PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

Indian Institute of Technology Tirupati

Week - 11

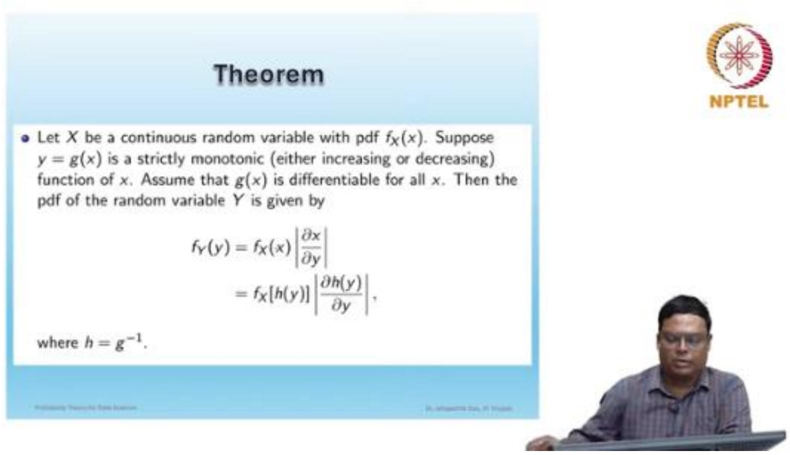
Lecture - 54

Theorem on Transformation of Random Variables

Let X be a continuous random variable with a probability density function f_X(x), and suppose Y = g(x) is a strictly monotonic function of x, either increasing or decreasing. Additionally, assume that g(x) is a differentiable function. Under these conditions, we can directly find the probability density function (PDF) of Y from the PDF of X using the following formula:

f_Y(y) = f_X(g^{-1}(y)) * |d/dy[g^{-1}(y)]|, where g^{-1}(y) is the inverse of g(x) and |d/dy[g^{-1}(y)]| is the absolute value of the derivative of the inverse function.

Let us now discuss the conditions. Let X be a continuous random variable with a probability density function f_X, and suppose g is a strictly monotonic and differentiable function.



Specifically, g is a strictly increasing or decreasing function. This means that if $X_1 < X_2$, for a strictly increasing function, $g(X_1)$ will be strictly less than $g(X_2)$. For a strictly decreasing function, $g(X_1)$ will be strictly greater than $g(X_2)$. The function g will not be constant anywhere. If we consider that Y = g(X), the inverse of g exists, denoted as h, where h is the inverse of g.

This means that X = h(Y), or equivalently, $X = g^{-1}(Y)$. If we take the derivative of h with respect to y, we have dx/dy = dh(y)/dy. This derivative will either be greater than 0 or less than 0 for all y, and it will never be equal to 0. Using this condition, the probability density function of Y can be found using the relationship:

f_Y(y) = |dx/dy| * f_X(x), where x = h(y).

Here, the absolute value of dx/dy is taken, ensuring that the result is always positive. This allows us to compute the probability density function of Y based on the probability density function of X. Before proving this theorem, let us first look at a numerical example to understand the process more clearly. As we can see, the process is quite involved and lengthy. First, suppose this is the probability density function, and this is the transformation.

You first find the cumulative distribution function, then find the cumulative distribution function of Y, take the derivative, and finally find the probability density function of Y. Here, we will use this theorem directly to find the probability density function of Y from the probability density function of X. Let X be a continuous random variable with the probability density function f_X(x) = 2x, for $0 \leq x \leq 1$. We are asked to find the probability density function of Y = g(X) = 3X + 1. Since the transformation Y = g(X) is strictly increasing, we can find the inverse.

To find the inverse, set Y = 3X + 1. Solving for X, we get X = (Y - 1) / 3. This is the inverse function h(Y) = (Y - 1) / 3. Next, we compute the derivative of X with respect to Y: dx/dy = 1/3. Using the formula for the probability density function of Y, we have:

f_Y(y) = |dx/dy| * f_X(h(y)).

Substitute h(y) = (y - 1) / 3 into f_X:

f_Y(y) = (1/3) * f_X((y - 1) / 3).

