Moment generating function

(and the central limit theorem)

Lecture 11b (STAT 24400 F24)

Moments

Definition The rth **moment** of random variable X is (when exists)

$$\mathbb{E}(X^r)$$

We have already looked at the first moment $\mathbb{E}(X)$ and second moment $\mathbb{E}(X^2)$.

Definition The rth **central moment** of random variable X is (when exists)

$$\mathbb{E}[(X - \mathbb{E}(X))^r]$$

- 1st central moment: $\mathbb{E}[X - \mathbb{E}(X)] = \mathbb{E}(X - \mu_X) = 0$
- 2nd central moment: $\mathbb{E}[(X - \mathbb{E}(X))^r] = \mathrm{Var}(X)$
- ...
- (3rd and 4th central moments are related to skewness and kurtosis)

Moment-generating function (MGF)

There is a neat way to deal with all moments:

Definition The **moment-generating function** (MGF) of random variable X is

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

for $t \in \mathbb{R}$ where the expectation exists.

- Continuous case:

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

- Discrete case:

$$M(t) = \mathbb{E}(e^{tX}) = \sum_x e^{tx}\, \mathbb{P}(X = x) = \sum_x e^{tx} p_X(x)$$

Derivatives of MGF (continuous case)

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

Taking 1st derivative (if exists):

$$M'(t) = \frac{d}{dt} \int_{-\infty}^{\infty} e^{tx} f_X(x) dx = \int_{-\infty}^{\infty} x\, e^{tx} f_X(x) dx$$

$$\Rightarrow \quad M'(0) = \int_{-\infty}^{\infty} x\, f_X(x) dx = \mathbb{E}(X)$$

Taking 2nd derivative (if exists):

$$M''(t) = \frac{d^2}{dt^2} \int_{-\infty}^{\infty} e^{tx} f_X(x) dx = \int_{-\infty}^{\infty} x^2 e^{tx} f_X(x) dx$$

$$\Rightarrow \quad M''(0) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \mathbb{E}(X^2)$$

Derivatives of MGF (discrete case)

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

Taking 1st derivative (if both series converge):

$$M'(t) = \frac{d}{dt} \sum_x e^{tx} p_X(x) = \sum_x x\, e^{tx} p_X(x)$$

$$\Rightarrow \quad M'(0) = \sum_x x\, p_X(x) = \mathbb{E}(X)$$

Taking 2nd derivative (if both series converge):

$$M''(t) = \frac{d^2}{dt^2} \sum_x e^{tx} p_X(x) = \sum_x x^2 e^{tx} p_X(x)$$

$$\Rightarrow \quad M''(0) = \sum_x x^2 p_X(x) = \mathbb{E}(X^2)$$

Moments and uniqueness of MGF

It turns out (proof omitted) that if the MGF $M(t)$ exists in an open interval containing $t = 0$, then

$$M^{(r)}(0) = \mathbb{E}(X^r)$$

Moreover, if the moment-generating function exists for t in an open interval containing zero, it uniquely determines the probability distribution.

So if needed, we can work with MGF instead of cumulative distribution function CDF or probability mass function PMF or probability density function PDF:

$$\text{MGF} \quad \Longleftrightarrow \quad \text{CDF} \quad \Longleftrightarrow \quad \text{PDF or PMF}$$

Example (Poisson MGF)

If $X \sim Poisson(\lambda)$, then $p_X(x) = \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$,

$$M(t) = \mathbb{E}(e^{tX}) = \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

where the sum converges for any $t \in \mathbb{R}$.

$$M'(t) = \frac{d}{dt} M(t) = \lambda e^t\, e^{\lambda(e^t-1)} \quad \Rightarrow \quad \mathbb{E}(X) = M'(0) = \lambda$$

$$M''(t) = \lambda e^t\, e^{\lambda(e^t-1)} + \lambda^2 e^{2t}\, e^{\lambda(e^t-1)} \quad \Rightarrow \quad \mathbb{E}(X^2) = M''(0) = \lambda + \lambda^2$$

Consequently, $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda$.

For a Poisson r.v.,

$$mean = variance$$

Useful properties of MGF

- Suppose X has MGF $M_X(t)$, and $Y = a + bX$, for $a, b \in \mathbb{R}$. Then

  $$M_Y(t) = \mathbb{E}(e^{tY}) = \mathbb{E}(e^{t(a+bX)}) = \mathbb{E}(e^{at} e^{btX}) = e^{at}\, \mathbb{E}(e^{btX})$$

  Thus

  $$M_{a+bX}(t) = e^{at} M_X(bt)$$

- X and Y are independent random variables with MGF's M_X and M_Y. Let $Z = X + Y$. Then

  $$M_Z(t) = \mathbb{E}(e^{tZ}) = \mathbb{E}(e^{tX+tY}) = \mathbb{E}(e^{tX} e^{tY}) = \mathbb{E}(e^{tX})\, \mathbb{E}(e^{tY})$$

  because X and Y are indep.

