Vecteurs aléatoires

BCPST 1, 26/09/2018

Notations du chapitre — *Dans tout ce chapitre,* $(\Omega, \mathscr{T}, P)$ *est un espace probabilisé.*

I — Couple de variables aléatoires, lois

Définition 1.1 — Couple de variables aléatoires
On appelle couple de variables aléatoires réelles la donnée de deux variables aléatoires X et Y sur le même espace probabilisé $(\Omega, \mathscr{T}, P)$.
On parle aussi de vecteur aléatoire réel, noté (X, Y).

Remarque I.1

- Si (X, Y) est un vecteur aléatoire et si $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$, alors $g(X, Y)$ est une variable aléatoire ;
- comme d'habitude, nous nous limiterons aux espaces probabilisés finis.

Exemple Considérons l'expérience « on lance deux dés à 6 faces honnêtes, de couleurs différentes ».
On peut définir plusieurs couples de variables aléatoires, comme X « la somme des résultats obtenus » et Y, « le plus grand des résultats obtenus ».
On peut définir d'autres couples de variables aléatoires, comme par exemple X_1 le résultat du lancer du premier dé, X_2 le résultat de l'autre dé.
On peut aussi étudier des vecteurs aléatoires tels que (X_1, X_2, Y) (à plus de 2 composantes, donc).

Définition 1.3 — Loi conjointe, lois marginales
Soit (X, Y) un vecteur aléatoire. On note

$$X(\Omega) = \{x_1, x_2, \dots, x_n\} \qquad \textit{et} \qquad Y(\Omega) = \{y_1, y_2, \dots, y_m\}.$$

- *On appelle loi conjointe de X et de Y la famille de réels*

$$\forall i \in [\![1 ; n]\!], \ \forall j \in [\![1 ; m]\!], \qquad P\big((X = x_i) \cap (Y = y_j)\big)$$

- *On appelle loi marginale de X la loi de la variable aléatoire (et de même la loi marginale de Y).*

Propriété 1.4 — Lien entre loi conjointe et loi marginale

D'après la formule des probabilités totales sur le système complet d'évènements $(Y = y_j)_{1\leqslant j\leqslant m}$

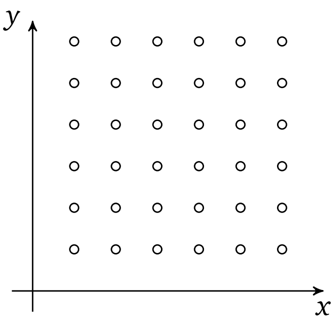
$$\forall i \in [\![1\,;n]\!], \qquad P(X = x_i) = \sum_{j=1}^{m} P(X = x_i \cap Y = y_j)$$

De même pour la loi marginale suivant Y

$$\forall j \in [\![1\,;m]\!], \qquad P(Y = y_j) = \sum_{i=1}^{n} P(X = x_i \cap Y = y_j)$$

Concrètement, on peut représenter la loi conjointe de (X,Y) dans un tableau à deux entrées. L'ensemble des valeurs considérées est ici $X(\Omega)\times Y(\Omega)$. Comme cela arrive souvent avec des fonctions, cet ensemble est un ensemble d'arrivée possible. Mais ce n'est pas forcément l'ensemble des valeurs prises par le couple.

Exemple A Lancer de deux dés



$y \backslash x$	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

FIGURE I.1 — *Deux représentations de* $\Omega_{X,Y}$.

