PROBABILITY THEORY

5° LECTURE - 06/12/2018

The distribution of a r.v completely characterizes its behavior.
However, if we don't know the law of a r.v., we can still derive information by means of inequalities or through the study of suitable "generating functions".

) Inequalities:

- Cauchy-Schwarz inequality: X, Y r.v., $g_1, g_2 : \mathbb{R} \to \mathbb{R}$. Then

$$\mathbb{E}(g_1(X) g_2(Y))^2 \leq \mathbb{E}(g_1(X)^2)\mathbb{E}(g_2(Y)^2)$$

- Jensen inequality: X r.v. $f : \mathbb{R} \to \mathbb{R}$ convex function ($f'' \geq 0$)

$\Longrightarrow$ (if $\mathbb{E}(|X|) < \infty$)

$$\mathbb{E}(f(X)) \geq f(\mathbb{E}(X))$$

[Also, if f is concave, $f'' \leq 0$, then $\mathbb{E}(f(X)) \leq f(\mathbb{E}(X))$

Proof: Set $\mu = \mathbb{E}(X)$. By Taylor expansion of f around μ:

$$f(x) = f(\mu) + f'(\mu)(x-\mu) + \underset{\geq 0}{f''(\xi)}\frac{(x-\xi)^2}{2}, \quad \text{for } \xi \in (\mu, x)$$

$$\geq f(\mu) + f'(\mu)(x-\mu)$$

Then $x \to X$ and averaging

$$\mathbb{E}(f(X)) \geq f(\mu) + f'(\mu)\underbrace{\mathbb{E}(X-\mu)}_{=0} = f(\mu) = f(\mathbb{E}(X)) \quad \#$$

(Convex functions: $f(x) = x^2$, $f(x) = x^{2k}$, $f(x) = e^{ax}$ with $a > 0$
or $f(x) = x^a$, for $x \geq 0$ and $a > 1$)

★ Application of Jensen inequality:

For $p \leq q < \infty$

$$\mathbb{E}(|X|^q) = \mathbb{E}\left(\left(|X|^p\right)^{q/p}\right) \geq \left(\mathbb{E}(|X|^p)\right)^{q/p}$$

$\frac{q}{p} > 1,\ |X|^p \geq 0$

$\Longrightarrow$ If $\mathbb{E}(|X|^q) < \infty \Longrightarrow \mathbb{E}(|X|^p) < \infty \quad \forall p \in (0,q]$.

- Markov inequality: X r.v. positive (≥ 0). Then $\forall \varepsilon > 0$

$$\mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}(X)}{\varepsilon}$$

Proof: $X \geq \varepsilon\, \mathbb{1}_{(X \geq \varepsilon)} \Longrightarrow \mathbb{E}(X) \geq \varepsilon \cdot \mathbb{E}(\mathbb{1}_{(X\geq\varepsilon)}) = \varepsilon\, \mathbb{P}(X \geq \varepsilon)$ #

- Chebychev inequality: X r.v. with finite mean and variance

then, $\forall \varepsilon > 0$

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geq \varepsilon\right) \leq \frac{Var(X)}{\varepsilon^2}$$

Proof: (Application of Markov inequality) let $Y = (X - \mathbb{E}(X))^2 \geq 0$.

Then $\mathbb{P}(Y > \varepsilon^2) \leq \frac{\mathbb{E}(Y)}{\varepsilon^2} = \frac{Var(X)}{\varepsilon^2}$

$$\mathbb{P}\left((X - \mathbb{E}(X))^2 > \varepsilon^2\right) = \mathbb{P}\left(|X - \mathbb{E}(X)| > \varepsilon\right) \quad \#$$

- Chernoff bounds (Exponential Markov inequality): X r.v. $\forall \varepsilon > 0$

1. $\mathbb{P}(X \geq \varepsilon) \leq e^{-t\varepsilon}\, \mathbb{E}(e^{tX}), \quad \forall t > 0$
2. $\mathbb{P}(X \leq \varepsilon) \leq e^{-t\varepsilon}\, \mathbb{E}(e^{tX}), \quad \forall t < 0$

Proof: $\mathbb{P}(X \geq \varepsilon) \overset{t>0}{=} \mathbb{P}(e^{Xt} \geq e^{\varepsilon t}) \leq e^{-t\varepsilon}\, \mathbb{E}(e^{Xt})$

$$\mathbb{P}(X \leq \varepsilon) \overset{t<0}{=} \mathbb{P}(e^{Xt} \geq e^{\varepsilon t}) \leq e^{-t\varepsilon}\, \mathbb{E}(e^{Xt})$$

#

Ex: Let $X \sim Po(20)$. Provide an estimate from above of $\mathbb{P}(X \geq 26)$ using Markov, Chebychev, Chernoff inequalities.

