

STAT 25100 Lecture 19
Section 7.7 Moment Generating Functions (MGF)

Yibi Huang
Department of Statistics
University of Chicago

Moment Generating Function (MGF)

The *moment generating function (MGF)* $M(t)$ of the random variable X is defined to be

$$M(t) = \mathbb{E}(e^{tX}) = \begin{cases} \sum_x e^{tx} p_X(x) & \text{if discrete} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x) dx & \text{if continuous} \end{cases}$$

where $p_X(x)$ and $f_X(x)$ are the PMF/PDF of X .

- ▶ MGF $M(t)$ is NOT a single value but a function of t
- ▶ $M(0) = \mathbb{E}(e^{0X}) = \mathbb{E}[1] = 1$ always

Example: MGF for Geometric

For **Geometric**(p), the PMF is

$$p(x) = (1 - p)^{x-1}p, \quad x = 1, 2, \dots,$$

Its MGF is

$$\begin{aligned} M(t) &= \sum_{x=1}^{\infty} e^{tx} p(x) = \sum_{x=1}^{\infty} e^{tx} (1 - p)^{x-1} p \\ &= p e^t \sum_{x=1}^{\infty} (e^t (1 - p))^{x-1} \\ &= \frac{p e^t}{1 - (1 - p) e^t}, \quad \text{since } \sum_{x=0}^{\infty} r^x = \frac{1}{1 - r}. \end{aligned}$$

The last step is valid only when $(1 - p)e^t < 1$, or $t < -\log(1 - p)$.

Thus the MGF is defined when $(1 - p)e^t < 1$.

Example: MGF for Binomial

For **Binomial**(n, p), the PMF is

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad k = 0, 1, 2, \dots, n.$$

Its MGF is

$$\begin{aligned} M(t) &= \sum_{x=0}^n e^{tx} p(x) = \sum_{x=0}^n e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^n \binom{n}{x} (pe^t)^x (1-p)^{n-x} \\ &= (pe^t + (1-p))^n \quad \text{valid for } -\infty < t < \infty. \end{aligned}$$

The last step comes from the Binomial expansion

$$(a+b)^N = \sum_{x=0}^N \binom{N}{x} a^x b^{N-x}, \quad \text{for } \begin{matrix} a = pe^t \\ b = 1-p \end{matrix} \text{ and } N = n.$$

Example: MGF for Exponential

For Exponential(λ), the PDF is

$$f(x) = \lambda e^{-\lambda x}, \quad 0 \leq x < \infty.$$

Its MGF is

$$\begin{aligned} M(t) &= \int_0^{\infty} e^{tx} f(x) dx = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx \\ &= \int_0^{\infty} \lambda e^{-(\lambda-t)x} dx \\ &= \frac{\lambda}{\lambda - t}, \end{aligned}$$

The integral is finite only when $\lambda - t > 0$.

Thus the MGF is defined only when $-\infty < t < \lambda$.

MGF for Standard Normal $N(0, 1)$

The PDF for $N(0, 1)$ is $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ for $-\infty \leq x < \infty$.

Its MGF is thus

$$M(t) = \int_{-\infty}^{\infty} e^{tx} \phi(x) dx = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\overbrace{-x^2/2+tx}^{\text{see below}}} dx$$

Using the technique of **completing the square**, as

$$-\frac{x^2}{2} + tx = -\frac{1}{2}(x^2 - 2tx + t^2) + \frac{t^2}{2} = -\frac{1}{2}(x - t)^2 + \frac{t^2}{2},$$

the integral equals

$$M(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-t)^2/2+t^2/2} dx = e^{t^2/2} \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-t)^2/2} dx}_{=1} = e^{t^2/2}.$$

The last integral is 1 since it integrates over the PDF of $N(t, 1)$.

The MGF is defined for all $-\infty < t < \infty$.

Moments and Moment Generating Functions

As $M(t) = E(e^{tX})$, its first derivative is

$$M'(t) = \frac{d}{dt} E[e^{tX}] = E \left[\frac{d}{dt} e^{tX} \right] = E[Xe^{tX}].$$

Its second derivative is

$$M''(t) = \frac{d}{dt} M'(t) = \frac{d}{dt} E[Xe^{tX}] = E \left[\frac{d}{dt} Xe^{tX} \right] = E[X^2 e^{tX}].$$

In general, the k th derivative of the MGF is

$$M^{(k)}(t) = \frac{d^k}{dt^k} M(t) = E[X^k e^{tX}].$$

Plugging in $t = 0$, we get

$$M'(0) = E(X), \quad M''(0) = E(X^2), \quad M^{(k)}(0) = E(X^k), \dots,$$

Moment generating functions got the name since the moments of X can be obtained by successively differentiating $M(t)$.