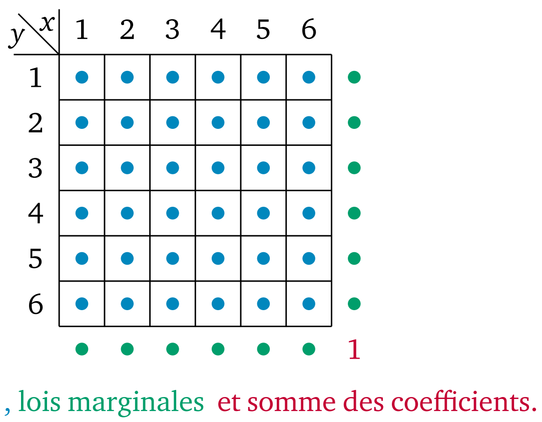


FIGURE I.2 — *Lois*

Propriété 1.5 — Somme des coefficients
Avec les notations précédentes

$$\sum_{i=1}^{n}\sum_{j=1}^{m} P\big((X=x_i)\cap(Y=y_j)\big) = \sum_{i=1}^{n} P(X=x_i) = \sum_{j=1}^{m} P(Y=y_j) = 1$$

II — Loi conditionnelle, indépendance de deux variables aléatoires

Définition 2.1 — Lois conditionnées
Soit (X,Y) un vecteur aléatoire et $x \in X(\Omega)$ tel que $P(X=x) \neq 0$. On appelle loi de Y conditionnée par l'évènement $(X=x)$s la famille de réels

$$\forall j \in [\![1 \,;\, m]\!], \qquad P_{(X=x)}(Y=y_j) = \frac{P\big((X=x)\cap(Y=y_j)\big)}{P(X=x)}$$

De même, on définit la loi de X conditionnée par l'évènement $(Y=y)$ tel que $P(Y=y)\neq 0$.

Exemple B Reprendre l'exemple des dés.

Définition 2.2 — Couple de variables aléatoires indépendantes
On dit que deux variables aléatoires X et Y sont indépendantes si et seulement si

$$\forall x \in X(\Omega),\ \ \forall y \in Y(\Omega), \qquad P(X=x\cap Y=y) = P(X=x)\times P(Y=y)$$

Comme toujours avec l'indépendance, c'est une propriété de calcul, ou un hypothèse du modèle.

Propriété 2.3 — *Si deux variables aléatoires X et Y sont indépendantes alors les lois de X conditionnées aux évènements $(Y=y)$ sont identiques à la loi de X (et de même pour Y).*

Propriété 2.4 — Fonctions de deux variables aléatoires indépendantes
Soit X et Y sont deux variables aléatoires indépendantes et f et g deux fonctions de $\mathbb{R}$ dans $\mathbb{R}$.
Alors les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

III — Fonction de deux variables aléatoires

III.1 — Loi

Propriété 3.1 — *Soit (X,Y) un vecteur aléatoire, g une fonction de $\mathbb{R}$ dans $\mathbb{R}$ et $Z = g(X,Y)$. On note*

$$X(\Omega) = \{x_1, x_2, \dots, x_n\} \qquad \text{et} \qquad Y(\Omega) = \{y_1, y_2, \dots, y_m\}$$

Le support de Z est alors

$$Z(\Omega) = \left\{g(x_i, y_j) \ ; \ i \in [\![1 \, ; n]\!], j \in [\![1 \, ; m]\!]\right\}.$$

et la loi de Z est donnée par la formule

$$\forall z \in Z(\Omega), \qquad P(Z = z) = \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega) \\ g(x,y) = z}} P(X = x \cap Y = y)$$

Si X et Y sont indépendantes, on a plus simplement

$$\forall z \in Z(\Omega), \quad P(Z = z) = \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega) \\ g(x,y) = z}} P(X = x)\, P(Y = y)$$

Application — Sommes de deux variables aléatoires
Dans un cas particulier utile, la formule donnant la loi de Z se simplifie : supposons que

$$X(\Omega) = [\![0 \, ; n]\!] \qquad \text{et} \qquad Y(\Omega) = [\![0 \, ; m]\!]$$

alors la loi de $Z = X + Y$ est

$$Z(\Omega) = [\![0 \, ; n + m]\!], \qquad \forall k \in [\![0 \, ; n + m]\!],$$

$$\boxed{P(Z = k) = \sum_{k=0}^{n+m} P(X = i \cap Y = k - i)}$$

Exemple

- Somme de deux variables indépendantes suivant une loi uniforme. (en exercice)
- Somme de deux variables indépendantes suivant une loi binomiale de paramètres n et p et m et p.