Sol: Recall that $\mathbb{E}(X) = Var(X) = 20$. Then

(MI) $\mathbb{P}(X \geq 26) = \frac{\mathbb{E}(X)}{26} = \frac{20}{26} \simeq 0.77$

(Cheb.I) $\mathbb{P}(X \geq 26) = \mathbb{P}(X - 20 \geq 6) \leq \mathbb{P}(|X-20| \geq 6) \leq \frac{Var(X)}{36} = \frac{20}{36} \simeq 0.$

(Cher I) $\mathbb{P}(X \geq 26) \leq e^{-26t}\mathbb{E}(e^{Xt}) \quad \forall t > 0$ ⊛

Compute $\mathbb{E}(e^{Xt}) = \sum_{k=0}^{\infty} e^{kt} \cdot \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!}$
for $X \sim Po(\lambda)$
$$= e^{-\lambda} \cdot e^{+\lambda e^t} = e^{\lambda(e^t - 1)}$$

Then ⊛ : $\mathbb{P}(X \geq 26) \leq e^{20(e^t-1)-26t} \quad \forall t > 0$

and we minimize over t: Set $h(t) = 20(e^t - 1) - 26t$

then $h'(t) = 20e^t - 26 \geq 0 \iff t \geq \ln\frac{26}{20} \implies t = \ln\frac{26}{20}$ minimum

$$\implies \mathbb{P}(X \geq 26) \leq e^{-26 \ln\frac{26}{20} + 6} \simeq 0.44$$

Generating function

We first state that the law of a r.v. X can be characterized by the average $\mathbb{E}(g(X))$ for any bounded and cont. function g. that is:

Theorem: $X \stackrel{d}{=} Y \iff \mathbb{E}(g(X)) = \mathbb{E}(g(Y))$ ⊛

$\forall g$ continuous and bounded

The above theorem applies also to d-dimensional r.v. (or random vector). However the class of continuous and bounded fcts is huge and property ⊛ may be difficult to check. It turns out, that the law of a r.v. can be characterized by the average of a single function:

) Characteristic function

Def: For X r.v. (real), and $\forall\, t \in \mathbb{R}$, the characteristic function of X is

$$\varphi_X(t) = \mathbb{E}\left[e^{itX}\right] = \int_{\mathbb{R}} e^{itx}\, dF_X(x) .$$

Remarks: $e^{itX} = \cos tX + i \sin tX$, then

- $|e^{itX}| = 1 \implies \varphi_X(t)$ is well defined $\forall\, t \in \mathbb{R}$
- $\varphi_X(0) = 1$ and $Re(\varphi_X(t)) \leq 1$, $Im(\varphi_X(t)) \leq 1$ $\forall t \in \mathbb{R}$

Examples

a. $X \sim Bi(n,p)$ $\qquad \mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \forall\, k \in \{0,\dots,n\}$

$$\varphi_X(t) = \mathbb{E}(e^{itX}) = \sum_{k=0}^{n} e^{itk} \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (pe^{it})^k (1-p)^{n-k} = (1-p+pe^{it})^n$$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

b. $X \sim N(0,1)$ $\quad f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

$$\varphi_X(t) = \int_{-\infty}^{+\infty} e^{itx} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx = e^{-\frac{t^2}{2}} \underbrace{\int_{-\infty}^{+\infty} \frac{e^{-\frac{1}{2}(x-it)^2}}{\sqrt{2\pi}} dx}_{=1} = \underline{e^{-\frac{t^2}{2}}}$$

c. $X \sim \Gamma(\alpha, \lambda)$ $\quad f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot e^{-\lambda x} \cdot \mathbb{1}(x \geq 0)$

$$\varphi_X(t) = \int_0^{+\infty} e^{itx} \frac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x} dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \underbrace{\int_0^{+\infty} e^{-(\lambda - it)x} \cdot x^{\alpha-1} dx}_{= \frac{\Gamma(\alpha)}{(\lambda - it)^\alpha}}$$

$$= \left(\frac{\lambda}{\lambda - it}\right)^\alpha \quad \#$$

inverse normalization constant of $\Gamma(\alpha, \lambda - it)$.

