Probabilités générales

Laurent Rouvière

Université Rennes 2
Place du Recteur H. le Moal
CS 24307 - 35043 Rennes
Tel : 02 99 14 18 21
Mel : laurent.rouviere@univ-rennes2.fr

Table des matières

1 Rappels et notations **5**
1.1 L'espace $(\Omega, \mathcal{A}, \mathbf{P})$ . . . 5
1.2 Variable aléatoire . . . 6
1.3 Intégration . . . 7
1.4 Mesure définie par une densité . . . 9
1.5 Espérance mathématique . . . 10

2 Variables aléatoires absolument continues **13**
2.1 Fonction de répartition, fonction de masse, fonction de densité . . . 13
2.1.1 Fonction de répartition d'une v.a.r. . . . 13
2.1.2 Fonction de masse et classification des v.a.r. . . . 14
2.1.3 Loi absolument continue - densité de probabilité . . . 15
2.2 Espérance - moments d'une v.a.r. . . . 16
2.2.1 Définition . . . 16
2.2.2 Inégalités faisant intervenir les moments . . . 19
2.3 Médiane - quantiles d'une v.a.r. . . . 19
2.4 Calcul de lois . . . 20
2.4.1 Cas discret . . . 20
2.4.2 Cas absolument continu . . . 20

3 Vecteurs aléatoires **23**
3.1 Généralités . . . 23
3.2 Fonctions de répartition - densités . . . 24
3.2.1 Fonctions de répartition . . . 24
3.2.2 Cas discret . . . 24
3.2.3 Cas absolument continu . . . 25
3.3 Espérance - moments d'un vecteur aléatoire . . . 28
3.3.1 Espérance . . . 28
3.3.2 Variance - covariance . . . 29
3.4 Corrélation . . . 31
3.4.1 Le coefficient de corrélation linéaire . . . 31
3.4.2 Interprétation hilbertienne . . . 32
3.5 Variables (ou vecteurs) aléatoires indépendant(e)s . . . 32
3.6 Calcul de loi . . . 35
3.6.1 Cas discret . . . 35
3.6.2 Cas absolument continue . . . 35

4 Lois usuelles dans $\mathbb{R}^n$ **39**
4.1 La loi multinomiale . . . 39
4.2 Vecteurs gaussiens - loi multinormale . . . 40
4.2.1 Rappels sur la loi normale dans $\mathbb{R}$ . . . 40
4.2.2 Définition - premières propriétés . . . 40
4.2.3 Vecteurs gaussiens et loi du Chi-Deux . . . 42

5 Fonctions génératrices, fonctions caractéristiques **43**
5.1 Fonction génératrice des moments factoriels . . . 43
5.2 Fonction génératrice des moments . . . 44
5.3 Fonction caractéristique . . . 45
5.3.1 Cas d'une variable aléatoire réelle . . . 45
5.3.2 Cas d'un vecteur aléatoire . . . 47

6 Conditionnement, espérance et variance conditionnelles **49**
6.1 Rappels de calcul de probabilités conditionnelles . . . 49
6.2 Cas discret . . . 49
6.3 Cas absolument continue . . . 51
6.4 Interprétation géométrique de l'espérance conditionnelle . . . 53
6.5 Espérance conditionnelle : le cas général . . . 54
6.6 Probabilités conditionnelles . . . 56
6.7 Généralisation au conditionnement pas des sous-tribus . . . 57

7 Convergences de suites de variables aléatoires **59**
7.1 Les différents types de convergence . . . 60
7.1.1 Convergence presque sûre ou convergence forte . . . 60
7.1.2 La convergence en probabilité . . . 61
7.1.3 La convergence en moyenne d'ordre $p > 0$ . . . 62
7.1.4 La convergence en loi . . . 62
7.2 La loi des grand nombres . . . 65
7.2.1 Lois faibles des grands nombres . . . 66
7.2.2 Lois fortes des grands nombres . . . 66
7.3 Le théorème central limite . . . 66

A Rappels de cours sur la loi normale **69**
A.1 Loi normale centrée réduite . . . 69
A.2 La loi normale $\mathcal{N}(\mu, \sigma^2)$ . . . 71
A.3 Somme et moyenne de lois normales . . . 72

B Lois usuelles dans $\mathbb{R}$ **73**
B.1 Lois discrètes . . . 74
B.2 Lois absolument continues . . . 76

C Annales **79**
Partiel décembre 2009 . . . 80
Examen janvier 2010 . . . 84
Partiel décembre 2010 . . . 90
Examen janvier 2011 . . . 95

Partiel décembre 2011 102
Examen janvier 2012 107
Partiel décembre 2012 115
Examen janvier 2013 120
Partiel décembre 2013 128
Examen janvier 2014 133
Examen janvier 2015 139

Bibliographie **146**

Chapitre 1

Rappels et notations

1.1 L'espace $(\Omega, \mathcal{A}, \mathbf{P})$

On désigne par Ω l'ensemble des épreuves ou évènements élémentaires ω et par $\mathcal{A}$ une tribu sur Ω, c'est-à-dire une classe de parties de Ω qui vérifie :

— $\Omega \in \mathcal{A}$;

— $\forall A \in \mathcal{A}, \bar{A} \in \mathcal{A}$ (stabilité par complémentation) ;

— $\forall A_n \in \mathcal{A}, n \in \mathbb{N}^\star, \bigcup_{n\in\mathbb{N}^\star} A_n \in \mathcal{A}$ (stabilité par union dénombrable).

L'ensemble des parties de Ω, $\mathcal{P}(\Omega)$, est un exemple de tribu sur Ω. Une élément A d'une tribu $\mathcal{A}$ sera appelé *évènement aléatoire* (un évènement aléatoire est un élément de la tribu mais c'est un sous-ensemble de Ω). Un couple $(\Omega, \mathcal{A})$ où Ω est un ensemble et $\mathcal{A}$ une tribu sur Ω est appelé *espace mesurable.*

Définition 1.1
Soit $\mathcal{C}$ une partie de Ω. La tribu engendrée par $\mathcal{C}$ est la plus petite tribu contenant $\mathcal{C}$. Cette tribu peut également être définie comme l'intersection de toutes les tribus contenant $\mathcal{C}$. On la note $\sigma(\mathcal{C})$.

Définition 1.2
On appelle tribu borélienne de $\mathbb{R}$, notée $\mathcal{B}_\mathbb{R}$, la plus petite tribu contenant tous les intervalles de $\mathbb{R}$. Cette tribu est la tribu engendrée par les intervalles ouverts de $\mathbb{R}$.

Définition 1.3

1. *On appelle mesure sur $(\Omega, \mathcal{A})$ toute application définie sur $\mathcal{A}$ à valeurs dans $\bar{\mathbb{R}}^+ = \mathbb{R}^+ \cup \{+\infty\}$ vérifiant la propriété de σ-additivité, c'est-à-dire que pour toute suite $(A_n)_{n\in\mathbb{N}^\star}$ d'éléments de $\mathcal{A}$ deux à deux disjoints :*
$$\mu\left(\bigcup_{n\in\mathbb{N}^\star} A_n\right) = \sum_{n\in\mathbb{N}^\star} \mu(A_n).$$
2. *L'espace $(\Omega, \mathcal{A}, \mu)$ est alors appelé espace mesurable.*
3. *Si on a de plus $\mu(\Omega) = 1$ alors on dira que μ est une mesure de probabilité et que l'espace $(\Omega, \mathcal{A}, \mu)$ est un espace de probabilité.*

Exemple 1.1

1. *Mesure de comptage.* Elle est définie sur l'espace $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ par
$$\mu_c = \sum_{x\in\mathbb{N}} \delta_x$$
où δ_x est la masse de Dirac en x : $\delta_x(a) = 1$ si $a = x$, 0 sinon.

2. *Mesure de Lebesgue.* Elle est définie sur l'espace $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ pour tout intervalle $]a, b]$ par
$$\lambda\left(]a, b]\right) = b - a.$$

1.2 Variable aléatoire

Nous reprenons l'exemple de Montfort (1996), page 30. L'expérience aléatoire consiste à jeter n fois une pièce de monnaie. L'espace $\Omega = \{P, F\}^n$ est muni de la tribu $\mathcal{P}(\Omega)$. On s'intéresse ici simplement au nombre de piles sortis au cours des n jets. On définit $\Omega' = \{0, 1, \ldots, n\}$ et on considère la fonction

$$\begin{aligned} X : \Omega &\to \Omega' \\ \omega &\mapsto X(w) = \text{nombre de piles dans } \omega. \end{aligned}$$

On munit $(\Omega, \mathcal{P}(\Omega))$ d'une mesure de probabilité uniforme $\mathbf{P}$, *i.e.*, $\mathbf{P}(\omega) = \frac{1}{2^n}$ $\forall \omega \in \Omega$. Afin de caractériser le phénomène d'intérêt (nombre de piles), il parait intéressant de définir sur $(\Omega', \mathcal{P}(\Omega'))$ une mesure de probabilité $\mathbf{P}'$ par

$$\forall A \in \mathcal{P}(\Omega'),\ \mathbf{P}'(A') = \mathbf{P}\left(X^{-1}(A')\right).$$

Cependant, pour que cette définition ait un sens il est nécessaire que $X^{-1}(A') \in \mathcal{P}(\Omega)$. C'est clairement le cas dans cet exemple mais il est possible d'envisager des cas où cette condition n'est pas vérifiée.

Définition 1.4
Soit $(\Omega, \mathcal{A})$ un espace de probabilité et $(\Omega', \mathcal{A}')$ un espace mesurable. On appelle variable aléatoire une fonction X définie sur Ω à valeurs dans Ω' telle que :

$$\forall A' \in \mathcal{A}',\ X^{-1}(A') \in \mathcal{A},$$

où $X^{-1}(A') = \{\omega \in \Omega : X(\omega) \in A'\} = \{X \in A'\}$.

Définition 1.5
Soit $(\Omega, \mathcal{A})$ et $(\Omega', \mathcal{A}')$ deux espaces mesurables. Une fonction f définie sur Ω à valeurs dans Ω' est dite mesurable si :

$$\forall A' \in \mathcal{A}',\ f^{-1}(A') \in \mathcal{A}.$$

Si $(\Omega', \mathcal{A}') = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, f sera dite borélienne.

Une variable aléatoire est donc une fonction mesurable définie sur un espace de probabilité.

Définition 1.6
Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace de probabilité et X une variable aléatoire définie sur Ω et à valeurs dans Ω'. On appelle loi de probabilité $\mathbf{P}_X$ de X la mesure image de $\mathbf{P}$ par X :

$$\forall A' \in \mathcal{A}' : \mathbf{P}_X(A') = \mathbf{P}(X^{-1}(A')) = \mathbf{P}(X \in A').$$

Proposition 1.1
Soit X une variable aléatoire réelle définie sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$ à valeurs dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Alors $X^{-1}(\mathcal{B}_{\mathbb{R}})$ est une tribu sur Ω : c'est la tribu des évènements engendrés par X. Elle est notée $\sigma(X)$.

1.3 Intégration

Définition 1.7
Soit f borélienne définie sur un espace mesurable $(\Omega, \mathcal{A})$. f est appelée fonction $\mathcal{A}$-étagée (ou étagée) sur Ω si elle est combinaison linéaire finie de fonctions indicatrices de parties $\mathcal{A}$-mesurable de Ω, i.e.,

$$f = \sum_{i=1}^{n} \alpha_i \mathbf{1}_{A_i}$$

où les α_i sont des réels et les A_i des éléments de $\mathcal{A}$. On désigne par ξ^+ l'ensemble des fonctions mesurables étagées positives.

Théorème 1.1
Soit f une application définie sur $(\Omega, \mathcal{A})$ à valeurs dans $(\bar{\mathbb{R}}, \mathcal{B}_{\bar{\mathbb{R}}})$.

1. *Si f est positive, elle est mesurable si et seulement si elle est limite d'une suite croissante d'applications mesurables étagées positives.*
2. *f est mesurable si et seulement si elle est limite d'une suite de fonctions mesurables étagées.*

Définition 1.8
Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et f une application de ξ^+. On appelle intégrale de f par rapport à μ, l'application définie sur ξ^+ à valeurs dans $\overline{\mathbb{R}^+}$ par

$$I(f) = \int f \,\mathrm{d}\mu = \int_\Omega f(\omega)\,\mathrm{d}\mu(\omega) = \sum_{i=1}^{n} \alpha_i \mu(A_i).$$

Définition 1.9
Soit μ une mesure sur $(\Omega, \mathcal{A})$ et $f \in \mathcal{M}^+$ l'ensemble des applications mesurables positives définies sur $(\Omega, \mathcal{A})$à valeurs dans $\overline{\mathbb{R}^+}$. On appelle intégrale de f la quantité

$$I(f) = \int f\,\mathrm{d}\mu = \sup\left\{\int h\,\mathrm{d}\mu : h \in \xi^+, h \leq f\right\}.$$

Proposition 1.2
Soit $f \in \mathcal{M}^+$ et $(h_n)_{n\in\mathbb{N}^\star}$ une suite croissante d'éléments de ξ^+ telle que $f = \lim_{n\to\infty} h_n$. Alors :

$$\int f\,\mathrm{d}\mu = \lim_{n\to\infty} \int h_n\,\mathrm{d}\mu.$$

Théorème 1.2 (Beppo-Lévi)
Soit $(f_n)_{n\in\mathbb{N}^\star}$ une suite croissante d'éléments de $\mathcal{M}^+$ et $f = \lim_{n\to\infty} f_n$. Alors

$$\lim_{n\to\infty} \int f_n\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu.$$

On note $\mathcal{M}$ l'ensemble des application mesurables à valeurs dans $\bar{\mathbb{R}}$, $f^+ = \max(0, f)$ et $f^- = \max(0, -f)$.

Définition 1.10
Soit $f \in \mathcal{M}$. f est dite μ-intégrable si :

$$\int f^+ \,\mathrm{d}\mu < +\infty \quad et \quad \int f^- \,\mathrm{d}\mu < +\infty.$$

L'intégrale de f par rapport à μ est la quantité :

$$I(f) = \int f \,\mathrm{d}\mu = \int f^+ \,\mathrm{d}\mu - \int f^- \,\mathrm{d}\mu.$$

On note $\mathcal{L}_1(\Omega, \mathcal{A}, \mu)$ l'ensemble des fonctions de $\mathcal{M}$ μ-intégrables.

Proposition 1.3
f est μ-intégrable si et seulement si $\int |f| \,\mathrm{d}\mu < +\infty$.

Exemple 1.2

1. **Intégrale par rapport à une mesure discrète.** Une mesure μ est discrète si elle est de la forme :
$$\mu = \sum_{n=1}^{+\infty} p_n \delta_{\omega_n}$$
où les p_n sont positifs et $\omega_n \in \Omega$. Alors $\forall f \in \mathcal{M}$:
$$\int_\Omega f(\omega) \,\mathrm{d}\mu(\omega) = \sum_{n=1}^{+\infty} p_n f(\omega_n).$$
f est donc intégrable si et seulement si la série de terme général $p_n f(\omega_n)$ est absolument convergente.

2. **Intégrale par rapport à la mesure de Lebesgue.**
$$\int \mathbf{1}_{]a,b]} \,\mathrm{d}\lambda = \lambda(]a,b]) = b - a = \int_a^b \mathrm{d}x.$$
L'intégrale de $\mathbf{1}_{]a,b]}$ par rapport à la mesure de Lebesgue est égale à l'intégrale de Riemann. Plus généralement, si f est Riemann-intégrable sur $]a,b[$, elle est Lebesgue-intégrable sur $]a,b[$ et les deux intégrales coïncident :
$$\int_{]a,b[} f(x) \,\mathrm{d}\lambda(x) = \int_a^b f(x) \,\mathrm{d}x.$$

Définition 1.11

1. *Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Un élément A de $\mathcal{A}$ est dit μ-négligeable s'il existe $B \in \mathcal{A}$ tel que $A \subset B$ et $\mu(B) = 0$.*
2. *Une propriété Π définie sur Ω est dite vraie μ-presque partout si l'ensemble $\{\omega \in \Omega : \Pi(\omega)$ est fausse$\}$ est μ-négligeable. Si μ est une probabilité, on dit que la propriété est vraie presque sûrement.*

Théorème 1.3 (Théorème de convergence dominée ou de Lebesgue)
Soit g une application μ-intégrable et $(f_n)_{n\in\mathbb{N}^\star}$ une suite d'applications mesurables telles que :

$$\forall n \in \mathbb{N}^\star,\ |f_n| \leq g \ \ \mu - \text{p.p.}$$

On suppose en outre que $(f_n)_{n\in\mathbb{N}^\star}$ converge $\mu - \text{p.p.}$ vers une application mesurable f. Alors

1. *f est μ-intégrable ;*
2. *On a*

$$\lim_{n\to+\infty}\int |f_n - f|\,\mathrm{d}\mu = 0 \quad \textit{ou encore} \quad \lim_{n\to+\infty}\int f_n\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu.$$

Théorème 1.4

1. *Soit $(f_n)_{n\in\mathbb{N}^\star}$ une suite d'applications mesurables positives. Alors*

$$\int \sum_{n\in\mathbb{N}^\star} f_n\,\mathrm{d}\mu = \sum_{n\in\mathbb{N}^\star}\int f_n\,\mathrm{d}\mu.$$

2. *Si $(f_n)_{n\in\mathbb{N}^\star}$ est une suite d'applications intégrables vérifiant*

$$\sum_{n\in\mathbb{N}^\star}\int |f_n|\,\mathrm{d}\mu < +\infty,$$

alors la série $\sum_{n\in\mathbb{N}^\star} f_n$ converge $\mu-\mathrm{p.p.}$, sa somme est intégrable et

$$\int \sum_{n\in\mathbb{N}^\star} f_n\,\mathrm{d}\mu = \sum_{n\in\mathbb{N}^\star}\int f_n\,\mathrm{d}\mu.$$

1.4 Mesure définie par une densité

Définition 1.12
Soit $(\Omega,\mathcal{A},\mu)$ un espace mesuré et $f\in\mathcal{M}^+$. Alors l'application $\nu:\mathcal{A}\to\overline{\mathbb{R}^+}$ définie par

$$\nu(A) = \int_A f\,\mathrm{d}\mu$$

est une nouvelle mesure sur $(\Omega,\mathcal{A})$ appelée mesure de densité f par rapport à μ.

Théorème 1.5
Soit ν une mesure de densité par rapport à μ. Alors

$$\forall A\in\mathcal{A},\ \mu(A)=0\Rightarrow\nu(A)=0.$$

Lorsque la propriété précédente est vérifiée, on dit que ν est absolument continue par rapport à μ ($\nu\ll\mu$).

Théorème 1.6 (Radon-Nikodym)
Soit μ une mesure σ-finie et ν une mesure sur $(\Omega,\mathcal{A})$ absolument continue par rapport à μ. Alors il existe une application f positive, unique à une μ-équivalence près telle que :

$$\forall A\in\mathcal{A},\ \nu(A) = \int_A f\,\mathrm{d}\mu.$$

Le résultat demeure vrai si ν est σ-finie et si ν est finie alors f est μ-intégrable. f est appelée dérivée de Radon-Nikodym de ν par rapport à μ et sera notée

$$f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \quad \textit{ou} \quad \mathrm{d}\nu = f\mathrm{d}\mu.$$

Exemple 1.3 (Loi absolument continue)
Soit X une v.a. absolument continue de loi $\mathbf{P}_X$. D'après le théorème précédent, X admet une densité f par rapport à la mesure de Lebesgue. f est de plus λ-intégrable et on a

$$\forall B \in \mathcal{B}_{\mathbb{R}},\ \mathbf{P}_X(B) = \int_B \mathrm{d}\mathbf{P}_X = \int_B f\,\mathrm{d}\lambda.$$

Exemple 1.4 (Loi discrète)
Soit X une v.a. discrète à valeurs dans $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. $\mathbf{P}_X$ est finie, la mesure de comptage sur $\mathbb{N}$ $\mu_c = \sum_{n\in\mathbb{N}} \delta_n$ est σ-finie. Par conséquent, il existe une densité μ_c-intégrable valant pour tout $k \in \mathbb{N}$

$$\frac{\mathrm{d}\mathbf{P}_X}{\mathrm{d}\mu_c} = \mathbf{P}(X = k).$$

On retrouve bien

$$\forall k \in \mathbb{N},\ \ \mathbf{P}_X(\{k\}) = \mathbf{P}(X = k) \sum_{k\in\mathbb{N}} \delta_n(\{k\}) = \mathbf{P}(X = k).$$

1.5 Espérance mathématique

Définition 1.13
Soit X une v.a. définie sur $(\Omega, \mathcal{A}, \mathbf{P})$ et à valeurs dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. On définit l'espérance mathématique de X comme l'intégrale de X par rapport à $\mathbf{P}$:

$$\mathbf{E}[X] = \int_\Omega X(\omega)\,\mathrm{d}\mathbf{P}(\omega)$$

sous réserve que $\mathbf{E}[|X|] < +\infty$.

Un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ est souvent abstrait et dans de nombreux cas, il est difficile (voir impossible) d'expliciter Ω ou $\mathcal{A}$. Il est alors pratique d'effectuer les calculs sur l'espace image (souvent $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$). Le théorème de transfert justifie le passage de l'espace abstrait à l'espace image.

Définition 1.14
Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré, $(\Omega', \mathcal{A}')$ un espace mesurable et f une application mesurable de $(\Omega, \mathcal{A})$ dans $(\Omega', \mathcal{A}')$. On appelle mesure image de μ par f la mesure ν sur $(\Omega', \mathcal{A}')$ définie par

$$\forall A' \in \mathcal{A}',\ \ \nu(A') = \mu\left(f^{-1}(A')\right).$$

On remarque que la loi d'une variable aléatoire X définie sur $(\Omega, \mathcal{A}, \mathbf{P})$ est la mesure image de $\mathbf{P}$ par X.

Théorème 1.7 (Théorème de transfert)
On reprend les notation de la définition précédente. Soit h une application mesurable définie sur $(\Omega', \mathcal{A}')$. Alors h est ν-intégrable si et seulement si $h \circ f$ est μ-intégrable et on a de plus

$$\int_{\Omega'} h\,\mathrm{d}\nu = \int_\Omega h \circ f\,\mathrm{d}\mu.$$

Corollaire 1.1
Soit X une v.a. définie sur $(\Omega, \mathcal{A}, \mathbf{P})$ et à valeurs dans $(\Omega', \mathcal{A}')$ et de loi $\mathbf{P}_X$. Soit h une application de $(\Omega', \mathcal{A}')$ dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Alors l'espérance mathématique de $h \circ X$ existe si et seulement si h est $\mathbf{P}_X$-intégrable et on a

$$\mathbf{E}[h \circ X] = \mathbf{E}[h(X)] = \int_{\Omega'} h \, \mathrm{d}\mathbf{P}_X.$$

En particulier, si $(\Omega', \mathcal{A}') = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, on a

$$\mathbf{E}[h(X)] = \int_{\mathbb{R}} h(x) \, \mathrm{d}\mathbf{P}_X.$$

Exemple 1.5 (Loi discrète)
Si $\mathbf{P}_X \ll \mu_c$ sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ alors $\mathbf{E}[h(X)]$ existe si et seulement si $\sum_{n\in\mathbb{N}} p_n|h(n)| < +\infty$ et

$$\mathbf{E}[h(X)] = \sum_{n\in\mathbb{N}} p_n h(n).$$

Exemple 1.6 (Loi continue)
Si $\mathbf{P}_X \ll \lambda$ sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ alors $\mathbf{E}[h(X)]$ existe si et seulement si $h.f$ est λ-intégrable (f désigne la densité de X). On a alors

$$\mathbf{E}[h(X)] = \int_{\mathbb{R}} h(x) f(x) \, \mathrm{d}\lambda(x).$$

Chapitre 2

Variables aléatoires absolument continues

Une variable aléatoire réelle (v.a.r.) est une application $X : (\Omega, \mathcal{A}, \mathbf{P}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ telle que pour tout borélien B de $\mathcal{B}_{\mathbb{R}}$ l'ensemble $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$ appartient à $\mathcal{A}$. La tribu Borélienne étant engendrée par les intervalles $]-, \infty, b[$, $b \in \mathbb{R}$, on peut également définir une variable aléatoire réelle comme une application $X : (\Omega, \mathcal{A}, \mathbf{P}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ telle que pour tout réel b, $X^{-1}(]-\infty, b[) \in \mathcal{A}$. Il découle que les propriétés algébriques usuelles telles que la composition par une fonction réelle mesurable conservent la notion de variable aléatoire (si X et Y sont deux v.a.r. et $\lambda \in \mathbb{R}$ alors $\lambda X, X + Y, XY, |X|, X^n, \exp(X)$ sont des v.a.r.).

Si X est une v.a.r, on rappelle que l'application $\mathbf{P}_X : \mathcal{B}_{\mathbb{R}} \to [0, 1]$ définie par $\mathbf{P}_X(B) = \mathbf{P}\left(X^{-1}(B)\right)$ est une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Cette mesure de probabilité est appelée *loi de probabilité* de X. On dit également que X suit la loi de probabilité $\mathbf{P}_X$ et on note $X \sim \mathbf{P}_X$.

Remarque 2.1
Pour toute mesure de probabilité Q sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, il existe un espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$ et une variable aléatoire $X : (\Omega, \mathcal{A}, \mathbf{P}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ telle que $\mathbf{P}_X = Q$. On peut ainsi parler (c'est ce qui est fait en général) de v.a.r. X ayant une loi de probabilité $\mathbf{P}_X$ sans spécifier l'espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$ sur lequel X est défini. En revanche, il est clair que X n'est pas l'unique v.a.r. vérifiant $\mathbf{P}_X = Q$: une v.a.r. n'est donc pas déterminée par sa loi.

2.1 Fonction de répartition, fonction de masse, fonction de densité

2.1.1 Fonction de répartition d'une v.a.r.

Définition 2.1
Soit $X : (\Omega, \mathcal{A}, \mathbf{P}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ une v.a.r. On appelle fonction de répartition de X la fonction $F_X : \mathbb{R} \to [0, 1]$ définie par

$$F(x) = \mathbf{P}_X\left(]-\infty, x]\right) = \mathbf{P}(X \leq x).$$

Proposition 2.1
La fonction de répartition F d'une v.a.r. X satisfait les propriétés suivantes :

1. $\forall x \in \mathbb{R}$, $0 \leq F(x) \leq 1$;
2. F_X *est une fonction croissante, continue à droite en tout point* $x \in \mathbb{R}$;
3. $\lim_{x \to -\infty} F(x) = 0$ *et* $\lim_{x \to +\infty} F(x) = 1$.

Proposition 2.2
Pour toute fonction F vérifiant les 3 assertions de la proposition précédente, il existe une unique loi de probabilité Q sur $(\mathbb{R}, \mathcal{B}_R)$ telle que $F(x) = Q\left(]-\infty, x]\right)$.

On déduit de cette proposition que la loi d'une v.a.r. est entièrement déterminée par sa fonction de répartition.

2.1.2 Fonction de masse et classification des v.a.r.

Définition 2.2 (fonction de masse d'une v.a.r.)
Soit $X : (\Omega, \mathcal{A}, \mathbf{P}) \to (\mathbb{R}, \mathcal{B}_R)$ une v.a.r. On appelle fonction de masse de X la fonction

$$\begin{aligned} \pi_X : \mathbb{R} &\to [0,1] \\ x &\mapsto \mathbf{P}(X = x) = F_X(x) - F_X(x-) \end{aligned}$$

Proposition 2.3
Si X est une v.a.r., alors l'ensemble $\mathcal{D}_X$ des points de discontinuité de sa fonction de répartition vérifie

$$\mathcal{D}_X = \{x \in \mathbb{R} : \pi_X(x) > 0\}.$$

De plus cet ensemble est fini ou dénombrable.

Définition 2.3
1. *Si $\mathcal{D}_X = \emptyset$, on dit que la loi de X est diffuse.*
2. *Si $\mathbf{P}(X \in \mathcal{D}_X) = \sum_{x \in \mathcal{D}_X} \pi_X(x) = 1$, on dit que la loi de X est discrète (son support est $\mathcal{S}_X = \mathcal{D}_X$).*

On peut montrer aisément que pour $a, b \in \mathbb{R}$, on a

$$\mathbf{P}(X \in]a,b]) = F_X(b) - F_X(a), \quad \mathbf{P}(X \in [a,b]) = F_X(b) - F_X(a) + \pi_X(a),$$

$$\mathbf{P}(X \in]a,b[) = F_X(b) - \pi_X(b) - F_X(a), \quad \mathbf{P}(X \in [a,b[) = F_X(b) - \pi_X(b) - F_X(a) + \pi_X(a).$$

En particulier, si la loi de X est diffuse, on obtient

$$\mathbf{P}(X \in]a,b]) = \mathbf{P}(X \in [a,b]) = \mathbf{P}(X \in]a,b[) = \mathbf{P}(X \in [a,b[) = F_X(b) - F_X(a).$$

Proposition 2.4
Pour toute v.a.r. de loi de probabilité Q, il existe $\alpha \in [0,1]$, Q_1 une loi discrète et Q_2 une loi diffuse tels que $Q = \alpha Q_1 + (1-\alpha) Q_2$. Si $\alpha \in]0,1[$ la loi est dite mixte.

Exemple 2.1
1. Si X suit une loi uniforme sur $[0,1]$ alors

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

   Par conséquent $\mathcal{D}_X = \emptyset$ et la loi de X est donc diffuse.

2. Si X suit une loi de Bernoulli de paramètre $p \in]0,1[$.

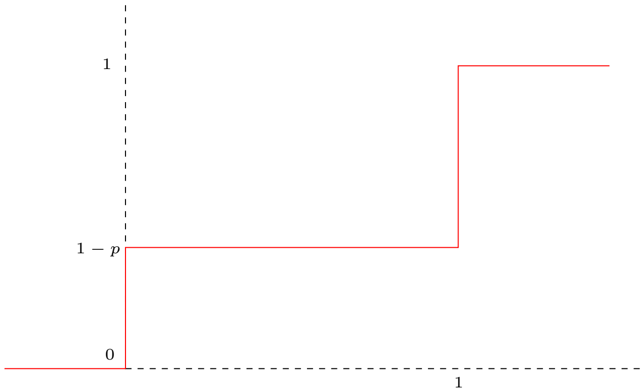


FIGURE 2.1 – Fonction de répartition de la loi de Bernoulli $\mathcal{B}(p)$.

On a clairement $\mathcal{D}_X = \{0,1\}$ et $\mathbf{P}(X \in \mathcal{D}_X) = 1$. X est donc une loi discrète.

3. Soit X une v.a.r. dont la fonction de répartition est donnée par

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 3x/4 & \text{si } 0 \leq x < 1/2 \\ 1/4 + 3x/4 & \text{si } 1/2 \leq x \leq 1 \\ 1 & \text{si } x \geq 1. \end{cases}$$

On a $\mathcal{D}_X = \{1/2\}$ et $\mathbf{P}(X \in \mathcal{D}_X) = 1/4 \neq 1$. X est une loi mixte :

$$P_X = \frac{1}{4}\delta_{1/2} + \frac{3}{4}U_{[0,1]}.$$

Les variables aléatoires discrètes ont été traitées en détail dans le cours de “Statistique 1”. Dans ce chapitre, nous allons nous intéresser à des v.a.r. de lois diffuses particulières : les v.a.r. de lois absolument continues. Les lois diffuses non absolument continues sont dites singulières.

2.1.3 Loi absolument continue - densité de probabilité

Avec un léger abus de langage, on dit que la loi de probabilité $\mathbf{P}_X$ d'une v.a.r. est absolument continue lorsqu'elle est absolument continue par rapport à la mesure de Lebesgue λ. D'après le théorème de Radon-Nikodym, pour ce type de loi de probabilité, il existe une fonction f_X positive et intégrable telle que pour tout $x \in \mathbb{R}$,

$$F_X(x) = \int_{]-\infty,x]} \mathrm{d}\mathbf{P}_X = \int_{]-\infty,x]} f_X(t)\,\mathrm{d}\lambda(t).$$

Cette fonction f_X est appelée *densité* de la loi P_X (ou par abus densité de X).

Proposition 2.5
Soit X une v.a.r absolument continue de fonction de répartition F_X et de densité f_X. Alors

1. *F_X est $\lambda-$p.p. dérivable avec $F'_X = f_X$ $\lambda-$p.p. ;*
2. *$\int_{\mathbb{R}} f_X(t)\,\mathrm{d}\lambda(t) = 1$ et $\mathbf{P}(X \in [a,b]) = \int_a^b f_X(t)\,\mathrm{d}\lambda(t)$;*

3. *f_X est unique à une λ-équivalence près. f_X caractérise donc la loi de X à une λ-équivalence près ;*
4. *Si F_X admet en tout point de $\mathbb{R}$ une dérivée continue, alors la loi de X est absolument continue. La réciproque est fausse : on peut juste dire que si $\mathbf{P}_X$ est absolument continue, alors $F(x)$ est continue.*
5. *La loi de X est diffuse et son support est défini par $\mathcal{S}_X = \{x \in \mathbb{R} : f_X(x) > 0\}$ (il existe des lois diffuses non absolument continues, voir par exemple l'escalier de Cantor).*

Proposition 2.6
Toute fonction réelle positive, d'une variable réelle, intégrable au sens de Lebesgue et telle que $\int_{\mathbb{R}} f(t)\,\mathrm{d}\lambda(t) = 1$ est la densité d'une loi de probabilité.

On rappelle que si μ_c désigne la mesure de comptage sur $\mathbb{Z}$ alors pour une loi $\mathbf{P}_X$ de support inclu dans $\mathbb{Z}$ on a $\mathbf{P}_X(B) = \sum_{x\in B} \pi_X(x) = \int_B \pi_X\,\mathrm{d}\mu_c$. Il est donc possible de regrouper l'étude des lois discrètes et celle des lois absolument continues. Cependant, d'un point de vue pratique (notamment calculatoire), l'étude de ces deux types de lois diffèrent (calcul d'espérance et de variance par exemple). C'est pourquoi on donnera à la suite de chaque résultat les techniques calculatoires selon que la variable est discrète ou absolument continue.

2.2 Espérance - moments d'une v.a.r.

2.2.1 Définition

On considère un espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$, X une v.a.r. définie sur cet espace de loi de probabilité $\mathbf{P}_X$.

Définition 2.4 (Espérance mathématique)
Soit X une v.a.r. $\mathbf{P}$-intégrable. On appelle espérance mathématique de X, notée $\mathbf{E}[X]$ l'intégrale de X par rapport à $\mathbf{P}$:

$$\mathbf{E}[X] = \int X\,\mathrm{d}\mathbf{P} = \int_\Omega X(\omega)\,\mathrm{d}\mathbf{P}(\omega).$$

Le théorème de transfert permet de calculer l'espérance mathématique sans connaître l'espace de probabilité sous-jacent.

Proposition 2.7
Soit X une v.a.r. $\mathbf{P}$-intégrable. On a alors

$$\mathbf{E}[X] = \int_{\mathbb{R}} x\,\mathrm{d}\mathbf{P}_X(x).$$

Remarque 2.2

1. L'espérance mathématique vérifie toutes les propriétés de l'intégrale, en particulier la linéarité et la monotonie ;
2. L'espérance mathématique s'interprète souvent comme un indicateur de centralité. On parle souvent de "valeur moyenne" de X ;
3. Si $A \in \mathcal{A}$, on a $\mathbf{P}(A) = \mathbf{E}[\mathbf{1}_A]$;

4. Soit φ une fonction réelle mesurable définie sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Alors $\varphi(X)$ est aussi une v.a.r. définie sur $(\Omega, \mathcal{A}, \mathbf{P})$. Par suite, si $\varphi(X)$ est $\mathbf{P}$-intégrable, elle possède une espérance

$$\mathbf{E}\left[\varphi(X)\right] = \int \varphi(X)\,\mathrm{d}\mathbf{P} = \int_{\mathbb{R}} \varphi(x)\,\mathrm{d}\mathbf{P}_X(x).$$

Exemple 2.2 (Calcul de l'espérance pour une loi discrète)
Si $\mathbf{P}_X$ est une loi discrète de fonction de masse π_X et de support S_X ($\mathbf{P}_X = \sum_{x\in\mathcal{S}_X} \pi_X(x)\delta_x$) et si $\sum_{x\in\mathcal{S}_X} |x|\pi_X(x) < +\infty$, alors $\mathbf{E}[X]$ existe et

$$\mathbf{E}[X] = \sum_{x\in\mathcal{S}_X} x\pi_X(x).$$

Exemple de la loi de Bernoulli ou du dé à 6 faces équiprobables.

Exemple 2.3 (Calcul de l'espérance pour une loi absolument continue)
Si $\mathbf{P}_X$ est une loi absolument continue, de densité f_X et si $x \mapsto |x|f_X(x)$ est λ-intégrable, alors $\mathbf{E}[X]$ existe et

$$\mathbf{E}[X] = \int_{\mathbb{R}} x f_X(x)\,\mathrm{d}\lambda(x).$$

Exemple de la loi uniforme ou de la loi normale.

Définition 2.5
Si $|\mathbf{E}[X]| < +\infty$, on dit que X a une espérance finie. De plus

1. *Si $\mathbf{E}[X]$ existe et $\mathbf{E}[X] = 0$, la v.a.r. X est dite centrée ;*
2. *Si $\mathbf{E}[X]$ existe, on dit qu'on centre ou recentre X lorsqu'on lui retranche $\mathbf{E}[X]$.*

Il est clair que $|\mathbf{E}[X]| \leq \mathbf{E}\left[|X|\right]$. La proposition suivante nous donne une inégalité plus générale.

Proposition 2.8 (Inégalité de Jensen)
Soit X une v.a.r. à valeurs dans un intervalle $]a, b[$ d'espérance finie et φ une fonction réelle convexe sur $]a, b[$. Alors

$$\varphi\left(\mathbf{E}[X]\right) \leq \mathbf{E}\left[\varphi(X)\right].$$

Rappels :

1. φ est convexe si et seulement si $\forall x_1, x_2$ le segment $[A_1, A_2]$ ($A_i = (x_i, \varphi(x_i)), i = 1, 2$) est situé au dessus de la courbe représentative de φ, c'est-à-dire $\forall x_1, x_2, \forall\lambda \in [0, 1]$

$$\varphi(\lambda x_1 + (1-\lambda)x_2) \leq \lambda\varphi(x_1) + (1-\lambda)\varphi(x_2).$$

2. φ est dite concave si et seulement si $-\varphi$ est convexe. On déduit que si φ est concave alors

$$\varphi\left(\mathbf{E}[X]\right) \geq \mathbf{E}\left[\varphi(X)\right].$$

On introduit maintenant d'autres caractéristiques d'une v.a.r. X qui permettent de rendre compte, entre autre, de sa dispersion.

Proposition 2.9
Soit r et $r' \in \mathbb{R}$ avec $0 < r < r'$. Si $\mathbf{E}\left[|X|^{r'}\right] < +\infty$ alors $\mathbf{E}\left[|X|^r\right] < +\infty$.

Définition 2.6 (Moments)

1. *Soit* $a, r \in \mathbb{R}$. *Si* $\mathbf{E}\left[|X|^r\right] < +\infty$, *on appelle moment d'ordre* r *de* X *la quantité* $\mathbf{E}\left[X^r\right] < +\infty$ *;*
2. *Lorsque* $\mathbf{E}\left[|X-a|^r\right] < +\infty$, *on appelle moment centré en* a *d'ordre* r *de* X *la quantité* $\mathbf{E}\left[(X-a)^r\right]$ *;*
3. *Lorsque* $\mathbf{E}\left[|X|\right] < +\infty$ *et* $a = \mathbf{E}[X]$, *on parle de moment centré de* X.

Définition 2.7 (Variance - écart-type)
Le moment centré d'ordre 2 de X *est appelé la variance de* X *et est noté* $\mathbf{V}[X]$ *:*

$$\mathbf{V}[X] = \mathbf{E}\left[(X - E[X])^2\right].$$

Sa racine carrée positive est appelée l'écart-type de X, *noté* $\sigma[X]$.

Exemple 2.4

1. Loi de Bernoulli : $\mathbf{V}[X] = p(1-p)$;
2. Loi uniforme sur $[0, 1]$: $\mathbf{V}[X] = 1/12$;
3. Loi uniforme sur $[1/4, 3/4]$: $\mathbf{V}[X] = 1/48$;

La variance peut s'interpréter comme une mesure de dispersion d'une v.a.r. On retrouve bien que la loi uniforme sur $[1/4, 3/4]$ est moins dispersée que la loi uniforme sur $[0, 1]$.

Proposition 2.10
Il découle de la définition précédente les assertions suivantes.

1. $\forall \alpha \in \mathbb{R}$, $\mathbf{V}[\alpha X] = \alpha^2 \mathbf{V}[X]$ *;*
2. $\forall a \in \mathbb{R}$, $\mathbf{V}[a + X] = \mathbf{V}[X]$ *;*
3. $\mathbf{V}[X] = 0$ *si et seulement si* X *est une v.a.r. presque sûrement constante (*$X = \mathbf{E}[X]$ *p.s.).*

Définition 2.8
Si $\mathbf{E}[X] = 0$ *et* $\mathbf{V}[X] = 1$, *on dit que* X *est centrée réduite.*

On remarque que lorsque $\mathbf{E}\left[|X|\right] < +\infty$ et $\mathbf{V}[X] < +\infty$, alors la v.a.r.

$$\frac{X - \mathbf{E}[X]}{\sigma[X]}$$

est centrée réduite.

La proposition suivante nous donne une autre manière (souvent utilisée) de calculer la variance.

Proposition 2.11
X *a un moment d'ordre 2 fini si et seulement si son espérance mathématique et sa variance existent et sont finies. On a alors*

$$\mathbf{V}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2.$$

2.2.2 Inégalités faisant intervenir les moments

Les moments permettent de donner une indication sur la dispersion d'une variable. Cette dernière peut par exemple être précisée à l'aide des inégalités suivantes.

Théorème 2.1 (Inégalité de Markov)
Si X est une v.a.r. positive, on a pour tout réel $a > 0$

$$\mathbf{P}(X \geq a) \leq \frac{\mathbf{E}[X]}{a}.$$

Preuve. On a

$$\mathbf{E}[X] = \int_{\mathbb{R}^+} x \, \mathrm{d}\mathbf{P}_X(x) = \int_{[0,a[} x \, \mathrm{d}\mathbf{P}_X(x) + \int_{[a,+\infty[} x \, \mathrm{d}\mathbf{P}_X(x).$$

D'où

$$\mathbf{E}[X] \geq \int_{[a,+\infty[} x \, \mathrm{d}\mathbf{P}_X(x) \geq a \int_{[a,+\infty[} \mathrm{d}\mathbf{P}_X(x).$$

■

Théorème 2.2 (Inégalité de Bienaymé-Chebychev)
Si $\mathbf{E}[X^2] < +\infty$, alors on a pour tout réel $a > 0$

$$\mathbf{P}(|X - \mathbf{E}[X]| > a) \leq \frac{\mathbf{V}[X]}{a^2}.$$

Preuve. Il suffit d'appliquer l'inégalité de Markov à la variable aléatoire $(X - \mathbf{E}[X])^2$.

■

La majoration de l'inégalité de Bienaymé-Chebychev n'est généralement pas très fine. Il est souvent préférable d'utiliser l'inégalité de Chernoff (voir TD1).

2.3 Médiane - quantiles d'une v.a.r.

L'espérance est une mesure de tendance centrale d'une v.a.r. Cependant, elle se révèle peu informative dans certains cas (il est par exemple assez facile de voir que l'espérance est sensible aux valeurs aberrantes). Il est souvent nécessaire d'étudier d'autres mesures de tendance centrale.

Définition 2.9 (Médiane)
On appelle médiane d'une v.a.r. X tout réel M qui vérifie

$$\mathbf{P}(X \leq M) \geq \frac{1}{2} \quad \textit{et} \quad \mathbf{P}(X \geq M) \geq \frac{1}{2}.$$

En particulier, lorsque la fonction de répartition F_X est strictement croissante, la médiane M est l'unique solution de $F_X(M) = \frac{1}{2}$.

Exemple 2.5

1. Si $\mathbf{P}_X = 1/4\delta_0 + 1/2\delta_1 + 1/4\delta_2$ alors $M = 1$;
2. Si $\mathbf{P}_X = 1/4\delta_0 + 1/4\delta_1 + 1/2\delta_2$ alors $M \in [1, 2[$;

3. Si $\mathbf{P}_X = U_{[0,1]}$ alors $M = 1/2$.

Définition 2.10 (Quantiles)
Pour $\alpha \in]0,1[$, on appelle quantile d'ordre α de X tout réel q_α qui vérifie

$$\mathbf{P}(X \leq q_\alpha) \geq \alpha \quad \textit{et} \quad \mathbf{P}(X \geq q_\alpha) \geq 1-\alpha.$$

On remarque que la médiane correspond au quantile d'ordre $\frac{1}{2}$. De même les quartiles de X correspondent aux quantiles d'ordre $\frac{1}{4}$ et $\frac{3}{4}$. De plus, lorsque la fonction de répartition F_X est strictement croissante, q_α est l'unique solution de $F_X(x) = \alpha$.

2.4 Calcul de lois

Etant donné une v.a.r. X de loi $\mathbf{P}_X$ et φ une fonction réelle mesurable, on se pose dans cette partie le problème de trouver la loi $\mathbf{P}_{\varphi(X)}$ de la v.a.r. $\varphi(X)$.

2.4.1 Cas discret

Pour une variable discrète, on cherche directement à exprimer $\mathbf{P}(\varphi(X) = y)$ pour tout $y \in \varphi(\mathcal{S}_X)$ à l'aide des propriétés de base des probabilités.

Exemple 2.6
Si la loi de X est donnée par le tableau 2.1

S_X	-1	0	1	2
π_X	1/5	1/5	2/5	1/5

TABLE 2.1 – Loi de X.

alors la loi de $Y = X^2$ est donnée par le tableau 2.2

S_Y	0	1	4
π_Y	1/5	3/5	1/5

TABLE 2.2 – Loi de $Y = X^2$.

2.4.2 Cas absolument continu

Utilisation de la fonction de répartition

La fonction de répartition détermine entièrement la loi d'une v.a.r. L'approche proposée consiste à exprimer la fonction de répartition $F_{\varphi(X)}$ en fonction de F_X. La fonction de densité s'obtient par $f_X = F'_X \ \lambda - \text{p.p.}$. On peut aussi utiliser cette approche pour déterminer la fonction de masse d'une variable discrète.

