260

Espérance, variance et moments de variables aléatoires

$(\Omega, \mathcal{F}, \mathbb{P})$ sur lequel X, Y sont des variables aléatoires réelles

I - Espérance d'une variable aléatoire

1) Propriétés de l'espérance

Déf 1 Si X est intégrable $\left(\int_\Omega |X(\omega)| d\mathbb{P}(\omega) < \infty\right)$, l'espérance de X est le réel $\mathbb{E}(X) = \int_\Omega X(\omega) d\mathbb{P}(\omega) =: \int_\Omega X d\mathbb{P}$

Rmq 2 • Si $\mathbb{E}(X) = 0$, on dit que X est centrée.
• $\forall a, b \in \mathbb{R}$, $\mathbb{E}[aX + bY] = a\mathbb{E}(X) + b\,\mathbb{E}(Y)$

Ex 3 Si $X = c^{ste}$ ps, alors $\mathbb{E}(X) = c^{ste}$.

Th 4 (transfert) Soit $\varphi : \mathbb{R} \to \mathbb{R}$ borélienne, alors
(i) si φ positive, alors $\mathbb{E}[\varphi(X)] = \int_{\mathbb{R}} \varphi(x)\, d\mathbb{P}_X(x)$
(ii) si φ à valeurs quelconques, $\varphi(X) \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \Leftrightarrow \varphi \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$ et l'égalité précédente est vérifiée.

Ex 5 On a $\mathbb{E}[\mathbb{1}_A(X)] = \mathbb{P}(X \in A)$, pour $A \in \mathcal{B}(\mathbb{R})$

Th 6 (Jensen) Soit $\Phi : \mathbb{R} \to \mathbb{R}$ convexe et X variable aléatoire réelle telle que X et $\varphi(X)$ soient intégrables, alors $\mathbb{E}[\Phi(X)] \geq \Phi(\mathbb{E}(X))$.

Ex 7 On a $\mathbb{E}[|X|] \geq |\mathbb{E}(X)|$ pour $X \in L^1$
et $\mathbb{E}(X^2) \geq (\mathbb{E}(X))^2$ pour $X \in L^2$

Th 8 (Markov) Soit $a > 0$ et $X \in L^1$, on a
$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

2) Calculs d'espérance

Déf 9 Pour $X = (X_1, \dots, X_d) \in \mathbb{R}^d$ vecteur aléatoire, son espérance est le vecteur $\mathbb{E}(X) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_d))$.

Th 10 Une famille quelconque de variables aléatoires réelles $(X_i)_I$ est indépendante ssi pour tout $J \subset I$, fini, et pour toute famille $(\Phi_j)_J$ de fonctions boréliennes telles que $\forall j \in J$, $\Phi_j(X_j) \in L^1$, on a
$$\mathbb{E}\left[\prod_{j\in J} \Phi_j(X_j)\right] = \prod_{j \in J} \mathbb{E}[\Phi_j(X_j)].$$

Cor 11 Si $X_1, \dots, X_n$ sont mutuellement indépendantes, alors
$$\mathbb{E}[X_1 \cdots X_n] = \mathbb{E}[X_1] \cdots \mathbb{E}[X_n]$$

Prop 12 Si X est discrète, alors $\mathbb{E}[X] = \sum_{x \in Val(X)} x\, \mathbb{P}(X = x)$

Ex 13 • $X \sim b(p)$, $\mathbb{E}(X) = p$ • $X \sim \mathcal{B}(n,p)$, $\mathbb{E}(X) = np$
• $X \sim \mathcal{G}(p)$, $\mathbb{E}(X) = 1/p$ • $X \sim \mathcal{P}(\lambda)$, $\mathbb{E}(X) = \lambda$

Prop 14 Si X admet une densité f_X et si $x \mapsto |x| f_X(x)$ est dans $L^1(\mathbb{R})$, alors $\mathbb{E}(X) = \int_{\mathbb{R}} f_X(x)\, dx$.

Ex 15 • $X \sim \mathcal{U}([a,b])$, $\mathbb{E}(X) = \frac{b-a}{2}$
• $X \sim \mathcal{N}(m, \sigma^2)$, $\mathbb{E}(X) = m$
• $X \sim \mathcal{E}(\lambda)$, $\mathbb{E}(X) = \lambda^{-1}$

C-ex 16 La loi de Cauchy de densité $f(x) = \frac{1}{\pi(1+x^2)}$ n'admet pas d'espérance.

3) Espérance conditionnelle

Déf 17 Soit $\mathcal{B}$ une sous-tribu de $\mathcal{F}$ et $X \in L^1$. Alors, il existe une variable aléatoire, ps unique, notée $\mathbb{E}[X|\mathcal{B}]$ telle que
• $\omega \mapsto \mathbb{E}[X|\mathcal{B}](\omega)$ est $\mathcal{B}$-mesurable
• $\forall B \in \mathcal{B}$, $\int_B \mathbb{E}[X|\mathcal{B}]\, d\mathbb{P} = \int_B X\, d\mathbb{P}$.