Basic properties

a. φ_X is continuous in t.

b. If $Y = aX + b \Rightarrow \varphi_Y(t) = e^{ibt} \cdot \varphi_X(at)$

c. If X and Y indep $\Rightarrow \varphi_{X+Y}(t) = \varphi_X(t) \cdot \varphi_Y(t)$

$\hookrightarrow$ generalized to sum $X_1 + \dots + X_n$ of indep. r.v.:

$$\varphi_{X_1 + \dots X_n}(t) = \varphi_{X_1}(t) \cdot \dots \cdot \varphi_{X_n}(t)$$

Proof: a. follows from continuity of e^{itx}

$$|\varphi_X(t+h) - \varphi_X(t)| \leq \int_{\mathbb{R}} |e^{ix(t+h)} - e^{ixt}| \, dF_X(x) = \int_{\mathbb{R}} \underbrace{|e^{ixh} - 1|}_{\leq 2} \, dF_X(x)$$

Since $|e^{ixh} - 1| \xrightarrow[h \to 0]{} 0$, the result follows by dominated convergence theorem.

b. proved by trivial check.

c. $\varphi_{X+Y}(t) = \mathbb{E}\left[e^{it(X+Y)}\right] \overset{\text{by independence}}{=} \mathbb{E}\left[e^{itX}\right]\mathbb{E}\left[e^{itY}\right] = \varphi_X(t)\varphi_Y(t)$ #

Examples:

- Let $X \sim N(\mu, \sigma^2)$ and compute φ_X.

  For $Z \sim N(0,1)$, we know: $X = \sigma Z + \mu$; $\varphi_Z(t) = e^{-\frac{t^2}{2}}$

  then by b.: $\varphi_X(t) = e^{it\mu} \cdot \varphi_Z(\sigma t) = e^{it\mu - \frac{\sigma^2}{2}t^2}$

- Let $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$ independent and compute φ_{X+Y}:

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) = e^{it\mu_1 - \frac{\sigma_1^2}{2}t} \cdot e^{it\mu_2 - \frac{\sigma_2^2}{2}t}$$

$$= e^{it(\mu_1+\mu_2) - \frac{(\sigma_1^2+\sigma_2^2)}{2}t}$$

#

Remark: If X is absolute continuous r.v with density f_X, then $\varphi_X(t) = \int_{\mathbb{R}} e^{itx} \cdot f_X(x)\,dx$

$\Rightarrow$ φ_X corresponds to the Fourier transform of f_X

$\Rightarrow$ by the inversion theorem, for any point x of continuity of f_X, one has

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \lim_{T\to\infty} \int_{-T}^{T} \varphi_X(t) e^{-itx}\,dt.$$

This can be generalized to r.v. as follows:

<u>Theorem 2</u> Let X r.v. with distribution F_X.
then for all x continuity-point of F_X, with $z > x$, it holds

$$F(z) - F(x) = \frac{1}{2\pi} \lim_{T \to \infty} \int_{-T}^{T} \varphi_X(t) \cdot \frac{e^{-itz} - e^{-itx}}{it} \, dt$$

#

In particular, the law of a r.v is completely characterized by its <u>characteristic function</u>:

<u>Theorem 3</u>: $X \stackrel{d}{=} Y \iff \varphi_X = \varphi_Y$ #

Consequences (examples):

1. If $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$ independent

$$\implies X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

Indeed, whe have shown (by independence) that

$$\varphi_{X+Y}(t) = \varphi_X(t) \cdot \varphi_Y(t) = e^{i(\mu_1+\mu_2)t - \frac{(\sigma_1^2+\sigma_2^2)}{2}t^2}$$

But this is indeed the characteristic fct of $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

2. Let $X \sim \chi^2(1) \equiv \Gamma(\frac{1}{2}, \frac{1}{2})$

then $\varphi_X(t) = \left(\frac{1}{1-2it}\right)^{1/2} = \frac{1}{(1-2it)^{1/2}}$

If $X_1, \dots, X_n$ iid $\sim \chi^2(1)$

$$\implies \varphi_{X_1 + \dots + X_n}(t) = (\varphi_{X_i}(t))^n = \frac{1}{(1-2it)^{n/2}} \equiv \varphi_W(t)$$