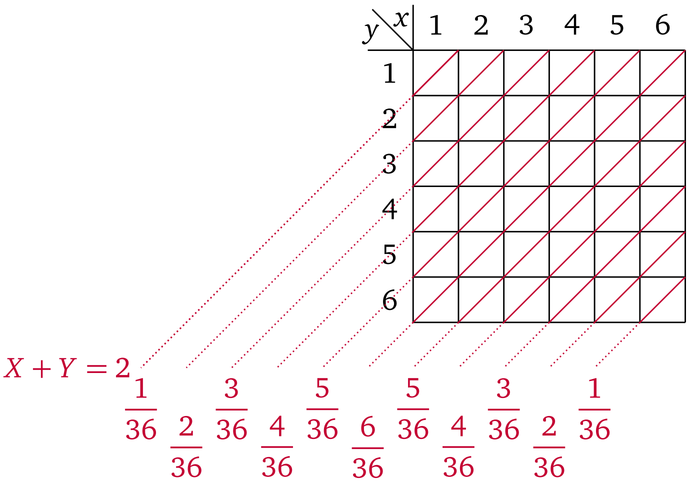


FIGURE I.3 — *Loi de la somme de deux lois uniformes*

III.2 — Espérance – Formule de transfert

Il est plus simple d'obtenir directement l'espérance de Z grâce à la

Théorème 3.4 — Formule de transfert
Soit (X, Y) un vecteur aléatoire, g une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ et $Z = g(X, Y)$.

$$E(Z) = \sum_{i=1}^{n} \sum_{j=1}^{m} g(x_i, y_j)\, P(X = x_i \cap Y = y_j)$$

Exemple Calculer directement l'espérance de $\sup(X_1, X_2)$.

Théorème 3.6 — Linéarité de l'espérance
Soit (X, Y) un vecteur aléatoire et $(a, b) \in \mathbb{R}^2$

$$E(aX + bY) = aE(X) + bE(Y) \tag{I.1}$$

DÉM. La première par la formule de transfert. La seconde est une conséquence de la première. □

Proposition 3.7 — *Soit (X, Y) un vecteur aléatoire.*
Si X et Y sont indépendantes, alors $E(XY) = E(X) \times E(Y)$.

DÉM. Par la formule de transfert. □

IV — Covariance, coefficient de corrélation linéaire

IV.1 — Covariance

Définition 4.1 — Covariance de deux variables aléatoires

Soit X et Y deux variables aléatoires.

On appelle covariance de X et Y le réel

$$\begin{aligned}\operatorname{cov}(X,Y) &= E\Big[\big(X - E(X)\big) \times \big(Y - E(Y)\big)\Big] \\ &= E(XY) - E(X)\,E(Y)\end{aligned}$$

Théorème 4.2 — Covariance de deux variables aléatoires indépendantes

Si X et Y sont deux variables indépendantes, alors leur covariance est nulle.

DÉM. Découle immédiatement de la proposition 3.7. □

Attention ! La réciproque est fausse !

Par exemple, soit X de loi ci-dessous et $Y = X^2$

X	-1	0	1
$P(X=x)$	$1/4$	$1/2$	$1/4$

On trouve $E(X) = 0$ et $E(XY) = E(X^3) = 0$ donc $\operatorname{cov}(X,Y) = 0$.

Mais $P(X = 0 \cap Y = 1) = 0 \neq P(X = 0)\ P(Y = 1)$, donc ces variables ne sont pas indépendantes.