Exemple 2.7
Soit X une v.a.r. de loi uniforme sur $[0,1]$. On s'intéresse à la loi de $Y = X^2$. On a pour tout $t \in [0,1]$,

$$F_Y(t) = \mathbf{P}(Y \leq t) = \mathbf{P}(X^2 \leq t) = \mathbf{P}(X \leq \sqrt{T}) = \sqrt{t}.$$

La densité de Y est donc

$$f_Y(t) = \frac{1}{2\sqrt{t}} \mathbf{1}_{[0,1]}(t).$$

Utilisation de la mesure image

On suppose que le support de X est un ouvert U de $\mathbb{R}$, la densité de X est donc de la forme $f_X(x)\mathbf{1}_U(x)$. On suppose de plus que $\varphi : U \to V$ est un difféomorphisme (fonction bijective et différentiable). On a alors le théorème de changement de variable.

Théorème 2.3 (changement de variable)
Sous les hypothèse présentées ci dessus on a :

$$\int_V f(y)\,\mathrm{d}\lambda(y) = \int_U f(\varphi(x))|\varphi'(x)|\,\mathrm{d}\lambda(x)$$

et

$$\int_U f(x)\,\mathrm{d}\lambda(x) = \int_V f(\varphi^{-1}(y))|\left(\varphi^{-1}\right)'(y)|\,\mathrm{d}\lambda(y).$$

On obtient la densité de Y à l'aide du théorème suivant.

Théorème 2.4
La loi $\mathbf{P}_Y$ de Y est la mesure image de $\mathbf{P}_X$ par φ ($\mathbf{P}_Y = \mathbf{P}_X \circ \varphi^{-1}$). On déduit du théorème de changement de variable

$$f_Y(y) = f_X(\varphi^{-1}(y))|\left(\varphi^{-1}\right)'(y)|\mathbf{1}_V(y).$$

Exemple 2.8
Refaire l'exemple 2.7 en utilisant le théorème précédent.

Utilisation de la fonction muette

Comme pour toute fonction mesurable bornée h on a

$$\mathbf{E}[h(X)] = \int_{\mathbb{R}} h(x)f_X(x)\,\mathrm{d}\lambda(x),$$

il suffit de trouver une fonction g telle que pour toute fonction borélienne bornée h on ait

$$\mathbf{E}\left[h\left(\varphi(X)\right)\right] = \int_{\mathbb{R}} h(y)g(y)\,\mathrm{d}\lambda(y).$$

Par identification, on pourra alors affirmer que g est la densité de $\varphi(X)$. Cette fonction g n'est pas toujours facile à trouver. Cependant, dans le cas où le support de X est un ouvert U et φ est un difféomorphisme de U sur un ouvert V, la fonction g s'obtient “facilement” à l'aide du théorème de changement de variable :

$$\mathbf{E}\left[h\left(\varphi(X)\right)\right] = \int_U h\left(\varphi(x)\right)f_X(x)\,\mathrm{d}\lambda(x) = \int_V h(y)f_X\left(\varphi^{-1}(y)\right)\left|\left(\varphi^{-1}\right)'(y)\right|\,\mathrm{d}\lambda(y).$$

Par identification, la densité de $Y = \varphi(X)$ est donnée par $y \mapsto f_X\left(\varphi^{-1}(y)\right)\left|\left(\varphi^{-1}\right)'(y)\right|\mathbf{1}_V(y)$. On retrouve bien la densité du théorème 2.4.

Exemple 2.9
On se replace dans le cas de l'exemple 2.7. On a

$$\mathbf{E}\left[h(X^2)\right] = \int_0^1 h(x^2)\,\mathrm{d}x = \int_0^1 h(y)\frac{1}{2\sqrt{y}}\,\mathrm{d}y.$$

Chapitre 3

Vecteurs aléatoires

On généralise les notions introduites au chapitre 2 à des vecteurs aléatoires.

3.1 Généralités

On appelle variable aléatoire à valeurs dans $\mathbb{R}^n$ ou vecteur aléatoire à n dimensions toute application $X : (\Omega, \mathcal{A}, \mathbf{P}) \to (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ telle que pour tout borélien B de $\mathbb{R}^n$, $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$. Pour tout $i \in \{1, \dots, n\}$, on note π_i la projection canonique

$$\begin{aligned}\pi_i : \mathbb{R}^n &\to \mathbb{R}\\ x = (x_1, \dots, x_n)' &\mapsto x_i.\end{aligned}$$

On note $X_i = \pi_i(X), i = 1, \dots, n$. On a alors $X = (X_1, \dots, X_n)'$ et X est mesurable si et seulement si les v.a.r. X_i le sont. Les X_i sont des variables aléatoires réelles appelées variables aléatoires marginales.

Comme pour les v.a.r., l'application

$$\begin{aligned}\mathbf{P}_X : \mathcal{B}_{\mathbb{R}^n} &\to [0, 1]\\ B &\mapsto \mathbf{P}_X(B) = \mathbf{P}(X^{-1}(B))\end{aligned}$$

est une mesure de probabilité sur $\mathbb{R}^n$ appelée loi de probabilité ou loi conjointe du vecteur X. L'espace $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \mathbf{P}_X)$ est un nouvel espace de probabilité. De plus, pour toute mesure de probabilité Q sur $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$, il existe un espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$ et un vecteur aléatoire $X : (\Omega, \mathcal{A}, \mathbf{P}) \to (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ tel que $\mathbf{P}_X = Q$. Par conséquent, comme pour les v.a.r. on peut parler de vecteur aléatoire X ayant une loi conjointe $\mathbf{P}_X$ sans spécifier l'espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$ sur lequel X est défini.

Proposition 3.1
Pour tout borélien B de $\mathcal{B}_{\mathbb{R}}$, on a

$$\mathbf{P}_{X_i}(B) = \mathbf{P}_X(\pi_i^{-1}(B)), \quad i = 1, \dots, n.$$

Ainsi la loi conjointe du vecteur aléatoire X permet de déterminer les lois $\mathbf{P}_{X_1}, \dots, \mathbf{P}_{X_n}$ des variables aléatoires marginales $X_1, \dots, X_n$. Ces lois sont appelées lois marginales de X. La réciproque est cependant fausse : la connaissance des lois marginales ne déterminent pas la loi conjointe $\mathbf{P}_X$.

3.2 Fonctions de répartition - densités

3.2.1 Fonctions de répartition

Dans la suite, $X : (\Omega, \mathcal{A}, \mathbf{P}) \to (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ désigne un vecteur aléatoire à n dimensions.

Définition 3.1 (fonction de répartition conjointe)
On appelle fonction de répartition de X la fonction

$$\begin{aligned} F_X : \mathbb{R}^n &\to [0,1] \\ (x_1,\dots,x_n) &\mapsto \mathbf{P}_X(]-\infty,x_1]\times\ldots\times]-\infty,x_n]) = \mathbf{P}(X_1 \leq x_1,\dots,X_n \leq x_n). \end{aligned}$$

Proposition 3.2
La fonction de répartition F_X satisfait les propriétés suivantes :

1. *Pour tout $x \in \mathbb{R}^n$, $0 \leq F_X(x) \leq 1$;*
2. *F_X est une fonction croissante (au sens large), continue à droite en chacun de ses arguments ;*
3. *F_X tend vers 0 lorsque un de ses arguments x_i tend vers $-\infty$;*
4. *F_X tend vers 1 lorsque tous ses arguments x_i tendent vers $+\infty$.*

Proposition 3.3
La fonction de répartition conjointe caractérise la loi conjointe.

Proposition 3.4
Pour $i = 1,\dots,n$, on a

$$F_{X_i}(x_i) = \lim_{x_j\to\infty, j\neq i} F_X(x_1,\dots,x_n).$$

Ainsi, la fonction de répartition conjointe d'un vecteur aléatoire $X = (X_1,\dots,X_n)'$ détermine les fonctions de répartition $F_{X_1},\dots,F_{X_n}$ des lois marginales (ces fonctions de répartition sont appelées fonctions de répartition marginales).

Par commodité d'écriture, on ne considère dans la suite que des couples aléatoires ($n = 2$). Les notions s'étendent aisément au cas général.

3.2.2 Cas discret

Définition 3.2
La loi conjointe P_X d'un couple aléatoire $X = (X_1, X_2)'$ est dite discrète s'il existe un sous-ensemble $\mathcal{S}_X$ de $\mathbb{R}^2$ fini ou dénombrable tel que pour tout $(x_1, x_2) \in \mathcal{S}_X$ on a

$$\mathbf{P}_X(\{x_1,x_2\}) > 0 \quad \text{et} \quad \sum_{(x_1,x_2)\in\mathcal{S}_X} \mathbf{P}_X(\{x_1,x_2\}) = 1.$$

L'ensemble $\mathcal{S}_X$ est appelé le support de X.

Remarque 3.1
Un couple aléatoire a une loi discrète si et seulement si ses lois marginales sont discrètes.

Proposition 3.5 (Calcul des lois marginales)
Soit $X=(X_1,X_2)'$ un couple aléatoire de loi $\mathbf{P}_X$ discrète. On désigne par $\mathcal{S}_{X_i}$ le support des lois marginales X_i, $i=1,2$. Alors $\mathcal{S}_X \subset \mathcal{S}_{X_1} \times \mathcal{S}_{X_2}$ et

$$\forall x_1 \in \mathcal{S}_{X_1},\ \mathbf{P}_{X_1}(\{x_1\}) = \sum_{x_2 \in \mathcal{S}_{X_2}} \mathbf{P}_X(\{x_1,x_2\}) \quad \text{et} \quad \forall x_2 \in \mathcal{S}_{X_2},\ \mathbf{P}_{X_2}(\{x_2\}) = \sum_{x_1 \in \mathcal{S}_{X_1}} \mathbf{P}_X(x_1,x_2).$$

Exemple 3.1

1. Soit $X=(X_1,X_2)'$ un couple aléatoire dont la loi $\mathbf{P}_X$ est uniforme sur le support $\mathcal{S}_X = \{(-1,0),(0,1),(0,-1),(1,0)\}$, c'est-à-dire,

$$\mathbf{P}_X = \frac{1}{4}\delta_{(-1,0)} + \frac{1}{4}\delta_{(0,1)} + \frac{1}{4}\delta_{(0,-1)} + \frac{1}{4}\delta_{(1,0)}.$$

On obtient les lois marginales en sommant les lignes et colonnes du tableau 3.1.

X_2 \ X_1	-1	0	1	$\mathbf{P}_{X_2}$
-1	0	1/4	0	1/4
0	1/4	0	1/4	1/2
1	0	1/4	0	1/4
$\mathbf{P}_{X_1}$	1/4	1/2	1/4	

Table 3.1 – Loi conjointe et lois marginales de X.

2. On considère maintenant un couple aléatoire $Y=(Y_1,Y_2)'$ dont la loi est donnée par le tableau 3.2.

Y_2 \ Y_1	-1	0	1	$\mathbf{P}_{Y_2}$
-1	1/16	1/8	1/16	1/4
0	1/8	1/4	1/8	1/2
1	1/16	1/8	1/16	1/4
$\mathbf{P}_{Y_1}$	1/4	1/2	1/4	

Table 3.2 – Loi conjointe et lois marginales de Y.

On a clairement $\mathbf{P}_{X_1} = \mathbf{P}_{Y_1}$ et $\mathbf{P}_{X_2} = \mathbf{P}_{Y_2}$. Or $\mathbf{P}_X \neq \mathbf{P}_Y$, on retrouve bien sur ces deux exemples que les lois marginales ne caractérisent pas la loi conjointe.

3.2.3 Cas absolument continu

Définition 3.3
La loi conjointe $\mathbf{P}_X$ d'un couple aléatoire $X=(X_1,X_2)'$ est dite absolument continue si elle est absolument continue par rapport à la mesure de Lebesgue sur $(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2})$.

Théorème 3.1 (Radon-Nikodym)
La loi conjointe d'un couple aléatoire $X=(X_1,X_2)'$ est dite absolument continue si et seulement si il existe une fonction f_X positive et intégrable telle que pour tout borélien B de $\mathbb{R}^2$,

$$\mathbf{P}_X(B) = \int\int_B f_X(x_1,x_2)\,\mathrm{d}\lambda(x_1)\,\mathrm{d}\lambda(x_2). \tag{3.1}$$

Définition 3.4
On appelle densité conjointe de X toute fonction f_X qui vérifie (3.1).

Proposition 3.6

1. *Une densité intègre à 1 :*
$$\int_{\mathbb{R}^2} f_X(x_1, x_2)\,\mathrm{d}\lambda(x_1)\,\mathrm{d}\lambda(x_2) = 1;$$
2. *La fonction de répartition s'obtient en intégrant la densité comme suit :*
$$F_X(x_1, x_2) = \int_{]-\infty, x_2]} \int_{]-\infty, x_1]} f_X(x_1, x_2)\,\mathrm{d}\lambda(x_1)\,\mathrm{d}\lambda(x_2);$$
3. *Si f_X est continue en (x_1, x_2) alors*
$$f_X(x_1, x_2) = \frac{\partial F_X}{\partial x_1 \partial x_2}(x_1, x_2);$$
4. *Si F_X est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$, alors la loi de X est absolument continue et*
$$f_X(x_1, x_2) = \frac{\partial F_X}{\partial x_1 \partial x_2}(x_1, x_2);$$
5. *A une équivalence près pour la mesure de Lebesgue sur $\mathbb{R}^2$, la fonction f_X caractérise la loi jointe $\mathbf{P}_X$.*

Proposition 3.7
Toute fonction f_X de $\mathbb{R}^2$ dans $\mathbb{R}$, positive, intégrable et telle que
$$\int_{\mathbb{R}^2} f_X(x_1, x_2)\,\mathrm{d}\lambda(x_1)\,\mathrm{d}\lambda(x_2) = 1$$
est la densité d'une loi absolument continue sur $(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2})$.

Proposition 3.8 (Calcul des lois marginales)
Si le couple $X = (X_1, X_2)'$ admet une loi conjointe absolument continue, alors ses lois marginales sont absolument continues et sa densité détermine les densités des lois marginales (appelées densités marginales) :
$$f_{X_1}(x_1) = \int_{\mathbb{R}} f_X(x_1, x_2)\,\mathrm{d}\lambda(x_2), \quad f_{X_2}(x_2) = \int_{\mathbb{R}} f_X(x_1, x_2)\,\mathrm{d}\lambda(x_1).$$

Exemple 3.2

1. Soit X un couple aléatoire de loi uniforme sur $\mathcal{D} = [0,1]^2$. On a alors $f_X(x) = \mathbf{1}_{\mathcal{D}}(x)$. Déterminons la fonction de répartition ainsi que les lois marginales de X.
   — La fonction de répartition est donnée par
$$F_X(x) = \begin{cases} 0 & \text{si } x_1 < 0 \text{ ou } x_2 < 0 \\ x_1 x_2 & \text{si } 0 \leq x_1 \leq 1 \text{ et } 0 \leq x_2 \leq 1 \\ x_1 & \text{si } 0 \leq x_1 \leq 1 \text{ et } x_2 > 1 \\ x_2 & \text{si } 0 \leq x_2 \leq 1 \text{ et } x_1 > 1 \\ 1 & \text{si } x_1 > 1 \text{ et } x_2 > 1. \end{cases}$$

— On obtient les marginales à l'aide de la proposition 3.8 :

$$f_{X_1}(x_1)=\int_{\mathbb{R}}\mathbf{1}_{\mathcal{D}}(x_1,x_2)\,\mathrm{d}x_2=\mathbf{1}_{[0,1]}(x_1)\int_R\mathbf{1}_{[0,1]}(x_2)\,\mathrm{d}x_2=\mathbf{1}_{[0,1]}(x_1),$$

et $f_{X_2}(x_2)=\mathbf{1}_{[0,1]}(x_2)$.

2. Soit $Z=(X,Y)'$ un couple aléatoire dont la loi est donnée par la densité

$$f_Z(z)=\frac{xy}{2}\mathbf{1}_{0\leq y\leq x\leq 2}.$$

— Déterminons la fonction de répartition au point (x_0,y_0). Il faut distinguer plusieurs cas. Par exemple, si $0\leq y_0\leq x_0\leq 2$. La fonction de répartition s'obtient en intégrant la densité sur le domaine hachuré.

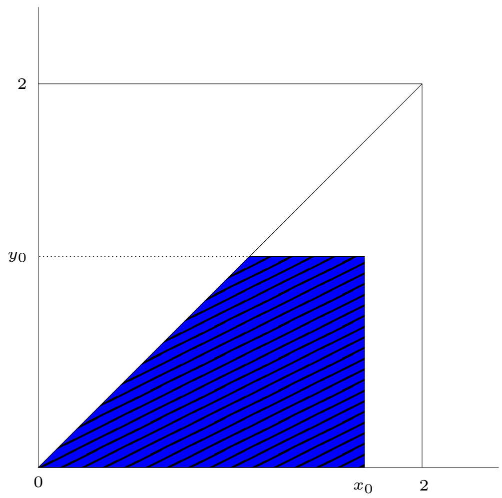


FIGURE 3.1 – Domaine d'intégration.

On peut calculer cette intégrale de deux manières différentes (au moins) :

$$F_Z(x_0,y_0)=\int_0^{y_0}\int_y^{x_0}\frac{xy}{2}\,\mathrm{d}x\,\mathrm{d}y=\frac{{y_0}^2}{16}(2x_0^2-y_0^2)$$

ou

$$F_Z(x_0,y_0)=\int_0^{y_0}\int_0^{x}\frac{xy}{2}\,\mathrm{d}y\,\mathrm{d}x+\int_{y_0}^{x_0}\int_0^{y_0}\frac{xy}{2}\,\mathrm{d}y\,\mathrm{d}x=\frac{{y_0}^2}{16}(2x_0^2-y_0^2).$$

En répétant ces calculs sur les différents ensembles, on obtient

$$F_Z(x,y)=\begin{cases}0 & \text{si } x<0 \text{ ou } y<0\\ \frac{y^2}{16}(2x^2-y^2) & \text{si } 0\leq y\leq x\leq 2\\ \frac{y^2}{16}(8-y^2) & \text{si } 0\leq y\leq 2\leq x\\ \frac{x^4}{16} & \text{si } 0\leq x\leq y \text{ et } x\leq 2\\ 1 & \text{si } x\geq 2 \text{ et } y\geq 2.\end{cases}$$

— Les marginales sont

$$f_X(x)=\frac{x^3}{4}\mathbf{1}_{[0,2]}(x)\quad\text{et}\quad f_Y(y)=\frac{y}{4}(4-y^2)\mathbf{1}_{[0,2]}(y).$$

— Calcul de $\mathbf{P}(Y > X^2)$. On a

$$\mathbf{P}(Y > X^2) = \int_{\mathbb{R}} \int_{\mathbb{R}} f_Z(x,y) \mathbf{1}_{y \geq x^2} \, \mathrm{d}x \, \mathrm{d}y.$$

Il suffit donc d'intégrer la densité sur le domaine représenté sur la figure 3.2.

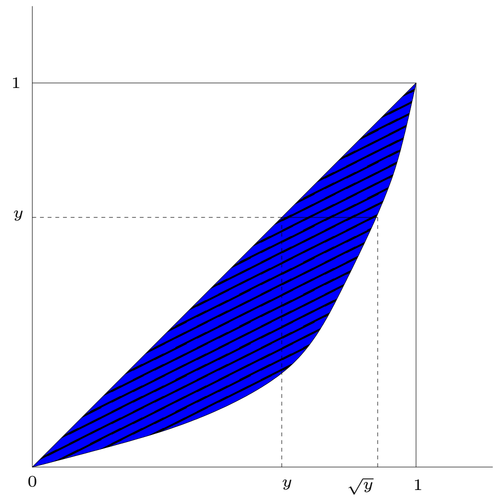


FIGURE 3.2 – Domaine d'intégration.

On obtient

$$\mathbf{P}(Y > X^2) = \int_0^1 \int_y^{\sqrt{y}} \frac{xy}{2} \, \mathrm{d}x \, \mathrm{d}y = \frac{1}{48}.$$

3.3 Espérance - moments d'un vecteur aléatoire

3.3.1 Espérance

Définition 3.5

On dit que $X = (X_1, \dots, X_n)'$ possède une espérance mathématique si ses variables marginales $X_1, \dots, X_n$ possèdent une espérance mathématique. On pose alors

$$\mathbf{E}[X] = (\mathbf{E}[X_1], \dots, \mathbf{E}[X_n]).$$

Les propriétés de l'espérance mathématiques pour les vecteurs aléatoires découlent donc directement de celles de l'espérance mathématiques des v.a.r. En particulier, on a le résultat de linéarité suivant.

Théorème 3.2

Soit A une matrice réelle $p \times n$. Alors

$$\mathbf{E}[AX] = A\mathbf{E}[X].$$

Le théorème de transfert pour les vecteurs aléatoires permet de calculer l'espérance d'une v.a.r. $\varphi(X)$.

Théorème 3.3 (Théorème de transfert)
Soit φ une fonction mesurable de $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Alors $\varphi(X)$ est une v.a.r. $\mathbf{P}$-intégrable si et seulement si φ est $\mathbf{P}_X$-intégrable. On a alors

$$\mathbf{E}[\varphi(X)] = \int_{\Omega} \varphi(X(\omega))\, \mathrm{d}\mathbf{P}(\omega) = \int_{\mathbb{R}^n} \varphi(x_1, \ldots, x_n)\, \mathrm{d}\mathbf{P}_X(x_1, \ldots, x_n).$$

Exemple 3.3
Soit $X = (X_1, X_2)'$ qui suit une loi uniforme sur $[0,1] \times [0,1]$. Calculons $\mathbf{E}[X_1 + X_2]$. D'après le théorème précédent, on a

$$\mathbf{E}[X_1 + X_2] = \int_0^1 \int_0^1 (x_1 + x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2 = 1.$$

On aurait aussi pu faire directement $\mathbf{E}[X_1 + X_2] = \mathbf{E}[X_1] + \mathbf{E}[X_2] = 1$.

3.3.2 Variance - covariance

Interprétation de la variance $\mathbf{V}[X_1]$. Pour simplifier on se place dans le cas $n = 2$. Si X_1 admet un moment d'ordre 2, on rappelle que

$$\mathbf{V}[X_1] = \mathbf{E}\left[([X_1] - \mathbf{E}[X_1])^2\right].$$

On note M le points de coordonnées (X_1, X_2) et G celui de coordonnées $(\mathbf{E}[X_1], \mathbf{E}[X_2])$.

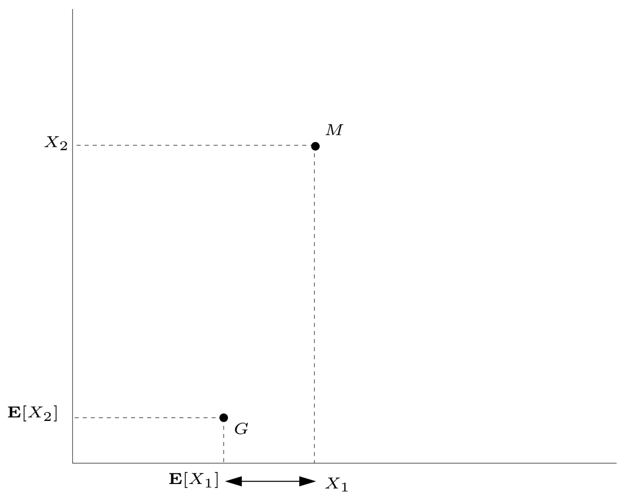


FIGURE 3.3 – Interprétation de $\mathbf{V}[X_1]$.

Nous voyons sur la figure 3.3 que $\mathbf{V}[X_1]$ mesure la dispersion de la distance entre les projetés des points G et M sur le premier axe. Il est nécessaire de chercher à caractériser la dispersion dans toutes les directions.

Soit u un vecteur unitaire de coordonnées (α, β). On cherche à caractériser la dispersion entre M et G sur la direction engendrée par u. On note H le projeté orthogonal de M sur $\mathsf{Vect}(u)$ (voir figure 3.4).

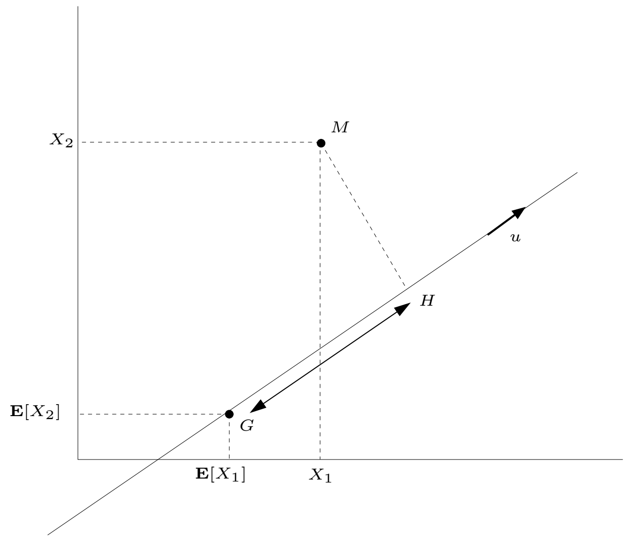


FIGURE 3.4 – Dispersion dans la direction engendrée par u.

On a par construction :

$$\overline{GH} = \alpha(X_1 - \mathbf{E}[X_1]) + \beta(X_2 - \mathbf{E}[X_2]).$$

Par conséquent $\mathbf{E}[\overline{GH}] = 0$ et

$$\begin{aligned}\mathbf{V}[\overline{GH}] =& \mathbf{E}\left[(\alpha(X_1 - \mathbf{E}[X_1]) + \beta(X_2 - \mathbf{E}[X_2]))^2\right] \\ =& \alpha^2\mathbf{V}[X_1] + \beta^2\mathbf{V}[X_2] + 2\alpha\beta\mathbf{Cov}(X_1, X_2) \\ =& \begin{pmatrix}\alpha & \beta\end{pmatrix}\begin{pmatrix}\mathbf{V}[X_1] & \mathbf{Cov}(X_1, X_2) \\ \mathbf{Cov}(X_1, X_2) & \mathbf{V}(X_2)\end{pmatrix}\begin{pmatrix}\alpha \\ \beta\end{pmatrix}\end{aligned}$$

avec $\mathbf{Cov}(X_1, X_2) = \mathbf{E}\left[(X_1 - \mathbf{E}[X_1])(X_2 - \mathbf{E}[X_2])\right]$. Cette espérance existe puisque

$$|(X_1 - \mathbf{E}[X_1])(X_2 - \mathbf{E}[X_2])| \leq \frac{1}{2}\left[(X_1 - \mathbf{E}[X_1])^2 + (X_2 - \mathbf{E}[X_2])^2\right].$$

La matrice permet de mesurer la dispersion du projeté de M dans toutes les directions de $\mathbb{R}^2$.

Remarque 3.2
En prenant $u = (1, 0)'$, on retrouve bien $\mathbf{V}[\overline{GH}] = \mathbf{V}[X_1]$.

Définition 3.6

1. *Un vecteur aléatoire $X = (X_1, \ldots, X_n)'$ dont toutes les variables marginales possèdent un moment d'ordre 2, i.e. $\mathbf{E}[X_i^2] < +\infty$, est dit du second ordre.*
2. *On appelle alors matrice des moments du second ordre de X la matrice carrée M_X de taille $n \times n$ de terme général $m_{ij} = \mathbf{E}[X_iX_j]$.*
3. *On appelle matrice des variance-covariance de X la matrice carrée Σ_X de taille $n \times n$ de terme général*

$$\sigma_{ij}^2 = \mathbf{E}[(X_i - \mathbf{E}[X_i])(X_j - \mathbf{E}[X_j])] = \mathbf{E}[X_iX_j] - \mathbf{E}[X_i]\mathbf{E}[X_j].$$

σ_{ij}^2 s'appelle la covariance entre X_i et X_j et est noté $\mathbf{Cov}(X_i, X_j)$.

Remarque 3.3

1. Les matrices M_X et Σ_X sont des matrices symétriques ;
2. Si toutes les variables marginales X_i sont centrées alors $M_X = \Sigma_X$;
3. $\mathbf{V}(X_i) = \mathbf{Cov}(X_i, X_i)$, $i = 1, \ldots, n$, donc les éléments de la diagonale de Σ_X sont les variances des variables marginales.
4. Si $\Sigma_X = \mathsf{Id}$, on dit que le vecteur aléatoire X est réduit.
5. Les matrices M_X et Σ_X peuvent s'écrire
$$M_X = \mathbf{E}[XX'] \quad \text{et} \quad \Sigma_X = \mathbf{E}[(X - \mathbf{E}[X])(X - \mathbf{E}[X])'].$$
6. Si A est une matrice réelle de dimension $p \times n$, on déduit de la linéarité de l'espérance
$$M_{AX} = AM_XA' \quad \text{et} \quad \Sigma_{AX} = A\Sigma_X A'.$$

Proposition 3.9
Une matrice symétrique Σ est la matrice de variance-covariance d'un vecteur aléatoire si et seulement si elle est positive ($u'\Sigma u \geq 0\ \forall u$).

Proposition 3.10
Soit X un vecteur aléatoire dont la matrice de variance-covariance Σ_X est inversible. Alors le vecteur $\Sigma_X^{-1/2} X$ est réduit et le vecteur $\Sigma_X^{-1/2}(X - \mathbf{E}[X])$ est centré réduit.

Proposition 3.11
Soit $X = (X_1, \ldots, X_n)'$ un vecteur aléatoire du second ordre. Alors
$$\mathbf{E}[X_1 + \ldots + X_n] = \mathbf{E}[X_1] + \ldots + \mathbf{E}[X_n]$$
et
$$\mathbf{V}[X_1 + \ldots + X_n] = \sum_{i=1}^{n} \mathbf{V}[X_i] + \sum_{1 \leq i \neq j \leq n} \mathbf{Cov}(X_i, X_j) = \sum_{1 \leq i,j \leq n} \mathbf{Cov}(X_i, X_j).$$
En particulier, si $\forall i \neq j, \mathbf{Cov}(X_i, X_j) = 0$, on a
$$\mathbf{V}[X_1 + \ldots + X_n] = \sum_{i=1}^{n} \mathbf{V}[X_i]. \tag{3.2}$$

3.4 Corrélation

Il est facile de voir que $\mathbf{Cov}(X_1 + a, X_2 + b) = \mathbf{Cov}(X_1, X_2)$. Par contre, $\mathbf{Cov}(aX_1, bX_2) = ab\mathbf{Cov}(X_1, X_2)$, il est ainsi difficile d'interpréter la valeur d'une covariance.

3.4.1 Le coefficient de corrélation linéaire

Définition 3.7

1. *Soit $(X_1, X_2)'$ un couple de variables aléatoires du second ordre tel que $\mathbf{V}(X_1)\mathbf{V}(X_2) > 0$. On appelle coefficient de corrélation (linéaire) du couple $(X_1, X_2)'$ le réel*
$$\rho(X_1, X_2) = \frac{\mathbf{Cov}(X_1, X_2)}{\sqrt{\mathbf{V}(X_1)\mathbf{V}(X_2)}}.$$

2. *Si* $\mathbf{Cov}(X_1, X_2) \neq 0$*, on dit que* X_1 *et* X_2 *sont corrélées.*
3. *Si* $\mathbf{Cov}(X_1, X_2) = 0$*, on dit que* X_1 *et* X_2 *sont non corrélées.*

Théorème 3.4

1. *Le coefficient de corrélation linéaire est compris entre -1 et 1, i.e.,* $|\rho(X_1, X_2)| \leq 1$.
2. *Si* $|\rho(X_1, X_2)| = 1$ *alors les variables* X_1 *et* X_2 *sont liées par une relation affine, i.e., il existe* $(a, b) \in \mathbb{R}^2$ *tel que* $X_2 = aX_1 + b$.
3. *Le coefficient de corrélation linéaire est invariant par transformation affine* $x \mapsto ax+b$ *avec* $a > 0$.

La première assertion du théorème est une conséquence directe de l'inégalité de Cauchy-Schwartz

$$(\mathbf{E}[X_1X_2])^2 \leq \mathbf{E}[X_1^2]\mathbf{E}[X_2^2] \Longrightarrow \mathbf{Cov}(X_1, X_2)^2 \leq \mathbf{V}[X_1]\mathbf{V}[X_2].$$

Remarque 3.4
Le coefficient de corrélation linéaire permet, comme son nom l'indique, de mettre en évidence une relation *linéaire* (et uniquement linéaire !) entre deux variables. En effet, si $X \sim \mathcal{N}(0, 1)$ et $Y = X^2$ alors

$$\mathbf{Cov}(X, Y) = \mathbf{E}[X^3] - \mathbf{E}[X]\mathbf{E}[X^2] = 0.$$

Par conséquent $\rho(X, Y) = 0$ alors qu'il existe une relation entre X et Y $(Y = X^2)$.

3.4.2 Interprétation hilbertienne

On considère l'espace $L_2(\Omega)$ des fonctions de carré intégrable muni du produit scalaire $< X_1, X_2 >= \mathbf{E}[X_1X_2]$. Cet espace est un espace de Hilbert (on montre qu'il est complet pour la norme associée $(\mathbf{E}[X_1^2])^{1/2}$). On a alors

$$\mathbf{Cov}(X_1, X_2) =< X_1 - \mathbf{E}[X_1], X_2 - \mathbf{E}[X_2] > \quad \text{et} \quad \rho(X_1, X_2) = \cos(X_1 - \mathbf{E}[X_1], X_2 - \mathbf{E}[X_2]).$$

Cette structure hilbertienne permet de retrouver les résultats suivants :

- $\mathbf{Cov}(X_1, X_2) = \mathbf{Cov}(X_2, X_1)$ et $\mathbf{Cov}(\lambda X_1 + X_2, X_3) = \lambda\mathbf{Cov}(X_1, X_3) + \mathbf{Cov}(X_2, X_3)$;
- $\rho(X_1, X_2) \leq 1$ et $|\rho(X_1, X_2)| = 1$ si X_1 et X_2 sont liés par une relation affine.
- $\mathbf{Cov}(X_1, X_2) = 0 \Longleftrightarrow (X_1 - \mathbf{E}[X_1]) \perp (X_2 - \mathbf{E}[X_2])$ et l'équation (3.2) s'interprète comme le théorème de Pythagore. Par abus de langage, on dit dans ce cas que X_1 et X_2 sont orthogonales et on note $X_1 \perp X_2$.

3.5 Variables (ou vecteurs) aléatoires indépendant(e)s

Définition 3.8 (Rappels)

1. *Deux évènements* A_1 *et* A_2 *de* $\mathcal{A}$ *sont indépendants si* $\mathbf{P}(A_1 \cap A_2) = \mathbf{P}(A_1)\mathbf{P}(A_2)$.
2. *Soit* $(A_i)_{i \in I}$ *une suite (finie ou infinie) d'évènements distincts. Les évènements* $A_i, i \in I$ *sont dits (mutuellement) indépendants si pour toute collection finie* $A_{i_1}, \ldots, A_{i_k}$ *d'évènements distincts on a*

$$\mathbf{P}(A_{i_1} \cap \ldots \cap A_{i_k}) = \mathbf{P}(A_{i_1}) \ldots \mathbf{P}(A_{i_k}).$$

3. *Deux classes d'évènements $\mathcal{C}_1$ et $\mathcal{C}_2$ sont indépendantes si pour tout évènement A_1 de $\mathcal{C}_1$ et A_2 de $\mathcal{C}_2$, A_1 et A_2 sont indépendants.*
4. *Soit $(\mathcal{C}_i)_{i\in I}$ une suite (finie ou infinie) de classes d'évènements de $\mathcal{A}$. On dit que les classes $\mathcal{C}_i, i \in I$ sont (mutuellement) indépendantes si pour toute suite finie $\mathcal{C}_{i_1}, \dots, \mathcal{C}_{i_k}$ extraite de la suite $\mathcal{C}_i, i \in I$ et pour toute suite d'évènements $A_{i_1} \in \mathcal{C}_{i_1}, \dots, A_{i_k} \in \mathcal{C}_{i_k}$, on a*

$$\mathbf{P}(A_{i_1} \cap \ldots \cap A_{i_k}) = \mathbf{P}(A_{i_1}) \ldots \mathbf{P}(A_{i_k}).$$

Théorème 3.5
Si les classes $\mathcal{C}_i, i \in I$ sont indépendantes et stables par intersection alors les tribus engendrées par $\mathcal{C}_i, i \in I$ sont indépendantes.

On rappelle que la tribu engendrée par un vecteur aléatoire X de $\mathbb{R}^n$ est $\sigma(X) = X^{-1}(\mathcal{B}_{\mathbb{R}^n})$. Nous pouvons donc prolonger la notion d'indépendance aux vecteurs aléatoires.

Définition 3.9

1. *Deux vecteurs aléatoires réels X et Y définis sur $(\Omega, \mathcal{A}, \mathbf{P})$ sont dits indépendants si les tribus $\sigma(X)$ et $\sigma(Y)$ sont indépendantes. En particulier, lorsque X est à n dimensions et Y à m dimensions, X et Y sont dits indépendants si pour tout $A \in \mathcal{B}_{\mathbb{R}^n}$, $B \in \mathcal{B}_{\mathbb{R}^m}$, on a*

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

2. *Soit $(X_i)_{i\in I}$ une suite de vecteurs aléatoires définis sur $(\Omega, \mathcal{A}, \mathbf{P})$. Les vecteurs $X_i, i \in I$ sont dits (mutuellement) indépendants si les tribus $(\sigma(X_i))_{i\in I}$ engendrées par ces vecteurs sont (mutuellement) indépendantes. En particulier, $X_i, i \in I$ sont dits (mutuellement) indépendants si pour toute suite finie $X_{i_1}, \dots, X_{i_k}$ extraite de $(X_i)_{i\in I}$ et pour toute suite finie de boréliens $B_{i_1}, \dots, B_{i_k}$, on a*

$$\mathbf{P}(X_{i_1} \in B_{i_1}, \dots, X_{i_k} \in B_{i_k}) = \mathbf{P}(X_{i_1} \in B_{i_1}) \ldots \mathbf{P}(X_{i_k} \in B_{i_k}).$$

Théorème 3.6
Soit $X_1, \dots, X_n$ n v.a.r. mutuellement indépendantes. Alors :

1. *Les vecteurs*
$$(X_1, \dots, X_{n_1}), (X_{n_1+1}, \dots, X_{n_2}), \dots, (X_{n_k+1}, \dots, X_n)$$
sont mutuellement indépendants ;
2. *Soit $g_1, \dots, g_n$ n fonctions réelles mesurables. Alors les v.a.r. $g_1(X_1), \dots, g_n(X_n)$ sont mutuellement indépendantes.*

Par conséquent si X_1, X_2 et X_3 sont des variables indépendantes, on déduit :

$$X_1 \perp\!\!\!\perp X_2, \qquad (X_1, X_2)' \perp\!\!\!\perp X_3, \qquad (X_1 + X_2) \perp\!\!\!\perp X_3, \qquad \ldots$$

Dans la suite de cette partie, on désigne par $X = (X_1, \dots, X_n)'$ un vecteur aléatoire à n dimensions, de loi conjointe $\mathbf{P}_X$ et de fonction de répartition F_X. On note $\mathbf{P}_{X_1}, \dots, \mathbf{P}_{X_n}$ les lois marginales et $F_{X_1}, \dots, F_{X_n}$ les fonctions de répartition des lois marginales. Les théorèmes suivants donnent des conditions simples pour vérifier que les v.a.r. $X_1, \dots, X_n$ sont indépendantes.

Théorème 3.7
Les variables marginales $X_1, \ldots, X_n$ sont (mutuellement) indépendantes si et seulement si pour tous réels $x_1, \ldots, x_n$, on a

$$\mathbf{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = \mathbf{P}(X_1 \leq x_1) \ldots \mathbf{P}(X_n \leq x_n),$$

ou encore, si et seulement si

$$F_X(x_1, \ldots, x_n) = F_{X_1}(x_1) \ldots F_{X_n}(x_n).$$

Théorème 3.8
Les variables marginales $X_1, \ldots, X_n$ sont (mutuellement) indépendantes si et seulement si la loi conjointe $\mathbf{P}_X$ est le produit des lois marginales $\mathbf{P}_{X_1} \otimes \ldots \otimes \mathbf{P}_{X_n}$. Dans ce cas, si $B_1, \ldots, B_n$ sont des boréliens de $\mathbb{R}$ et φ est une fonction $\mathbf{P}_X$-intégrable, on a d'après Fubini

$$\int_{B_1 \times \ldots \times B_n} \varphi(x_1, \ldots, x_n) \,\mathrm{d}\mathbf{P}_X(x_1, \ldots, x_n) = \int_{B_n} \ldots \int_{B_1} \varphi(x_1, \ldots, x_n) \,\mathrm{d}\mathbf{P}_{X_1}(x_1) \ldots \mathrm{d}\mathbf{P}_{X_n}(x_n).$$

Dans le cas particulier où les vecteurs aléatoires sont de loi absolument continue, on peut montrer à l'aide du théorème de Fubini-Tonelli le résultat suivant.

Théorème 3.9
Si les variables marginales $X_1, \ldots, X_n$ sont indépendantes, de lois marginales absolument continues et si l'on note $f_{X_1}, \ldots, f_{X_n}$ les densités marginales respectives, alors la densité conjointe f_X de X est de la forme

$$f_X(x) = f_{X_1}(x_1) \ldots f_{X_n}(x_n).$$

Réciproquement, si le vecteur aléatoire $X = (X_1, \ldots, X_n)'$ possède une densité conjointe de la forme

$$f_X(x_1, \ldots, x_n) = f_1(x_1) \ldots f_n(x_n),$$

alors les variables marginales sont indépendantes et leurs densités respectives sont égales, à une constante multiplicative près, à $f_1, \ldots, f_n$.

Exemple 3.4

1. Soit $X = (X_1, X_2)'$ de loi uniforme sur le carré $\mathcal{D} = [-1, 1]^2$. On a alors

$$f_X(x) = \frac{1}{4}\mathbf{1}_{\mathcal{D}}(x) = \frac{1}{4}\mathbf{1}_{[-1,1]}(x_1)\mathbf{1}_{[-1,1]}(x_2) = f_1(x_1)f_2(x_2)$$

   avec $f_1(x_1) = \frac{1}{2}\mathbf{1}_{[-1,1]}(x_1)$ et $f_2(x_2) = \frac{1}{2}\mathbf{1}_{[-1,1]}(x_2)$.

2. Pour le deuxième vecteur de l'exemple 3.2, on a clairement $f_Z(x, y) \neq f_X(x)f_Y(y)$. X et Y ne sont donc pas indépendantes.

Théorème 3.10
Deux v.a.r. X_1 et X_2 sont indépendantes si et seulement si pour toutes fonctions mesurables g_1 et g_2, on a

$$\mathbf{E}[g_1(X_1)g_2(X_2)] = \mathbf{E}[g_1(X_1)]\mathbf{E}[g_2(X_2)].$$

On déduit le corollaire suivant.

Corollaire 3.1 (Indépendance $\Longrightarrow$ covariance nulle)
Soit X_1 et X_2 deux variables aléatoires réelles indépendantes admettant des espérances mathématiques. Alors le produit X_1X_2 admet aussi une espérance mathématique et

$$\mathbf{E}[X_1X_2] = \mathbf{E}[X_1]\mathbf{E}[X_2].$$

Proposition 3.12
Si X_1 et X_2 sont deux variables aléatoires réelles indépendantes ayant des moments d'ordre 2 finis, alors $\mathbf{Cov}(X_1, X_2) = 0$.

Remarque 3.5
On déduit que deux variables indépendantes sont non corrélées. Mais **attention** la réciproque est fausse. On a vu que pour $X \sim \mathcal{N}(0,1)$ et $Y = X^2$, on a $\mathbf{Cov}(X,Y) = 0$. Or X et Y ne sont clairement pas indépendantes, il suffit par exemple de voir que pour $x > 0$ on a

$$\mathbf{P}(Y \leq x^2, |X| \leq x) = \mathbf{P}(|X| \leq x) \neq \mathbf{P}(Y \leq x^2)\mathbf{P}(|X| \leq x).$$

On déduit facilement que la variance de la somme de v.a.r. indépendantes est égale à la somme des variances.

Proposition 3.13
Soit $X_1, \dots, X_n$ n v.a.r. indépendantes ayant des moments d'ordre 2 finis, alors

$$\mathbf{V}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbf{V}(X_i). \tag{3.3}$$

On peut remarquer que (3.3) reste vrai si les n v.a.r. sont (seulement) deux à deux non corrélées.

3.6 Calcul de loi

Soit $X = (X_1, \dots, X_n)'$ un vecteur aléatoire de loi $\mathbf{P}_X$ et de fonction de répartition F_X. On considère $\varphi : \mathbb{R}^n \to \mathbb{R}^p$ une fonction mesurable. On se pose la problème de calculer la loi de $Y = \varphi(X)$.

3.6.1 Cas discret

Pour une loi discrète, l'approche s'effectue généralement en deux temps :

1. détermination du support $\mathcal{S}_Y$ de Y ;
2. calcul des probabilités $\mathbf{P}(Y = y) = \mathbf{P}(\varphi(X) = y)$ pour tout $y = (y_1, \dots, y_p) \in \mathcal{S}_Y$.

3.6.2 Cas absolument continue

Utilisation de la fonction de répartition

L'idée consiste à exprimer la fonction de répartition F_Y en fonction de F_X :

$$\begin{aligned} F_Y(y) &= \mathbf{P}(\varphi(X) \in]-\infty, y_1] \times \ldots \times]-\infty, y_p]) \\ &= \int \mathbf{1}_{\varphi(x) \in]-\infty, y_1] \times \ldots \times]-\infty, y_p]} \, \mathrm{d}\mathbf{P}_X(x_1, \dots, x_n). \end{aligned}$$

Une fois cette intégrale calculée, on obtient la densité de Y en dérivant F_Y par rapport à ses p variables :

$$f_Y(y) = \frac{\partial F_Y}{\partial y_1 \dots \partial y_p}(y_1, \dots, y_p).$$

Le calcul de F_Y est souvent long. On préfère utiliser les méthodes suivantes.

