# Chapitre 2: Vecteurs aléatoires gaussiens Lois conditionnelles

LAKHEL El Hassan

Département: Informatiques et Réseaux-Télécom.
Filière G.I.I.A.
ENSA, A. U. 2022-2023

Page Web: https://sites.google.com/view/lakhelelhassan/modélisation stochastiques

# Plan :

- Vecteurs aléatoires à valeurs dans $\mathbb{R}^n$
- Fonction caractéristique
- Cas $n = 2$, Couples aléatoires
- Distributions conditionnelles
- Espérance, Covariance et Coefficient de corrélation linéaire
- Vecteurs aléatoires Gaussiens à valeurs dans $\mathbb{R}^n$.

### Définition

*Une application de $(\Omega, \mathcal{F})$ à valeurs dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ est un vecteur aléatoire si chacun de ses composantes est une v. a. réelle.*

Notation : Un vecteur de v. a. est noté par :

$$X : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$$

$$\omega \longmapsto X(\omega) = \begin{pmatrix} X_1(\omega) \\ X_2(\omega) \\ \vdots \\ X_n(\omega) \end{pmatrix}$$

Il est d'usage d'identifier les vecteurs de $\mathbb{R}^n$ à des matrices-colonnes.
Ainsi

$$^tX = (X_1 \, X_2 \ldots X_n)$$

est une matrice-ligne.

## Définition

*Une application de $(\Omega, \mathcal{F})$ à valeurs dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ est un vecteur aléatoire si chacun de ses composantes est une v. a. réelle.*

Notation : Un vecteur de v. a. est noté par :

$$\begin{aligned} X : (\Omega, \mathcal{F}) & \longrightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \\ \omega & \longmapsto X(\omega) = \begin{pmatrix} X_1(\omega) \\ X_2(\omega) \\ \vdots \\ X_n(\omega) \end{pmatrix} \end{aligned}$$

Il est d'usage d'identifier les vecteurs de $\mathbb{R}^n$ à des matrices-colonnes.
Ainsi

$$^tX = (X_1\ X_2 \dots X_n)$$

est une matrice-ligne.

Produit scalaire : Si $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{R}^n$ et $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \in \mathbb{R}^n$, le p.s. s'écrira comme le produit des matrices :

$$< u|X >= \ ^tX.u = \ ^tu.X = \sum_{i=1}^{n} u_i.X_i.$$

# Loi d'un vecteur aléatoire

## Définition

1. *La loi du vecteur $X$ est une mesure de probabilité sur $\mathbb{R}^n$ muni de sa tribu de Borel $\mathcal{B}_{\mathbb{R}^n}$. Elle est notée $\mathbb{P}_X$ et se définit de la manière suivante :*

$$\forall B \in \mathcal{B}_{\mathbb{R}^n}, \quad \mathbb{P}_X(B) = \mathbb{P}(X \in B).$$

2. *On dira que le vecteur $X$ admet une densité s'il existe une fonction positive définie sur $\mathbb{R}^n$ et d'intégrale égale à 1. Cette densité est appelé e la densité (conjointe) sera alors notée $f_{(X_1,\ldots,X_n)}$. On écrira dans ce cas*

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \int_B f_{(X_1,\ldots,X_n)}(x_1, x_2, ..., x_n)dx_1...dx_n.$$

3. *La loi de la coordonnée $X_1$ (loi marginale) est alors la loi de densité :*

$$f_{X_1}(x_1) = \int_{\mathbb{R}^{n-1}} f(x_1, x_2, ..., x_n)dx_2...dx_n.$$

4. *La loi marginale de la coordonnée i admet une densité que l'on calcule, en général, en fixant la variable $x_i$ et en intégrant $f$ suivant les autres variables.*

# Fonction de répartition

### Définition

*On appelle fonction de répartition (conjointe) du vecteur aléatoire $X = (X_1, ..., X_n)$ l'application $F_X$ définie sur $R^n$ et à valeurs dans $[0, 1]$ par :*

$$F_X(x) = F_X(x_1, ..., x_n) = \mathbb{P}_X(]-\infty, x_1] \times ... \times ]-\infty, x_n]) = \mathbb{P}(\{X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n\}.$$

### Proposition

*La fonction de répartition conjointe d'un vecteur aléatoire $X = (X_1, ..., X_n)$ permet de déterminer les fonctions de répartition de toutes les marges.*

### Démonstration.

On a, pour tout $i = 1, ....; n$,

$$F_{X_i}(x_i) = lim_{x_1,...x_{i-1},x_{i+1},...,x_n \to +\infty} F_X(x) = F_X(+\infty, ...., +\infty, x_i, +\infty, ..., +\infty)$$

De même :

$$F_{X_1,X_2}(x_1, x_2) = lim_{\forall j>2, x_j \to +\infty} F_X(x) = F_X(x_1, x_2, +\infty, ..., +\infty),$$

et ainsi de suite pour les autres marges.

# Espérance et covariance

## Définition

1. *On dira que* $X = (X_1, \ldots, X_n)$ *est intégrable (resp. de carré intégrable.) si chacunes de ses composantes* $X_1, \ldots, X_n : \Omega \longrightarrow \mathbb{R}$ *l'est.*
2. *Espérance : Le vecteur moyen de* $X$ *est défini par*

$$\mathbb{E}(X) = (\mathbb{E}(X_1), \ldots, \mathbb{E}(X_n)),$$

*ou, matriciellement :*

$$\mathbb{E}(X) = \begin{pmatrix} \mathbb{E}(X_1) \\ \mathbb{E}(X_2) \\ \vdots \\ \mathbb{E}(X_n) \end{pmatrix}$$

# Espérance et covariance

## Définition

1. *On dira que* $X = (X_1, \dots, X_n)$ *est intégrable (resp. de carré intégrable.) si chacunes de ses composantes* $X_1, \dots, X_n : \Omega \longrightarrow \mathbb{R}$ *l'est.*
2. *Espérance : Le vecteur moyen de* $X$ *est défini par*

$$\mathbb{E}(X) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_n)),$$

*ou, matriciellement :*

$$\mathbb{E}(X) = \begin{pmatrix} \mathbb{E}(X_1) \\ \mathbb{E}(X_2) \\ \vdots \\ \mathbb{E}(X_n) \end{pmatrix}$$

## Exercice

*Soit* $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *une application linéaire, et* $X : \Omega \longrightarrow \mathbb{R}^n$ *un vecteur aléatoire. Montrer que*

$$\mathbb{E}(AX) = A\mathbb{E}(X).$$

Solution :
Il suffit d'écrire

$$AX = (\sum_{j=1}^{n} a_{i,j} X_j)_{1 \leq i \leq m}.$$

Donc

$$\mathbb{E}(AX) = (\sum_{j=1}^{n} a_{i,j} \mathbb{E}(X_j))_{1 \leq i \leq m} = A\mathbb{E}(X).$$

En particulier, pour tout $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{R}^n$ et $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \in \mathbb{R}^n$, l'espérance du p.s. s'érira comme le produit des matrices :

$$\mathbb{E} < u|X > = < u|\mathbb{E}(X) > .$$

Solution :
Il suffit d'écrire

$$AX = (\sum_{j=1}^{n} a_{i,j} X_j)_{1 \leq i \leq m}.$$

Donc

$$\mathbb{E}(AX) = (\sum_{j=1}^{n} a_{i,j} \mathbb{E}(X_j))_{1 \leq i \leq m} = A\mathbb{E}(X).$$

En particulier, pour tout $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{R}^n$ et $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \in \mathbb{R}^n$, l'espérance du p.s. s'érira comme le produit des matrices :

$$\mathbb{E} < u|X > = < u|\mathbb{E}(X) > .$$

En effet, l'application $A : x \longrightarrow^t u.x$ est une forme linéaire : Ainsi

$$\mathbb{E}AX = A\mathbb{E}X = {}^t u.\mathbb{E}X.$$

La notion de matrice de covariance remplace celle de variance pour les variables aléatoires réelles.

