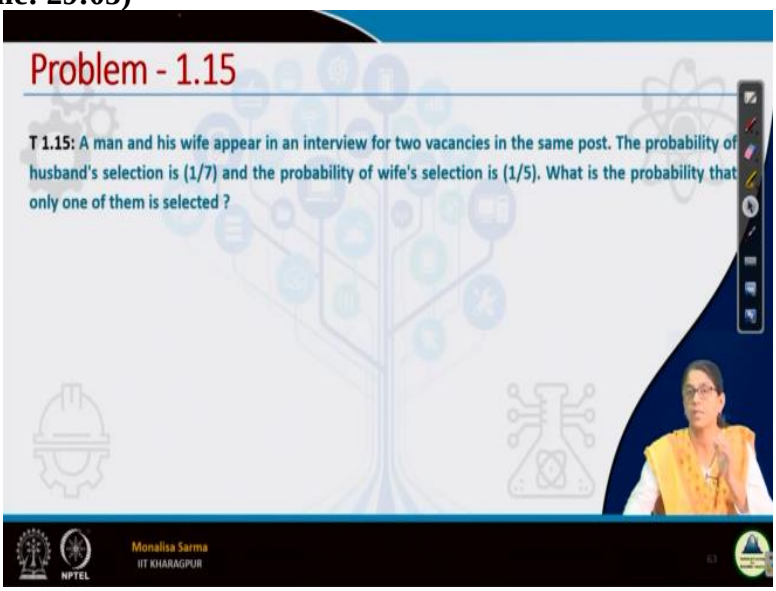


So, S is the sample space, E be the event of getting both the balls of same colour. So, now, what is the how do we calculate the sample space size of the sample space is as I told you, there are total 10 from 10 we have to select 2 where ordering is not important. So it will be simple combination. So it is 10 C 2 is the sample space. Now, how to find both the event of getting both the balls of the same colour? How do I get both the balls are the same colour?

Contains 6 white and 4 black balls. How do I get a ball? The ball so the same the same colour that means to get both the ball so, the same colour either have I have to get 2 from the 6 white balls 2 white balls from the 6 white balls or 2 black balls from the 4 black balls. So, that will be number of ways 2 balls out of 6 or 2 balls out of 4 and both are mutually exclusive again it can both have both cannot happen together.

So, it is this simple or means it is simple plus we have seen a or b is if it is mutually exclusive event so, it is probability of A + probability of B. So, it is what is the probability of A getting 2 balls out of 6, 6 C 2 when 2 balls out of 4 is 4 C 2, so n of E 6 C 2 + 4 C 2 this is the number now finding the probability this n of E upon n of S and total number of events is n of E and sample size is n of S probability is n of E upon n of S.

(Refer Slide Time: 29:05)

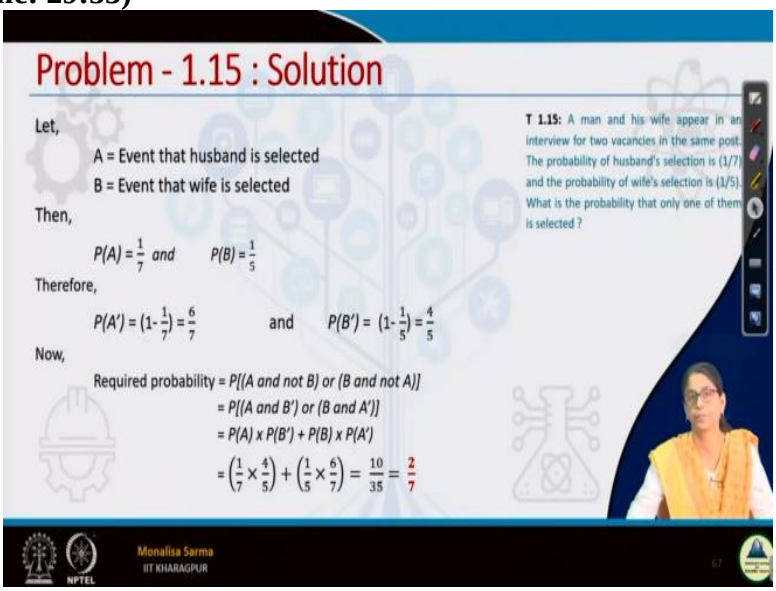


A man and his wife appeared in an interview for 2 vacancies in the same post the probability of husband selection is 1 by 7 probability of wife selection is 1 / 5 what is the probability that only one of them is selected? We have to find out what is the probability that one of them is selected and appears in an interview for 2 vacancies in the same post. So for that first

probability of husband selection is given probability of wife selection is given first we will write down the events then we will find out what is the probability of husband not selected.

What is the probability of wife not selected? So what is the probability that only one of them is selected? Only one of them is selected means husband selected wife not selected or wife selected husband not selected. These are the only 2 ways.

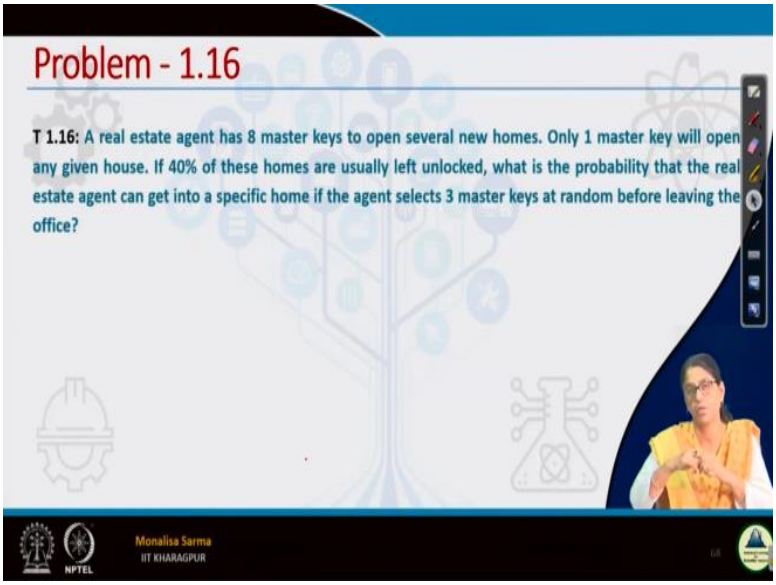
(Refer Slide Time: 29:53)



So A be the event that husband is selected B be the event that wife is selected. So, this is already given probability of A is given probability B is given. So, and then from there we found a probability of A complement that is probability of husband not getting selected and probability of wife not getting selected that is probability of B complement. Now, what is the probability that only one of them they selected.

What are the only one of them selected maybe 2 cases either husband selected or wife selected husband selected then wife is not selected while selected husband is not selected. So, required properties this probability of A and not B or B and not A, not A means complement. So, just put it in the formula all the values you have just put it in a formula and you get the result.

(Refer Slide Time: 30:47)

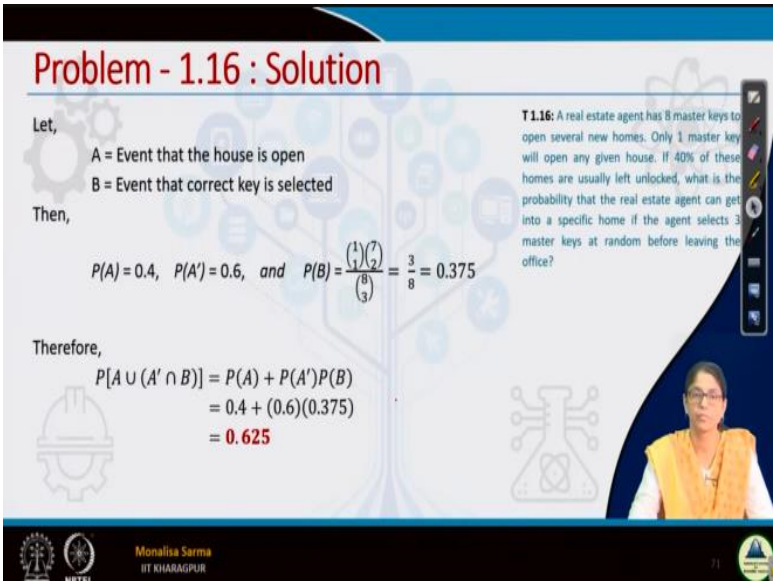


So, this is the last question in this tutorial session. So, a real estate agent has 8 master keys to open several new homes. He has 8 master keys there are not many new homes there are total 8 master keys to open those new homes, only 1 master key will open any given house. A given house can be opened by only 1 master key, though there are 8 master keys we cannot use other master keys. So if 40% of these homes are usually left unlocked what is the probability that a real estate as it can get into a specific home?

If agent select 3 master key at random before leaving the office? So what happened? What is the probability that he can enter a specific home the question is basically what is the property that a real estate agent can enter the specific home out of the 8 master keys he has selected this 3 master keys 3 keys and now and it is also given 40% of the homes are left unlocked. So, if it is the home which you want to go if it is unlocked and he does not need the key at all.

Next is suppose if the key home is locked then he has to have that particular key which will open the door whether that particular key because he has he did not take all the 8 keys here he has selected 3 master keys. So it has to be if the key falls within that 3 then only he will be able to enter that house.

(Refer Slide Time: 32:18)



So, what now first even is even though the house is open. So under what condition he can enter the house? He can enter the house under 2 situations one is on the houses open another is when the house is locked and he has the proper key. These are the 2 situations and these 2 situations are mutually exclusive. So basically if the probability of entering his home it will be just basically a probability of 2 mutually exclusive events. So, what are the 2 events?

One event is that a house is unlocked so houses what is the probability the houses unlock? It is 0.4 what is the; another event and house is locked and he has taken the proper master key. These are 2 independent events houses lock and he has taken the proper master key. So, what is the probability of taking the proper master key first is what is the size of the sample one? For finding probability always we need to find a sample space that we have seen.

So, what will be the sample space out of 8 here selected 3 to sample space will be 8 C 3 and what is the probability of taking that particular key which will open the house which he wants to enter? So that is that particular house will be opened by only 1 key. So from 1 he has to select 1 and from the rest 7, he will be selecting 2. So basically probability of B is for that if he select 1 particular key from that because each house has 1 key is not it is house can be opened by just 1 key that 1 he has selected he has taken total he has taken 3 keys.

What are the 3 keys? 1, he has taken the key for that house and another 2 any 2 from the rest 7 basically. So, this is the probability for this even property that the correct key is selected. So, probability of B we can write it this way that 1 C 1 into 7 C 2 upon 8 C 3. So now what is the probability that he can enter the house? Probability that he can enter the house is that a

probability that the house is open, plus, probability that the house is closed and he has taken the proper key. So proper key means probability of B.

So that is it therefore probability A union probability A intersection B, so probability is 0.4 and probability A complement x probability B. So that is all in today's tutorial class and what I will suggest, it is really not feasible to discuss many, many problems in one tutorial class in this sort, of course, but as to get more hands on your probability by seeing any problem you will be able to solve it. Like it is only possible if you practice more and more problems.

So it will get probability question in many books and statistics and probably there is not a single book there are many books available in the net, you can browse and in a book, you do not have to buy the books there are many free versions available in the net you can just download the books practice the problem. So essentially the more problems you solve more confidence you will get.

(Refer Slide Time: 35:47)

So, these are the references I have used. Thank you.

Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology, Kharagpur

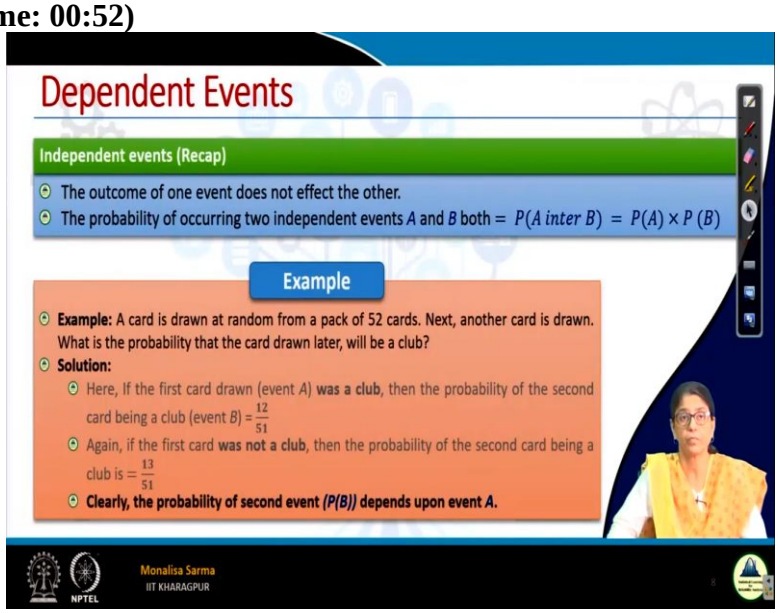
Lecture - 05
Conditional, Total and Reverse Probability

Hello everyone, so in continuation of our earlier lecture and probability theory to be will be again discussing from other concept of probability basically what to say total probability.

(Refer Slide Time: 00:38)

Will basically discussing dependent events, independent events, conditional probability, total probability and then we will also discuss Bayes theorem. So now first, we will discuss what is dependent event and independent event.

(Refer Slide Time: 00:52)



Independent event of course, we have discussed in our last class on probability that is in first class. And still we will repeat it here so that it becomes easier to correlate between dependent events. Now, when we talk of independent events, independent events are those events for which occurrence of one does not have an effect on any other, on the occurrence of the second event, like let me give you a very simple example.

Suppose to today is raining. So, probability that today is raining and what is the probability that you are attending the class. Now, if you are attending an offline class, so you have to come from the hostel to attend the class, so probability, if it rains, so this event will definitely may affect your attending the class. So, this becomes a dependent event. But however, now since all of your attending your class in the comfort of your home.

So, whether it is raining or it is not raining, so it does not make any difference. So, now, it is even that it is raining an event of your attending class, these 2 events then it becomes totally independent, that is the occurrence of one event does not have any effect on the other events then we call this as a independent event and vice versa, the other is the dependent event, if the occurrence of one has a has an effect on the occurrence of other events then we call that the dependent event.

Now, if the events are independent, so probability of occurring 2 events A and B with we have seen our last class that is probability of A intersection B. A intersection B is the probability of offering 2 events. So, that is we have seen that is probability of A cross probability of B. So, small example, like a card is drawn at random from a pack of 52 cards, from a deck of cards there total in a deck of cards total 52 cards from a deck of cards, we have drawn one card next we have drawn another one card. What is the probability that the card drawn later will be a club? So, now, there is 2 things first of all, if the first card that we have drawn, if it is a club then the second card will be a club does a different probability because it will be dependent on the first one there was number of clubs limited, right? And then it is a fixed number.

So, again, the first card that we have drawn is not a club then, the second that we have the second card, we will draw probability that it is a club that also be a different probability. So, we will see in what way it is different see here, here in the first card drawn was a club then

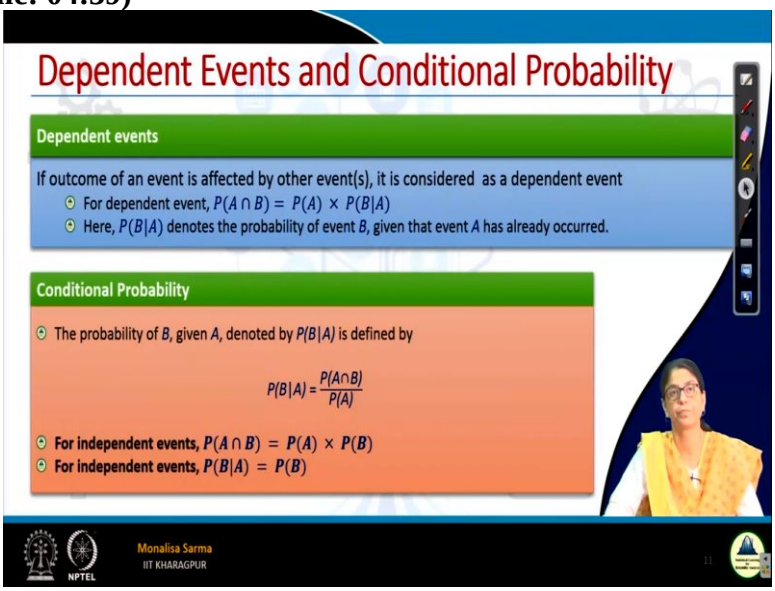
the probability of the second card being drawn will be 12 / 51, why 51 because earlier the number of cards is 52.

Now, already we have drawn the one card, so the remaining card is 51 and already their total how many clubs in that pack of card total 13 clubs already if the first card is a club, then the probability that the second card is club is it becomes 12 / 51. So, again it the first club was card was not a club, first card is not a club then what happens there are total 13 clubs in the pack and you the first slot that you have drawn is not a club.

Then what is the probability that the second card is a club that is again that will be 13 / 51 fine, it will not be 52 here also it will be 51 because already one card we have drawn that is only total non-remaining card is 51. However, if we keep back the card, we have drawn one card then we have again put back the card in the deck we have shuffle it then we draw the card then even becomes independent.

This is an example of a dependent event. Clearly the probability of your second event depends upon event A obvious right?

(Refer Slide Time: 04:39)



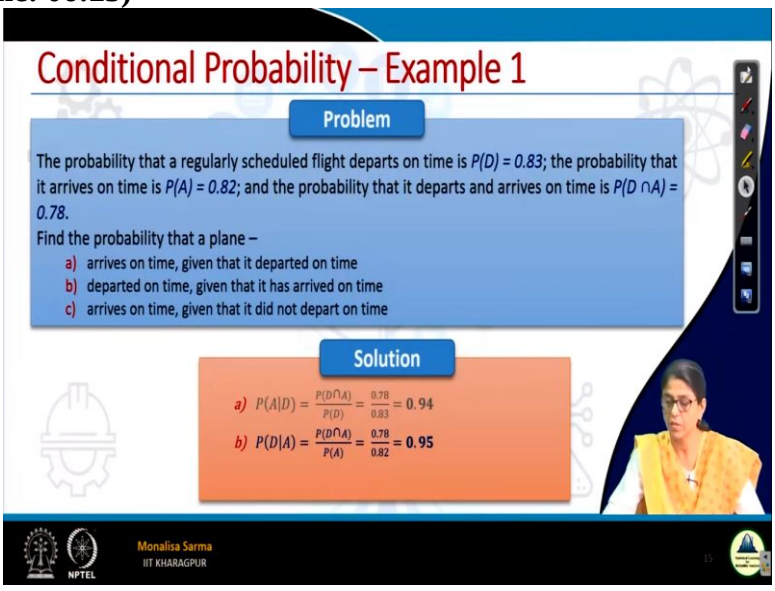
So, if the outcome of an event is affected by other events, then it is considered a dependent event. That is what I told. For dependent event, B intersection B that is probability of P probability of A and B. So, for dependent event, it is what it is not P probability of A cross probability of B, but it is probability of A cross probability of B | A, probability of A first A

as occurred then probability of B | A what is the probability that B has occurred given that A has occurred.

So, it is probability A intersection B is probability of A cross probability of B | A. So, here probability of B | A denotes the probability of event B given that event A has already occurred. So now if you are interested in finding out; so, this is the expression. Probability of A intersection B is also probability of A x probability of B | A now, if we are interested in finding out what is the probability of B | A.

Just from this expression only we can find a probability of B | A is nothing but probability of A intersection B / probability of A. So, for independent events, we have seen probability of A intersection B is follow a cross probability of B. So, what will be probability of B | A for independent event from the B | A will be probability of B it is not dependent on A so, what is the probability of B | that A has occurred, it is same because it is not dependent on A so, it will be probability of B only. Now, we will see some examples.

(Refer Slide Time: 06:23)



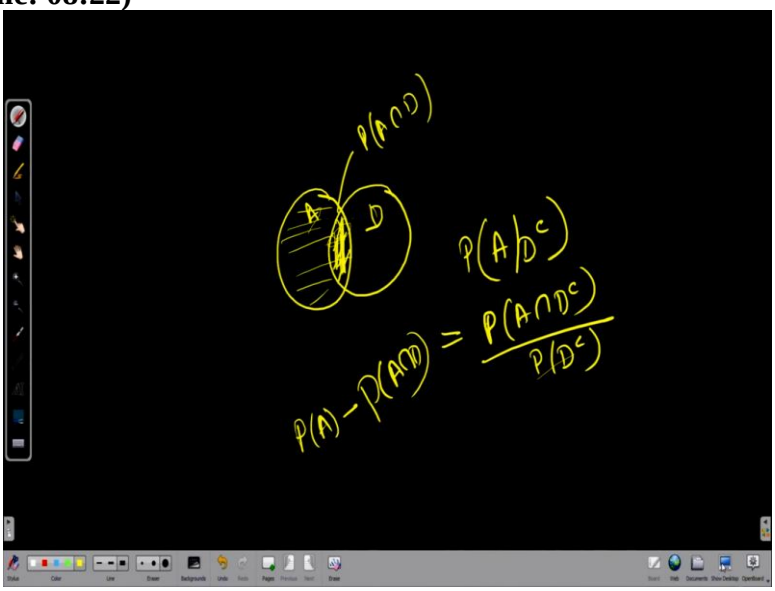
So, these examples you try to do it yourself of course, I will be explaining it. So, if you do it yourself your concept of conditional probability will be quite clear. So, the first question the probability that a regularly scheduled flight departs on time is given that is probability of D is 0.83 and the probability that it arrives on time that is probability of A is given a 0.82 and the probability that it departs and arrive on time.

That is departs on time and arrives on time that is probability of D intersection A it is given as 0.78. So, we have to find the probability that a plane arrives on time given that it has departed on time, given that it has departed on time what is the probability that it has arrived on time that means, we need to find out probability of A | D. So, what is given to us ? We need to find probability of A | D.

How what is somewhat to say equation for probability of A | D what is the expression that is probability of D intersection A / probability of D and so, probability of D intersection A is given 0.78 probability of D is given 0.83 so, we could compute probability of A | D. Now, similarly, departed on time given that it has arrived on time that means probability of D | A given that it has arrived on time what is the probability that it has departed on time.

So, same probability of D intersection A / probability of A simple formula you do not have to remember the formula if you understand the concept you will you do not need to remember the formula. So, now, see the next one that what it is asking is the probability that the plane arrives on time given that it did not depart on time. So, what is I am telling that has arrived on time given that it has not it did not depart on time we have to find that probability.

(Refer Slide Time: 08:22)

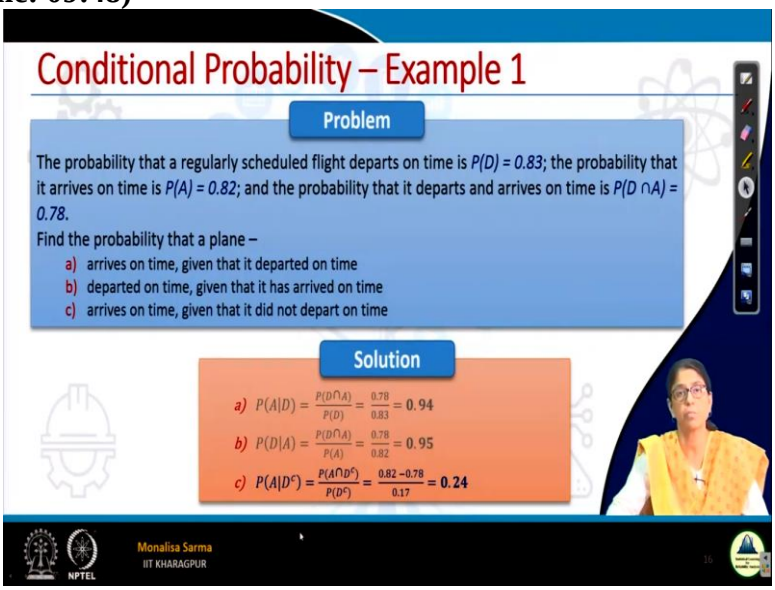


So, for that let us draw the Venn diagram first. So, if we draw the Venn diagram, see, if this is the event for the that plane arrives on time and if this is the event D that the plane has departed on time, now, this portion is probability of A intersection D that means, it has arrived on time and departed on time. Now, what it is asking probability that it has arrived on time given that it did not depart on time.

So, for that what we have to find, it is ask that probability that it has arrived on time given that it did not depart on time. So, what is this? This is equal to probability of A intersection D complement by probability of D complement. So, this is so, now, what is probability of A intersection D complement. So, that means probability of A intersection D complement is that means this whole area, this area is A this whole area. This whole area is A from this portion basically; we have to remove this portion, this portion, right?

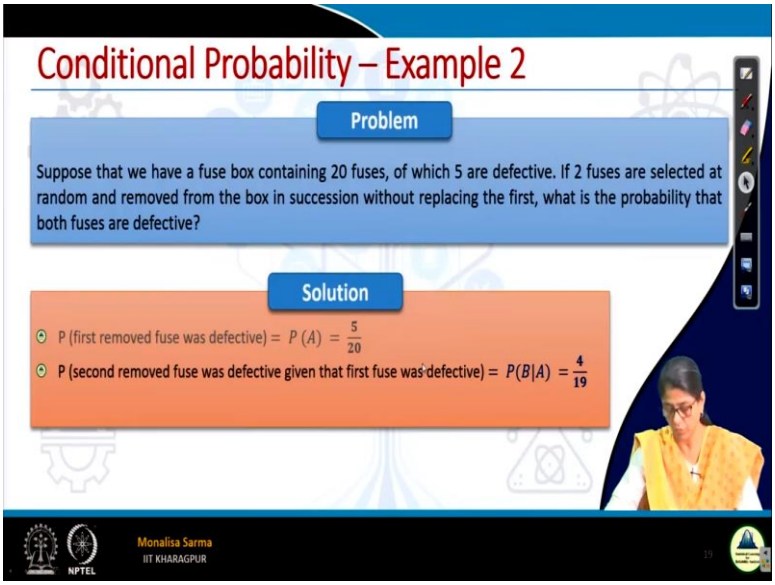
So, it will be what is probability of A - probability of A intersection D then we will be getting probability of A intersection D complement. So, from the Venn diagram, it will be able to visualise it so, it is easier if we go from the Venn diagram.

(Refer Slide Time: 09:48)



So, this is a probability of A that it has a probability that it has arrived on time given that it did not depart on time. So, this is probability of A intersection D complement / probability of D complement. So, putting the values we will get the result.

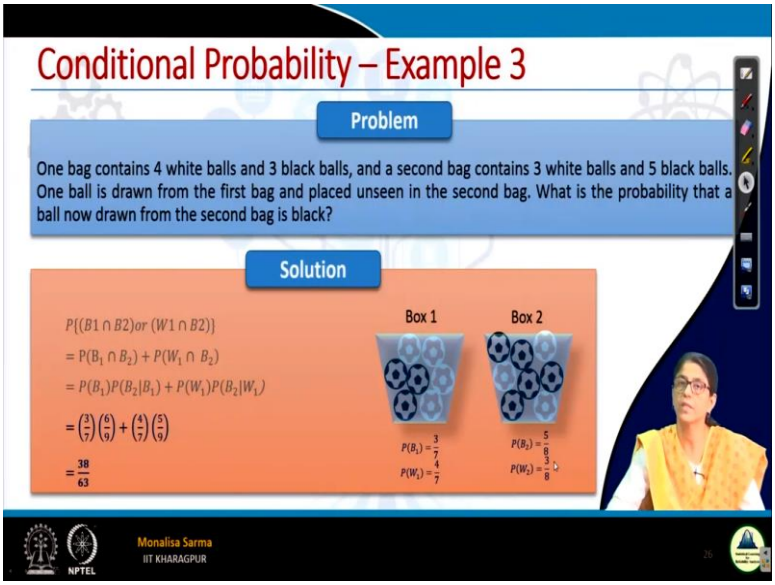
(Refer Slide Time: 10:07)



So, next example; suppose that we have a fuse box containing 20 fuses of which 5 are defective. So, my sample fuse is 20 of each 5 of defective if 2 fuses are selected at random and remove from the box in succession without replacing the first. We have selected one fuse and we did not replace it again we are taking the second fuse what is the probability that both fuses are defective.

When we are taking the first case first event probability that the first fuse is defective what is the probability the first fuses defected total there are 20 fuses and after them 5 are defective what will be the probability that the first use is defective it will be 5 upon 20. Now, when we are taking the second fuse second fuse what happened already we have taken 1 fuse that sample space has reduced from 22 it has become 19 and as well as our defective has also reduced if the first one is defective. That was from 5 it has come down to 4. So, what second probability that second remove fuse was defective given that the first fuse was defective is 4 / 19. So, then what is the probability that both fuse are defective 5 / 20 x 4 / 19. So, this is the answer.

(Refer Slide Time: 11:20)



So, next question. A bag contains 4 white and 3 black balls and a second bag contains 3 white balls and 5 black balls. This is the first box which contains 4 white and 3 black, this contains 5 black and 3 white 1 ball is drawn from the first bag and place and change in the second bag from here I am drawing 1 ball and I am not seeing what is the colour that is white or black I am just putting it here on the box 2.

Initially in a box to how many balls are there 1 2 3 4 5 6 7 8 balls there it is 8 balls here it is 7 balls. So, what is the probability that a ball now drawn from the second bag is black. What is the question I have first I have drawn 1 ball from the box 1 I did not see the colour what it is and I have put it in the box 2. Now, what is the probability that the ball now drawn from the second bag now I have drawn one ball from the second bag.

So, what is the probability that it is black. There are 2 scenarios here if the first time first draw in my first draw, if I have drawn a white ball from here and I have put it here of course my number of balls has reduced increased, but then my white balls remain same isn't it, but in the first but however if I have taken a picked a black ball in my first draw and I have put it in here, then number of balls increase along with that number of black balls also increase.

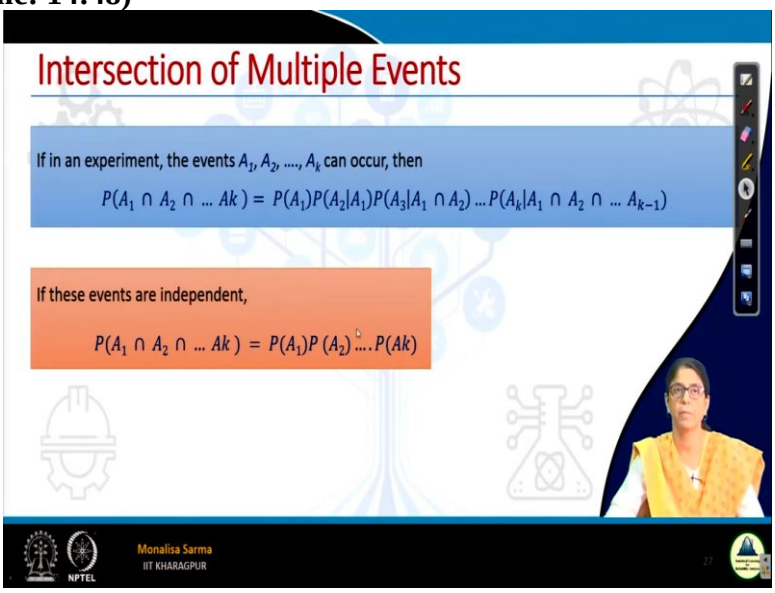
So, the accordingly probability will change. So, what I want you to this is probability that my first draw is black and second draw is also black probability that I have in the first draw I have drawn a black which I put it here without seeing it what is the probability that it is that black and second one is also black or probability the first time I have drawn a white ball and second time I have drawn a black ball.

So, what is this probability? So, this is once you know this once you can write this now it will be very easy for you to so what is probability a B 1 intersection B 2 is nothing the same it is or both it cannot be it is a mutually exclusive event both cannot happen together. So, it is just + probability or B 1 intersection B 2 + P of W 1 intersection B 2. Now, you just need to find out what is probability of B 1.

Now, here see probability of B 1 is I am drawing it first time, so that is independent. So, it will be there what is the probability that the first ball is black? So, what is the probability of the first ball is black here how many balls there are 7 balls out of the 3 black balls probability that it is black is 3 / 7. So, probability of B 1 is 3 / 7. Now given that the first one is black. What is the probability with that second one is black? Given that the first 1 is black.

So, now here my black balls earlier it was 5 now my black box has become 6 and total how many balls it was earlier it was 8 now it has become 9. So, it is 6 / 9. So, my probability of B 2 | B 1 has become 6 / 9. Similarly what is the probability of W 1, probability of W 1 is 4 / 7 and what is a probability of B 2 | W 1 initially if I picked a white then my black balls remain same. So, what is how many black balls? 5 black balls and total black balls is now 9 that is 5 / 9. So that is it.

(Refer Slide Time: 14:48)

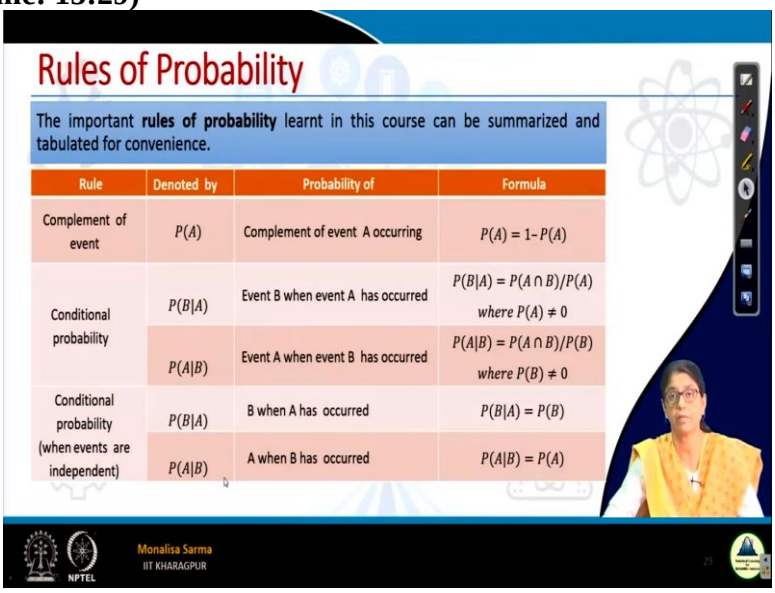


So, now if your intersections are interested in the intersection of multiple events, even experiment the events that A 1 A 2 A k can occur, then what is the occurrences of it is multiple events like probability of A 1 and A 2 and A 3 and A 4. So, if it is an independent

event then it is just a multiplication of all the events. If it is an dependent event then of course, we have to go for dependent probability that is a conditional probability.

Probability of A 1 probability of A 2 | A 1 probability of A 3 | A 1 intersection A 2 and likewise, but if it is independent simply multiplication of all the probabilities.

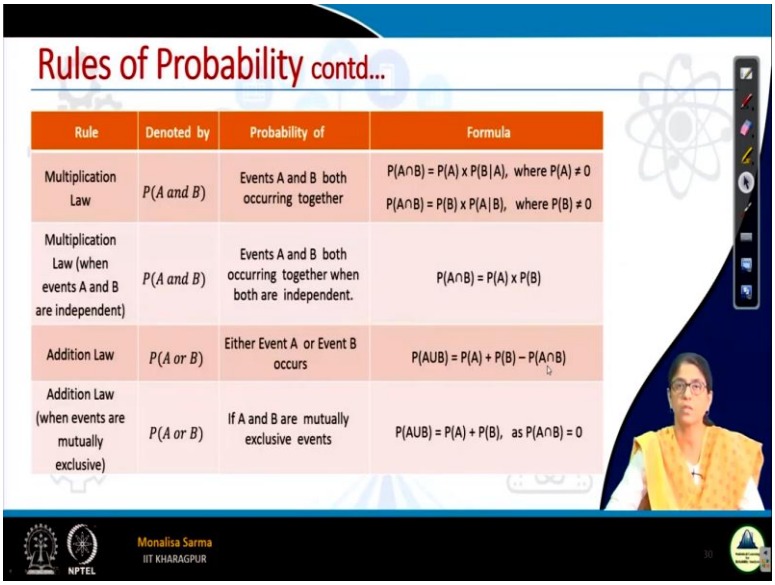
(Refer Slide Time: 15:29)



Now, here, I will just summarise the probability rule still know what we have learned I am not going to introduce anything new, but follow whatever we have learned till now, that is complement of the probability if it is P(A) is the probability its complement is ~P(A) that is 1 - P (A) conditional probability of B | A what is probability of B | A? Probability or B | A is probability of A intersection B divided by probability of A what we have already seen.

But where, where probability condition is a probability of A is not equals to 0. Similarly, probability of A | B same thing it will be probability of A intersection B upon probability of B here where P of B is not equals to 0. Now, when does events are independent, what is the probability of B | A it is probability of B. What is a probability of A | B that is probability of A if it is independent events.

(Refer Slide Time: 16:21)

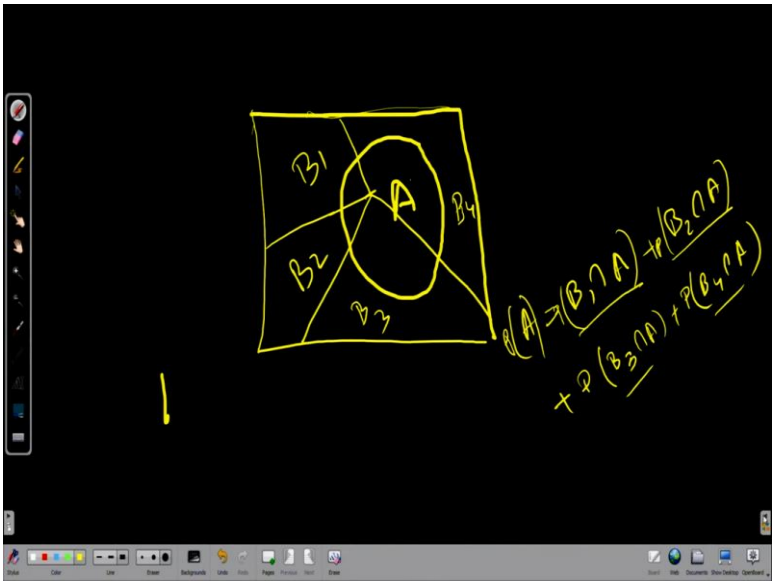


So, then multiplication rules what we have seen that is probability of A and B. Probability of A and B is nothing that probability of A x what is B probability of B | A if it is dependent, if it is independent simple probability of A cross probability of B what we have just seen it is just an I have summarised it in this table, we have even see probability of A or B also that is A union B probability of A + probability of B – probability of A intersection B and if it is mutually exclusive event, do you remember what is in case a mutually exclusive event. So, probability of A intersection B this portion will be 0. So, problem A or B will be simply probability of A + probability of B.

Now, total probability, it is something many people find it very complicated. If you understand it properly, then you will see it as very easier if you can conceptualize it, if you can understand it, what is actually total probability where we are and applying it. So first, so I am requesting you please pay attention and try to understand it properly.

So, when we talk of total probability first, before we going to the slide, let me first explain you with the help of this small example.

(Refer Slide Time: 17:33)



So, suppose I am interested in finding an event A, I do not know the occurrence of this event, what is the probability of occurrence of this event see what I know is that A is covered by some disjoint events. So, A is covered by some disjoint events so let me take the pen A is says suppose this is the universe sample space universal sample space. So, A is covered by some disjoint event.

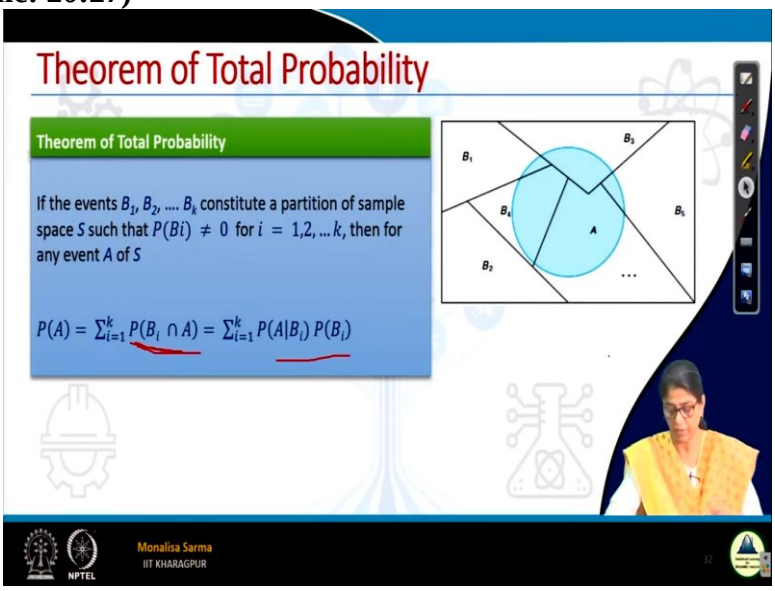
So, some disjoint event say suppose this is B 1, this is B 2 B 3 B 4. So, A is I do not know that occurrence of probability of A but I know that some disjoint events collectively cover A these are the disjoint and this is the universal sample space, this is the outer one is a universal sample space. So, what are the disjoint events B 1 B 2 B 3 B 4 these areb the disjoint events. Disjoin event means there is intersection between the time is 0.

Now, what is that, I know these disjoint events collectively forms A and this scenario if I know the probability of this scenario, how much there is some probability of this how this A is formed by what? A will be formed by summation of all this. So, basically A is B 1 intersection A + B 2 intersection A + B 3 intersection A likewise all the + B 4 intersection A, I know this probability, this is probability fine let me write it a bit clearly plus probability of B 3 intersection A + probability of B 4 intersection A.

So, A is formed by this disjoint element. A formed by this disjoin events B 1 B 2 B 3 and B 4 and I know the occurrence of this probability of occurrence of this, probability of this scenario. So, if I know this then I can find a probability of event A using this is not it, is it

just a summation of all these disjoint events. So, that is only total probability nothing else nothing else. So, it is that easy.

(Refer Slide Time: 20:27)

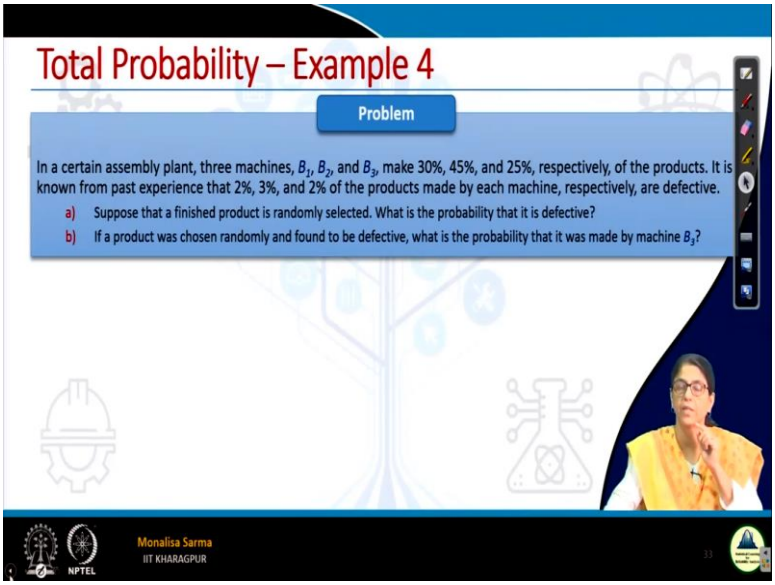


So, so, we are interested in this finding out what is this event A and we know A is formed by this disjoint event B 1 B 2 B 3 and up to B 6 it is there. So, what is probability of A? Probability of A is summation of all this. So, what is this intersection probability of B 1 intersection A what is that it is nothing but probability of A | B 1 x probability of B 1 simple conditional probability what we have seen for in this is when we are interested in finding out the A intersection B.

So, this is the formula for total probability. So, probability of A is summation of simple summation of this fellow, this summation of all these disjoint events so, now I know what is the probability of A intersection B. Probability of A intersection B is what? Probability A intersection B is probability of A | B x probability B so, that is what so, this instead of A intersection B I have written it in this form that is all that is total probability.

Now, this here again definitely we will be doing some example, but total when we talk about total probability again other extension is the Bayes theorem. So, what is Bayes theorem now, we know that some even A has occurred, what is the probability that this event has occurred because of set event B 1. A is occurred what is the probability that it has occurred because a B 1 or what is the probability that it is occurred because of B 2 that is Bayes theorem that is all I will see it more.

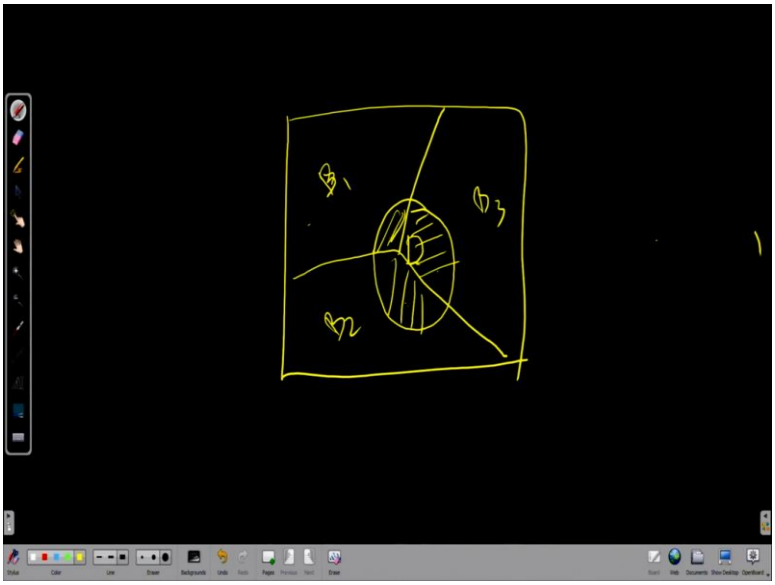
(Refer Slide Time: 22:11)



So, first let us do an example on total probability then we will go to Bayes theorem see here in a certain assembly plant 3 machines B 1, B 2, B 3 make 30%, 45% and 25% respectively of the product. So, in a universe of space total 3 years 100%, so, we have I am telling it is a universal sample space. So, it is B 1, B 2, B 3 3 machines that makes a whole all type of products like so, it is B 1 makes 30% a product, 45% is made by B 2 and 20% is made by B 3.

The 3 machines made in that assembly plant the 3 machines make all the product so, that is the universal samples was basically, it is known from the past experience that 2%, 3% and 2% of the product made by each machine respectively are defective. So, 2% of 30% is defective, 3% of 45% is defective, 2% of 25% is defective fine. So, I am interested in finding out so, what is suppose a finished product is randomly selected. What is the probability that it is defective, a finished product is randomly selected what is the probability that is defective.

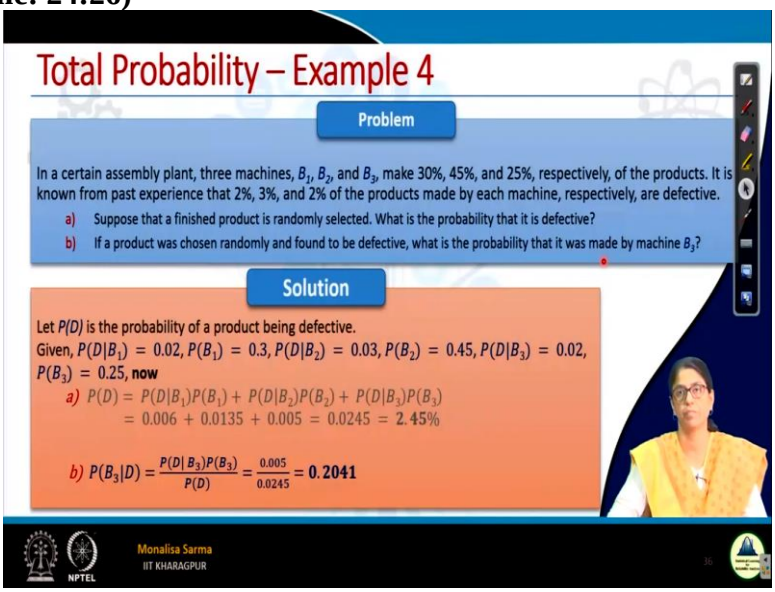
(Refer Slide Time: 23:24)



Here also basically is finished probability is randomly selected what is the probability say this is the sample space for defective product. Now finish, what are the different product and a universe of sample space, we have total 3 machines B 1, B 2 and B 3. So, portion of B 1 is defective means this portion we have seen is not a 2 portions some 3 portions and this figure is not okay not accurate some 2 portion and this is defective because of B 2 this portion is defective because of B 3.

So, we know that in this the disjoint areas which forms D so, then what is this whole D this nothing but summations of all this will bring give me D fine.

(Refer Slide Time: 24:26)



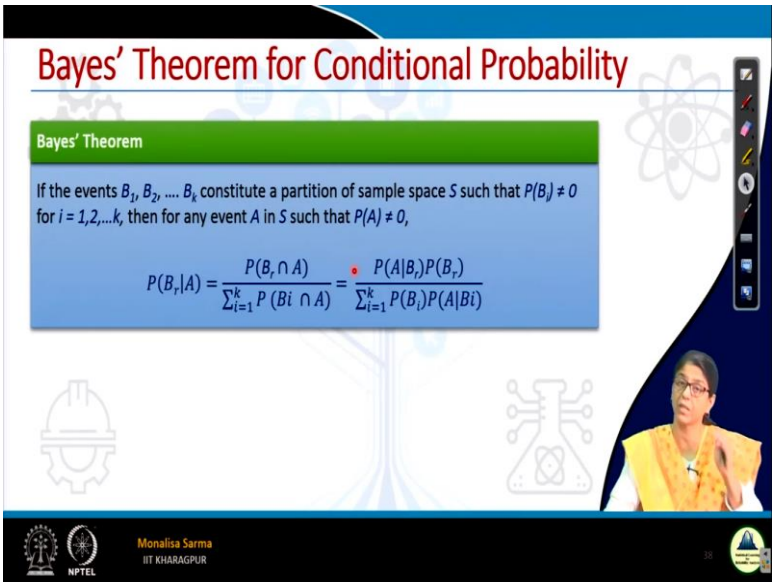
So, probability a P D is probability of probability being defective. So, what is given to us probability that is defective given that it is from B 1 is 2%. So, 0.02 probability of B 1 what is the probability of B 1? Probability of B 1 is 30% and probability that it is defective given that

it is from B 2 is 3% and what is the probability that B 2 is 0.45 45%, similarly what is the probability that it is defective given that it is from B 3 again 2%.

And probability of B 3 is 0.25 now, we need to find out what is probability of D, probability of D is nothing but the summation of all the intersection part probability of B 1 intersection D + probability of B 2 intersection D + probability B 3 intersection D. So, probability of B 1 intersection D is this portion. Then probability of B 2 intersection D is this portion, for B 3 intersection D is this portion then put the values are given and you will get the result.

Now, next either product is chosen randomly and found to be defective what is the probability that it was made by we should not be this is based your at chosen a product and it is found to be defective now, what is the probability that it is from a machine B 3 or this from machine B 1 or from machine B 2 if you understand that is nothing but that is the Bayes theorem. So, first let me go to the next slide.

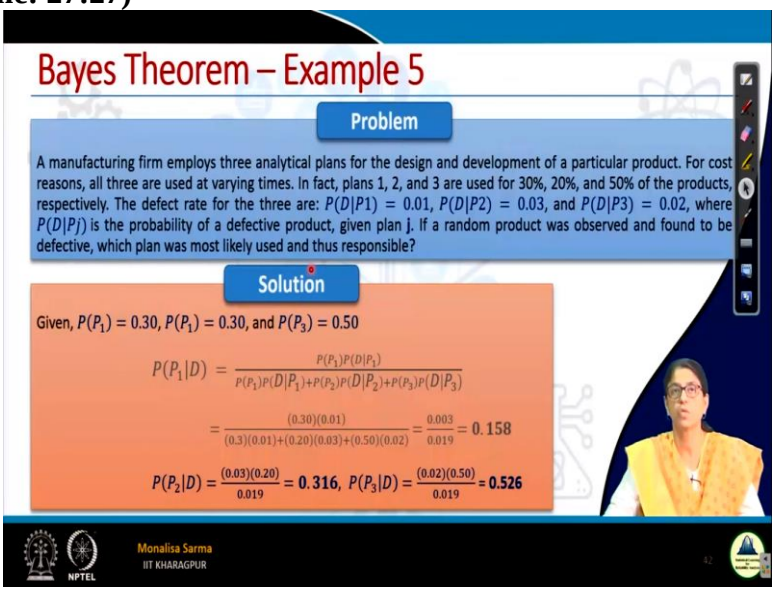
(Refer Slide Time: 26:20)



Then we will come back to this question again. So, what is Bayes theorem if the event B 1 B 2 be a constitute a partition of a sample space as such that probability B i is not equals to 0 and then for any event A in S such that P(A) ≠ 0 what is the probability that given that A has occurred what is the probability of B_r. So, it is nothing but probability of A | B_r x probability of B_r and divided by whole sample space what is whole sample space means the total probability.

Because we have found that it is defective so, what is the sample space? Sample space means that the whole defective portion which is the whole defective portion here in our question, the whole effective portion is the big circle D that we have drawn. So, that is the sample space that was the total probability. So, in the denominator is our whole sample space that is a total probability and a numerator is probability of B_r | A how we can find that so, that means probability that it was defective given that it was B_r x probability of B_r.

(Refer Slide Time: 27:27)



So, now so, we can we are coming back to the same question again if the product was chosen randomly and found to be defective, what is the probability that it was made by machine B 3. So, simple this was probability of B 3 given D so, this is nothing but the uses of Bayes theorem. If you have still doubt and total quality and Bayes theorem I suggest you please go through different examples and while trying to solve the total probability question on total probability a question and based on always try to draw the Venn diagram.

Whatever event you are interested in finding out find to make a circle for that and make it universal sample space give the what to say designed partition, and then try to find out the required probability of the particular event which are interested to find out. So, always draw out draw the figure or a Venn diagram and try to do it. If still, if you have doubt, we can clear this in the doubt clearing session.

But I do not think you will be having any doubt solve as many practice as many problems as you can. So, we will solve one more problem here. So, a manufacturing firm employs 3 analytical plans for the design and development of a particular product for cost region all 3

are used at varying times, plan 1, 2 and 3 are used for 30% 20% and 50% of the product respectively.

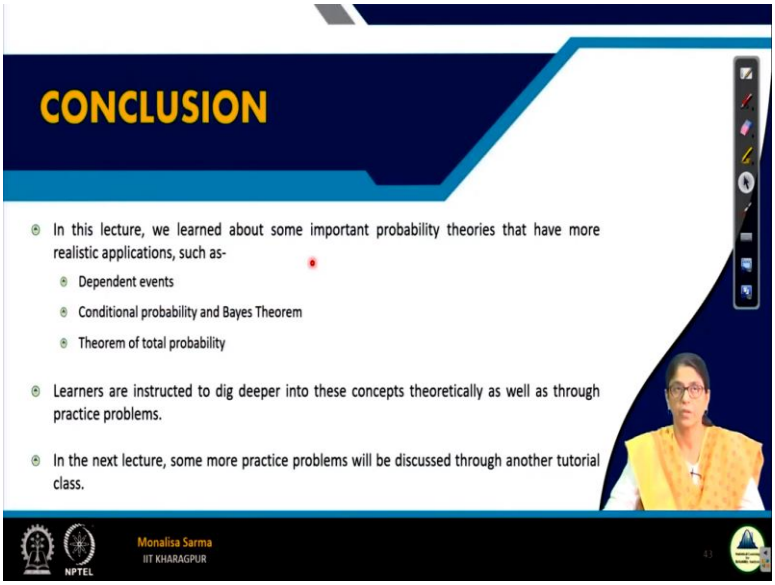
The defect rate of the P_r are it is given probability that it is defective given that it is P1 is 1% probability that it is defective given that it is from P2 it is 3% and probability that it is defective given from P3 is 2% where probability of $D \mid P_j$ is probability of a defective product | Plan j. If a random product was observed and found it to be defective, which plan was most likely used and thus responsible?

So, I have selected an entire product and I found it to be defective. Now, I need to find out which plan has most likely to be chosen, so for this what I will have, it is definitely a question Bayes theorem and what is what I will try to do, I will try to find out the probability of all the 3 options probability that it is the defective given that it is defective probability that it is from plan 1.

Since then, I will find out given that it is defective what is the problem is that it is from plan P 2 third I again I will find out given that it is defective, what is the probability that it is from plan 3 P 3. So, when I will calculate all these 3, whichever probability that is highest that is the most likely and thus responsible. So, these are given probability of P 1 probability of P 2 probability of P 3 given the 3 plans.

And so, what is the probability P 1 given the simple use of Bayes theorem denominator is a total theorem. So, we got probability of P 1 given the similarly we will find a probability P 2 given the probability P 3 given the and we found a probability P 3 given the is the highest so that means it is the which plan was most likely use so plan 3 was most likely use and that is responsible.

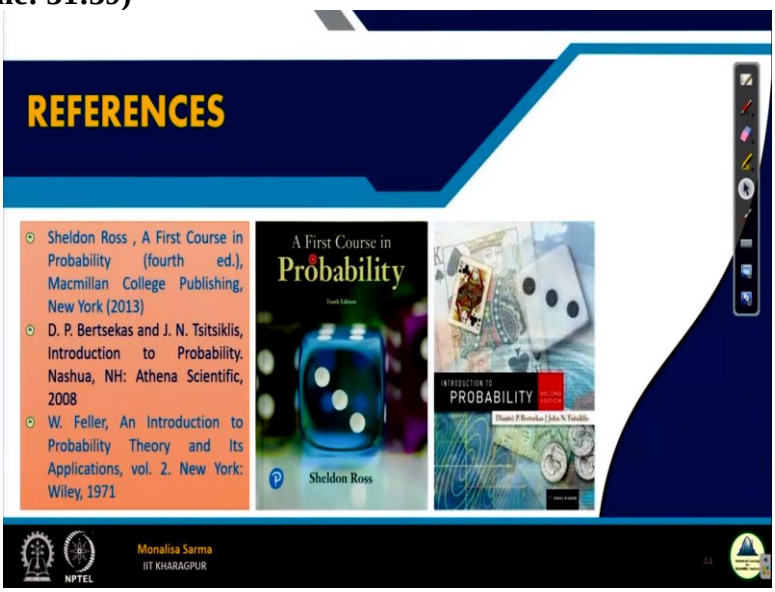
(Refer Slide Time: 30:51)



So to conclude this lecture, in this lecture we learned basically some important probability theories that have more realistic applications like based theorems total probability has many realistic application in many research studies in many industrial applications, you will find various application of Bayes theorem then we have this what to say we have learned as conditional probability as well dependent what is dependent events.

And learners are also, request to instruct to dig deeper into the concept theoretically as well as through practice problems, then your concepts will be more in clear. So, in the next lecture, definitely, I will be doing a tutorial class we will be practising some of the problems.

(Refer Slide Time: 31:39)



Then, these are the references which you can refer for some practising a problem, but these are not the only ones you can come to refer any book on probability and thank you.

Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology - Kharagpur

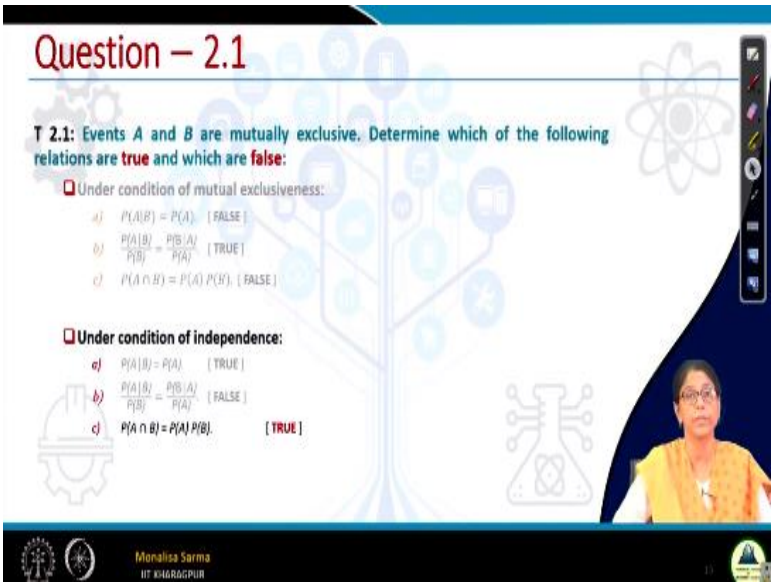
Lecture – 06
Tutorial on Conditional Probability and Total Probability

Hello all of you. So, today we will be discussing some problems which the probability theory what we have discussed in the last 2 classes, last 3 lectures in fact, one lecture on tutorial and we have discussed probabilities on 2 lectures. Now, the last lecture we have discussed probability on conditional probability and total probability today, we will be doing tutorial on those topics.

(Refer Slide Time: 00:49)

So, I will also solve we will later we will see some objective type questions, and mainly we will be doing some problems which will really help in building a solvable problem, problem solving aptitude. So, now first objective type question.

(Refer Slide Time: 01:04)



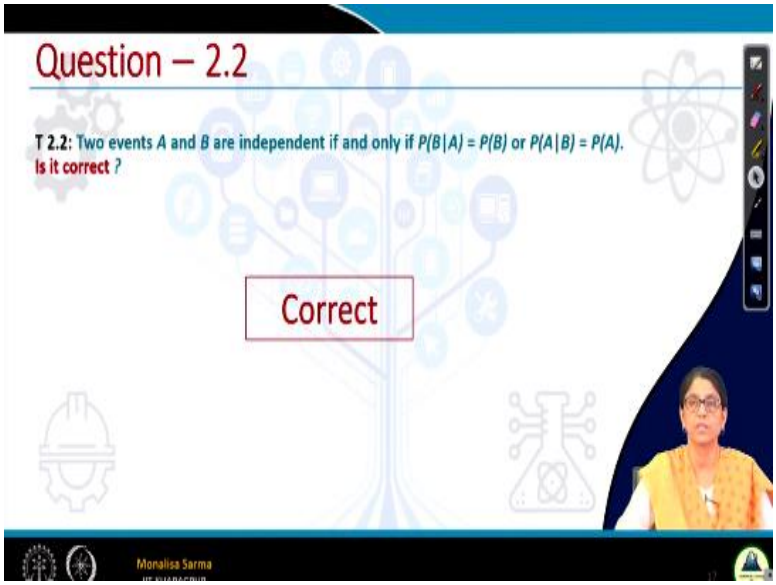
So, very simple objective type question all of you will be able to answer it just see very carefully, what it is given us events A and B are mutually exclusive determine which of the following relations are true and which are false. First one is given that A and B are mutually exclusive. So, even if A and B are mutually exclusive, so then the first question that is A, what it is given probability of A given B it is equals to probability A it is given. So, it is true or false what it is telling that event B has occurred, what is the probability that A has occurred?

Now, since A and B are mutually exclusive, what will be the probability of A given B it will be definitely it will be 0 it will not be probability A so, it is false. Now, coming to the second one what it is given? Probability of A | B / probability of B = probability of B | A / probability of A. So, I can now we will be able to answer this question what it will be because probability of A given B that is 0, 0 by PB is definitely 0, right hand side also it is 0. So, this is true.

Then a third question, third question what is probability of A intersection B is equal to probability of A into probability of B again mutually exclusive events we have seen probability A intersection B is 0. So, this is false. Now, we will see under condition of independence if the event A and B are independent, then under that condition a probability of A given B is also probability of A it is true or false. Of course, it is true.

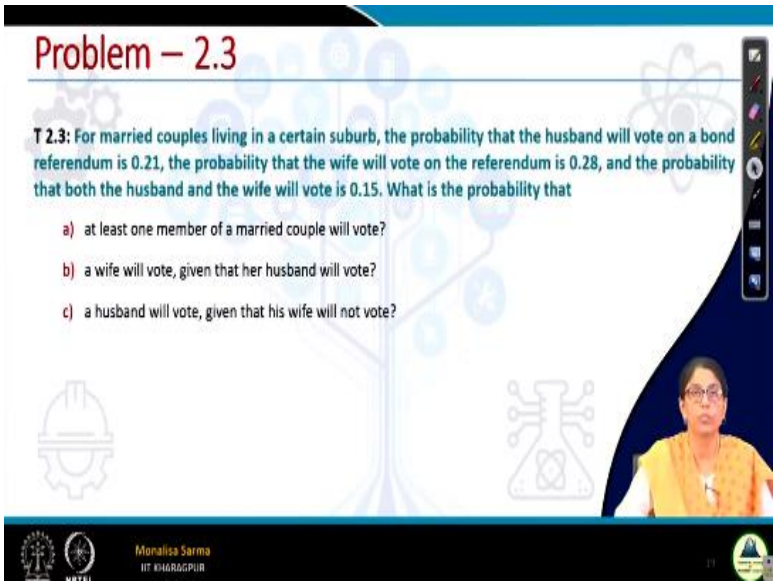
And the condition of independence, it does not affect whether the event B has occurred or not. So, probability of A given B is probability of A then what is probability of A given B upon probability of B is it equals to probability B given A by probability of A that will be able to do it by yourself and that is obviously that is false. Next is probability of A intersection B, probability A intersection B is also probability of A into probability B and that is true in case of independent event.

(Refer Slide Time: 03:15)



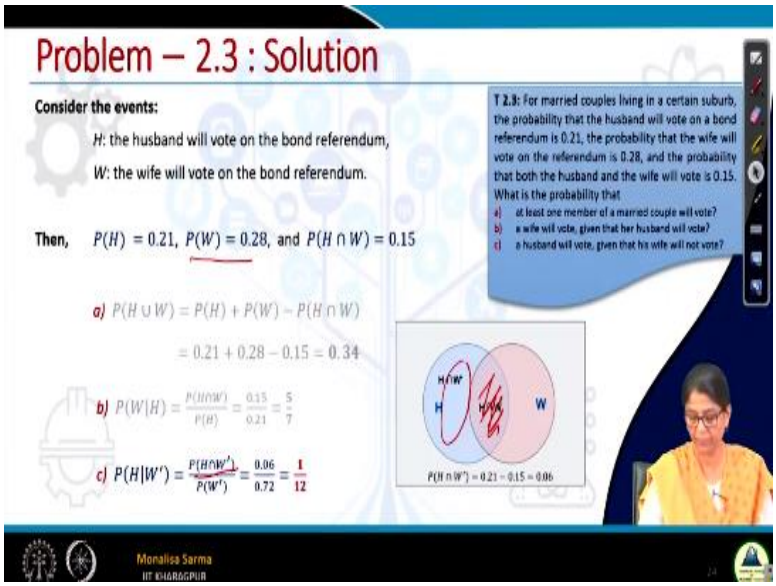
Second questions, 2 events, A and B are independent if and only if probability of B given A is P of B and probability of A given B is probability of A is it correct? Obviously that is correct. So, it is a very simple question. Now, we will be solving some big problems, which will really help you to understand the concept of conditional probability as well as the total probability.

(Refer Slide Time: 03:40)



First question, for married couples living in a certain suburbs, the probability that the husband will vote on a bond referendum is 0.21. The probability that a wife will vote on the referendum is 0.28. And the probability that both the husband and wife will vote is 0.15. So, what is the probability that a at least 1 member of a married couple will vote? Now the question is that here in my last tutorial, I have also mentioned that whenever any probability problem always first try to write down the events. If you write down the events and solving it will be become very easier. It will be able to picture it actually clearly.

(Refer Slide Time: 04:20)



So, now what is it? So, see here the Venn diagram, from the Venn diagram it will be clear because what it is given the probability that the husband will vote that is okay, let me show it for

a pen, this portion, this whole portion, this whole portion is the probability that the husband will vote and the probability that the wife will vote is the other portion is the W this circle.

Now and what it is given probability that both the husband and wife will vote. Both of them will vote means that is the intersection part. So, which one is the intersection part the slight pink this one, this one is the intersection part. So, all these probabilities are given. So, and this is given to us probability of H is given, probability of W is given, probability of H intersection W both husband and wife will vote that is H intersection W and both is independent.

So, now, what it is given, what is the probability that at least 1 member of married couple will vote at least 1 member means what we need to find out at least 1 member means either husband will vote or wife will vote or both will vote at least 1. At least 1 means it is union so, it is probability of H union W. So, what is probability of H union W we know the formula probability of H plus holder W and we will have to subtract the H intersection W part remember why we have subtracted it.

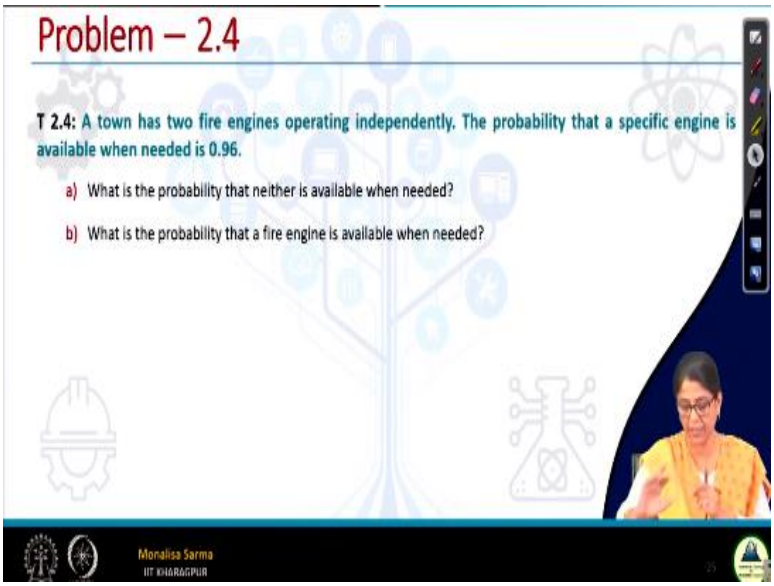
Because we have calculated that part twice once while we have calculated it for H we have taken that part again when we have done for W we have taken that part. So, we have taken it taking that part twice. So, we have to subtract it once. So, this problem H intersection W H union W yes we will be doing by using this formula. Then I guess next question a wife will vote given that the husband will vote? A wife will vote given that the husband will vote so, what is that means what given that has been that means probability that probability of W given H.

So, probability of W given H that is probability of W given H probability of H intersection W probability of H probability of H intersection W is given to us probability of H is given so, we will be able to solve it. So now, then similarly the next question a husband will vote given that his wife will not vote see, the third is a bit complicated not very complicated, if you draw the Venn diagram, so, always draw the write down the events and then draw the Venn diagram then things will become very clear.

Now, from the picture you can understand it has given a husband will vote given that the wife will not vote. So, husband will vote in this picture is a circle of H. This is the circle for H where husband will vote and this portion, this light pink portion, this is the portion where both the husband and wife will vote. Now, it is given that a wife will not vote that means from the husband part from this whole circle, we will have to subtract this portion.

Whole circle is probability of H from this whole circle, we will have to subtract this portion listen it? So, what is this probability of H given W bar that is given the wife will not vote is a probability of H intersection W bar by probability of W bar, W bar is from P of W we can easily find from P of W we can easily find P of W bar. Now, what is H intersection? H intersection W is from H we will minus this part then we will get H intersection W.

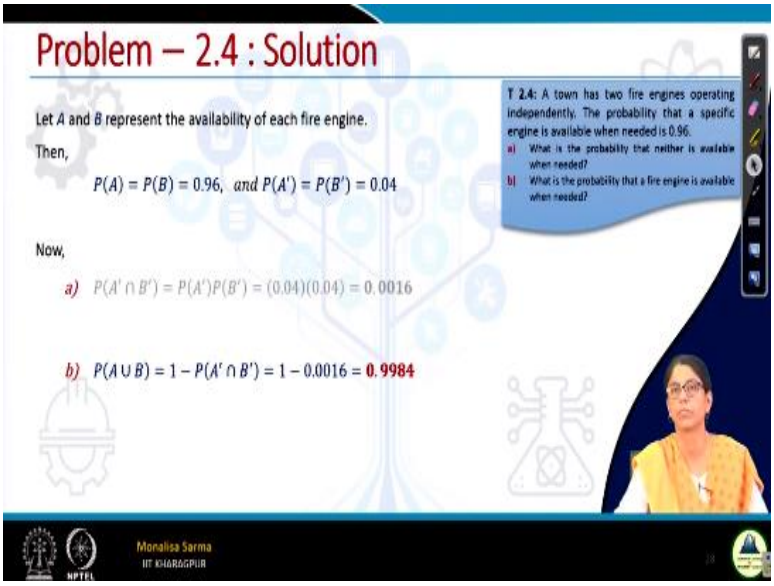
(Refer Slide Time: 08:18)



Now next sum, a town has 2 fire engines operating independently. That is independent events. The probability that a specific engine is available when needed is 0.96. Both the fire engines are what independently there are no dependency between them. So, what is the probability that neither is available when needed? So, first we will have to find out what is the probability that lable, then what is the probability of that the other one is not available?

Because there is independent events on when we try to find out when both is not available it is just a multiplication of both the events. That is all, is not it? When it is independent event what is probability A intersection B is probability A into probability of B.

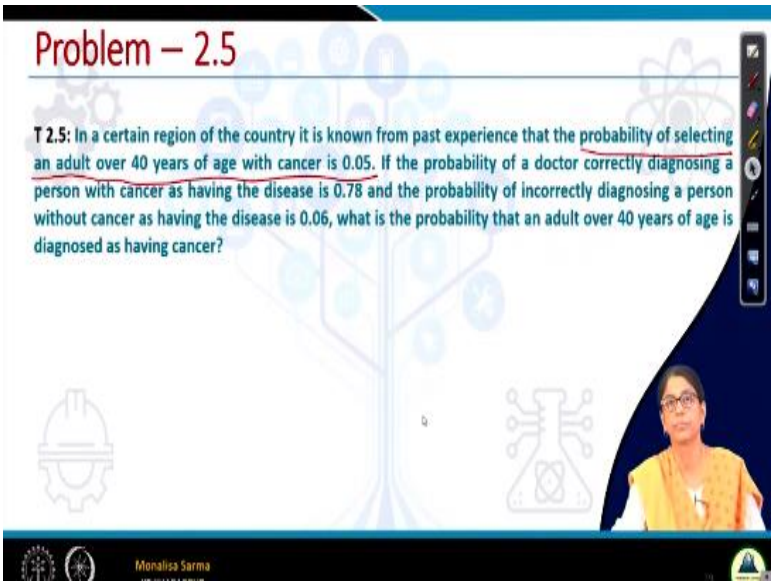
(Refer Slide Time: 09:05)



So, we have what is P of A, we have what is P of B and from that we can find out that ~A ~B it is just 1 - P (A) will give us one ~P(A). Similarly ~P (B). So, once we got that to find out value A intersection ~A intersection ~B, it is just a product of disturbance and then what is the probability that the fire engine is available when needed?

See here what is the probability that the fire engine is available when needed that means fire engine is available when needed that is can be 1 engine may be available, 2 engine may be available means either 1 is available or both are available that means when I need fire engine that is 1 that means either 1 is available or both are available. So that is A union B probability of A union B, then just put it in the formula.

(Refer Slide Time: 10:06)

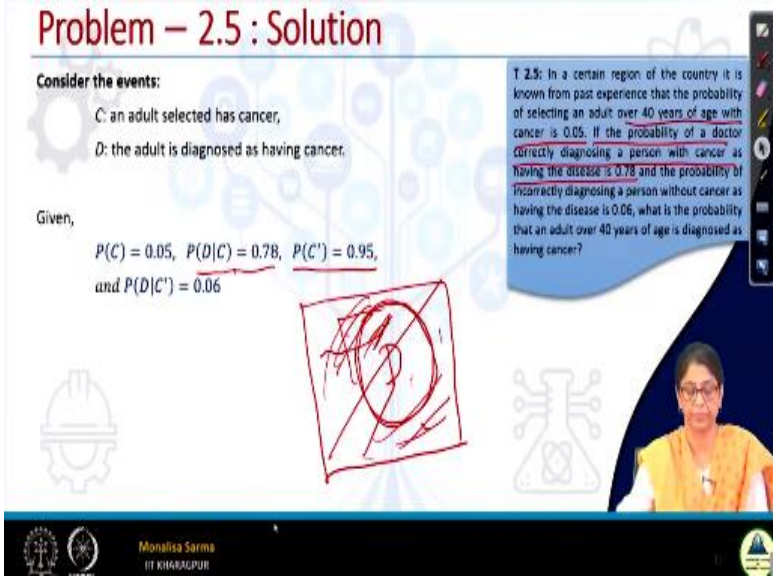


So, draw the Venn diagram, write down the events. For all the case, you do not need to write down the event or draw the Venn diagram of course, but events you will write for all the cases and even if you draw a Venn diagram, then things becomes more clear. So, next in a certain region of the country; it is known that from past experience, that the probability of selecting an adult over 40 years of age with cancer is 0.05, so what it is given the probability of selecting an adult over 40 years of age with cancer is 0.05.

If the probability of a doctor correctly diagnosing a person with cancer as having the disease is 0.78 that means the person has cancer and the doctor has diagnosed that is he has cancer that means given that the person has cancer, the doctor has diagnosed that he has cancer, conditional probability. And the other one is and the probability of incorrectly diagnosing a person without cancer is having the disease that means given that the person does not have cancer, but a doctor has diagnosed that he has cancer.

So now, what is the question what is the probability that an adult over 40 years of age is diagnosed as having cancer. So, this is clearly a question of what can you understand.

(Refer Slide Time: 11:35)



So, it is a question of total probability. So, first we have drawn, we have written down the events first season and adult selected has cancer that in their adult over 40 years of age, I have just removed age, because age is just a redundant information here, this is not very important information here sometimes in some question, you will see some very redundant information means which are not very much required. So unnecessary, we do not need to drawn those while writing down the events or sometimes we totally need to ignore some information.

