

# Probabilités II



STEP, MINES ParisTech\*

5 novembre 2025 (#994c40b)

## Table des matières

<b>Objectifs d'apprentissage</b>	<b>3</b>
<b>Moments d'une variable aléatoire</b>	<b>4</b>
Espérance d'une variable aléatoire réelle . . . . .	4
Interprétation . . . . .	4
Propriétés de base . . . . .	4
Rappel : cas discret . . . . .	5
Variance . . . . .	5
Propriétés . . . . .	5
La variance est positive . . . . .	6
Variable centrée réduite . . . . .	6
Covariance . . . . .	6
Inégalité de Cauchy-Schwarz . . . . .	7
Espérance de $g(X)$ . . . . .	7
Cas particulier . . . . .	8
Inégalité de Jensen . . . . .	8
Exemples de loi à densité . . . . .	9
<i>Loi uniforme</i> . . . . .	9
<i>Loi exponentielle</i> . . . . .	9
<i>Loi gamma</i> . . . . .	10
<i>Loi normale</i> . . . . .	10
Estimation statistique . . . . .	11
<b>Rappels sur les intégrales multiples</b>	<b>11</b>
Théorème de Fubini (pour les probabilités) . . . . .	11
Changement de variables . . . . .	12
<b>Vecteurs aléatoires</b>	<b>13</b>
Définitions . . . . .	13

---

\*Ce document est un des produits du projet  paulinebernard/CDIS issu de la collaboration de (P)auline Bernard (CAS) et (T)homas Romary (GEOSCIENCES). Il dérive du projet  boisgera/CDIS, initié par la collaboration de (S)ébastien Boisgérault (CAOR), (T)homas Romary et (E)milie Chautru (GEOSCIENCES), (P)auline Bernard (CAS), avec la contribution de Gabriel Stoltz (Ecole des Ponts ParisTech, CERMICS). Il est mis à disposition selon les termes de la licence Creative Commons “attribution – pas d'utilisation commerciale – partage dans les mêmes conditions” 4.0 internationale.

Vecteur aléatoire à densité . . . . .	14
Fonction d'un vecteur aléatoire . . . . .	14
Densités marginales . . . . .	14
Démonstration . . . . .	15
Réciproque ? . . . . .	15
Moments d'un vecteur aléatoire . . . . .	15
Vecteur espérance et covariance . . . . .	15
Matrice de covariance . . . . .	15
<b>Variables aléatoires indépendantes</b>	<b>16</b>
Vecteurs aléatoires indépendants . . . . .	17
Caractérisation de l'indépendance – cas à densité . . . . .	17
Fonctions de vecteurs aléatoires indépendants . . . . .	18
Cas général . . . . .	18
Covariance de variables aléatoires indépendantes . . . . .	18
Remarque . . . . .	18
<b>Identification de densité</b>	<b>18</b>
Méthode de la fonction muette . . . . .	19
<b>Modélisation statistique</b>	<b>21</b>
Divergence de Kullback-Leibler . . . . .	22
Remarque . . . . .	22
Positivité de la divergence de Kullback-Leibler . . . . .	22
<b>Exercices complémentaires</b>	<b>23</b>
Durée de vie . . . . .	23
Crues centennales de la Seine . . . . .	24
Indépendance et vecteurs gaussiens . . . . .	24
Symétrie de la gaussienne . . . . .	25
Loi du $\chi^2$ . . . . .	25
Combinaisons linéaires de variables Gaussiennes indépendantes . . . . .	25
Loi de Maxwell . . . . .	26
Partie 1 . . . . .	26
Partie 2 . . . . .	27
<b>Solutions</b>	<b>27</b>
Exercices essentiels . . . . .	27
Moments d'une variable aléatoire de loi uniforme sur $[a, b]$ . . . . .	28
Durée de vie . . . . .	34
Crues centennales de la Seine . . . . .	34
Indépendance et vecteurs gaussiens . . . . .	35
Symétrie de la gaussienne . . . . .	37
Loi du $\chi^2$ . . . . .	39
Combinaisons linéaires de variables Gaussiennes indépendantes . . . . .	40
Loi de Maxwell . . . . .	42
\$	\$

# Objectifs d'apprentissage

## Moments d'une variable aléatoire

- • connaître la définition de l'espérance d'une variable aléatoire
- • connaître l'interprétation de l'espérance
- • connaître les propriétés de base de l'espérance
- • connaître la définition de la variance d'une variable aléatoire
- • savoir que la variance est toujours positive
- • savoir ce que signifie variable centrée-réduite
- • connaître la définition de la covariance
- • savoir que la covariance est une forme bilinéaire
- • connaître et savoir manipuler l'inégalité de Cauchy-Schwarz
- •• connaître les conditions d'existence de l'espérance d'une fonction (mesurable) d'une variable aléatoire
- •• savoir exprimer la probabilité d'un événement de la forme  $X \in A$  sous la forme d'une espérance
- •• connaître et savoir calculer les moments des lois usuelles

## Vecteurs aléatoires

- • savoir que les composantes d'un vecteur aléatoire sont des variables aléatoires définies sur le même espace de probabilité
- • connaître la définition d'un vecteur aléatoire à densité
- •• connaître les conditions d'existence de l'espérance d'une fonction (mesurable) d'un vecteur aléatoire à densité
- •• savoir calculer les densités marginales d'un vecteur aléatoire à densité
- •• savoir qu'un couple de variables aléatoires à densité n'admet pas nécessairement de densité
- • connaître la définition de l'espérance d'un vecteur aléatoire à densité
- • connaître la définition de la matrice de covariance d'un vecteur aléatoire à densité
- • savoir que la matrice de covariance est semi-définie positive

## Variables et vecteurs aléatoires indépendantes

- • connaître la définition de l'indépendance de deux vecteurs/variables aléatoires
- • savoir caractériser l'indépendance de deux vecteurs/variables aléatoires à densité
- •• savoir que si  $X$  et  $Y$  sont indépendantes, alors  $g(X)$  et  $h(Y)$  sont aussi indépendantes
- •• savoir que si  $X$  et  $Y$  sont indépendantes et si  $g(X)$  et  $h(Y)$  sont intégrables alors  $\mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X))\mathbb{E}(h(Y))$
- • savoir que la covariance de deux variables indépendantes est nulle
- •• savoir que la réciproque est fausse et connaître un contre-exemple

## Identification de densité

- •• savoir identifier la densité d'une fonction d'une variable aléatoire par la méthode de la fonction muette
- •• savoir identifier la densité d'une fonction d'un vecteur aléatoire par la méthode de la fonction muette

## Moments d'une variable aléatoire

La loi de probabilité d'une variable aléatoire définie sur l'espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$  va nous permettre de calculer aisément des grandeurs caractéristiques telles que sa valeur moyenne et sa variance (lorsqu'elles existent). Elles sont définies ci-dessous.

### Définition – Espérance d'une variable aléatoire réelle

La variable aléatoire  $X : \Omega \rightarrow \mathbb{R}$  de loi  $\mathbb{P}_X$  est dite *intégrable* si l'intégrale  $\int_{\mathbb{R}} |x| d\mathbb{P}_X(x)$  est définie, autrement dit si  $x$  est  $\mathbb{P}_X$ -intégrable. On définit alors son *espérance* par

$$\mathbb{E}(X) = \int_{\mathbb{R}} x d\mathbb{P}_X(x) = \int_{\Omega} X(\omega) d\mathbb{P}(\omega)$$

Si la variable aléatoire  $X$  admet une densité  $f$  par rapport à la mesure de Lebesgue, elle sera *intégrable* si l'intégrale  $\int_{\mathbb{R}} |x| f(x) dx$  est définie, autrement dit si le produit  $xf(x)$  est intégrable par rapport à la mesure de Lebesgue. Son *espérance* vaut alors

$$\mathbb{E}(X) = \int_{\mathbb{R}} xf(x) dx$$

### Remarque – Interprétation

$\mathbb{E}(X)$  est un nombre réel qui donne une valeur moyenne résumant la variable aléatoire  $X$ .

L'espérance mathématique d'une variable aléatoire est un concept fondamental de la théorie des probabilités. La dénomination d'espérance pour cette quantité fait référence aux problèmes de jeux et d'espérance de gain. Cette terminologie imagée a été introduite par Pascal.

On note  $\mathcal{L}^1$  l'ensemble de toutes les variables réelles  $X$  intégrables. Les propriétés suivantes découlent directement des propriétés de l'intégrale.

### Proposition – Propriétés de base

—  $\mathcal{L}^1$  est un espace vectoriel et  $\forall X, Y \in \mathcal{L}^1, \forall a, b \in \mathbb{R}$

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y).$$

—  $X \in \mathcal{L}^1 \Leftrightarrow |X| \in \mathcal{L}^1$ , et dans ce cas

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|).$$

— Si  $X \geq 0$  et  $X \in \mathcal{L}^1$ , alors  $\mathbb{E}(X) \geq 0$ .

— Si  $X, Y \in \mathcal{L}^1$  sont telles que  $X \leq Y$ , alors

$$\mathbb{E}(X) \leq \mathbb{E}(Y).$$

— L'espérance d'une variable presque-sûrement constante est égale à cette constante :

$$\text{Si } \mathbb{P}(X(\omega) = a) = 1, \text{ alors } \mathbb{E}(X) = a.$$

— Si  $\exists b \in \mathbb{R}_+$  tel que  $|X| \leq b$ , alors  $X \in \mathcal{L}^1$  et  $\mathbb{E}(X) \leq b$ .

### Exercice – Démontrer ces propriétés (•)

#### Remarque – Rappel : cas discret

Dans le cas d'une variable aléatoire discrète  $X$  à valeurs dans  $\mathbb{N}^*$ , son espérance est définie par la quantité  $\mathbb{E}(X) = \sum_{i \in \mathbb{N}^*} i\mathbb{P}(X = i)$ , pourvu que celle-ci soit finie. On voit immédiatement que les propriétés ci-dessus sont également vérifiées.

Outre l'espace  $\mathcal{L}^1$ , nous pouvons définir l'espace  $\mathcal{L}^2$  des variables aléatoires réelles dont le carré  $X^2$  est dans  $\mathcal{L}^1$ .

#### Définition – Variance

La variable aléatoire  $X : \Omega \rightarrow \mathbb{R}$  de loi  $\mathbb{P}_X$  est dite *de carré intégrable* si  $\mathbb{E}(X^2) = \int_{\mathbb{R}} x^2 \mathbb{P}_X(dx) = \int_{\Omega} X^2(\omega) \mathbb{P}(d\omega) < +\infty$ . Sa *variance* est définie par

$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$$

Si la variable aléatoire  $X$  admet une densité  $f$  par rapport à la mesure de Lebesgue, elle sera *de carré intégrable* si  $\mathbb{E}(X^2) = \int_{\mathbb{R}} x^2 f(x) dx$  est définie, autrement dit si le produit  $x^2 f(x)$  est intégrable par rapport à la mesure de Lebesgue.

#### Proposition – Propriétés

$\mathcal{L}^2$  est un sous-espace vectoriel de  $\mathcal{L}^1$ , et si  $X \in \mathcal{L}^2$ ,

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|) \leq \sqrt{\mathbb{E}(X^2)}$$

**Démonstration** Soient  $X$  et  $Y$  deux variables aléatoires réelles de  $\mathcal{L}^2$  et  $a \in \mathbb{R}$ . Comme  $(aX + Y)^2 \leq 2a^2 X^2 + 2Y^2$ , alors  $aX + Y \in \mathcal{L}^2$ . Ainsi,  $\mathcal{L}^2$  est un espace vectoriel.

L'inclusion  $\mathcal{L}^2 \subset \mathcal{L}^1$  découle de  $|X| \leq 1 + X^2$  et de la proposition précédente (p. 4) (linéarité).

La première inégalité a déjà été vue ci-dessus (p. 4). Pour la seconde, nous pouvons nous limiter au cas où  $X$  est positive. Soit alors  $a = \mathbb{E}(X)$  et  $Y = X - a$ . Par linéarité, on a

$$\mathbb{E}(Y^2) = \mathbb{E}(X^2) - 2a\mathbb{E}(X) + a^2 = \mathbb{E}(X^2) - a^2.$$

Et  $\mathbb{E}(Y^2) \geq 0$  par le troisième point de la proposition ci-dessus (p. 4). Par conséquent,  $\mathbb{E}(X)^2 = a^2 \leq \mathbb{E}(X^2)$  ce qui est le résultat recherché. ■

**Exercice – Approximation par une constante (•)** Soit  $X$  une v.a. de carré intégrable. Montrer que  $\mathbb{E}((X - a)^2)$  atteint son minimum lorsque  $a = \mathbb{E}(X)$ . (Solution p. 27.)

### Remarque – La variance est positive

En vertu de cette proposition (p. 5),  $\mathbb{V}(X)$  est **positive** et sa racine carrée  $\sigma_X$  s'appelle l'*écart-type* de  $X$ . L'écart-type est une grandeur qui mesure la moyenne (en un certain sens) de l'écart des valeurs de  $X$  à sa moyenne, d'où son nom.

On a également

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

que l'on obtient en développant  $(X - \mathbb{E}(X))^2$ . Cette manipulation anodine est fort utile dans la pratique. On retiendra que "La variance est égale à la moyenne des carrés moins le carré de la moyenne". On désigne le terme  $\mathbb{E}(X^2)$  par l'expression *moment d'ordre deux* tandis que la variance est parfois appelée *moment centré d'ordre deux*.

### Remarque – Variable centrée réduite

D'après ce qui précède, si  $X$  est une variable aléatoire de carré intégrable, d'espérance  $\mathbb{E}(X)$  et d'écart-type  $\sigma_X > 0$ , alors la variable aléatoire

$$\frac{X - \mathbb{E}(X)}{\sigma_X}$$

est d'espérance nulle et de variance 1. On dira qu'une telle variable aléatoire est *centrée et réduite*.

On peut remarquer que si  $X$  et  $Y$  sont dans  $\mathcal{L}^2$ , la variable aléatoire  $XY$  est dans  $\mathcal{L}^1$ , puisque  $|XY| \leq \frac{1}{2}(X^2 + Y^2)$ . On peut ainsi définir la *covariance* de deux variables aléatoires :

### Définition – Covariance

Si  $X$  et  $Y$  sont dans  $\mathcal{L}^2$ , la variable aléatoire  $(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))$  est intégrable. On appelle la *covariance* de  $X$  et  $Y$  l'espérance de cette variable aléatoire et on la note :

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

Le *coefficient de corrélation* des variables aléatoires  $X$  et  $Y$  est le nombre

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\mathbb{V}(X)}\sqrt{\mathbb{V}(Y)}}$$

qui est bien défini lorsque  $\mathbb{V}(X) > 0$  et  $\mathbb{V}(Y) > 0$ .

**Exercice – Calcul d'une covariance (•)** Soit  $U \sim \mathcal{U}_{[0,1]}$  et  $V = 1 - U$ . Calculer  $\text{Cov}(U, V)$  et  $\rho(U, V)$ . (Solution p. 27.)

Du fait de la linéarité de l'espérance, on a

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

et d'ailleurs, on voit que la formule de calcul de la variance donnée plus haut est un cas particulier de cette formulation car  $\mathbb{V}(X) = \text{Cov}(X, X)$ . La linéarité de l'espérance nous donne encore pour  $a, a', b, b' \in \mathbb{R}$

$$\begin{aligned}\mathbb{E}((aX + b)(a'Y + b')) &= aa'\mathbb{E}(XY) + ab'\mathbb{E}(X) + a'b\mathbb{E}(Y) + bb' \\ \mathbb{E}(aX + b)\mathbb{E}(a'Y + b') &= aa'\mathbb{E}(X)\mathbb{E}(Y) + ab'\mathbb{E}(X) + a'b\mathbb{E}(Y) + bb'\end{aligned}$$

On en déduit que la covariance est une forme bilinéaire sur l'espace vectoriel des variables aléatoires de carré intégrable, et nous avons

$$\text{Cov}(aX + b, a'Y + b') = aa' \text{Cov}(X, Y)$$

En particulier, on a

$$\mathbb{V}(aX) = a^2 \mathbb{V}(X)$$

On en déduit que les coefficients de corrélation de  $X$  et  $Y$  et de  $aX + b$  et  $a'Y + b'$  sont égaux lorsque  $aa' > 0$ .

