PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

Indian Institute of Technology Tirupati

Week - 04

Lecture - 19

Applications of Binomial Distribution

Now we will discuss some applications of binomial distribution. Many instances of binomial distribution can be found in real life. We have already discussed that this is related to random phenomena, success or failure. Whenever you see some random phenomena, the output is related to success or failure, which can be modeled using binomial distribution. For example, if a new drug is introduced to cure a disease, it either cures the disease, which is a success, or it does not cure the disease, which is a failure.

If you purchase a lottery ticket, you are either going to win money or you are not. Basically, anything you can think of that can only be a success or failure can be represented by a binomial distribution. Let us discuss some examples. We have already discussed binomial distribution in detail, so let us consider one problem. Suppose 10 coins are thrown simultaneously.

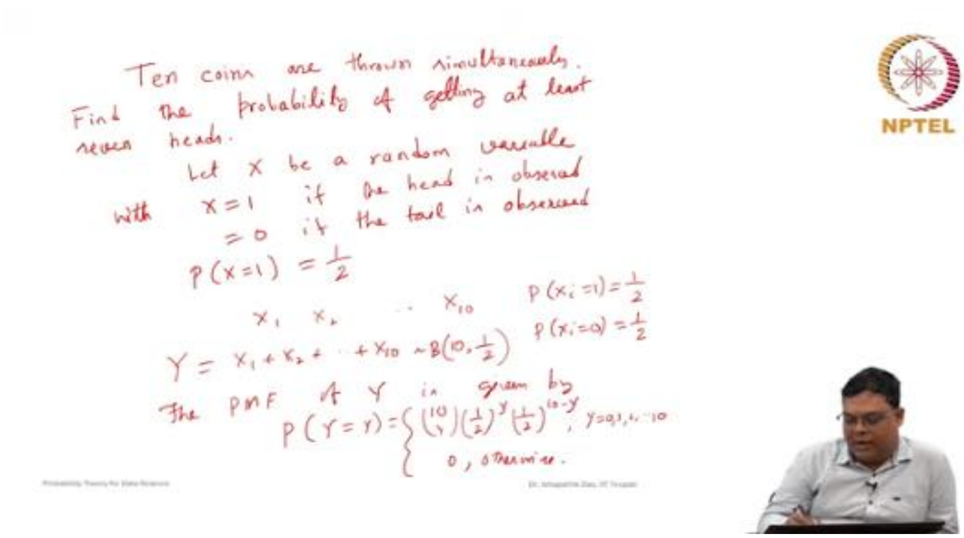
Find the probability of getting at least 7 heads. In this problem, you can see that the probability of heads for a particular coin when tossing a coin once is not given. Whenever it is not given, we will assume that it is an unbiased coin. So, let X be a random variable, where X = 1 if heads is observed and X = 0 if tails is observed. The probability of heads here will be considered as 1/2 because this is not mentioned.

So, whenever you see in a problem that it is not given, we can assume that the probability of heads is equal to 1/2. Now, 10 coins are thrown simultaneously. That means it is repeated 10 times. This is $X_1, X_2, ..., X_n$, all have the same probability mass function. This is why it is called identically distributed random variables $X_1, X_2, ..., X_n$. Here, n = 10, and the probability of $X_i = 0$ is 1 - p, which is also 1/2.

Now, Y = X1 + X2 + ... + X10, and it has a binomial distribution with parameters n = 10 and p = 1/2. You can remember that you can go back to the previous slides where we started discussing Bernoulli distribution. Y = X1, X2, ..., Xn, where X1, X2, ..., Xn are independent random variables, each having a Bernoulli distribution with probability p. Here, p = 1/2, and 1 - p = 1/2 again. If you are taking n number of sums, then it will be a binomial distribution with parameters n and p.

In this example, n = 10 and p = 1/2. So, what is the probability mass function? The probability mass function of Y is given by the probability that Y = y. This is (n choose y). You can remember the probability mass function of binomial distribution that we already discussed. Then p^y and (1 - p)^(n - y).