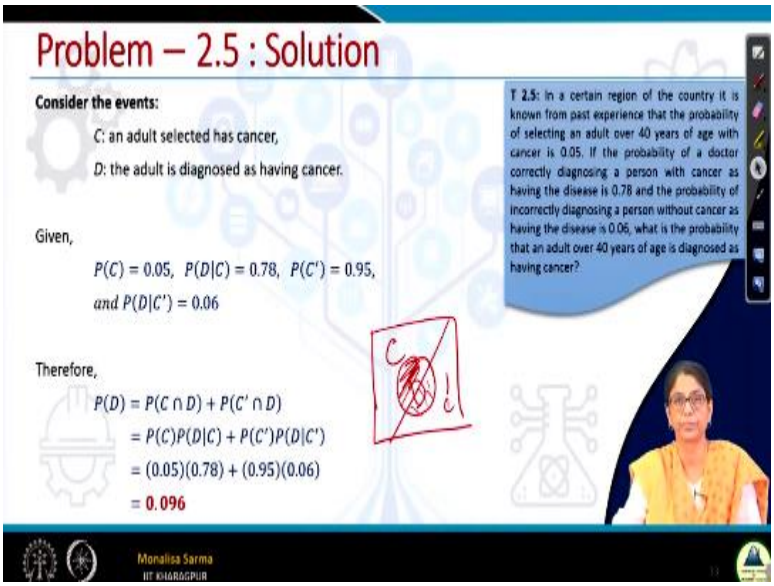
Which is some information are just given so as to make the common problems complicated, like here that adult is whatever age it is immaterial in the questions basically. So, here event C is an adult selected has cancer and the D is the adult is diagnosed as having cancer. Now, what is given say, first thing is given probability of C is 0.05 adult lower 40 years of age, if cancer is 0.05 here it is given. Next it is given a probability of a doctor correctly diagnosis a person with cancer is having the disease is 0.78 this is the thing.

Given that the person has cancer and the doctor has diagnosed that he has cancer and the other one C is 0.05 than probability of C bar is 0.95, 1 – 0.05. And the other one, the next part, you see and the probability that incorrectly the person does not have cancer, but the doctor has diagnosed he has cancer. So, he is not a very good doctor maybe or maybe he was not very much attentive at the time of reviewing the reports. So, this is the thing so we have all this information.

Now, what we have to find out what is the probability that an adult is diagnosed as having cancer, this is basically a question of total probability, but now, if I want to draw the diagram basically, I need to find out this is my D I need to find out what is the probability of this D and what information I have say this is my inverse of sample space, I have how many disjointed now disjoint events yet 2 disjoint events. I have 2 disjoint events here. One is this portion that is the person has cancer and one is this portion the person does not have cancer.

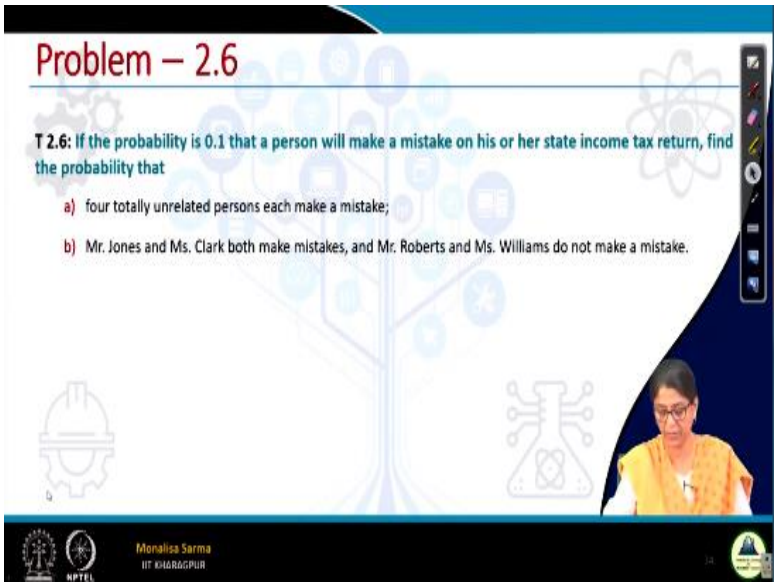
This is the portion where a person does have cancer, maybe this is the person that person does not have cancer and this D is the circle in the D circle all the person are diagnosed with cancer. So, some portion of the people who has cancer are diagnosed with cancer, some portion of this thing who does not have cancer, but still diagnosed as having us cancer. So, now that means, what we will do to find out probability of D is basically intersection of this part and this part intersection of what intersection of this with D intersection of this with D what we have seen yesterday.

(Refer Slide Time: 14:52)



So, what is probability of D? The probability of D is the people that person has cancer C intersection D + C bar intersection D that is the second part. So, this is C, this is C bar and this is D. So, this portion is C intersection D and this portion, this portion is C bar intersection D. So, and we know what is this we have the formula we have the value for all this just put it in a value and we will get a result.

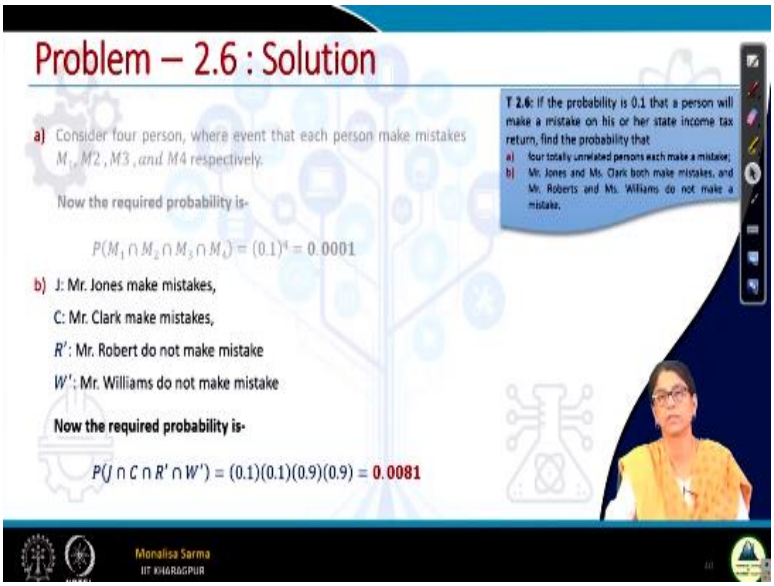
(Refer Slide Time: 15:33)



So now, next question if the probability is 0.1 that a person will make a mistake on his or her state income tax return find the probability of that 4 totally unrelated person each make a mistake. Now, is this a question of independent events, dependent events? Definitely, it is a question of independent events unrelated people. So, there cannot be any dependency if one is making this mistake means that does not mean the other will make any mistake totally unrelated people.

So, for totally unrelated person, it just unrelated people means it is just a multiplication of all the things.

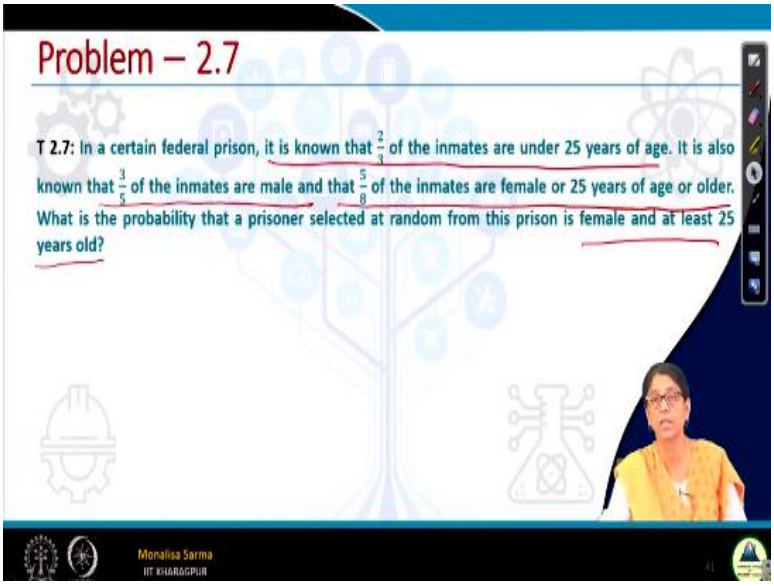
(Refer Slide Time: 16:09)



So, this is that goal then second is Mr. Jones and Miss Clark both makes mistakes and Mr. Roberts and Mr. William do not make a mistake. 2 person make mistake 2 person does not make mistake out of the 4 person may be and there 4 person out of this 4 person 2 person make mistake 2 person does not make mistake what is the probability of the not making mistake, it is just 1 - of probability of making mistake complimentary event.

So, again to compute B again since, all are independent events just a multiplication of all these events. So; now, the required probability this just the multiplication of this 0.1 person making mistake, 0.9 person do not making mistakes.

(Refer Slide Time: 17:05)

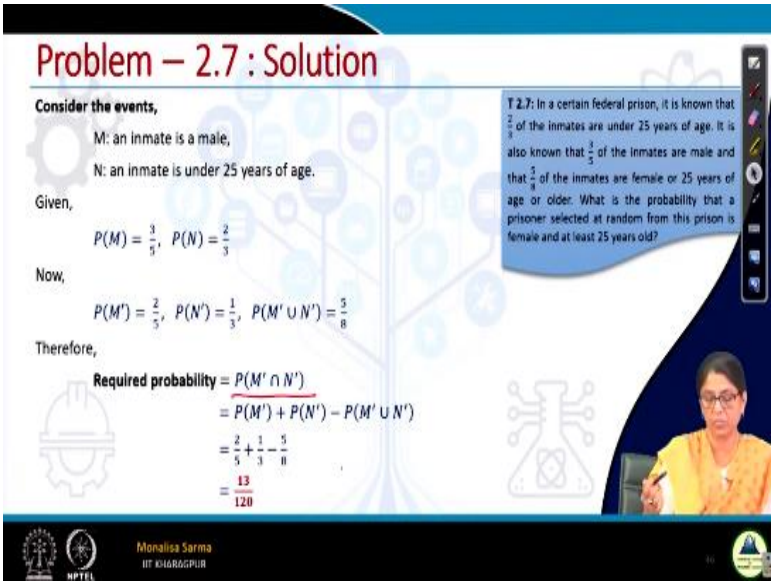


Now, in a federal prison, it is known that 2/3rd of the inmates are under 25 years of age. It is also known that 3/5th of the inmates are male and then 5 / 8 of the inmates are female or 25 years of age or older. What is the probability that a prisoner selected at random from this prison is female and at least 25 years old. Now, here age is not a redundant information here we age is necessary like in the last question for detecting cancer, the person is 40 years old and all that was like unnecessary information, but see here age is an important information.

So, we just cannot ignore it. So first, what is given we will see one by one it is known that 2/3rd of the inmates are under 25 years of age. So, this is one important information 2/3rd of the people are under 25 years of age. So, we will write it even corresponding to this next, it is also known that 3 / 5th of the inmates are male. So, male is 3 / 5 then what is female 1 - 3 / 5 because it is just a complimentary event. And 2 /3rd of the inmates are under 25 years. So, what is the percentage of above 25 years, so it will be 1 minus that right?

That is under 25 years, it is less than 25 not equal to 25, then what is another information is it a 5 by 8 of the inmates are female or 25 years of age or older? Here 5 by 8 are female or 25 years or older. Now, what we need to find out probability that the prisoner selected at random from this prison is female and at least 25 years old. Now write down the events first when you write down the events things you will see it is very easy.

(Refer Slide Time: 19:04)

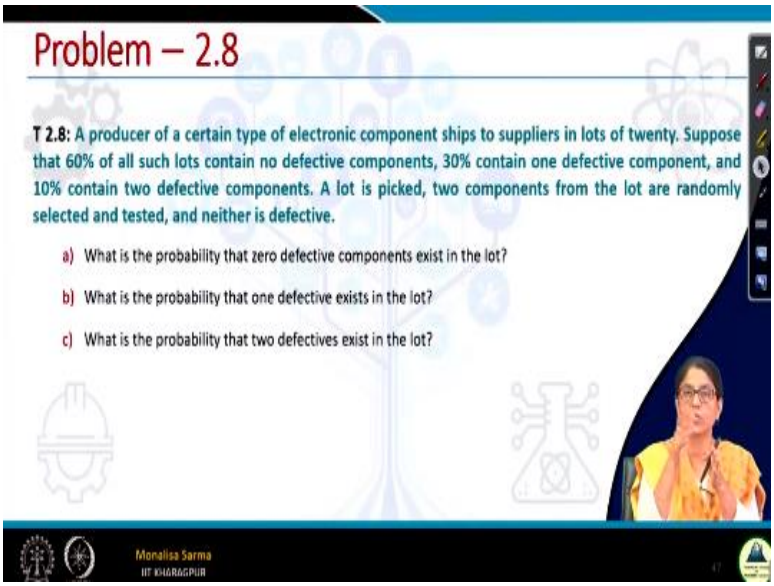


First is we will write that an inmate is a male and inmate is under 25 years of age that is the first information that we had and this was male is 3 / 5 and was 2 / 3, then again we had one more information is 5 / 8, 5 / 8 of the inmates are female or 25 years of age. So, what is that? Female or 25; if male we under writing by M than female we can write the ~M. Male we are the denoting the male event we are denoting by M than female event denoted by ~M.

And is event I am denoting that event inmate is under 25. So, N bar is an inmate which is 25 or above just a complimentary so my probability of M bar union N bar it is given it is 5 / 8. So, now what is this, what is the probability that a prisoner selected at random from this prison is female and at least 25 years, is female and at least 25 years or older. So, that means I have to find intersection. So, what do I need to find?

Probability of M intersection ~N, ~M denotes female and ~N denotes inmate is 25 or above at least 25 years old at least 25 means 25 or above. So, what is this formula corresponds to we have seen it before just put it a value and we will get it so, any probability problem always write events properly, clearly if you write events you will be able to solve it that I guarantee.

(Refer Slide Time: 20:48)



And for total probability always draw the Venn diagram which portion you have to find out, which total probability of what and what are the disjoint events always specify the disjoint events which disjoint in total probability about that different disjoint events cover up the required

probability, but is not it so, now, this required probability is nothing but a summation of all these disjoint events. So, once you do that and whatever value is given you will be able to calculate it.

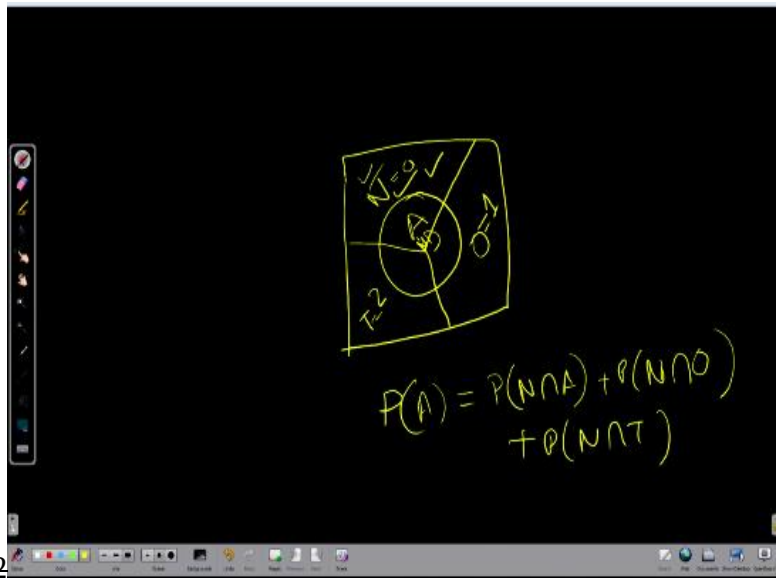
So, problem is if you could find draw the Venn diagram properly then it is done. Next a producer of a certain type of electronic component ships to supplier in lots of 20, suppose that 60% of all lots contain no defective components, 30% contains 1 defective component and 10% contains 2 defective components. 60% of the lots content no defective, 30% contain 1 defective component and 10% contained 2 defective component means out of 20, 60% is no defective again other different lot.

Lot means different categories and again in another lot 30% content 1 defective and another lots 10% contains 2 defective. A lot is picked 2 components from the lot are randomly selected and tested and neither is defective, 2 components from the lot are randomly selected and tested and neither is defective, what is the probability that 0 defective components exist in a lot? 2 components are picked and both are found to be not defective.

Now, what is the probability that it has come from that lot, which has 0 defective. So, this is a question of Bayes theorem if I would have asked what is the probability that getting no defective components. Then it is a question of total probability. Now, I have asked 0 defects we have 2 components from the lot we have selected and found that it is not defective what is the probability that it is come from a particular one disjoint area? So, it is a question of Bayes theorem.

(Refer Slide Time: 23:06)

~~`RD2~~



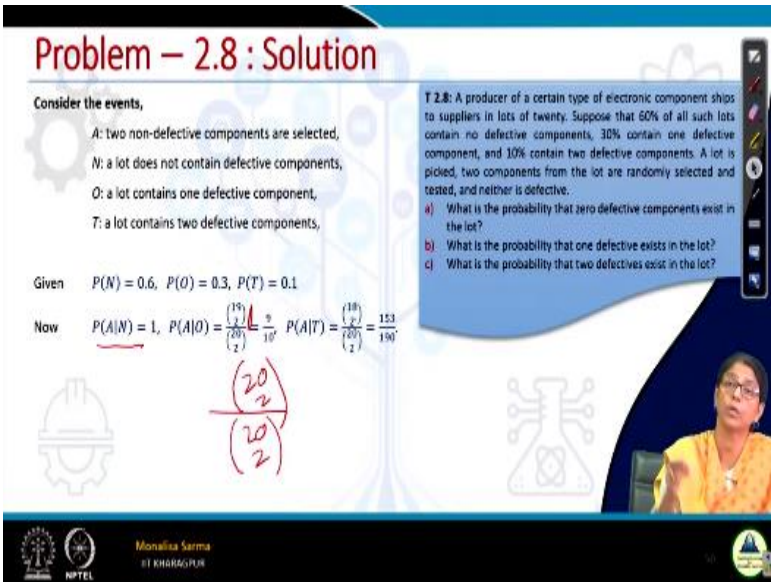
So, now this is the law which no defective component no defective and so, there are 3 different categories here, 3 different disjoint categories here, what are the 3 disjoint component? One with 0 defective say this is one category which is where there is 0 defective another say it is what we have say whatever event I have named I forgot here 0 is it is say with 1 defective another say it is T say it is 2 defective, here it is 2 defective. Now, here what I found? I have picked 2 component from a lot I have picked 2 component I found both the components are not defective.

So, what is the probability that I have picked from this lot or I means I need to find out when I found that both the components are not defective what is the probability that I have picked this lot from this one where there is no defective component. So, first we will for that first we will find out the total probability. Total probability this one, this whole circle, this whole circle maybe let me consider this whole circle as this okay this I am one second.

This I am writing is N then it is easier to differentiate N is no defective, 0 is 1 defective, T 2 T for 2 is 2 defective and this one with 2 defective components with no defective I am writing as A suppose this circle. So, in my what will be my probability of A? Probability of A will be what probability of N intersection A probability of N intersection O and probability of N intersection T this is my probability of A. That is the total probability.

Now, I am interested in finding out that my defect if it is some and I found that it is non defective what is the probability that I have picked it from this no defective component?

(Refer Slide Time: 25:25)



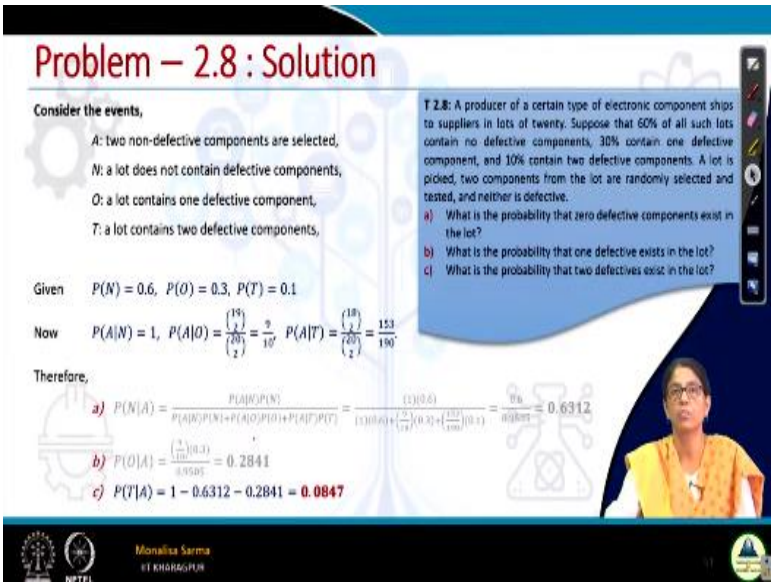
Let me come back to this. So, these are the things which are given probability of N is 0.6, 0 is 0.3 and probability of T that is contains 2 defective component 0.1. So, these are the things given what is see here. Probability of A given N what is this? Probability of A given N means what is this means given that the lot contains no defective component what is the probability that I have picked up 2 defective components.

What is the probability that I have picked up 2 non defective components what is my A? A is to non defective component this how will it find out what is the total sample space? Total sample spaces is from I am picking up from 20, 20 C 2 is my total sample space and out of these 20 when there is no defective components, how many defects it is no defective components. So, all are non defective components. So, that is from 20 again from non defective components I am picking 2. So, this is my probability of A | N is 1.

Probability of A | O; that both the components are not defective when I have picked it from the lot which has 1 defective component. So, if it is 1 defective component you can see here it has given this one 19 from 19 I am picking 2 because 1 defective components that I am removing it, so from 19 if I am picking 2. So, this is my probability of probability of A | O similarly, what is a

probability of A | T that is there 2 defective components. So, I am picking 2 from the rest 18. So, from 18 I am picking 2 that is 18 C 2 divided by 20 C 2.

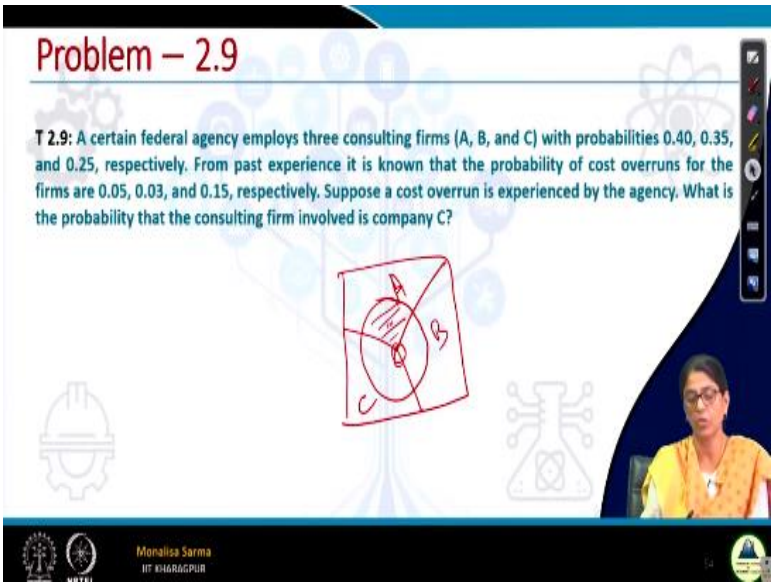
(Refer Slide Time: 27:17)



Now, what I need to find out is probability of N | A given that 2 non defective components are pick what is the probability that I have picked from N that is no defective components, this denominator is the total probability which I have just now I have shown into a whiteboard that a black board which I have shown in the blackboard how we have calculated the total probability this is the total probability, probability of N given A what is this probability A given N.

Probability of A that both the components are not defective given that we have picked from lot which are not defective into probability of N. So, by this the total probability that is probability of A this denominator is nothing but probability of A. So, now putting the formula we will be getting. Similarly, I can next question what is the probability that 1 defective exists in a lot same way, then again what is the probability that 2 defective exist in a lot same way just probability of T given A.

(Refer Slide Time: 28:19)

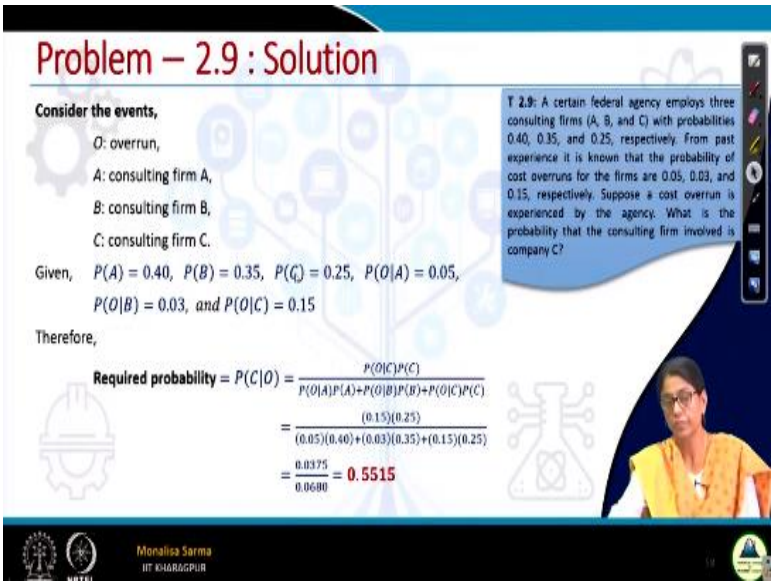


Next this question, if a certain federal agency employs 3 consultants in consulting firms A, B and C with probabilities of 0.40, 0.35 and 0.25 respectively, it employs 3 consulting firms each firm's the probabilities of each firm is given and from past experience it is known that probability of cost overrun for these firms are point 0.25, 0.03 and 0.15 each firm has a cost overrun. So, each firm is used some first one is used 40% of the time B is 35% of the time C is 25% of the time.

However, there is a cost overrun for A is 5%, cost overrun for B is 3%, cost overrun for C is how much 1.5% sorry 15%. Now, then, suppose a cost overrun is experienced by the agency what is the probability that a consulting firm involves in company see again this is a question of Bayes theorem cost overrun is involved. So, this is maybe the cost overrun. So, maybe overrun I can say it is let me take it as O, let me take this event as O.

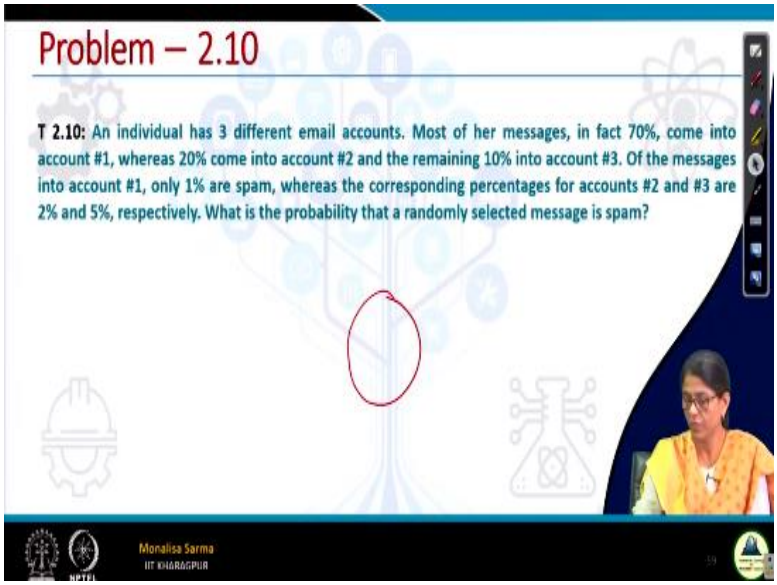
Then, the universal sample space, there are 3 that is A, B and C. A is 0.40, B is 0.35, C is 0.25 and this portion this intersection, intersection of A and O disjoint diameters O intersection of A and O is 0.05, intersection of B and O is 0.03 intersection of C and O is 0.15. Now, we have to we know the things now, what is the probability of consulting a firm involves can you see just same as previous question.

(Refer Slide Time: 30:14)



So, these information are given. So, required probability is this denominator is the total probability first calculate the total probability and what is, it is given what is the probability that a consulting firm involves company C. So, we have calculated probability of O that is the total probability now, what is given O what is the probability that it has cost overrun has happened because of the company C. That is all.

(Refer Slide Time: 30:49)

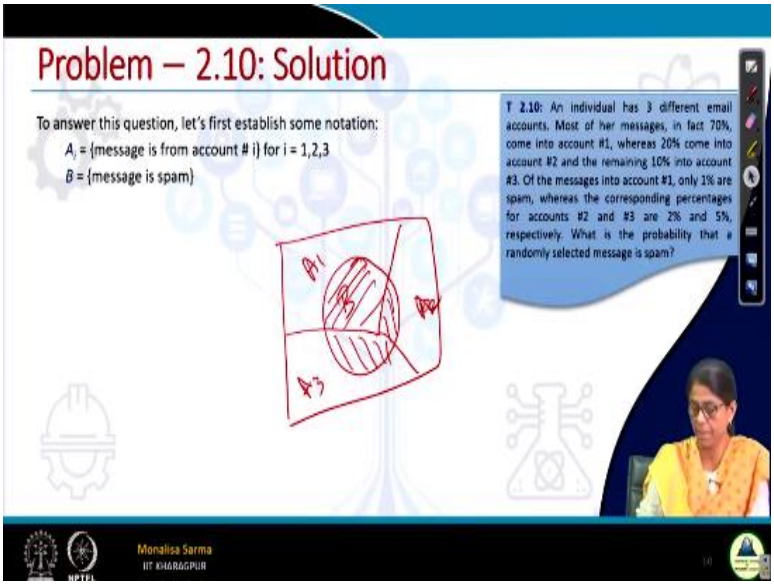


There is one more problem; I think the last problem in this tutorial. So, an individual has 3 different email accounts most of our messages in fact 70% come into account 1 whereas 20% come into account 2 and remaining 10% into account 3 the particular person has 3 accounts, all

the messages into account 1 person only 1 person has spam account 1, one person on the mail as spam mail whereas the corresponding person is for account 2 and 3 are 2 person and 5 person.

What is the probability the randomly selected messages just spam? Again what is the probability the randomly selected message you send spam. So, this is a question on what total probability not Bayes theorem we just want to find out the circle probability of the circle.

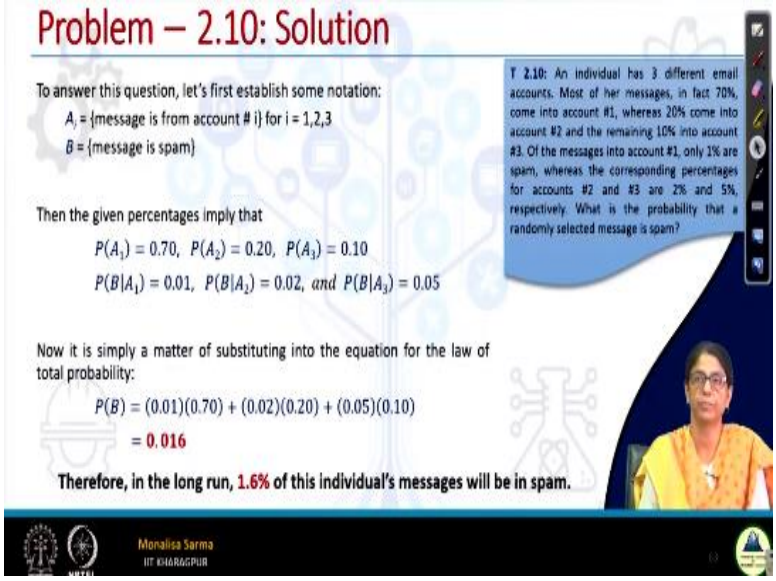
(Refer Slide Time: 31:37)



So, first let me write down the thing what to say event so, I have written as A 1 is messages from account 1, A2 is message from account 2, A 3 is message from account 3. So, this is now, this is my circle is B again this, this is A 1, this is A 2, this is A 3. A 1, A 2, A 3 and what is this? This intersection of A 1, intersection of A 1 and B what is given intersection of A 1 and B is 1 person, intersection of A 2 and B how much it is given A 2 and B is given 2 persons.

A 3 and B is given 5 person this intersection is given, this whole A 1 is how much given, whole A 2 is given how much whole A 3 is given how much to finding the total probability is easy, is not it?

(Refer Slide Time: 32:37)



So, that is it. So, therefore in the long run what we have got 1.6% of this individual message will be in spam. So, what again like In last tutorial I have suggested this again I am suggesting you solve as many problems as you can try to find out any book on probability you will just Google will be getting many free online books. Just go and browse those books you will get many problems, solve those problems in any books the answers are also given.

Do not directly see the answers for solving than try to see where it ends such a character not the solving this, it will be more which will give you more confident with that I end this tutorial. So, these are the references and thank you guys.

(Refer Slide Time: 33:28)

Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology - Kharagpur

Lecture - 07
Introduction to Probability Distributions

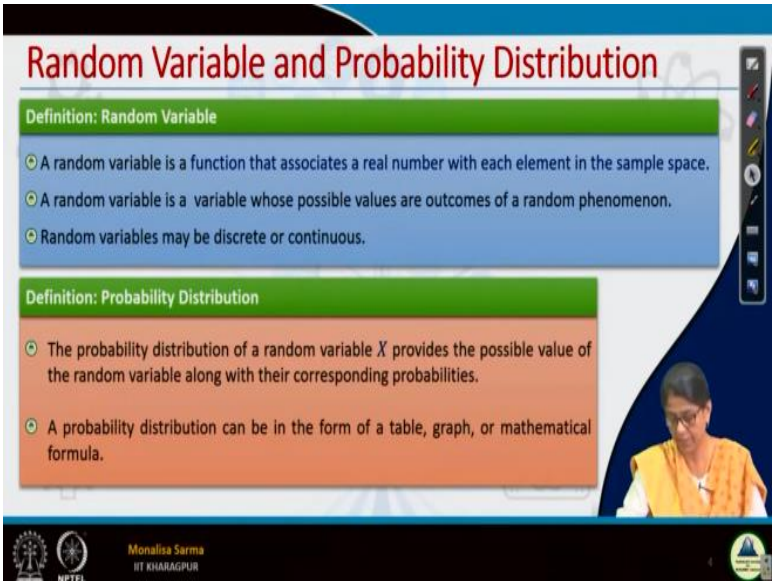
Hello once again so now, today we will be discussing probability distribution we have discussed probability but in probability theory we have in probability theory, we have seen different types of events then we have seen conditional probability total probability and Bayes theorem and now going a bit further. Now we will be seeing what is probability distributions.

(Refer Slide Time: 00:49)

So, for to in this probability distribution first I will be explaining what is a random variable then I will go on to explain what is a probability distribution? We will discuss discrete probability distribution what is a continuous probability distribution what is a discrete probability distribution we will also see what are the mean? We have seen what is the mean and variance given a set of numbers.

Now, we will see mean and variance with respect to a random variable mean and variance of a probability distribution essentially, then we will discuss what is a Chebyshev's theorem with that we will end this class.

(Refer Slide Time: 01:21)



Now, first is what is a random variable? So, random variable if you see the definition random variable is a function that associates a real number with each element in the sample space. Suppose in samples suppose of course, I have toss 2 coins, and when I toss 2 coins we are doing it let me take the pen, when I toss 2 coin, what is my sample space my sample space will be say, my sample space will be say head head, head tail.

Then again tail head, tail, tail this is my sample space. So what happened if I am interested in finding out from number of heads when I toss 2 coins so what the number of heads I get this is number of heads is 2 number of head is 1 here number of head is 1 here number of head is 0. So this is the random variable essentially. So it is a function that associated real number with each element, in the sample space.

A random variable is a variable whose possible values are outcomes of random phenomena. That is also we can define the variable so random variable now, this random variable may be discrete maybe continuous. So, now, with the knowledge of random variable, now, we can define what is a probability distribution when we do an X when any random experiment we will be getting different sort of outcomes. What are these outcomes?

These outcomes are the values of the random variable. Now, this random variable does values they have a certain probability of occurrence like, if you toss a coin, what is the 2 outcomes head and tail we have a probability of occurrence of head and tail if I let me take it in this way. What is the probability of getting head? So probability of getting either we will get 1

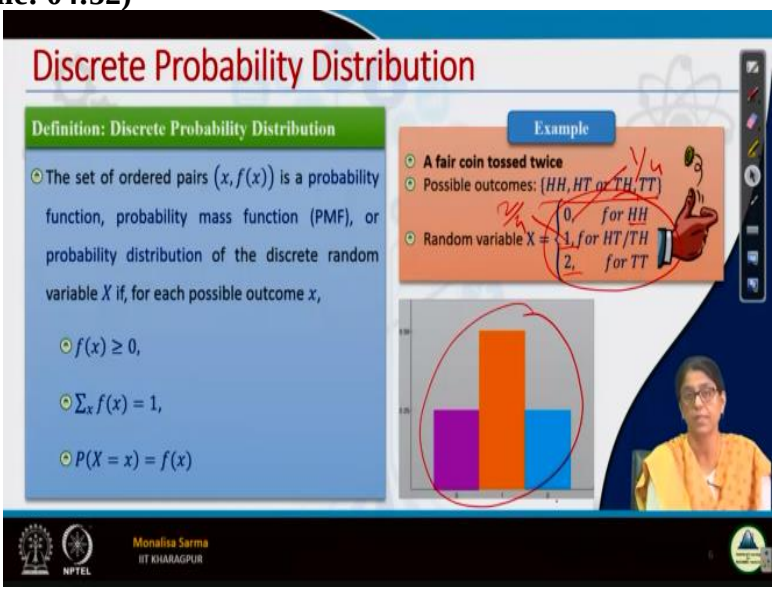
head or we will get 0 head if it is a tail that is a 0 head and what is the probability of getting 1 head is 0.5 and probably of getting 0 head is 0.5.

So, this that means the values of the random variable when we do an experiment the values that we get the random variable and the probability of occurring of those values these 2 things put together is probably distribution. So, what is probability distribution? Again I am repeating probability distribution is nothing but the values of the random variable as well as the probabilities of occurrence of these random values all the random variable in an experiment. So, how we define it?

The probability distribution of the random variable X provides the possible values of the random variable all possible values of the random variable along with their corresponding probabilities that is the probability distribution. So, we will be having since random variable may be discrete random variable may be continuous, discrete means it will take some infinite values 1, 2, 3, 4 some finite values may not be one to other but it will take some finite value and what it is continuous?

Continuous means it can take any value within an interval. So, if the random variable is discrete then we call it discrete probability distribution if the random variable is continuous, we call it a continuous probability distribution. So now a probability distribution we can describe a probability distribution in the form of a table in a form of a graph or in the form of a mathematical formula, we will see.

(Refer Slide Time: 04:52)



Now, first we will define we will see the condition of a probability distribution function. So what is the probability distribution? First when I tell that is random, suppose my random variable is X, then what is the probability of X? That I will define it as an f(x), what is the probability that X stays that I can let me define it as f(x) or you can tell P(X) anywhere and it is just a function of X basically the probability of X, f(x) is nothing but the quality of those the set of ordered pairs x f(x).

The set of ordered pairs f(x), f(x) is a probability function or we can call it a probability mass function or probability distribution we can call it a probability distribution function or probability mass function of the discrete random variable x. Now, here one thing please note it see I have mentioned here of the random variable x and here I have used capital X in here I have used small x, this is a general convention when we talk about the random variable, we always use capital letters.

And when we talk the value that the random variable takes we usually denote it by small letters, now, this x is the value that this capital X takes my random variable is capital X, and x is the value that a random variable takes. Now, when I am tossing a coin tossing 2 coins, I was getting 2 head 1 head, 0 head. So, this value is 2, 1, 0 that I will denote it with a small letter, but the random variable is denote it with a capital letter it is a general convention followed in all the books.

So, if for each possible outcome x so for a discrete x f(x)to a when we can say it is a probability mass function or a probability distribution function it satisfies this following conditions what is this first of f of x should be greater or equal to 0, what is f of x? f of x is nothing but f of x is the probability of x so it should be greater or equal to 0 in fact, it should be less or equal to 1 as well.

And what is summation of x f(x)? Or sorry sorry summation of f(x), x summation of f(x)for all values of x will be always 1 probabilities will be always 1 is not it so the probability of getting 2 head probability of getting 1 head probability of getting 0 head all probability put together summation should be always 1. And probability of x = x is nothing but we designated by f(x) so f(x)we can tell it is a probability density function.

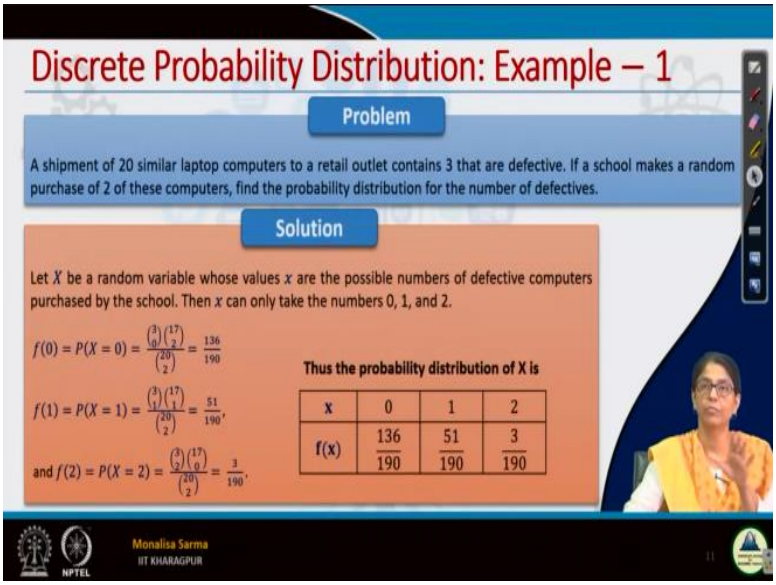
sorry probability distribution function when it satisfies these 2 condition basically f(x)will be greater or equal to 0 and f(x)summation of f(x)would be 1. So here is an example so a coin I have tossed a coin twice the example what just now that I have given, so possible, these are the possible outcomes. So HH, HT, TH and TT. So even if I am considering is the getting heads.

So, what is my getting 0 sorry getting maybe getting T so, getting 0 tail is when I get HH that means I got 0 tail both I got head then getting 1 tail is either HT, TH getting 2 is TT. Now, what is the probability of this says how do we find a probability? So, sample space size is 4. So, I got 0 head how much I got? What is the just only how much is favourable to this event of getting 0 head that is 1. So, what will be my probability? 1 / 4.

So, probability for this will be my 1 1 / 4, now, for 1 getting 1 head, my favourable outcomes are 2 HT and TH so, what it will know what will be my probability I am writing here probability will be 2 1 by 4 and getting 2 head sorry getting 2 tail is just 1 again it is 1 1 by 4. So, this if I can do when in the last slide I told this probability distribution I can define it in the form of a table in the form of a graph in the form of a function.

So, this is if I express this maybe in the form of a function this is what it is shown here. And it is this one this maybe it is in the form of a function, this is in the form of a graph. See and y axis I have given a probability x axis is the different values that random variable take random variables may take the 0 value random variable is with respect to an event always. So that what is my event? Getting a tail. So, the random variable maybe 0, 1 and 2 and this side it is the y axis is the probability and this is if I just explain it with the help of a graph.

(Refer Slide Time: 09:54)



A small example now to understand this concept, what it is given? First go to revision shipment of 20 similar laptop computers to a retail outlet. A shipment of 20 similar laptop computers to a retail outlet contains 3 that are defective out of 23, 3 defective if a school makes a random purchase of 2 of these computers find a probability distribution for the number of defectives.

Say we here it is asked that we have to find the probability distribution for the number of defective if the school has purchased 2 computers, what may be the number of defects in this case? That is purchased 2 computers. So, what may be the number of defectives? The number of defectives may be there may be 0 defective there may be 1 defective there may be 2 defective atmost 2 defective.

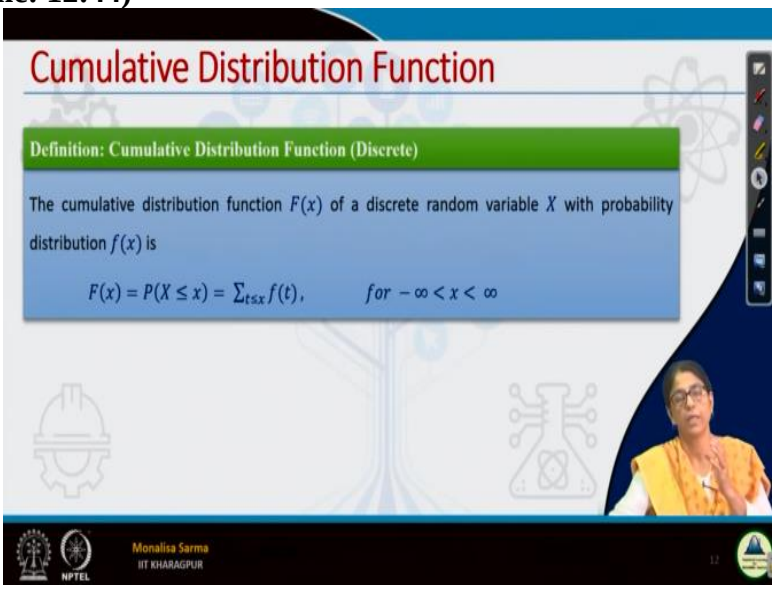
So, number of defective will be 0, 1, 2, that means, the random variable will take the value 0, 1 and 2 now we have to find out what is the probability of 0? What is the probability of 1? What is the probability of 2? This is the probability distribution is not it? So, that is what probability distribution is the value of the random variable all the possible values of the random variable along with this probability of occurrence.

So, how do we find out what is the probability that 0 defective? So, what is the total sample space? Total sample space is the total 20 computers out of 20 I am picking 2. So, it will be 20 C 2 we have seen remember what is where we will use permutation where we will use combination now, here 2, these are indistinguishable is not it? So, it will be 20 C 2 the whole sample space is 20 C 2.

So, and how much what we want to see? 0 defective that means, in 20 there are 3 defectives, how many non defectives are there? Non defects are 17 and we want to find out what is the probability of picking 0 defective that means, the 2 computers that we have purchased that means, we have we got it from the good ones good 17, 17s are the good ones. So, 3 C 0 x 17 C 2 see this one.

Similarly, for 1 same denominator is 20 C 2 and that means, we have 1 defective means from that 3 defective we got 1 and 2 another one we got it from the 17 similarly, 2 from 3 we got 2 and from the 17 we got 0. So, we write in tabular form this is the tabular form in the tabular form the probably distribution of x so, we have seen how we write in a tabular form we have seen how we write it in expression form how we write it in a figure graph.

(Refer Slide Time: 12:44)



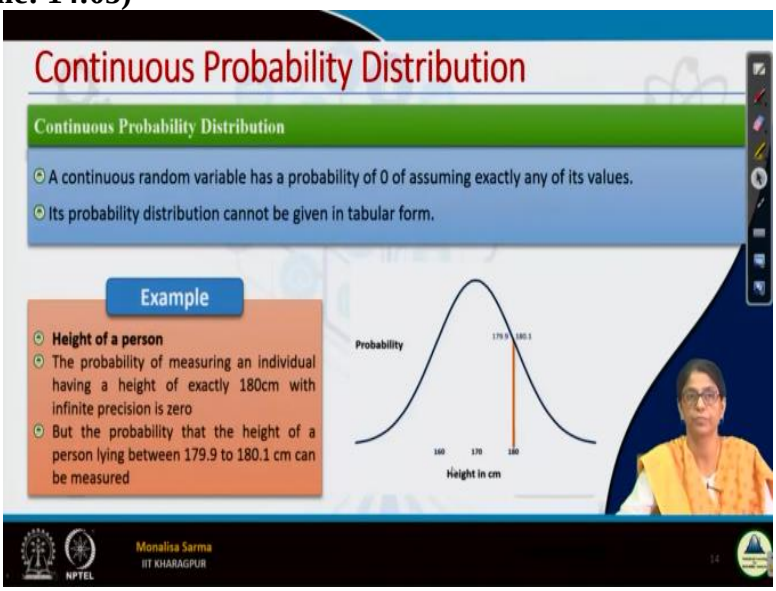
So, now, what is cumulative distribution function? What we have seen here more probability of getting 0 defective, 1 defective, 2 defective we are concerned of a particular number. Now, if I am asking this question of head and tail only 1 question of getting when I am tossing 2 coins what is the probability of getting either 0 tail, 1 tail or 2 tail maximum 2 tail I can get. So, I cannot obviously get 3 tails now, if my question is what is the probability of getting at most 1 tail?

That means, at most 1 tail means there may be 0 tail as a 0 and 1. So, I have to find the probability of occurrence 1, 0 probability of occurrence of 0 and probability of occurrence of 1. So, probability of occurrence of 0 plus probability of occurrence of 1 that will give me at

most 1. So, what is that? That is the cumulative distribution function we have seen cumulative frequency distribution remember, so, does the same portion same concept.

So, cumulative distribution function is F of x this probability that x that is the random variable is less than or equal to x is this summation of f(t) where t is range less than equals to x. So, here at most 1 means f (0) + f (1) that is cumulative.

(Refer Slide Time: 14:05)



Now, continuous probability distribution, continuous probability distribution when the random variable is continuous and a random variable contains any value in the range than the distribution that we get is a continuous probability distribution that will not be a discrete probability distribution. Now, taking any value in the range and for the when we tell that so, it can be any infinite value any value within the range it can.

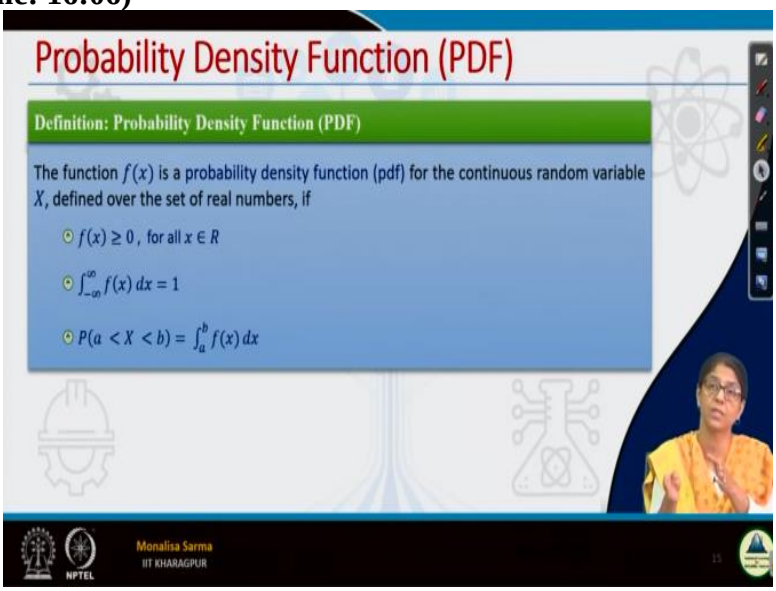
So, now like suppose if I asked a height of a person, you can tell this height of a person is a 5.8, but will it be exactly equal to 5.8 or it is equals to 5.800001 or 5.812345 something like that, you it is very difficult to tell exactly exact value it is almost the probability of getting an exact value is almost equal to 0, you cannot it is such a minute scale measurement is not there to measure.

So that is why what we do in case of continuous probability distribution, what we tell is that a continuous random variable has a probability of 0 assume exactly any of its value. And when we talk about continuous probability distribution, we have a probability of 0 assuming exactly of any of its value like in the discrete we are telling what is the probability of 0, 0

head, 1 head, 2 head exactly, but when you are talking about continuous probability distribution.

I have probability of taking exactly a particular value we cannot define so, that is basically 0 and since that is the situation so we cannot write a probability distribution table we cannot represent it in the form of a tabular form, in tabular form we have to give specific value. So, in continuous probability distribution, we can only represent it either using a function or using a graph. So, this example whatever example I told you this now.

(Refer Slide Time: 16:06)



So, now, like probability distribution function what we have seen our probability mass function for discrete random variables similarly, for continuous random variable we call it a probability density function. So, for probability density function the condition it has to satisfy saying like the that is the f of x is greater or equal to 0 for x belongs to any real number and integration of f(x) dx is 1 here it is integration.

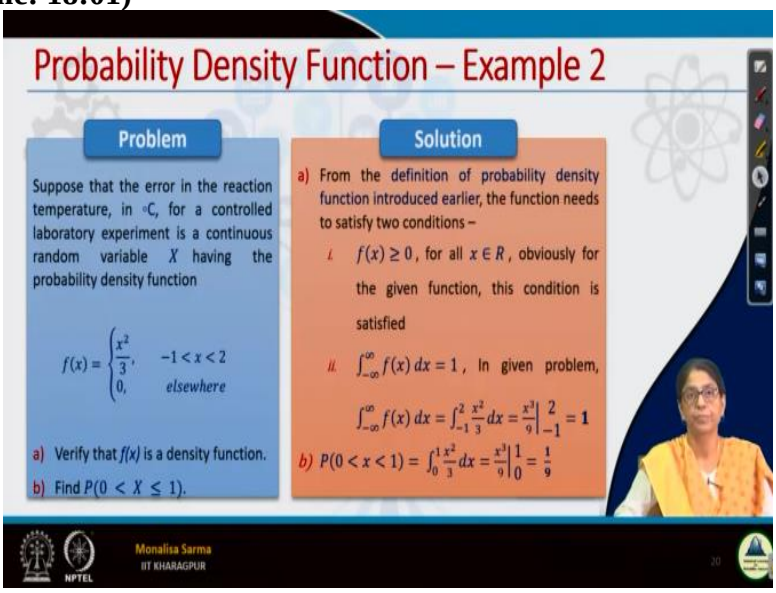
Now, here one thing which is very important to remember to note that as in when I am talking of discrete probability distribution, then value of f(x)cannot be greater than 1 because in case of discrete probability distribution, what is f(x)? f(x)is the probability of the occurrence of x , f(x)is nothing but the f(x)is the probability of x so, probability of x cannot be greater than 1. Any probability means that it has to be within 0 and 1.

So, in case of discrete f(x)is the probability of x, but in case of continuous probability distribution f(x)is not the probability of x it is the probability density function its value can be

variable it value can be variable greater than 1 as well. So, it is a density function it just in a curve when we draw the curve, it shows how much probability is concentrated at that point near that point.

Basically, f(x) basically will give us the shape of the curve and when we have to find out the area at a particular interval, it will always be the area in a particular interval, so, we will have to find the area so, that is why it is integration. So, remember for this probability density function my effects can be greater than 1. So do not get confused why is if this is probability, why it is greater than 1, it is not probability it is a probability density how much it is concentrated at that point, how much probabilities are concentrated at that point.

(Refer Slide Time: 18:01)



So, now, this is a small example, suppose that error in the reaction temperature is degrees centigrade for a control laboratory experiment is a continuous random variable X having the probability density function, probability density function is given. So, it is given f(x) = x square by 3 for a particular range what is the range? x varies from - 1 to 2 and in other places it is 0. So, we have to verify that f(x) is a density function.

See, when x takes a value very near to 2, then we will get f(x)greater than 1 you see, that is what I have told in the last slide, that is we have already discussed now, we have to verify that f(x) is a density function, how will you verify that f(x) is a density function? To verify that f(x) is a density function it has to satisfy those 2 conditions what we have seen what are the 2 condition?

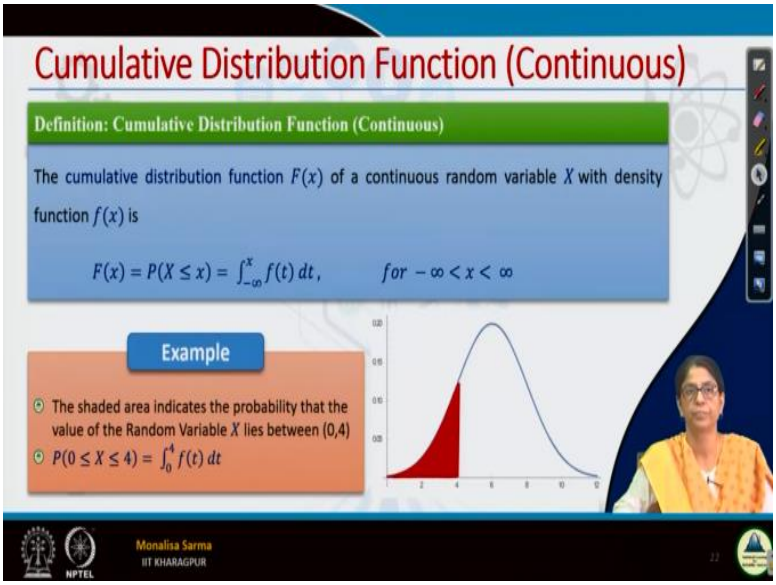
First is f(x) should be greater or equal to 0 and integration of f(x) what the range should be equals to 1. So, returning from this expression, what is given it is obvious that it is greater or equal to 0 from - 1 to 2 it is x^2 1 / 3 square definitely it is not less than 0 it cannot be less than 0 and elsewhere it is 0. So, it is greater than or equal to 0 first question condition is already satisfied second condition from infinity to - ∞ to ∞ it should be 1 that is the second condition.

And if we just do the integration, then it will just see the integration it is given here, then we will get 1 so, it satisfies both condition then we can say f(x) is a density function. Next question, we have to find out the probability between x of 0 to 1 nothing just f(x) integration from 0 to 1. So, what is f(x) ? f(x) is x^2 1 / 3 we have to just do the integration from 0 to 1. So, see, gradually we will be getting many such problems with your knowledge integration and differentiation will be necessary.

So, please you have already learned it in your school level and high secondary level may or may not be in school level of high secondary will definitely please brush up your knowledge on differentiation and integration for many problems like here we have done this a simple integration of course, all of you will remember that, but if any of your forgotten please brush up that so else it will be difficult for you to understand the problem.

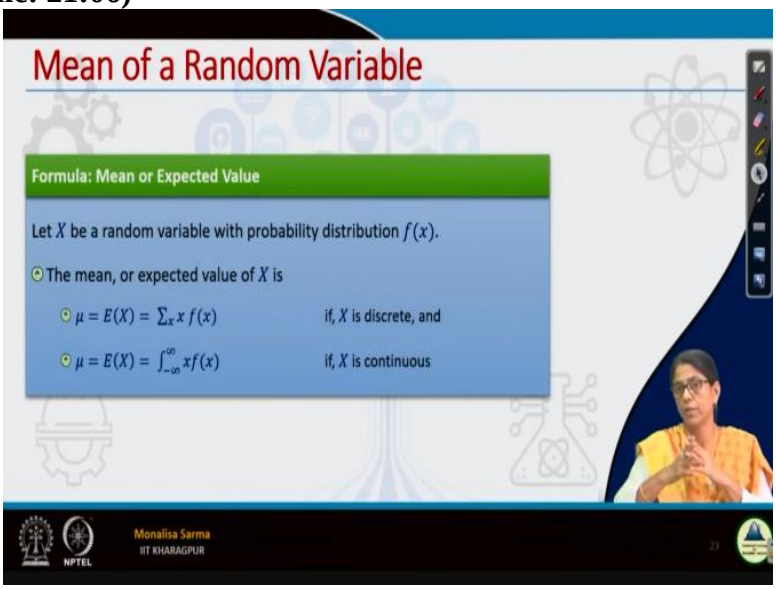
And as I have promised in my first introductory lecture that I will be taking lots of examples see for all each concept I will explain and accordingly I will give 1 example.

(Refer Slide Time: 20:38)



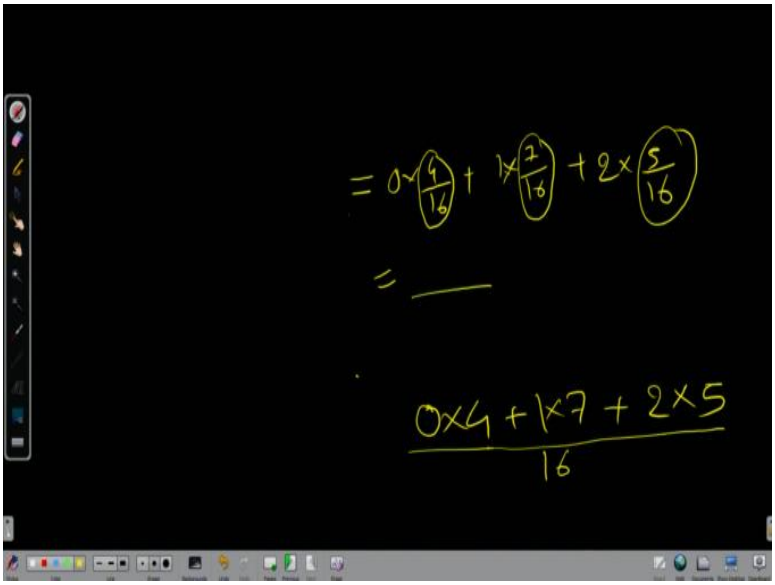
Now, what we have discussed for cumulative distribution function for discrete case same is the case cumulative distribution function for continuous case cumulative means from till x whatever x we specify from minus infinity to till x that is the cumulative distribution function same the concept of discrete.

(Refer Slide Time: 21:06)



Now, how do we find the mean of a random variable we have seen how to find out the mean of a set of numbers. Now, instead of taking let me give you another example suppose, I have 2 coins and suppose I have toss the let me take the graph paper I have toss 2 coins.

(Refer Slide Time: 21:32)



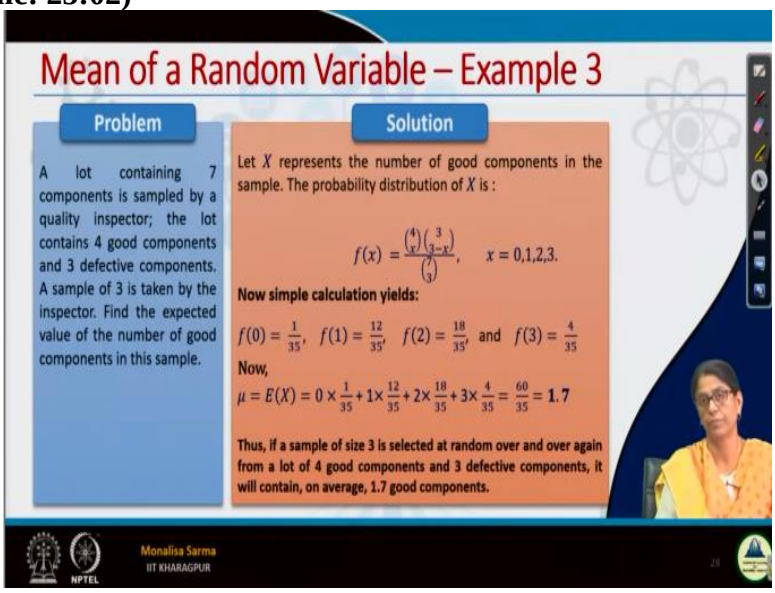
Suppose I have toss 2 coins 16 times and 4 times when they say 4 times I got no head only tail, tail, tail, tail, tail, tail. And suppose 7 times I got 1 head and suppose 16 so it is it will be 5, 5 times I got 2 head. So now what is the average number of heads per toss mean is average, is not it? So what is the average number of head per toss? I am tossing 2 coins, so what is the average number of heads per toss? Average number of head how will you find out?

Per toss is so 4, is not it? So this 4 I got 0 head 0 x 4 + 1 x 7 + 2 x 5 divided by how many? 16. This, of course, let me rub it, this thing I am writing in top this can I not write it, 0 x 4 1 / 16 + 1 x 7 1 / 16 + 2 x 5 1 / 16. And out of 16 I got 4 times 0 out of 16 I got 7 times 0, 7 times 1 head out of 16 I got 5 times 2 head what is this? This is relative frequency distribution relative frequency distribution is nothing but probability we have discussed it remember in our second lecture, so this same expression, I can write it this way. So what if I calculate it, I will be getting something whatever you find, calculate it and find it out. So this is the average of that so that is how we find a mean. Similarly, so, what how we found out the mean of what that means, the problem what are the different values that a random variable takes as well as the probability of that value 0 into what was that 0 x 4 / 16, 4 times we got 0 head.

So 0 x 4 / 16, 0 is a value of the random variable and 4 x 1 / 16 is a probability that is the relative frequency that I got 0 here. So what is 4 x 1 / 16 is nothing but a probability. So, similarly, so when I am interested in the mean or a mean is also called the expected value, it is also termed as E mean of a random variable it is designated as E is described as μ sometimes we call it as expected value, it has written it as mean it is also written as E.

So what is μ or E of X? It is nothing but x summation of x f(x) what we have just now we have seen, but now you have seen remember this one, this is 0, 1, 2 these are what? x And what is this? This is f(x)summation of x f(x). This is if x is discrete summation of x f(x). If x is continuous, then just integration of x f(x) is the same thing. So, that is how we find the mean of an expected value.

(Refer Slide Time: 25:02)



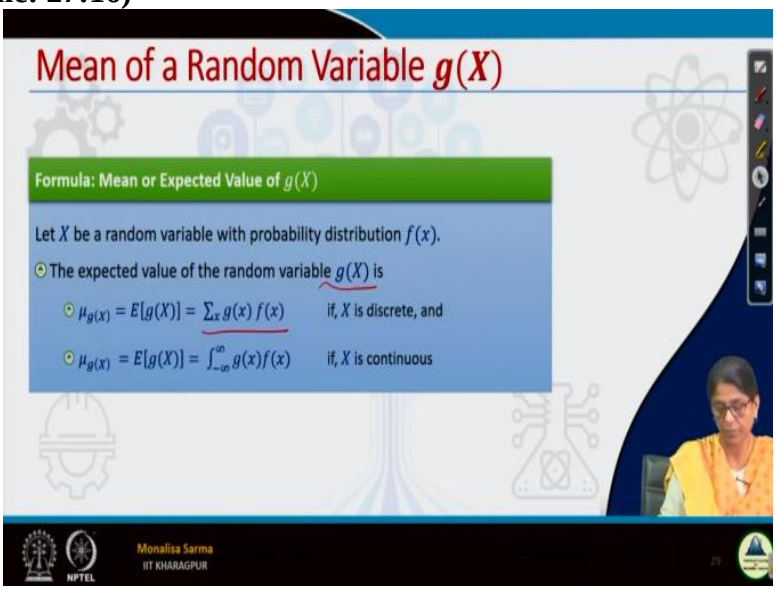
So it is a simple example is there a lot containing sample 7 components is sampled by a quality inspector and the lot contains 4 good components and 3 defective components they have total 7 components out of them 4 is good and 3 is defective a sample of 3 is taken by the inspector find the expected value of the number of good component in the sample see, X represents the number of good components assembled.

So, we will have to find out the probability distribution of X how many good components or what the value what value that a random variable X can take then a variable where we are interested in number of good components. So, how many good components are there? There are 4 good components, but we have taken 3 so, the random variable can take either 0 good component 0, 1, 2, 3 is not it?

So, basically, x can take the values you do 1, 2, 3 see here x can take. So, what is f of x? f of x is the 7 is this one, 7 C 3 is the total sample space and out of 7 how many good components? Good component is 4 we are interested in finding a good component. So, we are taking all the x components from 4 and remaining how many are remaining? 7 - 4 is 3. And from this 3 how many were picking total we are taking 3 components out of 3.

This x is the number of good components and from this 3 we are picking the other component that is 3 - x. So this is the function called the distribution function. So we can calculate this so we found out for f0, f1, f2, f3, we can calculate it. So now we need to find out the expected value expected value is summation of x f(x). So, this X is 0 into 1 1 / 35 + 1 x 12 1 / 35, 2 x 18 1 / 35 + 3 x 4 1 / 35, just put it next question and get developed you can check it again. So, on an average, it will contain 1.7 good components.

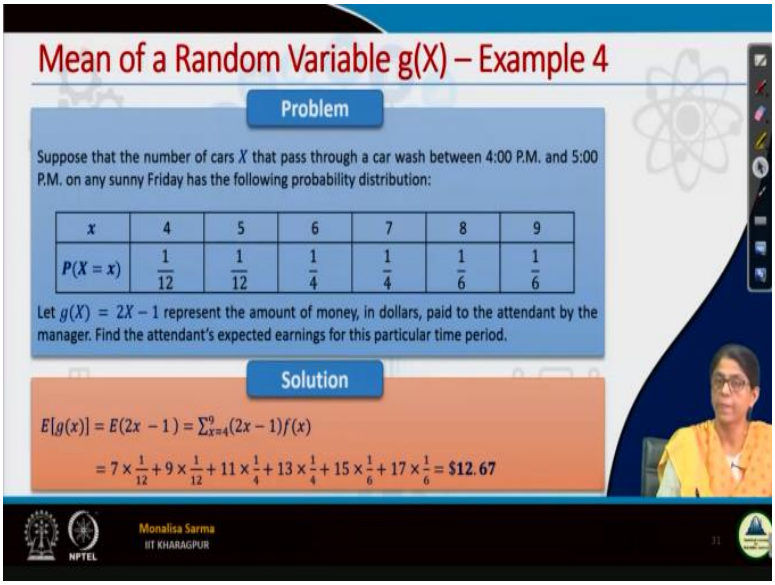
(Refer Slide Time: 27:16)



Now, mean our expected value of g(x) if we are interested in finding the expected value of a function of x, we have seen how to find out when X is a random variable, we have seen how to find the expected value of expected value of x is nothing but summation of x f(x)if it is discrete. Now, we are interested in finding out the expected value of a function of x, say x, maybe 3x x maybe x^2, x^3, or 3x + 2 anything.

So if you are interested in finding out the expected value of a function of x with a function of x, I am terming it as a g(x) here. So it is same, there is just summation of g(x) instead of x. I am writing g(x) that is, that is the difference.

(Refer Slide Time: 28:02)

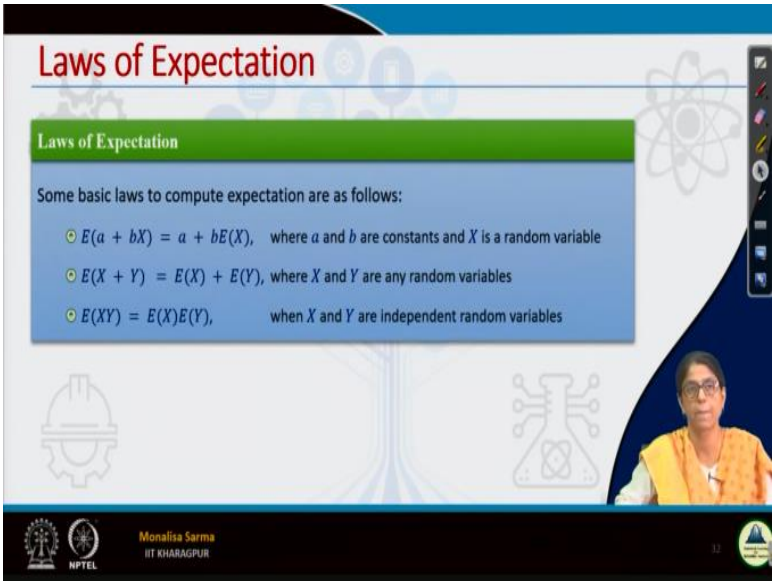


So, using this, this one simple example is there, it is very simple, you will be able to do it yourself only, just the weekly I will go through it, suppose that the n number of cars just pass through a car between 4pm and 5pm on a sunny Friday as the following probability distribution value added a random variable can take is 4, 5, 6, 7, 8, 9 and this is the second row is the probability of this particular thing.

So, now, we have to find out the attendant's expected earning g(x) = twice x - 1 represent the amount of money in dollars paid to the attendant by the manager. So, this represent amount of money in dollars paid to the attendant by the manager. So, we have to find out the attendant's expected earning. So that means the average earning average earning how will I find so this is g(x) f(x), so g(x) 2x - 1 is g(x). So, what x will take what value?

x will take 4 first actual day 5x will take 6, 7, 8 and 9 and this into and into f(x) whenever x is taking value for we will multiply it by 1 / 12 and by x is taking value 5 we will take multiply it by 1 / 12, x is taking 6 will multiply by 1 / 4 likewise see, just we have substitute different values for X and accordingly whichever values we have taken for x, that probability will be considering that is all.

(Refer Slide Time: 29:35)



So now there are simple laws of expectation some basic laws we have. So what is the expected value of a + bX? Always remember the expected value of a constant is the value itself the constant itself. So here expected value of a + bX expected value of a is a itself. Expected value of 2 is 2, expected value of 10 is 10. So, there is no error we cannot find an expected value that is average value of a constant number.

So, here it is E of a + bX so, obviously, it will be a + b into E of X. So clearly expected value of X + Y if 2 random variable X +Y what is the very interested in finding out the expected value so, it will be expected value of X + expected value of Y, expected value of E of XY = E of X x E of Y. However, this is applicable when X and Y are independent random variable. So, now I think I will stop this lecture here.

And though initially I have to told that I will be covering variance also and Chebyshev's theorem also in this lecture but time may not permit. So, I am stopping this lecture here, in the next lecture I will start from the variance of the random variable. Thank you.

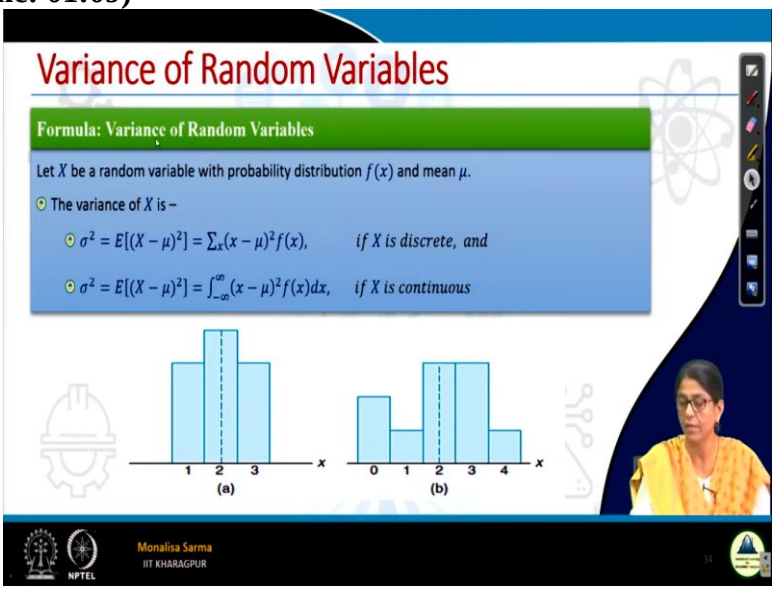
Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology, Kharagpur

Lecture - 08
Introduction to Probability Distributions (Contd.,)

Hello everyone warm welcome. So, now, in today's class we will continue from the last class like we were discussing probably distribution so, and then we have discussed the mean of an expected variable not we have discussed the mean of a random variable that we also call it as the expected value. So, and then today, we will start after that and that is we will proceed from there.

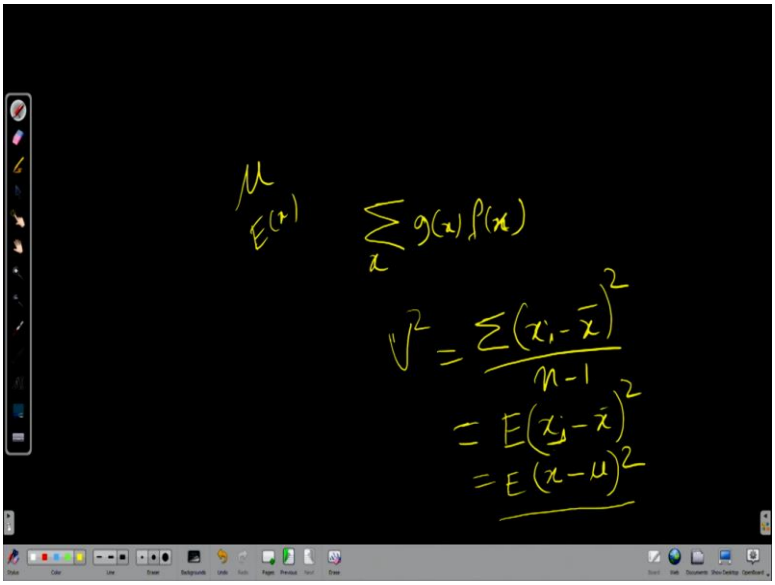
That is we will discuss now variants of random variables, how to calculate the variance of random variable. We have also seen while calculating the mean of a random variable, we have also seen that how to calculate the mean of g(x), what is g(x)?

(Refer Slide Time: 01:09)



Remember last class we have discussed that what is g x? g(x) is a function of x. So, how we calculate the mean of g x, mean of g(x) was same instead of x we have just written it g(x). So, let me just take the board and then explain it.

(Refer Slide Time: 01:24)



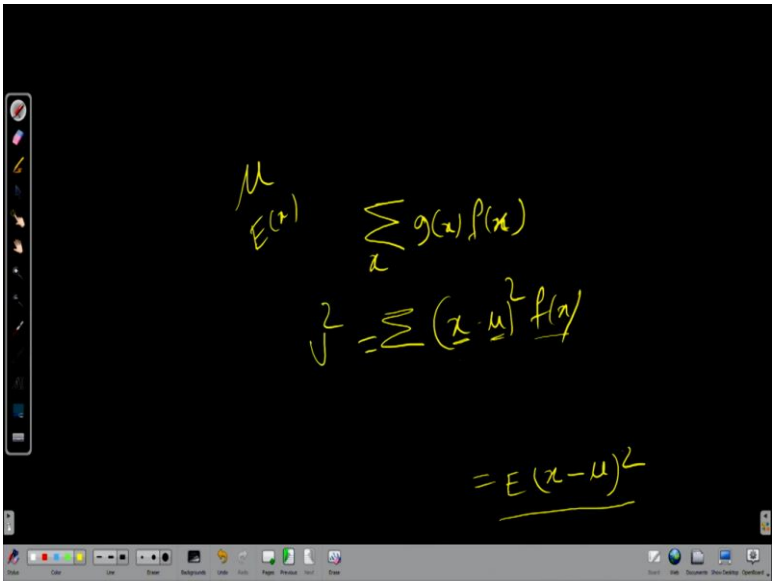
So, how we have calculated the mean of g(x)? Mean of g(x) is summation of g(x) f(x) and summation over all the values of x. So, what is g(x) is a function of x. So, now, when we are interested in finding the variance, variance of a random variable, what is variance basically? Basically when we calculate variance how do we calculate variance we calculate variance but it is v square we calculate variance as summation of $(x_i$ - ~x)^2/ n – 1.

This is the formula for variance now, this cannot I write it this way this is nothing but expect this is mean is not it when I am debating with n - 1 I am just finding the mean of all the values, all the values what I have get I just divided because total this much value that I will get and I divided it by that so, this I can just write it as mean of $(x_i$ - ~x)^2. This now, what is x_i? x_i means all the different variables that are random variability.

So that means, I can write it instead of x_i I can write it as x that has the different values that are random variability. And instead of x bar, so, we have already discussed when we talk of mean over random variable, either we use the term μ or we use the term μ E of x. So, let me call it as use mu. So, $(x_i - \mu)^2$ so, that means, when I am interested in finding the variance of a random variable.

