

VII / Moments and moment generating function

Let (Ω, \mathcal{F}, P) be a prob. space

X a random variable, F_x its distribution function

$$F_x(x) = P(X \leq x)$$

Thm : (decomposition thm of Lebesgue)

$$\exists ! F_{pp}, F_{ac}, F_{sc} : \mathbb{R} \rightarrow [0, 1]$$

$$\text{s.t } F_x = \alpha_{pp} F_{pp} + \alpha_{sc} F_{sc} + \alpha_{ac} F_{ac}$$

$$\alpha_j \in [0, 1]$$

$$\alpha_{pp} + \alpha_{ac} + \alpha_{sc} = 1$$

$$\text{with } F_{pp}(x) = \sum_{y \in \text{Im}(X), y \leq x} P(y), \quad \sum_{y \in \text{Im}(X)} P(y) = 1$$

$$\text{with } F_{ac}(x) = \int_{-\infty}^x f(y) dy, \quad \int_{-\infty}^{\infty} f(y) dy = 1$$

• with F_{sc} is differentiable function almost everywhere

$$F_{sc}'(x) = 0 \text{ whenever it is defined}$$

$$\text{and } \lim_{x \rightarrow -\infty} F_{sc}(x) = 0 \text{ and } \lim_{x \rightarrow +\infty} F_{sc}(x) = 1$$

Remark : Given F_x and given $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ measurable, we can define $\mathbb{E}(\varphi(X))$ as a limit of expression of the form:

$$\sum_j (F_x(x_{j+1}) - F_x(x_j)) \varphi(y_j)$$

If F_x is abs. cont, this converges to $\int_{-\infty}^{\infty} F_x(x) \varphi(x) dx$
and in the discrete case, one gets $\sum_{y \in \text{Im}(X)} P_x(y) \varphi(y)$.

For any $k \in \mathbb{N}$, we define the moments $\mathbb{E}(X^k)$ whenever the integral converges absolutely (i.e replace φ by $|t|^k$)

Remark : Assume that $\mathbb{E}(X^k)$ exists for all $k \in \mathbb{N}$ and that $G(t) = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \mathbb{E}(X^k)$ exists for $t \in (-\delta, \delta)$ for $\delta > 0$, then this function defines uniquely all the moments, and also F_x .

Recall that $\text{var}(X) = E((X - E(X))^2)$, it satisfies

$$\text{var}(ax + b) = a^2 \text{var}(X)$$

Since it is not linear in a , one often defines

$\sqrt{\text{var}(X)}$:= standard deviation

If X, Y are r.v., then:

$$\begin{aligned}\text{var}(X+Y) &= E((X+Y - E(X+Y))^2) \\ &= E((((X - E(X)) + (Y - E(Y)))^2)\end{aligned}$$

$$= \text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y)$$

$$\begin{aligned}\text{with } \text{cov}(X, Y) &= E((X - E(X))(Y - E(Y))) \\ &= E(XY) - E(X)E(Y)\end{aligned}$$

$\Rightarrow \text{cov}(X, Y) = 0$ if X and Y are independent

Def. We set $\rho'(X, Y) := \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)} \cdot \sqrt{\text{var}(Y)}}$, called correlation coeff

Property: $\rho'(ax+b, cY+d) = \rho'(X, Y)$
Scaling invariance

Lemma: $-1 \leq \rho'(X, Y) \leq 1$

Lemma: Cauchy-Schwarz inequality

For any r.v. X, Y : $E(XY)^2 \leq E(X^2)E(Y^2)$

Sketch of the proof: Set $W = X + SY$ with $S \in \mathbb{R}$ and study

$E(W^2) \geq 0 \Rightarrow$ discriminant: $E(XY)^2 - E(X^2)E(Y^2) \leq 0$

Proof: set $X' := X - E(X)$

$Y' := Y - E(Y)$

By C.S.I: $E(X'Y')^2 \leq E(X'^2)E(Y'^2)$

$\Rightarrow \text{cov}(X', Y')^2 \leq \text{var}(X) \cdot \text{var}(Y)$. \square

If we look more carefully, we get:

$\rho'(X, Y) = 0$ when X, Y are ind.