Définition

- *Covariance : Si $X \in \mathbb{R}^n$ et $Y \in \mathbb{R}^m$. La matrice de covariance $K_{X,Y} \in \mathbb{R}^{n,m}$ est définie par*

$$K_{X,Y} = \mathbb{E}[(X - \mathbb{E}(X)).^t(Y - \mathbb{E}(Y))].$$

- *Les éléments de la matrice de covariance sont donnés par : pour $1 \leq i \leq n$ et $1 \leq j \leq m$*

$$K_{X,Y}(i,j) = Cov(X_i, Y_j) = \mathbb{E}[(X_i - E[X_i])(Y_j - E[Y_j])]$$

Exercice

1. *Donner $K_X := K_{X,X} \in \mathbb{R}^{n,n}$ lorsque $X = Y$.*
2. *Expliciter les cas $n = 1$, $n = 2$ et $n = 3$ lorsque $X = Y$.*

Une matrice de variance/covariance est une matrice carrée qui comporte les variances et les covariances associées à plusieurs variables. Les éléments de diagonale de la matrice contiennent les variances des variables, tandis que les éléments hors diagonale contiennent les covariances entre toutes les paires possibles de variables.

### Définition

*On appelle matrice de variance/covariance du vecteur aléatoire $X$ la matrice carrée $K_X$ de taille $n$ :*

$$K_X = Var(X) = \mathbb{E}[(X - E[X])\ {}^t(X - E[X])] = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \dots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Var(X_2) & \dots & Cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \dots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \dots & Var(X_n) \end{pmatrix}$$

*les coefficients de $K_X$ sont données par :*

$$K_X(i,j) = Cov(X_i, X_j) = \mathbb{E}[(X_i - E(X_i))(X_j - E(X_j))]$$

### Exemples

- *Si $n = 1$, on obtient $K_X = Var(X_1)$, avec $X = X_1$*
- *Si $n = 2$, on obtient*

$$K_X = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Var(X_2) \end{pmatrix},$$

*avec $X = (X_1, X_2)$*

# fonction caractéristique : Propriété fondamentale

**Définition**

*Soit $X$ un vecteur aléatoire. Sa fonction caractéristique est définie par*

$$\phi_X(u) = \mathbb{E}e^{i\,{}^t uX} = \mathbb{E}e^{i\sum_{j=1}^{n} u_j X_j} \quad \forall u \in \mathbb{R}^n.$$

Le résultat suivant assure que la loi d'une variable aléatoire est déterminée par sa fonction caractéristique

**Théorème**

*Deux v.a. possédant la même fonction caractéristique ont la même loi : si $X$ et $Y$ sont deux v.a. à valeurs dans $R^n$ telles que $\phi_X = \phi_Y$ , alors $\mathbb{P}_X = \mathbb{P}_Y$ .*

**Preuve.** La transformée de Fourier est une application injective de $\mathcal{M}_b(\mathbb{R}^n)$ (ensemble des mesures bornées) vers $\mathcal{C}_b(\mathbb{R}^n, \mathbb{C})$ (ensemble des fonctions continues bornées).

Dans ce qui suit, nous allons expliquer les notions précédentes pour les couples de variables aléatoires. Le dernier paragraphe sera consacré aux vecteurs gaussiens de dimension supérieure.

### Définition

1. *On dit qu'un couple aléatoire est discret s'il prend ses valeurs dans un ensemble* $\{(x_i, y_j)_{(i,j)\in I\times J}, x_i, y_j \in \mathbb{R}\}$ *avec* $I \times J \subset \mathbb{N}^2$ *ou dans* $\mathbb{Z}^2$.
2. *La loi conjointe de* $X$ *et* $Y$ *(i.e. la loi du couple* $(X, Y)$*) est la donnée des nombres positifs :*

$$p_{ij} = P(\{X = x_i\} \cap \{Y = y_j\}), \qquad (i,j) \in I \times J$$

*avec*

$$\sum_{(i,j)\in I\times J} p_{ij} = 1.$$

3. *Si l'on s'intéresse uniquement au comportement de l'une des deux variables, on introduit les lois marginales de* $X$ *et* $Y$ *définies par*

$$p_{i.} = P(X = x_i) = \sum_{j\in J} p_{ij} \qquad \text{et} \qquad p_{.j} = P(Y = y_j) = \sum_{i\in I} p_{ij}.$$

### Remarque

*Les deux événements* $(X = x_i)$ *et* $(\cup_{j\in J} Y = y_j, X = x_i)$ *sont identiques.*
*Donc*

$$p_{i.} = \sum_{j\in J} p_{ij}.$$

*De plus,*

$$\sum_{i\in I} p_{i.} = \sum_{i\in I}\sum_{j\in J} p_{ij} = \sum_{(i,j)\in I\times J} p_{ij} = 1.$$

### Exemples

*On tire, avec remise, deux boules d'un sac contenant 3 boules numérotées 1, 2, et 3.*
*On considère la v. a. $X$ qui est la somme des points obtenus et soit $Y$ le maximum des points obtenus.*

1. *Calculer la loi conjointe de $X$ et $Y$.*
2. *Déduire les lois marginales de $X$ et $Y$.*

### Exemples

*On tire, avec remise, deux boules d'un sac contenant 3 boules numérotées 1, 2, et 3.*
*On considère la v. a. $X$ qui est la somme des points obtenus et soit $Y$ le maximum des points obtenus.*

1. *Calculer la loi conjointe de $X$ et $Y$.*
2. *Déduire les lois marginales de $X$ et $Y$.*

*On a*

$$X(\Omega) = \{2, 3, 4, 5, 6\} \qquad \textit{et} \qquad Y(\Omega) = \{1, 2, 3\}.$$

*Il s'agit d'évaluer $P(X = i, Y = j)$ lorsque $2 \leq i \leq 6$ et $1 \leq j \leq 3$.*
*Les résultats sont donnés dans le tableau suivant :*

| $y_j \| x_i$ | 2 | 3 | 4 | 5 | 6 | Loi de $Y$ |
|---|---|---|---|---|---|---|
| 1 | $\frac{1}{9}$ | 0 | 0 | 0 | 0 | $\frac{1}{9}$ |
| 2 | 0 | $\frac{2}{9}$ | $\frac{1}{9}$ | 0 | 0 | $\frac{3}{9}$ |
| 3 | 0 | 0 | $\frac{2}{9}$ | $\frac{2}{9}$ | $\frac{1}{9}$ | $\frac{5}{9}$ |
| Loi de $X$ | $\frac{1}{9}$ | $\frac{2}{9}$ | $\frac{3}{9}$ | $\frac{2}{9}$ | $\frac{1}{9}$ | 1 |

Les cas possibles sont :

$$\begin{array}{l}
(1,1) \longrightarrow X = 2 \\
(1,2),(2,1) \longrightarrow X = 3 \\
(1,3)(3,1)(2,2) \longrightarrow X = 4 \\
(3,2)(2,3) \longrightarrow X = 5 \\
(3,3) \longrightarrow X = 6.
\end{array}$$

La formule donnant l'espérance de $h(X)$ se généralise au contexte d'un couple de v. a. de $X$ et $Y$. Soit $h : \mathbb{R}^2 \to \mathbb{R}$ telle que $\mathbb{E}|h(X, Y)| < \infty$, alors

$$E(h(X, Y)) = \sum_{i \in I} \sum_{j \in J} h(x_i, y_j) p_{i,j},$$

où $h$ est une fonction continue bornée de deux variables.

## Définition

1. *On dit qu'un couple aléatoire* $(X, Y)$ *est à densité s'il existe une fonction* $f_{(X,Y)} : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *mesurable positive telle que pour tout domaine* $D$ *de* $\mathbb{R}^2$*, on a*

$$P_{(X,Y)}(D) = P((X, Y) \in D) = \int\int_D f_{(X,Y)}(x, y)dxdy.$$

*avec* $\int_{\mathbb{R}^2} f_{(X,Y)}(x, y)dxdy = 1.$

2. *Les densités marginales de* $X$ *et* $Y$ *sont données par :*

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x, y)dy \qquad et \qquad f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x, y)dx.$$

### Exercice

*On pose*

$$f(x,y) = cexp(-\frac{x^2+y^2}{2}).$$

1. *Déterminer c pour que f soit une densité de probabilité d'un vecteur aléatoire* $Z = (X, Y)$ *de* $\mathbb{R}^2$.
2. *Calculer les densités des lois marginales de* $Z$.