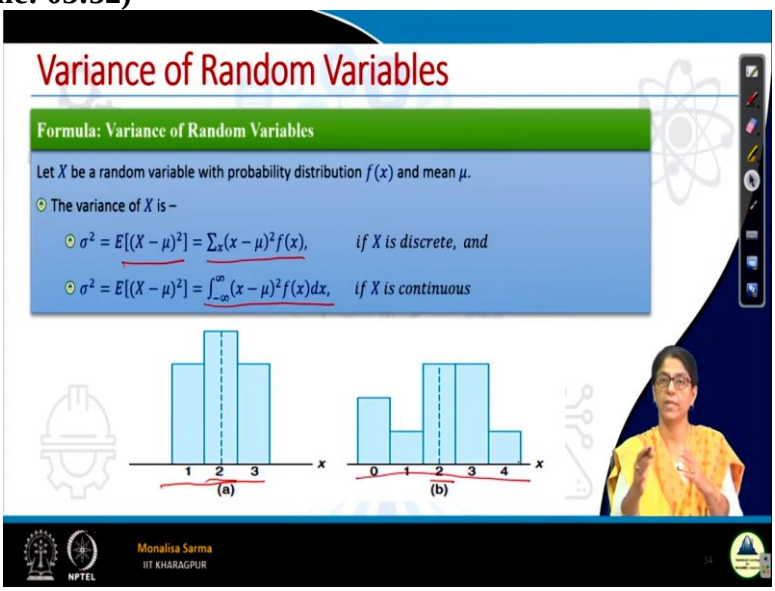
So that essentially I need to find expected value of $(x_i - \mu)^2$ now, what is x - μ also it is nothing but a function of x. So, what is how do I find the expected value of x - μ whole square just putting it in this formula.

(Refer Slide Time: 03:17)



I will just put it in this formula. So, if I put it in this formula what I get what is my g(x)? g(x) is $(x_i - \mu)^2$ f(x) so, what information I have? I have all the values of f(x) what the values are random variability I have all the values of f(x) that is probability of f(x) then of course, I can find out mu from the formula what we have learned and then with this I can find out my variance now my variance is this so, that is what.

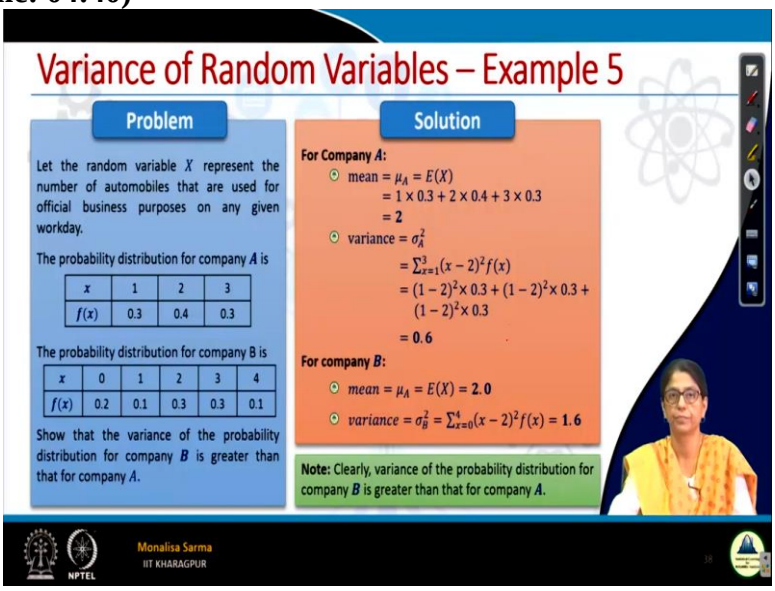
(Refer Slide Time: 03:52)



Here we see, you see this is the variance, variance is expected value of x - mu whole square if it is discrete we use summation, if it is continuous we use integration, so, what is see here this here in the example I have given 2 distribution so, distribution here mean is same for both the distribution mean is here also means is 2, here also mean is 2, here it is mean is 2 here a mean is 2, but the variation is different for both the cases.

This is here a variation is very less, here a variation is quite more it is more spread out variance means how spread out it is? It is more spread out.

(Refer Slide Time: 04:40)



So, here is an example. So, what it is given let the random variable X represents the number of automobiles that are used for official business purpose on any given workday, number of automobiles that if used in any given workday, in any given workday company may use 1 automobiles 2 or 3. What is the probability of using 1 is 0.3 probability of using 2 is 0.4 and probably using 3 is 0.3.

And again there is another company another company which may use in any particular day it can it may use no automobiles at all that is 0 or 1 2 3 or 4 and the probability of all this values are given for 0 is 0.2 and likewise it is given in this table this table. Now, what we have to see is so, that the variance or the probability distribution for company B is greater than for company A.

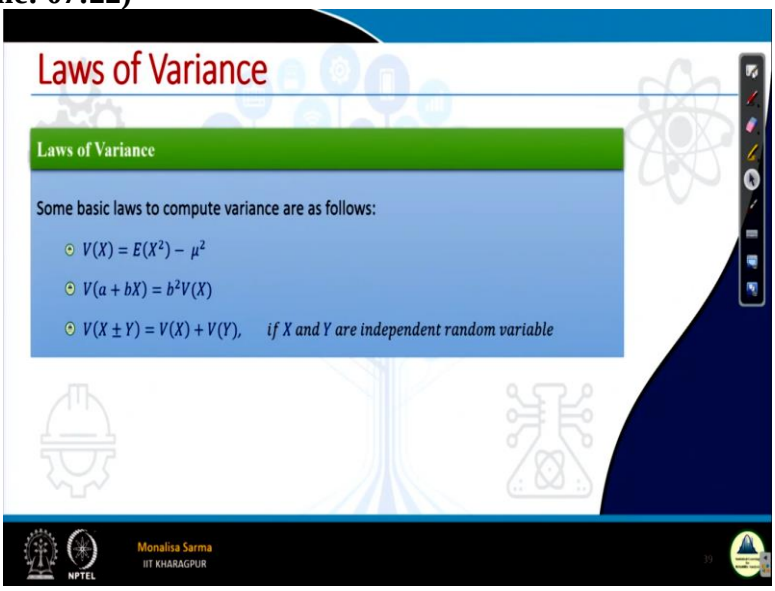
We have to show that the variance for company B is probability distribution of company B is better than company A that means, what we have to find out? We have to find out a variance of company A then we have to find out the variance of company B and from that we will see that the variance of company B is greater how we will do that now, how to find out the variance to find out a variance? First we need to find out the mean.

Mean is nothing but what is the formula for mean? Mean is summation of xf(x). So, what is the x value? x is 1 into 0.3 2 into 0.4 3 into 0.3. Now, what is variance? Similarly, variance is

$(x - 2)^2$ that is variances $(x - \mu)^2$ f(x). So, my μ is 2 $(x - 2)^2$ f(x), so, what is this x minus what value w will take? x will take 1, 2 and 3. So, 1 - 2 into probability of 3, so, this is 2.

So, 2 - 2 into probability of 2 there is a slight mistake in the slide I am sorry for that and then next is that it will take a value of 3. So, this value is 3 so $(3 - 2)^2$ into that probability that is 0.3. So, what we got this will cut out to be 0 2 - 2 is 0 and this from this and this will be getting 0.6. Similarly, we will find out the variance for company B as well. So, once we find out a company variance for company B same way we found out 1.6. So, the variance A is 0.6 variance B is 1.6. So, definitely company B's variances higher.

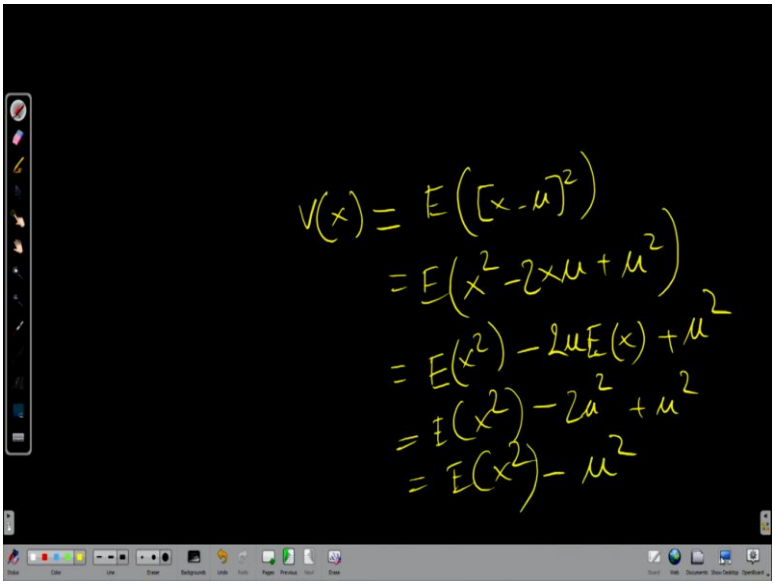
(Refer Slide Time: 07:22)



So, now, some basic laws to compute variance, variance that is we can tell it as VX, VX = E of $x^2 - \mu^2$, V of a + bX = b^2 V(X). V (X) + V(- Y) means 2 random variable X - Y which has 2 different distributions. So, whether these 2 random variables, we have added or sum it the variance what we get is V(X) + V(Y). This is a very important conclusion with we may need it many other problems when we are trying to find out the variance of 2 different populations.

And we are trying to variance of 2 different population if even it your earlier it was V (X + Y)or X - Y always it will be always the variance or added variants are never subtracted. So, now, how to we can also see how we can deduce this, I will be showing you 1 or 2 rest you can do it by yourself this meant a similar manner. So, first we will see the first one what is the first one?

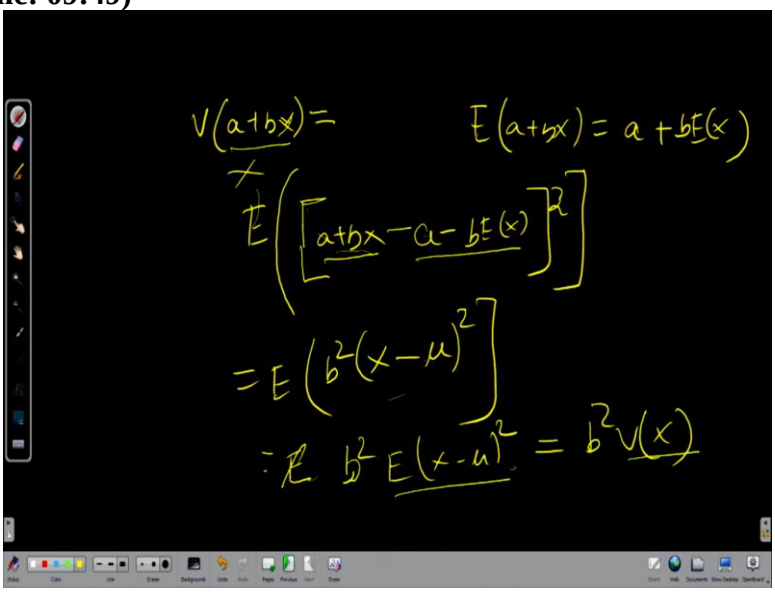
(Refer Slide Time: 08:32)



First one is your V of X.V of X how I can write V of X I can write as E of X – μ whole square. So, if I break it, how does X square minus twice X μ + μ^2, so, I am using the rule for expected value we have seen. So, this is E of X square this is 2 of X μ. So, this is 2 of μ into E of X because μ is constant. So, its expected value will be just μ only and then μ^2 what is mu square E of X? μ square E of X is E of μ^2 will be simple μ^2 that is a constant value.

So, what we got E(X^2) - E (X) is what E of X is μ. So, we got 2 μ^2 + μ^2. So, we got E of X^2 - μ^2. So, this is the first one V(X) is V (X^2 - μ^2). Similarly, second one also I will show you but third one you please do it yourself. So, V a + bX = b^2 VX. How will you find this?

(Refer Slide Time: 09:49)



So, for this V of we need to find out V (a + bX), so what first we will see here, we already know what is E (a + bX)? E (a + bX) we already know we have seen in our last lecture that it will be a + b of EX. So, now, when we are interested in finding out V (a + bX) what is this

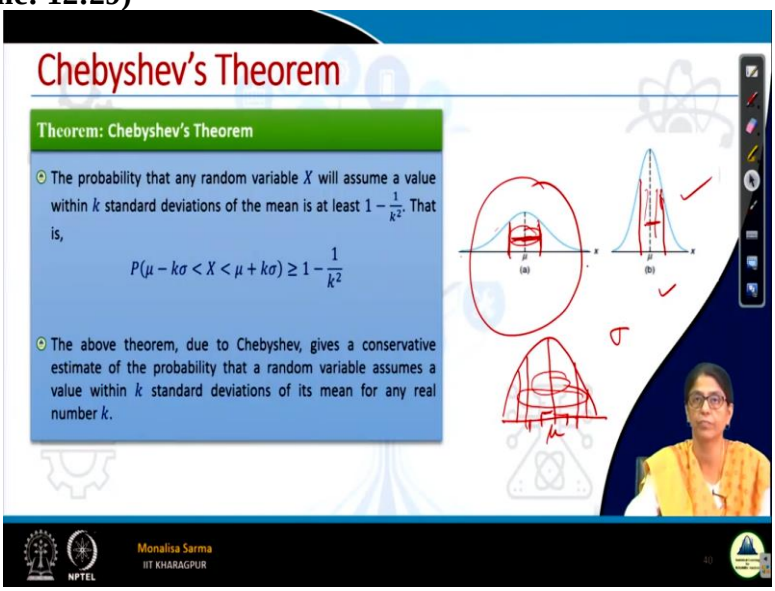
this will be equals to E of what X - μ ^{2} square this equal to E of this is if I consider this as X so, a + bX this is X – μ μ of means mu of this what is mu of this I have done it here that is - a - b of EX whole square.

So, what we got this a a get cancel b square remaining, what is there, what remains? So, it will be X - E of X is what E of x I can write as $\mu(X - \mu)^2$. So, what is this E of b square I can bring out let me cut this. So, this is b square b square E of $(X - \mu)^2$what is $(X - \mu)^2$? $(X - \mu)^2$ is nothing but V. So, this I can write as b square V of X.

Again I am repeating how I have done so, I am interested in finding out V of a + b what is V of a + bX? V of variance of X is $(X - \mu)^2$. So, this is X so, (X - μ), μ means E of X so, this is X and this portion is E of X E of this I have seen it is this. So, I have subtracted this a a get cancel what remains is b into X - E of X, E of X is μ (X - μ) whole square, so b square and bring it outside.

So, I got b ^{2} E of (X - μ)whole square and now, what is this this is nothing but equals to V of x. So, similarly, you can find out try and do the third one.

(Refer Slide Time: 12:29)



Now, there is an important theorem, important theorem in the sense like see now, you understood what is a variance? Variance means how spread out my data is, now this you see in this figure and this figure both the figure may have the same μ but say this figure is more spread out the second figure, this figure is more spread out than compared to this figure, this figure.

This figure is more spread out, the spread out means its variance is more and suppose I am interested in finding out what is the in any random experiment for any particular event, I am interested in finding out what is the probability that my value will lie in this interval some distance from μ similarly, I have considered the same distance from μ what is the probability of a value line in this interval.

So, in this case the probability of lying in this in this case my probability of lying when this interval will be high compared to this my probability of lying in this interval, lying in this interval will be high compared to this probability distribution, why this probability is nothing but in case of continuous probability it is nothing but the area if you consider the area this area is because I hold it this whole area corresponds to one.

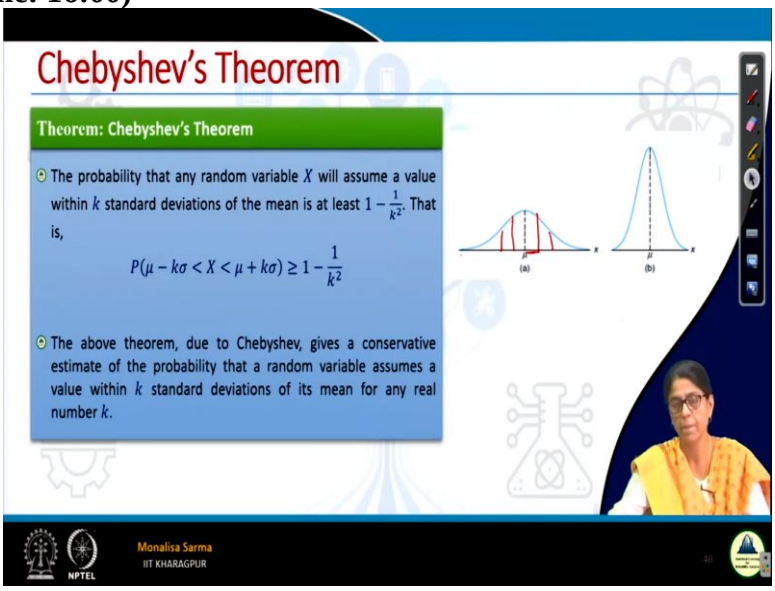
So, definitely this area will be less compared to this area. So, that means my probability will be higher lying in this range. So, Chebyshevs theorem basically gives a very conservative estimate of that. So if I, in any random exponent for a particular event, the probability of lying within a particular distance from mu what is the probability of lying getting a value within a particular distance from μ that Chebyshevs theorem gives that.

So, what is that the probability that any random variable X will assume a value within k standard deviation of the mean, k standard deviation means one standard deviation. So, suppose I have drawn a figure suppose this is my figure, suppose this is mean and suppose this is one standard deviation, this distance is standard deviation, standard deviation is variance route overall variance this my standard division This is my standard deviation.

So, this is one standard deviation then 2 standard deviation maybe this point what is the probability of lying this region within this region, 3 standard deviation maybe this point what is the probability of lying within this region. So, what the Chebyshevs theorem says the probability that any random variable X will assume a value within k standard deviation is at least $1 - 1 / k^2$.

But here k, this is applicable when $k > 1$ it is not applicable for $k = i$, so, above theorem due to Chebyshevs gives a conservative estimate of the probability that a random variable assumes a value within k standard deviation of its mean for any real number k, it is clear?

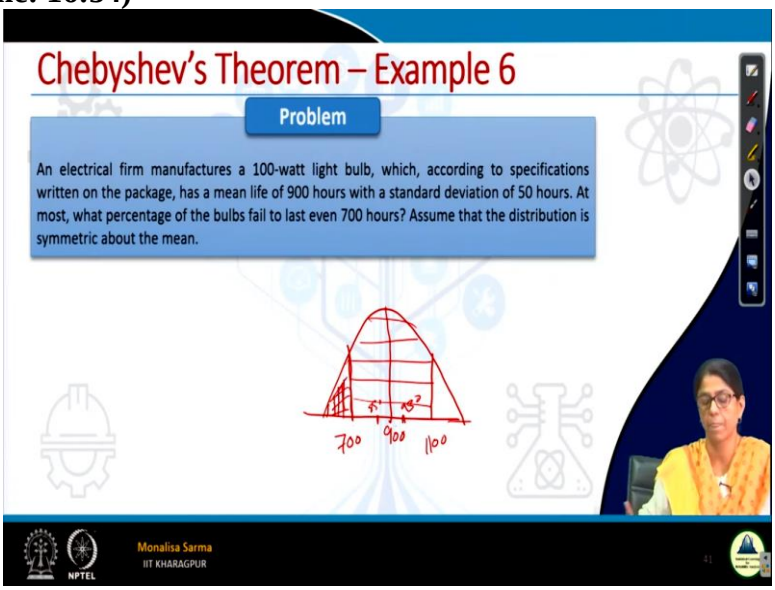
(Refer Slide Time: 16:00)



Essentially what the Chebyshevs theorem tells? Chebyshevs theorem tells what is my probability of getting a value in a particular range, this range is specified by the standard deviation what is the probability that I will run in a random experiment and I will be getting a value how much standard deviation from the mean, mean is the middle portion, from this mean how much standard deviations?

If it is one standard deviation means maybe this, one standard deviation maybe this mean, this may be one standard deviation, maybe 2 standard deviation maybe this, this is what Chebyshevs theorem says, this gives very conservative, conservative estimate of course. But still this gives us an idea in an experiment based on a variance we can find out what is the probability of getting a particular getting a value within that range.

(Refer Slide Time: 16:54)



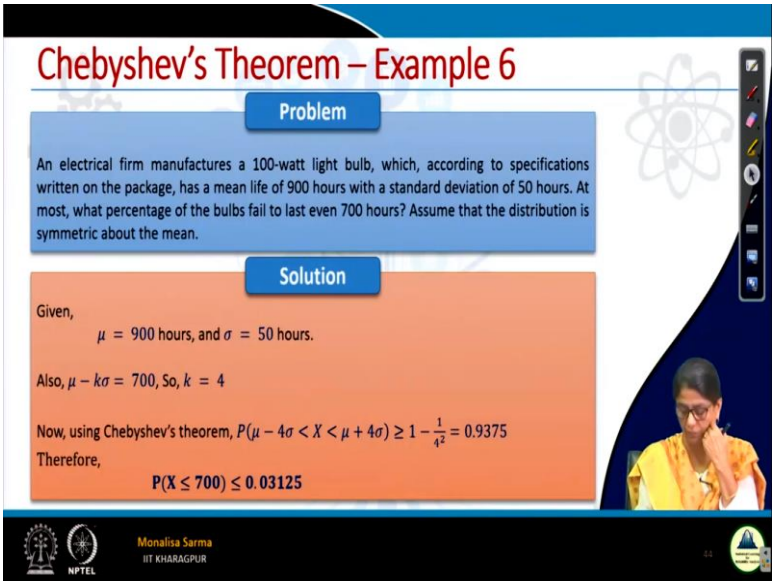
So, one example will make it more clear you see an example here an electrical firm manufactures 100 watt light bulb which according the specification written on the package has a mean life of 900 hours and a standard deviation of 50 hours. So, this electrical bulb, so it has suppose I am using a simple distribution suppose it is this type of distribution. Till now, I have not explained the type of continuous probability distribution of this problem.

So, let us assume that normal distribution so, this is what is given mean is 900 hours, so, this is maybe the mean, mean in standard hours this side it will be greater than 900, this side will be less than 900. And standard deviation of 50 maybe is 50 is this point that means, it is 950 and this is 850 here that is the standard deviation, standard deviation is 50. So, if I am considering one standard deviation, one standard deviation maybe 950 at this point.

This is 850 one standard deviation, 2 standard deviation this side it is 1000, this side is 800 within 2 standard deviation. So, now, what is the question at most what percentage of the bulb fail to last even 700 hours, failed to last even so, this 900 maybe this much is 700. So, the question is what percentage of the bulb failed to last even 700 hours. So, that means, my bulb but the value that I got that means that my working life of the bulb is all in this range lesser than 700.

It is not in this range greater than 700 is in this range. So, what I will do is that I will try to find out the probability within range from this 700, this side 700 means, it is from 900 to 700 how many standard deviations it is? It is 50, it is 4. So, here so 1100, so, what is the probability of finding in this range, I will find out using Chebyshevs theorem and then I will subtract it from that 1 minus I will get this area is not it.

(Refer Slide Time: 19:17)



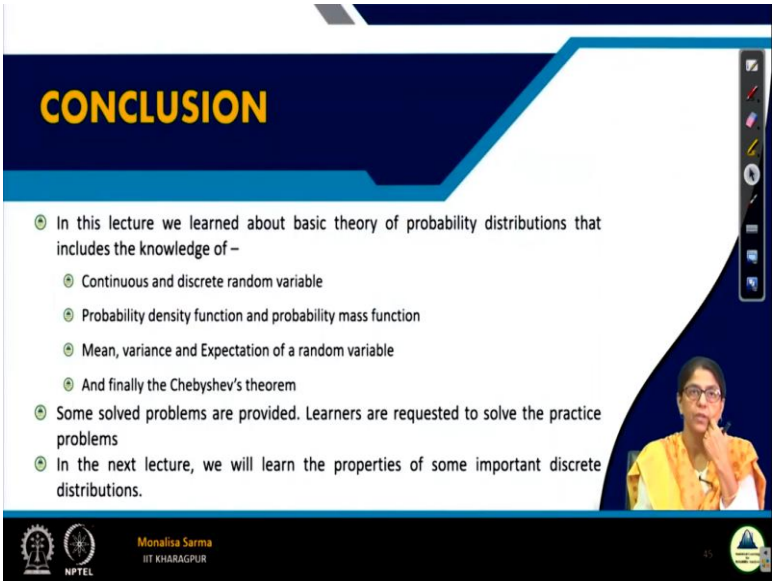
So, see here my 900 is, μ is 900 σ is 50. So, μ - k σ is 700. So, definitely k goes to 4 that all without solving also we can say because it is sigma is 50 it is very easier to tell. So, now, using Chebyshevs theorem, but we get probability of this is at least this value that means greater than equals to 1 minus this probability that it will lie in the range of 700 to 1100. It is this much 0.9375.

So, what is the probability that it will lie that it lesser than 700, this is 900 so, this is 700, this correspond to 700, this corresponds to 1100. So, the probability that it will lie in this range is just 0.9375 Now, I am interested in finding out this probability. So, 1 minus of this probability I will be getting this and this together. So, it will be lesser than this that will be divided by 2.

So, it is this much, we can just do it and one minus of this and then divided by 2 because it is this lesser than this as well as greater than 1100, it can greater to both the portion because we have just found out a property of the middle portion and we want to find out what is the property that has appeared in the lesser than 700. So, then what we got one minus of this it will greater to both the portion, lesser than 700 as well as greater than 1100.

1 minus so what that value 0.9375 will give that and we are interested in only lesser than 700. So, whatever value we will get probably that will be divided by 2, then we will get 0.3125.

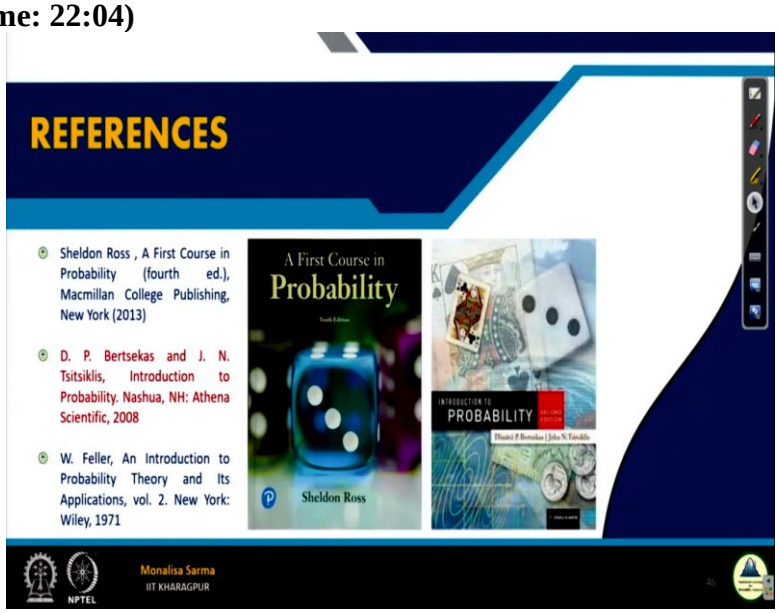
(Refer Slide Time: 21:13)



So, I conclude here. So, in this lecture let I should not say in this lecture, in this lecture and a previous lecture that was lecture 7, but we learned about the basic theory of probability distributions earlier we have learned probability now in this last 2 lectures we have learned what is the probability distribution, what is a random variable then we have now learned what is continuous probability distribution, discrete probability distribution, we have learned how to find out the mean of a distribution, variance of a distribution.

And also we have discussed Chebyshevs theorem like last time again I am repeating solve as many problems as possible. Now, in my next lecture, we will learn different discrete probability distributions. With that I end this class thank you. Here the references of the different books and thank you guys.

(Refer Slide Time: 22:04)



Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology - Kharagpur

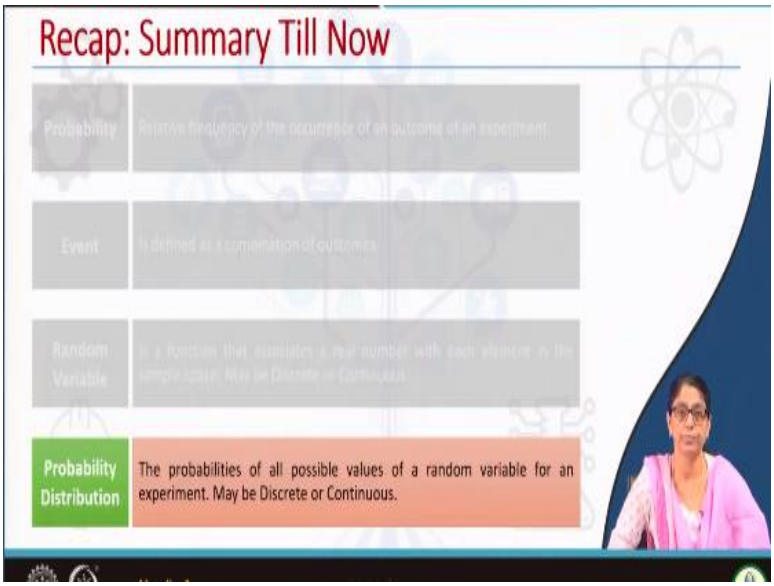
Lecture – 09
Discrete Probability Distribution (Part 1)

Hello guys, so continuing with our discussion on probability distribution, today we will be discussing some discrete probability distribution. So, there are many different types of discrete probability distributions. So, in today's lecture, we will be discussing few of them. And next few of them we will be discussing in our next lecture.

(Refer Slide Time: 00:42)

So, today's lecture will cover basically uniform distribution, binomial distribution, multinomial and hypergeometric distribution. There are some other distributions as well like Poisson negative binomial geometric I will be covering that in my next lecture.

(Refer Slide Time: 00:58)

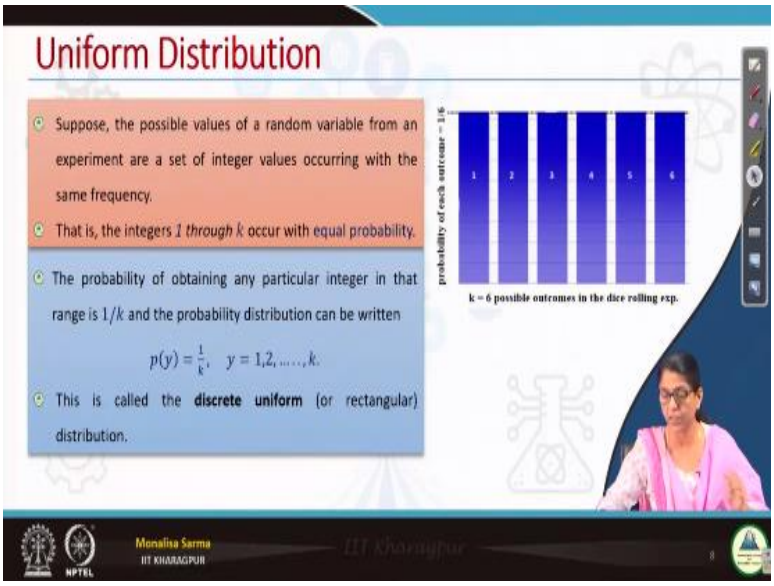


So, before going to discuss the different probability distribution, let us take a quick recap of what we have learned till now, just a few terminologies that we have learned that is the first is probability, what is probability? The relative frequency of the occurrence of an outcome of an experiment we have seen that even while trying to calculate the mean also have a stress in this factor, how will we take the relative frequency while calculating the probability of an event, remember that.

And then what is an event it is defined as a combination of outcomes, then what is a random variable, it is a function that associates a real number with each element in the sample space. So, the random variable may be discrete, it may be continuous, so, if the random variable is discrete, the corresponding probability of the different values that a random variable will take in an action that will give us a discrete probability distribution.

In a random variable is a continuous in an experiment that different values that are random variable can take along with the probabilities that is a continuous probability distribution.

(Refer Slide Time: 02:07)



So then next, first we will start with the uniform probability distribution; this is a discrete probability distribution. However, before discussing on this uniform discrete probability distribution I would like to mention here this distribution is not very much used in any practical applicants or access. Of course, it is not that it is not used at all, but its applicability is less compared to some other distribution like when we were discussing binomial distribution, Poisson distribution, although the quite used in many applications.

Like the uniform distribution is not very much used, but for completeness, you should know this also as well. So, what is uniform distribution? So, uniform distribution suppose the possible values of a random variable from an experiment are a set of integers from an experiment the probability values that a random variable can take and we have discussed this in an experiment random variable can take some possible values. So, possible values these are a set of integers and they occur with same frequency.

Like when we toss a die what are the value the random variables may take this 1 2 3 4 5 6. So, what is and this in case of uniform probability distribution what is the condition is that this all this values occur with the same frequency. So, that is in the case of die we can say it is 1 to 6 let us tell that it is k where k is 6 that is the integers one through k occur with equal probability. So, if it equals if it occurs with equal probability, then what is the probability of getting any n number say p (n) let me tell it as p (y).

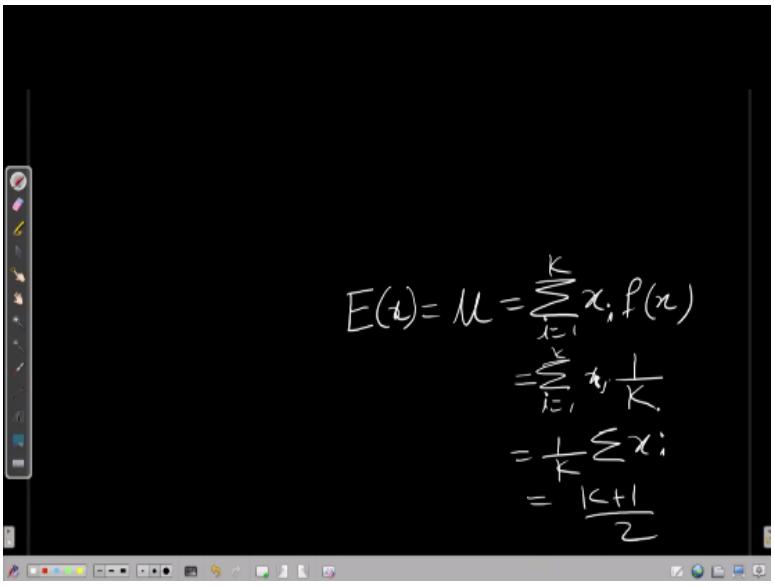
So, p (y) will be equal to 1 / k. So, this is the discrete probability distribution function p (y)or we can tell it as f(y) also. Now, when we describe any distribution any maybe continuous distribution or discrete probability distribution any distribution when we describe there are a few characteristics which you call it as parameters also which basically characterize the distributions, which having the information of these parameters will help us to find out what is the shape of the distribution.

Helps us to give us different information of the distribution so, like for different parameters that are that, we can also say that different characteristics that are one is the mean of a distribution, variance of a distribution and then symmetry, symmetry it is called and another one we can call it a kurtosis. But when we talk of a mean and a variance definitely mean and variance we will be discussing in this lecture of course, mean and variance is very much applicable when we talk of discrete in discrete probability distribution.

It is also applicable for continuous probability distribution definitely. And but when we talk of asymmetry, how symmetry the curve is. So, this is not very much relevant from the discrete probability distribution persue. So, we will not be discussing here, when I will be discussing continuous I will specifically mentioned about that, then again kurtosis, kurtosis is how flat the curve is that is also a delivered from the not very much I should not say it not very much relevant from the discrete probability distribution contexts.

So, the different parameters that we will be seeing in the discrete probability distribution are mean and variance. Now, how to calculate the mean of this distribution already we have seen how to calculate the mean of register and remember mean of the distribution is x f(x) summation of x all the values of x into f(x) what is f(x)? f(x) is the probability of x we have seen that. So, now, if we calculate them, so, like mean of a distribution we have seen it is a summation of x f(x), x is the all the values that the random variable will take. That is we call it x i basically i range from y 1 to n and then f(x), f(x) is the probability that x will take.

(Refer Slide Time: 06:17)

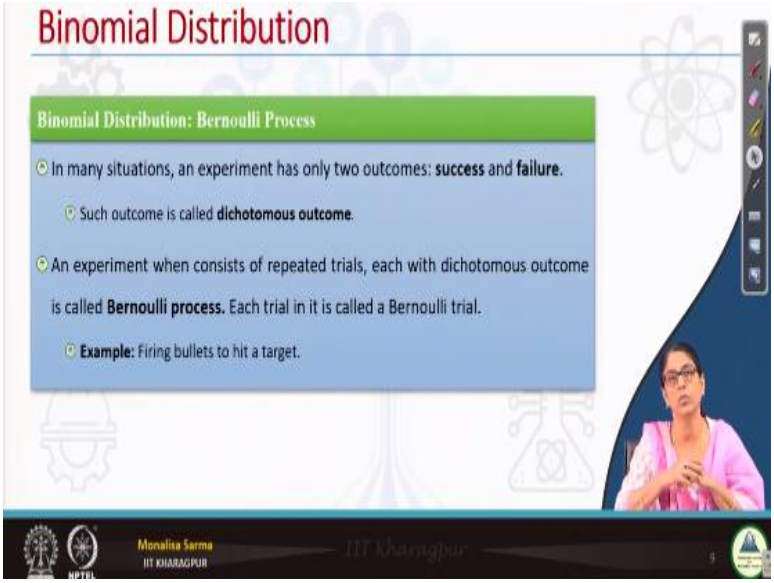


So, how let us see in case of uniform distribution what happens. So, what is we call it expected value remember we have done this expected value we could call it f of x also we can also call it μ and what is this? It is summation of x_i f (x f (x)) where i = 1 to k suppose here the values goes from 1 to k now, in this uniform distribution what is that i = 1 to k x of i and then what is f (x)? f (x) is nothing but 1 / k so, 1 by k it will just come out 1 by k so, it is summation of x i so, what it will give, what is summation of x I, what is k into k + 1 by 2?

So, what we got is k + 1 / 2, this is the mean of the uniform distribution. Mean of the uniform distribution is k + 1 / 2. Similarly, we can find out the variance also. So, if we find out a variance similarly, remember the formula for variance, formula for variance is summation of $(x - \mu)^2$ f (x), x – μ, μ is the expected value which we got just now got $(x - \mu)^2$ f(x).

Now, similarly, if you can simplify it, I will be showing you that we did a very easy actually, if you simplify, then the variance will be getting k^2 - 1 / 12 that is the variance of the uniform probability distribution.

(Refer Slide Time: 07:57)



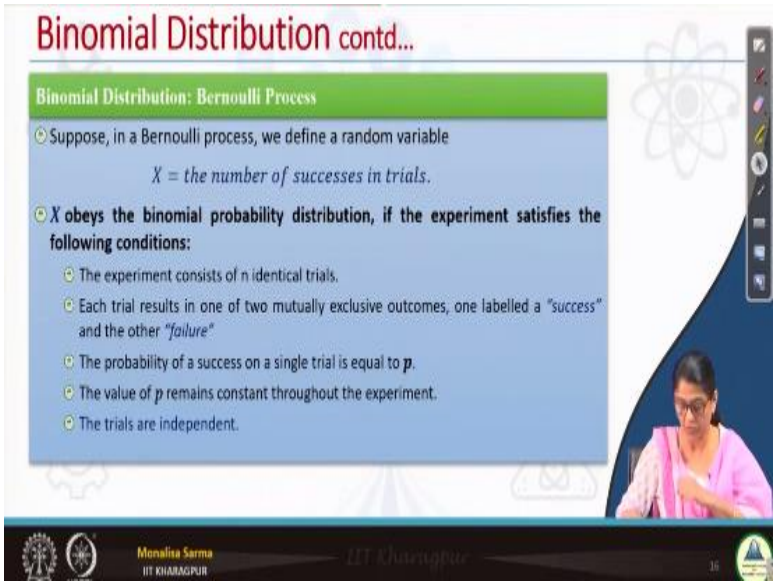
So, now, next distribution that we will be discussing is binomial distribution it is also called Bernoulli process. Now, in some situation an experiment has only 2 outcomes like we are trying to hit a target. So, hit or miss that is like we are doing an test by some say blood tests we are looking for some disease it is present or not. So, it is positive negative like now, we can support COVID positive negative. So, these are the in some experiment has just 2 outcomes.

Such outcomes is called dichotomous outcome. So, an experiment which consists of repeated trials, we do many trials each with dichotomous outcome is called a Bernoulli process, now, you are firing a target. So, we are doing it we are repeatedly doing it, there is a target fixed we are trying to hit the target repeatedly we are doing the experiment that is called repeated trials. Now, a blood test we are not suppose we are not convinced we are doing the experiment again and again. So, that is a repeat we call it repeated trials.

So, when the experiment when it is consists of repeated trials each for each of the trial we will get a dichotomous outcome that means, either success or failure success or failure are hit or miss that is in generalized form we call it as a success or failure it is not necessarily as success or failure only like for if we hit the target we call it a hit or miss. So, it is not a question of success or failure. So, but in generalized form we call it either we got a success or we call it failure.

So, this type of process is called a Bernoulli process and each such trial is called a Bernoulli trial. Example, but I have already just told firing bullets to hit a target.

(Refer Slide Time: 09:46)



So, now, suppose in a Bernoulli process we define a random variable. Probability distribution means there has to be a random variable, then only we can find a probability decision isnt it? Means what is the value that random variables will take as well as the probabilities of those random variables. I am repeating this again and again because what is probability distribution many students find it difficult to answer this question.

So, suppose in a Bernoulli process, we define a random variable. So, what is a random variable? Random variables is X, X is the number of success in trials. So, X is the number of success in trial. So, suppose that means let us take that we are doing total n trials, we are trying to hit a target, suppose in front of me there is a target, I am trying to hit the target, I am totally I am doing this n times, out of n time, suppose my number of success maybe 4, so that is X out of n my number is success maybe 1 2 3 4 up to n listen it, it can take any value from 0 to n basically.

It can be 0 as well, there cannot be any there may be no success at all. So, X is equal to the is represents the number of success in trials. So, X definitely X obeys the binomial probability distribution, because X will take the either X will, it will either hit or it will be missed. So, it follows the binomial probability distribution. So, if the experiment set is now for a random

variable, to call it as a binomial, that it obeys the binomial probability distribution, the experiment has to satisfy certain conditions.

That we are going to experiment now, this way trying to hit the target that is an experiment we are doing a blood test that is an experiment anything. So, now what to say now to tell this experiment is a binomial process, it needs to satisfy certain condition. So, what are the condition first is that the experiment will consist of n identical trials, n identical trial means I am trying to hit it from here. Next time also, I will try to hit it from here it is not that I am trying to hit it from here this time.

Next time I am going near the target and I am trying to hit it or next time I am trying to hit it from a different angle no, it should be the experiment consists of n identical trial the experiment each trial should be same. Second, each trial results in 1 or the 2 mutually exclusive outcome, definitely the outcomes are either success or failure both cannot happen success and failure but cannot happen together. So, each trial results in 1 or 2 mutually exclusive outcome, which we can label it as a success and the other as failure.

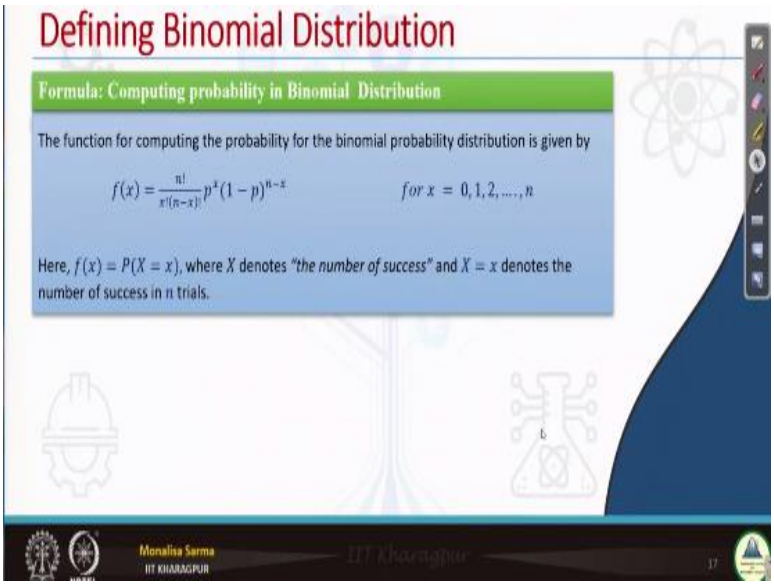
Then the probability of success on a single trial is p the probability of success out of n success n trial suppose your number of success is X, so what is the probability, probability is X / p. So, probability of success on a single trial is equals to p. And the value of p remains constant. Obviously, when the experiment consists of identical trial, the value of p will remain same, isnt it. So that is what the value of p remained constant throughout the experiment. And the trials are independent.

Each trial is independent means 1 trial, the results of 1 trial will not affect the other trial. So, the trials are independent. Now, let me give you a simple example of this drawing a car now we have in a deck of cards, we have 52 cards out of this 52 cards, what is the probability of in this pack of cards now my experiment is getting say black king. So, getting a black king, so out of this, I am picking it what is the probability of getting a Black King in my deck of card? What is the probability? How many black kings are there?

So, what do you find the probability? So, now suppose, what you do, you found a particular card, whatever car and you kept it, you do not put it back into us sub, you do not put it back in the deck of cards. And now again, when you try to pick one than, then it done my experiment will not be identical, why it will not be identical sorry, then my experiment will not be independent, why it will not be independent? Because I have now the number of cards is reduced. Early it is 52. Now, I have picked 1 card and I kept it aside, I did not mix it together.

So, now my trials are not independent. So, this is the concept of dependent independent according to the different situations different subjects.

(Refer Slide Time: 14:11)

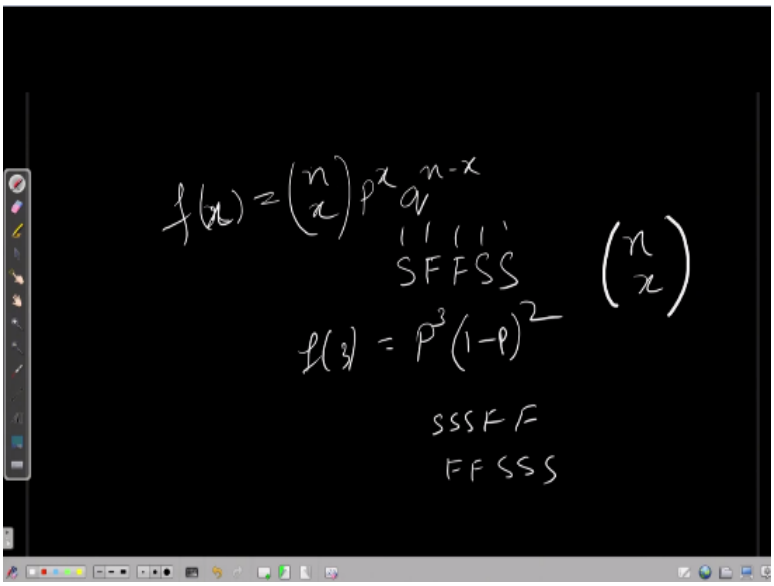


So, now, what is the, so any probability distribution we know we need to know what is f(x)? First we need to know what is a random variable, what the random variable represents in an experiment, then once we know what is the random variable, what the random variable represents, then we need to know what is the probability of f(x), what is the probability that is f(x). So, what is the in a binomial distributions what is f(x), for uniform distribution we have seen what is f(x)? f(x) = 1 / k in uniform distribution, now in binomial distribution, what is f(x)?

So, this is the formula basically nCx p ^ x 1 - p ^ n - x. That means probability of x success in n trials total the n trials out of n trials what is the probability of n success is n C x p ^x 1 - p ^ n - x. Let us see how we got this formula suppose we have total 5 trials suppose 5 trials out of 5 trials

suppose we are interested in finding out what is the probability of step 3 success, what is the probability of 3 success and 5 trails.

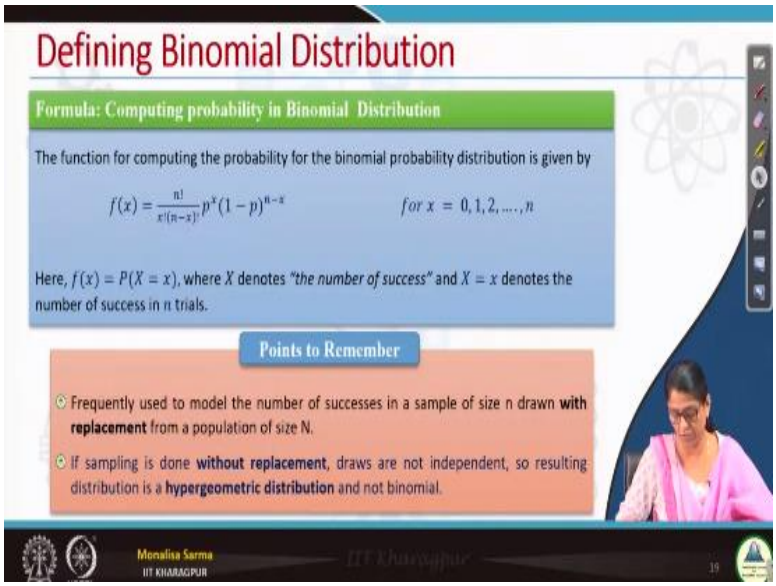
(Refer Slide Time: 15:23)



3 success means, suppose I got success, failure, failure, success, success, 5 trails this is 1 2 3 4 5. So, this suppose this is my sequence, so, then what will be my probability? Probability is independent so, my probability is p how many p is there? So, p^3 into how many failure that is 1 - p x 2 that is 5 - 3 this is 2. So, this is my f(x), f (3). Now, this is one sequence there may be another sequence like I may get S S S F F this may be 1 sequence, I may get F F S S S.

So, similarly, there may be different types of patterns where I may get 3 success. So, that is why when there are 3, when we have to select the x object from n objects, where we are not this distinguishing between the objects, what is the probability we have seen listen that is n C x. So, that is what. So, that is why my probability now f (x) = 2 n C x p ^ x q 1 - p let me write it as q q ^ n - x this is how we got the formula. So, that is the formula for probability binomial probability distribution random variables representing a binomial probability distribution.

(Refer Slide Time: 16:53)



So, now, some one important point to remember, this binomial distribution is frequently used to model the number of successes in a sample of size n drawn with replacement from a population of size N I have given the example of card where we need to replace it then only it will be independent and industry or weather in from a reliability perspective or quality perspective where this sort of thing where we use suppose in a production line in a pipeline, there are some products are coming out, some products are coming out in a pipeline.

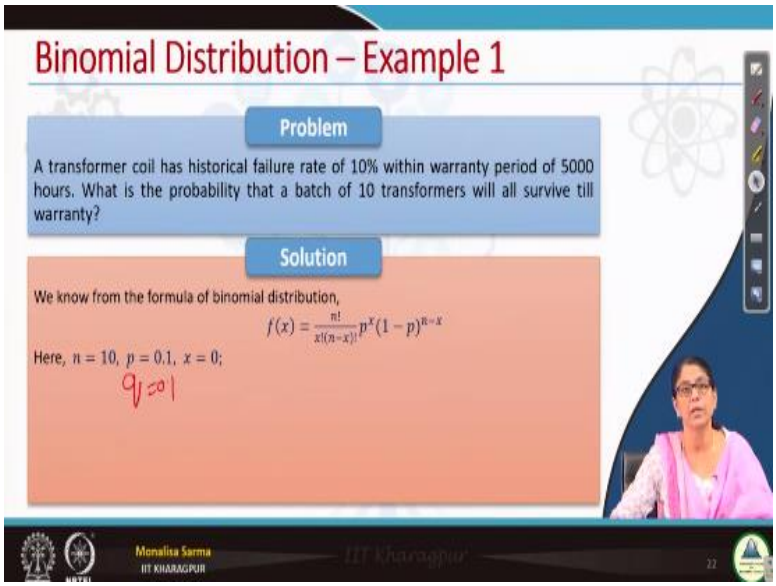
Now, from there, we want to check how many defective products are there. So, what we do? The products are coming out in the pipeline, we are just picking up 1 product, and we are checking whether it is defective or not, we are putting it back. Again, we are picking another way of putting it back. So, this is we are sampling with replacement, we are checking it on we are putting it back. So, this is in this sort of case, we use binomial distribution.

Again, there are some situations where, like, if you are interested in finding out the life of a bulb, so again, from the given a population of bulbs, we are interested in finding out the life of a bulb. So, definitely to find the life of a bulb, what we will do? We will try to do we will take some bulbs and we will see how long it works and from that basically, we will try to infer that what is the life of the bulk of the whole population.

Definitely we will not check for the whole population, we will take a smaller sample from the smaller samples, we will use those products till it becomes bad. So that means in that case, we cannot replace it back we have taken it, we will use it till it becomes bad. And then only we can find out the life of the bulk till when it lost. So, we could not replace it back. So, that is not a case of binomial distribution that is a different distribution we have for that that is called hypergeometric distribution.

When sampling is done without replacement, so that is hypergeometric distribution, it is not binomial difference mean there is a difference between both this is the only difference between the binomial and hypergeometric actually.

(Refer Slide Time: 19:09)

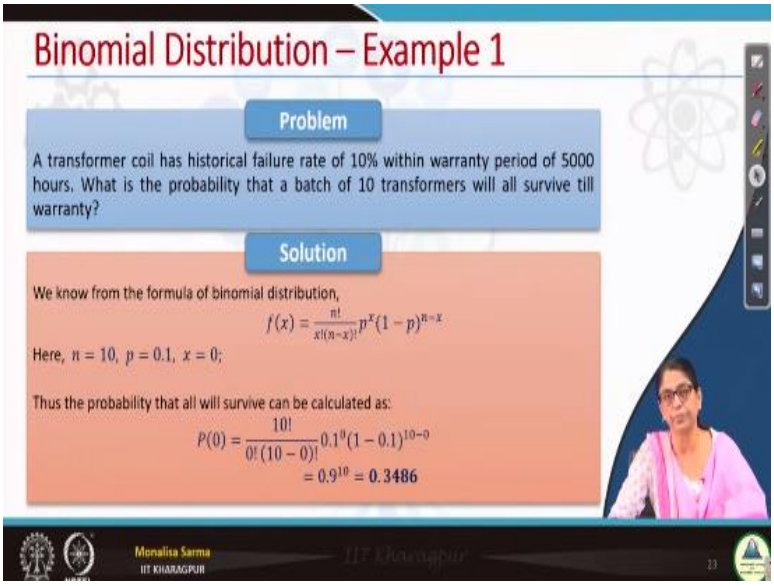


So, now, a small example before proceeding further. A transformer coil has historical failure rate of 10% within warranty period of 5000 hours, what is the probability that a batch of 10 transformer will all survive till warranty? So, for a failure rate of 10% within warranty period of 5000 hours what is the probability that a batch of 10 will all survive till warranty will all survive remains there is no failure, what is the failure rate is given failure probability is given 0.1 if the 10% means 0.1 failure probability is 0.1 so, what is the success probability successful is 1 – 0.1.

And what it is given what is the probability that a batch of 10 transform will all survive till warranty means there is no failure what is the probability? Probability of no failure in 10

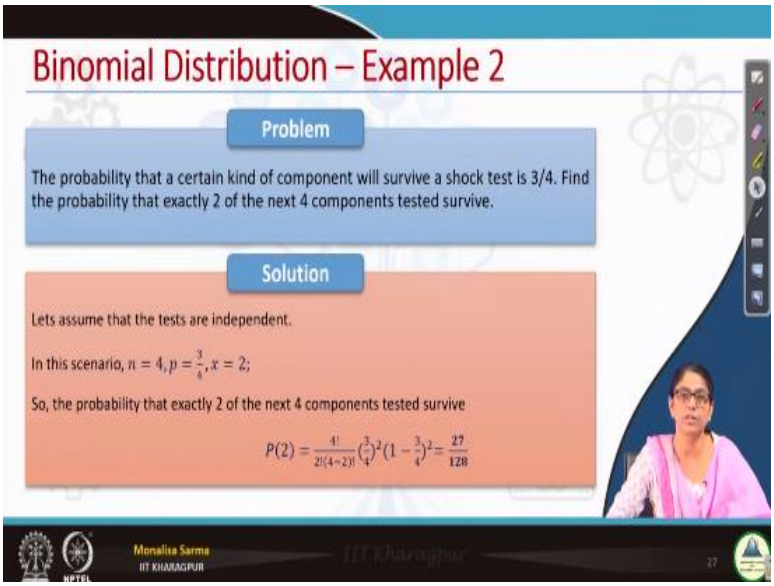
transformers that basically we have to find out that means we have to find out f of 0. So, your n = 10 p = 0.1 and x = 0 sorry here p is not equals to 0 rather I should say a p is because q = 0.1. So, there is a slight mistakes here this is this one can you see this is not p this is but q, this is not p this is q or sorry, this is not a question. Yeah, I am sorry for the confusion. So, this is not p this is q, q = 0.1. So, if q = 0.1 that means, p = 1 - 0.1 so, it is 0.9.

(Refer Slide Time: 20:50)



So, here the problem I have solved with taking p = 0.1. So, there is a mistake here accordingly you guys please change it. So, this is P is equal. So, according to equation p is equals to not 0.1, but 0.1 But here I have solved using p = 0.9. So, that is not an issue anyway whatever is p accordingly p value we will be putting here and whatever is then q will be 1 minus of p that is all.

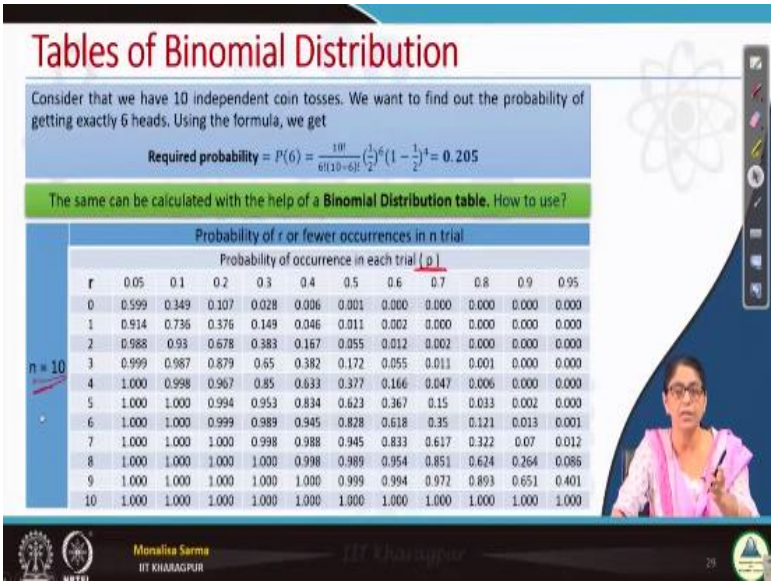
(Refer Slide Time: 21:15)



So, next the probability that a certain kind of component will survive a shock test is 3 4th find the probability that exactly 2 of the next 4 component is that survive. So, here probability is 3 by 4 and we have to find out of 4 components 2 will survive. So, it is again how we will be using binomial test thing distribution that means, here my random variable is a problem of survivability the number of component that has survived so, here I want to find out f of 2. So, f of 2 so, just put it in the formula that is all, nothing.

Not rocket science is here just putting into a formula and solving that. So, when you have the value of n, you have the value of p you have the value of x I am just put it in that formula.

(Refer Slide Time: 22:07)



r	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
0	0.599	0.349	0.107	0.028	0.006	0.001	0.000	0.000	0.000	0.000	0.000
1	0.914	0.736	0.376	0.149	0.046	0.011	0.002	0.000	0.000	0.000	0.000
2	0.988	0.93	0.678	0.383	0.167	0.055	0.012	0.002	0.000	0.000	0.000
3	0.999	0.987	0.879	0.65	0.382	0.172	0.055	0.011	0.001	0.000	0.000
4	1.000	0.998	0.967	0.85	0.633	0.377	0.166	0.047	0.006	0.000	0.000
5	1.000	1.000	0.994	0.953	0.834	0.623	0.367	0.15	0.033	0.002	0.000
6	1.000	1.000	0.999	0.989	0.945	0.828	0.618	0.35	0.121	0.013	0.001
7	1.000	1.000	1.000	0.998	0.988	0.945	0.833	0.617	0.322	0.07	0.012
8	1.000	1.000	1.000	1.000	0.998	0.989	0.954	0.851	0.624	0.264	0.086
9	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.972	0.893	0.651	0.401
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Now, here I will show you one more, I do not want to discuss one more example, but then I want to show you one concept like sometimes binomial distribution, we can also instead of calculating we can also use a table. So, but in the table we do not get a binomial distribution value, but what we get is a cumulative distribution remember, when we are discussing distribution probability distribution, we have also discussed cumulative distribution.

Cumulative distribution means we want to find out suppose in case of binomial distribution, suppose we are interested in finding out what is the probability of at least at most 2 success say at most to success means, so, there will be 0 success 1 success 2 success. So, it is cumulative, cumulative means we need to find a probability of 0 success + probability of 1 success + probability of 2 success. So, it will be finding many standard many tables standard tables where it gives the cumulative distribution.

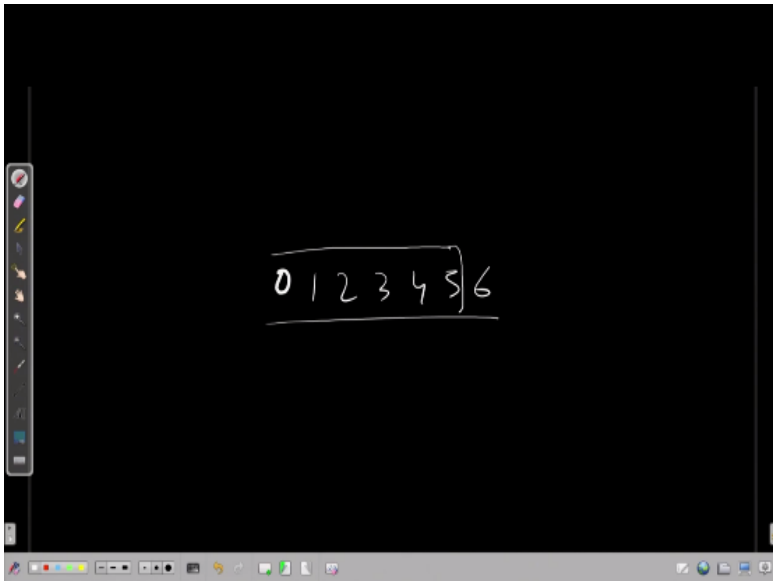
So, suppose some from the cumulative distribution table also you can use for finding out distribution at 1 point probability at 1 point you can use the table how you can use so, see this question consider that we have 10 independent coin tosses, we want to find out the probability of exactly 6 heads. So, here we want to find out a probability of exactly 6 heads. So, what is the probability? Probability of getting a head is half. So, 6 that means 10 C 6 half ^ 6 1 - 1 / 2 ^ 4.

Now, instead of solving this we can see this from the table how we can see this from the table brought exactly 6 heads how do we get exactly 6 at that table I told you it is cumulative. So, in the table that means I will get the value at most 6, then again I will get the value at most 5 heads. So, from at most 6 head if I subtract at most 5 heads, I will be getting the value exactly 6. That is what instead of calculating I can use the table. So, I will show you how we have so, this sort of probability tables are there.

It will find in any standard textbook in the appendix this sort of tables are given. So, this is a binomial thing here in this n = 10, n = 10 is the number of trials, this is the number of trials. And here we get this, this side you get the probability you get the probability of an r is the number of successes, where what we have used here x number of success, 0 success 1 success 2 success. So, r is the number of success and this is for this I have this table is only for n = 10.

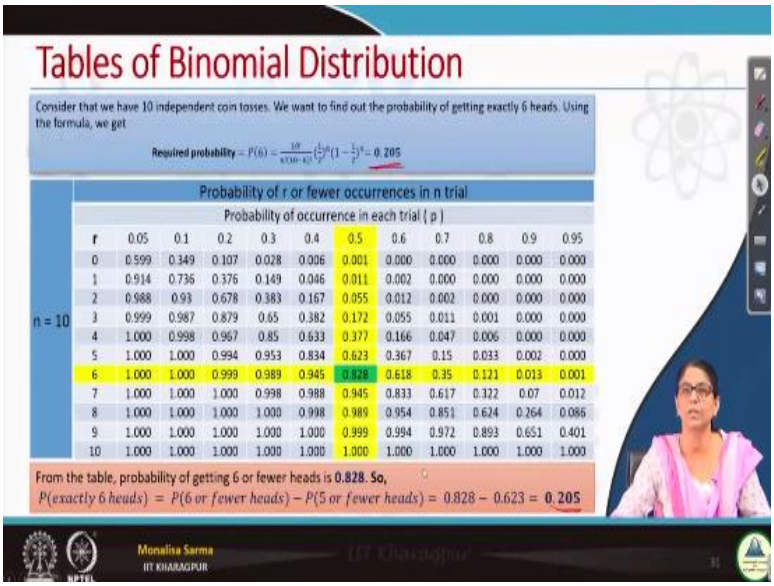
Similarly, there are different tables will find tables for n is equal to 11 12 many numbers and for r also for probability values also, today only after going through this lecture I request please go through this table you will see this table available in the net. So, how do we find out here for this question. So, for this question, we want to find out exactly 6 heads. So, how do we get exactly 6 head at most 6 whatever value we get, we will minus it from at most 5.

(Refer Slide Time: 25:10)



So, I will just show you suppose it is 0 1 2 3 4 5 6 at no 6 is, till this portion is in probability of 0 probability of one, 0 + 1 + 2 + 3 + 4 + 5 + 6. Now, if I did at most 5, so this so at most 6, if I subtract at most 5 from at most 6, I will be getting exactly 6.

(Refer Slide Time: 25:35)

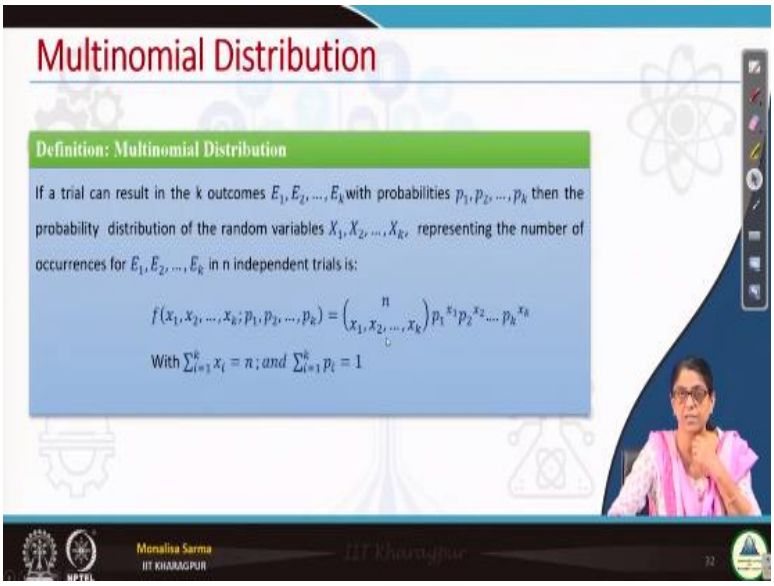


r	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
0	0.599	0.349	0.107	0.028	0.006	0.001	0.000	0.000	0.000	0.000	0.000
1	0.914	0.736	0.376	0.149	0.046	0.011	0.002	0.000	0.000	0.000	0.000
2	0.988	0.93	0.678	0.383	0.167	0.055	0.012	0.002	0.000	0.000	0.000
3	0.999	0.987	0.879	0.65	0.382	0.172	0.055	0.011	0.001	0.000	0.000
4	1.000	0.998	0.967	0.85	0.633	0.377	0.166	0.047	0.006	0.000	0.000
5	1.000	1.000	0.994	0.953	0.834	0.623	0.367	0.15	0.033	0.002	0.000
6	1.000	1.000	0.999	0.989	0.945	0.828	0.618	0.35	0.121	0.013	0.001
7	1.000	1.000	1.000	0.998	0.988	0.945	0.833	0.617	0.322	0.07	0.012
8	1.000	1.000	1.000	1.000	0.998	0.989	0.954	0.851	0.624	0.264	0.086
9	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.972	0.893	0.651	0.401
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

So, this is the value this is how this is the way we check. So, this is this yellow colour line you can say for n = 10 for r value is equals to 6 and probability is equals to half that is 0.5. So, where it matches so probability value is 0.828. Now, what is the value for at most 5? At most 5, you see, so for just above the horizontal line above the 6, so, at most 5 probability of 0.5 is 0.623 just above 0.28. So, we subtract this 0.623 from a 0.828 we will get the value see this in the top thing here, see you got 0.05.

So, we subtract this 0.623 from a 0.828 we will get the value see this in the top thing here, see you got 0.05. We calculated it manually here we calculate it from the table we got the same value. So, why unnecessarily will calculate it by hand we will consult the table the tables available but just you need to know how to see the table.

(Refer Slide Time: 26:34)

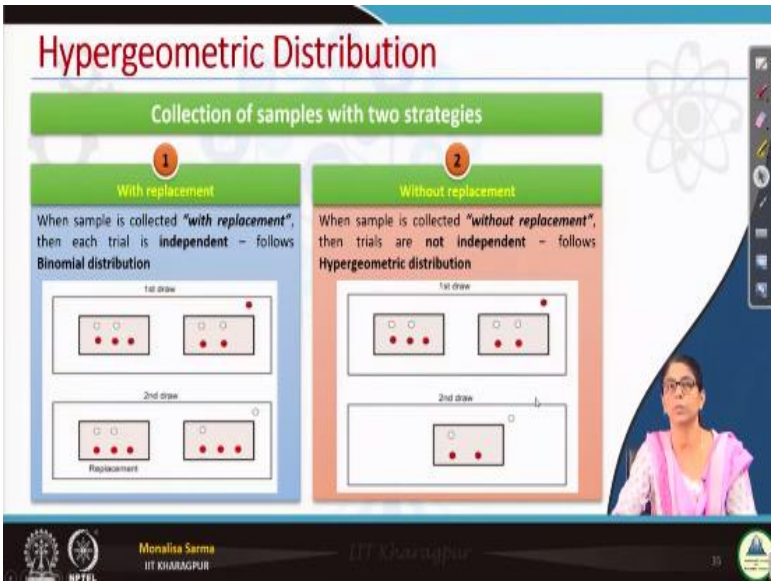


So, that is binomial distribution. Similarly, in the same line, we have another one type of distribution that is called multinomial distribution, multinomial distribution is also exactly same the properties are exactly same what we have seen the properties that the probability is same for all the trials, the trials are independent the trials are identical. So, here also the same but the only thing is that and they are with each trial, we get any one of the 2 outcomes.

So, here each trial can result in k outcomes mutually exclusive of course, they are also mutually exclusive 2 outcomes here are mutually exclusive k outcomes. So, from E 1 E 2 E k and each with different probability E 1 with probably p 1 E 2 probability P 2. However, the probability is remains same in all the trials. So, this is called multinomial distribution same quality characteristics exactly the same, but in binomial we have 2 outcomes in multinomial we have more than 2 outcomes, it may be 3 maybe 4 it may be 5 and even same exactly same.

So, exploration is also same there it was n C x via this n C x 1 x 2 up to x k how many different outcomes you have.

(Refer Slide Time: 27:46)

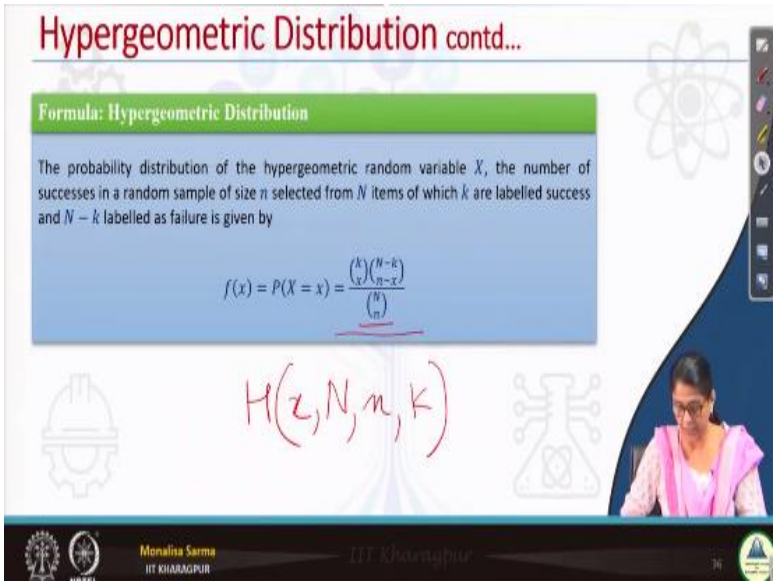


So, now hypergeometric distribution. So, what is hypergeometric already I have specimens and what is hypergeometric distribution when I was talking about binomial distribution to so, we replace without replacement what we get. So, binomial distribution is with replace, the trials are identity independent, but in hypergeometric distributions, so, what we did we do it without replacement from a sample of cards, we pick a card we do not replace it back. So, from in our next draw our probability changes because the number of card changes.

So, you can see this example. So, this week replacement suppose in a box there are some 5 balls, so, what we have done we have just picked 1 then the number of remaining ball is 4. So, now, we have replaced it again the number of balls became 5 then again we picked if we try to pick 1, so, our probability will not change because the number of the original number of balls remains same when we try to pick 1. But here we do not replace it initially suppose if we change them identify the color also so, suppose there are 3 red and 2 white balls.

And repeat 111 red ball. So, initially, here what will be the probability of picking a red ball picking a red ball what will be the probability, anyway I will not tell you should find it yourself. I have discussed this lots. So, now again now in the second row if I want to pick a white balls what happens earlier there was how many there was totally 5 balls now my number of balls is 4 from 4 I am picking 1 ball. So, according the my probability also changes. So, this there is dependency that characteristics not satisfied in hypergeometric distribution.

(Refer Slide Time: 29:30)



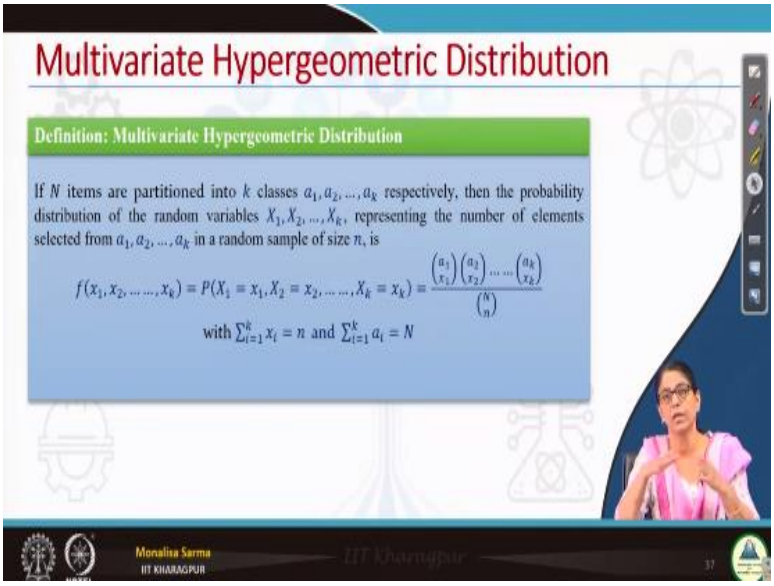
So, what is the expression for it affects so similarly, how we have done for binomial we can do it for my hypergeometric also, but I will not be solving it you do not need to solve that also, if you can, there is a very simple way by doing it again and again it will be in your brain or you do not have to it just in the tip of your tongue I should say like, so what is the hypergeometric distribution the formula for hypergeometric distribution as first here, one more thing I think we need to write it hypergeometric distribution.

How we write it basically H, H from number of successes x is number of successes here also we can tell it is a concept of success and failure, x than total number of elements, and is the total number of elements. And is this n is the sample that we have picked out number of elements that we picked out. And out of this n elements, how many successes are they, out of n, total n, that is N, and a N, how many successes were there, that is their total k success out of this.

You are trying to pick out from N you are you are trying to sample n number of items in n, x is the number of success. So, this is the formula for this h of x and k is this; what is given here this is my sample space is, this is my sample space. This is my sample space. Then from N, I am bringing n and our how many successes are there in N total k success there from k, my interest is picking out x axes. So, how many total item remains if I take out k, this n - k remains.

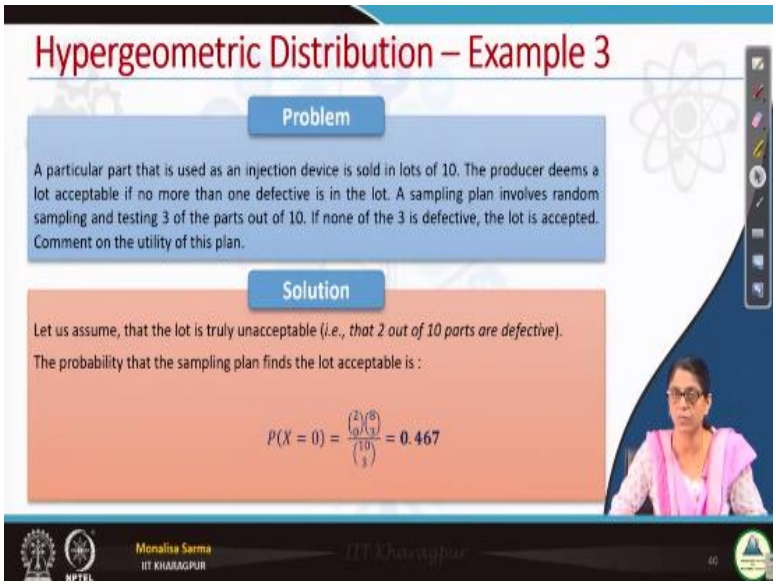
And from n - k how many I am picking because I am picking total n from n my success is x. So, how many is remaining and minus x is remaining. In fact, this sort problem we have done the 2 3 problems we have solved while solving probability problems that time we did not know about hypergeometric distribution, but we have done it in this way. Now, when you since you know hypergeometric distribution, solving that type of problem will be very easier, I suppose, listen it.

(Refer Slide Time: 31:43)



So, similarly multivariate hypergeometric distribution like what we have for multinomial binomial this multinomial distribution similarly, multivariate hypergeometric distribution, if we are interested in finding out suppose in N, there are different classes a 1 a 2 up to a n there are different classes and we are interested in picking up the x 1 item from a 1 class x 2 from a 2 class x 3 from a 3 class and that way. So, similarly to what is the formula for hypergeometric distribution same case that a 1 C x 1 a 2 C x 2 out of a k C x k. And upon what is the total number sample size? Total number of sample sizes N C r.