Propriété 4.3 — Propriétés de la covariance

Soit X, Y, X_1 et X_2 des variables aléatoires et $(a,b) \in \mathbb{R}^2$

- $\operatorname{cov}(X,X) = V(X)$
- $\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)$
- $\operatorname{cov}(aX,bY) = ab\operatorname{cov}(Y,X)$
- $\operatorname{cov}(X_1 + X_2, Y) = \operatorname{cov}(X_1,Y) + \operatorname{cov}(X_2,Y)$

Propriété 4.4 — Propriétés de la Variance
Soit X et Y deux variables aléatoires et $(a,b)\in\mathbb{R}^2$

- $V(aX+bY)=a^2V(X)+b^2V(Y)+2ab\operatorname{cov}(X,Y)$
- $V(X+Y)=V(X)+V(Y)+2\operatorname{cov}(X,Y)$
- $V(X-Y)=V(X)+V(Y)-2\operatorname{cov}(X,Y)$
- $\operatorname{cov}(X,Y)=\dfrac{1}{4}(V(X+Y)-V(X-Y))$

DÉM. Par des calculs à partir des définitions. □

V — Généralisation au cas de *n* variables aléatoires

Définition 5.1 — Vecteur aléatoire — *n* variables
Soit $n\in\mathbb{N}^*$ et $X_1,X_2,\dots,X_n$ n variables aléatoires.
On appelle vecteur aléatoire le n–uplet de variables aléatoires $(X_1,X_2,\dots,X_n)$.

V.1 — Espérance, variance, covariance

Par généralisation des résultats précédents, nous pouvons en déduire :

Théorème 5.2 — Linéarité de l'espérance — *n* variables
Soit $(X_1,X_2,\dots,X_n)$ un vecteur aléatoire, et $a_1,a_2,\dots,a_n\in\mathbb{R}^n$:

$$E(a_1X_1+a_2X_2+\cdots+a_nX_n)=a_1E(X_1)+a_2E(X_2)+\cdots+a_nE(X_n)$$

V.2 — Indépendance

Définition 5.3 — Indépendance mutuelle de plusieurs variables aléatoires
On dit que les n variables aléatoires $X_1,X_2,\dots,X_n$ sont mutuellement indépendantes si et seulement si

$$\begin{aligned}\forall(x_1,x_2,\dots,x_n)\in X_1(\Omega)\times X_2(\Omega)\times\cdots\times X_n(\Omega),&\\ P(X_1=x_1\cap X_2=x_2\cap\cdots\cap X_n=x_n)&\\ =P(X_1=x_1)\times P(X_2=x_2)\times\cdots\times P(X_n=x_n)&\end{aligned}$$

Propriété 5.4 — *Si $X_1,X_2,\dots,X_n$ sont mutuellement indépendantes, alors toute sous-famille de $X_1,X_2,\dots,X_n$ l'est aussi.*

L'indépendance mutuelle entraîne donc l'indépendance deux à deux, mais la réciproque est fausse.

Propriété 5.5 —
Soit $X_1, X_2, \ldots, X_n$ *n variables aléatoires mutuellement indépendantes, soit* $p in [\![1\,;\,n]\!]$ *et* $f : \mathbb{R}^p \longrightarrow \mathbb{R}$, $g : \mathbb{R}^{n-p} \longrightarrow \mathbb{R}$.
Les variables aléatoires $f(X_1, X_2, \cdots, X_p)$ *et* $g(X_{p+1}, \cdots X_n)$ *sont indépendantes.*

Propriété 5.6 —
Soit $X_1, X_2, \ldots, X_n$ *n variables aléatoires mutuellement indépendantes, soit* u_1, u_2, ..., u_n *des fonctions de* $\mathbb{R}$ *dans* $\mathbb{R}$.
Les variables aléatoires $u_1(X_1)$, $u_2(X_2)$, ... $u_n(X_n)$ *sont mutuellement indépendantes.*

Théorème 5.7 — Variance de la somme de variables aléatoires mutuellement indépendantes
Si les n variables $X_1, X_2, \ldots, X_n$ *sont mutuellement indépendantes, alors*

$$V(X_1 + X_2 + \cdots + X_n) = V(X_1) + V(X_2) + \cdots + V(X_n)$$

Application 1 Loi de la somme de n variables de Bernoulli indépendante et de même loi $\mathscr{B}(1, p)$.