STA 130A: Mathematical Statistics: Brief Course

Lecture 17: Moment Generating Functions

Dogyoon Song

Fall 2026, UC Davis

Announcements

Midterm 2: Mon, Nov 9, 11:00–11:50 AM PT (in class)

- Arrive early; the exam ends at 11:50 AM sharp.
- One hand-written, letter-sized (8.5×11 inches), double-sided sheet of formulas/notes.
- A simple non-graphing scientific calculator is allowed; no other reference materials or electronic devices.
- **SDC accommodations:** Confirm scheduling with AES online.

Preparation and support

- Primary coverage: Lectures 12–18; review on Fri, Nov 6 (Lecture 19).
- Earlier foundations may be assumed, although not to be actively asked.
- Check Canvas and the course webpage for practice materials.
- Office hours this week:
  - Instructor: Wed, 1–3 PM (1–2 PM shared with STA 35C).
  - TA: Thu, 8:30–10:30 AM.

Agenda

Last time

- Covariance, correlation, and variance of sums
- Conditional variance and the law of total variance

Today: Moment generating functions

- Encode moments in one function
- Identify distributions from MGFs
- Handle affine transformations
- Identify sums of independent random variables

Recap: Conditional variance and total variance

Conditional variance: spread around the conditional mean

$$\begin{aligned}\mathrm{Var}(X \mid Y = y) &= \mathbb{E}\big[(X - \mathbb{E}[X \mid Y = y])^2 \mid Y = y\big] \\ &= \mathbb{E}[X^2 \mid Y = y] - (\mathbb{E}[X \mid Y = y])^2.\end{aligned}$$

- For fixed y, this is a number; $\mathrm{Var}(X \mid Y)$ is a random variable whose value depends on Y.

Law of total variance

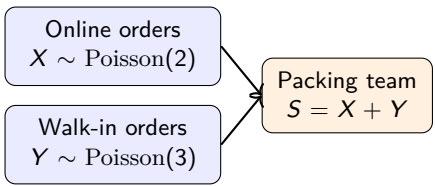
$$\mathrm{Var}(X) = \underbrace{\mathbb{E}[\mathrm{Var}(X \mid Y)]}_{\text{average within-group variance}} + \underbrace{\mathrm{Var}(\mathbb{E}[X \mid Y])}_{\text{variance of group means}}.$$

Machine example: $\frac{7}{2} = \frac{5}{2} + 1$.

- Conditioning cannot increase variance on average, but a particular conditional variance may be larger.

Motivation: Will total demand exceed capacity?

A shop can fill **8 orders in the next hour**. Online and walk-in orders share one packing team; assume their counts are independent.



Question: How likely is demand to exceed capacity?

$$P(S > 8) = ?$$

Knowing only $\mathbb{E}[S] = \text{Var}(S) = 5$ does not determine this probability.

Goal: Find the *distribution* of $S = X + Y$.

MGF idea: Exponentials turn sums into products.

$$\underbrace{\mathbb{E}[e^{tS}]}_{M_S(t)} = \mathbb{E}[e^{tX}e^{tY}] \overset{\text{independence}}{=} \underbrace{\mathbb{E}[e^{tX}]}_{M_X(t)}\underbrace{\mathbb{E}[e^{tY}]}_{M_Y(t)}.$$

Multiply MGFs $\longrightarrow$ identify the distribution of S $\longrightarrow$ calculate $P(S > 8)$.

Moment generating function

Definition (Moment generating function)

The **moment generating function** (or **MGF**) of a random variable X is

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

for values of $t \in \mathbb{R}$ where this expectation is finite.

- Discrete case:
$$M_X(t) = \sum_x e^{tx} p_X(x).$$
- Continuous case:
$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx.$$

The MGF is a function of the parameter t, not a single number.

- Fix t and average over X; varying t gives the function M_X.
- $M_X(0) = 1$ always.