(Refer Slide Time: 32:31)



So, just you can see this as an example, a particular part that is used as an injection device is so, sold in a lots of 10. That is particular part that is used in injection device we are the it is a total around 10 10 parts are put together and in a packet that means in a packet we can say there are total 10 parts and the producer deems a lot acceptable if no more than one defective is not a lot the producer he can say that a lot is acceptable.

If there is no more than 1 if there is 1 defective the producer will accept it Chullora 1 defective we can accept it, but there should be no more than 1 if there are 2 defectives, 3 defected that means a lot is not acceptable and 1 defective lot we can accept a lot. That is the view of the producers. Now, what he will do he cannot check if form the 10 per lot in a packet there 10 he cannot check all this all the items are all the 10 items and find out whether how many are defective.

So, what he use, so, he has devised a sampling plan. So, what is sampling plan and the sampling plan involves random sampling and testing 3 of the parts out of 10 from 10 he will pick up 3 items and if none of the 3 is defective, the lot is accepted. So, what he thought this guy does guy is an intelligent guy, he at least he taught himself to be intelligent, he thought, I will not check this 10 items, let me check out of his 10 item let me just pick 3 items.

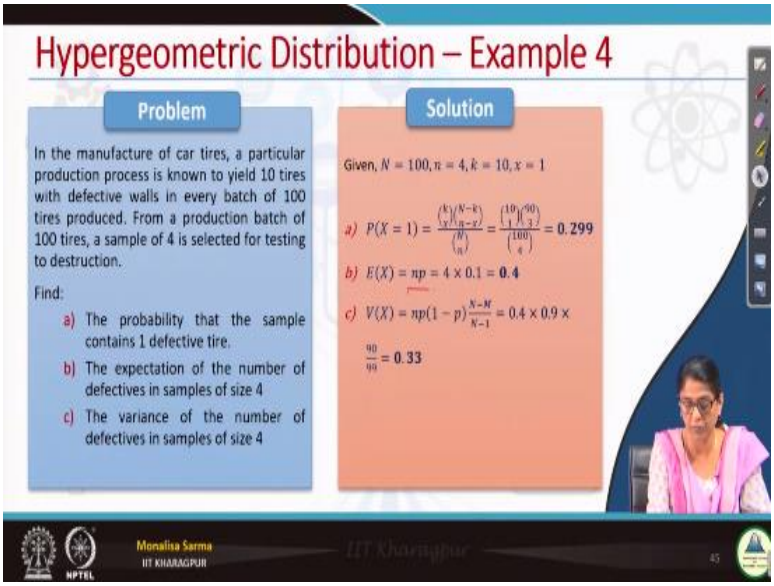
If in this 10 items at the most 1 defective is there, then if I pick 3 items, that means I should not get any defective that was his intention. So, if in 3 item if I do not get any defective, then I can be convinced that a lot is acceptable lot. Acceptable lot means that it is no more than 1 defective. So, comment on the utility of this plan. So, we have to find out whether this plan is whatever plan this person, this person has found whatever whether this plan is really good, let us find out the utility of this plan.

So, what we will do is that in this problem first let us assume that a lot is not acceptable, when the lot is not acceptable. There are more than 1 listen it. So, let us actually assume that there are 2 defective in this lot. And a lot of times suppose there are 2 defects in the lot and what he has done, he will pick 3 items. And what is the probability of getting 0 defective in the 3 item when the lot has 2 defective that is what we will find.

So, probability of x = 0 using how will from 10 he has picked 3 there are 2 defective from these 2 defective he did not pick anyone because 0 defective. What is the probability you are getting 0 defective that means 0 defective means we have not picked any defective. So, from the 8, he is picking the rest. So, what is the probability? Probability is 46% see. So, that means does even if the lot is unacceptable, it has 2 defective items that means a lot is unacceptable.

Even if the lot is unacceptable, but his sampling plan what he has done that if I pick 3 and if I do not get any defective then I can assume that a lot is acceptable, but his plan is not a good plan. His plan is a very poor plan. And because it is unacceptable lot also with his plan he will accept it 46% of the time, if this percentage is would have been very less than we can tell, his plan is good. Now, this is 46% is not a very less small number, it is a big number. So, we can tell this plan is not good. That is the common on this infinity of this plan.

(Refer Slide Time: 36:08)



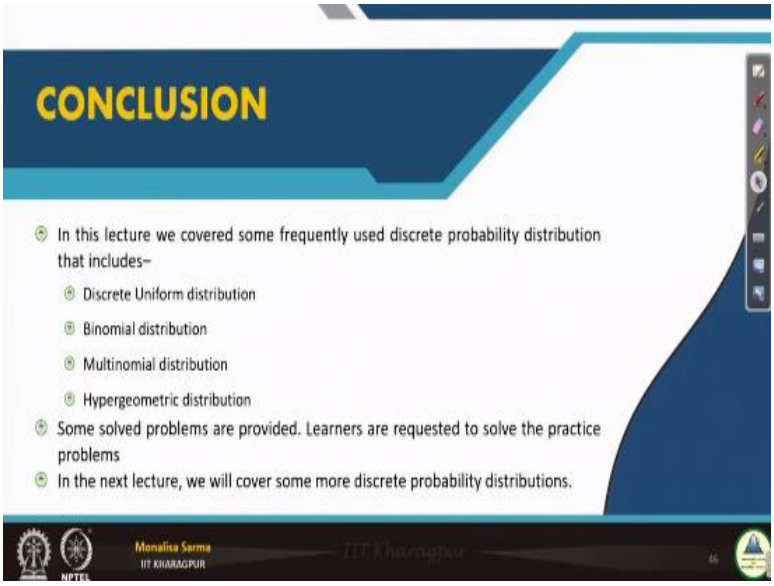
In a manufacture of car tires, a particular production process is known to yield 10 tires with defective walls in every batch of 100 tires produced from a production batch of 100 tires sample of 4 is selected for testing to destruction, this is very simple whatever we have done hypergeometric distribution same type of problem I will not be discussing it you guys will be able to solve it yourself. If you cannot solve it the answer is given here you can check it also it is the same similar type of problem, those solutions are given.

Now, here once the only one thing is that I forgot to mention the in case of binomial distribution, what is the mean what is the variance binomial distribution mean is p into q variance is root over and p and q I mean sorry, standard deviation is root over npq similarly for hypergeometric distribution when we talk about mean and variance. So, similar for hypergeometric distribution also what is mean and variance it is actually given in this formula only here.

Hypergeometric distribution mean is np, and variances np into 1 – p n – m, n - m means n you can take divided by n - 1 that is the formula for variance, you can calculate it by the formula x - mu whole square f(x). If you calculate it, you will get this or if you can remember than that is well and good. You do not have to calculate it again and again for each and every distribution. That is you understood how to calculate the mean how to calculate the variance mean is x f(x), variance is x - mu squared x - mu whole square f(x).

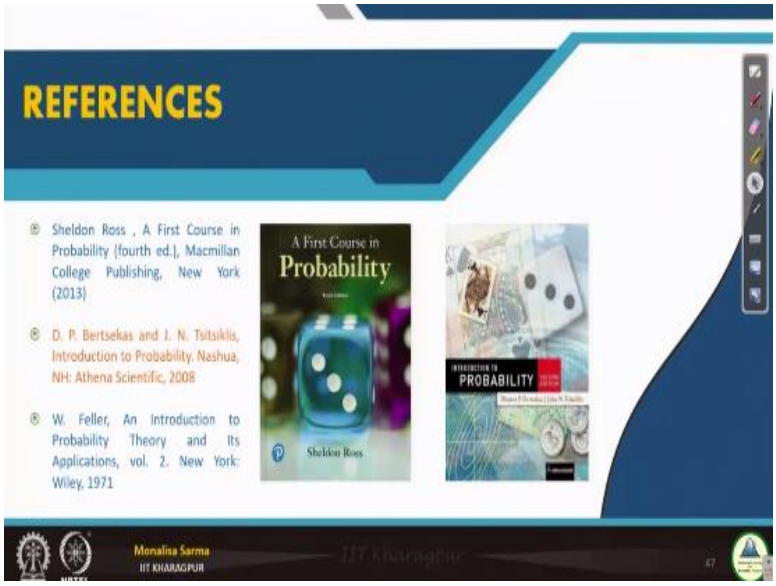
So, that is the way to calculate now, you can calculate it or you can remember what is whatever it is. So, this is the formula for mean this is the formula for here this is the formula for mean of hypergeometric distribution, this is the formula for variance of hypergeometric distribution.

(Refer Slide Time: 38:06)



So, with this I conclude this lecture, so, here we have discussed discrete uniform distribution, binomial distribution, multinomial and hypergeometric distribution. We have also solved few problems. Few more problems we will be solving in one of our 10 tutorial classes which I will be taking after completion of all the problems the distributions.

(Refer Slide Time: 38:28)



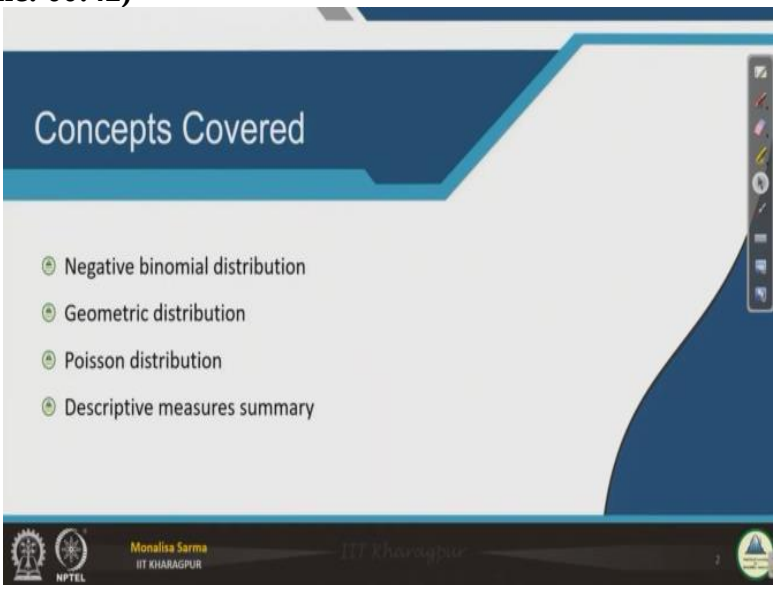
And again I request you solve as many problems as possible. These are the references which I had mentioned in my earlier lectures also and thank you.

Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology - Kharagpur

Lecture - 10
Discrete Probability Distribution (Part 2)

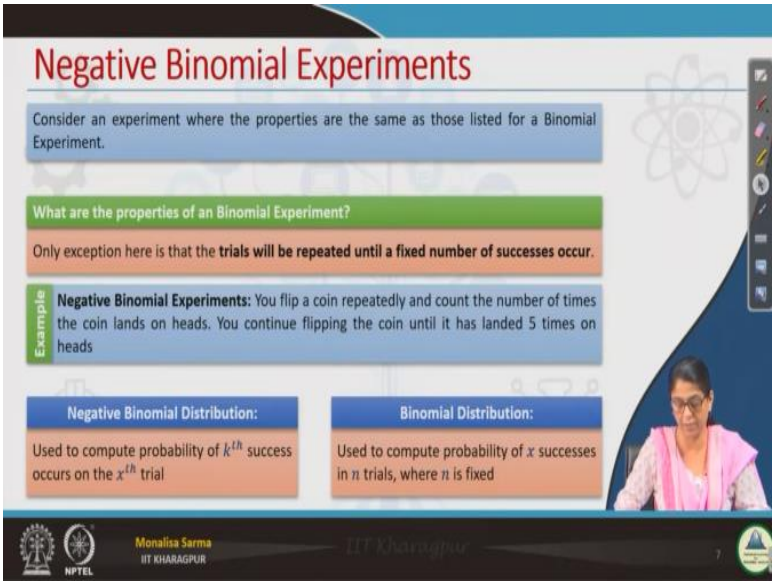
So, hello once again so in continuation to our earlier lecture on discrete probability distribution, where we have discussed binomial distribution, uniform distribution, binomial distribution, multinomial, and hypergeometric distribution.

(Refer Slide Time: 00:42)



In today's lecture, we will be discussing few other discrete probability distribution. So what are those we will be discussing negative binomial distribution, geometric distribution, Poisson distribution. And also finally we will discuss the descriptive measures of all these distributions.

(Refer Slide Time: 00:54)



So, first is what is negative binomial distribution? In negative binomial distribution if you guys can remember binomial distribution what was that our number of trials are fixed that was n we have specified it by small letter n remember that was the number of trials or number of trials are fixed in this number of trials, we our random variable was the number of successes.

So, variable is something which changes so, we have taken the random variable as the number of successes. So in n trial, whether the number of success is 0 success, 1 success, and it can go for X can take value from 0 to n, that was our random variable is the number of successes that was changing, but variance in some other experiment, but we get to see is that where the trials are repeated till we get a fixed number of success, meaning here.

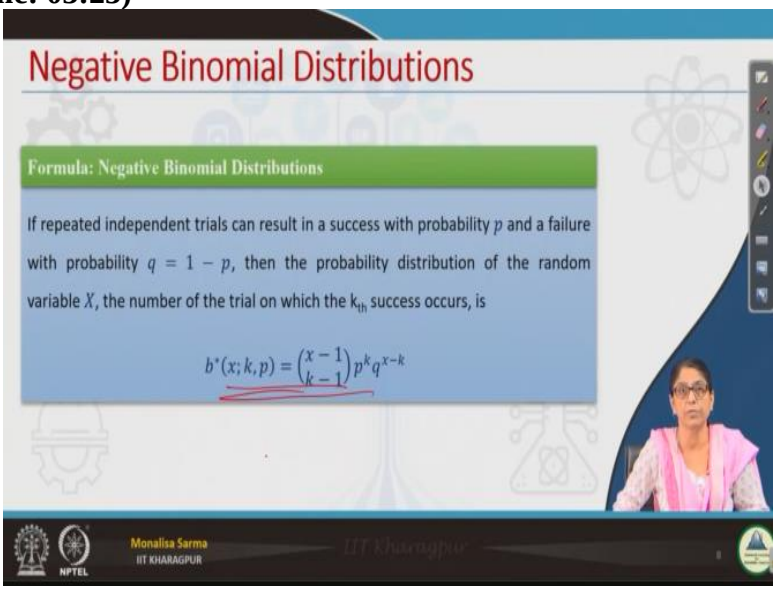
Now number of trials is not fixed that is number trials is not n the number of tails can be anything, but we are interested in getting a fixed number of successes that means suppose we are interested in getting total 5 success. So, our number of successes cannot be a random variable. So number of success is fixed here, is not it? It cannot be random variable now what is the random variable here, our random variable is the number of trials.

So now my X is number of trials, rather number of successes, number of success is fixed that is the negative binomial distribution, that is the only difference between binomial and negative binomial, there is not a very small difference, but it is a big difference, but that is the only difference between binomial and negative binomial distributions. So, negative binomial distribution and example you flip a coin repeatedly and count the number of times the coin

lands on head, you continue flip the flipping the coin until it has landed 5 times on the head, there continues my interest is I should get 5 heads.

So I am continually prepping it till I get 5 heads when I got 5 heads I stopped it. So, my number of trials is how many trials I needed to get 5 heads so that is my random variable. So, negative binomial distribution used to compute probability of the kth success occurs on the xth trial k is fixed x is a random variable, whereas in binomial distribution we use to compute probability of x success in n trial, where n is fixed. So, that is the only difference between the negative binomial and binomial distribution.

(Refer Slide Time: 03:23)

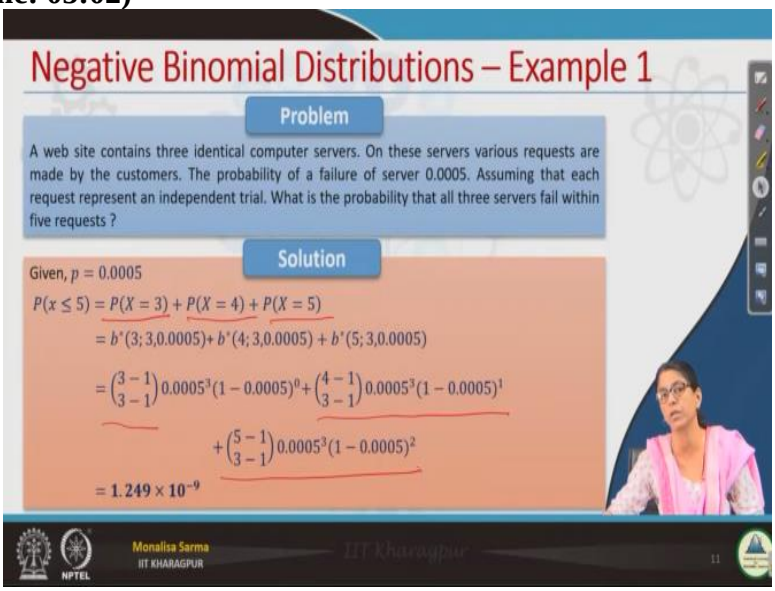


So, similarly, our definitely if this is the only difference. So our formula for calculating fx will also be same almost same the only difference here you will see since we are interested in finding out kth success depends on my last trial, last trial will be the k number of success. So trial before to that, if my last trial is x, trial before to that, that means an x - 1 in x - 1 I will get k - 1 success this can be in any pattern as we have seen in binomial distribution.

How it can be in different pattern success success, failure, failure, failure, failure, success success I have shown 1 example if you can remember, if you do not remember please go again go back to the lecture again, you will see that I have worked in a blue board showing you how we have done so similarly here, that means on the xth trial, I got my kth success. That means the all the trials before these xth, before this point is the x - 1 and this x - 1 trials, how many success I got? I got k - 1 success it can be in any order.

So that is why it is x - 1 C k - 1. And total how many successes are there? Definitely I am looking for k success. So total that will be k p^k. And how many non-success that is 1 - p that is q, q is x - q, q^{x-q}. This is the formula for binomial distribution negative binomial distribution.

(Refer Slide Time: 05:02)



So we will see an example a website contains 3 identical computer servers on this servers various requests are made by the customer the probability of a failure of server is 0.0005. The probability of there is a probability failure assuming that each request represents an independent trial here also trials are independent like what we have seen in binomial distribution, what is the probability that all 3 servers fails within 5 requests?

See an interesting question, but we have to find out? We have to find a probability that all 3 servers fail within 5 requests. So, what is this means? Our what to say we are interested in finding out, our number of failures are fixed, what is the number of failures? Number of failures is a random variable here. So, we are interested in finding out x is that is my, what to say my number of failures.

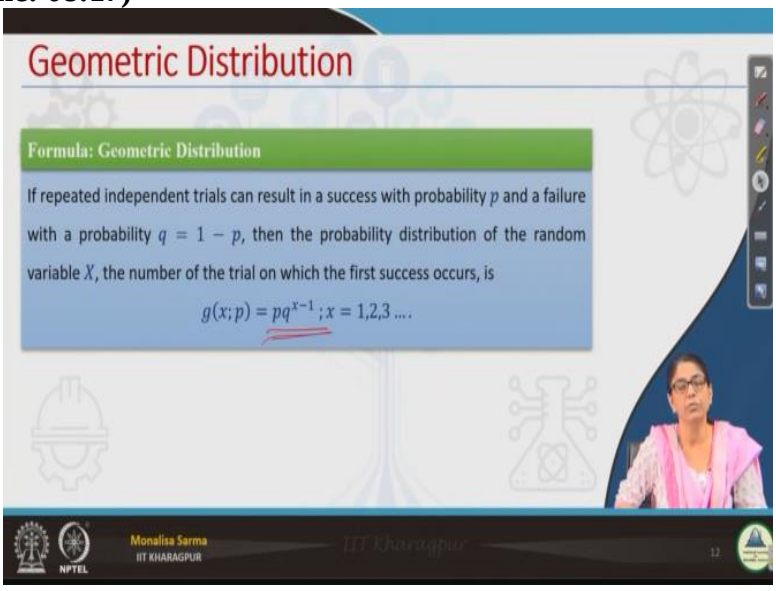
It is what it is given one thing assuming that each request representative, what is the probability that all 3 servers fail within 5 requests? So, within before 5 my all 3 servers would fail before or on 5. So that means I am in my number of failures are fixed that is 3 failures and how many trials maximum my trials will be 5. So, x is the so, my 3 servers may fail in that third request.

My third failure I may get in a third request my third failure I may get in a 4th request? My third failure I am getting a 5th request whenever you get the third failure stop it. So my question was what is the probability that all 3 servers fail within 5 requests? So it has to fail within 5 requests within 5 requests means it may fail within 3rd requests it may fail within 4th requests in may fail within 5 requests, is not it?

It definitely it cannot be less than 3 because then 1 request there can be at the most 1 failure. So, it might my third failure may be in the third requests. So that is p (x) = 3, my third failure maybe in the 4th request my third failure maybe in the 5th request. So x is the number of request that is the random variable. So, what is the formula what we have seen, so, this putting it in the formula, that is all nothing else, just if we put it there, how do you put it in a formula if we might 3rd request in a 3rd fail, so, that is x - 1 / k - 1.

So x - 1 / k - 1 x p ^{k} 1 - p how much? k - x so, here it is 0. Similarly, my third failure in the 4th request, this is in the 5th request, so that is negative binomial distribution.

(Refer Slide Time: 08:17)



So, now, this another type of distribution which is also in the same angle, same direction, I should not say angle same direction that is geometric distribution in geometric distribution, what we are interested? We are interested in finding out the number of trials we need to get the first success. Here also, my success is fixed success is, 1 just 1 in negative binomial my success can be anything but the success was fixed.

Here also my success is fixed and that is only 1. I am interested in finding out how many trials do I need to get the first success? The applicability, first let me tell you the applicability of this type of geometric distribution. Suppose, you are installing some company is installing some telephone and connectivity. And maybe we are interested in finding out when we dial in number after how many dial we get the how the line gets connected.

For dialling how many times on my line gets connected? I have tried it once the line did not connect, get connected in a busy time maybe I have tried it again it did not get connected, I tried it again it did not get connected. So suppose my line got connected on my 5th trial. That means the design there has to be something some problem with the design, no one will try wait for 5 times dialling and then getting the connection.

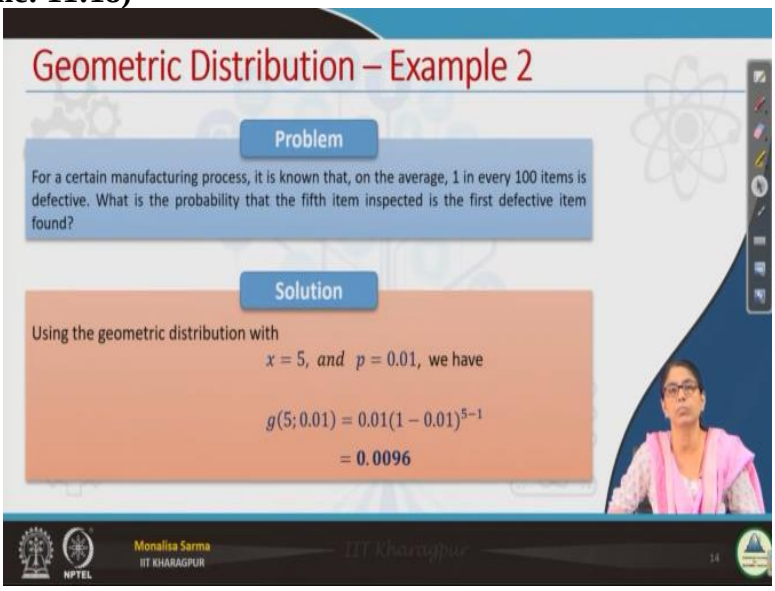
So, this sort of thing the geometric distribution is essential based on this people can find out so that there is problem in the line and accordingly steps can be taken. So there are many such cases where geometric distribution has great applicability. So my first success, here also the random variable is the number of trials remember always the random variable is that, that changes.

So, in any problem first you need to find out what is the random variable if you know and if you find out what is a random variable, then finding the probability will not be different than you will have to then finding out which it falls in which broader distribution, whether it falls in binomial, multinomial, hypergeometric, negative binomial, geometry what Poisson and whatever it is, first find out what is the random variable.

So, here random variable is the number of trials, and we were interested in number of trials we get the first success. So, the formula for this is quite simple, it is simple p (q) so, there are total x trials, is not it? So, in x trial, we will get 1 successor the last one is success, is not it? So, 1 success means that is p and in rest x - 1 that is the failure and failure is not distinguishable.

So, you do not have to do any convenience or something like that. So, it will be so, simple q to depart x - 1. So, simple this is the formula p into q x - 1. So, when x can take any value from 1, 2, 3 like means, we can get the first success in the first trial itself we can get the first x in the second trial. So, x can take anywhere from 1 to 3 so, that is geometric distribution.

(Refer Slide Time: 11:18)



So, problem for a certain manufacturing process it is known that on average 1 in every 100 item is defective what is that 1 in every 100 item is defective that means this question this particular statement what it gives? This particular statement gives us the probability of defective items what is the probability of defective items? 1 / 100 is the relative proportion called the defective item is 1 / 100.

So, what is the probability that a 5th item inspected is the first defective item found, this this reminds me we can also this sort of geometric distribution is also used in many kinds of drug survey when we are number before bringing a drug to the market, if we are interested in finding out suppose if a drug we should get for a drug to be to bring into the market, it should at least prove that it is 80% of the cases it is successful.

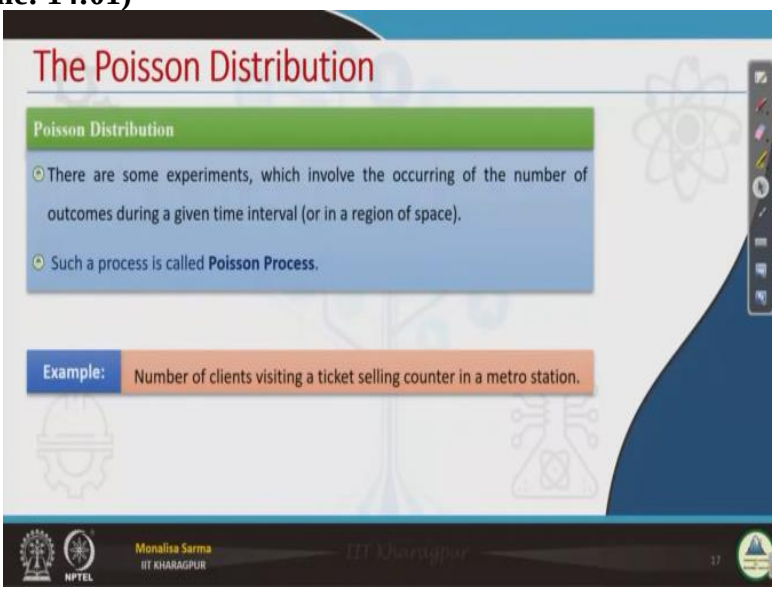
So, we will do go on doing the trial till we find that it is showing 80% successful now, this is people forgetting the 80% successful how many patients we have used this drug. So, if this number of persons is quite large to get 80% successful than this drug, there is some wrong with this drug, then it should not be brought to market we would need some more work on it. So, this geometric distribution has also it is used in that type of situation as well.

Now, coming back to this question, I just remember that example I thought of mentioning it now coming back to this example. So, here the defective rate is given for a certain manufacturing process it is known that on average 1 in every 100 item is defective. So, how much is defective it is given what is the probability that the 5th item inspected is the first

defective item found, we 5th item inspected first item we have inspected, it is not defective second we found it is not defective 5th item we found it is defective, what is that probability?

Probability is geometric distribution is simple formula p x q. There is a simple formula p into q, q^{x-1}, where x can take any value x can take any value 1, 2, 3 any value. So, we can find the first effective in the first trial itself. So, this is so here x = 5, p = 0.01, 1 in 10, 100. So, putting in the formula that is what that is the probability of getting 50 item inspected is the first defective found.

(Refer Slide Time: 14:01)



Now, there is one more this is the last probability distribution that we will be discussing under the umbrella of discrete quality distribution. So, Poisson distribution there are some experiments which involve the occurring of the number of outcomes during a given time interval here as in what to say, in contrary to binomial distribution here what we are number of outcomes in a given time interval, there we have trying in binomial distribution or its related negative geometry we are trying to find out the number of successes in so many trials here what is that which is occurring, the number of outcomes during a given interval. A very good example is toll where does what to say the toll collection way for highway we have this toll plaza is where we have to when the car passes, we have to give the toll value so we can say in a span of say 1 hour how many car has passed the toll or in a particular highway in a span of some few hours how many vehicle has passed.

So, or in a ticket counter in a particular time, how many people are there in the counter in a particular time. So, this is similarly suppose in a particular material and material a particular

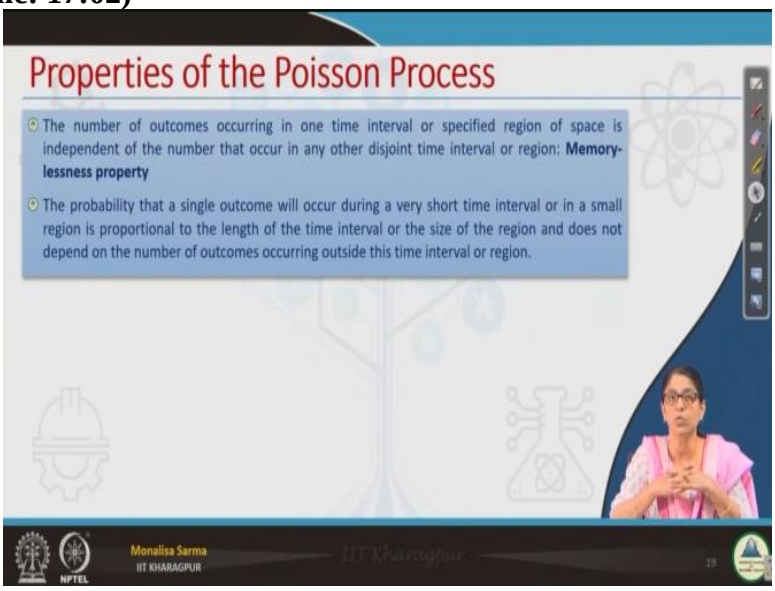
size 1 metre in a material of size 1 metre, how many flaws are there maybe flaws maybe in terms of design flaws, maybe in terms of density of the material whatever in a flaw material of 1 metre, how many flaws are there.

So, that is what which involves the occurring of the number of outcomes during a given time interval or in a region of space. So, such experiments are called Poisson experiment. So, example number of clients visiting ticket selling counter in a metro station as I already mentioned, like here one more thing like number of calls coming to a call centre at a particular time that is also a Poisson distribution.

But again there is something like if you talk about going to ticket counter ticket counter the people arriving and people in the queue for ticket counter and during early morning or during peak time it varies is not it? So, there are again different perception of Poisson distribution to model this where it is dependent on time with a number of people with the average number of people coming to a counter it is dependent on time.

So, the average number of events occurring in a particular time interval. So, if it is dependent on time, then it is not as simple Poisson distribution that we call it is as non homogeneous Poisson distribution, which we will be not be discussing this is a bit higher level. So, just for your knowledge, I just thought of mentioning it now, we are just considering the simple Poisson process.

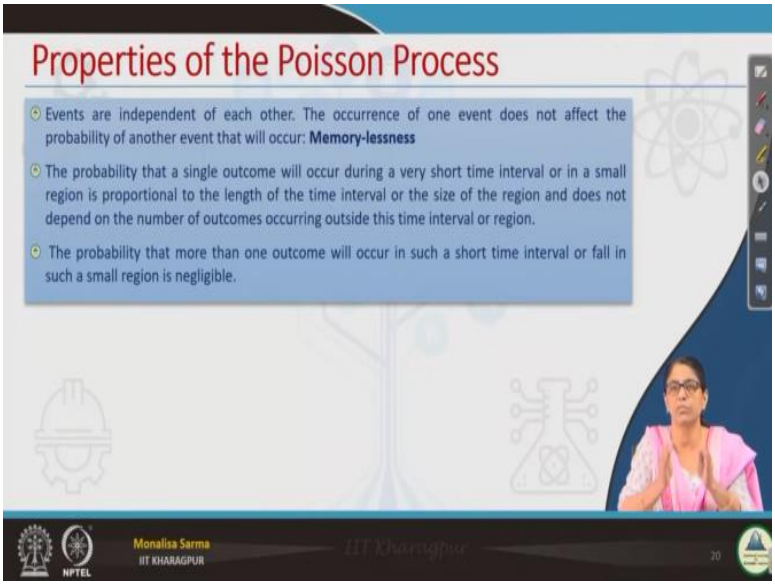
(Refer Slide Time: 17:02)



So, Poisson process had certain characteristics certain quality we can say. So, what are those? The number 1 the number of outcomes occurring in 1 time interval or specified region of space is independent of the number that occur in any other disjoint time interval or region what is it? The number of outcomes occurring in 1 time interval or specified region of space is independent of the number that occur in any other disjoint time interval or region.
In a highway if in 1 hour total 100 car has passed that does not mean in the next hour also 100 car will pass or that does not mean already 100 car has passed that means now, only 2 or 3 car will only pass that is not that it is not dependent on whatever what happened in one time interval that will not dictate what will happen in the other time interval. So this property basically this is as if it does not have memory.

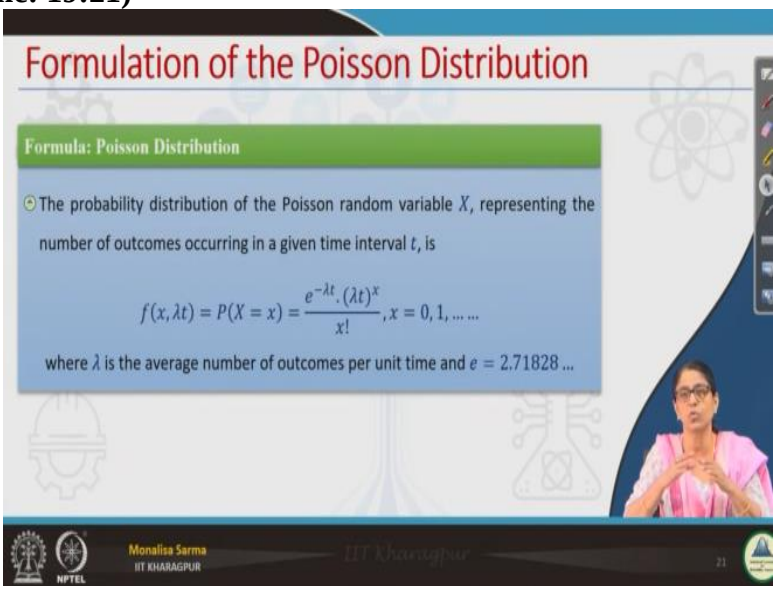
So, this property is called memory lessness property. So, again, I am repeating see number of outcomes occurring in one time interval is independent of the number that occurs in the previous time in time, so it does not have any memory. So, next property, the probability that a single outcome will occur during a short time interval or in a small region and in a very short time interval it is proportional to the length of the time interval.

If there is a probability that a single outcome will occur in this highway probability that a single car will pass it is what did that probability actually it is proportional to the length of the time interval there is no other thing that will dictate that it is not that because 1000 cars are passed in the last hour. So, just one 1 will pass it is no not that it is proportional to the length of the time interval or in case of space it is proportional to the size of the region and does not depend on the number of outcomes occurring outside this time interval or range or region.
(Refer Slide Time: 19:06)



The third property the probability that more than one outcome will occur in a very short time interval or fall in such a small region is negligible. If we consider a very small time interval the more than one outcome will occur is very negligible. So, these are the 3 properties of Poisson distribution.

(Refer Slide Time: 19:21)



Now, what is the function of Poisson distribution and Poisson distribution we have a parameter that is we call it lambda where lambda is the average number of outcomes per unit time, average number of cars passing through that highway in 1 hour. average number of cars passing the toll station in 1 hour or in 24 hours, whatever it is, so that is the parameter this is lambda.

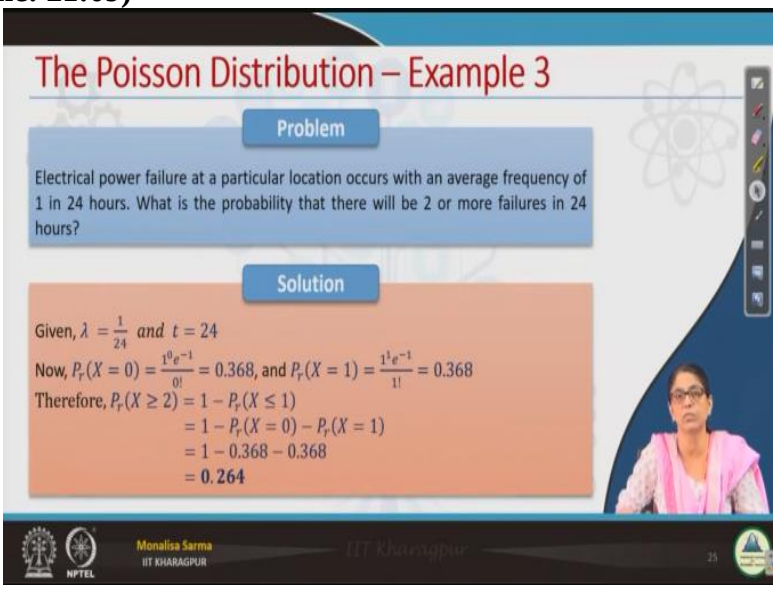
So x is the number of cars that pass the toll station or number of cars that pass the highway or number of flaws in the material of 1 metre, whatever it is, that is the X that is the random

variable, X is a random variable, λ is the average number of outcomes per unit time and t is the total time interval that we are considering. So, λ is given suppose in terms of 1 kilometre, and suppose we need to find out in terms of 10 kilometre.

So, that is t so, what is the formula for that $e^{-\lambda t} \lambda t^{x} / x!$ this also like how we have calculated for uniform probability distribution binomial whatever what we have seen we could usually can easily calculate it from our knowledge of probability theory is not it? What we have learned from probability theory binary permutation combination for finding out the sample space we could easily find out the function.

Similarly, for Poisson distribution also if we go step by step you can easily find out deduce this expression. But that is not necessarily if you are interested you can do it by yourself there is no issue, but this is the function for Poisson distribution f (x λt) is nothing but $e^{-\lambda t} \lambda t^{x} / x!$ these things are so, thing it is so, common it will be the tip of your tongue always once you learn it, you will never forget it.

(Refer Slide Time: 21:09)

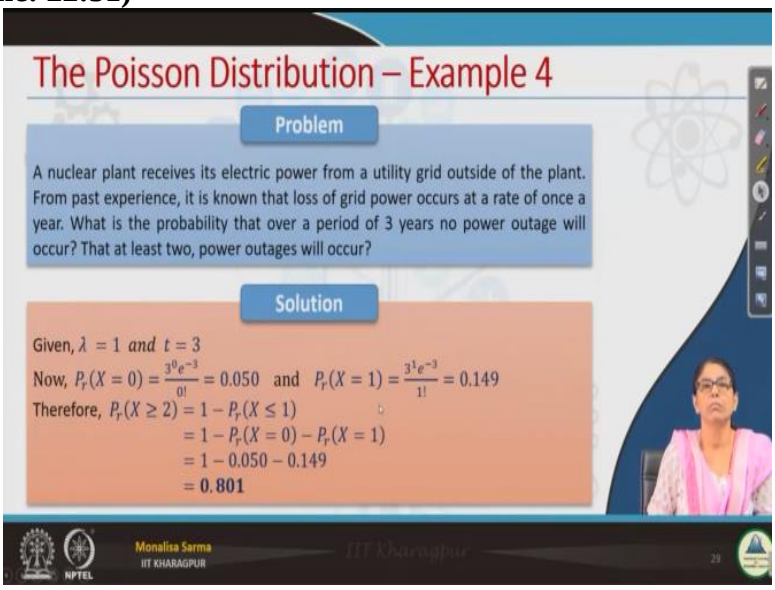


So, now, 1 small example, electrical power failure at a particular location occurs with an average frequency of 1 in 24 hours, what is the probability that there will be 2 or more failures in 24 hours? What it is given every frequency is 1 in 24 hours that means λ is 1 per 24 hours. Now, what we need to find out what is the probability that there will be 2 or more failure in 24 hours? My t is 24 hours and I have found in λ also 1 in 24 hours.

So my λt is just 1 λt is 1. Now my X is a random variable, what is the X? X is the number of failures, I am interested in finding out 2 or more failures that means, what is the probability of finding 2 or more failures? And it is basically same as finding out 1 minus of 0 failure + 1 failure, 0 failure, 1 failure means at most 1 failures, that is 0 + 1, if I subtract 1 minus of that, then I will get 2 or more failures.

Because or else I have to find out 2 failures, 3 failures, 4 failures, it might go to infinity, so that is difficult to find out instead of that I will take the easy path that is 1 - 0 + 1. So what is probability of X = 0? Putting in the formula $e^{-\lambda t} \lambda t^{x} / x!$. So, we found out X = 0, X = 1 we found that X = 1 so for X greater than 2 is 1 minus of these 2 values that is all.

(Refer Slide Time: 22:51)

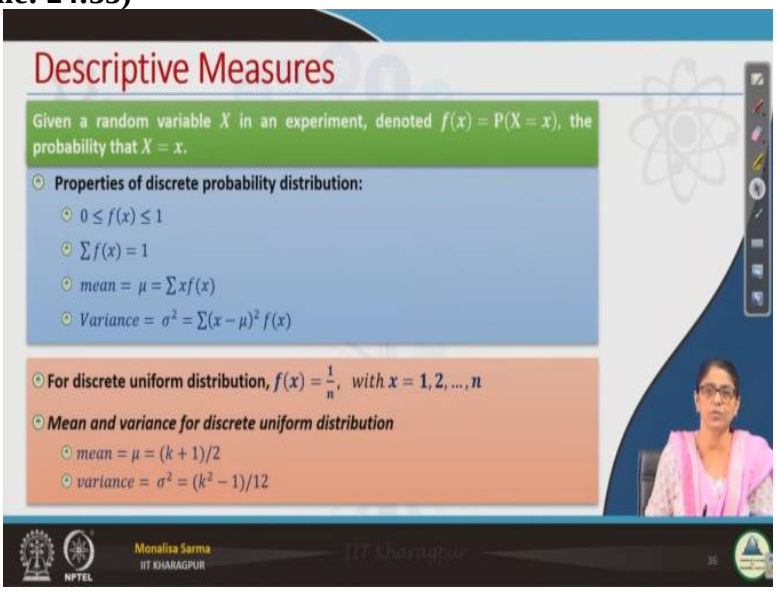


So, 1 more example here we have a nuclear plant receives this electric power from a utility grid outside of the plant. From past experience, it is known that loss of grid powers occurs at a rate of once a year. A nuclear plant receives its electric power from a utility grid outside of the plant, from past experience, it is known that loss of grid occurs at a rate of once a year. So that is λ average is once a year λis 1 per year.

What is the probability that over a period of 3 years no power outage will occur? Now my t is 3 years, λ is 1 per year my t is 3 years. I am interested in finding out what is the probability of a period this is no power test. That means I am interested in finding out probability of X = 0, given λ is 1 and t is 3 simple just put it in a formula λ is 1, t = 3, X = 0.

Now next question is at least 2 power outage will occur at least 2 that means it can 2, 3, 4 anything. So similarly now 1 - 0 + 1 that will give me at least 2 is not it? So that is why I am here see, we found out X = 0 we found out what is X = 1 so 1 - X = 0 - X = 1 will give me X >= 2.

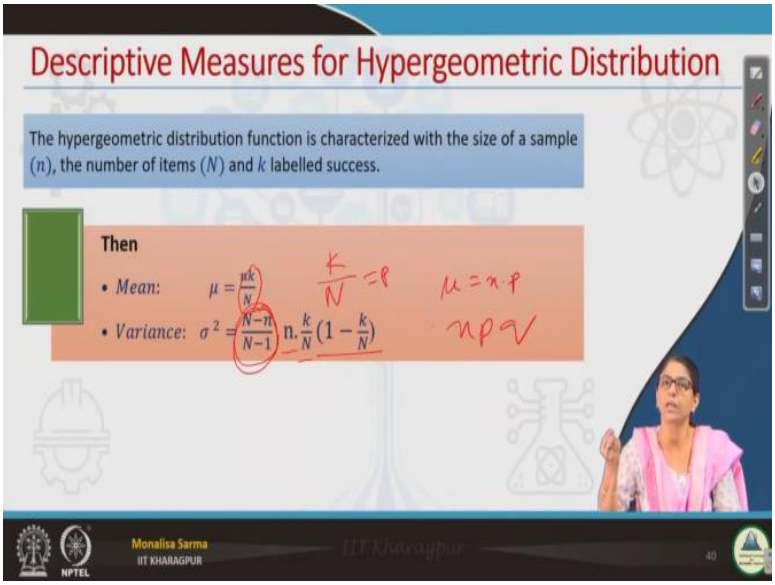
(Refer Slide Time: 24:33)



So, now to summarise the descriptive measures, but of all the probability distribution that we have seen, first properties of the discrete probability distribution what we have seen a value of fx lies from 0 to 1 that we have seen what is f(x)? f(x) is the probability at point x at value probability of X. Then summation of f(x) = 1 mean = x f(x) $\sum$ f(x) variances = $\sum (x - \mu)^2$ f(x) we have already seen all this now for discrete uniform distribution we have seen f(x) = 1 / n.

If the value that x can take is 1 / n, then if we find out μ= k + 1 / 2 variance is k^2 - 1 / 12 this putting into formula you will get it and again and again you do not have to put it in a formula and evaluate it you can just remember it that is this like a + b^2 we need to do it and solve it and do it we can remember a + b^2 = a^2 + 2 a b + b^2. Similarly, you can dissolve these things also it will be in your memory always.

(Refer Slide Time: 25:35)



Similarly, for binomial distribution, what is the mean? Mean is np σ variance is np q 1 - p is q. So, for hypergeometric distribution mean is I see here means hypergeometric we have also done 1 problem also mean is what? n x k / N what this k / N symbolises? k / N I am writing it here what is k / N? k / N is nothing but it will be out of n out of capital N k are success. So, what is this? This is p basically is not it?

So, n into so, μ is equal to basically n x p so, hyper geometric distribution mu and binomial distribution μ is same binomial distribution also μ is np here also μ is np because k / N is nothing but p. And in σ also same for binomial distribution what was my σ^2? σ^2 is npq here also you see this is n this is p this is q 1 - k / N is q, but we have just 1 extra term here that is n - n small n / N - 1.

Now, one thing now, if our size of N is very big, size of N is very big capital N and small n is very small, what happen like from a from a what to say, bucket of water if I bring out 1 glass of water does will it make any difference? It will not make any difference. It will be hardly it will make any difference from a bucket of water I am just bringing out one glass of water. But if in a small bowel where I am putting say 2 glasses of water or 1 glasses of water from there, if I pick up 1 glasses of water there, it will make difference.

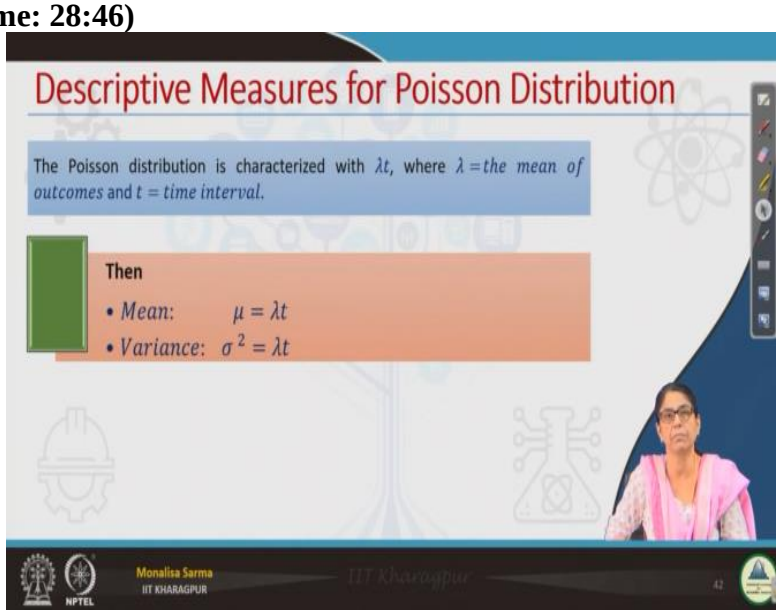
So, that is what here you can see in the hypergeometric distribution, this value N - n this value, n – N / N - 1 this will make a difference when my capital N is not very large and small n is also not very small, then only it will make a difference else it might capital N is very

large and small n is very small, that means N - n / N - 1 it will be almost equivalent to 1 only it will not make much of a difference.

Then it turns out to be same as binomial distribution, why I mentioned this for very large N, large capital N total size of the sample and very small n. So, instead of using hypergeometric distribution we can use binomial distribution as well because we have seen it hardly makes any difference μ is seen σ we have just another 1 extra factor here that is N - n / N - 1, but when N is large or small n is small it differences not reflected at all.

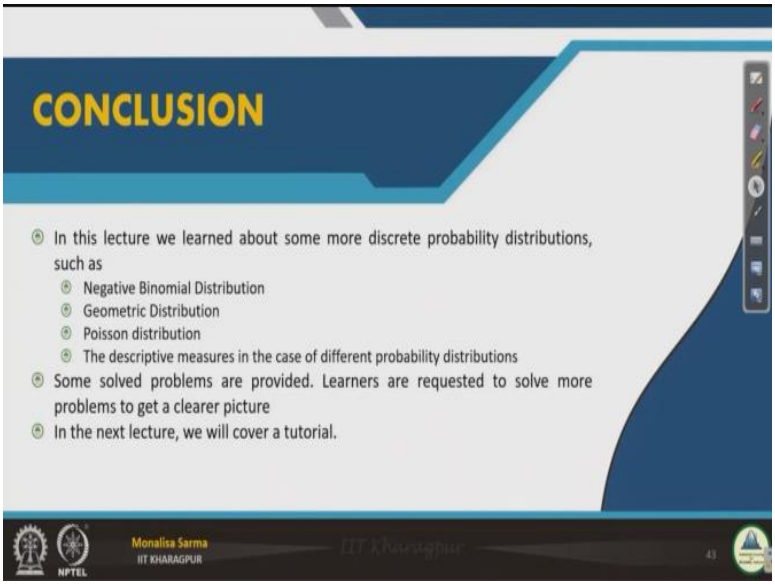
So, we can for very large N and very small value of the sample size that is a small n we can use instead of hypergeometric we can use binomial distribution as well, when the case is suitable for hypergeometric that means without replacement, but still we can use binomial that is an example from a bigger bucket of water bringing one glass will not make any difference.

(Refer Slide Time: 28:46)



So, similarly for Poisson distribution mean is λt and variance for positive is the mean and variance both are same that is λt.

(Refer Slide Time: 28:56)



So, now to conclude this lecture in this lecture, we have discussed negative binomial distribution, geometric distribution, Poisson distribution we have also seen the descriptive measure of different probability distribution. Again I request you to solve as many problems as possible, so, that you become confident in all these distributions. And in this next lecture, we will cover a tutorial on this discrete probability distribution and the references and thank you guys.

(Refer Slide Time: 29:20)

Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury of Quality and Reliability
Indian Institute of Technology - Kharagpur

Lecture – 11
Tutorial on Discrete Probability Distributions

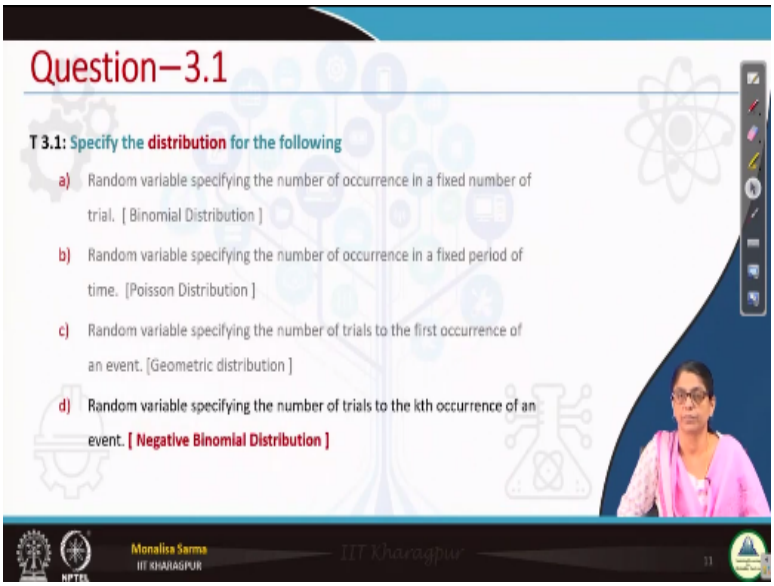
So, welcome once again. So, we have discussed discrete probability distribution, we have seen different type of discrete probability distribution uniform probability distribution, binomial probability distribution, multinomial than we have seen hyper geometric distribution, negative binomial geometric distribution and Poisson distribution, we have this discuss all this discrete probability distribution. Now, we will be doing a tutorial session on this distribution, the thing is here, I have problems catering to all the distribution.

But what I want is that before immediately jumping to the solution, because I have the solution also here, immediately jumping into the solution, what a request is that once you read the question, and then from the question, you try to find out what is the random variable here and it will fall in which distribution then once you are sure that it will fall in this distribution, then you can check for the answer check for a solution, that way your learning will be proper.

(Refer Slide Time: 01:20)

So, first we will be solving few objective type questions which will help you in understanding the concepts and then we will be solving some problems. So, first to discuss the objective type questions.

(Refer Slide Time: 01:33)



So, here I will give certain statements from this we need to find out what is the distribution for the following like first one is random variable, if my random variable specifies the number of occurrence in a fixed number of trials, random variable specifying the number of occurrences in a fixed number of trials what is that? This is binomial distribution, we have seen binomial distribution weather number of occurrences we are interested in finding out the number of successes and our trials are fixed.

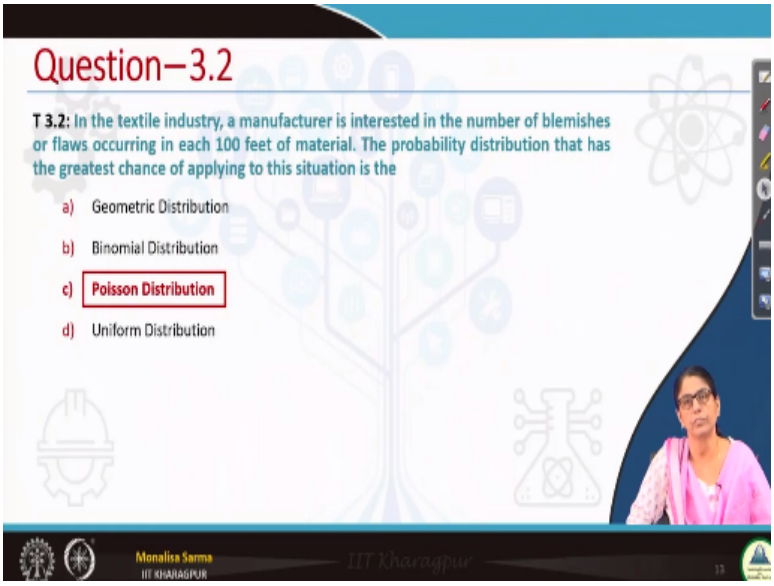
Which are x, which is actually the trials are independent as well, is not it? So, this is a binomial distribution. Now, the second is random variable specifying the number of occurrences in a fixed period of time and a fixed period of time, what are the number of total number of occurrences probability of a number of 0 occurrence, probability of number 1 occurrence in a fixed period of time, or in a fixed region of space we have seen that is not it? What is that? That is your Poisson distribution.

The next random variable specifying the number of trials to the first occurrence of an event, so here the number of trials are not fixed, but what is fixed, we want to find out the random variable

specifying it specifies the number of trials and we want to find out in what number of trials we will get the first occurrence that is maybe the first success or maybe the first failure, what is that that is geometric distribution.

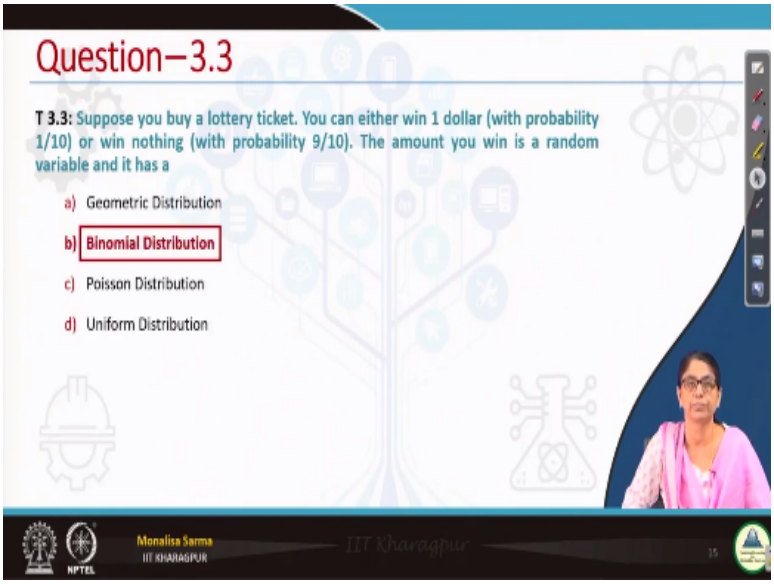
Next random variable specifying the number of trials to kth occurrence of an event. Here also my random variables is a number of trials and what I am interested in finding out the kth occurrence of an event, maybe the kth success of an event, kth failure of an event so, what is that that is negative binomial distribution.

(Refer Slide Time: 03:23)



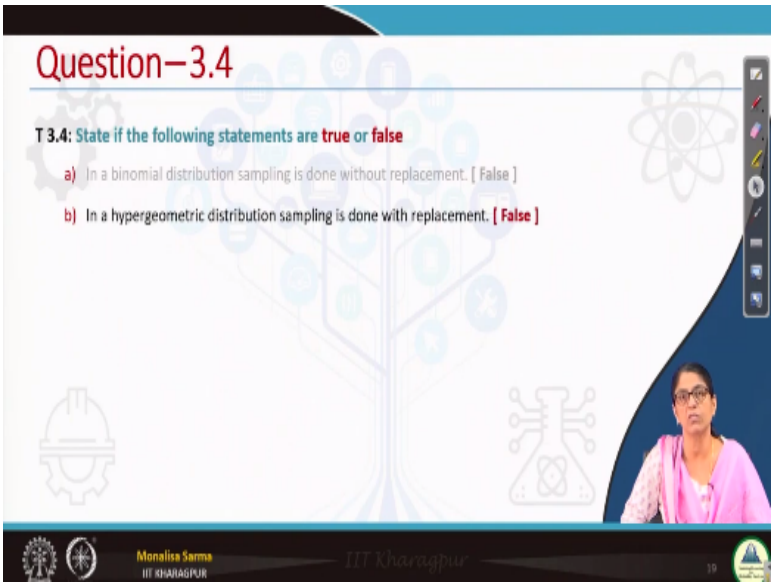
Again, another one question in a textile industry, a manufacturer is interested in the number of blemishes or flaws occurring in each 100 feet of material, the probability distribution that has the greatest chance of applying to this distribution. So, here it is number of blemishes of flaws in a particular length of material. So, what is that this is undoubtedly Poisson distribution.

(Refer Slide Time: 03:52)



Then suppose, you buy a lottery ticket, you can either win 1 dollar with probability of 1 / 10 or win nothing means you can win or you can lose that is where you are not winning is lose means not you did not win anything, you can say the success or a failure with probability 9 / 10. The amount you win is a random variable and it has a what distribution? That is a binomial distribution.

(Refer Slide Time: 04:20)

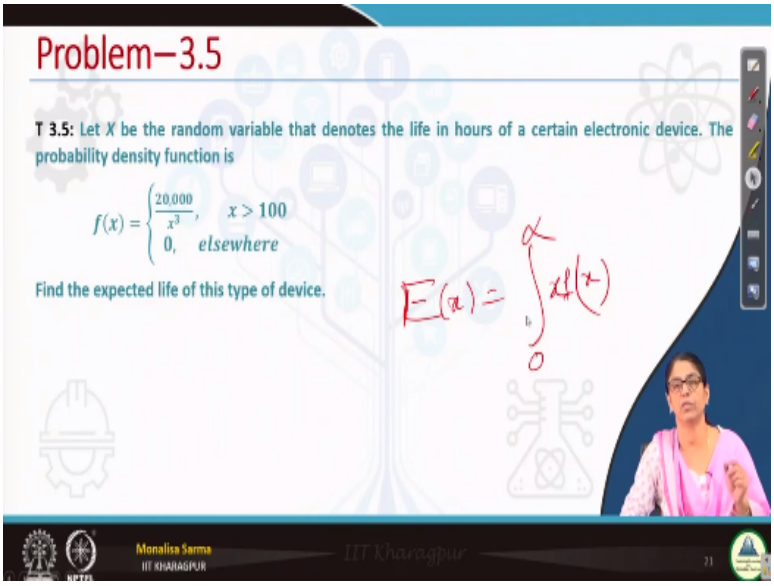


Then some statements to find out whether it is true or false. In a binomial distribution sampling is done without replacement. In a binomial distribution, do we do the sampling without replacement? No, we do the sampling with replacement in a deck of cards, we pick a card and we

put it back. So that if we do without replacement, then it will be dependent it will not be trials will not be independent, so it is false. In a hypergeometric distribution sampling is done with replacement.

On the contrary, in hypergeometric distribution sampling is done without replacement. So, this is again false. But those were very simple objective type questions. Now, we will come to some problems.

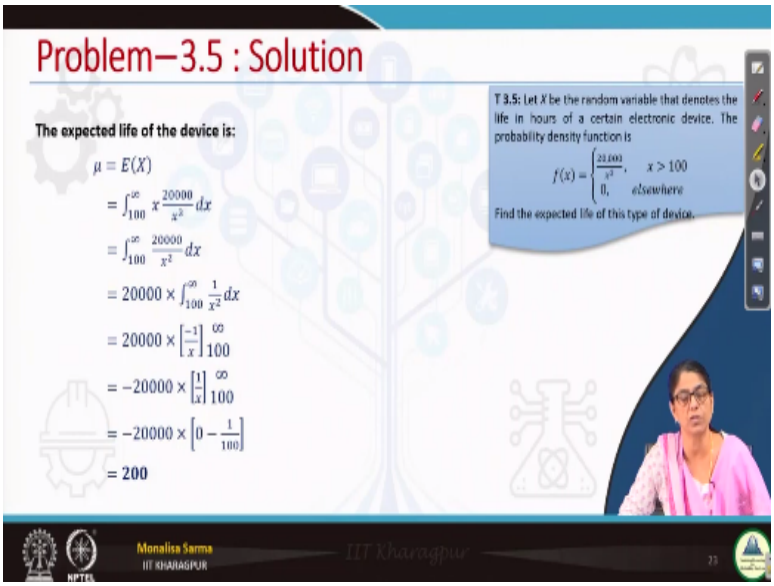
(Refer Slide Time: 05:06)



So, first question, if X is the random variable that denotes the life in hours of a certain electronic device, the probability density function, the probability density function is given here that is f(x) in a particular range where x > 100 it is given 20000 / x^3 and in other portion this means for values of x <= 100 it is 0, we have to find the expected life of this type of device. What is the formula for expected value you remember formula for expected value is E(x) = ∑ ∫ xf(x).

Since this is a continuous random variable, continuous random variable we use integration for discrete we use summation. So, e (x) is x / f(x), x f(x) so, it will be from 0 to infinity or you can say it from -∞ to +∞ as well. So, it is basically we do not take 0 it is from -∞ to +∞.

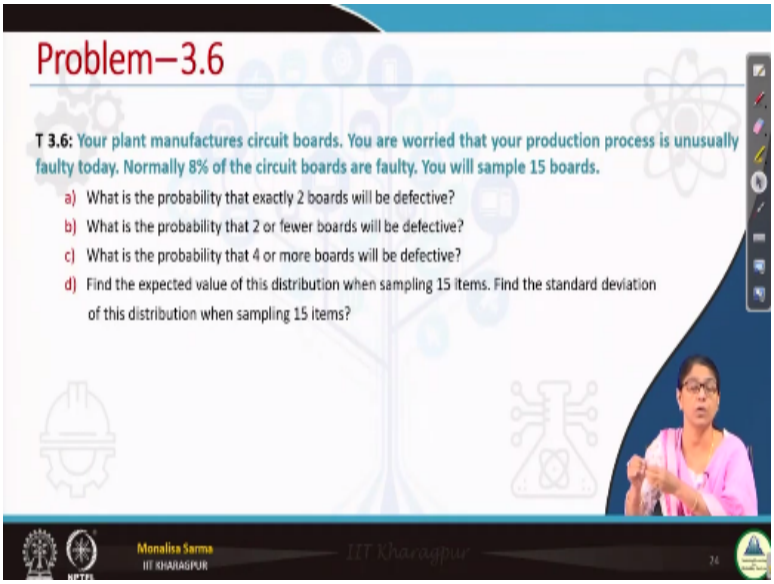
(Refer Slide Time: 06:20)



So, how do we find out E of x is here why we have specified from 100 to ∞ because an elsewhere it is given 0 so, we know that will be 0 only y instead of taking this we are taking it from 100 above x > 100 it is given so, we are taking from 100 to ∞, you can write it from 0 to ∞ also no issue but from 0 to ∞ from 0 to 100 it will be 0. So, and then it is simple question of simple integrating it.

As I already told you one of the class you will have to brush up your knowledge on integration and differentiation for different problems that we will be encountering in this course, we will not go for very complicated integration, but simple integration you will definitely need it also integration by parts also we will need it, so please brush up your knowledge on that.

(Refer Slide Time: 07:13)

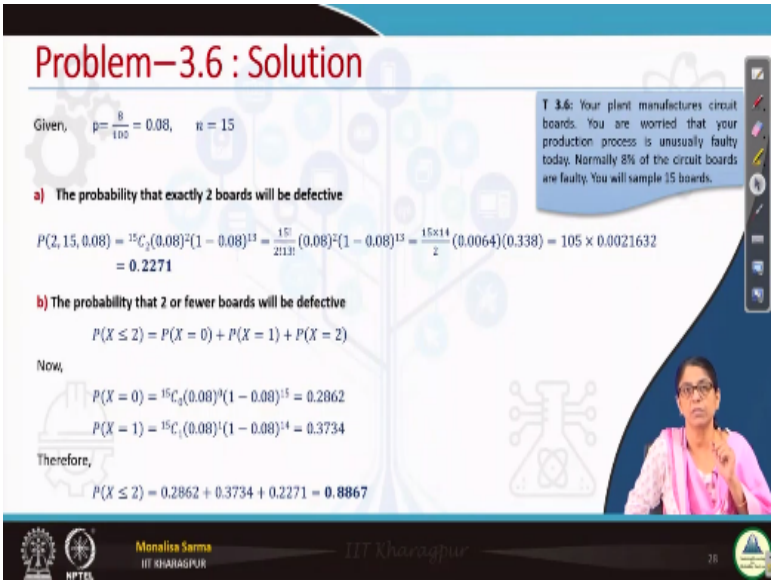


So, that is all now, next question. The earlier question is the simple question of continuous probability distribution where we are trying to find out the expected value. Now see this vision your plant manufactures circuit boards you are worried that your production process is unusually faulty today normally 8% of the circuit boards are faulty you will sample 15 boards. 8% of the circuit boards are faulty that means my probability of faulty is given, probability of fault is 8 / 100.

So, and whatever we will find out, we will sample total 15 boards now, the question is what is the probability that exactly 2 boards will be defective? Now, first one, this is a question of which probability distribution? It is definitely a question of binomial probability distribution, you are just what you are doing in a production line, you are just finding out one whether it is faulty or not, you are keeping it back taking one keeping it back. So, it is not that we are taking it and you are doing it finding it out till destruction, and then you cannot keep it back.

So, it is not that sort of sampling process, you are just picking a circuit board and you are testing it on a probably again picking another testing and putting it back. So, it is definitely a question of binomial probability distribution.

(Refer Slide Time: 08:31)

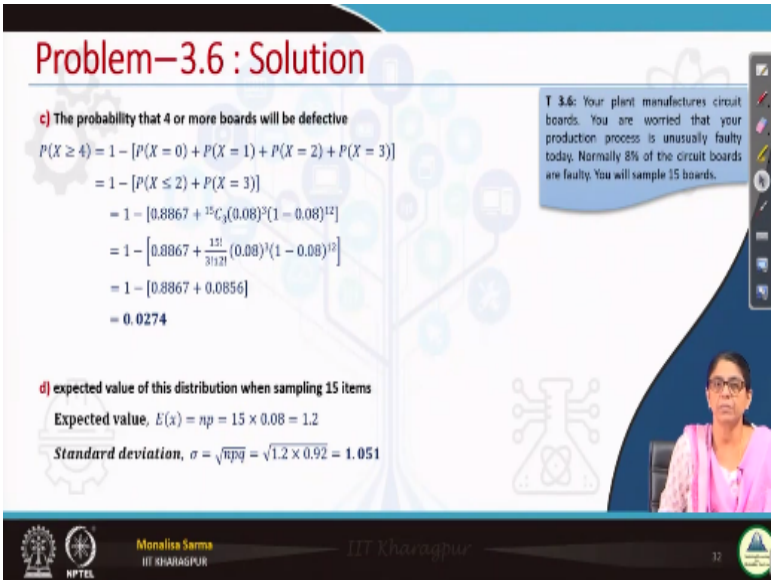


So, probability of exactly 2 boards it is what is my n? n is 15 so 15 C 2 we have this formula just put it in a formula that is all, once you know it is binomial distribution, you know the values of the parameter you know what is n, what is p, what is x then just putting it in a formula. Similarly, the next question probability that 2 or fewer boards will be defective 2 or fewer that means, we have to find 0 defective, 1 defective and 2 defective.

So, for 0 defective, 1 defective and 2 defective in fact, instead of doing it manually, you can consult the table also which I have discussed in the class that is binomial distribution table because binomial distribution table give us the cumulative value here we are interested in this question problem we are interested in a cumulative less than equals to 2 means from 0 to 2. Here I have solved it without solving the consulting table, but you can very well consult a table in the table if this particular probability value is given.

But, in a table there it has some limited number of probability values actually all this probability it can take any value. So, table it does not contain any probability values. If first you find out whether it is probability concerning probability value equals to 0.08 if it is there, then very well write it on the table.

(Refer Slide Time: 09:57)

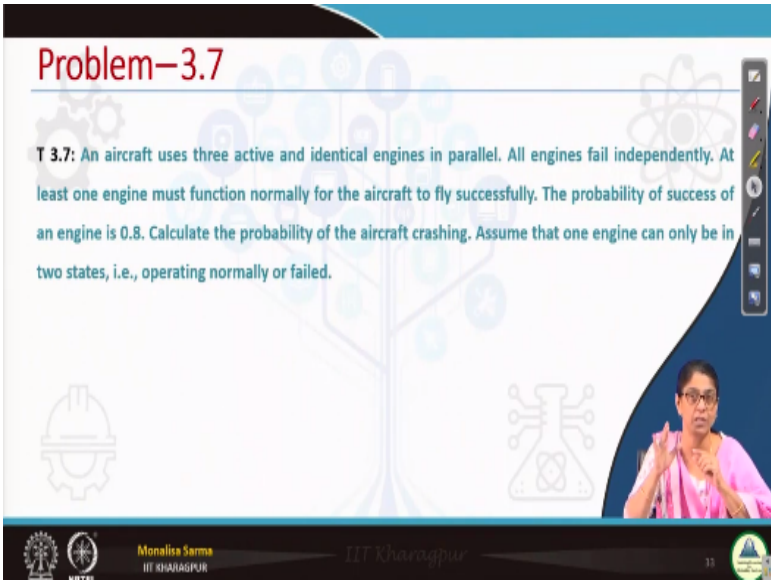


Then next question, what is the probability that 4 or more boards will be defective 4 or more that means probability that 4 will be defective total there are 15 boards, 4 or more will be defective means 4 may be defective 4 + 5 + 6 + up to it will go to 15. So, it will be well if it is in a table we can find out but if it is not at the table, then it will be very difficult to find out from 4 to 15. And the table also we do not get it we get it from 0 to 4, we will not get it the other way around. So, one way is that finding it out 1 minus of 0 to 4.

It greater equal to 4 that means 1 minus of 0 to 3 so that means 0 defective, 1 defective 2 defective and 3 defective 1 minus of that will give us the data is equal to 4. You will see it here, probability that 1 minus probability that x = 0, probability of x = 1, probability of x = 2, probability of x = 3 and probability of x = 0 and x = 1 these 2 we have found out in our last question. So, we found out this value that is where we will not solve it again for x = 2 and 3 we will find it out.

Sorry we find it out till this portion on the 3 is remaining with 3 we will find it out and we will get the value. Then expected value binomial probability distribution what is the expected value? Expected value is n x p what is n here? n is 15, p is what? p is 0.08. Standard deviation is √ np, variances npq, mean standard deviation is √ npq.

(Refer Slide Time: 11:36)



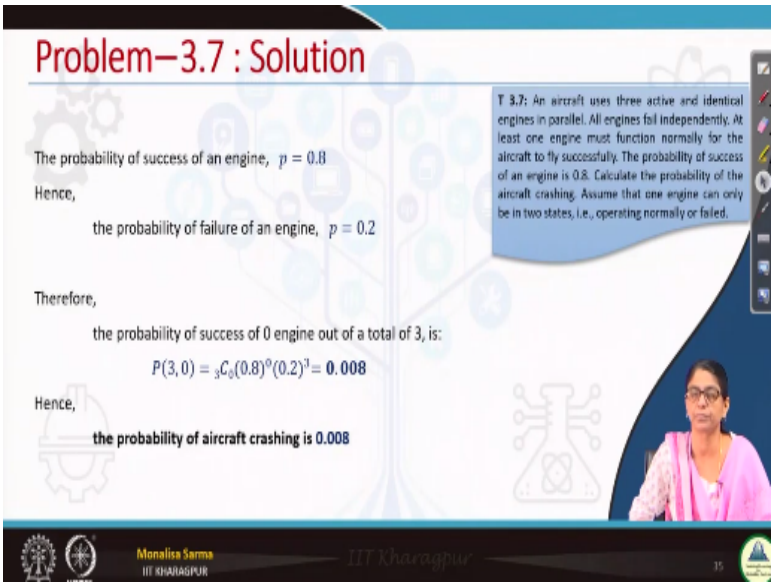
Now, one more question an aircraft uses 3 active and identical engineers parallel, all engine can fail independently it uses 3 identical engine at least 1 engineers function normally for the aircraft to fly successfully. It usually happens in a normal case in aircraft we usually have 3 different machines and at least 1 machine should perform so that aircraft will fly normally. So, at least once a function the probability of success of an engine is 0.8.

Calculate the probability of the aircraft crashing, when the aircraft will crash when there is not a single machine is successfully operating if all the machines are failing then the aircraft will crash even if 1 machine is working this aircraft will not crash. So, when we have to find out probability of the aircraft crashing that means we have to find a probability of 0 success, is not it? When there is not a single successful machine then it will crash so we will have to find a probability of 0 success.

Assume that 1 engine can only be in order of 2 states operating normally or fail. Why this question is why this point is given assume that engine can only be in 2 states operating normally or fail there are some machines where there are 3 different states that you do not need to know at this point just the 3 different states means failing as 1 state and working maybe through 2 different like some like circuit all of you may know circuit short circuit failure, open circuit failure 2 different failure modes 1 success and 2 different failure modes.

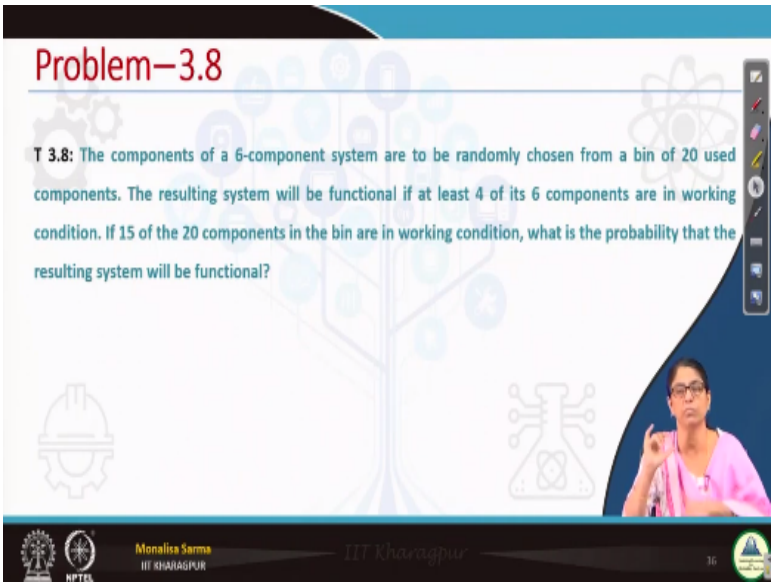
So that is why this particular statement is given. So, it is it has 2 states operating or fail but that is not very that is redundant information for description for your knowledge I have just mentioned it.

(Refer Slide Time: 13:23)



So, now how to find out a probability of success is given probability of success is 0.8 we have to find out 0 success in a total of 3, simple it is simply a case of binomial distribution very simple case of binomial distribution.

(Refer Slide Time: 13:38)



So, now distribution those components of 6, 6 component system are to be randomly chosen from a bin of 20 used component. 6 component system we have to choose randomly from a bin of 20 components the resulting system will be functional if at least 4 of its 6 components are in working condition. From this 20 we have to form a 6 components system there is a system which is consisting of 6 component system we have to pick from this 20.

