PROBABILITY THEORY FOR DATA SCIENCE

Prof. Ishapathik Das

Department of Mathematics and Statistics

Indian Institute of Technology Tirupati

Week - 12

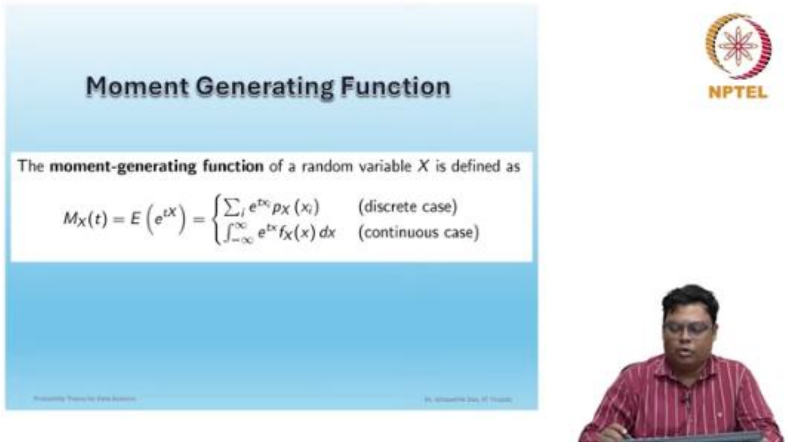
Lecture - 59

Moment Generating Functions

Let us discuss the moment generating function, which is an important function corresponding to a random variable X. We have learned that for a random phenomenon or sample space, X is a random variable, meaning X is a measurable function from the sample space S to $\mathbb{R}$. To identify this, we use different measures. One such measure is the cumulative distribution function (CDF) of X. If X is discrete, we use the probability mass function (PMF), and if X is continuous, we use the probability density function (PDF).

With these, we can perform computations such as calculating moments, mean, and variance. We have also learned about the transformation of random variables. Now, another important function for a random variable is the moment generating function. The moment generating function, denoted $M_x(t)$, is defined as the expected value of e^{tx}. For the discrete case, this is expressed as $\Sigma\ e^{tx} * p_x(X_i)$, where $p_x(X_i)$ is the probability mass function.

For the continuous case, it is defined as $\int_{-\infty}^{+\infty} e^{tx} * f_x(x)\, dx$, where $f_x(x)$ is the probability density function. Let us discuss the moment generating function. Let X be a random variable. The moment generating function, often abbreviated as MGF, is defined as the expected value of e^{tx}. Now, if X is a random variable, it can either be discrete or continuous. If X is discrete, it has a probability mass function, and if X is continuous, it has a probability density function.



When we apply a transformation, such as Y = g(X), we can find the expected value of Y. For a discrete random variable, the expected value of g(X) is Σ $g(X_k)$ * $P_x(X_k)$. For a continuous random variable, the expected value of g(X) is ∫ g(X) * $f_x(X)$ dX. In this case, we are interested in the expected value of e^{tx}. Here, g(X) = e^{tx}, with t being a real number.

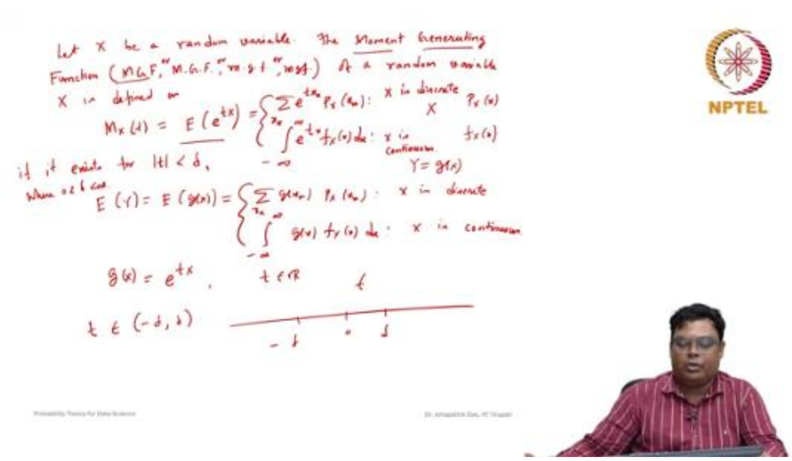
If t = 0, then e^{tx} = 1, meaning that g(X) = 1. If t = 1, then it becomes e^{x}. If t = 2, it becomes e^{2x}. If t = -3, it becomes e^{-3x}. Based on this, we need to find the expected value of e^{tx}.

From this general formula, if we replace e^{tx}, it becomes Σ (X_k * $e^{tx}{}_k$ * $P_x(X_k)$) when X is discrete. For the continuous case, it becomes $\int_{-\infty}^{+\infty}$ e^{tx} * $f_x(X)$ dX.