Example: Bernoulli MGF

Example (Bernoulli MGF)

Let $X \sim \mathrm{Bernoulli}(p)$, with $0 \le p \le 1$. Then there are only two terms:

x	0	1
$P(X = x)$	$1 - p$	p
e^{tx}	1	e^t

$$M_X(t) = \mathbb{E}[e^{tX}] = (1-p) \cdot 1 + p \cdot e^t = 1 - p + pe^t.$$

Concrete check: If $p = 0.3$, then $M_X(t) = 0.7 + 0.3e^t$.

$$M_X(0) = 0.7 + 0.3 = 1.$$

This is an ordinary weighted average, applied to the values of e^{tX}.

Example: Poisson MGF

Example (Poisson MGF)

If $X \sim \mathrm{Poisson}(\lambda)$, with $\lambda > 0$, then

$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots.$$

Hence

$$\begin{aligned} M_X(t) &= \sum_{k=0}^{\infty} e^{tk} p_X(k) = \sum_{k=0}^{\infty} e^{tk} e^{-\lambda}\frac{\lambda^k}{k!} \\ &= e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^t)^k}{k!} = e^{-\lambda}e^{\lambda e^t} \\ &= \exp\{\lambda(e^t - 1)\}. \end{aligned}$$

The last step uses $e^u = \sum_{k=0}^{\infty} u^k/k!$. This MGF is finite for every real t.

Self-study: Exponential MGF

Example

Let $X \sim \text{Exponential}(\lambda)$, so

$$f_X(x) = \lambda e^{-\lambda x}, \qquad x \geq 0.$$

For $t < \lambda$,

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

For $t \geq \lambda$, the integral diverges. Thus the MGF for $\text{Exponential}(\lambda)$ exists only for $t < \lambda$.

Self-study: Standard normal MGF

Example (Standard normal MGF)

Let $Z \sim N(0,1)$, so

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \qquad z \in \mathbb{R}.$$

Then

$$\begin{aligned} M_Z(t) &= \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(z-t)^2}{2} + \frac{t^2}{2} \right] dz \\ &= \exp\left(\frac{t^2}{2}\right) \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(z-t)^2}{2} \right] dz}_{=1,\ \text{normal density with mean } t} \\ &= e^{t^2/2}. \end{aligned}$$

Pattern to remember: Standard normal MGF $= e^{t^2/2}$.

MGFs generate moments

From MGFs to moments

When the MGF is finite on an open interval containing 0, differentiating under the expectation gives

$$M_X^{(k)}(0) = \frac{d^k}{dt^k} M_X(t)\Big|_{t=0} = \mathbb{E}[X^k].$$

Example: In particular, for t in that interval,

$$M_X'(t) = \mathbb{E}[Xe^{tX}], \qquad M_X''(t) = \mathbb{E}[X^2 e^{tX}].$$

Setting $t = 0$,

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

Therefore,

$$\mathrm{Var}(X) = M_X''(0) - (M_X'(0))^2.$$

Interpretation: Differentiation extracts moments from one generating function.

Example: Moments from Bernoulli MGFs

Example (Bernoulli)

For $X \sim \text{Bernoulli}(p)$,

$$M_X(t) = 1 - p + pe^t.$$

Thus,

$$M_X'(t) = pe^t, \qquad M_X''(t) = pe^t.$$

Therefore,

$$\mathbb{E}[X] = M_X'(0) = p$$

and

$$\begin{aligned} \text{Var}(X) &= M_X''(0) - \left(M_X'(0)\right)^2 \\ &= p - p^2 \\ &= p(1-p). \end{aligned}$$

Self-study: Moments from Poisson MGFs

Example (Poisson)

For $X \sim \text{Poisson}(\lambda)$,

$$M_X(t) = \exp\big[\,\lambda(e^t - 1)\,\big].$$

Differentiating this, we obtain

$$M_X'(t) = \lambda e^t M_X(t),$$

and

$$M_X''(t) = \left(\lambda e^t + \lambda^2 e^{2t}\right) M_X(t).$$

Thus,

$$M_X'(0) = \lambda, \qquad M_X''(0) = \lambda + \lambda^2.$$

Hence

$$\mathbb{E}[X] = \lambda, \qquad \text{Var}(X) = \lambda.$$

Self-study: Moments from the exponential MGF

Example

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

$$M_X(t) = \frac{\lambda}{\lambda - t}, \qquad t < \lambda.$$

Differentiate:

$$M_X'(t) = \frac{\lambda}{(\lambda - t)^2}, \qquad M_X''(t) = \frac{2\lambda}{(\lambda - t)^3}.$$

