

# Probabilités II


STEP, MINES ParisTech\*

30 août 2020 (#59ca161)

## Table des matières

<b>Variables aléatoires réelles</b>	<b>3</b>
Remarque . . . . .	3
Notation . . . . .	4
Proposition . . . . .	4
Définition – variable aléatoire réelle . . . . .	5
Définition – loi d’une variable aléatoire réelle . . . . .	5
Proposition . . . . .	5
Démonstration (idée) . . . . .	5
Proposition . . . . .	5
Définition – variable aléatoire réelle à densité . . . . .	5
 <b>Moments d’une variable aléatoire à densité</b>	 <b>6</b>
Définition . . . . .	6
Remarque . . . . .	6
Proposition . . . . .	7
Rappel : cas discret . . . . .	7
Définition . . . . .	7
Proposition . . . . .	7
Remarque . . . . .	8
Remarque . . . . .	8
Définition . . . . .	9
Inégalité de Cauchy-Schwarz . . . . .	9
Remarque . . . . .	10
Proposition . . . . .	10
Remarque . . . . .	11
Exemples . . . . .	11

---

\*Ce document est un des produits 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.

<i>Loi uniforme</i> . . . . .	11
<i>Loi exponentielle</i> . . . . .	11
<i>Loi gamma</i> . . . . .	12
<i>Loi normale</i> . . . . .	12
Remarque . . . . .	13
<i>Loi de Cauchy</i> . . . . .	13
<b>Vecteurs aléatoires à densité</b>	<b>13</b>
Définitions . . . . .	14
Définition . . . . .	14
Proposition . . . . .	14
Proposition . . . . .	15
Démonstration . . . . .	15
Remarque . . . . .	15
Exemple . . . . .	15
Moments d'un vecteur aléatoire . . . . .	16
Définition . . . . .	16
Proposition . . . . .	16
Variables aléatoires indépendantes . . . . .	17
Définition . . . . .	17
Proposition . . . . .	18
Proposition . . . . .	18
Remarque . . . . .	19
Corollaire . . . . .	19
Remarque . . . . .	19
<b>Identification de densité</b>	<b>19</b>
Proposition . . . . .	20
Exemples . . . . .	21
<b>Exercices</b>	<b>22</b>
Loi de vie et de mort . . . . .	22
Loi bêta . . . . .	23
Crues centennales de la Seine . . . . .	23
Indépendance et vecteurs gaussiens . . . . .	24
Symétrie de la gaussienne . . . . .	24
Loi du $\chi^2$ . . . . .	24
Combinaisons linéaires de variables Gaussiennes indépendantes . . . . .	25
<b>Solutions</b>	<b>26</b>
Loi de vie et de mort . . . . .	26
Loi bêta . . . . .	26
Crues centennales de la Seine . . . . .	27
Indépendance et vecteurs gaussiens . . . . .	28
Symétrie de la gaussienne . . . . .	30
Combinaisons linéaires de variables Gaussiennes indépendantes . . . . .	32

## Variables aléatoires réelles

En théorie moderne des probabilités, on préfère prendre un point de vue fonctionnel plutôt qu'ensambliste, et utiliser les variables aléatoires plutôt que les événements. Une variable aléatoire est une grandeur qui dépend du résultat de l'expérience. Par exemple,

- le nombre de 6 obtenus dans un lancer de 3 dés,
- le nombre d'appels dans un central téléphonique pendant une heure,
- la distance du point d'atteinte d'une flèche au centre de la cible,
- la valeur maximale d'un prix d'actif sur un intervalle de temps donné,

sont des variables aléatoires.