Example: Calculating Moments Using MGF — Exponential

The MGF for Exponential(λ) is

$$M(t) = \frac{\lambda}{\lambda - t}.$$

The derivatives and the moments are thus

$$M'(t) = \frac{\lambda}{(\lambda - t)^2} \Rightarrow E(X) = M'(0) = \frac{1}{\lambda}$$

$$M''(t) = \frac{2\lambda}{(\lambda - t)^3} \Rightarrow E(X^2) = M''(0) = \frac{2}{\lambda^2}$$

\vdots

$$M^{(k)}(t) = \frac{k!\lambda}{(\lambda - t)^{k+1}} \Rightarrow E(X^k) = M^{(k)}(0) = \frac{k!}{\lambda^k}.$$

Example: Calculating Moments Using MGF — $N(0, 1)$

The MGF for $N(0, 1)$ is

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

The derivatives and the moments are thus

$$\begin{aligned}M'(t) &= te^{t^2/2} && \Rightarrow E(X) = M'(0) = 0 \\M''(t) &= e^{t^2/2} + t^2e^{t^2/2} && \Rightarrow E(X^2) = M''(0) = 1 \\M^{(3)}(t) &= 3te^{t^2/2} + t^3e^{t^2/2} && \Rightarrow E(X^3) = M^{(3)}(0) = 0 \\M^{(4)}(t) &= (3 + 3t + 3t^2 + t^3)e^{t^2/2} && \Rightarrow E(X^4) = M^{(4)}(0) = 3\end{aligned}$$

MGF for $a + bX$

If X has the MGF $M_X(t)$ and $Y = a + bX$, then the MGF for Y is

$$\begin{aligned}M_Y(t) &= \mathbf{E}(e^{tY}) \\&= \mathbf{E}(e^{at+btX}) \\&= \mathbf{E}(e^{at}e^{btX}) \\&= e^{at} \mathbf{E}(e^{btX}) \\&= e^{at} M_X(bt)\end{aligned}$$

MGF for $N(\mu, \sigma^2)$

If $X \sim N(0, 1)$, we know in L10 that

$$Y = \mu + \sigma X \sim N(\mu, \sigma^2).$$

As the MGF for X is known to be $M_X(t) = e^{t^2/2}$, we can obtain the MGF for $Y = \mu + \sigma X$ from $M_X(t)$ to be

$$M_Y(t) = e^{\mu t} M_X(\sigma t) = e^{\mu t} e^{\sigma^2 t^2/2} = e^{\mu t + \sigma^2 t^2/2}.$$

Cauchy Distribution Has No MGF

The Cauchy Distribution has the PDF

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty \leq x < \infty.$$

Its MGF would be

$$M(t) = \int_{-\infty}^{\infty} \frac{e^{tx}}{\pi(1+x^2)} dx$$
$$> \begin{cases} \int_0^{\infty} \frac{e^{tx}}{\pi(1+x^2)} dx = \infty & \text{since } \lim_{x \rightarrow \infty} \frac{e^{tx}}{\pi(1+x^2)} = \infty \text{ if } t > 0, \\ \int_{-\infty}^0 \frac{e^{tx}}{\pi(1+x^2)} dx = \infty & \text{since } \lim_{x \rightarrow -\infty} \frac{e^{tx}}{\pi(1+x^2)} = \infty \text{ if } t < 0 \end{cases}$$

Cauchy Distribution Has No MGF

The Cauchy Distribution has the PDF

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty \leq x < \infty.$$

Its MGF would be

$$M(t) = \int_{-\infty}^{\infty} \frac{e^{tx}}{\pi(1+x^2)} dx$$
$$> \begin{cases} \int_0^{\infty} \frac{e^{tx}}{\pi(1+x^2)} dx = \infty & \text{since } \lim_{x \rightarrow \infty} \frac{e^{tx}}{\pi(1+x^2)} = \infty \text{ if } t > 0, \\ \int_{-\infty}^0 \frac{e^{tx}}{\pi(1+x^2)} dx = \infty & \text{since } \lim_{x \rightarrow -\infty} \frac{e^{tx}}{\pi(1+x^2)} = \infty \text{ if } t < 0 \end{cases}$$

Remark

- ▶ If X has an MGF $M(t)$ that exists for t in an open interval containing 0, then all the moments $E(X^k)$ exist.
- ▶ If X doesn't have all the moments $E(X^k)$, then X has a **heavier tail** than those with all the moments.

