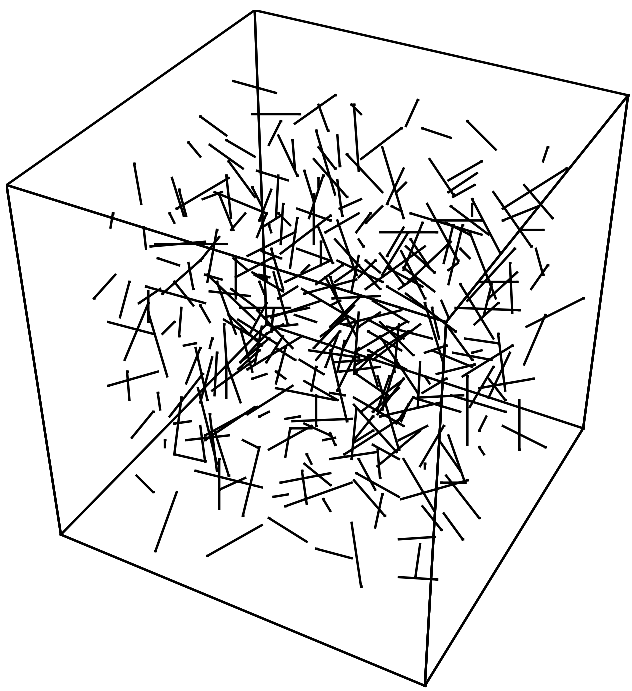
Chapitre 5
Vecteurs aléatoires

5.1.Introduction. En pratique, on est souvent confronté à des mesures simultanées caractérisant les propriétés physiques du phénomène aléatoire étudié . Par exemple,

- On peut caractériser un individu d'une population par son poids, sa taille, son âge,…

- Si on s'intéresse aux ménages ce peut être le revenu, la consommation,…

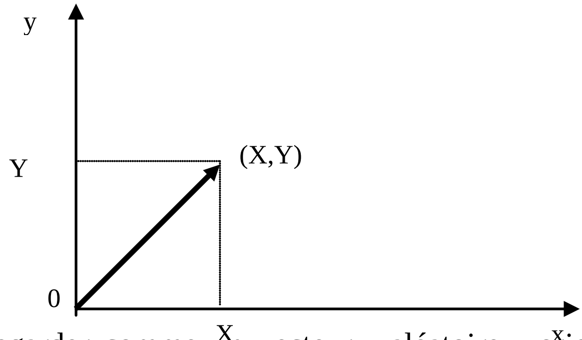
- en vision ou traitement d'images, l'image est caractérisée par le pixel (un point dans l'espace 3D avec la position , la luminosité,….



Nous ne considérerons que des vecteurs aléatoires dans 2D. La fonction de répartition

$$F_X(x) = P(X_1 < x_1, X_2 < x_2, ..., X_n < x_n) \quad \text{où} \quad x = (x_1, x_2, ..., x_n)$$

Et dans ce cas, le couple de variables aléatoires (X, Y) peut être interprété comme un point aléatoire de coordonnées (X, Y) dans le plan cartésien x0y.



On peut également le regarder comme un vecteur « aléatoire » orienté vers le point (X, Y).

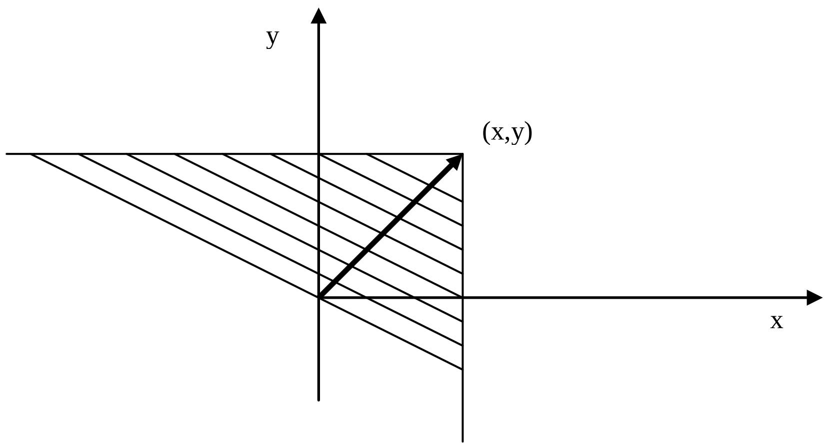


Figure 1. Représentation d'un vecteur aléatoire dans le plan 2D.