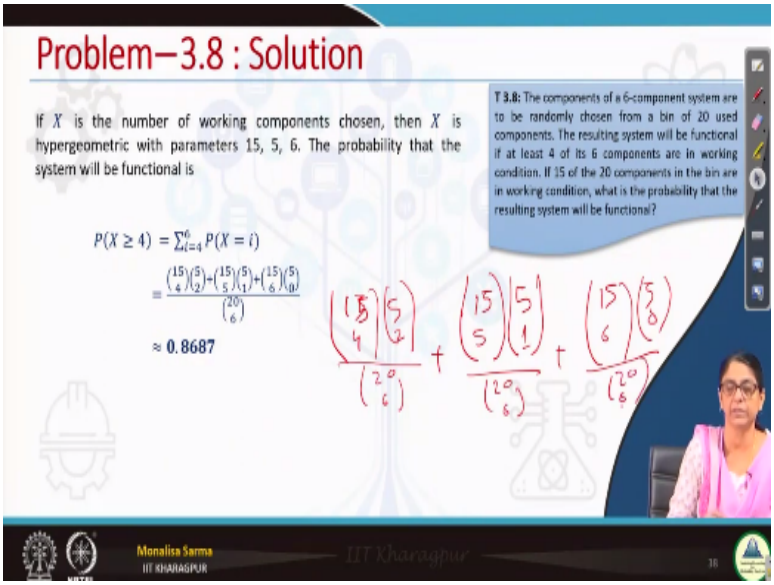
And form gradually will pick 1, will pick the second one, will pick the third one and that way we have to form a 6 component system. That directly implies we are bringing a component means we are not putting it back there. We are not replacing it back isnt that directly implied, because we are bringing a component and we are making the system and this will take this whole system and take 6 components one by one we bought 6 components. So, directly implies that we are not replacing it back.

So now next, the resulting system will be functional if at least 4 of its 6 components are in working condition for it to function, at least 4 of the 6 components should be in working condition. If 15 of the 20 components in a bin are in working condition in the bin there are 20 components out of this 20 we know that 15 are in working condition, 5 bad components are there. In a lot of 20 there are 5 bad components are there 15 good components are there.

And for the system that I have made the 6 component system picking component from the 20 components, I should get at least 4 working component than only my system will work. Now, what is the probability that the resulting system will be functional? Resulting system will be functional if I pick 4 good components, if I pick 5 good components, if I pick 6 good components, then my resulting system will work.

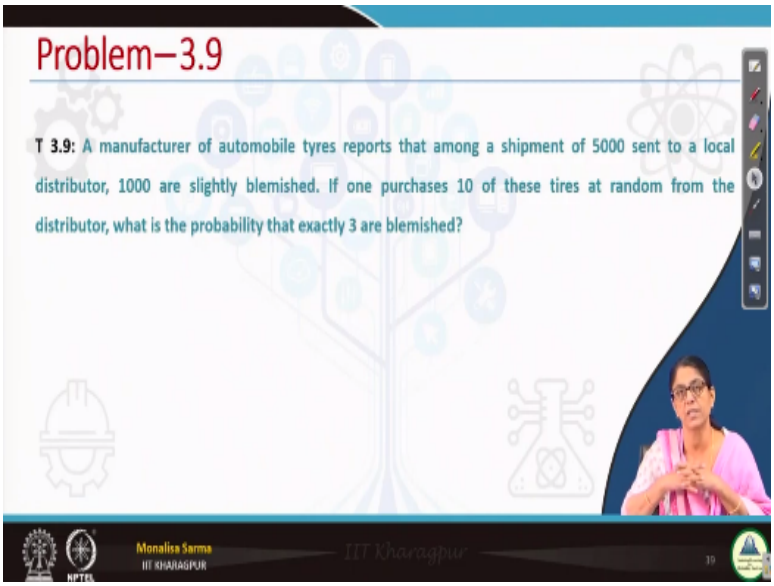
So, I will have to take these 3 scenarios picking 4 good components, picking 4 good components, picking 5 good components and picking 6 good components then only my system will work. So, definitely it is a question of hypergeometric distribution we are not putting it back and we have 2 different categories one is 20 components are working condition out of 20, 20 components out of these 20 15 are in working condition and is from 20 how many we have to pick we have to pick 6.

(Refer Slide Time: 15:56)



So, my sample space is 20 / 6 here I have written it together. So, basically if I write it separately it is 15 4 5 of 2 from 15 I got from 15 good ones I got 4 total I got 6 from rest 5 I got 2 and this is 20 of 6 + from 15 good got component and component I got 5 from the remaining 5 I got just 1 20 6 plus from 15 I got 6 then from remaining 5 I got 0 divided by 20 6 this I am writing it in this form here basically.

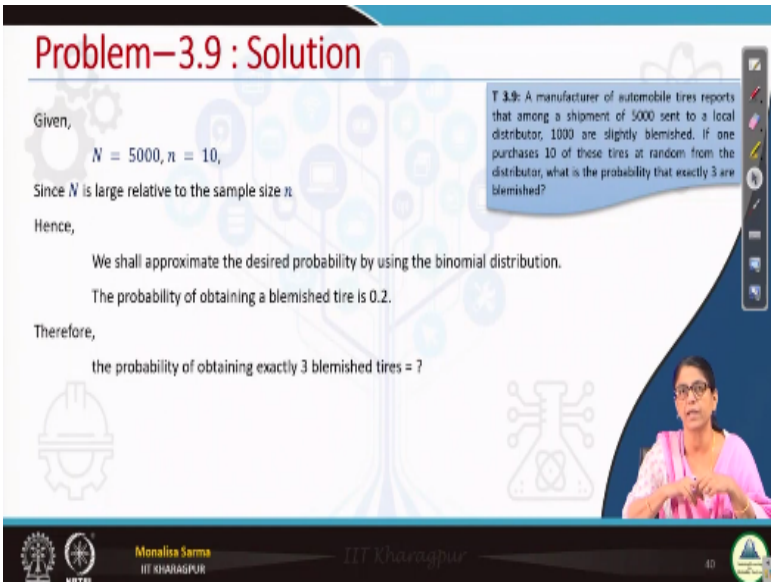
(Refer Slide Time: 16:54)



So, a manufacture automobile tyres report that among a shipment of 5000 sent to a local distributor 1000 are slightly blemished. Out of a shipment of 5000 1000 are blemished, what is

the probability a blemish that means 1000 / 5000 1 by 5 if 1 purchased 10 of these tires at random from the distributor, what is the probability that exactly 3 are blemished? Fine if one purchased 10 of these tires at random, from the distributor what is the probability that exactly 3 are blemished?

(Refer Slide Time: 17:37)

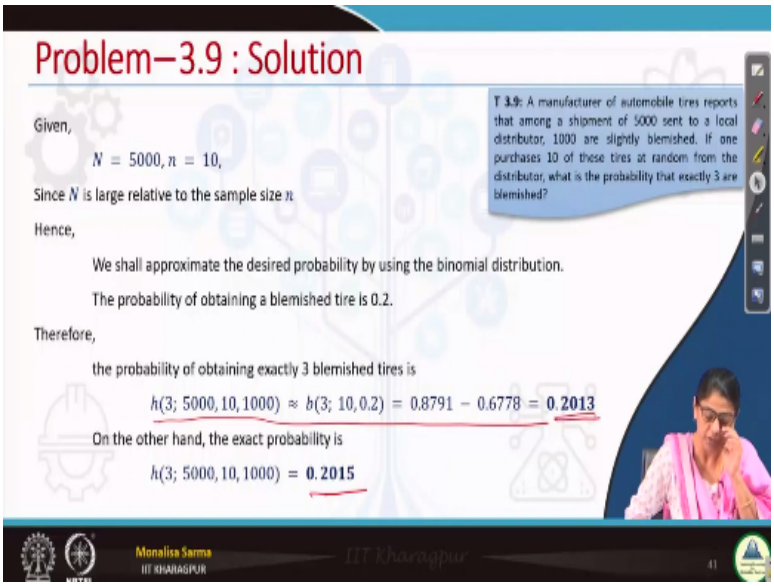


So, can you guess what is this? This is also a simple case of hypergeometric distribution, is not it?

If 1 purchase 10 of this tire, so I am purchasing 1 tire, and I am not giving it back again, is not it? Again I am purchasing another tire I am not giving it back again. So, if I am purchasing 10 tires out of 10 tires what is the probability that exactly 3 are blemished? And I know from this 5000 1000 are blemished total, there are 5000 tires out of them 1000 are blemishes.

And I am purchasing 10 one out of this 10 what is the probability that exactly 3 are blemished. So, it is directly a question of hypergeometric distribution.

(Refer Slide Time: 18:22)



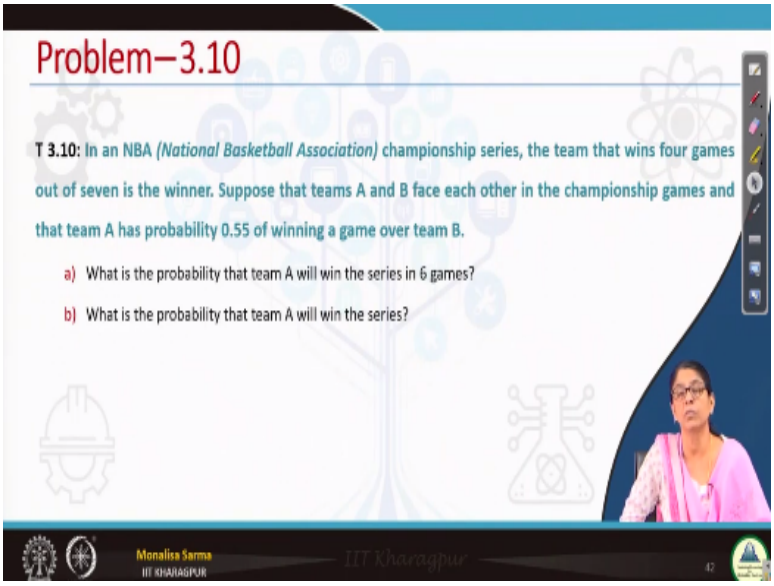
But here there is a twist here, see here; I can very well use the hypergeometric distribution here. But since N is very large relative to the sample size, as I told you from a bucket of water, if I pick one glass of water, it will hardly make it dangerous, N is very large what is N? N is 5000. And what I am picking out my sample size is only 10 out of 5000 I am picking out 10. From 5000 I am picking out 10 if I am not replacing it back also it will hardly make any difference it will not make much of a difference.

When N is very large relative to the sample size that is small and then instead of hypergeometric distribution, we can also approximate it using binomial distribution that I have already mentioned. But now the question is why we will do, when hyper it is fitting hypergeometric distribution why we will unnecessarily use binomial distribution. The main reason behind is that hypergeometric distribution the calculation is a bit tough, is not it?

We have to do from numerator and denominator the calculation is a bit complicated compared to the binomial distribution. And hypergeometric distribution you we do not have a table. So, instead of hypergeometric if we can approximate it binomial distribution, we can also consult the table and our calculation is also easier. So, if it satisfies this condition and is quite large compared to the sample size, we can approximate binomial distribution.

So, this is the question see, if we approximate if we do solve it by hypergeometric distribution, this is the thing we will get this value 0.2013 you can solve it and see. Now same thing, if we have used binomial distribution from by using binomial distribution we got this, but if we solve it by hypergeometric distribution we will get this value almost same slightly different. So, you can very well approximate by binomial this question that so, that is the objective for giving this question to bring you to your notice that how sometimes in some of the case we can approximate it using binomial distribution.

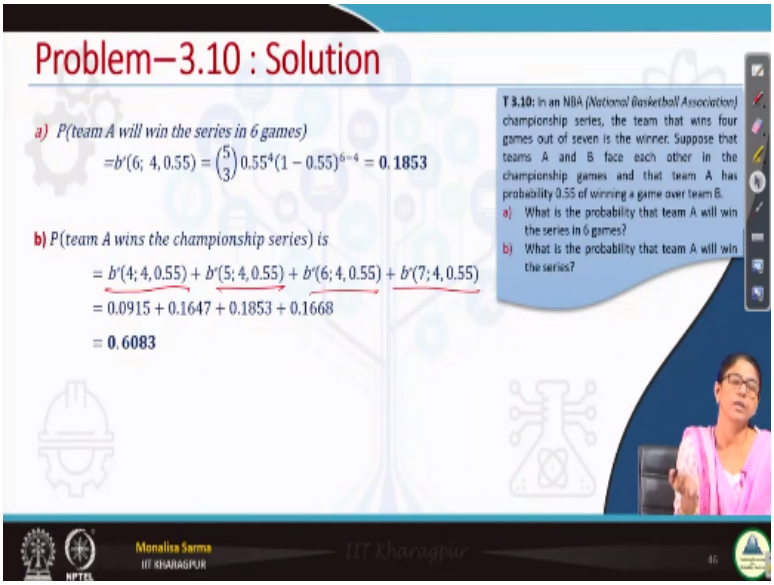
(Refer Slide Time: 20:30)



Now this in NBA National Basketball Association championship in the team that wins 4 games out of 7 is a winner. So, my number of success is 6 I need to have 4 success out of 7 games. That means number of 7 is success is fixed. Suppose the team A and B face each other in this championship games and a team A has probability of 0.55 of winning a game over team B that is probability A winning is given 0.55 what is the probability that team A will win the series in 6 games.

What is the probability that a team A will be in the series in 6 games that means the team A has for winning the series it needs to be in total 4 games, the 4th success it got in the 6th game 4th success, it is basically the question it is asking what is the probability that it got the 4th success in a 6 game. So that means it is a question of negative binomial distribution our success is fixed and number of games is the random variable.

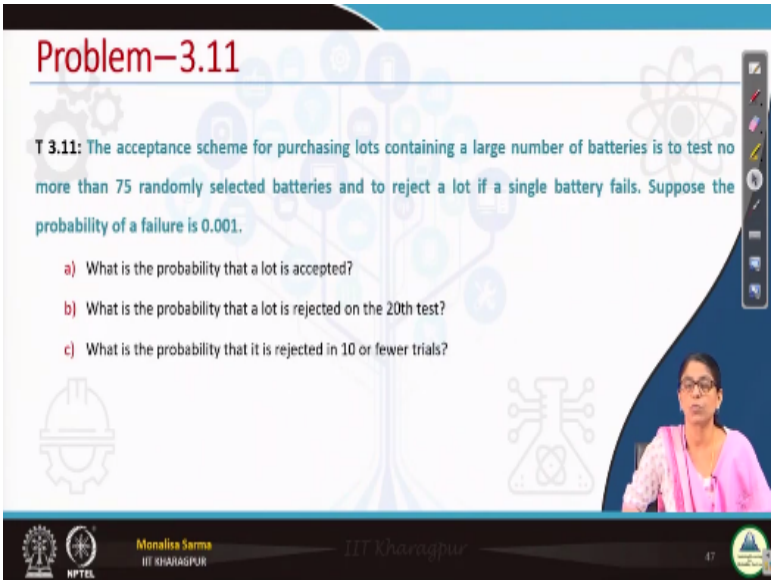
(Refer Slide Time: 21:36)



So, it is the 4th success he got in the 6th game. So that is just once we know this, you know what is n what is p, what is k, what is x than just put it in the formula. Second question, what is the probability that team A will win the series when a team A win the series for winning the series it has to win 4 games, so it can win again in the 4th game. The 4th success it can has it in the 4th game, it can have the 4th success in the 5th game, it can have the 4th success in the 6th game.

So, it will be total having the 4th success in the 4th plus having the 4th success in a 5th game plus having the 4th success in a 6th game. That is a total 7 games. So, here having the 4th success in the 4th game, having the 4th success in the 5th game, having the 4th success in the 6th game, having the 4th success in the 7th game then put it in a formula x - 1 C k – 1 x ^{k} x p $^{x-k}$.

(Refer Slide Time: 22:52)

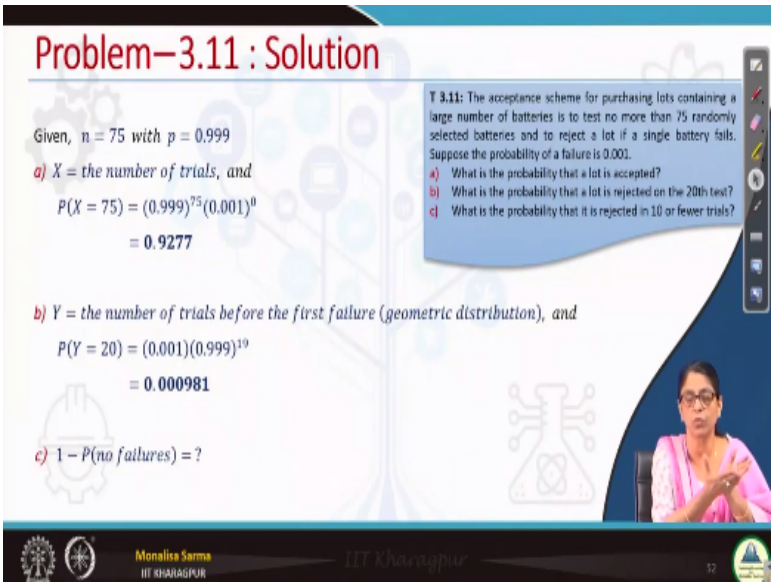


So, the acceptance sample for purchasing lots containing a large number of batteries is to test no more than 75 randomly selected batteries. And to reject a lot if a single battery fails. So, this is our acceptance sampling means is a large we want to purchase some batteries and a large number. And we will purchase the batteries for that what sampling plan we will see is we will pick at the most 75 batteries, at the most 75 batteries will fix and we will reject the lot in while picking the 75 battery if we get a single failure, we will reject a lot.

Let us see if we get a failure in the at the most 75 we will pick we will get a failure if suppose we are picking gradually 1 2 we are picking and we got the failure while picking the 10 batteries then also we will reject a lot we will go on suppose we got the first failure in that 10 to 10 also will reject a lot. So, at the most we will try till 75, till 75 if you do not get any failure then we will not see then we will assume that a lot is acceptable and we will accept it.

So, that is the question what it is asking containing a large sum is to test no more than 75 randomly selected lot is and to reject a lot if a single battery fails, suppose the probability of failure is 0.001 What is the probability that a lot is accepted when a lot is accepted? Till 75 there is no failure then a lot is accepted because it will check maximum till 75.

(Refer Slide Time: 24:27)

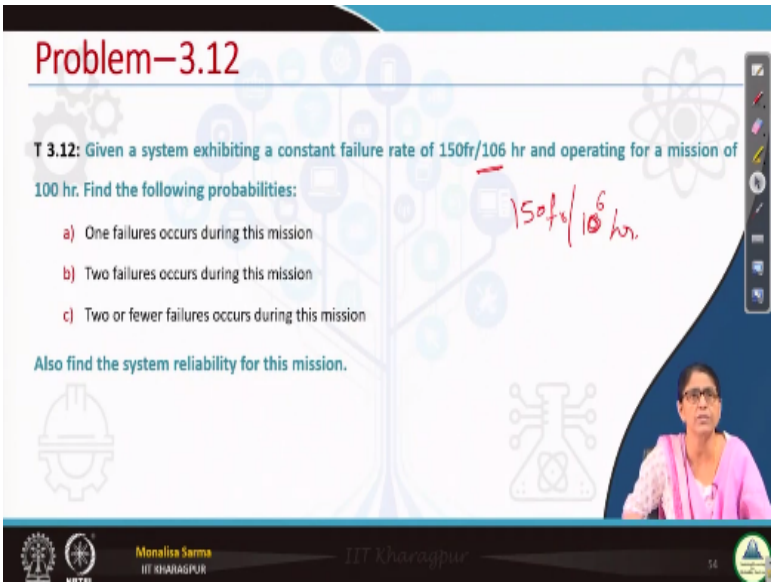


So, it is a question of geometric distribution. So, till 75 there is no failure. So, tell what is the probability of what to say probability of failure is 0.001. What is the probability of, then probability have success is 0.99 so we are using, 1 - probability of failures probability all till all 75 we got them success. So, this p to the 75 and this is q ^{0}, next is what is the probability that a lot is rejected on the 20th test, lot is rejected on the 20th test means the first failure that is the first failure we got on the 20th test mean still 19 it was all success.

The first failure we got in a 28 test. So, first failure this all this trial 19 all success then we got the first failure here geometric distribution. What is the probability that is rejected in 10 or fewer trials now distribution what is the probability that it is rejected in 10 or fewer trials, what does that means? That means, the first failure we may get in the first trial or in the second trial, third trial that means, the first release from 1 to 10 we may get it in any that is what we have to find out what is a probability that it is rejected in 10 or fewer trials.

So, that instead of finding it that we are getting the first failure in the first trial plus getting the first failure in the second trial plus getting the first failure in the third trial up to up to 10 instead of finding that we can just find out a compliment with 1minus probability of no failure that means still 10, what is till 10 we got no failure 1 minus probability of 10 is getting failure between 1 to 10 is not it?

(Refer Slide Time: 26:27)



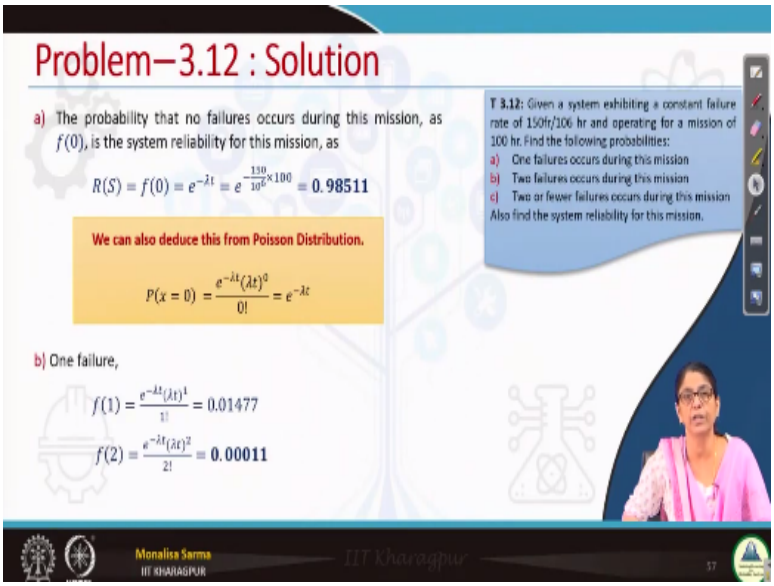
So, now, given a system exhibiting a constant failure that of 150 failure for it is not 106 hours is 10 to the power 6 hour, there is a mistake here. So, in an operating for a mission of 100 hours find a following probability. So, this here is 150 failure by 10 to the power 6, 10 over 6 hours this one this was a mistake here. Of course, when I will be giving the slide I will correct this and give it and for me to find a following probability is one failure occurs during this mission.

So, in this 100 hour time, we need to find out what is the probability of 1 failure what it is given in 10^6 hours total 150 failures this is the average failure, average failure intended for 6 hours is 150 failure and we need to find out what is the probability of 1 failure in this 100 hours 2 failures in this hour because mission time is 100 hours we have studied what is mission time in a first lecture. So, this is again a question of can you will find out? Yes, this is a question of Poisson distribution.

Where we are interested in a number of occurrences in a particular time interval or number of occurrences in a region of space, here number of occurrences in a time interval what is the time interval time interval is 100 hour. What is the λ? λ is 150 per 10^6 hours. So, what will be λ t 150 / 10^6 x 100 that is the λt so, first we are interested in 1 failure then 2 failures or 2 or fewer failures, 2 or fewer failures means 0 failure plus 1 failure plus 2 also find the system reliability for this mission.

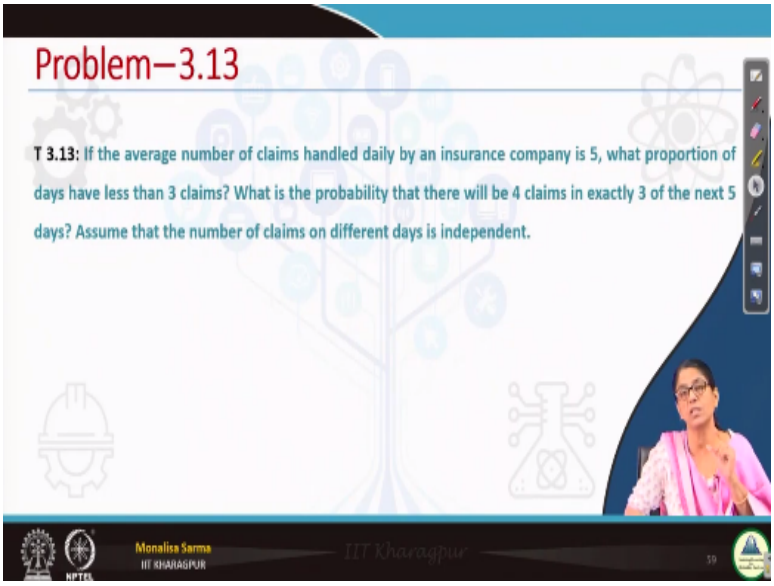
How do we find a system reliability? We talked up system level what is reliability? Reliability means failure free operation that means no failure that means 0 failures. So, what is the probability of 0 failures that is the system reliability.

(Refer Slide Time: 28:34)



First we are doing here probability finding 0 failure 0 failure is in Poisson distribution what is the formula e - λt, λt ^{x} !. So, what is x here x is 0 λt / 0 is 1 0 ! is 1 so, e to the power 1 - λt for zero failure. Similarly, we can do it for more than 1 failure, 2 failure and 2 or few failures is 0 1 2 summation on 0, 1 and 2, simple question of Poisson distribution.

(Refer Slide Time: 29:08)



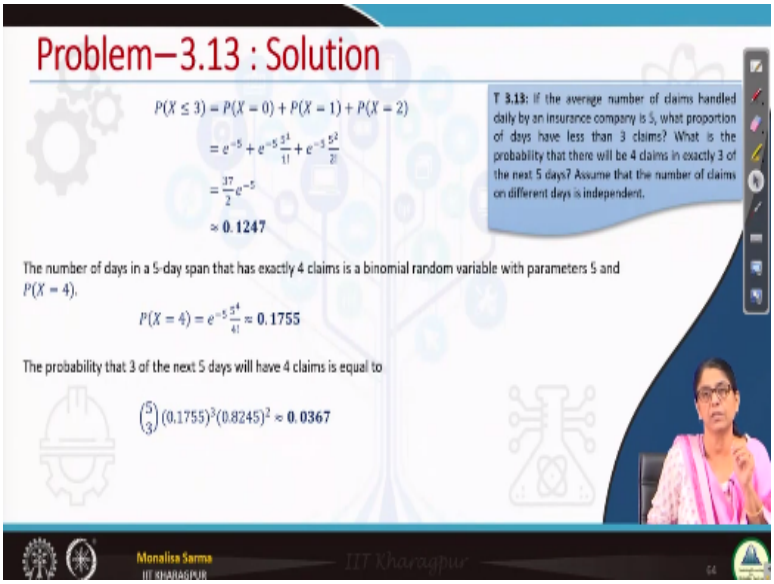
There is one more question here. This is an interesting question to see if the average number of claims handled daily by an insurance company is 5. Average number of claims is insurance companies 5 in a day, in a day average number of claims is 5 directly it is a question of Poisson distribution. What proportion of days have less than 3 claims? We have to find what proportion of days have less than 3 claims probability of less than 3, less than 3 means 0 claim, 1 claim, 2 claim that is one question.

Next is what is the probability that there will be 4 claims in exactly 3 of the next 5 days? We have to find out 3 here we have to find what is the probability that there will be exactly 4 claims in exactly 3 of the next 5 days, out of 5 days, in 3 days we will get 4 claims. So, what is this first question directly? First question is directly a question of Poisson distribution average number is given what proportion of day we will have less than 3 claims.

Probability of less than 3 probability of 0 claim what is the probability of 0 claim what is the probability of 1 claim what is a probability of 2 claim given that my lambda is 5 per day. Now, what I need to find out in exactly in the next 5 days, in exactly 3 days what is the probability that I will be getting 4 claims? So, for that, I need to find out next portion. Next portion is you see it is a binomial distribution and binomial distribution means out of 5, 5 trials my 3 success out of 5 it is a question of as if I can put it out of 5 what is the success? 3 successes.

Now I need the probability of success. So, here success means 4 claims. Here success is what success means not success, basically, for this question, I am just drawing an analogy here to success is basically the 4 claims. So, I need to find out what is the probability of 4 claims. Probability of 4 claims how do I get it? Probability of 4 claims by using this Poisson distribution, I will be getting the probability of 4 claims. Now, this probability I will be using it in a binomial distribution. So, this question is a combination of Poisson plus binomial distribution.

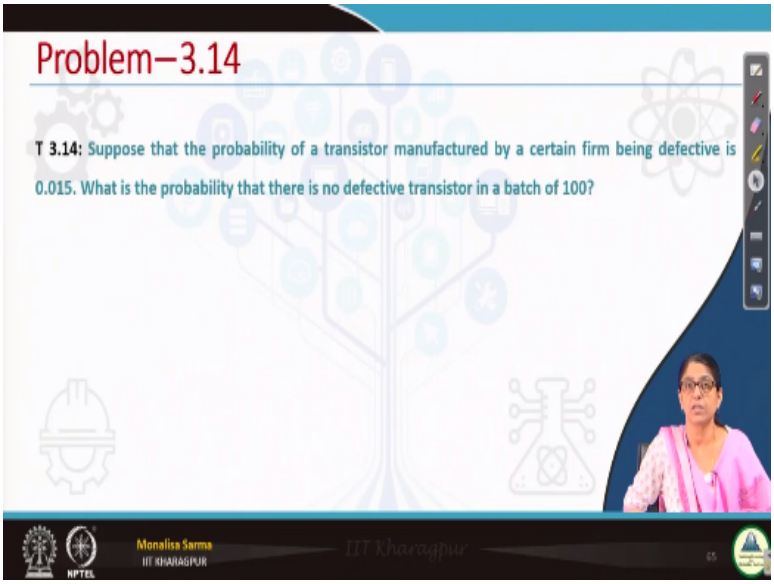
(Refer Slide Time: 31:30)



First you find a probability of x less than 3 that is for a first portion of the equation. As I told you, this, we have done probability of x = 0, x = 1, x = 2 by using the formula $e^{-\lambda t} \lambda t^{x}/x!$ then in a number of days in a 5 day span that is exactly 4 claim is a binomial random variable with parameters 5, 5 is the total number of trials that is m and p what is the p? p of x = 4 we found a probability distribution Poisson distribution.

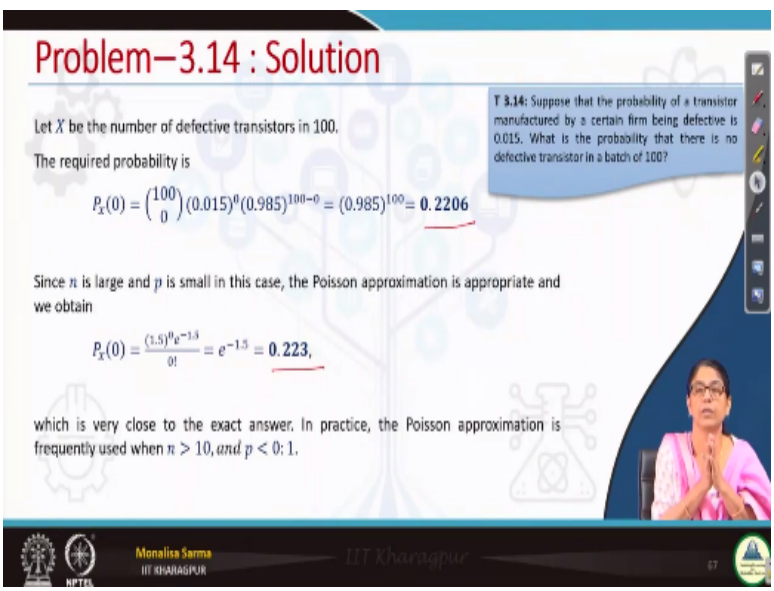
And then to find compute this probability of next 3 of the next 5 days using simple binomial distribution. So, this question we had Poisson distribution also as well as we found out the probability using Poisson distribution and then this probability we have used to calculate the value of the random variable binomial random variable.

(Refer Slide Time: 32:22)



Suppose the probability of a transistor manufactured by certain firm being defective is 0.015. What is the probability that there is no defective transistor in a batch of 100.

(Refer Slide Time: 32:34)



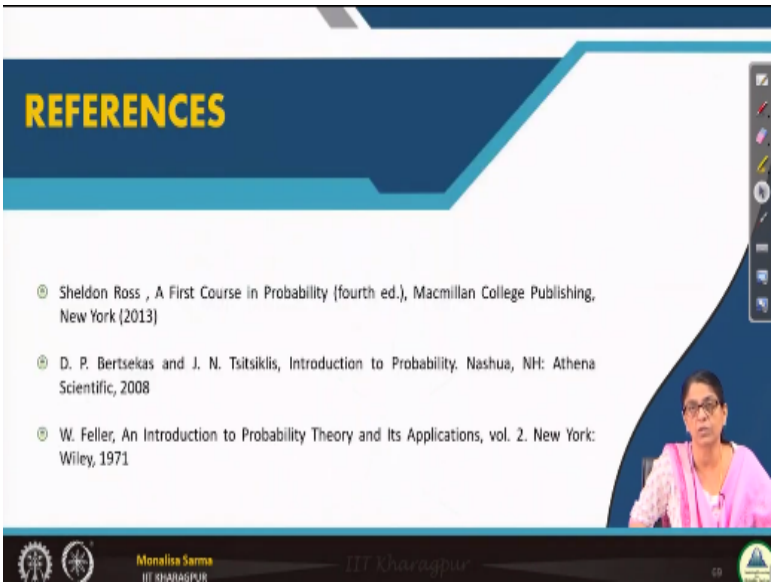
Very simple question of binomial distribution, but this question you can use, because n is very large and p is very small. Here again you can use approximation when n is very large and p is very small instead of binomial distribution, you can also use Poisson distribution. So, we have seen 2 approximation in a hypergeometric distribution can be approximated by binomial distribution when the number of population that is population size that is N is very large and sample size that is n is very small.

Then, instead of hypergeometric distribution we can use binomial distribution that we have seen. Now, again we have seen another one another one approximation in case of binomial distribution when what to say now, n is very large that is the number of trials, number of trials is very large and the probability is very small here to see if there is a batch of 100 that means, the number of trials basically it is 100, out of 100 how much we are picking up what is the probability is given? Probability is given only 0.015 p is very small.

When n is very large, p is very small instant when was under sufficient where binomial distributions would be used, instead of binomial we can approximate it by Poisson distribution as well which will not give a much of a difference in a result, you see both the result here, here we got the result is 0.2206. And here we got this not a very different high difference. So, there is a rule of term also in practice Poisson approximation is frequently used when n is greater than 10 and p is less than 0.1 and both the condition has to be satisfied.

Not that only when n is was greater than 10 we will go for Poisson and even a binomial is necessary in order both have satisfied n is greater than 10 and p is less than 0.1. So that is all and again, I request you to solve as many problems as you get which will give you give you a confidence.

(Refer Slide Time: 34:43)



So, these are the references and thank you. Thank you guys.

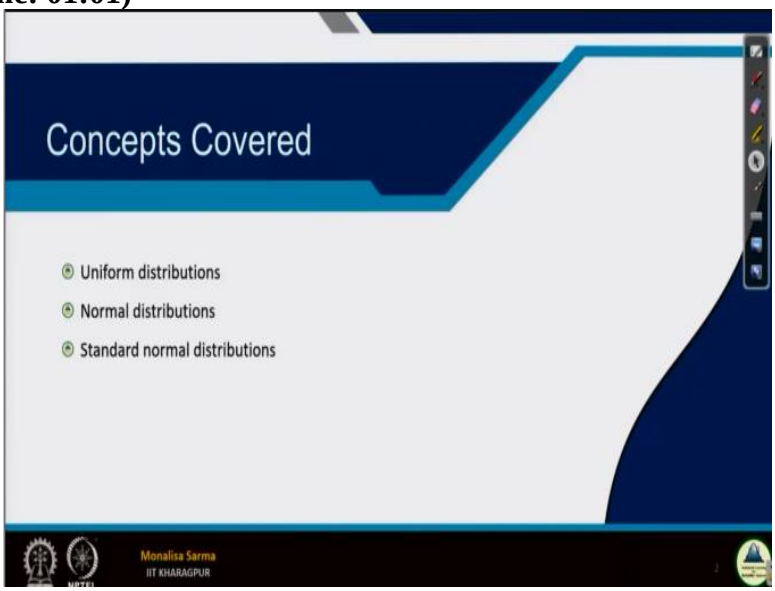
Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology - Kharagpur

Lecture - 12
Continuous Probability Distributions (Part 1)

Hello everyone, so in continuation of our discussion of probability distribution, today, we will be discussing continuous probability distribution. Last few classes, we have discussed discrete probability distribution, as well as we have done some tutorials also on the discrete probability distribution. Now, we will be discussing continuous probability distribution, this continuous probability distribution, I will be taking 2 lectures for it.

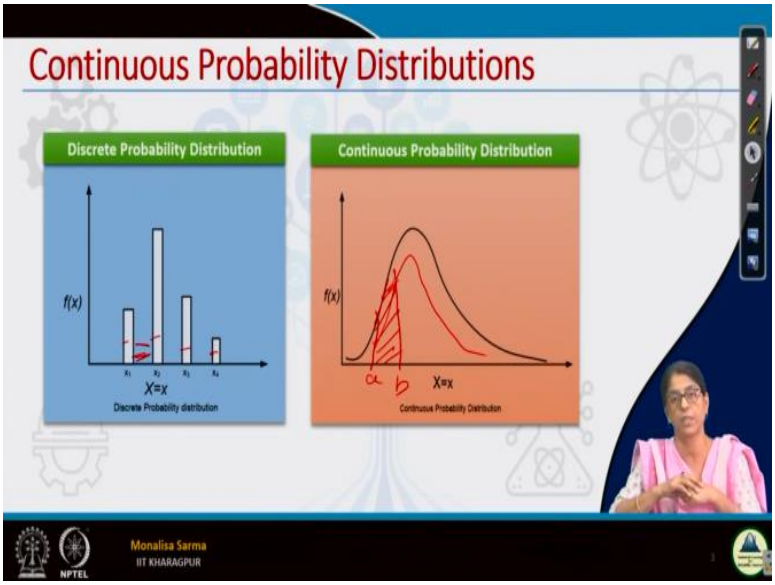
In the first lecture, we will be covering around 3 continuous probability distribution, and other 3 I will be covering in the next lecture, followed by a tutorial class.

(Refer Slide Time: 01:01)



So in this lecture, I will be covering uniform distribution, normal distribution, and standard normal distribution.

(Refer Slide Time: 01:07)



Now, before covering going to uniform distribution, let us first take a quick recap of what we have learned of discrete probability distribution, the left side figure so the discrete probability distribution remember how like, we can express a discrete probability distribution in the form of a tabular form, in the form of a graph, as well as in the form of expression, the figure here shows that we have expressed the discrete probably distribution in the form of a figure.

So now, can you remember what figure is this? It is a bar chart so where, why how I can distinguish that it is a bar chart, not an histogram, because in a histogram, there will not be any gap between the this rectangles, let me take the pen, and then I will be showing you. So can you see this gap? This gap in a histogram, this gap will not be there basically, they will not be this gaps, it will be just close together.

Because the width of this rectangle in a histogram, it refers to the in the range of the values. So we cannot put any gap in between 2 rectangles, so this directly shows that is a bar chart. Anyway, that is not a main concern, main concern is here and trying just trying to say you how we basically can have a figure of a discrete probability distribution. Similarly, for continuous probability distribution, the figure is in a continuous curve.

So here we see this is a continuous curve. So in this continuous probability distribution, the figure is basically a continuous curve. So now, the thing is that first, again, I will just, we will just go back a bit and we will see what was f(x) in case of a discrete probability distribution, what was f(x)? f(x) is the probability of the random variable X, the probability that X takes

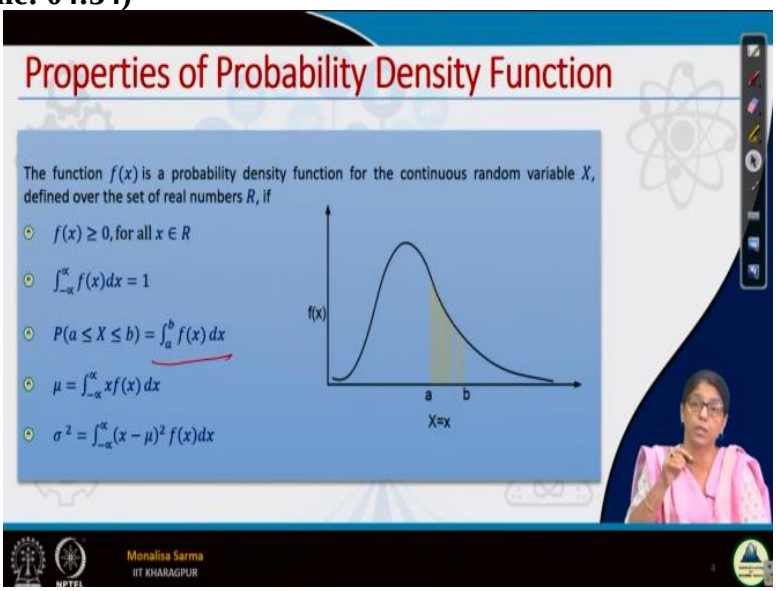
that a random variable X takes that is f(x), so value of f(x) cannot be greater than 1, it cannot be less than 0. So f(x) has to be between 0 and 1.

That is the case in case of discrete probability distribution. But in case of continuous probability distribution, I have already mentioned if you can remember, of course, f(x) cannot be less than 0, it has to be greater or equal to 0, but f(x) can be greater than 1 also, because here f(x) does not denote the probability it is here f(x) is a density function how μch the probability concentration is there in it within a certain interval.

When you want to find out the probability in for continuous, we will have to find out the area within that interval. Suppose I am considering this area suppose this area, so this area, I will have to find out the area of this area, I have to find out the area of this region. Suppose this is a this is b, I will have to find out the area of this region to find out the probability at this point. Sorry, it cannot be found when speaking of continuous probability distribution, it will never be at a single point.

I already mentioned that probability or a single point is equals to 0, it is almost equals to 0. So when we talk about probability and continuous domain, it has to be within an interval, maybe a very small interval, maybe a bigger interval. And, of course, for a random variable for the function to satisfy a continuous probability distribution rather, I should say function to satisfy a probability density function, we can say it is a probability density function, when it satisfies 2 of the properties if you remember.

(Refer Slide Time: 04:34)

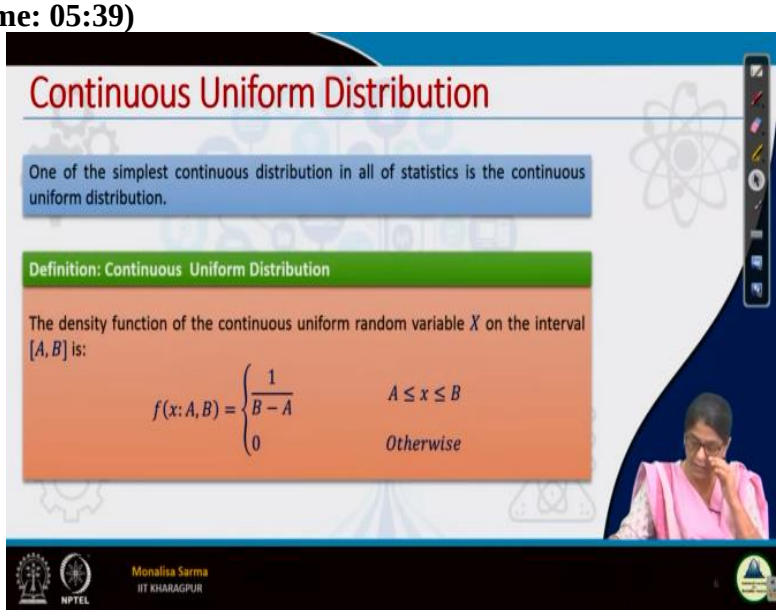


I am just putting it here again, it will it is called the probability density function. The first condition is that it f(x) will be greater than equal to 0 and integration of f(x) from all the range from - ∞ to + ∞ it should be equal to 1 these are the 2 conditions. So, now if I want to find out the probability between 2 values that a and b in the figure it is given this a and b.

So how we can find out it is just the integration of f(x) within the value a to b that is how we do in continuous random variables when the variable is continuous. So, for finding out the μ and that is the mean and the variance mean or I should say the expected value on the variance, that is same as what we have done for discrete probability distribution, but instead of summation here they are, we have used summation here we will be using $\int \mu f(x)$ $\int x\, f(x)$ and for σ this $\int (x - \mu)^2 f(x)$.

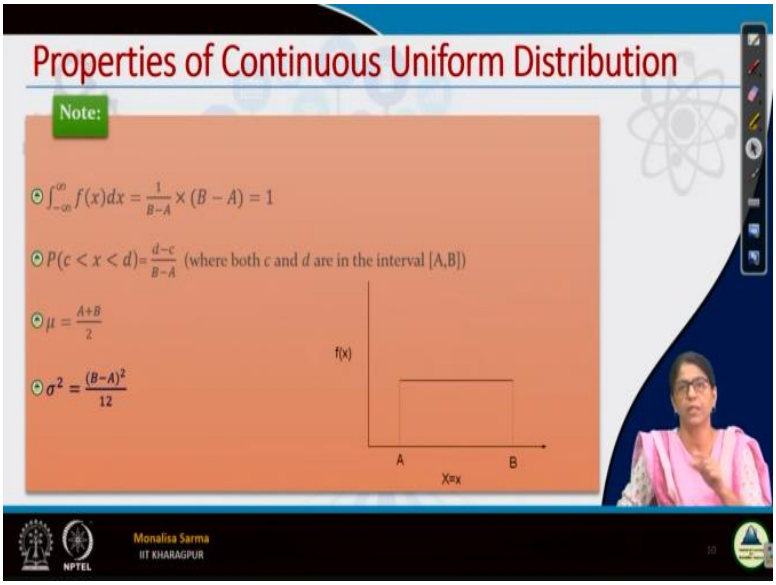
It is the same just summation get replaced by integration because here we are concerned of the forμla, the area within that range. So, we will be basically integrating it.

(Refer Slide Time: 05:39)



So, one of the simplest continuous probability distribution is the uniform distribution. So, we talk about uniform distribution, where the probability at each point is same.

(Refer Slide Time: 05:55)



So, basically, if I can tell you about the graph of a probability continuous uniform distribution that will be this sort of a graph here. So, we are trying to find out within the range the A to B, so, this is the graph here. So, it is equal probable in all the areas and of course, since it is a probability distribution, so, some ∫f(x) would be equals to 1. So, if ∫ f(x) = 1 so, what will be the probability of this probability of any value means probability in this range what is the probability it will be 1 / B - A is not it?

So, probability will so my f(x) will be as what it is given here f(x) will be 1 / B - A directly we can get ε since the area if we find out the area of this whole rectangle area of the rectangle is what? Length x breadth if now my this is my I have to find out my probability in this area this whole area has to be 1. So, what it will be to when will I get 1 because my total length is B - A that means, to get it one my I have to µltiply it 1 by B - A that directly gives me the probability is not it?

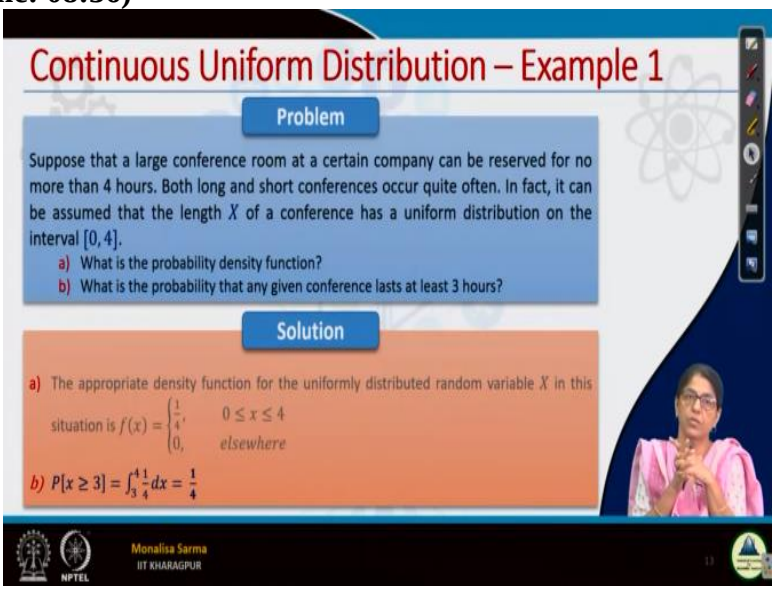
That I should not say probability that directly give me the density function that is f(x) it is not probability mind it that is the f(x) so, my f(x) is 1 / B - A, 1 / B - A from the range X ranging from A to B and in other areas it is 0. So, now, if we are interested in finding out any value within this range suppose we are interested in finding out what is the probability in this range c to d. So, what will be the probability in this range c to d what is my length of this area?

Length of these areas this d - c. So, d - c into what is my probability density function density function is 1 / B - A. So my probability become d - c / B - A. So, that is the probability within the range c to d probability always in continuous only distribution and I am repeating it again

when we talk about continuous probability distribution, the probability is always within an interval we do not talk of a single point.

So, here by using the forμla same forμla integration of x f(x) for expected value, we will be getting A + B / 2 for σ^2 we will be getting B - A / 2 where B and A are the range of the random variable where X is a random variable X random variable can take value within the range of A to B.

(Refer Slide Time: 08:36)



Now, one example to illustrate our continuous uniform distribution remember the uniform distribution we had it for discrete also similar, we have in uniform distribution both for discrete and continuous. Now, coming to this example, see this example if you go to the example you will find it is very easily what it is given. Suppose that a large conference room at a certain company can be reserved for no more than 4 hours.

The large conference room we cannot book it for more than 4 hours. That means my value of my random variable, it ranges from maxiμm 0 to 4 booking cannot be negative, though, when it our continuous probability distribution. So, it can be - ∞ to + ∞ but here booking cannot be negative so, it will be from 0 to 4 rest every all other places have a random variable the probability of the density function will take the value 0.

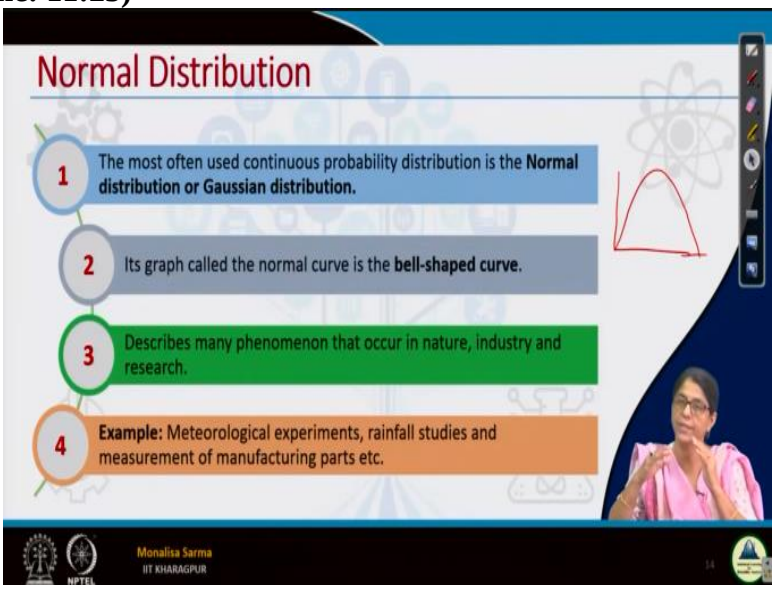
So, then both long and short conference can occur quite often. In fact, it can be assumed that the length X of a conference has any form distribution on the interval 0 to 4. It is specified that the length X where X is a random variable, X is a uniform distribution in the range of 0

to 4. If the X has a uniform distribution in the range 0 to 4. Then what is the probability density function that means what is my f(x)? My f(x) = 1 / 4 is not it?

1 by 4 - 0, I have to find it from this range, this is 0, this is 4, this is my range. So my probability density function will be 1 / 4 - 0 that is 1 / 4. So that is my f(x), so first question what is the probability density function that is all it is 1 by 4 from which range from 0 to 4 and rest are all other area, what is the range? It is 0, what is the probability density function other than this range, that is from 0 to 4, it is 0.

That is how we define the density function. See here, the density function it is 1 by 4 from 0 is elsewhere from x ranging from 0 to 4 and 0, it is elsewhere. Now, what is the next question? What is the probability that any given conference last at least 3 hours? So it is given that any given conference last at least 3 hours so that means my random variable will go from 3 and above? So I will integrate it from 3 to 4 so that is a p(x) >= 3 at simple integration from 3 to 4 1 before dx, that is all that is very easy problem.

(Refer Slide Time: 11:25)



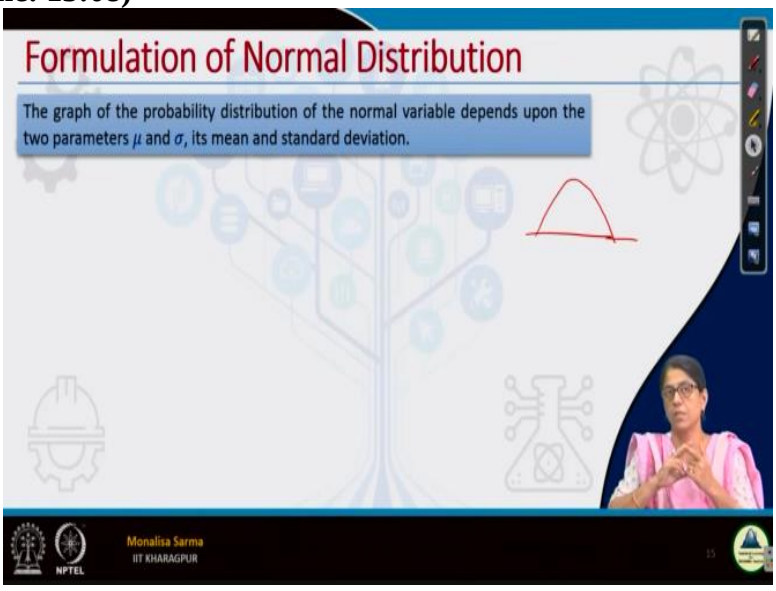
Now, we will discuss the next important distribution that is normal distribution is also Gaussian distribution this is a very important distribution in the sense that this distribution is it has wide application you can see this application of this in nature you can see application of design industry, research, business many places, you can see this application of this normal distribution this distribution was discovered by one professor,

Professor Carl Gauss, that is why by his name it is this distribution is also called Gaussian distribution and this distribution it has a basically a bell shaped. So, this is called a bell shaped curve, and describes any phenomenon on that occurs in nature industry and research. Some of the example is metrological example rainfall studies, measurements of manufacturing parts, etc.

Let me tell you, what made this professor Gauss to discover this distribution actually, he was trying to measure the movement of celestial body more he was trying to measure a moment of celestial body again and again of one celestial body, he was trying to measure and its moment on each measurement, he was getting different, different value. So, when he was trying to see there is an error in each measurement.

He saw that it fits into a distribution that distribution that he has come up with this distribution based on his error and measurement that is this bell shaped distribution that is a Gaussian distribution.

(Refer Slide Time: 13:08)



So, as I when I was discussing discrete probability distribution, I have already specified that when we talk about distribution, there are a few parameters which specify a distribution likewise, uniform distribution also we saw the parameters that describe the uniform distributions are μ and variants μ and σ. So, σ if I tell σ is a standard deviation like.

Similarly, for normal distributions, the parameters that are μ and σ here also when μ is the mean, and σ is the standard deviation, let me tell it in a different way actually, when we talk

about distribution and distribution are characterized by a few parameters, which are called maybe scale parameter, shape parameter, location parameter. So, instead of telling as μ or standard deviation, that a parameter is defined by scale, what is the scale of the distribution?

So, then parameter is defined by the location where it is centered? Basically the expected value the μ, where it is centered, that is the location then we have called something called a shape parameter, what is the shape of the distribution? Shape of the distribution like as I talked about discrete probability distribution, I told you we are not very μch concerned about the state of the distribution while talking about discrete probability distribution.

Actually, we do not bother that μch of what to say. But in case of continuous probability distribution, when we are talking about the distribution of a continuous random variable, shape plays an important role. So, there is a shape parameter as well. Now, this parameters, these are also called moment there, I have not kept it in a slide because this is you do not need to consider as a part of this course.

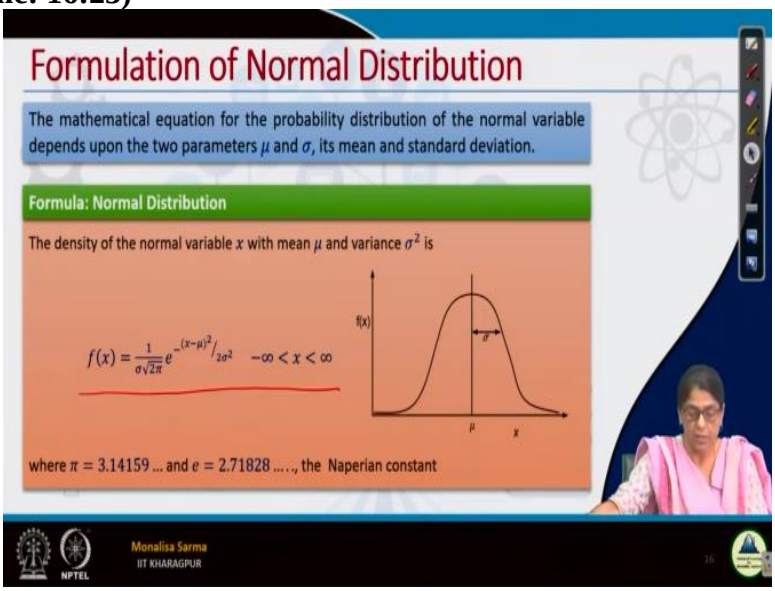
But for your knowledge for your information, I am just telling this we can also call the moments there are different moments which characterize the distribution, the first moment is the expected value that is the mean that is the location. Then the second moment is the scale parameter that is the standard deviation that is the spread how spread out my distribution is what the variants gives how spread out is.

So, that is the second moment then third moment is symmetric, how symmetrical is my distribution, whether it is skewed towards left it is skewed towards right or it is centered it is symmetrical to the mean. So, that is the third, what to say third moment, then fourth moment is the kurtosis, kurtosis gave the peakedness of the distribution and that is the height of the distribution.

So, these are the there are many other moments over these 4 moments are very necessary to describe characterize the distribution, but we will not go on discuss the moments here, but we will just stick to the parameters which are necessary to describe a distribution. So, talking of normal distributions, the parameters that describe a normal distribution basically, it is the mean and a σ mean is the location parameter and σ is the scale parameter here we do not have shape parameter why?

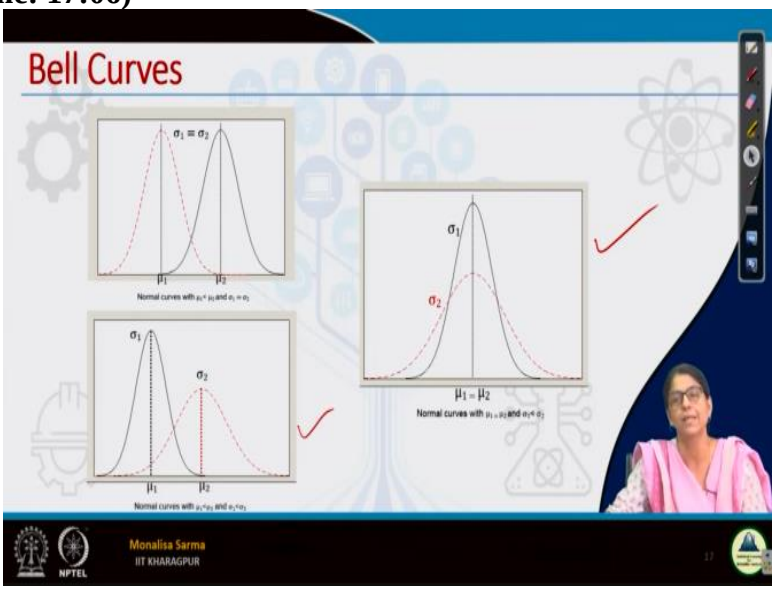
We do not have shape parameter because when we talk about normal distribution, normal distribution just has 1 shape that is the bell shaped just this one same. So, we do not have a special parameter to differentiate that different shapes geometric since normal distribution has this one set no point. So, normal distribution has this 2 parameter that is μ and σ it is also called the location parameter as well as the scale parameter.

(Refer Slide Time: 16:23)



Now, this is the probability density function for the normal random variable if x is a normal random variable. So, this is the density function like for uniform distribution, what is the density function? Density function is the f (x) = 1 / B - A where x range from B to A here x ranging from minus infinity to plus infinity my f (x) is this particular expression. Now, I will come to this expression again.

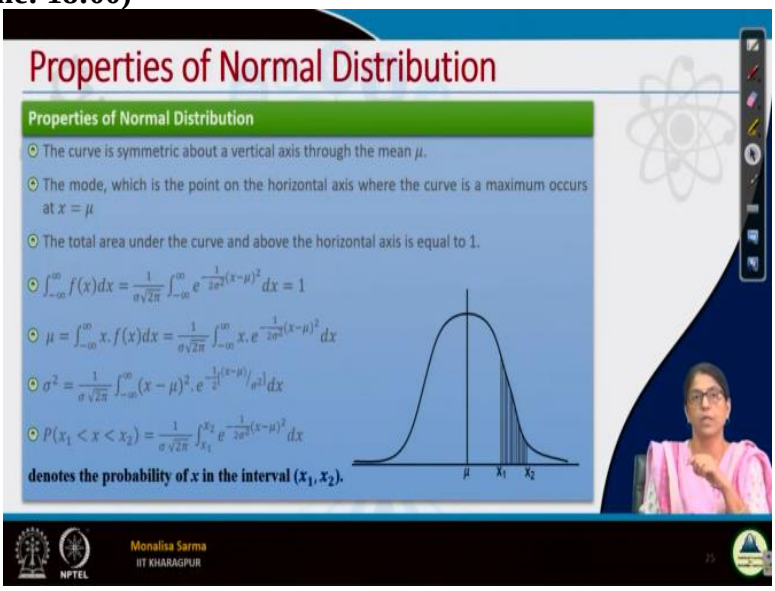
(Refer Slide Time: 17:06)



But before that, let us see some shapes of the normal distribution like here you can see different bell curve first one you see where it my σ 1 = σ 2 both you have seen here we have seen 2 bell curve where σ is same for both that case, but that means there my spread is same, but the location is different one location is μ 1 the another location is μ 2 you see how it is, then you see this one here the location is same that is μ1 = μ 2.

But both the scale is different for where here what is σ 1 is smaller than σ 2, σ 2 is more that is it is more spread out. Similarly, here we see the third one where μ 1 is less than μ 2 and σ 1 is less than σ 2 this this type of figure we basically get for different values of μ and σ.

(Refer Slide Time: 18:00)



So, now, for the normal distribution what the properties are the curve is symmetric about the vertical axis through the mean μ the curve is symmetric to the mean μ whatever may be the mean if this is my mean suppose, this is my y axis suppose this is my mean my curve is symmetric towards the mean and we also know what is mode see the mode is the point which is the point in the horizontal axis where the curve is maxiμm in a normal distribution curve as maxiμm mean at the expected value at μ.

So, mode is also μ here and at normal distribution my median is also μ here means what is median? At median 50% of the world's population falls in the left of the median with under 50% falls to the right of the median. So, here a normal distribution μ is also the median also in a normal distribution mean is equals to mode is equals to mean this is a very important characteristics please remember.

The total area under the curve of course, it has to be 1 in any continuous probability distribution we have seen f(x) gives the shape of the curve and the total area under the curve is equals to 1 that is why it we integrate f(x) from integration from minus infinity to plus infinity we get 1. So, this is the integration we get this is this whole is equals to 1. So, now what is μ? μ is integration x f(x) and σ^2 is integration $(x - \mu)^2$ and f(x).

Just put it into value and then we have to simplify. So, if we are interested in finding out what is the probability in this range x 1 to x 2 it is given here. If you want to find out what is the probability within this interval x 1 to x 2, just we have to integrate the f(x) value, whatever f(x) expression we have, we have to integrate f(x) from x 1 to x 2. Now see the problem here, when we are interested in finding the probability of this range x 1 to x 2.

Maybe any other range whatever range maybe we find you are interested in finding out the probability, then see, we will have to evaluate this expression, for this particular range evaluating this expression is not a very easy task we will have to do this is a complex integration solving this complex integration again and again for when trying to find out a probability of any interval it is a tedious job.

So, this job is made easy it can be I should say it can be made easy if we have a standard lookup table like what we see in the binomial distribution, I assume I told you that we have a table I also showed you remember similarly, if we can have a lookup table, then instead of doing this computing disintegration again and again we can refer to the table and we can come to the value that will be very easy.

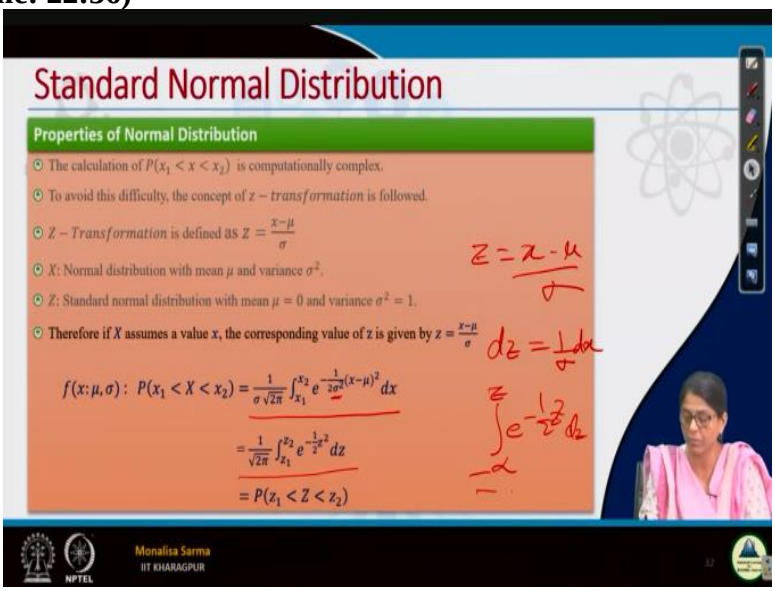
So, now, the problem is that how many tables will have because this value this expression will have different values for different μ and different σ square my mean and σ can be anything, it can be any value mean can be starting from 1, 2, 3, 4 y 1 it can be 0, so, it can be any value, even σ also it can be any value, so, we will have to have different tables for different values, is not it? So, that is not possible.

So, it is not really not impossible to have a normal table for such a case, but there is a rescue when we can the rescue is that we can transform this normal random variable, we can transform this normal random variable to is z distribution it is called z distribution and under

corresponding distribution is called standard normal distribution, how when we can transform this x to z distribution what happens then our mean is always 0 and a variance is 1.

So, that means, if we have a table for mean 0 and variance 1 for different probability values, then it is very µch visible we can have that we can look up the table. So, there are standard lookup table where we can find the value of the normal distribution, but this random variable, normal random variable first we have to convert it to z value. So, now once we can convert it to z value there are its corresponding distribution of z value has a mean of 0 and variance of 1.

(Refer Slide Time: 22:36)



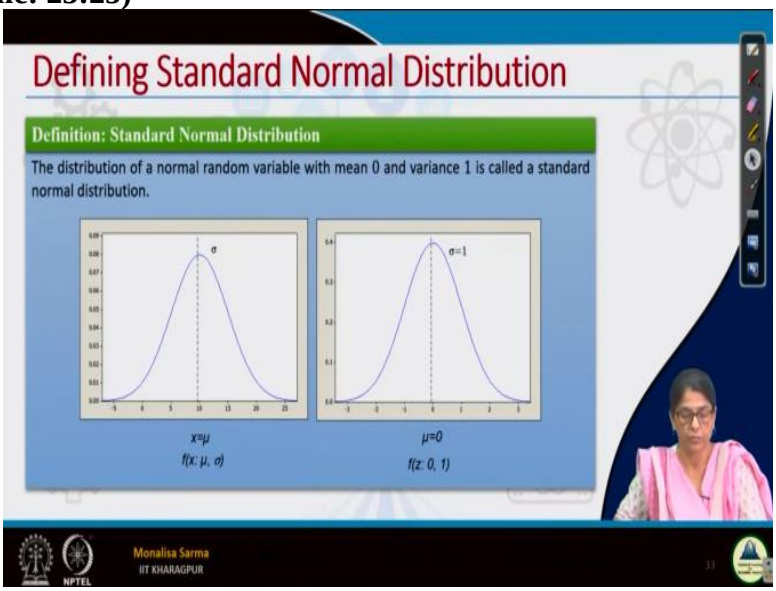
So, this is just how we can convert this z to x normal random variable x to z this is the forµla z = x - µ by σ. So, we need to convert x our random variable is x is not it? We need to convert x to z how we will convert, so, z = x - µ by σ. So, what we got is a z distribution or it is also called standard normal distribution, it has a mean of 0, always, its mean will be 0, this is the y axis, it will always mean will be 0 and standard deviation is 1.

So, we need just a couple of lookup tables with mean 0 and standard deviation 1, but having different because the probability might be different, is not it? So, for the different probability of occurrence, we can have different tables that is all. So, we will see some examples. So, that is what, see here remember this was the expression for f(x) if we are interested in finding out the values between in the interval x 1 to x 2 so, this was our expression.

Now, what we have taken we have taken z = x - μ by σ is not it? Now what will be dz? dz is equals to what? 1 by σ dx sorry 1 by σ dx, this is dx 1 by σ dx that is dz. So, here in this if I can substitute x - μ by σ square x - μ whole square by σ square this σ square is here if I can substitute by z and then substitute in dx by σ this dz so, I will be getting this and this value I can get it from the table but again one more thing there is one more point to that the table has the cuμlative value.

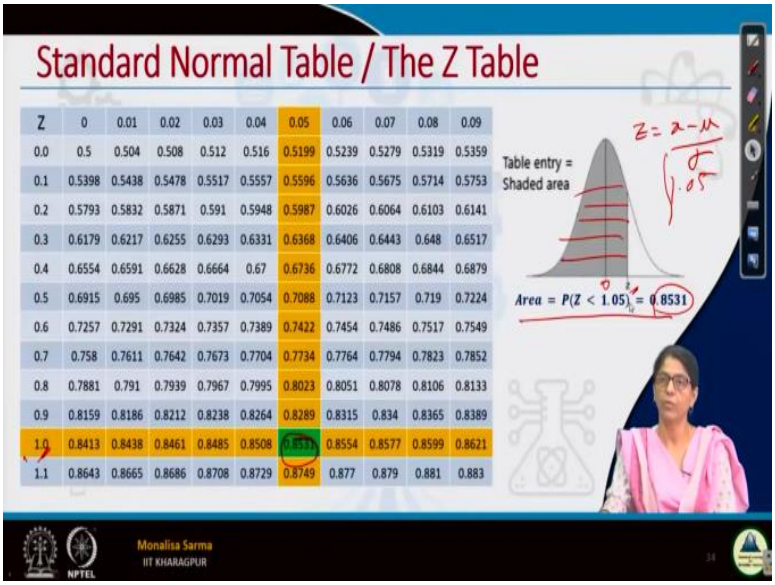
The table what we will see like in binomial distribution, binomial distribution table had the cuμlative value. Cuμlative probability distribution function values Similarly, here also normal table it has the cuμlative distribution value. So, cuμlative means I will have the value say from minus infinity to a particular value set to a particular value. So, z whatever is z and this expression $e^{-1/2 z^2}$ dz this value will get it from the table. So, once we can get the value from the table our job is done.

(Refer Slide Time: 25:25)



That is what the figure shows the difference between the standard normal distribution and normal distribution what I have already discussed.

(Refer Slide Time: 25:33)

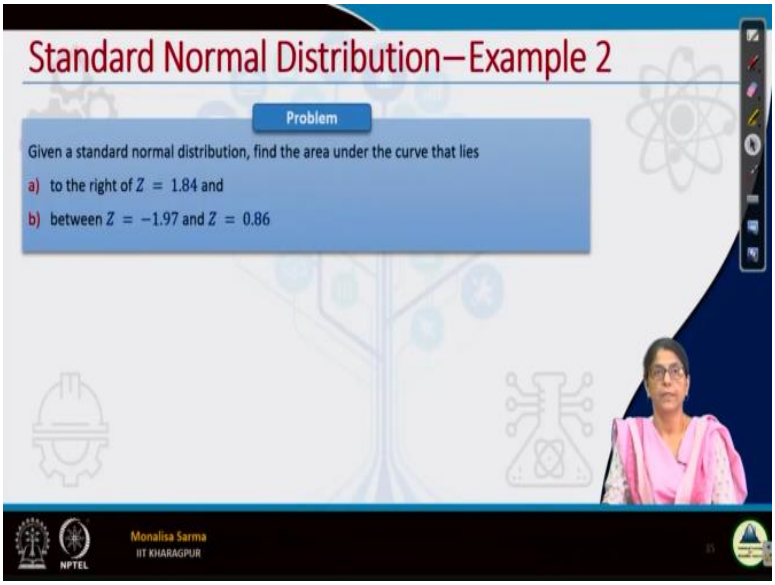


Now, see how we can look up the table see a table, this is the table suppose in the figure see in this figure there was some x I have found out from what is this z = x - μ / σ by substituting μ and σ value suppose I got my value of z as 1.05. Now, I will have to find out the probability corresponding to that means probability corresponding to 1.5 means integration of this value from - ∞ to a particular value.

So, whatever the to this value of z that is 1.05. As I told you it always gives the cuμlative value. So, how do I look the table see here is so, in the vertical you will see this is z value vertical and horizontal both are z value how we find in a vertical you see this yellow colour this is 1 I need 1.05. So, this is 1 this is 0.05 so, that is total together is 1.05 and where it meets this is my probability 0.8531.

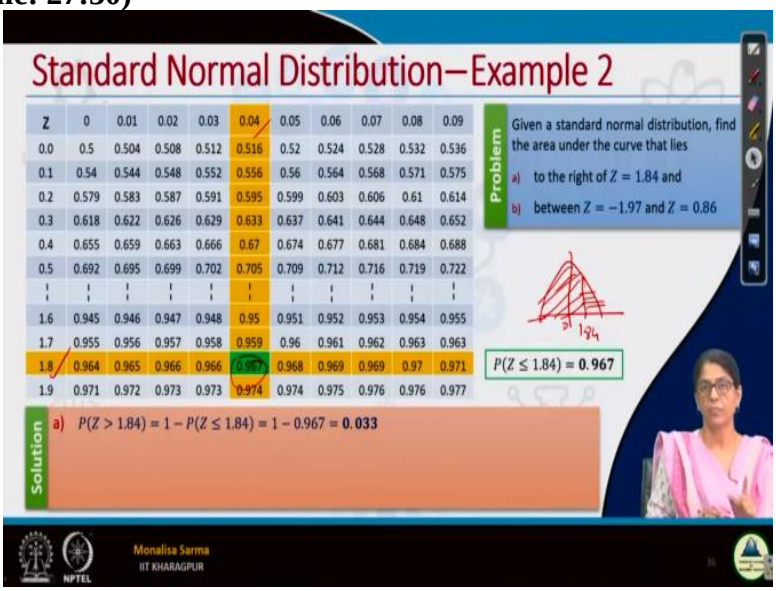
So, if as I told you standard normal distribution median is 0. So, this is 0 if this is the curve this is 0, so, 1.05 that will be 1.05 it will be to the towards the right of 0. So, this is my z value 1.05 and what is this area? This area is basically the probability that my value will lie in this range this area is nothing but the probability so, this area is 0.8531. This is how we look under the table. In the tutorial class we will be doing here also, of course, we will be doing some problem material first we will be doing more problems and things will be more clear.

(Refer Slide Time: 27:21)



Then say given a standard normal distribution find the area under curve that lies to the right of z = 1.84.

(Refer Slide Time: 27:30)

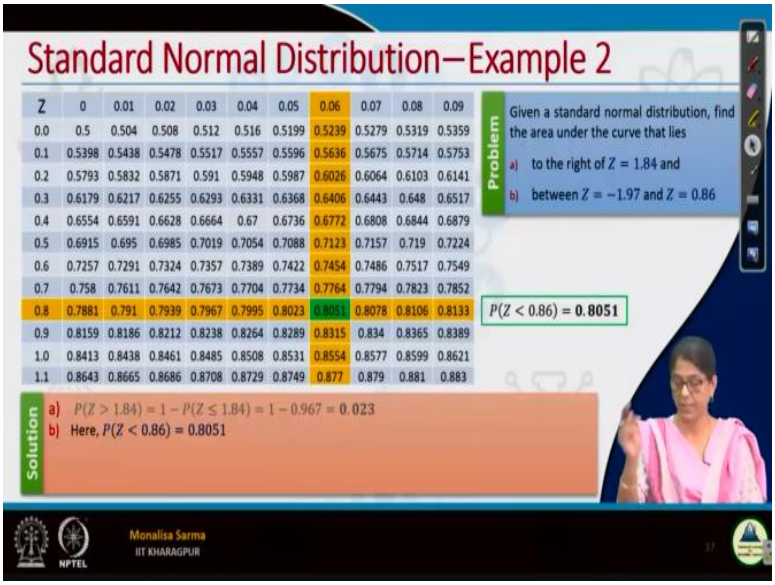


Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.504	0.508	0.512	0.516	0.52	0.524	0.528	0.532	0.536
0.1	0.54	0.544	0.548	0.552	0.556	0.56	0.564	0.568	0.571	0.575
0.2	0.579	0.583	0.587	0.591	0.595	0.599	0.603	0.606	0.61	0.614
0.3	0.618	0.622	0.626	0.629	0.633	0.637	0.641	0.644	0.648	0.652
0.4	0.655	0.659	0.663	0.666	0.67	0.674	0.677	0.681	0.684	0.688
0.5	0.692	0.695	0.699	0.702	0.705	0.709	0.712	0.716	0.719	0.722
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1.6	0.945	0.946	0.947	0.948	0.95	0.951	0.952	0.953	0.954	0.955
1.7	0.955	0.956	0.957	0.958	0.959	0.96	0.961	0.962	0.963	0.963
1.8	0.964	0.965	0.966	0.966	0.967	0.968	0.969	0.969	0.97	0.971
1.9	0.971	0.972	0.973	0.973	0.974	0.974	0.975	0.976	0.976	0.977

To the right of z = 1.84 so, right of that is was to 1.84 means, if this is my curve z = 1.84 because this is 0 my z value will be somewhere here 1.84 right means I need to find out this area. So, what but what my graph will give my table will give? My table will give this area left of this table is the cuμlative it will start on cuμlative means from minus infinity to 1.84.

