Leia Donaway

3. Moment Generating Functions

A) • Define MGF for r.v. X, $M_X(t) = \mathbb{E}[e^{tX}]$, for values of t only where the expectation exists some open interval containing zero, t is a dummy variable.

Continuous case: $M_X(t) = \int_{-\infty}^{\infty} e^{tX} \underbrace{f(x)}_{\text{pdf}} dx$

• Note: When $t=0$, $M_X(0) = \mathbb{E}[e^0] = \mathbb{E}[1] = 1$

• The k^{th} derivative of $M_X(t)$ is $M_X^{(k)}(t) = \frac{d^k}{dt^k}\mathbb{E}[e^{tX}]$

$= \mathbb{E}[X^k e^{tX}]$

and taken at $t=0$, $M_X^{(k)}(0) = \mathbb{E}[X^k e^{0 \cdot X}] = \mathbb{E}[X^k]$

AKA the k^{th} moment of the distribution

• General Uniqueness Property: If $M_X(t)$ exists for t in an open interval containing zero, it uniquely determines the probability function

→ That is, if $M_X(t) = M_Y(t)$ $\forall t$ near 0, X and Y have the same distribution.

• This is difficult to prove, but it is suggested intuitively by Moment Generation Property, which is that the Moment Generation Function encodes all moments $\mathbb{E}[X^k]$, which tells us that all these several moments are the same, so obviously they cannot come from significantly different distributions. However, it's difficult to say something more concrete, as the MGF doesn't always exist (special case), which brings us to the idea of the characteristic function.

• Lastly for part A) The characteristic function of X, $\psi(t) = \mathbb{E}[e^{itX}]$, $i = \sqrt{-1}$ always exists, even when $M_X(t)$ doesn't. Also, it is invertible (complex), and such a 1:1 function once inverted must always have the same MGF. It also generates moments (using calculus), but because it is a transformed version of the MGF / distribution, it maintains and hopefully helps motivate the intuition for the uniqueness property.

B) • Find MGF for Normal $Z \sim N(0,1)$, Normal pdf $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$

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

Complete the square

$$\left[\frac{z^2}{2} - tz = \frac{1}{2}(z^2 - 2tz + t^2) - \frac{t^2}{2}\right] = \frac{1}{2}(z-t)^2 - \frac{t}{2}$$

$$M_Z(t) = e^{t^2/2} \quad \forall t \in \mathbb{R}$$

• Find the MGF for $X = \mu + \sigma Z \sim N(\mu, \sigma^2)$

$$M_X(t) = \mathbb{E}[e^{tX}] = \mathbb{E}\left[e^{t(\mu+\sigma Z)}\right] = \underbrace{e^{t\mu}}_{\text{constant}} \mathbb{E}\left[e^{t\sigma Z}\right]$$

$M_Z(t) = e^{t^2/2}$ as we know, so substitute

$$M_X(t) = e^{t\mu} M_Z(t\sigma) = e^{t\mu} e^{(t\sigma)^2/2} = e^{t\mu + \frac{1}{2}\sigma^2 t^2} \quad \forall t \in \mathbb{R}$$

• Give the general formula for the MGF of a linear transformation of a variable.

take linear transformation of x, $Y = aX + b$, then their MGFs are related

$$M_Y(t) = \mathbb{E}\left[e^{t(aX+b)}\right] = e^{tb}\,\mathbb{E}\left[e^{(ta)X}\right] = e^{tb} M_X(at)$$

(Useful fact:)

• Additionally, if X, Y independent, the MGF of their sum is the product of their MGFs

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

Which relates to the first part of C)...

