Probability-2 Lecture-18

19 March 2024 11:25

Moment Generating Functions (MGFs):

X — a r.v.

For each $t \in \mathbb{R}$, e^{tX} is a non-ve r.v.

$E(e^{tX}) = m_X(t)$: moment generating function of X.

$I = \{t \in \mathbb{R} : m_X(t) < \infty\}$.

$0 \in I$.

Fact: I is a convex set.

Proof:

$t_1, t_2 \in I$. $m_X(t_1), m_X(t_2) < \infty$

$\Rightarrow \forall \ 0 < \alpha < 1$.

$$E\left(e^{(\alpha t_1 + (1-\alpha) t_2)X}\right)$$

$$= E\left(e^{\alpha t_1 X} \cdot e^{(1-\alpha) t_2 X}\right)$$

$$\leq \left(E\left(e^{\alpha t_1 X}\right)^p\right)^{\alpha} \cdot \left(E\left(e^{(1-\alpha) t_2 X}\right)^q\right)^{1-\alpha}$$

Use Holder's inequality with $\alpha = 1/p$

$\therefore 1-\alpha = 1/q$

$$= (m_X(t_1))^{\alpha} \cdot (m_X(t_2))^{1-\alpha} < \infty$$

$\left[\begin{array}{l}\alpha \cdot p = 1 \\ (1-\alpha) \cdot q = 1\end{array}\right]$

$m_X(t_1) < \infty$, $m_X(t_2) < \infty$

$\therefore t_1, t_2 \in I \Rightarrow \underbrace{\alpha t_1 + (1-\alpha) t_2}_{\text{convex combination of } t_1 \,\&\, t_2} \in I$, $\alpha \in (0,1)$

$\therefore I$ is a convex set. $\blacksquare$

Corollary: $\log(m_X(t))$ on I is a convex function.

Note: I is an interval containing 0.

$I = \{0\}$ is possible.

$P(X = \pm n) = c \cdot \frac{1}{n^2}$, $n \in \mathbb{N}$

Special Case:

0 is an interior point of I.

Special Case:

0 is an interior point of I.

Equivalently, $[-t_0, t_0] \subset I$ for some $t_0 > 0$.

Observe: for any $t \in [-t_0, t_0]$,

$$e^{|tX|} \leq e^{t_0 X} + e^{-t_0 X}$$

$$\therefore E(e^{|tX|}) < \infty \quad \forall \; t \in [-t_0, t_0].$$

Fact: for any $\theta > 0$, $\exists$ a finite constant $C (= C_\theta)$ such that, $x < C \cdot e^{\theta x}$ $\forall$ $x \geq 0$. $\left[\text{Why? } \because \underbrace{\frac{x}{e^{\theta x}}}_{\downarrow\, 0 \text{ as } x \to \infty} < C\right]$

or, $|x| < C \cdot e^{\theta |x|}$

$\Rightarrow |x|^k < C e^{\theta |x|}$.

$\therefore \forall \; k \geq 1$, $\exists$ C s.t.

$$|X|^k \leq \underbrace{C \cdot e^{|t_0 X|}}_{\text{finite in the interval } [-t_0, t_0].} \Rightarrow E(|X|^k) < \infty.$$

$\therefore$ For any $t \in [-t_0, t_0]$,

$$\underbrace{\left| \sum_{k=0}^{n} (tX)^k / \lfloor k \right|}_{} \leq e^{|t_0 X|}.$$

$$\sum_{k=0}^{n} (tX)^k / \lfloor k \longrightarrow e^{tX}.$$

$$Y_n \longrightarrow Y.$$

So, by DCT, $E\left(\sum_{k=0}^{n} (tX)^k / \lfloor k \right) \longrightarrow E(e^{tX})$

$$= m_X(t)$$

$$\Rightarrow \sum_{k=0}^{n} \frac{E(X^k) \cdot t^k}{\lfloor k} \longrightarrow m_X(t).$$

$$\Rightarrow \boxed{m_X(t) = \sum_{n=0}^{\infty} E(X^n) \cdot \frac{t^n}{\lfloor n}}$$

* X has all moments finite.
* For all $t \in [-t_0, t_0]$, $m_X(t) = \sum_{n=0}^{\infty} E(X^n) \cdot \frac{t^n}{\lfloor n}$.
* Thus, $m_X(t)$ has a power series expansion on an interval around 0 of positive radius of convergence.

$$\Rightarrow E(X^n) = m_X^{(n)}(0)$$

n^{th} (raw) moment $=$ n^{th} derivative of mgf evaluated at 0.