## Exercice

*On pose*

$$f(x,y) = cexp(-\frac{x^2+y^2}{2}).$$

1. *Déterminer $c$ pour que $f$ soit une densité de probabilité d'un vecteur aléatoire $Z = (X, Y)$ de $\mathbb{R}^2$.*
2. *Calculer les densités des lois marginales de $Z$.*

**Solution :**

1. 
   - $f$ est continue de $\mathbb{R}^2$ dans $\mathbb{R}$, donc mesurable.
   - $f \geq 0$ si $c \geq 0$.
   - pour l'intégrale :

$$\begin{aligned} 1 &= \int\int_{\mathbb{R}^2} f(x,y)dxdy = c\int\int_{\mathbb{R}^2} e^{-\frac{x^2+y^2}{2}}dxdy \\ &= c(\int_{\mathbb{R}} e^{-\frac{x^2}{2}}dx)(\int_{\mathbb{R}} e^{-\frac{y^2}{2}}dy) \\ &= c\sqrt{2\pi}\sqrt{2\pi}. \end{aligned}$$

   Donc

$$c = \frac{1}{2\pi}.$$

2.

$$f_X(x) = \int_{\mathbb{R}} \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} (\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Donc

$$X \sim \mathcal{N}(0,1).$$

De même :

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}.$$

Donc

$$Y \sim \mathcal{N}(0,1).$$

Deux variables aléatoires $X$ et $Y$ quelconques (discrètes ou continues) peuvent être entièrement caractérisées par leur fonction de répartition conjointe, qui est l'extension au cas de deux variables de la notion de fonction de répartition déja présentée au chapitre précédent.

### Définition

*Pour un couple $(X, Y)$, la fonction de répartition conjointe $F(x, y)$ est donnée par :*

$$F(x, y) = \mathbb{P}(X \leq x, Y \leq y).$$

### Remarque

1. *La fonction $F(x, y)$ prend ses valeurs dans l'intervalle $[0, 1]$, avec :*

$$F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0 \qquad F(\infty, \infty) = 1.$$

   *En effet, posons les événements $A = (X \leq x)$ et $B = (Y \leq y)$, on $F(x, y) = \mathbb{P}(A \cap B)$. L'événement $A = (X \leq -\infty)$ est l'événement impossible, donc $F(-\infty, y) = \mathbb{P}(\emptyset \cap B) = 0$. Les autres égalités se démontrent de façon similaire.*

2. *La fonction $F(x, y)$ est croissante par rapport à $x$ et $y$ :*

$$F(x_i, y_i) \leq F(x_j, y_j) \quad \forall x_i \leq x_j, y_i \leq y_j.$$

En général, la donnée des lois de probabilités de deux variables aléatoires $X$ et $Y$ ne permet pas de calculer la loi du couple aléatoire $(X, Y)$ (alors que la réciproque est vraie). C'est toutefois possible dans un cas particulier très important, celui des variables aléatoires indépendantes.

### Définition

1. *Soit $X$ et $Y$ deux v. a. discrètes à valeurs respectivement dans $\{a_i\}$ et $\{b_j\}$. Les v. a. $X$ et $Y$ sont indépendantes si et seulement si :*

$$P_{ij} = P(\{X = a_i\} \cap \{Y = b_j\}) = P(X = a_i)P(Y = b_j).$$

2. *Soit $X$ et $Y$ deux v. a. à densité. Les v. a. $X$ et $Y$ sont indépendantes si et seulement si :*

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

**Conséquence :**

$$X \amalg Y \Longrightarrow E(XY) = E(X)E(Y).$$

En effet,

- Cas continu : le théorème de transfert nous permis de montrer que $XY$ est intégrable

$$\begin{aligned} E(|XY|) &= \int\int_{\mathbb{R}^2} |xy| f_{(X,Y)}(x,y)dxdy \\ &= \int_{\mathbb{R}} |x| f_X(x)dx \int_{\mathbb{R}} |y| f_Y(y)dy. \\ &= E(|X|)E(|Y|) < \infty. \end{aligned}$$

- Maintenant, on a

$$\begin{aligned} E(XY) &= \int\int_{\mathbb{R}^2} xy f_{(X,Y)}(x,y)dxdy \\ &= \int_{\mathbb{R}} x f_X(x)dx \int_{\mathbb{R}} y f_Y(y)dy. \\ &= E(X)E(Y). \end{aligned}$$

- De même pour le cas discret.

**Attention** :
Piége à éviter : $\mathbb{E}(\frac{X}{Y})$ n'est pas égal à $\frac{\mathbb{E}(X)}{\mathbb{E}(Y)}$ même si $X$ et $Y$ sont indépendantes.
si $X$ et $Y$ sont indépendantes. Alors

$$\mathbb{E}(\frac{X}{Y}) = \mathbb{E}(X) \times \mathbb{E}(\frac{1}{Y}) \neq \frac{\mathbb{E}(X)}{\mathbb{E}(Y)}.$$

Car en général, $\mathbb{E}(\frac{1}{Y}) \neq \frac{1}{\mathbb{E}(Y)}$.
Soit $Y \sim \mathcal{U}_{(1,3)}$.

**Attention** :
Piége à éviter : $\mathbb{E}(\frac{X}{Y})$ n'est pas égal à $\frac{\mathbb{E}(X)}{\mathbb{E}(Y)}$ même si $X$ et $Y$ sont indépendantes.
si $X$ et $Y$ sont indépendantes. Alors

$$\mathbb{E}(\frac{X}{Y}) = \mathbb{E}(X) \times \mathbb{E}(\frac{1}{Y}) \neq \frac{\mathbb{E}(X)}{\mathbb{E}(Y)}.$$

Car en général, $\mathbb{E}(\frac{1}{Y}) \neq \frac{1}{\mathbb{E}(Y)}$.
Soit $Y \sim \mathcal{U}_{(1,3)}$.
Alors par définition

$$\mathbb{E}(Y) = \frac{1+3}{2} = 2 \quad \Longrightarrow \frac{1}{\mathbb{E}(Y)} = \frac{1}{2}.$$

Or,

$$\mathbb{E}(\frac{1}{Y}) = \int_1^3 \frac{1}{x}\frac{1}{3-1}dx = \frac{1}{2}\ln 3 \neq \frac{1}{2}.$$

**Remarque**

*Si $X$ et $Y$ sont deux variables aléatoires indépendantes et si $f$ et $g$ sont deux fonctions continues $\mathbb{R} \longrightarrow \mathbb{R}$, alors, sous réserve d'existence de ces quantités, on a :*

$$\mathbb{E}(f(X)g(Y)) = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$$

En effet,

1. Le cas continu : On utilisera la définition précédente et le théorème de Fubini pour les intégrales doubles.

$$\begin{aligned}
\mathbb{E}(f(X)g(Y)) &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x)g(y)f_{(X,Y)}(x,y) = \int_{\mathbb{R}}\int_{\mathbb{R}} f(x)g(y)f_X(x)f_Y(y)dxdy \\
&=^{\text{var. séparées}} \int_{\mathbb{R}} f(x)f_X(x))dx \int_{\mathbb{R}} g(y)f_Y(y)dy \\
&= \mathbb{E}[f(X)]\mathbb{E}[g(Y)]
\end{aligned}$$

2. De même pour le cas discret.

## Exercice

*Pour éviter l'inconvénient majeur que représente la panne d'une machine fragile et indisponsable au bon fonctionnement d'une chaîne de production, on envisage de prévoir une machine de secours, identique en tout point, qui se met automatiquement en marche lorsque la première machine tombe en panne.*
*La machine est mise en marche le matin à 8h00 et est arrêtée le soir à 18h00, et on ne peut envisager de réparer une machine qu'après 18h00 le soir.*
*On suppose que les durées de fonctionnement sans l'occurrence d'une panne à partir de la mise en marche sont indépendantes et de même loi exponentielle de paramètre* $\lambda = 0,01h^{-1}$.

1. *quelle est la probabilité que la chaîne de production s'arrête s'il y a une seule machine ?*
2. *quelle est la probabilité que la chaîne de production s'arrête s'il y une seconde machine de secours ?*

**Solution**

1. S'il y a une seule machine dont le temps de foctionnement $X$ suit $Exp(0,01)$. On a

$$\mathbb{P}(X < 10) = F_X(10) = 1 - e^{-0,1} = 0,0952.$$

2. S'il y a une machine de secours, son temps de fonctionnement est $Y \sim Exp(0,01)$. On suppose que $X \perp Y$. On a donc par le théorème de transfère

$$\begin{aligned}
\mathbb{P}(X + Y < 10) &= \mathbb{P}_{(X,Y)}(x + y < 10) \\
&= \int_{\mathbb{R}^2} 1_{\{x+y<10\}} f_{(X,Y)}(x,y)dxdy \\
&= \int_0^{10} \int_0^{10-x} f_X(x) f_Y(y)dxdy \\
&= \int_0^{10} \lambda e^{-\lambda x}(1 - e^{\lambda x - 10\lambda})dx \\
&= [-e^{-\lambda x} - x\lambda e^{10\lambda}]_0^{10} \\
&= 0,0047.
\end{aligned}$$

Le risque de voir la chaîne de production s'arrêter est donc 20 fois moindre si l'on prévoit une machine de secours.