Utilisation de la mesure image

On se place ici dans le cas où $\mathcal{S}_X$ est un ouvert U inclu dans $\mathbb{R}^n$ (la densité s'écrit $f_X(x)\mathbf{1}_U(x)$) et on suppose que $\varphi : U \to V$ est une C^1-difféomorphisme (différentiable, bijective et dont les dérivées partielles sont continues). On a alors le théorème de changement de variable suivant.

Théorème 3.11

Sous les hypothèses ci-dessus, on a pour toute fonction f intégrable

$$\int_V f(y)\,\mathrm{d}\lambda(y) = \int_U f(\varphi(x))|\det(J_{\varphi_x})|\,\mathrm{d}\lambda(x)$$

et

$$\int_U f(x)\,\mathrm{d}\lambda(x) = \int_V f(\varphi^{-1}(y))|\det(J_{\varphi_y^{-1}})|\,\mathrm{d}\lambda(y)$$

où J_{φ_x} désigne la matrice Jacobienne au point x

$$J_{\varphi_x} = \begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1}(x) & \dots & \frac{\partial \varphi_1}{\partial x_n}(x) \\ \vdots & \vdots & \vdots \\ \frac{\partial \varphi_n}{\partial x_1}(x) & \dots & \frac{\partial \varphi_n}{\partial x_n}(x) \end{pmatrix}.$$

On déduit alors la densité de Y.

Théorème 3.12

$\mathbf{P}_Y$ *est la mesure image de* $\mathbf{P}_X$ *par* φ *(*$\mathbf{P}_Y = \mathbf{P}_X \circ \varphi^{-1}$*). On a donc*

$$f_Y(y) = f_X(\varphi^{-1}(y))|\det(J_{\varphi_y^{-1}})|\mathbf{1}_V(y).$$

Remarque 3.6

Il est indispensable que φ soit bijective. Si par exemple V est un ouvert de $\mathbb{R}^p$ avec $p < n$, on ne peut utiliser directement le théorème précédent. Un moyen de pallier à cette difficulté consiste à "compléter" φ en une fonction bijective : on définit $\tilde{\varphi} = (\varphi_1, \dots, \varphi_p, \tilde{\varphi}_{p+1}, \dots, \tilde{\varphi}_n)$ un C^1-difféomorphisme. Le théorème 3.12 nous permet de calculer la densité $f_{\tilde{Y}}$ de $\tilde{Y} = (Y_1, \dots, Y_p, \tilde{Y}_{p+1}, \dots, \tilde{Y}_n)$. La densité de Y s'obtient en intégrant f_Y par rapport à ses $n - p$ dernières variables.

Exemple 3.5

Soit $X = (X_1, X_2)'$ de loi uniforme sur le carré $\mathcal{D} =]0,1[^2$. On cherche la loi de $X_1 + X_2$. On pose $Y = (X_1 + X_2, X_2)'$ et

$$\begin{aligned} \varphi : \mathcal{D} &\to \Delta = \{(y_1, y_2), 0 < y_2 < 1 \text{ et } y_2 < y_1 < 1 + y_2\} \\ (x_1, x_2) &\mapsto (y_1 = x_1 + x_2, y_2 = x_2). \end{aligned}$$

D'après le théorème 3.12, la densité de Y est donnée par

$$f_Y(y) = \mathbf{1}_{\mathcal{D}}(\varphi^{-1}(y))|\det(J_{\varphi_y^{-1}})|\mathbf{1}_\Delta(y) = \mathbf{1}_{0<y_1-y_2<1}\mathbf{1}_{0<y_2<1}.$$

La densité de $Y_1 = X_1 + X_2$ s'obtient en intégrant f_Y par rapport à y_2 :

$$f_{Y_1}(y_1) = \int_{\mathbb{R}} f_Y(y)\,\mathrm{d}y_2 = y_1 \mathbf{1}_{0<y_1<1} + (2 - y_1)\mathbf{1}_{1<y_1<2}.$$

Utilisation de la fonction muette

L'approche est identique à celle effectuée pour les v.a.r. Il s'agit de trouver une fonction g mesurable telle que pour toute fonction h mesurable on ait

$$\mathbf{E}[h(\varphi(X))] = \int_U h(\varphi(x))f_X(x)\,\mathrm{d}\lambda(x) = \ldots = \int_V h(y)g(y)\,\mathrm{d}\lambda(y).$$

Par identification, une telle fonction g est une densité de Y. Cette fonction s'obtient généralement par un changement de variable.

Exemple 3.6
On reprend l'exemple précédent. On effectue le changement de variable

$$\begin{aligned}\phi : \mathbb{R}^2 &\to \mathbb{R}^2\\ (y_1, y_2) &\mapsto (y_1 - y_2, y_2).\end{aligned}$$

On obtient

$$\begin{aligned}\mathbf{E}[h(X_1+X_2)] &= \int\int_{\mathbb{R}^2} h(x_1+x_2)\mathbf{1}_{\mathcal{D}}(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2\\ &= \int\int_{\mathbb{R}^2} h(y_1)\mathbf{1}_{\mathcal{D}}(y_1-y_2,y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2\\ &= \int_{\mathbb{R}} h(y_1)\int_{\mathbb{R}}\mathbf{1}_{\mathcal{D}}(y_1-y_2,y_2)\,\mathrm{d}y_2\,\mathrm{d}y_1.\end{aligned}$$

Il vient par identification

$$f_{X_1+X_2}(y_1) = \int_{\mathbb{R}}\mathbf{1}_{\mathcal{D}}(y_1-y_2,y_2)\,\mathrm{d}y_2.$$

Produit de convolution

Le produit de convolution permet de déterminer la densité de la somme de deux variables aléatoires réelles **indépendantes**.

Définition 3.10
Soient f et g deux fonction intégrables. On appelle convolution de f et g la fonction $f * g$ définie par

$$(f*g)(x) = \int f(x-t)g(t)\,\mathrm{d}\lambda(t).$$

Proposition 3.14
Soient f, g et h trois fonction intégrables. Alors

$$f*g = g*f \qquad \text{et} \qquad (f*g)*h = f*(g*h)\ \lambda - pp.$$

Théorème 3.13
Soient X et Y deux v.a.r. **indépendantes** *de densité f_X et f_Y. Alors $f_X * f_Y$ est la densité de X_1+X_2.*

On peut recalculer la densité de X_1+X_2 de l'exemple précédent en faisant le produit de convolution entre les deux fonction indicatrices sur $]0,1[$.

Chapitre 4

Lois usuelles dans $\mathbb{R}^n$

4.1 La loi multinomiale

On rappelle que la loi binomiale permet de modéliser le nombre de succès d'une expérience à 2 issues (succès ou échec) répétées n fois de façon indépendante. La loi multinomiale généralise la loi binomiale. Plus précisément, suite à n répétitions indépendantes d'une expériences à k issues de probabilités respectives $p_1, \dots, p_k$, la loi multinomiale permet de modéliser le nombre de réalisations X_i de l'issue i.

Définition 4.1
On dit que le vecteur $X = (X_1, \dots, X_k)'$ défini précédemment suit une loi multinomiale de paramètres $(n, p_1, \dots, p_k)$ et on note $X \sim \mathcal{M}(n, p_1, \dots, p_k)$. Pour $x = (x_1, \dots, x_k) \in \mathbb{N}^k$ avec $\sum_{i=1}^k x_i = n$, on a

$$\mathbf{P}((X_1, \dots, X_k) = (x_1, \dots, x_k)) = \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k}.$$

Exemple 4.1

1. On effectue un sondage en choisissant au hasard n individus (avec remise) dans une population partagée en k catégories et on note X_i le nombre d'individus sélectionné dans la $i^{\text{ème}}$ catégorie. Le vecteur $X = (X_1, \dots, X_k)'$ suit une loi multinomiale ;
2. On retrouve la loi multinomiale dans diverses parties de la statistique telles que la régression logistique multiclasse ou le test du Chi-Deux.

Proposition 4.1

1. *Les variables marginales $X_1, \dots, X_k$ sont dépendantes et par construction $\sum_{i=1}^k X_i = n$. On a bien évidemment $\sum_{i=1}^k p_i = 1$.*
2. *Chaque variable marginale X_i suit une loi binomiale $\mathcal{B}(n, p_i)$.*
3. *$\mathbf{E}[X] = (np_1, \dots, np_k)$ et la matrice de variance-covariance de X est égale à*

$$\Sigma_X = \begin{pmatrix} np_1(1-p_1) & -np_1p_2 & \dots & -np_1p_k \\ -np_1p_2 & np_2(1-p_2) & \dots & -np_2p_k \\ \vdots & \vdots & \vdots & \vdots \\ -np_1p_k & -np_2p_k & \dots & np_k(1-p_k) \end{pmatrix}.$$

Cette matrice est non inversible.

Preuve. Les points 1 et 2 sont évidents. Montrons 3. Il suffit de montrer $\mathbf{Cov}(X_1, X_2) = -np_1p_2$. On a pour $i = 1, \ldots, k$, $X_i = \sum_{j=1}^n X_{ij}$ où X_{ij} suit une loi de Bernoulli de paramètre p_i. Il vient

$$\mathbf{E}[X_1X_2] = \sum_{i=1}^n \sum_{j=1}^n \mathbf{E}[X_{1i}X_{2j}] = \sum_{1\leq i\neq j\leq n} \mathbf{E}[X_{1i}X_{2j}] + \sum_{i=1}^n \mathbf{E}[X_{1i}X_{2i}].$$

Or $X_{1i}X_{2i} = 0 \quad p.s.$ car les issues 1 et 2 ne peuvent pas se réaliser au cours de la même expérience et $\mathbf{E}[X_{1i}X_{2j}] = p_1p_2$ pour $i \neq j$. On déduit $\mathbf{E}[X_1X_2] = (n^2 - n)p_1p_2$ et

$$\mathbf{Cov}(X_1, X_2) = (n^2 - n)p_1p_2 - n^2p_1p_2 = -np_1p_2.$$

■

4.2 Vecteurs gaussiens - loi multinormale

4.2.1 Rappels sur la loi normale dans $\mathbb{R}$

Nous renvoyons le lecteur à l'annexe A.

4.2.2 Définition - premières propriétés

La définition d'un vecteur aléatoire gaussien n'est pas très intuitive.

Définition 4.2
On dit que le vecteur aléatoire $X = (X_1, \ldots, X_d)'$ est un vecteur gaussien si toute combinaison linéaire de ses variables marginales $\alpha_1X_1 + \ldots + \alpha_dX_d$ est une v.a.r. gaussienne.

Proposition 4.2
La loi d'un vecteur aléatoire gaussien $X = (X_1, \ldots, X_d)'$ est entièrement déterminée par la donnée de son espérance $m_X \in \mathbb{R}^d$ et de sa matrice de variance-covariance Σ_X (de dimension $d \times d$). On note alors $X \sim \mathcal{N}(m_X, \Sigma_X)$.

Proposition 4.3 (Vecteur gaussien ⇒ Composantes gaussiennes)
Si le vecteur $X = (X_1, \ldots, X_d)'$ est gaussien, alors chaque variable aléatoire $X_i, i = 1, \ldots, d$ est gaussienne.

Remarque 4.1 (Le réciproque est fausse !)
Soit $X \sim \mathcal{N}(0, 1)$ et ε une variable aléatoire de Rademacher indépendante de X , *i.e.*, $\mathbf{P}(\varepsilon = 1) = \mathbf{P}(\varepsilon = -1) = 1/2$. On considère la variable $Y = \varepsilon X$ et le vecteur $V = (X, Y)'$. On a alors

$$\begin{aligned} F_Y(u) &= \mathbf{P}(Y \leq u) = \mathbf{P}(\varepsilon X \leq u|\varepsilon = 1)\mathbf{P}(\varepsilon = 1) + \mathbf{P}(\varepsilon X \leq u|\varepsilon = -1)\mathbf{P}(\varepsilon = -1) \\ &= \frac{1}{2}[\mathbf{P}(X \leq u) + \mathbf{P}(X \geq -u) = \mathbf{P}(X \leq u) = F_X(u). \end{aligned}$$

Y est donc une v.a.r. gaussienne centrée réduite. Les composantes du vecteur $(X, Y)'$ suivent ainsi des lois gaussiennes alors que $(X, Y)'$ n'est pas un vecteur gaussien. En effet, il suffit de voir que $Z = X + Y$ n'est pas une v.a.r. gaussienne :

$$\mathbf{P}(Z = 0) = \mathbf{P}((1 + \varepsilon)X = 0) = \mathbf{P}(\varepsilon = -1) = \frac{1}{2}.$$

Or si Z est une v.a.r. gaussienne on a $\mathbf{P}(Z = 0) = 0$ (ou 1 dans le cas d'une v.a.r. centrée dégénérée).

La réciproque de la proposition 4.3 est en revanche vraie si les marginales sont indépendantes.

Proposition 4.4 (Composantes gaussiennes indépendantes ⇒ Vecteur gaussien)
Soit $X_1, \ldots, X_d$ d variables aléatoires indépendantes. Le vecteur $X = (X_1, \ldots, X_d)'$ est gaussien si et seulement si pour tout $i \in \{1, \ldots, d\}$ la variable aléatoire X_i est gaussienne.

Proposition 4.5

1. *$X = (X_1, \ldots, X_d)' \sim \mathcal{N}(m_X, \Sigma_X)$ si et seulement si X peut se décomposer sous la forme $X = AY + m_X$, avec $Y \sim \mathcal{N}(0, I)$, A matrice $d \times d$ satisfaisant $AA' = \Sigma_X$ et $\mathrm{rang}(A) = \mathrm{rang}(\Sigma_X)$.*
2. *Si $X = (X_1, \ldots, X_d)' \sim \mathcal{N}(m_X, \Sigma_X)$, A est une matrice $\ell \times d$ et $b \in \mathbb{R}^\ell$, alors $AX + b \sim \mathcal{N}(Am_X + b, A\Sigma_X A')$. En particulier, si Σ_X est inversible le vecteur $\Sigma_X^{-1/2}(X - m_X)$ est gaussien standard.*

On a vu dans la chapitre 3 que dans un cadre général l'indépendance implique la non corrélation mais que la réciproque est fausse. La proposition suivante montre que cette réciproque est valable pour les vecteurs gaussiens.

Proposition 4.6 (Vecteur gaussien : Indépendance ⇔ Décorrélation)

1. *Soit $X = (X_1, \ldots, X_d)'$ un vecteur gaussien de loi $\mathcal{N}(m_X, \Sigma_X)$. Les variables aléatoires $X_1, \ldots, X_d$ sont indépendantes si et seulement si elles sont non corrélées, c'est-à-dire si et seulement si la matrice de variance-covariance Σ_X est diagonale.*
2. *Soit $Z = (X_1, \ldots, X_d, Y_1, \ldots, Y_p)'$ un vecteur gaussien. Alors, les vecteurs $(X_1, \ldots, X_d)'$ et $(Y_1, \ldots, Y_p)'$ sont indépendants si et seulement si ils sont non corrélés.*

Proposition 4.7 (Densité)
Soit $X = (X_1, \ldots, X_d)' \sim \mathcal{N}(m_X, \Sigma_X)$. X admet une densité f_X si et seulement si $\det(\Sigma_X) \neq 0$. On a alors

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

où $x = (x_1, \ldots, x_d)'$.

Preuve. Comme Σ_X est inversible, le vecteur $Y = \Sigma^{-1/2}(X - m_X) \sim \mathcal{N}(0, I)$. Sa densité est donnée par

$$f_Y(y) = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_i^2}{2}\right) = \frac{1}{(\sqrt{2\pi})^d} \exp\left(-\frac{1}{2}\sum_{i=1}^{d} y_i^2\right).$$

Le changement de variable

$$\begin{aligned} \varphi : \mathbb{R}^d &\to \mathbb{R}^d \\ x &\mapsto y = \Sigma_X^{-1/2}(x - m_X) \end{aligned}$$

nous donne la densité de X

$$f_X(x) = f_Y\left(\Sigma_X^{-1/2}(x - m_X)\right) |\det J_{\varphi_x}|.$$

φ étant une application affine, sa matrice jacobienne vaut $\Sigma_X^{-1/2}$, d'où le résultat.

■

Définition 4.3
Une vecteur gaussien X dont la matrice de variance-covariance a un déterminant nul est dit dégénéré.

4.2.3 Vecteurs gaussiens et loi du Chi-Deux

Proposition 4.8
Si $X = (X_1, \dots, X_d)' \sim \mathcal{N}(m_X, \Sigma_X)$ avec Σ_X inversible alors $(X - m_X)'\Sigma_X^{-1}(X - m_X)$ suit une loi $\chi^2(d)$.

Proposition 4.9
Soit $X = (X_1, \dots, X_d)' \sim \mathcal{N}(0, I)$.

1. *Soit A une matrice symétrique $d \times d$. $X'AX$ suit une loi du χ^2 si et seulement si A est une matrice de projection orthogonale.*
2. *Soit A et B deux matrices de projection orthogonale. $X'AX$ et $X'BX$ sont indépendantes si et seulement si $AB = 0$.*

Nous pouvons maintenant énoncer le théorème de Cochran.

Théorème 4.1 (Théorème de Cochran)
Soit $X = (X_1, \dots, X_d)' \sim \mathcal{N}(0, I)$. Soit $A_1, \dots, A_p$ p matrices symétriques de dimension $d \times d$ telles que $\sum_{j=1}^p X'A_jX = X'X$. Alors les trois conditions suivantes sont équivalentes :

1. $\sum_{j=1}^p \mathrm{rang}(A_j) = d$;
2. *Pour tout $j = 1, \dots, p$, $X'A_jX \sim \chi^2(\mathrm{rang}(A_j))$;*
3. *Les variables $X'A_jX$ sont indépendantes.*

Cette version est la version générale du théorème de Cochran. En statistique, on utilise souvent le corollaire suivant. Ce corollaire est d'ailleurs parfois énoncé sous le nom de "Théorème de Cochran".

Corollaire 4.1
Soit $X = (X_1, \dots, X_d)' \sim \mathcal{N}(0, I)$ et $\mathcal{F}$ un sous-espace vectoriel de $\mathbb{R}^d$ de dimension p. On note $P_{\mathcal{F}}$ la matrice de projection orthogonale sur $\mathcal{F}$ et $P_{\mathcal{F}^\perp} = I - P_{\mathcal{F}}$ la matrice de projection orthogonale sur $\mathcal{F}^\perp$. On a les propriétés suivantes :

1. *$P_{\mathcal{F}}X \sim \mathcal{N}(0, P_{\mathcal{F}})$ et $P_{\mathcal{F}^\perp}X \sim \mathcal{N}(0, P_{\mathcal{F}^\perp})$;*
2. *Les vecteurs aléatoires $P_{\mathcal{F}}X$ et $P_{\mathcal{F}^\perp}X$ sont indépendants ;*
3. *$\|P_{\mathcal{F}}X\|^2 \sim \chi^2(p)$ et $\|P_{\mathcal{F}^\perp}X\|^2 \sim \chi^2(d - p)$.*

Remarque 4.2

1. Ce résultat sera très utilisé dans le cours de régression linéaire ;
2. On peut voir ce théorème comme un analogue (en loi) du théorème de Pythagore. L'identité $\|x\|^2 = \|P_F x\|^2 + \|P_{F^\perp}x\|^2$ pour $x \in \mathbb{R}^n$ devient $\|X\|^2 \stackrel{\mathcal{L}}{=} \|P_F X\|^2 + \|P_{F^\perp}X\|^2$ (la notation $\stackrel{\mathcal{L}}{=}$ signifiant une égalité en loi). On dispose de plus de la loi des deux termes de la somme.
3. La preuve s'écrit assez facilement en se plaçant dans une base orthonormée $\mathcal{B} = (u_1, \dots, u_d)$ de $\mathbb{R}^d$ telle que $(u_1, \dots, u_p)$ soit une base orthonormée de $\mathcal{F}$ et $(u_{p+1}, \dots, u_d)$ soit une base orthonormée de $\mathcal{F}^\perp$.

Chapitre 5

Fonctions génératrices, fonctions caractéristiques

Etant donné X une v.a.r. définie sur $(\Omega, \mathcal{A}, \mathbf{P})$, nous cherchons dans ce chapitre des fonctions ψ qui permettent :

1. de caractériser la loi de X (comme peuvent le faire la fonction de répartition, la densité et la fonction de masse) ;
2. de calculer "aisément" les moments de X (si ils existent).

5.1 Fonction génératrice des moments factoriels

Cette fonction est définie uniquement pour les v.a.r. discrètes à valeurs dans $\mathbb{N}$. Dans cette section, nous considérons donc X une v.a.r. discrète à valeurs dans $\mathbb{N}$, de loi $\mathbf{P}_X$ et de fonction de masse π_X.

Définition 5.1
On appelle fonction génératrice des moments factoriels de X la fonction

$$G_X : s \mapsto \mathbf{E}[s^X] = \sum_{n\in\mathbb{N}} s^n \pi_X(n).$$

Exemple 5.1
Si X suit une loi de Bernoulli de paramètre p alors $G_X(s) = sp + 1 - p$.

Proposition 5.1

1. *G_X est définie pour tout réel s tel que $\mathbf{E}[s^X] < +\infty$;*
2. *La série $\sum_{n\in\mathbb{N}} s^n \pi_X(n)$ étant absolument convergente dans $[-1, 1]$, G_X est continue sur $[-1, 1]$;*
3. *G_X admet des dérivées de tout ordre sur $]-1, 1[$.*

Théorème 5.1
La fonction génératrice des moments factoriels de X détermine la loi de X. On a de plus $G_X^{(k)}(0) = k!\,\pi_X(k)$.

Exemple 5.2
Pour la loi de Bernoulli de paramètre p, on retrouve bien

$$G_X(0) = 1 - p = \pi_X(0) \quad \text{et} \quad G'_X(0) = p = \pi_X(1).$$

Proposition 5.2
La fonction génératrice des moments factoriels G_X admet une dérivée à gauche de 1 d'ordre k $G_X^{(k)}(1)$ si et seulement si le moment factoriel d'ordre k de X défini par $\mathbf{E}[X(X-1)\dots(X-k+1)]$ existe et est fini. On a alors

$$\mathbf{E}[X(X-1)\dots(X-k+1)] = G_X^{(k)}(1).$$

On déduit de la proposition précédente

$$\mathbf{E}[X] = G_X'(1), \qquad \mathbf{E}[X(X-1)] = G_X''(1) \quad \text{d'où} \quad \mathbf{V}(X) = G_X''(1) + G_X'(1) - (G_X'(1))^2.$$

Proposition 5.3
Si X et Y sont des v.a.r. indépendantes, de lois discrètes à valeurs dans $\mathbb{N}$, alors on a pour tout s pour lequel $G_{X+Y}(s), G_X(s)$ et $G_Y(s)$ existent

$$G_{X+Y}(s) = G_X(s)G_Y(s).$$

Théorème 5.2
Soit $(X_k)_{k\in\mathbb{N}}$ une suite de v.a.r. à valeurs dans $\mathbb{N}$ de fonctions de masse respectives $(\pi_{X_k})_{k\in\mathbb{N}}$ et X une v.a.r. à valeur dans $\mathbb{N}$ de fonction de masse π_X. Alors les deux propriétés suivantes sont équivalentes :

1. *$\forall n \in \mathbb{N}$,$\lim_{k\to+\infty} \pi_{X_k}(n) = \pi_X(n)$ (on dit que X_k converge en loi vers X) ;*
2. *$\forall s \in]0,1[$, on a $\lim_{k\to+\infty} G_{X_k}(s) = G_X(s)$.*

5.2 Fonction génératrice des moments

Nous cherchons ici à généraliser la fonction génératrice des moments factoriels G_X à des v.a.r. quelconque. Soit X une v.a.r. (quelconque) définie sur $(\Omega, \mathcal{A}, \mathbf{P})$ de loi $\mathbf{P}_X$. On considère la fonction $g_X : u \mapsto \mathbf{E}[e^{uX}]$ définie que $\{u \in \mathbb{R} : \mathbf{E}[e^{uX}] < +\infty\}$. On peut déjà remarquer que cette fonction est toujours définie en 0 et que $g_X(0) = 1$.

Définition 5.2 (fonction génératrice des moments)
Si la fonction $g_X : u \mapsto \mathbf{E}[e^{uX}]$ est définie sur un voisinage ouvert de 0, alors elle est appelée fonction génératrice des moments de la loi de X.

Proposition 5.4

1. *$\forall (a,b) \in \mathbb{R}^2$, on a $g_{aX+b}(u) = e^{bu} g_X(au)$;*
2. *Si la loi de X est symétrique (X a la même loi que $-X$), alors la fonction g_X est paire ;*
3. *La fonction g_X est convexe ;*
4. *g_X détermine la loi de X.*

Théorème 5.3
Soit X une v.a.r. admettant une fonction génératrice des moments g_X définie pour tout u dans un intervalle ouvert $]u_1, u_2[$ $(u_1 < 0 < u_2)$. Alors :

1. *pour tout $k \in \mathbb{N}$, $\mathbf{E}[|X|^k] < +\infty$;*

2. *pour tout* $u \in]-u_0, u_0[$, *avec* $0 < u_0 < \min(-u_1, u_2)$, *on a*

$$g_X(u) = 1 + \mathbf{E}[X]u + \mathbf{E}[X^2]\frac{u^2}{2!} + \ldots + \mathbf{E}[X^n]\frac{u^n}{n!} + \ldots$$

3. *pour tout* $k \in \mathbb{N}$, $\mathbf{E}[X^k] = g_X^{(k)}(0)$.

Exemple 5.3
Soit X une v.a.r. de loi exponentielle de paramètre $\lambda > 0$. On a pour tout $u \in]-\lambda, \lambda[$:

$$g_X(u) = \int_0^{+\infty} e^{ut}\lambda e^{-\lambda t}\,\mathrm{d}t = \frac{\lambda}{\lambda - u}.$$

On retrouve bien que

$$\mathbf{E}[X] = g_X'(0) = \frac{1}{\lambda}, \quad \mathbf{E}[X^2] = g_X''(0) = \frac{2}{\lambda^2} \quad \text{d'où} \quad \mathbf{V}(X) = \frac{1}{\lambda^2}.$$

Remarque 5.1
D'après le théorème précédent, l'existence de g_X implique que Xadmet des moments de tout ordre. La réciproque est fausse (considérer par exemple la loi log-normale). Ainsi, la suite des moments d'une v.a.r. ne détermine pas sa loi.

Proposition 5.5
Soit X et Y deux v.a.r. indépendantes admettant une fonction génératrice des moments, alors pour tout u tel que $g_{X+Y}(u), g_X(u)$ et $g_Y(u)$ existent, on a

$$g_{X+Y}(u) = g_X(u)g_Y(u).$$

Remarque 5.2
Les fonctions g_X et G_X caractérisent la loi de X et permettent de calculer ses moments. Cependant, en pratique il est difficile de retrouver explicitement la loi de X à partir de g_X. C'est pourquoi pour les v.a.r. discrètes à valeurs dans $\mathbb{N}$, on préfére souvent utiliser G_X (on peut remarquer que $g_X(u) = G_X(e^u)$). Pour les variables absolument continues, on préférera utiliser la fonction caractéristique.

5.3 Fonction caractéristique

5.3.1 Cas d'une variable aléatoire réelle

On considère X une v.a.r. définie sur $(\Omega, \mathcal{A}, \mathcal{P})$.

Définition 5.3
On appelle fonction caractérisique de X la fonction $\varphi_X : \mathbb{R} \to \mathbb{C}$ définie comme la transformée de Fourier de sa loi de probabilité

$$\varphi_X(t) = \mathbf{E}[e^{itX}].$$

Si X est discrète de support $\mathcal{S}$ et de fonction de masse π_X alors

$$\varphi_X(t) = \sum_{x \in \mathcal{S}} \pi_X(t)e^{itx}.$$

De même si X est absolument continue de densité f_X alors

$$\varphi_X(t) = \int_{\mathbb{R}} e^{itx} f_X(x)\,\mathrm{d}x.$$

Exemple 5.4
Le tableau 5.1 donnes les fonctions caractéristiques de quelques loi usuelles :

Loi	**Fonction caractéristique**
Bernoulli $\mathcal{B}(p)$	$pe^{it} + (1-p)$
Binomiale $\mathcal{B}(n,p)$	$(pe^{it} + (1-p))^n$
Poisson $\mathcal{P}(\lambda)$	$e^{\lambda(e^{it}-1)}$
Géométrique $\mathcal{G}(p)$	$pe^{it}/(1-(1-p)e^{it})$
Uniforme $\mathcal{U}([-a,a])$	$\sin(at)/(at)$
Exponentielle $\xi(\lambda)$	$\lambda/(\lambda - it)$
Gaussienne (m,σ^2)	$e^{im}e^{-\sigma^2t^2/2}$

Table 5.1 – Fonctions caractéristiques de quelques lois usuelles.

Proposition 5.6

1. *φ_X est définie et continue pour tout nombre réel t ;*
2. *φ_X est bornée et $\forall t\ |\varphi_X(t)| \leq 1$;*
3. *$\forall (a,b) \in \mathbb{R}^2, \varphi_{aX+b}(t) = e^{ibt}\varphi_X(at)$;*
4. *Si la loi de X est symétrique alors φ_X est une fonction réelle paire ;*
5. *Toute combinaison linéaire convexe de fonctions caractéristiques est une fonction caractéristique.*

Théorème 5.4
φ_X détermine le loi de X (comme son nom l'indique...). Si F_X désigne la fonction de répartition de X alors on a $\forall (a,b) \in \mathbb{R}^2$,

$$F_X(a) - F_X(b) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \varphi_X(t) \frac{e^{-ita} - e^{-itb}}{it}\,\mathrm{d}t.$$

De plus, si $\int_{\mathbb{R}} |\varphi_X(t)|\,\mathrm{d}t < +\infty$, alors X admet une densité f_X et

$$f_X(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \varphi_X(t) e^{-itx}\,\mathrm{d}t.$$

Ce théorème permet de retrouver explicitement la densité de X connaissant sa fonction caractéristiques. De plus, tout comme la fonction génératrice des moments, le fonction caractéristique permet un calcul aisé des moments de tout ordre d'une v.a.r.

Théorème 5.5
Si il existe $n \in \mathbb{N}^\star$ tel que $\mathbf{E}[|X|^n] < \infty$, alors

1. *φ_X est continument dérivable jusqu'à l'ordre n inclu ;*
2. *$\forall k = 0,1,\ldots,n$, $\varphi_X^{(k)}(0) = i^k\mathbf{E}[X^k]$.*

3. *On a le développement*
$$\varphi_X(t) = \sum_{k=0}^{n} \frac{(it)^k}{k!}\mathbf{E}[X^k] + \mathrm{o}(|t|^n)$$
lorsque $t \to 0$.

Proposition 5.7
Si X et Y sont deux v.a.r. **indépendantes** *alors on a pour tout t*
$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t).$$

Attention la réciproque est fausse. Un contre-exemple est donné par X qui suit une loi de Cauchy et $Y = X$.

Exemple 5.5
La proposition précédente nous permet de retrouver la fonction caractéristique d'une variable aléatoire X binomiale de paramètres n et p :
$$\varphi_X(t) = \left[pe^{it} + (1-p)\right]^n.$$

5.3.2 Cas d'un vecteur aléatoire

Dans cette partie $X = (X_1, \dots, X_n)'$ désigne un vecteur aléatoire défini sur $(\Omega, \mathcal{A}, \mathcal{P})$.

Définition 5.4
On appelle fonction caractéristique de X la fonction $\varphi_X : \mathbb{R}^n \to \mathbb{C}$ définie par
$$\varphi_X(t) = \mathbf{E}[e^{i<t,X>}] = \mathbf{E}[e^{i(t_1X_1+\dots+t_nX_n)}].$$

Exemple 5.6
Si $X \sim \mathcal{N}(m_X, \Sigma_X)$ alors
$$\varphi_X(t) = e^{i<m_X,t>}e^{\frac{1}{2}t'\Sigma_X t}.$$

Théorème 5.6
La fonction caratéristique φ_X détermine la loi de X.

Une conséquence de ce théorème est que la loi de X est entièrement déterminée par les lois de toutes les combinaisons linéaires de ses marginales. La fonction caractéristique permet de montrer l'indépendance de n v.a.r.

Théorème 5.7
Les v.a.r. $X_1, \dots, X_n$ sont indépendantes si et seulement si la fonction caractéristique de $X = (X_1, \dots, X_n)'$ est égale au produit des fonctions caractéristiques de ses marginales, c'est-à-dire,
$$\varphi_X(t) = \varphi_{X_1}(t_1)\dots\varphi_{X_n}(t_n).$$

Chapitre 6

Conditionnement, espérance et variance conditionnelles

6.1 Rappels de calcul de probabilités conditionnelles

Soit A et B deux évènements d'un espace $(\Omega, \mathcal{A}, \mathcal{P})$ avec $\mathbf{P}(B) > 0$. On s'intéresse à la réalisation de l'évènement A sachant que l'évènement B est réalisé. L'univers des possibles n'est alors plus Ω tout entier : il est restreint à B d'où la définition de probabilité conditionnelle

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}.$$

On retrouve une vision intuitive de l'indépendance : A et B sont indépendants si et seulement si $\mathbf{P}(A|B) = \mathbf{P}(A)$ (la connaissance de B n'influe pas sur la probabilité de A). On rappelle également

— La formule de Bayes

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(A|B)\mathbf{P}(B)}{\mathbf{P}(A)};$$

— La formule des probabilités totales : soit $\{B_i, i \in \mathcal{I}\}$un ensemble d'évènements disjoints tels que $\Omega = \bigcup_{i\in\mathcal{I}} B_i$, alors

$$\mathbf{P}(A) = \sum_{i\in\mathcal{I}} \mathbf{P}(A|B_i)\mathbf{P}(B_i).$$

Nous allons dans ce chapitre étendre cette notion de probabilités conditionnelles pour des évènements à des variables aléatoires. Cependant la notion générale de conditionnement n'est pas simple. C'est pourquoi le choix est fait de présenter l'idée en plusieurs étapes et de façon intuitive : cas discret, cas absolument continu, interprétation géométrique dans L_2 et enfin extension dans L_1.

6.2 Cas discret

Nous introduisant la notion de conditionnement dans le cas discret par deux exemples.

Exemple 6.1

1. On considère ici deux variables aléatoires discrètes X et Y dont la loi jointe est donnée dans le tableau 6.1.

X \ Y	0	1
0	2/8	3/8
1	1/8	2/8

TABLE 6.1 – Loi jointe de $(X,Y)'$.

Considérons que l'évènement $X=0$ a été réalisé et déterminons la loi de Y. On a :

$$\mathbf{P}(Y=0|X=0)=\frac{\mathbf{P}(Y=0,X=0)}{\mathbf{P}(X=0)}=\frac{2}{5}\quad \text{et}\quad \mathbf{P}(Y=1|X=0)=\frac{3}{5}.$$

On vient ainsi de définir une nouvelle variable aléatoire notée $Y|X=0$ qui suit une loi de Bernoulli de paramètre $\frac{3}{5}$. De même il est facile de voir que la variable aléatoire $Y|X=1$ suit une loi de Bernoulli de paramètre $\frac{2}{3}$. Ces lois sont appelées lois conditionnelles de Y sachant X.

2. Soit $Y\sim\mathcal{P}(\alpha)$ et $Z\sim\mathcal{P}(\beta)$ deux variables aléatoires de loi de Poisson indépendantes. Alors la loi de $X=Y+Z$ suit également une loi de Poisson de paramètre $\alpha+\beta$. On s'intéresse ici à la loi de Y sachant X. Soit $n\in\mathbb{N}$, déterminons la loi de Y sachant $X=n$. Puisque $X=Y+Z$, il est clair que, sachant que $X=n$, Y est à valeurs dans $\{0,1,\dots,n\}$. Soit donc $k\in\{0,1,\dots,n\}$, on a

$$\mathbf{P}(Y=k|X=n)=\frac{\mathbf{P}(Y=k,X=n)}{\mathbf{P}(X=n)}=\frac{\mathbf{P}(Y=k,Z=n-k)}{\mathbf{P}(X=n)}=\frac{\mathbf{P}(Y=k)\mathbf{P}(Z=n-k)}{\mathbf{P}(X=n)}.$$

On obtient alors grâce aux fonctions de masse des lois de Poisson

$$\mathbf{P}(Y=k|X=n)=\binom{k}{n}\left(\frac{\alpha}{\alpha+\beta}\right)^k\left(\frac{\beta}{\alpha+\beta}\right)^{n-k}.$$

Ainsi, sachant $X=n$, Y suit une loi binomiale $\mathcal{B}\left(n,\frac{\alpha}{\alpha+\beta}\right)$.

Revenons au cas général et supposons que Y soit intégrable. Si X est "figé" à x_i, il est naturel de considérer la valeur moyenne de la variable aléatoire Y lorsque $X=x_i$. Cette valeur est appelée l'espérance conditionnelle de Y sachant $X=x_i$. Elle est notée $\mathbf{E}[Y|X=x_i]$.

Définition 6.1 (Espérance conditionnelle)
Supposons Y intégrable. La variable aléatoire qui prend les valeurs $\mathbf{E}[Y|X=x_i]$ avec les probabilités $\pi_X(x_i)$ est appelée espérance conditionnelle de Y sachant X et notée $\mathbf{E}[Y|X]$.

Remarque 6.1
Il est important de noter que l'espérance conditionnelle $\mathbf{E}[Y|X]$ est en général une **variable aléatoire** et non pas une quantité déterministe. On peut l'interpréter comme la valeur moyenne prise par Y lorsque l'on connaît X. Elle pourra donc s'écrire en fonction de X.

Exemple 6.2
On reprend l'exemple précédent. L'espérance de Y sachant $X=n$ est l'espérance d'une loi binomiale $\mathcal{B}\left(n,\frac{\alpha}{\alpha+\beta}\right)$. On a donc pour tout $n\geq 0$:

$$\mathbf{E}[Y|X=n]=\frac{\alpha n}{\alpha+\beta}.$$

Puisque ceci est vrai pour tout n, l'espérance conditionnelle de Y sachant X est :

$$\mathbf{E}[Y|X] = \frac{\alpha X}{\alpha+\beta},$$

qui est bien une fonction de X et donc une variable aléatoire.

On peut se poser la question de prendre l'espérance de l'espérance conditionnelle.

Théorème 6.1
Si Y est intégrable, alors la variable $\mathbf{E}[Y|X]$ est également intégrable et on a :

$$\mathbf{E}\left[\mathbf{E}[Y|X]\right] = \mathbf{E}[Y].$$

Exemple 6.3
Toujours sur l'exemple précédent, on a

$$\mathbf{E}\left[\mathbf{E}[Y|X]\right] = \frac{\alpha}{\alpha+\beta}\mathbf{E}[X],$$

pour l'espérance d'une loi de Poisson de paramètre $\alpha+\beta$ est $\alpha+\beta$, on retrouve donc bien

$$\mathbf{E}\left[\mathbf{E}[Y|X]\right] = \frac{\alpha}{\alpha+\beta}(\alpha+\beta) = \alpha = \mathbf{E}[Y].$$

On vient de dire que dans le cas général, l'espérance conditionnelle $\mathbf{E}[Y|X]$ est une variable aléatoire. Il existe cependant un cas particulier où ce n'est pas le cas : lorsque X et Y sont indépendantes.

Proposition 6.1
Si Y est intégrable et si X et Y sont indépendantes alors la variable aléatoire $\mathbf{E}[Y|X]$ est constante et égale à $\mathbf{E}[Y]$.

6.3 Cas absolument continue

Nous étendons dans cette partie la notion d'espérance conditionnelle au cas de variables absolument continues. Nous commençons par présenter la notion de densité conditionnelle à travers un exemple simple.

Exemple 6.4
On considère $(X,Y)'$ un couple aléatoire de loi uniforme sur le triangle $\mathcal{T} = \{(x,y) : 0 < y < x < 1\}$. $(X,Y)'$ admet donc comme densité $f_{X,Y}(x,y) = 2\mathbf{1}_{\mathcal{T}}(x,y)$. Les densités marginales sont données par

$$f_X(x) = 2x\mathbf{1}_{]0,1[}(x) \quad \text{et} \quad f_Y(y) = 2(1-y)\mathbf{1}_{]0,1[}(y).$$

On déduit $\mathbf{E}[X] = 2/3$ et $\mathbf{E}[Y] = 1/3$.

On considère $x_0 \in [0,1]$ une valeur fixée dans de la variable X et on s'intéresse à la loi de Y sachant que $X = x_0$. Nous voyons sur la Figure 6.1 que Y prend ses valeurs sur le segment $]0,x_0[$, de plus la loi du couple (X,Y) étant uniforme, tous les segments inclus de même longueur inclus dans dans $]0,x_0[$ ont la même probabilité.

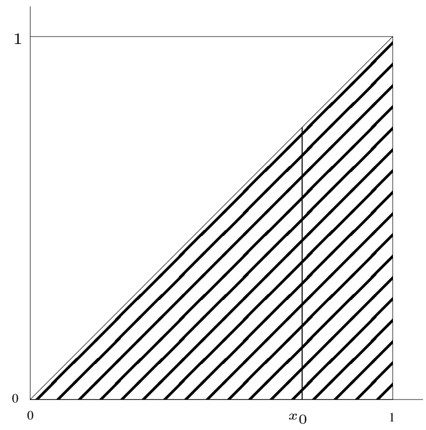


FIGURE 6.1 – Densité conditionnelle de $Y|X = x_0$.

On déduit que la variable aléatoire $Y|X = x_0$ suit une loi uniforme sur $]0, x_0[$:

$$f_{Y|X=x_0}(y) = \frac{1}{x_0}\mathbf{1}_{]0,x_0[}(y) = \frac{2\mathbf{1}_{0<y<x_0}}{2x_0} = \frac{f_{X,Y}(x_0, y)}{f_X(x_0)}.$$

La densité conditionnelle de $Y|X = x$ est ainsi définie comme un rapport entre la densité jointe et la densité marginale (associée à X).

Définition 6.2
Soit $(X, Y)'$ un couple aléatoire de densité $f_{X,Y}(x, y)$. La densité conditionnelle de $Y|X = x$ est définie par

$$f_{Y|X=x}(y) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_X(x)} & si\ f_X(x) > 0 \\ 0 & si\ f_X(x) = 0 \end{cases}$$

Remarque 6.2

1. Pour tout x telle que $f_X(x) \neq 0$ la fonction $f_{Y|X}(.|x)$ est une densité de probabilité, elle est notamment positive et somme à 1 ;
2. La relation $\mathbf{P}(Y = y) = \sum_{x\in\mathcal{S}_X} \mathbf{P}(Y = y|X = x)\mathbf{P}(X = x)$ vue dans le cas discret se transpose au cas absolument continue
$$f_Y(y) = \int_{\mathbb{R}} f_{Y|X=x}(y) f_X(x)\,\mathrm{d}x\,;$$
3. Si les variables X et Y sont indépendantes, on a bien évidemment $f_{Y|X=x}(y) = f_Y(y)$ et $f_{X|Y=y}(x) = f_X(x)$.

On souhaite maintenant définir l'espérance conditionnelle. Pour x fixé, l'espérance conditionnelle de Y sachant $X = x$ est

$$\mathbf{E}[Y|X = x] = \int_{\mathbb{R}} y f_{Y|X=x}(y)\,\mathrm{d}y.$$

La fonction $\varphi : x \mapsto \varphi(x) = \mathbf{E}[Y|X = x]$ est une fonction réelle d'une variable réelle. $\varphi(X)$ est donc une variable aléatoire : c'est l'espérance conditionnelle de Y sachant X.

Définition 6.3
La variable aléatoire qui prend les valeurs $\mathbf{E}[Y|X=x]$ *avec la densité* $f_X(x)$ *est appelée espérance conditionnelle de* Y *sachant* X*. On la note* $\mathbf{E}[Y|X]$*.*

Proposition 6.2
Si Y *est intégrable, alors la variable aléatoire* $\mathbf{E}[Y|X]$ *l'est aussi et on a*

$$\mathbf{E}[\mathbf{E}[Y|X]] = \mathbf{E}[Y].$$

Exemple 6.5
En reprenant l'exemple précédent, on a

$$\mathbf{E}[Y|X=x] = \int_{\mathbb{R}} y f_{Y|X=x}(y)\,\mathrm{d}y = \frac{x}{2},$$

d'où $\mathbf{E}[Y|X] = X/2$. On retrouve bien l'espérance de Y par le théorème précédent

$$\mathbf{E}[\mathbf{E}[Y|X]] = \frac{1}{2}\mathbf{E}[X] = \frac{1}{3} = \mathbf{E}[Y].$$

6.4 Interprétation géométrique de l'espérance conditionnelle

On reste tout d'abord dans le cadre où le couple $(X,Y)'$ est absolument continue. Etant donné une fonction $u:\mathbb{R}\to\mathbb{R}$, il est possible de montrer que

$$\mathbf{E}[(Y-u(X))^2] = \mathbf{E}[(Y-\mathbf{E}[Y|X])^2] + \mathbf{E}[(u(X)-\mathbf{E}[Y|X])^2]. \tag{6.1}$$

Ainsi la quantité $\mathbf{E}[(Y-u(X))^2]$ est minimale lorsque $u(X) = \mathbf{E}[Y|X]$.

On se replace dans le cadre de la section 3.4.2 : on considère l'espace $L_2(\Omega)$ des fonctions de carré intégrable muni du produit scalaire $< X_1, X_2 >= \mathbf{E}[X_1X_2]$. Ce produit scalaire induit la norme $\|X\|^2 = \mathbf{E}[X^2]$. Considérons maintenant une variable aléatoire X et posons

$$L_2(X) = \{u(X) \text{ avec } u:\mathbb{R}\to\mathbb{R} \text{ borélienne telle que } \mathbf{E}[u^2(X)] < +\infty\}.$$

$L_2(X)$ est un sous-espace fermé de $L_2(\Omega)$. Si $Y \in L_2(\Omega)$, on a donc d'après le théorème de projection orthogonale qu'il existe une unique variable aléatoire $\pi(Y)$ dans le sous-espace $L^2(X)$ qui soit à plus courte distance de Y. Dans le cas absolument continu, l'équation 6.1 montre que cette variable aléatoire n'est autre que l'espérance conditionnelle. Il paraît dont naturel d'étendre cette définition à n'importe quelle variable aléatoire de $L_2(\Omega)$.