### Proposition – Inégalité de Cauchy-Schwarz

Soit  $X$  et  $Y$  deux variables aléatoires de carré intégrable, alors on a l'inégalité de Cauchy-Schwarz :

$$|\mathbb{E}(XY)| \leq \mathbb{E}(|XY|) \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$$

avec égalité si et seulement si  $X$  et  $Y$  sont presque-sûrement proportionnelles.

**Démonstration** La première inégalité a été démontrée plus haut. Pour la seconde, on a  $\forall x \in \mathbb{R}$  d'après la linéarité de l'espérance :

$$x^2 \mathbb{E}(X^2) + 2x \mathbb{E}(XY) + \mathbb{E}(Y^2) = \mathbb{E}((xX + Y)^2) \geq 0.$$

Mais ceci n'est possible que si ce trinôme en  $x$  n'a au plus qu'une seule racine réelle ; son discriminant

$$4((\mathbb{E}(XY))^2 - \mathbb{E}(X^2)\mathbb{E}(Y^2))$$

doit donc être négatif ou nul ce qui donne le résultat.

Le discriminant est nul si et seulement si le trinôme admet une racine double  $x_0$  et dans ce cas,  $Y(\omega) = -x_0 X(\omega)$  pour presque tout  $\omega$ . ■

**Exercice – Coefficient de corrélation (•)** Montrer que le coefficient de corrélation de  $X$  et  $Y$  vérifie

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

(Solution p. 28.)

Enfin, il peut être intéressant de pouvoir calculer l'espérance d'une fonction mesurable d'une variable aléatoire réelle qui est une variable aléatoire en vertu de la proposition de composition.

### Proposition – Espérance de $g(X)$

Soit  $X$  une variable aléatoire réelle de loi  $\mathbb{P}_X$ , et  $g$  une fonction mesurable de  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  dans  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ . Alors  $g(X)$  est  $\mathbb{P}_X$  - intégrable si et seulement si l'intégrale

$$\int_{\mathbb{R}} |g(x)| d\mathbb{P}_X(x),$$

est définie et dans ce cas

$$\mathbb{E}(g(X)) = \int_{\mathbb{R}} g(x) d\mathbb{P}_X(x) = \int_{\Omega} g(X(\omega)) d\mathbb{P}(\omega)$$

Si  $\mathbb{P}_X$  admet une densité  $f$  par rapport à la mesure de Lebesgue, alors  $g(X)$  est  $\mathbb{P}_X$  - intégrable si et seulement si le produit  $fg$  est intégrable par rapport à la mesure de Lebesgue. On a alors :

$$\mathbb{E}(g(X)) = \int_{\mathbb{R}} g(x) d\mathbb{P}_X(x) = \int_{\mathbb{R}} g(x) f(x) dx = \int_{\Omega} g(X(\omega)) d\mathbb{P}(\omega)$$

**Démonstration** On rappelle que la loi de probabilité  $\mathbb{P}_X$  est définie pour  $B \in \mathcal{B}(\mathbb{R})$  par  $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$ . Par conséquent

$$\mathbb{E}(1_B(X)) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}_X(B) = \int_{\mathbb{R}} 1_B(x) d\mathbb{P}_X(x).$$

Ainsi, si  $g$  est une fonction étagée positive ( $g \in \mathcal{E}^+$ ), on obtient le résultat voulu par linéarité.

Si  $g$  est positive, soit  $(g_n)_{n \in \mathbb{N}^*}$  une suite croissante de fonctions étagées positives convergeant simplement vers  $g$ . On a alors

$$\begin{aligned} \mathbb{E}(g(x)) &= \mathbb{E}(\lim_{n \rightarrow \infty} g_n(X)) \\ &= \lim_{n \rightarrow \infty} \mathbb{E}(g_n(X)) \quad \text{par le théorème de convergence monotone} \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}} g_n(x) d\mathbb{P}_X(x) \\ &= \int_{\mathbb{R}} \lim_{n \rightarrow \infty} g_n(x) d\mathbb{P}_X(x) \quad \text{par le théorème de convergence monotone} \\ &= \int_{\mathbb{R}} g(x) d\mathbb{P}_X(x). \end{aligned}$$

Enfin, si  $g$  n'est pas positive, on utilise la décomposition  $g = g^+ - g^-$  et on déduit le résultat par soustraction. ■

### Remarque – Cas particulier

L'espérance et la variance sont des cas particuliers de ce résultat.

### Proposition – Inégalité de Jensen

Soit  $X$  une variable aléatoire réelle intégrable de loi  $\mathbb{P}_X$ , et soit  $g : \mathbb{R} \rightarrow \mathbb{R}$  une fonction convexe telle que  $g(X)$  soit  $\mathbb{P}_X$  - intégrable. On a alors

$$g(\mathbb{E}(X)) \leq \mathbb{E}(g(X)).$$



**Démonstration** Puisque  $g$  est convexe, pour tout  $a \in \mathbb{R}$  il existe  $\lambda_a \in \mathbb{R}$  tel que pour tout  $x \in \mathbb{R}$  on a

$$g(x) \geq g(a) + \lambda_a (x - a).$$

C'est une conséquence directe de la caractérisation de la convexité par les inégalités des pentes. C'est vrai en particulier pour  $x = X(\omega)$ ,  $\omega \in \Omega$ , et  $a = \mathbb{E}(X)$  : pour tout  $\omega \in \Omega$ ,

$$g(X(\omega)) \geq g(\mathbb{E}(X)) + \lambda_{\mathbb{E}(X)} (X(\omega) - \mathbb{E}(X)).$$

En intégrant de chaque côté de l'inégalité, on obtient bien

$$\begin{aligned} \int_{\Omega} g(X(\omega)) \mathbb{P}(d\omega) &= \mathbb{E}(g(X)) \geq g(\mathbb{E}(X)) + \lambda_{\mathbb{E}(X)} \left( \int_{\Omega} X(\omega) \mathbb{P}(d\omega) - \mathbb{E}(X) \right) \\ &= g(\mathbb{E}(X)) + \lambda_{\mathbb{E}(X)} (\mathbb{E}(X) - \mathbb{E}(X)) \\ &= g(\mathbb{E}(X)). \end{aligned}$$

■

## Exemples de loi à densité

Nous donnons ici quelques exemples de densités de probabilité. Nous reprenons en particulier les trois exemples de densités donnés au premier cours :

### *Loi uniforme*

sur  $[a, b]$ , où  $a < b$  et on note  $X \sim \mathcal{U}_{[a,b]}$  si  $X$  est de densité

$$f(x) = \frac{1}{b-a} 1_{[a,b]}(x).$$

### **Exercice – Moments d'une variable aléatoire de loi uniforme sur $[a, b]$**

(●) Soit  $X \sim \mathcal{U}_{[a,b]}$ . Montrer que son espérance vaut

$$\mathbb{E}(X) = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$$

et que sa variance vaut

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{(b-a)^2}{12}.$$

(Solution p. 28.)

### *Loi exponentielle*

de paramètre  $\theta > 0$  et on note  $X \sim \mathcal{E}(\theta)$  si  $X$  est de densité

$$f(x) = \theta e^{-\theta x} 1_{\{x>0\}}(x).$$

**Exercice – Moments d’une v.a. exponentielle (•)** Montrer que son espérance et sa variance valent

$$\mathbb{E}(X) = \frac{1}{\theta} \text{ et } \mathbb{V}(X) = \frac{1}{\theta^2}$$

(Solution p. 28.)

### **Loi gamma**

On rappelle tout d’abord que la fonction gamma est définie pour  $\alpha \in ]0, +\infty[$  par

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx.$$

En intégrant par partie, on obtient la relation  $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ , et on a  $\Gamma(1) = 1$ . On en déduit que  $\Gamma(n) = (n-1)!$ .

Une variable aléatoire  $X$  suit une loi gamma de paramètre d’échelle  $\theta$  et d’indice  $\alpha$  et on note  $X \sim \Gamma(\alpha, \theta)$ , si sa loi admet la densité

$$f(x) = \frac{1}{\Gamma(\alpha)} \theta^\alpha x^{\alpha-1} e^{-\theta x}.$$

**Exercice – Moments d’une v.a. de loi gamma (•)** Soit  $X \sim \Gamma(\alpha, \theta)$ . Montrer que son espérance et sa variance valent :

$$\mathbb{E}(X) = \frac{\alpha}{\theta} \text{ et } \mathbb{V}(X) = \frac{\alpha}{\theta^2}.$$

On remarquera que  $\Gamma(1, \theta)$  est la loi exponentielle de paramètre  $\theta$ .

Lorsque  $\alpha$  est entier, la loi gamma permet de modéliser le temps d’attente avant la  $n$ -ième occurrence d’événements indépendants de loi exponentielle de paramètre  $\theta$ . (Solution p. 28.)

### **Loi normale**

de paramètres  $\mu$  et  $\sigma^2$  et on note  $X \sim \mathcal{N}(\mu, \sigma^2)$  si  $X$  est de densité

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Son espérance et sa variance valent

$$\mathbb{E}(X) = \mu \text{ et } \mathbb{V}(X) = \sigma^2$$

**Exercice – Moments d’une v.a. gaussienne (•)** Après avoir vérifié que  $f$  est bien une densité, calculer l’espérance et la variance de  $X \sim \mathcal{N}(\mu, \sigma^2)$  (Solution p. 29.)

### Remarque – Estimation statistique

Dans les exemples ci-dessus, on peut remarquer que les densités sont paramétrées par des nombres réels qui sont liés directement aux valeurs de l'espérance et de la variance de la variable. C'est très utile en statistique où l'on cherchera à estimer la valeur de ces paramètres à partir des observations disponibles. Dans le cas de la loi normale, la moyenne et la variance (empiriques) des échantillons fourniront ainsi directement des estimateurs des paramètres.

Il existe des variables aléatoires qui n'ont pas d'espérance, comme le montre l'exercice suivant.

**Exercice – Loi de Cauchy (••)** Un gyrophare envoie un flash lumineux dans une direction aléatoire uniforme d'angle  $\theta$ . On cherche la distribution de l'abscisse  $X$  du point d'impact du rayon lumineux sur un écran plan infini situé à distance 1 du gyrophare. Donner la densité de  $X$ . Calculer son espérance. (Solution p. 29.)

## Rappels sur les intégrales multiples

Dans la suite, nous aurons besoin de quelques résultats à propos des intégrales multiples que nous reformulons ici dans le contexte des probabilités.

### Théorème – Théorème de Fubini (pour les probabilités)

Soient  $\mathbb{P}_X$  et  $\mathbb{P}_Y$  deux probabilités définies respectivement sur  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  et  $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ .

1. Soient  $A \in \mathcal{B}(\mathbb{R}^n)$  et  $B \in \mathcal{B}(\mathbb{R}^m)$ , alors

$$\mathbb{P}(A \times B) = \mathbb{P}_X(A)\mathbb{P}_Y(B)$$

définit de manière unique une probabilité sur  $(\mathbb{R}^{n+m}, \mathcal{B}(\mathbb{R}^{n+m}))$  (que l'on note aussi  $\mathbb{P}_X \otimes \mathbb{P}_Y$ ).

2. Soit  $f$  une fonction  $\mathcal{B}(\mathbb{R}^{n+m})$ -mesurable positive ou  $\mathbb{P}_X \otimes \mathbb{P}_Y$ -intégrable, alors la fonction  $x \mapsto \int f(x, y) \mathbb{P}_Y(dy)$  est  $\mathcal{B}(\mathbb{R}^n)$ -mesurable, la fonction  $y \mapsto \int f(x, y) \mathbb{P}_X(dx)$  est  $\mathcal{B}(\mathbb{R}^m)$ -mesurable et

$$\int f d\mathbb{P}_X \otimes \mathbb{P}_Y = \int \left( \int f(x, y) \mathbb{P}_Y(dy) \right) \mathbb{P}_X(dx) = \int \left( \int f(x, y) \mathbb{P}_X(dx) \right) \mathbb{P}_Y(dy)$$

**Exercice – Calcul de l'aire d'un triangle (•)** Considérons le triangle

$$T = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0 \text{ et } x + y \leq 1\}.$$

En supposant l'intégrale ci-dessous bien définie, calculer :

$$a(T) = \int_{\mathbb{R}^2} 1_T(x, y) dx dy.$$

(Solution p. 29.)

**Exercice – Contre-exemple (••)** Comparer les valeurs des intégrales multiples

$$\int_0^1 \left[ \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx \right] dy \quad \text{et} \quad \int_0^1 \left[ \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dy \right] dx,$$

puis expliquer le résultat. Indication : on remarquera que

$$\frac{x^2 - y^2}{(x^2 + y^2)^2} = \frac{\partial}{\partial x} \left( -\frac{x}{x^2 + y^2} \right)$$

(Solution p. 30.)

Pour pouvoir appliquer le théorème de Fubini, il faut savoir a priori que la fonction est intégrable et pas seulement que ses intégrales multiples sont bien définies (cf. “Contre-exemple” ci-dessus (p. 12)). Si la fonction est à valeurs positives toutefois, l’examen de ses intégrales multiples permet de s’assurer de l’intégrabilité ; c’est le théorème de Fubini-Tonelli vu dans le cours de calcul intégral.

Autrement dit, en pratique, on appliquera Fubini-Tonelli à la valeur absolue de la fonction pour vérifier que celle-ci est intégrable, puis Fubini, s’il est applicable, pour effectuer le calcul de l’intégrale.

### Théorème – Changement de variables

Soient  $D_1$  et  $D_2$  des ouverts de  $\mathbb{R}^n$  et  $h : D_1 \rightarrow D_2$  un  $C^1$ -difféomorphisme de  $D_1$  sur  $D_2$ , c’est-à-dire une fonction continûment différentiable et bijective dont l’inverse  $h^{-1} : D_2 \rightarrow D_1$  est également continûment différentiable. La matrice jacobienne associée à la différentielle de  $h$  étant notée  $J_h$ , la fonction  $f : D_2 \rightarrow \mathbb{R}$  est intégrable si et seulement si la fonction  $(f \circ h)|\det J_h| : D_1 \rightarrow \mathbb{R}$  est intégrable et dans ce cas,

$$\int_{D_2} f(y) dy = \int_{D_1} f(h(x)) |\det J_h(x)| dx.$$

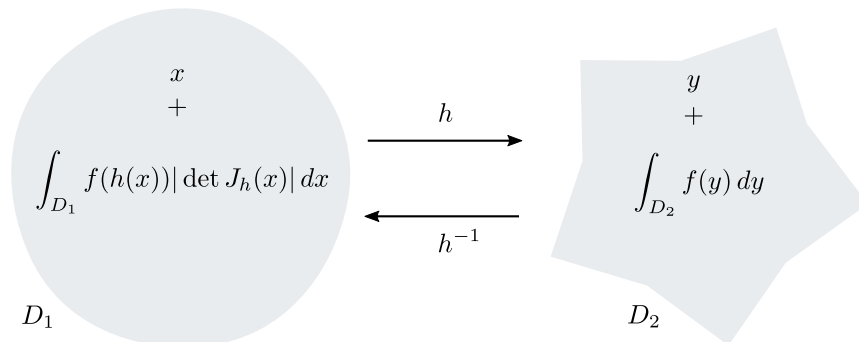


FIGURE 1 – Changement de variables

**Exercice – Homothétie (•)** Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction intégrable. Montrer que pour tout coefficient  $\alpha > 0$ , l'intégrale

$$\int_{\mathbb{R}^n} f(\alpha x) dx$$

est bien définie et la calculer en fonction de l'intégrale de  $f$  sur  $\mathbb{R}^n$ . (Solution p. 30.)

**Exercice – Volume et translation (•)** Montrer que si  $A$  est mesurable et de volume fini dans  $\mathbb{R}^3$  l'image de  $A$  par une translation est également mesurable et de même volume. (Solution p. 31.)

**Exercice – Coordonnées polaires (••)** Soit

$$C = \{(x, y) \in \mathbb{R}^2 \mid y \neq 0 \text{ ou } x > 0\} \text{ et } P = \{(r, \theta) \in \mathbb{R}^2 \mid r > 0 \text{ et } -\pi < \theta < \pi\}.$$

On note  $h$  la fonction de  $P$  dans  $C$  définie par  $h(r, \theta) = (r \cos \theta, r \sin \theta)$ . Montrer que pour toute fonction  $f : C \rightarrow \mathbb{R}$  intégrable, si l'on pose  $g(r, \theta) = f(x, y)$  où  $(x, y) = h(r, \theta)$ , alors

$$\int_C f(x, y) d(x, y) = \int_P g(r, \theta) r d(r, \theta).$$

(Solution p. 31.)