La définition formelle d'une variable aléatoire fait intervenir des éléments de la théorie de la mesure qui nous font pour l'instant défaut. On s'intéressera dans un premier temps au cas d'une variable réelle dont on donne une définition partielle :

Soit  $\Omega$  l'espace fondamental muni de sa tribu  $\mathcal{A}$ . Une *variable aléatoire*  $X$  est une application de  $(\Omega, \mathcal{A})$  dans un ensemble  $E$ ,

$$\omega \in \Omega \mapsto X(\omega) \in E$$

En pratique, l'ensemble  $E$  pourra être un ensemble fini ou dénombrable ou  $\mathbb{R}$  ou  $\mathbb{R}^d$  ou encore un espace plus sophistiqué tel que l'ensemble  $C(\mathbb{R}_+, \mathbb{R}^d)$  des fonctions continues de  $\mathbb{R}_+$  dans  $\mathbb{R}^d$ .

### Remarque – Remarque

La terminologie, consacrée par l'usage, peut être trompeuse. Une variable aléatoire n'est pas une variable (au sens de l'analyse) mais une fonction. Cette terminologie est apparentée à la notion de variable en physique ou en sciences humaines où on désigne volontiers par "variable" la valeur prise par une fonction de l'état du système étudié.

L'intérêt principal de travailler avec des variables aléatoires est de pouvoir substituer à l'espace abstrait  $\Omega$  des résultats de l'expérience l'espace  $E$ , mieux connu dans la pratique. Ainsi, grâce à une variable aléatoire  $X$ , nous pouvons transporter la structure abstraite du modèle probabiliste  $(\Omega, \mathcal{A}, \mathbb{P})$  sur l'espace d'arrivée  $E$ , en posant pour  $B \subset E$

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega, X(\omega) \in B\})$$

Cette formule définit une nouvelle probabilité, notée  $\mathbb{P}_X$  et définie sur  $E$ , qui s'appelle la *loi de la variable  $X$* .

### Remarque – Notation

Il est usuel de noter l'ensemble  $X^{-1}(B) = \{\omega \in \Omega, X(\omega) \in B\}$  par  $\{X \in B\}$ , ce qui allège les écritures. On se rappellera néanmoins que cette notation désigne un sous-ensemble de  $\Omega$ .

Comme  $\mathbb{P}(A)$  n'est définie que pour les  $A$  de la tribu  $\mathcal{A}$ , la formule ci-dessus ne permet de définir  $\mathbb{P}_X(B)$  que pour les ensembles  $B$  tels que  $X^{-1}(B) \in \mathcal{A}$ , d'où l'importance de la proposition suivante :

### Proposition – Proposition

- a) La famille  $\mathcal{E}$  des parties  $B$  de  $E$  telles que  $X^{-1}(B) \in \mathcal{A}$  est une tribu de  $E$ .
- b) L'application  $\mathbb{P}_X$  définie pour  $B \in \mathcal{E}$  par

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$$

définit une probabilité sur le couple  $(E, \mathcal{E})$ .

**Démonstration** Les 3 propriétés de la définition d'une tribu (p. ??) pour  $\mathcal{E}$  ainsi que les deux propriétés de la définition de la probabilité (p. ??) pour  $\mathbb{P}_X$  découlent immédiatement des mêmes propriétés pour  $\mathcal{A}$  et  $\mathbb{P}$ , une fois remarquées les propriétés élémentaires suivantes :

$$\begin{aligned} X^{-1}(\emptyset) &= \emptyset, X^{-1}(E) = \Omega, X^{-1}(B^c) = X^{-1}(B)^c \\ X^{-1}(\cap_i A_i) &= \cap_i X^{-1}(A_i), X^{-1}(\cup_i A_i) = \cup_i X^{-1}(A_i) \end{aligned}$$

■

$\mathbb{P}_X$  sera plus facile à caractériser que  $\mathbb{P}$  puisque  $E$  est un ensemble connu (on pourra en particulier utiliser ses propriétés topologiques) alors que  $\Omega$  est un espace abstrait. Les variables que nous rencontrerons dans ce cours seront soit à valeurs dans un ensemble dénombrable, soit à valeurs dans  $\mathbb{R}$  ou dans  $\mathbb{R}^d$ . Nous les appellerons respectivement des variables aléatoires discrètes, réelles ou des vecteurs aléatoires. Leurs lois seront alors des probabilités respectivement sur un ensemble dénombrable, sur  $\mathbb{R}$  ou sur  $\mathbb{R}^d$ . Le cas discret est considéré connu.

