PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

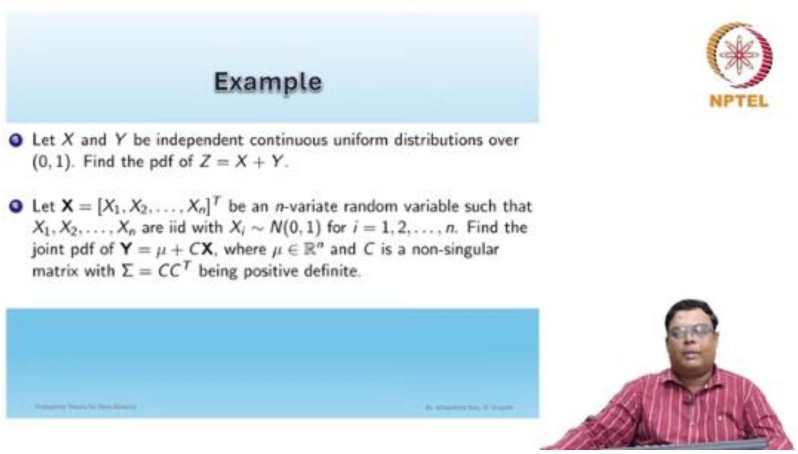
Indian Institute of Technology Tirupati

Week - 12

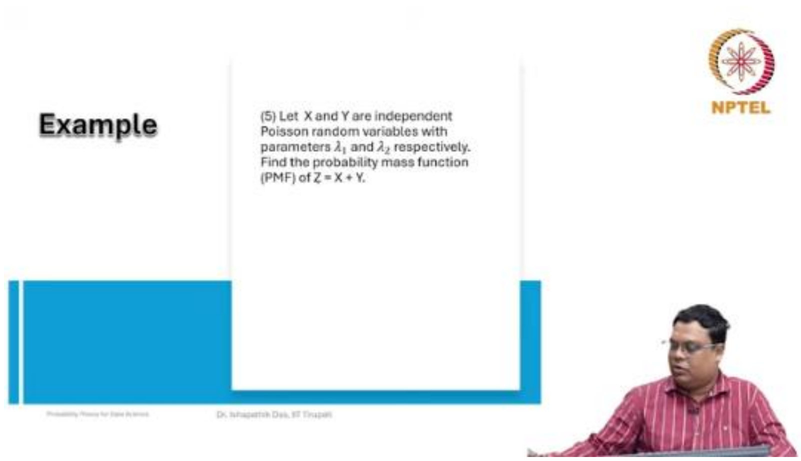
Lecture - 58

Transformation of Discrete Random Variables

So, I hope you have followed and understood this concept of transforming random variables. However, note that all of these theorems apply only to continuous random variables. These concepts will also be applicable in the case of convolution. In the discrete case, we cannot use this theorem because we cannot apply the derivative approach.



To demonstrate this, let's go through an example. We need to find the probability mass function of Z, where Z = X + Y. We previously studied this example, but for the purpose of understanding the transformation of random variables, let's go through it again. Suppose X and Y are independent Poisson random variables with parameters λ_1 and λ_2, respectively. Since X and Y are discrete random variables, we cannot use the theorem here. Instead, we will directly find the probability without using the theorem. Let X_1 and X_2 be independent Poisson random variables with parameters λ_1 and λ_2, respectively. We want to find the probability mass function of Z, where $Z = X_1 + X_2$. Since this involves discrete random variables, we cannot use the theorem from earlier. Instead, we will directly compute it.

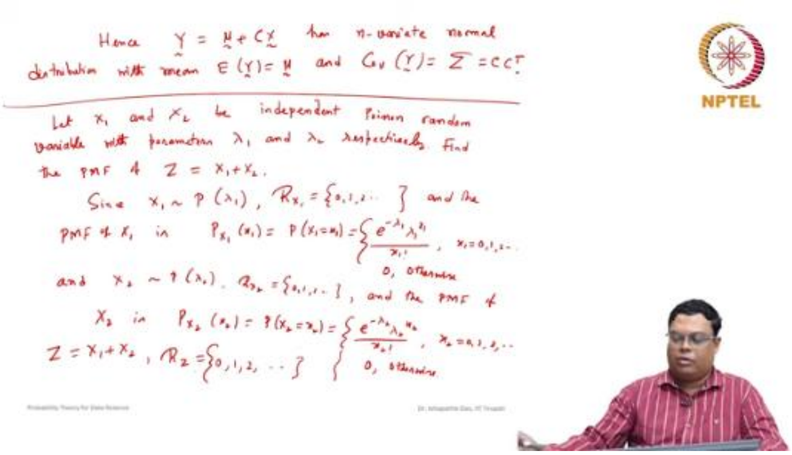


To start, since X_1 has a Poisson distribution with parameter λ_1, the range of X_1 will be 0, 1, 2, and so on. The probability mass function of X_1 can be written as $P(X_1 = x) = e^{(-\lambda_1)} * \lambda_1^{(X_1)} / X_1!$, for $X_1 = 0, 1, 2$, and so on. For other values, $P(X_1 = x) = 0$.

