

Engineering Statistics
Professor Manjesh Hanawal
Industrial Engineering and Operations Research
Indian Institute of Technology, Bombay
Week 3
Lecture no. 14
Moment Generating Functions

(Refer Slide Time: 00:18)

Previous Lecture:

- ▶ Joint distribution of Random Variable
- ▶ Marginal PMF and PDF
- ▶ Independence of Random Variables
- ▶ Correlation of Random Variables

This Lecture:

- ▶ Joint distribution of function of RVs
- ▶ Moment Generating Functions (MGFs)
- ▶ Conditional PMF and PDF
- ▶ Markov's and Chebyshev's inequalities
- ▶ Limit theorems: Law of Large Numbers (LLN)
- ▶ Limit theorems: Central Limit Theorem (CLT)

So, in the previous lectures, we have covered about joint distribution of random variables marginals, marginals of PMF, PDF, independence of random variables correlation of random variables. I think we also discussed joint distributions of functions of random variables. And we stopped at moment generating functions. Any questions so far on whatever we discussed so far? And where are you people in IE621?

Student: Conditional Expectations

Professor: Conditional expectations. I think and all these are law of large numbers, central limit theorem reached and momentum generating functions.

Student: Yes

Professor: So, we have conditional probability mass function in PDF, but I will be going through them very quickly today. So, this is the last lecture on the necessary probability we need and whatever the 2 theorems we are going to discuss today, law of large numbers and central limit theorem here they are the basically bridge between our probability and statistics.

(Refer Slide Time: 01:38)

Moment Generating Functions

Moment Generating Function (MGF) of a random variable X , denoted $\phi_X : \mathcal{R} \rightarrow \mathcal{R}_+$, is defined as

$$\phi_X(t) = \mathbb{E}(e^{tX}) = \begin{cases} \sum_{i=1}^{\infty} P_X(x_i) e^{tx_i} & \text{if } X \text{ is discrete} \\ \int_{\mathcal{R}} f_X(x) e^{tx} dx & \text{if } X \text{ is continuous.} \end{cases}$$

$Y = e^{tX} \quad E[Y]$

So, moment generating functions, what is abbreviated as MGFs is a function phi which goes from real values to positive real numbers and how it is defined? $\phi(X)$ of a random variable X at point t is defined as $E(e^{tX})$. So, X is a random variable and now, we are going to take expectation of this. So, what you are basically doing is we are defining a new random variable Y which is e^{tX} . And now, we are trying to find expectation of Y . And this is we are defining for every t that is why the argument of this function is the entire real line. And now, this e^{tX} this is always going to be positive.

So, this random variable Y I have defined this is a positive valued random variable. Even if my X takes negative values, Y will always take positive value. So, that is why the range is always going to be positive value. Now, its expectation you can just see that and if you recall, we had something called LOCUS

Law of unconscious statisticians. If you just apply this, I really do not need to go and find out what is the PMF or PDF of Y , from PMF or PDF of X . All I just need to do is compute these probabilities and use this function. Then we will directly get this moment generating function.

So, to find this moment and generating function that said like if I to find a moment generating function of X , if it is discrete, just use this X value to compute this expectation and if it is continuous, just use PDF of X to compute this X this expectation.

(Refer Slide Time: 03:46)

Moment Generating Functions

Moment Generating Function (MGF) of a random variable X , denoted $\phi_X : \mathcal{R} \rightarrow \mathcal{R}_+$, is defined as

$$\phi_X(t) = \mathbb{E}(e^{tX}) = \begin{cases} \sum_{i=1}^{\infty} P_X(x_i) e^{tx_i} & \text{if } X \text{ is discrete} \\ \int_{\mathcal{X}} f_X(x) e^{tx} dx & \text{if } X \text{ is continuous.} \end{cases}$$

$\phi_X'(t) = \mathbb{E}(Xe^{tX}) \rightarrow \phi_X'(0) = \mathbb{E}(X)$ (first moment)
 $\phi_X''(t) = \mathbb{E}(X^2 e^{tX}) \rightarrow \phi_X''(0) = \mathbb{E}(X^2)$ (second moment)
 $\phi_X^{(n)}(t) = \mathbb{E}(X^n e^{tX}) \rightarrow \phi_X^{(n)}(0) = \mathbb{E}(X^n)$ (nth moment)

From MGF of a RV all its moment can be generated, specifically, mean and variance.

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

$\phi_X(t) = \int_{\mathcal{X}} f_X(x) e^{tx} dx$
 $\frac{d}{dt} \phi_X(t) = \frac{d}{dt} \int_{\mathcal{X}} f_X(x) e^{tx} dx$
 $= \int_{\mathcal{X}} \frac{d}{dt} f_X(x) e^{tx} dx$
 $= \int_{\mathcal{X}} f_X(x) x e^{tx} dx$
 $\frac{d}{dt} \phi_X(t) \Big|_{t=0} = \mathbb{E}[X e^{tX}]_{t=0}$

Now, why this moment generating functions are useful? Now, look into this for time being that does not matter, let us assume so, notice that if you look into this $\phi(X)$, is this function is differentiable in t , so $\phi(X)$ is taking t as an argument. Can this function be differentiated at t ? Let us say we can differentiate. So, what is this, for time being just focus on the continuous case.