La proposition ci-dessus (p. 4) implique que l'ensemble  $X^{-1}(B)$  soit un évènement, pour tout  $B$  dans  $\mathcal{E}$ . Dans le cas où  $E = \mathbb{R}$ , on notera  $\mathcal{E}_{\mathbb{R}}$  la tribu associée<sup>1</sup>. Cela nous conduit à poser :

---

1. Nous n'avons pas les outils permettant de caractériser cette tribu pour le moment. On verra par la suite que, dans le cas des variables aléatoires réelles à densité, elle est très similaire à la tribu des ensembles mesurables de  $\mathbb{R}$ , à une collection d'ensembles négligeables près.

### Définition – Définition – variable aléatoire réelle

Soit l'espace d'état  $\Omega$  muni de la tribu  $\mathcal{A}$  des évènements. Une application  $X$  de  $\Omega$  dans  $\mathbb{R}$  est une *variable aléatoire réelle* si  $X^{-1}(B) \in \mathcal{A}$  pour tout  $B \in \mathcal{E}_{\mathbb{R}}$ .

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

La probabilité  $\mathbb{P}_X$ , définie sur  $(\mathbb{R}, \mathcal{E}_{\mathbb{R}})$  par  $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$  pour  $B \in \mathcal{E}_{\mathbb{R}}$  est appelée *loi de la variable  $X$* , ou *distribution* de  $X$ .

On a alors le résultat très utile suivant :

### Proposition – Proposition

Si  $X_1, \dots, X_n$  sont des variables aléatoires réelles et si  $g$  est une fonction mesurable de  $\mathbb{R}^n$  dans  $\mathbb{R}$ , alors  $Y = g(X_1, \dots, X_n)$  est une variable aléatoire réelle.

### Démonstration (idée)

Puisque  $g$  est mesurable, le critère de l'image réciproque implique que  $\forall A \in \mathcal{E}_{\mathbb{R}}$ ,  $g^{-1}(A) \in \mathcal{E}_{\mathbb{R}^n}$ . Par composition, on en déduit que  $Y = g(X_1, \dots, X_n)$  est une variable aléatoire.

Comme application de ce résultat, on a les propriétés suivantes :

### Proposition – Proposition

Soient  $X, Y$  et  $(X_n)_{n \in \mathbb{N}^*}$  des variables aléatoires réelles. On a

1.  $X + Y, XY, \frac{X}{Y}$  si  $Y \neq 0$ , sont des variables aléatoires.
2.  $\sup_{1 \leq p \leq n} X_p, \inf_{1 \leq p \leq n} X_p$ , sont des variables aléatoires.
3.  $\sup_{n \geq 1} X_n, \inf_{n \geq 1} X_n$ , sont des variables aléatoires.
4. Si  $X_n(\omega) \xrightarrow[n \rightarrow \infty]{} Z(\omega), \forall \omega$ , alors la limite  $Z$  est une variable aléatoire.
5.  $Z = 1_A$  est une variable aléatoire  $\Leftrightarrow A \in \mathcal{A}$ .

### Définition – Définition – variable aléatoire réelle à densité

Soit  $X$  une variable aléatoire. On dit que  $X$  a une *loi de densité  $f$*  (ou par abus de langage “est de densité  $f$ ”), si  $\mathbb{P}_X$  admet la densité  $f$  et donc si pour tout réel  $x$ ,

$$\mathbb{P}(X \leq x) = \int_{-\infty}^x f(y) dy.$$

**Exemple – Exemple** On suppose que la durée de fonctionnement, en heures, d’un ordinateur avant sa première panne est une variable aléatoire positive de loi exponentielle de paramètre  $1/100$ , de densité donnée par

$$f(x) = \begin{cases} \frac{1}{100} \exp\left(-\frac{x}{100}\right) & x \geq 0, \\ 0 & x < 0. \end{cases}$$

