Stat 2911 Lecture Notes

Class 33, 2017

Uri Keich

33. MGFs: of a sum of RVs, uniqueness, weak convergence, Continuity Theorem, CLT, Characteristic Functions (Rice 4.5, 4.6, 5.3)

Moment Generating Function (MGF)

The MGF of a RV X is defined as

$$M(t) = E e^{tX} \quad \forall t \text{ for which the RHS} < \infty$$

$M(0) = 1$ for all X.

(i) $X \sim Pois(\lambda)$ $\qquad M(t) = e^{\lambda(e^t - 1)} \quad \forall t \in \mathbb{R}$

(ii) $X \sim N(0,1)$ $\qquad M(t) = e^{t^2/2} \quad \forall t \in \mathbb{R}$

(iii) $X \sim \Gamma(\alpha, \lambda)$ $\qquad M(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha} \quad \forall t < \lambda$

<u>Theorem M1.</u> If $\exists \delta > 0$ s.t. $M(t) < \infty$ $\forall t \in (-\delta, \delta)$ then $\forall n \in \mathbb{N}$ $\quad M^{(n)}(0) = E(X^n)$ (and both exist).

<u>Theorem M2.</u> Let F and G be CDFs and suppose $\exists \delta > 0$ s.t. $\forall t \in (-\delta, \delta)$ $\quad M_F(t) = M_G(t)$ $(< \infty)$. Then $F \equiv G$.

<u>Claim.</u> If X and Y are ind. RVs with MGFs M_X, M_Y then

$$M_{X+Y} = M_X \cdot M_Y.$$

That is, $M_{X+Y}(t) < \infty$ iff $M_X(t)$ and $M_Y(t) < \infty$, and in this case $M_{X+Y}(t) = M_X(t)\, M_Y(t)$.

<u>Proof.</u> If $Z, W \in L'$ are ind. then $E(ZW) =$.

The same proof can extend the identity for $Z, W \geq 0$, where the equality possibly holds between ∞'s.

$$\begin{aligned} M_{X+Y}(t) &= E\, e^{t(X+Y)} \\ &= E\left(e^{tX} e^{tY}\right) \\ &\overset{?}{=} E\left(e^{tX}\right) E\left(e^{tY}\right) \\ &= M_X(t)\, M_Y(t). \end{aligned}$$

Example. $X \sim \Gamma(\alpha, \lambda)$ is ind. of $Y \sim \Gamma(\beta, \lambda)$ where $\alpha, \beta, \lambda > 0$. For $t < \lambda$,

$$\begin{aligned} M_{X+Y}(t) &= M_X(t) M_Y(t) \\ &= \left(\frac{\lambda}{\lambda - t}\right)^{\alpha} \left(\frac{\lambda}{\lambda - t}\right)^{\beta} \\ &= \left(\frac{\lambda}{\lambda - t}\right)^{\alpha+\beta} \end{aligned}$$

$$\Rightarrow X + Y \sim \Gamma(\alpha + \beta, \lambda) \quad \text{(why?)}$$

Suppose $Y = aX + b$ where $a, b \in \mathbb{R}$ and X is a RV with MGF M_X. Then

$$\begin{aligned} M_{aX+b}(t) = M_Y(t) &= E\, e^{tY} \\ &= E\, e^{t(aX+b)} \\ &= e^{tb}\, E\, e^{(ta)X} \\ &= e^{tb} M_X(ta). \qquad \forall t \in \mathbb{R} \end{aligned}$$

Examples. 1) $Y \sim N(\mu, \sigma^2)$ $\Rightarrow Y = \sigma \cdot Z + \mu$ where $Z \sim N(0,1)$

$$\begin{aligned} \Rightarrow M_Y(t) &= e^{t\mu} M_Z(t \cdot \sigma) \\ &= e^{t\mu} e^{t^2\sigma^2/2} = e^{t\mu + t^2\sigma^2/2} \qquad \forall t \in \mathbb{R}. \end{aligned}$$

2) If $X \sim N(\mu,\sigma^2)$ and $Y \sim N(\nu,\rho^2)$ are ind. then

$$M_{X+Y}(t) = M_X(t) M_Y(t)$$
$$= e^{t\mu + t^2\sigma^2/2} \cdot e^{t\nu + t^2\rho^2/2}$$
$$= e^{t(\mu+\nu) + t^2(\sigma^2+\rho^2)/2} \qquad \forall t \in \mathbb{R}$$

why?

$$\Longrightarrow X+Y \sim N(\mu+\nu, \sigma^2+\rho^2).$$

Alternatively use convolution.

