

Lecture 12:

Generating functions and moments

(1)

We will see that when working with probability measures over a complicated space (Ω, \mathcal{F}) it will be useful to be able to characterize a probability P on (Ω, \mathcal{F}) by analyzing the value of $\int_{\Omega} f dP$ computed over a range of test functions $f: \Omega \rightarrow \bar{\mathbb{R}}$.

Characteristic functions and moment generating functions are an example of this.

MGF's, CF's and Complex generating functions

Definition:

Let (Ω, \mathcal{F}, P) be a probability space and X be a random variable (taking values in $\bar{\mathbb{R}}$).

For $t \in \mathbb{R}$ and $z \in \mathbb{C}$ define

$$M_X(t) := E(e^{tzX}) \quad \leftarrow \text{Moment generating function of } X \text{ (MGF).}$$

$$G_X(z) := E(e^{zX}) \quad \leftarrow \text{Complex generating function of } X.$$

$$\phi_X(t) := E(e^{itX}) \quad \leftarrow \text{Characteristic function for } X \text{ (CF).}$$

In general, if μ is a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ define

$$M_{\mu}(t) := \int_{\mathbb{R}} e^{tx} d\mu(x)$$

$$G_{\mu}(z) := \int_{\mathbb{R}} e^{zx} d\mu(x).$$

$$\phi_{\mu}(t) := \int_{\mathbb{R}} e^{itx} d\mu(x)$$

Since $x \mapsto e^{zx}$ takes values in \mathbb{C} we need (2) to say a word about integrating complex valued functions.

If $f: \Omega \rightarrow \mathbb{C}$ then one can decompose it into real and imaginary parts:

$$f(w) = \underbrace{\operatorname{Re} f(w)} + i \underbrace{\operatorname{Im} f(w)}$$

functions mapping $\mathbb{R} \rightarrow \mathbb{R}$.

If μ is a measure on (Ω, \mathcal{F}) then

$$\int_{\Omega} f(w) d\mu(w) := \underbrace{\int_{\Omega} \operatorname{Re} f(w) d\mu(w)} + i \underbrace{\int_{\Omega} \operatorname{Im} f(w) d\mu(w)}$$

all the properties of $\int_{\Omega} f(w) d\mu(w)$ extend to the complex case with minor changes

when these two are defined i.e.
 $\operatorname{Re}, \operatorname{Im} \in Q(\Omega, \mathcal{F}, \mu)$.

The usefulness of these generating functions come from 3 facts:

1) ϕ_X & G_X (and M_X sometimes) characterizes the distribution of X .

E.g. if you have two r.v.s X & Y then $X = Y$ iff $\phi_X(t) = \phi_Y(t) \forall t \in \mathbb{R}$.

Note: This is analogous to c.d.f.s and densities.

2) The generating functions for sums of independent r.v.s is easy to calculate. i.e. If X_1, \dots, X_n are independent r.v.s all defined on (Ω, \mathcal{F}, P) then

$$\phi_{X_1 + \dots + X_n}(t) = \phi_{X_1}(t) \cdots \phi_{X_n}(t).$$

Note: The corresponding operation for densities is hard, i.e. the density of $X_1 + \dots + X_n$ is a n -fold convolution of the densities of each X_i .

3) If you know $M_X(t)$, $\phi_X(t)$ or $G_X(z)$ you can compute the moments $E(X^k)$ by differentiating.

Note: if μ is a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ s.t. (3)

$$d\mu = S dx$$

then $\phi_\mu(t)$ is just the Fourier transform of S . It is actually more natural to think of $\phi_\mu(t)$ as inverse Fourier transform of $s(x)$ where x represents frequency.

This explains fact 3) since the FT & inverse FT diagonalizes convolution.

Relating ϕ_x , M_x and G_x

There are two examples of generating functions that are useful to keep in your mind.