Calculons la probabilité que cette durée de fonctionnement  $X$  soit comprise entre 50 et 150 heures, elle vaut

$$\mathbb{P}(X \in [50, 150]) = \int_{50}^{150} \frac{1}{100} \exp\left(-\frac{x}{100}\right) dx = \exp(-1/2) - \exp(-3/2) \approx 0,38.$$

Calculons la probabilité que l’ordinateur fonctionne moins de 100 heures :

$$\mathbb{P}(X \leq 100) = \int_0^{100} \frac{1}{100} \exp\left(-\frac{x}{100}\right) dx = 1 - e^{-1} \approx 0,63.$$

## Moments d’une variable aléatoire à densité

La densité de probabilité d’une variable aléatoire va nous permettre de calculer aisément des grandeurs caractéristiques telles que sa valeur moyenne et sa variance définies ci-dessous :

### Définition – Définition

La variable aléatoire  $X : \Omega \rightarrow \mathbb{R}$  de densité  $f$  est dite *intégrable* si l’intégrale  $\int_{\mathbb{R}} |x|f(x)dx$  est définie, autrement dit si le produit  $xf(x)$  est absolument intégrable<sup>2</sup>. On définit alors son *espérance* par

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

### Remarque – Remarque

$\mathbb{E}(X)$  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

---

2. Comme  $f$  est positive, on peut en fait se convaincre que  $xf(x)$  intégrable équivaut à  $xf(x)$  absolument intégrable : si  $xf(x)$  est intégrable, ses “restrictions”  $g(x) = xf(x)1_{\mathbb{R}_-}(x)$  et  $h(x) = xf(x)1_{\mathbb{R}_+}(x)$  sont intégrables (passer par la restriction aux intervalles  $\mathbb{R}_-$  et  $\mathbb{R}_+$  puis par le critère qui étend par 0 à  $\mathbb{R}$ ; les deux opérations préservent l’intégrabilité). Or  $|x|f(x) = h(x) - g(x)$  (sauf en 0), donc elle est intégrable.

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  $X$  à densité intégrables. Les propriétés suivantes découlent directement des propriétés de l'intégrale.

**Proposition – Proposition**

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

**Remarque – Rappel : cas discret**

Dans le cas d'une variable aléatoire discrète  $Y$  à valeurs dans  $\mathbb{N}^*$ , son espérance est définie par la quantité  $\mathbb{E}(Y) = \sum_{i \in \mathbb{N}^*} i\mathbb{P}(Y = i)$ , 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 – Définition**

La variable aléatoire  $X : \Omega \rightarrow \mathbb{R}$  de densité  $f$  est dite *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. Sa *variance* est définie par

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

**Proposition – Proposition**

$\mathcal{L}^2$  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  $(aX + Y)^2 \leq 2a^2X^2 + 2Y^2$ , 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. 7) (linéarité).

La première inégalité a déjà été vue ci-dessus (p. 7). 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. 7). Par conséquent,  $\mathbb{E}(X)^2 = a^2 \leq \mathbb{E}(X^2)$  ce qui est le résultat recherché. ■

### Remarque – Remarque

En vertu de cette proposition (p. 7),  $\mathbb{V}(X)$  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 – Remarque

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 – Définition

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  $\mathbb{V}(Y) > 0$ .

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,  $Y(\omega) = -x_0X(\omega)$  pour presque tout  $\omega$ . ■

### Remarque – Remarque

On déduit de l'inégalité de Cauchy-Schwarz (p. 9) que le coefficient de corrélation de  $X$  et  $Y$  vérifie

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

Enfin, il peut être intéressant de pouvoir calculer l'espérance d'une fonction mesurable d'une variable aléatoire réelle à densité qui est une variable aléatoire en vertu de la proposition vue plus haut (p. 5).