Un des intérêts majeurs qu'il y a à définir le couple $(X, Y)$ est d'obtenir des informations sur l'une des variables lorsque des informations sont disponibles sur l'autre variable, en exploitant le lien existant entre elles.
La loi de probabilité conditionnelle de $X$ sachant l'événement $\{Y = y\}$ est donnée par le fait que c'est une loi sur $X(\Omega)$ ainsi que par les probabilités conditionnelles $\mathbb{P}(X = x/Y = y)$ pour tout $x \in X(\Omega)$.
On pourra donc écrire :

### Définition

*Soit $y_j$ fixé, tel que $p_Y(y_j) := \mathbb{P}(Y = y_j) > 0$. Pour un couple $(X, Y)$ discret, la loi de probabilité de $X$ conditionnelle à $Y = y_j$ est donnée par :*

$$p_X(x/y_j) = \mathbb{P}(X = x/Y = y_j) = \frac{p(x, y_j)}{p_Y(y_j)}.$$

*On peut, bien sûr, de manière symétrique définir la loi de probabilité de $Y$ conditionnelle à $X = x_i$.*

### Exercice

*soient* $X \sim \mathcal{P}(\lambda)$ *et* $Y \sim \mathcal{P}(\mu)$ *indépendantes. Montrer que la loi de probabilité conditionnelle de* $X$ *sachant l'événement* $\{X + Y = n\}$ *est égale à la loi binomiale* $\mathcal{B}(n, \frac{\lambda}{\lambda+\mu})$.

Etant donne que $X+Y \sim \mathcal{P}(\lambda+\mu)$, on a pour tout $0 \leq i \leq n$

$$\begin{aligned}
\mathbb{P}(X=i/X+Y=n) &= \frac{\mathbb{P}([X=i]\cap[X+Y=n])}{\mathbb{P}(X+Y=n))} = \frac{\mathbb{P}([X=i]\cap[Y=n-i])}{\mathbb{P}(X+Y=n))} \\
&= \frac{\mathbb{P}([X=i])\mathbb{P}([Y=n-i])}{\mathbb{P}(X+Y=n))} = \frac{e^{-\lambda}.\frac{\lambda^i}{i!}e^{-\mu}.\frac{\mu^{n-i}}{(n-i)!}}{e^{-(\lambda+\mu)}.\frac{(\lambda+\mu)^n}{n!}} \\
&= C_n^i(\frac{\lambda}{\lambda+\mu})^i.(1-\frac{\lambda}{\lambda+\mu})^{n-i} = \mathbb{P}_Z(\{i\})
\end{aligned}$$

avec $Z$ est de loi binomiale $\mathcal{B}(n, \frac{\lambda}{\lambda+\mu})$. Ainsi, la loi conditionnelle de $X$ à $[X+Y=n]$ est la loi binomiale $\mathcal{B}(n, \frac{\lambda}{\lambda+\mu})$.

Dans le cas d'un couple de variables aléatoires continues, les fonctions de densité de probabilités conditionnelles sont données par :

### Définition

*Pour un couple $(X, Y)$ continu,la formule qui donne la fonction de densité conditionnelle de $X$ sachant que $Y = y$ est donnée par :*

$$f_X(x/y) = \frac{f(x,y)}{f_Y(y)} \qquad \textit{avec} \qquad f_Y(y) > 0.$$

*Bien sûr, on définit de manière tout à fait similaire la loi de $Y$ sachant que $X = x$.*

**Exercice :** Soit la fonction $f(x,y)$ donnée par :

$$f(x,y) = \begin{cases} 6x & \textit{pour} \quad 0 \le x \le y \le 1 \\ 0 & \textit{autrement} \end{cases}$$

1. Calculer les fonctions de densités marginales $f_X(x)$ et $f_Y(y)$.
2. Trouver les distributions conditionnelles : $f_X(x/y)$ et $f_Y(y/x)$.

**Solution :**

1. Les distributions marginalessont données par :

$$f_X(x) = \begin{cases} \int_x^1 6xdy = 6x(1-x) & x \in [0,1] \\ 0 & \textit{ailleurs} \end{cases}$$

$$f_Y(y) = \begin{cases} \int_0^y 6xdx = 3y^2 & y \in [0,1] \\ 0 & \textit{ailleurs} \end{cases}$$

2. Les distributions conditionnelles sont données par :

$$f_X(x/y) = \begin{cases} \frac{6x}{3y^2} = \frac{2x}{y^2} & \textit{pour} \quad 0 \leq x \leq y \leq 1 \\ \\ 0 & \textit{ailleurs} \end{cases}$$

$$f_Y(y/x) = \begin{cases} \frac{6x}{6x(1-x)} = \frac{1}{1-x} & \textit{pour} \quad 0 < x \leq y < 1 \\ \\ 0 & \textit{ailleurs} \end{cases}$$

### Définition

*Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Soit $Z = (X, Y)$ un couple aléatoire. On note quand cette quantité existe,*

$$\mathbb{E}Z = (\mathbb{E}X, \mathbb{E}Y).$$

Pour mesurer le lien entre deux grandeurs aléatoires, on fait appel à la notion de covariance :

### Définition

*Soit $(X, Y)$ un couple aléatoire tel que $X$ et $Y$ admettent toutes les deux un moment d'ordre 2. On appelle covariance de $X$ et $Y$ et on note $Cov(X, Y)$ la quantité :*

$$Cov(X, Y) = \mathbb{E}([X - \mathbb{E}(X)][Y - \mathbb{E}(Y)]).$$

### Remarque

*On peut montrer qu'il existe une autre formule pour calculer la covariance :*

$$cov(X, Y) = E(XY) - E(X)E(Y).$$

### Définition

*La matrice de covariance $K_Z$ du vecteur $Z = (X, Y)$ est définie par*

$$K_Z = \mathbb{E}\left[\underbrace{\begin{pmatrix} X - \mathbb{E}(X) \\ Y - \mathbb{E}(Y) \end{pmatrix}}_{Z-\mathbb{E}(Z)} \underbrace{\begin{pmatrix} X - \mathbb{E}(X) & Y - \mathbb{E}(Y) \end{pmatrix}}_{{}^t(Z-\mathbb{E}(Z))}\right] = \begin{pmatrix} Var(X) & Cov(X, Y) \\ Cov(Y, X) & Var(Y) \end{pmatrix}.$$

## Remarque

1. *Si* $X \amalg Y$, *alors* $Cov(X, Y) = 0$, *et la matrice de covariance est diagonale.*
2. *Mais la réciproque est fausse.*
   *En effet, Soit une v. a.* $X$ *pour laquelle* $P_X(\{x\}) = \frac{1}{3}$ *pour* $x = -1, 0, 1$. *Si l'on définit une nouvelle variable* $Y = X^2$, *la fonction* $p(x, y)$ *du couple est donc :*

| $y \backslash x$ | -1 | 0 | 1 |
|---|---|---|---|
| 0 | 0 | $\frac{1}{3}$ | 0 |
| 1 | $\frac{1}{3}$ | 0 | $\frac{1}{3}$ |

## Remarque

1. *Si $X \amalg Y$, alors $Cov(X, Y) = 0$, et la matrice de covariance est diagonale.*
2. *Mais la réciproque est fausse.*
   *En effet, Soit une v. a. $X$ pour laquelle $P_X(\{x\}) = \frac{1}{3}$ pour $x = -1, 0, 1$. Si l'on définit une nouvelle variable $Y = X^2$, la fonction $p(x, y)$ du couple est donc :*

| $y \backslash x$ | -1 | 0 | 1 |
|---|---|---|---|
| 0 | 0 | $\frac{1}{3}$ | 0 |
| 1 | $\frac{1}{3}$ | 0 | $\frac{1}{3}$ |

*Il est facile de voir que*

$$\sum_i \sum_j x_i y_j p(x_i, y_j) = 0; \qquad \mathbb{E}(X) = 0 \qquad et \qquad \mathbb{E}(Y) = \frac{2}{3}.$$

*La covariance vaut donc $Cov(X, Y) = 0 - 0(\frac{2}{3}) = 0$ bien que ces deux variables ne soient pas indépendantes ; si l'événement $X = x$ est réaliser, alors l'événement $Y = x^2$ est forcément réalisé.*

$$\mathbb{P}_{(X,Y)}(\{(-1, 1)\}) = \frac{1}{3} \neq \mathbb{P}_X(\{-1\})\mathbb{P}_Y(\{1\}) = \frac{1}{3} \times \frac{2}{3} = \frac{2}{9}.$$

La proposition suivante donne quelques propriétés de la covariance.