Définition 6.4 (Espérance conditionnelle dans $L_2(\Omega)$)
Soit $(X,Y)'$ *un couple aléatoire avec* $Y \in L_2(\Omega)$*. L'espérance conditionnelle de* Y *sachant* X*, notée* $\mathbf{E}[Y|X]$*, est la projection orthogonale de* Y *sur le sous-espace* $L_2(X)$*.*

L'espérance conditionnelle de $Y|X$ admet ainsi une interprétation géométrique très simple (voir Figure 6.2). Cette interprétation permet de retrouver sans effort certaines propriétés usuelles de l'espérance conditionnelle.

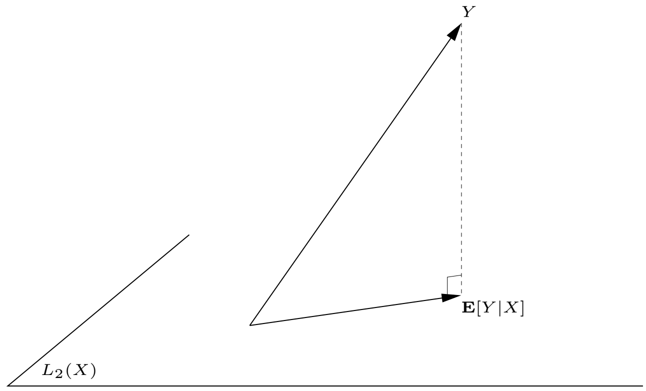


Figure 6.2 – L'espérance conditionnelle comme projection orthogonale.

Proposition 6.3

Soit $(X,Y)'$ un couple aléatoire avec $Y \in L_2(\Omega)$. On a alors

- *Distance minimale :* $\forall Z \in L_2(X), \|Y - \mathbf{E}[Y|X]\| \leq \|Y - Z\|$;
- *Orthogonalité :* $\forall Z \in L_2(X), \langle Y - \mathbf{E}[Y|X], Z\rangle = 0$;
- *Orthogonalité (bis) :* $\forall Z \in L_2(X), \langle Y, Z\rangle = \langle \mathbf{E}[Y|X], Z\rangle$;
- *Pythagore :* $\|Y\|^2 = \|\mathbf{E}[Y|X]\|^2 + \|Y - \mathbf{E}[Y|X]\|^2$;
- *Pythagore (bis) :* $\|\mathbf{E}[Y|X]\| \leq \|Y\|$ *avec égalité si et seulement si Y est une fonction de X.*

Les propriétés précédentes ont été par commodité énoncées en termes de produits scalaires et de normes. Il faut cependant savoir les lire aussi bien en termes d'espérances et d'espérances conditionnelles. De même, chaque fois que l'on écrit $Z \in L_2(X)$, il faut lire $Z = u(X)$, avec $u(X) \in L_2(\Omega)$. Par exemple, la propriété d'orthogonalité (bis) se lit : pour toute fonction u telle que la variable $u(X)$ soit de carré intégrable, on a :

$$\mathbf{E}[u(X)Y] = \mathbf{E}[u(X)\mathbf{E}[Y|X]].$$

6.5 Espérance conditionnelle : le cas général

Dans l'étude des cas discrets et continus, pour définir l'espérance conditionnelle, nous avons supposer seulement Y intégrable ($Y \in L_1(\Omega)$). C'est pourquoi l'interprétation géométrique présentée dans $L_2(\Omega)$ n'est pas complètement satisfaisante. Néanmoins, c'est celle qu'il faudra garder en tête pour se souvenir des propriétés usuelles de l'espérance conditionnelle.

Définition 6.5

Soit $(X,Y)'$ un couple aléatoire avec Y intégrable. On appelle espérance conditionnelle de $Y|X$ une variable aléatoire fonction de X, notée $\mathbf{E}[Y|X]$, telle que pour toute fonction bornée $u : \mathbb{R} \to \mathbb{R}$, on ait

$$\mathbf{E}[u(X)Y] = \mathbf{E}[u(X)\mathbf{E}[Y|X]]. \tag{6.2}$$

Il existe ainsi une fonction φ mesurable telle que $\mathbf{E}[Y|X] = \varphi(X)$. La fonction $x \mapsto \mathbf{E}[Y|X = x]$ est appelée fonction de régression.

Remarque 6.3

Il n'existe pas forcément une unique fonction de X qui vérifié (6.2). Cependant, si il existe deux fonctions qui vérifient (6.2) alors elles sont presque sûrement égales. On confond souvent l'ensemble des variables aléatoires qui vérifient (6.2) avec l'une quelconque de ses versions.

L'espérance conditionnelle vérifie les propriétés suivantes.

Proposition 6.4
Soit (X,Y) un couple aléatoire avec $Y \in L_1(\Omega)$ On a :

- *Cas d'égalité : si $Y = g(X)$ est fonction de X, alors $\mathbf{E}[Y|X] = g(X)$. En particulier $\mathbf{E}[X|X] = X$;*
- *Linéarité : soit Y_1 et Y_2 intégrables, α et β deux réels, alors :*
$$\mathbf{E}[\alpha Y_1 + \beta Y_2|X] = \alpha\mathbf{E}[Y_1|X] + \beta\mathbf{E}[Y_2|X]\,;$$
- *Linéarité (bis) : si $u : \mathbb{R} \to \mathbb{R}$ est bornée, alors $\mathbf{E}[u(X)Y|X] = u(X)\mathbf{E}[Y|X]$;*
- *Positivité : si $Y > 0$, alors $\mathbf{E}[Y|X] > 0$;*
- *Positivité (bis) : si Y_1 et Y_2 sont intégrables avec $Y_1 \leq Y_2$, alors $\mathbf{E}[Y_1|X] \leq \mathbf{E}[Y_2|X]$;*
- *Calcul d'espérance par conditionnement : $\mathbf{E}[\mathbf{E}[Y|X]] = \mathbf{E}[Y]$;*
- *Si X et Y sont indépendantes, alors $\mathbf{E}[Y|X] = \mathbf{E}[Y]$;*
- *Inégalité de Jensen : si Y prend ses valeurs dans un intervalle I de $\mathbb{R}$, si φ est convexe sur I et $\mathbf{E}[|\varphi(Y)|] < +\infty$, alors*
$$\varphi\left(\mathbf{E}[Y|X]\right) \leq \mathbf{E}[\varphi(Y)|X].$$

Remarque 6.4
Supposons maintenant que $Y \in L_2(\Omega)$, on a alors d'après Jensen

$$(\mathbf{E}[Y|X])^2 \leq \mathbf{E}[Y^2|X] \Longrightarrow \mathbf{E}\left[(\mathbf{E}[Y|X])^2\right] \leq \mathbf{E}[Y^2] < +\infty$$

donc $\mathbf{E}[Y|X] \in L_2(X)$. De plus pour tout $u(X) \in L_2(\Omega)$, on a

$$\langle Y - \mathbf{E}[Y|X], u(X)\rangle = \mathbf{E}[Yu(X)] - \mathbf{E}[\mathbf{E}[Y|X]u(X)] = 0$$

et donc $Y - \mathbf{E}[Y|X]$ est orthogonal à tout $u(X)$. On retrouve bien que $\mathbf{E}[Y|X]$ est la projection de Y sur $L_2(X)$. La définition vu dans cette partie $(L_1(\Omega))$ généralise donc bien celle de la section précédente $(L_2(\Omega))$.

Remarque 6.5
Dans de très nombreux cas, la définition vue dans cette partie n'est pas utilisée pour le calcul de l'espérance conditionnelle. En effet,

- si le couple (X,Y) est discret, on détermine dans un premier temps les lois conditionnelles $Y|X = x$ pour tout $x \in \mathcal{S}_X$ afin d'en déduire $\mathbf{E}[Y|X = x]$ puis $\mathbf{E}[Y|X]$ (voir section 6.2) ;
- si le couple (X,Y) est absolument continu, l'espérance conditionnelle se déduit de la densité conditionnelle $f_{Y|X=x}$ (voir section 6.3).

Nous terminons cette partie en donnant la définition de la variance conditionnelle ainsi qu'une de ses propriétés.

Définition 6.6
Soit $(X,Y)'$ un couple aléatoire avec $Y \in L_2(\Omega)$. La variance conditionnelle de Y sachant X est définie par
$$\mathbf{V}(Y|X) = \mathbf{E}[(Y - \mathbf{E}[Y|X])^2|X] = \mathbf{E}[Y^2|X] - (\mathbf{E}[Y|X])^2.$$

Théorème 6.2
Soit $(X,Y)'$ un couple aléatoire avec $Y \in L_2(\Omega)$. Alors

$$\mathbf{V}(Y) = \mathbf{E}[\mathbf{V}(Y|X)] + \mathbf{V}(\mathbf{E}[Y|X]).$$

6.6 Probabilités conditionnelles

Soit $(X, Y)'$ un couple aléatoire admettant une densité de probabilité $f_{X,Y}$ et A un évènement qui s'exprime en fonction de X et Y. Prenons par exemple $A = \{X < Y\}$. On peut écrire sa probabilité comme l'espérance d'une indicatrice :

$$\mathbf{P}(A) = \int_{\mathbb{R}^2} \mathbf{1}_{x<y}(x, y) f_{X,Y}(x, y) \,\mathrm{d}x \,\mathrm{d}y = \mathbf{E}[\mathbf{1}_A].$$

Il est souvent plus facile de calculer cette quantité en commençant par "geler" l'une des variables et d'intégrer par rapport à l'autre. C'est le principe du conditionnement.

Définition 6.7
Le probabilité conditionnelle de l'évènement A sachant $X = x$ est définie par

$$\mathbf{P}(A|X = x) = \mathbf{E}[\mathbf{1}_A|X = x].$$

La probabilité conditionnelle de A sachant X, notée $\mathbf{P}(A|X)$, est la variable aléatoire prenant pour valeurs $\mathbf{P}(A|X = x)$ avec densité $f_X(x)$.

Remarque 6.6
Il faut noter que, tout comme l'espérance conditionnelle $\mathbf{E}[Y|X]$, la probabilité conditionnelle $\mathbf{P}(A|X)$ est une variable aléatoire.

Proposition 6.5 (Calcul de probabilité par conditionnement)
Soit $(X, Y)'$ un couple aléatoire de loi absolument continue. On a alors

$$\mathbf{P}(A) = \int_{\mathbb{R}} \mathbf{P}(A|X = x) f_X(x) \,\mathrm{d}x.$$

Exemple 6.6
Reprenons la deuxième partie de l'exemple 3.2 page 26 : $(X, Y)'$ est un couple aléatoire de densité

$$f_{X,Y}(x, y) = \frac{xy}{2}\mathbf{1}_{0<y<x<2}.$$

Nous souhaitons recalculer la probabilité $\mathbf{P}(Y > X^2)$. Comme $f_X(x) = \frac{x^3}{4}\mathbf{1}_{]0,2[}(x)$, on déduit, pour $x \neq 0$, la densité conditionnelle

$$f_{Y|X=x}(y) = 2\frac{y}{x^2}\mathbf{1}_{0<y<x<2}.$$

Il vient donc

$$\mathbf{P}(Y > X^2) = \int \mathbf{P}(Y > x^2|X = x) f_X(x) \,\mathrm{d}x.$$

Or

$$\begin{aligned}\mathbf{P}(Y > X^2|X = x) &= \int \mathbf{1}_{y>x^2} f_{Y|X=x}(y) \,\mathrm{d}y = \int 2\frac{y}{x^2}\mathbf{1}_{0<x^2<y<x<1} \,\mathrm{d}y \\ &= \frac{2}{x^2}\mathbf{1}_{0<x<1}\int_{x^2}^{x} y \,\mathrm{d}y = (1 - x^2)\mathbf{1}_{0<x<1}.\end{aligned}$$

On a donc

$$\mathbf{P}(Y > X^2|X) = (1 - X^2)\mathbf{1}_{0<X<1},$$

et

$$\mathbf{P}(Y > X^2) = \int_0^1 (1 - x^2)\frac{x^3}{4} \,\mathrm{d}x = \frac{1}{48}.$$

6.7 Généralisation au conditionnement pas des sous-tribus

L'espérance conditionnelle définie dans la section 6.5 peut-être vue comme un cas particulier du conditionnement par une sous-tribu. On considère ici une variable aléatoire Y définie sur un espace $(\Omega, \mathcal{A}, \mathbf{P})$ telle que $\mathbf{E}[|Y|] < +\infty$ et $\mathcal{S}$ une sous-tribu de $\mathcal{A}$. On suppose qu'on ne s'intéresse qu'aux évènements de la tribu $\mathcal{S}$ et on se demande si on peut trouver une version simplifiée de Y, qui résumerait toute l'information contenue dans Y, sachant que l'on se restreint aux évènements de $\mathcal{S}$.

Définition 6.8
On appelle espérance conditionnelle de Y sachant $\mathcal{S}$ la classe des variables aléatoires Z telles que :

1. *Z est $\mathcal{S}$-mesurable ;*
2. *pour tout $A \in \mathcal{S}$: $\mathbf{E}[Y\mathbf{1}_A] = \mathbf{E}[Z\mathbf{1}_A]$.*

Cette classe de variables aléatoires est notée $\mathbf{E}[Y|\mathcal{S}]$. Une variable aléatoire Z qui vérifie les deux propriétés ci-dessus est appelée une version de l'espérance conditionnelle.

Remarque 6.7
Deux versions de l'espérance conditionnelle sont presque sûrement égales. C'est pourquoi on confond souvent la classe $\mathbf{E}[Y|\mathcal{S}]$ avec l'une quelconque de ses versions.

Proposition 6.6
Soit $(X, Y)'$ un couple aléatoire tel que $\mathbf{E}[|Y|] < +\infty$. L'espérance conditionnelle présentée dans la définition 6.5 coïncide avec l'espérance conditionnelle de Y sachant la sous-tribu $\sigma(X)$.

Les propriétés énoncées dans la section 6.5 sur l'espérance conditionnelle restent vraies pour l'espérance conditionnelle de Y sachant une sous-tribu. Pour plus de détails sur cette notion d'espérance conditionnelle, on pourra se référer à Montfort (1996) et Fromont (2008).

Chapitre 7

Convergences de suites de variables aléatoires

Nous motivons la nécessité de définir des modes de convergences particuliers pour des suites de variables aléatoires en reprenant l'exemple de Jacod & Protter (2003), chapitre 17. On rappelle qu'on dit qu'une suite de fonctions $f_n : \mathbb{R} \to \mathbb{R}$ converge simplement vers $f : \mathbb{R} \to \mathbb{R}$ si $\lim_{n\to\infty} f_n(x) = f(x)$ pour tout $x \in \mathbb{R}$. Une variable aléatoire X étant une fonction $X : \Omega \to \mathbb{R}$, il semble naturel d'élargir cette notion : une suite de variables aléatoires X_n converge simplement vers X si $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ pour tout $\omega \in \Omega$. Bien que naturelle, cette définition est, de manière surprenante, à peu près inutile en probabilités.

Exemple 7.1

Considérons $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires i.i.d. suivant une loi de Bernoulli de paramètre $1/2$ (penser $X_i = 1$ si on obtient "face" au $i^{\text{ème}}$ lancé, 0 sinon). Lorsque n est grand, on s'attend à ce que la proportion de faces obtenue soit à peu près égale à $1/2$, ce qui mathématiquement devrait se traduire par

$$\lim_{n\to\infty} \frac{X_1(\omega) + \ldots + X_n(\omega)}{n} = \frac{1}{2} \quad \forall \omega \in \Omega.$$

Il est cependant évident qu'on ne peut espérer obtenir un tel résultat. Il suffit en effet de considérer $\omega_0 = \{f, f, f, f, f, ...\}$ la suite ne contenant que des faces, on a alors

$$\lim_{n\to\infty} \frac{X_1(\omega_0) + \ldots + X_n(\omega_0)}{n} = 1.$$

Plus généralement

$$\lim_{n\to\infty} \frac{X_1(\omega) + \ldots + X_n(\omega)}{n} = 0$$

pour tout ω dans $A = \{\omega :$ il n'y a qu'un nombre fini de faces$\}$. Plus généralement encore, on peut trouver des ω pour lesquels la fréquence converge vers n'importe quel nombre fixé dans $[0, 1]$, et d'autres pour lesquels la fréquence ne converge pas. Bien évidemment, l'évènement A est plutôt invraisemblable. Il est en effet possible de montrer, lorsque $n \to \infty$, que $\mathbf{P}(A) = 0$. Il est donc nécessaire de définir de nouveaux modes de convergences, spécifiques aux variables aléatoire. Nous verrons à la fin de ce chapitre que, pour cet exemple, nous avons

$$\mathbf{P}\left(\left\{\omega : \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i(\omega) = \frac{1}{2}\right\}\right) = 1.$$

Ce type de convergence, pour lequel on n'a pas convergence pour tout ω, mais seulement pour *presque tout* ω, est appelé *convergence presque sûre*.

7.1 Les différents types de convergence

On désigne par $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbf{P})$ et à valeurs dans $\mathbb{R}$.

7.1.1 Convergence presque sûre ou convergence forte

Définition 7.1
On dit que $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers une variable aléatoire X si l'ensemble N des ω tels que la suite numérique $(X_n(\omega))_{n\in\mathbb{N}}$ ne converge pas vers $X(\omega)$ est négligeable (c'est-à-dire vérifie $\mathbf{P}(N) = 0$). On note

$$\lim_{n\to\infty} X_n = X \quad p.s. \qquad ou \qquad X_n \stackrel{p.s.}{\to} X.$$

Remarque 7.1
On peut aussi dire que $X_n \stackrel{p.s.}{\to} X$ si et seulement si

$$\mathbf{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) \neq X(\omega)\right\}\right) = 0$$

ou encore

$$\mathbf{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1.$$

Proposition 7.1
1. *Si $X_n \stackrel{p.s.}{\to} X$ et si $\varphi : \mathbb{R} \to \mathbb{R}$ est une fonction continue sur $\mathbb{R}$ alors $\varphi(X_n) \stackrel{p.s.}{\to} \varphi(X)$.*
2. *Si $X_n \stackrel{p.s.}{\to} X$ et $Y_n \stackrel{p.s.}{\to} Y$ alors*
   - *pour tout réels a et b, $aX_n + bY_n \stackrel{p.s.}{\to} aX + bY$;*
   - *$X_nY_n \stackrel{p.s.}{\to} XY$.*
   - *$X_n/Y_n \stackrel{p.s.}{\to} X/Y$ si $\mathbf{P}(Y = 0) = 0$.*

On utilise rarement la définition pour montrer la convergence presque sûre. On a souvent recourt à l'un des critères suivants.

Théorème 7.1
La suite de v.a.r. $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers X si et seulement si pour tout $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbf{P}(\sup_{m\geq n} |X_m - X| > \varepsilon) = 0.$$

Proposition 7.2 (Lemme de Borel-Cantelli)
1. *Si pour tout $\varepsilon > 0$,*
$$\sum_{n\in\mathbb{N}} \mathbf{P}(|X_n - X| > \varepsilon) < +\infty$$
alors $X_n \stackrel{p.s.}{\to} X$.
2. *Si les $(X_n)_{n\in\mathbb{N}}$ sont mutuellement indépendantes, alors $X_n \stackrel{p.s.}{\to} 0$ si et seulement si*
$$\forall \varepsilon > 0, \quad \sum_{n\in\mathbb{N}} \mathbf{P}(|X_n| > \varepsilon) < +\infty.$$

Proposition 7.3
Si il existe $p > 0$ tel que $\sum_{n\in\mathbb{N}} \mathbf{E}[|X_n - X|^p] < +\infty$, alors $X_n \stackrel{p.s.}{\to} X$.

7.1.2 La convergence en probabilité

Définition 7.2
On dit que $(X_n)_{n\in\mathbb{N}}$ *converge en probabilité vers* X *si pour tout* $\varepsilon > 0$*, on a*

$$\lim_{n\to\infty} \mathbf{P}(|X_n - X| > \varepsilon) = 0.$$

On note $X_n \stackrel{\mathbf{P}}{\to} X$*.*

La définition peut se réécrire : $X_n \stackrel{\mathbf{P}}{\to} X$ lorsque pour tout $\varepsilon > 0$ et $\eta > 0$ il existe $N = N(\varepsilon, \eta)$ tel que pour tout $n \geq N$ on a

$$\mathbf{P}(|X_n - X| > \varepsilon) < \eta.$$

Exemple 7.2
Soit $X_n, n \geq 1$ des v.a.r. non corrélées deux à deux telles que $\mathbf{E}[X_n] = 0$ et $\mathbf{V}(X_n) = \sigma^2$. On note $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. D'après Bienaymé-Tchebytchev, on a

$$\mathbf{P}(|\bar{X}_n| > \varepsilon) \leq \frac{1}{n^2\varepsilon^2}\mathbf{V}\left(\sum_{i=1}^n X_i\right) = \frac{\sigma^2}{n\varepsilon^2}.$$

On a donc $\bar{X}_n \stackrel{\mathbf{P}}{\to} 0$.

Exemple 7.3
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires dont la loi est définie par

$$\mathbf{P}(X_n = \sqrt{n}) = \frac{1}{n} \quad \text{et} \quad \mathbf{P}(X_n = 0) = 1 - \frac{1}{n}.$$

On a pour $\varepsilon > 0$ fixé,

$$\mathbf{P}(|X_n| > \varepsilon) = \mathbf{P}(|X_n| > \varepsilon \cap X_n = \sqrt{n}) + \mathbf{P}(|X_n| > \varepsilon \cap X_n = 0) = \mathbf{P}(|X_n| > \varepsilon \cap X_n = \sqrt{n}).$$

Or, pour n assez grand, $\{|X_n| > \varepsilon\} \subset \{X_n = \sqrt{n}\}$, donc

$$\lim_{n\to\infty} \mathbf{P}(|X_n| > \varepsilon) = \lim_{n\to\infty} 1/n = 0.$$

On déduit $\bar{X}_n \stackrel{\mathbf{P}}{\to} 0$.

Proposition 7.4 (Opérations sur les convergences en probabilité)

1. *Si* $X_n \stackrel{\mathbf{P}}{\to} X$ *et si* $\varphi : \mathbb{R} \to \mathbb{R}$ *est une fonction continue sur* $\mathbb{R}$ *alors* $\varphi(X_n) \stackrel{\mathbf{P}}{\to} \varphi(X)$*.*
2. *Si* $X_n \stackrel{\mathbf{P}}{\to} X$ *et* $Y_n \stackrel{\mathbf{P}}{\to} Y$ *alors*
   — *pour tout réels* a *et* b*,* $aX_n + bY_n \stackrel{\mathbf{P}}{\to} aX + bY$ *;*
   — $X_nY_n \stackrel{\mathbf{P}}{\to} XY$*.*
   — $X_n/Y_n \stackrel{\mathbf{P}}{\to} X/Y$ *si* $\mathbf{P}(Y = 0) = 0$*.*

Théorème 7.2
Si $X_n \stackrel{p.s.}{\to} X$ *alors* $X_n \stackrel{\mathbf{P}}{\to} X$*.*

La réciproque est fausse, un contre exemple est donné dans Jacod & Protter (2003), page 152.

7.1.3 La convergence en moyenne d'ordre $p > 0$

Définition 7.3
Soit $p > 0$. On dit que $(X_n)_{n\in\mathbb{N}}$ converge en moyenne d'ordre p (ou dans L_p) vers X si les X_n et X sont dans L_p ($\mathbf{E}[|X_n|^p] < +\infty$ et $\mathbf{E}[|X|^p] < +\infty$), et si on a

$$\lim_{n\to\infty} \mathbf{E}[|X_n - X|^p] = 0.$$

On note $X_n \stackrel{L_p}{\to} X$.

Les cas les plus importants sont $p = 1$ (convergence en moyenne) et $p = 2$ (convergence en moyenne quadratique). En particulier si $X_n \stackrel{L_1}{\to} X$, alors

$$\lim_{n\to\infty} \mathbf{E}[X_n] = \mathbf{E}[X] \quad \text{et} \quad \lim_{n\to\infty} \mathbf{E}[|X_n|] = \mathbf{E}[|X|].$$

Concernant la convergence L_2, on a

$$\mathbf{E}[(X_n - a)^2] = (\mathbf{E}[X_n] - a)^2 + \mathbf{V}[X_n]. \tag{7.1}$$

On déduit

$$X_n \stackrel{L_2}{\to} a \iff \begin{cases} \lim_{n\to\infty} \mathbf{E}[X_n] = a \\ \lim_{n\to\infty} \mathbf{V}[X_n] = 0 \end{cases}$$

La décomposition (7.1) est très utilisée en statistique. Elle permet d'étudier l'erreur quadratique d'estimateurs en se ramenant à l'étude du biais et de la variance de ces derniers.

Enfin on peut remarquer que

$$X_n \stackrel{L_2}{\to} X \quad \Longrightarrow \quad X_n \stackrel{L_1}{\to} X.$$

Il suffit d'écrire

$$\mathbf{E}|X_n - X| = \mathbf{E}\sqrt{(X_n - X)^2} \leq \sqrt{\mathbf{E}|X_n - X|}$$

d'après l'inégalité de Jensen. On termine cette partie par un résultat qui prouve que la convergence L_p est plus forte que la convergence en probabilité.

Théorème 7.3
Si $X_n \stackrel{L_p}{\to} X$ alors $X_n \stackrel{\mathbf{P}}{\to} X$.

Une fois de plus la réciproque est fausse. En effet, pour l'exemple 7.3. En effet $\mathbf{E}[X_n^2] = 1$ donc $(X_n)_{n\in\mathbb{N}}$ ne converge pas vers 0 dans L_2 alors qu'elle converge vers 0 en probabilité.
Par ailleurs, même la convergence presque sûre n'implique pas la convergence en moyenne d'ordre p. On pourra se référer à Barbe & Ledoux (2007), chapitre 5, pour des contre-exemples.

7.1.4 La convergence en loi

Bien que différent, les trois modes de convergence vus précédemment sont de même nature et peuvent être abordés comme des variantes de la convergence habituelle. Il existe un autre mode de convergence, différent des précédents mais très utiles en probabilité : la convergence *en loi*, ou convergence *faible* ou encore convergence *étroite*. Dans cette partie, nous donnons la définition ainsi que les principales propriétés de ce nouveau mode de convergence. Pour plus de détails, ainsi que pour les preuves des résultats, le lecteur pourra consulter Carbon (2007) et Jacod & Protter (2003).

On considère X une variable aléatoire et $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires définis sur $(\Omega, \mathcal{A}, \mathbf{P})$ et à valeurs dans $(\mathbb{R}, \mathcal{B}_\mathbb{R})$. On souhaite donner un sens à l'idée suivante : "pour n grand, la loi de X et la loi de X_n sont voisines". Une définition naturelle serait d'écrire que pour tout borélien A, on a

$$\lim_{n\to\infty} \mathbf{P}(X_n \in A) = \mathbf{P}(X \in A).$$

Cependant, l'examen de certains cas particuliers montre que cette définition n'est pas satisfaisante.

Exemple 7.4
Prenons le cas où X_n est de loi uniforme sur $[-1/n, 1/n]$ et X est de loi *p.s.* nulle. Il est facile de voir que X_n converge vers X selon les différents modes déjà étudiés. On a cependant

$$\mathbf{P}(X_n \leq 0) = \frac{1}{2} \neq 1 = \mathbf{P}(X \leq 0) \quad \text{et} \quad \mathbf{P}(X_n > 0) = \frac{1}{2} \neq 0 = \mathbf{P}(X > 0).$$

En revanche, pour tout intervalle $[a, b]$ tel que $a \neq 0$ et $b \neq 0$, on a

$$\lim_{n\to\infty} \mathbf{P}(X_n \in [a, b]) = \mathbf{P}(X \in [a, b]).$$

Le problème évoqué dans cet exemple vient du fait que la loi de X charge le point 0, ou encore que la fonction de répartition de la loi de X est discontinue en 0. D'où la définition suivante.

Définition 7.4 (Convergence en loi)
On dit que la suite $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X si, en tout point de continuité de F_X, on a $\lim_{n\to\infty} F_{X_n}(x) = F(x)$. On note $X_n \stackrel{\mathcal{L}}{\to} X$.

Remarque 7.2
1. On a bien $X_n \stackrel{\mathcal{L}}{\to} X$ pour l'exemple 7.4. En effet,

$$F_{X_n}(x) = \begin{cases} 0 & \text{si } x \leq -1/n \\ n/2(x + 1/n) & \text{si } -1/n < x \leq 1/n \\ 1 & \text{si } x > 1/n. \end{cases}$$

On a donc

$$\begin{cases} \lim_{n\to\infty} F_{X_n}(x) = 0 & \text{si x<0} \\ \lim_{n\to\infty} F_{X_n}(x) = 1 & \text{si x>0.} \end{cases}$$

Comme F_X est discontinue en 0, on conclut que $X_n \stackrel{\mathcal{L}}{\to} X$.

2. Si $X_n \stackrel{\mathcal{L}}{\to} X$ et si F_X est continue en tout point de $\mathbb{R}$, alors pour tout $a, b \in \mathbb{R}$

$$\lim_{n\to\infty} \mathbf{P}(X_n \in [a, b]) = \mathbf{P}(X \in [a, b]).$$

3. On a déjà vu que $X_n \stackrel{\mathcal{L}}{\to} X$ n'implique pas $\lim_{n\to\infty} \mathbf{P}(X_n \in A) = \mathbf{P}(X \in A)$ pour tout borélien. On a également $X_n \stackrel{\mathcal{L}}{\to} X$ n'implique pas $\lim_{n\to\infty} \mathbf{E}[X_n] = \mathbf{E}[X]$ (voir Foata & Fuchs (2003), chapitre 16).

4. De même, $X_n \xrightarrow{\mathcal{L}} X$ n'implique pas $X_n - X \xrightarrow{\mathcal{L}} 0$. Il suffit de prendre $X \sim \mathcal{N}(0,1)$ et $X_n = (-1)^n X_n$.
5. La définition de convergence en loi se généralise au cas vectoriel (voir Jacod & Protter (2003) et Carbon (2007)).

La convergence en loi peut également se montrer ou même se définir (selon les ouvrages) à l'aide des fonctions caractéristiques.

Théorème 7.4
Les trois assertions suivantes sont équivalentes :

1. $X_n \xrightarrow{\mathcal{L}} X$;
2. *Pour toute fonction* $f : \mathbb{R} \to \mathbb{R}$ *continue bornée, on a* $\lim_{n\to\infty} \int f \,\mathrm{d}\mathbf{P}_{X_n} = \int f \,\mathrm{d}\mathbf{P}_X$;
3. *Pour tout* $t \in \mathbb{R}$, *on a* $\lim_{n\to\infty} \varphi_{X_n}(t) = \varphi_X(t)$.

La dernière assertion est une conséquence directe du théorème de Paul Levy (voir Jacod & Protter (2003)). L'utilisation des fonctions caractéristiques est très souvent utilisée pour montrer la convergence en loi.

Exemple 7.5
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variable aléatoire de loi $\mathcal{B}(n, p_n)$. On suppose que $np_n \to \lambda$ lorsque $n \to \infty$. On a

$$\varphi_{X_n}(t) = [p_n e^{it} + (1 - p_n)]^n.$$

On montre à l'aide d'un développement limité

$$\varphi_{X_n}(t) \sim \left[1 + (e^{it} - 1)\frac{\lambda}{n}\right]^n \to e^{\lambda(e^{it}-1)}.$$

On déduit que $X_n \xrightarrow{\mathcal{L}} X$ où X est une v.a.r. qui suit une loi de Poisson de paramètre λ.

Exemple 7.6
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires de loi de Poisson de paramètre λ_n avec $\lambda_n \to \infty$ lorsque $n \to \infty$. De la même manière que dans l'exemple précédent on peut montrer que

$$\frac{X_n - \lambda_n}{\sqrt{\lambda_n}} \xrightarrow{\mathcal{L}} X$$

où $X \sim \mathcal{N}(0,1)$. On note également cette convergence

$$\frac{X_n - \lambda_n}{\sqrt{\lambda_n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Dans les cas discret et absolument continue, la convergence en loi peut également se montrer à partir des fonctions de masse et de densité.

Théorème 7.5

1. *Soit* X_n *et* X *des v.a.r. à valeurs dans un espace* E *fini ou dénombrable. Alors* $X_n \xrightarrow{\mathcal{L}} X$ *si et seulement si*

$$\forall j \in E, \quad \lim_{n\to\infty} \mathbf{P}(X_n = j) = \mathbf{P}(X = j).$$

2. *Soit X_n et X des v.a.r. dont les lois admettent pour densité (par rapport à la mesure de Lebesgue) f_n et f. Si pour presque tout x de $\mathbb{R}$ on a $\lim_{n\to\infty} f_n(x) = f(x)$, alors $X_n \xrightarrow{\mathcal{L}} X$.*

La convergence en loi est préservée par certaines opérations arithmétiques.

Proposition 7.5
Soit $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ deux suites de v.a.r., X une v.a.r. et a un réel. On a :

1. *Si $X_n \xrightarrow{\mathcal{L}} X$ et $Y_n \xrightarrow{\mathcal{L}} a$ alors*
$$X_n + Y_n \xrightarrow{\mathcal{L}} X + a, \quad X_nY_n \xrightarrow{\mathcal{L}} aX \quad \text{et} \quad \frac{X_n}{Y_n} \xrightarrow{\mathcal{L}} \frac{X}{a} \quad (\text{si } a \neq 0).$$
2. *Si $g : \mathbb{R} \to \mathbb{R}$ est continue en tout point de $\mathbb{R}$ alors $g(X_n) \xrightarrow{\mathcal{L}} g(X)$.*

Remarque 7.3
1. La première assertion de ce théorème est appelée théorème de Slutsky ;
2. Ces opérations algébriques se généralisent au cas vectoriel.

Nous terminons en énonçant les relations entre la convergence en loi et les autres modes de convergence.

Théorème 7.6
Si $X_n \xrightarrow{\mathbf{P}} X$ alors $X_n \xrightarrow{\mathcal{L}} X$.

Remarque 7.4
1. Là encore la réciproque est fausse. Un contre-exemple est donné par $X \sim \mathcal{N}(0,1)$ et $X_n = (-1)^n X$. La réciproque devient vraie lorsque X_n converge en loi vers une variable constante a. On a
$$X_n \xrightarrow{\mathcal{L}} a \Longleftrightarrow X_n \xrightarrow{\mathbf{P}} a.$$
2. Les convergences presque sûre et en moyenne d'ordre p impliquant la convergence en probabilité, on en déduit que ces modes impliquent également la convergence en loi. La convergence en loi est donc le mode de convergence le plus "faible". On peut résumer les différents modes de convergence par le diagramme suivant :

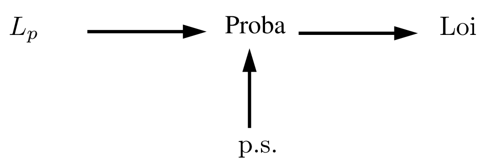


FIGURE 7.1 – Relations entre les différents mode de convergence.

7.2 La loi des grand nombres

Soit $X_1, \ldots, X_n$ n v.a.r. indépendantes de loi Bernoulli de paramètre p. On a alors $S_n = \sum_{i=1}^n X_i \sim \mathcal{B}(n,p)$ et l'inégalité de Bienaymé-Chebychev nous donne :

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right) < \frac{p(1-p)}{n\varepsilon^2} \to 0 \quad \text{quand } n \to \infty.$$

On déduit $S_n/n \xrightarrow{\mathbf{P}} p$. Ce résultat se généralise à d'autres types de loi.

Dans la suite, étant donnée n v.a.r. $X_1, \ldots, X_n$, on désignera alors par S_n la somme $\sum_{i=1}^n X_i$. Lorsque l'on obtient pour S_n une convergence en probabilité du même genre que sur l'exemple précédent, on parle de loi faible des grands nombres. Si la convergence est presque sûre, on emploie le terme, loi forte des grands nombres.

7.2.1 Lois faibles des grands nombres

Nous énonçons 2 lois faibles des grands nombres.

Théorème 7.7 (Loi faible des grands nombres dans L_1)
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. de $L_1(\Omega, \mathcal{A}, \mathbf{P})$ 2 à 2 indépendantes et de même loi. On note $\mathbf{E}[X_1] = \mu$. On a

$$\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{L_1} \mu.$$

Théorème 7.8 (Loi faible des grands nombres dans L_2)
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. de $L_2(\Omega, \mathcal{A}, \mathbf{P})$ 2 à 2 non corrélées et de même loi. On note $\mathbf{E}[X_1] = \mu$. On a

$$\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{L_2} \mu.$$

On parle de lois faibles des grands nombres pour ces deux théorèmes car les convergences ont bien évidemment également lieu en probabilité. On pourra consulter Foata & Fuchs (2003), chapitre 17, pour la preuve de ces résultats.

7.2.2 Lois fortes des grands nombres

La convergence presque sûre s'obtient supposant l'indépendance mutuelle dans le théorème 7.7.

Théorème 7.9
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. de $L_1(\Omega, \mathcal{A}, \mathbf{P})$ indépendantes et de même loi. On note $\mathbf{E}[X_1] = \mu$. On a

$$\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{p.s.} \mu.$$

Remarque 7.5
Si on suppose de plus que $\mathbf{E}[X_1^2] < \infty$, alors la convergence a également lieu dans L_2.

7.3 Le théorème central limite

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. indépendantes et de même loi $\mathcal{N}(\mu, \sigma^2)$. Si on note $S_n = \sum_{i=1}^n$ et $\bar{X}_n = S_n/n$, on rappelle que

$$\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

Le théorème central limite stipule que, sous des hypothèses très faibles, on peut étendre ce résultat (asymptotiquement) à "n'importe quelle" suite de variables aléatoire indépendantes. C'est l'un des résultats les plus impressionnants et les plus utilisés en probabilités et statistique. L'idée est très simple : étant donnée une suite i.i.d. $X_1,\dots,X_n$ de variables aléatoires de moyenne μ et de variance σ^2, on a pour n assez est grand, $\mathcal{L}(S_n)\approx\mathcal{N}(n\mu,n\sigma^2)$. Et ceci, quelle que soit la loi des X_i. La seule hypothèse requise est l'existence d'une variance.

Théorème 7.10 (Théorème central limite)
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes, de même loi, et telles que $\mathbf{E}[X_i^2]<+\infty$. On note $\mathbf{E}[X_i]=\mu$, $\mathbf{V}[X_i]=\sigma^2$ et $\bar{X}_n=\frac{1}{n}\sum_{i=1}^n X_i$. On a alors

$$\sqrt{n}\frac{\bar{X}_n-\mu}{\sigma}\xrightarrow{\mathcal{L}}\mathcal{N}(0,1)\quad \textit{quand } n\to\infty.$$

La preuve de ce résultat est relativement simple, elle repose sur un développement limité de la fonction caractéristique.

Preuve (Voir Jacod & Protter (2003)). On note φ la fonction caractéristique des variables aléatoires $X_i-\mu$ et

$$Y_n=\sqrt{n}\frac{\bar{X}_n-\mu}{\sigma}.$$

En combinant les propositions 5.6 et 5.7, on obtient

$$\varphi_{Y_n}(t)=\left(\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n.$$

De plus, il est facile de voir que

$$\varphi(0)=1,\quad \varphi'(0)=0\quad \text{et}\quad \varphi''(0)=-\sigma^2.$$

Il suffit d'utiliser le théorème 5.5. On a donc

$$\varphi(u)=1-\frac{\sigma^2u^2}{2}+\mathrm{o}(u^2)$$

et

$$\varphi_{Y_n}(t)=\exp\left(n\log(1-t^2/2+\mathrm{o}(1/n))\right).$$

Par conséquent

$$\lim_{n\to\infty}\varphi_{Y_n}(t)=\exp(-t^2/2)$$

et $t\mapsto\exp(-t^2/2)$ est la fonction caractéristique de la loi $\mathcal{N}(0,1)$. On déduit du théorème 7.4 que $Y_n\xrightarrow{\mathcal{L}}\mathcal{N}(0,1)$.

■

Exemple 7.7
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. i.i.d. de loi de Bernoulli de paramètre $p\in]0,1[$. On note $S_n=\sum_{i=1}^n X_i$. On a d'après la loi des grands nombres

$$\frac{S_n}{n}\xrightarrow{p.s.}p\quad \text{quand } n\to\infty$$

et d'après le théorème central limite

$$\frac{S_n-np}{\sqrt{np(1-p)}}\xrightarrow{\mathcal{L}}\mathcal{N}(0,1)\quad \text{quand } n\to\infty.$$

Le théorème central limite se généralise au cas vectoriel.

Théorème 7.11 (Théorème central limité pour des vecteurs aléatoires)
Soit $(X_n)_{n\in\mathbb{N}}$ *une suite de vecteurs aléatoires de* $\mathbb{R}^d$ *indépendants, du second ordre, de même loi, d'espérance* $\mu \in \mathbb{R}^d$ *et de matrice de variance-covariance* Σ *(matrice* $d \times d$*). Alors*

$$\sqrt{n}\left(\frac{\sum_{i=1}^n X_i}{n} - \mu\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,\Sigma) \quad \textit{quand } n \to \infty.$$

Annexe A

Rappels de cours sur la loi normale

A.1 Loi normale centrée réduite

Définition A.1
On dit qu'une variable aléatoire X suit une loi normale centrée réduite si elle admet comme densité de probabilité la fonction

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

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

Proposition A.1
Si $X \sim \mathcal{N}(0,1)$, alors $\mathbf{E}(X) = 0$ et $\mathbf{V}(X) = 1$.

Définition A.2
La fonction de répartition d'une variable aléatoire suivant une loi normale $\mathcal{N}(0,1)$ est :

$$F(x) = \mathbf{P}(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

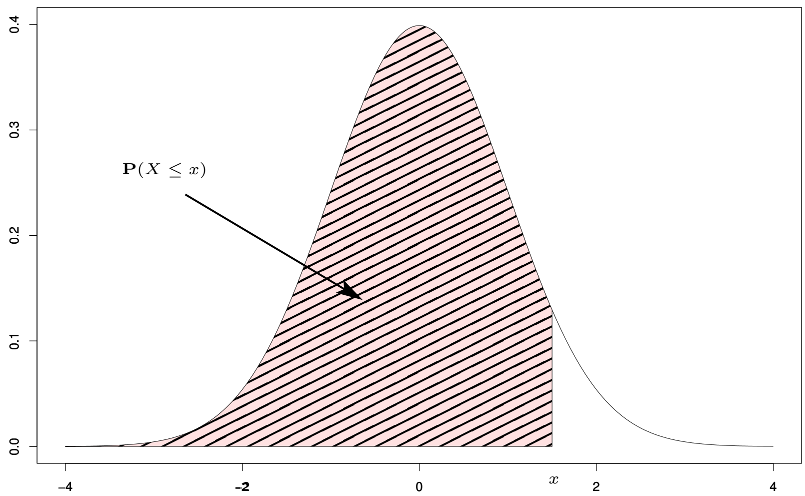


FIGURE A.1 – Fonction de répartition.

Cette fonction ne s'exprime pas à l'aide des fonctions usuelles, on peut trouver ses valeurs dans une table.

Proposition A.2
$\forall x \in \mathbb{R}$, on a $F(-x) = 1 - F(x)$.

Exemple A.1 (Lecture de table)
Soit X une variable aléatoire qui suit une loi $\mathcal{N}(0,1)$. On a

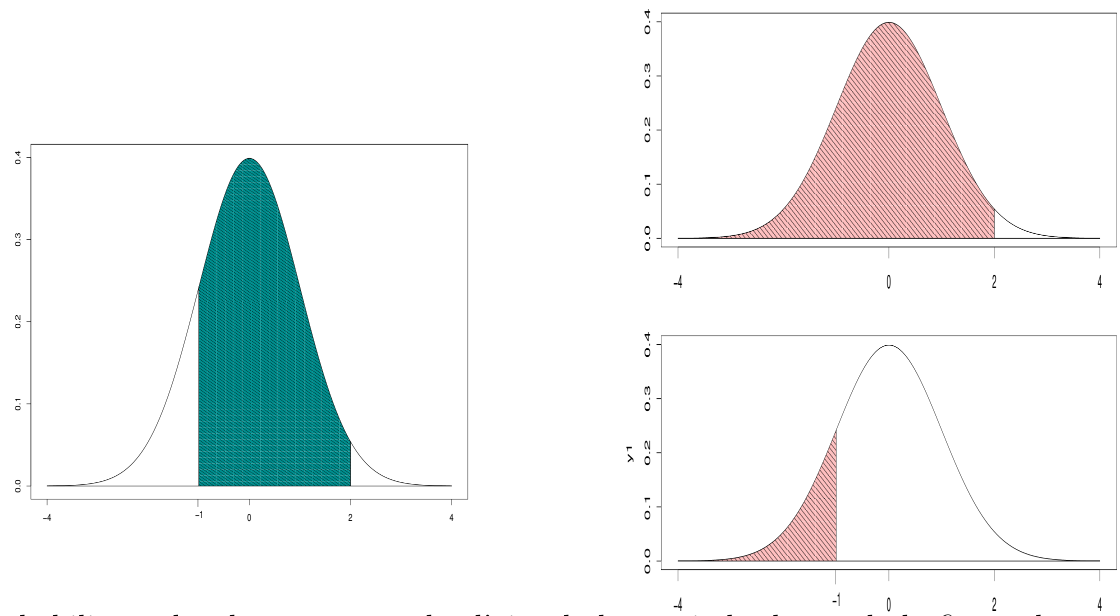
$$\begin{aligned}
&\mathbf{P}(X \leq 2.11) = F(2.11) = 0.9826 \\
&\mathbf{P}(X \leq -2.11) = F(-2.11) = 1 - F(2.11) = 1 - 0.9826 = 0.0174 \\
&\mathbf{P}(X \geq 2.11) = 1 - \mathbf{P}(X \leq 2.11) = 1 - F(2.11) = 1 - 0.9826 = 0.0174
\end{aligned}$$

Pour calculer une probabilité du genre $\mathbf{P}(-1 \leq x \leq 2)$ on peut raisonner de deux manières différentes :

1. **Analytiquement :**

$$\begin{aligned}
\mathbf{P}(-1 \leq x \leq 2) &= \int_{-1}^{2} f(t)dt = \int_{-1}^{-\infty} f(t)dt + \int_{-\infty}^{2} f(t)dt = -\int_{-\infty}^{-1} f(t)dt + \int_{-\infty}^{2} f(t)dt \\
&= -F(-1) + F(2) = F(2) - (1 - F(1)) = 0.9772 - 1 + 0.8413 = 0.8185.
\end{aligned}$$

2. **Géométriquement :** cette approche est plus naturelle.