So $W \sim \chi^2(n) \equiv \Gamma(\frac{n}{2}, \frac{1}{2})$ #

At last, notice that from the expansion

$$e^{itx} = \sum_{j=0}^{\infty} \frac{(itx)^j}{j!} = 1 + \sum_{j=1}^{\infty} \frac{(itx)^j}{j!}$$

replacing $x \to X$ and averaging, we get

<u>Theorem 4</u>: Let X r.v. and $m \in \mathbb{N}$ s.t. $\mathbb{E}(X^m) < \infty$.
Then φ_X is m-times differentiable at 0, with

$$\varphi_X^{(m)}(0) = \frac{d^m}{dt^m}(\varphi_X(t))\Big|_{t=0} = (i)^m \cdot \mathbb{E}(X^m)$$

and in particular $\varphi_X(t) = 1 + \sum_{j=1}^{m} \frac{(it)^j}{j!} \mathbb{E}(X^j) + o(t^m)$ as $t \to 0$

| taylor formula

Viceversa, if $\varphi_X^{(2m)}(0)$ exists $\Rightarrow \mathbb{E}(X^{2m}) < \infty$ for $m \in \mathbb{N}$.

<u>Examples</u>

1. Let $X \sim N(\mu, \sigma^2)$ with $\varphi_X(t) = e^{i\mu t - \frac{\sigma^2}{2}t^2}$

$\Rightarrow \varphi_X$ is m-differentiable at 0 $\forall m \in \mathbb{N}$

$\Rightarrow$ All moments of X are finite!

$$\mathbb{E}(X) = \frac{\varphi'(0)}{i} = \frac{(i\mu - \sigma^2 t)e^{i\mu t - \frac{\sigma^2}{2}t^2}}{i}\Big|_{t=0} = \mu$$

$$\mathbb{E}(X^2) = \frac{\varphi''(0)}{i^2} = -\varphi''(0) = \dots = (\mu^2 + \sigma^2)$$

and so on.

2: Let $X \sim \Gamma(\alpha, \lambda)$ with $\varphi_X(t) = \frac{\lambda^\alpha}{(\lambda - it)^\alpha}$

$\Rightarrow \varphi_s$ is m-differentiable at 0 $\forall m \in \mathbb{N}$

$\Rightarrow$ All moments of X are finite!

It holds $\varphi^{(m)}(t) = (i)^m [\alpha(\alpha+1) \cdots (\alpha+m-1)] \cdot \frac{\lambda^\alpha}{(\lambda - it)^{\alpha+m}}$

and then $\varphi^{(m)}(0) = (i)^{m+1} \frac{(\alpha+m-1)!}{(\alpha-1)!} \cdot \frac{1}{\lambda^m}$

$$\Rightarrow \mathbb{E}(X^m) = \frac{(\alpha+m-1)!}{(\alpha-1)!} \frac{1}{\lambda^m}$$

Generalization to random vectors

Def: Let $X : \Omega \to \mathbb{R}^d$, $d \geq 1$.

Then $\forall\, t = (t_1, \dots, t_d) \in \mathbb{R}^d$, define the characteristic fct

$$\varphi_X(t) := \mathbb{E}\left[e^{i t \cdot X}\right] = \mathbb{E}\left[e^{i(t_1 X_1 + \dots + t_d X_d)}\right]$$

for $X = (X_1, \dots, X_d)$.

All the results seen for r.v., generalize to r. vector:

- $|\varphi_X(t)| \leq 1$ and it is continuous in all components
- if $Y = AX + b$, for $A \in M_{d \times d}$, $b \in \mathbb{R}^d$, then
  $\varphi_Y(t) = e^{i t \cdot b} \cdot \varphi_X(At)$
- if X, Y indep. $\Rightarrow \varphi_{X+Y}(t) = \varphi_X(t) \varphi_Y(t)$
- $X \overset{d}{=} Y \Leftrightarrow \varphi_X = \varphi_Y$

#

<u>Example</u>: let $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ indep,
and consider $X = (X_1, X_2) \in \mathbb{R}^2$

then $\varphi_X(\underline{t}) = \mathbb{E}\left[e^{it_1X_1 + it_2X_2}\right] = \mathbb{E}\left[e^{it_1X_1}\right]\mathbb{E}\left[e^{it_2X_2}\right]$

$\underline{t} = (t_1, t_2) \in \mathbb{R}^2$

$$= e^{i(t_1\mu_1 + t_2\mu_2) - \frac{1}{2}(t_1^2\sigma_1^2 + t_2^2\sigma_2^2)}$$

$$= e^{i\underline{t}\cdot\mu - \frac{1}{2}A\underline{t}\cdot\underline{t}}$$

with $\mu = (\mu_1, \mu_2)$ and $A = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$.

the random vector X is a Gaussian vector in $\mathbb{R}^2$.