### Proposition

*Soit $X$, $Y$ et $Z$ trois v. a. r. admettant un moment d'ordre 2. on a :*

1. $cov(X, X) = Var(X)$,
2. $cov(aX + bY, Z) = acov(X, Z) + bcov(Y, Z), \qquad \forall a, b \in \mathbb{R}.$
3. $Var(aX + bY) = a^2 Var(X) + 2abcov(X, Y) + b^2 Var(Y).$
4. $cov(X, Y)^2 \leq Var(X)Var(Y)$ *(Inégalité de Cauchy-Schwartz.)*

### Remarque

- *Plus généralement, si $X_1, ..., X_n$ sont $n$ variables aléatoires réelles définies sur un même espace probabilisé et $a_1, ..., a_n$ des réels :*

$$Var(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 Var(X_i) + 2 \sum_{1 \leq i < j \leq n} a_i a_j Cov(X_i, X_j).$$

- *En particulier, la variance d'une somme de variables aléatoires deux à deux non corrélées est la somme de leurs variances. (On rappelle qu'en général deux variables aléatoires réelles non-corrélées i.e. $\rho = 0$, ne sont pas indépendantes).*

**Preuve.**

1. Il suffit d'appliquer la définition de la variance.
2. C'est immédiat en utilisant la linéarité de l'espérance.
3. "
4. Il suffit de remarquer que $(X, Y) \longrightarrow Cov(X, Y)$ est une forme bilinéaire symétrique sur l'ensemble des v. a. r. de carrée intégrable.

# Coefficient de corrélation

### Définition

*Soit $(X, Y)$ un couple aléatoire tel que $X$ et $Y$ soient non constantes et admettent toutes les deux un moment d'ordre 2. On appelle coefficient de corrélation linéaire de $X$ et $Y$, et on note $\rho(X, Y)$, la quantité :*

$$\rho(X,Y) = \frac{cov(X,Y)}{\sigma(X)\sigma(Y)} = \mathbb{E}\left[\frac{(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))}{\sqrt{Var(X)}\sqrt{Var(Y)}}\right],$$

*où $\sigma(X)$ et $\sigma(Y)$ sont les écarts-types de $X$ et $Y$.*

### Définition

*Soit $(X, Y)$ un couple aléatoire défini sur $(\Omega, \mathcal{F}, \mathbb{P})$, du second ordre. On dit que $X$ et $Y$ sont corrélées si $Cov(X, Y) \neq 0$, non corrélées si $Cov(X, Y) = 0$.*

### Proposition

$-1 \leq \rho(X, Y) \leq +1.$

**Preuve.** Découle de l'inégalité de Cauchy-Schwartz.

### Remarque

1. *Le coefficient $\rho_{XY}$ est une mesure de l'association entre deux variables. On remarquera que si deux variables sont liées entre elles par une relation affine pour laquelle*

$$Y = aX + b \qquad a \neq 0$$

*alors $\rho_{XY} = +1$ lorsque $a > 0$ (relation linéaire positive), tandis que $\rho_{XY} = -1$ lorsque $a < 0$ (relation linéaire négative).*
*En effet, si $Y = aX + b$, alors $\sigma_Y^2 = a^2\sigma_X^2$, ce qui donne $\sigma_Y = a\sigma_X$, quand $a > 0$ et $\sigma_Y = -a\sigma_X$, quand $a < 0$.*
*On a donc, $\sigma_X\sigma_Y = |a|\sigma_X^2$.*

Plus généralement on a :

### Exercice

$$\rho_{XY} = \pm 1 \Longleftrightarrow Y = aX + b$$

Solution :
On pose $g(\alpha) = \mathbb{E}[(X - \mathbb{E}(X)) + \alpha(Y - \mathbb{E}(Y))]^2 \geq 0 \quad \forall \alpha.$
Mais
$$g(\alpha) = Var(X) + 2\alpha Cov(X, Y) + \alpha^2 Var(Y) \geq 0 \qquad \forall \alpha.$$
Donc
$$\Delta' = Cov(X, Y)^2 - Var(X)Var(Y)$$
Lorsque $\rho(X, Y) = \mp 1$, alors
$$\Delta' = 0.$$
Par suite
$$g(\alpha) = Var(Y)(\alpha - \alpha_0)^2.$$
Or,
$$g(\alpha_0) = \mathbb{E}[(X - \mathbb{E}(X)) + \alpha_0(Y - \mathbb{E}(Y))]^2 = 0$$
Donc
$$(X - \mathbb{E}(X)) + \alpha_0(Y - \mathbb{E}(Y)) = 0 \quad p.s.$$
$X$ et $Y$ sont liées par une relaion affine. La réciproque est Vraie.

Rappelons la définition des variables aléatoires gaussiennes réelles.

### Définition

*Une v. a. r. X est dite gaussienne centée réduite si elle admet pour densité :*

$$f(x) = \frac{1}{\sqrt{2\pi}} exp(-\frac{x^2}{2}), \quad x \in \mathbb{R}.$$

*Dans ce cas, on note* $X \sim \mathcal{N}(0,1)$.

Rappelons la définition des variables aléatoires gaussiennes réelles.

**Définition**

*Une v. a. r. $X$ est dite gaussienne centée réduite si elle admet pour densité :*

$$f(x) = \frac{1}{\sqrt{2\pi}} exp(-\frac{x^2}{2}), \quad x \in \mathbb{R}.$$

*Dans ce cas, on note $X \sim \mathcal{N}(0,1)$.*

**Définition**

*Une variable aléatoire réelle est dite gaussienne s'il existe $X \sim \mathcal{N}(0,1)$ et deux réels $\sigma$ et $m$ tels que $Y = \sigma X + m$.*

Identification des paramètres : on a

$$\mathbb{E}(Y) = m, \qquad Var(Y) = \sigma^2 Var(X) = \sigma^2.$$

Rappelons la définition des variables aléatoires gaussiennes réelles.

### Définition

*Une v. a. r. $X$ est dite gaussienne centée réduite si elle admet pour densité :*

$$f(x) = \frac{1}{\sqrt{2\pi}} exp(-\frac{x^2}{2}), \quad x \in \mathbb{R}.$$

*Dans ce cas, on note $X \sim \mathcal{N}(0,1)$.*

### Définition

*Une variable aléatoire réelle est dite gaussienne s'il existe $X \sim \mathcal{N}(0,1)$ et deux réels $\sigma$ et $m$ tels que $Y = \sigma X + m$.*

Identification des paramètres : on a

$$\mathbb{E}(Y) = m, \qquad Var(Y) = \sigma^2 Var(X) = \sigma^2.$$

### Remarque

*Dans le cas $\sigma = 0$ et $Y = m$ p.s., on parle de lois gaussiennes dégénérées. Dans ce cas, la v.a. $Y$ n'admet pas de densité.*

Une variable gaussienne est caractérisée par sa fonction caractéristique, donnée par le théorème suivant :

**Théorème**

*La fonction caractéristique : soit* $Y \sim \mathcal{N}(m, \sigma^2)$ *est*

$$\mathbb{E}[exp(itY)] = exp(itm - \frac{t^2}{2}\sigma^2), \qquad t \in \mathbb{R}.$$

Une variable gaussienne est caractérisée par sa fonction caractéristique, donnée par le théorème suivant :

### Théorème

*La fonction caractéristique : soit* $Y \sim \mathcal{N}(m, \sigma^2)$ *est*

$$\mathbb{E}[exp(itY)] = exp(itm - \frac{t^2}{2}\sigma^2), \qquad t \in \mathbb{R}.$$

**Preuve.** La fonction caractéristique de $X \sim \mathcal{N}(0,1)$ vérifie l'équation différentielle linéaire suivante :

$$\phi_X(t)' = -t\phi_X(t).$$

### Définition

*Soit* $X \in \mathbb{R}^n$. $X$ *est un vecteur aléatoire gaussien ssi pour tout* $\alpha \in \mathbb{R}^n$
$< \alpha | X > = {}^t\alpha . X = \sum_{i=1}^{n} \alpha_i . X_i$ *est une v.a. gaussienne réelle.*

### Définition

*Soit* $X \in \mathbb{R}^n$. *X est un vecteur aléatoire gaussien ssi pour tout* $\alpha \in \mathbb{R}^n$ $< \alpha | X >= {}^t\alpha.X = \sum_{i=1}^{n} \alpha_i.X_i$ *est une v.a. gaussienne réelle.*

### Remarque

1. $X$ *est un vecteur aléatoire gaussien ssi la v.a.r.* $\varphi(X) : \Omega \longrightarrow \mathbb{R}$ *est gaussienne pour toute forme linéaire* $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$.

