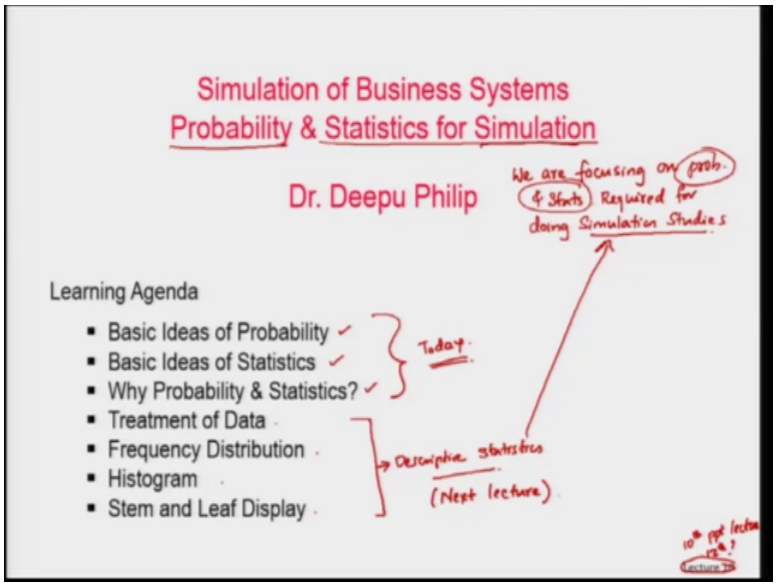
Simulation of Business Systems
Prof. Deepu Philip
Department of Industrial & Management Engineering
Indian Institute of Technology, Kanpur

Lecture - 13
Probability and Statistics for Simulation

Good morning students. Welcome to yet another lecture of Simulation of Business Systems Course.

(Refer Slide Time: 00:20)



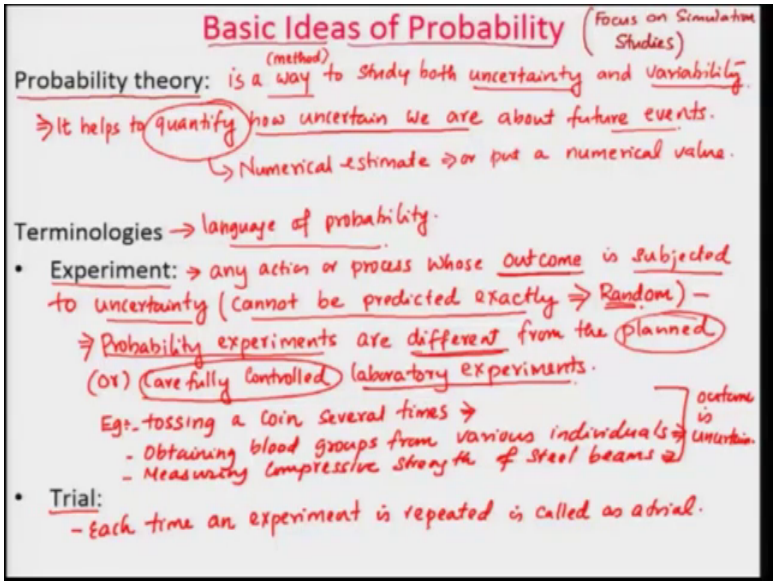
I am Dr. Deepu Philip. I am from IIT, Kanpur. And this course is you have already seen quite a lot of things about the use of Simulation; why it is necessary for studying business systems; how do you model open ended systems and what are different simulation programs available, software's available. We already seen lit bit of arena and other things and today's lecture the number written here is lecture 10.

This means it is the 10th PPT lecture ok; the power point lecture. The actual would be 10th or 12th or something like that. I think I am not sure whether it is 12th or not. But this is the 10th power point based lecture and today, what we are going to study is Probability and Statistics for Simulation. You might have studied Probability and Statistics as part of your other courses, but here the focus is on Simulation or what we are going to see?

We are focusing on probability and Stats required for doing Simulation studies ok. Our focus is how to do simulation studies and what are the probability and statistics tools that are required to do various simulation studies for us ok. And our basic learning agenda includes the Basic Ideas of Probability; Basic Ideas of Statistics and Why Probability and Statistics?

So, today we will try to cover these 3 ok, in today's lecture. Then, we will talk about the Treatment of Data; Frequency Distribution; Histogram; Stem and Leaf Display which are some data. So, then these typically comes under the Descriptive Statistics aspects which are required for us to sometimes do simulation studies. So, that part we will see in we will say next lecture ok. So, our focus is these 3 things today right ok.

