260

Espérance, variance et moments d'une variable aléatoire

I Notion d'espérance et de variance

1) espérance et moments d'ordre p.

def : soit X une variable aléatoire (v.a) réelle définie sur $(\Omega, \mathcal{A}, \mathbb{P})$. Si X est intégrable, on appelle espérance ou espérance mathématique de X (sous la probabilité $\mathbb{P}$) le nombre réel

$$\mathbb{E}[X] = \int_\Omega X \, d\mathbb{P}_X$$

rem : on dit que X est centrée si X intégrable et $\mathbb{E}[X]=0$.

def : plus généralement, si $X \in L^p$, $p>0$, on définit le moment absolu d'ordre p de X par $\mathbb{E}[|X|^p] = \int_\Omega |X|^p \, d\mathbb{P}_X$. Si $p \in \mathbb{N}$, on peut aussi définir le moment d'ordre p par $\mathbb{E}[X^p] = \int_\Omega X^p \, d\mathbb{P}_X$

th : de transfert (ou transport) soit X un vecteur aléatoire à valeurs dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ et ϕ une fonction borélienne de $\mathbb{R}^d \to \mathbb{R}$. Si ϕ est à valeurs positives,

$$\mathbb{E}[\phi(X)] = \int_\Omega \phi \circ X(\omega) \, d\mathbb{P}(\omega) = \int_{\mathbb{R}^d} \phi(x) \, d\mathbb{P}_X(x)$$

Si ϕ est à valeurs quelconques mais L^1 on a aussi l'égalité.

prop : • X vecteur aléatoire à valeurs dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ $A \in \mathcal{B}(\mathbb{R}^d)$ alors $\mathbb{1}_A$ est mesurable et

$$\mathbb{E}[\mathbb{1}_A(X)] = \mathbb{P}(X \in A)$$

• Soit X un vecteur aléatoire à valeurs dans $\mathbb{R}^d$ admettant une densité f. Soit h une bijection sur $\mathbb{R}^d$, de classe C^1, de jacobien $J_h(x) \neq 0 \ \forall x$. Alors le vecteur $Y = h(X)$ a pour densité :

$$g(y) = |J_h(h^{-1}(y))|^{-1} f \circ h^{-1}(y)$$

rem : en pratique, le calcul de $\mathbb{E}[\phi(X)]$ ne nécessite pas le calcul de la loi de $\phi(X)$.

ex : • $X \sim \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1 \Rightarrow \mathbb{E}[X] = \frac{1}{2}$

• $X \sim P_n = \frac{1}{n} \sum_{1 \le i \le n} \delta_{x_i} \Rightarrow \mathbb{E}[X] = \frac{1}{n} \sum_{i=1}^n x_i$

• $X \sim \mathcal{B}(n,p) \Rightarrow \mathbb{E}[X] = pn$

• $X \sim \mathcal{P}(\lambda) \Rightarrow \mathbb{E}[X] = \lambda$

• $X \sim \exp(\lambda) \Rightarrow \mathbb{E}[X] = \frac{1}{\lambda}$.

2) variance et covariance

def : X v.a réelle de carré intégrable. On appelle variance de X notée $Var(X)$, la quantité :

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

$\sqrt{Var(X)}$ est appelé l'écart-type noté σ

rem : une v.a d'écart type 1 est dite réduite.

prop : $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

ex : • $Var X = 0 \Rightarrow X = C$ ps $C = \mathbb{E}[X]$

• $X \sim \mathcal{B}(n,p) \Rightarrow Var(X) = np(1-p)$

• $X \sim \mathcal{N}(0,1) \Rightarrow \mathbb{E}[X] = 0$, $Var(X) = 1$

• $\alpha \in \mathbb{R}$, $Var(X+a) = Var(X)$, $Var(\alpha X) = \alpha^2 Var(X)$

def : $(X_1, \dots, X_d) \in \mathbb{R}^d$ vecteur aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ on dit que X est de puissance p-ième intégrable $(p>0)$ si chacune de ses composantes l'est. Sa (matrice carrée de) covariance est :

$$Cov(X) = \left(\mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])] \right)_{1 \le i,j \le d}$$

rem : $Cov(X)$ est symétrique défini, positive.