La probabilité recherchée correspond à l'aire de la partie hachurée de la figure de gauche. Il est évident que cette aire est égale à l'aire hachurée de la figure "droite-haut" moins celle de "droite-bas". En terme de probabilité, ceci s'écrit :

$$\begin{aligned}
\mathbf{P}(-1 \leq X \leq 2) &= \mathbf{P}(X \leq 2) - \mathbf{P}(X \leq -1) = F(2) - F(-1) = F(2) - (1 - F(1)) \\
&= 0.9772 - 1 + 0.8413 = 0.8185.
\end{aligned}$$

Définition A.3
Etant donné un réel α dans $[0,1]$, le quantile u_α d'ordre α (voir Figure A.2) est défini par la relation

$$\mathbf{P}(X \leq u_\alpha) = \alpha.$$

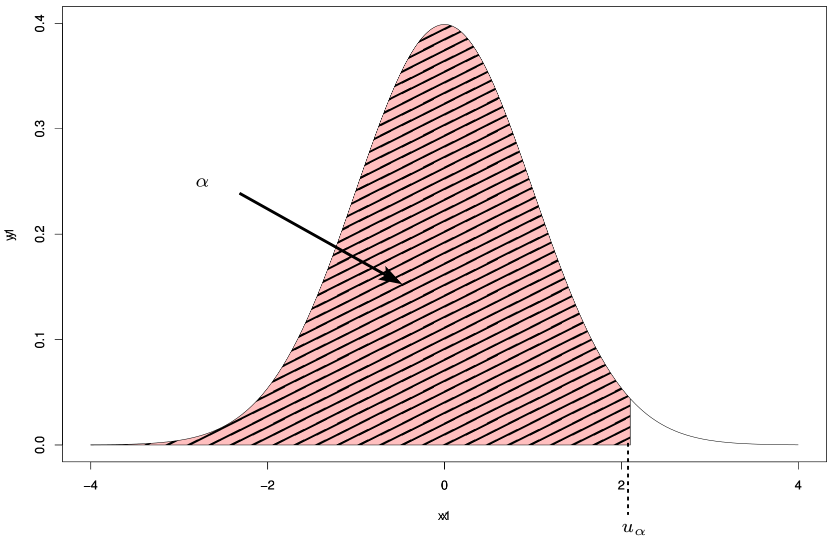


FIGURE A.2 – Représentation du quantile d'ordre α.

Tout comme la fonction de répartition, les quantiles de la loi normale centrée réduite ne s'expriment pas à l'aide des fonctions usuelles et il faut les lire sur des tables. Il faut toujours penser à vérifier que les quantiles dont l'ordre est inférieur à 0.5 sont des valeurs négatives.

Exemple A.2
On lit sur la table :

$$u_{0.95} = 1.64\,; \quad u_{0.05} = -1.64\,; \quad u_{0.57} = 0.18\,; \quad u_{0.32} = -0.47\,.$$

A.2 La loi normale $\mathcal{N}(\mu, \sigma^2)$

Définition A.4
On dit qu'une variable Y suit une loi $\mathcal{N}(\mu, \sigma^2)$ si elle admet comme densité de probabilité la fonction :

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

Proposition A.3
Si $Y \sim \mathcal{N}(\mu, \sigma^2)$ alors

1. $\mathbf{E}(X) = \mu$ *et* $\mathbf{V}(X) = \sigma^2$.
2. $\forall (a, b) \in \mathbb{R}^2, \quad aY + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

Pour obtenir les valeurs de la fonction de répartition et des quantiles d'une loi $\mathcal{N}(\mu, \sigma^2)$, on se ramène à l'utilisation des tables de la loi $\mathcal{N}(0, 1)$ à l'aide du théorème suivant.

Théorème A.1
Si $Y \sim \mathcal{N}(\mu, \sigma^2)$, alors la variable aléatoire

$$X = \frac{Y - \mu}{\sigma}$$

suit une loi $\mathcal{N}(0, 1)$.

Exemple A.3
Soit $Y \sim N(2,9)$, déterminons $\mathbf{P}(Y \leq 1.4)$ et le second quartile de Y. On pose $X = \frac{Y-2}{3}$, d'après le théorème précédent, X suit une loi $\mathcal{N}(0,1)$. On a alors :

$$\begin{aligned}\mathbf{P}(Y \leq 1.4) &= \mathbf{P}\left(\frac{Y-2}{3} \leq \frac{1.4-2}{3}\right) = \mathbf{P}(X \leq -0.2) \\ &= F(-0.2) = 1 - F(0.2) = 1 - 0.579 = 0.421.\end{aligned}$$

Soit $q_{0.75}$ le second quartile de Y, c'est à dire :

$$\mathbf{P}(Y \leq q_{0.75}) = 0.75,$$

ou encore

$$\mathbf{P}\left(\frac{Y-2}{3} \leq \frac{q_{0.75}-2}{3}\right) = 0.75 \quad \Longleftrightarrow \quad \mathbf{P}\left(X \leq \frac{q_{0.75}-2}{3}\right) = 0.75$$

On en déduit que $\dfrac{q_{0.75}-\mu}{\sigma}$ est la quantile d'ordre 0.75 de la loi $\mathcal{N}(0,1)$, par conséquent

$$\frac{q_{0.75}-2}{3} = 0.6745 \quad \Longleftrightarrow \quad q_{0.75} = 4.0235.$$

A.3 Somme et moyenne de lois normales

Théorème A.2
Soient Y_1 et Y_2 deux variables aléatoires **indépendantes** *de lois respectives $\mathcal{N}(\mu_1,\sigma_1^2)$ et $\mathcal{N}(\mu_2,\sigma_2^2)$. Alors*

$$Y_1 + Y_2 \sim \mathcal{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2).$$

Corollaire A.1
Soient $Y_1,\dots,Y_n$ n variables aléatoires **indépendantes** *et identiquement distribuées de lois normales $\mathcal{N}(\mu,\sigma^2)$. Alors*

$$\overline{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right),$$

ou encore

$$\frac{\overline{Y}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1).$$

Annexe B

Lois usuelles dans $\mathbb{R}$

B.1 Lois discrètes

Loi	Paramètres	Fonction de masse	Espérance	Variance	Modélisation	Observations
Dirac δ_a	a	$\mathbf{1}_a(x),\ x \in \mathbb{R}$	a	0	X est une v.a. prenant la valeur a quel que soit le résultat de l'expérience.	Fonction caratéristique (F.c.) $\varphi(t) = e^{ita}$.
Uniforme de support $\{1, \dots, n\}$		$\frac{1}{n}$, $x \in \{1, \dots, n\}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	X est une v.a. désignant un nombre entier compris entre 1 et n choisi au hasard (avec équiprobabilité).	
Bernoulli $\mathcal{B}(p)$	$p \in [0, 1]$	$p^x(1-p)^{1-x}$, $x \in \{0, 1\}$	p	$p(1-p)$	X est le résultat d'une exprérience à deux issues : 1 avec probabilité p, 0 avec probabilité $1-p$.	Si $X_i \sim \mathcal{B}(p)$, $i = 1, \dots, n$ indépendantes alors $\sum_{i=1}^n X_i \sim \mathcal{B}(n, p)$. F.c. $\varphi(t) = pe^{it} + (1-p)$.
Binomiale $\mathcal{B}(n, p)$	$n \in \mathbb{N}^\star, p \in [0, 1]$	$\binom{n}{x} p^x (1-p)^{n-x}$ $x \in \{0, 1, \dots, n\}$	np	$np(1-p)$	X est la somme de n expériences de loi de Bernoulli indépendantes et de même paramètre.	Si $X_1 \sim \mathcal{B}(n_1, p)$ et $X_2 \sim \mathcal{B}(n_2, p)$ avec X_1 et X_2 indépendantes alors $X_1 + X_2 \sim \mathcal{B}(n_1 + n_2, p)$. F.c. $\varphi(t) = (pe^{it} + (1-p))^n$.
Poisson $\mathcal{P}(\lambda)$	$\lambda > 0$	$\frac{\lambda^x}{x!} e^{-\lambda}$, $x \in \mathbb{N}$	λ	λ	Loi limite d'une binomiale $\mathcal{B}(n, p)$ lorsque $n \to \infty$, $p \to 0$ et $np \to \lambda$.	Si $X_1 \sim \mathcal{P}(\lambda_1)$, $X_2 \sim \mathcal{P}(\lambda_2)$ avec X_1 et X_2 indépendantes alors $X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$. F.c. $\varphi(t) = e^{\lambda(e^{it}-1)}$.
Géométrique $\mathcal{G}(p)$	$p \in [0, 1]$	$(1-p)^{x-1}p$, $x \in \{1, 2, \dots\}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	X est la v.a. correspondant au nombre d'essais nécessaires pour obtenir un succés lorsque l'on répète plusieurs fois une expérience de Bernoulli.	F.c. $\varphi(t) = \frac{pe^{it}}{1-(1-p)e^{it}}$.

Loi	Paramètres	Fonction de masse	Espérance	Variance	Modélisation	Observations
Binomiale négative	$n \in \mathbb{N}^\star, p \in [0,1]$	$\binom{x-1}{n-1} p^n (1-p)^{x-n}$, $x \in \{n, n+1, \ldots\}$	$\frac{n}{p}$	$\frac{n(1-p)}{p^2}$	X est la v.a. correspondant au nombre de fois où il a fallu répéter une expérience de Bernoulli pour obtenir n succès.	Si $n = 1$, on retrouve la loi géométrique. Si X_1 et X_2 suivent des loi binomiales négatives de paramètres (n_1, p) et (n_2, p) avec X_1 et X_2 indépendantes, alors $X_1 + X_2$ suit une loi binomiale négative de paramètre $(n_1 + n_2, p)$.
Hypergéométrique	$N \in \mathbb{N}^\star$, $N_1 \in \{1, \ldots, N\}$, $n \in \{1, \ldots, N\}$	$\frac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}}$, $x \in \mathbb{N}$, $x \geq \max(0, n - N + N_1)$, $x \leq \min(n, N_1)$	$\frac{nN_1}{N}$	$\frac{nN_1(N-N_1)(N-n)}{N^2(N-1)}$	On tire n boules sans remise dans une urne contenant N_1 boules blanches, $N - N_1$ boules noires. X est la v.a. correspondant au nombre de boules blanches tirées.	On retrouve la loi binomiale lorsque $N_1 \to \infty$, $N \to \infty$ et $\frac{N_1}{N} \to p$.

B.2 Lois absolument continues

Loi	Paramètres	Densité	Espérance	Variance	Modélisation	Observations
Uniforme $\mathcal{U}([a,b])$	$a, b \in \mathbb{R}$, $a < b$	$\frac{1}{b-a}\mathbf{1}_{[a,b]}(x)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	X est la variable aléatoire représentant un nombre tiré au hasard entre a et b.	F.c. $\varphi(t) = \frac{\sin(at)}{at}$ pour une loi $\mathcal{U}([-a,a])$
Exponentielle $\xi(\lambda)$	$\lambda > 0$	$\lambda e^{-\lambda x}\mathbf{1}_{[0,+\infty[}(x)$	$1/\lambda$	$1/\lambda^2$		$\xi(\lambda) = \gamma(1,\lambda)$. F.c. $\varphi(t) = \lambda/(\lambda - it)$
Gamma $\gamma(p,\lambda)$	$p > 0,\ \lambda > 0$	$\frac{\lambda}{\Gamma(p)}e^{-\lambda x}(\lambda x)^{p-1}\mathbf{1}_{[0,+\infty[}(x)$	p/λ	p/λ^2		Si $X_1 \sim \gamma(p_1,\lambda)$, $X_2 \sim \gamma(p_2,\lambda)$ et X_1 et X_2 indépendantes, alors $X_1+X_2 \sim \gamma(p_1+p_2,\lambda)$. Pour $n \in \mathbb{N}^\star$, $\Gamma(n) = (n-1)!$
Beta I $\beta_1(a,b)$	$a > 0$, $b > 0$	$\frac{1}{\beta(a,b)}(1-x)^{b-1}x^{a-1}\mathbf{1}_{[0,1]}(x)$	$\frac{a}{a+b}$	$\frac{ab}{(a+b+1)(a+b)^2}$	Si $X \sim \gamma(a,1)$ et $Y \sim \gamma(b,1)$ avec X et Y indépendantes, alors $\frac{X}{X+Y} \sim \beta_1(a,b)$	$\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$. Les statistiques d'ordre d'une loi $\mathcal{U}([0,1])$ suivent des lois β_1
Beta II $\beta_2(a,b)$	$a > 0$, $b > 0$	$\frac{1}{\beta(a,b)}\frac{x^{a-1}}{(1+x)^{a+b}}\mathbf{1}_{[0,+\infty[}(x)$	$\frac{a}{b-1}$ si $b > 1$	$\frac{a(a+b-1)}{(b-1)^2(b-2)}$ si $b > 2$	Si $X \sim \gamma(a,1)$ et $Y \sim \gamma(b,1)$ avec X et Y indépendantes, alors $\frac{X}{Y} \sim \beta_2(a,b)$	Si $X \sim \beta_1(a,b)$ alors $\frac{X}{1+X} \sim \beta_2(a,b)$ et réciproquement.
Weibull $W(a,\lambda)$	$a > 1$, $\lambda > 0$	$a\lambda x^{a-1}e^{-\lambda x^a}\mathbf{1}_{[0,+\infty[}(x)$	$\frac{\Gamma(1+\frac{1}{a})}{\lambda^{1/a}}$	$\frac{\Gamma\left(1+\frac{2}{a}\right)-\Gamma^2\left(1+\frac{1}{a}\right)}{\lambda^{2/a}}$	$X \sim W(a,\lambda)$ si $X^a \sim \xi(\lambda)$	
Gaussienne ou normale $\mathcal{N}(m,\sigma^2)$	$m \in \mathbb{R}$, $\sigma^2 > 0$	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-m)^2}{2\sigma^2}}$	m	σ^2	Loi limite de théorème central limite	Si $X \sim \mathcal{N}(m,\sigma^2)$, alors $\frac{X-m}{\sigma} \sim \mathcal{N}(0,1)$. Si $X_1 \sim \mathcal{N}(m_1,\sigma_1^2)$ et $X_2 \sim \mathcal{N}(m_2,\sigma_2^2)$ avec X_1 et X_2 indépendantes alors $X_1 + X_2 \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$. F.c. $\varphi(t) = e^{imt}e^{-\sigma^2t^2/2}$.

Loi	Paramètres	Densité	Espérance	Variance	Modélisation	Observations
Cauchy		$\frac{1}{\pi(1+x^2)}$			Si X_1 et X_2 sont indépendantes de même loi $\mathcal{N}(0,1)$, si $U \sim \mathcal{U}(]-\pi/2, \pi/2[)$, alors X_1/X_2 et $\tan U$ suivent des lois de Cauchy.	F.c. $\varphi(t) = e^{-\lvert t\rvert}$.
Log-normale	$m \in \mathbb{R}$, $\sigma^2 > 0$	$\frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln x - m)^2}{2\sigma^2}} \mathbf{1}_{[0,+\infty[}(x)$	$e^{m+\sigma^2/2}$	$e^{2m+\sigma^2}(e^{\sigma^2}-1)$	X suit une loi log-normale de paramètres (m, σ^2) si $\ln X \sim \mathcal{N}(m, \sigma^2)$.	
Chi-Deux $\chi^2(n)$	$n \in \mathbb{N}^\star$	$\frac{1}{2\Gamma(n/2)} e^{-x/2} \left(\frac{x}{2}\right)^{n/2-1} \mathbf{1}_{[0,+\infty[}(x)$	n	$2n$	Si $X_1, \ldots, X_n$ sont i.i.d. de loi $\mathcal{N}(0,1)$, alors $X = X_1^2 + \ldots + X_n^2 \sim \chi^2(n)$.	Si $X_1 \sim \chi^2(n_1)$ et $X_2 \sim \chi^2(n_2)$ avec X_1 et X_2 indépendantes, alors $X_1 + X_2 \sim \chi^2(n_1 + n_2)$. Si $X \sim \chi^2(n)$, $X \sim \gamma(n/2, 1/2)$.
Student $\mathcal{T}(n)$	$n \in \mathbb{N}^\star$	$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$	0 si $n > 1$	$\frac{n}{n-2}$ si $n > 2$	Si $X \sim \mathcal{N}(0,1)$, $Y \sim \chi^2(n)$ avec X et Y indépendantes, alors $\frac{X}{\sqrt{Y/n}} \sim \mathcal{T}(n)$.	Si $X \sim \mathcal{T}(n)$ alors $\frac{X^2}{n} \sim \beta_2\left(\frac{1}{2}, \frac{n}{2}\right)$.
Fisher-Snedecor $\mathcal{F}(m,n)$	$m \in \mathbb{N}^\star$, $n \in \mathbb{N}^\star$	$\frac{m^{\frac{m}{2}} n^{\frac{n}{2}}}{\beta\left(\frac{m}{2}, \frac{n}{2}\right)} x^{\frac{m}{2}-1} (n + mx)^{-\frac{n+m}{2}}$	$\frac{n}{n-2}$ si $n > 2$	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ si $n > 4$	Si $X \sim \chi^2(m)$, $Y \sim \chi^2(n)$ avec X et Y indépendantes, alors $\frac{X/m}{Y/n} \sim \mathcal{F}(m,n)$.	Si $X \sim \mathcal{F}(m,n)$, alors $1/X \sim \mathcal{F}(n,m)$.

Annexe C

Annales

Cette annexe comporte les sujets d'examens des dernières années acompagnés de quelques corrigés.

Examen partiel Statistique 2

9 décembre 2009
Aucun document, pas de calculatrice

Le barème est donné à titre indicatif

Exercice 1 (Question de cours, 3 points)

1. Enoncer l'inégalité de Markov (sans oublier de donner les hypothèses).
2. Enoncer l'inégalité de Bienaymé-Chebychev (sans oublier de donner les hypothèses).
3. Démontrer l'inégalité de Bienaymé-Chebychev.

Exercice 2 (Questionnaire à choix multiple, 5 points)
Pour chaque question, on reportera sur la copie le numéro de la question avec celui de la réponse exacte. Le choix devra être justifié en 2 lignes maximum.

1. Soit X une variable aléatoire réelle de denstié f_X. La variabe aléatoire $Y = X + 1$ admet pour densité :
   (a) $f_X(y + 1)$
   (b) $1 + f_X(y)$
   (c) $1 - f_X(y)$
   (d) $f_X(y - 1)$
   (e) une autre
2. Soit X une variable aléatoire réelle qui suit une loi normale. Alors $2X + 1$ suit également une loi normale
   (a) Oui
   (b) Non (ou dans certains cas seulement)
3. Si X et Y sont deux variables aléatoires réelles telle que $\rho(X, Y) = 0$ (ρ désigne le coefficient de correlation), alors X et Y sont indépendantes.
   (a) Oui
   (b) Non (ou dans certains cas seulement)
4. Soit X et Y deux variables aléatoires indépendantes. Alors $X + Y$ et Y sont indépendantes.
   (a) Oui
   (b) Non (ou dans certains cas seulement)
5. On note F_X la fonction de répartition de la loi normale centrée réduite (pour une variable aléatoire réelle). On a pour tout $x \in \mathbb{R}$:
   (a) $F_X(-x) = -F_X(x)$
   (b) $F_X(-x) = 1 - F_X(x)$
   (c) $F_X(-x) = F_X(x)$
   (d) aucune de ces égalités n'est correcte

Exercice 3 (3,5 points)
Soit $X = (X_1, X_2)'$ un vecteur aléatoire de loi uniforme sur le carré $]-1,1[\times]-1,1[$.

1. Déterminer la densité de X ainsi que de ses marginales.
2. Calculer le coefficient de correlation entre X_1 et X_2.
3. Déterminer la densité de $Y = X_1 - X_2$ (on pourra par exemple utiliser le produit de convolution).

Exercice 4 (5 points)
Soit Y une variable aléatoire de loi exponentielle de paramètre $\lambda > 0$:

$$f_Y(y) = \lambda e^{-\lambda y}\mathbf{1}_{[0,+\infty[}(y).$$

On considère la variable aléatoire $X = e^Y$.

1. Montrer que X admet pour densité
$$f_X(x) = \frac{\lambda}{x^{\lambda+1}}\mathbf{1}_{[1,+\infty[}(x).$$
2. Calculer $\mathbf{E}[X]$ en fonction de λ.
3. On suppose que $\lambda > 1$. Soit Z une variable aléatoire réelle indépendante de X et de loi uniforme sur $]0,1[$. Déterminer la densité de la variable aléatoire Z/X.

Exercice 5 (3,5 points)
Soit $Z = (X, Y)'$ un couple aléatoire dont la densité est donnée par

$$f_Z(x,y) = \begin{cases} \frac{k}{\sqrt{xy}} & \text{si } 0 < x \leq y < 1 \\ 0 & \text{sinon.} \end{cases}$$

1. Montrer que $k = \frac{1}{2}$.
2. Calculer les densités marginales.
3. On note $A =]-2, 1/4[\times]-1, 3/2[$. Déterminer $\mathbf{P}(Z \in A)$.
4. Les variables aléatoires X et Y sont-elles indépendantes ? Justifier.

CORRECTION

Exercice 1

1. Si X est une v.a.r. positive, on a pour tout réel $a > 0$

$$\mathbf{P}(X \geq a) \leq \frac{\mathbf{E}[X]}{a}.$$

2. Si $\mathbf{E}[X^2] < +\infty$, alors on a pour tout réel $a > 0$

$$\mathbf{P}(|X - \mathbf{E}[X]| > a) \leq \frac{\mathbf{V}[X]}{a^2}.$$

3. Il suffit d'appliquer Markov à la v.a.r. positive $|X - \exp[X]|^2$:

$$\mathbf{P}(|X - \mathbf{E}[X]| > a) = \mathbf{P}((X - \mathbf{E}[X])^2 > a^2) \leq \frac{\mathbf{E}\left[(X - \mathbf{E}[X])^2\right]}{a^2} = \frac{\mathbf{V}[X]}{a^2}.$$

Exercice 2

1. d par un changement de variable
2. a la loi normale est stable par transformation affine
3. b contre exemple (vu en cours) $X \sim \mathcal{N}(0,1)$ et $Y = X^2$
4. b on peut trouver plein de contre exemple
5. b la densité est paire

Exercice 3

1. On a clairement

$$f_X(x_1, x_2) = \frac{1}{4}\mathbf{1}_{]-1,1[^2}(x_1, x_2)$$

et les marginales sont de même loi uniforme sur $]-1, 1[$.

2. $\rho(X_1, X_2) = 0$ car X_1 et X_2 sont indépendantes.
3. La loi de $Y = X_1 - X_2$ est donnée par le produit de convolution entre X_1 et X_2. Soit $y \in [-2, 2]$, on a

$$f_Y(y) = \frac{1}{4}\int_{\mathbb{R}} \mathbf{1}_{]-1,1[}(y-t)\mathbf{1}_{]-1,1[}(t) = \int_{\max(y-1,-1)}^{\min(1,1+y)} \frac{1}{4}\,\mathrm{d}t = \frac{1}{4}(2 - |y|).$$

On a donc

$$f_Y(y) = \frac{1}{4}(2 - |y|)\mathbf{1}_{[-2,2]}(y).$$

Exercice 4

1. On peut utiliser la fonction muette. Soit h une fonction mesurable. On a

$$\mathbf{E}[h(X)] = \int_0^{\infty} h(e^y)\lambda e^{-\lambda y}\,\mathrm{d}y = \int_1^{+\infty} h(x)\lambda e^{-\lambda \ln(x)}\frac{1}{x}\,\mathrm{d}y = \int_1^{+\infty} h(x)\frac{\lambda}{x^{\lambda+1}}\,\mathrm{d}x$$

2. On a
$$\mathbf{E}[X] = \int_1^{+\infty} \frac{\lambda}{x^\lambda} \, \mathrm{d}x.$$
On déduit $\mathbf{E}[X] = +\infty$ si $\lambda \leq 1$, $\mathbf{E}[X] = \frac{\lambda}{\lambda-1}$ sinon.
3. On pose $U = Z/X$ et $V = Z$. Cherchons la loi du couple $T = (U, V)'$. On considère le changement de variable
$$\begin{aligned} \varphi :]1, +\infty[\times]0, 1[&\to \Delta = \{(u, v) : 0 < v < u < 1\} \\ (u, v) &\mapsto (v/u, v) \end{aligned}$$
Le déterminant de la matrice jacobienne vaut $-v/u^2$. Il vient donc
$$f_T(u, v) = \lambda v^{-\lambda} u^{\lambda-1} \mathbf{1}_{0 \leq u \leq v \leq 1}.$$
On obtient la densité de U en intégrant par rapport à v
$$f_U(u) = \lambda u^{\lambda-1} \int_u^1 v^{-\lambda} \, \mathrm{d}v = \frac{\lambda}{\lambda - 1}(1 - u^{\lambda-1}).$$

Exercice 5

1. L'intégrale sur $\mathbb{R}^2$ doit être égale à 1 :
$$k \int_0^1 \frac{1}{\sqrt{x}} \int_x^1 \frac{1}{\sqrt{y}} \, \mathrm{d}y \, \mathrm{d}x = 2k.$$
2. On les obtient par intégration de la densité jointe :
$$f_X(x) = \left(\frac{1}{\sqrt{x}} - 1 \right) \mathbf{1}_{]0,1[}(x) \quad \text{et} \quad f_Y(y) = \mathbf{1}_{]0,1[}(y).$$
3. On a
$$\mathbf{P}(X \in A) = \frac{1}{2} \int_0^{1/4} \int_x^1 \frac{1}{\sqrt{xy}} \, \mathrm{d}y \, \mathrm{d}x = \frac{3}{4}.$$
4. Non car $f_Z(x, y) \neq f_X(x) f_Y(y)$.

Examen de Statistique 2
25 janvier 2010, deux heures
Document autorisé : une feuille A4 manuscrite, pas de calculatrice

---

Le barème est donné à titre indicatif

Exercice 1 (Questionnaire à choix multiple, 5 points)
Pour chaque question, on reportera sur la copie le numéro de la question avec celui de la réponse exacte. Le choix devra être justifié en 2 lignes maximum.

1. Soit X et Y deux variables aléatoires réelles indépendantes qui suivent une loi normale centrée réduite. Alors le vecteur aléatoire $(X, Y)'$ est un vecteur gaussien.
   (a) Oui
   (b) Non (ou dans certains cas seulement)
2. Si $X = (X_1, \ldots, X_d)'$ est un vecteur gaussien, alors les variables aléatoires $X_1, \ldots, X_d$ sont des variables aléatoires gaussiennes.
   (a) Oui
   (b) Non (ou dans certains cas seulement)
3. Soit X une variable aléatoire réelle de densité f_X. La variable aléatoire $Y = 2X$ admet pour densité :
   (a) $\frac{1}{2} f_X(x)$
   (b) $\frac{1}{2} f_X(\frac{x}{2})$
   (c) $2 f_X(x)$
   (d) $2 f_X(2x)$
   (e) une autre
4. Soit X et Y deux variables aléatoires réelles de densités respectives $f_X(x)$ et $f_Y(y)$. Alors la densité de $X + Y$ se calcule en faisant le produit de convolution entre f_X et f_Y.
   (a) Oui
   (b) Non (ou dans certains cas seulement)
5. Soit X_n une suite de variables aléatoires réelles qui converge en probabilité vers une variable aléatoire X et Y_n une suite de variables aléatoires réelles qui converge en loi vers un réel a. Alors
   (a) $X_n + Y_n$ converge en loi vers $X + a$ mais pas en probabilité
   (b) $X_n + Y_n$ converge en probabilité vers $X + a$ mais pas en loi
   (c) $X_n + Y_n$ converge en probabilité et en loi vers $X + a$
   (d) $X_n + Y_n$ ne converge ni en probabilité ni en loi vers $X + a$

Exercice 2 (6 points)
Rappels : *On rappelle que la densité d'une variable aléatoire X qui suit une loi exponentielle de paramètre $\lambda > 0$ est donnée par*
$$f_X(x) = \lambda e^{-\lambda x} \mathbf{1}_{]0,+\infty[}(x).$$
De plus, on admettra que son moment d'ordre n est donné par $\mathbf{E}[X^n] = \frac{n!}{\lambda^n}$.

Soit $(X, Y)'$ un couple aléatoire de densité jointe :
$$f(x, y) = e^{-y} \mathbf{1}_{\{0<x<y\}}.$$

1. Montrer que $f(x, y)$ définit bien une densité.
2. Montrer que X suit une loi exponentielle de paramètre 1 et que la densité marginale de Y vaut
$$f_Y(y) = ye^{-y} \mathbf{1}_{]0,+\infty[}(y).$$
3. Les variables aléatoires X et Y sont elles indépendantes ?
4. Calculer $\mathbf{E}[X]$, $\mathbf{E}[Y]$ et $\mathbf{Cov}(X, Y)$.
5. Déterminer la densité conditionnelle $f_{Y|X=x}(y)$ de Y sachant $X = x$.
6. Donner deux manières différentes de calculer la probabilité
$$\mathbf{P}\left(Y < \frac{1 + X}{2}\right)$$
(juste écrire les formules, les calculs ne sont pas demandés).
7. Déterminer l'espérance conditionnelle $\mathbf{E}[Y|X]$ (on exprimera $\mathbf{E}[Y|X]$ en fonction de X).
8. On considère le couple (Z, T) défini par
$$\begin{cases} Z = X + Y \\ T = Y - X. \end{cases}$$
Déterminer la densité jointe $f_{Z,T}(z, t)$ du couple (Z, T).
9. En déduire les densités marginales de Z et T.

Exercice 3 (6 points)
Rappel : *Si $X = (X_1, \dots, X_d)'$ est un vecteur gaussien d'espérance m_X, de matrice de variance-covariance Σ_X et qui admet une densité f_X alors*
$$f_X(x) = \frac{1}{(\sqrt{2\pi})^d \sqrt{\det(\Sigma_X)}} \exp\left(-\frac{1}{2}(x - m_X)'\Sigma_X^{-1}(x - m_X)\right).$$

Soit $(X, Y)'$ un vecteur gaussien centré (d'espérance $(0, 0)'$) et de matrice de variance-covariance
$$\Sigma = \begin{pmatrix} 1 & u \\ u & 4 \end{pmatrix}.$$

1. Déterminer le coefficient de corrélation linéaire $\rho(X, Y)$ de X et Y en fonction de u. En déduire les valeurs de u pour lesquelles $|\rho(X, Y)| \leq 1$.
2. Pour quelle(s) valeurs de u, les variables marginales sont-elles indépendantes ?
3. Pour quelle(s) valeurs de u, le vecteur gaussien $(X, Y)'$ admet-il une densité ?
4. On suppose maintenant que $u = 1$.
   (a) Donner la densité conjointe du vecteur gaussien $(X, Y)'$.
   (b) Calculer la densité conditionnelle de Y sachant $X = x$ et en déduire la loi de $Y|X = x$.
   (c) Exprimer $\mathbf{E}[Y|X]$ en fonction de X et en déduire la loi de $\mathbf{E}[Y|X]$.
   (d) Quelle est la loi du couple aléatoire $(U, V)' = (-X, X + 2Y)'$? Les variables aléatoires $-X$ et $X + 2Y$ sont elles indépendantes ?

Exercice 4 (6 points)

Soit $X_1, \dots, X_n$ n variables aléatoires indépendantes et de même loi de Bernoulli de paramètre $p \in]0, 1[$. On note

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad T_n = \frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{p(1-p)}}.$$

1. Parmi les différents modes de convergence vus en cours, préciser ceux pour lesquels $\bar{X}_n$ converge vers p.
2. A l'aide de l'inégalité de Bienaymé-Tchebychev, montrer que
$$\mathbf{P}\left(\left|\bar{X}_n - p\right| \leq \sqrt{\frac{p(1-p)}{0.05n}}\right) \geq 0.95.$$
3. En déduire que
$$\mathbf{P}(-4.48 \leq T_n \leq 4.48) \geq 0.95$$
(on pourra utiliser l'approximation $1/\sqrt{0.05} \approx 4.48$).
4. Montrer que T_n converge en loi vers une loi normale centrée réduite.
5. On admet que si X suit un loi normale centrée réduite on a $\mathbf{P}(|X| \leq 1.96) = 0.95$. Déduire de la question précédente que
$$\lim_{n\to\infty} \mathbf{P}(-1.96 \leq T_n \leq 1.96) = 0.95.$$
Comparer ce résultat avec celui de la question 3. Commenter.
6. Est-ce que la variable aléatoire
$$\frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{\bar{X}_n(1-\bar{X}_n)}}$$
converge en loi vers une loi normale centrée réduite ? Justifier votre réponse.

CORRECTION

Exercice 1

1. a car X et Y sont indépendantes.
2. a toute combinaison linéaire des marginales est une v.a.r. gaussienne
3. b changement de variable
4. b il faut que X et Y soient indépendantes
5. c on peut additionner les convergences en loi lorsque l'une des suites converge vers une constante.

Exercice 2

1. $f(x, y)$ est mesurable, positive et intègre à 1 sur $\mathbb{R}^2$, c'est donc bien une densité de probabilité.
2. On calcule les densités marginales. Pour $x > 0$, on a
$$f_X(x) = \int_x^{+\infty} e^{-y}\, \mathrm{d}y = e^{-x}.$$
On reconnaît la densité d'une loi exponentielle de paramètre 1. De même, pour $y > 0$, on a
$$f_Y(y) = \int_0^y e^{-y}\, \mathrm{d}x = ye^{-y}$$
3. Les variables ne sont pas indépendantes puisque $f(x, y) \neq f_X(x)f_Y(y)$.
4. On a d'après l'indication $\mathbf{E}[X] = 1$ et $\mathbf{E}[Y] = 2! = 2$. Pour la covariance :
$$\mathbf{E}[XY] = \int\int_{\mathbb{R}^2} xyf(x, y)\, \mathrm{d}x\, \mathrm{d}y = \frac{1}{2}\int_0^{+\infty} y^3 e^{-y}\, \mathrm{d}y = \frac{1}{2}3! = 3.$$
On déduit $\mathbf{Cov}(X, Y) = 1$.
5. Pour tout $x > 0$, le densité conditionnelle $f_{Y|X=x}(y)$ vaut
$$f_{Y|X=x}(y) = \frac{f(x, y)}{f_X(x)} = e^{x-y}\mathbf{1}_{\{x<y\}}.$$
6. Pour la première méthode on utilise la loi jointe :
$$\mathbf{P}\left(Y < \frac{1+X}{2}\right) = \int\int_{\mathbb{R}^2} \mathbf{1}_{\{y<\frac{1+x}{2}\}} f(x, y)\, \mathrm{d}x\, \mathrm{d}y.$$
Pour la deuxième méthode, on utilise le conditionnement :
$$\mathbf{P}\left(Y < \frac{1+X}{2}\right) = \int_{\mathbb{R}} \mathbf{P}\left(Y < \frac{1+x}{2}|X = x\right) f_X(x)\, \mathrm{d}x,$$
avec
$$\mathbf{P}\left(Y < \frac{1+x}{2}|X = x\right) = \int_{\mathbb{R}} \mathbf{1}_{\{y<\frac{1+x}{2}\}} f_{Y|X=x}(y)\, \mathrm{d}y.$$
Les deux calculs doivent donner $1 + e^{-1} - 2e^{-1/2} \approx 0.155$.

7. On a
$$\mathbf{E}[Y|X=x] = \int_x^{+\infty} y e^{x-y}\,\mathrm{d}y = x+1.$$
On a donc $\mathbf{E}[Y|X] = X+1$.
8. On considère le changement de variable
$$\begin{aligned}\varphi : U &\to V\\ (x,y) &\mapsto (z,t) = (x+y, y-x)\end{aligned}$$
avec $U = \{(x,y)\in\mathbb{R}^2 : 0<x<y\}$ et
$$\begin{aligned}V =&\{(z,t)\in\mathbb{R}^2 : \exists (x,y), 0<x<y, z=x+y, t=y-x\}\\ =&\{(z,t)\in\mathbb{R}^2 : \exists (x,y), 0<x<y, x=(z-t)/2, y=(z+t)/2\}\\ =&\{(z,t)\in\mathbb{R}^2 : 0<z-t<z+t\}\\ =&\{(z,t)\in\mathbb{R}^2 : 0<t<z\}.\end{aligned}$$
On obtient grâce au théorème de changement de variable
$$f_{Z,T}(z,t) = f\left(\varphi^{-1}(z,t)\right)|\det(J_{\varphi^{-1}})| = \frac{1}{2}e^{-(z+t)/2}\mathbf{1}_{\{0<t<z\}}.$$
9. Les densités marginales sont
$$f_Z(z) = e^{-z/2}\mathbf{1}_{\{z>0\}} - e^{-z}\mathbf{1}_{\{z>0\}} \quad \text{et} \quad f_T(t) = e^{-t}\mathbf{1}_{\{t>0\}}.$$

Exercice 3

1. On a $\rho(X,Y) = u/2$. Par conséquent $|\rho(X,Y)|\leq 1$ si et seulement si $u\in[-2,2]$.
2. Les marginales sont indépendantes si et seulement si la $\mathbf{Cov}(X,Y)=0$ (car $(X,Y)'$ est un vecteur gaussien), donc si et seulement si $u=0$.
3. $(X,Y)'$ admet une densité si et seulement si $\det(\Sigma)\neq 0$, c'est-à-dire si et seulement si $u\neq -2$ et $u\neq 2$.
4. (a) La densité de $(X,Y)'$ est
$$f(x,y) = \frac{1}{2\pi\sqrt{3}}e^{-1/6(4x^2-2xy+y^2)}.$$
(b) Comme $X\sim\mathcal{N}(0,1)$, on déduit
$$f_{Y|X=x}(y) = \frac{1}{\sqrt{2\pi}\sqrt{3}}e^{-1/2((y-x)^2/3)}.$$
On déduit que $Y|X=x$ suit une loi $\mathcal{N}(x,3)$.
(c) On a donc $\mathbf{E}[Y|X=x]=x$ et donc $\mathbf{E}[Y|X]=X$. Par conséquent $\mathbf{E}[Y|X]\sim\mathcal{N}(0,1)$.
(d) On a
$$\begin{pmatrix}U\\V\end{pmatrix} = \begin{pmatrix}-1 & 0\\ 1 & 2\end{pmatrix}\begin{pmatrix}X\\Y\end{pmatrix}.$$
Par conséquent $(U,V)'$ est un vecteur gaussien centré et de matrice de variance-covariance
$$\begin{pmatrix}-1 & 0\\ 1 & 2\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & 4\end{pmatrix}\begin{pmatrix}-1 & 0\\ 1 & 2\end{pmatrix}' = \begin{pmatrix}1 & -3\\ -3 & 21\end{pmatrix}.$$
On déduit que les variables $-X$ et $X+2Y$ ne sont pas indépendantes.

Exercice 4

1. Les X_i admettent un moment d'ordre 2 et sont indépendantes, on peut donc utiliser toutes les lois des grands nombres vues en cours. En particulier on a convergence presque sûre et dans L_2, ce qui implique qu'on a également convergence en probabilité, en loi et dans L_1.
2. On applique Bienaymé-Tchebychev avec $\varepsilon = \sqrt{\frac{p(1-p)}{0.05n}}$:

$$\mathbf{P}\left(|\bar{X}_n - p| > \sqrt{\frac{p(1-p)}{0.05n}}\right) \leq \mathbf{V}(\bar{X}_n)\frac{0.05n}{p(1-p)},$$

comme $\mathbf{V}(\bar{X}_n) = \frac{p(1-p)}{n}$, il vient

$$\mathbf{P}\left(|\bar{X}_n - p| > \sqrt{\frac{p(1-p)}{0.05n}}\right) \leq 0.05$$

d'où le résultat.
3. Il suffit de récrire l'inégalité de la question précédente

$$\mathbf{P}\left(|\bar{X}_n - p| \leq \sqrt{\frac{p(1-p)}{0.05n}}\right) = \mathbf{P}\left(-\frac{1}{\sqrt{0.05}} \leq T_n \leq \frac{1}{\sqrt{0.05}}\right) \geq 0.95.$$

4. C'est exactement le théorème central limite.
5. On note F_n la fonction de répartition de T_n et F celle de la loi $\mathcal{N}(0,1)$. Comme T_n converge en loi vers la loi normale centrée réduite, F_n converge vers F en tout point de continuité de F. F étant continue sur $\mathbb{R}$, la convergence a donc lieu sur $\mathbb{R}$. On a donc

$$\lim_{n\to\infty} \mathbf{P}(-1.96 \leq T_n \leq 1.96) = \lim_{n\to\infty} (F_n(1.96) - F_n(-1.96)) = F(1.96) - F(-1.96) = 0.95.$$

L'inégalité de la question 3 (Bienaymé-Tchebychev), donne une approximation moins précise de l'écart entre $\bar{X}_n$ et p. Cependant, l'avantage de cette inégalité est qu'elle est valable pour tout n, contrairement au théorème central limite qui fournit une approximation asymptotique.
6. Comme $\bar{X}_n \xrightarrow{\mathbf{P}} p$, on déduit

$$Y_n = \frac{\sqrt{\bar{X}_n(1-\bar{X}_n)}}{\sqrt{p(1-p)}} \xrightarrow{\mathbf{P}} 1.$$

On obtient grâce au lemme de Slutsky

$$\frac{T_n}{Y_n} = \frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{\bar{X}_n(1-\bar{X}_n)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Examen partiel Statistique 2

8 décembre 2010
Aucun document, pas de calculatrice

Le barème est donné à titre indicatif

Exercice 1 (Question de cours, 4.5 points)

1. Donner un exemple de lois discrète, absolument continu et mixte. Tracer les fonctions de répartitions de ces 3 lois.
2. Etant donné $\alpha \in]0,1[$ on désigne par q_α le quantile de la loi $\mathcal{N}(0,1)$. Soit Y une variable aléatoire de loi normale d'espérance 2 et de variance 9. Exprimer le quantile d'ordre α de Y en fonction de q_α. Justifier votre réponse.
3. Soit $X = (X_1, \ldots, X_n)'$ un vecteur aléatoire suivant une loi multinomiale de paramètre $(n, p_1, \ldots, p_n)$:
   (a) Quelle est la loi de X_1 ?
   (b) Que vaut $\mathbf{Cov}(X_1, X_2)$?
   (c) Montrer le résultat de la question précédente.
4. Soit $X = (X_1, \ldots, X_d)'$ un vecteur gaussien de loi $\mathcal{N}(m_X, \Sigma_X)$. On désigne par M une matrice de dimension $p \times d$ et ℓ un vecteur de $\mathbb{R}^p$. Quelle est la loi de $MX + \ell$?

Exercice 2 (Questionnaire à choix multiple, 2 points)

Pour chaque question, on reportera sur la copie le numéro de la question avec celui de la réponse exacte. Le choix devra être justifié en 2 lignes maximum.

1. Soit X une variable aléatoire réelle de fonction de répartition F_X. La variable aléatoire $Y = 2X + 1$ admet pour fonction de répartition :
   (a) $F_X(2y+1)$
   (b) $\frac{F_X(y)-1}{2}$
   (c) $2F_X(y) + 1$
   (d) $F_X(\frac{y-1}{2})$
   (e) une autre
2. Soit X_1 et X_2 deux variables aléatoires réelles qui suivent une loi normale. Alors $X = (X_1, X_2)'$ est un vecteur gaussien.
   (a) Oui
   (b) Non (ou dans certains cas seulement)
3. Soit X, Y et Z trois variables aléatoires (mutuellement) indépendantes. Alors $X + Z$ et Y sont indépendantes.

(a) Oui

(b) Non (ou dans certains cas seulement)

4. On note F_X la fonction de répartition de la loi normale centrée réduite (pour une variable aléatoire réelle). On a pour tout $x \in \mathbb{R}$:

   (a) $F_X(-x) = -F_X(x)$

   (b) $F_X(-x) = 1 - F_X(x)$

   (c) $F_X(-x) = F_X(x)$

   (d) aucune de ces égalités n'est correcte

Exercice 3 (4,5 points)
Soit $X = (X_1, X_2)'$ un vecteur aléatoire de loi uniforme sur l'ensemble $\{]0,1[\times]0,1[\} \cup \{]1,2[\times]0,2[\}$.

1. Représenter le support de X sur un graphique.
2. Quelle est la densité de X ?
3. Les variables X_1 et X_2 sont elles indépendantes ?
4. Calculer l'espérance X_1.
5. Calculer la médiane de X_1.

Pour les exercices 4 et 5 on rappelle que étant donné $\lambda > 0$, la densité d'une loi exponentielle de paramètre λ est

$$f(x) = \lambda \exp(-\lambda x) \mathbf{1}_{[0,+\infty[}(x).$$

On a de plus $\mathbf{E}[X] = \frac{1}{\lambda}$, $\mathbf{E}[X^2] = \frac{2}{\lambda^2}$ et $\mathbf{V}(X) = \frac{1}{\lambda^2}$.

Exercice 4 (4.5 points)
On considère deux variables aléatoires indépendantes X_1 et X_2 de loi exponentielle de paramètre λ_1 et λ_2 ($\lambda_1 > 0$ et $\lambda_2 > 0$).

1. Quelle est la densité du vecteur aléatoire $X = (X_1, X_2)'$?
2. Soit $Y = \min(X_1, X_2)$ le minimum de ces deux variables aléatoires. Montrer que
$$\mathbf{P}(Y = X_1) = \mathbf{P}(X_1 < X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$
3. Calculer la fonction de répartition de Y et en déduire que Y suit une loi exponentielle de paramètre $\lambda_1 + \lambda_2$.

Exercice 5 (4.5 points)
Soit $(X, Y)'$ un couple aléatoire de densité jointe

$$f(x, y) = c \exp(-y) \mathbf{1}_{0<2x<y}.$$

1. Montrer que $c = 2$.
2. Calculer les densités marginales. Les variables X et Y sont-elles indépendantes ? Justifier.
3. Calculer $\mathbf{E}[X]$, $\mathbf{E}[Y]$ et $\mathbf{Cov}(X, Y)$.
4. On pose $T = 2X - Y$. Calculer la densité de T.