Let $Z \sim N(0,1)$. Then

$$\phi_z(t) = e^{-t^2/2}$$

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

$$G_z(z) = e^{z^2/2}$$

Also let Y be a r.v. with density $s(x) = \frac{1}{\pi(1+x^2)}$ w.r.t. dx . Then

$$\phi_y(t) = e^{-|t|}$$

$$M_y(t) = \begin{cases} 1 & \text{if } t=0 \\ \infty & \text{o.w.} \end{cases}$$

$$G_y(z) = \begin{cases} e^{-|z|} & \text{if } \operatorname{Re} z = 0 \\ \infty & \text{if } \operatorname{Re} z \neq 0 \& \operatorname{Im} z = 0 \\ \text{Not defined} & \text{o.w.} \end{cases}$$

Note: Y is a Cauchy r.v..

Looking at the case of $Z \sim N(0,1)$ we have (4)

$M_z(it) = \phi_z(t)$. However this can't hold in general since $M_y(it) \neq \phi_y(it)$.

To understand the difference we need to analyze G_x .

Definition: If \mathcal{S} is a metric space, with metric d , and $A \subset \mathcal{S}$ define

$A^\circ := \text{the open interior of } A \leftarrow \text{union of all open sets } C \subset A$

$\bar{A} := \text{the closure of } A \leftarrow \text{intersection of all closed sets containing } A$

$$\partial A := \bar{A} - A^\circ$$

$$d(x, A) := \inf \{ d(x, y) : y \in A \}$$



Also for any subset of \mathbb{C} let

$$\operatorname{Re} A := \{ \operatorname{Re} z : z \in A \}$$

$$\operatorname{Im} A := \{ \operatorname{Im} z : z \in A \}$$

Definition:

For any r.v. X let

$$\mathcal{D}_X := \{ u + iz \in \mathbb{C} : E(e^{uX}) < \infty \}$$

= the cylinder in \mathbb{C} with base $\{u \in \mathbb{R} : M_X(u) < \infty\}$

Theorem ($\operatorname{Re} \mathcal{D}_X$ is an interval)

If X is a r.v. then $\operatorname{Re} \mathcal{D}_X$ is an interval containing 0 (closed, open or half open) and M_X is convex on $\operatorname{Re} \mathcal{D}_X$.

Remark: This thm is true for $\operatorname{Re} \mathcal{D}_\mu$ & M_μ when μ is a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ but now the interval can be empty (in which case it will not contain 0).

Proof:

Since $t \mapsto e^{tx}$ is convex we have

$$e^{[\alpha t_1 + (1-\alpha)t_2]x} \leq \alpha e^{t_1 x} + (1-\alpha)e^{t_2 x}$$

$\forall t_1, t_2 \in \mathbb{R}$ & $\alpha \in [0, 1]$.

$$\begin{aligned} M_X(\alpha t_1 + (1-\alpha)t_2) &= E(e^{[\alpha t_1 + (1-\alpha)t_2]X}) \\ &\stackrel{\text{by } 3}{\leq} \alpha E(e^{t_1 X}) + (1-\alpha) E(e^{t_2 X}) \\ &= \alpha M_X(t_1) + (1-\alpha) M_X(t_2) \end{aligned}$$

$\therefore M_X$ is convex.

Now suppose $t_1, t_2 \in \text{Re } \mathcal{D}_X$. Then $M_X(t_1) < \infty$,

$M_X(t_2) < \infty$ and

$$t_1 \leq t \leq t_2 \implies M_X(t) \leq \alpha M_X(t_1) + (1-\alpha) M_X(t_2) < \infty$$