### Proposition – Proposition

Soit  $X$  une variable aléatoire réelle admettant la densité  $f$ , et  $g$  une fonction mesurable de  $\mathbb{R}$  dans  $\mathbb{R}$ . Alors  $g(X)$  est intégrable si et seulement si l'intégrale

$$\int_{\mathbb{R}} |g(x)|f(x)dx,$$

est définie et dans ce cas

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

Nous n'avons pas tous les éléments permettant de démontrer ce résultat, mais l'argument heuristique suivant permet de comprendre pourquoi il est vrai : supposons  $g$  continue telle que  $|g|f$  soit intégrable. Alors il existe une jauge  $\gamma(t)$ , sur  $[-\infty, +\infty]$  telle que, si la subdivision pointée (totale ou partielle)  $\mathcal{D} = \{(t_i, I_i)\}_i$  est subordonnée à  $\gamma$ , on a

$$S(gf, \mathcal{D}) = \left| S(gf, \mathcal{D}) - \int_{\mathbb{R}} g(x)f(x)dx \right| \leq \varepsilon.$$

Posons  $X_n = t_i$  si  $X \in I_i$ , pour  $i \in \{0, \dots, n-1\}$ . Ainsi, pour tout  $\omega$ ,  $X_n(\omega) \xrightarrow{n \rightarrow \infty} X(\omega)$  et par continuité de  $g$ , on a  $g(X_n) \xrightarrow{n \rightarrow \infty} g(X)$ . Comme  $X_n$  est une variable aléatoire discrète, on a

$$\mathbb{E}(g(X_n)) = \sum_{i=0}^{n-1} g(t_i) \mathbb{P}(X \in I_i) l(I_i) \approx \sum_{i=0}^{n-1} g(t_i) f(t_i).$$

### Remarque – Remarque

L'espérance et la variance sont des cas particulier de ce résultat. On de plus pour  $A \in \mathcal{E}_{\mathbb{R}}$  :

$$\mathbb{E}(1_A(X)) = \int_A f(x) dx = \mathbb{P}(X \in A)$$

### Exemples

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

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

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

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

#### *Loi exponentielle*

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

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

Son espérance et sa variance se calculent aisément et valent

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

### ***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

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

Son espérance et sa variance s'obtiennent en utilisant le changement de variable  $x \mapsto \theta x$  et la définition de la fonction gamma :

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

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  $\theta$ .

### ***Loi normale***

de paramètres  $\mu$  et  $\sigma^2$  et on note  $X \sim \mathcal{N}(\mu, \sigma^2)$  si  $X$  est de densité<sup>3</sup>

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

Son espérance et sa variance valent

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

Pour le voir, 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  $x \mapsto \frac{x-\mu}{\sigma}$ .

---

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

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

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

### Remarque – Remarque

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'exemple suivant.

### 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.

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.

## Vecteurs aléatoires à densité

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

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

Mais 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.

### Définition – Définition

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. 10), on a :

### Proposition – Proposition

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 – Proposition

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 pour  $1_{]-\infty, x] \times \mathbb{R}}(u, v)f(u, v)$  intégrable. 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 – Remarque

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.

### Exemple

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 lancés qui atteignent la cible).

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.

## Moments d'un vecteur aléatoire

### Définition – Définition

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 – Proposition

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

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

■

**Exemple – Exemple : Vecteur Gaussien  $n$ -dimensionnel** Un exemple de vecteurs aléatoires est celui des vecteurs gaussiens, que nous étudierons 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  $C_X = C$ .