3) propriétés fondamentales

prop : inégalité de Markov $X \in L^1$, $t > 0$
Alors $\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[|X|]}{t}$

cor : inégalité de Tchebytchev $X \in L^2$, $t > 0$
$\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq \frac{Var(X)}{t^2}$

th : • inégalité de Jensen ϕ convexe, $X, \phi \in L^1$
Alors $\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)]$
• inégalité de Hölder : $X \in L^p$, $Y \in L^q$ $\frac{1}{p} + \frac{1}{q} = 1$
$p, q \geq 1$ Alors $XY \in L^1$ et $\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{1/p} \mathbb{E}[|Y|^q]^{1/q}$
• $p \mapsto \mathbb{E}[|X|^p]^{1/p}$ est croissante
• $\|.\|_p = \mathbb{E}[|.|^p]^{1/p}$ norme sur $L^p(\Omega, \mathcal{A}, \mathbb{P})$ $p \geq 1$
• $\|X\|_\infty = \lim_{p \to \infty} \|X\|_p$ est une norme sur $L^\infty(\Omega, \mathcal{A}, \mathbb{P})$

II Variables aléatoires et fonction caractéristique

1) transformée de Laplace et de Fourier

def : X un vecteur aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}^d$. On appelle fonction caractéristique de X, ou transformée de Fourier, notée
$\varphi_X : t \in \mathbb{R}^d \mapsto \mathbb{E}[e^{i\langle t, X\rangle}]$

th : X, Y deux vecteurs aléatoires de loi $\mathbb{P}_X$ et $\mathbb{P}_Y$, $\varphi_X = \varphi_Y \Leftrightarrow \mathbb{P}_X = \mathbb{P}_Y$

th : formule d'inversion de Fourier : soit φ_X la fonction caractéristique d'un vecteur aléatoire X, $\varphi_X \in L^1(\mathbb{R}^d)$ Alors la loi de X admet une densité continue, bornée f_X sur $\mathbb{R}^d$ avec :

$$f_X(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\langle t, x\rangle} \varphi_X(t)\, dt$$

app : X va. réelle de fonction caractéristique φ_X et de loi $\mathbb{P}_X$
• $\mathbb{E}[|X|^n] < \infty$ Alors φ_X est n fois dérivable et $\varphi_X^{(k)}(t) = i^k \mathbb{E}[X^k e^{itX}]$ $k \leq n$
$\varphi_X^{(k)}(0) = i^k \mathbb{E}[X^k]$
• réciproquement si n pair et si φ est n-fois dérivable en 0, alors X admet tout moment d'ordre plus petit ou égal à n.

app : théorème des moments : X, Y va sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans un intervalle borné $[a,b]$
$\mathbb{E}[X^k] = \mathbb{E}[Y^k]$ $\forall k \in \mathbb{N} \Rightarrow X \sim Y$

def : X un vecteur aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}^d$, on appelle la transformée de Laplace $L_X(s) = \mathbb{E}[e^{\langle s, X\rangle}]$ définie pour les valeurs de s pour lesquelles $e^{\langle s, X\rangle}$ est intégrable.

prop : X va réelle, e^{tX} intégrable pour t dans un intervalle ouvert contenant 0. Alors la transformée de Laplace L_X est défini sur un ouvert contenant 0 (intervalle). De plus elle est analytique dans un voisinage de 0 et

$$L_X(t) = \sum_{n \in \mathbb{N}} \frac{t^n}{n!} \mathbb{E}[X^n] \quad \forall t \text{ dans ce voisinage}$$

app : étude de la loi de $\gamma(a, \lambda)$ [DeP I] [Cot]