<u>Fact</u>: $h : I \times \mathbb{R} \longrightarrow \mathbb{R}$, h - continuous in both variables.

I ↑ open intervals.

Let X be a real random variable.

Suppose,

$\exists$ r.v. Y with $E(Y) < \infty$.

s.t. $|h(t, X)| \leq Y \quad \forall \quad t \in I$ & $\forall\, \omega$.

Then, $g(t) = E(h(t, X))$ is continuous on I

<u>Proof</u>: Exercise (use DCT)

$\forall \{t_n\}_{n \geq 1}$ s.t $t_n \longrightarrow t$

then, $h(t_n, X) \longrightarrow h(t, X)$ [by continuity of X]

$\therefore |h(t_n, X)| \leq Y$.

$\therefore$ By DCT, $E(h(t_n, X)) \longrightarrow E(h(t, X))$.

$\Rightarrow g(t_n) \longrightarrow g(t)$

$\therefore g$ is continuous. ▨

<u>Fact</u>:

Suppose $\forall x$,

$\frac{\partial}{\partial t}(h(t, x))$ exists at all $t \in I$, and

$\exists$ r.v. Z with $E(Z) < \infty$,

$\frac{\partial}{\partial t}(h(t,x))$ exists at all $t \in I$, and

$\exists$ a r.v. Z with $E(Z) < \infty$,

s.t. $\left| \frac{\partial}{\partial t}(h(t,x)) \right| \leq Z$

Then, $g(t) = E(h(t,X))$ is differentiable

& $g'(t) = E\left(\frac{\partial}{\partial t} h(t,X)\right)$

[ie, in layman's words, expectation & differentiation can be swapped.]

Proof: (Exercise). Use DCT & MVT.

$\forall \{t_n\}_{n \geq 1}$ s.t. $t_n \to t$,

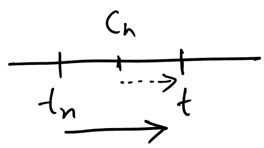
$$\frac{h(t_n,X) - h(t,x)}{t_n - t} = \frac{\partial}{\partial t} h(c_n, X) \quad \text{①}$$

[for some $c_n \in (t_n, t)$ By MVT]

taking E on both sides—

$$E\left(\frac{h(t_n,X) - h(t,X)}{t_n - t}\right) = E\left(\frac{\partial}{\partial t} h(c_n, X)\right) \quad \text{②}$$

As $\frac{\partial}{\partial t}(h(t,X))$ exists $\forall t \in I$,



$$\lim_{n \to \infty} \frac{\partial}{\partial t}(h(c_n, X)) = \frac{\partial}{\partial t}(h(t,X))$$

$\&, \because \left| \frac{\partial}{\partial t}(h(c_n, X)) \right| < Z \quad \forall n, \quad \& \quad E(Z) < \infty$

$\therefore$ By DCT, $E\left(\frac{\partial}{\partial t}(h(c_n,X))\right) \longrightarrow E\left(\frac{\partial}{\partial t}(h(t,X))\right)$

using ②

$$E\left(\frac{h(t_n,X) - h(t,X)}{t_n - t}\right) \longrightarrow E\left(\frac{\partial}{\partial t}(h(t,X))\right)$$

linearity of Expectation.

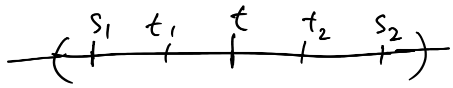
$$\frac{E(h(t_n,X)) - E(h(t,X))}{t_n - t} \longrightarrow E\left(\frac{\partial}{\partial t}(h(t,X))\right)$$

as $n \to \infty$ $\therefore$ $\frac{\partial}{\partial t}(E(h(t,X))) = E\left(\frac{\partial}{\partial t}(h(t,X))\right)$ ∎

Suppose J is an open interval $m_X(t) < \infty \quad \forall\, t \in J$.

