Lesson 18: Moment Generating Functions

Meike Niederhausen and Nicky Wakim

2025-12-03

Learning Objectives

1. Learn the definition of a moment-generating function.
2. Find the moment-generating function of a random variable.
3. Use a moment-generating function to find the mean and variance of a random variable.

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

Where are we?

Basics of probability

- Outcomes and events
- Sample space
- Simulations
- Probability axioms
- Probability properties
- Counting
- Rules of probability

Probability for random variables

- Probability functions: pmfs/pdfs
- Cumulative distribution functions (CDFs)
- Important distributions

Two random variables

- Transformations
- Joint distributions
- Independence and conditioning

Expected values

- Means / expected values
- Linear combinations and conditioning
- Variance

Advanced probability

- Central limit theorem
- Functions: moment generating functions

What are moments?

Definition 1

The j^{th} moment of a r.v. X is $\mathbb{E}[X^j]$

Okay, but what are they?

Gamma(α, β)

Example 1

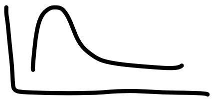
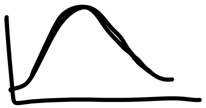
$1^{st} - 4^{th}$ moments   jth moment : $E(X^j)$

1. 1st moment: $E(X)$ mean

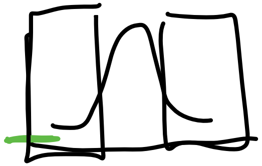
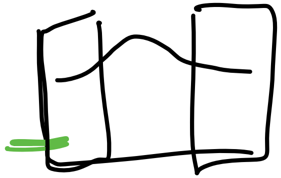
2nd mom - 1st mom^2

2. 2nd moment: $E(X^2)$ → variance : $E(X^2) - [E(X)]^2$

3. 3rd moment: $E(X^3)$ → skewness

 vs 

4. 4th moment: $E(X^4)$ → Kurtosis

 vs 

$\hookrightarrow E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$ OR $\sum_{\text{all } x} g(x)\,p(x)$

What is a *moment generating function* (MGF)??

Definition 3

If X is a r.v., then the **moment generating function (MGF)** associated with X is:

$$M_X(t) = \mathbb{E}[\underbrace{e^{tX}}_{g(x)}]$$

Remarks

- For a discrete r.v., the MGF of X is

$$M_X(t) = \mathbb{E}[e^{tX}] = \sum_{all\ x} e^{tx} p_X(x)$$

- For a continuous r.v., the MGF of X is

$$M_X(t) = \mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx$$

- The MGF $M_X(t)$ is a function of t, not of X, and it might not be defined (i.e. finite) for all values of t. We just need it to be defined for $t = 0$.

Example

Example 4

What is $M_X(t)$ for $t = 0$?

$$M_X(t) = E(e^{tX})$$

$$M_X(t=0) = E(e^{0 \cdot X}) = E(\underbrace{e^0}_{=1})$$

$$= E(1) = 1$$

when $t=0$, MGF is 1 for ALL RVs

How do MGFs give us moments?

Theorem 5

The moment generating function uniquely specifies a probability distribution. AKA **all moments can be found from the MGF through its derivatives at $t = 0$.**

Theorem 6

$$\rightarrow \left[\mathbb{E}[X^r] = M_X^{(r)}(0) \right]$$

(r) in this equation is the rth derivative with respect to t. We calculate the derivative at $t = 0$

M'' M''' $M^{(4)}$

- When $r = 1$, we are taking the first derivative $\rightarrow M'_X(0)$ $M'_X(t=0)$
- When $r = 4$, we are taking the fourth derivative $\rightarrow M_X^{(4)}(0)$

$$Var(X) = E(X^2) - [E(X)]^2 = M''(0) - [M'(0)]^2$$

Using the MGF to uniquely describe a probability distribution

Example 7

Let $X \sim Poisson(\lambda)$

1. Find the MGF of X
2. Find $\mathbb{E}[X]$
3. Find $Var(X)$

$$P_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

① $$M_X(t) = E[e^{tX}] = \sum_{x=0}^{\infty} e^{tx}\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x e^{tx}}{x!}$$

$$= e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda e^t)^x}{x!} = e^{-\lambda}\left(e^{e^t\lambda}\right)$$

$$e^{-\lambda + e^t\lambda}$$

$$M_X(t) = e^{\lambda(e^t-1)}$$

RULE

$$\sum_{n=0}^{\infty}\frac{(z)^n}{n!} = e^z$$

② $$E(X) = M'(t=0)$$

$$M'_X(t) = \frac{d}{dt}e^{\lambda(e^t-1)}$$

$$= e^{\lambda(e^t-1)}\frac{d}{dt}\lambda(e^t-1) = e^{\lambda(e^t-1)}\lambda e^t = \lambda e^t e^{\lambda(e^t-1)}$$

$$\lambda e^t - \lambda$$

$$M'(t=0) = \lambda \underbrace{e^0}_{=1} \underbrace{e^{\lambda(\underbrace{e^0 - 1}_{=0})}}_{=1} = \lambda$$