**Exercice – Absence du déterminant jacobien (••)** Supposons  $D_1, D_2, h$  et  $f$  conformes aux hypothèses du théorème de changement de variables (p. 12). On suppose de plus que  $f \circ h$  est intégrable sur  $D_1$ . Exprimer l'intégrale

$$\int_{D_1} f(h(x)) dx$$

comme une intégrale sur  $D_2$ . (Solution p. 31.)

## Vecteurs aléatoires

Nous allons généraliser ici la notion de variable aléatoire en considérant qu'elle peut prendre ses valeurs dans  $\mathbb{R}^n$ . Les vecteurs aléatoires se rencontrent naturellement lorsqu'on s'intéresse à plusieurs quantités conjointement, par exemple dans le cas de la météo, la température, la pluviométrie et la vitesse et la direction du vent.

### Définitions

Une variable aléatoire  $X$  à valeurs dans  $\mathbb{R}^n$  (ou vecteur aléatoire) est simplement une collection de  $n$  variables réelles définies sur le même espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ , qui sont les *composantes* de  $X$ . On écrit  $X = (X_1, \dots, X_n)$ .

De même qu'en dimension 1, la loi de  $X$  est caractérisée par la fonction de répartition multi-dimensionnelle  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  définie par

$$F(x_1, \dots, x_n) = \mathbb{P}_X(X_1 \leq x_1, \dots, X_n \leq x_n)$$

Cependant, caractériser les fonctions de répartition sur  $\mathbb{R}^n$  est délicat, de sorte que cette notion est rarement utilisée. Nous allons plus particulièrement nous intéresser aux vecteurs aléatoires à densité. On gardera tout de même à l'esprit que les lois de probabilités définies sur  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  n'admettent pas toutes une densité !

### Définition – Vecteur aléatoire à densité

On dit que  $X$  admet la densité  $f$  si la fonction réelle  $f$  sur  $\mathbb{R}^n$  est positive, intégrable et vérifie

$$\int_{\mathbb{R}^n} f(x) dx = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x_1, \dots, x_n) dx_1 \dots dx_n = 1$$

et si

$$\mathbb{P}_X(X_1 \leq x_1, \dots, X_n \leq x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(x_1, \dots, x_n) dx_1 \dots dx_n.$$

De la même manière que dans le cas unidimensionnel (p. 7), on a :

### Proposition – Fonction d'un vecteur aléatoire

Soit  $X$  un vecteur aléatoire de densité  $f$ , et soit  $g$  une fonction de  $\mathbb{R}^n$  dans  $\mathbb{R}$ , mesurable. On a alors  $g(X) \in \mathcal{L}^1$  si et seulement si

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} |g(x_1, \dots, x_n)| f(x_1, \dots, x_n) dx_1 \dots dx_n,$$

est définie et dans ce cas, on a

$$\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} g(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 \dots dx_n.$$

Pour revenir à la densité d'une composante d'un vecteur aléatoire, on intègre par rapport aux autres variables. On le présente ici dans le cas d'un couple  $Z = (X, Y)$  de variables aléatoires. On généralise aisément à une dimension supérieure.

### Proposition – Densités marginales

Supposons que  $Z$  admette une densité  $f$ . Alors  $X$  et  $Y$  admettent les densités  $f_X$  et  $f_Y$  données par

$$f_X(x) = \int_{\mathbb{R}} f(x, y) dy, \quad f_Y(y) = \int_{\mathbb{R}} f(x, y) dx.$$

Les fonctions  $f_X$  et  $f_Y$  s'appellent les *densités marginales* de  $f$ .

### Démonstration

Pour tout  $x \in \mathbb{R}$ , on a par définition

$$\mathbb{P}(X \leq x) = \mathbb{P}(Z \in ]-\infty, x] \times \mathbb{R}) = \mathbb{E}(1_{]-\infty, x] \times \mathbb{R}}(Z)) = \int_{-\infty}^x du \left( \int_{\mathbb{R}} f(u, v) dv \right),$$

où l'on a utilisé le théorème de Fubini-Tonelli pour  $1_{]-\infty, x] \times \mathbb{R}}(u, v)f(u, v)$  qui est bien positive. Donc si  $f_X$  est définie par  $f_X(x) = \int_{\mathbb{R}} f(x, y) dy$ , on obtient que  $\mathbb{P}(X \leq x) = \int_{-\infty}^x f_X(u) du$ , ce qui montre que  $f_X$  est la densité de  $X$ . Le raisonnement est analogue pour  $Y$ .

### Remarque – Réciproque ?

La réciproque de cette proposition est fautive en revanche : les variables aléatoires  $X$  et  $Y$  peuvent avoir des densités sans que le couple  $Z = (X, Y)$  en ait une.

Supposons par exemple que  $X = Y$ . Si  $\Delta = \{(x, x); x \in \mathbb{R}\}$  est la diagonale de  $\mathbb{R}^2$ , nous avons évidemment  $\mathbb{P}_Z(\Delta) = 1$  mais si la proposition précédente (p. 14) était valide pour  $\mathbb{P}_Z$ , on aurait  $\mathbb{P}_Z(\Delta) = \mathbb{E}(1_\Delta) = \int_{\mathbb{R}^2} 1_\Delta f_Z(z) dz = 0$  car  $\Delta$  est de volume nul dans  $\mathbb{R}^2$ .

En particulier, il faut faire attention au fait que dans le cas général, la densité d'un couple de variables aléatoires n'est pas le produit des densités.

**Exercice – Lancer de fléchette (••)** On lance une fléchette sur une cible circulaire de rayon unité. Le joueur est suffisamment maladroit pour que le point  $M$  d'impact de la fléchette soit supposé uniformément distribué sur la cible (On décide de n'observer que les lancers qui atteignent la cible). Exprimer la densité du couple formé par les coordonnées cartésiennes  $(X, Y)$  de  $M$  puis leurs densités marginales. (Solution p. 32.)

### Moments d'un vecteur aléatoire

#### Définition – Vecteur espérance et covariance

Si les composantes  $X_i$  du vecteur aléatoire  $X = (X_1, \dots, X_n)$  sont intégrables, on peut définir le *vecteur espérance*

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

Si les composantes  $X_i$  du vecteur aléatoire  $X = (X_1, \dots, X_n)$  sont de carré intégrable, la *matrice de covariance* de  $X$  est la matrice  $C_X = (c_{i,j})_{1 \leq i \leq n, 1 \leq j \leq n}$  de taille  $n \times n$  et dont les éléments valent

$$c_{i,j} = \text{Cov}(X_i, X_j)$$

#### Proposition – Matrice de covariance

La matrice de covariance est symétrique non-négative (ou encore semi-définie positive).

**Démonstration** La symétrie est évidente. Non-négative signifie que pour tous réels  $a_1, \dots, a_n$ , on a  $\sum_{i=1}^n \sum_{j=1}^n a_i a_j c_{i,j} \geq 0$ . Un calcul simple montre que

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j c_{i,j} = \mathbb{V}(\sum_{i=1}^n a_i X_i) \geq 0.$$

■

**Exemple – Vecteur Gaussien  $n$ -dimensionnel** Un exemple important de vecteurs aléatoires est celui des vecteurs gaussiens, que nous étudierons plus en détail au chapitre suivant. Soient  $m \in \mathbb{R}^n$  et  $C$  une matrice symétrique définie positive (c'est-à-dire telle que pour tout  $x \in \mathbb{R}^n$  non identiquement nul  $x^t C x > 0$  où  $^t$  désigne la transposée). Le vecteur  $X \in \mathbb{R}^n$  est un vecteur aléatoire gaussien d'espérance  $m$  et de matrice de covariance  $C$  si sa densité s'écrit

$$f(x) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(C)}} \exp\left(-\frac{1}{2}(x - m)^t C^{-1}(x - m)\right)$$

On a alors  $\mathbb{E}(X) = m$  et  $C_X = C$ .

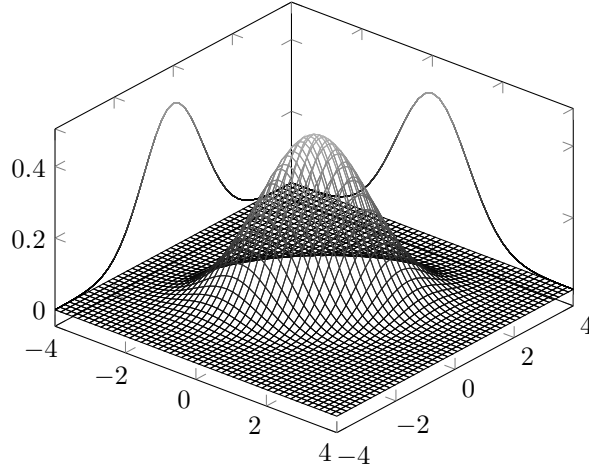


FIGURE 2 – Densité de probabilité de la loi normale bvariée centrée, réduite, de coefficient de corrélation 1/2

## Variables aléatoires indépendantes

Lorsqu'on modélise plusieurs variables conjointement, une hypothèse importante est celle de l'indépendance. Ce caractère traduit l'absence de lien de causalité entre les variables. Par exemple, on fait naturellement l'hypothèse d'indépendance lorsque l'on considère une répétition d'une même expérience dans les mêmes conditions.

Dans ce paragraphe, on considère un couple  $(X, Y)$  de vecteurs aléatoires respectivement à valeurs dans  $\mathbb{R}^m$  et  $\mathbb{R}^n$ . Les résultats s'étendent sans peine à une famille finie quelconque.



On peut se ramener aux évènements pour caractériser l'indépendance de deux vecteurs aléatoires. En effet, considérons le vecteur aléatoire  $Z = (X, Y)$ ,  $A$  et  $B$  deux ensembles dans  $\mathcal{B}(\mathbb{R}^m)$  et  $\mathcal{B}(\mathbb{R}^n)$ . On a vu que les évènements  $X \in A$  et  $Y \in B$  sont indépendants si et seulement si  $\mathbb{P}_Z(X \in A, Y \in B) = \mathbb{P}(X^{-1}(A) \cap Y^{-1}(B)) = \mathbb{P}(X^{-1}(A))\mathbb{P}(Y^{-1}(B)) = \mathbb{P}_X(X \in A)\mathbb{P}_Y(Y \in B)$ . Pour que deux vecteurs aléatoires soient indépendants, on va donc demander que ceci soit valable quels que soient  $A$  et  $B$ .

### Définition – Vecteurs aléatoires indépendants

Les vecteurs aléatoires  $X$  et  $Y$  sont *indépendants* si pour tous ensembles  $A$  et  $B$  des tribus correspondantes,

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

En termes de mesures de probabilité, cette égalité s'écrit

$$\mathbb{P}_{(X,Y)}(A \times B) = \mathbb{P}_X(A)\mathbb{P}_Y(B) = (\mathbb{P}_X \otimes \mathbb{P}_Y)(A \times B).$$

Autrement dit,  $X$  et  $Y$  sont indépendantes si et seulement si la mesure de probabilité du couple  $(X, Y)$  est la mesure produit  $P_X \otimes P_Y$  dont l'existence et l'unicité sont garanties par le théorème de Fubini (p. 11).

Cette définition se traduit dans le cas à densité dans la proposition suivante que l'on énonce sans perte de généralité pour un couple de variables aléatoires.

### Proposition – Caractérisation de l'indépendance – cas à densité

Soient  $X$  et  $Y$  deux variables aléatoires réelles de densités  $f_X$  et  $f_Y$ .  $X$  et  $Y$  sont indépendantes si et seulement si le couple  $Z = (X, Y)$  a pour densité (sur  $\mathbb{R}^2$ ) :

$$f_Z(x, y) = f_X(x)f_Y(y).$$

**Démonstration** S'il y a indépendance, la définition (p. 17) implique

$$P(X \leq x, Y \leq y) = P(X \leq x)\mathbb{P}(Y \leq y) = \int_{-\infty}^x f_X(u)du \int_{-\infty}^y f_Y(v)dv$$

ce qui montre que  $\mathbb{P}_Z$  vérifie la définition d'un vecteur aléatoire à densité (p. 14) avec  $f_Z = f_X f_Y$ .

Inversement, si  $f_Z = f_X f_Y$ , on a pour tous  $A, B$  de  $\mathcal{B}(\mathbb{R})$

$$\mathbb{P}(X \in A, Y \in B) = \int_A \int_B f_X(x)f_Y(y)dxdy = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

où on a utilisé le théorème de Fubini-Tonelli. ■

**Exercice – Lancer de fléchette (suite) (•)** Les coordonnées cartésiennes  $(X, Y)$  de  $M$  sont-elles indépendantes? (Solution p. 32.)

Considérons maintenant deux fonctions mesurables  $g$  et  $h$  définies sur  $\mathbb{R}^m$  et  $\mathbb{R}^n$ .

### Proposition – Fonctions de vecteurs aléatoires indépendants

Avec les notations précédentes, si  $X$  et  $Y$  sont indépendantes de densités respectives  $f_X$  et  $f_Y$ , les variables aléatoires  $g(X)$  et  $h(Y)$  sont aussi indépendantes. Si de plus  $g(X)$  et  $h(Y)$  sont intégrables, alors le produit  $g(X)h(Y)$  est aussi intégrable, et on a

$$\mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X))\mathbb{E}(h(Y))$$

**Démonstration** La première assertion est évidente par définition de l'indépendance. Par ailleurs, si  $g(X)$  et  $h(Y)$  sont intégrables, en notant  $f_{(X,Y)}$  la densité du couple  $(X, Y)$ , et en utilisant le théorème de Fubini, on a

$$\begin{aligned}\mathbb{E}(g(X)h(Y)) &= \int_{\mathbb{R}^{m+n}} g(x)h(y)f_{(X,Y)}(x,y)dx dy \\ &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^n} g(x)h(y)f_X(x)f_Y(y)dx dy \\ &= \left( \int_{\mathbb{R}^m} g(x)f_X(x)dx \right) \left( \int_{\mathbb{R}^n} h(y)f_Y(y)dy \right) \\ &= \mathbb{E}(g(X))\mathbb{E}(h(Y))\end{aligned}$$

■

### Remarque – Cas général

Ce résultat est encore valable dans le cas général (sans densité). Il suffit de remplacer  $f_{(X,Y)}(x,y)dx dy$  par  $d\mathbb{P}_{X,Y}(x,y) = d\mathbb{P}_X(x)d\mathbb{P}_Y(y)$  dans la démonstration précédente.

On déduit de ce résultat et de la définition de la covariance (p. 6) que :

### Corollaire – Covariance de variables aléatoires indépendantes

Si les variables aléatoires réelles  $X$  et  $Y$  sont indépendantes et de carré intégrable, alors  $\text{Cov}(X, Y) = 0$  et  $\rho(X, Y) = 0$ .

**Exercice – Démonstration (•)** Démontrer ce résultat. (Solution p. 32.)

### Remarque

Attention, la réciproque est fausse. Par exemple, si  $X \sim \mathcal{U}_{[-1,1]}$  et  $Y = X^2$ .  $X$  et  $Y$  ne sont clairement pas indépendantes mais on a

$$\text{Cov}(X, Y) = \text{Cov}(X, X^2) = \mathbb{E}(X^3) - \mathbb{E}(X)\mathbb{E}(X^2) = 0$$

## Identification de densité

Un problème important est le suivant. Soit  $X$  une variable aléatoire réelle, admettant la densité  $f_X$ . Soit  $g$  une fonction mesurable, de sorte que  $Y = g(X)$  est aussi une variable aléatoire. Est-ce que  $Y$  admet une densité, et si oui, comment la calculer ?

On peut déjà remarquer que cette densité n'existe pas toujours. Si par exemple  $g(x) = a$  pour tout  $x$ , la loi de  $Y$  est la mesure de Dirac en  $a$ , qui n'a pas de densité par rapport à la mesure de Lebesgue.

Pour résoudre ce problème, l'idée consiste à essayer de mettre  $E(h(Y)) = E(h \circ g(X))$  sous la forme  $\int h(y)f_Y(y)dy$  pour une fonction convenable  $f_Y$ , et une classe de fonctions  $h$  suffisamment grande. La fonction  $f_Y$  sera alors la densité cherchée.

