

Chapter 6B: Moment Generating Functions

MATH/STAT 394: Probability I

Arman Jahangiri

University of Washington
Department of Mathematics

Summer 2026

By the end of this lecture, students should be able to:

- ▶ define the moment generating function (MGF);
- ▶ compute MGFs of common distributions;
- ▶ extract moments from MGFs;
- ▶ understand how MGFs determine distributions;
- ▶ compute distributions of sums using MGFs;
- ▶ derive moments of Exponential and Normal distributions using MGFs.

Definition

The moment generating function (MGF) of X is

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

provided the expectation is finite in some open interval around 0.

Equivalently,

$$M_X(t) = \begin{cases} \sum_x e^{tx} p_X(x), & \text{if } X \text{ is discrete,} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x) dx, & \text{if } X \text{ is continuous.} \end{cases}$$

Quick check

Every valid MGF satisfies

$$M_X(0) = 1.$$

$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_x e^{tx} p_X(x).$$

Bernoulli

If $X \sim \text{Bern}(p)$, then

$$M_X(t) = e^{t \cdot 1} p + e^{t \cdot 0} (1 - p) = pe^t + q.$$

Discrete Uniform

If $X \sim \text{DUnif}\{1, \dots, m\}$, then

$$M_X(t) = \frac{1}{m} \sum_{k=1}^m e^{tk} = \frac{1}{m} (e^t + e^{2t} + \dots + e^{mt}) = \frac{e^t(e^{mt} - 1)}{m(e^t - 1)}.$$

Binomial

If $X \sim \text{Bin}(n, p)$, then

$$M_X(t) = \sum_{k=0}^n e^{tk} \binom{n}{k} p^k q^{n-k} = \sum_{k=0}^n \binom{n}{k} (pe^t)^k q^{n-k} = (q + pe^t)^n.$$

Poisson

If $X \sim \text{Pois}(\lambda)$, then

$$M_X(t) = \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}.$$

Geometric

If $X \sim \text{Geom}(p)$, where X counts failures before the first success, then

$$M_X(t) = \sum_{k=0}^{\infty} e^{tk} q^k p = p \sum_{k=0}^{\infty} (qe^t)^k = \frac{p}{1 - qe^t}, \quad qe^t < 1.$$

Negative Binomial

If $X \sim \text{NegBin}(r, p)$, where X counts failures before the r -th success, then

$$M_X(t) = \sum_{k=0}^{\infty} e^{tk} \binom{k+r-1}{k} p^r q^k = p^r \sum_{k=0}^{\infty} \binom{k+r-1}{k} (qe^t)^k = \left(\frac{p}{1 - qe^t} \right)^r.$$

Hypergeometric

If $X \sim \text{HGeom}(w, b, n)$,

then

$$\mathbb{P}(X = k) = \frac{\binom{w}{k} \binom{b}{n-k}}{\binom{w+b}{n}},$$

where

$$\max(0, n - b) \leq k \leq \min(w, n).$$

Therefore,

$$M_X(t) = \sum_{k=\max(0, n-b)}^{\min(w, n)} e^{tk} \frac{\binom{w}{k} \binom{b}{n-k}}{\binom{w+b}{n}}.$$

Note: Unlike Binomial or Poisson, this MGF usually does not simplify to a nice elementary closed form.

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

Uniform

If $X \sim \text{Unif}(a, b)$, then

$$M_X(t) = \int_a^b e^{tx} \frac{1}{b-a} dx = \frac{1}{b-a} \left[\frac{e^{tx}}{t} \right]_a^b = \frac{e^{tb} - e^{ta}}{t(b-a)}, \quad t \neq 0.$$

Also, $M_X(0) = 1$.

Exponential

If $X \sim \text{Exp}(\lambda)$, then

$$M_X(t) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{-(\lambda-t)x} dx = \frac{\lambda}{\lambda-t}, \quad t < \lambda.$$

Gamma

If $X \sim \text{Gamma}(\alpha, \lambda)$, with shape α and rate λ , then

$$M_X(t) = \int_0^{\infty} e^{tx} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha-1} e^{-(\lambda-t)x} dx.$$

Let $u = (\lambda - t)x$. Then $x = u/(\lambda - t)$, $dx = du/(\lambda - t)$, and

$$M_X(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{1}{(\lambda - t)^\alpha} \underbrace{\int_0^{\infty} u^{\alpha-1} e^{-u} du}_{\Gamma(\alpha)} = \frac{\lambda^\alpha}{(\lambda - t)^\alpha} = \left(\frac{\lambda}{\lambda - t} \right)^\alpha, \quad t < \lambda.$$

Standard Normal

Let $Z \sim N(0, 1)$. Then

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

Complete the square:

$$tz - \frac{z^2}{2} = -\frac{1}{2}(z^2 - 2tz) = -\frac{1}{2}((z - t)^2 - t^2) = -\frac{(z - t)^2}{2} + \frac{t^2}{2}.$$