Prop 18 L'espérance conditionnelle est linéaire et positive.

Prop 19 Si $\mathcal{B}$ est indépendante de $\sigma(X)$, alors $\mathbb{E}[X|\mathcal{B}] \overset{ps.}{=} \mathbb{E}(X)$

Prop 20 Si Y est $\mathcal{B}$-mesurable et $XY \in L^1$, alors
$\mathbb{E}[XY|\mathcal{B}] = Y\mathbb{E}[X|\mathcal{B}]$.

Ex 21
• Si $X_1 \sim \mathcal{B}(n_1, p)$, $X_2 \sim \mathcal{B}(n_2, p)$ et $X_1 \perp\!\!\!\perp X_2$, $\mathbb{E}[X_1 | X_1 + X_2] \overset{ps.}{=} \frac{n_1(X_1+X_2)}{n_1+n_2}$
• Si $X_1 \sim \mathcal{P}(\lambda_1)$, $X_2 \sim \mathcal{P}(\lambda_2)$, $X_1 \perp\!\!\!\perp X_2$, alors $\mathbb{E}[X_1|X_1+X_2] \overset{ps.}{=} \frac{\lambda_1(X_1+X_2)}{\lambda_1+\lambda_2}$

II - Moments d'ordre supérieur

1) Variance

Déf 22 Pour $X \in L^2$, sa variance est : $Var(X) := \mathbb{E}[(X - \mathbb{E}(X))^2]$.

Rmq 23 On peut écrire : $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.
$= \|X - \mathbb{E}(X)\|^2_{L^2(\Omega, \mathcal{F}, \mathbb{P})}$.

Ex 24 · $X \sim b(p)$, $\mathrm{Var}(X) = p(1-p)$ · $X \sim \mathcal{N}(m, \sigma^2)$, $\mathrm{Var}(X) = \sigma^2$
· $X \sim \mathcal{P}(\lambda)$, $\mathrm{Var}(X) = \lambda$ · $X \sim \mathcal{E}(\lambda)$, $\mathrm{Var}(X) = 1/\lambda^2$.

Prop 25 Pour tout $\alpha \in \mathbb{R}$, $\mathrm{Var}(\alpha X) = \alpha^2 \mathrm{Var}(X)$
$\mathrm{Var}(X+\alpha) = \mathrm{Var}(X)$.

C-ex 26 La loi de Cauchy n'admet pas de variance.

Prop 27 $\mathrm{Var}(X) = 0 \iff X = c^{ste}$ ps

Th 28 (Tchebychev) Soit $X \in L^2$ et $a > 0$, on a
$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}.$$

Th 29 (Cauchy-Schwarz) Si $X, Y \in L^2$ alors $XY \in L^1$ et
$$\mathbb{E}[|XY|]^2 \leq \mathbb{E}(X^2)\,\mathbb{E}(Y^2).$$

2) Variance et indépendance

Déf 30 $X, Y \in L^2$ sont dites non-corrélées ssi $\mathrm{Cov}(X,Y) = 0$,
où $\mathrm{Cov}(X,Y) := \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))]$.

Rmq 31 L'indépendance implique la non-corrélation.

C-ex 32 $X \sim \mathcal{N}(0,1)$ et $Y = X^2$ sont non-corrélées, mais pas indépendantes.

Prop 33 Si $X_1, \dots, X_n$ sont deux à deux non-corrélées, alors
$$\mathrm{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathrm{Var}(X_i).$$

App 34 $X \sim \mathcal{B}(n,p)$, alors $\mathbb{E}[(X - \mathbb{E}(X))^2] = \mathrm{Var}(X) = np(1-p)$.

App 35 (Weierstraß) Soit $f : [0,1] \to \mathbb{C}$ continue et $\omega : h \mapsto \sup\{|f(x) - f(y)| : |x-y| \leq h\}$. Pour $n \in \mathbb{N}^*$, on pose $B_n = \sum_{k=0}^n \binom{n}{k} X^k (1-X)^{n-k} f(k/n)$, alors

i) $B_n \xrightarrow[{[0,1]}]{\text{cvu}} f$ et il existe $C > 0$ tel que $\|f - B_n\|_\infty \leq C\,\omega\left(\frac{1}{\sqrt{n}}\right)$.

ii) Cette majoration est optimale au sens où pour une certaine fonction f lipschitzienne, il existe $\delta > 0$ tel que pour tout $n \in \mathbb{N}^*$, $\|f - B_n\|_\infty \geq \delta\,\omega\left(\frac{1}{\sqrt{n}}\right)$.