  $$\Rightarrow \quad M_{X+Y}(t) = M_X(t)\, M_Y(t), \qquad for \quad X \perp\!\!\!\perp Y$$

MGF for $N(0,1)$

$X \sim N(0,1)$,

$$M(t) = \mathbb{E}(e^{tX}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x^2/2 - tx)} dx$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-t)^2/2 \; + \; t^2/2} dx = e^{t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-t)^2/2} dx$$

Thus for standard normal distribution,

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

where we used

- $\frac{x^2}{2} - tx = \frac{1}{2}(x^2 - 2tx + t^2) - \frac{t^2}{2} = \frac{1}{2}(x-t)^2 - \frac{t^2}{2}$
- $\frac{1}{\sqrt{2\pi}} e^{-(x-t)^2/2}$ is the density of $N(t,1)$ thus $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-t)^2/2} dx$.

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

If $Y \sim N(\mu, \sigma^2)$, then we may write

$$Y = \mu + \sigma Z, \qquad Z \sim N(0,1)$$

By the property

$$M_{a+bX}(t) = e^{at} M_X(bt)$$

and

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

we obtain

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

MGF for $i.i.d.$ sums

$X_1, \cdots, X_n$ are i.i.d. random variables with MGF $M(t)$,

$$\mathbb{E}(X_i) = \mu = 0, \qquad \mathrm{Var}(X_i) = \sigma^2, \qquad i = 1, \cdots, n.$$

Let

$$S_n = \sum_{i=1}^{n} X_i, \qquad Z_n = \frac{1}{\sigma\sqrt{n}} S_n$$

By the independence of X_i's, the MGF's

$$M_{S_n}(t) = [M(t)]^n, \qquad M_{Z_n}(t) = \left[M\left(\frac{t}{\sigma\sqrt{n}} \right) \right]^n$$

MGF for $i.i.d.$ sums (Taylor expansion)

Expand $M(t) = \mathbb{E}(e^{tX_i})$ in a Taylor series around 0,

$$M(t) = M(0) + t\,M'(0) + \frac{1}{2} t^2 M''(0) + \cdots + \frac{1}{k!} t^k M^{(k)}(0) + \cdots$$
$$= 1 + \mu t + \frac{1}{2}(\sigma^2 + \mu^2) t^2 + ct^3 + \cdots$$
$$= 1 + 0 + \frac{1}{2}\sigma^2 t^2 + ct^3 + \cdots$$

Replacing t with $\frac{t}{\sigma\sqrt{n}}$,

$$M\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 + \frac{1}{2}\sigma^2 \left(\frac{t}{\sigma\sqrt{n}}\right)^2 + c\left(\frac{t}{\sigma\sqrt{n}}\right)^3 + \cdots$$
$$= 1 + \frac{t^2}{2}\frac{1}{n} + c' t^3 \frac{1}{n^{3/2}} + \cdots$$

MGF for *i.i.d.* sums (n large)

Consider fixed $t \neq 0$ and large n. Then

$$M_{Z_n}(t) = \left[M\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n = \left(1 + \frac{1}{2}\sigma^2\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + \cdots\right)^n$$
$$= \left(1 + \frac{t^2}{2}\frac{1}{n} + \cdots\right)^n$$

where "$\cdots$" terms all contain (constant multiple of) higher powers of $\frac{1}{n}$.

For large n,

$$M_{Z_n}(t) \approx \left(1 + \frac{t^2}{2n}\right)^n$$

MGF for *i.i.d.* sums ($n \to \infty$)

Recall

$$\lim_{n\to\infty}\left(1 + \frac{\alpha}{n}\right)^n = e^\alpha, \qquad \forall \alpha \in \mathbb{R}$$

and $M_{Z_n}(t) \approx \left(1 + \frac{t^2/2}{n}\right)^n$ for large n.

We may conclude (non-rigorously) that,

$$M_{Z_n}(t) \xrightarrow{n\to\infty} e^{t^2/2}$$

Note that $e^{t^2/2}$ is the MGF of $N(0,1)$, the standard normal distribution.

MGF for sums and CLT

Therefore, for

$$Z_n = \frac{S_n}{\sigma\sqrt{n}} = \frac{\sum_{n=1}^{n} X_i}{\sigma\sqrt{n}}$$

we have

$$Z_n \xrightarrow{n\to\infty} Z \qquad \text{(in distribution)}$$

and the limiting random variable

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

We just derived the Central Limit Theorem (non-rigorously, for *i.i.d.* case).

MGF for sums and CLT (remarks)

- This derivation helps us to better understand why the normal distribution is so omnipresent and important.
- A rigorous proof of CLT requires showing that the MGFs exist, and that their convergence leads to convergence in distribution.
- MGF is often used to simplify proofs.
- Other functions playing similar roles:
  - the characteristic function $\phi(t) = \mathbb{E}(e^{itX})$, exists for all $t \in \mathbb{R}$
  - the probability-generating function $G_X(s) = \mathbb{E}(s^X)$, for discrete r.v.'s.