Transcriber's Name Maria Priya

Applied Econometrics

Prof. Tutan Ahmed

Vinod Gupta School of Management

Indian Institute of Technology, Kharagpur

Module No. # 02

Lecture No. # 20

Bernoulli Distribution

Hello, and welcome back to the lecture on Applied Econometrics. uh We have been talking about probability distribution. And in this lecture, we are going to see uh one such probability distribution which is famously known as Bernoulli distribution. (refer time: 00:37) This is a disc discrete random; this is a dis distribution that comes from a discrete random variable and it has a very specific properties. And we will see the Bernoulli distribution to be a foundational, a building block for many other distributions going forward.

uh So, we remember uh uh a discrete random variable, for example, tossing a coin; we will always get either a head or a tail, or success or a failure. So, this way, uh Bernoulli distribution is always representing, say, 1 trial; it is always a distribution that comes from 1 trial, 1 trial. It always represents the outcome of the trial as success and failure, like in coin toss, it can be head or tail, you can either pass in the exam or you fail, or you can either, you know, win a lottery or you will not win the lottery.

So, these kinds of examples where you have 2 possible outcomes, 2 possible outcomes, and where the outcomes are actually mutually exclusive, they are mutually exclusive, so, that represents a Bernoulli distribution, Bernoulli distribution. And here, another important criteria for Bernoulli distribution that we need to keep in mind is that the probability, the probability of success, you know; I so also should write that the outcomes are represented as success and failure.

So, the probability of success, the third criteria for Bernoulli distribution, the probability of success is always going to be constant. So, that is a very, very important criteria that we need to keep in mind that the probability, the p or for that matter 1 minus p; essentially, if p is constant, then 1 minus p is also constant. p is constant; the probability will not change. Okay. So, that is what constitute a Bernoulli trial. Now, the way we talk about success and failure, the way we assign the value of (pro) p will exceeds ma matter of convention, but still we we, it is good to have a grip on that.

So, let us say I write (refer time: 02:51) success, failure. Now, which one I would attribute as success? which one I will attribute as failure? Now, it totally depends on you, how you want to attribute. But in general, like let us say, I say that head is a success and tail is a failure. Okay. You can very well write tail is a success, head is a failure. How do I assign values to this variable? So, u usually, I, my random variable X, it can take 2 values as I said. So, it is 0 or 1. So, you can uh usually represent success as 1, failure as 0.

And the probability value, the p, again the p is absolutely up to you, whether you want to represent success as 1 minus p or failure as uh p or whichever way. But usually we write success as p and failure as 1 minus p. Okay. It is just a matter of convention, you feel free to change it. Okay. uh Now, if I want to actually derive the PMF, probability mass function; that is what we do. When we want to talk about a distribution, we always talk in terms of the probability mass function. So, how we will, how we write that?

So, we know that probability mass function we write X is equal to x is equal to here, I know that it is only 1 trial, and I have 2 possible outcomes, either it is a success or it is a failure. So, if it is a success, I write it like this, and failure, I write it like 1 minus x. Okay. So, how did I get that? And if I see, if I actually put the values of x and p or or x and 1 minus x, so, we will actually get the value of probability X equal to x. So, now, X can take 2 values, 0 and 1. So, I can write probability X is equal to 1 is equal to p of p to the 1 1 minus p 1 minus 1. Okay.

And which means only, only p. All right. Now, if I write P X is equal to 0, so, I will have p power 0 1 minus p; now, x is equal to 0, which is 1 minus 0; so, it is going to give me 1 minus p. Okay. So, that is basically my probability of failure here. And this is the probability of success here. Okay. So, essentially, this uh this probability mass function here, that represents a Bernoulli trial or a Bernoulli distribution. Remember, when I talk about a Bernoulli distribution, I am talking about only 1 trial, okay, there is only 1 trial; I am not adding any term here, any term here; there is no term here, because it is just 1 trial. All right.

Now, what are the examples? What are the examples of Bernoulli trial? (refer time: 05:41) Let me actually give you some examples, and let me ask you what do you, what do you think, whether it is a, whether it qualifies to be a Bernoulli trial or not. Okay. So, that is, that is, let us think it together. Let us say 1 example. I say that, well, so, the stock market; I am actually looking at stock market. And let us say whether the stock market will go up or down, go up or down has equal probability; probability of stock market going up or down is equal to 0.5.

So, at any point of time, whether your stock market is going to go up or going to go down, that is basically, uh the probability is 0.5. Okay. So, uh in this case, do I have a Bernoulli trial? So, that is a, that is a question. Another another example, another question; let us say probability of winning a lottery, okay, winning a lottery. Let us say it is a constant value, let us say 0.0001. And let us say I ask you that, let us say probability number of people entering in bank every second or every minute, let us say, every minute, okay.

Now, the probability of that is not constant, because the number of people who can come in the bank in the morning and who can come in the bank during lunchtime; so, depending on what time it is, the number will actually vary. So, the probability is not constant, non-constant probability. All right. Okay. So, essentially, when I have a constant probability here or here, so, I have, if I just look at 1 trial, let us say I just look at 1 observation, 1 trial here, and I see that whether the market has gone up or down, so, that is going to give me a fixed uh probability.

And I am going to have 2 possible outcomes; either the stock can go up or it can go down; I have p constant. Okay. So, then it satisfies the criteria Bernoulli distribution. Similarly, winning a lottery; as long as this value is constant, so, I can draw al, I can buy a lo, a lottery ticket. And that is 1 trial and I have 2 possible outcomes, either I can win or I can lose and I have 2 possibilities. Basically or, you know, lose or (suc) and I can be I can be successful or I can fail, and the probability is constant. Okay.