MGFs for Common Discrete Distributions

Name and range	PMF at k	Mean	Variance	MGF
Bernoulli(p) on $\{0, 1\}$	$\begin{cases} 1-p & \text{if } k=0 \\ p & \text{if } k=1 \end{cases}$	p	$p(1-p)$	$pe^t + 1 - p$
Binomial(n, p) on $\{0, 1, \dots, n\}$	$\binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$	$(pe^t + 1 - p)^n$
Geometric(p) on $\{1, 2, 3, \dots\}$	$(1-p)^{k-1} p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1 - (1-p)e^t}$
Negative Binomial(r, p) on $\{r, r+1, r+2, \dots\}$	$\binom{k-1}{r-1} p^r (1-p)^{k-r}$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\left(\frac{pe^t}{1 - (1-p)e^t} \right)^r$
Poisson(λ) on $\{0, 1, 2, \dots\}$	$e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ	$\exp(\lambda(e^t - 1))$

► MGF for the Hypergeometric distribution exists but is complicated

MGFs for Common Continuous Distributions

Name	PDF $f(x)$	Range	Mean	Variance	MGF $M(t)$
Exponential(λ)	$\lambda e^{-\lambda x}$,	$0 \leq x < \infty$	$1/\lambda$	$1/\lambda^2$	$\frac{\lambda}{\lambda-t}$, $t < \lambda$
Gamma(α, λ)	$\frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$,	$0 \leq x < \infty$	α/λ	α/λ^2	$(\frac{\lambda}{\lambda-t})^\alpha$, $t < \lambda$
Normal(μ, σ^2)	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$,	$-\infty < x < \infty$	μ	σ^2	$\exp(\mu t + \frac{\sigma^2 t^2}{2})$
Cauchy	$\frac{1}{\pi(1+x^2)}$,	$-\infty < x < \infty$	not exist	not exist	does not exist

- MGF for the Beta distribution exists but is complicated.

MGF for Sum of Independent R.V.'s

If X and Y are **independent** r.v.'s with MGF's $M_X(t)$ and $M_Y(t)$, then $M_{X+Y}(t) = M_X(t)M_Y(t)$ on the common interval where both MGF's exist.

Proof.

$$\begin{aligned}M_{X+Y}(t) &= \mathbb{E}(e^{t(X+Y)}) = \mathbb{E}(e^{tX}e^{tY}) \\ &= \mathbb{E}(e^{tX}) \mathbb{E}(e^{tY}) \quad \text{since } X, Y \text{ are indep.} \\ &= M_X(t)M_Y(t)\end{aligned}$$

More generally, if X_1, \dots, X_n are **independent** with corresponding MGF $M_{X_i}(t)$'s, then the MGF for $T = \sum_{i=1}^n X_i$ is

$$M_T(t) = \mathbb{E}[e^{t\sum_{i=1}^n X_i}] = \prod_{i=1}^n \mathbb{E}[e^{tX_i}] = \prod_{i=1}^n M_{X_i}(t),$$

and it exists on the common interval where all MGF's exist.

The MGF Uniquely Determines the Distribution (★★★★★)

If the moment-generating function $M(t)$ exists for t in an open interval containing 0, like $(-t_0, t_0)$, for some $t_0 > 0$, then it **uniquely determines the probability distribution**.

That is, if X and Y have identical MGF

$$M_X(t) = M_Y(t) \quad \text{for all } t \text{ in an open interval containing } 0,$$

then X and Y have the same probability distribution.

The MGF Uniquely Determines the Distribution (★★★★★)

If the moment-generating function $M(t)$ exists for t in an open interval containing 0, like $(-t_0, t_0)$, for some $t_0 > 0$, then it **uniquely determines the probability distribution**.

That is, if X and Y have identical MGF

$$M_X(t) = M_Y(t) \quad \text{for all } t \text{ in an open interval containing } 0,$$

then X and Y have the same probability distribution.

► **Ex1.** If the MGF of X is

$$M(t) = (1/2)^{10}(e^t + 1)^{10} \quad \Rightarrow \quad X \sim \text{Bin}(n = 10, p = 1/2).$$

► **Ex2.** If the MGF of X is $M(t) = \exp(3(e^t - 1))$, what's the distribution of X ?

The MGF Uniquely Determines the Distribution (★★★★★)

If the moment-generating function $M(t)$ exists for t in an open interval containing 0, like $(-t_0, t_0)$, for some $t_0 > 0$, then it **uniquely determines the probability distribution**.