Ici, aléatoire signifie que pour « mesurer » les évènements liés à ce vecteur on utilisera la probabilité. La fonction de répartition du vecteur (X, Y) $F_{X,Y}(x, y) = P(X < x, Y < y)$ représente la probabilité pour que le point aléatoire tombe dans le carré de sommet (x,y), hachuré dans le plan

5.2.Propriétés. Les vecteurs aléatoires réels possèdent des propriétés très similaires à celles des variables aléatoires. Les propriétés spécifiques concernent surtout les relations d'indépendance, covariance (corrélation) et influence mutuelle.

(i) $\lim_{x \to -\infty} F_{(X,Y)}(x, y) = F_{(X,Y)}(-\infty, y) = 0$

$$\lim_{y \to -\infty} F_{(X,Y)}(x,y) = F_{(X,Y)}(x,-\infty) = 0$$

$$\lim_{x \to -\infty, y \to -\infty} F_{(X,Y)}(x,y) = F_{(X,Y)}(-\infty,-\infty) = 0$$

(ii) $\lim_{x \to +\infty} F_{(X,Y)}(x,y) = F_{(X,Y)}(+\infty, y) = F_Y(y)$: c'est la *fonction de répartition marginale* de la seule variable Y .

(iii) $\lim_{y \to +\infty} F_{(X,Y)}(x,y) = F_{(X,Y)}(x,+\infty) = F_X(x)$: f.r. de la v.a. X.

(iv)

$$\lim_{x \to +\infty, y \to +\infty} F_{(X,Y)}(x,y) = F_{(X,Y)}(+\infty,+\infty) = 1$$

(v) la fonction $F_{(X,Y)}(x,y)$ est croissante au sens large par rapport à chacun de ses arguments.

(vi) La fonction $F_{(X,Y)}(x,y)$ est continue à gauche par rapport à chacun de ses arguments. Notons que si on utilise la seconde définition de la fonction de répartition, nous aurons la continuité à droite.

(vii) On peut calculer la probabilité pour que le point aléatoire tombe dans le rectangle de côté parallèle aux axes de coordonnées $R = [a,b] \times [c,d]$ (dont l'aire est hachurée sur la figure ci-dessus) :

$$P(a \le X < b,\ c \le Y < d) = F_{X,Y}(a,c) + F_{X,Y}(b,d) - F_{X,Y}(a,d) - F_{X,Y}(b,c)$$

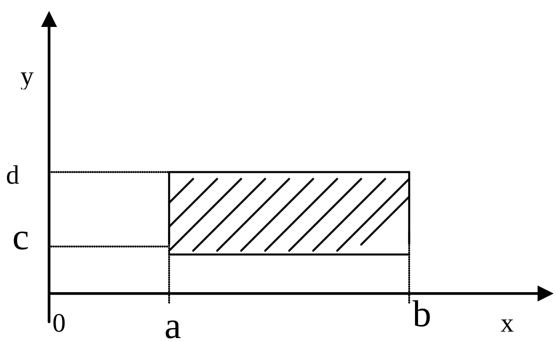


Figure 2. P(tomber dans le carré unité (a,b)(c,d)

5.3. Cas discret. Sa densité nécessite l'énumération de toutes les probabilités $p_{ij} = P(X = x_i, Y = y_j)$, $1 \le i \le N < \infty, 1 \le j \le M < \infty$ vérifiant $\sum_{i=1}^{N}\sum_{j=1}^{M} p_{ij} = 1$.

Les lois marginales s'écrivent

$$p_{i.} = P(X = x_i) = \sum_{i} p_{ij}$$

$$p_{.j} = P(Y = y_j) = \sum_{j} p_{ij}$$

La densité conditionnelle de X sachant Y est

$$P\{X = x_i / Y = y_j\} = \frac{P\{X = x_i, Y = y_j\}}{P\{Y = y_{.j}\}} = \frac{p_{ij}}{p_{.j}}$$

Exemple : Considérons le lancement de deux dés, et les variables aléatoires discrètes liées à ces expériences. X =nombre d'occurrence du «6 » et Y =nombre d'occurrence d'un nombre pair ».
On peut s'intéresser aux questions suivantes :

(i) Quelle est la loi du vecteur (X,Y) ?
(ii) Est-ce que les variables X et Y sont dépendantes ou indépendantes ?
(iii) Quelle est la loi de chacun des vecteurs X et Y ?
(iv) Calculer certaines caractéristiques telles que $P(X \ge Y)$, $E(X)$,...

