PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

Indian Institute of Technology Tirupati

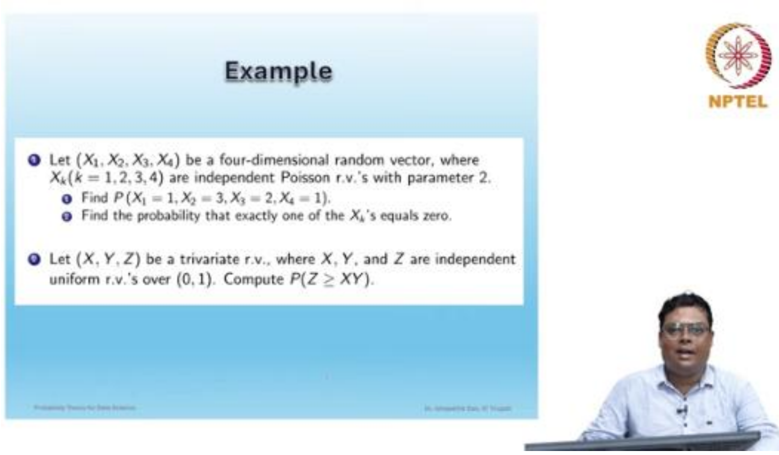
Week - 10

Lecture - 50

Numerical Examples on Joint Probability Mass Functions

Let us discuss some numerical examples using multiple random variables. Let X_1, X_2, X_3, and X_4 be a four-dimensional random vector, where each X_k is an independent Poisson random variable with a parameter of 2. Let X_1, X_2, X_3, and X_4 be four random variables, where X_3 and X_4 are independent Poisson random variables, each with a Poisson distribution with parameter $\lambda = 2$. Therefore, X_1, X_2, X_3, and X_4 form a 4-dimensional random vector, with each X_k being an independent Poisson random variable with parameter 2. Since they are independent, we can state that X_1, X_2, X_3, and X_4 are independent and identically distributed (i.i.d) random variables. The probability mass function (PMF) for each X_i is given by the formula for the Poisson distribution:

$P(X_i = x_i) = (e^{-\lambda} * \lambda^{x_i}) / (x_i!)$, where $\lambda = 2$ and $x_i = 0, 1, 2, ...$



Since all the random variables are i.i.d., the PMF is the same for each X_i, where $\lambda = 2$:

$P(X_i = x_i) = e^{\wedge}(-2) * 2^{\wedge}x_i / x!$, for $x_i = 0, 1, 2, ...$

Now, let's calculate the probability that $X_1 = 1$, $X_2 = 3$, $X_3 = 2$, and $X_4 = 1$. Since the variables are independent, the joint probability is the product of the individual probabilities:

$P(X_1 = 1, X_2 = 3, X_3 = 2, X_4 = 1) = P(X_1 = 1) * P(X_2 = 3) * P(X_3 = 2) * P(X_4 = 1)$

We now calculate each individual probability using the Poisson PMF:

For $X_1 = 1$:
$P(X_1 = 1) = e^{\wedge}(-2) * 2^1 / 1! = e^{\wedge}(-2) * 2$

For $X_2 = 3$:
$P(X_2 = 3) = e^{\wedge}(-2) * 2^3 / 3! = e^{\wedge}(-2) * 8 / 6 = e^{\wedge}(-2) * 4/3$

For $X_3 = 2$:
$P(X_3 = 2) = e^{\wedge}(-2) * 2^2 / 2! = e^{\wedge}(-2) * 4 / 2 = e^{\wedge}(-2) * 2$

For $X_4 = 1$:
$P(X_4 = 1) = e^{\wedge}(-2) * 2^1 / 1! = e^{\wedge}(-2) * 2$

Now, multiply all the individual probabilities:

$P(X_1 = 1, X_2 = 3, X_3 = 2, X_4 = 1) = e^{\wedge}(-2) * 2 * e^{\wedge}(-2) * 4/3 * e^{\wedge}(-2) * 2 * e^{\wedge}(-2) * 2$