That is, if X and Y have identical MGF

$$M_X(t) = M_Y(t) \quad \text{for all } t \text{ in an open interval containing } 0,$$

then X and Y have the same probability distribution.

► **Ex1.** If the MGF of X is

$$M(t) = (1/2)^{10}(e^t + 1)^{10} \quad \Rightarrow \quad X \sim \text{Bin}(n = 10, p = 1/2).$$

► **Ex2.** If the MGF of X is $M(t) = \exp(3(e^t - 1))$, what's the distribution of X ?
Poisson($\lambda = 3$)

Finding the Distribution of the Sum of Indep. R.V.'s Using MGF

Ex1. $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$ are independent, what's the distribution of $X + Y$?

Finding the Distribution of the Sum of Indep. R.V.'s Using MGF

Ex1. $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$ are independent, what's the distribution of $X + Y$?

Sol. The MGF for $X + Y$ is

$$\begin{aligned}M(t) &= M_X(t)M_Y(t) = \exp\left(\mu_x t + \frac{\sigma_x^2 t^2}{2}\right) \exp\left(\mu_y t + \frac{\sigma_y^2 t^2}{2}\right) \\ &= \exp\left((\mu_x + \mu_y)t + \frac{(\sigma_x^2 + \sigma_y^2)t^2}{2}\right)\end{aligned}$$

which is the MGF for $N(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$, meaning

$$X + Y \sim N(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2).$$

Finding the Distribution of the Sum of Indep. R.V.'s Using MGF

Ex2. What's the distribution of $\sum_{i=1}^n X_i$ for independent $X_i \sim \text{Poisson}(\lambda_i)$?

Sol. The MGF for X_i is $M_{X_i}(t) = \exp(\lambda_i(e^t - 1))$. The MGF for $\sum_{i=1}^n X_i$ is

$$M(t) = \prod_{i=1}^n \exp(\lambda_i(e^t - 1)) = \exp\left((e^t - 1) \sum_{i=1}^n \lambda_i\right),$$

which is the MGF for $\text{Poisson}(\sum_{i=1}^n \lambda_i)$, meaning

$$\sum_{i=1}^n X_i \sim \text{Poisson}\left(\sum_{i=1}^n \lambda_i\right).$$

Sum of i.i.d. Random Variables

If X_1, X_2, \dots, X_n are i.i.d. with MGF $M_X(t)$, the MGF for $\sum_{i=1}^n X_i$ would be

$$M(t) = \mathbb{E}[e^{t \sum_{i=1}^n X_i}] = \prod_{i=1}^n \mathbb{E}[e^{t X_i}] = \prod_{i=1}^n M_X(t) = (M_X(t))^n .$$

Sum of i.i.d. Random Variables

If X_1, X_2, \dots, X_n are i.i.d. with MGF $M_X(t)$, the MGF for $\sum_{i=1}^n X_i$ would be

$$M(t) = \mathbb{E}[e^{t \sum_{i=1}^n X_i}] = \prod_{i=1}^n \mathbb{E}[e^{t X_i}] = \prod_{i=1}^n M_X(t) = (M_X(t))^n .$$

Ex3. If X_1, X_2, \dots, X_n are i.i.d. $\text{Exponential}(\lambda)$, what's the distribution of $\sum_{i=1}^n X_i$?

Sum of i.i.d. Random Variables

If X_1, X_2, \dots, X_n are i.i.d. with MGF $M_X(t)$, the MGF for $\sum_{i=1}^n X_i$ would be

$$M(t) = \mathbb{E}[e^{t \sum_{i=1}^n X_i}] = \prod_{i=1}^n \mathbb{E}[e^{t X_i}] = \prod_{i=1}^n M_X(t) = (M_X(t))^n.$$

Ex3. If X_1, X_2, \dots, X_n are i.i.d. $\text{Exponential}(\lambda)$, what's the distribution of $\sum_{i=1}^n X_i$?

Sol. The MGF for $\text{Exponential}(\lambda)$ is $M_X(t) = \frac{\lambda}{\lambda - t}$.

The MGF for $\sum_{i=1}^n X_i$ would be

$$M(t) = (M_X(t))^n = \left(\frac{\lambda}{\lambda - t} \right)^n,$$

which is the MGF for $\text{Gamma}(\alpha = n, \lambda)$, meaning

$$\sum_{i=1}^n X_i \sim \text{Gamma}(\alpha = n, \lambda).$$

Sum of i.i.d. Random Variables

Ex4. If X_1, X_2, \dots, X_n are i.i.d. Geometric(p), what's the distribution of $\sum_{i=1}^n X_i$?