This is called a moment generating function because it is used to generate the moments of the distribution, such as the mean and variance, by taking derivatives of the function. The moment generating function may or may not exist because the sum must be absolutely summable for the expected values to exist. The integration must also exist. If it does not converge or is infinite, then we say that the moment generating function does not exist.

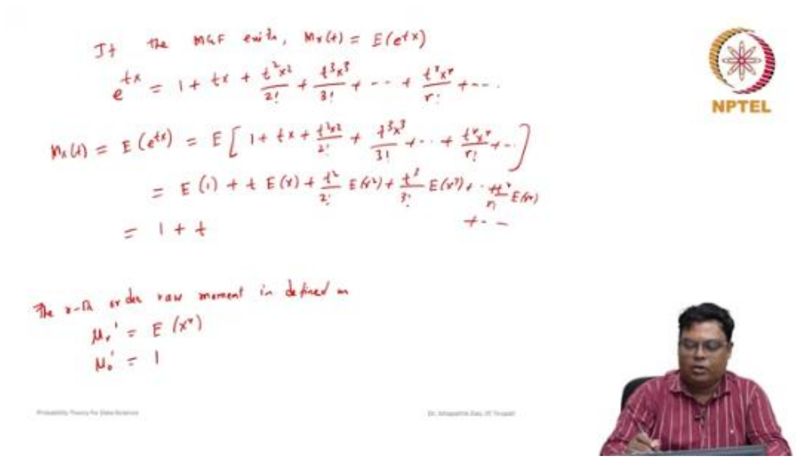
It is called the moment generating function if it exists for some values of t, less than δ, where δ is a real number greater than 0. The moment generating function may or may not exist, but if it does exist for values of t within the interval from -δ to +δ, then we say that the moment generating function exists. For example, if δ is some number, the interval would range from -δ to +δ. If the moment generating function exists for all values of t within this interval, then we can say the moment generating function exists.

Now, the question is, if the moment generating function exists, why is it called the moment generating function? It is called this because it seems to generate moments. If we assume that the moment generating function exists, we define it as the expected value of e^{tx}. We also know that e^{tx} can be represented as an infinite sum. The expression e^{tx} can be represented as $1 + tX + t^2X^2 / 2! + t^3X^3 / 3!$, and so on. In general, it can be written as $t^r * X^r / r!$, where r is a natural number.



Now, for the moment generating function $M_x(t)$, we have the expected value of e^{tx}, which is equivalent to the expected value of the sum: $1 + tX + t^2X^2 / 2! + t^3X^3 / 3! + \ldots$, with the general term being $t^r * X^r / r!$, Using the properties of expected values, this can be written as the sum of expected values: $1 + t * E(X) + t^2 / 2! * E(X^2) + t^3 / 3! * E(X^3) + \ldots$, with the general term being $t^r / r! * E(X^r)$. The expected value of 1 is 1, and for other terms, we recall that these are the raw moments of the distribution. The r-th order raw moment is denoted as μ_r', which is the expected value of X^r.

For example, when r = 0, μ_0' is the expected value of X^0, which is 1. Therefore, we can write the moment generating function as the sum of infinite moments, each multiplied by the appropriate coefficient. The sum represents the moment generating function as a series of moments with their corresponding coefficients. We can see that the expected values of e^{tx} can be represented as an infinite series: $1 + tX + t^2X^2 / 2!$, and so on. If we assume that all the moments, such as the r-th order moments, exist, we can proceed with the following.



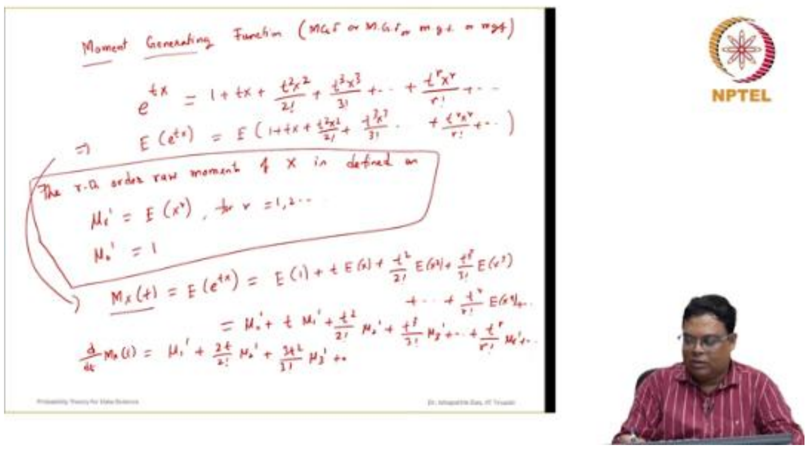
The r-th order moment of a random variable X is defined as the raw moment, which we denote as μ_r'. For example, the expected value of X raised to the power of 0, μ_0', is 1. Similarly, the expected value of X, which is μ_1', is the first raw moment. We can follow this notation for r = 1, 2, and so on. Now, if we assume that all the moments exist, we can take the expected value on both sides.