Then, $m_X(t)$ is infinitely differentiable on J, &

$$m_X^{(n)}(t) = E(X^n \cdot e^{tX}) \quad \forall \; t \in J.$$



Proof: Fix points $s_1 < t_1 < t_2 < s_2$ in J.

We'll prove: $m_X(t)$ is differentiable at $t \in (t_1, t_2)$ &

$$m_X'(t) = E(X e^{tX}) \quad \forall\, t \in (t_1, t_2)$$

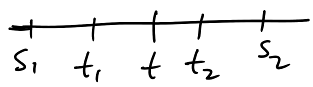
(the rest of the proof follows by induction.).

We can get a constant C s.t.

$$|x| \leq C e^{\underbrace{(s_2 - t_2)}_{\geq 0} x} \cdot \mathbb{1}_{x \geq 0} + C e^{\overbrace{(s_1 - t_1)}^{\leq 0} \cdot \overset{-ve}{x}} \cdot \mathbb{1}_{x<0}.$$

$\overset{\geq 0}{}$

$\Rightarrow$ for $t_1 < t < t_2$:



$$\left|\frac{\partial}{\partial x}(e^{tx})\right| = |x| \cdot e^{tX}$$

$$\leq C e^{\underbrace{(s_2 - \underbrace{t_2 + t}_{\leq 0})}_{< s_2} \cdot x} \cdot \mathbb{1}_{x \geq 0} + C e^{\underbrace{(s_1 - \underbrace{t_1 + t}_{\geq 0})}_{> s_1} x} \cdot \mathbb{1}_{x<0}$$

Substitute $|x|$ by this.

$$\leq C e^{s_2 x} \cdot \mathbb{1}_{x \geq 0} + C e^{s_1 x} \cdot \mathbb{1}_{x<0} \rightarrow$$

this bound is free of t. Hence, can be used as Z.

$$\therefore \left|\frac{\partial}{\partial x}(e^{tX})\right| \leq \underbrace{C e^{s_2 X} \cdot \mathbb{1}_{x \geq 0} + C e^{s_1 X} \cdot \mathbb{1}_{x<0}}_{\text{this is a r.v with finite expectation.}}$$

$\therefore$ By DCT, (2nd part of the "Fact":)

$$E\left(\frac{\partial}{\partial t}(e^{tX})\right) = \frac{\partial}{\partial t} E(e^{tX}) \text{ exists.}$$

& is equal to $E(X e^{tX}) \quad \forall\, t \in (t_1, t_2)$

for higher order derivatives, just use the same approach and conclude by induction.

for higher order derivatives, just use the same approach and conclude by induction.

Special Case:

$X \geq 0$.

This $I \supset (-\infty, 0]$

$m_X(t) < \infty \quad \forall\, t \leq 0$.

Denote $L_X(t) = m_X(-t)$

$= E(e^{-tX})$

$X \geq 0,\ t < 0$

$\therefore tX \leq 0$

$e^{tX} \leq 1$

$E(e^{tX}) < \infty$. ✓

Hence, $I \supset (-\infty, 0]$.

L_X is called the "Laplace Transform" of X.

We'll prove: (in the problem set-5)

① L_X is continuous on $[0, \infty)$

② L_X is infinitely differentiable on $(0, \infty)$

& , $L_X^{(n)}(t) = (-1)^n \cdot E(X^n e^{-tX})$.

Further, the right limit $L_X^{(n)}(0^+)$ exists & is finite $\iff E(X^n) < \infty$.

Digression:

Let $Z_{nx} \sim Poi(nx)$ $\quad n \geq 1$, $x \geq 0$.

$P(Z_{nx} \leq na) \to\ ?$

for $a > 0$.

Case-I: $(x < a)$:

$P(Z_{nx} \leq na)$

$= P\left(Z_{nx} - E(Z_{nx}) \leq n(a-x)\right)$

$= 1 - \underbrace{P\left(Z_{nx} - E(Z_{nx}) > n(a-x)\right)}$

$\leq \dfrac{V(Z_{nx})}{(n(a-x))^2}$ [By Chebyshev's Inequality]

$\because X \sim Poi(\lambda) \Rightarrow E(X) = \lambda,\ V(X) = \lambda$

$\therefore E(Z_{nx}) = nx$,

$V(Z_{nx}) = nx$.