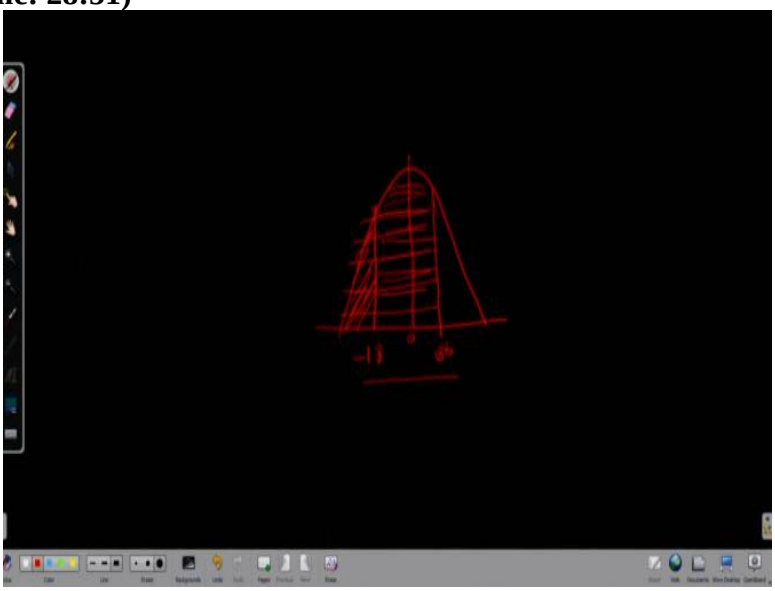
I will get this value, but I need this value that means one minus of that I will be getting this value right. So, here you see, so, it is 1.84 this is 1.8 and 0.04 will make it 1.84. So, I got this is the value then one minus of that is whatever value I needed.

(Refer Slide Time: 28:26)



Similarly between z = - 1.97 and z = 0.86. So, it is asking between z = 1.97 and 0.86 how we will do here suppose let me take the table then it will be better.

(Refer Slide Time: 28:51)



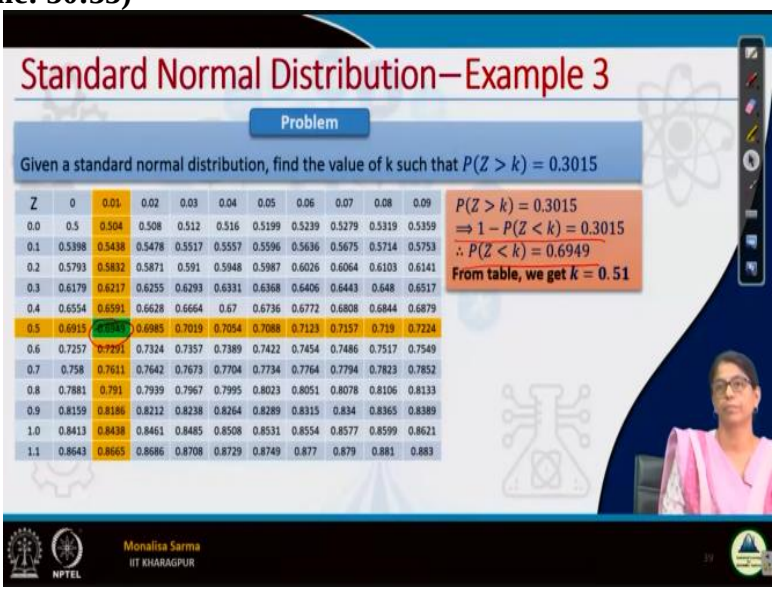
So, it is asking some minus one point I do not remember what was it say minus 1.5 suppose whatever it was in the question I do not remember and it was asking from minus 1 sorry it is not 1.5 will not be this, I will erase it here so, minus one point will be something here say minus 1.8 whatever it is, because this is 0 minus will be this side and between 0.86 so my 0.86 may be say this value, 0.86 so, I need this area 0, I need to find out this area.

That means basically I need to find out this area means what it is asking? I need to find out what is the probability that my random variable will live in this range 1.86 to 0.86. So when I am interested in finding this area, my cuμlative table will have all the values to the left of it.

So if I see 0.86, I will be getting this value from minus infinity to this value. And then when I am looking for 1.8 from here this portion.

But my intention is finding out this portion with between this. So, how will I find between this whatever value I got from 0.86 from this value if I subtract this value I will get this value. So, that is how we look it from the table okay fine.

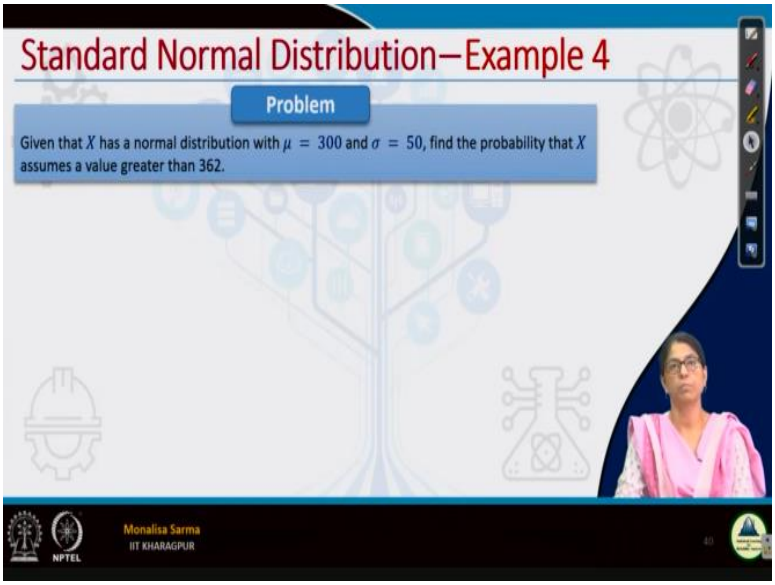
(Refer Slide Time: 30:33)



Similarly, given a standard normal distribution find a value of k such that probability of z greater than k is 0.315, now, we have to find out what is k here. So, it is given greater than k in the cuμlative table it is always less from - ∞ to a particular value. So, first of all when it is that greater than k what we have done first we have tried to find out what is less than k so, z less than k is this value 0.6949.

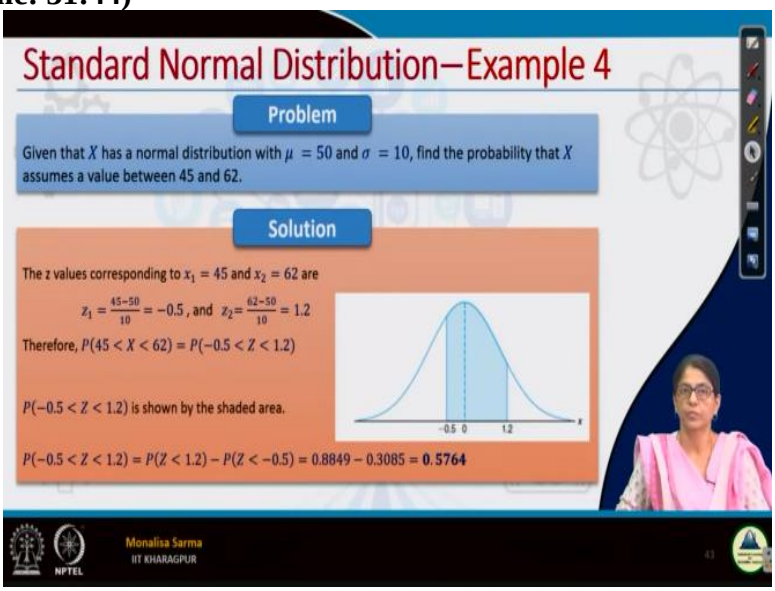
Now, from what is the value of k corresponding to this area this is the all this here in a table all the entries are nothing but the area within this curve so, here it is 0.6949. So, I found 0.6949 is this value. So, what is corresponding to 0.6949 it is 0.5 here it is 0.1 that means 0.51, my k value is 0.51.

(Refer Slide Time: 31:30)



So, one problem given that X is a normal distribution with mean μ= 300 and σ = 50 find a probability that X assumes a value greater than 262.

(Refer Slide Time: 31:44)



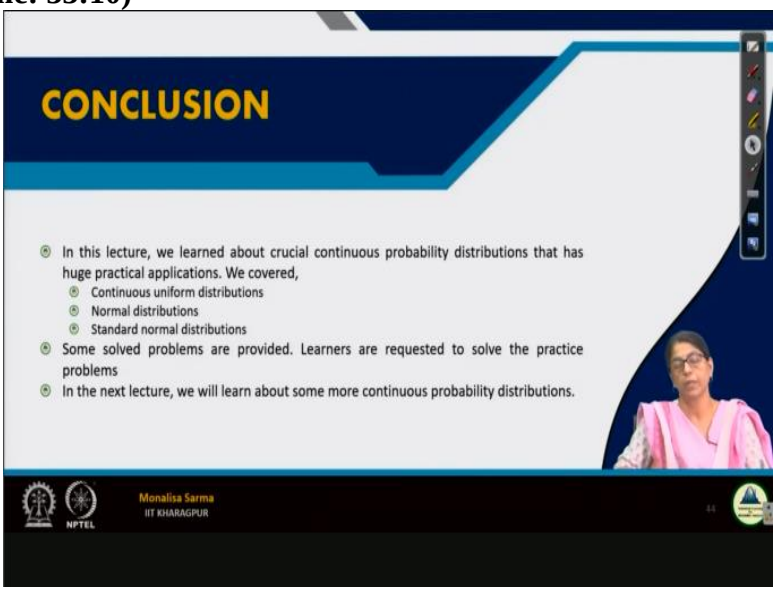
No, sorry, this is there is a slight mistake. That is the question given that X is a normal distribution with μ = 50. And σ is we will put them find a probability that X assumes a value between 45 and 62. So first, between 45 and a random value of the random variable is 45. Value of the random variable is 62 this X, I will have to first convert it to z. So, what is corresponding to x 45 what is the z value I found that value is - 0.5.

Corresponding to x value 62 my z value is 1.2. So, if we can see the figure that means I am interested in this blue shaded area, I want to find out what is the probability that my random variable will take the value from minus 0.5 to 1.2. But in question it was between 45 and 62,

45 and 62 get transformed to minus 0.5 to 1.2. So, from the table I can get this value, what is the probability of getting 1.2 - ∞1.2.

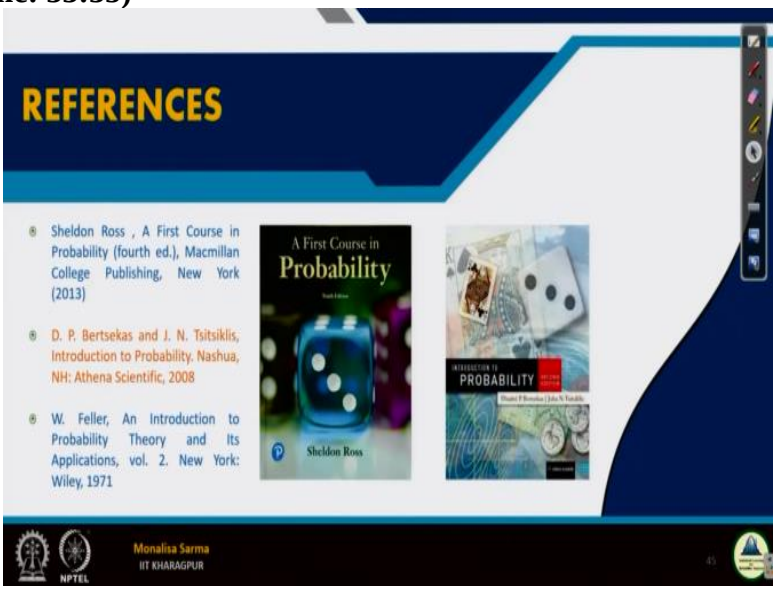
Then again, I can get what is the probability of getting the value minus 0.5. I will get this value. So, from this whole value, if I minus this value, I will get the blue portion that is the area of this blue portion that is the probability. So, that is how we solve this sort of problem.

(Refer Slide Time: 33:10)



So, we will be doing some real application of normal distribution in our tutorial classes. So now in this lecture, basically, we learned continuous uniform distribution, we learned what is normal distribution, what is standard normal distribution. And we have solved some problems though more problems we will be solving in our tutorial classes. And in the next lecture, we will discuss some more continuous probability distribution with that.

(Refer Slide Time: 33:35)



These are the references. Thank you guys.

Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury of Quality and Reliability
Indian Institute of Technology – Kharagpur

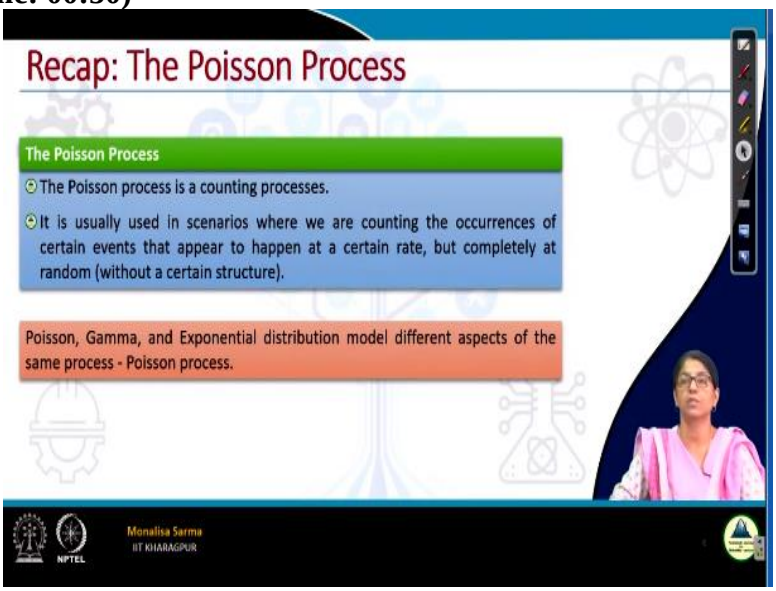
Lecture – 13
Continuous Probability Distributions (Part 2)

Once again hello all of you, so in continuation to our earlier discussion on continuous probability distributions where we have studied uniform probability distributions, normal distribution and standard normal distribution. Now we will be studying some more continuous probability distributions.

(Refer Slide Time: 00:42)

So, in this lecture basically we will be studying gamma distribution, exponential distribution and lognormal distribution.

(Refer Slide Time: 00:50)



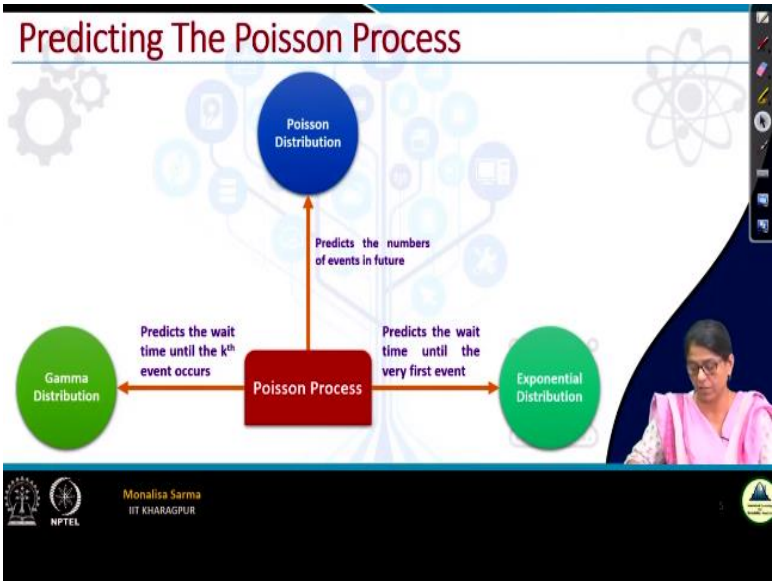
Now, before going to gamma distribution, let us take a quick recap on the Poisson process. Now, why am I suddenly bringing the Poisson process you will see? So, what was the Poisson process if you can remember Poisson processes was nothing but it was a counting process. Basically, I can say it is usually used in scenarios where we are counting the occurrence of certain events that appear to happen at a certain rate where the average occurrence rate is given.

So, when average occurrence rate is given, we are interested in the occurrence of a certain number of events within a time interval or within a region of space is not it? But this occurrence is completely at random. Like if let me take you an example suppose we know the occurrence of earthquake in a particular area, so particular area say maybe in square meters or whatever it is in occurrence of earthquake is given at a particular rate, say around in one year say around 5 earthquake has happened. So that is the average rate.

Now, we know in average occurrence rate is 5% 5 earthquake, but does not mean that we will know exactly this when earthquake will happen within an interval within in 1 year. Moreover, if in 1 year 6 earthquake happened, that does not mean that in the next interval, since the indices or 6 earthquake has happened that does not mean in the next interval next year, maybe we will get less number of earthquake it is not that as well.

So that we have seen that was the property of memory-lessness so that is what this occurrence is totally random, that was Poisson process we have seen. Now, why I bought out this poisson distribution here because Poisson distribution, gamma distribution and exponential distribution, they model different aspects of the same process that is a Poisson process, that is the counting process, they model the same aspect but in a different way.

(Refer Slide Time: 02:46)



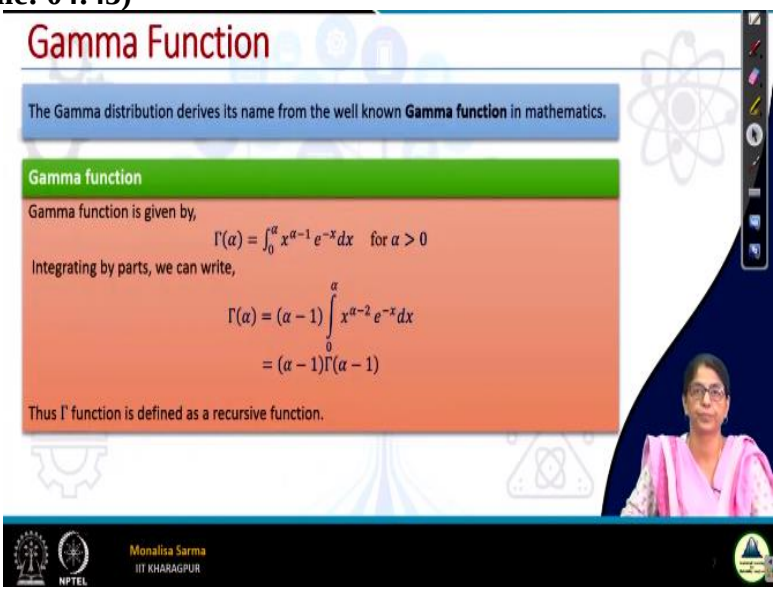
How and Poisson distribution as we have seen, we predict the number of events in the future, given the average occurrence rate, we know what is the how many events is on an average the number of events that occur within a time interval, then based on that we can predict a number of events that may happen within it in a time interval. That is what we do in Poisson distribution. Whereas exponential distribution we say exponential distribution, we try to predict the occurrence of an event.

Maybe the occurrence of a random event, maybe the occurrence of a failure when you talk in reliability studies, I maybe talk a failure occurrence a failure of a component or maybe the random event maybe like my bus coming in coming / bus at a particular time. So, basically it predicts the occurrence of an event or rather I should say it predicts the wait time until the very first event.

So, when is the time that the first event has occurred or we can also the data say time between 2 events. These, all these are can be modelled by exponential distribution. Now, what is the gamma distribution? Gamma distribution predicts a wait time until the kth event occurs. So, see the difference Poisson distributions given the average rate, we can predict a number of events that can happen in the future exponential distributions, we model the wait time between for the first event or the wait time between 2 events.

Or we can also say occurrence of an event it models the occurrence of any random event, this gamma distribution, it predict the wait time until the k event occurs. The time that pass so that k event occurs. Now first we will be discussing gamma distribution.

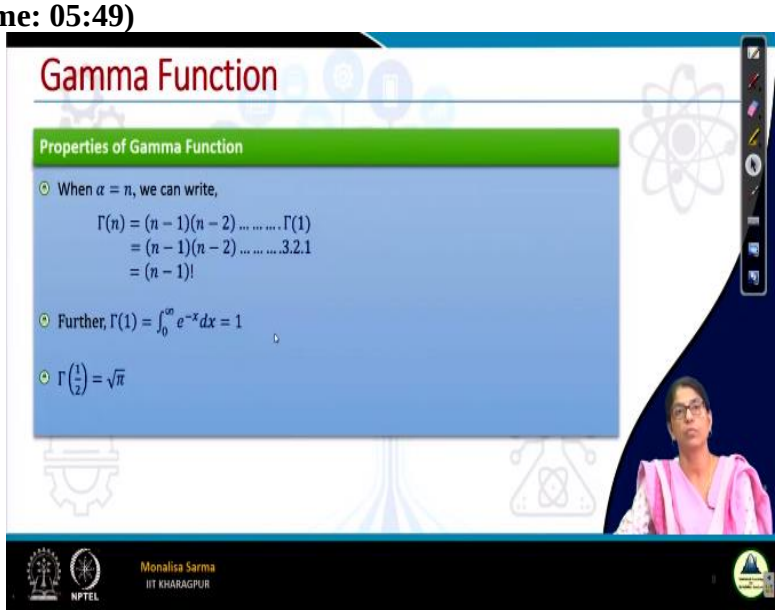
(Refer Slide Time: 04:43)



Gamma distribution, this name it has got from the gamma function, gamma function most of you I think you already know what is a gamma function. It is basically what we have factorial for integer numbers, this gamma function it is a representation of that only, but it is in the complex numbers and real number for factorial is always for integer number. So, gamma function is a representation of that, but in complex and real numbers.

So, gamma function we can specify this is the gamma function. So, this is integration means, if we will have to do integration by parts I told you on one of the class will have to learn integration so, how to do integration the part so, if integration the part so, we get gamma of alpha is this is the thing that means, if you are interested in finding Γ of 3 that means what is Γ of 3 that is 2 into Γ of 2 and what is Γ 2? Γ 2 is 2 into Γ 1.

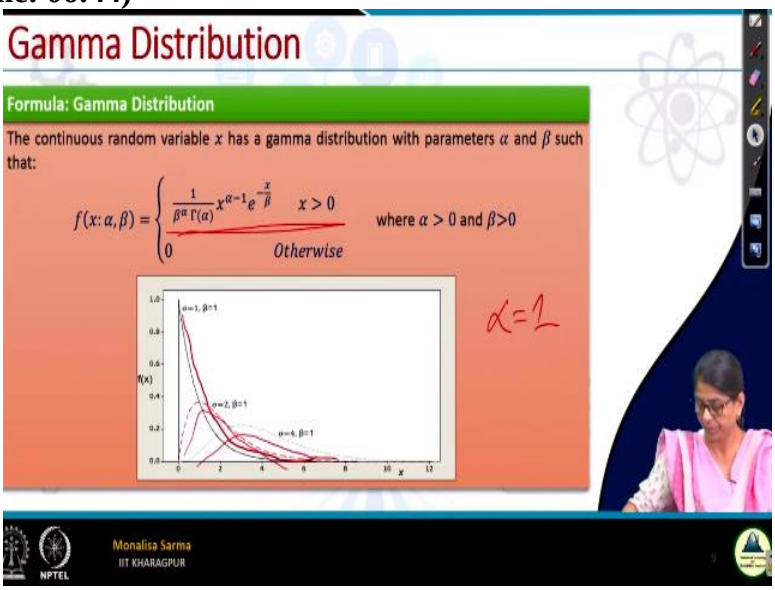
(Refer Slide Time: 05:49)



So, I have also here some of the properties of the gamma function. So, if α = n than n means, if it is an integer then we can write it is n - 1 Γ of n is n - 1 factorial whereas, factorial of n is n factorial. So, Γ of n is n - 1 !, for the gamma of n is nothing if you can put there instead of n if we put the value we will get in the expression the last expression what we got here if we put the value α = 1.

So, we will get Γ of 1 is 1, Γ(1/2) is √π these are some standard things which you can remember and of course, you can find out a gamma function for any given value. So, from this my gamma distribution has this function in it the gamma density function that density function corresponding to gamma function has this function in it and so, we call it as a gamma distribution function.

(Refer Slide Time: 06:44)



So, gamma distribution function has 2 parameters again here what are the 2 parameters? A and β. So, α is basically the shape parameter like till now, we did not get the shape parameter normal distribution did not have the shape parameter here we have to separate the way you see the see for different values of α we get differentiate. So, here for α = 1 this shape α = 2 we know got this sort of shape this one. This α = 2 and α = 4 we got this shape.

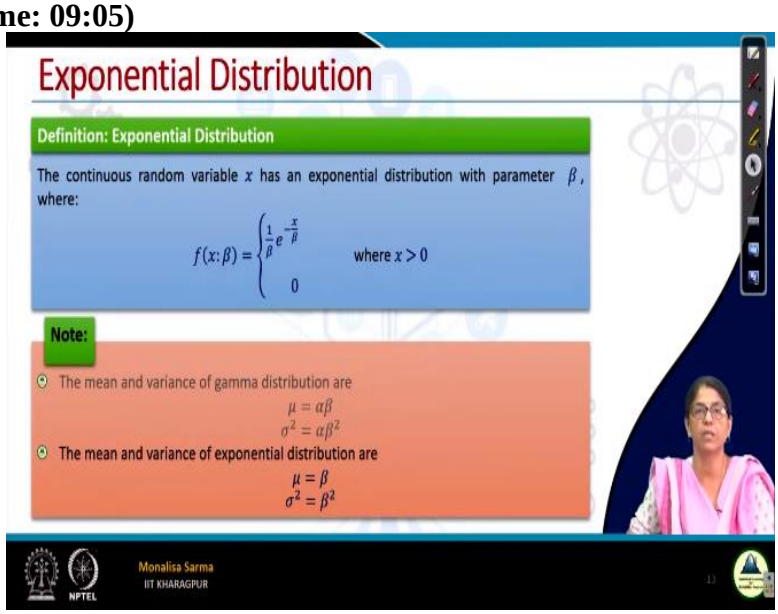
So, we got different shape for different values of α. So, α is that is why it is called a shape parameter. So, now, what is β here? B is the rate of event occurrence. So, if the β is rate of even occurrence that means, what is in terms of λ, what is λ in this Poisson distribution we have λ. Λ is a number of events that occur within a time average number of events that occur

within a time interval. So, if the average number of events that occur within a time it is suppose x events occurs within a time interval.

So, what is the thing rate of the occurrence of event? Rate of the occurrence of event will be 1 / x is not it? So, β is nothing but β = 1 / α. So, here like exponential distribution so, this is the density function for exponential distribution basically, now here are sorry this is the density function for gamma distribution. Now, exponential distribution is a special case of gamma distribution special case in the sense like in this gamma distribution.

If we put α = 1 resulting distribution is exponential distribution. So, here in this figure this α = 1 this is exponential distribution this is the curve corresponding to exponential distribution. So, this is the exponential distribution we call it this special case of gamma distribution. So, exponential distribution again it is, there is just 1 shape so, there is no shape parameter value of α is just 1.

(Refer Slide Time: 09:05)



So, exponential distribution what is the formula same in the gamma distribution probably density function putting α = 1 whatever we get that is definitely see here if I put α = 1 what do I get α = 1 Γ (α) Γ (1) is 1, so, β α that is β then x α^{-1} that is 1 – 1 x tends to the 0, so, $e^{-x/\beta}$ so, what I got 1 / β $e^{-x/\beta}$ is not it? Here if I put α = 1 in this expression if I put α = 1 what do I get?

I get 1 / β $e^{-x/\beta}$ is not it? That is the probability density function for exponential distribution. See, that is the probability density function for exponential distribution now what is β? B is

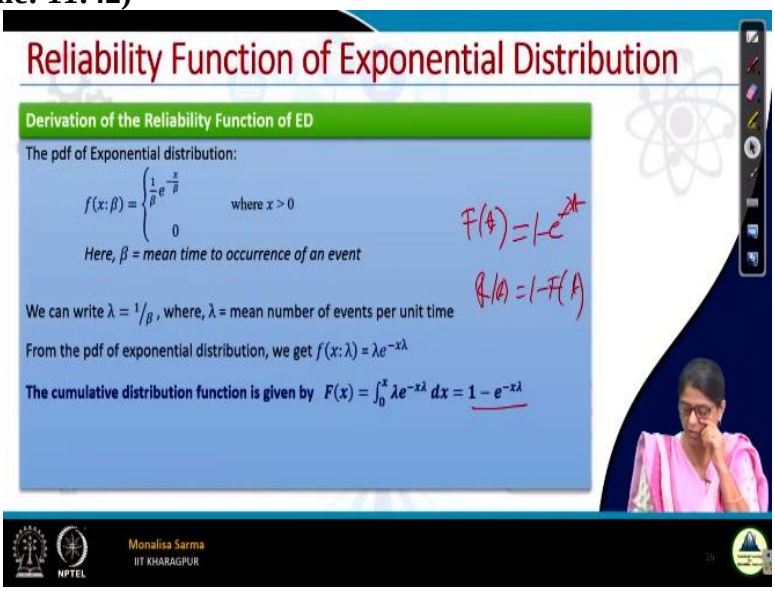
nothing but it is the rate of occurrence or event rate of occurrence of event means, if the average number of events that occurs within a time and time interval is λ so that means β is 1 / λ.

So, this if I write it in terms of λ that is the parameter of Poisson distribution λ was the parameter Poisson distribution is not it? Λ gives the average number of events that happens within a time interval. So, if I write this expression in terms of λ, so, what it will be my expression becomes $\lambda e^{-\lambda x}$ if I write in terms of λ. So, now the mean and variance of gamma distribution is mean is α β variance is $\alpha \beta^2$.

Now, if this is the mean and variance of gamma distribution, so what will be the mean and variance of exponential distribution then, from this exponential we will be able to say how putting α = 1. So, if we put α = 1 we get μ = β, σ^2 as well a $beta^2$ that is all. So, mean and variance of exponential distribution is β, β or we can tell also 1 / α and variance is β^2 that means, standard deviation σ is β.

So, mean and this is also one important characteristic which is to remember mean and standard deviation of exponential distribution is seen that is β or 1 / λ mean and standard deviation and variance mean and standard deviation of exponential distribution is 1 / λ or we can say β.

(Refer Slide Time: 11:42)



So, see here I have mentioned where λ is mean number of events per unit time. So, we can write λ is equal to where β means mean time to occurrence of an event, this is mean number

of events per unit time, say remember the difference. Now, this is F(x) this is the probability density function. So, if you are interested in finding the cumulative distribution function like how we have found it for normal.

Normal from what to say integrating from - ∞ to the particular value we are interested similarly, we can find out the cumulative distribution function for gamma also gamma distribution also exponential, exponential distribution also now, let us see for exponential distribution, if you are interested in finding out the cumulative distribution function. So, what it will be F (x) is integration of 0 to x $\lambda e^{-x} \lambda$. So, what we got this is the value.

Now, one thing this exponential distribution normal distribution, all this distribution has gamma distribution, this is where it used in reliability studies. So, when we use in reliability studies, we basically do not use x as a random variable. The general convention basically instead of x we write t because in reliability and reliability studies, we basically try to model the, our random variable is the time, time to failure time to repair whatever it is. So, when we talk of time, so in the variable, instead of using the variable x, we use the variable t that is all.

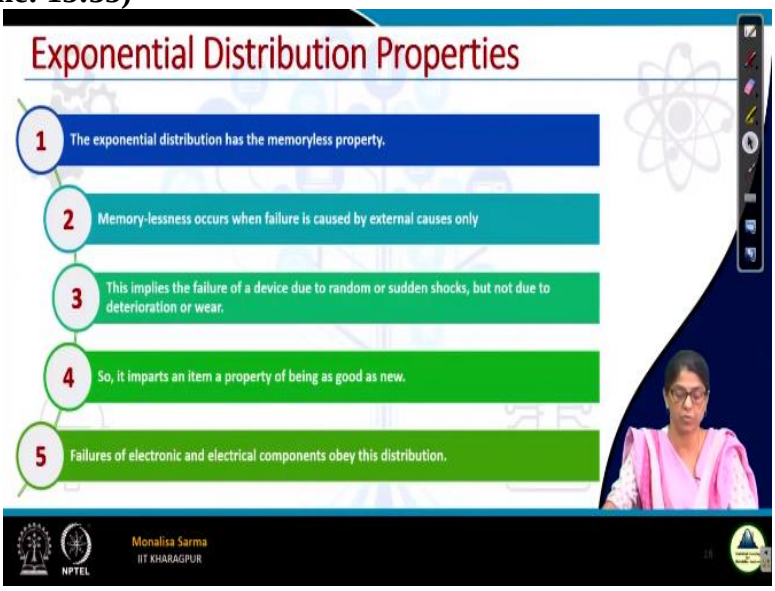
So, that is just a general convention that is all. So, now, this is the cumulative distribution function $1 - e^{x} \lambda$ just simple integrating the probability density function, so, if this is F(x), now, let us what is reliability? Reliability is the probability of failure free operation is not it? So, what is F(x)? F(x) is the probability that it will fail from -∞ to x probability that it will fail till xth time what is x? x is nothing but if I talk about reliability it will be in terms of t in terms of time.

So, what is F(x) instead of writing F(x), let me write this F(t), what is F(t)? F(t) the integration from 0 to t means humility failure probability of failure till time t. So, probability of failure till time t I got $1 - e^{-} \lambda^{t}$, this is my probability of failure. Now, if this is my probability failure then what is the reliability? Reliability is probability of failure free operation this is Ft normal reliability is nothing but 1 minus of Ft.

Similarly, this here I have shown in exponential distribution same case you can apply for all normal distribution. Normal distribution also you can find out Ft then Rt is 1 minutes of Ft. Ft how same integration from 0 to t using the probability density function that is Ft then R(t) is 1 minus Ft this of course, I mentioned it here, but you can use the same process what will be

the difference and nothing different. So, it is the same process we can apply for all the other distribution as well. So, my reliability is $e^{-x}\lambda$.

(Refer Slide Time: 15:35)



Now, this exponential distribution it has a property one of the most important property is that it has a memory-lessness property what is this memory-lessness property in case of exponential distribution like, what does exponential distribution failure? Exponential distribution it has a constant failure meaning like suppose if I talk about vehicle, vehicle is not a very in the normal useful remember when I talked about the failure rate that is the fact of that.

First usually it is the quality failure we say that is the infant failure rate then is the normal working life as a normal failure then is the wearout failure is not it? In a normal working life the failure that happens is maybe due to random reasons, this may not be because of some problems in the machine may not be some quality related failure it may not be because of wearout failures in a normal working life when a failure happens it failed due to some random reasons.

So, in this time the failure that happens is because of it has a constant failure rate. So, constant failure rate meaning is that so, suppose if I talk of a vehicle only in its useful life that suppose a vehicle that has failed at time t = 100. So, now at t = 100 a vehicle that has failed, it will not be different if I consider that the vehicle has run for 1000 kilometre before that, or the vehicle has run 1000 kilometre before that just because the vehicle has run 1000 kilometre before that does will have any affect that.

If I tell that I am at a certain point of time, if I consider what will be the probability that it will fail after 100 hours. Suppose the vehicle has run 1000 kilometre and I am asking so what is the probability that it will fail after 100 hours? Again, there is another vehicle which has run say 100 kilometres only then I am asking what is the probability that it will fail after 100 kilometres if this failure was a failure can be represented by exponential distribution.

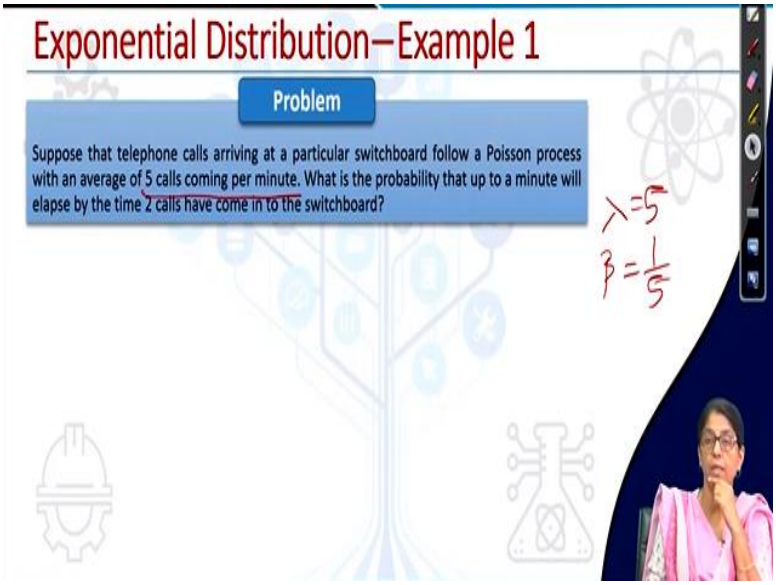
So, if this failure shows a constant failure rate, then whether it run 100 kilometre or a it run 1000 kilometre it makes no difference probability that it will fail after 100 hours in both the case will be same this happens why this sort of failure will only be seen in case of random events so random events may not unless it is written here. So, memory-lessness this failure x what is it exponential distribution like this constant failure rate?

It happens when failure is caused by external causes only maybe due to some external causes, but maybe the external causes. So, this implies that failure of a device due to random or sudden shocks but not due to deterioration or wear maybe due to sudden shock. Then so, it imparts an item the property of being as good as new whether it has 1000 kilometre or it has an 2000 kilometre does not make any difference.

Failure of electronics and electrical components obey this type of distribution, this exponential distribution it always electronic and electrical components obey this distribution even for software failure also software failure we see this type of distribution when we talk about software here or only we are talking of hardware of course software also we can think of software reliability. So, when we took me think of software failure.

Its software failure also what to say it can be model using exponential distribution here also the failure rate is due to some random event here random for a software random event not been may not be due to environment maybe some other causes, which I will not dwell into details.

(Refer Slide Time: 19:27)



So, we will see some examples here. Suppose that and telephone calls arriving at a particular switchboard follow a Poisson process see here, it follows a Poisson process. Poisson process means it is a counting process, basically. How many calls arrived at a particular switchboard? That is why is calling the Poisson process. So, we are interested in the events, different events. So, what is with an average of 5 calls coming per minute, so per minute, it is 5 calls that means my λ is 5 calls.

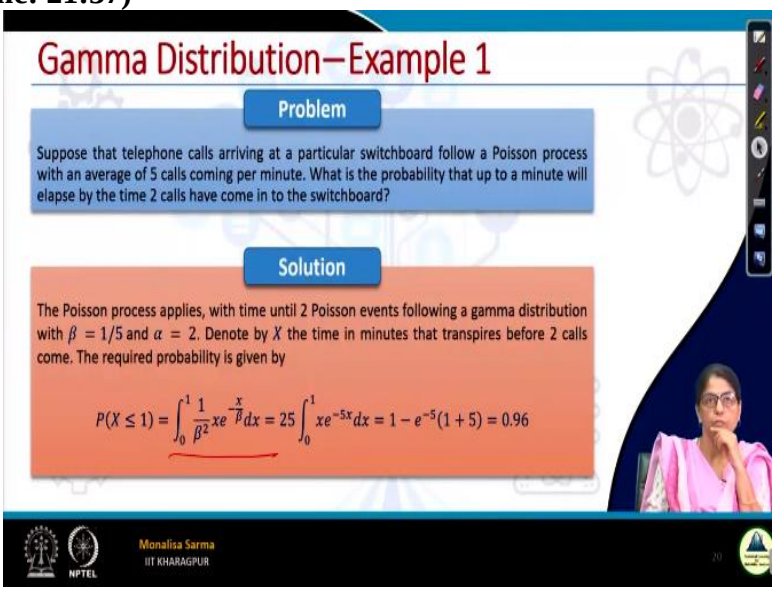
It my λ is 5 then what is my β? B will be equals to 1 / 5. So, here it is given 5 calls coming per minutes that means my λ = 5 and Poisson distribution we have seen this is not it? So, now what is my β? B will be 1 / 5. So, what is the probability that up to a minute will elapse by the time 2 calls have come into the switchboard? See here what it is asking, what is the random variable here?

Random variable is definitely the time always I am repeating again which I repeated in my last few classes also always try to find out what is the random variable. So, what is the random variable here? The random variable is the time but it is asking time up to a minute with the last by the time we need, what is the probability that 2 calls have come? So, my number of calls is 2 my number of call is 2 and it is telling up to a minute what is the probability that 2 calls will come?

If number of calls is 2 that means see if you see the first slide well I have to explain the difference between the Poisson process let me come to those that slide this one see here. Poisson exponential and gamma, exponential is the first event order time between 2 events.

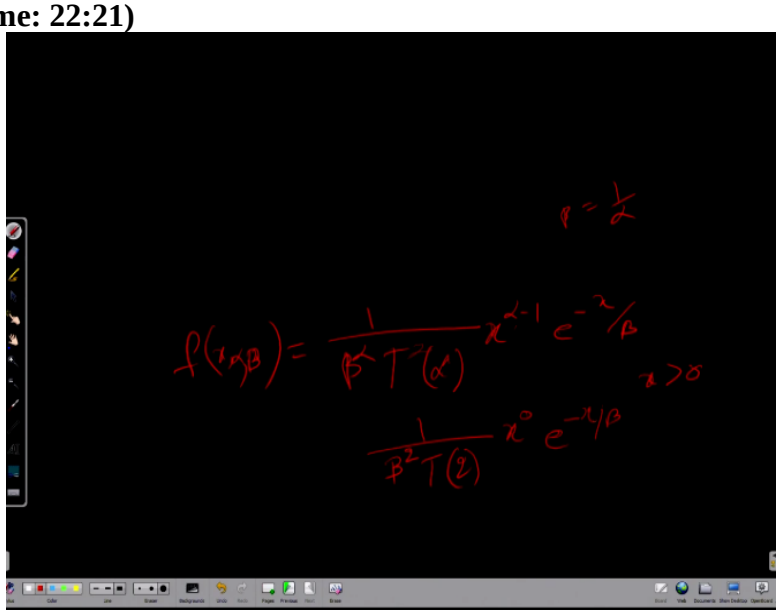
What does gamma says predict the wait time until the kth event occurs? So, wait time and kth event so here kth event from 0 to 2 events. So that means this question is a question of gamma distribution.

(Refer Slide Time: 21:37)



So, it is a question of gamma distribution. So, simple whatever is given it is now it is a simple question we know that it is equation on gamma distribution, we know the value of β what is the value of β? B value is 1 / 5 we know what is the value of α? A = 2 just simple put it in the expression; put it in what to say expression for gamma distribution?

(Refer Slide Time: 22:21)



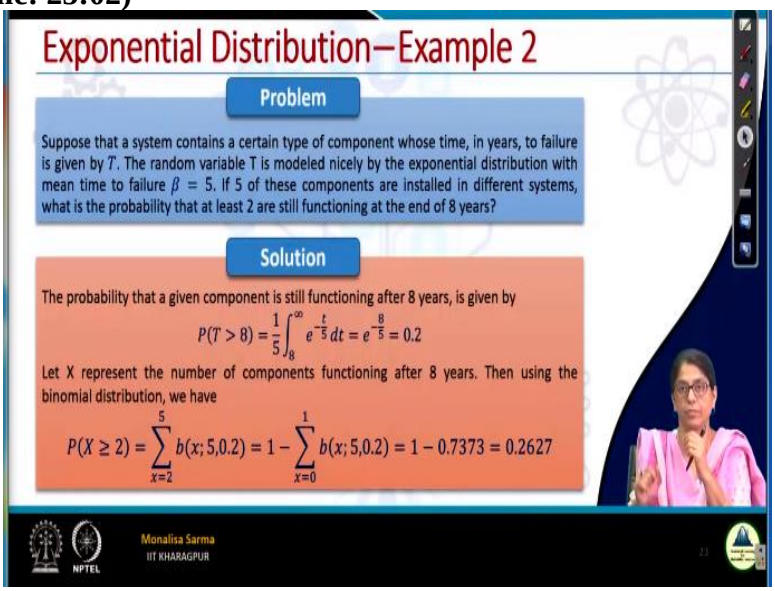
So, you remember what was the expression for gamma distribution it is $1 / \beta^{\alpha}\, \Gamma(\alpha) \times \alpha^{-1}$ then it is $e^{-x/\beta}$ where x > 0 this is f (x α β) this is the formula is not it? Now, here α is given 2. So, what is Γ(2)? So, α is given 2 β is we got β is equals to 1 / α, α is 5 so β is what? β = 1 / α. So, α is 5 so β is 1 / 5 is not it?

So, afterwards we will take substitute a value for β or first let us have to value of 2 α and that is α is 2 and Γ(2) what will be to Γ of 2? Γ (2) will be 1 Γ(2) will give us 2 into 1 !. So, it is basically 1 into 1 !, so Γ(2) will be 1 so x of α - 1 so 2 - 1 that is x of 0, x of 0 is 1. So, $e^{-x/\beta}$ so, this is the expression. So, here this is the expression.

So, simply by solving it whatever we got that is the probability that up to a minute will elapse by the time which we will get 2 events. See the difference between the Poisson distribution and gamma distribution. In Poisson distribution what we tried to find out what was the probability that 3 events have occurred, what is the probability that 4 events have occurs in a particular time? There are random, what is the random number? Random number was the number of events.

Time was fixed and we wanted to find out the number of events within that time interval. Now, here what is the random number here? Random number is time t, in time t we are interested in finding out what is the probability that in time t 2 events is occur. The random number is the time and in the interval.

(Refer Slide Time: 25:02)



So, next is another one question. So, what it is given suppose that a system contains a certain type of component whose time in years to failure is given by T. The random variable T so, T is the random variable it is mentioned only its modelled nicely by exponential distribution here our job has become very easy it has already explained that it is exponential distribution

we do not need to worry last question it was not given what distribution is that we need to find out by seeing the problem.

So, here it is already mentioned that it is exponential distribution. So, T is modelled nicely by exponential distribution with mean time to failure that is $\beta = 5$ at a specified β this is one part of the question then it is if 5 of these components are installed in different system, what is the probability that at least 2 are still functioning out of 5 2 are functioning at the end of 8 years. See the last part last sentence if 5 of these components are installed in different system, what is the probability that at least 2 are still functioning at the end of 8 years.

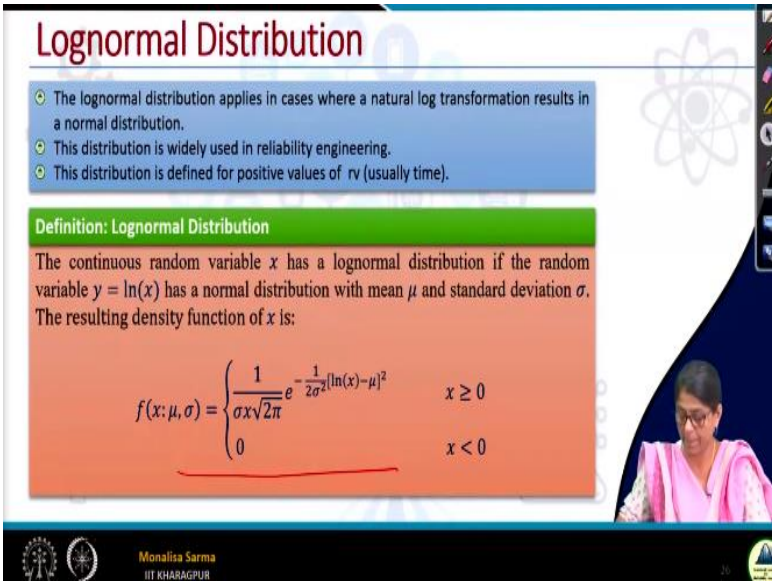
Does not it sound similar to binomial distribution the last portion out of 5 trials 2 successes is not it? Out of 5 trials if there are 2 successes that is what the 5 of these components are installed in different system what is the probability that these 2 are still functionally at the end of 8 years. Now out of 5 2 are still functioning this is a binomial distribution. Now, the binomial distribution means what is the probability of success and binomial distribution to probability of success that means here.

What is the probability success probability that the component is still functioning at the end of 8 years that is my probability of success here probability that the component is working till at the end of 8 years. Now, how do I find a probability that a component is working at the end of 8 years as it is already told that it is exponentially distributed, since it is exponentially distributed, I can just put it in a formula and then I can find out what is this probability?

Probability of $T > 8$ is nothing but putting just simply putting it in a formula of exponential distribution $1 / \beta^{-t} \beta$ dt, β is 5 so $1 / \beta\ e^{-3/5}\ 1 / \beta$ is not it? dt, so, I got this is my probability. Now, my probability of success means probability that it is functioning at the end of 8 years is 0.2. Now simple binomial distribution at least 2 are still functioning means 2 are functioning, 3 are functioning for 2 + 3 + 4 + 5 thus instead of doing that.

I can do it 1 - 0 to 1, the 0 to 1 means does a cumulative from 0 to 1 it is cumulative. So, cumulative means I can see it from the table. So, next one is lognormal distribution.

(Refer Slide Time: 28:06)



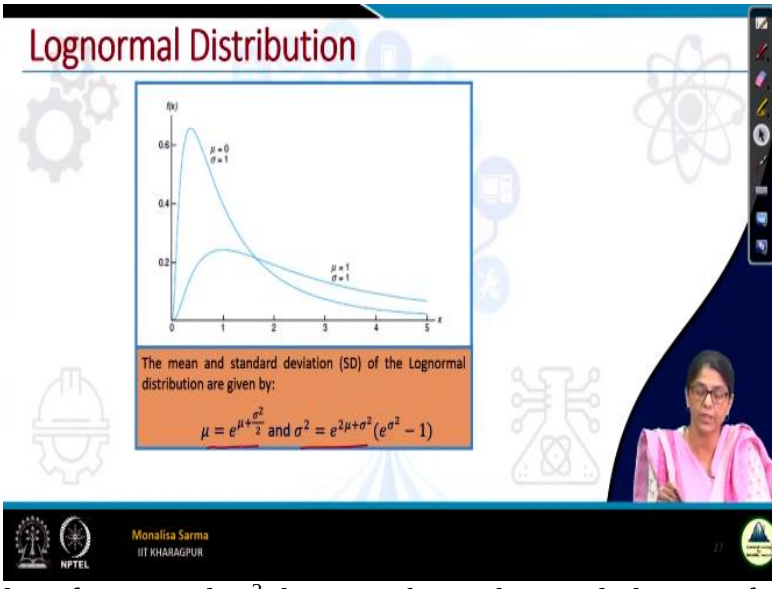
Lognormal distribution also like exponential distribution and gamma distribution is a widely used distribution. Till now, the distribution that we have studied exponential, gamma and normal are widely used distribution normal mean standard normal basically. So, now lognormal distribution is also has also it is used in reliability studies. So, this mainly it is used to model the repair times not the failure times when you talk a failure times, we do not use lognormal distribution, we use normal distribution or maybe exponential distribution.

But lognormal distribution we use when we try to model the failed repair time of a component. So, and lognormal distribution applies in cases where natural log transformation result in a normal distribution. So, if x is log normally distributed, what the person means, if x is log normally distributed, then log of x is normally distributed, the meaning of the first statement is that if x is log normally distributed, then log of x is normally distributed.

So, this distribution is widely used in reliability known mainly for modelling the repair times, this distribution is defined for positive values of random variables only usually 10. Definitely when we talk about repair time's repentance cannot be negative. So, this distribution is defined for positive values of random variables only. So, this is the probability density function. This is the probability density function for lognormal distribution.

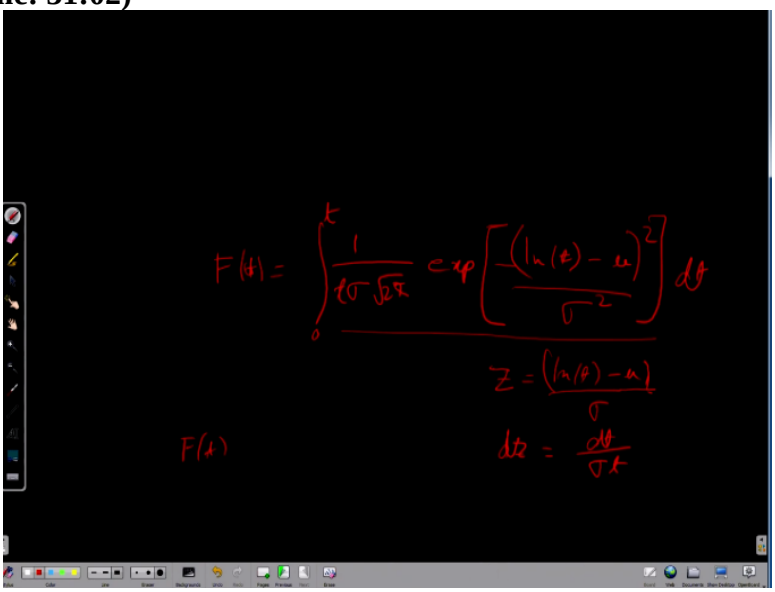
Similarly, we can find a cumulative, if you do the cumulative distribution function that is from - ∞ to the particular value x whatever it is, then we see when we can easily use the standard normal distribution table to find out the values are lognormal as well.

(Refer Slide Time: 30:02)



So, this is the value of mean and σ^2 this just what is the simply how we find out the mean? Mean is $\int f(x)$. Variance is $x - \mu^2 f(x)$ by solving that if we know by putting that in this value and by solving it, we will be getting this 2 value. Now, there is one problem which you have to solve, before solving this problem, we will this let us see this question how as I told you this lognormal distribution for when we try to find out a cumulative lognormal distribution, we can use the standard normal table we can very well use the standard normal table for that, let us see how we can do that.

(Refer Slide Time: 31:02)



So, what is the lognormal if I write it in terms of cumulative, so, it is F of t is equals to I can write 0 to t F (t) means, cumulative means always it is from - infinity to a particular value. So, here since I am talking of what to say in the lognormal, it is always a positive value. So, I am not a minus infinity I am taking 0. So, 0 to t that is 1 by what is the expression $1 / t \sigma \sqrt{2}$

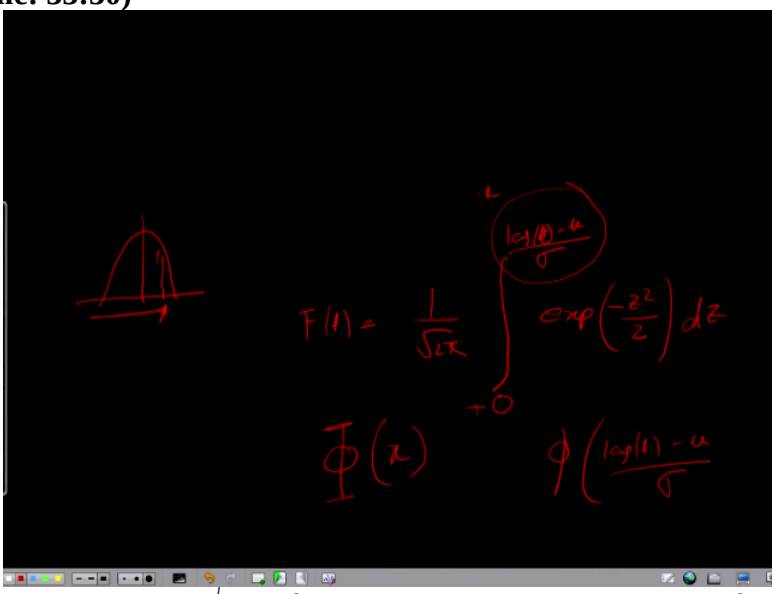
pi, it is a very big expression, it is difficult to remember so, it is always a better to use the lookup table.

So, it is exp (- ln (t – μ / σ ^{2}) and this is whole square dt this was the expression for Ft probability density function. So, now, when we want to cumulative that will be integration from 0 to t. So, this is the expression for cumulate. Now, if I substitute say Z is equals to earlier case I substituted Z = x - μ / σ ^{2} here since, F t is log normally distributed log t is normally distributed.

So, I can substitute Z = ln t - μ / σ ^{2} - μ whole square by saying sorry I will not use the square term here that made up I will not use this square here. So, this is my Z what will be my dz? dz = dt / σ t is not it? So, now, substituting this values like how we have done for normal similarly, substituting this ln t - μ / σ ^{2} if I substituted by Z then what will be and similarly substitute dt / dz σ t then what I will get?

I will get the same whatever I got it for the normal distribution so, I will get it in my Ft firstly, let me take a separate space from this.

(Refer Slide Time: 33:30)

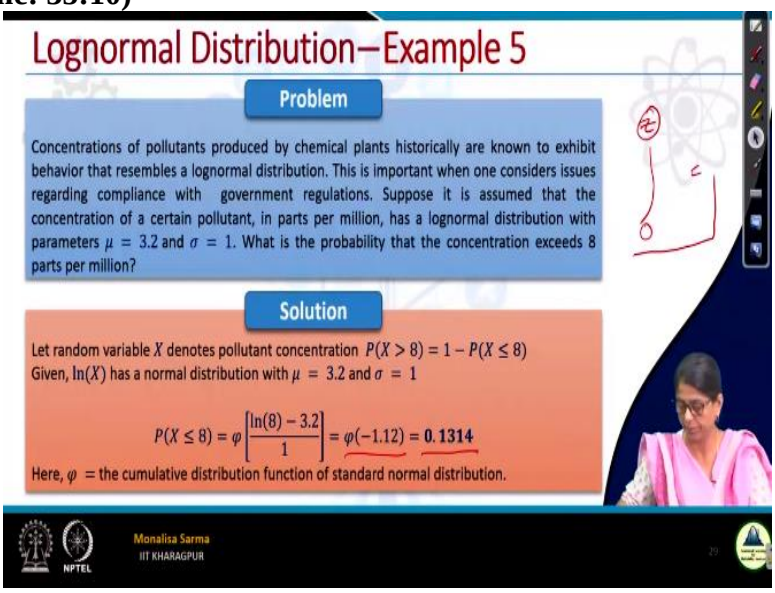


So, my Ft will be equals to 1 / √2 π after substitution whatever I got it for standard normal similarly, this is that it will be from sorry it is not minus 0 to log of t - μ / σ because here the range is also changes earlier it was from x. Now, what is x = log of t - μ / σ I have substituted that so, it becomes exp – Z^2 / 2 dz this is the same thing, but we got it when we have done for standard normal distribution instead of this value here we had what here we had x.

Now, x or t whatever it is now, here we are having log of t - μ / σ that is all. So, now, for this we have referred to the table that is standard lookup table at what point you have to refer to the table. Table usually we call this value this is the symbol we use earlier case we have referred to this value it is phi of x means at portion x is not it? If this was the normal distribution suppose, this is my x, this is phi of x this gives me this value now, here so, what I will find out from that normal table?

I will find a phi this log of t - μ / σ that is all, that will give me my value of this area probability of this area from 0 to less whatever it is solving this problem will be clear.

(Refer Slide Time: 35:10)



So, concentration of pollutants produced by chemical plants historically are known to exhibit behaviour that resembles a lognormal distribution, this is information for us concentration of pollution produced by chemical plant it resembles a lognormal distribution. This is important when one considers issues regarding compliance with government regulation and other information. Suppose it is assumed that the concentration of a certain pollutant in parts per million has a lognormal distribution with parameter μ = 3.2, σ = 1.

So, it is given what it is given that the concentration of the certain pollutant in parts per million of course, it has a lognormal distribution with this is these are the parameters μ and σ. So, what is the probability that a concentration exceeds 8 parts per million. So, that is basically we have to find a probability of x greater equal to 8. And problem is greater equals

to 8 means, if I say this is if I convert if my x is a lognormal distribution, my log of x will be normal distribution. So, what I need to find out?

So, first this I will have to convert it to standard normal variable how log x - μ / σ. And basically, I need to find a, it is greater than what value, we will be getting this value suppose my Z value is this log of x - μ / σ, this is whatever is Z value. So, this is Z value from the table I can get this value. But what I needed is greater than this, I need this portion. So, it will be 1 minus of this I will be getting the value. See here P of X greater than 8 this now, simple put just finding out the Z value.

But I told you we just find out the Z value ln of 8 - 3.2. Because why we are finding out is Z value? Because integration is from 0 to this Z value. What was the Z value? ln of x = μ / σ and this whatever expression that is probability density function and that we do not have to calculate it because we know what the value corresponding to this just value, this probability, I am in the area for this value that we can get it from the lookup table. So, this we got Z value we got -1.12.

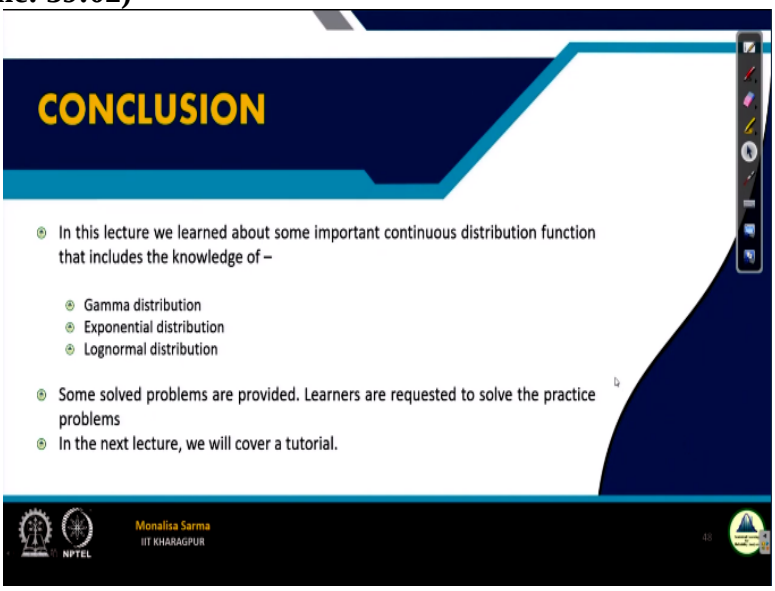
So, again from the normal distribution table, what is the value corresponding to 1.12 is this, we are interested in finding out greater than 8. So, 1 minus this will give us that value that is the value greater than 8. So, now, next distribution, that is there is one important distinction that is that Weibull distribution, but the Weibull distribution, I will not be discussing in this class, though this Weibull distribution, I have not included in the course also.

But I want I would like to mention that it is a very important discussion for reliability studies. But because this course audiences not only students who are from reliability, maybe none of them are from reliability background, you do not have any knowledge on reliability. So, Weibull distribution will be very a bit complicated for them to understand it is a bit high standard. So, what I am doing is that people who are interested I have kept the slides here, people who are interested you please go through this Weibull distribution.

And if you have any doubt we can clear in a doubt clearing session, but I will not be included in this course. That means when I am telling I am not including in this course you will not be you cannot expect an equation from Weibull distribution in your exam as well plus I will not be discussing here also, but if you have any doubt we can if anyone is interested in going

through Weibull distribution, then we can discuss it then any doubt we can discuss in the doubt clearing session.

(Refer Slide Time: 39:02)



So, now, I end this lecture on continuous probability distribution in this class we have discussed gamma distribution, exponential distribution and lognormal distribution. Of course, we have solved your problem also. And in the next class we will cover a tutorial.

(Refer Slide Time: 39:18)

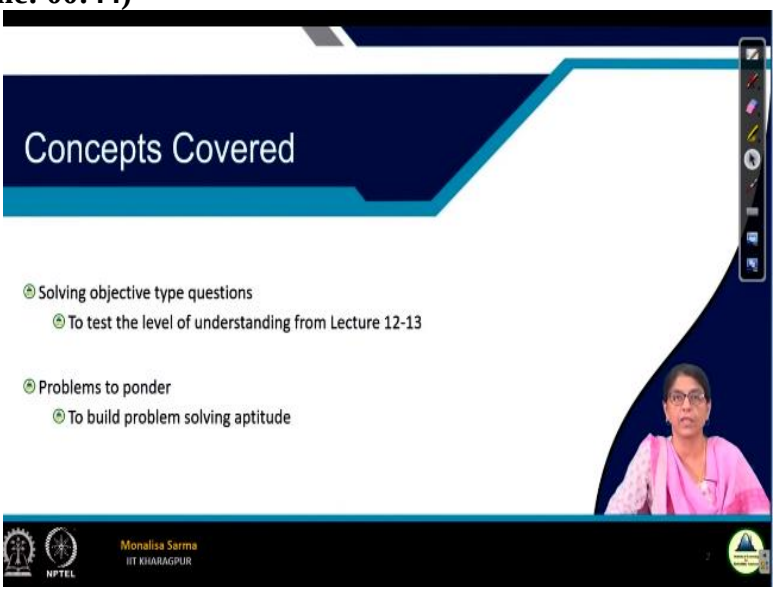
With this and I here I specify the references and thank you guys.

Statistical Learning for Reliability Analysis
Dr. Monalisa Sarma
Subir Chowdhury of Quality and Reliability
Indian Institute of Technology – Kharagpur

Lecture – 14
Tutorial on Continuous Probability Distributions Functions (Part 1)

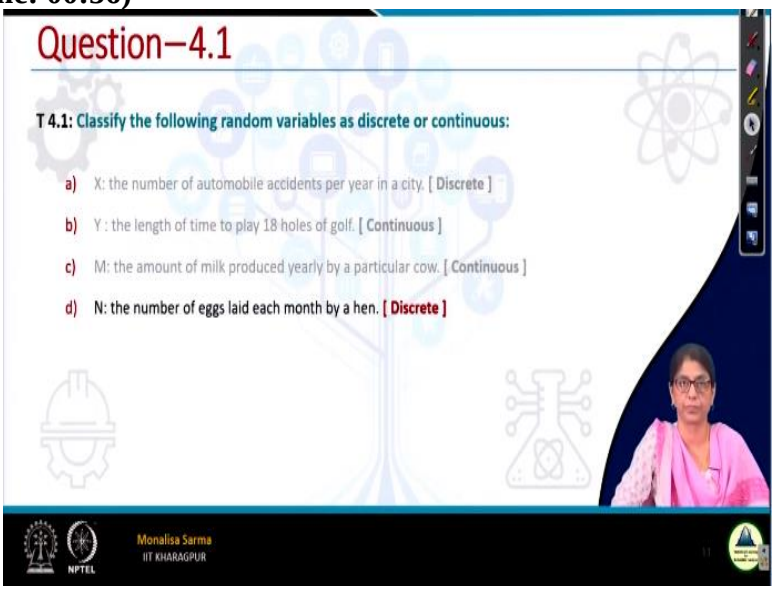
Hello guys. So, today we will be discussing some problems on the topics which we have covered in the last 2 lectures basically, that is lecture 12 and lecture 13 that is, we have discussed the different continuous probability distribution based on that we will be covering some of the problems so it is a tutorial class.

(Refer Slide Time: 00:44)



So, first we will be solving some of the objective type question which will be essentially test the level of understanding and then we will solve few problems.

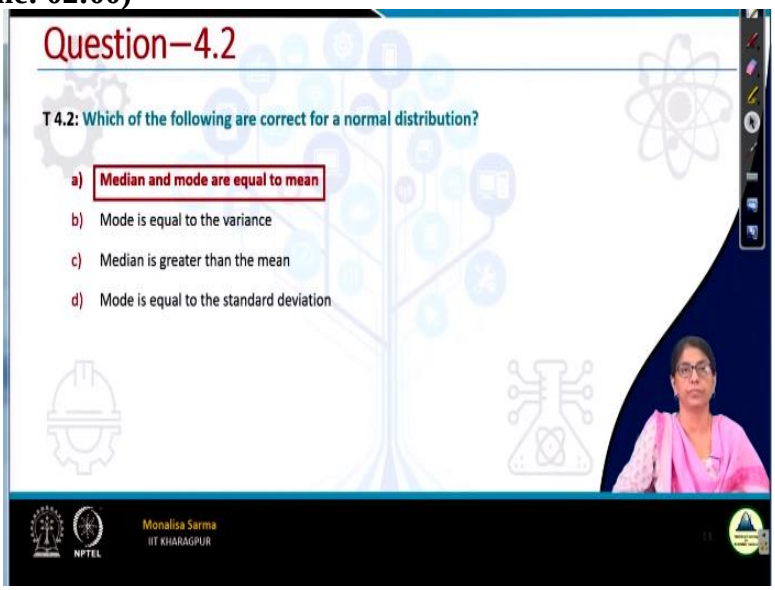
(Refer Slide Time: 00:56)



So, first the objective type equations the very easy questions, I am sure all of you will be able to answer it, but again, my request is that like in all the tutorial classes be do not see the answer immediately first try to answer yourself if you cannot, then go to the answer. So, classify the following random variable as discrete or continuous X the number of automobile accidents per year in a day.

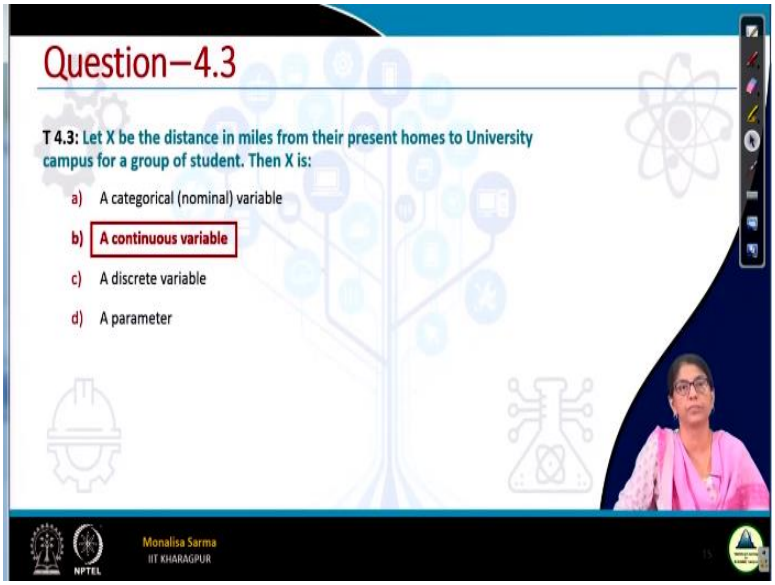
So, it is a discrete variable or continuous variable number of automobile accidents. So, number of automobile accident, it cannot take any possible values within an interval. So, it has to be a discrete variable. Then next is length of the time to play 18 holes of golf. So, when we talk of a length of time so, time it can take any value in an interval, so it has to be a continuous variable. Similarly, the amount of milk produced merely by a particular cow that also it is a continuous variable. Then the number of eggs laid each month by a hen that number of e m x that has to be an integer value. So, it has to be a discrete variable.

(Refer Slide Time: 02:00)



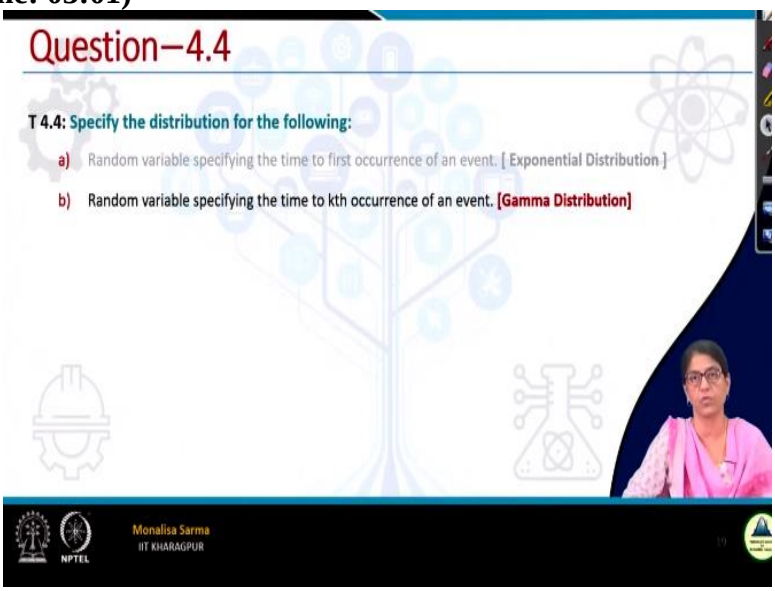
So, which of the following are correct for a normal distribution, for a normal distribution we have seen the mean is equals to the median is equals to the mode. Mode is always the highest value and in normal distribution, we have seen it is symmetric on towards the mean and 50% of the value is on the left side of the mean, 50% of the value is on the right side of the mean. So, that is why that is also the median. So, we have already seen at median we get the high speed. So, median and mode are equal to mean that is the correct answer.

(Refer Slide Time: 02:34)



Then let X be a distance in miles from their present homes to university campus for a group of students. Then X is what is X? X basically is the distance from their space wherever they stay the hostel or whatever it is from the hostel to the university campus. So, if X is a distance. Distance means it can take any value within a range. So, it has to be a continuous variable.

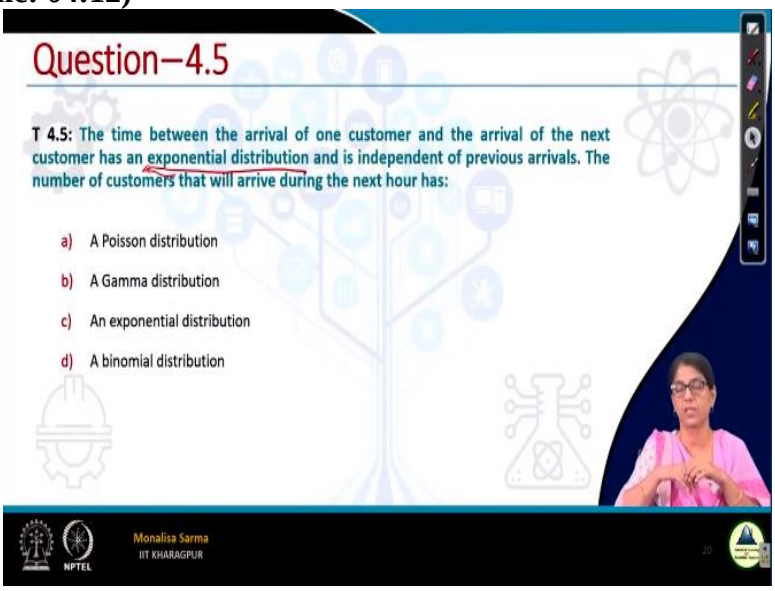
(Refer Slide Time: 03:01)



Specify the distribution for the following, random variable specifying the time to first occurrence of an event time to first occurrence of an event we have seen what distribution it is this? This is a exponential distribution than random variables specifying the time to kth occurrence of an event, kth occurrence of an event mentioned in a time we know that we are interested in finding out a specific number of events that will happen, when a specific number of events that will happen at done it has to be a gamma distribution.

Similar kind of thing we have seen in negative binomial distribution is not it? Very negative binomial distribution, it was the number of trials were fixed and we are interested in a fixed number of events, fixed number of success or failure whatever it is in a fixed number of trials are sorry number of trials are not fixed and negative binomial number of trials are not fixed, but the number of events are fixed. So, here also random variable is the time so here the time is not fixed, but we are interested till we see kth event k events so that is the gamma distribution.

(Refer Slide Time: 04:12)



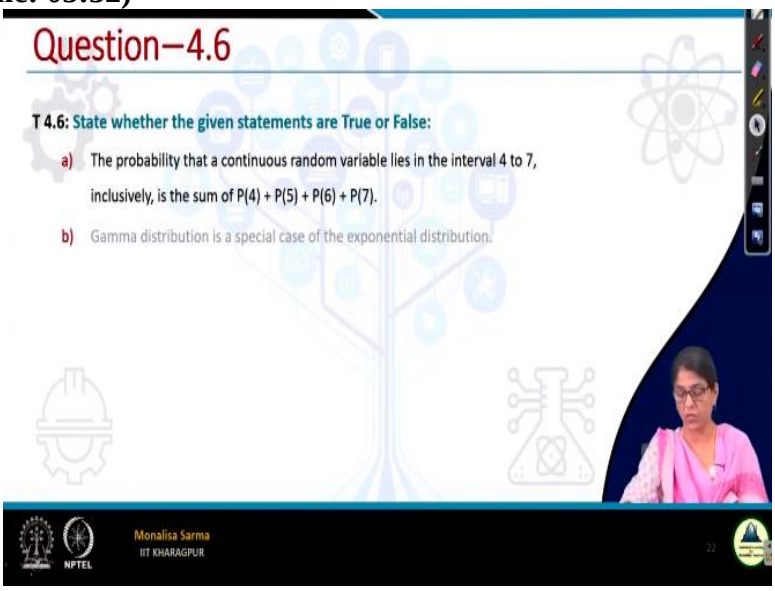
Interesting question the time between the arrival of 1 customer and arrival of the next customer has an exponential distribution and is independent of the previous arrivals. That is given the time between the arrival of 1 customer and arrival of the next customer it has an exponential distribution, that is an exponential distribution arrival between 2 events. We already know arrival between 2 events or time for arrival of the very first event that is basically we consider rate of the event it is all these are the model by exponential distribution.

We already know this and what it is asking the number of customers that will arrival during the next hour question is something else it has nothing to do with exponential distribution. It has specified data time between 2 events as it follows exponential distribution the number of customer that will arrive during the next hour. So, the number of customers that will arrive during the next hour that we have. So, what this is definitely Poisson distribution?

We know Poisson distribution we have seen given the average arrival rate we are interested in the number of events that will arrive in a particular length of time or in a space whatever it

is so that is Poisson distribution here also, it is the same question the number of customers that will arrive during the next hour, so it is a Poisson of distribution.

(Refer Slide Time: 05:32)

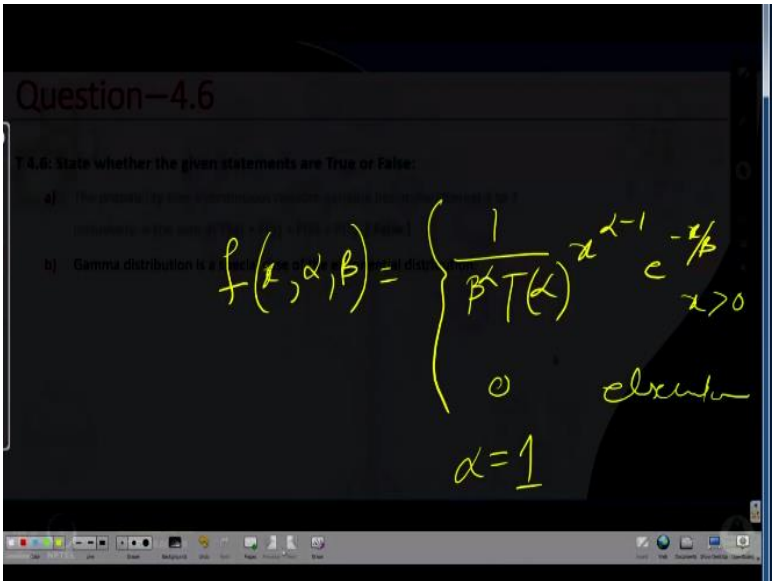


Here we to say true or false. The probability that a continuous random variable lies in the interval 4 to 7 inclusively is the sum of P 4 + P 5 + P 6 + P 7 it true or this false true? No, it is false why when we are talking about continuous random variable the probability cannot be in a continuous random variable we cannot specify the probability at a specific point I get the probability at P 4 will be equal to 0, probability at P 5 will be equal to 0, we cannot specify probability at a specific point whenever we have to specify probability it has to be within an interval.