So, now $\phi(X)$ of t is f of $X \times dx$ of, sorry e to the power $t \times dx$ and this is entire real line. Now, what I am asking is d by dt of ϕ of X at t ? And now I need to, now what I am doing here? First I am doing the integration operation and then I am doing the differentiation operations. Can this integration and differentiation operation can be swapped? So, instead of this I want to do this now, I want to do differentiation first d/dx , are they still the same if I do this? Yes. All the time.

Student: (()) (5:45)

Professor: When this is possible, it need to be you said somebody say uniformly continuous, what is that? Anyway check this when this is possible.

Student: (()) (6:02)

Professor: They are integrating x , we are getting integrating over x that is fine, but is still always possible to chain this integration and differentiation. It is in general not true, but it works under some conditions. You can look for those conditions. Now, if you do this here this

becomes simpler only thing that is differentiable since, I have to differentiate with respect to t this is simply going to be $t \cdot e^{tx}$ dx I know if I do this, this is going to be t times, what is this?

Now, t has come out.

Student: It should be x .

Professor: Right. If you look into this expression one thing you will notice is had this term not been there e^{tx} . This is exactly equal to expectation of x and when this term will not be there? One possibility is when t equals to 0. So, if I take do this d by $dt \phi_x(t)$ and compute it at t equals to 0, I get expectation of x . So, like this if you do it, take that second derivation you will end up with this expression and now, see like this one here what I basically did here this I can simply write as this is nothing but $E(Xe^{tX})$.

Now, simply now, if I so, that is why this is X to the power t X . Now, if you do a double differentiation you will get and again if you put t equals to 0, you will get expectation of X squared and similarly if you keep on doing take the n th derivative, you will get this and when you compute that value t equals to 0, you will get this. So, now notice that you got expectation of X , expectation of X square and expectation of X power n . These are called basically the moments. Expectation of X is the first moment which we also call as mean and this is the second moment and like that and when you do it n th this is called a n th moment

So, if you somehow know the moment generating function of a random variable, basically you can generate all the moments. Just you need to do if you need to find the n th moment you need to differentiate that many times. And after that plugin t equals to 0 in the n th derivative. This is true. I am doing this another assumption that MGF is differentiable at a point t equal. And now if I know the first moment and second moment I know the relation for computing the variance, the variance of X is nothing but second moment minus square of the first moment, you can even go and collect the variance in this case.

(Refer Slide Time: 10:12)

Properties of MGF

- MGF of sum of independent RVs: (X_1, X_2, \dots, X_n) are independent, then

$$\phi_Y(t) = \phi_{X_1+X_2+\dots+X_n}(t) = \prod_{i=1}^n \phi_{X_i}(t)$$

$$E[e^{t(X_1+X_2+\dots+X_n)}] = E[e^{tX_1} e^{tX_2} \dots e^{tX_n}] = \prod_{i=1}^n E[e^{tX_i}] = \prod_{i=1}^n \phi_{X_i}(t)$$

- MGF uniquely determines the distributions one-to-one correspondence between the MGF and distribution.

So, another thing why moment generating functions are used is with them operations becomes many times useful. For example, let us say if you have this bunch of random variables, and I say that they are independent. Now in that case, and I will be interested in this new random variable, which is basically the sum of all these random variables and I am interested in finding $\phi_Y(t)$. So, now, that is why it is nothing but summation and by definition, before I write this up by definition, this quantity is nothing but expectation of $e^{t(X_1+X_2+\dots+X_n)}$ which is nothing but expectation of $e^{tX_1} \cdot e^{tX_2} \dots e^{tX_n}$.

And now, if this X_1, X_2 are independent, can I write the expectation of the product as product of the expectations of each owner this term? Now, we know that this is nothing but this product i equals to 1 to n expectation of e^{tX_i} and by definition, this is nothing but $\phi_{X_i}(t)$ and that is what we obtain. So, when I have this bunch of random variable, and if they are independent, if I know the moment generating function of each one of them, then I can easily compute the moment generating function of their sum. All I need to do is take the product.

And another good thing about MGF is moment generating functions uniquely determine the distribution. If you give a distribution and if I find out its moments generating function, there is a one to one map between them.

(Refer Slide Time: 12:27)

Properties of MGF

- MGF of sum of independent RVs: (X_1, X_2, \dots, X_n) are independent, then

$$\phi_{X_1+X_2+\dots+X_n}(t) = \prod_{i=1}^n \phi_{X_i}(t)$$
- MGF uniquely determines the distributions
one-to-one correspondence between the MGF and distribution.