Using the properties of expected values, we can break this down into the sum of expected values: $1 + t * E(X) + t^2 / 2! * E(X^2) + \ldots$, continuing with the general term. This infinite series can be represented as the moment generating function, provided the moments exist. The moment generating function, if it exists for certain values of t within an interval, can be represented as a sum of terms involving the moments of the random variable. Therefore, all the moments must exist for this representation to hold. We can write it as $\mu_0' + t * \mu_1' + t^2 / 2! * \mu_2' + t^3 / 3! * \mu_3' + \ldots$, continuing with $t^r / r! * \mu_r'$.

This is how we represent the moment-generating function. Now, why is it called a moment-generating function? Suppose we know the moment function of a random variable, but we do not know its distribution function or probability mass function in the case of discrete variables, or the probability density function in the case of continuous variables. Using the moment-generating function, we can derive the moments of the random variable. How can we do this? We differentiate with respect to t. Taking the derivative of the moment-generating function, we get $\mu_1' + 2t / 2! * \mu_2' + 3t^2 / 3! * \mu_3' + \ldots$, continuing with $r * t^{r-1} / r! * \mu_r'$.

If we differentiate a second time, we can proceed in the same manner. Let's write this again on the next page. Thus, we have the moment-generating function as e^{tx}, assuming this expression exists. So, this is $\mu_0' + t * \mu_1' + t^2 / 2! * \mu_2' + t^3 / 3! * \mu_3' + \ldots$, continuing with $t^r / r! * \mu_r'$.

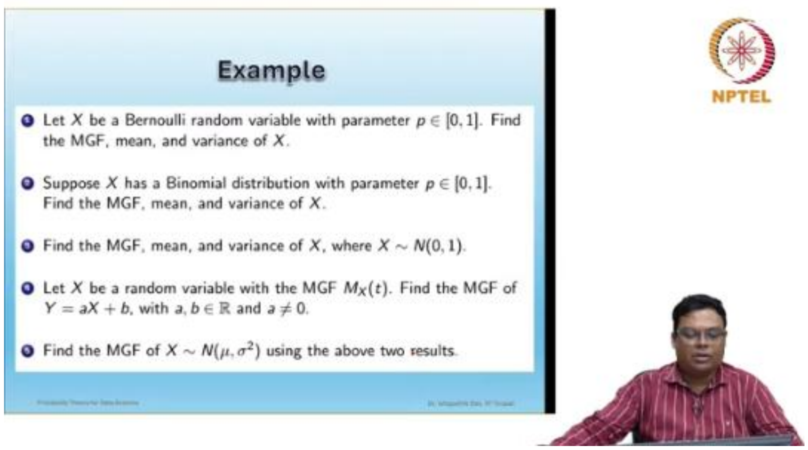
Now, let's check if this is correct. If you take the first derivative of $M_x(t)$ with respect to t, the result is $\mu_1' + 2t / 2! * \mu_2' + 3t^2 / 3! * \mu_3' + \ldots$, continuing with $r * t^{r-1} / r! * \mu_r'$. Next, if you take the second derivative of $M_x(t)$ with respect to t, you get $\mu_2' + 3 * 2t^2 / 3! * \mu_3' + \ldots$, continuing with $r * (r-1) * t^{r-2} / r! * \mu_r'$.



Now, if you differentiate this again with respect to t, the result is $\mu_3' + 4 * 3 * 2 * t / 4! * \mu_4' + \ldots$, continuing with $r * (r - 1) * (r - 2) * t^p / r! * \mu_r'$. By differentiating repeatedly, you can observe a developing pattern. The terms follow a similar structure, continuing with higher-order derivatives of the moments. So, we can now understand this pattern. For a particular r, if you differentiate the moment-generating function $M_x(t)$, the result will start from μ_r' and it will become $t * \mu_{r+1}' / (r + 1)! + \ldots$.

This pattern holds for any value of r, such as 1, 2, and so on. You can verify that this is true for r = 1 as well. Now, from this expression, what we gain is that if you differentiate the moment-generating function for the first time with respect to t and then evaluate it at t = 0, you will get μ_1'. Similarly, if you differentiate it a second time and evaluate the result at t = 0, you will get μ_2'. In general, if you continue differentiating the moment-generating function r times and then evaluate the result at t = 0, you will obtain μ_r', which represents the r-th order moment.