$= e^{\wedge}(-8) * 2^4 * 4 / 3$

$= e^{\wedge}(-8) * 16 * 4 / 3$

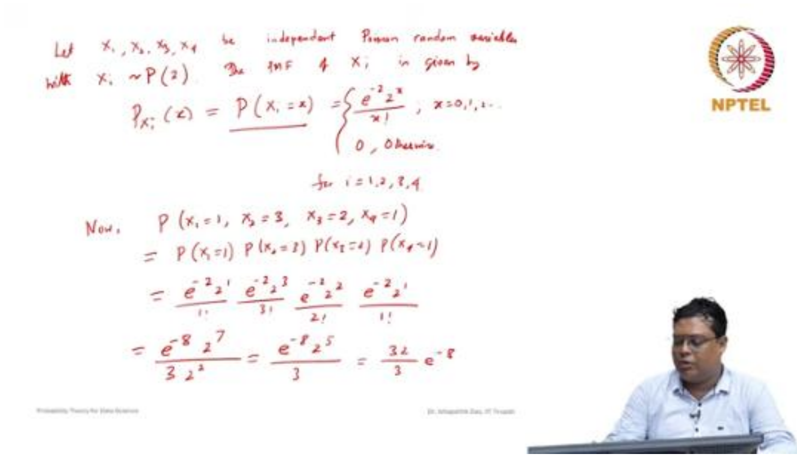
$= e^{\wedge}(-8) * 32 / 3$

Thus, the joint probability is:
$P(X_1 = 1, X_2 = 3, X_3 = 2, X_4 = 1) = (32 / 3) * e^{\wedge}(-8)$

This is the exact probability in its simplest form. You can also use a calculator to find the decimal value if needed. Otherwise, this expression is sufficient as the answer.

Now, let's move on to the next question, which asks to find the probability that exactly one of the X_k's equals 0. The random variables involved are X_1, X_2, X_3, and X_4, and we need to find the probability that exactly one of these variables is equal to 0. This means one of the

X_k's will take the value 0, and the other three will take values greater than 0. Since there are four random variables, there are four possible ways this can happen:



$X_1 = 0, X_2 \neq 0, X_3 \neq 0, X_4 \neq 0$
$X_1 \neq 0, X_2 = 0, X_3 \neq 0, X_4 \neq 0$
$X_1 \neq 0, X_2 \neq 0, X_3 = 0, X_4 \neq 0$
$X_1 \neq 0, X_2 \neq 0, X_3 \neq 0, X_4 = 0$

These are mutually exclusive events, so we can sum the probabilities of each of them. The probability of each event can be written as the product of the marginal probabilities:

$P(X_1 = 0, X_2 \neq 0, X_3 \neq 0, X_4 \neq 0) = P(X_1 = 0) * P(X_2 \neq 0) * P(X_3 \neq 0) * P(X_4 \neq 0)$
$P(X_1 \neq 0, X_2 = 0, X_3 \neq 0, X_4 \neq 0) = P(X_1 \neq 0) * P(X_2 = 0) * P(X_3 \neq 0) * P(X_4 \neq 0)$
$P(X_1 \neq 0, X_2 \neq 0, X_3 = 0, X_4 \neq 0) = P(X_1 \neq 0) * P(X_2 \neq 0) * P(X_3 = 0) * P(X_4 \neq 0)$
$P(X_1 \neq 0, X_2 \neq 0, X_3 \neq 0, X_4 = 0) = P(X_1 \neq 0) * P(X_2 \neq 0) * P(X_3 \neq 0) * P(X_4 = 0)$

Since the random variables are independent, we can use their individual probabilities to compute the joint probability.

Now, let's break down how to calculate these probabilities:

$P(X_i = 0)$ for any random variable X_i is the probability that the Poisson random variable equals 0.

From the Poisson distribution with parameter $\lambda = 2$, we know: $P(X_i = 0) = e^{(-2)} * 2^0 / 0! = e^{(-2)}$

$P(X_i \neq 0)$ is the complement of $P(X_i = 0)$, so: $P(X_i \neq 0) = 1 - P(X_i = 0) = 1 - e^{(-2)}$

Since all X_i's are i.i.d., we can apply these probabilities to each term in the sum:

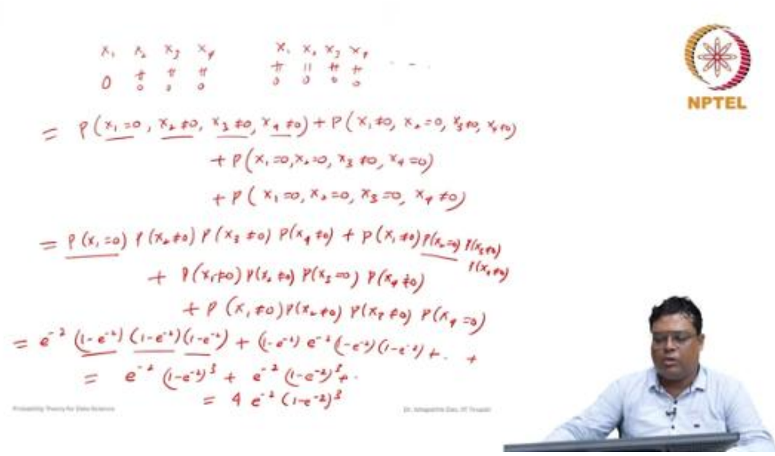
$P(X_1 = 0, X_2 \neq 0, X_3 \neq 0, X_4 \neq 0) = e^{(-2)} * (1 - e^{(-2)})^3$
$P(X_1 \neq 0, X_2 = 0, X_3 \neq 0, X_4 \neq 0) = (1 - e^{(-2)}) * e^{(-2)} * (1 - e^{(-2)})^2$
$P(X_1 \neq 0, X_2 \neq 0, X_3 = 0, X_4 \neq 0) = (1 - e^{(-2)})^2 * e^{(-2)} * (1 - e^{(-2)})$
$P(X_1 \neq 0, X_2 \neq 0, X_3 \neq 0, X_4 = 0) = (1 - e^{(-2)})^3 * e^{(-2)}$