---

3) Moment generating function

<u>Def</u>: X r.v. (real). For all $t \in \mathbb{R}$: $M_X(t) : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ s.t.

$$M_X(t) := \mathbb{E}(e^{tX}) = \int_{\mathbb{R}} e^{tx}\, dF_X(x).$$

<u>Remarks</u>:
- $M_X(0) = 1$
- $M_X(t) = \varphi_X(-it)$

In spite of the characteristic function, $M_X(t)$ may be ∞.

<u>Def</u>: If $\exists\ \delta_X$ s.t. $M_X(t) < +\infty$ $\forall\ |t| < \delta_X$, we say that M_X has radius of convergence δ_X.

<u>Theorem 5</u>: Let X r.v. s.t. M_X has radius of convergence $\delta_X > 0$.

then:
i. $\mathbb{E}(X^n) = M_X^{(n)}(0)$ $\forall\ n \in \mathbb{N}$
ii. $M_X(t) = \sum_{n=0}^{\infty} \mathbb{E}(X^n) \cdot \frac{t^n}{n!}$

\#

Properties of M_X (special case of properties of φ_X)

1. If $|t| < \delta_X$, radius of convergence of M_X, and $Y = aX + b$

$$\Rightarrow M_Y(t) = e^{at} M_X(bt) \quad , \quad \forall |t| < \frac{\delta_X}{|b|}$$

2. If X and Y are independent

$$\Rightarrow M_{X+Y}(t) = M_X(t) M_Y(t) \quad \forall |t| = \min\{\delta_X, \delta_Y\}$$

↳ generalised to sum of indep. $X_1, \dots, X_n$.

3. $X \overset{d}{=} Y \iff \exists \, \delta > 0$ s.t. $M_X(t) = M_Y(t) \quad \forall |t| < \delta$

Examples

1. $X \sim \text{Exp}(\lambda)$. Then

$$M_X(t) = \int_0^{+\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{+\infty} e^{-(\lambda - t)x} dx$$

$$= \begin{cases} +\infty & \text{if } t \geq \lambda \\ \frac{\lambda}{\lambda - t} & \text{if } t < \lambda \end{cases} \Rightarrow \delta_X = \lambda$$

2. Let $X \sim N(\mu, \sigma^2)$ and $Y = e^X$

Notice that $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}\, y} e^{-\frac{1}{2\sigma^2}(\log y - \mu)^2} \mathbb{1}(y > 0)$

Y has <u>log-normal distribution</u>.

In particular: $\mathbb{E}(Y^k) = \mathbb{E}[e^{k \cdot X}] = M_X(k) = e^{k\mu + \frac{\sigma^2}{2}k^2}$
$\forall k \in \mathbb{N}$

that is, Y has finite k-moment $\forall k \in \mathbb{N}$.

However, $M_Y(t) = +\infty \quad \forall |t|>0$, that is $\delta_Y = 0$.

Indeed (Consider $X \sim N(0,1)$ to simplify)

$$M_Y(t) = \mathbb{E}[e^{tY}] = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{ty} \cdot \frac{1}{y} e^{-\frac{1}{2}(\lg y)^2} dy \quad \overset{\frac{1}{y} = e^{-\lg y}}{=}$$

$$= \underbrace{\frac{e^{1/2}}{\sqrt{2\pi}}}_{c} \int_0^{+\infty} e^{ty} e^{-\frac{1}{2}(\lg y + 1)^2} dy = c \int_0^{+\infty} \underbrace{e^{y\left(t - \frac{1}{2}\frac{(\lg y+1)^2}{y}\right)}}_{\geq 0} dy \overset{*}{\geq}$$

Notice that $\frac{1}{2}\frac{(\lg y + 1)^2}{y} \xrightarrow[y \to +\infty]{} 0$, that is

$\forall \varepsilon > 0$, $\exists y_0$ s.t. $\frac{1}{2}\frac{(\lg y + 1)^2}{y} < \varepsilon \quad \forall y \geq y_0$

$$\overset{*}{\geq} c \int_{y_0}^{+\infty} e^{y\left(t - \frac{1}{2}\frac{(\lg y+1)^2}{y}\right)} dy \geq c \int_{y_0}^{+\infty} e^{y(t-\varepsilon)} dy = +\infty \quad \forall t > \varepsilon$$

$$\Longrightarrow \mathbb{E}(e^{tY}) = +\infty \quad \forall t > 0 \quad (\text{taking } \varepsilon \to 0) \;\#$$