C) • Show that for independent random variables $X_1, \ldots, X_n$ with MGFs $M_{X_i}(t)$, the MGF for $Y = \sum X_i$ is $Y = \prod M_{X_i}(t)$

$M_{X_i}(t) = \mathbb{E}\left[e^{tX_i}\right]$, $Y = X_1 + X_2 + \ldots + X_n$

so

$M_Y(t) = \mathbb{E}\left[e^{tX_1} e^{tX_2} e^{tX_3} \ldots e^{tX_n}\right]$ but since X_i are independent

$M_Y(t) = \mathbb{E}\left[e^{tX_1}\right]\mathbb{E}\left[e^{tX_2}\right] \ldots \mathbb{E}\left[e^{tX_n}\right]$

$M_Y(t) = M_{X_1}(t) M_{X_2}(t) \ldots M_{X_3}(t) = \prod_{i=1}^{n} M_{X_i}(t)$

• Use this to show that the sum of independent Normals is Normal

Now say $X_i \sim N(\mu_i, \sigma_i^2)$ are independent, we've just shown that $M_{X_i}(t) = \exp\left[t\mu_i + \frac{1}{2}\sigma_i^2 t^2\right]$ so the MGF of their sum, $Y = X_1 + X_2 + \ldots + X_n$ is

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

Which we can recognize as a Normal distribution in the form

$Y \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$

• Show that for iid Normals $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

Now $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$ and their sum, $S = X_1 + X_2 + \ldots + X_n$ is distributed, $S \sim N(n\mu, n\sigma^2)$ (from the above result) and has MGF $M_S(t) = \exp\left[t(n\mu) + \frac{1}{2}t^2(n\sigma^2)\right]$

The sample mean $\bar{X} = \frac{S}{n}$ has an MGF that is a linear transformation, so $M_{\bar{X}}(t) = M_S(t/n)$. Substitute:

$M_{\bar{X}}(t) = \exp\left[(t/n)(n\mu) + \frac{1}{2}(t/n)^2(n\sigma^2)\right] = e^{t\mu + \frac{1}{2}\sigma^2 t^2/n}$

which we see is

$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

★ This implies that the Sample Mean of a Normal population is also Normal, with variance decreasing as sample size increases.

D) · State the Central Limit Theorem (Blitz 10.3) for reference

If $X_1, X_2, \ldots, X_n$ iid r.v.s with mean μ and var σ^2, then for large n, the distribution of $\bar{X}_n$ converges to a Normal dist. after standardization. So, as $n \to \infty$

$X_i \overset{iid}{\sim} [\mu, \sigma^2]$ (unspecified distrib.)

$$\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \to N(0,1) \text{ dist.}$$

… meaning that for large n, $\bar{X}_n$ distributed approximately normal, regardless of X_i's distribution (fitting conditions).

· Sketch a CLT proof for distributions that have an MGF

not always true

*See Phil's notes on a proof of CLT, Blitzstein Hwang

Let $M(t) = \mathbb{E}(e^{tX_i})$. WLOG $\mu = 0$ because otherwise we'd subtract it off.

Let $S_n = \sum_{i=1}^{n} X_i$ $\quad Z_n = \frac{S_n}{\sqrt{n\sigma^2}}$ (standardized). Then

$$M_{S_n}(t) = (M_X(t))^n \quad \text{and} \quad M_{Z_n}(t) = \left(M_X\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n$$

Use Taylor expansion on $M_X(s) = M_X(0) + sM_X'(0) + \frac{s^2 M_X''(0)}{2} + \varepsilon_s$

$M_X(s) = 1 + 0 + \frac{1}{2}s^2\sigma^2 + \varepsilon_s$, Recognize error $\varepsilon_s = \sum_{k=3}^{\infty} \frac{M^{(k)}(0)}{k!} s^k$ has an exponent of at least three, so $\lim_{n\to\infty} s = 0$, $\lim_{n\to\infty} n\varepsilon_n = 0$, $\lim_{s\to 0} \frac{\varepsilon_s}{s^2} \to 0$

Let $s = \frac{t}{\sigma\sqrt{n}}$, $\lim_{n\to\infty} s = 0$, $M_{Z_n}(t) = \left(1 + \frac{\sigma^2}{2}\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + \varepsilon_n\right)^n = \left(1 + \frac{t^2}{2n} + \varepsilon_n\right)^n$