This is equal whenever y = 0, 1, 2, ..., 10, and this is equal to 0 otherwise. So now it can be simplified. The probability mass function of Y, P(Y), this is nothing but the probability that Y = y. It can be further simplified like this, because you can see that here (10 choose y), and since p = 1/2, that is why 1 - p is also coming to 1/2. So, this can be Y = 10 - Y.



This is (1/2) to the power of 10. It is coming out to be 10 choose Y. This is 10 choose Y multiplied by (1/2) to the power of 10. Whenever Y = 0, 1, 2, ..., 10, this is equal to 0 otherwise. Now, what is the question?

This is the distribution of the 10 coins whenever you are throwing them. It is a binomial distribution. We need to find the probability of getting at least 7 heads. So, this is actually asking what is the probability that $Y = Y$, which means what is the probability of $Y = 0$, which is 0 heads; $Y = 1$, which is 1 head; and $Y = 2$, which is 2 heads out of 10 thrown.

So here it is asked what is the probability of getting at least 7 heads. At least 7 heads means we want the probability of $Y \geq 7$.

This can be expressed as the probability that $Y = 7$ plus the probability that $Y = 8$ plus the probability that $Y = 9$ plus the probability that $Y = 10$. We just have to simplify it to find the final probability. For $Y = 7$, we will find from this expression. So this is nothing but (10 choose 7) × (1/2)^10 + (10 choose 8) × (1/2)^10 + (10 choose 9) × (1/2)^10 + (10 choose 10) × (1/2)^10

because this is the probability mass function. Now, (1/2)^10 is common. So, (10 choose 7) can be written as (10! / (7! × 3!)). Plus, (10 choose 8) can be written as (10! / (8! × 2!)). Plus, (10 choose 9), we know this is nothing but 10, and (10 choose 10) is nothing but 1. Now, we just need to simplify these terms. We can simplify 10!, 7!, and 3! as follows:

$(10! / (7! \times 3!)) = (10 \times 9 \times 8 \times 7!) / (7! \times 3 \times 2 \times 1)$

So, 7! cancels out, leaving us with $(10 \times 9 \times 8) / (3 \times 2 \times 1) = 120$.

So, we will replace that term, and if I have done any mistakes, please check and make the correct computation. So, this is nothing but (1/2)^10 × 120. Now, for the next term, (10 choose 8) can be found as follows:

$(10! / (8! \times 2!)) = (10 \times 9 \times 8!) / (8! \times 2)$

This cancels out to give us

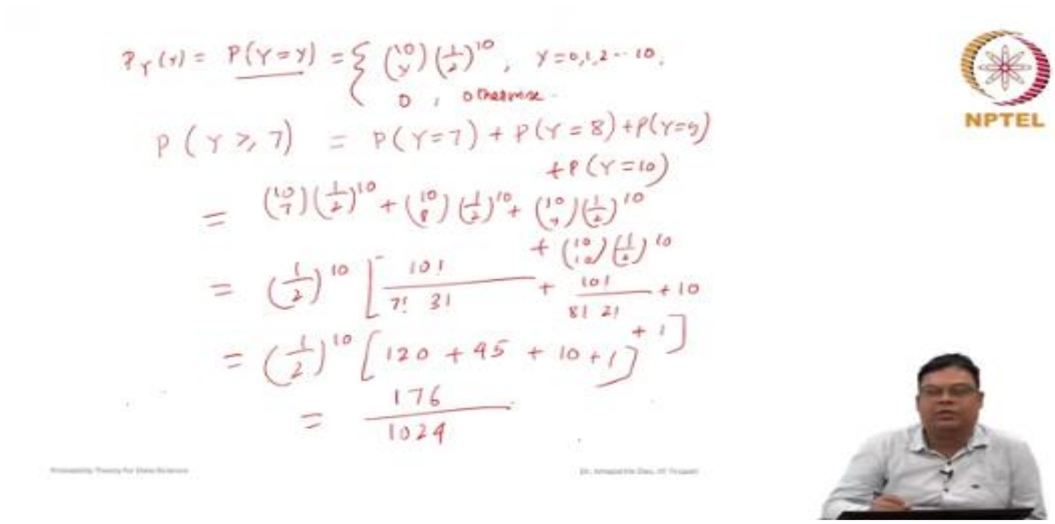
$(10 \times 9) / 2 = 45$.