### Définition

*Soit* $X \in \mathbb{R}^n$. *X est un vecteur aléatoire gaussien ssi pour tout* $\alpha \in \mathbb{R}^n$
$< \alpha | X >= {}^t\alpha.X = \sum_{i=1}^{n} \alpha_i.X_i$ *est une v.a. gaussienne réelle.*

### Remarque

1. *X est un vecteur aléatoire gaussien ssi la v.a.r.* $\varphi(X) : \Omega \longrightarrow \mathbb{R}$ *est gaussienne pour toute forme linéaire* $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$.
   *Il suffit de remarquer que la matrice de* $\varphi$ *dans les bases canoniques est de la forme* $(\alpha_1, ..., \alpha_n)$.
2. *X vecteur gaussien*
   $\Longrightarrow$ *Chaque composante* $X_i$ *de X est une v.a. rélle gaussienne.*

### Définition

*Soit $X \in \mathbb{R}^n$. $X$ est un vecteur aléatoire gaussien ssi pour tout $\alpha \in \mathbb{R}^n$*
*$< \alpha|X >= {}^t\alpha.X = \sum_{i=1}^n \alpha_i.X_i$ est une v.a. gaussienne réelle.*

### Remarque

1. *$X$ est un vecteur aléatoire gaussien ssi la v.a.r. $\varphi(X) : \Omega \longrightarrow \mathbb{R}$ est gaussienne pour toute forme linéaire $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$.*
   *Il suffit de remarquer que la matrice de $\varphi$ dans les bases canoniques est de la forme $(\alpha_1, ..., \alpha_n)$.*
2. *$X$ vecteur gaussien*
   *$\Longrightarrow$ Chaque composante $X_i$ de $X$ est une v.a. rélle gaussienne.Puisque les applications $(x_1, ..., x_n) \longrightarrow x_k$ sont des forme linéaires sur $\mathbb{R}^n$.*

### Définition

*Soit $X \in \mathbb{R}^n$. $X$ est un vecteur aléatoire gaussien ssi pour tout $\alpha \in \mathbb{R}^n$ $< \alpha|X >= {}^t\alpha.X = \sum_{i=1}^{n} \alpha_i.X_i$ est une v.a. gaussienne réelle.*

### Remarque

1. *$X$ est un vecteur aléatoire gaussien ssi la v.a.r. $\varphi(X) : \Omega \longrightarrow \mathbb{R}$ est gaussienne pour toute forme linéaire $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$.*
   *Il suffit de remarquer que la matrice de $\varphi$ dans les bases canoniques est de la forme $(\alpha_1, ..., \alpha_n)$.*
2. *$X$ vecteur gaussien*
   *$\Longrightarrow$ Chaque composante $X_i$ de $X$ est une v.a. rélle gaussienne.Puisque les applications $(x_1, ..., x_n) \longrightarrow x_k$ sont des forme linéaires sur $\mathbb{R}^n$.*
3. *L'inverse est faux. On peut construire un exemple de $X = (X_1, X_2) \in \mathbb{R}^2$ tel que :*
   - i) *$X_1$, $X_2$ gaussiennes réelles*
   - ii) *$X$ n'est pas un vecteur gaussien.*

### Définition

*Soit $X \in \mathbb{R}^n$. $X$ est un vecteur aléatoire gaussien ssi pour tout* $\alpha \in \mathbb{R}^n$
$< \alpha | X > = {}^t\alpha.X = \sum_{i=1}^{n} \alpha_i.X_i$ *est une v.a. gaussienne réelle.*

### Remarque

1. *$X$ est un vecteur aléatoire gaussien ssi la v.a.r. $\varphi(X) : \Omega \longrightarrow \mathbb{R}$ est gaussienne pour toute forme linéaire $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$.*
   *Il suffit de remarquer que la matrice de $\varphi$ dans les bases canoniques est de la forme $(\alpha_1, ..., \alpha_n)$.*
2. *$X$ vecteur gaussien*
   *$\Longrightarrow$ Chaque composante $X_i$ de $X$ est une v.a. rélle gaussienne.Puisque les applications $(x_1, ..., x_n) \longrightarrow x_k$ sont des forme linéaires sur $\mathbb{R}^n$.*
3. *L'inverse est faux. On peut construire un exemple de $X = (X_1, X_2) \in \mathbb{R}^2$ tel que :*
   i) *$X_1$, $X_2$ gaussiennes réelles*
   ii) *$X$ n'est pas un vecteur gaussien.*

### Exercice (**Vecteur non gaussien à marginales gaussiennes**)

*Soit $X$ une variable aléatoire de loi $\mathcal{N}(0,1)$ et $T$ une v. a. indépendante de $X$ telle que $\mathbb{P}[T = 1] = \mathbb{P}[T = -1] = 1/2$. Montrer que $Y = TX$ suit une loi $\mathcal{N}(0,1)$ et que le vecteur aléatoire $(X, Y)$ n'est pas gaussien.*

### Exercice

*Montrer que si* $X \in \mathbb{R}^n$ *est un vecteur gaussien et si* $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *est une application linéaire, alors* $AX : \Omega \longrightarrow \mathbb{R}^m$ *est encore un vecteur gaussien.*

### Exercice

*Montrer que si $X \in \mathbb{R}^n$ est un vecteur gaussien et si $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ est une application linéaire, alors $AX : \Omega \longrightarrow \mathbb{R}^m$ est encore un vecteur gaussien.*

Il suffit de remarquer que $\psi \circ A$ est une forme linéaire sur $\mathbb{R}^n$, pour toute f.l. $\psi$ sur $\mathbb{R}^m$.

# Exercice : exemple fondamental :

Considérons n variables aléatoires $X_1$, ..., $X_n$, indépendantes et de loi respectivement $N(m_1, \sigma_1^2)$,...., $N(m_n, \sigma_n^2)$.
Pour $i = 1, ..., n$, la variable aléatoire $X_i$ est donc de densité

$$f(x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} exp - \frac{1}{2} \frac{(x - m_i)^2}{\sigma_i^2}$$

1. Montrer que l'on peut écrire la densité conjointe du vecteur $(X_1, ..., X_n)$ de la forme :

$$f_X(x_1, ..., x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} det(\Gamma_X)^{\frac{1}{2}}} e^{-\frac{1}{2} {}^t(x-m).\Gamma_X^{-1}.(x-m)},$$

avec $m$ est le vecteur espérance du vecteur $X = (X_1, ..., X_n)$ et $\Gamma_X$ est sa matrice de covariance :

$$m = \mathbb{E}(X) = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{pmatrix} \quad et \quad \Gamma_X = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}.$$

2. On rappelle que la fonction caractéristique (f.c.) d'une v.a.r. $X_j$ de loi $N(m_j, \sigma_j^2)$ est

$$\mathbb{E}[exp(itX_j)] = exp(itm_j - \frac{t^2}{2}\sigma_j^2), \qquad t \in \mathbb{R}.$$

En déduire que la fonction caractéristique du vecteur $X = (X_1, ..., X_n)$ vaut

$$\phi_X(u) = exp(i\,{}^t u.\mathbb{E}(X) - \frac{1}{2}\,{}^t u\Gamma_X u) = \phi_{<u,X>}(1),\ \forall u \in \mathbb{R}^n,$$

où $\Gamma_X$ est la matrice de covariance de $X$.