La proposition qui assure que  $g(X)$  est intégrable pour une variable aléatoire réelle  $X$  (p. 7) implique

$$\mathbb{E}(h(Y)) = \mathbb{E}(h \circ g(X)) = \int_{\mathbb{R}} h \circ g(x)f_X(x)dx$$

et on fait le changement de variable  $y = g(x)$  dans cette intégrale. Cela nécessite que  $g$  soit dérivable et bijective "par morceaux", et il faut faire très attention aux domaines où  $g$  est croissante ou décroissante. Puisque la fonction  $h$  est arbitraire, on appelle couramment cette technique la *méthode de la fonction muette*. Cette approche résulte en fait de la proposition suivante :

**Proposition – Méthode de la fonction muette**

Si il existe une fonction  $f$  positive telle que pour toute fonction mesurable  $h$  telle que  $h(x)f(x)$  soit intégrable,

$$\mathbb{E}(h(X)) = \int_{\mathbb{R}} h(x)f(x)dx$$

alors la loi de  $X$  admet la densité  $f$  par rapport à la mesure de Lebesgue.

**Démonstration** Supposons qu'un tel  $f$  existe. Soit  $B \in \mathcal{B}(\mathbb{R})$ , alors  $1_B$  est mesurable. En particulier, si  $B = ]-\infty, x]$ ,  $x \in \mathbb{R}$ , on a

$$\mathbb{P}(X \leq x) = \mathbb{P}(X \in B) = \mathbb{P}_X(B) = \mathbb{E}(1_B(X)) = \int_{-\infty}^x f(y)dy$$

Par ailleurs, prenant  $B = \mathbb{R}$ , on a  $1 = \mathbb{P}(X \in \mathbb{R}) = \int_{\mathbb{R}} f(y)dy$ . Par conséquent,  $\mathbb{P}_X$  admet la densité  $f$ . ■

Nous donnons ici quelques exemples d'application de cette méthode :

- Soit  $Y = aX + b$ , où  $a$  et  $b$  sont des constantes. Si  $a = 0$ , alors  $Y = b$  et la loi de  $Y$  est la masse de Dirac en  $b$  (sans densité). Si  $a \neq 0$ , on fait le changement de variable  $y = ax + b$ , ce qui donne

$$\mathbb{E}(h(Y)) = \int_{\mathbb{R}} h(ax + b)f_X(x)dx = \int_{\mathbb{R}} h(y)f_X\left(\frac{y-b}{a}\right)\frac{1}{|a|}dy$$

Donc

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right)\frac{1}{|a|}$$

- Soit  $Y = X^2$ . La fonction  $g$  est décroissante sur  $\mathbb{R}_-$  et croissante sur  $\mathbb{R}_+$ .  
Le changement de variable  $y = x^2$  donne alors

$$\begin{aligned}\mathbb{E}(h(Y)) &= \int_{-\infty}^0 h(x^2) f_X(x) dx + \int_0^{+\infty} h(x^2) f_X(x) dx \\ &= \int_0^{+\infty} h(y) f_X(-\sqrt{y}) \frac{1}{2\sqrt{y}} dy + \int_0^{+\infty} h(y) f_X(\sqrt{y}) \frac{1}{2\sqrt{y}} dy\end{aligned}$$

et on en déduit

$$f_Y(y) = (f_X(-\sqrt{y}) + f_X(\sqrt{y})) \frac{1}{2\sqrt{y}} 1_{]0, +\infty[}$$

Dans le cas des vecteurs aléatoires, l'idée est la même. Soit  $X = (X_1, \dots, X_n)$ , un vecteur aléatoire de densité  $f_X$  sur  $\mathbb{R}^n$ ,  $g$  une fonction de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$  et  $Y = g(X)$ . Plusieurs cas sont à considérer :

1.  $m > n$ , le vecteur  $Y$  n'admet pas de densité.
2.  $m = n$ , on utilise comme dans le cas unidimensionnel le changement de variable  $y = g(x)$  dans

$$\mathbb{E}(h(Y)) = \mathbb{E}(h \circ g(X)) = \int_{\mathbb{R}^n} h \circ g(x) f_X(x) dx$$

Supposons d'abord que  $g$  soit une bijection continûment différentiable de  $A$  dans  $B$ , ouverts de  $\mathbb{R}^n$ . Sous l'hypothèse que  $h \circ g(x) f_X(x)$  soit intégrable, le théorème de changement de variable nous assure :

$$\int_A h \circ g(x) f_X(x) dx = \int_B h(y) f_X \circ g^{-1}(y) \frac{1}{|\det J_g(g^{-1}(y))|} dy,$$

où  $J_g$  désigne la matrice de Jacobi associée à la différentielle de  $g$ . Dans le cas où  $f_X(x) = 0$  en dehors de  $A$ , on obtient que  $Y$  admet la densité

$$f_Y(y) = 1_B(y) f_X \circ g^{-1}(y) \frac{1}{|\det J_g(g^{-1}(y))|}.$$

Lorsque  $g$  est simplement continûment différentiable, il existe souvent une partition finie  $(A_i)_{1 \leq i \leq n}$  de l'ensemble  $\{x; f_X(x) > 0\}$ , telle que  $g$  soit bijective sur chaque  $A_i$ . On note alors  $B_i = g(A_i)$  l'image de  $A_i$  par  $g$ . On découpe alors l'intégrale selon les  $A_i$ , on applique la formule précédente à chaque morceau et on somme pour obtenir :

$$f_Y(y) = \sum_{i=1}^n 1_{B_i}(y) f_X \circ g^{-1}(y) \frac{1}{|\det J_g(g^{-1}(y))|},$$

où  $g^{-1}$  est bien définie sur chaque  $B_i$  comme image réciproque de la restriction de  $g$  à  $A_i$ .

3.  $m < n$ , on commence par "compléter"  $Y$ , en essayant de construire une application  $g'$  de  $\mathbb{R}^n$  dans  $\mathbb{R}^n$  dont les  $m$  premières composantes coïncident avec les composantes de  $g$  et pour laquelle on peut appliquer l'une des deux formules précédentes. On obtient ainsi la densité  $f_{Y'}$  de  $Y' = g'(X)$  puis on obtient la densité de  $Y$  en calculant sa loi marginale :

$$f_Y(y_1, \dots, y_m) = \int_{\mathbb{R}^{n-m}} f_{Y'}(y_1, \dots, y_m, y_{m+1}, \dots, y_n) dy_{m+1} \dots dy_n.$$

**Exemple – Coordonnées polaires** Soit  $X = (U, V)$  un vecteur aléatoire de  $\mathbb{R}^2$ , et  $Y = (R, \Theta)$  ses coordonnées polaires. La transformation  $g$  est un difféomorphisme de  $A = \mathbb{R}^2 \setminus \{0\}$  dans  $B = ]0, +\infty[ \times ]0, 2\pi]$ , et son inverse  $g^{-1}$  s'écrit :  $u = r \cos \theta$ ,  $v = r \sin \theta$ . Le Jacobien de  $g^{-1}$  au point  $(r, \theta)$  vaut  $r$ , donc le point 2. ci-dessus entraîne que

$$f_Y(r, \theta) = r f_X(r \cos \theta, r \sin \theta) 1_B(r, \theta)$$

Par exemple, si  $U$  et  $V$  sont indépendantes et de loi  $\mathcal{N}(0, 1)$ , on a  $f_X(u, v) = \frac{1}{2\pi} \exp\left(-\frac{u^2+v^2}{2}\right)$  et donc

$$f_Y(r, \theta) = \frac{1}{2\pi} r \exp\left(-\frac{r^2}{2}\right) 1_{]0, +\infty[}(r) 1_{]0, 2\pi]}(\theta).$$

En particulier, on remarque que  $R$  et  $\Theta$  sont indépendantes de densités respectives  $r \exp\left(-\frac{r^2}{2}\right) 1_{]0, +\infty[}(r)$  et  $1_{]0, 2\pi]}(\theta)$ .

**Exercice – Lancer de fléchette (fin) (•)** Donner la densité du couple  $(R, \Theta)$  des coordonnées polaires de  $M$ . Sont-elles indépendantes ? (Solution p. 32.)

**Exemple – Somme de deux variables aléatoires indépendantes** Soient  $U$  et  $V$  indépendantes et admettant les densités  $f_U$  et  $f_V$ , on cherche la densité de la somme  $Z = U + V$ . On commence par compléter  $Z$  en le couple  $T = (U, Z)$  (par exemple), correspondant à la bijection  $g(u, v) = (u, u + v)$  sur  $\mathbb{R}^2$  dont le jacobien est 1 et d'inverse  $g^{-1}(x, y) = (x, x - y)$ . Appliquant le point 3. ci-dessus, on obtient :

$$f_Z(z) = \int_{\mathbb{R}} f_U(u) f_V(z - u) du = \int_{\mathbb{R}} f_U(z - v) f_V(v) dv.$$

La fonction  $f_Z$  est appelée *le produit de convolution* des des fonctions  $f_U$  et  $f_V$ .

**Exercice – Loi bêta (•••)** Soit  $X = (U, V)$  un vecteur aléatoire de  $\mathbb{R}^2$ , avec  $U$  et  $V$  indépendantes de lois  $\Gamma(\alpha, \theta)$  et  $\Gamma(\beta, \theta)$ . Identifier la densité de  $Y = \frac{U}{U+V}$ , puis de  $Z = U + V$ . (Solution p. 32.)

## Modélisation statistique

On a vu un certain nombre d'exemples de lois de probabilité à densité. Celles-ci font intervenir des paramètres qui permettent d'ajuster la forme de la densité aux données observées. Par exemple, la loi normale fait intervenir deux paramètres : l'espérance et la variance. Dans cette section, nous formalisons l'approche statistique qui consiste à choisir une loi de probabilité pour modéliser un phénomène aléatoire, puis à estimer les paramètres de cette loi à partir d'observations.

La première hypothèse que nous faisons est que le phénomène étudié peut être modélisé par une variable aléatoire  $X$  de densité  $f$  inconnue. La première étape de la modélisation consiste à choisir une famille de densités  $(f_\theta)_{\theta \in \Theta}$ , où  $\Theta$  est un ensemble de paramètres. Par exemple, pour la loi normale, on peut choisir  $\Theta = \mathbb{R} \times \mathbb{R}_+^*$  et  $f_{(m, \sigma)}$  la densité de la loi  $\mathcal{N}(m, \sigma^2)$ .

La seconde étape consiste à estimer le paramètre  $\theta$  de manière à ce que la densité  $f_\theta$  soit la plus proche possible de la densité  $f$  inconnue. On a donc besoin d'une notion de proximité entre deux densités. Plusieurs choix sont possibles, mais le plus courant est celui de la divergence de Kullback-Leibler.

### Définition – Divergence de Kullback-Leibler

Soient  $f$  et  $g$  deux densités de probabilité sur  $\mathbb{R}$ , telles que  $g(x) > 0$  pour tout  $x$  tel que  $f(x) > 0$ . La *divergence de Kullback-Leibler* de  $g$  par rapport à  $f$  est définie par

$$D_{KL}(f||g) = \int f(x) \log \left( \frac{f(x)}{g(x)} \right) dx.$$

### Remarque – Remarque

La divergence de Kullback-Leibler n'est pas une distance au sens mathématique du terme. En effet, elle n'est pas symétrique et ne satisfait pas l'inégalité triangulaire. En revanche, on a toujours  $D_{KL}(f||g) \geq 0$ , avec égalité si et seulement si  $f = g$  presque partout.

### Proposition – Positivité de la divergence de Kullback-Leibler

Soient  $f$  et  $g$  deux densités de probabilité sur  $\mathbb{R}$ , telles que  $g(x) > 0$  pour tout  $x$  tel que  $f(x) > 0$ . Alors

$$D_{KL}(f||g) \geq 0,$$

avec égalité si et seulement si  $f = g$  presque partout.

**Démonstration** La fonction  $\log$  est strictement concave, donc  $-\log$  est strictement convexe. On peut donc appliquer l'inégalité de Jensen (p. 8) :

$$\begin{aligned} D_{KL}(f||g) &= \int_{\mathbb{R}} f(x) \log \left( \frac{f(x)}{g(x)} \right) dx = - \int_{\mathbb{R}} f(x) \left( -\log \left( \frac{g(x)}{f(x)} \right) \right) dx \\ &= -\mathbb{E} \left( -\log \left( \frac{g(X)}{f(X)} \right) \right) \geq - \left( -\log \left( \mathbb{E} \left( \frac{g(X)}{f(X)} \right) \right) \right) = -\log(1) = 0. \end{aligned}$$

On a égalité si et seulement si  $\frac{g(x)}{f(x)}$  est presque sûrement constant, ce qui revient à dire que  $f = g$  presque partout. ■

### Exercice – Divergence de Kullback-Leibler entre deux gaussiennes

Soient  $f$  et  $g$  les densités de deux lois normales  $\mathcal{N}(m_1, \sigma_1^2)$  et  $\mathcal{N}(m_2, \sigma_2^2)$ . Montrer que  $D_{KL}(f||g) = \log \left( \frac{\sigma_2}{\sigma_1} \right) + \frac{\sigma_1^2 + (\mu_1 - \mu_2)^2}{2\sigma_2^2} - \frac{1}{2}$ .

Revenons à notre problème d'estimation de paramètres. Soit  $X$  une variable aléatoire de densité inconnue  $f$ . On choisit une famille de densités  $(f_\theta)_{\theta \in \Theta}$ , et on cherche à minimiser la divergence de Kullback-Leibler entre  $f$  et  $f_\theta$  :

$$\hat{\theta} = \arg \min_{\theta \in \Theta} D_{KL}(f||f_\theta).$$

On peut réécrire la divergence de Kullback-Leibler sous la forme suivante :

$$\begin{aligned} D_{KL}(f||f_\theta) &= \int f(x) \log \left( \frac{f(x)}{f_\theta(x)} \right) dx \\ &= \int f(x) \log(f(x)) dx - \int f(x) \log(f_\theta(x)) dx \\ &= -H(f) - \mathbb{E}_f(\log f_\theta(X)), \end{aligned}$$

où  $H(f) = -\int f(x) \log f(x) dx$  est l'**entropie** de la densité  $f$ . Celle-ci s'interprète comme Ce terme ne dépend pas de  $\theta$ , donc minimiser  $D_{KL}(f||f_\theta)$  revient à maximiser  $\mathbb{E}_f(\log f_\theta(X))$ . En statistique,  $f_\theta(X)$  est appelée la *vraisemblance* de l'observation  $X$  pour le paramètre  $\theta$ , et  $\log f_\theta(X)$  la *log-vraisemblance*. L'estimateur  $\hat{\theta}$  est ainsi appelé *estimateur du maximum de vraisemblance*.

En pratique, on dispose d'un échantillon  $X_1, \dots, X_n$  de variables aléatoires indépendantes et identiquement distribuées selon la loi de densité  $f$ . Le choix de la forme paramétrique est généralement guidé par des considérations théoriques, portant sur la physique du phénomène étudié, ou empiriques, en regardant par exemple l'histogramme des données. On peut ensuite estimer  $\mathbb{E}_f(\log f_\theta(X))$  par la moyenne empirique

$$\frac{1}{n} \sum_{i=1}^n \log f_\theta(X_i),$$

et l'estimateur du maximum de vraisemblance s'écrit

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^n \log f_\theta(X_i).$$

Cela revient à remplacer  $f$  par la mesure empirique de l'échantillon  $\tilde{p} = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$  dans la définition de la divergence de Kullback-Leibler.

## Exercices complémentaires

### Durée de vie

La durée de vie, exprimée en années, d'un circuit électronique est une variable aléatoire  $T$  dont la fonction de répartition  $F$  est définie par :

$$F(t) = (1 - e^{-t^2/2}) 1_{t \geq 0}$$

**Question 1** Donner la densité de probabilité  $f$  de  $T$ . Calculer  $\mathbb{E}(T)$ . (Solution p. 34.)

**Question 2** Sachant que le circuit a déjà fonctionné durant 1 an, quelle est la probabilité qu'il continue à fonctionner encore durant au moins 2 ans ? La loi est-elle sans mémoire ?

Un équipement électronique  $E$  est composé de 10 circuits identiques et indépendants. Au circuit  $i$  ( $1 \leq i \leq 10$ ) est associée la variable aléatoire  $X_i$ , avec  $X_i = 1$  si la durée de vie du circuit  $i$  est inférieure à un an et  $X_i = 0$  sinon. (Solution p. 34.)

**Question 3** Quelle est la loi du nombre  $N$  de circuits de  $E$  dont la durée de vie est inférieure à 1 an ? (Solution p. 34.)