Solution. Pour décrire la loi du vecteur discret (X,Y), on doit décrire l'ensemble des valeurs possibles (x_i, y_j) et les probabilités de ces valeurs p_{ij}. Les résultats peuvent être résumés dans un tableau et peuvent être aisément stockées dans la mémoire de votre ordinateur.

$\downarrow x_i \quad y_j \rightarrow$	0	1	2	$p_{i.} = P(X = x_i)$
0	1/4	1/3	1/9	25/36
1	0	1/6	1/9	10/36
2	0	0	1/36	1/36
$p_{.j} = P(Y = y_j)$	1/4	1/2	1/4	$\sum_{i=0}^{2}\sum_{j=0}^{2} pij = 1$

La première colonne représente les valeurs possibles de la variable X et la première ligne celles de la variable Y . Dans la dernière ligne, on trouve la loi marginale de la variable Y, et dans la dernière colonne celle de la variable X . La somme des éléments de la dernière ligne (ainsi que de celle de la dernière colonne) doit être égale à l'unité.

L'ensemble des évènements élémentaires peut être obtenu de deux manières :
-soit une seule expérience avec ensemble des évènements élémentaires $\Omega=\{(i,j): i=1,2,...,6;\ j=1,2,...,6\}$, card $\Omega=6^2=36$ (tirage de deux éléments avec remise)
-soit deux expériences indépendantes avec $\Omega=\Omega_1\times\Omega_2$, $\Omega_1=\Omega_2=\{1,2,3,4,5,6\}$

En considérant le premier procédé de construction de l'ensemble des évènements élémentaires, la probabilité $p_{11}=P(X=1,Y=1)=\frac{6}{36}=\frac{1}{6}$; en effet, l'événement $\{X=1,Y=1\}=\{(6,1),(6,3),(6,5),(1,6),(3,6),(5,6)\}$.

En utilisant, maintenant la seconde manière de construction de l'ensemble construction de l'ensemble des évènements élémentaires, et la *formule de multiplication des probabilités*

$$P\{X=1,Y=1\}=P(X=1)P(Y=1/X=1)$$

Notons que $P(Y=1/X=1)$ =P(au cours du jet d'un dé il sort ?????

5.4. **Cas continu.**

Dans le cas continu,, la fonction de répartition s'exprime en fonction de la densité $f_{(X,Y)}(x,y)$ vérifiant

(i) $F_{(X,Y)}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_{(X,Y)}(s,t)dsdt$

(ii) $f_{(X,Y)}(x,y)\geq 0.$

(iii) $\int_{-\infty}^{x}\int_{-\infty}^{y} f_{(X,Y)}(s,t)dsdt=1$

(iv) $f_{(X,Y)}(x,y)=\dfrac{\partial^2 F_{(X,Y)}(x,y)}{\partial x\partial y}$

(v) loi marginales

$$f_X(x)=\int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy \quad ; \qquad f_Y(y)=\int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx$$

(vi) La probabilité que le vecteur (X,Y) appartienne à un certain domaine D de IR^2 ;

$$P\{(X,Y)\in D\}=\iint_{x\in D} f_{(X,Y)}(x,y)dy$$

5.5. Loi conditionnelle

$$f_X(x/y)=\frac{f_{X,Y}(x,y)}{f_Y(y)} \quad , \; x\in IR \; , \qquad f_Y(y/x)=\frac{f_{X,Y}(x,y)}{f_X(x)} \quad , \; y\in IR$$