Similarly, X_2 follows a Poisson distribution with parameter λ_2. The range of X_2 will also be the same: 0, 1, 2, and so on. The probability mass function of X_2 can be written as $P(X_2 = x) = e^{(-\lambda_2)} * \lambda_2^{(X_2)} / X_2!$, for $X_2 = 0, 1, 2$, and so on. For other values, $P(X_2 = x) = 0$.

Now, let's consider the range of Z, where $Z = X_1 + X_2$. Since the range of both X_1 and X_2 is 0, 1, 2, and so on, the range of Z will also be 0, 1, 2, and so on. The minimum values of X_1 and X_2 can be 0, 1, 2, or any value. So, the range of Z will be 0, 1, 2, and so on.

Next, we need to determine the probability mass function of Z. To find the probability mass function of Z, where $Z = X_1 + X_2$, we need to calculate $P(Z = z)$. Let's consider different values for Z within its range, which includes 0, 1, 2, and so on. First, let's look at the case where $Z = 0$. For Z to be 0, the sum of X_1 and X_2 must equal 0.

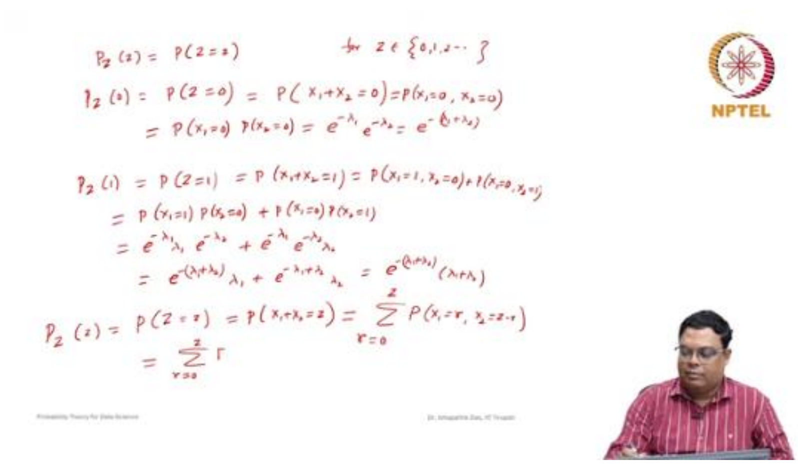


This is only possible if both X_1 and X_2 are 0, as their minimum values are 0. Since X_1 and X_2 are independent, we can multiply the probabilities for X_1 being 0 and X_2 being 0. The probability for X_1 being 0 is $e^{(-\lambda_1)}$, and the probability for X_2 being 0 is $e^{(-\lambda_2)}$. Therefore, the total probability is $e^{(-\lambda_1 + \lambda_2)}$.

Next, let's consider the case where $Z = 1$. For Z to be 1, the sum of X_1 and X_2 must equal 1. There are two possibilities for this: either $X_1 = 1$ and $X_2 = 0$, or $X_1 = 0$ and $X_2 = 1$. Since X_1 and X_2 are independent, we can add the probabilities of these two mutually exclusive events. The probability of X_1 being 1 is $e^{(-\lambda_1)} * \lambda_1^1 / 1!$, and the probability of X_2 being 0 is $e^{(-\lambda_2)} * \lambda_2^0 / 0!$. Similarly, the probability of X_1 being 0 and X_2 being 1 follows the same pattern. The total probability is the sum of the two possibilities: $e^{(-\lambda_1 + \lambda_2)} * \lambda_1 * \lambda_2$.

From these examples, we can see that there is a pattern emerging. To find the probability that $Z = z$, where Z is the sum of X_1 and X_2, we can express it as the summation of possible values. For any given value z, we can sum over the possible values of r, where r ranges from 0 to z. For each r, the probability that $X_1 = r$ and $X_2 = z - r$ is multiplied, as X_1 and X_2 are independent random variables.

For example, when $r = 0$, X_1 will be 0, and X_2 will have to be z. When $r = 1$, X_1 will be 1, and X_2 will be $z - 1$. This pattern continues for other values of r. So, in general, for each value of r from 0 to z, we compute the probability of $X_1 = r$ and the probability of $X_2 = z - r$. The final result is the summation of these probabilities. We will continue this process on the next page.