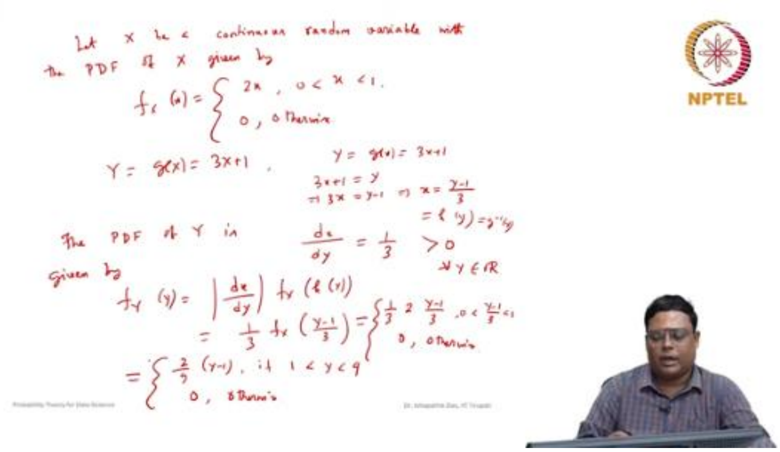
Now, substitute f_X(x) = 2x into the equation:

$f_Y(y) = (1/3) * 2 * ((y - 1) / 3) = 2/9 * (y - 1)$.

The probability density function of Y is:

$f_Y(y) = 2/9 * (y - 1)$ for $1 \leq y \leq 4$, and 0 otherwise.

This method is simple and can be applied to find the probability density function of any other transformation of random variables, as long as the condition of the transformation being strictly increasing or decreasing is satisfied. Let's discuss another example.



Let $Y = \tan(X)$, and we need to find the probability density function of Y if X is a uniform random variable over the interval [0, 1]. Since X is uniformly distributed over [0, 1], the probability density function of X is:

$f_X(x) = 1$ for $0 \leq x \leq 1$, and 0 otherwise.

Now, let's consider the transformation $Y = \tan(X)$. The function $\tan(X)$ is strictly increasing between 0 and 1.

We can express X in terms of Y using the inverse function: $X = \tan^{-1}(Y)$. However, we realize there is a mistake in the problem setup. The uniform distribution should be defined over the interval $[0, \pi/2]$, not [0, 1], since the tangent function is increasing in the interval $[0, \pi/2]$. So, let X be a uniform random variable between 0 and $\pi/2$.

The probability density function of X is:

$f_X(x) = 2/\pi$ for $0 \leq x \leq \pi/2$, and 0 otherwise.

This ensures that the distribution is uniform over the interval $[0, \pi/2]$. Next, we compute the derivative of the inverse transformation ($X = \tan^{-1}(Y)$) to find dx/dy. The derivative of $\tan^{-1}(Y)$ is:

$dx/dy = 1 / (1 + y^2)$.

Since the derivative is always positive, the transformation is valid. The range of Y is from 0 to infinity. When X = 0, Y = tan(0) = 0, and when $X = \pi/2$, Y approaches infinity. Therefore, Y takes values in the interval $[0, \infty)$.

To find the probability density function of Y, we use the following formula:

$f_Y(y) = |dx/dy| * f_X(h(y))$, where $h(y) = \tan^{-1}(y)$.

Substituting into the formula:

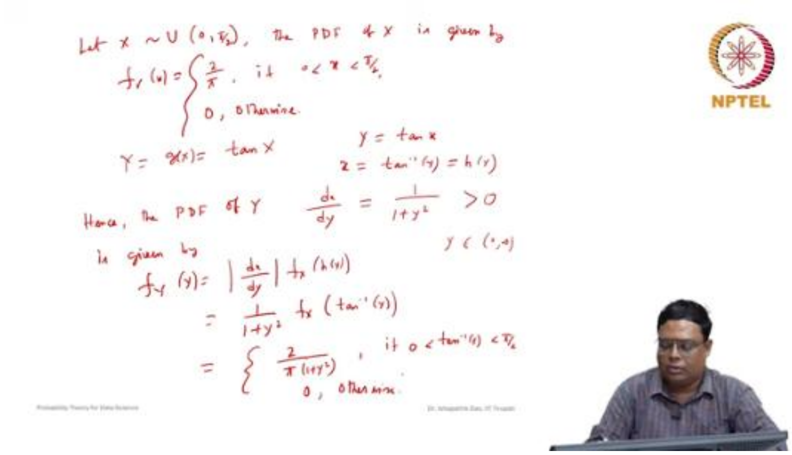
$f_Y(y) = (1 / (1 + y^2)) * (2/\pi)$ for $y \geq 0$.