Distribution	MGF $\phi(t)$	Distribution	MGF $\phi(t)$
Ber(p)	$pe^t + (1-p)$	Uni(a, b)	$\frac{e^{bt} - e^{at}}{t(b-a)}$
Bin(n, p)	$(pe^t + (1-p))^n$	Exp(λ)	$\frac{\lambda}{\lambda-t}$
Geo(n, p)	$\frac{pe^t}{1-(1-p)e^t}$	N(μ, σ^2)	$\exp\{\mu t + \frac{t^2 \sigma^2}{2}\}$
Poi(λ)	$e^{\lambda(e^t-1)}$	Gamma(n, λ)	$(\frac{\lambda}{\lambda-t})^n$

Handwritten notes: $E[X^n] < \infty$ for some n, $E[X^n] < \infty \forall j \leq n$, $\phi_X(t) = (0.3e^t + 0.7)^{10}$, $X \sim \text{Bin}(10, 0.3)$, $\phi_X(t) = e^{0.7t + \frac{t^2}{2}}$, $X \sim N(0.7, 1)$

Characteristic function: $\Phi_X(t) = \mathbb{E}(e^{itX})$, where $j = \sqrt{-1}$ (always exists)

Moment Generating Functions

Moment Generating Function (MGF) of a random variable X , denoted $\phi_X : \mathcal{R} \rightarrow \mathcal{R}_+$, is defined as

$$\phi_X(t) = \mathbb{E}(e^{tX}) = \begin{cases} \sum_{i=1}^{\infty} P_X(x_i) e^{tx_i} & \text{if } X \text{ is discrete} \\ \int_{\mathcal{X}} f_X(x) e^{tx} dx & \text{if } X \text{ is continuous.} \end{cases}$$

Handwritten notes: $\phi_X(t) = \int_{\mathcal{X}} f_X(x) e^{tx} dx$, $\frac{d}{dt} \phi_X(t) = \frac{d}{dt} \int_{\mathcal{X}} f_X(x) e^{tx} dx = \int_{\mathcal{X}} \frac{d}{dt} f_X(x) e^{tx} dx = \int_{\mathcal{X}} f_X(x) x e^{tx} dx = \mathbb{E}[X e^{tX}]$

- $\phi_X'(t) = \mathbb{E}(X e^{tX}) \rightarrow \phi_X'(0) = \mathbb{E}(X)$ (first moment)
- $\phi_X''(t) = \mathbb{E}(X^2 e^{tX}) \rightarrow \phi_X''(0) = \mathbb{E}(X^2)$ (second moment)
- $\phi_X^{(n)}(t) = \mathbb{E}(X^n e^{tX}) \rightarrow \phi_X^{(n)}(0) = \mathbb{E}(X^n)$ (nth moment)

From MGF of a RV all its moment can be generated, specifically, mean and variance.

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

So, here is I have directly put a table here computing moment generating functions of some of the distributions that we have studied. The moment generating functions of a Bernoulli random variable with probability p or a parameter p is this quantity and binomial is going to be like this. And geometric can be computed to be like this, like this Poisson are also computed and for the distributions like uniform, exponential, Gaussian and Gamma also computed and written here.

Now suppose let us say I give you a moment generating function which is of this form. This is what, this means this X is definitely has to be binomial. And let us take another example if I have another moment generating function which I said to be like this so, this has to be this is this distribution first normal 0.7 and 1 variance is going to be 1. So, like this, if I tell you any

moment generating functions, you should be able to uniquely map it to that distribution. And there is one more notion of characteristic function which is very analogous to moment generating function, but that involves this complex number instead of simply taking e^{tX} , we will take it as e^{jtX} if our random variable has some complex value in that maybe we can handle that.

And what is j here? j is the Euler's number. And by definition or like as a property, this characteristic function always exist. So, once we make it complex valued, there is something we can not present in real, we can always present in a complex domain. So, that is why this the good thing about character function is this always exist. And sometime moment generating functions may not exist. And as a remark like if we know that, let us say for some n , what does this mean?

Student: n th moment is finite

Professor: So, the n th moment of a random variable is finite. That means its n th moment exist. What does this mean is, if this implies I mean, this needs a proof, but I am just giving as a remark, this is going to be done, this implies that X^j is going to be less than infinity for all j less than or equals to n . So, if n th moment is finite, which means if exist all the lower moments also exist. But with this information, I cannot claim that expectation of X_{n+1} is finite. I cannot claim this. This does not imply. So, more on the momentum generating functions we will see in the tutorial session, and we just you have to remember this. Like the only thing is discrete you have to use the appropriate summation while a formula like and if it is continuous, you have to use the integration, use the PDF there.