**Question 4** L'équipement  $E$  est dit en série si la défaillance de l'un de ses circuits entraîne sa défaillance. Quelle est alors la probabilité qu'il soit défaillant avant 1 an ? (Solution p. 34.)

**Question 5** L'équipement  $E$  est dit en parallèle si sa défaillance ne peut se produire que si tous ses circuits sont défaillants. Quelle est alors la probabilité qu'il soit défaillant avant 1 an ? (Solution p. 34.)

## Crues centennales de la Seine

On suppose que la hauteur d'eau maximale au cours de l'année  $n$  est décrite par une variable aléatoire  $X_n$  de densité  $f(x)$  et de fonction de répartition  $F(x)$ , identique pour toutes les  $X_n$ , que l'on suppose également indépendantes.

**Question 1** Calculer la fonction de répartition du maximum sur  $n$  années  $Y = \max(X_1, \dots, X_n)$ . (Solution p. 34.)

**Question 2** En déduire la densité de  $Y$ . (Solution p. 35.)

**Question 3** Expliciter la densité de  $Y$  lorsque les  $X_i$  sont de loi uniforme sur  $[0, 1]$ . (Solution p. 35.)

**Question 4** Selon mêmes hypothèses, calculer la hauteur minimale des quais pour que la probabilité de voir la Seine déborder sur une période de  $n$  années soit inférieure à  $1/1000$ . (Solution p. 35.)

**Question 5** Toujours selon les mêmes hypothèses, calculer la médiane de  $Y$ . La comparer avec celle des  $X_i$ . Commenter. (Solution p. 35.)

## Indépendance et vecteurs gaussiens

Soit  $Z := (X, Y)$  un vecteur gaussien (i.e. qui suit une loi Normale bi-variée) d'espérance  $m$  et de matrice de covariance définie positive  $C$ . Notons  $f_Z$  sa densité.

**Question 1** Rappeler la forme de  $f_Z$  et en donner une expression non matricielle faisant apparaître le coefficient de corrélation noté  $\rho$ . (Solution p. 35.)

**Question 2** Expliciter les lois marginales de  $X$  et de  $Y$ . (Solution p. 36.)

**Question 3** Montrer que  $X$  et  $Y$  sont indépendantes ssi  $\text{Cov}(X, Y) = 0$ . (Solution p. 36.)



## Symétrie de la gaussienne

Soit  $X$  une variable aléatoire réelle de loi Normale centrée réduite, dont la densité est notée  $f$  et la fonction de répartition  $F$ . Pour tout  $c > 0$  on pose

$$X_c := \begin{cases} X & \text{si } |X| > c, \\ -X & \text{si } |X| \leq c. \end{cases}$$

**Question 1** Déterminer la loi de  $X_c$ . (Solution p. 37.)

**Question 2** Calculer la covariance de  $X$  et de  $X_c$  et montrer qu'il existe une valeur  $c_0$  de  $c$  qui l'annule. (Solution p. 38.)

**Question 3** Les deux variables  $X$  et  $X_{c_0}$  sont-elles indépendantes ? (Solution p. 38.)

**Question 4** Le vecteur aléatoire  $(X, X_{c_0})$  est-il gaussien ? (Solution p. 39.)

## Loi du $\chi^2$

On considère une variable aléatoire  $X$  gaussienne centrée réduite. On note  $f$  sa densité et  $F$  sa fonction de répartition.

### A 1 degré de liberté

1. Calculer la densité de la variable aléatoire  $Y := X^2$ . En donner une expression faisant apparaître la fonction gamma (on rappelle que  $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ ).

On dit que  $Y$  suit une *loi du  $\chi^2$  à 1 degré de liberté*, et on note  $Y \sim \chi^2$ . (Solution p. 39.)

**A n degrés de liberté** Soient maintenant  $n \in \mathbb{N}^*$  et  $X_1, \dots, X_n$  des copies indépendantes de  $X$ . Pour tout  $i \in \{1, \dots, n\}$  on pose  $Y_i := X_i^2$ .

2. Montrer que la variable aléatoire  $Y := \sum_{i=1}^n Y_i$  admet pour densité

$$f_Y : x \in \mathbb{R} \mapsto \begin{cases} \frac{x^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \exp\left\{-\frac{x}{2}\right\} & \text{si } x > 0, \\ 0 & \text{si } x \leq 0. \end{cases}$$

On pourra procéder par récurrence sur  $n$ .

On dit que  $Y$  suit une *loi du  $\chi^2$  à  $n$  degrés de liberté* et on note  $Y \sim \chi_n^2$ . (Solution p. 39.)

## Combinaisons linéaires de variables Gaussiennes indépendantes

Soit  $X$  une variable aléatoire de loi Normale centrée réduite, dont on note  $f$  la densité et  $F$  la fonction de répartition.

### Préliminaires

1. Rappeler la loi de la variable aléatoire  $sX + m$ , où  $s, m \in \mathbb{R}$ .

(Solution p. 40.)

**Combinaisons linéaires** Soient maintenant  $n \in \mathbb{N}^*$  variables aléatoires  $X_1, \dots, X_n$  indépendantes et de même loi que  $X$  (on dit que ce sont des *copies indépendantes* de  $X$ ). Pour tout vecteur  $a = (a_1, \dots, a_n) \in (\mathbb{R}^*)^n$  on pose  $S_n^a := \sum_{i=1}^n a_i X_i$ .

2. Montrer que pour  $S_n^a$  suit une loi Normale d'espérance nulle et de variance  $\sum_{i=1}^n a_i^2$  (on pourra raisonner par récurrence sur  $n$ ).
3. Soient  $a, b \in (\mathbb{R}^*)^n$ . Sous quelle condition a-t-on  $\text{Cov}(S_n^a, S_n^b) = 0$  ?

(Solution p. 41.)

### Loi de Maxwell

On désire déterminer la distribution des vitesses des molécules d'un gaz monoatomique parfait à l'équilibre (loi de Maxwell (1859)).

On représente la vitesse d'une molécule d'un gaz monoatomique parfait à l'équilibre dans un repère orthonormal par un vecteur aléatoire  $V = (V_1, V_2, V_3)$ . Le choix du repère étant arbitraire, il est naturel de supposer que la loi de  $V$  est invariante par rotation (autour de l'origine) et que les composantes de  $V$  sont indépendantes.

### Partie 1

Soit  $(X, Y, Z)$  un vecteur aléatoire à valeurs dans  $\mathbb{R}^3$  de densité  $f_{(X,Y,Z)}$ . Ce vecteur aléatoire est supposé invariant par rotation : il existe une fonction  $\phi : [0, +\infty[ \rightarrow ]0, +\infty[$  (strictement positive) telle que

$$f_{(X,Y,Z)}(x, y, z) = \phi(x^2 + y^2 + z^2) \text{ pour tout } (x, y, z) \in \mathbb{R}^3.$$

**Question 1** On suppose de plus les variables aléatoires  $X, Y$  et  $Z$  indépendantes. Montrer qu'il existe une fonction  $f : \mathbb{R} \rightarrow [0, +\infty[$  qui soit une densité telle que

$$f_{(X,Y,Z)}(x, y, z) = f(x)f(y)f(z) \text{ pour tout } (x, y, z) \in \mathbb{R}^3.$$

Montrer finalement que  $f(x) > 0$  pour tout  $x \in \mathbb{R}$ . (Solution p. 42.)

**Question 2** On suppose de plus que les densités marginales de  $X, Y$  et  $Z$  ainsi que la fonction  $\phi$  sont continûment différentiables (c'est-à-dire de classe  $C^1$ ). Montrer que la densité de chacune des composantes s'écrit sous la forme  $f(x) = ae^{cx^2/2}$ , avec  $a$  et  $c$  deux constantes réelles. (Solution p. 43.)

**Question 3** En déduire que le vecteur  $(X, Y, Z)$  suit une loi gaussienne d'espérance l'origine dont on précisera la matrice de covariance en fonction de la constante  $\sigma = 1/\sqrt{|c|}$ . (Solution p. 44.)

## Partie 2

On suppose que le vecteur aléatoire  $(X, Y, Z) = V$  vérifie les hypothèses des questions précédentes.

**Question 4** Calculer l'énergie cinétique moyenne d'un atome du gaz, c'est-à-dire l'espérance

$$E_c := \mathbb{E} \left( \frac{1}{2} m \|V\|^2 \right)$$

où  $m$  est la masse d'un atome du gaz. L'énergie cinétique moyenne d'un atome du gaz de masse  $m$  étant égale à  $\frac{3}{2} kT$  où  $k$  est la constante de Boltzmann et  $T$  la température du gaz, en déduire la valeur de  $\sigma^2$  en fonction de  $k$ ,  $T$  et  $m$ . (Solution p. 44.)

**Question 5** On rappelle que si  $X$  et  $Y$  sont deux variables aléatoires indépendantes de loi respective  $\Gamma(a, \lambda)$  et  $\Gamma(b, \lambda)$ , alors la loi de  $X + Y$  est la loi  $\Gamma(a + b, \lambda)$ . On rappelle également que la densité  $g$  de la loi  $\Gamma(a, \lambda)$  vérifie

$$g(x) = \frac{1}{\Gamma(a)} \lambda^a x^{a-1} e^{-\lambda x} 1_{]0, +\infty[}(x) \text{ pour tout } x \in \mathbb{R},$$

où  $\Gamma$  est la fonction définie par

$$\Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x} dx \text{ pour } a \in ]0, +\infty[.$$

On remarquera que  $\Gamma(1/2) = \sqrt{\pi}$  et que  $\Gamma(x+1) = x\Gamma(x)$ .

Calculer la loi de  $V_1^2$ . En déduire la loi de  $\|V\|^2$  puis la densité de  $\|V\| = \sqrt{V_1^2 + V_2^2 + V_3^2}$ . La probabilité associée est appelée loi de Maxwell.

---

(Solution p. 44.)

## Solutions

### Exercices essentiels

**Approximation par une constante** On a

$$\mathbb{E}((X - a)^2) = \mathbb{E}((X - \mathbb{E}(X) + \mathbb{E}(X) - a)^2) = \mathbb{E}((X - \mathbb{E}(X))^2) + (\mathbb{E}(X) - a)^2$$

le terme de gauche ne dépend pas de  $a$  tandis que celui de droite s'annule en  $a = \mathbb{E}(X)$ .

**Calcul d'une covariance** On a

$$\begin{aligned} \text{Cov}(U, V) &= \mathbb{E}((U - \mathbb{E}(U))(V - \mathbb{E}(V))) \\ &= \mathbb{E}(UV - \frac{U}{2} - \frac{U}{2} - \mathbb{E}(U)\mathbb{E}(V)) \\ &= \mathbb{E}(U(1 - U)) - 1/4 - 1/4 + 1/4 \text{ par linéarité et d'après la solution ci-dessous} \\ &= 1/2 - 1/3 - 1/4 = -1/12 \end{aligned}$$

et

$$\rho(U, V) = \frac{\text{Cov}(U, V)}{\sqrt{\mathbb{V}(U)}\sqrt{\mathbb{V}(V)}} = -1$$

**Coefficient de corrélation** On déduit de l'inégalité de Cauchy-Schwarz (p. 7) que

$$|\rho(X, Y)| = \left| \frac{\mathbb{E}((X - \mathbb{E}(X))(X - \mathbb{E}(Y)))}{\sqrt{\mathbb{V}(X)}\sqrt{\mathbb{V}(Y)}} \right| \leq \frac{\sqrt{\mathbb{E}((X - \mathbb{E}(X))^2)\mathbb{E}((X - \mathbb{E}(Y))^2)}}{\sqrt{\mathbb{V}(X)}\sqrt{\mathbb{V}(Y)}} = 1$$

**Moments d'une variable aléatoire de loi uniforme sur  $[a, b]$**

Soit  $X \sim \mathcal{U}_{[a, b]}$ . Son espérance vaut

$$\mathbb{E}(X) = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$$

et puisque

$$\mathbb{E}(X^2) = \int_a^b \frac{x^2}{b-a} dx = \frac{a^2 + ab + b^2}{3},$$

alors sa variance vaut

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{(b-a)^2}{12}.$$

**Moments d'une v.a. de loi exponentielle** Soit  $X \sim \mathcal{E}(\theta)$ , alors

$$\mathbb{E}(X) = \int_0^{+\infty} x\theta \exp(-\theta x) dx = \left[ \frac{x}{\theta} \exp(-\theta x) \right]_0^{+\infty} + \int_0^{+\infty} \exp(-\theta x) dx = \frac{1}{\theta}$$

**Moments d'une v.a. de loi gamma** Soit  $X \sim \Gamma(\alpha, \beta)$ . On a

$$\begin{aligned} \mathbb{E}(X) &= \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} \theta^\alpha x^\alpha e^{-\theta x} dx \\ &= \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} \frac{1}{\theta} u^\alpha e^{-u} du \text{ par le changement de variable } u = \theta x \\ &= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \frac{1}{\theta} = \frac{\alpha}{\theta} \end{aligned}$$

puis

$$\begin{aligned} \mathbb{E}(X^2) &= \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} \theta^\alpha x^{\alpha+1} e^{-\theta x} dx \\ &= \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} \frac{1}{\theta^2} u^{\alpha+1} e^{-u} du \text{ par le changement de variable } u = \theta x \\ &= \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \frac{1}{\theta^2} = \frac{\alpha(\alpha+1)}{\theta^2} \end{aligned}$$

d'où

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{\alpha}{\theta^2}$$

**Moments d'une v.a. gaussienne** Pour s'assurer qu'il s'agit bien d'une densité, on remarque que  $I = \int_{-\infty}^{+\infty} f(x)dx$  vérifie

$$I^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x)f(y)dx dy = \int_0^{2\pi} d\theta \int_0^{+\infty} \frac{1}{2\pi} r e^{-r^2/2} dr$$

(en passant en coordonnées polaires dans l'intégrale double). Le calcul est ensuite aisé et on obtient  $I = 1$ .

Pour les moments, on fait d'abord le calcul dans le cas centré réduit ( $\mu = 0$  et  $\sigma^2 = 1$ ) puis on s'y ramène par le changement de variable  $x \mapsto \frac{x-\mu}{\sigma}$ . Soit donc  $X \sim \mathcal{N}(0, 1)$ , on a

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{+\infty} \frac{x}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= \left[ -\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \right]_{-\infty}^{+\infty} = 0. \end{aligned}$$

et

$$\begin{aligned} \mathbb{V}(X) = \mathbb{E}(X^2) &= \int_{-\infty}^{+\infty} \frac{x^2}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= \left[ -\frac{x}{\sqrt{2\pi}} e^{-x^2/2} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= 0 + 1. \end{aligned}$$

**Loi de Cauchy** L'angle  $\theta$  est une variable aléatoire uniforme sur  $[-\pi/2, \pi/2]$ , de densité  $g(\theta) = \frac{1}{\pi} 1_{[-\pi/2, \pi/2]}(\theta)$ . L'abscisse  $X$  est donnée par  $X = \tan \theta$ , c'est donc une variable aléatoire, de fonction de répartition

$$\begin{aligned} F(x) &= \mathbb{P}(X \leq x) \\ &= \mathbb{P}(\theta \leq \arctan x) \\ &= \int_{-\infty}^{\arctan x} \frac{1}{\pi} 1_{[-\pi/2, \pi/2]}(\theta) d\theta \\ &= \frac{1}{\pi} \arctan x + \frac{1}{2}. \end{aligned}$$

$F$  est de classe  $C^1$  de dérivée

$$f(x) = \frac{1}{\pi(1+x^2)}, x \in \mathbb{R}$$

C'est la densité de la loi de Cauchy. Une variable aléatoire  $X$  de loi de Cauchy n'admet pas d'espérance. En effet,  $\frac{x}{\pi(1+x^2)} \sim_{x \rightarrow \infty} \frac{1}{x}$  n'est pas intégrable.