Simplifying, we get:

$f_Y(y) = 2 / (\pi(1 + y^2))$ for $y \geq 0$, and 0 otherwise.

This is the probability density function of Y. Hence, the probability density function of Y is given by:

$f_Y(y) = 2 / (\pi(1 + y^2))$ for y in the interval $[0, \infty)$, and 0 otherwise.



This is because the tangent function is increasing, with $\tan(0) = 0$ and $\tan(\pi/2)$ approaching infinity. Therefore, Y takes values between 0 and infinity. Now, let's see how we can prove the theorem we used earlier. The theorem states that if X is a continuous random variable with a known probability density function, and we apply a strictly increasing transformation, we can find the probability density function of Y. Let's prove this theorem step by step.

Without loss of generality, assume that $g: \mathbb{R} \to \mathbb{R}$ is strictly increasing. The process will be similar for a decreasing function. Let X be a continuous random variable with probability density function $f_X(x)$ and cumulative distribution function $F_X(x)$. Now, we want to find the cumulative distribution function of Y, where $Y = g(X)$. The cumulative distribution function of Y is:

$F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$.

Since g is strictly increasing, we can apply the inverse function g^{-1} to both sides:

$P(g(X) \leq y) = P(X \leq g^{-1}(y)) = P(X \leq h(y))$, where $h(y) = g^{-1}(y)$.

This is simply the cumulative distribution function of X evaluated at $h(y)$:

$F_Y(y) = F_X(h(y))$.

Now, to find the probability density function of Y, we differentiate the cumulative distribution function $F_Y(y)$ with respect to y:

$f_Y(y) = \frac{d}{dy} [F_X(h(y))]$.

Using the chain rule, we differentiate $F_X(h(y))$ with respect to y:

$f_Y(y) = f_X(h(y)) * dh(y)/dy$.

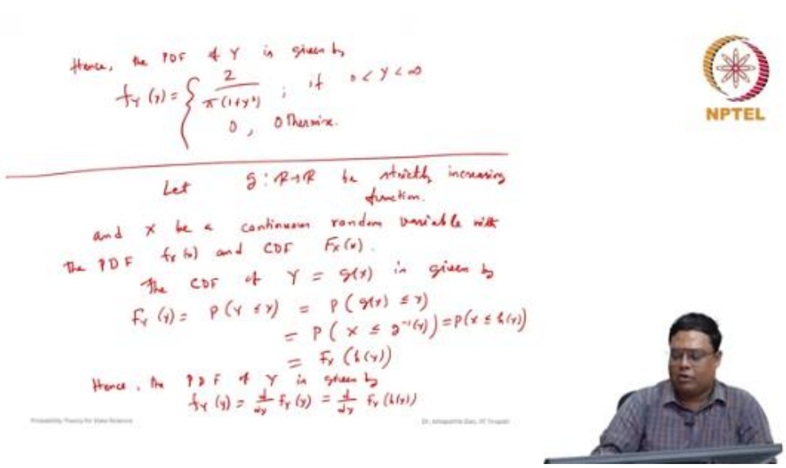
Since $h(y) = g^{-1}(y)$, the derivative $dh(y)/dy$ is the derivative of $g^{-1}(y)$, which is:

$dh(y)/dy = 1 / g'(h(y))$.

Thus, the probability density function of Y is:

$f_Y(y) = f_X(h(y)) * |1 / g'(h(y))|$.

This completes the proof for the general case of a strictly increasing transformation.



We found that the probability density function of Y is given by $f_Y(y)$, which is equal to the derivative of $f_X(h(y))$ with respect to y. This can be expressed as:

$d/dy\ [f_Y(y)] = d/dy\ [f_X(h(y))]$.

This is equivalent to the derivative of f_X with respect to x, evaluated at $x = h(y)$, multiplied by dx/dy. Therefore, we have:

$f_Y(y) = (dx/dy) * (d/dx\ f_X(x))$ evaluated at $x = h(y)$.

Now, if g is a strictly decreasing function, the cumulative distribution function of Y is given by:

$F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$.

Since g is a strictly decreasing function, this probability is the same as:

$P(X \geq g^{-1}(y))$.

