STSCI 5080
Probability Models and Inference
Lecture 14: MGF and LLN

October 16, 2018

Moment generating function

Definition

Suppose that $E(e^{\theta X}) < \infty$ for all $|\theta| < a$ for some $a > 0$. Then the function

$$\psi(\theta) = E(e^{\theta X}),\ |\theta| < a$$

is called the moment generating function (mgf) of X.

The mgf does not exist in the following cases:

(a) $E(e^{\theta X}) < \infty$ for $0 \leq \theta < a$ but $E(e^{\theta X}) = \infty$ otherwise.

(b) $E(e^{\theta X}) < \infty$ for $-a < \theta \leq 0$ but $E(e^{\theta X}) = \infty$ otherwise.

Recall that if the mgf exists for X, then $E(|X|^k) < \infty$ for any positive integer k. In Cases (a) and (b) above, however, it may happen that $E(|X|^k) = \infty$ for some k.

Example

Let Y have Cauchy density and let $X = |Y|$. Then $X > 0$ and so $E(e^{\theta X}) < \infty$ for any $\theta \leq 0$ because $\theta X \leq 0$ and so $e^{\theta X} \leq 1$.
But if $\theta > 0$, then since

$$e^x = 1 + x + \frac{x^2}{2} + \cdots \geq x \quad \text{for } x \geq 0,$$

we have

$$e^{\theta X} \geq \theta X.$$

Because $E(X) = E(|Y|) = \infty$, we have

$$E(e^{\theta X}) \geq \theta E(X) = \infty$$

for any $\theta > 0$. So in this case the mgf does not exist.

Do moments uniquely determine the cdf?

Example

If X and Y are such that $E(X^k) = E(Y^k)$ for all $k = 1, 2, \ldots$ (assuming that they are all finite), then do they have the same cdf?

Do moments uniquely determine the cdf?

Example

If X and Y are such that $E(X^k) = E(Y^k)$ for all $k = 1, 2, \ldots$ (assuming that they are all finite), then do they have the same cdf?

Answer: No.

Heyde's counterexample[1]

Let X be a log-normal random variable:

$$X = e^Z, \quad \text{where } Z \sim N(0, 1).$$

The mgf of Z is $\psi_Z(\theta) = e^{\theta^2/2}$ and so

$$E(X^k) = E(e^{kZ}) = \psi_Z(k) = e^{k^2/2}.$$

The pdf of X is

$$f_X(x) = \frac{1}{\sqrt{2\pi}x} e^{-(\log x)^2/2}, \ x > 0.$$

Next, consider a random variable Y with pdf

$$f_Y(y) = f_X(y)\{1 + \sin(2\pi \log y)\}, \ y > 0.$$

---

[1] Heyde, C.C. (1963) On a property of the lognormal distribution. *J. Royal. Stat. Soc. B.* **29** 392–393.

Is f_Y a pdf? – Yes. Because f_X is the pdf of X, we have

$$\int_0^\infty f_X(y)\sin(2\pi\log y)dy = E\{\sin(2\pi\log X)\}$$
$$= E\{\sin(2\pi Z)\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\underbrace{\sin(2\pi z)e^{-z^2/2}}_{\text{odd function}}dz = 0.$$

We will verify that

$$E(X^k) = E(Y^k)$$

for any positive integer k. To this end, it is enough (why?) to verify that

$$\int_{-\infty}^{\infty} y^k f_X(y)\sin(2\pi\log y)dy = 0.$$

$$\int_{-\infty}^{\infty} y^k f_X(y)\sin(2\pi\log y)dy = E\{X^k \sin(2\pi\log X)\}$$

$$= E\{e^{kZ}\sin(2\pi Z)\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sin(2\pi z)e^{kz-z^2/2}dz.$$

Using the identify

$$kz - z^2/2 = -(z-k)^2/2 + k^2/2$$

we have

$$= e^{k^2/2} \times \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sin(2\pi z)e^{-(z-k)^2/2}dz. \qquad (*)$$

Change the variables $w = z - k$, i.e., $z = w + k$. Because

$$\sin(2\pi z) = \sin\{2\pi(w+k)\} = \sin(2\pi w + 2\pi k) = \sin(2\pi w),$$

we have

$$(*) = \frac{e^{k^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\underbrace{\sin(2\pi w)e^{-w^2/2}}_{\text{odd function}}dz = 0.$$

$$\int_{-\infty}^{\infty} y^k f_X(y)\sin(2\pi\log y)dy = E\{X^k\sin(2\pi\log X)\}$$

$$= E\{e^{kZ}\sin(2\pi Z)\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sin(2\pi z)e^{kz-z^2/2}dz.$$

Using the identify

$$kz - z^2/2 = -(z-k)^2/2 + k^2/2$$

we have

$$= e^{k^2/2} \times \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sin(2\pi z)e^{-(z-k)^2/2}dz. \qquad (*)$$

Change the variables $w = z - k$, i.e., $z = w + k$. Because

$$\sin(2\pi z) = \sin\{2\pi(w+k)\} = \sin(2\pi w + 2\pi k) = \sin(2\pi w),$$

we have

$$(*) = \frac{e^{k^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\underbrace{\sin(2\pi w)e^{-w^2/2}}_{\text{odd function}}dz = 0. \quad \backslash(\geqq \nabla \leqq)/$$

- If X has mgf $\psi(\theta)$, then we can expand $\psi(\theta)$ in a power series:

$$\psi(\theta) = \sum_{k=0}^{\infty} \frac{\mu_k}{k!} \theta^k$$

  in a small open neighborhood of the origin, where $\mu_k = \psi^{(k)}(0) = E(X^k)$. So in this case the moments uniquely determine the mgf and therefore the cdf.