(Refer Slide Time: 02:54)



So, first we will get into what we call as the basic ideas of probability as we say ok; what is probability and again, as I said focus on simulation studies or we are discussing here is gear towards conducting simulation studies, we are not focusing much on doing general probability and statistics.

So, the probability theory, people here about this name probability theory and sometimes people talk about this also as a important thing. We change this colour and the reason this probability theory and people mention about this because probability theory is a way ok, it is a way to study it is a way or we can think about it as a method; it is a way or a

method to study both uncertainty and variability ok; these 2 things uncertainty and variability ok.

So, probability theory provides you a way or a method to study both uncertainty and variability; implies it helps to quantify how uncertain we are ok; how uncertain we are about future events ok. This quantification ok; quantify in a sense we say numerical estimate or put a numerical value or put a numerical value ok.

So, it allows you to quantify how uncertain are we about future events ok. We are interested in what will happen in future and how uncertain are we or how certain are we. So, to understand the concept of probability and a get into the more details of it, there are certain terminologies ok. So, this is the language of probability ok.

So, to speak the language of probability we need to learn what are the major terminologies associated with the probability. So, the first terminology that we learn today is called the term called experiment ok. We have already seen this term, already I mentioned this time multiple times in the class called as simulation experiment. But the probability experiment is also similar idea. Let us see what it actually means.

Experiment we define it as any action or process any action or process, whose outcome whose outcome is subject or subjected to uncertainty ok. So, a process whose outcome is subjected; whose outcome is subjected to uncertainty, which means it cannot be predicted. It cannot be predicted exactly ok. The outcome is subjected to uncertainty means you cannot predict it exactly; this implies it is random ok.

So, you how no control over how it is going to happen ok. So, any action or process whose outcome is subjected to uncertainty, which means, it cannot be predicted exactly or it's a random one ok. So, this is what we call as a experiment. Probability experiments, experiments are different from the planned or carefully controlled laboratory experiments laboratory experiments.

So, what we are saying here is the laboratory experiments that we are typically used to, they are different. So, the probability experiments are different. They are different from the laboratory experiments because laboratory experiments are well planned and they are carefully controlled ok. Whereas, a probabilistic experiment is outcome is subjected to uncertainty.

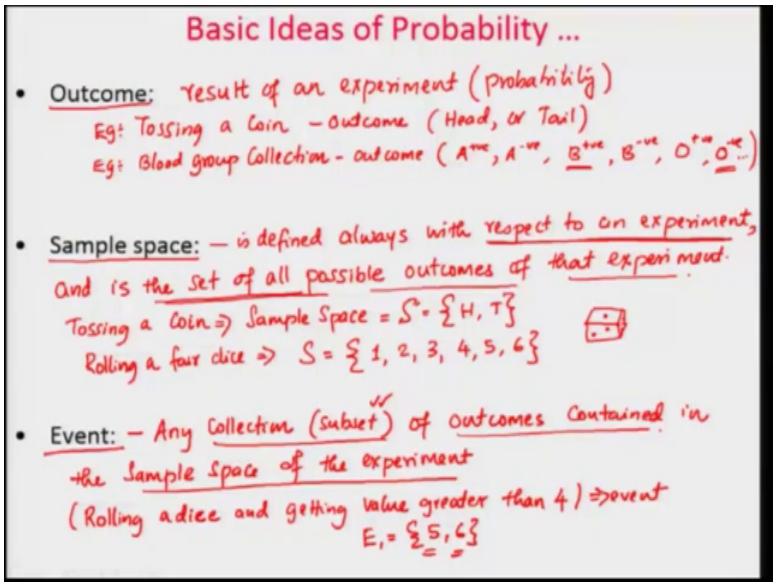
There is no assurity of what you are going to get. Because of this uncertainty or what we call as the randomness is the reason why we basically say that the probability experiment is different compared from the laboratory experiments. Some examples of this are tossing a coin several times. Everybody toss about coin several times and observe whether you get a head or a tail is one example ok.

Another example would be obtaining blood groups; obtaining blood groups from various individuals ok. So, somebody you find people individually and you ask their blood group or take blood from them and find out whether they are A positive, B negative, AB positive etcetera like that ok. Another example of this would be measuring compressive strength, compressive strength of steel beams ok.

So, you take different type of steel beams and using and UTM or universal testing machine try to measure the compressive strength of the steel beams. So, then you record those values. So, all these things, so you are not sure what would be the outcome; what would be the compressive strength of the steel beam in front of you. You are not we cannot be predicted exactly. You cannot predict exactly what will be the blood group of the individual. You cannot predict exactly whether the coin the next toss will be head or a tail.