**Calcul de l'aire d'un triangle** Si la fonction  $1_D$  est intégrable, alors le théorème de Fubini est applicable (p. ??). On a donc

$$a(T) = \int_{\mathbb{R}} \left[ \int_{\mathbb{R}} 1_T(x, y) dx \right] dy.$$

Or, si  $0 \leq y \leq 1$ ,

$$1_T(x, y) = \begin{cases} 1 & \text{si } 0 \leq x \leq 1 - y, \\ 0 & \text{sinon.} \end{cases}$$

et dans le cas contraire,  $1_T(x, y) = 0$ . On a donc

$$a(T) = \int_0^1 \left[ \int_0^{1-y} dx \right] dy = \int_0^1 (1 - y) dy = \left[ y - \frac{y^2}{2} \right]_0^1 = \frac{1}{2}.$$

**Contre-exemple** Pour tout  $y \in ]0, 1]$  (et donc presque tout  $y \in [0, 1]$ ), on a

$$\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx = \left[ x \mapsto -\frac{x}{x^2 + y^2} \right]_0^1 = -\frac{1}{1 + y^2},$$

par conséquent

$$\int_0^1 \left[ \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx \right] dy = - \int_0^1 \frac{dy}{1 + y^2} = -[y \mapsto \arctan y]_0^1 = -\frac{\pi}{4}.$$

En exploitant la relation

$$\frac{x^2 - y^2}{(x^2 + y^2)^2} = \frac{\partial}{\partial y} \left( \frac{y}{x^2 + y^2} \right)$$

on établit de façon similaire que

$$\int_0^1 \left[ \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dy \right] dx = \frac{\pi}{4}.$$

Or, si le théorème de Fubini (p. 11) était applicable, on pourrait intervertir l'ordre d'intégration des variables sans changer le résultat. Comme cela n'est pas le cas, on en déduit que l'hypothèse exigée par le théorème de Fubini ne tient pas : la fonction

$$(x, y) \in [0, 1] \times [0, 1] \mapsto \frac{x^2 - y^2}{(x^2 + y^2)^2}$$

n'est pas intégrable.

**Homothétie** Si l'on pose  $D_1 = \mathbb{R}^n$ ,  $D_2 = \mathbb{R}^n$ , l'application  $h : D_1 \rightarrow D_2$  définie par  $h(x) = \alpha x$  est un  $C^1$ -difféomorphisme. De plus, on a

$$|\det J_h(x)| = \left| \det \begin{bmatrix} \alpha & 0 & \cdots & 0 \\ 0 & \alpha & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha \end{bmatrix} \right| = |\alpha^n| = \alpha^n.$$

Par conséquent, le théorème de changement de variables (p. 12) fournit

$$\int_{\mathbb{R}^n} f(y) dy = \int_{\mathbb{R}^n} f(\alpha x) \alpha^n dx,$$

soit

$$\int_{\mathbb{R}^n} f(\alpha x) dx = \frac{1}{\alpha^n} \int_{\mathbb{R}^n} f(y) dy.$$

**Volume et translation** L'ensemble  $A$  mesurable et de volume fini a une fonction caractéristique  $1_A$  intégrable et

$$\lambda(A) = \int 1_A(x) dx.$$

Soit  $h(x) = x - u$  ou  $u \in \mathbb{R}^3$ . La fonction  $h$  est une bijection de  $\mathbb{R}^3$  sur lui-même, continûment différentiable ainsi que son inverse,  $h^{-1}(x) = x + u$  et  $J_h(x) = I$ , donc  $\det J_h(x) = 1$ . Par le théorème de changement de variables (p. 12), la fonction  $1_{h^{-1}(A)} = 1_A \circ h$  est donc intégrable sur  $\mathbb{R}^3$  et

$$\int 1_{h^{-1}(A)} dx = \int_{\mathbb{R}^3} (1_A \circ h) |\det J_h(x)| dx = \int_{\mathbb{R}^3} 1_A(x) dx = \lambda(A).$$

L'ensemble translaté  $A + u = h^{-1}(A)$  est donc mesurable, de volume fini égal au volume de  $A$ .

**Coordonnées polaires** Les ensembles  $P$  et  $C$  sont ouverts et la fonction  $h$  – qui permet de passer des coordonnées polaires aux coordonnées cartésiennes – est une bijection de  $P$  dans  $C$ . Elle est continûment différentiable ; sa matrice jacobienne en  $(r, \theta)$  vaut

$$J_h(r, \theta) = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix},$$

et son déterminant satisfait

$$\det J_h(r, \theta) = (\cos \theta)(r \cos \theta) - (\sin \theta)(-r \sin \theta) = r > 0.$$

Le jacobien est donc inversible et  $h^{-1}$  est continûment différentiable par le théorème d'inversion locale. On peut donc appliquer le théorème de changement de variables (p. 12) à la fonction  $f$ , ce qui fournit

$$\begin{aligned} \int_C f(x, y) d(x, y) &= \int_P f(h(r, \theta)) |\det J_h(r, \theta)| d(r, \theta) \\ &= \int_P g(r, \theta) r d(r, \theta). \end{aligned}$$

**Absence du déterminant jacobien** La façon la plus rapide de procéder consiste à considérer que  $f \circ h$  joue le rôle de  $f$  dans le théorème de changement de variables (p. 12), que  $h$  dans notre énoncé désigne  $h^{-1}$  dans ce théorème et que les rôles de  $D_1$  et  $D_2$  sont intervertis. Une fois que l'on a permuté ces notations, on réalise que l'intégrale que l'on souhaite calculer est le membre de gauche de l'équation du théorème de changement de variables, dont toutes les hypothèses sont par ailleurs satisfaites. Par conséquent on a

$$\int_{D_1} f(h(x)) dx = \int_{D_2} (f \circ h)(h^{-1}(y)) |\det J_{h^{-1}}(y)| dy.$$

On peut simplifier  $(f \circ h)(h^{-1}(y))$  en  $f(y)$  et éventuellement exprimer le jacobien de  $h^{-1}$  en fonction du jacobien de  $h : J_{h^{-1}}(y) = [J_h(h^{-1}(y))]^{-1}$  ; on obtient donc

$$\int_{D_1} f(h(x)) dx = \int_{D_2} f(y) |\det J_{h^{-1}}(y)| dy = \int_{D_2} f(y) \frac{1}{|\det J_h(h^{-1}(y))|} dy.$$

**Lancer de fléchette** Les coordonnées cartésiennes de  $M \in D = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}$  constituent un couple de variables aléatoires de densité

$$f_{(X,Y)}(x, y) = \frac{1}{\pi} 1_D(x, y)$$

uniforme sur le disque, par hypothèse. L'abscisse de  $X$  est distribuée selon la densité marginale

$$f_X(x) = \int f_{(X,Y)}(x, y) dy = \frac{2}{\pi} (1 - x^2)^{1/2} 1_{[-1,1]}(x).$$

La loi de  $Y$  a la même densité.

**Lancer de fléchette (suite)** La densité des coordonnées cartésiennes ne s'exprime pas comme le produit des densités marginales; elles ne sont pas indépendantes.

**Démonstration** Soit  $X$  et  $Y$  deux v.a. aléatoires indépendantes et de carré intégrable. On a d'après la proposition (p. 18)

$$\mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(X - \mathbb{E}(X))\mathbb{E}(Y - \mathbb{E}(Y)) = 0 = \rho(X, Y)$$

**Lancer de fléchette (fin)** En coordonnées polaires, le domaine  $D$  est décrit par  $D = \{(r, \theta) \in [0, 1] \times [0, 2\pi]\}$ . La densité de  $(R, \theta)$  s'écrit ainsi :

$$f_{(R,\theta)}(r, \theta) = \frac{r}{\pi} 1_D(r \cos(\theta), r \sin(\theta)) = 2r 1_{[0,1]}(r) \frac{1}{2\pi} 1_{[0,2\pi]}(\theta).$$

Ainsi  $R$  et  $\theta$  sont indépendantes de densités respectives  $2r 1_{[0,1]}(r)$  et  $\frac{1}{2\pi} 1_{[0,2\pi]}(\theta)$ .

**Loi bêta** On note d'abord que la dimension de  $Y$  est plus petite que celle de  $X$ . On va donc compléter  $Y$  en prenant par exemple  $Y' = (Y, Z)$ , avec  $Z = U + V$ , ce qui correspond à  $g(u, v) = \left(\frac{u}{u+v}, u+v\right)$ . Cette application est bijective de  $A = ]0, +\infty[^2$  dans  $B = ]0, 1[ \times ]0, +\infty[$ , d'inverse  $g^{-1}(y, z) = (yz, z(1-y))$ , qui a pour jacobien  $z$ .

Comme  $f_X(u, v) = \frac{\theta^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} u^{\alpha-1} v^{\beta-1} e^{-\theta(u+v)} 1_A(u, v)$ , on a

$$f_{Y'}(y, z) = \frac{\theta^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} z^{\alpha+\beta-1} y^{\alpha-1} (1-y)^{\beta-1} e^{-\theta z} 1_B(y, z).$$

On obtient alors la densité de  $Y$  en intégrant  $f_{Y'}(y, z)$  par rapport à  $z \in ]0, +\infty[$  :

$$\begin{aligned} f_Y(y) &= \int_0^{+\infty} f_{Y'}(y, z) dz \\ &= \frac{\theta^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1} (1-y)^{\beta-1} 1_{]0,1[}(y) \int_0^{+\infty} z^{\alpha+\beta-1} e^{-\theta z} dz \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1} (1-y)^{\beta-1} 1_{]0,1[}(y), \end{aligned}$$



où on a utilisé la définition de la fonction gamma et le changement de variable linéaire  $z \mapsto \theta z$ . On appelle loi bêta de paramètres  $\alpha$  et  $\beta$  la loi admettant cette densité. Admettant une grande variété de formes, elle permet de modéliser de nombreuses distributions à support fini.

La loi bêta apparaît naturellement dans une expérience d'urnes, donnée par George Pólya dans un article de 1930, *Sur quelques points de la théorie des probabilités*. Il décrit l'expérience suivante : on se donne une urne contenant initialement  $r$  boules rouges et  $b$  boules bleues, on tire une boule dans l'urne, puis on la remet dans l'urne avec une deuxième boule de même couleur. Alors la proportion de boules rouges tend vers une variable aléatoire de loi  $\text{Beta}(r, b)$ , et, inversement, la proportion de boules bleues tend vers une variable aléatoire de loi  $\text{Beta}(b, r)$ .

Nous obtenons aussi facilement la densité de  $Z$ . En effet, on a  $f_{Y'}(y, z) = f_Y(y)f_Z(z)$  ( $Y$  et  $Z$  sont donc indépendantes), où

$$f_Z(z) = \frac{\theta^{\alpha+\beta}}{\Gamma(\alpha+\beta)} z^{\alpha+\beta-1} e^{-\theta z} 1_{]0, +\infty[}$$

On a ainsi démontré que si  $U$  et  $V$  sont deux variables aléatoires indépendantes de lois respectives  $\Gamma(\alpha, \theta)$  et  $\Gamma(\beta, \theta)$ , alors  $U + V$  suit la loi  $\Gamma(\alpha + \beta, \theta)$  et est indépendante de  $\frac{U}{U+V}$  qui suit une loi bêta de paramètres  $(\alpha, \beta)$ .

**Divergence de Kullback-Leibler entre deux gaussiennes** On a

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

et

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}$$

On en déduit que

$$\begin{aligned} D_{KL}(f||g) &= \int_{-\infty}^{+\infty} f(x) \log \left( \frac{f(x)}{g(x)} \right) dx \\ &= \int_{-\infty}^{+\infty} f(x) \left[ \log \left( \frac{\sigma_2}{\sigma_1} \right) + \frac{(x-\mu_2)^2}{2\sigma_2^2} - \frac{(x-\mu_1)^2}{2\sigma_1^2} \right] dx \\ &= \log \left( \frac{\sigma_2}{\sigma_1} \right) + \frac{1}{2\sigma_2^2} \int_{-\infty}^{+\infty} f(x)(x-\mu_2)^2 dx - \frac{1}{2\sigma_1^2} \int_{-\infty}^{+\infty} f(x)(x-\mu_1)^2 dx \end{aligned}$$

On a :

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x)(x-\mu_2)^2 dx &= \int_{-\infty}^{+\infty} f(x)((x-\mu_1) + \mu_1 - \mu_2)^2 dx \\ &= \int_{-\infty}^{+\infty} f(x)(x-\mu_1)^2 dx + (\mu_1 - \mu_2)^2 \\ &= \sigma_1^2 + (\mu_1 - \mu_2)^2 \end{aligned}$$

On en déduit la formule annoncée :

$$D_{KL}(f||g) = \log \left( \frac{\sigma_2}{\sigma_1} \right) + \frac{\sigma_1^2 + (\mu_1 - \mu_2)^2}{2\sigma_2^2} - \frac{1}{2}$$

## Durée de vie

**Question 1** En dérivant, on obtient que la densité de  $T$  est  $f(t) = te^{-t^2/2}1_{t>0}$ .

L'espérance vaut :

$$\begin{aligned}\mathbb{E}(T) &= \int_0^{+\infty} t^2 e^{-t^2/2} dt \\ &= \left[ -te^{-t^2/2} \right]_0^{+\infty} + \int_0^{+\infty} e^{-t^2/2} dt \text{ par intégration par parties} \\ &= \frac{1}{2} \int_{-\infty}^{+\infty} e^{-t^2/2} dt \\ &= \frac{\sqrt{2\pi}}{2}\end{aligned}$$

**Question 2** La probabilité s'écrit :

$$\mathbb{P}(T \geq 3 | T \geq 1) = \frac{\mathbb{P}(T \geq 3, T \geq 1)}{\mathbb{P}(T \geq 1)} = \frac{\mathbb{P}(T \geq 3)}{\mathbb{P}(T \geq 1)} = \frac{e^{-9/2}}{e^{-1/2}} = e^{-4}$$

**Question 3** Les variables aléatoires  $X_1, \dots, X_{10}$  sont indépendantes et suivent une loi de Bernoulli de paramètre :

$$\mathbb{P}(T \leq 1) = F(1) = 1 - e^{-1/2}$$

On en déduit que la loi de  $N$  est la loi binomiale de paramètre  $(10, 1 - e^{-1/2})$ .

**Question 4** La probabilité que l'équipement en série soit défaillant avant 1 an vaut :

$$P(N \geq 1) = 1 - \mathbb{P}(N = 0) = 1 - e^{-10/2} \approx 9.9 \cdot 10^{-1} = 99\%$$

**Question 5** La probabilité que l'équipement en parallèle soit défaillant avant 1 an vaut :

$$\mathbb{P}(N = 10) = (1 - e^{-1/2})^{10} \approx 8.9 \cdot 10^{-5}$$

## Crues centennales de la Seine

**Question 1** Soit  $x \in \mathbb{R}$

$$\begin{aligned}F_Y(x) &= F(Y \leq x) = F(X_1 \leq x, \dots, X_n \leq x) \\ &= \prod_{i=1}^n F(X_i \leq x) \\ &= F(x)^n\end{aligned}$$

puisque les  $X_i$  sont indépendantes et de même loi

**Question 2** Soit  $x \in \mathbb{R}$

$$\begin{aligned} f_Y(x) &= \frac{d}{dx} F_Y(x) \\ &= \frac{d}{dx} F(x)^n \\ &= n f(x) F(x)^{n-1} \end{aligned}$$

**Question 3** On a alors  $f(x) = 1_{[0,1]}(x)$  et  $F(x) = x 1_{[0,1]}(x) + 1_{[1,+\infty]}$ . On en déduit

$$f_Y(x) = n x^{n-1} 1_{[0,1]}$$

**Question 4** On cherche à déterminer  $x$  tel que  $\mathbb{P}(Y > x) \leq \frac{1}{1000}$ . Or,

$$\mathbb{P}(Y > x) = 1 - \mathbb{P}(Y \leq x) = 1 - F_Y(x) = 1 - F(x)^n = 1 - x^n$$

Ainsi

$$1 - x^n \leq \frac{1}{1000} \Leftrightarrow (1 - \frac{1}{1000})^{1/n} \leq x$$

**Question 5** On rappelle que la médiane est le réel  $m$  tel que  $\frac{1}{2} = \mathbb{P}(Y \leq m) = F_Y(m)$ .

On cherche donc  $m$  tel que  $m^n = \frac{1}{2}$ , soit  $m = \frac{1}{2^{1/n}}$ .

On voit que cette valeur est plus élevée que la médiane des  $X_i$  qui vaut  $\frac{1}{2}$ . Elle tend même vers 1 lorsque  $n$  tend vers l'infini.