Thus,

$$\mathbb{E}[X] = M_X'(0) = \frac{1}{\lambda}, \qquad \mathbb{E}[X^2] = M_X''(0) = \frac{2}{\lambda^2}.$$

Therefore,

$$\text{Var}(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Pop-up quiz: Moments from an MGF

Suppose a random variable X has MGF

$$M_X(t) = \exp(2t + 3t^2).$$

Question: What are $\mathbb{E}[X]$ and $\mathrm{Var}(X)$?

A) $\mathbb{E}[X] = 2, \quad \mathrm{Var}(X) = 3$
B) $\mathbb{E}[X] = 2, \quad \mathrm{Var}(X) = 6$
C) $\mathbb{E}[X] = 5, \quad \mathrm{Var}(X) = 3$
D) $\mathbb{E}[X] = 5, \quad \mathrm{Var}(X) = 6$

Answer: B. We have

$$M_X'(t) = (2 + 6t)e^{2t+3t^2}, \qquad \text{and thus,} \qquad M_X'(0) = 2.$$

Similarly,

$$M_X''(t) = \left[6 + (2 + 6t)^2\right] e^{2t+3t^2}, \qquad \text{and thus,} \qquad M_X''(0) = 10.$$

Therefore,

$$\mathrm{Var}(X) = M_X''(0) - \left(M_X'(0)\right)^2 = 10 - 4 = 6.$$

MGFs can identify distributions

Inversion property

If $M_X(t)$ is finite for all t in some open interval $(-a, a)$ around 0, then the MGF M_X uniquely determines the distribution (CDF) of X.

If two random variables have the same MGF, then they have the same distribution.

$$M_X(t) = 0.7 + 0.3e^t \quad \Longrightarrow \quad X \sim \text{Bernoulli}(0.3),$$
$$M_Y(t) = \exp\{5(e^t - 1)\} \quad \Longrightarrow \quad Y \sim \text{Poisson}(5).$$

Important remarks:

- In general, matching only the mean and variance does not identify a distribution.
- There exist explicit inverse-transform formulas that recover the PMF or PDF from MGF.
- Here we identify distributions by "pattern matching" to known MGFs.

MGF of an affine transformation

MGF of an affine transformation

Let X be a random variable with MGF $M_X(t)$, and let $a, b \in \mathbb{R}$. Then

$$M_{aX+b}(t) = e^{bt} M_X(at),$$

whenever $M_X(at)$ is finite.

Proof. Let $Y = aX + b$. Then

$$\begin{aligned} M_Y(t) &= \mathbb{E}[e^{tY}] = \mathbb{E}[e^{t(aX+b)}] \\ &= e^{bt}\mathbb{E}[e^{(at)X}] \\ &= e^{bt} M_X(at). \end{aligned}$$

Interpretation:

- Adding b multiplies the MGF by e^{bt}.
- Scaling by a changes the input to the MGF from t to at.

Example: Affine transformation of a normal random variable

Example

Suppose $Z \sim N(0,1)$, whose MGF is

$$M_Z(t) = \exp\left(\frac{t^2}{2}\right).$$

Let $X = \sigma Z + \mu$, with $\mu \in \mathbb{R}$ and $\sigma > 0$. Then

$$\begin{aligned} M_X(t) &= e^{\mu t} M_Z(\sigma t) \\ &= e^{\mu t} \exp\left(\frac{\sigma^2 t^2}{2}\right) \\ &= \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right). \end{aligned}$$

This gives the MGF of

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

MGF of a sum of independent random variables

MGF of a sum of independent random variables

Let X and Y be independent. Then

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

Proof.

$$\begin{aligned} M_{X+Y}(t) &= \mathbb{E}[e^{t(X+Y)}] \\ &= \mathbb{E}[e^{tX}e^{tY}] \\ &= \mathbb{E}[e^{tX}]\,\mathbb{E}[e^{tY}] \qquad \text{(by independence)} \\ &= M_X(t)M_Y(t). \end{aligned}$$