So, all these things, the outcome cannot be predicted. So, the outcome is uncertain. You cannot predict it ok. Then, we get into the next word which is called as a trial ok. So, we already gone through experiment. Now, we get into Trial ok. Trial means each time an experiment is repeated is called as a Trial. So, when you toss a coin that is a trial; if you toss it 10 times, there 10 trials ok.

(Refer Slide Time: 11:03)



Then, the second terms out of this is the next set of terms is outcome. I already used this term before without telling you what it is ok. The outcome is used right here ok. So, the word outcome is in a simplistic sense, it is the result of an experiment and this experiment is the probability experiment not the laboratory experiment ok.

So, for example, is tossing a coin the outcomes will be head or tail. If the coin is a proper coin, then you either get a head or a tail ok. If you do the blood group collection, blood group collection; then, we could get outcomes as A positive, A negative, B positive, B negative, O positive, O negative like this ok. This will be the outcome.

So, you find a person randomly find a person, you take the blood and you check what is the blood what is the blood of that person and you say he is B positive, then the you do same thing for the next person and you get it as O negative and stuff like this. So, the result of an experiment is the probabilistic experiment is what we call as an outcome ok.

Then, we talk about the term called a Sample space and Sample space is defined always with respect to an experiment, to an experiment; always with respect to an experiment and is the set of all possible outcomes of that experiment ok.

So, a sample space can only be defined with respect to a specific experiment ok. Each experiment will have its own sample space and what it is? It is the set; it is a collection of all possible outcomes that are pertaining to that experiment. So, since it is a set, if we say

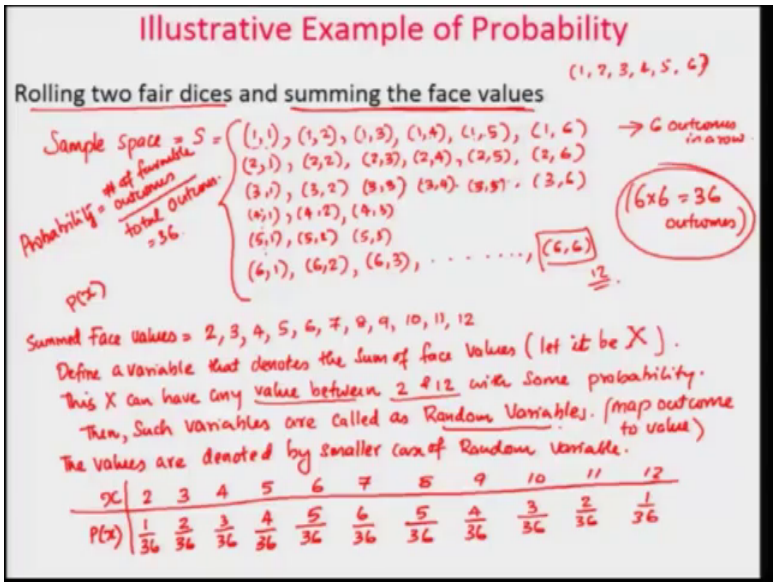
tossing a coin; then, sample space typically denoted by capital S uppercase S is a set. So, we denote it as set H and T which means head and tail; we denote a heads by H and tails by T ok.

Similarly, if you talk about rolling a die rolling a fair die or dice whatever you want to call it. The sample space is equal to the face values of 1, 2, 3, 4, 5 and 6. So, its 6 faced die and you will be able to get. So, what we are talking about here is having a die like this ok, somewhat like this and you have stuff like ok. So, you are rolling this and you get to see any one of those face and those are the face values you are writing here ok.

So, now, the event is another term that is important in the case of probability. An event is a it is a any collection any collection or what we call as a subset of outcomes; outcomes contained in the sample space of the experiment ok. So, what we are saying here is it is any collection or the right word here is subset; any subset of outcomes that are related to the sample space of a experiment.

So, if you talk about rolling a die ok; rolling a die dice and getting value greater than 4 ok. So, this is an event. This becomes an Event, which means the event E 1 is equal to set of greater than 4 will be 5 and 6. So, which is equivalent to getting the value face value of 5 and 6 ok, so, that is what we call as an event.

(Refer Slide Time: 16:34)



So, let us do a bigger example in this regard and it is little bit of a complicated example. Rolling 2 fair dice and summing the face values. So, you have 2 dice; each die has the faces 1, 2, 3, 4, 5 and 6 and what you trying to do is you rolling 2 of such dice together and you are summing the face values ok.