## Indépendance et vecteurs gaussiens

**Question 1** D'après le cours, en interprétant les éléments de  $\mathbb{R}^2$  comme des vecteurs colonnes, pour tout  $z = (x, y) \in \mathbb{R}^2$  on a

$$f_Z(z) = \frac{1}{2\pi \sqrt{\det(C)}} \exp \left\{ -\frac{1}{2} (z - m)^t C^{-1} (z - m) \right\}.$$

Or par construction  $C$  peut s'écrire

$$C = \begin{pmatrix} \sigma_X^2 & c \\ c & \sigma_Y^2 \end{pmatrix},$$

où  $\sigma_X^2 := \mathbb{V}(X)$ ,  $\sigma_Y^2 := \mathbb{V}(Y)$  et  $c = \text{Cov}(X, Y)$ . Comme  $C$  est définie positive elle est inversible, avec  $\det(C) = \sigma_X^2 \sigma_Y^2 - c^2 > 0$  et

$$C^{-1} = \frac{1}{\sigma_X^2 \sigma_Y^2 - c^2} \begin{pmatrix} \sigma_Y^2 & -c \\ -c & \sigma_X^2 \end{pmatrix}.$$

En outre,  $m = (m_X, m_Y)$  avec  $m_X := \mathbb{E}(X)$  et  $m_Y := \mathbb{E}(Y)$ . On en déduit que

$$\begin{aligned} f_Z(x, y) &= \frac{1}{2\pi \sqrt{\sigma_X^2 \sigma_Y^2 - c^2}} \exp \left\{ \frac{-1}{2 (\sigma_X^2 \sigma_Y^2 - c^2)} \times \right. \\ &\quad \left. (\sigma_Y^2 (x - m_X)^2 - 2c(x - m_X)(y - m_Y) + \sigma_X^2 (y - m_Y)^2) \right\} \\ &= \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1 - \rho^2}} \exp \left\{ \frac{-1}{2 (1 - \rho^2)} \times \right. \\ &\quad \left. \left( \frac{(x - m_X)^2}{\sigma_X^2} - 2\rho \frac{(x - m_X)}{\sigma_X} \frac{(y - m_Y)}{\sigma_Y} + \frac{(y - m_Y)^2}{\sigma_Y^2} \right) \right\}. \end{aligned}$$

**Question 2** Ici, pour  $x, y \in \mathbb{R}^2$ , on remarque que :

$$\begin{aligned} f_Z(x, y) &= \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1 - \rho^2}} \exp \left\{ \frac{-1}{2 (1 - \rho^2)} \times \right. \\ &\quad \left. \left( \left( \frac{y - m_Y}{\sigma_Y} - \rho \frac{x - m_X}{\sigma_X} \right)^2 + (1 - \rho^2) \frac{(x - m_X)^2}{\sigma_X^2} \right) \right\} \\ &= \frac{1}{\sqrt{2\pi} \sigma_X} \exp \left\{ -\frac{(x - m_X)^2}{2 \sigma_X^2} \right\} \times \\ &\quad \frac{1}{\sqrt{2\pi} \sigma_Y \sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2 (1 - \rho^2)} \left( \frac{y - m_Y}{\sigma_Y} - \rho \frac{x - m_X}{\sigma_X} \right)^2 \right\}. \end{aligned}$$

Ainsi, avec le changement de variable  $u = \frac{y - m_Y}{\sigma_Y}$  on obtient

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f_Z(x, y) dy \\ &= \frac{1}{\sqrt{2\pi} \sigma_X} \exp \left\{ -\frac{(x - m_X)^2}{2 \sigma_X^2} \right\} \times \\ &\quad \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi} \sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2 (1 - \rho^2)} \left( u - \rho \frac{x - m_X}{\sigma_X} \right)^2 \right\} du, \end{aligned}$$

et on reconnaît dans cette dernière intégrale la densité d'une loi Normale d'espérance  $\rho \frac{x - m_X}{\sigma_X}$  et de variance  $1 - \rho^2$ . Cette intégrale vaut donc 1 et on conclut que

$$f_X(x) = \frac{1}{\sqrt{2\pi} \sigma_X} \exp \left\{ -\frac{(x - m_X)^2}{2 \sigma_X^2} \right\},$$

qui correspond à la densité d'une loi Normale d'espérance  $m_X$  et de variance  $\sigma_X^2$ .

En procédant de manière symétrique, on obtient de même que  $Y$  suit une loi Normale d'espérance  $m_Y$  et de variance  $\sigma_Y^2$ .

**Question 3** Le premier sens est évident : si  $X$  et  $Y$  sont indépendantes, alors nous avons vu dans le cours que  $\text{Cov}(X, Y) = 0$ , et ce que l'on soit gaussien ou non.

Supposons maintenant que  $\text{Cov}(X, Y) = 0$ , et montrons l'indépendance entre  $X$  et  $Y$ . En reprenant la formule de la question 1 et en remplaçant  $\rho$  par 0, pour tout  $(x, y) \in \mathbb{R}^2$  on a :

$$\begin{aligned} f_Z(x, y) &= \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left\{-\frac{1}{2}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right)\right\} \\ &= \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left\{-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right\} \times \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left\{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right\} \\ &= f_X(x) \times f_Y(y) \end{aligned}$$

d'après la question 2. On en conclut que  $X$  et  $Y$  sont bien indépendantes.

## Symétrie de la gaussienne

**Question 1** Nous avons admis dans le cours que la fonction de répartition  $F_c$  de  $X_c$  caractérisait sa loi. Nous allons donc l'expliciter. Soit  $x \in \mathbb{R}$ . Par définition de la fonction de répartition d'une variable aléatoire et d'après la formule des probabilités totales, on a

$$\begin{aligned} F_c(x) &= \mathbb{P}(X_c \leq x) = \mathbb{P}(X_c \leq x, |X| > c) + \mathbb{P}(X_c \leq x, |X| \leq c) \\ &= \mathbb{P}(X \leq x, |X| > c) + \mathbb{P}(-X \leq x, |X| \leq c) \\ &= \mathbb{P}(X \leq x, X > c) + \mathbb{P}(X \leq x, X < -c) + \mathbb{P}(X \geq -x, -c \leq X \leq c). \end{aligned}$$

Or

$$\begin{aligned} \mathbb{P}(X \leq x, X > c) &= \begin{cases} \mathbb{P}(c < X \leq x) = F(x) - F(c) & \text{si } x > c, \\ 0 & \text{si } x \leq c, \end{cases} \\ \mathbb{P}(X \leq x, X < -c) &= \mathbb{P}(X \leq \min\{-c, x\}) = F(\min\{-c, x\}), \\ \mathbb{P}(X \geq -x, -c \leq X \leq c) &= \begin{cases} \mathbb{P}(-c \leq X \leq c) = F(c) - F(-c) & \text{si } x > c, \\ \mathbb{P}(-x \leq X \leq c) = F(c) - F(-x) & \text{si } -c \leq x \leq c, \\ 0 & \text{si } x < -c. \end{cases} \end{aligned}$$

Ainsi,

— si  $x > c$  alors  $\min\{-c, x\} = -c$  et

$$F_c(x) = F(x) - F(c) + F(\min\{-c, x\}) + F(c) - F(-c) = F(x),$$

— si  $-c \leq x \leq c$  alors  $\min\{-c, x\} = -c$  et

$$F_c(x) = F(\min\{-c, x\}) + F(c) - F(-x) = 1 - F(c) + F(c) - 1 + F(x) = F(x)$$

car la densité de la loi Normale centrée réduite est paire, ce qui implique que  $F(-x) = 1 - F(x)$  pour tout  $x \in \mathbb{R}$ ,

— si  $x < -c$  alors  $\min\{-c, x\} = x$  et

$$F_c(x) = F(\min\{-c, x\}) = F(x).$$

On en conclut que  $F_c = F : X_c$  a la même loi que  $X$ . Ce résultat est dû à la symétrie de la densité de la loi Normale centrée réduite. On remarque d'ailleurs que le résultat resterait inchangé si l'on prenait n'importe quelle autre loi de densité paire !

**Question 2** Tout d'abord, par définition,

$$\text{Cov}(X, X_c) = \mathbb{E}((X - \mathbb{E}(X))(X_c - \mathbb{E}(X_c))) = \mathbb{E}(X X_c)$$

puisque  $X$  et  $X_c$  sont gaussiennes centrées (d'espérance nulle) réduites. Or on remarque que l'on peut écrire  $X_c = X 1_{\mathbb{R} \setminus [-c, c]}(X) - X 1_{[-c, c]}(X)$ . On déduit par linéarité de l'espérance que

$$\begin{aligned} \mathbb{E}(X X_c) &= \mathbb{E}(X^2 1_{\mathbb{R} \setminus [-c, c]}(X) - X^2 1_{[-c, c]}(X)) \\ &= \mathbb{E}(X^2 (1_{\mathbb{R} \setminus [-c, c]}(X) + 1_{[-c, c]}(X) - 1_{[-c, c]}(X))) - \mathbb{E}(X^2 1_{[-c, c]}(X)) \\ &= \mathbb{E}(X^2) - 2 \mathbb{E}(X^2 1_{[-c, c]}(X)). \end{aligned}$$

Etant donné que  $X$  est de loi Normale centrée réduite, on a directement que  $\mathbb{E}(X^2) = 1$ ; il suffit donc de calculer la dernière espérance pour obtenir le résultat. Comme  $x \mapsto x^2 1_{[-c, c]}(x)$  est bornée et continue sur  $\mathbb{R} \setminus \{-c, c\}$  (donc continue presque partout), elle est intégrable et

$$\mathbb{E}(X^2 1_{[-c, c]}(X)) = \int_{-\infty}^{+\infty} x^2 1_{[-c, c]} f(x) dx = \int_{-c}^c x^2 f(x) dx.$$

Or on pourra remarquer que  $f'(x) = -x f(x)$  puis que  $f''(x) = (x^2 - 1) f(x)$ . Puisque  $f$ ,  $f'$  et  $f''$  sont continues et bornées (voir l'exercice *Densité et fonction de répartition d'une loi Normale* du chapitre Probabilités I) elles sont intégrables sur des segments. On en déduit que

$$\begin{aligned} \mathbb{E}(X^2 1_{[-c, c]}(X)) &= \int_{-c}^c (x^2 - 1) f(x) dx + \int_{-c}^c f(x) dx \\ &= [-x f(x)]_{-c}^c + F(c) - F(-c) \\ &= -c (f(c) + f(-c)) + F(c) - F(-c) \\ &= 2 (F(c) - c f(c)) - 1 \end{aligned}$$

par parité de  $f$ . On en conclut que  $\text{Cov}(X, X_c) = 3 + 4 (c f(c) - F(c))$ .

La fonction  $g : c \in \mathbb{R}_+^* \mapsto 3 + 4 (c f(c) - F(c))$  est continue, infiniment dérivable et pour tout  $c > 0$  on a  $g'(c) = -4 c^2 f(c) < 0$ . Par conséquent,  $g$  est strictement décroissante, avec  $\lim_{c \rightarrow 0^+} g(c) = 1$  et  $\lim_{c \rightarrow +\infty} g(c) = -1$ . Le théorème des valeurs intermédiaires nous permet de conclure qu'il existe bien un unique  $c_0 > 0$  tel que  $g(c_0) = 0$ .

**Question 3** Les variables aléatoires  $X$  et  $X_{c_0}$  sont indépendantes ssi pour tous ensembles  $A, B \in \mathcal{E}_{\mathbb{R}}$  on a

$$\mathbb{P}(X \in A, X_{c_0} \in B) = \mathbb{P}(X \in A) \mathbb{P}(X_{c_0} \in B).$$

Or si l'on prend  $A = B = ] - \infty, x]$  où  $x < -c_0$ , alors

$$\begin{aligned} \mathbb{P}(X \leq x, X_{c_0} \leq x) &= \mathbb{P}(X \leq x, X < -c_0) + \mathbb{P}(X \leq x, X > c_0) \\ &\quad + \mathbb{P}(X \leq x, -c_0 \leq X \leq c_0) \\ &= \mathbb{P}(X \leq x) + 0 + 0 = F(x) \\ &\neq F(x) F(x) = \mathbb{P}(X \leq x) \mathbb{P}(X_{c_0} \leq x). \end{aligned}$$

On en conclut que bien que  $\text{Cov}(X, X_{c_0}) = 0$ ,  $X$  et  $X_{c_0}$  ne sont pas indépendantes.

**Question 4** Supposons que  $(X, X_{c_0})$  soit un vecteur gaussien. Alors, puisque  $\text{Cov}(X, X_{c_0}) = 0$ ,  $X$  et  $X_{c_0}$  sont nécessairement indépendantes. Or nous venons de voir que ce n'était pas le cas; nous avons donc contradiction. On en conclut que  $(X, X_{c_0})$  n'est pas gaussien.

## Loi du $\chi^2$

### A 1 degré de liberté

1. Soit  $h : \mathbb{R} \rightarrow \mathbb{R}$  une fonction mesurable telle que  $h \times f$  est intégrable. Alors en opérant le changement de variable  $u = x^2$  on a

$$\begin{aligned}\mathbb{E}(h(Y)) &= \mathbb{E}(h(X^2)) = \int_{\mathbb{R}} h(x^2) f(x) dx \\ &= \int_{-\infty}^0 h(x^2) f(x) dx + \int_0^{+\infty} h(x^2) f(x) dx \\ &= \int_0^{+\infty} h(u) f(-\sqrt{u}) \frac{1}{2\sqrt{u}} du + \int_0^{+\infty} h(u) f(\sqrt{u}) \frac{1}{2\sqrt{u}} du.\end{aligned}$$

Or  $f$  est paire, donc pour tout  $u \in \mathbb{R}^+$  on a  $f(-\sqrt{u}) = f(\sqrt{u})$  et

$$\mathbb{E}(h(Y)) = \int_{\mathbb{R}_+} h(u) f(\sqrt{u}) \frac{1}{\sqrt{u}} du.$$

Par propriété  $Y$  possède donc une densité  $f_Y$  valant pour tout  $x \in \mathbb{R}$

$$f_Y(x) = \begin{cases} f(\sqrt{x}) \frac{1}{\sqrt{x}} = \frac{x^{-\frac{1}{2}}}{2^{\frac{1}{2}} \Gamma(\frac{1}{2})} \exp\left\{-\frac{x}{2}\right\} & \text{si } x > 0, \\ 0 & \text{si } x \leq 0. \end{cases}$$

**Remarque.** La fonction  $f_Y$  est bien une densité : elle est continue par morceaux donc mesurable, puis positive, et son intégrale sur  $\mathbb{R}$  vaut bien 1 :

$$\begin{aligned}\int_{\mathbb{R}} f_Y(x) dx &= \frac{1}{2\Gamma(\frac{1}{2})} \int_0^{+\infty} \left(\frac{x}{2}\right)^{\frac{1}{2}-1} \exp\left\{-\frac{x}{2}\right\} dx \\ &= \frac{1}{\Gamma(\frac{1}{2})} \int_0^{+\infty} u^{\frac{1}{2}-1} e^{-u} du = \frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2})} = 1.\end{aligned}$$

### A n degrés de liberté

2. On note  $(\mathcal{P}_n)$  la propriété à démontrer au rang  $n$ .

**Initialisation.** Nous avons vu à la question précédente que  $(\mathcal{P}_1)$  est vraie.

**Hérédité.** Soit maintenant  $n \geq 2$  et supposons  $(\mathcal{P}_{n-1})$  vraie. Alors  $Y = \sum_{i=1}^{n-1} Y_i + Y_n$  est la somme de deux variables aléatoires indépendantes, dont la première suit un  $\chi_{n-1}^2$  de densité notée  $f_{n-1}$  et la seconde un  $\chi_1^2$  de densité notée  $f_1$ . Elle possède donc une densité, égale au produit de convolution de  $f_{n-1}$  et  $f_1$  : pour tout  $x \in \mathbb{R}$

$$f_Y(x) = \int_{\mathbb{R}} f_1(x-u) f_{n-1}(u) du.$$

Or pour tout  $x, u \in \mathbb{R}$  on a  $f_1(x - u) = 0$  ssi  $x \leq u$  et  $f_{n-1}(u) = 0$  ssi  $u \leq 0$ . Donc leur produit est strictement positif ssi  $x > u > 0$ , nul sinon. La densité  $f_Y$  est par conséquent nulle sur  $\mathbb{R}^-$  et nous allons maintenant exhiber son expression sur  $\mathbb{R}_+^*$ . Soit  $x > 0$ , alors

$$\begin{aligned} f_Y(x) &= \int_0^x \frac{(x-u)^{-\frac{1}{2}}}{\sqrt{2}\Gamma(\frac{1}{2})} \exp\left\{-\frac{x-u}{2}\right\} \frac{u^{\frac{n-1}{2}-1}}{2^{\frac{n-1}{2}}\Gamma(\frac{n-1}{2})} \exp\left\{-\frac{u}{2}\right\} du \\ &= \frac{\exp\left\{-\frac{x}{2}\right\}}{2^{\frac{n}{2}}\Gamma(\frac{1}{2})\Gamma(\frac{n-1}{2})} \int_0^x (x-u)^{-\frac{1}{2}} u^{\frac{n-3}{2}} du. \end{aligned}$$