DEV 1

3) Moments d'ordre p

Déf 36 Le moment d'ordre $p \geq 1$ de X est $\mathbb{E}(X^p) = \int_\Omega X^p \, d\mathbb{P}$
(il est bien défini dès que $X \in L^p$)

Prop 37 (Hölder) Soit $X \in L^p$ et $Y \in L^q$, $p, q \geq 1$ conjugués, alors $XY \in L^1$ et $\mathbb{E}(|XY|) \leq \mathbb{E}(|X|^p)^{1/p}\,\mathbb{E}(|X|^q)^{1/q}$.

Prop 38 L'application $p \mapsto \mathbb{E}(|X|^p)^{1/p}$ est croissante.

App 39 $\forall 1 \leq p \leq q$, $L^q \subset L^p$.

Prop 40 Pour X variable aléatoire positive, alors pour tout $p \in]0, +\infty[$, $\mathbb{E}(X^p) = p \int_0^{+\infty} t^{p-1} \mathbb{P}(X > t)\, dt$.

Cor 41 Pour X variable aléatoire positive on a
$$X \in L^1 \iff \text{pour un, ou tout, } \varepsilon > 0, \begin{cases} \sum_{n \geq 0} \mathbb{P}(X > n\varepsilon) < \infty \\ \text{ou} \\ \sum_{n \geq 0} 2^n \mathbb{P}(X > 2^n \varepsilon) < \infty \end{cases}$$

III - Utilisation des moments

1) Fonction génératrice

Dans ce paragraphe, X est à valeurs dans $\mathbb{N}$.

Déf 42 Pour $s \in [1,1]$, $G_X(s) := \mathbb{E}[s^X]$ est la fonction génératrice de X.

Prop 43 1) $\forall s \in [-1,1]$, $|G_X(s)| \leq 1$ et $G_X(1) = 1$.
2) $\forall s \in [-1,1]$, $G_X(s) = \sum_{n=0}^\infty \mathbb{P}(X=n)\, s^n$.
3) G_X est continue sur $[-1,1]$ et C^∞ sur $]-1,1[$.

Th 44 G_X caractérise la loi de X et $\mathbb{P}(X=n) = \frac{G_X^{(n)}(0)}{n!}$, $\forall n \geq 0$.

Ex 45 · $X \sim b(p)$, $G_X(s) = 1 - p + ps$
· $X \sim \mathcal{G}(p)$, $G_X(s) = \frac{ps}{1-(1-p)s}$ · $X \sim \mathcal{P}(\lambda)$, $G_X(s) = \exp(\lambda(s-1))$

Prop 46 : Si $X \perp\!\!\!\perp Y$, alors $G_{X+Y} = G_X G_Y$

Ex 47 : $X \hookrightarrow \mathcal{B}(n,p)$, $G_X(s) = (ps+1-p)^n$, $\forall s \in [-1;1]$

Thm 48 : (Calcul des moments). X admet un moment d'ordre $p \in \mathbb{N}^*$ ssi G_X est p-dérivable en 1^-. Dans ce cas :

$G_X^{(p)}(1^-) = \sum_{k=p}^{+\infty} k(k-1)\dots(k-p+1)p_k = \mathbb{E}[X(X-1)\dots(X-p+1)]$

En particulier, $\mathbb{E}[X] = G_X'(1^-)$.

Exple 49 : Soient $X \hookrightarrow b(p)$ et $Y \hookrightarrow \mathcal{G}(a)$, $(a,p) \in (]0;1[)^2$, $X \perp\!\!\!\perp Y$. Soit U, va égale à 0 si $X=0$, Y si $X=1$. Alors $\mathbb{E}[U] = \frac{p}{a}$ et $\mathbb{E}[U^2] = \frac{p(2-a)}{a^2}$

2) Fonction caractéristique

Déf 50 : On appelle la fonction caractéristique de X

$\varphi_X : \begin{array}{l} \mathbb{R} \longrightarrow \mathbb{C} \\ t \longmapsto \mathbb{E}[e^{itX}] \end{array}$

Thm 51 : $\mathcal{L}(X) = \mathcal{L}(Y) \Leftrightarrow \varphi_X = \varphi_Y$

Exple 52 :
- $X \hookrightarrow \mathcal{N}(m,\sigma^2)$, $\varphi_X(t) = e^{itm - \sigma^2 t^2/2}$
- $X \hookrightarrow \mathcal{B}(n,p)$, $\varphi_X(t) = (1-p+pe^{it})^n$
- $X \hookrightarrow \mathcal{P}(\lambda)$, $\varphi_X(t) = \exp(\lambda(e^{it}-1))$
- $X \hookrightarrow \mathcal{E}(1)$, $\varphi_X(t) = \frac{1}{1-it}$

Thm 53 : (i) si $\mathbb{E}[|X|^n] < \infty$, alors φ est n-dérivable et $\varphi^{(k)}(t) = i^k \mathbb{E}[X^k e^{itX}]$ pour $k \leq n$.
- en particulier : $\varphi^{(k)}(0) = i^k \mathbb{E}[X^k]$

(ii) si n est pair et φ n-dérivable en 0, alors X admet un moment d'ordre r $\forall r \leq n$.