So, the sample space in this regard let us see what is the sample space ok, is denoted by S ok. So, the first die can get a face value of 1, second die can also get a face value of 1. That is be 1 roll ok. The other case is first die can get a face value of 1; second can get a face value of 2. So, this is 1 and 3. The first die getting 1; second die getting 4. First die getting 1; second die getting 5. First die getting 1; second die getting 6 ok. This is one case.

Now let us look at the second scenario, where the second die first die getting 2; second die getting 1. First die getting 2; second die getting 2. First die getting 2; second die getting 3. First die getting 2; second die getting 4. First die getting 2; second die getting 5. First die getting 2; second die getting 6.

Now, I can do the same way 3 1, 3 2 like this all the way to 3 6 and I can go all the way down to 6. First die getting 6; second die getting 1; 6 1, 6 2, 6 3 like this all the way to 6 and 6 both die giving you the values of 6. So, if you look into this, you have 1 2 3 4 5 6. So, there are 6 outcomes in a row ok. Same way you have 1 2 3 4 5 6 ok. There are 6 columns or 6 rows also ok. So, these are 6 columns and 6 rows. So, you have 6 times 6 equal to 36 outcomes in this case and these 36 outcomes what are you doing? You are summing the face values. So, what are those face values ok?

So, you get 2 1's. So, you sum them you get 1 and 1 as 2. So, the face values, the minimum value you can get is basically 2 right; 1 and 1. The second value you can get is 3 which is 1 and 2 and 2 and 1 can also give you 3. Same way, then you have is 4, Then, you have is 5, 6, 7, 8, 9, 10, 11 and 12 which is 6 plus 6 is 12. So, these are the face values, the summed the face values that you can get ok. So, if I define a variable ok. So, define a variable; variable that denotes the sum of face values ok. Let it be X; capital X, I am just putting X ok. So, then this X can have this X can have any value between 2 and 12 with some probability ok.

And so that then such variables are called as random variables. So, I have been using this time random variable early, but now you guys know what a random variable is random

variable is a variable that maps the outcome of a probability experiment to some values ok; mapping probability experiment outcome to a value ok; map outcome to value. So, then I can say it as the values are denoted by smaller case of random variable ok. So, I say that the x, the smaller value of x denotes the random variable.

So, the random variable values here are 2 3 5 6 7 8 9 10 11 12 ok. So, these little x or the smaller case x denotes the values the random variable can take and the random variable is x which maps the outcome; outcome is rolling two fair dice and summing the values to a numerical value, a real number in this regard. So, then how do you find the probability? So, probability is defined as probability can be defined as number of favourable outcomes by total outcomes divided by total outcomes. So, we know that the total outcomes is 36 ok, this is total 36.

Now, the favourable outcomes we have a numerate. So, the probability of x is typically denoted by p of x ok. So, if I say p of x then that means, probability of x with this x equal to 2; that means, what is the probability that x will take a value of 2 which means the sum of the face value is equal to 2 and there is only one way you can get it which is 1 and 1. So, it is 1 divided by 36, the 3, the sum is to be 3 can be 1 and 2 and 2 and 1. So, there are 2 favourable outcomes out of 36. So, it is 2 over 36. The 4 is 1 and 3, 3 and 1. So, 1 plus 3 is 4; 3 plus 1 is 4. Three is one more which is 2 plus 2 ok. So, there is 3 possible outcomes. So, it becomes 3 out of 36.

Now, 5 will be 1 plus 4 is 5 ok. So, 4 plus 1 is also 5. So, there will be somewhere here a not there actually. There will be 4 1 right here ok. So, that will be 1 4 and 4 1. So, that is 2. Then there is will be we have to get 5. So, will be 2 3 and 3 2 that is 2 ok. So, 2 more. So, it will be 2 plus 2 4. So, 4 by 36; 6 will be, how will how many there will be 6? There will be 1 plus 5, 6 ok. Then there is a 5 1 which is another one. So, that is 2. Then, there is a 2 4 which will be 6 and there will be a 4 2 as well here. So, another 6. So, there are 4 outcomes already and there is one more which is your 3 and 3 ok. 3 3 will also give you 6. So, there are 5 outcomes out of 36 which will give you the sum of 6 ok.

Now, how many will be there for 7. 7 will be 1 6, there will be 7; 6 1 that is 2 ok. Then, what is the other way to calculate 2 5, will be 7. Then, you also have 5 2 also other one that is also 7. Then, what is other way? There is a three 4 7, then there is a 4 2 which will

also give you 7. So, that is 6 right. So, all together you have 6 possible outcomes out of 36.