# Solution de l'exercice :

Considérons n variables aléatoires $X_1$, ..., $X_n$, indépendantes et de loi respectivement $N(m_1, \sigma_1^2)$,...., $N(m_n, \sigma_n^2)$.
Pour $i = 1, ..., n$, la variable aléatoire $X_i$ est donc de densité

$$f(x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} exp - \frac{1}{2}\frac{(x - m_i)^2}{\sigma_i^2}$$

par rapport à la mesure de Lebesgue sur $\mathbb{R}$.
En raison de l'indépendance des variables aléatoires $X_i$, la densité conjointe du vecteur $X_1$, ..., $X_n$ est :

$$f_{(X_1,\ldots,X_n)}(x_1, .., x_n) = \frac{1}{(\sqrt{2\pi})^n} . \frac{1}{\sqrt{\prod_{i=1}^n \sigma_i^2}} exp - \frac{1}{2}\sum_{i=1}^{n}\frac{(x - m_i)^2}{\sigma_i^2}.$$

le vecteur espérance du vecteur $X = (X_1, ..., X_n)$ et sa matrice de covariance sont :

$$\mathbb{E}(X) = m = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{pmatrix} \quad et \quad \Gamma_X = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}.$$

Notons que la matrice $\Gamma_X$ est diagonale en raison de l'indṕendance des v.a.r. $(X_i)_{i=1,\ldots,n}$.
Comme toutes les variances $\sigma_i^2$ sont strictement positives, on obtient aisément la matrice inverse :

$$\Gamma_X^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_n^2} \end{pmatrix}.$$

On peut alors réécrire la densité conjointe du vecteur $X = (X_1, ..., X_n)$ sous la forme

$$f_X(x_1, ..., x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} det(\Gamma_X)^{\frac{1}{2}}} e^{-\frac{1}{2}\, {}^t(x-m).\Gamma_X^{-1}.(x-m)},$$

puisque

$$\begin{aligned} {}^t(x-m).\Gamma_X^{-1}.(x-m) &= (x_1 - m_1, ..., x_n - m_n). \begin{pmatrix} \frac{1}{\sigma_1^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_n^2} \end{pmatrix} . \begin{pmatrix} x_1 - m_1 \\ \vdots \\ x_n - m_n \end{pmatrix}. \\ &= \sum_{i=1}^n \frac{(x_i - m_i)^2}{\sigma_i^2}. \end{aligned}$$

2. En raison de l'indépendance entre les composantes du vecteur $X$, on a, pour tout $u \in \mathbb{R}^n$

$$\varphi_X(u) = \varphi_{(X_1,\dots,X_n)}(u_1, \dots, u_n) = \prod_{j=1}^{n} \varphi_{X_j}(u_j)$$

d'où on tire

$$\varphi_X(u) = e^{i \sum_{j=1}^{n} u_j . m_j - \frac{1}{2} \sum_{j=1}^{n} u_j^2 . \sigma_j^2} = e^{i\, {}^t u . m - \frac{1}{2} {}^t u \Gamma_X . u}$$

# Propriétés des vecteurs gaussiens

**Théorème**

*Si $X$ est un vecteur gaussien, sa fonction caractéristique vaut*

$$\phi_X(u) = exp(i\,{}^t u.\mathbb{E}(X) - \frac{1}{2}\,{}^t u K_X u) = \phi_{<u,X>}(1),\ \forall u \in \mathbb{R}^n,$$

*où $K_X$ est la matrice de covariance de $X$.*

# Propriétés des vecteurs gaussiens

**Théorème**

*Si $X$ est un vecteur gaussien, sa fonction caractéristique vaut*

$$\phi_X(u) = exp(i\,{}^t u.\mathbb{E}(X) - \frac{1}{2}\,{}^t u K_X u) = \phi_{<u,X>}(1),\ \forall u \in \mathbb{R}^n,$$

*où $K_X$ est la matrice de covariance de $X$.*

**Preuve.**
L'application $x \longrightarrow^t u.x =< u|x >$ est une forme linéaire. La v. a. r. ${}^t u.X$ est donc gaussienne. Son espérance est $m = \mathbb{E}({}^t u.X) =^t u.\mathbb{E}(X)$, et sa variance est $\sigma^2 = Var({}^t u.X) =^t uK_X u$.

# Propriétés des vecteurs gaussiens

**Théorème**

*Si $X$ est un vecteur gaussien, sa fonction caractéristique vaut*

$$\phi_X(u) = exp(i\,{}^t u.\mathbb{E}(X) - \frac{1}{2}\,{}^t u K_X u) = \phi_{<u,X>}(1),\ \forall u \in \mathbb{R}^n,$$

*où $K_X$ est la matrice de covariance de $X$.*

**Preuve.**
L'application $x \longrightarrow {}^t u.x =< u|x >$ est une forme linéaire. La v. a. r. ${}^t u.X$ est donc gaussienne. Son espérance est $m = \mathbb{E}({}^t u.X) = {}^t u.\mathbb{E}(X)$, et sa variance est $\sigma^2 = Var({}^t u.X) = {}^t u K_X u$.

**Exercice**

*Vérifier que*

$$Var({}^t u.X) = {}^t u K_X u.$$

**Preuve.**

1. On a

$$ {}^t u.K_X.u = {}^t u.\mathbb{E}[(X-\mathbb{E}(X)).{}^t(X-\mathbb{E}(X))].u = Var < u|X >= Var({}^t u.X) \geq 0, $$

puisque

$$ {}^t u.[X-\mathbb{E}(X)] =< u|X-\mathbb{E}(X) >=< X-\mathbb{E}(X)|u >= {}^t(X-\mathbb{E}(X))).u, $$

et

$$ \begin{aligned} Var({}^t u.X) &= \mathbb{E}({}^t u.[X-\mathbb{E}(X)])^2 = \mathbb{E}({}^t u.[X-\mathbb{E}(X)].{}^t u.[X-\mathbb{E}(X)]) \\ &= \mathbb{E}({}^t u.[X-\mathbb{E}(X)].{}^t[X-\mathbb{E}(X)].u) \\ &=^t u.\underbrace{\mathbb{E}([X-\mathbb{E}(X)].{}^t[X-\mathbb{E}(X)])}_{K_X}.u =^t u.K_X.u \end{aligned} $$

# Remarque et notation

1. Théorème précédent
   $\hookrightarrow$ La loi d'un vecteur gaussien $X$ est caractérisée par sa moyenne $m$ et sa matrice de covariance $K_X$.
   $\hookrightarrow$ si $X$ et $Y$ sont deux vecteurs gaussiens, ayant même moyenne et même matrice de covariance, ils ont même loi.

2. Attention : Cette propriété concerne les vecteurs gaussiens, elle n'est pas vraie en général, et il n'y a aucune raison pour que deux variables aléatoires à valeurs réelles qui ont même moyenne et même variance aient même loi.

3. Notation : Si $X$ vecteur gaussien de moyenne $m$ et covariance $K_X$
   On écrit $X \sim \mathcal{N}(m, K_X)$.

**Exercice (Transformations affines)**

*Soient* $X \sim \mathcal{N}(m_X, K_X)$ , $A \in \mathbb{R}^{p,n}$ *et* $z \in \mathbb{R}^p$*. On pose* $Y = AX + z$*. Montrer que*

$$Y \sim \mathcal{N}(m_Y, K_Y), \quad \text{avec } m_Y = Am_X + z, \qquad K_Y = Ak_X\, {}^tA.$$

### Exercice (Transformations affines)

*Soient $X \sim \mathcal{N}(m_X, K_X)$ , $A \in \mathbb{R}^{p,n}$ et $z \in \mathbb{R}^p$. On pose $Y = AX + z$. Montrer que*

$$Y \sim \mathcal{N}(m_Y, K_Y), \quad \text{avec } m_Y = Am_X + z, \qquad K_Y = Ak_X\, {}^tA.$$

Solution :
Soit $\alpha \in \mathbb{R}^p$.
On veut montrer que ${}^t\alpha.Y$ est gaussienne. Or

$${}^t\alpha Y = {}^t\alpha z + {}^t\alpha AX = {}^t\alpha z + {}^tuX.$$

$u = {}^tA\alpha.$
$Y$ est donc bien un vecteur gaussien. De plus,

$$m_Y = z + Am_X, \qquad K_Y = \mathbb{E}(Y - \mathbb{E}Y)\,{}^t(Y - \mathbb{E}(Y)) = AK_X\,{}^tA.$$

# Positivité de la matrice de corrélation

**Proposition**

*Soit $X$ vecteur aléatoire de matrice de covariance $K$. Alors $K$ est une matrice symétrique et positive*

**Preuve.**

1. Symétrie :
$$K(i,j) = Cov(X_i, X_j) = Cov(X_j, X_i) = K(j,i)$$

2. Positivité : Soit $u \in \mathbb{R}^n$ et $Y = {}^t u.X$. Alors $< u, K.u > \geq 0$. En effet :
$$Var(Y) = {}^t u.K.u \geq 0.$$

# Construction de vecteur gaussien

**Lemme (Lemme d'Algèbre)**

*Soit* $K \in \mathbb{R}^{n,n}$ *symétrique et positive. Alors il existe une matrice* $A \in \mathbb{R}^{n,n}$ *telle que* $K = A\,A^t$.