$\overline{(n(a-x))^2}$ [Inequality]

$$\therefore \quad \geq 1 - \frac{nx}{n^2(a-x)^2}$$

$$= 1 - \frac{x}{n(a-x)^2} \to 1 \text{ as } n \to \infty.$$

Case – II : $(x > a)$.

$$P(Z_{nx} \leq na) = P(-Z_{nx} \geq -na)$$
$$= P(-(Z_{nx} - E(Z_{nx})) \geq n(x-a))$$
$$= P(|Z_{nx} - E(Z_{nx})| \geq n(x-a))$$

(Again, by Chebyshev's Inequality) $\leq \frac{x}{n(x-a)^2} \longrightarrow 0.$

$$\therefore \; P(Z_{nx} \leq na) \longrightarrow \begin{cases} 1, & x < a \\ 0, & x > a. \end{cases}$$

pmf of Poi(nx)

$$\therefore \sum_{k=0}^{na} e^{-nx} \frac{(nx)^k}{\lfloor k} \longrightarrow \begin{cases} 1 & \text{if } x < a \\ 0 & \text{if } x > a. \end{cases}$$

$\forall \omega, X(\omega) \in \mathbb{R}$. hence this holds.

$$\therefore \sum_{k=0}^{na} e^{-nX} \frac{(nX)^k}{\lfloor k} \xrightarrow[\text{(pointwise convergence)}]{} \begin{cases} 1 & \text{on } \{X < a\} \\ 0 & \text{on } \{X > a\} \end{cases}$$

Suppose $a > 0$ is such that,

$$P(X = a) = 0.$$

$$\therefore P(X > a) + P(X < a) = 1.$$

Also, $\left| \sum_{k=0}^{na} e^{-nX} \cdot \frac{(nX)^k}{\lfloor k} \right| \leq 1 \leftarrow$ ("Z" for using DCT) & , $E(1) = 1 < \infty$.

$\therefore$ By DCT, $\quad \lim_{n \to \infty} \sum_{k=0}^{na} e^{-nX} \cdot \frac{(nX)^k}{\lfloor k} = \mathbb{1}_{\{X < a\}}$

(taking E on both sides)

$$\Rightarrow P(X \leq a) = \lim_{n \to \infty} \sum_{k=0}^{na} \frac{n^k}{\lfloor k} \cdot E(X^k \cdot e^{-nX})$$

$F_X(a)$

as $P(X=a)=0$ ie, has no mass.

$$= \lim_{n \to \infty} \sum_{k=0}^{na} (-1)^k \cdot \frac{n^k}{\lfloor k} \cdot L_X^{(k)}(n).$$

mgf of Laplace transform

So, this works $\forall$ pts of continuity of the distribution of X.

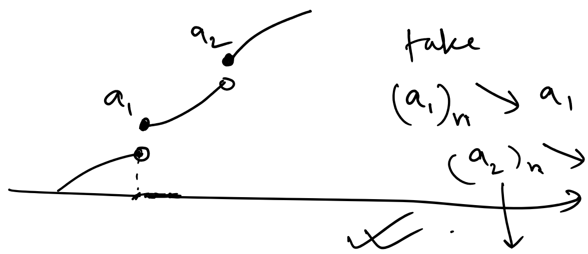
As for the pts of discontinuity, we can use right continuity of F_X.

So far we've shown, L_X determines the CDF of X at only those points where Probability Mass = 0.

for points where Probability Mass > 0 at such points there will be discontinuity in F_X.

&, F_X is monotonic. Hence, it can have atmost countably many discontinuities.

Hence, set of such points must be atmost countable.



take $(a_1)_n \searrow a_1$, $(a_2)_n \searrow a_2$.

By Right cont. of F_X.

Conclusion:

Laplace transform of a non-ve r.v. completely determines the distribution of that random variable.

(G. Ramachandran, ISI Delhi.)