Therefore,

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

Since the last integral equals $\sqrt{2\pi}$,

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

Normal

Let $X \sim N(\mu, \sigma^2)$. Write

$$X = \mu + \sigma Z, \quad Z \sim N(0, 1).$$

Then

$$M_X(t) = \mathbb{E}(e^{tX}) = \mathbb{E}(e^{t(\mu + \sigma Z)}) = e^{\mu t} \mathbb{E}(e^{\sigma t Z}).$$

Using

$$M_Z(s) = e^{s^2/2},$$

with $s = \sigma t$, we get

$$M_X(t) = e^{\mu t} M_Z(\sigma t) = e^{\mu t} e^{\sigma^2 t^2 / 2} = e^{\mu t + \frac{1}{2} \sigma^2 t^2}.$$

Moments from MGFs

Theorem

The n^{th} population moment is obtained by differentiating the MGF:

$$\mathbb{E}(X^n) = M_X^{(n)}(0).$$

Examples:

$$\mathbb{E}(X) = M_X'(0), \quad \mathbb{E}(X^2) = M_X''(0).$$

Proof: MGFs Generate Moments

Recall that the moment generating function is

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

Differentiate once:

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

Now plugging in $t = 0$,

$$M'_X(0) = \mathbb{E}(Xe^0) = \mathbb{E}(X).$$

Differentiate another time noting that X is a constant w.r.t. t .

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

Plugging in $t = 0$,

$$M''_X(0) = \mathbb{E}(X^2 e^0) = \mathbb{E}(X^2).$$

Similarly, differentiating n times gives

$$M_X^{(n)}(t) = \mathbb{E}(X^n e^{tX}).$$

Therefore, at $t = 0$,

$$M_X^{(n)}(0) = \mathbb{E}(X^n e^0) = \mathbb{E}(X^n).$$

Taylor Expansion Around a Point

Let f be infinitely differentiable in a neighborhood of a . Its n th-order Taylor polynomial around a is

$$T_n(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k.$$

Taylor's theorem gives

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k + R_n(x),$$

where, under the usual differentiability assumptions, the remainder can be written as

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - a)^{n+1}$$

for some ξ between a and x .

Infinite Taylor Series

If $R_n(x) \rightarrow 0$ as $n \rightarrow \infty$, then

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x - a)^k.$$

Taylor Expansion of the Exponential Function

For

$$f(u) = e^u,$$

we have

$$f^{(k)}(u) = e^u \quad \text{for every } k \geq 0.$$

Expanding around $a = 0$,

$$f^{(k)}(0) = 1,$$

so

$$e^u = \sum_{k=0}^{\infty} \frac{u^k}{k!} = 1 + u + \frac{u^2}{2!} + \frac{u^3}{3!} + \cdots.$$

The exponential function is analytic on all of $\mathbb{R}$, so this series converges to e^u for every $u \in \mathbb{R}$.

Now substitute

$$u = tX.$$

For every fixed realization of X ,

$$e^{tX} = \sum_{k=0}^{\infty} \frac{(tX)^k}{k!} = 1 + tX + \frac{t^2 X^2}{2!} + \frac{t^3 X^3}{3!} + \cdots.$$

Thus, the identity holds **pointwise** for every value of X .

Taking Expectations: Why Is It Valid?

The moment-generating function of X is

$$M_X(t) = \mathbb{E}\left(e^{tX}\right),$$

provided this expectation is finite.

Using the Taylor expansion,

$$M_X(t) = \mathbb{E}\left[\sum_{k=0}^{\infty} \frac{t^k X^k}{k!}\right].$$

To write

$$\mathbb{E}\left[\sum_{k=0}^{\infty} \frac{t^k X^k}{k!}\right] = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbb{E}(X^k),$$

we must justify interchanging the expectation and the infinite sum.

A sufficient condition is that, for some $r > |t|$,

$$\mathbb{E}\left(e^{r|X|}\right) < \infty.$$

Moments as Taylor Coefficients of the MGF

Therefore,

$$M_X(t) = \sum_{k=0}^{\infty} \frac{\mathbb{E}(X^k)}{k!} t^k$$

whenever the required interchange is valid.

Equivalently,

$$M_X(t) = 1 + t\mathbb{E}(X) + \frac{t^2}{2!}\mathbb{E}(X^2) + \frac{t^3}{3!}\mathbb{E}(X^3) + \cdots.$$

Comparing this with the Taylor expansion

$$M_X(t) = \sum_{k=0}^{\infty} \frac{M_X^{(k)}(0)}{k!} t^k,$$

we obtain

$$M_X^{(k)}(0) = \mathbb{E}(X^k)$$

whenever M_X exists in an open neighborhood of 0.

Main Idea

The k th raw moment of X is the k th derivative of its MGF evaluated at $t = 0$.

Example: Bernoulli Moments

For

$$M_X(t) = pe^t + q,$$

we compute:

$$M'_X(t) = pe^t.$$

Thus:

$$M'_X(0) = p.$$

So

$$\mathbb{E}(X) = p.$$

Theorem (Uniqueness of MGF)

Let X and Y be random variables. Suppose there exists some $h > 0$ such that

$$M_X(t) = \mathbb{E}(e^{tX}) < \infty, \quad M_Y(t) = \mathbb{E}(e^{tY}) < \infty$$

for every $t \in (-h, h)$.