Sum of i.i.d. Random Variables

Ex4. If X_1, X_2, \dots, X_n are i.i.d. Geometric(p), what's the distribution of $\sum_{i=1}^n X_i$?

Sol. The MGF for Geometric(p) is $M_X(t) = \frac{pe^t}{1 - (1-p)e^t}$.

The MGF for $\sum_{i=1}^n X_i$ would be

$$M(t) = (M_X(t))^n = \left(\frac{pe^t}{1 - (1-p)e^t} \right)^n,$$

which is the MGF for NegBin($r = n, p$), meaning

$$\sum_{i=1}^n X_i \sim \text{NegBin}(r = n, p).$$

Characteristic Functions

- ▶ Drawback of MGF: It may not exist.
- ▶ The *characteristic function* of a random variable X is defined to be

$$\phi(t) = \mathbb{E}(e^{itX}), \text{ where } i = \sqrt{-1} = \text{the imaginary number.}$$

- ▶ $\phi(t)$ always exists since $|e^{it}| = 1$, even for Cauchy distribution.
- ▶ $\phi_X(t) = M_X(it)$ if $M_X(t)$ exists (See next page)
- ▶ $\phi_{a+bX}(t) = e^{ait} \phi_X(bt)$
- ▶ $\phi_{X+Y}(t) = \phi_X(t) \phi_Y(t)$ if X and Y are independent
- ▶ The characteristic function uniquely determines the distribution

Characteristic Functions for Common Distributions

Name and range	PMF at k	MGF	Characteristic Function
Binomial(n, p) on $\{0, 1, \dots, n\}$	$\binom{n}{k} p^k (1-p)^{n-k}$	$(pe^t + 1 - p)^n$	$(pe^{it} + 1 - p)^n$
Geometric(p) on $\{1, 2, 3, \dots\}$	$(1-p)^{k-1} p$	$\frac{pe^t}{1 - (1-p)e^t}$	$\frac{pe^{it}}{1 - (1-p)e^{it}}$
Negative Binomial(r, p) on $\{r, r+1, r+2, \dots\}$	$\binom{k-1}{r-1} p^r (1-p)^{k-r}$	$\left(\frac{pe^t}{1 - (1-p)e^t}\right)^r$	$\left(\frac{pe^{it}}{1 - (1-p)e^{it}}\right)^r$
Poisson(λ) on $\{0, 1, 2, \dots\}$	$e^{-\lambda} \frac{\lambda^k}{k!}$	$\exp(\lambda(e^t - 1))$	$\exp(\lambda(e^{it} - 1))$

Name	PDF $f(x)$	Range	MGF	Characteristic Function
Exponential(λ)	$\lambda e^{-\lambda x}$,	$0 \leq x < \infty$	$\frac{\lambda}{\lambda - t}$	$\frac{\lambda}{\lambda - it}$
Gamma(α, λ)	$\frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$,	$0 \leq x < \infty$	$\left(\frac{\lambda}{\lambda - t}\right)^\alpha$	$\left(\frac{\lambda}{\lambda - it}\right)^\alpha$
Normal(μ, σ^2)	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$,	$-\infty < x < \infty$	$\exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$	$\exp\left(\mu it - \frac{\sigma^2 t^2}{2}\right)$
Cauchy	$\frac{1}{\pi(1+x^2)}$,	$-\infty < x < \infty$	does not exist	$e^{- t }$

Summary of Learning Objects of L19

1. Given a PMF or PDF, find the corresponding MGF

$$M(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} p_X(x) & \text{if discrete} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x) dx & \text{if continuous} \end{cases}$$

2. Using the MGF to find the moments of a random variable

$$M(0) = 1, \quad M'(0) = E(X), \quad M''(0) = E(X^2), \quad M^{(k)}(0) = E(X^k).$$

3. If $M_X(t)$ is the MGF of X , the MGF of $a + bX$ is $M_{a+bX}(t) = e^{at} M_X(bt)$
4. $M_{X+Y}(t) = M_X(t) M_Y(t)$ if X, Y are independent.
5. For $T = \sum_{i=1}^n X_i$, $M_T(t) = \prod_{i=1}^n M_{X_i}(t)$ if X_1, X_2, \dots, X_n are independent.
6. The MGF uniquely determines the distribution (if it exists in an open interval containing 0) One can thus use MGF to identify the distribution,