Fact: $\lim_{n\to\infty}\left(1 + \frac{a_n}{n}\right)^n = e^a$ if we know $\lim_{n\to\infty} a_n = a$. So we choose $a_n = \frac{t^2}{2} + n\varepsilon_n$

Since we can rearrange $\lim_{n\to\infty} M_{Z_n}(t) = \lim_{n\to\infty}\left(1 + \frac{t^2/2 + n\varepsilon_n}{n}\right)^n = e^{t^2/2}$

We recognize this as a Standard ($\mu = 0, \sigma^2 = 1$) Normal MGF

· Explain why the CLT isn't relevent when we assume $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

If $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$, then the sum $S \sim N(n\mu, n\sigma^2)$ and the sample mean $\bar{X}_n = \frac{S}{n} \sim N(\mu, \sigma^2/n)$ exactly for all n.

Phil's Notes

Proof of the Central Limit Theorem

Suppose $X_1, \ldots, X_n$ are i.i.d. (independent and identically distributed) random variables with mean 0, variance σ_x^2 and Moment Generating Function (MGF) $M_x(t)$. Note that this assumes an MGF exists, which is not true of all random variables.

Let $S_n = \sum_{i=1}^n X_i$ and $Z_n = S_n/\sqrt{n\sigma_x^2}$. Then

$$M_{S_n}(t) \;=\; (M_x(t))^n \;\text{ and }\; M_{Z_n}(t) \;=\; \left(M_x\left(\frac{t}{\sigma_x\sqrt{n}}\right)\right)^n.$$

Using Taylor's theorem, we can write $M_x(s)$ as

$$M_x(s) \;=\; M_x(0) + sM_x'(0) + \frac{1}{2}s^2M_x''(0) + e_s,$$

where $e_s/s^2 \to 0$ as $s \to 0$.

$M_x(0) = 1$, by definition, and with $E(X_i) = 0$ and $Var(X_i) = \sigma_x^2$, we know $M_x'(0) = 0$ and $M_x''(0) = \sigma_x^2$. So

$$M_x(s) = 1 + \frac{\sigma_x^2}{2}s^2 + e_s.$$

Letting $s = t/(\sigma_x\sqrt{n})$, we have $s \to 0$ as $n \to \infty$, and

$$M_{Z_n}(t) \;=\; \left(1 + \frac{\sigma_x^2}{2}\left(\frac{t}{\sigma_x\sqrt{n}}\right)^2 + e_n\right)^n \;=\; \left(1 + \frac{t^2}{2n} + e_n\right)^n,$$

where $n\sigma_x^2 e_n/t^2 \to 0$ as $n \to \infty$.

If $a_n \to a$ as $n \to \infty$, it can be shown that

$$\lim_{n\to\infty}\left(1 + \frac{a_n}{n}\right)^n \;=\; e^a.$$

It follows that

$$\lim_{n\to\infty} M_{Z_n}(t) \;=\; \lim_{n\to\infty}\left(1 + \frac{t^2/2 + ne_n}{n}\right)^n \;=\; e^{t^2/2},$$

which is the MGF of a standard Normal. If the MGF exists, then it uniquely defines the distribution. Convergence in MGF implies that Z_n converges in distribution to $N(0,1)$.

The practical application of this theorem is that, for large n, if $Y_1, \ldots, Y_n$ are independent with mean μ_y and variance σ_y^2, then

$$\sum_{i=1}^n\left(\frac{Y_i - \mu_y}{\sigma_y\sqrt{n}}\right) \mathrel{\dot\sim} N(0,1), \quad \text{or} \quad \bar{Y} \mathrel{\dot\sim} N(\mu_y, \sigma_y^2/n).$$

How large is "large" depends on the distribution of the Y_i's. If Normal, then $n = 1$ is large enough. As the distribution becomes less Normal, larger values of n are needed.