$\rho'(X, Y) = 1$ iff. $Y = aX + \beta$ with $a > 0$

" " = -1 "

" " < "

We say that X and Y are uncorrelated if $\text{cov}(X, Y) = 0$

Def: The moment generating function of a r.v. X is the function:

$$M_X: I \rightarrow \mathbb{R} \text{ given by}$$

$$M_X(t) = E(e^{tx}), t \in I \text{ (interval around 0)}$$

Remark: If X is a.c., then $M_X(t) = \int_{\mathbb{R}} e^{tx} f_X(x) dx$
(Laplace transform of f_X)

Thm: If M_X exists for $t \in (-s, s)$ with $s > 0$ then
 $E(X^k) = M_X^{(k)}(0)$

Properties: 1, $M_{ax+b}(t) = e^{tb} M_X(at)$

2, $M_{x+y}(t) = M_X(t) M_Y(t)$, if X, Y are ind.

Thm: Markov's inequality

X is a r.v. w.r.t (Ω, \mathcal{F}, P) with $E(X) < \infty$ and $\text{var}(X) > 0$

Then:

$$P(X > x) \leq \frac{E(X)}{x} \text{ for any } x \in \mathbb{R}_+$$

Proof: (Ω, \mathcal{F}, P) . For any $w \in \Omega$, one has:

$$\begin{aligned} X(w) &\geq \begin{cases} x & \text{if } X(w) > x \\ 0 & \text{if } X(w) \leq x \end{cases} \\ &= x \chi_A(w), \text{ where } A = \{w \in \Omega \mid X(w) > x\} \end{aligned}$$

It means: $X > x \chi_A$

$$\Rightarrow E(X) > x E(\chi_A) = x P(A) = x P(X > x) \quad \square$$

Application: For a random variable X , we call the median m , any $m \in \mathbb{R}$ s.t $P(X < m) \leq \frac{1}{2} \leq P(X \leq m)$

If X is non-negative, one has

$$\frac{1}{2} \leq P(X \geq m) \leq \frac{E(X)}{m}$$

$$\Rightarrow m \leq 2E(X)$$

Recall: a convex function $g: (a, b) \rightarrow \mathbb{R}$, satisfies

$$g((1-t)u + tv) \leq (1-t)g(u) + t g(v) \text{ for any } t \in (0, 1),$$

for any $a \leq u \leq v \leq b$

Thm: (Supporting tangent thm)

If $g: (a, b) \rightarrow \mathbb{R}$ is convex then for any $v \in (a, b)$, there exists $\alpha \in \mathbb{R}$ s.t

$$g(x) \geq g(u) + \alpha(x - u) \quad \forall x \in (a, b).$$

Thm: Jensen's inequality

Let X be a r.v s.t $E(X)$ exists

let $a, b \in \mathbb{R}$ s.t $E(X) \in [a, b]$

and let $g: (a, b) \rightarrow \mathbb{R}$ be convex and s.t $E(g(X))$ exists.

Then $E(g(X)) \geq g(E(X))$

Proof: By the previous thm for $u = E(X)$ there exist $\alpha \in \mathbb{R}$ s.t.

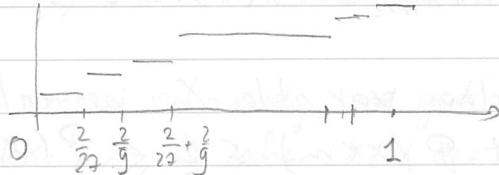
$$g(x) \geq g(E(X)) + \alpha(x - E(X)) \quad \forall x \in (a, b)$$

$$\Rightarrow g(X) \geq g(E(X)) + \alpha(X - E(X))$$

Take E : $E(g(X)) \geq g(E(X)) + 0 \quad \square$

A singular example: Play with a coin, and receive $\frac{2}{3^j}$ yen if it is a face at the j^{th} toss, and 0 if it a tail. What is the distribution function of the total gain after infinite sequence of tosses.

$$\sum_{j=1}^{\infty} \frac{2}{3^j} = 1$$



Cantor function