C-ex 54 : Soit X de loi $\sum_{k \in \mathbb{Z}} a_k \delta_k$, où $\begin{cases} a_k = a_{-k} = \frac{c}{k^2 \ln(k)}, & \forall k \geq 2 \\ a_1 = a_0 = a_{-1} = 0 \end{cases}$

$c = \left(\frac{1}{2} \sum_{k=2}^{+\infty} \frac{1}{k^2 \ln(k)}\right)^{-1}$ constante de normalisation

φ_X est $\mathcal{C}^1$ mais X n'admet pas d'espérance.

Prop 55 : Si X admet des moments de tout ordre et si $\varlimsup_n \frac{\|X\|_n}{n} < +\infty$, alors φ_X est DSE au voisinage de tout réel.

Prop 56 : Si φ_X est analytique, alors $\mathcal{L}(X)$ est caractérisée par $(\mathbb{E}[X^k])_{k \in \mathbb{N}}$

C-ex 57 : Soit $X \hookrightarrow \mathcal{N}(0;1)$, $f_{Z=e^X}(z) = \frac{e^{-(\ln z)^2/2}}{z\sqrt{2\pi}} \mathbb{1}_{\mathbb{R}_+^*}(z)$.

$\forall a \in [-1;1]$ $f_a(x) = f_Z(x)(1 + a\sin(2\pi \ln(x)))$

Z_a de densité f_a et Z ont même moments.

3) Transformée de Laplace

Déf 58 : La transformée de Laplace de X est $L_X(s) = \mathbb{E}[e^{sX}]$, pour s convenable, ie $e^{sX} \in L^1$.

App 59 : (Inégalité de Hoeffding)

Soit $(X_n)_{n \geq 1}$ suite de var centrées, indépendantes.

On suppose avoir $(c_n)_{n \geq 1} \in (\mathbb{R}_+^*)^{\mathbb{N}^*}$ telle que $|X_n| \leq c_n$ p.s. Alors

$\forall \varepsilon > 0, \forall n \in \mathbb{N}^*$ $\mathbb{P}(|S_n| \geq \varepsilon) \leq 2\exp\left(\frac{-\varepsilon^2}{2\sum_{k=1}^n c_k^2}\right)$

Coro 60 : Soit $\alpha > 0$. On suppose de plus avoir $\beta > 0$ tq $\forall n \in \mathbb{N}^*$, $\sum_{k=1}^n c_k^2 \leq n^{2\alpha - \beta}$. Alors $\frac{S_n}{n^\alpha} \xrightarrow{ps} 0$

DEV 2

Thm 61 : Si $L_X = L_Y$ dans un voisinage de 0, alors $\mathcal{L}(X) = \mathcal{L}(Y)$

Thm 62 : Soit X une va réelle telle que e^{tX} est intégrable pour tout t dans un intervalle ouvert contenant 0. Alors L_X est définie sur un voisinage de 0 et $L_X(t) = \sum_{n \in \mathbb{N}} \frac{t^n}{n!} \mathbb{E}(X^n)$.
- En particulier $L_X^{(n)}(0) = \mathbb{E}[X^n]$

4) Convergence

Thm 63 : $(X_i)_{i \geq 1}$ suite de var iid de même loi que X.
On note pour $n \geq 1$ $S_n = \sum_{i=1}^n X_i$

(LGN faible) si $\mathbb{E}[|X|] < \infty$, alors $\frac{S_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}[X]$

(LGN forte) $\mathbb{E}[|X|] < \infty \iff \frac{S_n}{n} \xrightarrow{ps} \mathbb{E}[X]$

App 64 : La moyenne empirique $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ est un estimateur fortement consistant de l'espérance.

Thm 65 : (TCL) si $\mathbb{E}[X^2] < \infty$ et $\mathrm{Var}(X) = \sigma^2$ alors $\frac{S_n - n\mathbb{E}[X]}{\sigma\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0;1)$

App 66 : (intervalle de confiance)

si $X_i \hookrightarrow b(p)$, alors pour tout $a < b$

$\lim_{n \to +\infty} \mathbb{P}\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) = \int_a^b \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt$

Références : Barbe-Ledoux, Probabilités

Ouvrard, probabilités 1 et 2,

Cottrell, exercices de probabilités

Zuily-Queffélec, Analyse pour l'agrégation

Cadre, statistique mathématique