If

$$M_X(t) = M_Y(t) \quad \text{for all } t \in (-h, h),$$

then

$$X \stackrel{d}{=} Y,$$

that is, X and Y have the same distribution.

Why is this useful? If we want to determine the distribution of a random variable X :

Find $M_X(t)$ $\rightarrow$ recognize it as the MGF of a known distribution $\rightarrow$ identify the distribution of X .

Using the MGF Uniqueness Theorem

Suppose we compute the MGF of a random variable X and obtain

$$M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right), \quad t \in (-h, h)$$

for some $h > 0$.

We recognize this as the MGF of a normal random variable:

$$Y \sim \mathcal{N}(\mu, \sigma^2) \quad \implies \quad M_Y(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right).$$

Therefore,

$$M_X(t) = M_Y(t) \quad \text{for all } t \in (-h, h).$$

a

By the **MGF Uniqueness Theorem**,

$$\boxed{X \stackrel{d}{=} Y} \quad \implies \quad \boxed{X \sim \mathcal{N}(\mu, \sigma^2)}.$$

Key Idea

If the MGF of an unknown random variable matches the MGF of a known distribution on an interval

MGFs of Sums

Theorem

If X and Y are independent, then

$$M_{X+Y}(t) = M_X(t) \times M_Y(t).$$

More generally, if $X_1, \dots, X_n$ are independent, then

$$M_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n M_{X_i}(t)$$

This is one of the main reasons MGFs are useful. Since using the MGF Uniqueness theorem, the MGF characterizes the distribution in full.

A Binomial random variable can be written as:

$$X = I_1 + \cdots + I_n,$$

where the I_j are independent Bernoulli(p) random variables.

Since each Bernoulli MGF is

$$pe^t + q,$$

we get

$$M_X(t) = (pe^t + q)^n.$$

If

$$X \sim \text{NBin}(r, p),$$

then

$$X = G_1 + \cdots + G_r,$$

where the G_j are independent $\text{Geometric}(p)$ random variables.

Thus,

$$M_X(t) = \left(\frac{p}{1 - qe^t} \right)^r.$$

Sums via MGFs

If

$$X \sim \text{Pois}(\lambda), \quad Y \sim \text{Pois}(\mu),$$

independently, then

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

and

$$M_Y(t) = e^{\mu(e^t - 1)}.$$

Therefore,

$$M_{X+Y}(t) = e^{(\lambda+\mu)(e^t - 1)}.$$

Hence:

$$X + Y \sim \text{Pois}(\lambda + \mu).$$

Remark: This can be generalized to n independent Poisson r.v.s using induction.

If

$$X_1 \sim N(\mu_1, \sigma_1^2), \quad X_2 \sim N(\mu_2, \sigma_2^2),$$

independently, then

$$M_{X_1+X_2}(t) = e^{(\mu_1+\mu_2)t + \frac{1}{2}(\sigma_1^2+\sigma_2^2)t^2}.$$

Thus,

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

Remark: This can be generalized to n independent Normal r.v.s using induction.

Distributions Closed Under Addition Under Independence

Distribution	Assumption	Sum Distribution
Bernoulli	$X_i \sim \text{Bern}(p)$	$\sum_{i=1}^n X_i \sim \text{Bin}(n, p)$
Binomial	$X_i \sim \text{Bin}(n_i, p)$	$\sum_{i=1}^n X_i \sim \text{Bin}\left(\sum n_i, p\right)$
Poisson	$X_i \sim \text{Pois}(\lambda_i)$	$\sum_{i=1}^n X_i \sim \text{Pois}\left(\sum \lambda_i\right)$
Negative Binomial	$X_i \sim \text{NegBin}(r_i, p)$	$\sum_{i=1}^n X_i \sim \text{NegBin}\left(\sum r_i, p\right)$
Geometric	$X_i \sim \text{Geom}(p)$	$\sum_{i=1}^n X_i \sim \text{NegBin}(n, p)$
Normal	$X_i \sim N(\mu_i, \sigma_i^2)$	$\sum_{i=1}^n X_i \sim N\left(\sum \mu_i, \sum \sigma_i^2\right)$
Gamma	$X_i \sim \text{Gamma}(\alpha_i, \lambda)$	$\sum_{i=1}^n X_i \sim \text{Gamma}\left(\sum \alpha_i, \lambda\right)$
Exponential	$X_i \sim \text{Exp}(\lambda)$	$\sum_{i=1}^n X_i \sim \text{Gamma}(n, \lambda)$

Summary

- ▶ MGFs encode moments through derivatives.
- ▶ MGFs determine distributions.
- ▶ Independent sums correspond to products of MGFs.
- ▶ MGFs simplify moment calculations dramatically.
- ▶ The Normal and Exponential MGFs are especially important.

Next lecture

Joint distributions.

Temporary page!

$\text{\LaTeX}$ was unable to guess the total number of pages correctly. As there was some unprocessed data have been added to the final page this extra page has been added to receive it.
If you rerun the document (without altering it) this surplus page will go away, because $\text{\LaTeX}$ now knows the number of pages to expect for this document.