Exercice 1

1. Loi discrète : P_{X_1} loi de Bernoulli de paramètre $p = 1/3$. Loi absolument continue : P_{X_2} loi uniforme sur $]-1,1[$. Loi mixte : $1/2P_{X_1} + 1/2P_{X_2}$. On a

$$F_{X_1}(t) = \begin{cases} 0 & \text{si } t < 0 \\ 2/3 & \text{si } 0 \leq t < 1 \\ 1 & \text{si } t \geq 1 \end{cases}, \quad F_{X_2}(t) = \begin{cases} 0 & \text{si } t < -1 \\ 1/2(t+1) & \text{si } -1 \leq t < 1 \\ 1 & \text{si } t \geq 1 \end{cases}$$

$$\text{et} \quad F_{X_3}(t) = \begin{cases} 0 & \text{si } t < -1 \\ 1/4(t+1) & \text{si } -1 \leq t < 0 \\ 7/12 + t/4 & \text{si } 0 \leq t < 1 \\ 1 & \text{si } t \geq 1 \end{cases}$$

2. On désigne par t_α le quantile d'ordre α de la loi de Y. On a par définition $\mathbf{P}(Y \leq t_\alpha) = \alpha$. Or
$$\mathbf{P}(Y \leq t_\alpha) = \mathbf{P}\left(\frac{Y-2}{3} \leq \frac{t_\alpha - 2}{3}\right) = \mathbf{P}\left(X \leq \frac{t_\alpha - 2}{3}\right) = \alpha.$$
On déduit $\frac{t_\alpha-2}{3} = q_\alpha$ ou encore $t_\alpha = 3q_\alpha + 2$.

3. (a) $X_1 \sim \mathcal{B}(n, p_1)$.
   (b) $\mathbf{Cov}(X_1, X_2) = -np_1p_2$.
   (c) voir page 40.

4. $X \sim \mathcal{N}(Mm_X + \ell, M\Sigma_X M')$.

Exercice 2

1. d car $\mathbf{P}(Y \leq t) = \mathbf{P}(X \leq (t-1)/2)$.
2. b ça devient vrai si X_1 et X_2 sont indépendantes (voir poly pour un contre-exemple).
3. a on a même $g_1(X, Z)$ indépendant de $g_2(Y)$ pour toute fonction réelle mesurable g_1 et g_2.
4. b car $\mathbf{P}(X \leq -x) = \mathbf{P}(X \geq x) = 1 - \mathbf{P}(X \leq -x) = 1 - F_X(x)$.

Exercice 3

1. Le support est représenté sur la Figure C.1.
2. On désigne par $\mathcal{D}$ le support de X. La densité est alors donnée par $f_{X_1,X_2}(x_1, x_2) = \frac{1}{3}\mathbf{1}_{\mathcal{D}}(x_1, x_2)$.
3. X_1 et X_2 ne sont clairement pas indépendantes. Il suffit de prendre l'ensemble $A = [0, 1/2]$ et $B = [1, 1.5]$. On a alors
$$\mathbf{P}(X \in A, Y \in B) = 0 < \mathbf{P}(X \in A)\mathbf{P}(Y \in B).$$
4. On peut calculer la densité marginale de X_1 :
$$f_{X_1}(x_1) = \int_{\mathbb{R}} \frac{1}{3}\mathbf{1}_{\mathcal{D}}(x_1, x_2)\,\mathrm{d}x_2 = \frac{1}{3}\mathbf{1}_{]0,1[}(x_1) + \frac{2}{3}\mathbf{1}_{]1,2[}(x_1).$$
On a alors $\mathbf{E}[X_1] = \frac{1}{3}\frac{1}{2} + \frac{2}{3}\frac{3}{2} = \frac{7}{6}$.

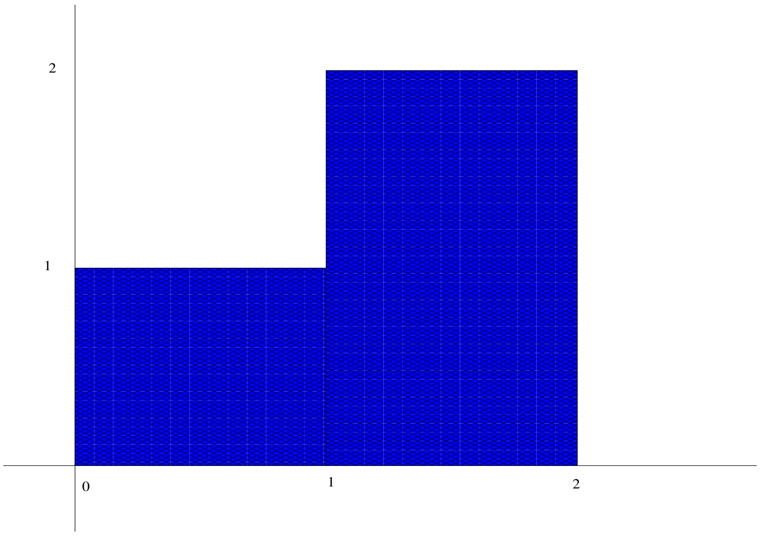


FIGURE C.1 – Support de X.

5. On obtient comme fonction de répartition

$$F_{X_1}(t) = \begin{cases} 0 & \text{si } t < 0 \\ t/3 & \text{si } 0 \le t < 1 \\ 2t/3 - 1/3 & \text{si } t \le 1 < 2 \\ 1 & \text{si } t \ge 2. \end{cases}$$

On déduit que la médiane vaut $5/4$.

Exercice 4

1. Comme X_1 et X_2 sont indépendantes, la densité de X est donnée par

$$f_{X_1,X_2}(x_1, x_2) = \lambda_1\lambda_2 \exp(-\lambda_1 x_1) \exp(-\lambda_2 x_2) \mathbf{1}_{[0,+\infty[}(x_1)\mathbf{1}_{[0,+\infty[}(x_2).$$

2. On a par définition

$$\begin{aligned}\mathbf{P}(Y = X_1) =& \mathbf{P}(X_1 < X_2) \\ =& \int\int \mathbf{1}_{\{x_1<x_2\}}\lambda_1\lambda_2 \exp(-\lambda_1 x_1) \exp(-\lambda_2 x_2)\mathbf{1}_{[0,+\infty[}(x_1)\mathbf{1}_{[0,+\infty[}(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\ =& \int_0^{+\infty}\int_0^{x_2} \lambda_1\lambda_2 \exp(-\lambda_1 x_1) \exp(-\lambda_2 x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = ... = \frac{\lambda_1}{\lambda_1 + \lambda_2}\end{aligned}$$

3. On a pour tout $t \ge 0$:

$$\begin{aligned}F_Y(t) =& \mathbf{P}(\min(X_1, X_2) \le t) = 1 - \mathbf{P}(\min(X_1, X_2) > t) = 1 - \mathbf{P}(X_1 > t, X_2 > t) \\ =& 1 - \mathbf{P}(X_1 > t)\mathbf{P}(X_2 > t) = 1 - \exp(-(\lambda_1 + \lambda_2)t).\end{aligned}$$

D'où $f_Y(t) = (\lambda_1 + \lambda_2) \exp(-(\lambda_1 + \lambda_2)t)$.

Exercice 5

1. L'intégrale sur $\mathbb{R}^2$ doit être égale à 1 :

$$c\int_0^{+\infty}\int_{2x}^{+\infty} \exp(-y)\,\mathrm{d}y\,\mathrm{d}x = c/2.$$

2. On les obtient par intégration de la densité jointe :

$$f_X(x) = 2\exp(-2x)\mathbf{1}_{]0,+\infty[}(x) \quad \text{et} \quad f_Y(y) = y\exp(-y)\mathbf{1}_{]0,+\infty[}(y).$$

   $f(1,3) = 2\exp(-3) \neq 6\exp(-6) = f_X(1)f_Y(3)$ donc X et Y ne sont pas indépendantes.

3. X suit un loi exponentielle de paramètre 2 donc $\mathbf{E}[X] = 1/2$. Pour Y, on a

$$\mathbf{E}[Y] = \int_0^{+\infty} y^2 \exp(-y)\,\mathrm{d}y = 2$$

   d'après le rappel (moment d'ordre 2 d'une loi exponentielle de paramètre 1). Pour la covariance on a

$$\mathbf{E}[XY] = \int\int xy2\exp(-y)\mathbf{1}_{\{0<2x<y\}}\,\mathrm{d}x\,\mathrm{d}y = ... = 3/2, \quad \text{d'où } \mathbf{Cov}(X,Y) = 1/2.$$

4. On pose $Z = Y$ et on cherche la loi du couple $W = (T,Z)'$ à l'aide du changement de variable :

$$\begin{aligned} \varphi : \mathcal{D} = \{(x,y) : 0 < 2x < y\} &\to \Delta = \{(t,z) : 0 < -t < z\} \\ (x,y) &\mapsto (2x-y, y) \end{aligned}$$

   On déduit $\varphi^{-1}(t,z) = ((z+t)/2, z)$. Le déterminant de la matrice Jacobienne de φ^{-1} vaut $1/2$. On obtient ainsi la densité de W :

$$f_{T,Z}(t,z) = \exp(-z)\mathbf{1}_{\{0<-t<z\}}.$$

   La densité cherchée s'obtient en intégrant par rapport à Z :

$$f_T(t) = \exp(t)\mathbf{1}_{\{t<0\}}.$$

Examen de Statistique 2

28 janvier 2011, deux heures

Document autorisé : une feuille A4 manuscrite, pas de calculatrice

Le barème est donné à titre indicatif

Rappel : *Si $X = (X_1, \ldots, X_d)'$ est un vecteur gaussien d'espérance m_X, de matrice de variance-covariance Σ_X et qui admet une densité f_X alors*

$$f_X(x) = \frac{1}{(\sqrt{2\pi})^d \sqrt{\det(\Sigma_X)}} \exp\left(-\frac{1}{2}(x - m_X)'\Sigma_X^{-1}(x - m_X)\right).$$

Exercice 1 (quelques applications "directes" de cours, 5 points)
Les réponses aux questions de cet exercice ne necéssitent pas de longs développements mathématiques (si elles sont bien abordées...). Les réponses données devront être courtes et précises (2 ou 3 lignes maximum).

1. Soit $X = (X_1, X_2)'$ un vecteur aléatoire suivant une loi multinomiale $\mathcal{M}(1, 1/3, 2/3)$. Calculer $\mathbf{P}(X_1 = 1)$ et $\mathbf{P}(X_2 = 0)$. Les variables X_1 et X_2 sont-elles indépendantes ?
2. Soit X_1 et X_2 deux variables aléatoires réelles indépendantes telles que X_1 suit une loi uniforme sur $]1, 2[$ et X_2 suit une loi normale de moyenne 3 et de variance 2. On considère le vecteur aléatoire $X = (X_1, X_2)'$. Donner l'espérance de X ainsi que sa matrice de variance-covariance.
3. Soit $(X, Y)'$ un vecteur aléatoire de $\mathbb{R}^2$ admettant pour densité
$$f(x, y) = \lambda \exp\left(-\frac{1}{2}(2x^2 - 2xy + y^2)\right).$$
Le vecteur (X, Y) est-il un vecteur gaussien ? Si oui, préciser son espérance, sa matrice de variance-covariance ainsi que la valeur de λ. Si non, justifier brièvement.
4. On désigne par q_α le quantile d'ordre alpha de la loi du chi-deux à un degré de liberté et par u_α le quantile d'ordre α de la loi $\mathcal{N}(0, 1)$. Montrer que $q_\alpha = u^2_{(1-\alpha)/2}$.

Pour les deux questions à suivre, on reportera sur la copie le numéro de la question avec celui de la réponse exacte. Le choix devra être justifié en 2 lignes maximum.

5. Soit X une variable aléatoire qui admet une densité f_X continue sur $\mathbb{R}$ et paire. On note F_X la fonction de répartition de X et t un réel positif.
   (a) $F_X(t) = 1 - F_X(-t)$
   (b) $F_X(t) = F_X(-t)$
   (c) $F_X(t) = -F_X(t)$

(d) $F_X(t) = 1/2 + F_X(-t)$

(e) $F_X(-t) = 1/2 + F_X(t)$

(f) aucune de ces égalités est correcte.

6. Si $X = (X_1, \ldots, X_d)'$ est un vecteur gaussien et Y est une variable aléatoire gaussienne, alors le vecteur $(X_1, \ldots, X_d, Y)'$ est un vecteur gaussien.

   (a) Oui

   (b) Non (ou dans certains cas seulement)

Exercice 2 (5 points)
Soit $Z = (X, Y)'$ un couple aléatoire dont la densité est donnée par

$$f_Z(x, y) = \begin{cases} \frac{k}{\sqrt{xy}} & \text{si } 0 < x \leq y < 1 \\ 0 & \text{sinon.} \end{cases}$$

1. Montrer que $k = \frac{1}{2}$.
2. Calculer les densités marginales. En déduire $\mathbf{E}[X]$ et $\mathbf{E}[Y]$.
3. Déterminer la densité conditionnelle de $Y|X = x$.
4. Déterminer l'espérance conditionnelle $\mathbf{E}[Y|X]$.
5. Retrouver la valeur de $\mathbf{E}[Y]$ en utilisant la question précédente.
6. Calculer $\mathbf{P}(Y < 2X|X = x)$.
7. En déduire $\mathbf{P}(Y < 2X)$.

Exercice 3 (5 points)
Soit $X = (X_1, X_2)'$ un vecteur gaussien centré (d'espérance $(0, 0)'$) et de matrice de variance-covariance

$$\Sigma_X = \begin{pmatrix} 1 & u \\ u & 1 \end{pmatrix} \quad \text{avec} \quad u \in [-1, 1].$$

1. On désigne par $\rho(X_1, X_2)$ le coefficient de corrélation linéaire entre X_1 et X_2. Pour quelle(s) valeur(s) de u a-t-on $|\rho(X_1, X_2)| \leq 1$?
2. Pour quelle(s) valeur(s) de u le vecteur X admet-il une densité ?
3. Pour quelle(s) valeur(s) de u les variables $Z = X_1$ et $T = X_1 + 2X_2$ sont-elles indépendantes ?
4. On suppose que $u \in]-1, 1[$.

   (a) Calculer la densité conditionnelle de $X_2|X_1 = x_1$ et en déduire la loi de $X_2|X_1 = x_1$.

   (b) Calculer l'espérance conditionnelle $\mathbf{E}[X_2|X_1]$ et en déduire la loi de $\mathbf{E}[X_2|X_1]$.

   (c) Quelle est la loi du couple aléatoire $(U, V)' = (X_1, X_2 - \mathbf{E}[X_2|X_1])'$? Pour quelle(s) valeur(s) de u les variables X_1 et $X_2 - \mathbf{E}[X_2|X_1]$ sont-elles indépendantes ?

Exercice 4 (5 points)
Soit X une variable aléatoire de densité f_θ définie par :

$$f_\theta(x) = \theta x^{\theta - 1} \mathbf{1}_{]0,1[}(x)$$

où θ est un nombre réel strictement positif fixé.

1. Montrer que
$$\mathbf{E}[X] = \frac{\theta}{\theta+1} \quad \text{et} \quad \mathbf{V}[X] = \frac{\theta}{(\theta+1)^2(\theta+2)}.$$
2. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et de même loi que X. On note
$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad U_n = \frac{\bar{X}_n}{1-\bar{X}_n}.$$
   (a) Montrer que $(\bar{X}_n)_{n\geq 1}$ converge en probabilité vers $g(\theta)$, où g est une fonction que vous préciserez.
   (b) En déduire que $(U_n)_{n\geq 1}$ converge en probabilité vers θ.
   (c) On pose
$$T_n = \frac{1}{1+U_n}\sqrt{\frac{U_n}{U_n+2}}.$$
   La suite $(T_n)_{n\geq 1}$ converge-t-elle en probabilité ? Si oui, déterminer sa limite.
   (d) Déterminer une suite $(a_n)_{n\geq 1}$ et un nombre réel m (fonction de θ) tels que la suite de variables aléatoires
$$\left(a_n \frac{\bar{X}_n - m}{T_n}\right)_{n\geq 1}$$
   converge en loi vers une limite à préciser.

CORRECTION

Exercice 1

1. $\mathbf{P}(X_1 = 1) = \mathbf{P}(X_2 = 0) = 1/3$. $\mathbf{Cov}(X_1, X_2) = -2/3$ donc X_1 et X_2 ne sont pas indépendantes.
2. On a
$$\mathbf{E}[X] = \begin{pmatrix} 1.5 \\ 3 \end{pmatrix} \qquad \text{et} \qquad \Sigma_X = \begin{pmatrix} 1/12 & 0 \\ 0 & 2 \end{pmatrix}.$$
3. On réécrit la densité
$$f(x, y) = \lambda \exp\left[-\frac{1}{2}\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\right].$$
On déduit que $(X, Y)'$ est un vecteur gaussien centré de matrice de variance covariance
$$\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \qquad \text{et} \qquad \lambda = \frac{1}{2\pi}.$$
4. Soit X une variable qui suit une loi $\mathcal{N}(0, 1)$. On a par définition $\mathbf{P}(X^2 \leq q_\alpha) = \alpha$. De plus
$$\mathbf{P}(X^2 \leq q_\alpha) = \mathbf{P}(-\sqrt{q_\alpha} \leq X \leq \sqrt{q_\alpha}) = 1 - 2\mathbf{P}(X \leq -\sqrt{q_\alpha}) = \alpha.$$
On déduit $\mathbf{P}(X \leq -\sqrt{q_\alpha}) = (1 - \alpha)/2$, ou encore $-\sqrt{q_\alpha} = u_{(1-\alpha)/2}$.
5. réponse a. En effet
$$\begin{aligned} F_X(t) =& 1 - \mathbf{P}(X > t) = 1 - \int_t^{+\infty} f_X(t)\,\mathrm{d}t = 1 - \int_{-\infty}^{-t} f_X(-t)\,\mathrm{d}t \\ =& 1 - \int_{-\infty}^{-t} f_X(t)\,\mathrm{d}t = 1 - F_X(-t). \end{aligned}$$
6. réponse b. Pour $d = 1$ on a vu le contre exemple $X \sim \mathcal{N}(0, 1)$ et $Y = \varepsilon X$ avec $\mathbf{P}(\varepsilon = 1) = \mathbf{P}(\varepsilon = -1) = 1/2$.

Exercice 2

1. L'intégrale sur $\mathbb{R}^2$ doit être égale à 1 :
$$k \int_0^1 \frac{1}{\sqrt{x}} \int_x^1 \frac{1}{\sqrt{y}}\,\mathrm{d}y\,\mathrm{d}x = 2k.$$
2. On les obtient par intégration de la densité jointe :
$$f_X(x) = \left(\frac{1}{\sqrt{x}} - 1\right)\mathbf{1}_{]0,1[}(x) \quad \text{et} \quad f_Y(y) = \mathbf{1}_{]0,1[}(y).$$
On a donc $\mathbf{E}[Y] = 0.5$. Pour $\mathbf{E}[X]$, on fait le calcul
$$\mathbf{E}[X] = \int_0^1 x\left(\frac{1}{\sqrt{x}} - 1\right)\,\mathrm{d}x = \frac{1}{6}.$$

3. Pour $0 < x < y < 1$, la densité conditionnelle vaut

$$f_{Y|X=x}(y) = \frac{1}{2}\frac{1}{\sqrt{xy}}\left(\frac{\sqrt{x}}{1-\sqrt{x}}\right) = \frac{1}{2\sqrt{y}}\left(\frac{1}{1-\sqrt{x}}\right).$$

On a donc que pour $x \in]0,1[$,

$$f_{Y|X=x}(y) = \frac{1}{2\sqrt{y}}\left(\frac{1}{1-\sqrt{x}}\right)\mathbf{1}_{0<x<y<1}.$$

4. On déduit pour $x \in]0,1[$,

$$\mathbf{E}[Y|X=x] = \frac{1}{2(1-\sqrt{x})}\int_x^1 \sqrt{y}\,\mathrm{d}y = \frac{x+\sqrt{x}+1}{3},$$

d'où $\mathbf{E}[Y|X] = \frac{X+\sqrt{X}+1}{3}$.

5. On a

$$\begin{aligned}\mathbf{E}[Y] =&\mathbf{E}[\mathbf{E}[Y|X]] = \frac{1}{3}\left[\frac{1}{6}+\mathbf{E}[\sqrt{X}]+1\right]\\ =&\frac{1}{3}\left[\frac{1}{6}+\int_0^1\sqrt{x}\left(\frac{1}{\sqrt{x}}-1\right)\mathrm{d}x+1\right] = \frac{1}{3}\left[\frac{1}{6}+\frac{1}{3}+1\right] = \frac{1}{2}.\end{aligned}$$

6. On a pour $0 < x < 1/2$:

$$\begin{aligned}\mathbf{P}(Y<2X|X=x) &= \int \mathbf{1}_{y<2x} f_{Y|X=x}(y)\,\mathrm{d}y = \int \mathbf{1}_{y<2x}\frac{1}{2\sqrt{y}}\left(\frac{1}{1-\sqrt{x}}\right)\mathbf{1}_{0<x<y\le 1}\,\mathrm{d}y\\ &= \left(\frac{1}{1-\sqrt{x}}\right)\int_x^{2x}\frac{1}{2\sqrt{y}}\,\mathrm{d}y = \frac{\sqrt{2x}-\sqrt{x}}{1-\sqrt{x}}.\end{aligned}$$

7. Il vient donc

$$\mathbf{P}(Y<2X) = \int_0^{1/2}\frac{\sqrt{2x}-\sqrt{x}}{1-\sqrt{x}}\,\frac{1-\sqrt{x}}{\sqrt{x}}\,\mathrm{d}x = \frac{\sqrt{2}-1}{2}.$$

Exercice 3

1. $\rho(X_1,X_2) = u$. Donc $|\rho(X_1,X_2)| \le 1$ si et seulement si $|u| \le 1$.
2. $\det\Sigma_X = 1-u^2$ donc X admet une desité si et seulement si $u \in]-1,1[$.
3. Le vecteur $(Z,T)'$ est gaussien (il existe une matrice A telle que $(Z,T)' = A(X_1,X_2)'$). Z et T sont donc indépendantes si et seulement si leur covariance est nulle. On a :

$$\mathbf{Cov}(Z,T) = \mathbf{V}[X_1] + 2\mathbf{Cov}(X_1,X_2) = 1+2u.$$

On déduit que Z et T sont indépendantes si et seulement si $u = -1/2$.

4. (a) On a

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sqrt{1-u^2}}\exp\left[-\frac{x_1^2+x_2^2-2ux_1x_2}{2(1-u^2)}\right]$$

et

$$f_{X_1}(x_1) = \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{x_1^2}{2}\right].$$

Par conséquent

$$f_{X_2|X_1=x_1}(x_2) = \frac{1}{\sqrt{2\pi}\sqrt{1-u^2}} \exp\left[-\frac{(x_2-ux_1)^2}{2(1-u)^2}\right].$$

On déduit que $X_2|X_1 = x_1$ suit une loi $\mathcal{N}(ux_1, (1-u)^2)$.

(b) Par conséquent $\mathbf{E}[X_2|X_1] = uX_1$ et $\mathbf{E}[X_2|X_1] \sim \mathcal{N}(0, u^2)$.

(c) On a

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -u & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}.$$

Par conséquent $(U, V)'$ est un vecteur gaussien centré et de matrice de variance covariance

$$\begin{pmatrix} 1 & 0 \\ -u & 1 \end{pmatrix} \begin{pmatrix} 1 & u \\ u & 1 \end{pmatrix} \begin{pmatrix} 1 & -u \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1-u^2 \end{pmatrix}$$

Les variables X_1 et $X_2 - \mathbf{E}[X_2|X_1]$ sont donc indépendantes pour tout $u \in]-1, 1[$.

Exercice 4

1. On a

$$\mathbf{E}[X] = \int_0^1 \theta x^\theta \, \mathrm{d}x = \frac{\theta}{\theta+1} \quad \text{et} \quad \mathbf{E}[X^2] = \int_0^1 \theta x^{\theta+1} \, \mathrm{d}x = \frac{\theta}{\theta+2}.$$

Par conséquent

$$\mathbf{V}[X] = \frac{\theta}{(\theta+1)^2(\theta+2)}.$$

2. (a) On obtient par la loi des grands nombres que $\bar{X}_n \xrightarrow{\mathbf{P}} g(\theta) = \frac{\theta}{\theta+1}$.

(b) La fonction

$$\begin{aligned} h :]0,1[&\to \mathbb{R} \\ y &\mapsto \frac{y}{1-y} \end{aligned}$$

est continue. Par conséquent $h(\bar{X}_n) = U_n \xrightarrow{\mathbf{P}} h(g(\theta)) = \theta$.

(c) De même, les fonctions

$$s(y) = \frac{1}{1+y} \quad \text{et} \quad t(y) = \sqrt{\frac{y}{y+2}}$$

étant continues sur $\mathbb{R}^+$, on déduit que

$$T_n \xrightarrow{\mathbf{P}} a = \frac{1}{1+\theta}\sqrt{\frac{\theta}{\theta+2}}.$$

(d) On a d'après le TCL

$$V_n = \sqrt{n}\left(\bar{X}_n - \frac{\theta}{\theta+1}\right) \xrightarrow{\mathcal{L}} V \sim \mathcal{N}\left(0, \frac{\theta}{(\theta+1)^2(\theta+1)}\right).$$

Par conséquent, d'après le théorème de Slutsky, $V_n/T_n \overset{\mathcal{L}}{\to} V/a$. Comme V/a est une variable gaussienne, centrée et de variance

$$\frac{(\theta+1)^2(\theta+2)}{\theta}\,\frac{\theta}{(\theta+1)^2(\theta+2)} = 1,$$

on obtient

$$\sqrt{n}\left(\frac{\bar{X}_n - \frac{\theta}{\theta+1}}{T_n}\right) \overset{\mathcal{L}}{\to} \mathcal{N}\left(0,1\right).$$

Examen partiel Statistique 2

7 décembre 2011
Aucun document, pas de calculatrice

Le sujet est composé de 4 exercices indépendants (il y a 2 pages !).

Le barême est donné à titre indicatif.

Exercice 1 (Question de cours, 5 points)

1. Donner un exemple de lois discrète, absolument continu et mixte. Tracer les fonctions de répartition de ces 3 lois.
2. Soit X une variable aléatoire de loi uniforme sur l'ensemble $]-1, 1/2[\cup[3, 8[$. Quelle est la densité de X ?
3. Soit X une variable aléatoire de fonction de répartition F_X. On pose $Y = aX + b$ où $a > 0$ et $b \in \mathbb{R}$. Exprimer la fonction de répartition de Y en fonction de F_X, a et b.
4. Soit X une variable aléatoire de loi uniforme sur $]0, 1[$. Calculer la densité de $Y = X^2$.
5. Etant donné $\alpha \in]0, 1[$ on désigne par q_α le quantile de la loi $\mathcal{N}(0, 1)$. Soit Y une variable aléatoire de loi normale d'espérance 2 et de variance 9. Exprimer le quantile d'ordre α de Y en fonction de q_α. Justifier votre réponse.
6. Soit $X = (X_1, X_2)'$ un vecteur gaussien centré réduit. On note
$$P = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \quad \text{et} \quad u = \begin{pmatrix} 5 \\ 8 \end{pmatrix}.$$
Quelle est la loi du vecteur $Y = PX + u$?
7. Soit X une variable aléatoire de loi de Bernoulli de paramètre p $(0 < p < 1)$. On note $Y = 1 - X$.
   (a) Quelle est la loi du vecteur $(X, Y)'$?
   (b) En déduire $\mathbf{Cov}(X, Y)$.

Exercice 2 (5 points)

Soit $X = (X_1, X_2)'$ un vecteur aléatoire de loi uniforme sur le triangle
$$\mathcal{T} = \{(x_1, x_2) \in \mathbb{R}^2 \text{ tel que } 0 < x_2 < x_1 < 1\}.$$

1. Représenter le support de X sur un graphique.

2. Quelle est la densité de X ?
3. Calculer les densités marginales.
4. Les variables X_1 et X_2 sont-elles indépendantes ?
5. Calculer l'espérance X_1.
6. Calculer la médiane de X_1.

Exercice 3 (5 points)
Soit X une variable aléatoire réelle admettant comme densité

$$f(x) = c\exp(-|x|), \quad x \in \mathbb{R}.$$

1. Montrer que $c = 1/2$.
2. Calculer la fonction de répartition de X.
3. Calculer $\mathbf{E}[X]$.
4. Pour tout $n \in \mathbb{N}$, on appelle I_n l'intégrale définie par

$$I_n = \int_0^{+\infty} x^n \exp(-x)\,\mathrm{d}x.$$

   (a) Combien vaut I_0 ?
   (b) Montrer que pour tout $n \in \mathbb{N}^\star$ $I_n = nI_{n-1}$. En déduire que pour tout $n \in \mathbb{N}$, $I_n = n!$.
5. Pout tout $n \in \mathbb{N}$, calculer $\mathbf{E}[X^{2n}]$. En déduire $\mathbf{V}[X]$.
6. Pour tout $n \in \mathbb{N}$, que vaut $\mathbf{E}[X^{2n+1}]$?

Exercice 4 (5 points)
Soit X une variable aléatoire réelle de densité f_θ définie par

$$f_\theta(x) = (1-\theta)\mathbf{1}_{[-1/2,0]}(x) + (1+\theta)\mathbf{1}_{]0,1/2]}(x),$$

où θ est un paramètre réel tel que $|\theta| \neq 1$.

1. Quelles conditions doit vérifier θ pour que f_θ soit bien une densité de probabilité par rapport à la mesure de Lebesgue sur $\mathbb{R}$?
2. Calculer l'espérance de X.
3. Soient $X_1, \ldots, X_n$ n variables aléatoires indépendantes et de même loi de densité f_θ. Soient U_n et V_n les variables aléatoires définies par

$$U_n = \sum_{i=1}^n \mathbf{1}_{]-\infty,0]}(X_i) \quad \text{et} \quad V_n = \sum_{i=1}^n \mathbf{1}_{]0,+\infty[}(X_i).$$

   (a) Montrer que U_n et V_n suivent des lois binomiales dont on précisera les valeurs des paramètres.
   (b) Calculer $\mathbf{E}[U_n]$, $\mathbf{E}[V_n]$ et en déduire $\mathbf{E}[(V_n - U_n)/n]$.
   (c) Calculer $\mathbf{E}[U_nV_n]$ et $\mathbf{Cov}(U_n, V_n)$.
   (d) Montrer que

$$\mathbf{V}\left[\frac{V_n - U_n}{n}\right]$$

   tend vers 0 lorsque n tend vers l'infini.

CORRECTION

Exercice 1

1. Loi discrète : P_{X_1} loi de Bernoulli de paramètre $p = 1/3$. Loi absolument continue : P_{X_2} loi uniforme sur $]-1,1[$. Loi mixte : $1/2P_{X_1} + 1/2P_{X_2}$. On a

$$F_{X_1}(t) = \begin{cases} 0 & \text{si } t < 0 \\ 2/3 & \text{si } 0 \leq t < 1 \\ 1 & \text{si } t \geq 1 \end{cases}, \quad F_{X_2}(t) = \begin{cases} 0 & \text{si } t < -1 \\ 1/2(t+1) & \text{si } -1 \leq t < 1 \\ 1 & \text{si } t \geq 1 \end{cases}$$

$$\text{et} \quad F_{X_3}(t) = \begin{cases} 0 & \text{si } t < -1 \\ 1/4(t+1) & \text{si } -1 \leq t < 0 \\ 7/12 + t/4 & \text{si } 0 \leq t < 1 \\ 1 & \text{si } t \geq 1 \end{cases}$$

2. On note $\mathcal{A} =]-1, 1/2[\cup [3, 8[$. La densité de la loi uniforme sur $\mathcal{A}$ est donnée par

$$f(x) = \frac{1}{\lambda(\mathcal{A})}\mathbf{1}_{\mathcal{A}}(x) = \frac{1}{3/2+5}\mathbf{1}_{\mathcal{A}}(x).$$

3. On a

$$F_Y(t) = \mathbf{P}(aX + b \leq t) = \mathbf{P}\left(X \leq \frac{t-b}{a}\right) = F_X\left(X \leq \frac{t-b}{a}\right).$$

4. Si $t \leq 0$, $F_Y(t) = 0$ et si $t \geq 1$, $F_Y(t) = 1$. Soit $t \in]0, 1[$, on a

$$F_Y(t) = \mathbf{P}(X^2 \leq t) = \mathbf{P}(-\sqrt{t} \leq X \leq \sqrt{t}) = \mathbf{P}(X \leq \sqrt{t}) = \sqrt{t}.$$

On a donc

$$f_Y(t) = \frac{1}{2\sqrt{t}}\mathbf{1}_{]0,1[}(t).$$

5. Soit t_α le quantile d'ordre α de Y. On a

$$\mathbf{P}(Y \leq t_\alpha) = \alpha \Longleftrightarrow \mathbf{P}\left(\frac{Y-2}{3} \leq \frac{t_\alpha - 2}{3}\right).$$

Comme $\frac{Y-2}{3} \sim \mathcal{N}(0, 1)$, on déduit

$$q_\alpha = \frac{t_\alpha - 2}{3} \Longleftrightarrow t_\alpha = 3q_\alpha + 2.$$

6. Y s'obtient par transformation affine de X, c'est donc un vecteur gaussien. On obtient son espérance et sa matrice de variance-covariance comme suit

$$\mathbf{E}[Y] = P\mathbf{E}[X] + u = u \quad \text{et} \quad \mathbf{V}[Y] = P\mathbf{V}[X]P' = \begin{pmatrix} 5 & 1 \\ 1 & 2 \end{pmatrix}.$$

7. (a) Y vaut 0 si $X = 1$ et 1 si $X = 0$. Le vecteur $(X, Y)'$ suit donc une loi multinomiale de paramère $(1, p, 1-p)$.

(b) On a donc $\mathbf{Cov}(X,Y) = -p(1-p)$. On peut retrouver cette covariance par le calcul :

$$\mathbf{Cov}(X,Y) = \mathbf{Cov}(X, 1-X) = -\mathbf{V}[X] = -p(1-p).$$

Exercice 2

1. Le support est représenté sur la Figure C.2.

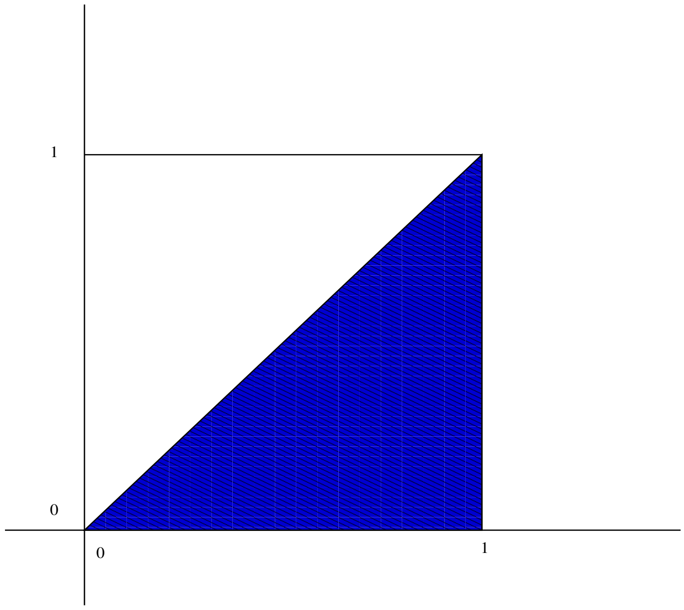


Figure C.2 – Support de X.

2. La densité de X est donnée par

$$f_X(x_1,x_2) = \frac{1}{\lambda_2(\mathcal{T})}\mathbf{1}_{\mathcal{T}}(x_1,x_2) = 2\mathbf{1}_{\mathcal{T}}(x_1,x_2).$$

3. X_1 et X_2 ont pour support $[0,1]$. Soit $x_1 \in]0,1[$, on a

$$f_{X_1}(x_1) = 2\int_{x_1}^{1} 1 \,\mathrm{d}x_2 = 2(1-x_1).$$

Donc $f_{X_1}(x_1) = 2(1-x_1)\mathbf{1}_{]0,1[}(x_1)$. De même $f_{X_2}(x_2) = 2x_2\mathbf{1}_{]0,1[}(x_2)$.

4. $f_X \neq f_{X_1}f_{X_2}$ donc X_1 et X_2 ne sont pas indépendantes.
5. On a

$$\mathbf{E}[X_1] = \int_0^1 x_1 2(1-x_1)\,\mathrm{d}x_1 = \frac{2}{3}.$$

6. Pour $x_1 \in]0,1[$, $F_{X_1}(x_1) = 2x_1 - x_1^2$. On résout l'équation $F_{X_1}(x_1) = 0.5$. La seule solution sur $]0,1[$ est $1-\sqrt{2}/2$, c'est la médiane.

Exercice 3

1. La fonction f est paire, on a donc

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 2c\int_0^{+\infty} \exp(-x)\,\mathrm{d}x = 2c = 1.$$

Par conséquent, $c = 1/2$.

2. Pour $x \leq 0$, on a
$$F_X(x) = \int_{-\infty}^{x} f(u)\,\mathrm{d}u = \frac{\exp(x)}{2}.$$
Pour $x > 0$,
$$F_X(x) = F_X(0) + \int_0^x \exp(-u)\,\mathrm{d}u = 1 - \frac{\exp(-x)}{2}.$$
3. La densité f est impaire et symétrique par rapport à 0, on a donc $\mathbf{E}[X] = 0$.
4. (a) $I_0 = 1$ d'après la question 1.
   (b) On obtient en intégrant par parties :
$$I_n = [-x^n \exp(-x)]_0^{+\infty} + \int_0^{\infty} n x^{n-1} \exp(-x)\,\mathrm{d}x = nI_{n-1}.$$
   On a ainsi pour tout $n \in \mathbb{N}$, $I_n = nI_{n-1} = n(n-1)I_{n-2} = \ldots = n!I_0 = n!$.
5. On déduit
$$\mathbf{E}[X^{2n}] = \int_{-\infty}^{+\infty} x^{2n} \frac{\exp(-|x|)}{2}\,\mathrm{d}x = \int_0^{+\infty} x^{2n} \exp(-x)\,\mathrm{d}x = I_{2n} = (2n)!.$$
Donc $\mathbf{V}[X] = \mathbf{E}[X^2] = 2$.
6. La fonction $x \mapsto x^{2n+1} \frac{\exp(-|x|)}{2}$ étant impaire, on déduit que pour tout $n \in \mathbb{N}$, $\mathbf{E}[X^{2n+1}] = 0$.

Exercice 4

1. $f_\theta(x)$ doit être positive et intégrer à 1. On déduit que $|\theta| < 1$.
2. On a
$$\mathbf{E}[X] = (1-\theta)\int_{-1/2}^{0} x\,\mathrm{d}x + (1+\theta)\int_0^{1/2} x\,\mathrm{d}x = \frac{\theta}{4}.$$
3. (a) U_n est la somme de n variables aléatoires indépendantes de loi de Bernoulli de paramètre $\mathbf{P}(X_1 \leq 0) = 1/2(1-\theta)$. On en déduit que U_n suit une loi binomiale de paramètres $(n, 1/2(1-\theta))$. De même on montre que V_n suit une loi binomiale de paramètres $(n, 1/2(1+\theta))$.
   (b) On a ainsi
$$\mathbf{E}[U_n] = \frac{n(1-\theta)}{2}, \quad \mathbf{E}[V_n] = \frac{n(1+\theta)}{2} \quad \text{et} \quad \mathbf{E}\left[\frac{V_n - U_n}{n}\right] = \theta.$$
   (c) On a
$$\begin{aligned}\mathbf{E}[U_n V_n] &= \sum_i \sum_j \mathbf{E}[\mathbf{1}_{]-\infty,0]}(X_i)\mathbf{1}_{]0+\infty[}(X_j)] = \sum_{i \neq j} \mathbf{E}[\mathbf{1}_{]-\infty,0]}(X_i)\mathbf{1}_{]0+\infty[}(X_j)] \\ &= (n^2 - n)\frac{1-\theta^2}{4}.\end{aligned}$$
   Par suite $\mathbf{Cov}(U_n, V_n) = -n\frac{1-\theta^2}{4}$.
   (d) On calcule la variance
$$\begin{aligned}\mathbf{V}\left[\frac{V_n - U_n}{n}\right] &= \frac{1}{n^2}\left(\mathbf{V}[U_n] + \mathbf{V}[V_n] - 2\mathbf{Cov}(U_n, V_n)\right) \\ &= \frac{1}{n^2}\left(2n\frac{1-\theta^2}{4} + n\frac{1-\theta^2}{2}\right) = \frac{1-\theta^2}{n} \to 0 \quad \text{quand } n \to \infty.\end{aligned}$$

Examen de Statistique 2

19 janvier 2012, 2h30

Aucun document, pas de calculatrice

Le barème est donné à titre indicatif

Rappels :

— La densité d'une variable aléatoire de loi exponentielle de paramètre $\lambda > 0$ est donnée par

$$f_X(x) = \lambda \exp(-\lambda x) \mathbf{1}_{]0,+\infty[}(x).$$

On a de plus $\mathbf{E}[X] = \frac{1}{\lambda}$ et $\mathbf{V}[X] = \frac{1}{\lambda^2}$.

— Si elle existe, la densité d'un vecteur gaussien $X = (X_1, \dots, X_d)'$ d'espérance m_X et de matrice de variance-covariance Σ_X est donnée par

$$f_X(x) = \frac{1}{(\sqrt{2\pi})^d \sqrt{\det(\Sigma_X)}} \exp\left(-\frac{1}{2}(x - m_X)' \Sigma_X^{-1} (x - m_X)\right).$$

Exercice 1 (quelques applications "directes" de cours, 6 points)
Les réponses aux questions de cet exercice ne necéssitent pas de longs développements mathématiques (si elles sont bien abordées...). Les réponses données devront être courtes et précises.

1. Soit $X = (X_1, X_2)'$ un vecteur aléatoire suivant une loi multinomiale $\mathcal{M}(1, 1/4, 3/4)$. Calculer $\mathbf{P}(X_1 = 0)$ et $\mathbf{P}(X_2 = 1)$. Les variables X_1 et X_2 sont-elles indépendantes ?
2. Soit X_1 et X_2 deux variables aléatoires réelles indépendantes telles que X_1 suit une loi uniforme sur $]1, 4[$ et X_2 suit une loi normale de moyenne -2 et de variance 1. On considère le vecteur aléatoire $X = (X_1, X_2)'$. Donner l'espérance de X ainsi que sa matrice de variance-covariance.
3. Soit X une variable aléatoire de loi de Bernoulli de paramètre p $(0 < p < 1)$. On note $Y = 1 - X$.
   (a) Quelle est la loi du vecteur $(X, Y)'$?
   (b) En déduire $\mathbf{Cov}(X, Y)$.
4. Soit X une variable aléatoire de loi uniforme sur $[-1, 1]$.
   (a) Quelle est la densité de X ?
   (b) Rappeler la définition de la fonction caractéristique d'une variable aléatoire réelle Y.
   (c) Calculer la fonction caractéristique de X.
   (d) Soit $X_1, \dots, X_n$ n variables aléatoires réelles indépendantes et de même loi que X. Que vaut la fonction caractéristique de $S_n = X_1 + \ldots + X_n$?
5. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes et de loi de Bernoulli de paramètre $p \in]0, 1[$. On note $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

(a) Sans utiliser la loi des grands nombres, montrer que $(\bar{X}_n)_{n\geq 1}$ converge en probabilité vers p (on pourra utiliser l'inégalité de Bienaymé Chebychev).

(b) A l'aide du théorème central limite, donner une suite réelle $(a_n)_{n\geq 1}$ et une fonction φ telles que la suite

$$a_n \frac{\bar{X}_n - p}{\varphi(p)}$$

converge en loi vers une limite à préciser.

(c) Déduire des questions précédentes une suite réelle $(b_n)_{n\geq 1}$ et une suite de variables aléatoires réelles $(Y_n)_{n\geq 1}$ telles que

$$b_n \frac{\bar{X}_n - p}{Y_n} \xrightarrow{\mathcal{L}} N(0,1) \quad \text{quand } n \to \infty.$$

6. Soit $(X,Y)'$ un couple aléatoire dont la loi jointe est donnée dans le tableau C.1

X \ Y	0	1
0	3/9	2/9
1	1/9	3/9

TABLE C.1 – Loi jointe de $(X,Y)'$.

On lit par exemple $\mathbf{P}(X=0,Y=1) = 2/9$.

(a) Calculer $\mathbf{E}[Y]$.

(b) Calculer $\mathbf{E}[Y|X=0]$ et $\mathbf{E}[Y|X=1]$.

(c) En déduire la loi de $\mathbf{E}[Y|X]$ (on pourra donner son support et sa fonction de masse).

(d) Retrouver le valeur de $\mathbf{E}[Y]$ à partir de la question précédente.

Exercice 2 (8 points)
Partie 1

Soit $X = (X_1, X_2)'$ un vecteur gaussien d'espérance $\mu_X = (1,0)'$ et de matrice de variance-covariance

$$\Sigma_X = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

1. Déterminer $\mathbf{V}[X_1]$, $\mathbf{V}[X_2]$ et $\mathbf{Cov}(X_1, X_2)$.
2. Calculer le coefficient de corrélation linéaire entre X_1 et X_2
3. On pose $U = 3X_1 + 1$ et $V = X_1 - 2X_2$.
   (a) Le vecteur $(U,V)'$ est-il gaussien ?
   (b) Calculer l'espérance et la matrice de variance covariance de $(U,V)'$.
   (c) Les variables aléatoires U et V sont-elles indépendantes ?
4. Calculer la densité conditionnelle de $X_2|X_1 = x_1$ et en déduire la loi de $X_2|X_1 = x_1$.
5. En déduire l'espérance conditionnelle $\mathbf{E}[X_2|X_1]$ ainsi que la loi de $\mathbf{E}[X_2|X_1]$.