Writing $t = \alpha t_1 + (1-\alpha)t_2$ for some $\alpha \in [0, 1]$

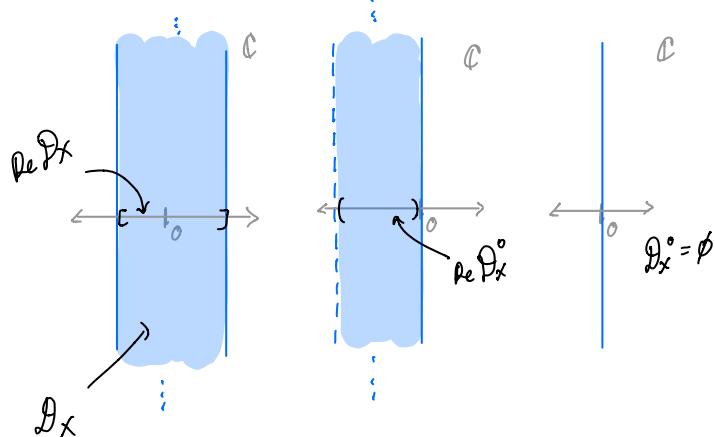
$$\implies t \in \text{Re } \mathcal{D}_X.$$

$\therefore \text{Re } \mathcal{D}_X$ is an interval containing 0,

since clearly $M_X(0) = 1$.

QED

So for any rv. X \mathcal{D}_X could look something like this:



(5)

Theorem: (The Analyticity of G_X over \mathcal{D}_X^o)

Let X be a r.v. (mapping into \mathbb{R}) such that $\mathcal{D}_X^o \neq \emptyset$. Then $\forall z \in \mathcal{D}_X^o$

i) $E|X^n e^{zx}| < \infty$ for $n = 0, 1, 2, \dots$

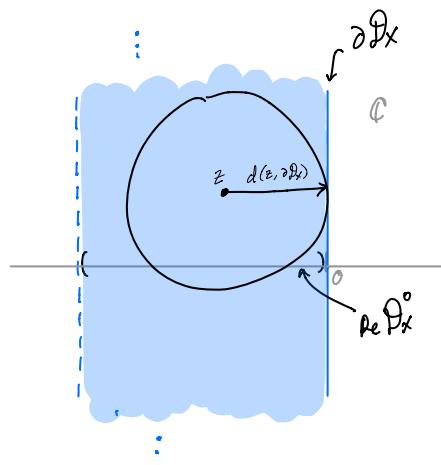
$$\text{ii) } G_X(z) = \sum_{n=0}^{\infty} E(X^n e^{zx}) \frac{(z-z)^n}{n!}$$

for all z in the open ball of \mathbb{C} centered at z with radius $d(z, \partial \mathcal{D}_X)$.

iii) G_X is infinitely differentiable on \mathcal{D}_X^o with complex derivative

$$\frac{d^n}{dz^n} G_X(z) = E(X^n e^{zx}).$$

e.g.



Proof:

First note that for any $z, \zeta \in \mathbb{C}$ we have

$$e^{\zeta X} = e^{zX} e^{(\zeta-z)X} = \sum_{n=0}^{\infty} e^{zX} \frac{X^n (\zeta-z)^n}{n!}$$

and

$$(*) \quad \sum_{n=0}^{\infty} \left| e^{zX} \frac{X^n (\zeta-z)^n}{n!} \right| \leq |e^{zX}| \sum_{n=0}^{\infty} \frac{|X(\zeta-z)|^n}{n!}$$

$$= e^{uX} e^{-r|X|}$$

$$\leq e^{(u-r)X} + e^{(u+r)X}$$

$$\text{since } e^{-r|X|} = \begin{cases} e^{-rX} & \text{if } X \geq 0 \\ e^{rX} & \text{if } X < 0 \end{cases}$$

(6)

Notice that when $z, \zeta \in \mathcal{D}_X$ and $r := |z - \zeta| < d(z, \partial \mathcal{D}_X)$ then

$$u+r := \operatorname{Re} z \pm i|z-\zeta| \in \operatorname{Re} \mathcal{D}_X^o$$

so that

$$E(e^{(u-r)X}) = M_X(u-r) < \infty$$

$$E(e^{(u+r)X}) = M_X(u+r) < \infty$$

$$\therefore E\left(\sum_{n=0}^{\infty} \left|e^{zX} \frac{X^n (\zeta-z)^n}{n!}\right|\right) < \infty \quad (**)$$

