2. Moment Generating Function Technique (MGF)

Monents

What is the Moments in Probability/ Statistics?

Let's say the random variable we are interested in is **X**.

The moments are the expected values of X, e.g., E(X), E(X^2), E(X^3), … etc.

The first moment is **E(X)**,

The second moment is **E(X^2)**,

The third moment is **E(X^3)**,

…

The n-th moment is **E(X^n)**.

We are pretty familiar with the first two moments, the mean $\mu = E(X)$ and the variance $E(X^2) - \mu^2$. They are important characteristics of **X**. The mean is the average value and the variance is how spread out the distribution is. But there must be **other features as well** that also define the distribution. For example, the third moment is about the asymmetry of a distribution. The fourth moment is about how heavy its tails are.

Moment	Uncentered	Centered
1st	$E(X) = \mu$	
2nd	$E(X^2)$	$E((X-\mu)^2)$
3rd	$E(X^3)$	$E((X-\mu)^3)$
4th	$E(X^4)$	$E((X-\mu)^4)$

Mean (X)	=	$E(X)$
Var (X)	=	$E((X-\mu)^2) = \sigma^2$
Skewness (X)	=	$E((X-\mu)^3) / \sigma^3$
Kurtosis (X)	=	$E((X-\mu)^4) / \sigma^4$

The **moments** tell you about the **features** of the distribution.

Therefore,

- A moment is a quantitative measure of the shape of a set of points.
- The first moment,r=1 is called the **MEAN** which describes the center of the distribution.
- The second moment,r=2 is the **VARIANCE** which describes the spread of the observations around the center.
- Third moment,r=3 describe other aspects of a distribution such as how the distribution is skewed from its mean or peaked.
- A moment designates the power to which deviations are raised before averaging them.

2. Then what is Moment Generating Function (MGF)?

As its name hints, MGF is literally the function that generates the moments — $E(X), E(X^2), E(X^3), \ldots, E(X^n)$.

$$MGF_X(t) := E[e^{tX}] = \begin{cases} \sum_x e^{tx} \cdot P(x) & X: \text{discrete (PMF)} \\ \int_x e^{tx} \cdot f(x)\, dx & X: \text{continuous (PDF)} \end{cases}$$

The definition of Moment-generating function

If you look at the definition of MGF, you might say…

> *"I'm not interested in knowing E(e^tx). I want E(X^n)."*

Take a derivative of MGF n times and plug t = 0 in. Then, you will get E(X^n).

$$E(X^n) = \frac{d^n}{dt^n} MGF_X(t) \Big|_{t=0}$$

($E(X^n)$: n-th moment)

e.g.

$$E(X) = \frac{d}{dt} MGF_X(t) \Big|_{t=0} = MGF'_X(0)$$

$$E(X^2) = \frac{d^2}{dt^2} MGF_X(t) \Big|_{t=0} = MGF''_X(0)$$

$$\vdots$$

This is how you get the moments from the MGF.

3. Show me the proof. Why is the n-th moment the n-th derivative of MGF?

We'll use Taylor series to prove this.

① $$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$$

then,

$$e^{tx} = 1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots + \frac{(tx)^n}{n!}$$

Then take the expected value.

② $$E(e^{tx}) = E\left(1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots + \frac{(tx)^n}{n!}\right)$$

$$= E(1) + tE(x) + \frac{t^2}{2!}E(x^2) + \frac{t^3}{3!}E(x^3) + \cdots + \frac{t^n}{n!}E(x^n)$$

Now, take a derivative with respect to **t**.

③ $$\frac{d}{dt}E(e^{tx}) = \frac{d}{dt}\left(E(1) + tE(x) + \frac{t^2}{2!}E(x^2) + \frac{t^3}{3!}E(x^3) + \cdots + \frac{t^n}{n!}E(x^n)\right)$$

plug $t=0$ in...

$$= 0 + E(x) + 0 + 0 + \cdots + 0$$

$$= E(x) \quad ■$$

If you take another derivative on ③ (therefore total twice), you will get **E(X²)**.

If you take another (the third) derivative, you will get $\mathbf{E(X^3)}$, and so on and so on…

but we can calculate moments using the definition of expected values. Why do we need MGF exactly?

Moments and Moment Generating Function Technique

A population can be identified through the complete sequence of its moments. If all the moments exist then there exists a function into which they are all summarized. If such a function exists then any required moment can be extracted.