So this is 45, and then we add the remaining terms: 10 plus 1. Now, just add those values together. This gives us (1/2)^10 × (120 + 45 + 10 + 1).

Now let’s check the values: (1/2)^10 = 1/1024, and the sum is 176.

Therefore, the final result is 176 / 1024.

So, this is the probability we are finding. The 10 coins are thrown simultaneously to find the probability of getting at least 7 heads. The probability of getting at least 7 heads is 176 / 1024.This is one problem.



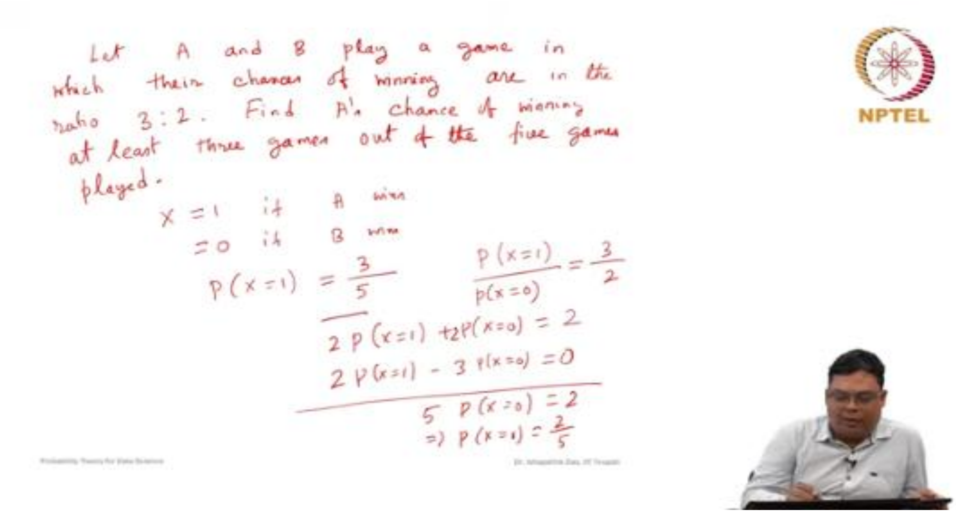
Now, let us do some other numerical examples. Let A and B play a game in which their chances of winning are in the ratio of 3 to 2. We want to find A's chance of winning at least 3 games out of the 5 games played. So, let’s discuss this problem again. Let A and B play a game where their chances of winning are in the ratio of 3 to 2.

We need to find A's chance of winning at least 3 games out of the 5 games. They are playing 5 times. Suppose we first discuss what happens if they play just once. Let X be the random variable that equals 1 if A wins. Then, the probability of X = 1 is equal to 0 if B wins. Since their chances of winning are given in the ratio of 3 to 2, this means if 5 games are played, the probability of X = 1 is the odds for A, which is 3 out of (3 + 2) = 3/5.

So, the probability of A winning, denoted as P(X = 1), is 3/5. You can also think of this ratio in terms of probabilities: P(X = 1) is the probability that A wins, and P(X = 0) is the probability that B wins. Thus, the ratio is 3 to 2. From this, we know that the sum of the probabilities, P(X = 1) + P(X = 0), equals 1. From here, we can also find that if P(X = 1) = 3/5, then P(X = 0) will be 2/5.