5.6. Indépendance

Les variables aléatoires X et Y sont indépendantes si

$$F_{X,Y}(x,y)=F_X(x).F_Y(y)$$

$$f_{X,Y}(x,y)=f_X(x).f_Y(y)$$

Dans le cas discret

$$P\{X=x,Y=y\}=P\{X=x\}\;.\;P\{Y=y\}$$

5.7. Espérance conditionnelle

$$\hat{y}_j=E[X/Y=y_j]=\sum_i x_i \;\mathrm{P}\{X=x_i\;/Y=y_j\}$$

On appelle espérance conditionnelle de la variable X sachant Y, la variable aléatoire $\hat{Y}=E(X/Y)$, dont les valeurs possibles $\hat{y}_j$ sont déterminées par la formule ( ) avec les probabilités correspondantes

$$P\{\hat{Y}=\hat{y}_j\}=P\{Y=y_j\}$$

Formule de l'espérance conditionnelle totale

$$E(X)=E[E[X/Y]]=\sum_j E[X/Y_j=y_j]P\{Y=y_j\}$$

Cas continu

$$E[X/Y=y]=\int_{-\infty}^{+\infty} xf_X(x/y)dx$$

où $f_x(x/y)$ est la densité conditionnelle

Formule de l'espérance conditionnelle totale

$$E(X) = E\big[E[X/Y]\big] = \int_{-\infty}^{+\infty} E\big[X/Y = y\big] f_Y(y)dy$$

5.8. Covariance et corrélation.

La covariance entre les variables X et Y est par définition

$$Cov(X,Y) = E\{(X - m_X)(Y - m_Y)\} = E(XY) - m_X m_Y$$

$$a_k = E(X_l, X_s) = \begin{cases} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^k y^s f_{X,Y}(x,y)dxdy & si \quad (X,Y) \quad vc.a.c \\ \sum_i \sum_j x_i^k y_j^s p_{ij} & si \quad (X,Y) \quad vc.a.d \end{cases}$$

$$\mu_k = \begin{cases} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x - m_x)^k (y - m_Y)^s f_{X,Y}(x,y)dxdy & si \quad (X,Y) \quad vc.a.c \\ \sum_i \sum_j (x_i - m_X)^k (y_j - m_Y)^s p_{ij} & si \quad (X,Y) \quad vc.a.d \end{cases}$$

5.9.Corrélation

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}, \text{ où } \sigma_X^2 = Var(X)$$

Propriétés :

(i) Le coefficient de corrélation est indépendant de l'unité de mesure

$$\rho(X,Y) = \rho(aX, bY) = \rho(X,Y), \quad \forall \quad a,b \in IR$$

(ii) $\mathrm{cov}(X,X) = Var(X)$ et $\rho(X,X) = 1$.

(iii) $\mathrm{cov}(X,Y) = \mathrm{cov}(Y,X)$

(iv) $\mathrm{cov}(aX + bY, Z) = a\,\mathrm{cov}(X,Z) + b\,\mathrm{cov}(Y,Z)$

(v) Si X et Y sont indépendantes, alors $\mathrm{cov}(X,Y) = 0$ et $\rho(X,Y) = 0$.

(vi) $\mathrm{cov}(X,Y)^2 \le Var(X).Var(Y)$

(vii) Le coefficient de corrélation $0 \le \rho(X,Y) \le 1$ (d'après l'inégalité de Schwartz).

(viii) $|\rho(X,Y)| = 1$ si et seulement si $\exists\ a,b \in IR \ /\ Y = aX + b$.

La dernière propriété signifie que le coefficient de corrélation mesure le degré de linéarité entre les deux variables.

Exercice : Montrons les propriétés (i), (ii) du coefficient de corrélation.

$$E\{(X - E(X) + t(Y - E(Y)\}^2 = Var(X) + 2t\,\mathrm{cov}(X,Y) + t^2 Var(Y)\ \} \geq 0 \quad \forall\ t$$

Le discriminant du trinôme en t

$$\Delta = \mathrm{cov}(X,Y)^2 - Var(X)Var(Y) \leq 0 \text{ d'où}$$

$$|\mathrm{cov}(X,Y)| \leq \sigma X\, \sigma Y \quad \text{, par conséquent } |\rho(X,Y)| \leq 1.$$

Dans le cas où $\rho = \pm 1$ (discriminant nul) si et seulement si les racines du trinôme sont réelles et égales $t_1 = t_2 = t_0$.

On a $E\{(X - E(X) + t_0(Y - E(Y)\}^2 = 0$, alors

$$P\{X - E(X) + t_0(Y - E(Y) = 0\} = 1$$

Par conséquent, ρ peut être considéré comme une mesure de dépendance linéaire entre X et Y.