So, when you are talking about probability that lies in the interval 4 to 7, it has to be integration of the probability density function whatever we have integration of probability density function is Fx. So, integration of Fx from 4 to 7 that will be the probability but not this whatever it is given here. So, this is false. Now, next is gamma distribution is a special case of exponential distribution.

Is gamma distribution is a special case of exponential distribution? No, gamma is not a special case of exponential distribution. In fact, exponential distribution is a special case of gamma distribution.

(Refer Slide Time: 06:45)

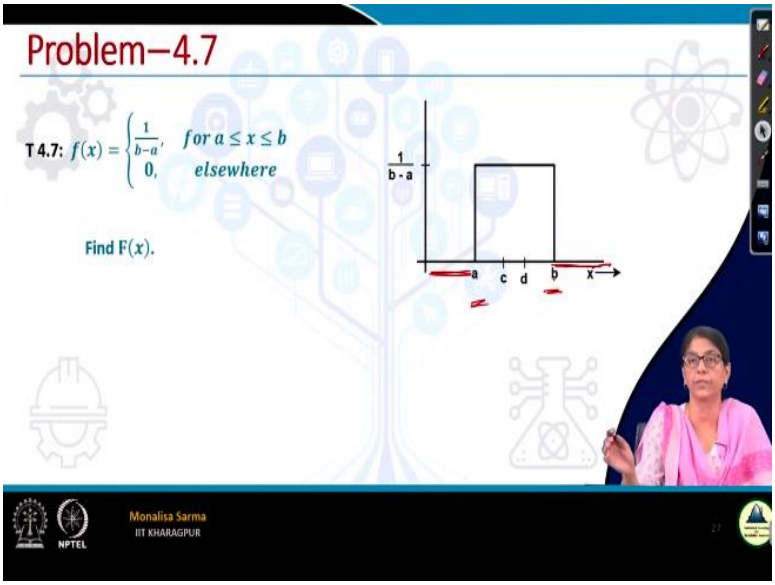


Let us see why if you can remember we have already done this, but still if you let us see whether you can remember it or not. So, like remember the formula for gamma distribution, so, we are interested in say time t and gamma distribution at 2 parameter α and β. So, what was the expression it is $1 / \beta^{\alpha}$ there is a gamma function of $\alpha\ x^{\alpha - 1}\ e^{-x/\beta}$. This is if you have considered here x then this will also be an x and let us make this x.

So, this is for x greater than 0 and it is 0 elsewhere. So, if x is the random variable and this is the gamma probability density function, so what is α here for gamma distribution we have 2 parameters α and β what is α here? A is the specific number of events that occur in time t α shape talks about a specific number of events that occurs in time t a fixed number of events as I told you we can draw an analogy from the negative binomial distribution so that is α.

And what is β? B is the rate of events rate of events we can say means time that was the very first event has occurred or the time between 2 events that is basically this is gamma distribution. Now, in this gamma distribution, if α = 1 if we specify α = 1 then the resulting distribution is nothing but an exponential distribution. So, we say exponential distribution is a special case of gamma distribution. So, this is also false

(Refer Slide Time: 08:35)

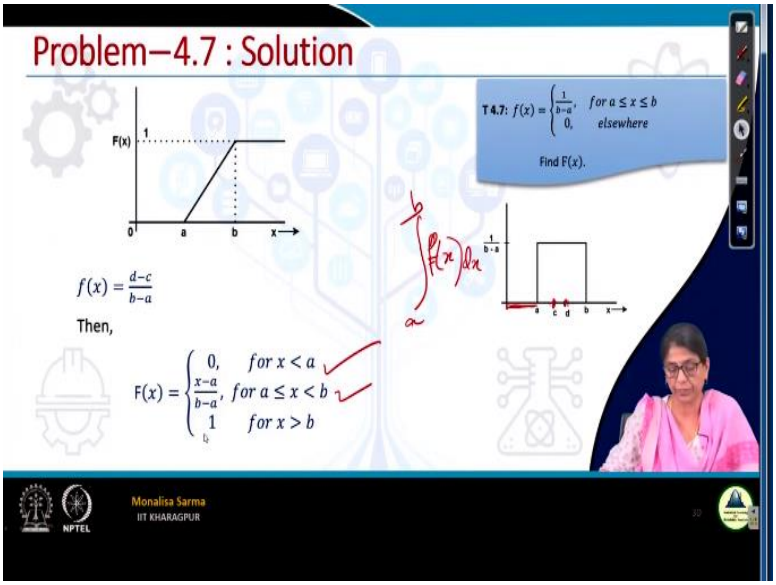


So, now we will be solving some bigger problems. So, here this is equation here fx is given that is called the probability density function is given this is a probability density function for uniform distribution we can see it is a uniform distribution and it is a range of our random variable that is here random variable is x random variable it ranges from a to b. So, in this range of a to b random variable can take values within the interval a to b and what is the probability density function? The probability the density function is 1 / b – a.

So, now, what we are asked? We are asked to find the F(x) that is the cumulative distribution function; we have already learned what is the cumulative distribution function? Cumulative distribution function means the value of the function t from - ∞ to that level to 2x when you are considering x F(x) means what is the probability of getting a value from - ∞ to x probability to getting a value from x taking a value from - ∞ to x so that is cumulative distribution function.

So, here there are see 3 different reasons one is before a that is the problem the density function F(x) = 0 before it is 0 and after this what happens? After this is also 0. Now we have to find out when you are considering effects so accordingly we will consider it in 3 different states. First we will find out till a what it will be?

(Refer Slide Time: 10:09)

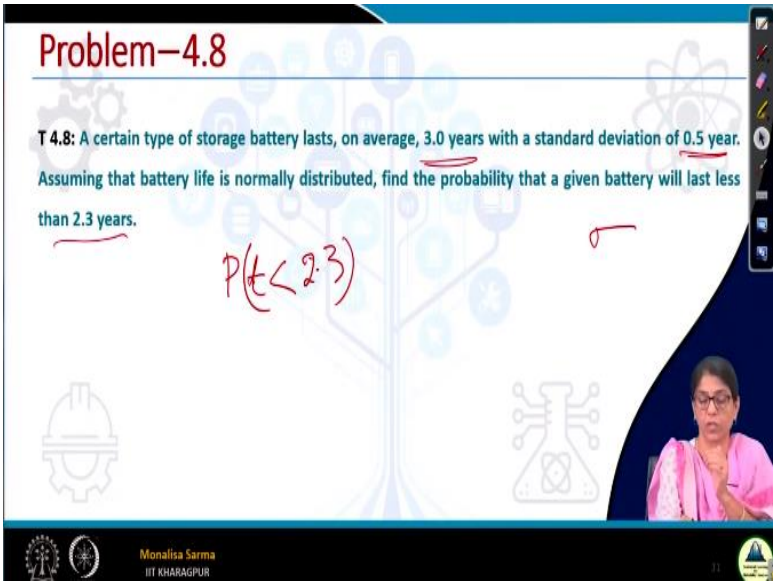


So, for that to find it out first thing is that it is better if we find out take an interval set from c to d, suppose if you are interested in finding out the probability density function probability of in the area of c to d or it is called the density function of the c to d how we will find out this is d - c / b - a is not it? So, now, bringing it in 3 different steps first service from till a to 0 definitely, because still all the values are 0. So, even if it were to get cumulative, this even if you make it cumulative 0 this whole reason is 0.

So, if we get 0 till x is less than a and then x values goes from till a to b what it will be x - a whatever x takes value whichever suppose x is value here than what it is x - a / b - a if it takes here whatever it is x - a / b - a, so, till b this will be the value of x – a / b - a this will be the Fx that is cumulative value and we know area of the curve is always one, area under the curve is always one that means, when we call this Fx dx.

So, if we take integration of fx dx and here it is ranging from a to b this will be equals to 1 other portion it is everything is 0 so this will be definitely equal to 1. So, that means, once it reaches b it has become 1. So, after b definitely it will be 1 only after b all the values are 0, but till b it is 1 so, when we add it up with b it becomes 1 so, x = 1 for x greater than b. So, here this is the figure for the cumulative distribution, this is a fixed a we got 0 than it raises this way this value is x - a / b - a and then from here it reaches 1 so and it continues.

(Refer Slide Time: 12:24)



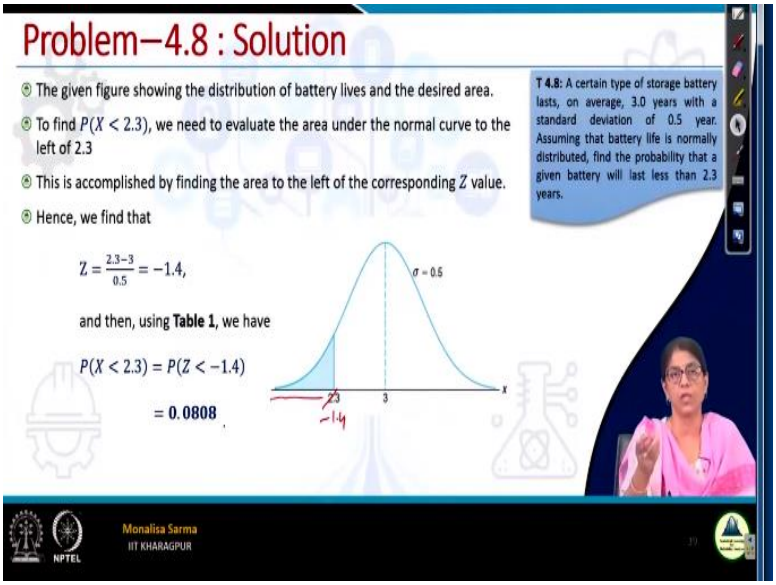
So, simple question on normal distribution already we have seen normal distribution 1, 2 problems in our class only. So, here again we will be seeing 1 2 problems. So, a certain type of storage but battery lasts on average 3 years and standard deviation is 0.5 that means my μ is 3, my μ is 3 years and my standard deviation that is this is 0.5 assuming that the battery life is normally distributed find the probability that a given battery will last less than 2.3 years.

So, I want to find probability that might if I take a random variable whatever random variable you take the x t y any can any variable can do if my take a random variable as say t so, my t is would be less than 2.3. This is the probability I want to find a probability of t less than 2.3. So, how do I find out I will have to take it now to finding this out what I have to take the normal distribution probability density function.

Remember probability density function normal distribution function was a complicated function then we have to integrate it. So, less than 2.3 years means we have to integrate it from 0 to 2.3 years. So, integration of that will be complex. So, what remedy we had? We had a remedy of converting the normal random variable 2 standard normal random variable that is just converting it to Z value.

And then we have the standard normal tables available in the literature books and net anywhere it is freely available downloadable, so we can cancel those lookup table and can get the value. So for that, we will have to find out the Z value corresponding to this 2.3.

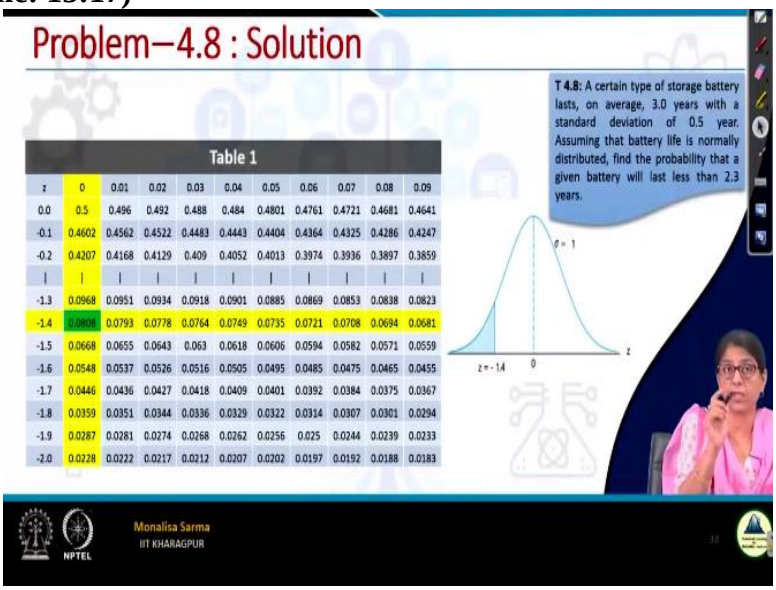
(Refer Slide Time: 14:15)



So, this is what this is the mean, this is the standard deviation 0.5. So, this 2.3 is essentially we want to find this probability that our value is less than 2.3 means essentially we want to find out this area, probability the atlases is this area probability is nothing but the area under this curve. So, this is we have to calculate the Z value how do we calculate the Z value? Z value is x a random variable is x, x - μ / σ. So, we got the Z value is - 1.4.

So, this 2.3 is nothing but if we draw a standard normal distribution, so standard normal distribution, this is 0 mean is 0 and standard deviation is 1. So, this -1.4 will be somewhere here, -1.4. So, we need to find out this area, so we can cancel it from the table. So, how do we get it from the table?

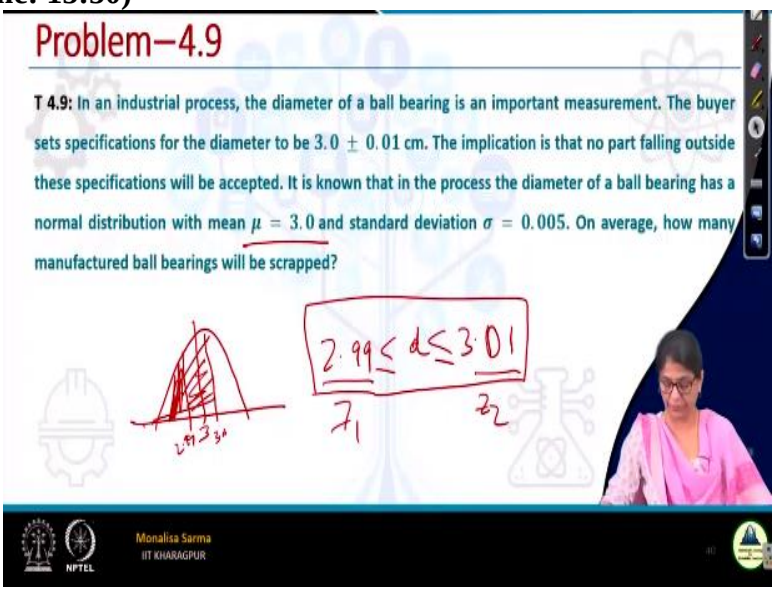
(Refer Slide Time: 15:17)



So, in the table, you will see 1 - 1.4 corresponding to - 1.4 this value is 0.0808. Again, I am telling you this table is for cumulative value that means when I am taking is -1.4 that means

from - ∞ to - 1.4 where probability is 0.0808. And that is what we wanted? We wanted this value this is from - ∞ to if this value corresponds to -1.4 - ∞ to -1.4 that is the area, this value from the table we are getting is 0.0808 so that is the probability.

(Refer Slide Time: 15:50)

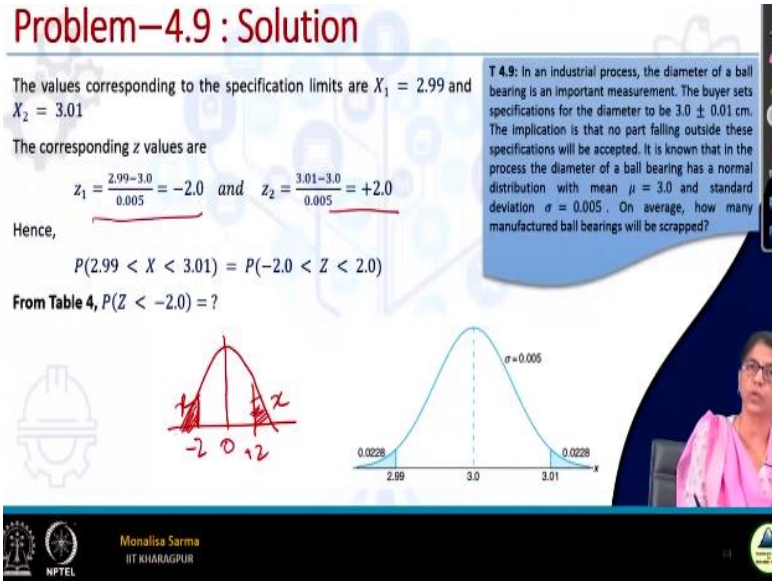


So, that is how we solve this problem? So, this is the second problem is also similar line I will not go into the details. So, just for what is practice, if you have this problem. So, here the industrial process, the diameter a ball bearing is an important measurement, the buyer set specification for the diameter to be 3 plus - 0.01. That means my diameters should be the buyer sets the specification that the diameter buyer sets that my diameter should be in the range of if I am telling its diameter is d.

So, my diameter will be between 3.01 to 2.99 so, my diameter should be in this range. So, essentially if it is normally distributed, so mean is here it is given mean is 3, so, this is mean is 3. So, I want basically what is the probability it will fall in this area this is 3.01 this is 2.99. So, now corresponding to 2.9 and what is the Z value corresponding to 3.01 what is the Z value if we can find a Z value then from the table we can find out the corresponding probability of this.

If I tell it says Z 1 if it is a Z 2 from the table if I can find out what is the Z value? Z 2 will be set to value will be this and Z 1 value will be dispersion. So, Z 2 - Z 1 will give me this probability.

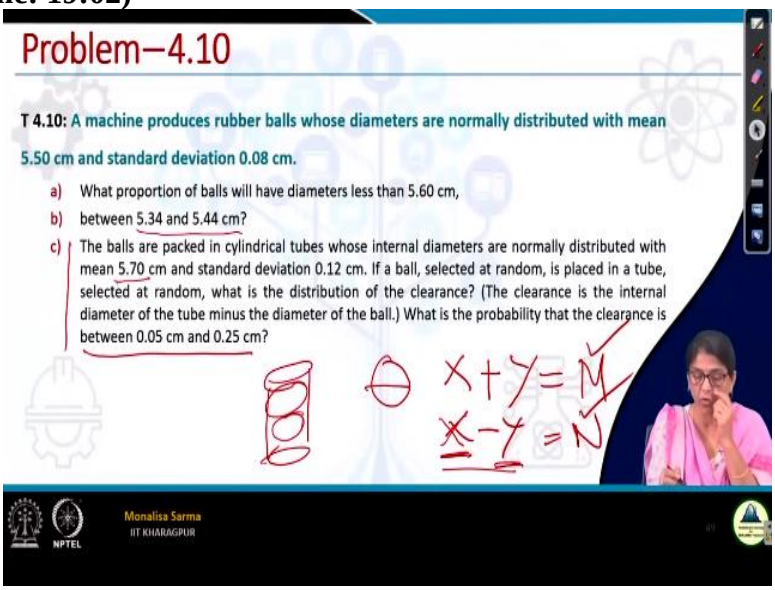
(Refer Slide Time: 17:40)



So, this is how it is been shown here, but I am not going into the details whatever I have discussed, it is a we will be able to solve it and show, here one thing I have just wanted to mention. See here we will calculating the Z value what I got Z 1? Z 1 when I use 2.99 it is to between 2.99 to 3.01 and I use 2.91 my Z 1 value is -2.0 and Z 2 value is 2.0. Here I mentioned one point like see normal distribution is very much a symmetric distribution.

So, if this is and say normal distribution if this is a standard normal distribution, so if I get whatever, say this is -2 and this is +2 from the table, I got whatever value is this whatever probability is this suppose this probability is x. If this probability is x, this probability will also be x because this is symmetric, so whatever value you got from by looking at the table form - 2.0, again you do not need to look up in a table to find out the value for 2.0 you will be same value since it is symmetric.

(Refer Slide Time: 19:02)



So that is what now the next question say a machine produces rubber balls whose diameters are normally distributed with mean of 5.50 centimetre and standard deviation of 0.08 centimetre. Some rubber balls are given was the mean and standard division are given, mean and standard deviation of the diameter, some rubber balls and the diameter. So, this is a rubber ball, it is diameters mean of rubber balls, mean of the diameters as well as the standard division of the diameters is become. What proportion of the balls will have a diameter less than 5.60 centimetre that will be able to solve it. We have done many questions.

So, I am not going to this, then what proportion of the balls will have between 5.34 and 5.4 per centimetre. That also you will be able to do into I have given the solution here, but I will not be discussing it will be able to do it. It is the same thing. Why I have kept this question is I want to discuss this portion. So, see the point c the balls are packed in cylindrical tubes whose internal diameters are normally distributed with mean of 5.70 centimetre and standard deviation of 0.12 centimetre.

So, this balls are kept in some central cylindrical tubes, the tubes diameter mean of this diameters are given 5.70 and standard deviation is 0.12 means, there are many tubes. So, when we find out what is the on an average what is the diameter of the tubes that way we find out the mean. So, the mean is 5.70 and standard deviation is 0.12 centimetre. If a ball selected at random is placed in a tube selected a random tube also have selected at random ball also have selected.

A random means balls diameter size is given mean diameter size is given on an average that is the diameter size similarly, there are many 2 tubes on an average the diameter size is given this is its diameter sizes pipe I mean average diameter size is 5.70 that is the mean. So, I have picked randomly I pick on ball and I picked what to say cylindrical tube also randomly and I have put the balls in the tube now mind it.

The cylindrical tube is also normally distributed the diameter is normally distributed and the balls have also normally distributed this we have already seen. Now, it is asking what is the distribution of the clearance? So, what is clearance when it is made, that clearance is the internal diameter of the tube minus the diameter of the ball. So, this is we have a tube so, here I am putting the balls the internal clearance, what is a clearance?

Clearance is this diameter when I put the ball there will be some clearance someplace. So that clearance I have to find out that clearance. What is the probability that a clearance is between 0.25 and 0.05 and 0.25 centimetre my main intention of giving this question is to bring to your notice one important point here like what I want to say is that here if suppose X is a normal random variable, pay attention if X is a normal random variable, if Y is a normal random variable X is a normal random variable Y is a normal random variable.

When I say X is a normal random variable, what do I mean by that X is a normal random variable means all the values that the x will take along with its probabilities, it will follow a normal distribution means when we plot it, it will follow a bell shaped curve, that is why we call it X is a normal distribution X is a normal random variable. Specifying a variable to be a normal random variable meaning is that only. So, it applies to all distributions similarly, so now X is a normal random variable Y is a normal random variable.

If then in such a case X + Y or X - Y both will be a normal random variable X + Y. So, let me take it as say another variable say M say this is another variable N. So, M is also a normal random variable N is also a normal random variable. Keep give notice to this if X is a normal random variable Y is another normal random variable, it is arithmetic operation where there may be plus or minus the resulting will be again in normal random variable.

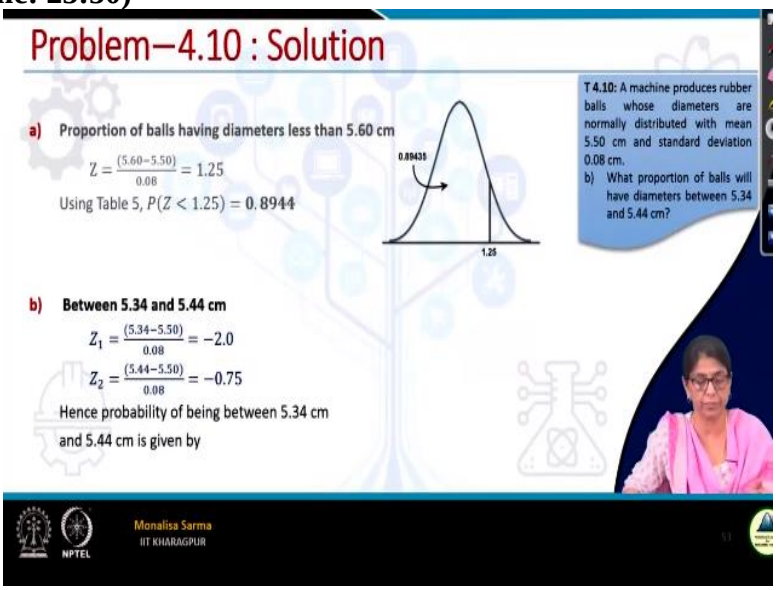
So, here the clearance is that means the clearance that value that we will get from the clearance that is also a normal random variable clearance will be minus. Suppose this is the x is the normal random variable their specifying the diameter of the cylinder, why is there normal random variables specifying the diameter of the ball. So, when we talk about the clearance, clearance we talked means that is X - Y. So, definitely X - Y is also a normal random variable.

So, this is first question, what is the distribution of the clearance? Distribution of the clearance is normal distribution that is the answer. Now secondly when we; found out the arithmetic combination of 2 normal variables now what about the mean and standard deviation. So, if X is a normal random variable, Y is normal random variable and then X + Y, that is M. M is a normal random variable, the mean of M will be mean of X + mean of Y. So, again, I mentioned the mean of M will be mean of X + mean of Y.

Standard Deviation of M will be or variance of M will be variance of X + variance of Y. what will be added for means it will be added for variance also it will be added, but see for N X is a normal random variable Y is a normal random variable X - Y is a normal random variable. Then means of N is μ of X that is mean of X - mean of Y since this is minus mean of X mean of N. Mean of n is mean of X - mean of Y that means, mean of X is taken is a μ X that means, μ X - μ Y.

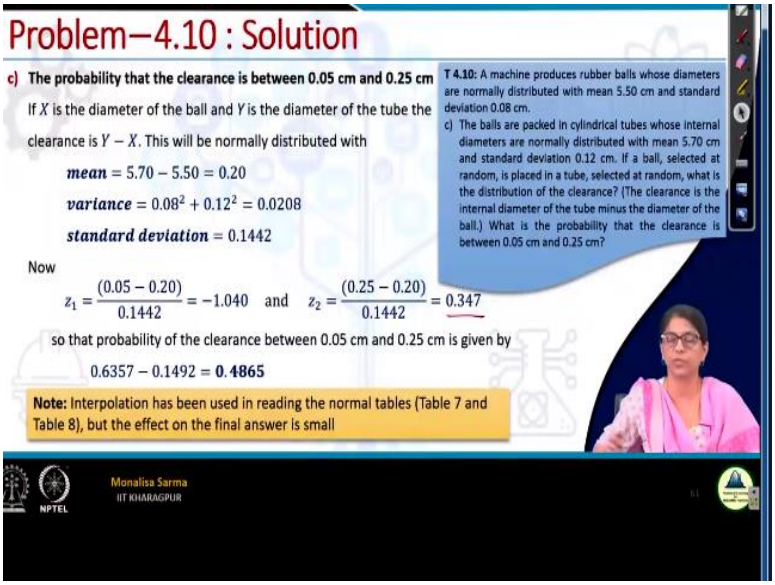
Whatever the variance of N variance of N will be variance of X + variance of Y it will not be subtracted it will be added though it is minus it will not be subtracted it will be added we have seen the laws of variance when we have discussed Poisson distribution if you have forgotten please refer to those laws variance is always added it is never subtracted even if the random variables are subtracted. So, to bring home this point, this is the intention of giving this question as it is the solving of this is same what we have learned till now.

(Refer Slide Time: 25:50)



So, this is I am just skipping you will be able to do it second b also am speaking, skipping it in be able to do it.

(Refer Slide Time: 26:01)



So, now coming to the third point, see here, the probability that a clearance is between we have to find out a probability that a clearance is between this and so now it is a question of simple normal distribution. So, if X is the diameter of the ball, Y is the diameter of the tube so, clearance Y - X this will be normally distributed. Any normal distribution we can to define a normal distribution we will have to need 2 parameters what are those 2 parameters mean and variance or you can say mean and standard deviation whatever it is.

So, to define a normal random variable normal distribution, we need 2 parameters mean and variance. So, what is the mean? Mean is mean of Y - mean of X that is mean this is my mean what is variance? Variance of Y + variance of X so, this is my variance. So, this is variance standard deviation is the square root of variance. So, now I got mean value I got standard deviation value then I need to find out what is the probability that it will the clearance is between this and this value.

Similar type of problems we have done many problems we will be able to do it and very show and even if you cannot this answer is given here solution is given here you can see, but I suggest please do it by yourself. So, in this problem basically we have used an interpolation why interpolation sometimes like here value of Z 2 is 0.347, but in the table, we will not get that 0.347.

(Refer Slide Time: 27:40)

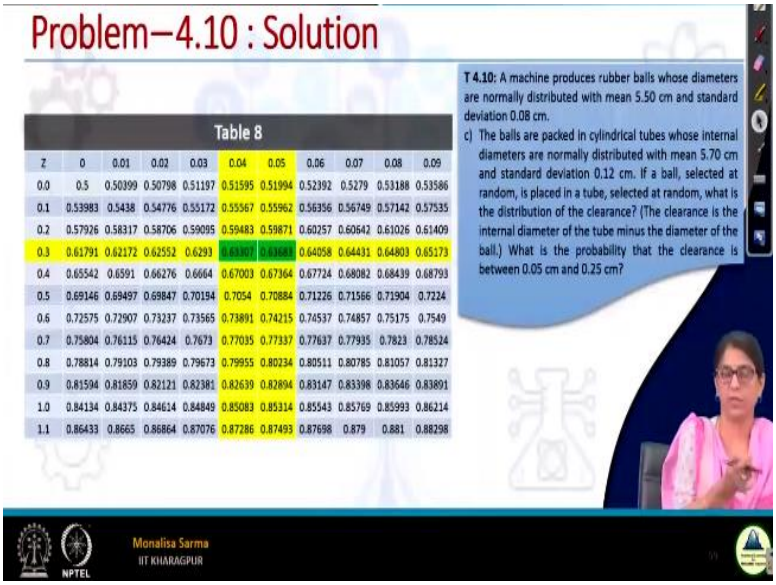
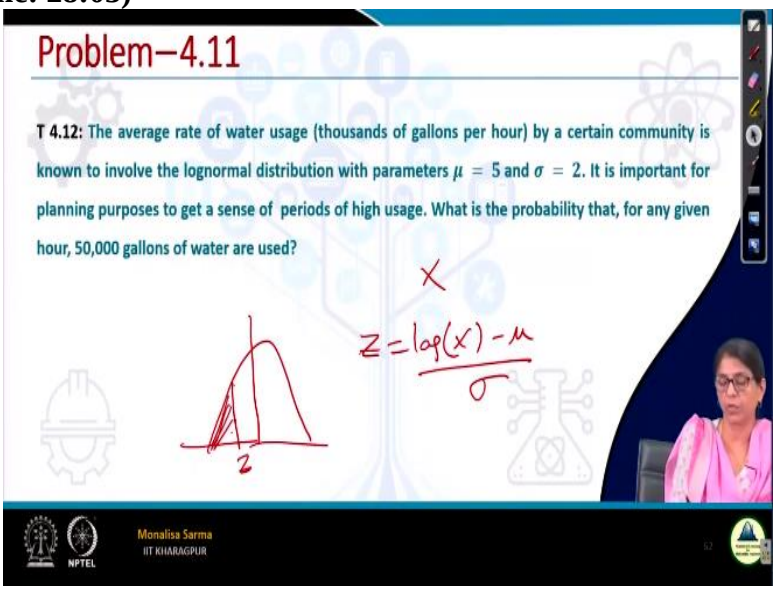


Table 8

Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.879	0.881	0.88298

You see in the table we will get 0.34 or 0.35 347 we would not get, so you will get the value of 0.34 or 0.35. So, in that case, what we can do? We can just add these 2 values divided by 2 that just interpolation that is all. So, that way we got the value that is all this is the solution of that problem

(Refer Slide Time: 28:03)



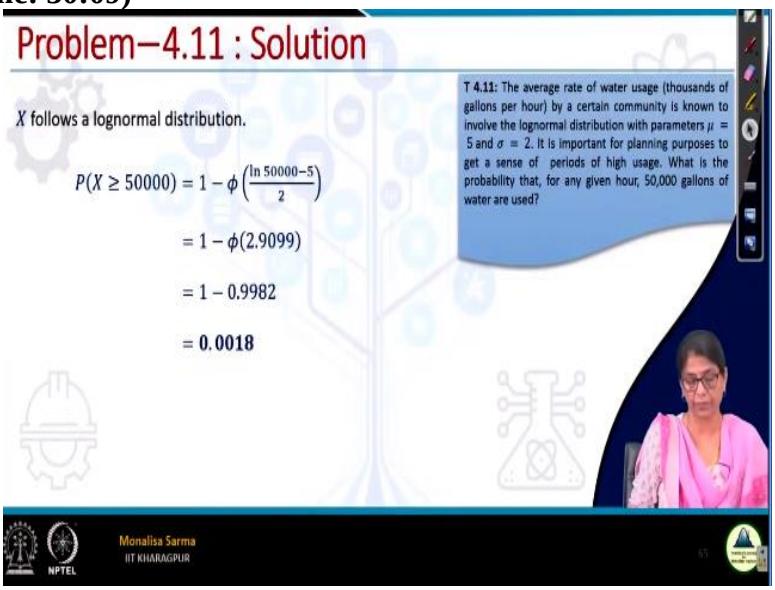
So, you will see if you have any issues you can see it there yourself and not discussing that in details, because many similar problems we have discussed. Now, the next question, the average rate of water usage 1000s of gallons per hour by a certain community is known to involve lognormal distribution with parameter μ and σ, μ is 5 σ is 2 remember lognormal distribution if X is log normally distributed, that means log of X is normally distributed with μ and with whatever μ is specified and whatever σ is specified.

So, when a variable is log normally distributed, the log of that variable will be normally distributed, since the log of that variable will be normally distributed, we have already seen in lecture 13. If it is log normally distributed, then how to find a log normal distribution we have a probability density function again we have to find out the probability of occurrence of a particular value what we have to do we have to either calculate that probability density function means we have to do the integration for particular range.

Or we can look into the standard normal distribution table for lognormal distribution also, we look into the standard normal distribution, how they are our Z value here Z value suppose my normal thing suppose my X is what is a log normally distributed then my log of X will be normally distributed. So, my Z value will be Z = log of x - μ / σ. So, from the standard normal distribution table whatever Z values suppose, this is the Z value whatever I got here if I get minus then it will be the left side of the 0 mean.

So, whatever value I get it from the table that will corresponds to the probability of whatever value is given.

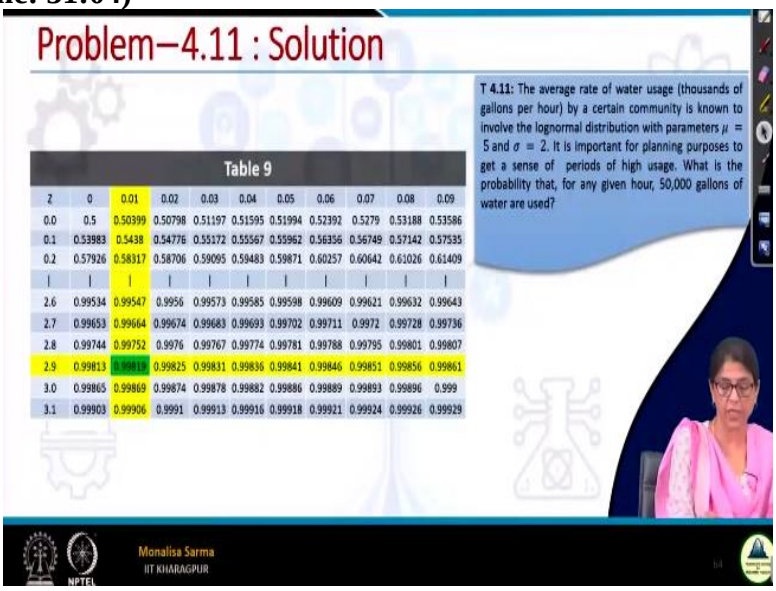
(Refer Slide Time: 30:09)



So, we are interested in finding out what is the probability that any given our 50000 gallons of water uses the probability that is called 50,000 gallons of water is used. So, at least 50,000 gallons of water that means property of X greater equals to 50,000, greater equals to 50,000 means, we have cumulative means we have from - ∞ to 50,000. So, in case of log normal it is not minus always it is not lognormal it is for positive scale we have seen that so, in the from the table we can get from one value from 0 to 50,000.

So, when we are interested in finding greater than so, it will be 1 minus of that, so 1 minus of we call it phi I told you yesterday in my last class. So, log of 50,000 - 5 / 2. So, whatever we got.

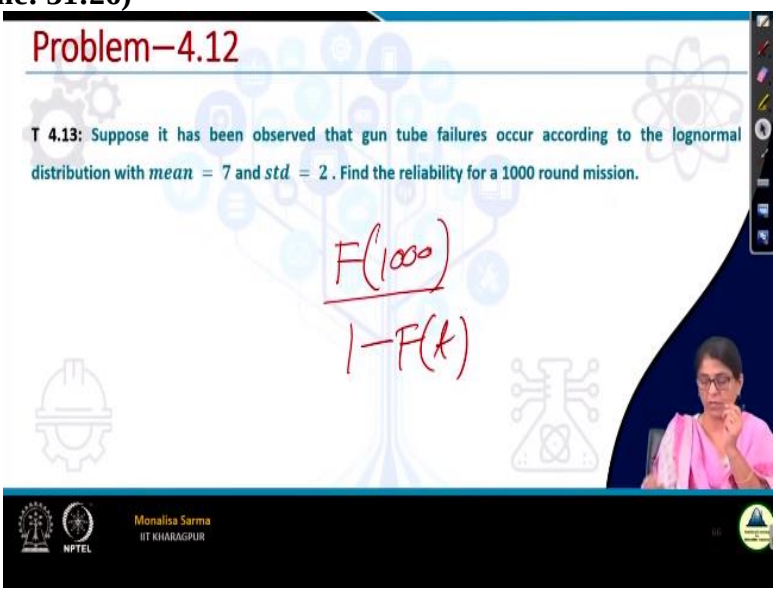
(Refer Slide Time: 31:04)



Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
|	|	|	|	|	|	|	|	|	|	|
2.6	0.99534	0.99547	0.9956	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.9972	0.99728	0.99736
2.8	0.99744	0.99752	0.9976	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.999
3.1	0.99903	0.99906	0.9991	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929

So, this value we can look at from the table. So, what was the value? Value was 2.9099 we can consider it is a 2.91 so, if you can consider it as 2.91 so, this is 2.91 corresponding this is value 0.99819. So, 1 minus of that will give me the value.

(Refer Slide Time: 31:26)



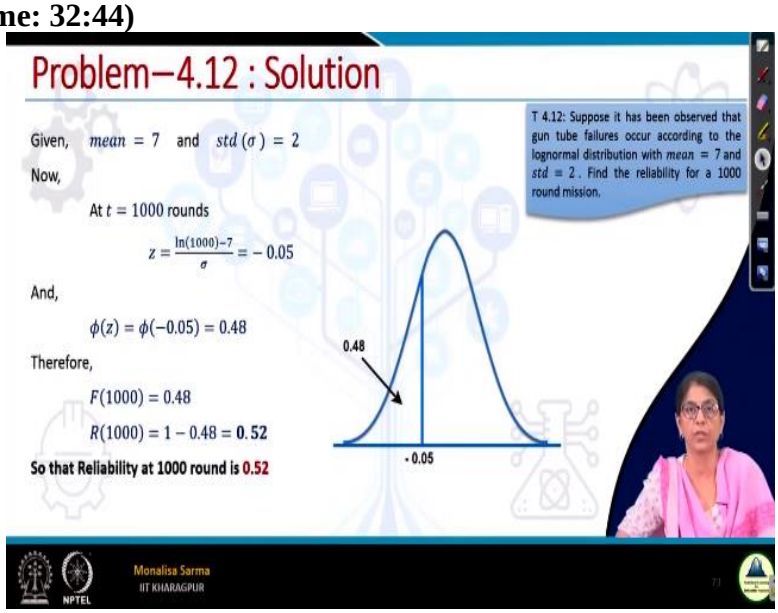
So, next question suppose it has been observed that the gun tube failure occur according to the lognormal distribution with mean = 7 standard division is 2 find the reliability for 1000 round mission. We have to find out the reliability for 1000 round mission how do we find out

the reliability for a 1000 round mission? So, it is also that means when we are talking of reliability for 1000 round mission.

First we will find out that cumulative probability that is F of 1000 that is the gun to failure this is cumulative failure till 1000 hours 1000 round mission we have to find out what is the probability of failure till 1000 round mission that is F of 1000. F of 1000 we have seen how we can find it out cumulative we can easily find it out from the lookup table. So, when we got to F of 1000 then reliability is nothing, reliability is what, reliability is just 1 - F (t). So, here t is 1000 so 1 – F(1000) will give me reliability.

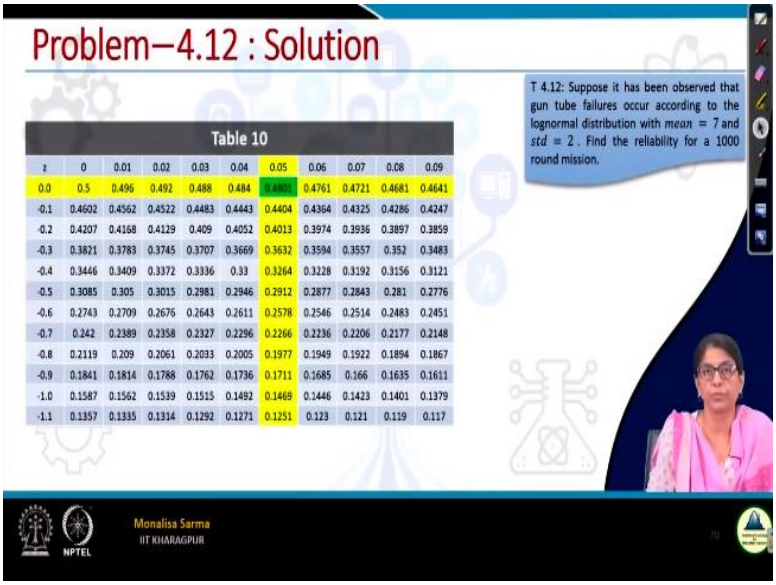
Reliability is failure free if till this if this is the failure question then what is the reliability? Reliability is 1 - F of 1000.

(Refer Slide Time: 32:44)



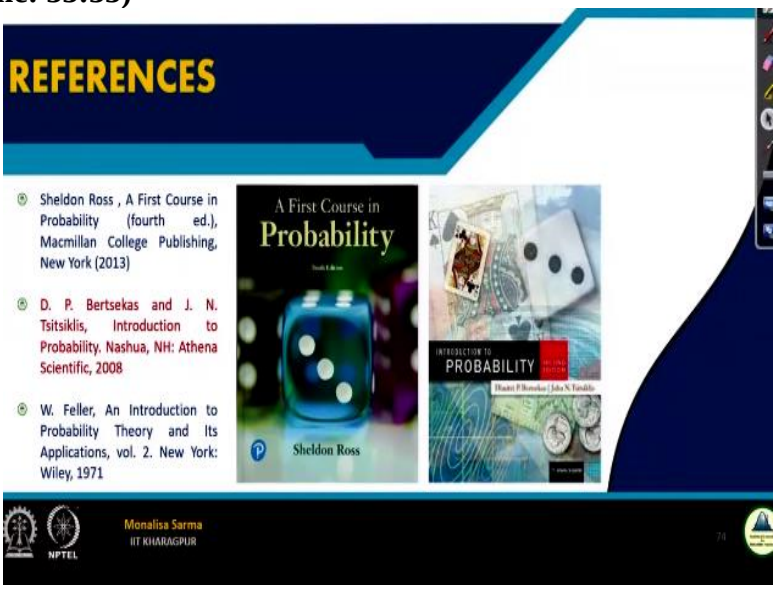
So, at 1000 round we got Z value we have calculated so, this is the Z value so, from the table we will find out what is the value of this.

(Refer Slide Time: 32:56)



So, 0.5 this is the value of 0.4801 this table you will be getting in any appendix of any product statistics book or if and you can download it standalone also you can just Google it, you will get it and get a link and you can download it this table are freely available so we got is 0.48. So, F of 1000 is 0.48. So, R of 1000 is 1 - 0.48, so reliability at 1000 round is 0.52.

(Refer Slide Time: 33:33)



So that is all in today's lecture there we will be discussing few more problems. And for that I will take one more tutorial class before going to our next topic. So, then these are the references which I have already mentioned before an earlier classes and thank you guys.

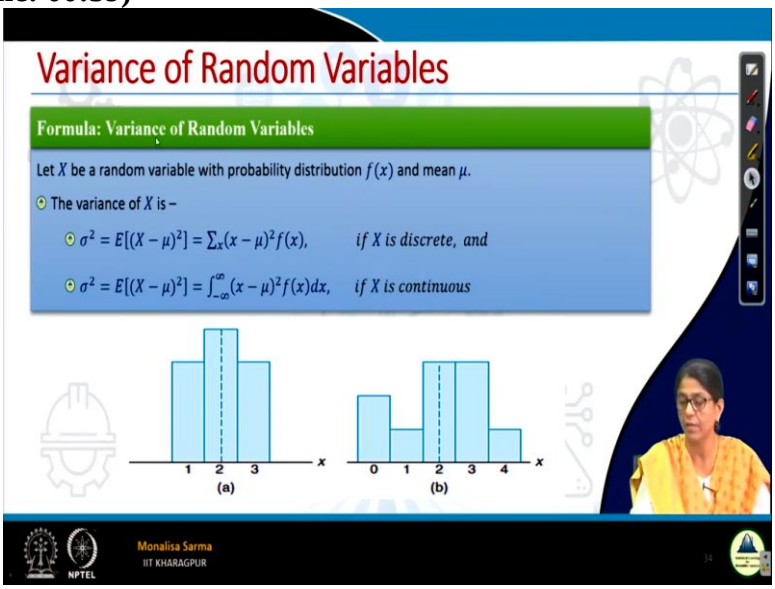
Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology, Kharagpur

Lecture - 15
Tutorial on Continuous Probability Distribution Functions (Part 2)

Welcome once again. So, last class, we have covered that tutorial on continuous probability distribution but we could not solve problems on all the distributions. We have only done some problems on normal distribution and lognormal distribution. In today's class we will be seeing exponential distribution, problems on exponential distribution, problems on Gaussian distribution and also again some more problems on normal distribution.

Because normal distribution is a very important distribution, not only from the reliability perspective from and in other application perspective as well.

(Refer Slide Time: 00:59)



So, directly coming to the problem say the first question the length of time for one individual to be served at a cafeteria is a random variable having an exponential distribution with a mean of 4 minutes, see interesting the length of time for one individual to be served in a cafeteria it is at this length of time that is the time that is not a random variable and it is an exponential distribution with a mean of 4 minutes.

Mean of the time it takes for a person to be served basically, how much minute it takes for the person to be that means we can say that is the rate of the event. So that the rate of the event it is given specific that it is 4 minutes. What is the probability that a person has served in less

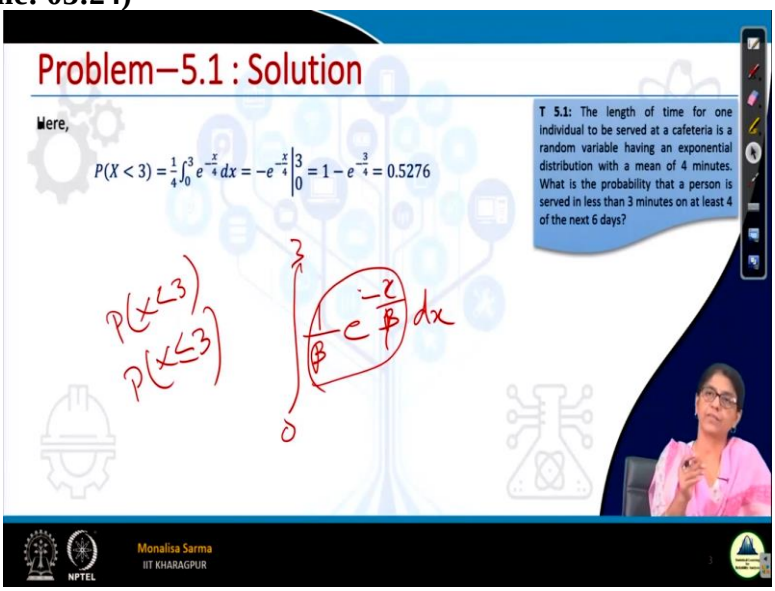
than 3 minutes or at least for the next 6 days. See the second part, what is the probability that a person is served in less than 3 minutes.

That means we have to find out probability that if you are taking t as a random variable, t is less than equal to 3, probability of t is less than equal to 3 when we need to find out the person is serving less than 3 minutes, on at least 4 of the next 6 days. 4 of the next 6 days, this directly implies a binomial distribution as if there are 6 trials, out of 6 trials I need 4 success, 4 success and what is the probability of success my probability of success is this a person has served in less than 3 minutes.

This is my probability of success. Trials 6 trials out of 6 trials 4 success in binomial distribution, what are the parameters first is number of trials, number of success, probability of success these are the 3 parameters to define a binomial distribution. So, this is number 6 is the number of trials, 4 is the number of success and that the person is serving less than 3 minutes that is the probability of success.

Now, how do I get the problem to find out a probability of success second part I understood it is a binomial distribution, how do I get the probability of success, that is, what is that is exponential distribution, is not it? The time it takes for the first event to happen, time will take for the bus to arrive, time it takes for the failure to happen, or the time it takes for the person to be served in a team in a cafeteria. So these are the occurrence of an event. So that is the rate of an event. So this is and this directly, it is an exponential distribution.

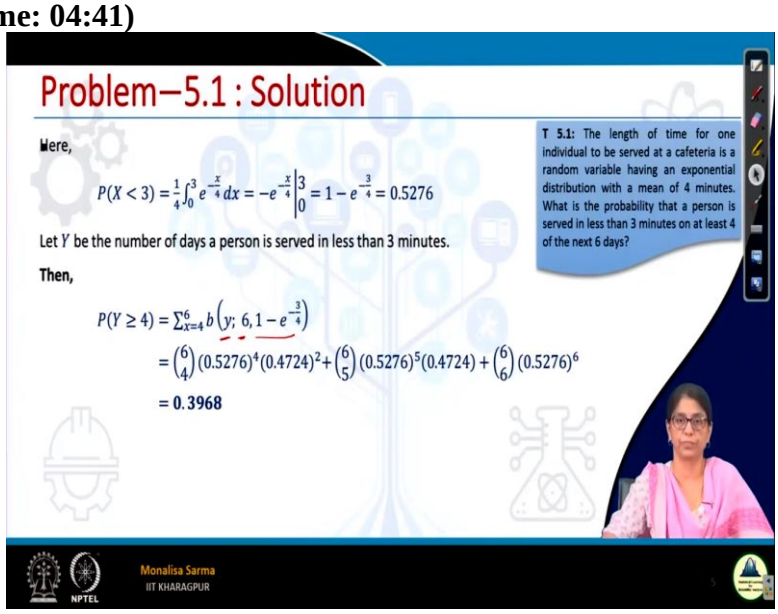
(Refer Slide Time: 03:24)



So, we know the formula for the exponential distribution, I mean, when I call it formula, I mean probability density function, I will not mention it again and again when I am telling it is a formula it means the probability density function. So, I know the expression for the probability density function, what is the exponential for exponential density function? It is 1 / β, 1 / β $e^{-x/\beta}$ this is my density function.

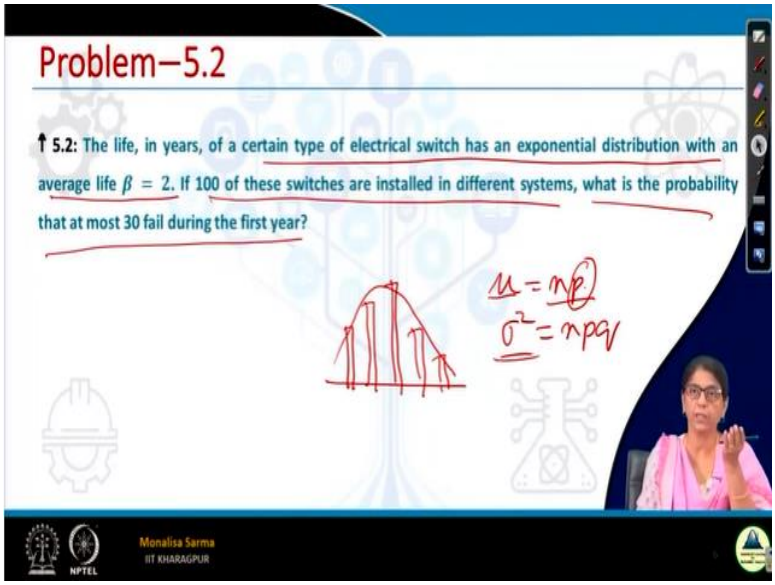
So, I need to find out what it is given certainly less than 3 minutes that has means I have to integrate it from 0 to 3 dx, this is my fx the feature to integrate it from 0 to 3 then I will get the value for probability of x less than equal to 3 whether I tell it is less than equal to 3 or in very important thing in continuous probability distribution. If I tell p of x 3 and p of x less than equal to 3, 2 are different things are same thing, 2 are same thing because that equal to 3 probability is 0. Similarly, if I tell greater 3 or greater equal to 3 both are same thing.

(Refer Slide Time: 04:41)



So Y be the number of days a person is serving less than 3 minutes, that is the Y be the number of successes what I told you. So I found the probability is 0.5276 probability of your success now simple binomial distribution, simple binomial distribution, specifying the value so, this is the probability, this is the number of successes, this is the number of trials simple put in probability distribution function or binomial and we will get the value.

(Refer Slide Time: 05:14)



Next question the life in years of a certain type of electrical switch has an exponential distribution when an average life β = 2, the life in years of a certain type of electrical switch has an exponential distribution with an average life of β is also equals to 2. Here β is specified as it is equals to 2. Life in years means the time it fails. So, rate of failure is that is β = 2 if 100 of these switches are installed in different system, say if 100 of these switches installed in different system what is the probability that at most 30 will fail during the first year.

So, I need to find out what is the probability that a switch will fail in one year, this in the first year, so, from 0 to what is the probability that a switch will fail in first year that we will find it by exponential distribution, then see the next portion 100 of a switches installed in different system, what is the probability atmost 30 will fail at most 30 when 0 can fail, 1 can fail, 2 can fail up to 30 it is again a binomial distribution.

Where n is the total number of 100 is the total number of trials and 30 is the number of successes at most 30. So, essentially it is a cumulative binomial distribution from 30 to 100. Now, there is one thing which you need to know is that when the n is quite large, like here, when n is quite large, then sometimes binomial distribution can a discrete distribution also can be estimated by approximated by in normal distribution.

Mainly in case of binomial distribution if the binomial distribution if it takes a bell shaped, binomial distribution, definitely we will our graph will not be a curve it will be either a bar chart or it will be a histogram whatever it is. So, if it is in the form of a, this sort of thing, so,

it forms what? Bell shaped it is not it? So, and then the binomial distribution, if the curve is, the resulting curves, resulting diagram I should not occur.

The resulting diagram is in the form of a bell, if it gives a bell shape, then we can approximate binomial distribution by normal distribution, why we will do that is that now see for this case, you would have to find out what 100 is n, n is 100, our x is 30. And you did have to find the cumulative value and for cumulative value of 30. And n = 100. If we have the corresponding table for that for the corresponding probability, of course, the probability is also important.

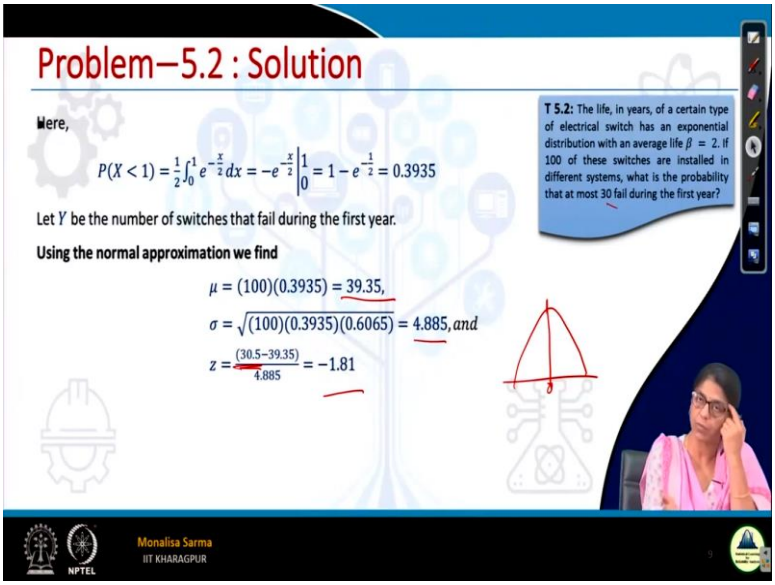
For the probability, if the corresponding probability if we can look it in a table it is fine, but if a probability in the table we do not have the table for all possible values of probability on all possible values are n and x, so, you do a table is not there to calculate such a big number it is a very, it is not a very easy task. So, in that case, if we can approximate it by binomial distribution things become very easy.

Or life becomes easy if you can approximate that is the end when you see if we approximate it, our value will not be significantly change, their value will change slightly, which it can be considered. So, that is why in such cases we can approximate by normal distribution. Now here in case we will be approximated by normal distribution. So, when is the normal distribution again, we need 2 parameters, what are the 2 parameters the parameters are μ and σ.

So, for normal distribution, what is μ? μ is n x p and what is σ? Σ is npq, sorry variances npq. So, to find out for normal distribution, we need μ and σ square here μ and σ square nothing is given. But here we will be able to find out the p probability that it will fail during the first year using the exponential distribution we will be able to find out that will be my p we have n, n is 100.

So, once I find my pq is 1 - p is q. So, I got my μ, I got my σ square, then I will be able to find out my, I will be able to find out the required probability using approximating by normal distribution.

(Refer Slide Time: 09:47)



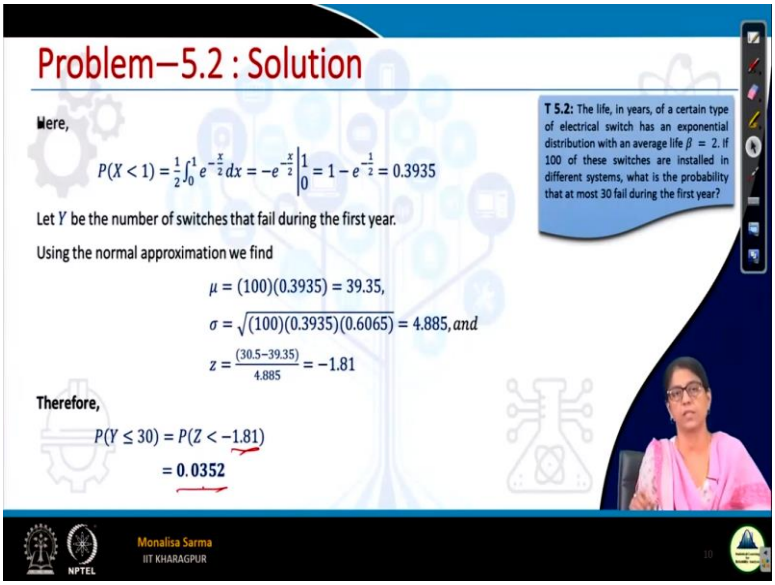
So you see how first we found out the probability using exponential distribution, my probability density function exponential distribution where β is 2 is given. So this is how this is the probability we got so, now we will be approximating using normal distribution. So, what we took μ is n x p we found is this then we found out the standard deviation we found this then we need to find out the z value corresponding to at most 30 fails.

So, it is this is 0 and something 30 means whatever, whatever value will come to for whatever value will come for 30, whatever z value now see when we are trying to approximate using a discrete number when you are trying to approximate by continuous variable, we will have to considered as what to say it is near about value like 30 because probability at exactly 30 will be 0.

So and when we are considering a 30 that means, it will be in a very small interval nearby 30. So, when it is given at most 30, it will fail at most 30. So, we cannot consider an exactly 30 here, see here, we have considered 30.5 instead of 30. Because when we approximate in for continuous, we will have to take a closer interval nearer to 30. So, I have taken is 30.5 here. So, if in other cases, if it is specified that it is less than 30.

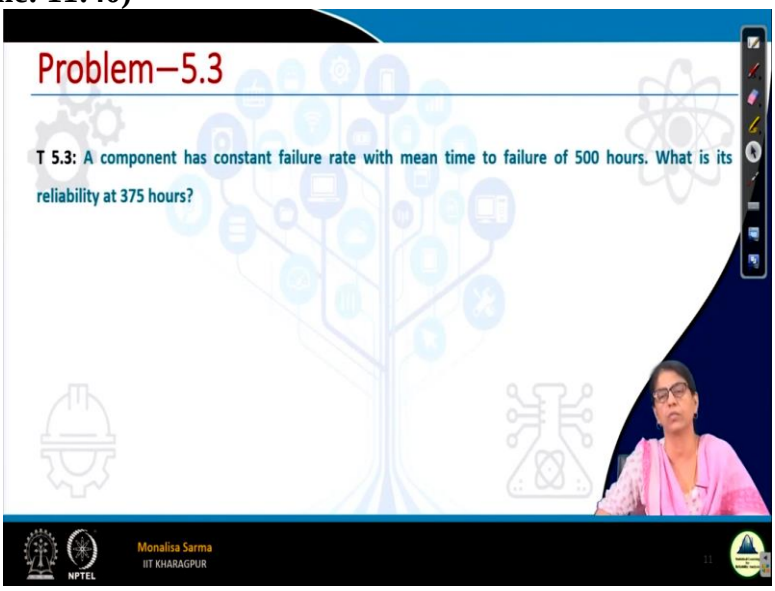
Then I would have considered, I would not have considered 29, I would have considered 29.5. So, I found the z value here -1.81. Once I found the z value, then getting the probability and just a matter of looking into the table that is all.

(Refer Slide Time: 11:30)



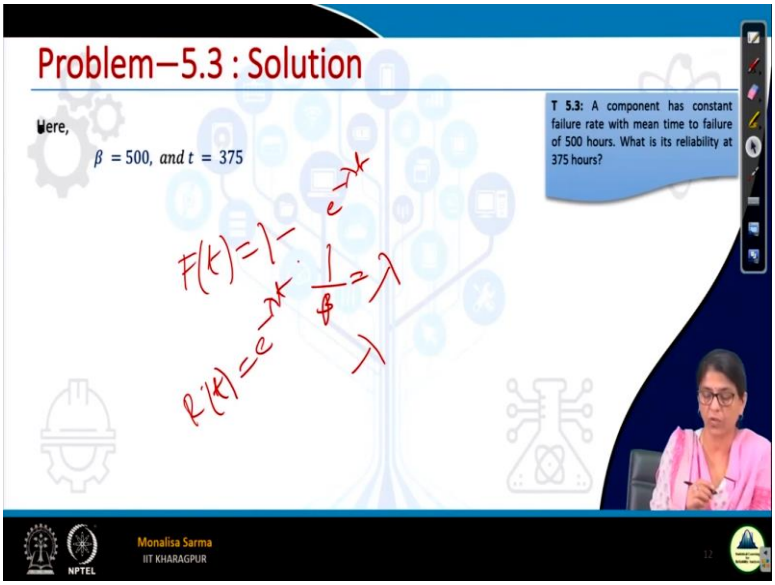
So that is the -1.81 there is a value from the table I got this. So this is the probability.

(Refer Slide Time: 11:40)



So, a come next component has a constant failure rate with mean time to failure of 500 hours what is its reliability at 375 hours similar type of problem.

(Refer Slide Time: 11:54)

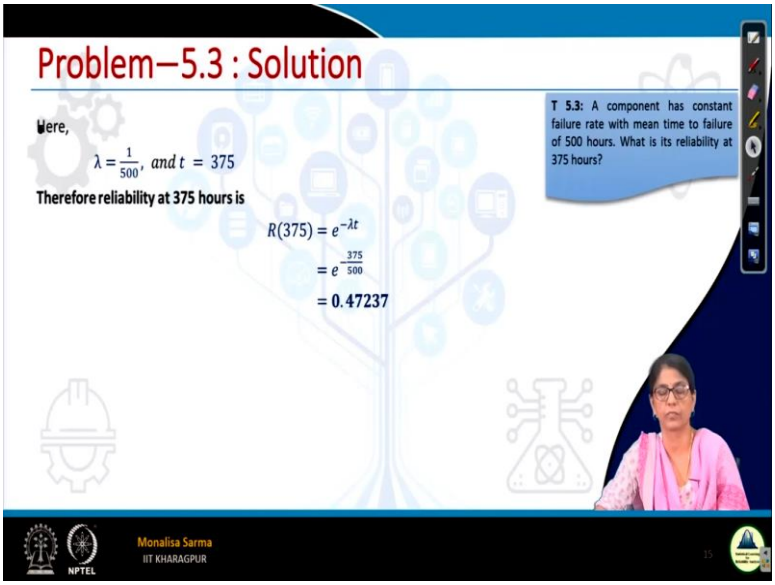


So, here β is given it is 500 hours, mean time to failure that is 500 hours. So, then we have to find out the reliability this is again an exponential distribution exponential distribution we know it is 1 / β e $^{-\text{ x square } \beta}$. Again one more thing the 1 / β is also equals to λ we have seen that is not it? 1 / β is nothing but it was λ we have already discussed.

So, instead of writing that, we can also write we have also instead of writing 1 / β we can also write λ. So, when we find try to find out the reliability of this first we will have to find out F of t. So F of t will be nothing but integration of the probability density function what we have said once we integrate that what we will get? We will get 1 - λ e - 1 - e $^{-\lambda t}$ not λ it is not there e $^{-\lambda t}$.

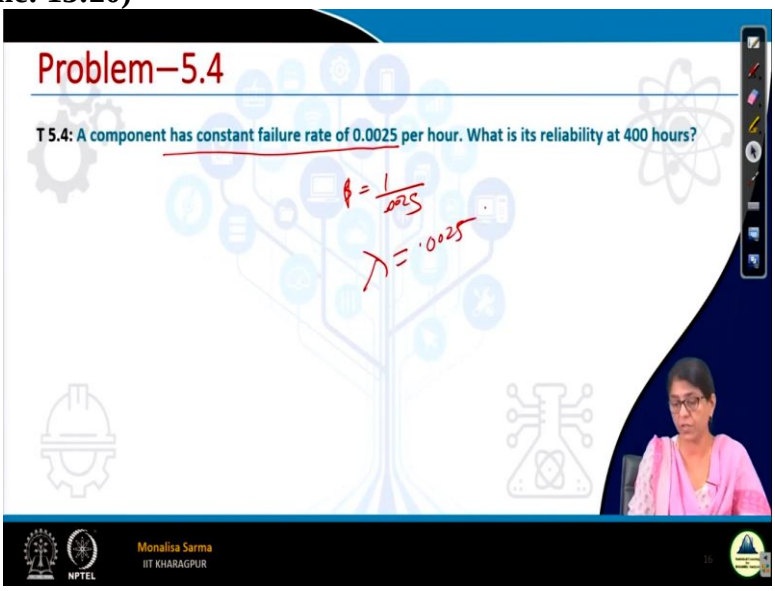
I can you can write it in terms of β or you can write it in terms of λ whatever it is. So, if this is Ft and what is my Rt? Rt will be just e $^{-\lambda t}$ or e $^{-1/\beta t}$ whatever you may write in terms of λ, you may write in terms of β.

(Refer Slide Time: 13:12)



So, that is all here it is given I have written in terms of λ.

(Refer Slide Time: 13:20)

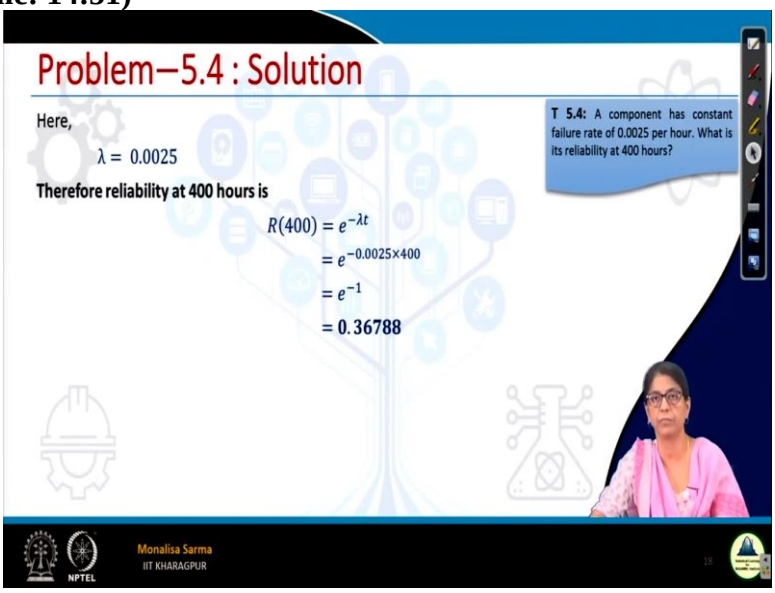


So, a component has a constant failure rate of 0.0025 per hour, failure rate of this much this much failure 0.0025 failure per hour see here, please notice the difference between this question and the previous question, it is given that component has a constant failure rate of 0.0025 per hour that means, 0.0025 failures per hour, if this is 0.0025 failures per hour that means, what will be my then that means, this is my λ value this is not β value.

That means, if I consider β, β will be equals to 1 / 0.0025 because this is considering 0.0025 per hour 0.0025 failures per hour when I consider β is? B is 1 failure per unit time. So, it will be β = 0.0025 whether I can say here it is specified α is specified sorry λ is specified. So, here λ is also 0.0025. See the difference between the previous question what does the previous question as constant failure with mean time to failure of 500 hours.

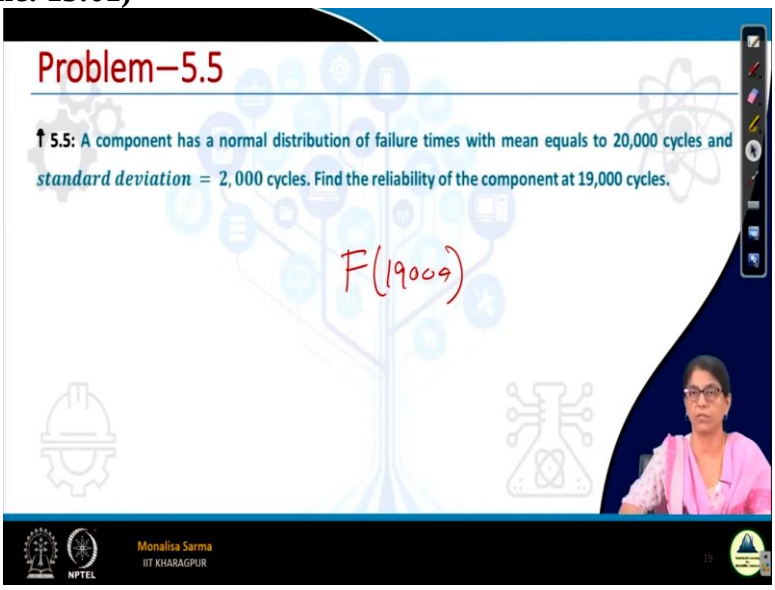
Mean time to failure means that is the time it has time for the first failure or the whatever the one failure is maybe has occurred the time it has taken for the next failure to happen. So, it is within 500 that is the first failure that has happened. So, that is β see the difference between the α and β here.

(Refer Slide Time: 14:51)



So, here sorry not α λ, λ and β here. So, this is λ is given so directly we can find out a reliability $e^{\lambda t}$.

(Refer Slide Time: 15:01)



Again, a component has a normal distribution of failure times with mean equals to 20000 cycles, standard deviations because the 2 cycles find the reliability of the component at 19,000 cycles. I have given lots of question on reliability, why mainly, we will have an idea

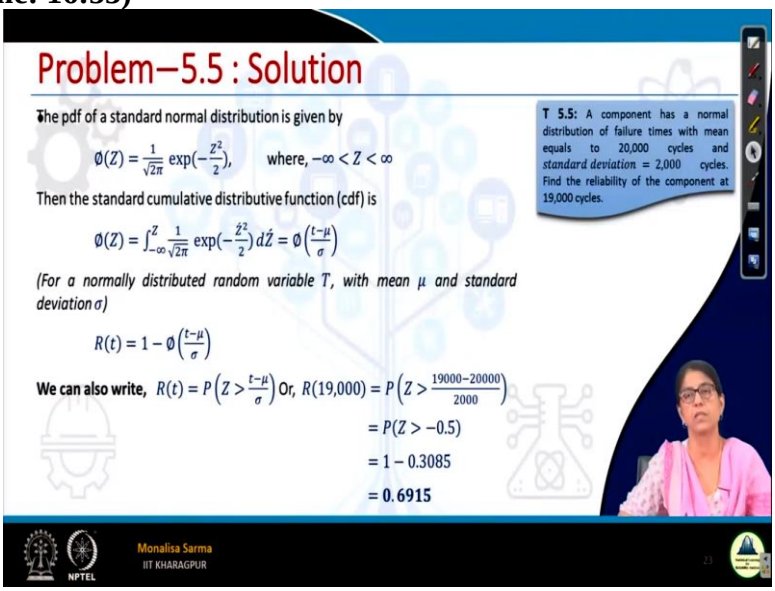
on reliability as well plus we will be practising more and more problems on different distribution.

So, when you talk about reliability, it is nothing just you find out a first you find a cumulative, if I consider the variable when we talk of reliability, then the random variable we consider easily in terms of time it is time to failure. So then if t is my time to failure than small f of t is my failure density function, then I get F of t, but this F of t implies F of t is the cumulative failure till time t.

That means, from - ∞ to t that is the cumulative failure. So, once I talk of cumulative failure, that is F of t from time ∞ to t that is the cumulative failure, then when I what will be my reliability? Reliability will be simply 1 - F of t. So that is just the difference between the F of t and R of t 1 - F of t is R of t, that is why I am telling that is the reliability of the component.

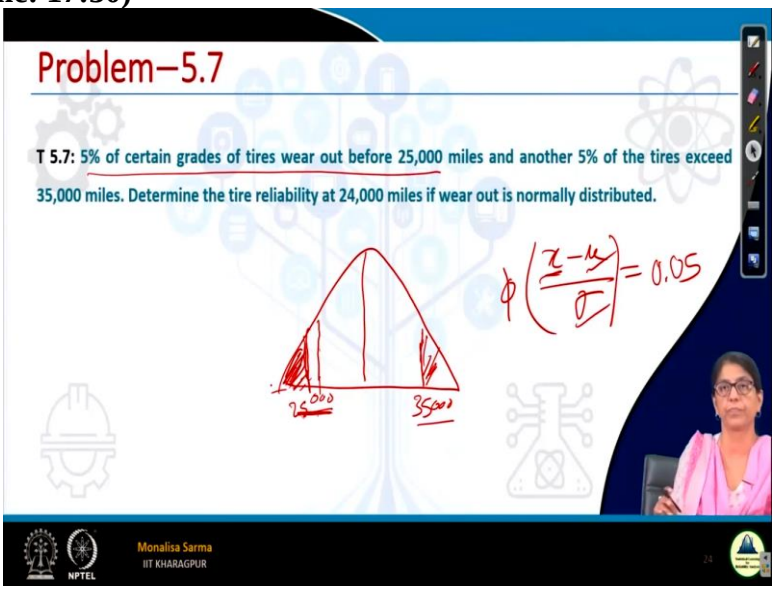
Now, when I am asking reliability of the component at 19,000 cycles. So, that means, first I have to find out what is the failure probability till 19,000 cycles that is F of 19,000, so that means, till 19,000 F of 19,000 means from - ∞ to 19,000 so, it is a cumulative so, we can look it from the table we can get the value and whatever my 1 - of that will give the value.

(Refer Slide Time: 16:53)



So it is, these are the expression but just for a quick recap I have given you this is the failure density function for normal distribution, if we find out the cumulative this is the value and this is the value will look up at the table that is the z value from the z value z is close to - 0.5. So, corresponding to -0.5 we will get the value from the table and what we get is 69.15

(Refer Slide Time: 17:30)



Similar type of question 5% of certain grades of tires were out before 25% 25,000 miles another 5% of tires exceed 35,000 miles determined the tire reliability at 24,000 miles the wear out is normally distributed, we given the wear out is normally distributed. So, first let me draw the normal distribution what it is given? 5% of the certain tire whereas before 25,000 5%, so, 5% means always this 5% means the probability that is in this area.

So, this is corresponds to 25,000 miles. So, 5% were out before 25,000 miles another 5% of tires exceed 35,000 miles exceed 5000 and exceed 35000 definitely it will be this side, definitely it will be this side another exceed 35000. So, this side is exceed, this side is less so, this is 35,000. So, what it is given what is the probability corresponding to this value it is given, probability corresponding to this value it is given.

In earlier cases till now, what we have solved what we have we did not know the probability corresponding to a particular z value, particular z value we could find out how if the random variable was x - μ / σ it gives us the z value from the z value we found out the corresponding probability from the table. Now, here the probability is given and my z value is given that means, my say x - μ / σ is equals to this phi of this I get the probability value.