Definition:-The K^{th} moment of a random variable *X* taken about the origin is defined as $E(X^r)$ and denoted by $\boldsymbol{M_k'}$ or μ_k'(if X is representing a population variable) or denoted by $\boldsymbol{m_k'}$(if X is representing a sample variable) .

Definition:- Let X be a random variable for which the mathematical expectation $E(e^{tx})$existis for every value of t in some interval$-\delta < t < \delta$, then $E(e^{tx})$ is called the moment generating function (m.g.f.) of X or equivalently m.g.f. for the distribution function of X and generally denoted by $M_X(t)$, i.e, m.g.f. $M_X(t) = E(e^{tx})$

Theorem(1):- Let the mathematical expectation corresponding to the function $g(x)=e^{tx}$ exists for every t, $-\delta < t < \delta$ for the random variable X, if $M_X(t)$ stands for mathematical expectation then

$$\mu_r' = \left[\frac{d^r M_X(t)}{dt^r}\right]_{t=0} \quad , \qquad r = 1,2,\dots$$

Proof:- Since $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^r}{r!} + \cdots$

$$e^{tx} = 1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \cdots + \frac{t^rx^r}{r!} + \cdots$$

and $M_X(t) = E(e^{tx}) = 1 + tE(x) + \frac{t^2}{2!}E(x^2) + \frac{t^3}{3!}E(x^3) + \cdots + \frac{t^r}{r!}E(x^r) + \cdots$

$$M_X(t) = E\left[\sum_{r=0}^{\infty} \frac{t^r x^r}{r!}\right] \quad \text{If X is discrete}$$

$$M_X(t) = E\left[\int_0^{\infty} \frac{t^r x^r}{r!}\, dx\right] \quad \text{If X is continuous}$$

For continuous random variable,

$$M_X(t) = 1 + \int_0^{\infty} tx\, dx + \int_0^{\infty} \frac{t^2}{2!}x^2\, dx + \int_0^{\infty} \frac{t^3}{3!}x^3\, dx + \cdots + \int_0^{\infty} \frac{t^r}{r!}x^r\, dx + \cdots$$

$$= 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \frac{t^3}{3!}\mu_3' + \cdots + \frac{t^r}{r!}\mu_r' + \cdots$$

$\left[\frac{dM_X(t)}{dt}\right]_{t=0} = \mu_1'$, $\left[\frac{d^2M_X(t)}{dt^2}\right]_{t=0} = \mu_2'$, …. , $\left[\frac{d^rM_X(t)}{dt^r}\right]_{t=0} = \mu_r'$, $r = 1,2,\dots$

Note:-

1- From the theory of mathematical analysis, it can be shown that the existence of $M_X(t)$ for $-\delta < t < \delta$ implies that derivatives of $M_X(t)$ of all orders exists at $t = 0$ (or summation of $m_X(t)$ for discrete type).

2- m.g.f. if exists is an effective method of finding the distribution function of several random variables.

Example:- Show that m.g.f. for binomial distribution is given by $m_X(t) = [1 + p(e^t - 1)]^n$
And find the mean and the variance for X.

Example:- Show that the m.g.f. for normal distribution with parameters μ and σ^2 is given by

$$M_X(t) = e^{\mu t + \frac{t^2\sigma^2}{2}}$$

Proof:

$$M(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx = \int_{-\infty}^{\infty} e^{tx}\left[\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}\right]dx$$

Simply expanding the term in the second exponent, we get:

$$M(t) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\exp\{tx\}\exp\left\{-\frac{1}{2\sigma^2}(x^2 - 2x\mu + \mu^2)\right\}dx$$

And, combining the two exponents, we get:

$$M(t) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}(x^2 - 2x\mu + \mu^2) + tx\right\}dx$$

Pulling the tx term into the parentheses in the exponent, we get:

$$M(t) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}(x^2 - 2x\mu - 2\sigma^2 tx + \mu^2)\right\}dx$$

And, simplifying just a bit more in the exponent, we get:

And, simplifying just a bit more in the exponent, we get:

$$M(t) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}(x^2 - 2x(\mu + \sigma^2 t) + \mu^2)\right\}dx$$

And, simplifying just a bit more in the exponent, we get:

Now, let's take a little bit of an aside by focusing our attention on just this part of the exponent:

$$(x^2 - 2(\mu + \sigma^2 t)x + \mu^2)$$

If we let:

$$a = \mu + \sigma^2 t \text{ and } b = \mu^2$$

then that part of our exponent becomes:

$$x^2 - 2(\mu + \sigma^2 t)x + \mu^2 = x^2 - 2ax + b$$

Now, complete the square by effectively adding 0:

$$x^2 - 2(\mu + \sigma^2 t)x + \mu^2 = x^2 - 2ax + a^2 - a^2 + b$$

And, simplifying, we get:

That is, because the integral is 1:

$$M(t) = \exp\left\{-\frac{1}{2\sigma^2}\left[-(\mu+\sigma^2 t)^2 + \mu^2\right]\right\}\left\{\int_{-\infty}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}\left[\left(x-(\mu+\sigma^2 t)\right)^2\right]\right\}dx\right.$$

$= 1$ since $N(\mu+\sigma^2 t, \sigma^2)$

our moment-generating function reduces to this:

$$M(t) = \exp\left\{-\frac{1}{2\sigma^2}\left[-\mu^2 - 2\mu\sigma^2 t - \sigma^4 t^2 + \mu^2\right]\right\}$$

Now, it's just a matter of simplifying:

$$M(t) = \exp\left\{\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2}\right\}$$

and simplifying a bit more:

$$M(t) = \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}$$

Our second messy proof is complete!

EX:-Show that the m.g.f. for Poisson distribution($i.e.,\ X\sim P(\lambda)$) is $M_X(t) = e^{\lambda(e^t-1)}$ and for Gamma distribution ($i.e.,\ X\sim G(\alpha,\beta)$) is $M_X(t) = (1-\beta t)^{-\alpha}$

Note:- Chi- square distribution is gamma distribution with $\alpha = \frac{r}{2}\ and\ \beta = 2$, so that the m.g.f. is defined as $M_X(t) = (1-2t)^{-\frac{r}{2}}$

To find the distribution function for the random variables $X_1, X_2, \dots, X_n$ from the distribution function for random variable we can use the following theorems:

Example: The normal distribution: Suppose $X_1, \dots, X_p$ independent, $X_i \sim N(\mu_i, \sigma_i^2)$ so that

$$\begin{aligned} M_{X_i}(t) &= \int_{-\infty}^{\infty} e^{tx} e^{-\frac{1}{2}(x-\mu_i)^2/\sigma_i^2} dx/(\sqrt{2\pi}\sigma_i) \\ &= \int_{-\infty}^{\infty} e^{t(\sigma_i z+\mu_i)} e^{-z^2/2} dz/\sqrt{2\pi} \\ &= e^{t\mu_i} \int_{-\infty}^{\infty} e^{-(z-t\sigma_i)^2/2+t^2\sigma_i^2/2} dz/\sqrt{2\pi} \\ &= e^{\sigma_i^2 t^2/2+t\mu_i}. \end{aligned}$$

Theorem: (Uniqueness theorem)

Let $M_{X_1}(t)$ and $M_{X_2}(t)$ be the m.g.f. for two random variables $X_1\ and\ X_2$ respectively, if both m.g.f.s exists and $M_{X_1}(t) = M_{X_2}(t)$ for all values of t, then $X_1\ and\ X_2$ have the same distribution function ($i.e., f_1(X_1) = f_2(X_2)$) whenever $X_1 = X_2$

Example:- Let $X\sim N(\mu,\sigma^2)$. Show that $Z=\frac{X-\mu}{\sigma}\sim N(0,1)$ using m.g.f. technique.

$$M_Z(t)=E(e^{tz})=E\left(e^{t(\frac{X-\mu}{\sigma})}\right)=E\left(e^{\frac{t}{\sigma}}e^{(x-\mu)}\right)=M_{(X-\mu)}\left(\frac{t}{\sigma}\right)=e^{\left(\frac{t}{\sigma}\right)^2(\frac{\sigma^2}{2})}=e^{\frac{t^2}{2}}$$

$\Rightarrow\ Z\sim N(0,1)$

Example:- Let $Z\sim N(0,1)$. Find the distribution function for Z^2 using m.g.f. technique.