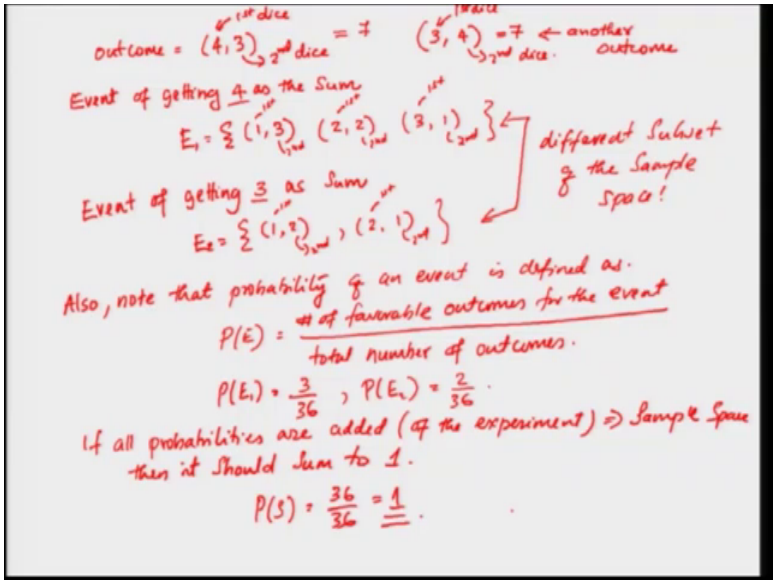
So, then 8 what happens is 6 plus 1 only will give you 7. So, you cannot have anything out of 1. The only other way to get 8 is 2 and 6. 6 plus 2 is 8. Then there is a so, this is 1, this is 2. Then the other option is to get 8 is 4 and 4 is 8 ok. There is another one which is 3, 4 plus 3, 7. 5 and 3 there is a 5 3 here; sorry this is 3 5 actually speaking, 3 5 and there is a 5 3 here. So, that is 4 and then, there is a 4 and a 4.

So, there all together you have 5 out of 36 outcomes ok. Same way, if you look at 9, you will only get 6 and 3 and this numbers will now start to reduce; it will be 4 out of 36. Then, 10 will be 3 out of 36. 11 will be 2 out of 36 and 12 will be 1 out of 36 as and this is basically just this one option. This will be 12 ok. So, 1 out of 36 will be the option.

So, this whole thing that we talk about is the probability of get rolling a die and getting and summing the face values and what is the probability of seeing this sum must 12, it is 1 out of 36 that will be the probability. If you keep on rolling this for the long time, then you will actually see one out of 36 times, the number the sum value of 12 will show up ok.

Now, if I say that ok; so, the sample space we already see what it is?

(Refer Slide Time: 27:24)



Now, let us talk about what is an outcome ok? So, one outcome of this is let say 4 3; that means, this is the first die and this is the second die ok. So, the sum will come to 7 ok. So, this is an outcome in this regard. Same way 3 4 which will also give you 7 is this is another outcome that is because this is the first die and this is the second die right now. So, these are the outcomes in this regard and if I say the event ok; event of getting 4 as the sum; so, what is that E 1 that E 1 is equal to I can get it as 1 3 will give me the first sum, this is the first die this is second. Then, other option is 2 and 2, first second; then, I have 3 1, this is first this is the second ok.

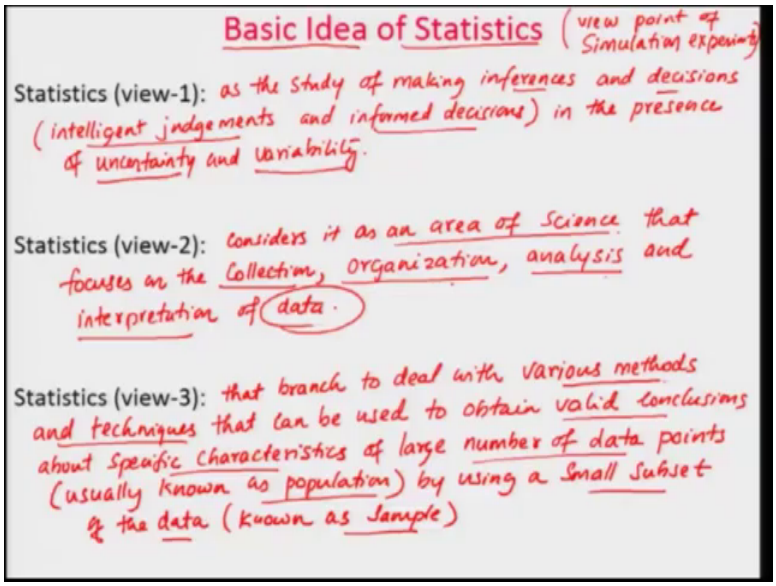
So, this is the event; in this case where I getting 4 as a sum, I have 3 possible outcomes which will give me the sum as the 4 alright. So, you have seen the outcome the event and all those kind of things. If I say at some point of time event of getting 3 as sum, then that will be E 2 will be equal to 3 will be 1 and 2. This is first die. This is the second die and 2 and 1 first and second ok. So, there are 2 possible outcomes out of this. So, that is the summing the die example. So, we can see that depending upon how these sum values change, you get different subsets or different subset of the sample space ok.