③ $Var(X) = \underbrace{E(X^2)}_{M''(t=0)} - [\underbrace{E(X)}_{\lambda}]^2$

$$M''_X(t) = \lambda e^{\underbrace{t+\lambda(e^t-1)}_{1}} \underbrace{(1+\lambda e^t)}_{1+\lambda}$$

$$M''(t=0) = \lambda(1+\lambda) = \lambda + \lambda^2$$

$$Var(X) = [\lambda + \lambda^2] - [\lambda]^2$$

$$= \lambda$$

Theorem

Remark: Finding the mean and variance is sometimes easier with the following trick

Theorem 8

Let $R_X(t) = \ln[M_X(t)]$. Then,

$M_X'(t=0)$

$$\mu = \mathbb{E}[X] = R_X'(0), \text{ and}$$

$$\sigma^2 = Var(X) = R_X''(0)$$

@ $t=0$

Proof.

$$R_X'(t) = \frac{d}{dt} \underbrace{\ln(M_X(t))}_{g(x)} = \frac{1}{M_X(t)} M_X'(t) = \frac{M_X'(t)}{M_X(t)}$$

$M_X(t) = 1$ ↳ @ $t=0$

$$R_X'(t=0) = \frac{M_X'(0)}{1} = E(X)$$

Using $R_X(t)$ to uniquely describe a probability distribution

Example 9

Let $X \sim Poisson(\lambda)$.

1. Find $\mathbb{E}[X]$ using $R_X(t)$
2. Find $Var(X)$ using $R_X(t)$

① $R_X(t) = \ln(M_X(t)) = \cancel{\ln}(\cancel{e}^{\lambda e^t - \lambda})$

$R_X(t) = \lambda e^t - \lambda$

$E(X) = R'_X(t=0)$

$R'_X(t) = \lambda \underline{e^t}$ $\quad R'_X(t=0) = \lambda$

$E(X) = \lambda$

$Var(X) = R''_X(t=0)$

② $R''_X(t) = \frac{d}{dt}(\lambda e^t) = \lambda e^t$

$Var(X) = R''_X(t=0) = \lambda$

Using the MGF to uniquely describe the standard normal distribution

pdf of std norm

Example 10

Let Z be a standard normal random variable, i.e. $Z \sim N(0,1)$.

1. Find the MGF of Z
2. Find $\mathbb{E}[Z]$
3. Find $Var(Z)$

① $M_Z(t) = E[e^{tz}] = \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \, dz$

$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tz - z^2/2} \, dz$

$tz - z^2/2 = \dfrac{t^2 - (z-t)^2}{2}$

$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \underbrace{e^{\frac{t^2-(z-t)^2}{2}}}_{e^{t^2/2} e^{-(z-t)^2/2}} \, dz$

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

$u = z - t$
$du = dz$

$= e^{t^2/2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \, du}_{=1}$

$M_Z(t) = e^{t^2/2}$

Using the MGF to uniquely describe the standard normal distribution

② $E(Z) = R_Z'(t=0)$

$R_Z(t) = \ln(M_X(t)) = \cancel{\ln}(\cancel{e}^{t^2/2})$
$= t^2/2$

$R_Z'(t) = \frac{d}{dt}\left(t^2/2\right) = \frac{2t}{2} = t$

$E(Z) = R_Z'(t=0) = 0 \rightarrow$ std normal $\mu = 0$

vs.
$E(Z) = \int_{-\infty}^{\infty} z f_Z(z)\, dz$

③ $Var(Z) = R_Z''(t=0)$

$R_Z''(t) = \frac{d}{dt}(t) = 1$

$Var(Z) = \int_{-\infty}^{\infty} (z-\mu)^2 f_Z(z)\, dz$

$Var(Z) = 1 \rightarrow$ var of std normal $= 1$

MGFs of sums of independent RV's

Theorem 9

If X and Y are independent RV's with respective MGFs $M_X(t)$ and $M_Y(t)$, then

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX}e^{tY}] = E[e^{tX}]E[e^{tY}] = M_X(t)M_Y(t)$$

$$M'_{X+Y}(t) = M'_X(t)\, M'_Y(t)$$

Main takeaways

- MGFs are a purely mathematically definition
  - We can't really relate it to our real world analysis
- They are helpful mathematically because they are unique to a probability distribution
  - We can find the unique MGF from for a probability distribution
  - And we can find a distribution from an MGF
- MGFs can *sometimes* make it easier to find the mean and variance of an RV
- MGFs are most helpful when we are finding a joint distribution that is a sum or transformation of two RV's
  - Make the calculation easier!
- MGFs are often used to prove certain distribution are sums of other ones!

More resources

- https://online.stat.psu.edu/stat414/book/export/html/676
- https://www.youtube.com/watch/ez_vq23xWrQ
- https://www.youtube.com/watch/2p9J9ChTeFI
- https://www.youtube.com/watch/A5bWU8xcQkE
- https://www.youtube.com/watch/QeUrTGFTFm4
- https://www.youtube.com/watch/HhrkwyyRtgI