The probability that $Z = z$ can be found by summing over all possible values of r, from 0 to z. This probability is equal to the summation of $P(X_1 = r) * P(X_2 = z - r)$, for each r. Specifically, for each value of r, we compute $P(X_1 = r)$ and $P(X_2 = z - r)$. For the Poisson random variables X_1 and X_2, we use their respective probability mass functions. The probability that $X_1 = r$, given it follows a Poisson distribution with parameter λ_1, is represented as:

$P(X_1 = r) = e^{(-\lambda_1)} * (\lambda_1^r / r!)$

Similarly, the probability that $X_2 = z - r$, given it follows a Poisson distribution with parameter λ_2, is represented as:

$P(X_2 = z - r) = e^{(-\lambda_2)} * (\lambda_2^{(z - r)} / (z - r)!)$

When we combine these terms, we can factor out $e^{(-(\lambda_1 + \lambda_2))}$, which is independent of the summation range. The result becomes a simplification of the sum of the remaining terms. Now, summation of r is from 0 to z. So:

Σ (from $r = 0$ to z) $[(\lambda_1^r * \lambda_2^{(z - r)}) / (r! * (z - r)!)]$

Now, how can we simplify this? Let us multiply by z!, because there is a z! factorial, and this is an independent constant. We divide by the outside z!, and then it cancels out. Now, what is this value? You can recognize that this is a binomial expression. For example, if you have $(a + b)^n$, it can be represented as:

$(a + b)^n = \Sigma$ (from $r = 0$ to n) $[(n! / (r! * (n - r)!)) * a^r * b^{(n - r)}]$

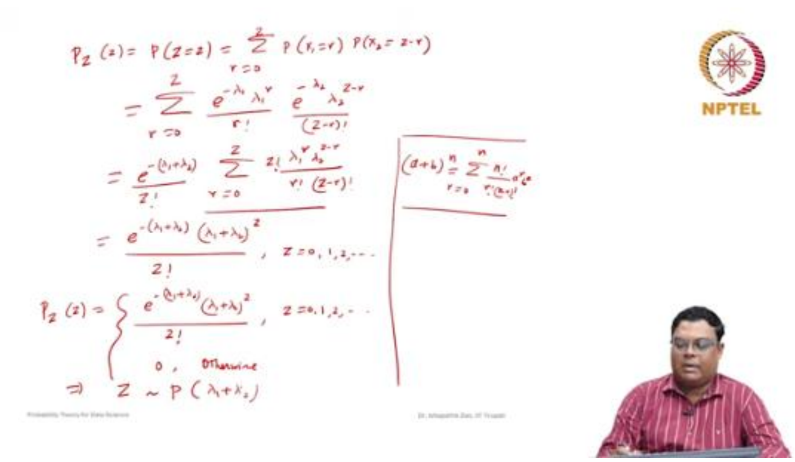
This represents (a + b)^n. Comparing this with the given expression, you can see that if you set n = z, the rest follows accordingly, with r! and (n - r)! as components. So, (z - r)! * a^r * λ_1^r * λ_2^(z - r) is essentially equal to:

e^(-(λ_1 + λ_2)) * ((λ_1 + λ_2)^z / z!)

This holds for any values of z = 0, 1, 2, and so on. For any other values, it is 0. Therefore, we can write:

P(Z = z) = e^(-(λ_1 + λ_2)) * ((λ_1 + λ_2)^z / z!)

whenever z = 0, 1, 2, and so on. For other values, the probability is 0. This distribution is also very similar to the Poisson distribution, which implies that Z follows a Poisson distribution with parameter (λ_1 + λ_2). But with this method, we can also find the transformed probability distribution function. However, it is a bit more complicated.



We did this example earlier, but in the context of transforming random variables, we just wanted to explain it again. This method is time-consuming and complicated, so we are looking for a simplified technique. Now, suppose Y_1, Y_2, and Y_n are independent Poisson random variables with parameters λ_1, λ_2, and λ_n, respectively. We need to find the probability mass function of Z, where Z = Y_1 + Y_2 + ... + Y_n. We can apply the previous approach, as we did for Y_1 + Y_2, and extend it to Y_1 + Y_2 + ... + Y_n.

It will be true that the sum of these variables will follow a Poisson distribution with the sum of the parameters, i.e., Z ~ Poisson(λ_1 + λ_2 + ... + λ_n). However, proving this is complicated, which is why we will demonstrate another technique. We need a different method. We will discuss another method to determine if there is a simplified approach to find the distribution more easily, without going through all these details. Next, we will explore a topic known as the moment generating function (MGF), which may be useful for finding the distribution of such transformed random variables.