II \leftarrow By monotone convergence in Big 3

$$\sum_{n=0}^{\infty} E\left|e^{zX} \frac{X^n (\zeta-z)^n}{n!}\right|$$

$$\therefore E\left|e^{zX} \frac{X^n (\zeta-z)^n}{n!}\right| = \frac{|\zeta-z|^n}{n!} E\left|e^{zX} X^n\right| < \infty$$

for all $n = 1, 2, \dots$ so i) holds

To show ii) notice that

$$\sum_{n=0}^N e^{zX} \frac{X^n (\zeta-z)^n}{n!} \xrightarrow{N \rightarrow \infty} \sum_{n=0}^{\infty} e^{zX} \frac{X^n (\zeta-z)^n}{n!} \text{ P-a.e.}$$

and DCT applies with upper bound given

by the LHS of (*) which is integrable by (**).

$$\therefore G_X(\zeta) = E(e^{\zeta X})$$

$$= E\left(\lim_N \sum_{n=0}^N e^{zX} \frac{X^n (\zeta-z)^n}{n!}\right)$$

$$\stackrel{\text{DCT}}{=} \lim_N \sum_{n=0}^N E\left(e^{zX} X^n\right) \frac{(\zeta-z)^n}{n!}$$

$$= \sum_{n=0}^{\infty} E\left(e^{zX} X^n\right) \frac{(\zeta-z)^n}{n!}$$

This gives ii).

Finally iii) follows directly from ii).

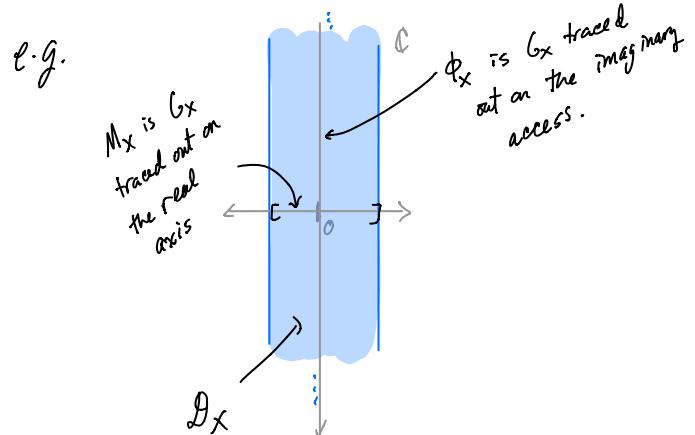
QED

(7)

Now we can understand the relationship (3)
btwn ϕ_X and M_X :

- $G_X(z)$ always exist and is finite on \mathcal{D}_X
since $u+iv \in \mathcal{D}_X \Rightarrow M_X(u) = E(e^{uX}) < \infty$
 $\Rightarrow E\left|e^{(u+iv)X}\right| = E(e^{uX}) < \infty$
 $\Rightarrow |G_X(u+iv)| < \infty$.

- \mathcal{D}_X always contains $\{it : t \in \mathbb{R}\}$
since $0 \in \operatorname{Re} \mathcal{D}_X$
- $\phi_X(it) = G_X(it) \quad \forall t \in \mathbb{R}$ and
 $M_X(it) = G_X(it) \quad \forall t \in \operatorname{Re} \mathcal{D}_X$



- $G_X(z) =$ the unique analytic extension of ϕ_X on $\{it : t \in \mathbb{R}\}$ to \mathcal{D}_X .
only when $\operatorname{Re} \mathcal{D}_X \neq \emptyset$ \downarrow the unique analytic extension of M_X on $\operatorname{Re} \mathcal{D}_X$ to \mathcal{D}_X

This follows by a complex analysis result:

Thm: Suppose $D \subset \mathbb{C}$ is open and connected.
If f and g are differentiable complex-valued functions defined on D which agree on distinct $z_1, z_2, \dots \in D$ s.t. $\lim_{n \rightarrow \infty} z_n \in D$ then

$$f(z) = g(z) \quad \forall z \in D.$$

Now suppose we have a formula for ⑨

$M_X(t)$ s.t. $\text{Re } \mathcal{D}_X \neq \{0\}$ and a extension $H(z)$ defined on \mathcal{D}_X s.t.

$$H(t) = M_X(t) \quad \forall t \in \text{Re } \mathcal{D}_X.$$

Then if H is complex differentiable on \mathcal{D}_X^o must be that

$$H(z) = G_X(z), \quad \forall z \in \mathcal{D}_X^o \text{ and}$$

$$H(it) = \phi_X(t), \quad \forall t \in \mathbb{R}.$$

defined by Continuity to $\partial \mathcal{D}_X$
in case $\text{Re } \mathcal{D}_X = [0, a)$ for e.g.

This also works if you can compute

$$\alpha_n := E X^n$$

when $\beta_n := E |X|^n$ decay fast enough so $\sum_{n=0}^{\infty} \beta_n \frac{t^n}{n!}$ has a non-zero radius of convergence. In which case

$$M_X(t) = \sum_{n=0}^{\infty} \alpha_n \frac{t^n}{n!}$$

for all t in an open neighborhood of 0 (use similar arguments for the thm on G_X).

This completely determines G_X , and thus ϕ_X , by analytic extension to \mathcal{D}_X .

To summarize:

Corollary:

i) M_X characterizes ϕ_X when $M_X(t)$ finite over some non-empty open interval of \mathbb{R} .

ii) The moments $\{E|X|^n\}_{n \geq 1}$ characterize ϕ_X only when

$$\sum_{n=0}^{\infty} E|X|^n \frac{t^n}{n!}$$

has a non-zero radius of convergence about $t=0$.

e.g. If $X \sim N(0, 1)$ then one can derive that $M_X(t) = e^{t^2/2}$ & $\text{Re } \mathcal{D}_X = \mathbb{R}$. Here are two extensions defined on $\mathcal{D}_X = \mathbb{C}$:

$$H_1(z) = e^{z^2/2} \leftarrow \text{not analytic}$$

$$H_2(z) = e^{z^2/2} \leftarrow \text{analytic}$$

$\therefore G_X(z) = H_2(z)$ but not $H_1(z)$ and

$$\phi_X(t) = H_2(it) = e^{-t^2/2} \text{ but not } H_1(it) = e^{t^2/2}$$

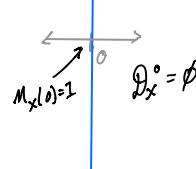
e.g. If X is a Cauchy R.V.

then $\text{Re } \mathcal{D}_X^o = \emptyset$ which

explains why we can't get

ϕ_X or G_X from M_X .

$$\begin{aligned} \phi_X(it) &= G_X(it) \\ &= e^{-|t|} \end{aligned}$$



e.g. Suppose $X \geq 0$ which satisfies

$$EX^n = E|X|^n = n!$$

Can we infer what ϕ_X is?

$$\text{Since } \sum_{n=0}^{\infty} n! \frac{t^n}{n!} < \infty \quad \forall t \in (-1, 1),$$

$M_X(t)$ is finite on $(-1, 1) = \text{Re } \mathcal{D}_X$

$$\begin{aligned} \therefore M_X(t) &= G_X(t) = \sum_{n=0}^{\infty} E(X^n) \frac{t^n}{n!} \\ &= \sum_{n=0}^{\infty} t^n = \frac{1}{1-t} \end{aligned}$$

when $t < 1$

Since $\frac{1}{1-t}$ is an analytic extension of M_X to \mathcal{D}_X we must have

$$\phi_X(t) = G_X(it) = \frac{1}{1-it} \quad \forall t \in \mathbb{R}.$$

This r.v. is an exponential r.v. where

$$dP_X^{-1} = e^{-x} I_{[0, \infty)}(x) dx$$

Another consequence of our analyticity theorem for G_X is that both ϕ_X and M_X give the moments of X . (11)