- However, if X is a log-normal random variable, then

$$\mu_k = e^{k^2/2}$$

  diverges too quickly, so that the power series does not converge in any open neighborhood of the origin except for the origin (in fact the mgf does not exist for a log-normal random variable).

Theorem

For a given sequence μ_k, if $\limsup_{k\to\infty} \mu_{2k}^{1/(2k)}/k < \infty$, then there is at least one cdf F such that $\mu_k = E(X^k)$ for all k for $X \sim F$.

See Theorem 3.3.11 in Durrett. [2]

---

[2] Durrett, R. *Probability: Theory and Examples* (3rd Edition). Cambridge University Press.

Chapter 5 Limit Theorems

- Suppose that there is a random sample from a (unknown) cdf F (called the population cdf),

$$X_1, \ldots, X_n \sim F \text{ i.i.d.}$$

  We will make statistical inference (estimation/testing/construction of confidence regions) for F based on the sample.

- We want to find distributions of statistics (=functions) of the sample, based on which we evaluate/derive inference procedures.

- Deriving exact distributions is hard. Instead, we often derive approximate distributions by letting $n \to \infty$.

LLN and CLT

- We focus on the sample mean

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

- Suppose that the cdf F has mean μ and variance $\sigma^2 > 0$.

LLN

Is $\overline{X}_n$ close to μ as $n \to \infty$? In what sense are they close?

CLT

If we normalize $\overline{X}_n$ as

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \tag{*}$$

which has mean 0 and variance 1, is the cdf of (*) close to that of $N(0, 1)$?

Convergence in probability

Definition

Random variables X_n converge in probability to another random variable X (that may be a constant) as $n \to \infty$, $X_n \xrightarrow{P} X$ in short, if

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$$

for any $\varepsilon > 0$.

Chebyshev's inequality

Theorem

Let X be a random variable with $E(X^2) < \infty$. Then

$$P(|X - E(X)| > x) \leq \frac{\mathrm{Var}(X)}{x^2}$$

for any $x > 0$.

Proof

Because of the equivalence

$$|X - E(X)| > x \Leftrightarrow |X - E(X)|^2 > x^2,$$

we have

$$P(|X - E(X)| > x) = P(|X - E(X)|^2 > x^2).$$

By Markov's inequality,

$$P(|X - E(X)|^2 > x^2) \leq \frac{E\{|X - E(X)|^2\}}{x^2} = \frac{\mathrm{Var}(X)}{x^2}.$$

LLN: setup

Suppose that we have a random sample from a cdf F:

$$X_1, \ldots, X_n \sim F \text{ i.i.d.}$$

and F has mean μ and variance $\sigma^2 > 0$.

Consider the sample mean

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Theorem

We have

$$E(\overline{X}_n) = \mu \quad \textit{and} \quad \mathrm{Var}(\overline{X}_n) = \frac{\sigma^2}{n}.$$

Proof

We note that

$$E(\overline{X}_n) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n} \sum_{i=1}^{n} \underbrace{E(X_i)}_{=\mu} = \mu.$$

Next, we have

$$\mathrm{Var}(\overline{X}_n) = \frac{1}{n^2} \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} \sum_{i=1}^{n} \underbrace{\mathrm{Var}(X_i)}_{=\sigma^2} = \frac{\sigma^2}{n}.$$

LLN

Theorem

We have $\overline{X}_n \xrightarrow{P} \mu$ *as* $n \to \infty$*, i.e.,*

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| > \varepsilon) = 0$$

for any $\varepsilon > 0$*.*

Proof

We know that $E(\overline{X}_n) = \mu$ and $\mathrm{Var}(\overline{X}_n) = \sigma^2/n$, so that by Chebyshev's inequality,

$$P(|\overline{X}_n - \mu| > \varepsilon) \leq \frac{\mathrm{Var}(\overline{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

The right hand side $\to 0$ as $n \to \infty$. In addition, the probability is non-negative, so that by the sandwich rule, we have

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| > \varepsilon) = 0.$$

Example: Monte Carlo Integration

Suppose that we want to numerically evaluate

$$E\{g(X)\} = \int_{-\infty}^{\infty} g(x)f(x)dx \quad \text{for } X \sim f \text{ (pdf)}.$$

Assume that $E\{g(X)^2\} < \infty$.

- Generate $X_1, \ldots, X_n \sim f$ i.i.d.
- Random variables $g(X_1), \ldots, g(X_n)$ are i.i.d. as well.
- By LLN,

$$\frac{1}{n}\sum_{i=1}^{n} g(X_i) \xrightarrow{P} E\{g(X)\}.$$

  as $n \to \infty$.