$$M_{Z^2}(t)=E\left(e^{tz^2}\right)=\int_{-\infty}^{\infty}e^{tz^2}f(z)dz=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{tz^2}e^{-\frac{z^2}{2}}dz=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}(1-2t)}dz$$

multiply the numerator and denominator by $(1-2t)^{-\frac{1}{2}}$

$$\Rightarrow\ M_{Z^2}(t)=\frac{1}{(1-2t)^{\frac{1}{2}}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}(1-2t)^{-\frac{1}{2}}}e^{-\frac{z^2}{2}(1-2t)}dz$$

$$\Rightarrow\ M_{Z^2}(t)=\frac{1}{(1-2t)^{\frac{1}{2}}}$$

which is m.f.g. of gamma distribution with $\alpha=\frac{1}{2}\ \ and\ \ \beta=2$

So, we conclude that

if$Z\sim N(0,1)$ then $Z^2\sim G(\frac{1}{2}.2)$ and this means that $Z^2\sim\chi^2_{(1)}$

Note:

One of the most useful facts about moment generating functions is that the moment generating function of a sum of independent variables is the product of the individual moment generating functions.

Theorem (2):- Let $X_1,X_2,\dots,X_n$ be independent random variables with m.g.f.s $M_{X_1}(t),M_{X_2}(t),\dots,M_{X_n}(t)$, if $u=x_1+x_2+\cdots+x_n$ then

$$M_u(t)=M_{X_1}(t)\,.\ M_{X_2}(t).\dots.M_{X_n}(t)=\prod_{i=1}^{n}M_{X_i}(t)$$

Proof:- EX

Example:- using m.g.f. technique to show that the binomial distribution with parameters p and n approaches to the poisson distribution with parameter $\lambda\ as\ \ n\to\infty$ and $p\to 0$ while $\lambda=np$ remains fixed.

Solution:-m.g.f. for $X\sim B(n,p)$ is $m_X(t)=[1+p(e^t-1)]^n$

Take the limit as $p\to 0$, $\lim\limits_{p\to 0}m_X(t)=\lim\limits_{p\to 0}[1+p(e^t-1)]^n$

Since $\lambda=np$ then $=\frac{\lambda}{p}$,

$$\lim_{p\to 0}m_X(t)=\lim_{p\to 0}[1+p(e^t-1)]^{\frac{\lambda}{p}}=\lim_{p\to 0}[1+p(e^t-1)]^{\frac{1}{p(e^t-1)}*\lambda(e^t-1)}$$

$\lim_{p\to 0} m_X(t) = e^{\lambda(e^t-1)}$[since$e = \lim_{t\to 0}(1+t)^{\frac{1}{t}}$]; Which is the m.g.f. for Poisson

Theorem (3):- Let $X_1, X_2, \dots, X_n$ be stochastically independent random variables that have respectively the Chi-square distribution $\chi^2_{(r_1)}, \chi^2_{(r_2)}, \dots, \chi^2_{(r_n)}$, then the random variable $Y = X_1 + X_2 + \cdots + X_n$ has a Chi-square distribution with $r_1 + r_2 + \cdots + r_n$ degree of freedom i.e., $Y \sim \chi^2_{(r_1+r_2+\cdots+r_n)}$

Proof:- If $X \sim \chi^2_{(r)}$ then $M_X(t) = (1-2t)^{-\frac{r}{2}}$

$M_Y(t) = E\left(e^{t(x_1+x_2+\cdots+x_n)}\right)$,but$X_i$ are independent $\Rightarrow M_Y(t) = E(e^{tx_1})E(e^{tx_2}) \dots E(e^{tx_n}) \quad \Rightarrow \quad M_Y(t) = (1-2t)^{-\frac{(r_1+r_2+\cdots+r_n)}{2}}, \quad t < \frac{1}{2}$

Exercises:

Solve the problems below using the moment-generating-function technique. Make sure to state the distribution and its parameters.

1. Let $X_1,\dots,X_n$ be independent random variables, such that $X_i \sim$ Exponential(θ), for $i = 1,\dots,n$. Find the distribution of: $Y = X_1 + \cdots + X_n$.
2. Let $X_1,\dots,X_n$ be independent random variables, such that $X_i \sim$ Poiss(λ_i), for $i = 1,\dots,n$. Find the distribution of: $Y = X_1 + \cdots + X_n$.
3. Let $X_1,\dots,X_n$ be independent random variables, such that $X_i \sim N(\mu_i, \sigma_i^2)$, for $i = 1,\dots,n$. Find the distribution of: $Y = a_1X_1 + \cdots + a_nX_n$..