Vecteurs aléatoires : à retenir

On se donne une probabilité P sur un ensemble Ω (muni d'évènements).

Définition-Proposition

On considère une v. a. $X = (X_1, ..., X_n)$ à valeurs dans $\mathbb{R}^n$.

(a) La *loi de* X (« loi jointe de $X_1, ..., X_n$ ») est l'unique probabilité P_X sur $\mathbb{R}^n$ telle que :

$$P_X\big([a_1, b_1] \times ... \times [a_n, b_n]\big) = P\big(a_1 \le X_1 \le b_1, ..., a_n \le X_n \le b_n\big)$$

pour $-\infty \le a_1 \le b_1 \le +\infty$, ..., $-\infty \le a_n \le b_n \le +\infty$.

(b) On dit que $\overbrace{X\ \textit{est discrète}}^{\text{ou « \textit{la loi de X est discrète} »}}$ s'il existe $D \subseteq \mathbb{R}^n$ finie ou dénombrable telle que $P_X(D) = 1$.

Pour toute application $(x_1, ..., x_n) \mapsto p_{x_1,...,x_n}$ de $\mathbb{R}^n$ dans $\mathbb{R}^+$ nulle hors d'une partie finie ou dénombrable, la v. a. X est discrète avec $P(X_1 = x_1, ..., X_n = x_n) = p_{x_1,...,x_n}$ pour $(x_1, ..., x_n) \in \mathbb{R}^n$ si et seulement si : $\mathrm{E}\big(\underbrace{h(X)}_{\text{discrète}}\big) = \sum\limits_{\substack{(x_1,...,x_n)\in\mathbb{R}^n \\ \text{tel que } p_{x_1,...,x_n} \neq 0}} h(x_1, ..., x_n)\, p_{x_1,...,x_n}$ $\underbrace{\text{pour toute } h\colon D \to \mathbb{R} \text{ bornée}}_{\text{ou pour toute } h\colon D \to \mathbb{R} \text{ positive}}$.

Une application $(x_1, ..., x_n) \in \mathbb{R}^n \mapsto p_{x_1,...,x_n} \in \mathbb{R}^+$ nulle hors d'une partie finie ou dénombrable est associée à une telle X pour une probabilité P si et seulement si $\sum\limits_{(x_1,...,x_n)\in\mathbb{R}^n} p_{x_1,...,x_n} = 1$.

On appellera « fonction de masse sur $\mathbb{R}^n$ » une telle fonction p.

(c) On dit que $\overbrace{X\ \textit{admet pour densité}}^{\text{ou « \textit{la loi de X admet pour densité} »}}$ une application $p\colon \mathbb{R}^n \to \mathbb{R}^+$ mesurable si :

$$P\big(a_1 \le X_1 \le b_1, ..., a_n \le X_n \le b_n\big) = \int_{a_1}^{b_1} ... \int_{a_n}^{b_n} p(x_1, ..., x_n)\, \mathrm{d}x_1...\mathrm{d}x_n$$

pour $-\infty \le a_1 \le b_1 \le +\infty$, ..., $-\infty \le a_n \le b_n \le +\infty$.

Pour toute $p\colon \mathbb{R}^n \to \mathbb{R}^+$ mesurable, la v. a. X admet pour densité p si et seulement si :

$$\mathrm{E}(h(X)) = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} h(x_1, ..., x_n)\, p(x_1, ..., x_n)\, \mathrm{d}x_1...\mathrm{d}x_n \quad \underbrace{\text{pour toute } h \in \mathscr{C}_b(\mathbb{R}^n, \mathbb{R})}_{\text{ou pour toute } h\colon \mathbb{R}^n \to \mathbb{R} \text{ mesurable positive}}.$$

Dans ce cas, on a : $P(X \in B) = \int ... \int_B p(x_1, ..., x_n)\, \mathrm{d}x_1...\mathrm{d}x_n$ pour tout $\underset{\text{évènement}}{B} \subseteq \mathbb{R}^n$ (choisir $h = \mathbb{1}_B$).
Donc deux telles applications mesurables p_0 et p_1 continues en $a \in \mathbb{R}^n$ vérifient $p_0(a) = p_1(a)$.

Une application $p\colon \mathbb{R}^n \to \mathbb{R}^+$ mesurable est associée à une telle X pour une certaine probabilité P si et seulement si $\int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} p(x_1, ..., x_n)\, \mathrm{d}x_1...\mathrm{d}x_n = 1$.

On appellera « densité de probabilité sur $\mathbb{R}^n$ » une telle fonction p.

Définition-Proposition

On considère une v. a. $X = (X_1, ..., X_n)$ à valeurs dans $\mathbb{R}^n$.