**Théorème**

*Soient* $m \in \mathbb{R}^n$ *et* $K \in \mathbb{R}^{n,n}$ *symétrique positive d'ordre n. Alors il existe un vecteur gaussien* $X \sim \mathcal{N}(m, K)$.

# Construction de vecteur gaussien

**Lemme (Lemme d'Algèbre)**

*Soit $K \in \mathbb{R}^{n,n}$ symétrique et positive. Alors il existe une matrice $A \in \mathbb{R}^{n,n}$ telle que $K = A\,A^t$.*

**Théorème**

*Soient $m \in \mathbb{R}^n$ et $K \in \mathbb{R}^{n,n}$ symétrique positive d'ordre n. Alors il existe un vecteur gaussien $X \sim \mathcal{N}(m, K)$.*

**Preuve.**
Soient $Y_1, Y_2, ..., Y_n$, i.i.d de loi $\mathcal{N}(0, 1)$. On pose ${}^tY = (Y_1, ..., Y_n) \Longrightarrow Y \sim \mathcal{N}(0, Id_n)$.
Soit $A \in \mathbb{R}^{n,n}$ telle que $A\,{}^tA = K$
On définit $X$ par :

$$X = m + AY.$$

Alors $X \sim \mathcal{N}(m, K_X)$, avec

$$K_X = \mathbb{E}[(X - m).\,{}^t(X - m)] = AK_Y\,{}^tA = AId_n\,{}^tA = A\,{}^tA = K.$$

### Théorème

*Si $X = (X_1, ..., X_n)$ est un vecteur gaussien, alors ses composantes sont indépendantes si et seulement si elles sont non corrélées :*

$$cov(X_j, X_k) = 0, \ \forall j \neq k.$$

*Autrement dit, la matrice de covariance $K_X$ est diagonale :*

$$K_X = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$$

Autrement dit, si et seulement si $X_1 - \mathbb{E}(X_1), \ldots, X_n - \mathbb{E}(X_n)$ sont orthogonales dans $\mathbb{L}^2(\mathbb{P})$.

**Preuve.** Nous savons que des v.a.r. indépendantes sont non corrélées. Inversement, si elles ne sont pas corrélées, la matrice de covariance de $X$ est diagonale :

$$K_X = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$$

avec $\sigma_k^2 = var(X_k)$. Notons $m_k = \mathbb{E}(X_k)$, on a $\mathbb{E}(X) = (m_1, ..., m_n)$. Comme $X$ est un vecteur gaussien, on a pour tout $u = (u_1, ..., u_n) \in \mathbb{R}^n$ :

$$\begin{aligned} \phi_X(u) &= exp(i < u|\mathbb{E}(X) > -\tfrac{1}{2}\, {}^t u K_X u) \\ &= exp(i \textstyle\sum_{j=1}^n u_j m_j - \tfrac{1}{2} \sum_{j=1}^n u_j^2 \sigma_j^2) \\ &= exp(iu_1 m_1 - \tfrac{1}{2} u_1^2 \sigma_1^2) \times .... \times exp(iu_n m_n - \tfrac{1}{2} u_n^2 \sigma_n^2) \\ &= \phi_{X_1}(u_1)...\phi_{X_n}(u_n). \end{aligned}$$

Alors,

$$\begin{aligned} \phi_X(u) &= \int_{\mathbb{R}^n} e^{i \sum_{j=1}^n u_j x_j} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) dx_1 \ldots dx_n \\ &= \phi_{X_1}(u_1)\ldots\phi_{X_n}(u_n) = \int_{\mathbb{R}^n} e^{i \sum_{j=1}^n u_j x_j} f_{X_1}(x_1)\ldots f_{X_n}(x_n) dx_1 \ldots dx_n \end{aligned}$$

Donc la loi de densité $f_{(X_1,\ldots,X_n)}$ a la même fonction caractéristique que loi de densité $f_{X_1} \ldots f_{X_n}$, et puisque la fonction caractéristique caractérise la loi, la densité conjointe est le produit des densités marginales, Ce qui prouve l'indépendance de $X_1, \ldots, X_n$.

### Remarque

*Il est indispensable de savoir au préalable que le vecteur $X$ est gaussien.*

### Corollaire

*Si le couple $(X, Y)$ est un vecteur gaussien, on a*

$$X \coprod Y \Longleftrightarrow Cov(X, Y) = 0.$$

**Preuve.** Immédiate.

### Remarque

*Nous attirons votre attention sur le fait que deux variables aléatoires réelles et non corrélées ne sont pas nécessairement indépendantes. Pour s'assurer qu'elles le soient il faut pour cela qu'elles constituent un couple gaussien.*

## Exercice

*Soit* $X$ *une v. a. de loi* $N(0,1)$ *et* $Y$ *une v. a. telle que* $\mathbb{P}(Y=1)=\mathbb{P}(Y=-1)=\frac{1}{2}$, *indépendante de* $X$. *Montrer que*

1. $XY \sim N(0,1)$.
2. $Cov(X, XY) = 0$.
3. *Le vecteur* $(X, XY)$ *n'est pas gaussien.*

## Exercice

*Soit $X$ une v. a. de loi $N(0,1)$ et $Y$ une v. a. telle que $\mathbb{P}(Y=1)=\mathbb{P}(Y=-1)=\frac{1}{2}$, indépendante de $X$. Montrer que*

1. $XY \sim N(0,1)$.
2. $Cov(X, XY) = 0$.
3. *Le vecteur $(X, XY)$ n'est pas gaussien.*

Solution :

1. Pour tout borélien $A$ de $\mathbb{R}$, on a

$$\mathbb{P}(XY \in A) = \mathbb{P}(Y=1, X \in A) + \mathbb{P}(Y=-1, -X \in A)$$

l'indépendance donne

$$\mathbb{P}(Y=1, X \in A) = \mathbb{P}(Y=1)\mathbb{P}(X \in A) = \frac{1}{2}\mathbb{P}(X \in A), \;\; et \; \mathbb{P}(Y=-1, -X \in A) = \frac{1}{2}\mathbb{P}(X \in A),$$

puisque $-X$ suit la même loi que $X$ ($\phi_X(t) = \phi_{-X}(t) = e^{\frac{-t^2}{2}} \; \forall t$). Donc $YX$ suit la même loi que $X$.

2. $cov(X, YX) = \mathbb{E}(X.XY) = \mathbb{E}(X^2 Y) \stackrel{\perp\!\!\!\perp}{=} \mathbb{E}(X^2)\mathbb{E}(Y) = 1 \times 0 = 0.$
3. Pourtant, les v.a.r. $X$ et $YX$ ne sont pas indépendantes (car sinon $X$ serait indépendante de $(YX)^2 = X^2$, ce qui n'est pas le cas). Il en résulte que le vecteur $(X, YX)$ n'est pas gaussien, bien que ses composantes soient des v.a.r. gaussiennes.

### Exercice

*Soit* $(X, Y)$ *un vecteur aléatoire gaussien dans* $\mathbb{R}^2$*, centré de matrice de convariance* $I_2$*. Soit* $(Z, Q)$ *le vecteur aléatoire donné par*

$$Z = \frac{X+Y}{2} \qquad \text{et} \qquad Q = \frac{X-Y}{2}.$$

*On pose*

$$U = \frac{1}{2}(X-Z)^2 + \frac{1}{2}(Y-Z)^2.$$

1. *Calculer la matrice de covariance du couple* $(Z, Q)$*.*
2. $Z$ *et* $Q$ *sont-elles indépendantes ?*
3. *Montrer que* $Z$ *et* $U$ *sont indépendantes.*

Démonstration :

1. On a $Cov(Z, Q) = 0$ , $Var(Z) = \frac{Var(X)+Var(Y)}{4} = \frac{1}{2}$ et aussi $Var(Q) = \frac{1}{2}$ . Donc la matrice de covariance est
$$\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$
2. Le vecteur $(Z, Q)$ est donné par une transformation linéaire de $(X, Y)$. Il est donc gaussien et étant $Cov(Z, Q) = 0$ les v.a. $Z$ et $Q$ sont indépendantes.
3. On a $U$ est fonction de $Q$ (car : $U = Q^2$) et $Q$ indépendant de $Z$. Donc $U$ et $Z$ sont indépendantes.