The key takeaway here is that whenever the moment-generating function is known, finding the moments directly from the probability mass function or probability density function may be complicated for certain random variables. In such cases, it can be simpler to use the moment-generating function. By differentiating it r times and evaluating at t = 0, you can easily obtain the r-th order moment, known as the raw moment, from the moment-generating function. So, that is why it is called the moment-generating function. Now, let's discuss an example of how this moment-generating function can be useful in finding certain properties. Consider the example where X is a Bernoulli random variable with parameter p, where p ∈ [0, 1]. We are asked to find the moment-generating function, the mean, and the variance of X. Although we have already computed these values earlier, in this case, we will use the moment-generating function to find them.



First, we will determine the moment-generating function for this case. After that, we will use it to find the mean and variance of X. Let's proceed with that. Let X be a Bernoulli random variable with parameter p, where p ∈ [0, 1]. The moment-generating function (MGF) is a helpful tool, and in this case, we will use its definition to find the MGF and, from there, compute the mean and variance. The probability mass function (PMF) of X is given as: P(X = k) = p^k * (1 - p)^(1 - k), where k ∈ {0, 1}. If X = 0, P(X = 0) = (1 - p), and if X = 1, P(X = 1) = p.

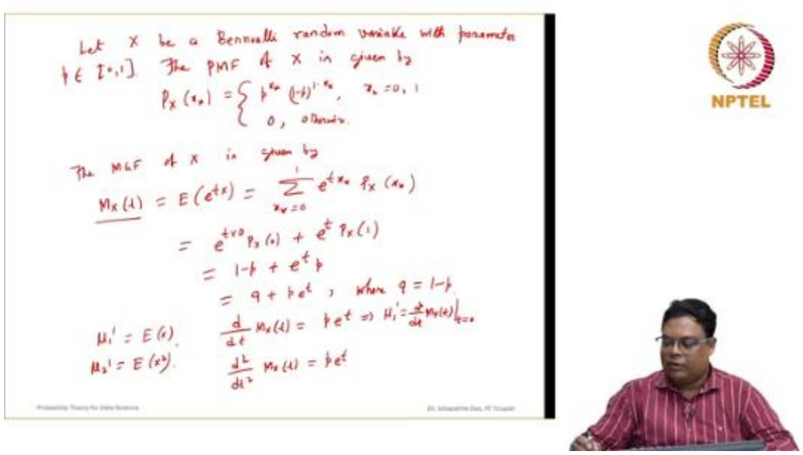
The moment-generating function of X is defined as E(e^{tx}). Since X is a discrete random variable, we sum over all the possible values of X. Thus, the MGF of X can be written as Σ e^{tx} * P(X = x). So, the MGF of X = (1 - p) * e^{t0} + p * e^{t1}. This simplifies to (1 - p) + p *

e^t. For simplicity, let q = (1 - p). This allows us to rewrite the moment-generating function as q + p * e^t.

This is the moment-generating function for a Bernoulli random variable X with parameter p. So, where q = 1 - p, we can see that the moment-generating function exists for any real number t. This means that the MGF is valid for all finite values of t, which ensures no issues with its use. Now, we aim to find the mean and variance using the moment-generating function. To find the mean, we first need to differentiate the moment-generating function.

The MGF is given as q + p * e^t. The first derivative of this with respect to t is p * e^t. To find the mean, we evaluate this derivative at t = 0, which gives us E(X), which is p. Therefore, the mean of X is p. Next, to find the variance, we need to compute the second moment. To do this, we differentiate the moment-generating function a second time.

The second derivative of the MGF is also p * e^t. When we evaluate this second derivative at t = 0, we again get p. Since the second moment is equal to p, we can calculate the variance of X using the formula for variance, Var(X) = E(X^2) - $(E(X))^2$. Substituting the values for the mean and second moment, we find that the variance of X is p * (1 - p). Thus, the mean of X is p, and the variance of X is p * (1 - p).



So, we found that the mean, denoted as μ_1, is equal to p. Now, we are finding the variance of X. The variance can be calculated as E(X) - μ_1^2. We know that the simplified formula for variance is $\mu_2 - \mu_1^2$. From earlier, we determined that $\mu_2 = p$.

Therefore, the variance of X is $p - p^2$, which simplifies to p * (1 - p), or p * q (where q = 1 - p). In this case, you can observe that for any r-th order moment, the derivative remains the same. If you take the r-th derivative of the moment-generating function, it will always result in p * e^t. This means that for any r-th order moment, the r-th derivative of the MGF evaluated at t = 0 will give p. This is an example showing that for a binomial random variable, we can find the moments using the moment-generating function, and the same principles apply for any order of moments.