(a) On suppose que X est discrète avec une fonction de masse $(x_1, ..., x_n) \mapsto p_{x_1,...,x_n}$.
Les lois de $X_1, ..., X_n$, appelées « lois marginales de X », sont discrètes avec :

$$P(X_i = x) = \sum_{\substack{(x_1,...,x_{i-1},x_{i+1}x_n)\in\mathbb{R}^{n-1} \\ \text{tel que } p_{x_1,...,x_{i-1},x,x_{i+1},...,x_n} \neq 0}} p_{x_1,...,x_{i-1},x,x_{i+1},...,x_n} \quad \text{pour } x \in \mathbb{R}\ (1 \le i \le n).$$

(b) On suppose que X a une densité $p\colon \mathbb{R}^n \to \mathbb{R}^+$.
Les lois de $X_1, ..., X_n$, appelées « lois marginales de X », ont des densités $p_1, ..., p_n$ définies par :

$$p_i(x) = \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} p(x_1, ..., x_{i-1}, x, x_{i+1}..., x_n)\, \mathrm{d}x_1...\mathrm{d}x_{i-1}\mathrm{d}x_{i+1}...\mathrm{d}x_n \quad \text{pour } x \in \mathbb{R}\ (1 \le i \le n).$$

Définition

On considère deux v. a. r. $X, Y\colon \Omega \to \mathbb{R}$ telles que : $\mathrm{E}(X^2) < +\infty$ et $\mathrm{E}(Y^2) < +\infty$.

(a) On a : $|\mathrm{E}(|XY|)| \le \sqrt{\mathrm{E}(X^2)}\sqrt{\mathrm{E}(Y^2)} < +\infty$ puis $\boxed{|\mathrm{E}(\underbrace{XY}_{\text{a un moment d'ordre 1}})| \le \sqrt{\mathrm{E}(X^2)}\sqrt{\mathrm{E}(Y^2)}}$.
On définit : $\mathrm{Cov}(X, Y) = \underbrace{\mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)}_{\mathrm{E}\left((X - \mathrm{E}(X)(Y - \mathrm{E}(Y))\right)} \in \mathbb{R}$.

On a donc la formule suivante : $\mathrm{Var}(X + Y) = \mathrm{Var}\, X + \mathrm{Var}\, Y + 2\,\mathrm{Cov}(X, Y)$.

(b) On dit que X et Y sont non-corrélées si $\mathrm{Cov}(X, Y) = 0$ (par exemple X et Y indépendantes).
Quand $\mathrm{Var}(X) \neq 0$ et $\mathrm{Var}(Y) \neq 0$, le coefficient de corrélation linéaire de X et Y est défini par :

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\,\sigma(Y)} \in [-1, 1].$$

Théorème (« théorème de Fubini »)

(a) On considère une application $f : \mathbb{R}^2 \to \mathbb{C}$ mesurable.

On a $\int_{-\infty}^{+\infty}\big(\int_{-\infty}^{+\infty} |f(x,y)|\,\mathrm{d}y\big)\mathrm{d}x = \int_{-\infty}^{+\infty}\big(\int_{-\infty}^{+\infty} |f(x,y)|\,\mathrm{d}x\big)\mathrm{d}y \ \leq\ +\infty$

et, quand le réel précédent est fini : $\int_{-\infty}^{+\infty}\big(\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y\big)\mathrm{d}x = \int_{-\infty}^{+\infty}\big(\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x\big)\mathrm{d}y$.

(b) On considère une suite d'applications $u_n : \mathbb{R} \to \mathbb{C}$ mesurables avec $n \geq 0$.

On a $\int_{-\infty}^{+\infty}\big(\sum\limits_{n=0}^{+\infty} |u_n(x)|\big)\mathrm{d}x = \sum\limits_{n=0}^{+\infty}\big(\int_{-\infty}^{+\infty} |u_n(x)|\,\mathrm{d}x\big) \ \leq\ +\infty$

et, quand le réel précédent est fini : $\int_{-\infty}^{+\infty}\big(\sum\limits_{n=0}^{+\infty} u_n(x)\big)\mathrm{d}x = \sum\limits_{n=0}^{+\infty}\big(\int_{-\infty}^{+\infty} u_n(x)\,\mathrm{d}x\big)$.

(c) On considère une suite double $(u_{p,q})_{p,q\geq 0}$ de nombres complexes.

On a $\sum\limits_{p=0}^{+\infty}\big(\sum\limits_{q=0}^{+\infty} |u_{p,q}|\big) = \sum\limits_{q=0}^{+\infty}\big(\sum\limits_{p=0}^{+\infty} |u_{p,q}|\big) \ \leq\ +\infty$

et, quand le réel précédent est fini : $\sum\limits_{p=0}^{+\infty}\big(\sum\limits_{q=0}^{+\infty} u_{p,q}\big) = \sum\limits_{q=0}^{+\infty}\big(\sum\limits_{p=0}^{+\infty} u_{p,q}\big)$.

Théorème (« théorème de changement de variable »)

On considère une partie U de $\mathbb{R}^n$ qui est ouverte et une application $\varphi = (\varphi_1, ..., \varphi_n)$ de U dans $\mathbb{R}^n$ qui a des dérivées partielles, toutes continues.

(a) Soit $x \in U$. On appelle *matrice jacobienne de φ en x* la matrice $J_\varphi(x) := \left(\dfrac{\partial \varphi_i}{\partial x_j}(x)\right)_{\substack{1\leq i\leq n\\ 1\leq j\leq n}}$.