Lorsque $\rho(X,Y)=0$, on dit que les vecteurs X et Y sont non corrélés. En particulier, si X et Y sont indépendants, alors la covariance est nulle (car $E(XY) = E(X)E(Y)$ par définition) et ils sont non corrélés. La réciproque est fausse

Exercice : Considérons un contre exemple : $g(x) = \begin{cases} e^{-x}, & si \quad x \geq 0 \\ 0, & si \quad x < 0 \end{cases}$

$$f(x,y) = \frac{g(\sqrt{x^2+y^2})}{2\pi\sqrt{x^2+y^2}}$$ densité d'un certain couple de variables aléatoires (X,Y).

•**Interprétation géométrique**. L'ensemble des variables aléatoires définies sur une même espace probabilisé, de moyenne nulle et de variance finie ($E(X)=0, Var(X)<\infty$) forme un espace vectoriel E(+,.). Toute variable aléatoire sera ainsi assimilée à un vecteur. On peut munir cet espace d'un produit scalaire de la manière suivante : $<X,Y>=E(XY)=\mathrm{cov}(X,Y)$. (E,<.,.>) est un espace euclidien. La longueur (norme) du vecteur X est $\sigma_X = \|X\|$ et $Var(X) = \|X\|^2$.

Si on considère maintenant deux vecteurs X et Y , alors

$$\rho(X,Y)=\frac{\operatorname{cov}(X,Y)}{\sigma_X\sigma_Y}=\frac{\|X\|^2}{\|X\|.\|Y\|}=\cos\varphi$$ où φ est l'angle entre les vecteurs X et Y.

•Si deux variables X et Y sont non corrélées, alors leur covariance est nulle $\rho(X,Y)=0$ ($\Rightarrow\varphi=\pi/2$ et le produit scalaire dans notre espace vectoriel est nul): ce qui signifie que le vecteur X est orthogonal à Y.

•si $\rho=\pm1$, alors φ=0 ou π et les vecteurs X et Y sont colinéaires (on a vu en effet que dans ce cas $X=t_0Y$). Si ρ=+1, ils ont même direction, si ρ=-1, ils ont des directions opposées.

•Quelques propriétés

(i) $E(X+Y)=E(X)+E(Y)$

(ii) $Cov(X,Y)=E(XY)-E(X).E(Y)$

(iii) $Var(X)=Cov(X,X)=E(X^2)-(E(X))^2$

(iv) $E(aX+b)=aE(X)+b$ pour toutes constantes $a,b\in IR$

(v) $Var(aX+b)=a^2Var(X)$ pour toutes constantes $a,b\in IR$

(vi) $Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$

(vii) En particulier si X et Y indépendantes, on a $Cov(X,Y)=0$.

La variable $U=(X-m)/\sigma$ s'appelle variable centrée réduite car $E(U)=0$ et $Var(U)=1$ (propriétés (iv) et (v)).

Exemple 2 : Loi normale dans le plan $f_{X,Y}(x,y)=\dfrac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}}\times$

$$\times \exp\left\{-\frac{1}{2(1-\rho_{X,Y}^2)}\left[\frac{(x-m_X)^2}{\sigma_X^2}-\frac{2\rho_{X,Y}\left(x-m_X\right)\left(y-m_Y\right)}{\sigma_X\sigma_Y}+\frac{(y-m_Y)^2}{\sigma_Y^2}\right]\right\}$$

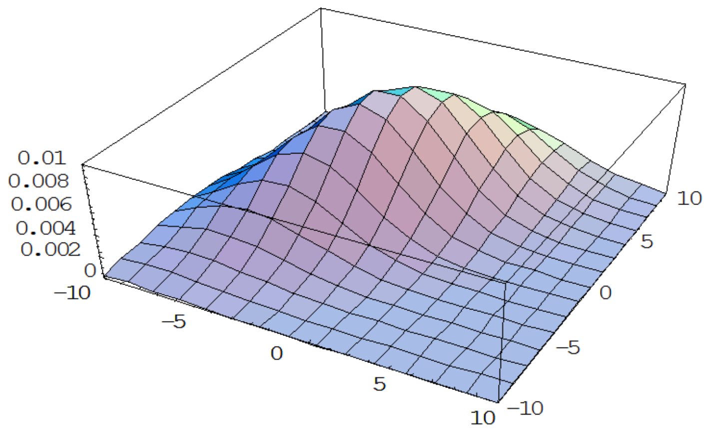


Figure 6. Densité de la loi Normale bidimensionnelle de moyenne et de matrice des covariances

$$K=\begin{pmatrix}16 & 12\\ 12 & 25\end{pmatrix}$$