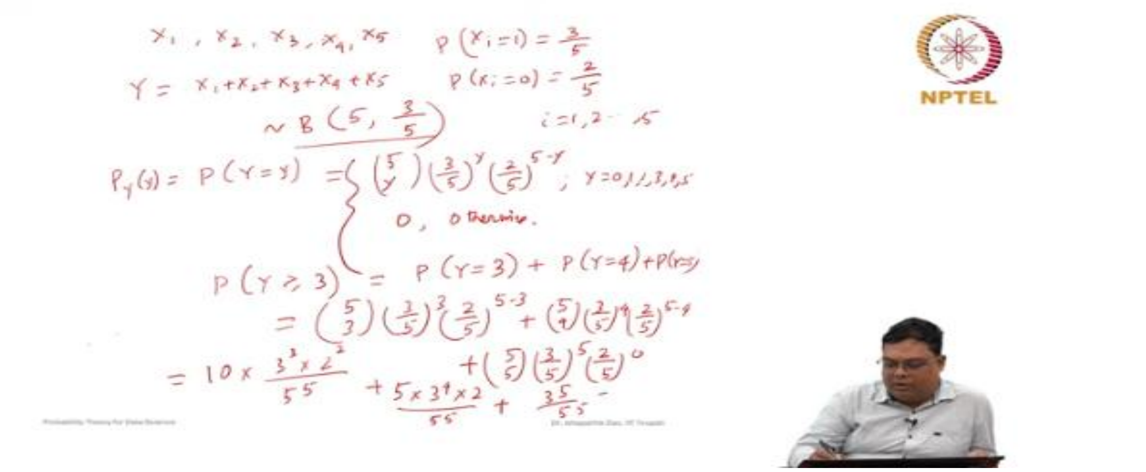
Now, if we set up the equations, we have 2 × P(X = 1) − 3 × P(X = 0) = 0. If we multiply by 2 and subtract, we find that P(X = 0) = 2/5. This implies that P(X = 1) = 1 − 2/5 = 3/5.

The probability that A wins when they play once is 3/5. Since they are playing 5 times, we can consider that X_1 represents whether A wins in the first game, X_2 in the second game, and so on, up to X_5.



Each random variable has the same probability, as they are identically distributed random variables with a probability of 3 out of 5 for A winning and 2 out of 5 for A not winning. If we consider the sum of these random variables, $X_1 + X_2 + X_3 + X_4 + X_5$, we have a distribution that follows a binomial distribution where n = 5 and p = 3/5. We want to find A's chance of winning at least 3 games out of the 5 played, which corresponds to the probability mass function of Y, representing the number of times A wins out of the 5 games. We need to find P(Y ≥ 3), which is equal to P(Y = 3) + P(Y = 4) + P(Y = 5). Calculating these probabilities, we find that 5C3 = 10, 5C4 = 5, and 5C5 = 1.

Substituting these values, we compute P(Y = 3) as 1080/3125, P(Y = 4) as 162/625, and P(Y = 5) as 243/3125. Converting 162/625 to a common denominator of 3125 gives us 648/3125. Therefore, we add P(Y = 3), P(Y = 4), and P(Y = 5) to find P(Y ≥ 3), resulting in (1080/3125) + (648/3125) + (243/3125) = (1971/3125). Thus, A's chance of winning at least 3 games out of the 5 games is 1971/3125. The value I am writing down can be checked for correctness after simplification; you can use a calculator to compute it, and then you will find that this probability is approximately 0.68.



This is just to show how we can utilize binomial distribution to solve this kind of problem. Consider this problem: in a precision bombing attack, there is a 50% chance that any one bomb will strike the target. Two direct hits are required to completely destroy the target. The question is, how many bombs must be dropped to give a 99% chance or better of completely destroying the target? Let X be the random variable defined as X = 1 if a bomb strikes the target and X = 0 if a bomb does not strike the target.

According to the given information, in a precision bombing attack, there is a 50% chance that any one bomb will strike the target. The probability that X = 1, meaning the bomb strikes the target, is 1/2. Similarly, the probability that X = 0 is also 1/2. Since two direct hits are required to destroy the target completely, we need to determine how many bombs must be dropped to achieve a 99% probability of at least two hits.