---

3. Let $Y = \sum_{k=1}^{N} X_k$, where $(X_k)_{k \in \mathbb{N}}$ are iid and N is r.v. with value on $\mathbb{N}^+$, indep. of X_k. Compute M_Y as a function of M_{X_k}'s.

$$M_Y(t) = \mathbb{E}\left(e^{t\sum_{k=1}^{N} X_k}\right) = \sum_{n=1}^{\infty} \mathbb{E}\left(e^{t\sum_{k=1}^{N} X_k} \mid N=n\right) \mathbb{P}(N=n)$$

$$= \sum_{n=1}^{\infty} \mathbb{E}\left[e^{t\sum_{k=1}^{n} X_k}\right] \cdot \mathbb{P}(N=n) = \sum_{n=1}^{\infty} (M_X(t))^n \, \mathbb{P}(N=n)$$

$$= \mathbb{E}\left[M_X(t)^N\right] . \quad \#$$

4) Stable distribution

By the properties of the characteristic function, if $X_1, \dots, X_n$ are iid then

$$\varphi_{\frac{X_1+\dots+X_n}{\sqrt{n}}}(t) = \varphi_{X_1+\dots+X_n}\left(t/\sqrt{n}\right) = \left(\varphi_{X_1}\left(t/\sqrt{n}\right)\right)^n$$

In particular if X_k's $\sim N(0,1)$ then

$$\varphi_{\frac{X_1+\dots+X_n}{\sqrt{n}}}(t) = \left(e^{-\frac{1}{2}\left(\frac{t}{\sqrt{n}}\right)^2}\right)^n = e^{-\frac{1}{2}t} = \varphi_{X_1}(t)$$

that implies that $\frac{X_1+\dots+X_n}{\sqrt{n}} \stackrel{d}{=} X_1$ (*)

there is a class of RV's, called α-stable RV's, that are such that (*) is replaced by the more general

$$\frac{X_1+\dots+X_n}{n^{1/\alpha}} \stackrel{d}{=} X_1 \qquad (*)_1$$

for $\alpha \in [0,2]$.

For example if X_k's are s.t. $\varphi_{X_1}(t) = e^{-c|t|^\alpha}$

$$\Rightarrow \varphi_{\frac{X_1+\dots+X_n}{n^{1/\alpha}}}(t) = \left(e^{-c\frac{|t|^\alpha}{n}}\right)^n = e^{-c|t|^\alpha} = \varphi_{X_1} \Rightarrow (*)_1$$

More formally:

<u>Def</u>: Let $(X_n)_{n\in\mathbb{N}}$ iid RV's. Their distribution is stable of parameter α if

(A) $$X_1+\dots+X_n \stackrel{d}{=} n^{1/\alpha} X_1 + b_n \qquad \left[\frac{X_1+\dots+X_n - b_n}{n^{1/\alpha}} \stackrel{d}{=} X_1\right]$$

where $(b_n)_{n\in\mathbb{N}}$ is a sequence

(when $b_n \equiv 0$, then the distribution is <u>strictly stable</u>)

Examples

- $N(0, \sigma^2)$ is (strictly) stable with $\alpha = 2$

  [and $N(\mu, \sigma^2)$ is stable with $\alpha = 2$, taking $b_n \equiv \mu(n - \sqrt{n})$]

- Cauchy distribution (with density $\frac{1}{\pi} \frac{1}{1+x^2}$ on $\mathbb{R}$)

  is stable with $\alpha = 1$

- In most of the cases the density is not explicit and the distribution is characterized by the characteristic function or through the identity in distribution (A).

Their importance is mainly related to the following

Theorem: Let $X_1, \dots, X_n, \dots$ iid RV's and assume that there exist sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ s.t.

$$\frac{(X_1 + \dots + X_n) - b_n}{a_n} \xrightarrow[n\to\infty]{d} F$$

where F is a non-trivial distribution.

Then F is an α-stable distribution and we say that the distribution of X_k belong to domain of attraction of the α-stable distribution.

For example, if X_k's are such that

$$\mathbb{P}(|X_k| > x) \simeq c \frac{1}{x^\alpha} \quad \text{as } x \to \infty, \quad \left(\begin{array}{c}\text{heavy tailed} \\ \text{RV's}\end{array}\right)$$

that (X_k) are in the domain of attraction of α-stable RV.