Reminder.

Let $\{X_n\}$ and X be jointly distributed RVs.

$X_n \xrightarrow[n]{} X$ in prob. if $\forall \varepsilon > 0$ $P(|X_n - X| > \varepsilon) \xrightarrow[n]{} 0$.

$X_n \xrightarrow[n]{} X$ strongly, or a.s., if $P(X_n \xrightarrow[n]{} X) = 1$.

$X_n \xrightarrow[n]{} X$ weakly, or in distribution, if $F_{X_n}(x) \xrightarrow[n]{} F_X(x)$ $\forall x$ where F_X is cont. at x.

Note: convergence in dist. is phrased in terms of the CDFs (no need for jointly distributed RVs).

You already saw examples of conv. in dist. – where?

Central Limit Theorem (CLT)

Suppose X_i are iid RVs with $\sigma^2 = V(X_i) < \infty$ and $\mu = E(X_i)$.
Then with $S_n = \sum_1^n X_i$, $\forall x \in \mathbb{R}$

$$P\left(\frac{S_n - n\mu}{\sqrt{n}\,\sigma} \le x\right) \xrightarrow[n]{} \phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt.$$

Let $Y_n = \frac{S_n - n\mu}{\sqrt{n}\,\sigma}$, then the CLT states that

$$F_{Y_n}(x) \xrightarrow[n]{} F_Z(x), \qquad \forall x \in \mathbb{R}$$

where $Z \sim N(0,1)$.
Since $\Phi \equiv F_Z$ is cont. everywhere this is equivalent to:

$$\frac{S_n - n\mu}{\sqrt{n}\,\sigma} = Y_n \xrightarrow[n]{} Z \text{ in dist.}$$

MGFs are useful for characterizing convergence in dist.

Continuity Theorem

Let F_n be CDFs with MGFs M_n, and let F be a CDF with MGF M, and suppose that $\exists \delta > 0$ s.t. $M_n(t) \xrightarrow[n]{} M(t)$ $\forall t \in (-\delta, \delta)$. Then $F_n(x) \longrightarrow F(x)$ $\forall x$ where F is cont. at x.

Note that the corollary of the continuity thm is exactly conv. in dist.

<u>Proof</u> (of CLT for the special case where M_{X_i} exists)

We assume here the MGF $M_{X_i} \equiv M$ exists $\forall t \in (-\delta, \delta)$ for some $\delta > 0$. It follows that $X_i \in L^2$. (why?)

WLOG we can assume $\mu = 0$ and $\sigma^2 = 1$: indeed, let $\hat{X}_i = \frac{X_i - \mu}{\sigma}$ and $\hat{S}_n = \sum_1^n \hat{X}_i$. Clearly

$$E(\hat{X}_i) = \frac{1}{\sigma} E(X_i - \mu) =$$

$$V(\hat{X}_i) = \frac{1}{\sigma^2} V(X_i - \mu) =$$

$$\frac{S_n - n\mu}{\sqrt{n}\,\sigma} = \frac{1}{\sqrt{n}} \sum_1^n \frac{X_i - \mu}{\sigma} =$$

So if $\frac{\hat{S}_n}{\sqrt{n}} \xrightarrow[n]{} Z$ in dist. so does $\frac{S_n - n\mu}{\sqrt{n}\,\sigma}$.

$\forall t \in (-\delta, \delta)$: $M_{S_n/\sqrt{n}}(t) = M_{S_n}(t/\sqrt{n})$

$$= [M(t/\sqrt{n})]^n \qquad (M_{X_i} \equiv M)$$

$$< \infty. \qquad (\text{since } t/\sqrt{n} \in (-\delta, \delta))$$

If we can show that $\forall t \in (-\delta, \delta)$

$$M_{S_n/\sqrt{n}}(t) \xrightarrow[n]{} e^{t^2/2} = M_Z(t),$$

then by the continuity theorem, $S_n/\sqrt{n} \longrightarrow Z$ in dist.