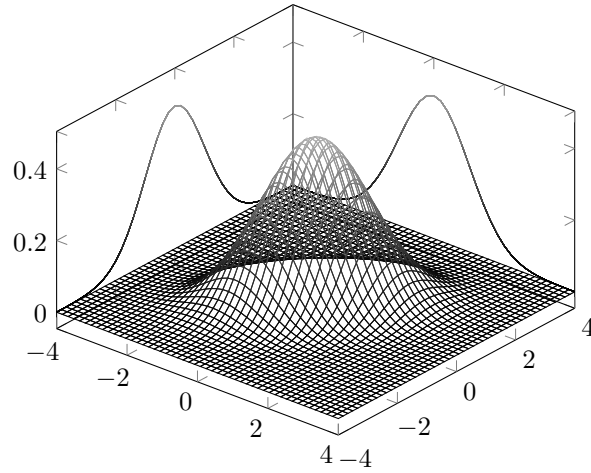


FIGURE 1 – Densité de probabilité de la loi normale bivariable centrée, réduite, de coefficient de corrélation  $1/2$

## Variables aléatoires indépendantes

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 événements pour caractériser l'indépendance de deux vecteurs aléatoires. En effet, considérons le vecteur aléatoire  $Z = (X, Y)$ ,  $A$  et  $B$  deux ensembles dans  $\mathcal{E}_{\mathbb{R}^m}$  et  $\mathcal{E}_{\mathbb{R}^n}$ . On a vu que les événements  $X \in A$  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 quelques soient  $A$  et  $B$ .

### Définition – Définition

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

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

Cette définition se traduit en termes de densités dans la proposition suivante que l'on énonce sans perte de généralité pour un couple de variables aléatoires.

### Proposition – Proposition

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  $\mathcal{E}_{\mathbb{R}}$

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

où on a utilisé le théorème de Fubini. ■

Considérons maintenant deux fonctions  $g$  et  $h$  définies sur  $\mathbb{R}^m$  et  $\mathbb{R}^n$  telles que  $g(X)$  et  $h(Y)$  soient aussi des variables aléatoires.

### Proposition – Proposition

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

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

### Remarque – Remarque

Ce résultat est encore valable  $X$  et  $Y$  dans le cas général (sans densité) mais nous ne disposons pas encore des outils de théorie de la mesure nécessaires à sa démonstration.

On déduit de ce résultat et de la définition de la covariance (p. 9) que :

### Corollaire – Corollaire

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

### Remarque

Attention, la réciproque est fautive. 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 réel, admettant la densité  $f_X$ . Soit  $g$  une fonction mesurable, de sorte que  $Y = g(X)$  soit 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 masse de Dirac en  $a$ , qui n'a pas de densité.

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. 10) 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 que nous ne démontrerons pas :

### Proposition – Proposition

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

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

alors la loi de  $X$  admet la densité  $f$ .

L'idée de la preuve repose sur le fait que parmi ces fonctions se trouvent les  $h = 1_{]-\infty, y]}$ , pour laquelle la formule précédente donne la fonction de répartition de  $f$ .

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 absolument intégrable, le théorème de changement de variable nous assure :

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

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 injective 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 :

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

## Exemples

1. **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  $N^o(0, 1)$ , on a  $f_X(u, v) = \frac{1}{2\pi} \exp\left(-\frac{u^2+v^2}{2}\right)$  et donc

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

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  $1_{]0, 2\pi]}(\theta)$ .

2. **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  $f_V$ .

## Exercices

### Loi de vie et de mort

La durée de vie d'un être vivant ou d'un matériel peut-être assimilée à une variable aléatoire strictement positive  $T$ . Dans ce cadre, on peut définir les notions suivantes :

- Loi de vie a priori : il s'agit de la loi du temps  $T$  caractérisée par fonction de répartition complémentaire

$$G(t) = \mathbb{P}(T > t)$$

- Loi de survie après  $t_0 \geq 0$  : il s'agit de la loi du temps  $T - t_0$  qu'il lui reste à vivre, sachant qu'il était encore en vie à  $t_0$ , de fonction de répartition complémentaire

$$G_{t_0}(t) = \mathbb{P}(T > t + t_0 | T > t_0)$$

On dira que la loi de vie satisfait la propriété de non vieillissement (ou d'absence de mémoire) si la loi de survie et la loi de vie sont égales :

$$\forall t \geq 0 \text{ et } \forall t_0 \geq 0, G_{t_0}(t) = G(t)$$

**Question 1** Exprimer la loi de survie à partir de la loi de vie a priori. (Solution p. 26.)

**Question 2** Montrer qu'une variable aléatoire  $T$  satisfait la propriété de non-  
vieillessement si et seulement si elle est de loi exponentielle. (Solution p. 26.)