2) le cas de l'indépendance

prop : $(X_i)_I$ famille de va réelles indépendante
$\Leftrightarrow$ $\forall J \subset I$ et toute famille de fonctions boréliennes ϕ_i, $i \in J$, $\phi_i(X_i)$ intégrable, $\mathbb{E}[\prod_{i \in J} \phi_i(X_i)] = \prod_{i \in J} \mathbb{E}[\phi_i(X_i)]$

cor : $(X_i)_{1\le i\le n}$ va réelles est indépendante $\Leftrightarrow \forall (t_i)_{1\le i\le n} \in \mathbb{R}^n \quad \varphi_{(X_1,\dots,X_n)}(t_1,\dots,t_n) = \varphi_{X_1}(t_1)\dots\varphi_{X_n}(t_n)$

def : $X, Y \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ v.a réelles sont non corrélées si $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

rem : $X, Y \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ et indépendantes $\Rightarrow X, Y$ non corrélées.

prop : identité de Bienaymé : $(X_i)_{1\le i\le n}$ deux à deux non corrélées, alors $Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i)$

III Applications à la convergence

1) convergence en loi

def, th : X, X_n des va réelles définies sur $(\Omega, \mathcal{A}, \mathbb{P})$. $X_n \xrightarrow{loi} X$ si l'une des conditions équivalentes suivantes est vérifiée :

- $\lim_{n\to\infty} \mathbb{E}[\phi(X_n)] = \mathbb{E}[\phi(X)] \ \forall \phi : \mathbb{R}\to\mathbb{R}$ continue bornée
- $\lim_{n\to\infty} \varphi_{X_n}(t) = \varphi_X(t) \quad \forall t\in\mathbb{R}$

app : théorème central limite : $S_n = \sum_{i=1}^n X_i$ IID

- $\mathbb{E}[X^2] < \infty \Rightarrow \dfrac{S_n - n\mathbb{E}[X]}{\sqrt{n}} \xrightarrow{loi} \mathcal{N}(0, Var(X))$
- $\dfrac{S_n}{n}$ converge en loi alors $\mathbb{E}[X] = 0$ et $\mathbb{E}[X^2] < \infty$ et la loi limite est normale centrée de variance $Var(X)$.

ex : • $X_i \sim \mathcal{B}(1,p)$ $S_n = X_1 + \dots + X_n$

$$\lim \mathbb{P}\left(a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b\right) = \int_a^b \frac{e^{-t^2/2}}{\sqrt{2\pi}}\,dt$$

2) convergence dans L^p

def : soient $(X_n)_{n\in\mathbb{N}}, X$ des va réelles dans $L^p(\Omega, \mathcal{A}, \mathbb{P})$ $0 < p < \infty$. On dit que $X_n \xrightarrow{L^p} X$ dans L^p si $\lim_{n\to\infty} \|X_n - X\|_p = 0$ ou de façon équivalente $\mathbb{E}[|X_n - X|^p] \xrightarrow[n\to\infty]{} 0$

th : de Bernstein $f : [0,1] \to \mathbb{C}$ continue [ZQ]

$\omega(h) = \sup\{|f(u) - f(v)| ; |u - v| \le h\}$. Soit $n \ge 1$

$B_n(f, x) = B_n(x) = \sum C_n^k x^k (1-x)^{n-k} f(\frac{k}{n})$ [DVP 2]

Alors :
- $B_n \xrightarrow{\|\cdot\|_\infty} f$ sur $[0,1]$
- $\|f - B_n\|_\infty \le C\omega(\frac{1}{\sqrt{n}})$ C constante
- l'estimation est optimale : $\exists f$ lipschitzienne pour laquelle $\|f - B_n\|_\infty \ge \frac{\delta}{\sqrt{n}}$ où δ constante

3) convergence en proba

def : X_n, X $n\in\mathbb{N}$ des va réelles sur $(\Omega, \mathcal{A}, \mathbb{P})$ on dit que $X_n \xrightarrow{P} X$ (en probabilité) si $\forall \varepsilon > 0 \lim_{n\to\infty} \mathbb{P}(|X_n - X| \ge \varepsilon) = 0$

ex : X_i $i \ge 1$ va. réelles non corrélées, $\mathbb{E}[X_i] = 0$ et $Var(X_i) = \sigma^2$ $\forall i \ge 1$ Alors $\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{P} 0$

prop : convergence $L^p \Rightarrow$ convergence $P \Rightarrow$ convergence en loi

Lesigne, Pile ou Face
Jeulin.