Recall the second order Taylor expansion

$$M(s) = M(0) + sM'(0) + \frac{s^2}{2}M''(0) + R_2(s)$$

where $R_2(s)/s^2 \xrightarrow[s\to 0]{} 0$. (how do you know it applies to M?)

In our case,

$M(0) =$, $M'(0) =$, $M''(0) =$

$$\Rightarrow \quad M(s) = 1 + s^2/2 + R_2(s).$$

$$\Rightarrow \quad M_{S_n/\sqrt{n}}(t) = [M(t/\sqrt{n})]^n \qquad s = t/\sqrt{n}$$

$$= [1 + t^2/2n + R_2(t/\sqrt{n})]^n$$

$$= \left[1 + \frac{t^2/2 + R_2(t/\sqrt{n})\cdot n}{n}\right]^n.$$

Let $a_n = t^2/2 + R_2(t/\sqrt{n})\cdot n$, then

$$M_{S_n/\sqrt{n}}(t) = \left(1 + \frac{a_n}{n}\right)^n = e^{n\log(1+\frac{a_n}{n})},$$

where

$$a_n = \frac{t^2}{2} + \frac{R_2(t/\sqrt{n})}{(t/\sqrt{n})^2}\cdot t^2 \xrightarrow[n]{}$$

Recall the first order Taylor expansion of $\log(1+x)$:

$$\log(1+x) = x + R_1(x) \quad \text{where } R_1(x)/x \xrightarrow[x\to 0]{} 0$$

$$\Rightarrow n\log\left(1+\frac{a_n}{n}\right) = n\left[\frac{a_n}{n} + R_1\left(\frac{a_n}{n}\right)\right]$$

$$= a_n + \frac{R_1\left(\frac{a_n}{n}\right)}{a_n/n}\cdot a_n \xrightarrow[n]{}$$

$$\Rightarrow M_{S_n/\sqrt{n}}(t) = e^{n\log\left(1+\frac{a_n}{n}\right)} \xrightarrow[n]{}$$

Since this holds $\forall t \in (-\delta,\delta)$, by the continuity theorem $S_n/\sqrt{n} \xrightarrow[n]{} Z$ in dist. $\square$

There are dists. for which the MGF exists only for $t=0$.

For example, $f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$ (Cauchy):

$$\int_{-\infty}^{\infty} e^{tx}\frac{1}{\pi}\frac{1}{1+x^2}\,dx = +\infty \quad \forall t\neq 0 .$$

For this reason we generally prefer to work with the characteristic function (CF)

$$\phi_X(t) = E\left[e^{itX}\right] \qquad i=\sqrt{-1}$$

$$= M_X(it)$$

$$E\left(e^{itX}\right) = E\left[\cos(tX) + i\sin(tX)\right]$$

$$= E\cos(tX) + iE\sin(tX)$$

Since $|\cos\alpha|, |\sin\alpha| \le 1$ these expectations are well-defined $\forall t \in \mathbb{R}$, so the CF is defined $\forall t$.

Example. If $Z \sim N(0,1)$, then

$$\phi_Z(t) = M_Z(it)$$
$$= e^{(it)^2/2}$$
$$= e^{-t^2/2}$$

If X and Y are ind. RVs then

$$\phi_{X+Y}(t) = Ee^{it(X+Y)}$$
$$= E(e^{itX} \cdot e^{itY})$$
$$\overset{?}{=} \phi_X(t)\,\phi_Y(t)$$

If $E(|X|^k) < \infty$ then $\phi^{(k)}(0) = i^k E(X^k)$

(there is some sort of a converse)

If $Y = aX + b$ $a, b \in \mathbb{R}$ then

$$\phi_Y(t) = e^{itb}\phi_X(at).$$

Example. If $X \sim N(\mu, \sigma^2)$ then $\phi_X(t) = e^{it\mu - \sigma^2 t^2/2}$.

There is an analog of the Continuity Thm using CFs.