More generally, for independent $X_1, \ldots, X_n$ with MGFs finite at t,

$$M_{X_1+\cdots+X_n}(t) = \prod_{i=1}^{n} M_{X_i}(t).$$

Example: Sum of Bernoulli variables

Example

Let $X_1, \ldots, X_n$ be independent $\mathrm{Bernoulli}(p)$ random variables, and let

$$S = X_1 + \cdots + X_n.$$

For one Bernoulli random variable,

$$M_{X_i}(t) = 1 - p + pe^t.$$

Therefore,

$$M_S(t) = \prod_{i=1}^{n} M_{X_i}(t) = (1 - p + pe^t)^n.$$

This is the MGF of a $\mathrm{Binomial}(n, p)$ random variable

$$S \sim \mathrm{Binomial}(n, p).$$

Message: MGFs recover the familiar fact that a sum of independent Bernoulli trials is binomial.

Example: Sum of independent Poisson variables

Example

Let

$$X \sim \text{Poisson}(\lambda_1), \qquad Y \sim \text{Poisson}(\lambda_2),$$

and suppose X and Y are independent.

Recall the MGF of $\text{Poisson}(\lambda)$ is

$$M(t) = \exp\big[\,\lambda(e^t - 1)\,\big].$$

Therefore,

$$\begin{aligned} M_{X+Y}(t) &= M_X(t) M_Y(t) \\ &= \exp\big[\,\lambda_1(e^t - 1)\,\big] \exp\big[\,\lambda_2(e^t - 1)\,\big] \\ &= \exp\big[\,(\lambda_1 + \lambda_2)(e^t - 1)\,\big]. \end{aligned}$$

Thus,

$$X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2).$$

Example: Sum of independent normal variables

Example

Suppose $X_1, \ldots, X_n$ are independent, and let

$$S = \sum_{i=1}^{n} X_i \qquad \text{where} \qquad X_i \sim N(\mu_i, \sigma_i^2), \qquad i = 1, \ldots, n,$$

Recall that $M_Z(t) = e^{t^2/2}$ when $Z \sim N(0,1)$. Thus, if $X = \sigma Z + \mu$, then

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

and therefore, we have

$$M_S(t) = \prod_{i=1}^{n} M_{X_i}(t) = \exp\left[\left(\sum_{i=1}^{n} \mu_i\right) t + \frac{\left(\sum_{i=1}^{n} \sigma_i^2\right) t^2}{2}\right].$$

By the uniqueness of MGFs (inversion property), $S \sim N\left(\sum_{i=1}^{n} \mu_i,\ \sum_{i=1}^{n} \sigma_i^2\right)$.

Pop-up quiz: MGFs and sums

Let $X \sim \text{Poisson}(4)$ and $Y \sim \text{Poisson}(1)$ independently.

Question: Which is the distribution of $X + Y$?

A) $\text{Poisson}(4)$

B) $\text{Poisson}(5)$

C) $\text{Poisson}(4) \times \text{Poisson}(1)$

D) Cannot be determined

Answer: B. The product of the MGFs is

$$e^{4(e^t-1)}e^{1(e^t-1)} = e^{5(e^t-1)}.$$

Wrap-up

Moment generating function

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

- Derivatives at 0 generate moments
- An MGF existing near 0 identifies the distribution

Transformations and independent sums

$$M_{aX+b}(t) = e^{bt} M_X(at), \qquad M_{X+Y}(t) = M_X(t) M_Y(t) \quad \text{under independence.}$$

- Sums of independent Bernoulli(p) variables are binomial
- Independent Poisson counts add their mean-count parameters
- Sums of independent normals are normal, with means and variances added

Suggested reading: [BT08, Sec. 4.4]

Next lecture: Probabilistic tail bounds (Markov and Chebyshev inequalities).

References


Dimitri Bertsekas and John N Tsitsiklis.
Introduction to probability, volume 1.
Athena Scientific, 2nd edition, 2008.