(b) On suppose que φ est injective et $\det J_\varphi(x) \neq 0$ pour tout $x \in U$.

Pour toute fonction continue bornée f de $\varphi(U)$ dans $\mathbb{R}$, on a :

$$\int\!\cdots\!\int_{\varphi(U)} f(y_1, ..., y_n)\, \underbrace{\mathrm{d}y_1...\mathrm{d}y_n}_{\text{penser à } |\mathrm{d}y_1| \text{ quand } n=1} = \int\!\cdots\!\int_U f(\varphi(x_1, ..., x_n))\, \underbrace{|\det J_\varphi(x_1, ..., x_n)|}_{\text{« jacobien de } \varphi \text{ en } (x_1,...,x_n) \text{ », noté parfois } \frac{D(y_1,...,y_n)}{D(x_1,...,x_n)}}\, \mathrm{d}x_1...\mathrm{d}x_n$$

Proposition

On considère une v. a. $X = (X_1, ..., X_n)$ à valeurs dans $\mathbb{R}^n$.

(a) On considère des fonctions de masse $p_1, ..., p_n$ sur $\mathbb{R}$.

Les v. a. r. $X_1, ..., X_n$ sont discrètes de fonctions de masse p_1, ..., p_n et sont indépendantes si et seulement si X est discrète de fonction de masse $p\colon (x_1, ..., x_n) \mapsto p_1(x_1)...p_n(x_n)$.

(b) On considère des densités de probabilité $p_1, ..., p_n$ sur $\mathbb{R}$.

Les v. a. r. $X_1, ..., X_n$ admettent respectivement les densités $p_1, ..., p_n$ et sont indépendantes si et seulement si X a pour densité $p\colon (x_1, ..., x_n) \mapsto p_1(x_1)...p_n(x_n)$.

Proposition

On considère des v. a. r. X et Y.

(a) On suppose que X et Y sont discrètes, et indépendantes.

La v. a. r. $X + Y$ est discrète et : $P(X + Y = u) \ =\ \sum\limits_{x\in\mathbb{R}} \underbrace{P(X = x)\, P(Y = u - x)}_{\text{lire 0 lorsque } P(X=x)=0}$ pour $u \in \mathbb{R}$.

(b) On suppose que X et Y ont des densités respectives p et q, et sont indépendantes.

La v. a. r. $X + Y$ a pour densité r avec : $r(u) \ =\ \int_{-\infty}^{+\infty} p(x)\, q(u - x)\, \mathrm{d}x \ \leq +\infty$ pour tout $u \in \mathbb{R}$.

Définition-Proposition

(a) Soient X et Y deux v. a. r. discrètes, avec $Y \geq 0$ ou $\mathrm{E}(|Y|) < +\infty$.

Lorsque $x \in \mathbb{R}$ et $P(X = x) \neq 0$, on introduit la « probabilité de $\{Y = y\}$ sachant $\{X = x\}$ » :

$$p_{Y|X=x}(y) = \tfrac{P(\{Y=y\}\cap\{X=x\})}{P(X=x)} \quad \text{pour } y \in \mathbb{R}.$$

Donc : $p_{Y|X=x}$ est une fonction de masse.

On a : $\mathrm{E}(Y) = \sum\limits_{x\in\mathbb{R}} \underbrace{\mathrm{E}(Y|X = x)\, P(X = x)}_{\text{lire 0 lorsque } P(X=x)=0}$ où $\mathrm{E}(Y|X{=}x) := \sum\limits_{y\in\mathbb{R}} y\, p_{Y|X=x}(y)$.

(b) Soit (X, Y) une v. a. à valeurs dans $\mathbb{R}^2$ qui a une densité $p_{X,Y}$, avec $Y \geq 0$ ou $\mathrm{E}(|Y|) < +\infty$.

Lorsque $x \in \mathbb{R}$ et $p_X(x) \neq 0$, où $p_X(x) := \int_{-\infty}^{+\infty} p_{X,Y}(x, y)\, \mathrm{d}y$, on pose :

$$p_{Y|X=x}(y) = \tfrac{p_{X,Y}(x,y)}{p_X(x)} \quad \text{pour } y \in \mathbb{R}$$

Donc : $p_{Y|X=x}$ est une $\underbrace{\text{densité de probabilité}}_{\text{de la loi de probabilité sur } \mathbb{R} \text{ qui envoie un évènement } A \text{ sur } \lim\limits_{h\to 0^+} P(Y\in A\,|\,x\leq X\leq x+h)}$, appelée « densité conditionnelle de Y sachant $X = x$ ».

On a : $\mathrm{E}(Y) = \int_{-\infty}^{+\infty} \underbrace{\mathrm{E}(Y|X = x)\, p_X(x)}_{\text{lire 0 lorsque } p_X(x)=0}\, \mathrm{d}x$ où $\mathrm{E}(Y|X = x) := \int_{-\infty}^{+\infty} y\, p_{Y|X=x}(y)\, \mathrm{d}y$.