So, if you list all the events possible, then it is also another way to find the probability ok. Also note that probability of an event of an event is defined as P of E is equal to number of favourable outcomes for the event divided by total number of outcomes ok. So, in this case probability of E 1 will be equal to 3 out of 36. There are 3 favourable outcomes, total 36. Probability of E 2 will be 2 out of 36 like this ok.

Also if you look into this case ok, if you add this ok, this is all the probabilities if you add ok. So, if all probabilities are added, of the experiment or another way is sample space ok, if you do that then, it should some to 1 ok.

So, an example is if you look in here, we can see that 1 out of the 36, 2 out of 36 like this. If we sum this up this is 1 plus 2 is 3; 3 plus 3 is 6; 6 plus 4 is 10; 10 plus 5 is 15; 15 plus 6 is 21; 21 plus 5 is 26; 26 plus 4 is 30; 30 plus 3 is 33; 33 plus 2 is 35 plus 1 is 36. So, the sum will come to probability will be coming to 36 over 36. So, in this case it will be probability of sample space will be 36 over 36 which will be equal to 1. So, the sum of the sample space, the probability of the sample space will always be sum to equal to 1 ok. You should also remember that.

(Refer Slide Time: 32:48)



Now, that I have explain to you the basic idea of probability. Now, let us look in the basic idea of statistics again from my viewpoint of simulation experiments alright. So, in the basic idea of statistics, there are 3 views of statistics that we need to take a look into in this ok. The first view is it is statistics as a study as the study of making inferences; making inferences and decisions. So, I would call in the business terms as intelligent judgements; intelligent judgements and informed decisions ok. Intelligent judgements and informed decisions ok.

So, as the study of making inferences and decisions, intelligent judgements and informed decisions in the presence in the presence of uncertainty and variability. So, statistics one view is that when you have uncertainty and variability; how do you make intelligent judgements and informed decisions ok. Typically people also referred to the mass inferences and decisions which are in nothing other than intelligent judgements and informed decisions. So, the uncertainty and variability in the phase of that how do you do that or the how do you make intelligent judgements and informed decisions is to some people is a view point of statistics.

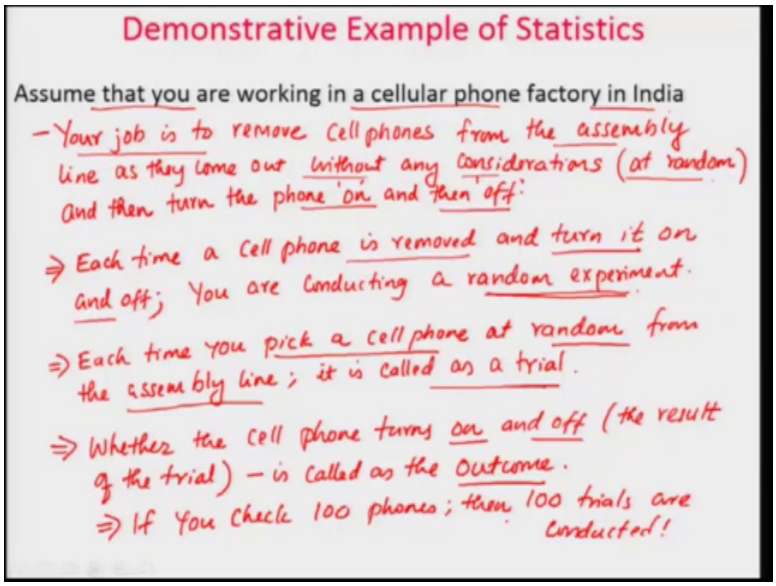
The second viewpoint of statistics considers it as considers it as an area of science. Area of science, considers it as an area of science that focuses on the collection organization analysis and interpretation of data ok. The other view point considers this as an area of science considers it as an area of science that area particularly focuses on the collection,

organization, analysis and interpretation of data ok. So, things that are pertaining to data, how to do collect data; how to organize data; how to analyze data; how to interpret data, these are also this is an that part that is branch of science that does all of these things is considered as statistics by another section of people.