Partie 2

On se place maintenant dans le cas général où $Z = (X, Y)'$ est un vecteur gaussien d'espérance $\mu \in \mathbb{R}^2$ et de matrice de variance-covariance Σ (on suppose que X et Y admettent des moments d'ordre 2 finis et que $\mathbf{V}[X] > 0$). Le but de cette partie est de montrer que

$$\mathbf{E}[Y|X] = \mathbf{E}[Y] + \frac{\mathbf{Cov}(X,Y)}{\mathbf{V}[X]}(X - \mathbf{E}[X]). \qquad \text{(C.1)}$$

Pour alléger les notations, on désigne par $u(X)$ la variable aléatoire $\mathbf{E}[Y] + \frac{\mathbf{Cov}(X,Y)}{\mathbf{V}[X]}(X - \mathbf{E}[X])$.

1. Montrer que le couple aléatoire $(X, Y - u(X))'$ est un vecteur gaussien.
2. Montrer que les variables aléatoires X et $Y - u(X)$ sont indépendantes.
3. On désigne par $L_2(X)$ l'ensemble des variables aléatoires qui s'expriment en fonction de X et de carré intégrable :

$$L_2(X) = \{f(X) \text{ avec } f : \mathbb{R} \to \mathbb{R} \text{ borélienne telle que } \mathbf{E}[f(X)^2] < +\infty\}.$$

   (a) Rappeler la définition de l'espérance conditionnelle $\mathbf{E}[Y|X]$ pour des variables aléatoires de carré intégrable (on utilisera la définition vue en cours qui fait intervenir les projecteurs).
   (b) Soit $T \in L_2(X)$. Les variables aléatoires T et $Y - u(X)$ sont-elles indépendantes ? Justifier.
   (c) En déduire (C.1).
4. **Application :** Retrouver le résultat de la question 5 de la partie 1.

Exercice 3 (6 points)

Soit $\theta > 0$ un nombre réel fixé et $Z = (X, Y)'$ un vecteur aléatoire dont la densité f est définie sur $\mathbb{R}^2$ par

$$f(x, y) = C \exp(-\theta y)\mathbf{1}_{0<x<y}.$$

1. Montrer que $C = \theta^2$.
2. Déterminer les lois marginales de Z ainsi que $\mathbf{E}[X]$ et $\mathbf{V}[X]$. Les variables aléatoires X et Y sont-elles indépendantes ?
3. Calculer la densité conditionnelle de $Y|X = x$.
4. En déduire $\mathbf{E}[Y|X]$ puis $\mathbf{E}[Y]$.
5. Calculer $\mathbf{P}(Y > 2X)$.
6. On pose $U = X$ et $V = Y - X$.
   (a) Déterminer la loi du couple $(U, V)'$.
   (b) Les variables aléatoires U et V sont-elles indépendantes ?
7. A l'aide de la question précédente calculer $\mathbf{Cov}[X, Y]$. En déduire que $\mathbf{V}[Y] = \frac{2}{\theta^2}$.
8. Soit $(X_i, Y_i)_{i\geq 1}$ une suite de couples aléatoires indépendants et de même loi de densité f. Pour $n \geq 1$, on pose $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ et $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i$.
   (a) Montrer que, pour une suite $(a_n)_{n\geq 1}$ à préciser, $a_n(\bar{Y}_n - \bar{X}_n - 1/\theta)$ converge en loi vers une loi à déterminer.
   (b) Montrer que, pour une suite $(b_n)_{n\geq 1}$ à préciser, $\dfrac{b_n}{\bar{X}_n}(\bar{Y}_n - \bar{X}_n - 1/\theta)$ converge en loi vers une loi à déterminer.

CORRECTION

Exercice 1

1. $\mathbf{P}(X_1 = 0) = \mathbf{P}(X_2 = 1) = 3/4$. $\mathbf{Cov}(X_1, X_2) = -3/4$ donc X_1 et X_2 ne sont pas indépendantes.
2. On a
$$\mathbf{E}[X] = \begin{pmatrix} 2.5 \\ -2 \end{pmatrix} \qquad \text{et} \qquad \Sigma_X = \begin{pmatrix} 3/4 & 0 \\ 0 & 1 \end{pmatrix}.$$
3. (a) Y vaut 0 si $X = 1$ et 1 si $X = 0$. Le vecteur $(X, Y)'$ suit donc une loi multinomiale de paramètre $(1, p, 1 - p)$.

   (b) On a donc $\mathbf{Cov}(X, Y) = -p(1 - p)$. On peut retrouver cette covariance par le calcul :
$$\mathbf{Cov}(X, Y) = \mathbf{Cov}(X, 1 - X) = -\mathbf{V}[X] = -p(1 - p).$$
4. (a) La densité est donnée par $f_X(x) = \frac{1}{2}\mathbf{1}_{]-1,1[}(x)$.

   (b) La fonction caractéristique est $\varphi_Y(t) = \mathbf{E}[\exp(-itY)]$.

   (c) On déduit
$$\varphi_X(t) = \frac{1}{2}\int_{-1}^{1} \exp(itx)\,\mathrm{d}x = \frac{1}{2}\int_{-1}^{1} (\cos(tx) + i\sin(tx))\,\mathrm{d}x = \frac{\sin(t)}{t}.$$

   (d) Les $X_i, i = 1, \ldots, n$ étant indépendantes, la fonction caractéristique de la somme est donnée par le produit des fonctions caractéristiques :
$$\varphi_{S_n}(t) = \left(\frac{\sin(2t)}{2}\right)^n.$$
5. (a) Soit $\varepsilon > 0$. On a d'après Bienaymé-Chebychev
$$\mathbf{P}(|\bar{X}_n - p| > \varepsilon) \leq \frac{\mathbf{V}[\bar{X}_n]}{\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2} \to 0 \text{ lorsque } n \to \infty.$$

   (b) Il suffit d'appliquer le théorème central limite
$$U_n = \sqrt{n}\frac{\bar{X}_n - p}{\sqrt{p(1-p)}} \xrightarrow{\mathcal{L}} N(0, 1).$$

   (c) On déduit des opérations sur la convergence en probabilité que
$$V_n = \sqrt{\frac{\bar{X}_n(1 - \bar{X}_n)}{p(1-p)}} \xrightarrow{\mathbf{P}} 1.$$

   On obtient ainsi par le théorème de Slutsky
$$\frac{U_n}{V_n} = \sqrt{n}\frac{\bar{X}_n - p}{\sqrt{\bar{X}_n(1 - \bar{X}_n)}} \xrightarrow{\mathcal{L}} N(0, 1).$$

6. (a) Y suit une loi de Bernoulli de paramètre 5/9, donc $\mathbf{E}[Y] = 5/9$.
   (b) $Y|X = 0 \sim B(2/5)$ et $Y|X = 1 \sim B(3/4)$, donc $\mathbf{E}[Y|X = 0] = 2/5$ et $\mathbf{E}[Y|X = 1] = 3/4$.
   (c) La loi de $\mathbf{E}[Y|X]$ est donnée dans le tableau suivant

Support	2/5	3/4
Fonction de masse	5/9	4/9

TABLE C.2 – Support et fonction de masse de la loi de $\mathbf{E}[Y|X]$.

   (d) On a ainsi

$$\mathbf{E}[Y] = \mathbf{E}[\mathbf{E}[Y|X]] = \frac{2}{5}\frac{5}{9} + \frac{3}{4}\frac{4}{9} = \frac{5}{9}.$$

Exercice 2
Partie 1

1. Il suffit de lire dans Σ_X :

$$\mathbf{V}[X_1] = \mathbf{V}[X_2] = 1 \text{ et } \mathbf{Cov}(X_1, X_2) = 0.5.$$

2. Le coefficient de corrélation vaut

$$\rho(X_1, X_2) = \frac{\mathbf{Cov}(X_1, X_2)}{\sqrt{\mathbf{V}[X_1]\mathbf{V}[X_2]}} = 0.5.$$

3. (a) On a

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

   $(U, V)'$ est donc un vecteur gaussien.
   (b) Son espérance vaut $(4, 1)'$. On a de plus

$$\mathbf{V}[U] = 9\mathbf{V}[X_1] = 9,\ \ \mathbf{V}[V] = \mathbf{V}[X_1] + 4\mathbf{V}[X_2] - 4\mathbf{Cov}(X_1, X_2) = 3$$

   et

$$\mathbf{Cov}(U, V) = 3\mathbf{V}[X_1] - 6\mathbf{Cov}(X_1, X_2) = 0.$$

   La matrice de variance covariance de (U, V) est

$$\begin{pmatrix} 9 & 0 \\ 0 & 3 \end{pmatrix}.$$

   (c) $(U, V)'$ est un vecteur gaussien et $\mathbf{Cov}(U, V) = 0$, on conclut que U et V sont indépendantes.

4. On a
$$f_{X_1,X_2}(x_1,x_2)=\frac{1}{2\pi\sqrt{3/4}}\exp\left(-\frac{1}{3}\left(2(x_1-1)^2+2x_2^2-2(x_1-1)x_2\right)\right)$$
et
$$f_{X_1}(x_1)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(x_1-1)^2\right).$$
Par conséquent
$$f_{X_2|X_1=x_1}=\frac{1}{2\pi\sqrt{3/4}}\exp\left(-\frac{1}{2}\frac{(x_2-(x_1-1)/2)^2}{3/4}\right).$$
Ainsi $X_2|X_1=x_1\sim\mathcal{N}((x_1-1)/2,3/4)$.
5. D'après la question précédente, on a $\mathbf{E}[X_2|X_1]=(X_1-1)/2$, d'où $\mathbf{E}[X_2|X_1]\sim\mathcal{N}(0,1/4)$.

Partie 2

1. Soient a te b tels que $u(X)=aX+b$. On a
$$\begin{pmatrix}X\\Y-u(X)\end{pmatrix}=\begin{pmatrix}1&0\\-a&1\end{pmatrix}\begin{pmatrix}X\\Y\end{pmatrix}+\begin{pmatrix}0\\-b\end{pmatrix},$$
le vecteur $(X,Y-u(X))'$ est donc gaussien comme transformée affine du vecteur gaussien $(X,Y)'$.
2. $(X,Y-u(X))'$ étant gaussien, il suffit de montrer que $\mathbf{Cov}(X,Y-u(X))=0$:
$$\begin{aligned}\mathbf{Cov}(X,Y-u(X))&=\mathbf{Cov}(X,Y)-\mathbf{Cov}(X,aX+b)\\&=\mathbf{Cov}(X,Y)-a\mathbf{V}[X]=\mathbf{Cov}(X,Y)-\mathbf{Cov}(X,Y)=0.\end{aligned}$$
3. (a) On munit $L_2(\Omega)$ du produit scalaire $\langle X,Y\rangle=\mathbf{E}[XY]$. Si $Y\in L_2(\Omega)$, l'espérance conditionnelle de Y sachant X est le projeté orthogonal de Y sur $L_2(X)$.
   (b) Ecrivons $T=f(X)$. Comme $Y-u(X)$ et X sont indépendantes, on déduit que $Y-u(X)$ et $f(X)$ sont indépendantes.
   (c) If faut montrer que
   — $u(X)\in L_2(\Omega)$;
   — $(Y-u(X))\perp f(X)\ \forall f(X)\in L_2(X)$.
   $u(X)$ étant une transformation affine d'une variable gaussienne, elle est dans $L_2(\Omega)$. Soit $f(X)\in L_2(X)$. On a
$$\langle f(X),Y-u(X)\rangle=\mathbf{E}[f(X)(Y-u(X))]=\mathbf{E}[f(X)]\mathbf{E}[Y-u(X)]=0,$$
   l'avant dernière inégalité est une conséquence de l'indepéndance entre $f(X)$ et $Y-u(X)$, la dernière s'obtient en remarquant que $Y-u(X)$ est centrée.
4. On a
$$\mathbf{E}[X_2|X_1]=\mathbf{E}[X_2]+\frac{\mathbf{Cov}(X_1,X_2)}{\mathbf{V}[X_2]}(X_1-\mathbf{E}[X_1])=0+\frac{0.5}{1}(X_1-1)=\frac{X_1-1}{2}.$$
On retrouve bien le résultat de la partie précédente.

Exercice 3

1. Il suffit de calculer l'intégrale double

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \exp(-\theta y)\mathbf{1}_{0<x<y}\,\mathrm{d}x\,\mathrm{d}y = \int_0^{+\infty}\int_0^y \exp(-\theta y)\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{\theta^2}.$$

2. On obtient

$$f_X(x) = \theta^2\int_x^{+\infty} \exp(-\theta y)\,\mathrm{d}y = \theta\exp(-\theta x),\ \text{ pour } x>0,$$

X suit donc une loi exponentielle de paramètre θ, d'où $\mathbf{E}[X] = 1/\theta$ et $\mathbf{V}[X] = 1/\theta^2$. La seconde marginale vaut

$$f_Y(y) = \theta^2\int_0^y \exp(-\theta y)\,\mathrm{d}y = \theta^2 y\exp(-\theta y),\ \text{ pour } y>0.$$

Ainsi $f(x,y) \neq f_x(x)f_Y(y)$, X et Y ne sont donc pas indépendantes.

3. Soit $x>0$, on a par définition

$$f_{Y|X=x} = \frac{f(x,y)}{f_X(x)} = \theta\exp(-\theta(y-x))\mathbf{1}_{0<x<y}.$$

4. Calculons $\mathbf{E}[Y|X=x]$ pour $x>0$:

$$\mathbf{E}[Y|X=x] = \int_x^{+\infty} y\theta\exp(-\theta(y-x))\,\mathrm{d}y = x+\frac{1}{\theta}.$$

On a donc $\mathbf{E}[Y|X] = X + 1/\theta$ et

$$\mathbf{E}[Y] = \mathbf{E}[\mathbf{E}[Y|X]] = \mathbf{E}[X] + \frac{1}{\theta} = \frac{2}{\theta}.$$

5. On calcule d'abord la probabilité conditionnelle $\mathbf{P}(Y>2X|X=x)$:

$$\mathbf{P}(Y>2X|X=x) = \int_{\mathbb{R}} \mathbf{1}_{y>2x} f_{Y|X=x}\,\mathrm{d}y = \theta\exp(\theta x)\int_{2x}^{+\infty}\exp(-\theta y)\,\mathrm{d}y = \exp(-\theta x).$$

Ainsi

$$\mathbf{P}(Y>2X) = \int_{\mathbb{R}} \mathbf{P}(Y>2X|X=x)f_X(x)\,\mathrm{d}x = \theta\int_0^{+\infty}\exp(-2\theta x)\,\mathrm{d}x = \frac{1}{2}.$$

6. (a) On considère le changement de variable

$$\begin{aligned}\varphi : D = \{(x,y)\in\mathbb{R}^2 : 0<x<y\} &\to \Delta = \{(u,v)\in\mathbb{R}^2 : u>0, v>0\}\\ (x,y) &\mapsto (x, y-x).\end{aligned}$$

φ est un C^1 difféomorphisme, d'après le théorème de changement de variable, la densité de $(U,V)'$ est donnée par

$$f_{U,V}(u,v) = f(\varphi^{-1}(u,v))|\det(J_{\varphi^{-1}_{u,v}})|\mathbf{1}_{\Delta}(u,v).$$

Or

$$\varphi^{-1} : \Delta \to D$$
$$(u,v) \mapsto (u, u+v),$$

d'où

$$\det(J_{\varphi^{-1}_{u,v}}) = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1.$$

On obtient

$$f_{U,V}(u,v) = f(u,u+v) = \theta^2 \exp(-\theta(u+v))\mathbf{1}_\Delta(u,v) = \theta \exp(-\theta u)\mathbf{1}_{u>0}\,\theta \exp(-\theta v)\mathbf{1}_{v>0}.$$

(b) D'après la question précédente, U et V sont indépendantes. Elles suivent de plus une loi exponentielle de paramètre θ.

7. Comme
$$\mathbf{Cov}(U,V) = \mathbf{Cov}(X, Y-X) = \mathbf{Cov}(X,Y) - \mathbf{V}[X] = 0,$$
on déduit $\mathbf{Cov}(X,Y) = \mathbf{V}[X] = 1/\theta^2$. De même,
$$\mathbf{V}[V] = \mathbf{V}[Y-X] = \mathbf{V}[Y] + \mathbf{V}[X] - 2\mathbf{Cov}(X,Y) = \frac{1}{\theta^2} \implies \mathbf{V}[Y] = \frac{2}{\theta^2}.$$

(a) On pose $V_i = Y_i - X_i$. Les variables $V_i, i = 1, \ldots, n$ sont indépendantes, de même loi d'espérance $1/\theta$ et de variance $1/\theta^2$. D'après le théorème central limite, on a donc
$$\sqrt{n}\left(\bar{V}_n - \frac{1}{\theta}\right) = \sqrt{n}\left(\bar{Y}_n - \bar{X}_n - \frac{1}{\theta}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{\theta^2}\right).$$

(b) D'après la loi des grands nombres $\bar{X}_n \xrightarrow{\mathbf{P}} 1/\theta$. Soit T une variable aléatoire de loi $\mathcal{N}\left(0, \frac{1}{\theta^2}\right)$, d'après le théorème de Slutsky, on a
$$\frac{\sqrt{n}}{\bar{X}_n}\left(\bar{Y}_n - \bar{X}_n - \frac{1}{\theta}\right) \xrightarrow{\mathcal{L}} \theta T,$$
c'est-à-dire
$$\frac{\sqrt{n}}{\bar{X}_n}\left(\bar{Y}_n - \bar{X}_n - \frac{1}{\theta}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Examen partiel Statistique 2

12 décembre 2012
Aucun document, pas de calculatrice

Le sujet est composé de 4 exercices indépendants (il y a 2 pages !).

Le barème est donné à titre indicatif.

Rappel : si elle existe, la densité d'un vecteur gaussien $X = (X_1, \ldots, X_d)'$ d'espérance m_X et de matrice de variance-covariance Σ_X est donnée par

$$f_X(x) = \frac{1}{(\sqrt{2\pi})^d \sqrt{\det(\Sigma_X)}} \exp\left(-\frac{1}{2}(x - m_X)'\Sigma_X^{-1}(x - m_X)\right).$$

Exercice 1 (Questions de cours, 6.5 points)

1. Donner la définition d'une variable aléatoire réelle ainsi que de sa loi de probabilité.
2. Soit X une variable aléatoire de loi uniforme sur $[-1, 1]$.
   (a) Quelle est la densité de X ?
   (b) Rappeler la définition de la fonction caractéristique d'une variable aléatoire réelle.
   (c) Calculer la fonction caractéristique de X.
   (d) Soit $X_1, \ldots, X_n$ n variables aléatoires réelles indépendantes et de même loi que X. Que vaut la fonction caractéristique de $S_n = X_1 + \ldots + X_n$?
3. Soit X une variable aléatoire de loi de Bernoulli de paramètre p $(0 < p < 1)$. On note $Y = 1 - X$.
   (a) Quelle est la loi de Y ?
   (b) Quelle est la loi du vecteur $(X, Y)'$?
   (c) En déduire $\mathbf{Cov}(X, Y)$.
4. Soit $X = (X_1, X_2)$ un vecteur gaussien centré réduit. On note

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

   (a) Donner l'espérance et la matrice de variance-covariance de X.
   (b) Calculer $\mathbf{P}((X_1, X_2) = (0, 0))$.
   (c) Les variables aléatoires X_1 et X_2 sont-elles indépendantes ?
   (d) Quelle est la loi du vecteur $Y = AX + b$?

5. (a) Rappeler la définition de la loi du χ^2 à n degrés de liberté.
   (b) Soit X une variable aléatoire réelle de loi $\chi^2(n)$ et Y une variable aléatoire réelle de loi $\chi^2(m)$. X et Y sont indépendantes. Quelle est la loi de $X+Y$? Justifier votre réponse.

Exercice 2 (3.5 points)
Soit $\theta > 0$ et soit X une variable aléatoire réelle de densité f_θ définie par

$$f_\theta(x) = 2\theta x \exp(-\theta x^2)\mathbf{1}_{[0,+\infty[}(x).$$

1. Montrer que f_θ est bien une densité de probabilité.
2. Déterminer la fonction de répartition de X.
3. Soit Y la variable aléatoire définie par $Y = \theta X^2$.
   (a) Calculer la fonction de répartition de Y.
   (b) En déduire la densité de Y ainsi que $\mathbf{E}[X^2]$.

Exercice 3 (5 points)
Soient X et Y deux variables aléatoires réelles indépendantes et de même loi de densité f définie par

$$f(z) = \exp(-z)\mathbf{1}_{[0,+\infty[}(z).$$

1. Quelle est la densité du vecteur aléatoire (X, Y) ?
2. Calculer $\mathbf{P}((X, Y) \in [0, 1]^2)$.
3. On pose $U = X + Y$ et $V = X/(X + Y)$.
   (a) Déterminer la loi du vecteur (U, V).
   (b) En déduire les lois des variables aléatoires U et V.
   (c) Les variables aléatoires U et V sont-elles indépendantes ?

Exercice 4 (5 points)
Soit (X, Y) un couple aléatoire de densité $f_{X,Y}$ définie par :

$$f_{X,Y}(x, y) = \frac{1}{4\pi} \exp\left(-\frac{1}{2}\left(\frac{x^2}{2} - xy + y^2\right)\right).$$

1. Montrer que (X, Y) est un vecteur gaussien, puis déterminer son espérance et sa matrice de variance covariance.
2. Déterminer les lois des marginales X et Y.
3. Soit $a \in \mathbb{R}$ et soit $U = Y - aX$. Quelle est la loi du vecteur (U, X) ?
4. En déduire que U et X sont des variables aléatoires indépendantes si et seulement si $a = 1/2$.

CORRECTION

Exercice 1

1. Une variable aléatoire est une fonction X définie sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$ à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que

$$\forall B \in \mathcal{B}(\mathbb{R}),\ X^{-1}(B) \in \mathcal{A}.$$

Sa loi de probabilité est l'application $\mathbf{P}_X$ définie par

$$\begin{aligned} \mathbf{P}_X : (\mathbb{R}, \mathcal{B}(\mathbb{R})) &\to [0,1] \\ B &\mapsto \mathbf{P}_X(B) = \mathbf{P}(X^{-1}(B)). \end{aligned}$$

2. (a) La densité est donnée par $f_X(x) = \frac{1}{2}\mathbf{1}_{]-1,1[}(x)$.

   (b) La fonction caractéristique est $\varphi_Y(t) = \mathbf{E}[\exp(-itY)]$.

   (c) On déduit

$$\varphi_X(t) = \frac{1}{2}\int_{-1}^{1} \exp(itx)\,\mathrm{d}x = \frac{1}{2}\int_{-1}^{1} (\cos(tx) + i\sin(tx))\,\mathrm{d}x = \frac{\sin(t)}{t}.$$

   (d) Comme $X_1, \dots, X_n$ sont indépendantes, on a $\varphi_{S_n}(t) = (\varphi_X(t))^n$.

3. (a) Y vaut 0 si $X = 1$ et 1 si $X = 0$. Y suit donc une loi de Bernoulli de paramètre $1 - p$.

   (b) Le vecteur $(X, Y)'$ suit une loi multinomiale de paramètre $(1, p, 1-p)$.

   (c) On a donc $\mathbf{Cov}(X, Y) = -p(1-p)$. On peut retrouver cette covariance par le calcul :

$$\mathbf{Cov}(X, Y) = \mathbf{Cov}(X, 1 - X) = -\mathbf{V}[X] = -p(1-p).$$

4. (a) On a $\mathbf{E}[X] = (0,0)'$ et $\Sigma_X = Id$.

   (b) Σ_X est inversible donc le vecteur X est absolument continu. Cette probabilité est donc nulle.

   (c) $\mathbf{Cov}(X_1, X_2) = 0$. X étant un vecteur gaussien, on déduit que X_1 et X_2 sont indépendantes.

   (d) Y est la transformée affine d'un vecteur gaussien, c'est donc un vecteur gaussien. Il suffit de déterminer son espérance et sa matrice de variance covariance. On a $\mathbf{E}[Y] = b$ et

$$\Sigma_Y = A\Sigma_X A' = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

5. (a) $X \sim \chi^2(n)$ si il existe n variables aléatoires $X_1, \dots, X_n$ indépendantes de loi normale centrée réduite telles que $X = X_1^2 + \dots + X_n^2$.

   (b) On a

$$X + Y = X_1^2 + \dots + X_n^2 + Y_1^2 + \dots + Y_m^2$$

   où les $X_i, i = 1, \dots, n$ et $Y_j, j = 1, \dots, m$ suivent des lois normales centrées réduites. Comme X et Y sont indépendantes, les variables $X_i, i = 1, \dots, n$ et $Y_j, j = 1, \dots, m$ sont indépendantes. On déduit que $X + Y$ suit une loi $\chi^2(n + m)$.

Exercice 2

1. f_θ est positive sur $[0,+\infty[$ et
$$\int_{\mathbb{R}} f_\theta(x)\,\mathrm{d}\lambda(x) = \int_0^\infty 2\theta x \exp(-\theta x^2)\,\mathrm{d}x = [-\exp(-\theta x^2)]_0^\infty = 1.$$
2. On a $F_X(t)=0$ si $t<0$ et $F_X(t)=1-\exp(-\theta t^2)$ si $t\geq 0$.
3. (a) On a pour $t \geq 0$
$$F_Y(t) = \mathbf{P}(X^2 \leq t/\theta) = \mathbf{P}(X \leq \sqrt{t/\theta}) = 1-\exp(-t).$$
   (b) On déduit $f_Y(t) = \exp(-t)\mathbf{1}_{[0,+\infty[}(t)$ et $\mathbf{E}[X^2] = 1/\theta$.

Exercice 3

1. La densité de (X,Y) est
$$f_{X,Y}(x,y) = \exp(-(x+y))\mathbf{1}_{[0,+\infty[}(x)\mathbf{1}_{[0,+\infty[}(y).$$
2. On a
$$\mathbf{P}((X,Y)\in[0,1]^2) = \int_0^1\int_0^1 \exp(-(x+y))\,\mathrm{d}x\,\mathrm{d}y = (1-\exp(-1))^2.$$
3. (a) Il suffit de considérer le changement de variable
$$\begin{aligned}\varphi : \mathbb{R}^{+\star}\times\mathbb{R}^{+\star} &\to \mathbb{R}^{+\star} + \times]0,1[\\ (x,y) &\mapsto (x+y, x/(x+y)).\end{aligned}$$
   La densité de $(U,V)=\varphi(X,Y)$ est alors donnée par
$$g(u,v) = f(\varphi^{-1}(u,v))|\det J_{\varphi^{-1}}|.$$
   Ici
$$\begin{aligned}\varphi^{-1} : \mathbb{R}^{+\star}\times]0,1[&\to \mathbb{R}^{+\star}\times\mathbb{R}^{+\star}\\ (u,v) &\mapsto (uv, u(1-v)),\end{aligned}$$
   on a donc $|\det J_{\varphi^{-1}}| = u$ et
$$g(u,v) = u\exp(-u)\mathbf{1}_{]0,+\infty[}(u)\mathbf{1}_{]0,1[}(v).$$
   (b) On déduit que V suit une loi uniforme sur $]0,1[$ et
$$g_U(u) = u\exp(-u)\mathbf{1}_{]0,+\infty[}(u).$$
   (c) Les variables aléatoires U et V sont clairement indépendantes.

Exercice 4

1. On a
$$f_{X,Y}(x,y) = \frac{1}{(\sqrt{2\pi})^2\sqrt{\det(\Sigma)}}\exp\left(-\frac{1}{2}(x-m)'\Sigma^{-1}(x-m)\right)$$
avec $m=(0,0)'$ et
$$\Sigma^{-1} = \frac{1}{2}\begin{pmatrix}1 & -1\\ -1 & 2\end{pmatrix}.$$
On déduit que (X,Y) est un vecteur gaussien centré et de matrice de variance covariance
$$\Sigma = \begin{pmatrix}4 & 2\\ 2 & 2\end{pmatrix}.$$

2. On déduit que $X \sim \mathcal{N}(0,4)$ et $Y \sim \mathcal{N}(0,2)$.
3. On a

$$\begin{pmatrix} U \\ X \end{pmatrix} = \begin{pmatrix} -a & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

   Donc (U, X) est un vecteur gaussien centré et de matrice de variance covariance

$$\begin{pmatrix} 4a^2+2 & -4a+2 \\ -4a+2 & 4 \end{pmatrix}.$$

Par conséquent, X et U sont indépendantes si et seulement si $a = 1/2$.

Examen de Statistique 2, janvier 2013
deux heures, sans document, pas de calculatrice

Le barème est donné à titre indicatif

Rappels :

— si elle existe, la densité d'un vecteur gaussien $X = (X_1, \dots, X_d)$ d'espérance m_X et de matrice de variance covariance Σ_X est donnée par

$$f_X(x) = \frac{1}{(\sqrt{2\pi})^d \sqrt{\det(\Sigma_X)}} \exp\left(-\frac{1}{2}(x - m_X)'\Sigma_X^{-1}(x - m_X)\right) ;$$

— la densité de la loi exponentielle de paramètre $\lambda > 0$ est $f(x) = \lambda \exp(-\lambda x)\mathbf{1}_{]0,+\infty[}(x)$.

Exercice 1 (6 points)

Les réponses aux questions de cet exercice ne necéssitent pas de longs développements mathématiques (si elles sont bien abordées...). Les réponses données devront être courtes et précises.

1. Enoncer la loi forte des grands nombres.
2. Enoncer le théorème central limite.
3. Donner la définition de l'espérance conditionnelle $\mathbf{E}[Y|X]$ pour des variables aléatoires de carré intégrable (on utilisera la définition vue en cours qui fait intervenir les projecteurs).
4. Soit X une variable aléatoire réelle.
   (a) Rappeler la définition de la fonction caractéristique de X. On la note φ_X.
   (b) Soit a et b deux réels. Montrer que la fonction caractéristique de la variable aléatoire $aX + b$ est donnée par $\exp(ibt)\varphi_X(at)$.
   (c) On rappelle que la fonction caractéristique de la loi normale centrée réduite est donnée par $\exp(-t^2/2)$. Déduire de la question précédente la fonction caractéristique la loi normale d'espérance μ et de variance σ^2.
5. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires, indépendantes, à valeurs dans $\{0, 2\}$ et de même loi définie par
$$\mathbf{P}(X_n = 0) = \mathbf{P}(X_n = 2) = \frac{1}{2}.$$
   On pose $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.
   (a) Rappeler la définition de la convergence en probabilité.
   (b) Montrer à l'aide de l'inégalité de Bienaymé-Tchebychev que la suite $(\bar{X}_n)_{n\geq 1}$ converge en probabilité vers une constante a que l'on précisera.
   (c) A l'aide du théorème central limite, donner une suite réelle $(a_n)_{n\geq 1}$ et un réel θ tels que la suite
$$a_n(\bar{X}_n - \theta)$$
   converge en loi vers une limite à préciser.

6. Soit $Z = (X, Y)$ un vecteur aléatoire de loi uniforme sur le triangle $\{(x, y) \in \mathbb{R}^2 : 0 < x < y < 1\}$.
   (a) Donner la densité de Z et représenter son support.
   (b) Calculer la probabilité $\mathbf{P}((X, Y) \in [1/2, 1] \times [1/2, 1])$.
   (c) Calculer les lois marginales de Z ainsi que $\mathbf{E}[X]$.
   (d) Calculer l'espérance conditionnelle $\mathbf{E}[Y|X]$.
   (e) Déduire des deux questions précédentes $\mathbf{E}[Y]$.

Exercice 2 (3 points)
On considère une variable aléatoire X de densité

$$f(x) = \begin{cases} cx^2 & \text{si } 0 \leq x \leq 3 \\ 0 & \text{sinon.} \end{cases}$$

1. Evaluer la constante c pour que f soit une densité de probabilité.
2. Déterminer la fonction de répartition de X.
3. Calculer $\mathbf{P}(1 < X < 2)$.
4. Déterminer l'espérance et la variance de X.
5. Pour tout $n \in \mathbb{N}^\star$, déterminer le moment d ordre n de X.

Exercice 3 (4 points)
On rappelle que si U suit une loi exponentielle de paramètre $\lambda > 0$ alors $\mathbf{E}[U^n] = \frac{n!}{\lambda^n}$.

Soit (X, Y) un couple de densité jointe

$$f_{X,Y}(x, y) = cx(y - x)\exp(-y)\mathbf{1}_{\{0<x\leq y\}}.$$

1. Montrer que $c = 1$.
2. Calculer la densité conditionnelle de X sachant $Y = y$.
3. En déduire $\mathbf{E}[X|Y]$.
4. Montrer que X admet pour densité

$$f_X(x) = x\exp(-x)\mathbf{1}_{\{x>0\}}.$$

5. Calculer la densité conditionnelle de Y sachant $X = x$.
6. En déduire $\mathbf{E}[Y|X]$.
7. Déduire des questions 3 et 6 les quantités $\mathbf{E}[X]$ et $\mathbf{E}[Y]$.

Exercice 4 (3 points)
Soit $(X_n)_{n\geq 1}$ une suite de v.a.r. indépendantes, de même loi d'espérance μ et de variance $\sigma^2 < +\infty$. Pour tout entier $n \geq 2$, on note

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad S_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2.$$

1. Calculer $\mathbf{E}[S_n]$.

2. Montrer que S_n converge presque surement vers une constante que l'on précisera.
3. En déduire que

$$T_n = a_n \frac{(\bar{X}_n - \mu)}{\sqrt{S_n}}$$

converge en loi vers une limite que l'on précisera (on précisera également la suite $(a_n)_n$).

Exercice 5 (4 points)
Soit (X, Y) un couple aléatoire sur $\mathbb{R}^2$ dont la densité est donnée par

$$f_{X,Y}(x, y) = \frac{1}{4\pi} \exp\left(-\frac{1}{2}\left(\frac{1}{2}x^2 - xy + y^2\right)\right), \quad (x, y) \in \mathbb{R}^2.$$

1. Quelle est la loi du vecteur (X, Y). Donner les lois marginales de ce vecteur.
2. Soit Z une variable aléatoire réelle définie par $Z = Y - aX$, où $a \in \mathbb{R}$. Pour quelle(s) valeur(s) de a les variables aléatoires Z et X sont-elles indépendantes ?
3. On considère maintenant $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes de loi normale centrée réduite. On pose

$$S_n = \sum_{i=1}^{n} X_i \quad \text{et} \quad T_m = S_m - \frac{m}{n} S_n,$$

avec $n \geq 1$ et $m \in \{1, \ldots, n-1\}$.

(a) Déterminer la loi du vecteur $(S_1, \ldots, S_n)$.

(b) Montrer que les variables aléatoires T_m et S_n sont indépendantes.

CORRECTION

Exercice 1

1. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. de $L_1(\Omega, \mathcal{A}, \mathbf{P})$ indépendantes et de même loi. On note $\mathbf{E}[X_1] = \mu$. On a
$$\frac{1}{n}\sum_{i=1}^{n} X_i \overset{p.s.}{\to} \mu \quad \text{quand } n \to \infty.$$
2. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes, de même loi, et telles que $\mathbf{E}[X_i^2] < +\infty$. On note $\mathbf{E}[X_i] = \mu$, $\mathbf{V}[X_i] = \sigma^2$ et $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. On a alors
$$\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \overset{\mathcal{L}}{\to} \mathcal{N}(0,1) \quad \text{quand } n \to \infty.$$
3. On considère l'espace $L_2(\Omega)$ des fonctions de carré intégrable muni du produit scalaire $< X_1, X_2 >= \mathbf{E}[X_1X_2]$. Soit (X, Y) un couple aléatoire avec $Y \in L_2(\Omega)$. L'espérance conditionnelle de Y sachant X, notée $\mathbf{E}[Y|X]$, est la projection orthogonale de Y sur le sous-espace
$$L_2(X) = \{u(X) \text{ avec } u : \mathbb{R} \to \mathbb{R} \text{ borélienne telle que } \mathbf{E}[u^2(X)] < +\infty\}.$$
4. (a) La fonction caractéristique est définie par $\varphi_X(t) = \mathbf{E}[\exp(itX)]$.

   (b) On a
$$\varphi_{aX+b} = \mathbf{E}[\exp(it(aX + b))] = \exp(itb)\mathbf{E}[\exp(itaX)] = \exp(itb)\varphi_X(at).$$
   (c) Soit Y de loi $\mathcal{N}(\mu, \sigma^2)$. Il existe X de loi normale centrée réduite telle que $Y = \mu + \sigma X$. Il vient
$$\varphi_Y(t) = \exp(i\mu t)\varphi_X(\sigma t) = \exp(i\mu t)\exp(-\sigma^2 t^2/2).$$
5. (a) Une suite $(X_n)_n$ converge en probabilité vers X si $\forall\varepsilon > 0$ on a
$$\lim_{n\to+\infty} \mathbf{P}(|X_n - X| > \varepsilon) = 0.$$
   (b) On a $\mathbf{E}[\bar{X}_n] = 1$ et $\mathbf{V}[\bar{X}_n] = 1/n$. Donc d'après Bienaymé-Tchébychev on a pour tout $\varepsilon > 0$
$$\mathbf{P}(|\bar{X}_n - 1| > \varepsilon) \leq \frac{\mathbf{V}[\bar{X}_n]}{\varepsilon^2} = \frac{1}{n\varepsilon^2}.$$
   On déduit que la suite $(\bar{X}_n)_n$ converge en probabilité vers 1.

   (c) Il suffit d'appliquer le TCL :
$$\sqrt{n}(\bar{X}_n - 1) \overset{\mathcal{L}}{\to} \mathcal{N}(0,1).$$
6. (a) La densité de Z est donnée par
$$f_Z(x,y) = 2\mathbf{1}_{\mathcal{T}}(x,y),$$
   où $\mathcal{T}$ désigne le support de Z.

(b) La probabilité cherchée vaut 1/4. En effet, il suffit de calculer

$$\int_{1/2}^{1}\int_{x}^{1} 2\,\mathrm{d}x\,\mathrm{d}y = 1/4.$$

(c) Les densités marginales sont données par

$$f_X(x) = 2(1-x)\mathbf{1}_{]0,1[}(x) \quad \text{et} \quad f_Y(y) = 2y\mathbf{1}_{]0,1[}(x).$$

On déduit $\mathbf{E}[X] = 1/3$.

(d) On calcule la densité conditionnelle

$$f_{Y|X=x}(y) = \frac{1}{1-x}\mathbf{1}_{\mathcal{T}}(x,y), \quad \text{pour } 0 < x < 1.$$

Il vient

$$\mathbf{E}[Y|X=x] = \int y f_{Y|X=x}(y)\,\mathrm{d}y = \frac{1+x}{2}, \quad \text{d'où } \mathbf{E}[Y|X] = \frac{1+X}{2}.$$

(e) Il suffit d'utiliser

$$\mathbf{E}[Y] = \mathbf{E}[\mathbf{E}[Y|X]] = \mathbf{E}\left[\frac{1+X}{2}\right] = \frac{2}{3}.$$

Exercice 2

1. Il suffit de calculer

$$\int_0^3 x^2\,\mathrm{d}x = 9.$$

On a donc $c = 1/9$.

2. La fonction de répartition est donnée par

$$F_X(t) = \mathbf{P}(X \leq t) = \begin{cases} 0 & \text{si } t \leq 0 \\ t^3/27 & \text{si } 0 < t \leq 3 \\ 1 & \text{si } t > 3. \end{cases}$$

3. Comme X est aboslument continue, on a

$$\mathbf{P}(1 < X < 2) = F_X(2) - F_X(1) = \frac{7}{27}.$$

4. On a

$$\mathbf{E}[X] = \int_0^3 x^3/9\,\mathrm{d}x = \frac{9}{4} \quad \text{et} \quad \mathbf{E}[X^2] = \int_0^3 x^4/9\,\mathrm{d}x = \frac{27}{5},$$

donc $\mathbf{V}[X] = 27/80$.

5. Le moment d'ordre n est donné par

$$\mathbf{E}[X^n] = \int_0^3 \frac{x^{n+2}}{9}\,\mathrm{d}x = \frac{3^{n+1}}{n+3}.$$

Exercice 3

1. On a

$$\int\int_{\mathbb{R}^2} f_{X,Y}(x,y) = \int_0^{+\infty}\left(\int_0^y cx(y-x)\exp(-y)\,\mathrm{d}x\right)\mathrm{d}y = c\int_0^{+\infty}\left[y\frac{x^2}{2}-\frac{x^3}{3}\right]_0^y \exp(-y)\,\mathrm{d}y$$
$$=\frac{c}{6}\int_0^{+\infty} y^3\exp(-y)\,\mathrm{d}y = \frac{c}{6}\mathbf{E}[U^3] = c.$$

2. Il faut au préalable calculer la densité marginale f_Y :

$$f_Y(y) = \exp(-y)\mathbf{1}_{\{y>0\}}\int_0^y x(y-x)\,\mathrm{d}x = \frac{y^3}{6}\exp(-y)\mathbf{1}_{\{y>0\}}.$$

On déduit

$$f_{X|Y=y}(x)\frac{6x(y-x)}{y^3}\mathbf{1}_{\{0<x<y\}}.$$

3. On a donc pour $y>0$:

$$\mathbf{E}[X|Y=y] = \frac{6}{y^3}\int_0^y x^2(y-x)\,\mathrm{d}x = \frac{y}{2},$$

d'où $\mathbf{E}[Y|X] = Y/2$.

4. On calcule la densité de X :

$$f_X(x) = x\mathbf{1}_{\{x>0\}}\int_x^{+\infty}(y-x)\exp(-y)\,\mathrm{d}y = x\mathbf{1}_{\{x>0\}}\int_0^{+\infty} u\exp(-u-x)\,\mathrm{d}u$$
$$= x\exp(-x)\mathbf{1}_{\{x>0\}}\mathbf{E}[U] = x\exp(-x)\mathbf{1}_{\{x>0\}}.$$

5. On a donc

$$f_{Y|X=x}(y) = (y-x)\exp(-(y-x))\mathbf{1}\{0<x<y\}.$$

6. Ainsi, pour $x>0$, on a

$$\mathbf{E}[Y|X=x] = \int_x^{+\infty} y(y-x)\exp(-(y-x))\,\mathrm{d}y = \int_0^{+\infty}(u+x)u\exp(-u)\,\mathrm{d}u = \mathbf{E}[U^2]+x\mathbf{E}[U] = 2+x,$$

d'où $\mathbf{E}[Y|X] = 2+X$.

7. On déduit des questions précédentes

$$\begin{cases}\mathbf{E}[X] = \mathbf{E}[Y]/2\\ \mathbf{E}[Y] = \mathbf{E}[X]+2\end{cases} \iff \begin{cases}\mathbf{E}[X] = 2\\ \mathbf{E}[Y] = 4.\end{cases}$$

Exercice 4

1. On a

$$S_n = \frac{1}{n-1}\sum_{i=1}^n (X_i-\bar{X}_n)^2 = \frac{1}{n-1}\sum_{i=1}^n X_i^2 - \frac{2}{n-1}\sum_{i=1}^n X_i\bar{X}_n + \frac{n}{n-1}\bar{X}_n^2$$
$$= \frac{1}{n-1}\sum_{i=1}^n X_i^2 - \frac{n}{n-1}\bar{X}_n^2,$$

d'où

$$\mathbf{E}[S_n] = \frac{n}{n-1}(\sigma^2+m^2) - \frac{n}{n-1}(\mathbf{V}(\bar{X}_n)+\mathbf{E}[\bar{X}_n]^2) = \frac{n}{n-1}(\sigma^2+m^2-\sigma^2/n-m^2) = \sigma^2.$$

2. On a
$$S_n = \frac{1}{n-1}\sum_{i=1}^{n} X_i^2 - \frac{n}{n-1}\bar{X}_n^2,$$
et la loi forte des grands nombres donne :
$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \stackrel{p.s.}{\to} \sigma^2 + m^2 \quad \text{et} \quad \bar{X}_n^2 \stackrel{p.s.}{\to} m^2.$$
En remarquant que
$$S_n = \frac{n}{n-1}\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}_n^2\right]$$
et en appliquant les propriétés des opérations classique pour la convergence presque sûre, en conclut que $S_n \stackrel{p.s.}{\to} \sigma^2$.

3. Par le théorème central limite, on a
$$\frac{\sqrt{n}(\bar{X}_n - m)}{\sigma} \stackrel{\mathcal{L}}{\to} \mathcal{N}(0,1)$$
et $S_n/\sigma^2 \stackrel{p.s.}{\to} 1$. Le lemme de Slutsky donne
$$\sqrt{n}\frac{(\bar{X}_n - m)}{\sqrt{S_n}} \stackrel{\mathcal{L}}{\to} \mathcal{N}(0,1).$$

Exercice 5

1. On a
$$f_{X,Y}(x,y) = \frac{1}{(\sqrt{2\pi})^2\sqrt{\det(\Sigma)}}\exp\left(-\frac{1}{2}(x-m)'\Sigma^{-1}(x-m)\right)$$
avec $m = (0,0)'$ et
$$\Sigma^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$
On déduit que (X,Y) est un vecteur gaussien centré et de matrice de variance covariance
$$\Sigma = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}.$$
Par conséquent $X \sim \mathcal{N}(0,4)$ et $Y \sim \mathcal{N}(0,2)$.

2. On a
$$\begin{pmatrix} Z \\ X \end{pmatrix} = \begin{pmatrix} -a & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}.$$
Donc (Z,X) est un vecteur gaussien centré et de matrice de variance covariance
$$\begin{pmatrix} 4a^2 - 4a - 2 & -4a+2 \\ -4a+2 & 4 \end{pmatrix}.$$
Par conséquent, X et Z sont indépendantes si et seulement si $a = 1/2$.

3. (a) On a

$$\begin{pmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 0 \\ \vdots & & & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

donc $(S_1, \dots, S_n)$ est un vecteur gaussien. Sont espérance vaut $(0, \dots, 0)$ et le terme de la $i$ème ligne et $j$ème colonne de la matrice de variance covariance est donné par $\min(i, j)$.