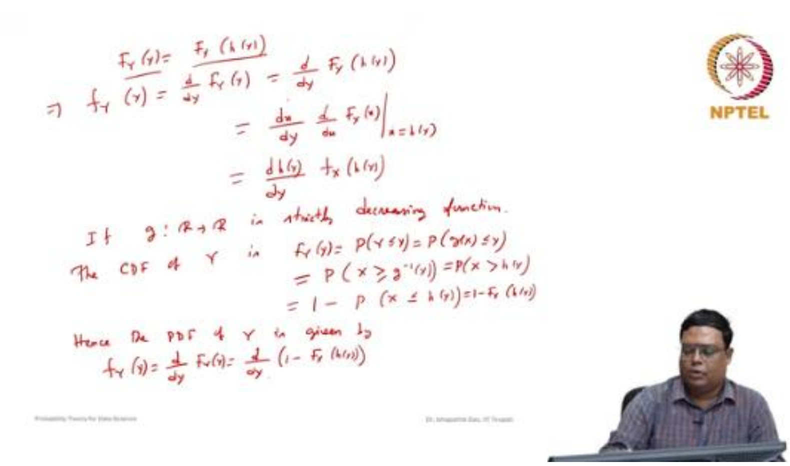
This can be written as:

$P(X \leq h(y)) = 1 - P(X \leq h(y)) = 1 - F\_X(h(y))$.

Thus, the probability density function of Y is:

f_Y(y) = -d/dy [F_X(h(y))] = - (dx/dy) * f_X(h(y)).

Since the derivative of the CDF is negative for decreasing functions, multiplying by a negative sign makes the probability density function positive.



So, we have combined the cases for strictly increasing and decreasing functions. Therefore, we can express the probability density function of Y as:

f_Y(y) = f_X(h(y)) * |dx/dy|,

where h(y) is the inverse function of g(x), and dx/dy is positive for strictly increasing functions and negative for strictly decreasing functions (with the negative sign making the result positive).

This theorem applies when X is a continuous random variable, and the transformation function g is either strictly increasing or decreasing, differentiable, and has an inverse.

Let's now apply this theorem to a specific example: let X be a normal distribution with mean 0 and variance 1, and let $Y = e^X$. We need to find the probability density function of Y. Given that X follows a normal distribution with mean 0 and variance 1, the probability density function of X is:

$f_X(x) = 1 / \sqrt{2\pi} * e^{-x^2/2}$, for x in the range from $-\infty$ to $+\infty$.

Now, consider the transformation $Y = e^X$, where $g(x) = e^x$. This is a strictly increasing function, and the inverse of this function is:

$x = \log(y)$,

which means $h(y) = \log(y)$.

Next, we compute dx/dy. Since $x = \log(y)$, we differentiate with respect to y to get:

$dx/dy = 1 / y$.

Thus, the probability density function of Y is:

$f_Y(y) = |dx/dy| * f_X(h(y))$,

which becomes:

$f_Y(y) = (1 / y) * f_X(\log(y))$.

Substituting the expression for $f_X(x)$, we get:

$f_Y(y) = (1 / y) * (1 / \sqrt{2\pi}) * e^{-(\log(y))^2 / 2}$.

This is the probability density function of Y, valid for $y > 0$.

For values of y outside this range, the probability density function is 0.

In summary, the probability density function of $Y = e^X$, where X follows a standard normal distribution, is:

$f_Y(y) = (1 / (y\sqrt{2\pi})) * e^{-(\log(y))^2 / 2}$, for $y > 0$, and 0 otherwise.

This concludes the discussion of the transformation of random variables in invariant cases. Hopefully, you have followed the concepts, and I encourage you to go through additional examples for further practice. You can refer to the textbooks mentioned earlier, which provide more numerical examples and detailed explanations. Additionally, you may want to verify the computations to ensure accuracy.

Some of the problems you'll encounter in the assignments will also cover this topic, providing an opportunity for further practice. As you work through the assignments, if you have any doubts, feel free to clarify them. The more you practice, the clearer the concept will become.

In this session, we focused on the transformation of random variables for single variables. Next, we will explore transformations for multiple random variables, such as bivariate or multivariate cases. We will discuss the related theories, results, and numerical examples specific to these scenarios.