Avec le changement de variable  $v = \frac{u}{x}$  on obtient

$$f_Y(x) = \frac{x^{\frac{n}{2}-1} \exp\left\{-\frac{x}{2}\right\}}{2^{\frac{n}{2}}\Gamma(\frac{1}{2})\Gamma(\frac{n-1}{2})} \int_0^1 v^{\frac{n-3}{2}} (1-v)^{-\frac{1}{2}} dv.$$

Or  $f_Y$  est une densité, son intégrale vaut donc 1 :

$$\begin{aligned} \int_{\mathbb{R}} f_Y(x) dx &= \int_0^{+\infty} \frac{x^{\frac{n}{2}-1} \exp\left\{-\frac{x}{2}\right\}}{2^{\frac{n}{2}}\Gamma(\frac{1}{2})\Gamma(\frac{n-1}{2})} \left( \int_0^1 v^{\frac{n-3}{2}} (1-v)^{-\frac{1}{2}} dv \right) dx \\ &= \left( \int_0^1 v^{\frac{n-3}{2}} (1-v)^{-\frac{1}{2}} dv \right) \left( \int_0^{+\infty} \frac{x^{\frac{n}{2}-1} \exp\left\{-\frac{x}{2}\right\}}{2^{\frac{n}{2}}\Gamma(\frac{1}{2})\Gamma(\frac{n-1}{2})} dx \right) \\ &= \left( \int_0^1 v^{\frac{n-3}{2}} (1-v)^{-\frac{1}{2}} dv \right) \left( \int_0^{+\infty} \frac{x^{\frac{n}{2}-1} \exp\left\{-\frac{x}{2}\right\}}{\Gamma(\frac{1}{2})\Gamma(\frac{n-1}{2})} dx \right) \\ &= \left( \int_0^1 v^{\frac{n-3}{2}} (1-v)^{-\frac{1}{2}} dv \right) \left( \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{n-1}{2})} \right) = 1. \end{aligned}$$

On en déduit que

$$\int_0^1 v^{\frac{n-3}{2}} (1-v)^{-\frac{1}{2}} dv = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})},$$

et finalement que pour tout  $x \in \mathbb{R}$

$$f_Y(x) = \begin{cases} \frac{x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \exp\left\{-\frac{x}{2}\right\} & \text{si } x > 0, \\ 0 & \text{sinon;} \end{cases}$$

$(\mathcal{P}_n)$  est donc vraie.

**Conclusion.** La propriété est vraie pour tout  $n \in \mathbb{N}^*$ .

## Combinaisons linéaires de variables Gaussiennes indépendantes

**Préliminaires** Cette question a été traitée de manière générale dans le cours. Nous en proposons une preuve alternative, basée sur le calcul de la fonction de répartition de la variable aléatoire  $sX + m$ , qui caractérise sa loi. Elle dépend clairement des valeurs de  $s$ .



- Si  $s = 0$ , alors  $sX + m$  est toujours égale à  $m$  : sa loi est une masse de Dirac en  $\{m\}$  et pour tout  $x \in \mathbb{R}$  on a  $\mathbb{P}(sX + m \leq x) = 1_{[m, +\infty[}(x)$ .
- Si  $s \neq 0$ , alors pour tout  $x \in \mathbb{R}$  on a

$$\mathbb{P}(sX + m \leq x) = \begin{cases} \mathbb{P}\left(X \leq \frac{x-m}{s}\right) = \int_{-\infty}^{\frac{x-m}{s}} f(u) du & \text{si } s > 0, \\ \mathbb{P}\left(X \geq \frac{x-m}{s}\right) = \int_{\frac{x-m}{s}}^{+\infty} f(u) du & \text{si } s < 0, \end{cases}$$

qui en posant le changement de variable  $v = sx + m$  donne

$$\mathbb{P}(sX + m \leq x) = \int_{-\infty}^x \frac{1}{|s|} f\left(\frac{u-m}{s}\right) dx.$$

Ainsi,  $sX + m$  admet une densité, qui pour tout  $x \in \mathbb{R}$  est égale à

$$\frac{1}{|s|} f\left(\frac{u-m}{s}\right) = \frac{1}{\sqrt{2\pi}|s|} \exp\left\{-\frac{(x-m)^2}{2s^2}\right\}.$$

On reconnaît la densité d'une loi Normale d'espérance  $m$  et de variance  $s^2$ .

## Combinaisons linéaires

2. Commençons par supposer que pour tout  $n \in \mathbb{N}^*$ ,  $a \in \mathbb{R}^n$  est tel qu'aucune de ses composantes n'est nulle. Nous allons montrer par récurrence sur  $n$  que  $S_n^a$  suit une loi Normale d'espérance nulle et de variance  $\sum_{i=1}^n a_i^2$ . On note cette propriété  $(\mathcal{P}_n)$ .

### Initialisation.

- Si  $n = 1$  et  $a_1 \neq 0$ , alors  $S_1^a = a_1 X_1$  suit une loi Normale centrée de variance  $a_1^2$  d'après la question 1;  $(\mathcal{P}_1)$  est donc vraie.
- Si  $n = 2$  et  $a_1, a_2 \neq 0$ , alors d'après le cours  $S_2^a = a_1 X_1 + a_2 X_2$  admet une densité, notée  $f_2^a$ , égale au produit de convolution des densités  $f_1$  de  $a_1 X_1$  et  $f_2$  de  $a_2 X_2$ . En outre, d'après la question 1, pour tout  $x \in \mathbb{R}$  on a  $f_1(x) = \frac{1}{|a_1|} f\left(\frac{x}{a_1}\right)$  et  $f_2(x) = \frac{1}{|a_2|} f\left(\frac{x}{a_2}\right)$ . Par conséquent, pour tout  $x \in \mathbb{R}$ ,

$$\begin{aligned} f_2^a(x) &= \int_{\mathbb{R}} f_1(x-u) f_2(u) du = \int_{\mathbb{R}} \frac{1}{|a_1||a_2|} f\left(\frac{x-u}{a_1}\right) f\left(\frac{u}{a_2}\right) du \\ &= \frac{1}{\sqrt{a_1^2 + a_2^2}} f\left(\frac{x}{\sqrt{a_1^2 + a_2^2}}\right) \int_{\mathbb{R}} \frac{\sqrt{a_1^2 + a_2^2}}{|a_1||a_2|} \frac{f\left(\frac{x-u}{a_1}\right)}{f\left(\frac{x}{\sqrt{a_1^2 + a_2^2}}\right)} f\left(\frac{u}{a_2}\right) du. \end{aligned}$$

Or pour tout  $x, u \in \mathbb{R}$  on a

$$\begin{aligned}
\frac{f\left(\frac{x-u}{a_1}\right)}{f\left(\frac{x}{\sqrt{a_1^2+a_2^2}}\right)} f\left(\frac{u}{a_2}\right) &= \frac{1}{\sqrt{2\pi}} \frac{\exp\left\{-\frac{(x-u)^2}{2a_1^2}\right\}}{\exp\left\{-\frac{x^2}{2(a_1^2+a_2^2)}\right\}} \exp\left\{-\frac{u^2}{2a_2^2}\right\} \\
&= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} \left( \frac{(x-u)^2}{a_1^2} + \frac{u^2}{a_2^2} - \frac{x^2}{a_1^2+a_2^2} \right)\right\} \\
&= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(a_2^2 x - (a_1^2+a_2^2)u)^2}{2a_1^2 a_2^2 (a_1^2+a_2^2)}\right\} \\
&= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\left(u - \frac{a_2^2 x}{a_1^2+a_2^2}\right)^2}{2 \frac{a_1^2 a_2^2}{a_1^2+a_2^2}}\right\},
\end{aligned}$$

qui multiplié par  $\frac{\sqrt{a_1^2+a_2^2}}{|a_1||a_2|}$  correspond à la densité d'une loi Normale d'espérance  $\frac{a_2^2 x}{a_1^2+a_2^2}$  et de variance  $\frac{a_1^2 a_2^2}{a_1^2+a_2^2}$ . La précédente intégrale vaut donc 1 et  $(\mathcal{P}_2)$  est vraie.

**Héritage.** Soit maintenant  $n \geq 2$ , et supposons  $(\mathcal{P}_{n-1})$  vraie. Alors  $S_n^a = S_{n-1}^{a_{-n}} + a_n X_n$ , où  $a_{-n} := (a_1, \dots, a_{n-1})$ . Or  $S_{n-1}^{a_{-n}}$  et  $a_n X_n$  sont des variables gaussiennes centrées, de variances respectives  $\sum_{i=1}^{n-1} a_i^2$  d'après  $(\mathcal{P}_{n-1})$  et  $a_n^2$  d'après la question 1. Par  $(\mathcal{P}_2)$ ,  $S_n^a$  suit donc une loi Normale centrée de variance  $\sum_{i=1}^n a_i^2$ .

**Conclusion.** On en conclut que  $(\mathcal{P}_n)$  est vraie pour tout  $n \in \mathbb{N}^*$ .

3. Calculons cette covariance :

$$\begin{aligned}
\text{Cov}(S_n^a, S_n^b) &= \text{Cov}\left(\sum_{i=1}^n a_i X_i, \sum_{j=1}^n b_j X_j\right) \\
&= \sum_{i=1}^n a_i b_i \mathbb{V}(X_i) + \sum_{1 \leq i \neq j \leq n} a_i b_j \text{Cov}(X_i, X_j).
\end{aligned}$$

Or par hypothèse  $\mathbb{V}(X_i) = 1$  pour tout  $i \in \{1, \dots, n\}$  et par indépendance on a  $\text{Cov}(X_i, X_j) = 0$  pour tous  $i, j \in \{1, \dots, n\}$  tels que  $i \neq j$ . Ainsi,

$$\text{Cov}(S_n^a, S_n^b) = \sum_{i=1}^n a_i b_i$$

qui est nulle ssi  $a$  et  $b$  sont orthogonaux.

## Loi de Maxwell

**Question 1** La densité du triplet  $(X, Y, Z)$  étant invariante par rotation, elle est de la forme  $f_{(X,Y,Z)}(x, y, z) = \phi(x^2 + y^2 + z^2)$ . De plus, les variables  $X, Y$  et  $Z$  étant indépendantes, on a

$$f_{(X,Y,Z)}(x, y, z) = f_X(x)f_Y(y)f_Z(z) \text{ pour tout } (x, y, z) \in \mathbb{R}^3.$$

Soit  $(x_0, y_0, z_0) \in \mathbb{R}^3$ ; comme  $\phi(x_0^2 + y_0^2 + z_0^2) > 0$  on a  $f_{(X,Y,Z)}(x_0, y_0, z_0) > 0$  et donc  $f_X(x_0) \neq 0$ ,  $f_Y(y_0) \neq 0$  et  $f_Z(z_0) \neq 0$ . Pour tout  $x \in \mathbb{R}$ , comme

$$f_X(x)f_Y(y_0)f_Z(z_0) = \phi(x^2 + y_0^2 + z_0^2) = \phi(y_0^2 + x^2 + z_0^2) = f_X(y_0)f_Y(x)f_Z(z_0),$$

on a l'égalité

$$f_X(x) = \frac{f_X(y_0)}{f_Y(y_0)} f_Y(x).$$

Comme on a par ailleurs

$$\int_{-\infty}^{+\infty} f_X(x) dx = \int_{-\infty}^{+\infty} f_Y(x) dx = 1,$$

il est nécessaire que le ratio  $f_X(y_0)/f_Y(y_0)$  soit égal à 1. On a donc  $f_X = f_Y$ ; la preuve que  $f_Y = f_Z$  s'obtient de manière similaire.

Pour finir, puisque pour tout  $x \in \mathbb{R}$  on a  $f(x)^3 = f_{(X,Y,Z)}(x, x, x) = \phi(3x^2) > 0$ , la fonction  $f$  est bien positive en tout point.

**Question 2** L'égalité

$$f(x)f(y)f(z) = \phi(x^2 + y^2 + z^2) \text{ pour tout } (x, y, z) \in \mathbb{R}^3$$

implique  $f(x)f'(y)f(z) = 2y\phi'(x^2 + y^2 + z^2)$  et  $f'(x)f(y)f(z) = 2x\phi'(x^2 + y^2 + z^2)$ . On en déduit que

$$xf(x)f'(y)f(z) = f'(x)yf(y)f(z) \text{ pour tout } (x, y, z) \in \mathbb{R}^3.$$

Soit  $(x, y_0, z_0) \in \mathbb{R}^3$  avec  $y_0 \neq 0$ ; on a  $f(x) > 0$ ,  $f(y_0) > 0$  et  $f(z_0) > 0$ . En posant

$$c = \frac{f'(y_0)f(z_0)}{y_0 f(y_0)f(z_0)},$$

on voit que

$$f'(x) = cx f(x) \text{ pour tout } x \in \mathbb{R}.$$

On peut réécrire cette équation différentielle sous la forme

$$(\ln f(x))' = cx.$$

En intégrant les deux membres de l'équation entre 0 et  $x$ , on obtient

$$\ln f(x) - \ln f(0) = c \frac{x^2}{2}$$

et donc

$$f(x) = f(0)e^{cx^2/2},$$

ce qui est bien la forme cherchée  $f(x) = ae^{cx^2/2}$ .

**Question 3** Comme  $\int_{\mathbb{R}} f(x)dx = 1$ , on a nécessairement  $c < 0$  et on en déduit que

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

Les variables aléatoires  $X$ ,  $Y$  et  $Z$  suivent donc loi gaussienne centrée et la covariance du vecteur est  $\sigma^2 I_3$ , où  $I_3$  est la matrice identité.

**Question 4** On a

$$\frac{1}{2}m\mathbb{E}(\|V\|^2) = \frac{1}{2}m\mathbb{E}(V_1^2 + V_2^2 + V_3^2).$$

Comme  $\mathbb{E}(V_i^2) = \sigma^2$ , on en déduit que  $E_c = \frac{3}{2}m\sigma^2$  et par conséquent que

$$\sigma^2 = \frac{kT}{m}.$$

**Question 5** À l'aide de la méthode de la fonction muette, on montre que  $V_1^2$  suit une loi gamma de paramètres  $(1/2, \frac{1}{2\sigma^2})$ . En effet, en appliquant le changement de variable  $y = x^2$  sur  $\mathbb{R}_-^*$  et  $\mathbb{R}_+^*$ , on obtient :

$$f_{V_1^2}(y) = (f(-\sqrt{y}) + f(\sqrt{y})) \frac{1}{2\sqrt{y}} 1_{]0, +\infty[}(x) = f(\sqrt{y}) \frac{1}{\sqrt{y}} 1_{]0, +\infty[}(x)$$

par parité de  $f$ , soit

$$\begin{aligned} f_{V_1^2}(y) &= \frac{1}{\sqrt{\pi}} \left(\frac{1}{2\sigma^2}\right)^{1/2} y^{-1/2} \exp\left(-\frac{y}{2\sigma^2}\right) \\ &= \frac{1}{\Gamma(1/2)} \left(\frac{1}{2\sigma^2}\right)^{1/2} y^{-1/2} \exp\left(-\frac{y}{2\sigma^2}\right) \end{aligned}$$

On déduit de l'indication que  $\|V\|^2 = V_1^2 + V_2^2 + V_3^2$  suit une loi gamma de paramètres  $(3/2, \frac{1}{2\sigma^2})$ . Sa densité est

$$f_{\|V\|^2}(x) = \frac{1}{\sqrt{2\pi\sigma^3}} \sqrt{x} e^{-x/2\sigma^2} 1_{]0, +\infty[}(x).$$

À l'aide de la méthode de la fonction muette, en appliquant le changement de variable  $v = \sqrt{x}$  sur  $\mathbb{R}_+^*$  et avec  $\sigma^2 = \frac{kT}{m}$ , on montre enfin que  $\|V\|$  est une variable de densité :

$$f_{\|V\|}(v) = \frac{\sqrt{2}}{\sqrt{\pi}} \left(\frac{m}{kT}\right)^{3/2} v^2 e^{-mv^2/2kT} 1_{]0, +\infty[}(v).$$

Il s'agit de la densité de la loi de Maxwell.