Now, adding all these terms together: Total probability $= 4 * e^{(-2)} * (1 - e^{(-2)})^3$

This is the probability that exactly one of the X_k's equals 0. Thus, the answer is: $P(\text{exactly one } X_k = 0) = 4 * e^{(-2)} * (1 - e^{(-2)})^3$

This is the final probability, expressed in terms of $e^{(-2)}$. You can compute the exact numerical value by evaluating this expression, or you can leave it in this form as the exact answer.

Now, let's redefine the random variable to make use of the binomial distribution concept. Let Y_i be a random variable defined as follows:
$Y_i = 1$ if $X_i = 0$,
$Y_i = 0$ if $X_i \neq 0$.



In this case, Y_i is a Bernoulli random variable. We know the following probabilities:
$P(X_i = 0) = e^{-2}$,
$P(X_i \neq 0) = 1 - e^{-2}$.

So, the probability mass function for Y_i is:
$P(Y_i = 1) = e^{-2}$ (since $Y_i = 1$ implies $X_i = 0$),
$P(Y_i = 0) = 1 - e^{-2}$ (since $Y_i = 0$ implies $X_i \neq 0$).

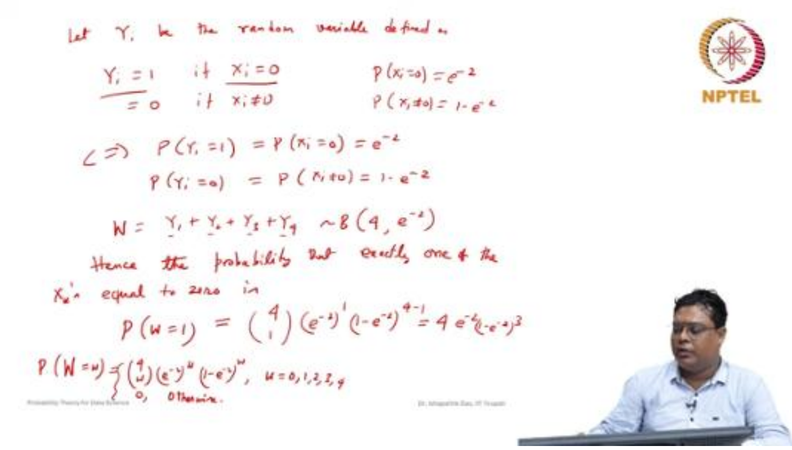
Now, consider the sum of these random variables:
$W = Y_1 + Y_2 + Y_3 + Y_4$.
This sum W follows a binomial distribution with parameters n = 4 (the number of trials) and $p = e^{-2}$ (the probability of success, i.e., $Y_i = 1$).
The binomial distribution gives the probability mass function:
$P(W = w) = \binom{4}{w} * (e^{-2})^w * (1 - e^{-2})^{(4 - w)}$, for w = 0, 1, 2, 3, 4.

Now, we are asked to find the probability that exactly one of the X_k's is equal to 0. This is equivalent to the probability that exactly one of the Y_1, Y_2, Y_3, Y_4 variables equals 1.
In other words, we want the probability that the sum W = 1. Thus, the probability is:
$P(W = 1) = \binom{4}{1} * (e^{-2})^1 * (1 - e^{-2})^3$.

We can now compute each part:
$\binom{4}{1} = 4$,
$(e^{-2})^1 = e^{-2}$,
$(1 - e^{-2})^3$ remains as it is.

So, the final probability is:
$P(W = 1) = 4 * e^{-2} * (1 - e^{-2})^3$.

This is the desired probability that exactly one of the X_k's is equal to 0. You can compute this value numerically using a calculator or leave it in this form as the exact result. This is a different approach to solving this problem.



We explicitly computed the probability, but this approach is more general. It involves adding multiple terms and finding their values. However, if the problem asks something different, such as the sum of two X_i being 1 and the remaining X_i being exactly 0, we would need to use the binomial concept. In such cases, we must apply transformation techniques. This is part of the concept of transformation of random variables, which we will explore in upcoming lectures.