**Question 3** On suppose que  $T$  admet une densité continue sur  $\mathbb{R}_+^*$ . Montrer  
que l'on peut définir le taux de mort à l'instant  $t$  par

$$D(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(t < T \leq t + \Delta t | T > t)$$

et exprimer  $G$  en fonction de  $D$ .

Quelles lois correspondent-elles à  $D$  constant ? (Solution p. 26.)

## 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  $\frac{U}{U+V}$ . (Solution p. 26.)

## 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. 27.)

**Question 2** En déduire la densité de  $Y$ . (Solution p. 28.)

**Question 3** Expliciter la densité de  $Y$  lorsque les  $X_i$  sont de loi uniforme sur  
 $[0, 1]$ . (Solution p. 28.)

**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. 28.)

**Question 5** Toujours selon les mêmes hypothèses, calculer la médiane de  $Y$ .  
La comparer avec celle des  $X_i$ . Commenter. (Solution p. 28.)

## 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. 28.)

**Question 2** Expliciter les lois marginales de  $X$  et de  $Y$ . (Solution p. 29.)

**Question 3** Montrer que  $X$  et  $Y$  sont indépendantes ssi  $\text{Cov}(X, Y) = 0$ . (Solution p. 30.)

## 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. 30.)

**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. 31.)

**Question 3** Les deux variables  $X$  et  $X_{c_0}$  sont-elles indépendantes ? (Solution p. 32.)

**Question 4** Le vecteur aléatoire  $(X, X_{c_0})$  est-il gaussien ? (Solution p. 32.)

## 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. 34.)

**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. 35.)

## 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. 32.)

**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. 33.)

## Solutions

### Loi de vie et de mort

**Question 1** Par définition, on a

$$G_{t_0}(t) = \mathbb{P}(T > t + t_0 | T > t_0) = \frac{\mathbb{P}(T > t + t_0, T > t)}{\mathbb{P}(T > t)} = \frac{G(t + t_0)}{G(t)}$$

**Question 2** D'après la question précédente,  $G$  vérifie alors  $G(t_0 + t) = G(t)G(t_0)$  pour tous  $t, t_0 > 0$ . Comme  $G$  est décroissante, continue à gauche et tend vers 0 à l'infini, on en déduit que  $G(t) = e^{-\theta t}$ , pour un  $\theta > 0$ . On reconnaît la fonction de répartition complémentaire d'une loi exponentielle de paramètre  $\theta$ .

**Question 3**

$$\begin{aligned} D(t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(t < T \leq t + \Delta t | T > t) \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \frac{\mathbb{P}(t < T \leq t + \Delta t)}{\mathbb{P}(T > t)} \\ &= \frac{1}{G(t)} \lim_{\Delta t \rightarrow 0} \frac{G(t) - G(t + \Delta t)}{\Delta t} \\ &= -\frac{g(t)}{G(t)} \end{aligned}$$

Ainsi,  $D(t) = \frac{d}{dt}(-\ln G(t))$  et comme  $G(0) = 1$ , alors on a pour  $t > 0$

$$G(t) = \exp\left(-\int_0^t D(s)ds\right)$$

Si  $D$  est constant, on retrouve une loi exponentielle.

### 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 Beta( $r, b$ ), et, inversement, la proportion de boules bleues tend vers une variable aléatoire de loi 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  $(\alpha, \beta)$ .

## 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

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

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

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

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.

## Combinaisons linéaires de variables Gaussiennes indépendantes

**Preliminaires** 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  $f_n^a$ , 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

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

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 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 absolument 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 :

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

On en déduit que

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

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

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

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

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