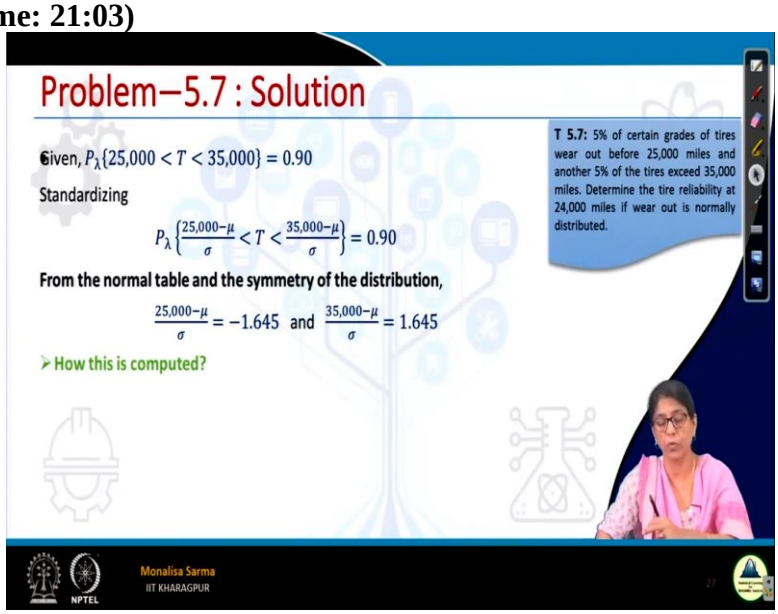
Now, this probability value is given it is 0. 5% and 0.05% 0.05. This is given and what I know? I know this x, x is 25,000. But I do not know μ I do not know σ. Similarly, this is also given, this is also same x - μ / σ this is also 5% 0.05 I do not know x, I do not know μ, I do

not know σ. So, 2 questions are there, 2 equation 2 unknown variables. So, I can find out the values of μ and σ that is by solving these 2 problems this 2 equation.

Now determine the tire reliability at 24 tires and 24,000 miles if wear out is normally distributed. So, I have to find out tire reliability at 24,000 miles. 24,000 miles means it will be somewhere here because let me rub this and it will be easier for me to show this was this, this was 25,000 so, I need 24,000 miles 24 it means this is 25,000 so 24,000 will be left of 25,000 so 24,000 will be somewhere here this value.

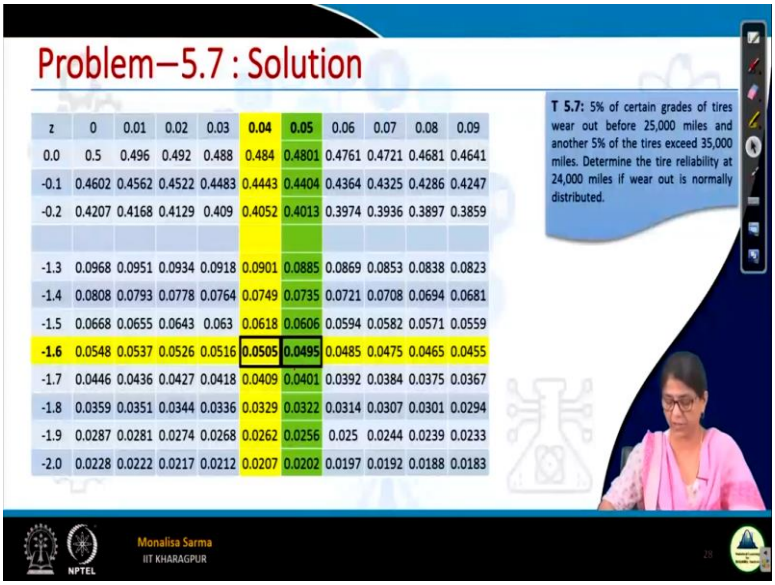
So, I need this reliability at 24,000 miles, to reliability of 24,000 miles this side which will give me F of t cumulative failure after that. So, reliability will be this whole thing is 1, 1 minus of this will give me reliability at 24,000 miles what is the reliability that what is the probability that it will not fail probability failure if so, it will be 1 minus of this same thing does the same concept everywhere.

(Refer Slide Time: 21:03)



So, here this is given so, it both side is 0.5, in this figure what we have seen this is 0.05. This is 0.05 total put together is 10% so, if this is 0.05, this is 0.05 what is the remaining this? Remaining this will be 0.90. So, this value is equal to 0.90, now we know what is the problem the corresponding to this, we know what is the probability corresponding to this then we found out the value of μ and σ, from the normal table similarly the same way now, this one is we got the value 1.645.

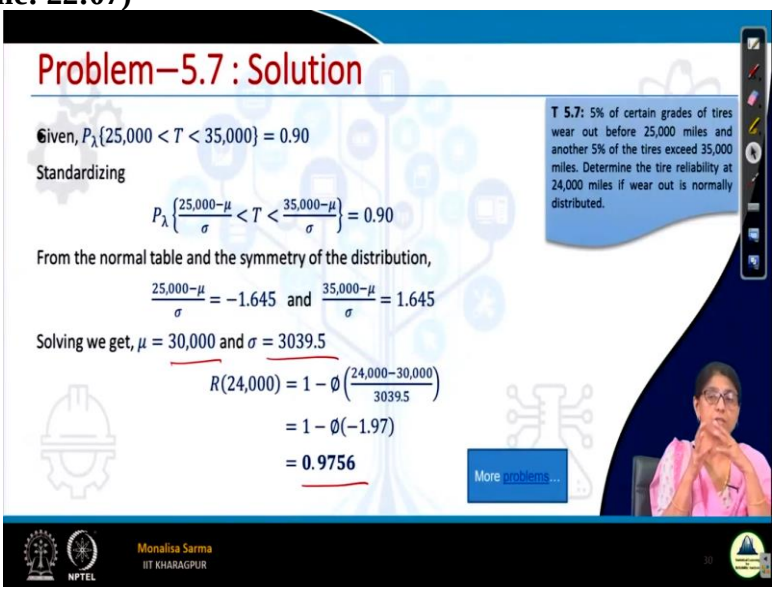
(Refer Slide Time: 21:55)



z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.496	0.492	0.488	0.484	0.4801	0.4761	0.4721	0.4681	0.4641
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.2	0.4207	0.4168	0.4129	0.409	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.5	0.0668	0.0655	0.0643	0.063	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.025	0.0244	0.0239	0.0233
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183

Again as I told you in the 1.645 it is not in the 1.64 is there 1.65 is there how we can get the value for 1.61 just we will add these 2 values and divide by 2 in simple interpolation.

(Refer Slide Time: 22:07)



So, that is how we got this value and by solving this we could solve this, we could get μ is equal to this, σ is equal to this now simply putting the values for that for finding F of t 1 - phi of F of t I will get reliable data 24,000 simple. So, there are some more problems which I have given a link here there is more problems if you see which are these are problems on Weibull distribution.

As I already mentioned in my 13th lecture that Weibull distribution I will not be discussing in this class because it is this distribution it is a bit difficult to understand for people with reliability background they find it more easier to understand. So, industry people if an industry people or even a professionals who are dealing with reliability or any quality or

reliability if they are interested they can see the reliability distribution slides also already I have kept there, already mentioned in the class and if they want to practice any problems.

They can also see the problems are here the link is here. They have around 2 3 problems solve problems here. So now but not only that any students who are interested and M.Tech students if they are interested even B.Tech students who are interested because Weibull distribution is widely used distribution, it is widely used in reliability and quality studies. So, if they are interested, they can go through the problems as well as the concept which I have given in my 13th lecture.

And any doubt of course, I will take up in the doubt clearing session. So it is but I am repeating it is not included in the course that means questions, when I will be setting questions for exams, question will not come from this portion you can be assured of that. So, so now, so from here we have almost come to end of these tutorial classes. Now, by solving so many problems by now, I think you could understand why we need probability distribution.

So, I thought I should ask this question at the end of this class, we have done many problems and problems in probability distribution. So by now, I am sure all of you could answer, all of you could get it, why we need a probability distribution? So, from this from the type of problems we can solve, do not you see that if given if a probability distribution is given then from the probability distribution, we will be able to find out probability of occurrence of certain values.

Suppose, the weight of students or weight of child in a particular locality, suppose weight of child in a particular locality weight of child of a certain age in a particular locality. Suppose in the age of say 15 to 20, 15 to 20 in a particular locality, the weight of the children is suppose it is mean weight is say 40 with a standard deviation of say some x whatever may be the standard deviation.

So, now if I am interested if that means, if it is given the probability distribution here the probability distribution is given what it is given, we know that it is normally distributed, we know the mean, we know the standard variation standard deviation. Now, if we are interested

what is the probability that of getting a child of particular weight probability that of getting with 30 or 20 or maybe 50 or 60 given the distribution we will be able to tell that.

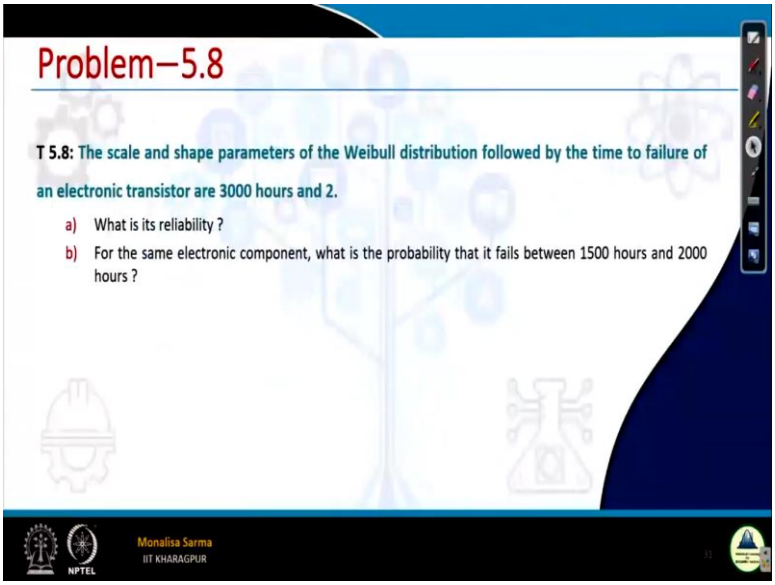
Now, the question is that how do we know given a setup data first initially when we are talking about this child weight in a particular area, who told me that distribution how do I know what distribution it is I just I will just have the data how will I how I will have the data there are some I have taken the weight of all the children in that all I cannot take I have taken the weight of the few students suppose in that areas there are around 1000 children out of one suppose, I have taken the weight of around 1000 children I have taken the weight.

So, from the weight how do I know this weight is normally distributed. So, for that there are again different techniques which I will not be discussing. So, given the set of data, so, from the set of data we can find out this set of data falls in which distribution and there are different techniques for that one of the techniques is probability plotting the many such techniques and power from, from the set of data.

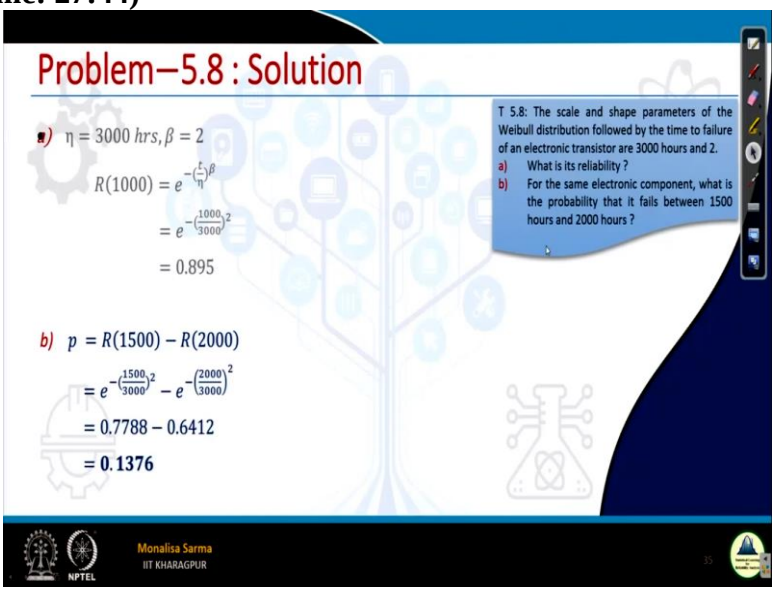
So, first step is that we have the data, from the data we come up with the probability distribution again from the once we have the probability distribution again from the set of data itself by parameter distribution, we can estimate the parameters that means, what are the parameters, if it is normal my parameters is μ and σ, if it is gamma my parameters is α and β, if it is exponential my parameters is only β, β or I can say λ or whatever I can say.

So, once we know that distributions, then the parameters are estimated of the distribution. So, we have the distribution, we have the parameters of the distribution, we have all the information. So, any probability for that we can find out from this distribution that is basically useful application of probability distribution. So, we are learning probability distribution mainly from this application perspective or not only in reliability studies and any studies and any analytics in any business decision and I think we will have quite a use quite application of this probability distribution.

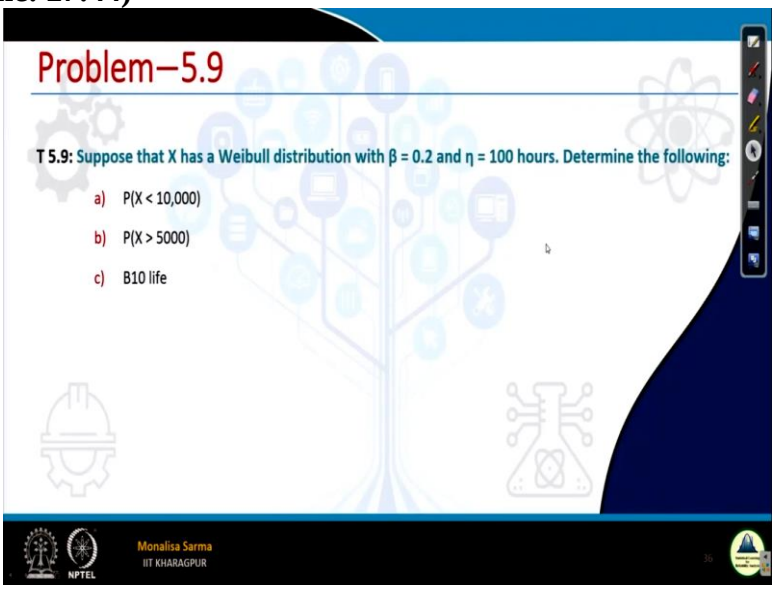
(Refer Slide Time: 27:41)



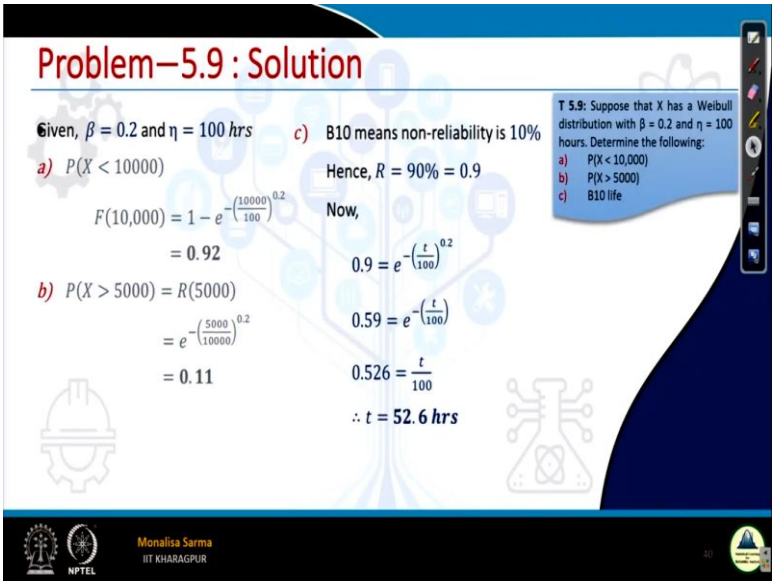
(Refer Slide Time: 27:44)



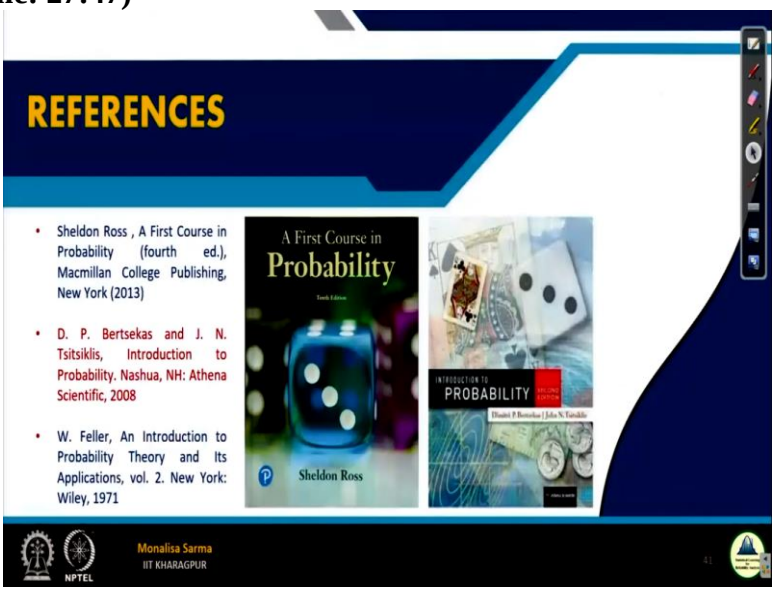
(Refer Slide Time: 27:44)



(Refer Slide Time: 27:46)



(Refer Slide Time: 27:47)



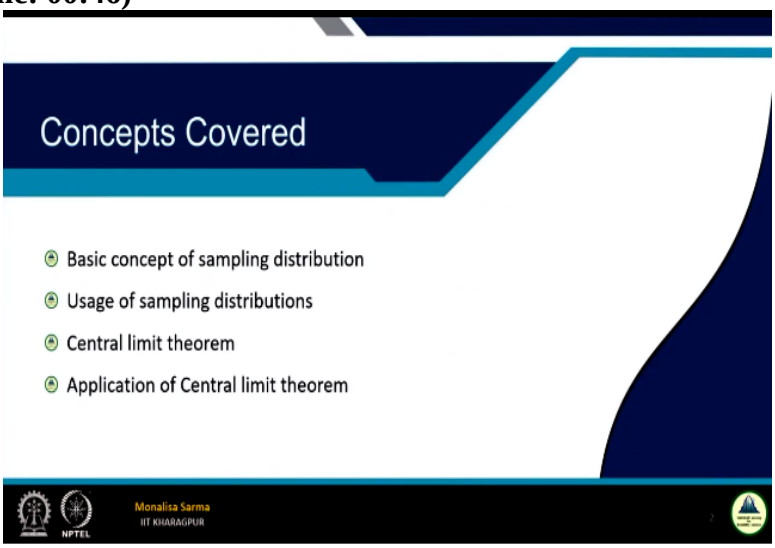
So, these are the questions for which are not be solving for this reliability problems. I mean for Weibull distribution and then I will conclude this tutorial class with the references and thank you guys. So, next class will start a new topic, a very interesting topic that is sampling distribution. Thank you.

Statistical Learning for Reliability Analysis
Dr. Monalisa Sarma
Subir Chowdhury of Quality and Reliability
Indian Institute of Technology – Kharagpur

Lecture – 16
Sampling Distributions (Part 1)

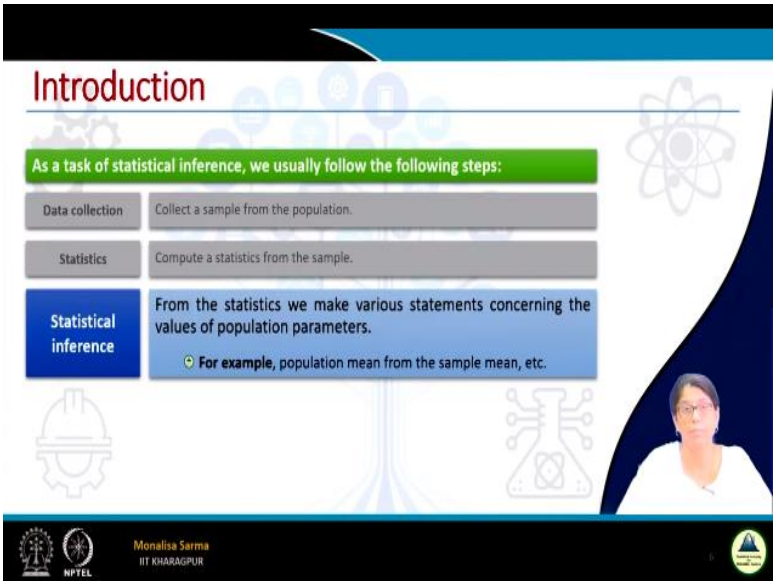
Hello, everyone. So today we are starting a new topic that is sampling distribution, it is a very interesting topic, I am sure all of you will enjoy doing this. So, the sampling distribution I will be taking in 3 different parts. So today I will be covering the first part and conjunctively the other parts.

(Refer Slide Time: 00:46)



So now in this lecture today, I will cover the basic concept of sampling distribution, then the uses why we need a sampling distribution, then I will talk about central limit theorem, a very well known theorem and of course, the application of the central limit theorem.

(Refer Slide Time: 01:02)



Now, as you remember, in our first class of statistics, latest from that is the second lecture, I have talked of statistical inference what is a statistical inference? If we can remember statistical inference if we talk about statistical inference, the main objective of statistical inference that what we use in most of our application is like given if we want some information about the population, we have definitely population is a big thing.

So, if we want some information from population, what we do is that we get a sample from the population of course, the sample has to be an unbiased sample and from the sample we try to extract the different information, which based on this information, we try to predict the information about the population that is recall the statistical inference in a very broad sense. So now, in the when we talk of statistical inference, what we need to do?

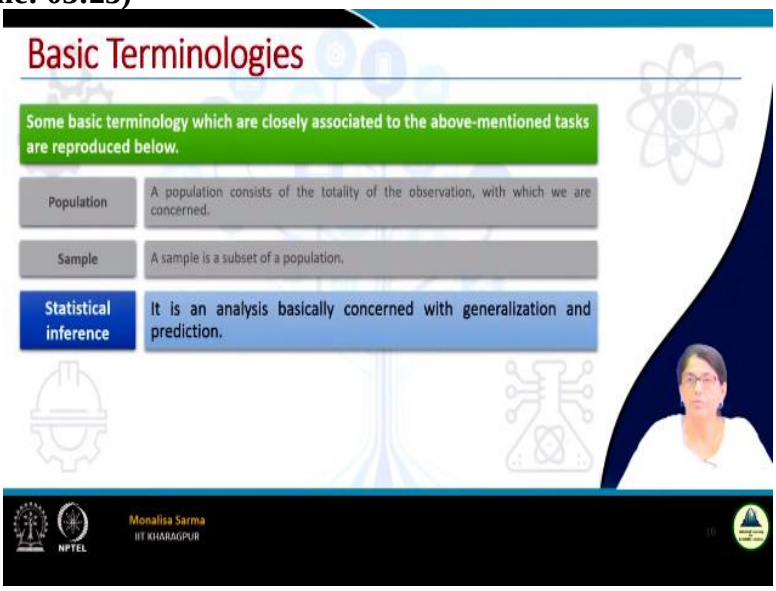
First we need to collect the data that is, as I told you, first we will have to collect a sample from the population what is sample, what is population this we have already discussed in my earlier lectures. Then from the sample, we will have to compute a statistics now, what is the statistics? In a population remember, we talked about the characteristics of the population when you talk about populations there, there are certain characteristics which define a population like it may be the mean it may be the variance.

So the mean and variance these are this we basically called a parameters of a population, please remember this we called a parameters of the population. Now the counterpart of parameters and sample we call it statistics we know longer call it parameter, we call it statistics. So mean of a sample variance of a sample, we call it the statistics of the sample,

rather than parameter of the sample but in case of population, we call it the parameter of the population.

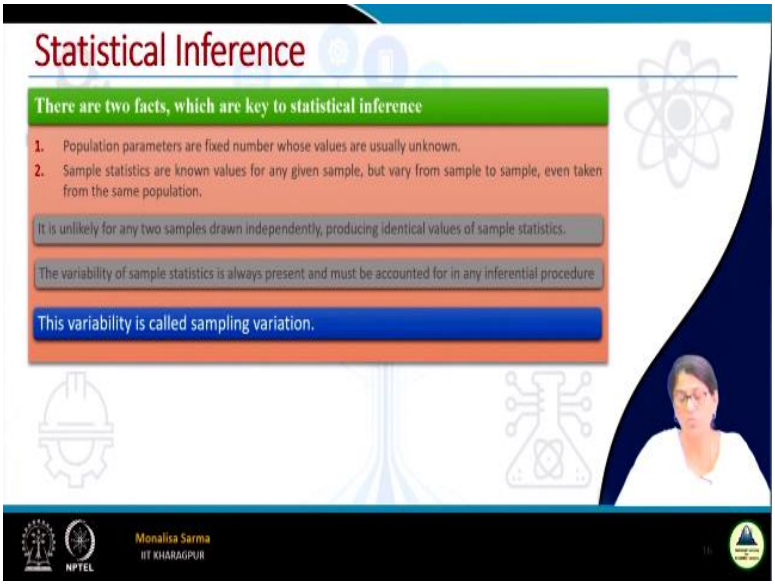
So once we collect the sample, the next job is we will compute the statistics from the sample. So, from the statistics, then we will make various statements concerning the different values of the population parameter. For example, we may talk about the population mean from the sample mean, if we get some we find out the mean of the sample from there, we will try to infer something about the population this is one sort of statistical inference. Likewise, there can be in many other statistical inferences, which you do for a population from a sample.

(Refer Slide Time: 03:23)



So, some of the basic terminology which are closely associated with the statistical inference basically, our population which I have already discussed and not be repeating here than sample then for it is statistical inferences already we know all this.

(Refer Slide Time: 03:40)



So, now, coming straight to this, there are 2 factors which are keys to statistical inferences, what are the first is population parameters are fixed numbers whose values are usually unknown. So population parameters, when you talk about suppose a mean of a population is a fixed parameter, but then it is really not possible to know the mean of a population like let me take a very small example suppose I am considering the some LED bulbs a company is producing a bulk of LED bulbs that light.

So now, if we consider it a whole lot that lot that is the population. Now, if I want to find out the mean what to say mean lifetime of those bulbs, then it is really not possible to find out the mean lifetime of those bulbs, by considering the whole lot, because there may be say in 1 lot blub, 1 lot bulb is produced in a blub. So if I want to find out the mean lifetime. So I need to know the lifetime of all the bulbs.

Then using whatever formula we have for mean you already known the formula for mean, using the formula for mean will have to tell the mean lifetime of the bulb, but that is nearly not possible. We cannot take the mean lifetime of all the bulbs. So, now so what is the remedy done sample statistics are known values for any given sample, but vary from sample to sample even taken from a very small population. So, now, we are interested in mean lifetime available.

So, we really not possible for us to calculate the lifetime of all the bulbs and then find out the mean. So, what we do is that the normal procedure is that we take a sample from the population a small sample maybe and from that sample we try to find out the mean lifetime

of the bulbs we take the lifetime of all the bulbs how do we take the lifetime of all the bulbs do a different experiment basically, maybe what is say we use the bulbs for quite a long till it goes bad.

So, that definitely we cannot do of all the population we will take a smaller sample and from the smaller sample we keep the light on till it goes back and then from that we can find out a mean lifetime of the bulbs this is just an example of course, this is not the way to find out the mean lifetime of the bulbs there are different other techniques also to find out these values these statistics, so, I will not valid into those.

So, now, if we take a sample from 1 sample suppose, I have taken a sample of 10 bulbs from 10 bulbs if I take the lifetime and then I find out the mean I get a particular value say x. Now, suppose I take a different sample and then again I take the lifetime of those value when I take the mean I get maybe that is x 1, this x is not equal to x 1 there are very less probability that x will be exactly equal to x 1 there will be some variation between this x and x 1.

So, this sample statistics so the second point that means, sample statistics we can know it we take a sample from the sample we can find out the sample size because this is a small number small amount, but then this small sample statistics vary from sample to sample even we have taken it from the same population. So, it is very unlikely that 2 samples drawn independently producing identical values of the sample statistics.

So, the variability of sample statistics is that means always there. So as I told you in the beginning when I talk of statistical inference. Statistical inference means we will take a sample from the sample we will try to find out the statistics based on the statistics we will try to infer about the parameters of the population. Now, see the dilemma here the samples that is when we talk about the sample statistics, the sample statistics vary from sample to sample, I have taken a sample of 10 numbers.

I have taken another sample of 10 numbers, the statistics we say get from this first sample, which I get from the second sample, it is very unlikely that I will get the same value. So, there are variations. So, if they are a variation then how do I do the statistical inference. So, that is the thing. So, this variation we must take into account this variation in this inferential

procedure. So, how do we do that? So, this variability is called sampling variation, because variation around the samples so it is called sampling variations.

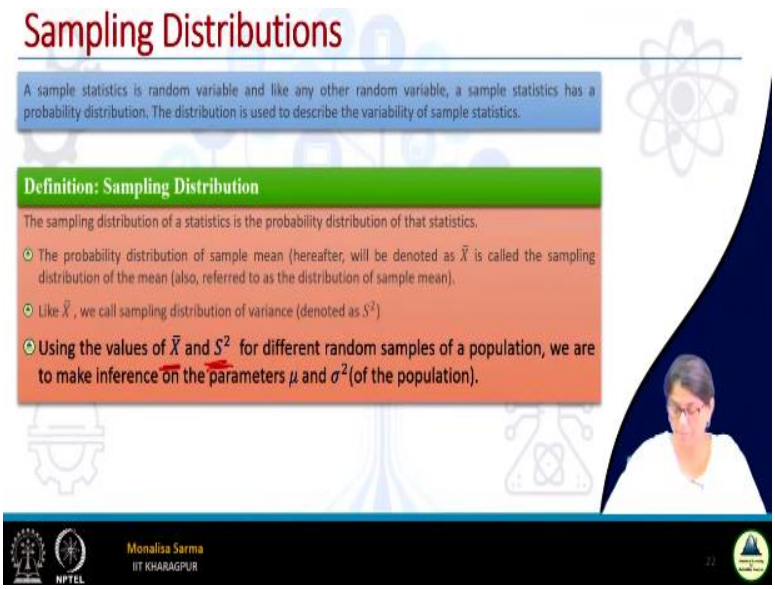
Now, see when we talk about this say population mean only sample mean so, we take a sample we get a mean another sample we get another mean say first sample mean I got a x 1, we take another sample, we got the same mean x 2, we take another sample say, we got mean x 3. So, this sample mean will be different and the probability of getting this value is different there it is also different.

So, now remember what is probably a distribution? Probability distribution is nothing but the value that a random variable will take. And the probability of those random variable these 2 things, if we put together it is the only problem the distribution Now, here what our sample mean or variance or whatever statistics we consider. So, this sample statistic is a random variable, because it may take different values, it is not that it will take a fixed value, so it will take different value. So, our sample mean is a random variable.

And since this is a random variable, we can also have its probability of occurrence. So, these 2 things put together this sample mean, it is a random variable. And since it is a random variable, it has different probability of occurrence. So, considering these 2 things together, we can have a probability distribution of that. So, this probability distribution is nothing but it is called a sampling distribution. So that is it that is the distribution sampling distribution is nothing but a probability distribution only.

But it is a probability distribution of the sample statistics that is also a random variable. Now what is this random variable? This random variable is a sample statistic that is why we call it a sampling distribution.

(Refer Slide Time: 10:09)



Now, this distribution so, when we have a distribution what is the use of probability distribution I have talked already remember why we need to study probability distribution one of them important objective of studying probability distribution is that we can find out the different values that are random variability and with this what we can find out? We can find out the probability of occurrence of a particular random value.

So, because now we have for the sample mean or the sample statistics, let us not take mean because there can be any other statistic for the sample statistics now, since when we have a distribution. So that means we from this distribution, we can use to describe the variability of the sample statistics what are the different values that sample may take as well as this probability of occurrence.

So, the definition of sampling distribution, the sampling distribution of a statistic is the probability distribution of that statistic. So, probability distribution of sample mean, we remember for the population mean, how do we designate the population mean we denote by mu you remember and population variance removed we denote by σ^2. So, similarly for sample mean we will no longer use mu rather we will use either X bar or Y bar general convention is using X bar in some books you will also find using Y bar.

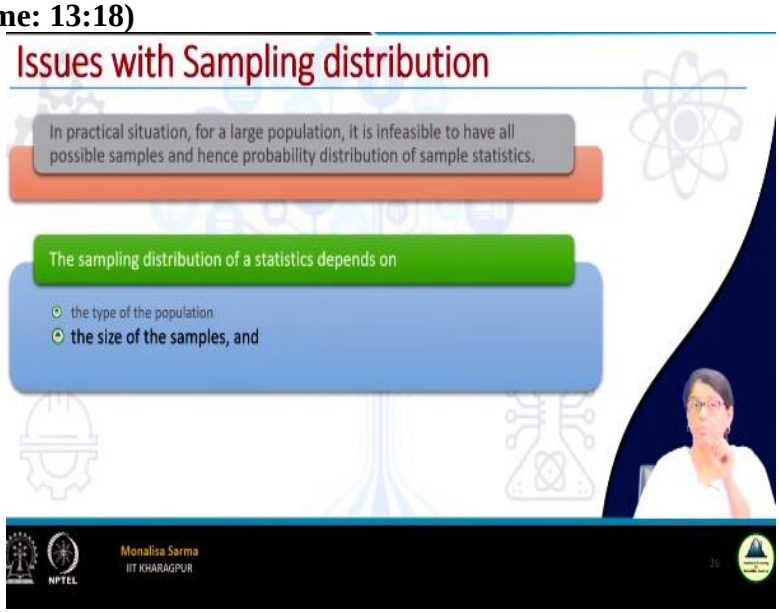
So that is the general convention of using denoting sample mean. So, now when we do the distribution of X bar that is a sample mean then we call it as a sampling distribution of the mean. So, the probability distribution so say this line the probability distribution of the sample mean is called the sampling distribution of the mean. So this if we find out the

probability distribution of sample mean that we call it a sampling distribution of the mean similarly, the sampling distribution of the variance also we will discuss later.

So, sampling distribution of variance, we will be discussing in the next class, this class we will continue our discussion on the under sampling distribution of mean. So, using the values of this X unlike for population variance, we use the term σ^2. Similarly, for sample variance, we use the term S^2 this is the population sample mean it is X bar Y bar for sample variance we use S^2 and all the books you will be finding it is denoted by S^2.

So, using the values of X bar and S^2 for different random sample of the population, we are to make inference on the parameters of population that is parameters μ and σ^2 so that is statistical inference.

(Refer Slide Time: 13:18)



Now, how we make the sampling distribution because infer about a population we need to have the sampling distribution because just taking 1 sample will not do because, by them as we have already seen the sample statistics differs from different samples. So, that is why we need the sampling distributions once we have the sampling distribution then only we can infer about the population. Now how to have the sampling distribution?

So to like be, we have learned about probability distribution remember, we have learned in to have a probability distribution what we need? We needed different values that a random variability as well as the probability of occurrence of those values, what is the probability of occurrence? Probability of occurrence is nothing but the relative frequency is not it? A

relative frequency of occurrence. So now similarly so how we will construct a sampling distribution so given a population.

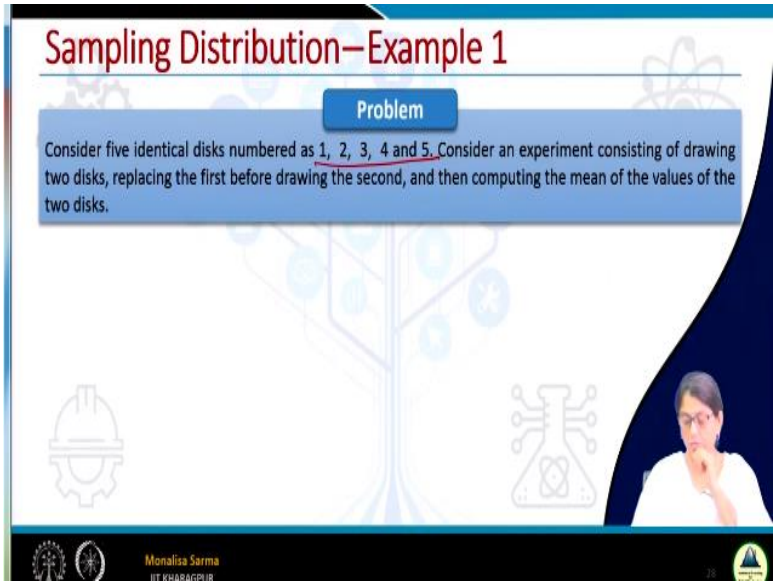
So, what we will do? We will take as many sample as possible given a say big population from there suppose, we take a population initially suppose, we took a population say first 1 sample from there suppose, we got a mean of X 1 bar then again X again we took another 1 sample and different set of sample say we took sample mean X 2 bar another again we take a third sample for the suppose we take with likewise we suppose took different samples from different samples we found different means.

So, it may be that again there may be some samples which made mean value maybe same So, based on this we can find out what is the relative occurrence of this sample mean relative occurrence of X 1 bar, relative occurrence of X 2 bar, relative occurrence X 3 bar. So, relative occurrence of that particular variable what is that that is nothing but the probability that so, with this we can calculate that we can, what to say develop the sampling distribution.

Now for a huge population, is it really feasible to take out all sub samples and then find out the sample statistics? It is really not possible we may miss out on some sample. So, we do not have to do that many researchers have already done that and based on that, they have already come up with some based on different empirical study they have already come up with some theorem. Now, I go to the theorem later, but firstly just to see how the researchers have come up with different theorems for that, I will just take a small example.

First of all, the sampling distribution of the statistics depends on what it depends on the type of the population, whether the population is normal distribution or totally not normal distribution or near normal distribution or some other distribution, why I am stressing on normal I will come to that again. It also depends on the sample values it also depends on the size of the sample, it also depends on a method of choosing the samples whether the sample is a bias sample or a non bias 1.

(Refer Slide Time: 16:35)

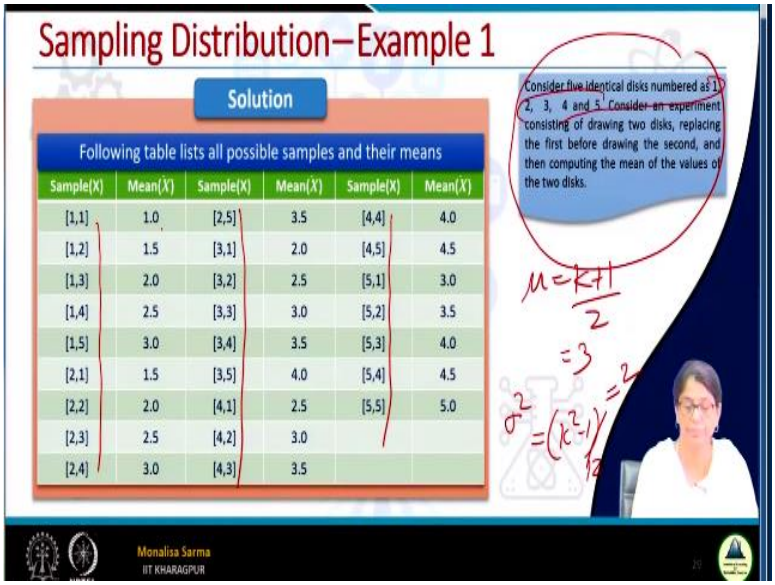


So, what is a bias sample, what is a non bias sample? I have already discussed. Now, coming to the example how we will find out the sampling distribution let us take a very small population suppose and we are considering a population where there are only 5 identical disks, this disks are numbered say 1, 2, 3 and 4 and 5 identity this is my population, that is very small population. So consider an experiment consisting of drawing 2 disks, replacing first before drawing the second what?

I will draw 1 disk I am replacing that and I will draw it again second one the maximum is what to say drawing 2 disks and then computing the mean of the values of the 2 disks. My interest is I want to compute the mean of the values of the 2 disks. Why I am doing that, my main objective is to find out the mean of the population. Of course, since this is just a 5 value, so we can I can very easily find out I mean, my if there is a population of huge number, it is really not possible to know the mean.

So here, let us not take it that way let us take this is a population, my intention is to find out the mean of the population somehow this mean of this population, I will try to find out by taking a sample. So what is my sample? Sample is 2 disks. So, how I am picking 2 disks, I am picking 1 disk and I am replacing it back again, I am picking the second disk, so I am replacing it back that is the way I am picking 2 disks.

(Refer Slide Time: 18:08)



So, this way, I am picking the disk see what these are the different samples I may get the total 5 samples the different possible samples, these are the samples you can see these are the different possible samples all possible here I have because this was a small population, it was possible for me to find out all possible values. So, I found that the all possible values of the sample these are my different way how I can pick 2 values 2 disks?

Now, what I have done is that so, if I pick knows see this, first of all consider does this 5 identical this picking 1 and replacing it. And again, I am picking 1, what is the probability of this picking, what distribution is first? This is definitely a uniform distribution is not it? It is same as rolling a die and coming which value is coming on the top it is the same thing. So it is very much a uniform distribution. So, uniform distribution, what is the probability f(x) is 1 / K where K is the total number of digits. So here what is the probability 1 / 5.

Since this is a uniform distribution, so now in a uniform distribution, can you remember what the how do we calculate the mean and the variance? So if considering this as the population, so if I need to calculate the mean, what is the uniform distribution calculation of mean if you can remember, μ= K + 1 / 2. Is not it? μ= K + 1 by 2. So what is here K? K is 5, 5 + 1 / 2 that is μ= 3 for this population that means my μ= 3 since a small populations I knew all about so I could find.

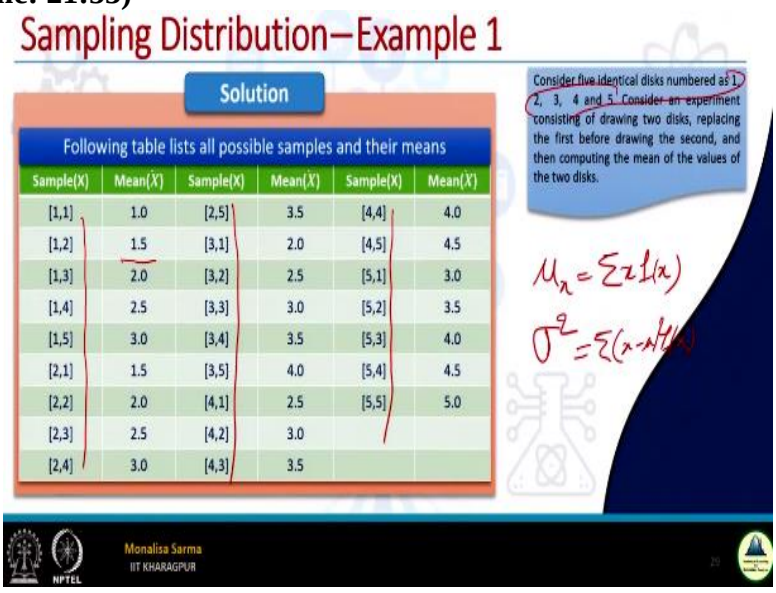
So, what is my variance of this population variance formula if you can remember K^2 - 1 / 12 variance formula we have discussed into when we have discussed any uniform distribution if we put the value of K as 5, so, what do I get? So, it is this I will be getting equals to 2. So, my

μ= 2 for this population my μ= 3 and my σ^2 is equal variances = 2, that is for this population μ= 3 σ^2, because this is a uniform probability distribution that is now come to the sample what sample we have taken?

We have picked out all possible values of the sample all possible samples taking 2 together, these are, there is no other way by which we can pick 2 samples, there is all the possible scenarios are jotted down here then from all of these 2 samples, I tried to find out the mean. So, 1 and 1 if I pick 1 and 1 that is the mean. Mean is definitely 1 if I pick 1 and 2 mean is definitely 1.5 1 + 2 divided by 2 is not it? It is definitely 1.5 definitely I found that the mean.

Now, if this is the case, if this is the sample this is the mean we know what is the formula for mean? In a probability distribution given a probability distribution when the distribution is specified, whether it is uniform distribution, binomial distribution or whatever distribution whenever distribution is specified, then we know for each specific distribution, what is the value of the formula, but actually what is the main formula? Main formula μ is x f(x).

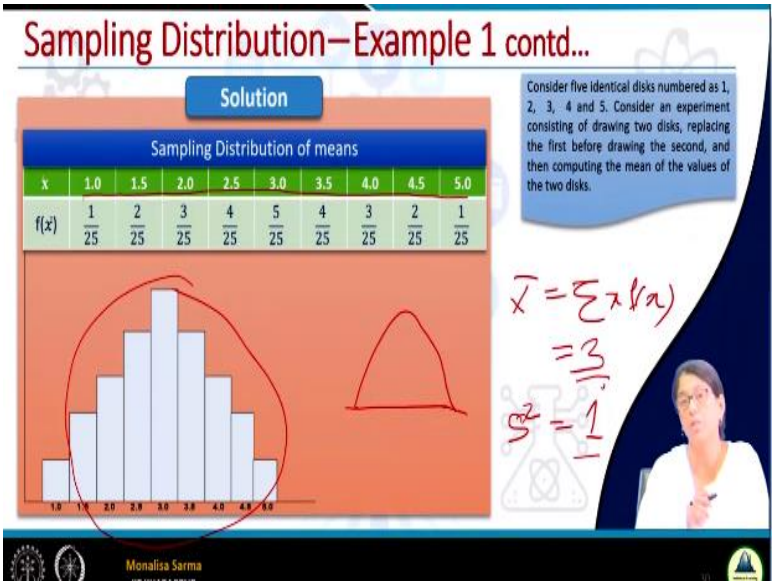
(Refer Slide Time: 21:53)



Sample(X)	Mean(X)	Sample(X)	Mean(X)	Sample(X)	Mean(X)
[1,1]	1.0	[2,5]	3.5	[4,4]	4.0
[1,2]	1.5	[3,1]	2.0	[4,5]	4.5
[1,3]	2.0	[3,2]	2.5	[5,1]	3.0
[1,4]	2.5	[3,3]	3.0	[5,2]	3.5
[1,5]	3.0	[3,4]	3.5	[5,3]	4.0
[2,1]	1.5	[3,5]	4.0	[5,4]	4.5
[2,2]	2.0	[4,1]	2.5	[5,5]	5.0
[2,3]	2.5	[4,2]	3.0		
[2,4]	3.0	[4,3]	3.5		

What was mu of X? μ of x is summation of x f(x) is not it? This is how we find out that mu and mu have a distribution. And by putting values in this x and f(x) that how we have simplified and I found out different formula for different distribution, but this is the generalized formula for μ x is this and similarly, we have the general formula for sigma square as well remember is equal to summation of $(x - \mu)^2$ f(x).

(Refer Slide Time: 22:46)



So, now, this whole thing this whole table, if we want to put it in a form of a thing, what is that histogram that means if you want to put it in a form of a probability table, see how we can put it say this is the different values of what we got here the different values the mean takes the different values, we found out the different values mean can take this all these are the different values that mean can take there are no other values.

So, what is the probability that mean takes value 1 we have taken the relative frequency. So, relative frequency how much is the time that a mean takes 1 see here mean takes 1 that is only 1 value is there only once and how many total combinations there are total 25 combination. So, what is the relative frequency distribution is 1 / 25. Similarly, for 1.5 we found what is 2 / 25 this is how we found out this is nothing but the probability distribution this is in a tabular form, is not it?

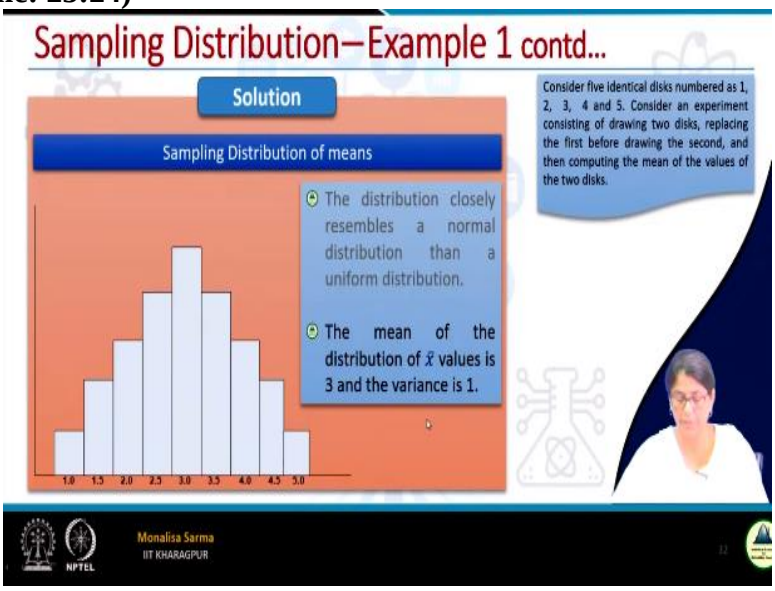
This is nothing but the probability distribution in a tabular form, this is the X bar is a random variable f(x) bar is nothing but the probability of X bar. This we can see it in the form of a histogram also, this is the histogram of this corresponding data. Now, given this distribution, if we find out the mean, because I do not know what distribution it is, I do not know whether it is uniform distribution, normal distribution, this distribution I do not know what distribution if I want to find out the mean.

Then what is the formula for mean the general formula? General formula as I told is mu is equal to summation of x f(x). So, if I do x f(x) $\mu = \sum x f(x)$, x = x goes from 1 to n. So, x of f(x) means 1 x 1 / 25 + 1.5 x 2 / 25 + 2 x 3 / 25 + 5 x 1 / 25. You can calculate this and you

will find that you will get 3. If you calculate this you will get 3 similarly, if you calculate the variance formula what this summation of $(x - \mu)^2 f(x)$.

So, if you calculate that σ^2 you will get that 1 I have done a mistake here. And since this is a we are talking of in terms of sample, so, I will not use the term μ and σ^2 what I will use remember and use X bar and here I will use S^2 because this is in terms of a sample, this is what I got in terms of for this probability distribution.

(Refer Slide Time: 25:24)

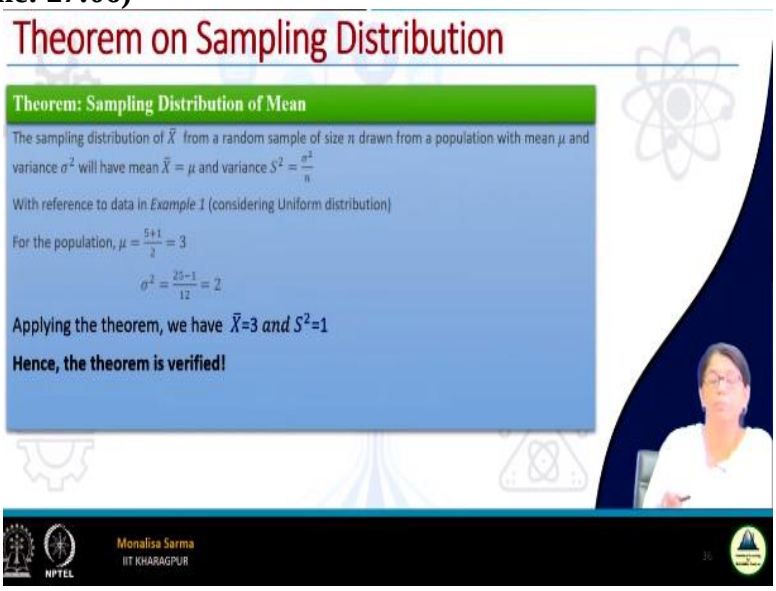


And this distribution you can also see from the table this is the table when we have drawn the histogram, but we have seen this distribution almost resembles a normal distribution, what is the normal distribution how is the normal distribution curve is this and this distribution also this figure also resembles a normal distribution. So, what we got from this calculation when we took all possible samples our sample of size 2 even if you take sample of size 3 whatever it is.

We know, for this explain suppose, we took all possible sample when we took all possible sample and we found out the probability distribution and this probability distribution very much resembles a normal distribution and thus the mean of this distribution we got 3 remember this population mean was also 3 K + 1 / 2 this is a uniform distribution that also is equals to 3 we found and this variance we found 1 and what was from this population, what was the variance we found variance 2 here we found variance 1.

So how is this link? Let us see here. So, this distribution closely resembles a normal distribution, our parent population has uniform distribution, is not it? Our parents population was very much a uniform distribution from this uniform distribution, we have taken a sample of 2 items and the resulting distribution that we got is our, this is a sampling distribution is not it? This resulting distribution is very much a normal distribution. And the mean of the distribution of X bar value is 3 and the variance is 1 if we calculate using the formula x of x, $(x - \mu)^2 f(x)$.

(Refer Slide Time: 27:06)



So, what we got? The sampling distribution of X bar from a random sample of size n drawn from a population with mean mu and variance sigma square will have mean X bar = μ and $S^2 = \sigma^2$ by n is not it? That is what we got for these experiment what we got? We got our X bar = μ, our populations mean was also 3 our sample mean also we got 3 that means our X bar = μand variance S^2 variance of the sample we got 1, what is the variance of the population of 2?

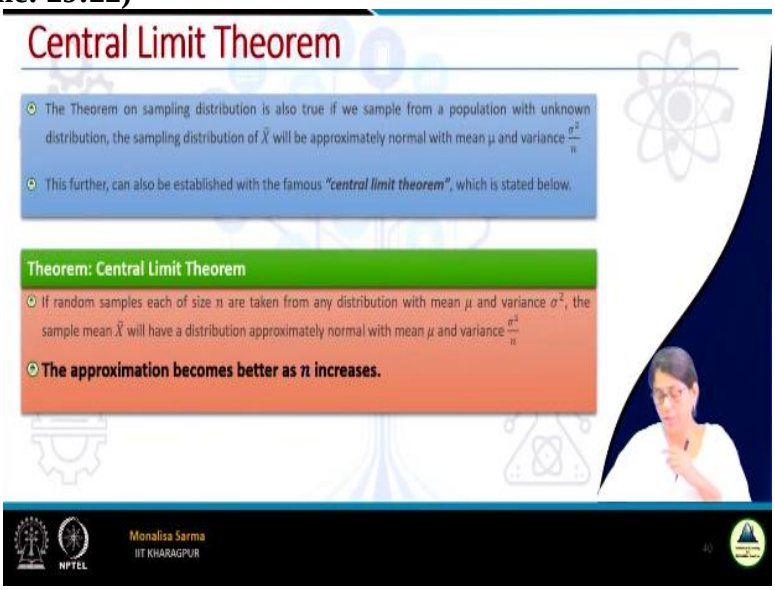
So, how we will get that means sigma square that is variance of the population divided by the size of the sample, what is the size of the sample? Size of the sample is 2 so 2 / 2 = 1. So, this is a very important theorem, this is sampling distribution of mean this is this we got for this experiment. Similarly, the researchers then experimenters they have done empirical study on many subpopulations.

Of course, infinite population, they cannot do they have done on very small population, medium population, large population they have done and from all those empirical

experiments, they could come up with this theorem this very important theorem, I am reading the theorem again the sampling distribution of X bar what is X bar? X bar is a sample mean X bar from a random sample of size n drawn from a population with mean mu and variance sigma square will have mean X bar = μ.

And variance sigma square S square = sigma square by n this is called a sampling distribution of mean. So, now, with reference to the data from what we have uniform distribution already I have shown you we got $\mu = 3$ $\sigma^2 = 2$. So, our theorem is basically proved we could prove this theorem basically, we already know that theorem and we have done an example to prove, but actually the researcher has done many experiments together and they could come up at last they could come up with this theorem, this is called sampling distribution of mean.

(Refer Slide Time: 29:22)



So, this theorem on sampling distribution is also true if we sample from a population with unknown distribution, this is what this population what we have considered this is a very much a known population, we know that it is a uniform population, even if we sample from a very much unknown population also the sampling distribution of X bar will be approximately normal we have seen it is approximately normal this size the sampling distribution of X bar will be approximately normal with mean μ and variance σ^2 by n.

So, this further can also be established with the famous central limit theorem this another theorems central limit theorem, this is basically again similar based on same set of experiment one most stronger theorem than the sampling distribution and sampling

distribution theorem, we see the sampling distribution theorem here it does not speak about the shape of the sampling distribution say this is the sampling distribution of mean.

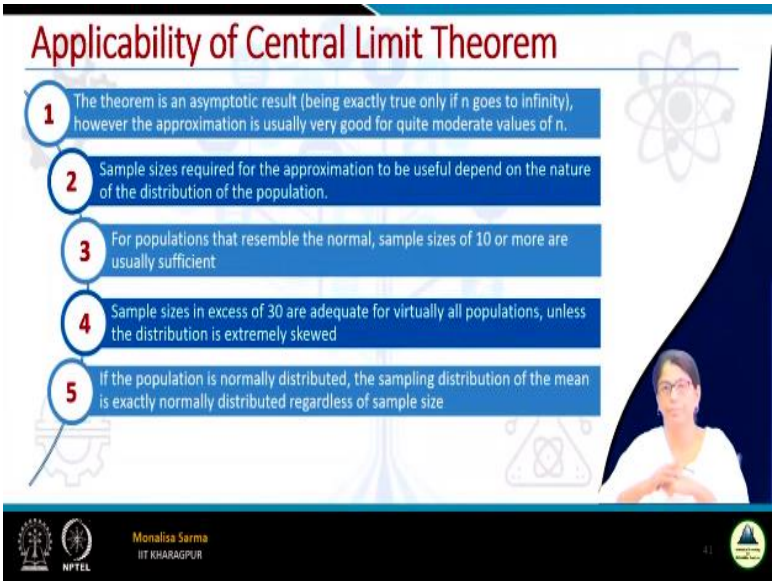
The sampling distribution of X bar from a random sample of size n drawn from a population with mean μ and variance σ^2 will have a mean X bar = μ and variance as far as well as σ^2 / n. It does not speak about the shape of the sampling distribution remains silent on that. So, but already we have seen we have seen the sampling distribution take same normal shape.

So, central limit theorem is a more stronger definition than sampling distribution, it also talks about that. So, what is the central limit theorem say this is a central limit theorem, the very important theorem, if random sample each of size n are taken from any distribution with mean μ and variance σ^2 μ and σ^2 is the parameter of the population the sample mean X bar will have a distribution approximately normal with mean μ and variance σ^2 / n, yes everything same.

But here also it talks that a distribution is approximately normal. So, it is irrespective of the parent population whether the population your normal population, binomial population, normal population whatever population it is, but this theorem it satisfies this theorem, but there is another twist to it that means, if the population this approximation becomes better as n increases this is approximation. This approximation mean is value of μ and value of X bar and S.

But the size or the shape that shape I told that errors distribution is approximately normal that distribution is approximately normal distribution becomes better as n increases for very small value of n, we may not get a normal distribution, when the n becomes more when n is becomes greater than definitely we will get a normal distribution.

(Refer Slide Time: 32:19)



But here, there is another twist to it also see the different points the see all the points are very important, please look carefully that this theorem is an asymptotic result that means being exactly true only if n goes to infinity when our sample size becomes large, this theorem is exactly true. However the approximation is usually very good for quite moderate values of n. Sample sizes required for approximation to be useful depend on the nature of the distribution of the population.

Now, when I told that is when the sample size goes to infinity, this approximation is very good again in fact, it is exactly true when sample size goes to infinity and even for a moderate values of n also this approximation is very good. Now question that arises what should be the sample size? So, this approximation depends on the distribution of their parent populations. If my parent population is normal with my parent population from where, I am taking the sample.

If my parent population is normal, then we do not need a very bigger sample size and a small sample size will also do but in the population is a bit away from normal it is not normal, but a bit totally not normal, but a bit away from normal then we need a moderate value of n, but it a population is very much skewed then we need a very big value of n. So, population that resembles a normal sample size of 10 or more are usually sufficient if it is normal or if it resembles normal not exactly normal.

Sample sizes in excess of 30 are adequate for virtually all population unless the distribution is extremely skewed. It does distribution is extremely skewed then we need a very big sample

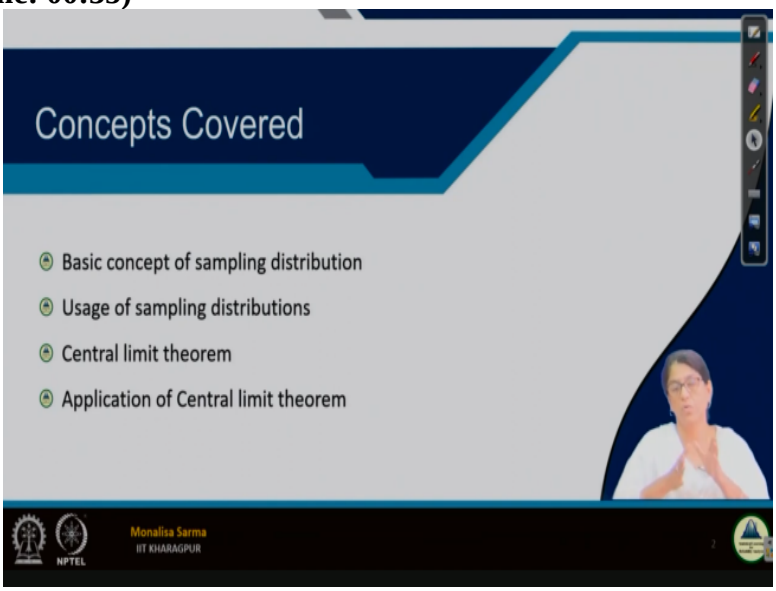
size maybe to tune of say 100 if the population is normally distributed, the sampling distribution of the mean is exactly normally distributed regardless of the sample size. So, I there are a few more points to be discussed about this sampling distribution about this central limit theorem basically. So I end this lecture here. In the next lecture, I will cover those things. Thank you.

Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology - Kharagpur

Lecture - 17
Sampling Distributions (Part 2)

Hi, welcome back again so, last class like we were discussing sampling distribution last class was the first class on sampling distribution.

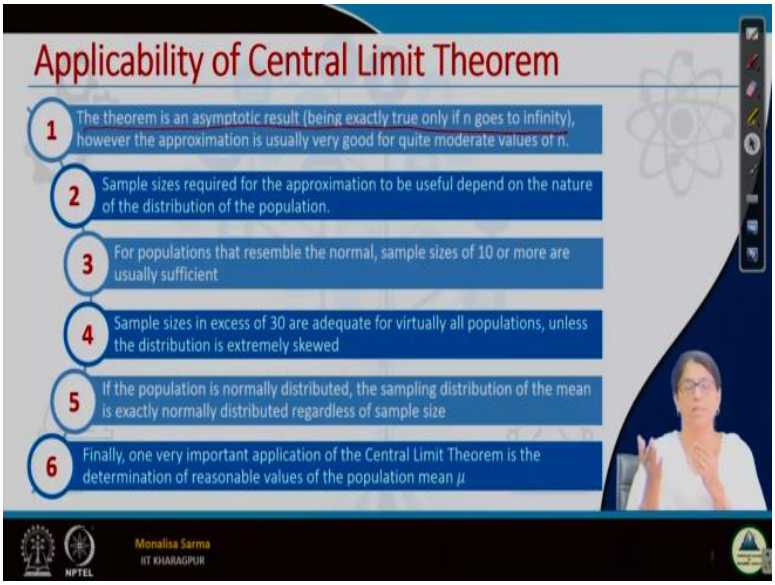
(Refer Slide Time: 00:35)



And then, so, in the last class, we have covered what we mean by sampling distribution. So, and then we have also seen the uses of sampling distribution, then we have discussed central limit theorem, but why I have given this as a concept covered because this class is basically a continuation of the last class, I could not cover the whole portion because of the time constraint.

So, I have given the same concepts I am showing it here as well. So, it is just basically the extension of the last class.

(Refer Slide Time: 01:06)



So, now, where we have stopped last class is that we were discussing the applicability of central limit theorem, remember, what was some central limit theorem. So, basically, when we have to infer something out of populations, because population is a big number, we cannot infer and infer something specifically we some summary statistics, like say, we want to know about the mean of the population.

We want to know the variance of a population, like say, we want to compare 2 different populations. So, these are the things when we want to know something about a population, since it is a big population is a big unit. So, what we do is that we take a sample out of it from the sample, we calculate the statistics. And from the statistics, we try to infer about the population parameter, I am reminding you again when we talk of sample we call it statistics.

And when we talk of populations, the characteristics of a population we call it parameter. So, statistics is the counterpart of parameter in sample basically. So, we have already seen why sampling distribution first, what is sampling distribution that I think you can remember sampling distribution is nothing but a population distribution only. And why we call it a sampling distribution?

Because, here the random variable that the random always when we have a population, probability distribution, it is always for a random variable, is not it? A probability distribution is always for an random variable, why? Because the; random variable can take any values, and then what is the probability of taking these different values. Now, here in case of a sampling distribution, my random variable is a sample statistics, is not it?

So, my random variable is a sample statistics. So, and this sample statistics, there are lots of variants to the sample statistics; if we take different samples, we will get different statistics value will be different, there is a variation since there is a sampling variation. So, to and address these variations, what we do is that, we find out its distribution. And that is why we call it as a sampling distribution.

We do not call it probability distribution; we call it sampling distribution and since, we have already seen in the last class, how to, what to say, basically, how to have the sampling distribution, how to construct the sampling distribution. To construct a sampling distribution, we need the different values of the random variable we will take that means different variables that this sample statistics will take different value as well as its probability of occurrence.

If I let me not say the probability of occurrence let me say is the relative frequency of occurrence. So, how do I find that? So, that is really not possible to find out the different value that the sample statistics will take as well as the frequency or the relative frequency observation and that is really not possible, because for a sample, we have to take all possible sample for a population will have to take all possible samples from all possible samples only then we can find out its relative occurrences is not it?

So, since that was not possible, so, we took help of some theorem. Now, what is that theorem? The first theorem that we have learned is a sampling theorem, is not it? So, sampling distribution of mean; that was the first theorem. So now, we have also proved that theorem how we prove that theorem for proving the theorem, we have taken a very small populations, our population.

That population if we can remember we are considered a what to say a uniform population from the uniform populations, we tried to take out the sample, sample of size 2, and for least sample, we try to find out what is the mean and then from that, we could basically calculate different statistics of the sample as well as since it was a uniform population, we have calculated the parameter values of the samples that was the mean and the variance.

And we could prove that, the theorem, which is there the theorem is correct, anyway, we do not need to prove it this big, big researchers they have done it, but for our own curiosity we saw it is really true, but we definitely cannot do it for a bigger population. So, for a smaller population, we have seen that is right, but the sampling distribution of mean that distribution, it did not speak about the shape of the distribution, what distribution is the, it just talks about the mean value as well as the variance value.

Then we infer that we have referred to a more stronger theorem, that is a very well known theorem, that is the central limit theorem, the central limit theorem not only talks about the mean and the variance of the sample statistics of the sampling distribution, but it also talks about the shape of the distribution. So, what was the shape? It is approximately normal. Now, then, we talk about the applicability of the central limit theorem here basically, we stopped.

So now, what was the applicability? Applicability also, I have discussed first of course, this theorem is exactly true when the sample sizes sample size goes to ∞, but then that is also not necessary if it has a moderate value that also it is good and it if our parent population is normal, then a sample size of around 10 is also sufficient a dependent population is it is not normal, it is not very much also skewed then a model size of 30 is enough, but the problem sample is very much skewed, then we need a very bigger sample size.

Now, finally, one very important application of central limit what is the important application of central limit theorem? Can you just tell me with the knowledge of probability distribution? From this probability distribution, where we have used probability distribution till now, I have discussed probability distribution in many numbers of classes. So, one of the main uses of probably distribution was given a probability distribution, if we have the distribution.

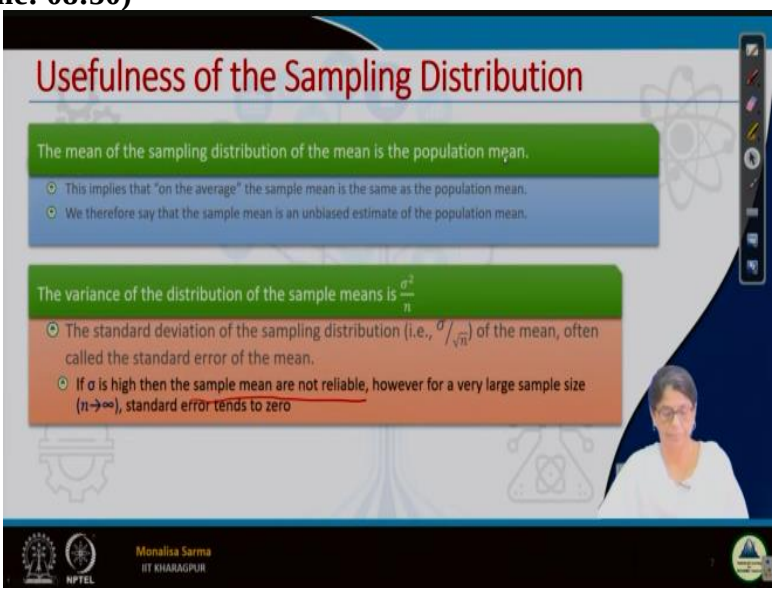
If we have known the mean of the distribution, if we know the variance of the distribution from that distribution, we could find out the probability of occurrence of a particular value, is not it? Like suppose we have a probability distribution of the lifetime of a bulb. So, if lifetime of a bulb follows a particular probability distribution, if we know that if we know the mean of the distribution, if we know the variance of the distribution, it is normal distribution, definitely central limit theorem means we are talking of normal distribution.

So, if we know the mean and the variance of this distribution, then everything is known to us like then from that we can find out the probability of occurrence of any value like if we are interested suppose, if you are interested in finding out what is the probability that the lifetime of the bulb is say around them, say 2 years. So, we have the distribution, we have the mean we are the variance from that we will be able to calculate.

We have done this sort of example, if you cannot remember please I suggest please just to have a quick recap of those probability distribution thing, because we will be needing it very much in this coming lectures as well. So, that is what so, one important application of central limit theorem is the determination of reasonable values of the population means that is the whole story, is not it?

That is why only we are doing that, why are we can I am developing the sampling distribution, we have to talk we want to talk something about the population. So, what we want to talk? We want to tell about the population we are mind it here, we are talking about sampling distribution of mean we are finding out the sampling distribution of mean that means why we are doing it, we want to talk something about the population mean. So, this is one of the important applications of central limit theorem.

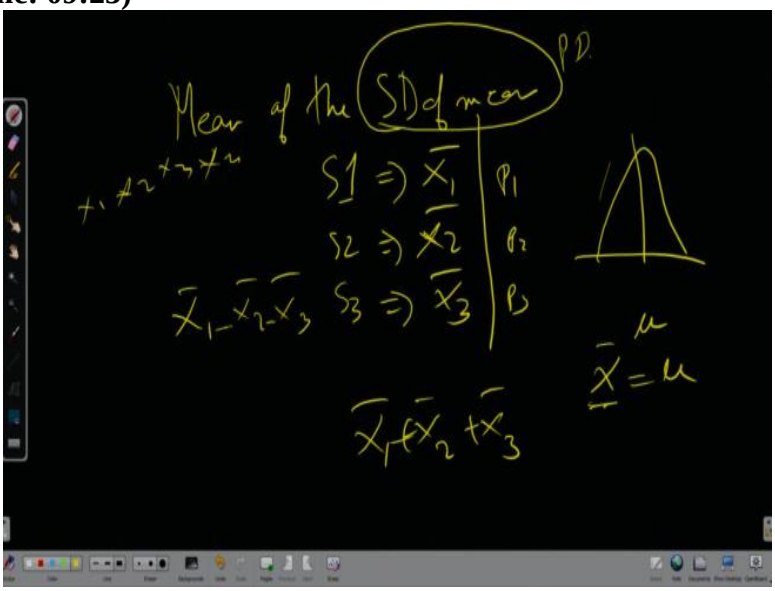
(Refer Slide Time: 08:50)



So, the sampling distribution of the mean what we have seen it is the population mean, is not it? Sampling distribution so, when we have drawn a sampling distribution, how we have drawn different means, different samples have different means, and it is different probability

of occurrence based on that we have drawn the sampling distribution, is not it? Now, what is this mean? The mean of the sampling distribution we found it is equals to the population mean.

(Refer Slide Time: 09:23)



So that means see so, 1 sample suppose I have taken sample 1 and my this population is say X 1 bar another sample my population is say X 2 bar another sample say my population is say X 3 bar. So, along with that, I have also probability of this occurrence, say some probability p 1 p 2 p 3, whatever it is from that I found my sampling distribution that is not an issue. Now, my issue here is that, what is the mean of the sampling distribution.

I already from the central limit theorem, I have known that my distribution shape is normal, and I have known my mean is equals 2 population means now the population mean is μ that means my sampling distribution mean is also it is μ. Now, what is this X bar? X bar is basically mean of all this is not it? Mean of X 1 bar + X 2 bar + X 3 bar, this is nothing this mean of all this is my X bar.

That is, that is when I talk about mean of the sampling distribution, mean of the sampling distribution of mean see the statement properly. What is sampling distribution of means? Sampling distribution means, means I am talking about the probability distribution. So that means the mean of the call the distribution, so what is that that is my X bar. So what is X bar? X bar = μ, that means what?

On an average this implies that, the sample mean, X bar is what mean means on an average, the sample mean is the same as the population mean. This is one so suppose I do not know the population mean, I took a sample from the sample, whatever mean I get, then I can say on an average population mean will be the sample mean only, is not it? So we therefore say that sample mean is an unbiased estimate of the population mean.

We will be seeing many applications from this application, the concept will be more and more clear. Now, even if you have some doubt, I am sure things will get clearer when we do the problem. So we therefore said a sample mean is an unbiased estimate of the population mean. Now, what is the variance of the distribution? See here, here, when I am talking about a variance; what I am talking about a variance?

Let me take this again, say this here, when I am talking about the variance, what is this? Variance among this X 1 bar X 2 bar X 3 bar variance among these how these values differ? When I am talking the variance of the sampling, distribution mind it, I am talking about how when you talk a variance of any number means what there are a few numbers X 1 X 2 X 3 X 4 when I talk of variance means how these numbers are varied, is not it?

So this is when I am talking about the variance of the sampling distribution, I am basically talking about a variance among the mean among the different means of the sample. So, the variance of the distribution of the sample mean is σ^2 / n that we have seen, so now that means, if the variance is σ^2 / n, so, what is the standard deviation? Standard deviation is $\sigma / \sqrt{n}$, this standard division is also called a standard error of the mean why it is called the standard error of the mean?

Means, this is the difference among different means of different samples, we suppose, we took 10 means, so, we found a 10 samples from 10 samples we got 10 means, so, the difference in this means, this is nothing we are calling it as a standard error of the mean. Now, if the population in the population σ is very high if my population is a very varied population variance is very high.

Variance is very high means standard deviation also be very high σ^2 / n σ will also be up in my population. In my population and variance is very high then what happens my when my sampling distribution what is the variance of the sampling distribution? Variance of the

sampling distribution is σ^2 / n variance of the population is high means my variance of the sampling mean will also be high is not it?

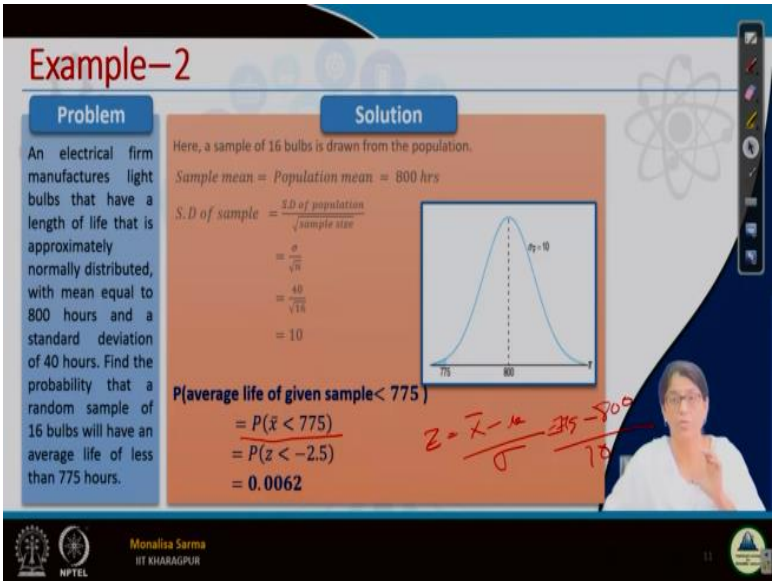
So, that means, 1 population sample mean suppose I maybe I got 5 another I got a 20 and other I got a 25 another I got a 41 I got an 80 is not it? If my population variance is high, my the variance of the distribution of sample means will be high. In such case, when my variance among the different sample mean is high, that means, this statement on an average the sample mean is same as the population mean this statement is not a very reliable statement.

That means, if my σ is higher than the sample mean are not a reliable estimate of the population mean. If the population variance is very high, suppose I am talking about the lifetime of the bulb, 1 bulb say may run 10 hours another bulb say more than 200 hours, another bulb say may run 100 hours, another bulb say run 500 hours. So, it is very much varied figure. In that case, when I take it from sample, from the sample, whatever I take the mean, that sample mean is not a reliable estimate of the population mean.

That is why I am telling when σ is higher than the sample mean are not reliable. However, for a very large sample size, because see and σ^2 / n, if n is the denominator, it is more than my sampling variance will also be less, is not it? So for the large sample size my standard error tends to 0, but if I take n is very large, then what happens, my difference becomes very, very less difference is what?

Variance of the distribution means the difference among the means that becomes that comes very less. So that is why my first statement on this central limit theorem being exactly true if n goes to ∞, that this theorem is being exactly true only if n goes to ∞. If n goes to ∞, that what happens? Sampling the variance among the sample mean is it tends to 0 then on an average sample mean is equals to population means that is a very reliable estimate.

(Refer Slide Time: 16:19)



So now we will just do 1 problem, definitely, we will be doing more problems in our tutorial class also, just 1 problem so as you understand the things a bit so see the problem. An electrical firm manufactures light bulbs that have a length of life that is approximately normally distributed with mean equals to 800 hours and a standard deviation of 40 hours. So what it is given an electrical firm, it manufactures light bulbs.

So and its life, it is normally distributed. It is this firm, it is claiming that the life of the bulbs is normally distributed. And it is also claiming that its mean is 800. That means my μ is 800 and standard deviation that is my σ is 40 hours. This farm is claiming mind it. So suppose I am the owner of the farm, I am claiming that the mean life is 800. And the standard deviation is 40 hours, what I am claiming may not be true.

Because I am just it may be an educated guess, basically, I am predicting something out of some past experience or whatever it is, but it is not possible for me to tell that is exactly true. Why? Because how can I find out the life of all the bulbs? I did not check off all the balls, how can I tell? So this is just a manufacturer that is claiming. Now find the probability if that is true assuming whatever he is telling is true.

Find the probability that a random sample of 16 bulbs will have an average life of less than 775 hours. So look the question other way around. I have gone to purchase some electric bulk from a manufacturer and the manufacturer has told me this my life is normally distributed. I am educated person I know about distribution and all. So, that is why he talked in terms of technical terms, he told me this, life for the bulb is normally distributed.