So, then, that way I can have I can have it as a Bernoulli uh distribution. Now, in this case, since the probability is not constant, so, then we will not have a Bernoulli distribution. All right. Okay. So, going forward, couple of things uh I should uh uh tell you. One is that, now, there is something called Bernoulli process. (refer time: 08:40) Bernoulli process. So, when you do this trial repeatedly, when you do this trial repeatedly, you actually constitute a Bernoulli process. Okay.

So, you can have like, you know, say, you know, say, you you represent head and tail in a coin toss experiment and you keep on, you know, tossing the coin, and you can get something like this. So, this kind of represents a Bernoulli process. All right. Now, what you need to remember in case of Bernoulli process is that, the trials are independent, trials are independent. So, what it means? It means that, whatever happened in the previous trial, it does not influence the next trial, right.

So, for example, if I got a head or tail, it doesn't matter; I mean, if I get a head or tail in this trial, so, it is not going to influence what I am going to get in the next trial, right. And since, it does not take into account the past events or whatever happened previously, so, we call this process as a memory, memoryless process. Okay. It does not keep anything in the memory. So, that is a very good thing, that is a memoryless process. And, of course, so, essentially, that they are independent, they are memoryless, and of course they are discrete, right?

So, that we said at the very beginning that it is a discrete; so, it is not a continuous thing. Okay. So, we will see using this points, using this kind of uh criteria. So, we know by now what is the Bernoulli trial and what is the Bernoulli process. And we will derive something very interesting going forward uh about the properties of the random variable. Okay. So, that property is going to be uh really crucial going forward. So, we will just talk about it in a while.

But before that, before I get into that, let me actually explain uh what you have learned previously about the expectation and variance of a Bernoulli process, okay. So, we will just try to derive the expectation and variance of a Bernoulli process. (refer time:10:43) So, if I want to have the expectation, I know expectation of X is sum total of X probability of X, right? Now, I can write small x. Okay. So, here, in case of Bernoulli process, I have, I can have, so basically 2 values; 0.

Now, I can go back to the probability mass function; if I have 0, I have this, which is 1 minus p; 1 minus p plus I can have 1. And I can have from this probability mass function, if x is equal to 1, this value is going to be p. So, I am going to write 1 into p. Okay. So, that means the expectation of x is going to be simply p. All right. That is a result. And the variance, the variance as we know, we are going to have V X is equal to; we are going to use the formula that you have derived earlier; expectation of X square minus expectation of X of whole square. All right.

So, if I want to have expectation of X square; now, we will use our uh learnings that you had previously. So, the X square is going to be, so the value of x is equal to; so, basically what you do, 1 square into the corresponding value is going to be p; p plus 0 square into 1 minus p, okay, and minus here; your expectation of X is, we have already derived p; so, p of whole square. So, essentially, it means, p minus p square. Essentially, that means p into 1 minus p. Okay. I hope you got that.

So, this one, what I did, I uh simply used the value X square and the probability remains the same, right, for for X square. So, that way, I derived the variance. So, this is a expectation and variance for a Bernoulli trial. Okay. And this has a very important connotation, uh the way uh you see this uh expectation and variance. And the most important thing is, they are constant, both the expectation and variance, they are actually constant terms, all right, they are non-changing.

Now, so, distribution, when you actually draw several trials, you you do several trials, repeated trials, you are having a constant mean and a constant variance. Okay. So, that means, the distributions are actually same. Okay. So, that is, that is a very important criteria. And as I said previously that they are independent, now, think about it. Think about a distribution, think about a distribution where you have (refer time: 13:59) mean constant, variance constant, and independent.

So, essentially, I can say when the mean and variance constant, so, basically, all the distributions that I get, since their mean and va variance is constant, they are basically, I can write identical, identical, and this is independent. Okay. So, this kind of vary uh this kind of distribution, this kind of random variables are known as independent and identically distributed or IID in short, random variable. This is very crucial, IID random variable. We need to keep in mind this properties of IID random variable, and we are going to see time and again uh how we use this concept of IID random variable where the distributions are basically identical, okay, the random variable distribution is basically identical. Okay.

So, last thing, we will just do a simple problem. (refer time: 15:14) And let us say that the problem is, let us say, so, there are, let us say 10 students; or let us say uh there is a probability that; or let us say; okay, let me write down this way. Only 1 student among 10 students can solve Rubik's, Rubik's; how do I spell Rubick's Cube? I think there is no k; Rubik's cube. So, only 1 student, 1 among 10 students uh can solve Rubik's Cube. Okay. Now, if I; so, let us say, if I draw 1 trial, if I do 1 trial, if I draw 1 student, what is the probability that, that particular student will be able to solve the Rubik's uh, you know, Rubik's Cube?

Now, we can very well say that, okay, of course, like 10 students, so, you know, if if the probability of uh each one succeeding is just 1 by 10. uh But if we actually try to plot that or i if we actually use the probability mass function for a Bernoulli trial where we wrote p of X is equal to x is nothing but p x 1 minus p 1 minus x. And if I use the p, here the p is 1 by 10. So, the value is going to be 1 by 10. And here it is going to be 1 by 1 minus 10. And x you have drawn 1, so, 1 minus 1. And you are going to get 1 by 10. Right.

So, exactly the same, but this is uh the same result, but this is how we actually represent the problem uh as a Bernoulli trial. So, that is the purpose of giving this example. So, with this, uh we will, we have come to an end to the discussion on uh Bernoulli trial or the Bernoulli process or the Bernoulli distribution. So, uh going forward, we are going to see their implication of or the application of Bernoulli trial uh when we are going to talk about ba uh binomial distribution. So, with this, we end the lecture here. Thank you.