Then, there is a third popular view point which is the statistics is the you know that branch to deal with branch to deal with or to deal with various methods and techniques, techniques that branch that deals with various methods and techniques that can be used that can be used to obtain to obtain valid conclusions, valid conclusions about specific characteristics, specific characteristics of large number of data points ok. These large number of data points usually known as population, by using a small subset of the data the data known as Sample ok.

So, it is that branch that has various methods and techniques that can be used to obtain valid conclusions and about specific characteristics of large number of data points typically known as population by using a small subset of data known as sample. So, it is that branch that has methods and techniques to obtain valid conclusions about a population using a sample. So, some people consider that as also a statistics ok.

(Refer Slide Time: 38:57)



So, let us look at a demonstrative example of statistics in. So, maybe with an example we might be able to understand some aspects of a statistics. Assume that you are working in a cellular phone factory in India ok. You are working in a factory and you have some job

ok. So, your job your job is to remove is to remove cell phones, is to remove cell phones from the assembly line.

You are removing cell phones from the assembly line as they come out, as they come out without any consideration or any considerations plural or I use word at random ok; no predispositions, just pick it at random and then, then turn the phone on and then off ok.

So, first you go to the assembly line without looking into without having any consideration, pick a phone and then turn it on and then turn it off. So, that is your job description. Your job is to remove cell phones from the assembly line as they come out without any consideration or you do this at random and then, turn the phone on and then off ok, that is what you are doing; that is your job description ok.

So, then if that is the case each time a cell phone, each time a cell phone is each time a cell phone is removed; each time a cell phone is removed and turn it on and off ok, you are conducting you are conducting a random experiment. So, this is another way of looking at how the probability comes into the statistics part.

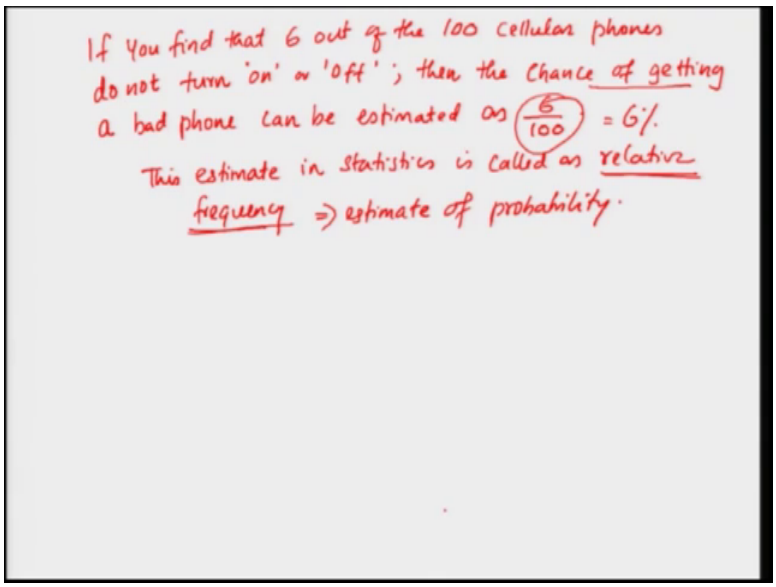
So, what you doing is you randomly picking a cell phone and you are turning it on and off. You are not sure whether the phone will turn on or turn off. So, in that regard if you look into this when you randomly pick the phone, you remove it from the assembly line, turn it on turn off; what you are actually doing is you are doing an experiment whose outcome is not a sure to you ok.

So, that is how you are conducting a random experiment. Now, each time, each time you pick a phone, pick a cell phone at random from the assembly line, from the assembly line; each time you pick a cell phone at random from the assembly line it is called as a trial. So, when somebody says it is a you pick a cell phone, when you pick a cell phone at random from the assembly line. It is called as a Trial; you randomly do the trial ok.

Now, the next part is whether the cell phone turns on and off which is the result of the trial. So, the trial is picking the phone at random and turn from the assembly line and turn it on and off ok. So, whether the phone will turn on and off, the result of the trial is called as the outcome.

Now, we can see how the Trial. How the outcome which are probabilistic ideas are coming into what we call as the statistics, where we are doing the data collection ok. So, an example is if you check 100 phones, then 100 trials are conducted. So, think about it that way, you pick 100 phones turn them on and off, then that implies you have conducted 100 trials ok.