Transformation of random variables can be useful for calculating probabilities. This is the concept we applied in the previous example, and now we will discuss another numerical example.

Let X, Y, and Z be independent uniform random variables over the interval [0, 1]. We are asked to compute the probability that $Z \geq XY$. In the previous example, we worked with discrete random variables, but here we are dealing with continuous random variables.

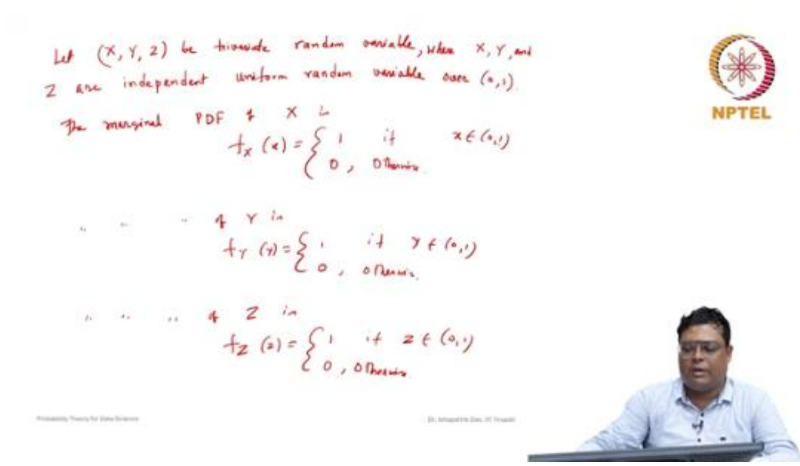
Let X, Y, and Z be independent uniform random variables over the interval [0, 1]. Since they are independent and identically distributed, their marginal probability density functions are the same. The probability density function for X is:
$f_X(x) = 1$ for $0 \leq x \leq 1$, and 0 otherwise.
This is true for Y and Z as well.

Thus, the marginal probability density function for Y is:
$f_Y(y) = 1$ for $0 \leq y \leq 1$, and 0 otherwise.
And similarly, the marginal probability density function for Z is:
$f_Z(z) = 1$ for $0 \leq z \leq 1$, and 0 otherwise.

Since the random variables are independent, we can calculate the joint probability density function by multiplying the individual PDFs.

Therefore, the joint probability density function is:
f_X,Y,Z(x, y, z) = 1 for $0 \leq x, y, z \leq 1$, and 0 otherwise.

The joint probability density function of X, Y, and Z is given by f_X,Y,Z(x, y, z). Since these are independent random variables, we know that the joint probability density function can be obtained by multiplying the individual probability density functions: f_X, f_Y, and f_Z. As we have already determined, each of these is 1 in the interval [0, 1] and 0 otherwise.



Therefore, the joint probability density function is: f_X,Y,Z(x, y, z) = 1 for $0 \leq x, y, z \leq 1$, and 0 otherwise. Now that we know the joint probability density function, we can compute the probability that $Z \geq XY$. To compute this, we need to evaluate the integral over the region where $Z \geq XY$ in three-dimensional space.

We will integrate over the region where X, Y, and Z satisfy the condition $Z \geq XY$. The integral we need to compute is: $P(Z \geq XY) = \iiint$ f_X,Y,Z(x, y, z) dz dy dx. Since we know that the joint probability density function is 1, this simplifies to: $P(Z \geq XY) = \iiint dz\, dy\, dx$.

The region of integration for Z is from XY to 1, and for X and Y, the limits are from 0 to 1. Therefore, we can write the integral as: $P(Z \geq XY) = \int_0^1 \int_0^1 \int_{xy}^1 dz\, dy\, dx$.

First, integrate with respect to z: $\int_{xy}^1 dz = 1 - XY$.

Now the integral becomes: $P(Z \geq XY) = \int_0^1 \int_0^1 (1 - XY)\, dy\, dx$.

Next, integrate with respect to y: $\int_0^1 (1 - XY)\, dy = y - Xy^2/2$ evaluated from 0 to 1 = 1 - X/2.

So the integral now is: $P(Z \geq XY) = \int_0^1 (1 - X/2)\, dx$.

Finally, integrate with respect to x: $\int_0^1 (1 - X/2)\, dx = x - X^2/4$ evaluated from 0 to 1 = 1 - 1/4 = 3/4.

Thus, $P(Z \geq XY) = 3/4$. Hopefully, you now understand how to compute such probabilities using the joint probability density function.

This is just one example, and you can refer to additional resources to practice more examples. By solving more problems, you will gain a clearer understanding. If you have any doubts, feel free to clarify them through practice. Next, we will discuss another numerical example.