Since ϕ_X is always defined and not every r.v. has finite moments of all orders it suggests that

The larger n is
s.t. $E|X|^n < \infty \iff \phi_X^{(n)}(t) \text{ is at } t=0$

Corollary: (Moments from M_X or ϕ_X)

If X is a r.v. s.t. $0 \in \text{Re } \phi_X^0$ then $E|X|^n < \infty$ for all n and

$$M_X^{(n)}(0) = (-i)^n \phi_X^{(n)}(0) = E(X^n).$$

Proof:

Note that $G_X(t) = M_X(t)$ when $t \in \mathbb{R}$ and the complex derivatives of an analytic function equal the "directional derivatives".

$$\therefore \frac{d^n}{dt^n} M_X(t) \Big|_{t=0} = \frac{d^n}{dz^n} G_X(z) \Big|_{z=0} = E(X^n).$$

A similar argument holds for ϕ_X . (2) \square

Getting Moments from ϕ_X (12)

What about the case when $0 \notin \text{Re } \phi_X^0$? Now either $E|X|^n$ doesn't decay fast enough to be summable, or they are ∞ for all large n .

We need a more fine tuned argument.

Studying Taylor's theorem gets

$$(*) \quad \left| e^{itx} - \sum_{n=0}^N \frac{(itx)^n}{n!} \right| \leq \frac{|tx|^{N+1}}{(N+1)!}$$

$t \in \mathbb{R}$ & $N \geq 0$.

Note: this bound gives a slightly sub-optimal bound on the regularity of $\phi_X(t)$ near $t=0$ when $E|X|^{N+1} = \infty$. See Billingsley p. 343 for the more fine tuned result.

Now (*) already gives

$$\left| \phi_X(t) - \sum_{n=0}^N E(X^n) \frac{(it)^n}{n!} \right| \leq \frac{|t|^{N+1}}{(N+1)!} E(|X|^{N+1})$$

Theorem: (Moments from ϕ_X)

If X is a r.v. that satisfies $E|X|^{n+1} < \infty$ for some $n \in \{1, 2, \dots\}$ then ϕ_X is n times differentiable and

$$\phi_X^{(m)}(t) = E((iX)^m e^{itX}).$$

If $m \leq n$.

Proof: Start with $m=1$.

$$\begin{aligned} \frac{\phi_X(t+\varepsilon) - \phi_X(t)}{\varepsilon} &= E\left(\frac{e^{itX} e^{i\varepsilon X} - e^{itX}}{\varepsilon}\right) \\ &= E\left(e^{itX} \frac{e^{i\varepsilon X} - 1}{\varepsilon}\right) \end{aligned}$$

(14)

(13)

Therefore

$$\frac{\phi_X(t+\epsilon) - \phi_X(t)}{\epsilon} - E(iX e^{itX}) = E\left(e^{itX} \underbrace{\frac{e^{i\epsilon X} - 1 - i\epsilon X}{\epsilon}}_{\text{bdd in magnitude}}\right)$$

$$\therefore \lim_{\epsilon \rightarrow 0} \text{RHS} = 0 \text{ by DCT.} \quad \text{by } \frac{|i\epsilon X|^2}{\epsilon^2 2!} \text{ from (*)}$$

$$\therefore \phi'_X(t) = E(iX e^{itX}).$$

Repeating the argument gives the result.

QED.

- In the lecture on Separating classes we will show that if $\phi_X(t) = \phi_Y(t) \ \forall t \in \mathbb{R}$ then $X \stackrel{d}{=} Y$
- In the lecture on Convergence in distribution we will show $\phi_{X_n}(t) \rightarrow \phi_X(t) \ \forall t$ then $X_n \xrightarrow{d} X$