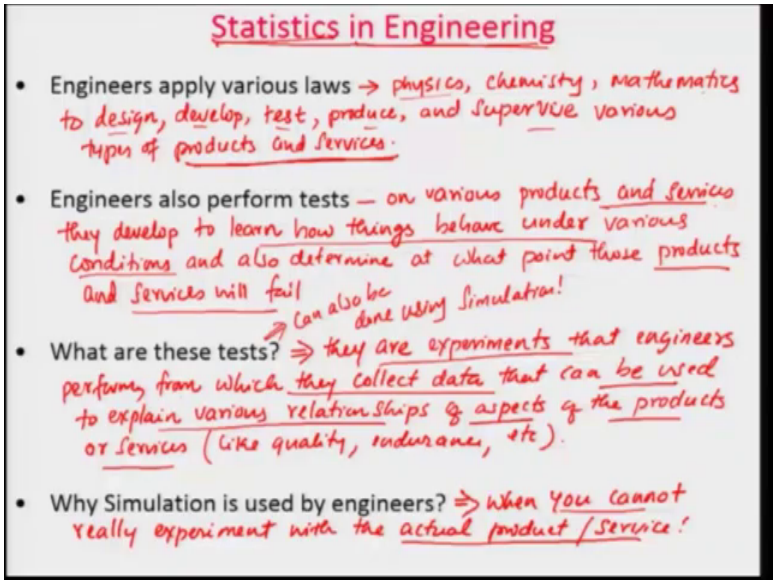
(Refer Slide Time: 44:07)



So, if you find that 6 out of the 100 cellular phones, do not do not turn on or off; then, the chance of getting a bad phone can be estimated by can be estimated as 6 by 100 or this estimate in statistics is called as relative frequency. So, in a way this chance of getting the bad phone which is 6 out of 100 or 6 percent ok, this is also known as the relative frequency in statistics. So, in a way it can be an estimate of probability ok. So, the estimate of probability can also be derived through statistics which is called as the relative frequency in this ok.

So, hopefully, now you understand how the concept of experiments, outcomes, trials, relative frequency and chance estimates all ties up together. So, what you are trying to do is if you do this experiment 1000 times and you got 67 phones, then the probability become 6.7 percent and stuff like that and more the experiments you do, then you get a different estimates out of it.

(Refer Slide Time: 46:29)



Finally, why is statistics necessary for engineering ok? This is important for us because most of us most of the people in this course that are engineers and people usually ask why statistics is important to this. So, the first thing is engineers applies various laws ok; laws of Physics, Chemistry, Mathematics. They apply various laws of Physics, Chemistry and Mathematics to design, develop, test, produce and supervise various types of products and services ok. So, they up they use various laws of Physics, Chemistry and Mathematics to design develop, test, produce and supervise various products and services. So, they are trying to they are making they are applying this laws to make something a product or a service.

So, they when they are applying these laws, then they need to have a mechanism to apply; they should have a way to do this; so, they should know what will be the outcome and if they apply the law. So, they should be able to be associate probabilistic laws along with this engineers also performs various test ok. This test on various products and services, they develop to learn how things behave under various conditions, conditions and also determine at what point those products, those products and services will fail.

So, they conduct various tests on products and services to learn how things behave under various conditions ok, for the operation conditions of the product and service and more than that they want to know when the products and services will fail; what are the conditions at which they will fail. So, they do test for this.

So, what are these tests ok? They are experiments; see how statistics and probability came into picture. These are experiments that engineers perform, engineers perform from which from which they collect data collect data that can be used to explain, that can be used to explain various relationships, various relationships of aspects of the products or services products or services ok.

So, the aspects like quality, endurance etcetera. So, what you saying is that these experiments, the engineers conduct these experiments to obtain data or they want to collect data and they want to use this data to explain various relationships of the product and service, so, relationship between various aspects of the product and service ok. So, that is the reason why engineers conduct these test.

Then, why simulation is used by engineers? It is when you cannot really experiment with the actual product or service. So, there are scenarios where you will not be able to actually conduct the experiment with the actual product or service. Then, that is the time you use simulation to do same type of test or experiments ok, so, can also be done using simulation ok. So, that is the idea of this aspect ok.

With this today, we reached the conclusion of our basic probability and statistics aspect of this course and I hope that you guys have an understanding of what is basic probability and what is statistics and how do engineers use both the concept of probability and statistics to make informed decisions about things that are uncertain and has variability in the future.

Thank you very much.