(b) On a

$$T_m = \sum_{i=1}^{m} X_i - \frac{m}{n} \sum_{i=1}^{n} X_i = \sum_{i=1}^{m} X_i \left(1 - \frac{m}{n}\right) - \frac{m}{n} \sum_{i=m+1}^{n} X_i.$$

T_m est donc une variable aléatoire gaussienne. Il suffit donc de montrer que $\mathbf{Cov}(T_m, S_n) = 0$. On a

$$\begin{aligned} \mathbf{Cov}(T_m, S_n) =& \mathbf{Cov}\left(\left(1 - \frac{m}{n}\right) \sum_{i=1}^{m} X_i - \frac{m}{n} \sum_{i=m+1}^{n} X_i, \sum_{i=1}^{n} X_i\right) \\ =& \left(1 - \frac{m}{n}\right) \sum_{i=1}^{m} \sum_{j=1}^{n} \mathbf{Cov}(X_i, X_j) - \frac{m}{n} \sum_{i=m+1}^{n} \sum_{j=1}^{n} \mathbf{Cov}(X_i, X_j) \\ =& \left(1 - \frac{m}{n}\right) \times m - \frac{m}{n} \times (n - m) = 0. \end{aligned}$$

Examen partiel Statistique 2

20 novembre 2013
Aucun document, pas de calculatrice

Le sujet est composé de 3 exercices indépendants (il y a 2 pages !).

Le barême est donné à titre indicatif.

Exercice 1 (Questions de cours, 10 points)

1. Donner la définition d'une variable aléatoire réelle ainsi que de sa loi de probabilité.
2. (a) Enoncer l'inégalité de Markov (sans oublier de donner les hypothèses).
   (b) Enoncer l'inégalité de Bienaymé Tchebychev (sans oublier de donner les hypothèses).
   (c) Démontrer l'inégalité de Bienaymé Tchebychev.
3. Soit X une variable aléatoire de fonction de répartition F_X. On pose $Y = aX + b$ avec $a > 0$ et $b \in \mathbb{R}$. Exprimer la fonction de répartition de Y en fonction de F_X, a et b.
4. Soit X une variable aléatoire de loi uniforme sur $[-1, 1]$.
   (a) Quelle est la densité de X ?
   (b) Rappeler la définition de la fonction caractéristique d'une variable aléatoire réelle.
   (c) Calculer la fonction caractéristique de X.
   (d) Soit $X_1, \ldots, X_n$ n variables aléatoires réelles indépendantes et de même loi que X. Que vaut la fonction caractéristique de $S_n = X_1 + \ldots + X_n$?
5. Soit $X = (X_1, X_2)$ un vecteur gaussien centré réduit. On note
$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$
   (a) Donner l'espérance et la matrice de variance-covariance de X.
   (b) Calculer $\mathbf{P}((X_1, X_2) = (1, 0))$.
   (c) Les variables aléatoires X_1 et X_2 sont-elles indépendantes ?
   (d) Quelle est la loi du vecteur $Y = AX + b$?
6. Soit X et Y deux variables aléatoires réelles telles que $\mathbf{Cov}(X, Y) = 0$. X et Y sont-elles indépendantes (si oui, justifier votre réponse, si non donner un contre exemple).
7. (a) Rappeler la définition d'un vecteur gaussien.
   (b) Soit $X = (X_1, X_2)$ un vecteur gaussien. On suppose que $\mathbf{Cov}(X_1, X_2) = 0$. Montrer que X_1 et X_2 sont indépendantes. On pourra utiliser que la fonction caractéristique d'un vecteur gaussien $Y = (Y_1, \ldots, Y_d)$ d'espérance m_Y et de matrice de variance-covariance Σ_Y est
$$\varphi_Y(t) = e^{i<m_Y,t>}e^{-0.5t'\Sigma_Y t}, \quad t \in \mathbb{R}^d.$$

Exercice 2 (5.5 points)
Soit X et Y deux variables aléatoires indépendantes de même loi de densité f définie par

$$f(x) = \exp(-x)\mathbf{1}_{\{x\geq 0\}}.$$

On pose $S = X + Y$.

1. Calculer $\mathbf{P}(0 < X < 3)$.
2. Calculer $\mathbf{E}[X]$ et $\mathbf{V}[X]$.
3. Déterminer la densité du couple (X, S).
4. En déduire la densité de S.
5. Retrouver la densité de S en utilisant le produit de convolution.
6. Les variables aléatoires X et S sont-elles indépendantes ? Justifier.
7. Calculer $\mathbf{P}((X, S) \in [0, 1] \times [0, 1])$.

Exercice 3 (4.5 points)
Soit $\theta > 0$ un nombre réel fixé et $Z = (X, Y)$ un vecteur aléatoire dont la densité f est définie sur $\mathbb{R}^2$ par

$$f(x, y) = C \exp(-\theta y)\mathbf{1}_{0<x<y}, \quad (x, y) \in \mathbb{R}^2.$$

1. Montrer que $C = \theta^2$.
2. Déterminer les lois marginales de Z ainsi que $\mathbf{E}[X]$ et $\mathbf{V}[X]$. Les variables aléatoires X et Y sont-elles indépendantes ?
3. On pose $U = X$ et $V = Y - X$.
   (a) Déterminer la loi du couple $(U, V)'$.
   (b) Les variables aléatoires U et V sont-elles indépendantes ?
4. A l'aide de la question précédente, calculer $\mathbf{Cov}[X, Y]$. En déduire $\mathbf{V}[Y]$.

CORRECTION

Exercice 1

1. Une variable aléatoire est une fonction X définie sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$ à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que
$$\forall B \in \mathcal{B}(\mathbb{R}),\ X^{-1}(B) \in \mathcal{A}.$$
Sa loi de probabilité est l'application $\mathbf{P}_X$ définie par
$$\begin{aligned}\mathbf{P}_X : (\mathbb{R}, \mathcal{B}(\mathbb{R})) &\to [0,1] \\ B &\mapsto \mathbf{P}_X(B) = \mathbf{P}(X^{-1}(B)).\end{aligned}$$

2. (a) Si X est une v.a.r. positive, on a pour tout réel $a > 0$
$$\mathbf{P}(X \geq a) \leq \frac{\mathbf{E}[X]}{a}.$$
   (b) Soit X une v.a.r. telle que $\mathbf{E}[X^2] < +\infty$, on a pour tout réel $a > 0$
$$\mathbf{P}(|X - \mathbf{E}[X]| \geq a) \leq \frac{\mathbf{V}[X]}{a^2}.$$
   (c) Il suffit d'appliquer Markov à la variable aléatoire $(X - \mathbf{E}[X])^2$.

3. On a
$$F_Y(y) = \mathbf{P}(Y \leq y) = \mathbf{P}(aX + b \leq y) = \mathbf{P}\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right).$$

4. (a) La densité est donnée par $f_X(x) = \frac{1}{2}\mathbf{1}_{]-1,1[}(x)$.
   (b) La fonction caractéristique est $\varphi_Y(t) = \mathbf{E}[\exp(-itY)]$.
   (c) On déduit
$$\varphi_X(t) = \frac{1}{2}\int_{-1}^{1} \exp(itx)\,\mathrm{d}x = \frac{1}{2}\int_{-1}^{1} (\cos(tx) + i\sin(tx))\,\mathrm{d}x = \frac{\sin(t)}{t}.$$
   (d) Comme $X_1, \dots, X_n$ sont indépendantes, on a $\varphi_{S_n}(t) = (\varphi_X(t))^n$.

5. (a) On a $\mathbf{E}[X] = (0,0)$ et $\Sigma_X = Id$.
   (b) Σ_X est inversible donc le vecteur X est absolument continu. Cette probabilité est donc nulle.
   (c) $\mathbf{Cov}(X_1, X_2) = 0$. X étant un vecteur gaussien, on déduit que X_1 et X_2 sont indépendantes.
   (d) Y est la transformée affine d'un vecteur gaussien, c'est donc un vecteur gaussien. Il suffit de déterminer son espérance et sa matrice de variance covariance. On a $\mathbf{E}[Y] = b$ et
$$\Sigma_Y = A\Sigma_X A' = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

6. Non, il suffit par exemple de prendre X de loi normale centrée réduite et $Y = X^2$.
7. (a) $X = (X_1, \ldots, X_d)$ est un vecteur gaussien si toute combinaison linéaire de ses marginales est une variable aléatoire réelle gaussienne.
   (b) Il suffit de montrer que $\varphi_X(t) = \varphi_{X_1}(t_1)\varphi_{X_2}(t_2)$.

Exercice 2

1. On a $\mathbf{P}(0 < X < 3) = 1 - \exp(-3)$.
2. X suit une loi exponentielle de paramètre 1 donc $\mathbf{E}[X] = \mathbf{V}[X] = 1$.
3. On détermine la densité du couple (X, Y) :
$$f_{X,Y}(x, y) = \exp(-(x + y))\mathbf{1}_{x>0}\mathbf{1}_{y>0},$$
et on effectue le changement de variable
$$\begin{aligned} \varphi : \mathbb{R}^+ \times \mathbb{R}^+ &\to \mathbb{R}^+ \times \mathbb{R}^+ \\ (x, y) &\mapsto (x, x + y). \end{aligned}$$
on déduit que la densité de (X, S) est donnée par
$$f_{X,S}(x, s) = \exp(-v)\mathbf{1}_{0<u<v}.$$
4. La densité de S est donnée par
$$f_S(s) = s\exp(-s)\mathbf{1}_{s>0}.$$
5. On peut également utiliser le produit de convolution :
$$f_S(s) = \int_{\mathbb{R}} \exp(-(s - t))\mathbf{1}_{s-t>0} \exp(-t)\mathbf{1}_{t>0} \,\mathrm{d}t = s\exp(-s)\mathbf{1}_{s>0}.$$
6. Clairement $f_{X,S} \neq f_X \times f_S$ donc X et S ne sont pas indépendantes.
7. Il suffit de calculer
$$\mathbf{P}((X, S) \in [0, 1]^2) = \int_{[0,1]^2} f_{X,S}(x, s) \,\mathrm{d}\lambda^2(x, s) = 1 - \exp(-1).$$

Exercice 3

1. Il suffit de calculer l'intégrale double
$$\int_{\mathbb{R}}\int_{\mathbb{R}} \exp(-\theta y)\mathbf{1}_{0<x<y} \,\mathrm{d}x \,\mathrm{d}y = \int_0^{+\infty}\int_0^y \exp(-\theta y) \,\mathrm{d}x \,\mathrm{d}y = \frac{1}{\theta^2}.$$
2. On obtient
$$f_X(x) = \theta^2 \int_x^{+\infty} \exp(-\theta y) \,\mathrm{d}y = \theta \exp(-\theta x), \text{ pour } x > 0,$$
X suit donc une loi exponentielle de paramètre θ, d'où $\mathbf{E}[X] = 1/\theta$ et $\mathbf{V}[X] = 1/\theta^2$. La seconde marginale vaut
$$f_Y(y) = \theta^2 \int_0^y \exp(-\theta y) \,\mathrm{d}y = \theta^2 y \exp(-\theta y), \text{ pour } y > 0.$$
Ainsi $f(x, y) \neq f_X(x)f_Y(y)$, X et Y ne sont donc pas indépendantes.

3. (a) On considère le changement de variable

$$\begin{aligned}\varphi : D = \{(x,y) \in \mathbb{R}^2 : 0 < x < y\} &\to \Delta = \{(u,v) \in \mathbb{R}^2 : u > 0, v > 0\} \\ (x,y) &\mapsto (x, y-x).\end{aligned}$$

φ est un C^1 difféomorphisme, d'après le théorème de changement de variable, la densité de $(U,V)'$ est donnée par

$$f_{U,V}(u,v) = f(\varphi^{-1}(u,v))|\det(J_{\varphi^{-1}_{u,v}})|\mathbf{1}_\Delta(u,v).$$

Or

$$\begin{aligned}\varphi^{-1} &: \Delta \to D \\ x(u,v) &\mapsto (u, u+v),\end{aligned}$$

d'où

$$\det(J_{\varphi^{-1}_{u,v}}) = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1.$$

On obtient

$$f_{U,V}(u,v) = f(u,u+v) = \theta^2 \exp(-\theta(u+v))\mathbf{1}_\Delta(u,v) = \theta \exp(-\theta u)\mathbf{1}_{u>0}\, \theta \exp(-\theta v)\mathbf{1}_{v>0}.$$

(b) D'après la question précédente, U et V sont indépendantes. Elles suivent de plus une loi exponentielle de paramètre θ.

4. Comme

$$\mathbf{Cov}(U,V) = \mathbf{Cov}(X, Y-X) = \mathbf{Cov}(X,Y) - \mathbf{V}[X] = 0,$$

on déduit $\mathbf{Cov}(X,Y) = \mathbf{V}[X] = 1/\theta^2$. De même,

$$\mathbf{V}[V] = \mathbf{V}[Y-X] = \mathbf{V}[Y] + \mathbf{V}[X] - 2\mathbf{Cov}(X,Y) = \frac{1}{\theta^2} \implies \mathbf{V}[Y] = \frac{2}{\theta^2}.$$

Examen de Statistique 2, janvier 2014
deux heures, sans document, pas de calculatrice

Le barème est donné à titre indicatif

Rappels :

— si elle existe, la densité d'un vecteur gaussien $X = (X_1, \dots, X_d)$ d'espérance m_X et de matrice de variance covariance Σ_X est donnée par

$$f_X(x) = \frac{1}{(\sqrt{2\pi})^d \sqrt{\det(\Sigma_X)}} \exp\left(-\frac{1}{2}(x - m_X)'\Sigma_X^{-1}(x - m_X)\right) ;$$

— la densité de la loi exponentielle de paramètre $\lambda > 0$ est $f(x) = \lambda \exp(-\lambda x)\mathbf{1}_{]0,+\infty[}(x)$.

Exercice 1 (6 points)

Les réponses aux questions de cet exercice ne necéssitent pas de longs développements mathématiques (si elles sont bien abordées...). Les réponses données devront être courtes et précises.

1. Enoncer les définitions des convergences presque sure, en probabilité et en loi.
2. Enoncer la loi forte des grands nombres et le théorème central limite.
3. Soit X une variable aléatoire réelle de loi normale $\mathcal{N}(0, 1)$. Soit ε une variable aléatoire indépendante de X de loi
$$\mathbf{P}(\varepsilon = -1) = \mathbf{P}(\varepsilon = 1) = \frac{1}{2}.$$
On pose $Y = \varepsilon X$.
   (a) Montrer que Y suit une loi $\mathcal{N}(0, 1)$.
   (b) Le vecteur (X, Y) est-il gaussien ? Justifier.
4. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles qui converge vers une variable aléatoire X dans L_2 (ou en moyenne quadratique). Est-ce que $(X_n)_{n\in\mathbb{N}}$ converge dans L_1 ? Si non, exhiber une contre exemple, si oui, prouver le.
5. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires, indépendantes, de loi exponentielle de paramètre $\lambda > 0$. On pose $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.
   (a) Montrer à l'aide de l'inégalité de Bienaymé-Tchebychev que la suite $(\bar{X}_n)_{n\geq 1}$ converge en probabilité vers une constante a que l'on précisera.
   (b) A l'aide du théorème central limite, donner une suite réelle $(a_n)_{n\geq 1}$ telle que la suite
$$a_n\left(\bar{X}_n - \frac{1}{\lambda}\right)$$
converge en loi vers une limite à préciser.
6. Soit (X, Y) un vecteur aléatoire gaussien tel que $\mathbf{V}[X] > 0$. On pose
$$Z = Y - \frac{\mathbf{Cov}(X, Y)}{\mathbf{V}[X]}X.$$

(a) Montrer que (X, Y, Z) est un vecteur gaussien.

(b) Les variables aléatoires réelles X et Z sont-elles indépendantes ? Justifier.

Exercice 2 (3 points)
Soit (X, Y) un vecteur gaussien admettant une densité f définie par $f(x, y) = Ce^{-\frac{3x^2+2xy+3y^2}{8}}$.

1. Déterminer la constante C, l'espérance et la matrice de variance-covariance de (X, Y).
2. Déterminer la loi des variables marginales X et Y. Ces variables sont-elles indépendantes ?
3. On pose $U = X + Y$ et $V = X - Y$. Quelle est la loi du couple (U, V) ? Conclure.

Exercice 3 (5 points)
Soit (X, Y) un couple aléatoire dont la densité jointe est donnée par

$$f_{X,Y}(x, y) = 4y(x - y) \exp(-(x + y))\mathbf{1}_{0<y<x}.$$

Afin de simplifier les calculs, on admettra que pour $u > 0$, on a

$$\int_u^{+\infty} t \exp(-t)\,\mathrm{d}t = (u + 1) \exp(-u) \quad \text{et} \quad \int_u^{+\infty} t^2 \exp(-t)\,\mathrm{d}t = (u(u + 2) + 2) \exp(-u).$$

1. Calculer la densité de Y.
2. Montrer que $\mathbf{E}[Y] = 1$.
3. Calculer l'espérance conditionnelle $\mathbf{E}[X|Y]$ et en déduire $\mathbf{E}[X]$.
4. Calculer la probabilité conditionnelle $\mathbf{P}[X < 1|Y]$.
5. Indiquer deux façons différentes de calculer $\mathbf{P}[X < 1]$ (il n'est pas demandé de faire les calculs).

Exercice 4 (6 points)
Soient n variables aléatoires $X_1, \ldots, X_n$ indépendantes, toutes distribuées selon la loi uniforme $\mathcal{U}_{[0,\theta]}$ sur l'intervalle $[0, \theta]$, où θ désigne un paramètre inconnu strictement positif.

1. Donner la densité de X_1.
2. Calculer $\mathbf{E}[X_1]$ et $\mathbf{V}[X_1]$.
3. En utilisant la loi forte des grands nombres, donner une suite de variables aléatoires $(T_n)_{n\in\mathbb{N}}$ qui converge presque sûrement vers θ.
4. On pose $Y_n = \max(X_1, \ldots, X_n)$.
   (a) Calculer la fonction de répartition de Y_n.
   (b) En déduire la densité de Y_n est donnée par
   $$f_{Y_n}(x) = n\frac{x^{n-1}}{\theta^n}\mathbf{1}_{]0,\theta[}(x).$$
   (c) Calculer $\mathbf{E}[Y_n]$ et $\mathbf{E}[Y_n^2]$.
   (d) Rappeler la définition de la convergence en moyenne quadratique et déduire de la question précédente que $(Y_n)_{n\in\mathbb{N}}$ converge en moyenne quadratique vers θ.
   (e) Soit $\varepsilon > 0$, calculer $\mathbf{P}(|Y_n - \theta| > \varepsilon)$.
   (f) En déduire que $(Y_n)_{n\in\mathbb{N}}$ converge presque surement vers une limite à préciser (on pourra utiliser Borel Cantelli).
   (g) Calculer la fonction de répartition de la variable aléatoire $Z_n = n(\theta - Y_n)$.
   (h) En déduire que la suite de variables aléatoires $(Z_n)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire dont on précisera la loi.

CORRECTION

Exercice 1

1. $(X_n)_{n\in\mathbb{N}}$ converge presque surement vers X si

$$\mathbf{P}\left(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) \neq X(\omega)\}\right) = 0.$$

$(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X si

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} \mathbf{P}(|X_n - X| > \varepsilon) = 0.$$

2. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. de $L_1(\Omega, \mathcal{A}, \mathbf{P})$ indépendantes et de même loi. On note $\mathbf{E}[X_1] = \mu$. On a

$$\frac{1}{n}\sum_{i=1}^{n} X_i \stackrel{p.s.}{\to} \mu \quad \text{quand } n \to \infty.$$

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes, de même loi, et telles que $\mathbf{E}[X_i^2] < +\infty$. On note $\mathbf{E}[X_i] = \mu$, $\mathbf{V}[X_i] = \sigma^2$ et $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. On a alors

$$\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \stackrel{\mathcal{L}}{\to} \mathcal{N}(0,1) \quad \text{quand } n \to \infty.$$

3. (a) Montrons que $F_Y(u) = F_X(u)$ pour tout $u \in \mathbb{R}$. On a

$$\begin{aligned} F_Y(u) &= \mathbf{P}(Y \leq u) = \mathbf{P}(\varepsilon X \leq u | \varepsilon = 1)\mathbf{P}(\varepsilon = 1) + \mathbf{P}(\varepsilon X \leq u | \varepsilon = -1)\mathbf{P}(\varepsilon = -1) \\ &= \frac{1}{2}[\mathbf{P}(X \leq u) + \mathbf{P}(X \geq -u)] = \mathbf{P}(X \leq u) = F_X(u). \end{aligned}$$

Y est donc une v.a.r. gaussienne centrée réduite.

(b) Les composantes du vecteur $(X,Y)'$ suivent des lois gaussiennes alors que $(X,Y)'$ n'est pas un vecteur gaussien. En effet, il suffit de voir que $Z = X + Y$ n'est pas une v.a.r. gaussienne :

$$\mathbf{P}(Z = 0) = \mathbf{P}((1+\varepsilon)X = 0) = \mathbf{P}(\varepsilon = -1) = \frac{1}{2}.$$

Or si Z est une v.a.r. gaussienne on a $\mathbf{P}(Z = 0) = 0$ (ou 1 dans le cas d'une v.a.r. centrée dégénérée).

4. Oui. En effet

$$\mathbf{E}|X_n - X| = \mathbf{E}\left[\sqrt{(X_n - X)^2}\right] \leq \sqrt{\mathbf{E}[(X_n - X)^2]}$$

d'après l'inégalité de Jensen (racine carrée est concave).

5. (a) Comme $\mathbf{E}[X_1] = 1/\lambda$ et $\mathbf{V}[X_1] = 1/\lambda^2$, on a d'après Bienaymé-Tchebychev que $\forall \varepsilon > 0$,

$$\mathbf{P}\left(\left|\bar{X}_n - \frac{1}{\lambda}\right| > \varepsilon\right) \leq \frac{\mathbf{V}[\bar{X}_n]}{\varepsilon^2} = \frac{1}{n\lambda^2\varepsilon^2}.$$

On déduit que $\bar{X}_n$ converge en probabilité vers $1/\lambda$.

(b) Il suffit d'appliquer le TCL :

$$\sqrt{n}\left(\bar{X}_n - \frac{1}{\lambda}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{\lambda^2}\right).$$

6. (a) Il suffit d'écrire (X, Y, Z) comme la transformée linéaire du vecteur gaussien (X, Y) :

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{\mathbf{Cov}(X,Y)}{\mathbf{V}[X]} & 1 \end{pmatrix}.$$

(b) (X, Y, Z) étant un vecteur gaussien, il suffit de montrer que $\mathbf{Cov}(X, Z) = 0$:

$$\mathbf{Cov}(X, Z) = \mathbf{Cov}(X, Y) - \frac{\mathbf{Cov}(X, Y)}{\mathbf{V}[X]}\mathbf{V}[X] = 0.$$

Exercice 2

1. Par identification avec l'expression de la densité d'un vecteur gaussien, on trouve l'espérance $m = (0, 0)$, et l'inverse de la matrice de variance-covariance

$$\Sigma^{-1} = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix}, \text{ d'où } \quad \Sigma = \begin{pmatrix} 3/2 & -1/2 \\ -1/2 & 3/2 \end{pmatrix},$$

et $\det(\Sigma) = 2$, d'où $C = (2\sqrt{2}\pi)^{-1}$.

2. Par définition du vecteur gaussien, les variables marginales X et Y suivent des loi normales, centrées, de variance $3/2$. Elles ne sont pas indépendantes car elles sont corrélées.
3. On a

$$\begin{pmatrix} U \\ V \end{pmatrix} = A \begin{pmatrix} X \\ Y \end{pmatrix} \quad \text{avec} \quad A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Donc (U, V) suit une loi normale centrée, de matrice de variance-covariance

$$A\Sigma A' = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

Puisque (U, V) est un vecteur gaussien et que les variables marginales U et V ne sont pas corrélées, U et V sont indépendantes.

Exercice 3

1. On a

$$f_Y(y) = 4y \exp(-2y)\mathbf{1}_{y>0}.$$

2. On déduit

$$\mathbf{E}[Y] = \int_0^{+\infty} 4y^2 \exp(-2y)\,\mathrm{d}y = 1.$$

3. Pour $y > 0$, on a

$$f_{X|Y=y} = \frac{f_{X,Y}(x, y)}{f_Y(y)} = (x - y)\exp(y - x)\mathbf{1}_{0<y<x}.$$

Ainsi

$$\mathbf{E}[X|Y = y] = \int x f_{X|Y=y}\,\mathrm{d}x = y + 2.$$

Par conséquent $\mathbf{E}[X|Y] = Y + 2$ et $\mathbf{E}[X] = \mathbf{E}[\mathbf{E}[X|Y]] = \mathbf{E}[Y] + 2 = 3$.

4. On a pour $0 < y < 1$,

$$\mathbf{P}[X < 1|Y = y] = \mathbf{E}[\mathbf{1}_{X<1}|Y = y] = \int \mathbf{1}_{x<1} f_{X|Y=y}\,\mathrm{d}x$$
$$= \int_y^1 (x - y)\exp(y - x)\,\mathrm{d}x = (1 + \exp(y - 1)(y - 2)).$$

D'où $\mathbf{P}[X < 1|Y] = (1 + \exp(Y - 1)(Y - 2))\mathbf{1}_{0<Y<1}$.

5. On peut utiliser le conditionnement

$$\mathbf{P}[X < 1] = \int \mathbf{P}[X < 1|Y = y] f_Y(y)\,\mathrm{d}y,$$

ou directement la loi de X

$$\mathbf{P}[X < 1] = \int_0^1 f_X(x)\,\mathrm{d}x.$$

Exercice 4

1. La densité de X_1 est donnée par $f(x) = 1/\theta\ \mathbf{1}_{[0,\theta]}(x)$.
2. L'espérance et la variance sont respectivement données par $\theta/2$ et $\theta^2/12$.
3. Il suffit de prendre $T_n = 2\bar{X}_n$.
4. (a) Par définition, la fdr vaut 0 si $t \leq 0$ et 1 si $t \geq \theta$. Si $0 \leq t \leq \theta$, on a

$$F_{Y_n}(t) = \mathbf{P}(Y_n \leq t) = \prod_{i=1}^n \mathbf{P}(X_i \leq t) = \prod_{i=1}^n \frac{1}{\theta}\int_0^t \mathrm{d}t = \frac{t^n}{\theta^n}.$$

(b) La densité de Y_n est ainsi donnée par

$$f_{Y_n}(t) = n\frac{t^{n-1}}{\theta^n}\mathbf{1}_{]0,\theta[}(t).$$

On peut ainsi calculer

$$\mathbf{E}[Y_n] = \frac{n}{n+1}\theta \quad \text{et} \quad \mathbf{E}[Y_n^2] = \frac{n}{n+2}\theta^2.$$

(c) La suite $(Y_n)_{n\in\mathbb{N}}$ converge en moyenne quadratique vers θ si

$$\lim_{n\to\infty} \mathbf{E}[(Y_n - \theta)^2] = 0.$$

On déduit de la question précédente

$$\mathbf{E}[(Y_n - \theta)^2] = \frac{2\theta^2}{(n+2)(n+1)} \to 0 \quad \text{quand} \quad n \to \infty,$$

donc $Y_n \overset{L_2}{\to} \theta$.

(d) Soit $\varepsilon < \theta$. On a

$$\mathbf{P}(|Y_n - \theta| > \varepsilon) = 1 - \mathbf{P}(\theta - \varepsilon \leq Y_n \leq \varepsilon + \theta) = 1 - \int_{\theta-\varepsilon}^{\theta} f_{Y_n}(t)\,\mathrm{d}t = \left(1 - \frac{\varepsilon}{\theta}\right)^n = u_n.$$

(e) Comme la série de terme général u_n est convergente, on déduit d'après Borel-Cantelli que Y_n converge presque surement vers θ.

(f) Si $t \leq 0$, $F_{Z_n}(t) = 0$ et si $t \geq n\theta$, $F_{Z_n}(t) = 1$. Sinon

$$F_{Z_n}(t) = 1 - F_{Y_n}\left(\theta - \frac{t}{n}\right) = 1 - \left(1 - \frac{t}{n\theta}\right)^n.$$

Or

$$\begin{aligned}\left(1 - \frac{t}{n\theta}\right)^n &= \exp\left(n \log\left(1 - \frac{t}{n\theta}\right)\right) \\ &= \exp\left(n\left(-\frac{t}{n\theta} + \mathrm{o}\left(\frac{t}{n\theta}\right)\right)\right) \to \exp\left(-\frac{t}{\theta}\right) \text{ quand } n \to \infty.\end{aligned}$$

On a donc $\lim_{n\to\infty} F_{Z_n}(t) = 1 - \exp(-t/\theta)$. Par conséquent $(Z_n)_{n\in\mathbb{N}}$ converge en loi vers une loi exponentielle de paramètre $1/\theta$.

Examen de Probabilités Générales, janvier 2015
deux heures, sans document, pas de calculatrice

Le barème est donné à titre indicatif

Rappel : si elle existe, la densité d'un vecteur gaussien $X = (X_1, \dots, X_d)$ d'espérance m_X et de matrice de variance covariance Σ_X est donnée par

$$f_X(x) = \frac{1}{(\sqrt{2\pi})^d \sqrt{\det(\Sigma_X)}} \exp\left(-\frac{1}{2}(x - m_X)'\Sigma_X^{-1}(x - m_X)\right).$$

Exercice 1 (6.5 points)

Les réponses aux questions de cet exercice ne necéssitent pas de longs développements mathématiques (si elles sont bien abordées...). Les réponses données devront être courtes et précises.

1. Enoncer les définitions des convergences presque sûre, en probabilité et en loi.
2. On considère $(X_n)_{n\in\mathbb{N}^\star}$ une suite de variables aléatoires réelles telle que X_n suit une loi uniforme sur $]-1/n; 1/n[$.
   (a) Est-ce que X_n converge en probabilité vers 0 ? Justifier.
   (b) Est-ce que X_n converge en loi vers 0 ? Justifier.
3. Soit $(X_n)_{n\in\mathbb{N}^\star}$ une suite de variables aléatoires réelles telle que
$$\mathbf{P}(X_n = \sqrt{n}) = \frac{1}{n} \quad \text{et} \quad \mathbf{P}(X_n = 0) = 1 - \frac{1}{n}.$$
   (a) Est-ce que $(X_n)_{n\in\mathbb{N}^\star}$ converge vers 0 dans L_2 ? Justifier.
   (b) Est-ce que $(X_n)_{n\in\mathbb{N}^\star}$ converge en probabilité vers 0 ? Justifier.
4. Enoncer la loi forte des grands nombres et le théorème central limite.
5. (a) Donner la définition d'un vecteur gaussien.
   (b) Rappeler la définition de la loi du χ^2 à n degrés de liberté.
   (c) Soit X une variable aléatoire réelle de loi $\chi^2(n)$ et Y une variable aléatoire réelle de loi $\chi^2(m)$. X et Y sont indépendantes. Quelle est la loi de $X + Y$? Justifier votre réponse.
   (d) Enoncer le théorème de Cochran.
   (e) Soit $X_1, \dots, X_n$ n variables aléatoires indépendantes et de même loi normale $\mathcal{N}(\mu, \sigma^2)$ où $\mu \in \mathbb{R}$ et $\sigma^2 > 0$. En utilisant le théorème de Cochran, donner la loi de $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.

Exercice 2 (3.5 points)

Soit (X, Y) un couple aléatoire de densité $f_{X,Y}$ définie par :

$$f_{X,Y}(x, y) = \frac{1}{4\pi} \exp\left(-\frac{1}{2}\left(\frac{x^2}{2} - xy + y^2\right)\right), \quad (x, y) \in \mathbb{R}^2.$$

1. Montrer que (X, Y) est un vecteur gaussien, puis déterminer son espérance et sa matrice de variance covariance.
2. Déterminer les lois des marginales X et Y.
3. Soit $a \in \mathbb{R}$ et soit $U = Y - aX$. Quelle est la loi du vecteur (U, X) ?
4. Pour quelle(s) valeur(s) de a les variables aléatoires U et X sont-elles indépendantes ?

Exercice 3 (4.5 points)
Soit $\theta > 0$ un nombre réel fixé et $Z = (X, Y)$ un vecteur aléatoire dont la densité f est définie sur $\mathbb{R}^2$ par

$$f(x,y) = C\exp(-\theta y)\mathbf{1}_{0<x<y}, \quad (x,y) \in \mathbb{R}^2.$$

1. Montrer que $C = \theta^2$.
2. Déterminer les lois marginales de Z ainsi que $\mathbf{E}[X]$ et $\mathbf{V}[X]$. Les variables aléatoires X et Y sont-elles indépendantes ?
3. On pose $U = X$ et $V = Y - X$.
   (a) Déterminer la loi du couple (U, V).
   (b) Les variables aléatoires U et V sont-elles indépendantes ?
4. A l'aide de la question précédente, calculer $\mathbf{Cov}[X, Y]$. En déduire $\mathbf{V}[Y]$.
5. Calculer $\mathbf{E}[Y|X]$. En déduire $\mathbf{E}[Y]$.
6. Indiquer deux façons différentes de calculer $\mathbf{P}(Y > X^2)$ (il n'est pas demandé de faire le calcul).

Exercice 4 (5.5 points)
Soit $X_1, \ldots, X_n$ un n échantillon composé de variables aléatoires indépendantes et identiquement distribuées selon la loi admettant pour densité

$$f(x) = \frac{1}{2\sqrt{x\theta}}\mathbf{1}_{]0,\theta]}(x)$$

où θ est un paramètre strictement positif.

1. Montrer que f est bien une densité de probabilité.
2. Montrer que $\mathbf{E}[X_1] = \theta/3$ et calculer la variance de X_1.
3. En utilisant le théorème central limite, proposer une suite $(U_n)_{n\in\mathbb{N}}$ telle que
$$\sqrt{n}(U_n - \theta) \xrightarrow{\mathcal{L}} N(0, \varphi(\theta))$$
   où φ est une fonction à préciser.
4. On pose $Y_n = \max(X_1, \ldots, X_n)$. Calculer la fonction de répartition de Y_n et en déduire la densité de Y_n.
5. Calculer l'espérance et la variance de Y_n.
6. Montrer que
$$\mathbf{E}[(Y_n - \theta)^2] = (\mathbf{E}[Y_n] - \theta)^2 + \mathbf{V}(Y_n).$$
7. Est-ce que $(Y_n)_{n\in\mathbb{N}}$ converge vers θ en moyenne quadratique (ou dans L_2) ? Justifier.
8. Montrer que $n(\theta - Y_n)$ converge en loi vers une loi à préciser.
9. Etudier les convergences en probabilité et presque sûre de Y_n vers θ.

CORRECTION

Exercice 1

1. $(X_n)_{n\in\mathbb{N}}$ converge presque surement vers X si

$$\mathbf{P}\Big(\{\omega\in\Omega : \lim_{n\to\infty} X_n(\omega)\neq X(\omega)\}\Big)=0.$$

$(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X si

$$\forall\varepsilon>0,\quad \lim_{n\to\infty}\mathbf{P}(|X_n-X|>\varepsilon)=0.$$

$(X_n)_{n\in\mathbb{N}}$ converge en loi vers X si $\lim_{n\to\infty} F_{X_n}(x)=F_X(x)$ en tout point de continuité de F_X.

2. (a) On a $\forall\varepsilon>0$

$$\mathbf{P}(|X_n|>\varepsilon)=1-\mathbf{P}(-\varepsilon\le X_n\le\varepsilon)=1-\frac{n}{2}[\min(1/n,\varepsilon)-\max(-1/n,-\varepsilon)]=0$$

pour n assez grand. On conclut $X_n\overset{\mathbf{P}}{\to}0$.

(b) On a

$$F_{X_n}(x)=\begin{cases}0 & \text{si } x\le -1/n\\ \frac{n}{2}\left(x+\frac{1}{n}\right) & \text{si } -1/n<x\le 1/n\\ 1 & \text{si } x>1/n.\end{cases}$$

On a donc $\lim_{n\to+\infty}F_{X_n}(x)=F_X(x)$ pour tout $x\in\mathbb{R}^\star$. F_X étant discontinue en 0, on conclut que $X_n\overset{\mathcal{L}}{\to}0$. On aurait aussi pu conclure en disant que la convergence en probabilité implique la convergence en loi.

3. (a) Comme $\mathbf{E}[X_n^2]=1$, on déduit que $(X_n)_n$ ne converge pas vers 0 dans L_2.

(b) Soit $\varepsilon>0$. On a

$$\mathbf{P}(|X_n|>\varepsilon)=\mathbf{P}(X_n=\sqrt{n})=\frac{1}{n}$$

pour n assez grand, donc $X_n\overset{\mathbf{P}}{\to}0$.

4. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. de $L_1(\Omega,\mathcal{A},\mathbf{P})$ indépendantes et de même loi. On note $\mathbf{E}[X_1]=\mu$. On a

$$\frac{1}{n}\sum_{i=1}^n X_i\overset{p.s.}{\to}\mu\quad\text{quand } n\to\infty.$$

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes, de même loi, et telles que $\mathbf{E}[X_i^2]<+\infty$. On note $\mathbf{E}[X_i]=\mu$, $\mathbf{V}[X_i]=\sigma^2$ et $\bar{X}_n=\frac{1}{n}\sum_{i=1}^n X_i$. On a alors

$$\sqrt{n}\frac{\bar{X}_n-\mu}{\sigma}\overset{\mathcal{L}}{\to}\mathcal{N}(0,1)\quad\text{quand } n\to\infty.$$

5. (a) $X=(X_1,\dots,X_n)$ est un vecteur gaussien si toute combinaison linéaire de ses marginales est une variable aléatoire réelle gaussienne.

(b) $X \sim \chi^2(n)$ si il existe n variables aléatoires $X_1, \ldots, X_n$ indépendantes de loi normale centrée réduite telles que $X = X_1^2 + \ldots + X_n^2$.

(c) On a
$$X + Y = X_1^2 + \ldots + X_n^2 + Y_1^2 + \ldots + Y_m^2$$
où les $X_i, i = 1, \ldots, n$ et $Y_j, j = 1, \ldots, m$ suivent des lois normales centrées réduites. Comme X et Y sont indépendantes, les variables $X_i, i = 1, \ldots, n$ et $Y_j, j = 1, \ldots, m$ sont indépendantes. On déduit que $X + Y$ suit une loi $\chi^2(n + m)$.

(d) Voir cours.

(e) Il suffit de centrer-réduire le vecteur $X = (X_1, \ldots, X_n)$ et d'appliquer le théorème de Cochran en considérant le sous-espace engendré par le vecteur $(1, \ldots, 1)$.

Exercice 2

1. On a
$$f_{X,Y}(x, y) = \frac{1}{(\sqrt{2\pi})^2 \sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(x - m)'\Sigma^{-1}(x - m)\right)$$
avec $m = (0, 0)'$ et
$$\Sigma^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$
On déduit que (X, Y) est un vecteur gaussien centré et de matrice de variance covariance
$$\Sigma = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}.$$
2. On déduit que $X \sim \mathcal{N}(0, 4)$ et $Y \sim \mathcal{N}(0, 2)$.
3. On a
$$\begin{pmatrix} U \\ X \end{pmatrix} = \begin{pmatrix} -a & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$
Donc (U, X) est un vecteur gaussien centré et de matrice de variance covariance
$$\begin{pmatrix} 4a^2 + 2 & -4a + 2 \\ -4a + 2 & 4 \end{pmatrix}.$$
4. (U, X) étant un vecteur gaussien, on déduit que U et X sont indépendantes si et seulement si $\mathbf{Cov}(U, X) = 0$, c'est-à-dire si et seulement si $a = -1/2$.

Exercice 3

1. Il suffit de calculer l'intégrale double
$$\int_{\mathbb{R}} \int_{\mathbb{R}} \exp(-\theta y)\mathbf{1}_{0<x<y} \, dx \, dy = \int_0^{+\infty} \int_0^y \exp(-\theta y) \, dx \, dy = \frac{1}{\theta^2}.$$
2. On obtient
$$f_X(x) = \theta^2 \int_x^{+\infty} \exp(-\theta y) \, dy = \theta \exp(-\theta x), \text{ pour } x > 0,$$
X suit donc une loi exponentielle de paramètre θ, d'où $\mathbf{E}[X] = 1/\theta$ et $\mathbf{V}[X] = 1/\theta^2$. La seconde marginale vaut
$$f_Y(y) = \theta^2 \int_0^y \exp(-\theta y) \, dy = \theta^2 y \exp(-\theta y), \text{ pour } y > 0.$$
Ainsi $f(x, y) \neq f_X(x) f_Y(y)$, X et Y ne sont donc pas indépendantes.

3. (a) On considère le changement de variable

$$\varphi : D = \{(x, y) \in \mathbb{R}^2 : 0 < x < y\} \to \Delta = \{(u, v) \in \mathbb{R}^2 : u > 0, v > 0\}$$
$$(x, y) \mapsto (x, y - x).$$

φ est un C^1 difféomorphisme, d'après le théorème de changement de variable, la densité de $(U, V)'$ est donnée par

$$f_{U,V}(u, v) = f(\varphi^{-1}(u, v))|\det(J_{\varphi^{-1}_{u,v}})|\mathbf{1}_\Delta(u, v).$$

Or

$$\varphi^{-1} : \Delta \to D$$
$$(u, v) \mapsto (u, u + v),$$

d'où

$$\det(J_{\varphi^{-1}_{u,v}}) = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1.$$

On obtient

$$f_{U,V}(u, v) = f(u, u + v) = \theta^2 \exp(-\theta(u + v))\mathbf{1}_\Delta(u, v) = \theta \exp(-\theta u)\mathbf{1}_{u>0}\, \theta \exp(-\theta v)\mathbf{1}_{v>0}.$$

(b) D'après la question précédente, U et V sont indépendantes. Elles suivent de plus une loi exponentielle de paramètre θ.

4. Comme
$$\mathbf{Cov}(U, V) = \mathbf{Cov}(X, Y - X) = \mathbf{Cov}(X, Y) - \mathbf{V}[X] = 0,$$
on déduit $\mathbf{Cov}(X, Y) = \mathbf{V}[X] = 1/\theta^2$. De même,

$$\mathbf{V}[V] = \mathbf{V}[Y - X] = \mathbf{V}[Y] + \mathbf{V}[X] - 2\mathbf{Cov}(X, Y) = \frac{1}{\theta^2} \implies \mathbf{V}[Y] = \frac{2}{\theta^2}.$$

5. On calcule d'abord la densité conditionnelle, pour $x > 0$,

$$f_{Y|X=x}(y) = \theta \exp(-\theta(y - x))\mathbf{1}_{0<x<y},$$

puis

$$\mathbf{E}[Y|X = x] = \theta \exp(\theta x) \int_x^\infty y \exp(-\theta y) = x + \frac{1}{\theta}.$$

On déduit $\mathbf{E}[Y|X] = X + \frac{1}{\theta}$ et $\mathbf{E}[Y] = 2/\theta$.

6. On peut passer par l'intégrale double

$$\mathbf{P}(Y > X^2) = \int \mathbf{1}_{y>x^2} f(x, y)\, \mathrm{d}\lambda_2(x, y)$$

ou par le conditionnement

$$\mathbf{P}(Y > X^2) = \int \mathbf{P}(Y > X^2|X = x) f_X(x)\, \mathrm{d}x.$$

Exercice 4

1. Il suffit de calculer
$$\int f(x)\,\mathrm{d}x = \int_0^\theta \frac{1}{2\sqrt{x\theta}}\,\mathrm{d}x = 1.$$
2. On a
$$\mathbf{E}[X_1] = \int_0^\theta \frac{\sqrt{x}}{2\sqrt{\theta}}\,\mathrm{d}x = \frac{1}{2\sqrt{\theta}}\left[\frac{2}{3}x^{3/2}\right]_0^\theta = \frac{\theta}{3}.$$
De même
$$\mathbf{E}[X_1^2] = \frac{\theta^2}{5} \quad \text{et} \quad \mathbf{V}[X_1] = \frac{4}{45}\theta^2.$$
3. On pose $U_n = \frac{1}{n}\sum_{i=1}^n 3X_i$. On a d'après le théorème central limite
$$\sqrt{n}(U_n - \theta) \xrightarrow{\mathcal{L}} N\left(0, \frac{4}{5}\theta^2\right).$$
4. Si $t < 0$, $F_{Y_n}(t) = 0$ et si $t > \theta$, $F_{Y_n}(t) = 1$. Si $0 \le t \le \theta$,
$$F_{Y_n}(t) = \prod_{i=1}^n \mathbf{P}(X_i \le t) = \prod_{i=1}^n \int_0^t \frac{1}{2\sqrt{x\theta}}\,\mathrm{d}x = \left(\frac{t}{\theta}\right)^{n/2}.$$
On déduit
$$f_{Y_n}(t) = \frac{n}{2\theta^{n/2}}t^{n/2-1}.$$
5. On a
$$\mathbf{E}[Y_n] = \frac{n}{2\theta^{n/2}}\int_0^\theta t^{n/2}\,\mathrm{d}t = \frac{n}{n+2}\theta \quad \text{et} \quad \mathbf{E}[Y_n^2] = \frac{n}{n+4}\theta^2,$$
d'où
$$\mathbf{V}[Y_n] = \frac{4n\theta^2}{(n+4)(n+2)^2}.$$
6. On a
$$\begin{aligned}\mathbf{E}[(Y_n-\theta)^2] &= \mathbf{E}[(Y_n - \mathbf{E}[Y_n] + \mathbf{E}[Y_n] - \theta)^2] \\ &= (\mathbf{E}[Y_n]-\theta)^2 + \mathbf{V}(Y_n) + 2\mathbf{E}[(Y_n - \mathbf{E}[Y_n])(\mathbf{E}[Y_n]-\theta)] \\ &= (\mathbf{E}[Y_n]-\theta)^2 + \mathbf{V}(Y_n)\end{aligned}$$
7. Il vient
$$\mathbf{E}[(Y_n-\theta)^2] = \left(\frac{n}{n+2}\theta - \theta\right)^2 + \frac{4n\theta^2}{(n+4)(n+2)^2} = \frac{4\theta^2}{(n+2)^2} + \frac{4n\theta^2}{(n+4)(n+2)^2} \to 0 \quad \text{quand} \quad n \to \infty$$
On déduit que $(Y_n)_n$ converge vers θ en moyenne quadratique.
8. Si $t < 0$, $F_{Y_n}(t) = 0$ et $F_{Y_n}(t) = 1$ si $t > n\theta$. Si $0 \le t \le n\theta$, alors
$$F_{Y_n}(t) = 1 - F_{Y_n}\left(\theta - \frac{t}{n}\right) = 1 - \left(1 - \frac{t}{n\theta}\right)^{n/2}.$$

Or

$$\left(1-\frac{t}{n\theta}\right)^{n/2} = \exp\left(\frac{n}{2}\log\left(1-\frac{t}{n\theta}\right)\right)$$
$$= \exp\left(\frac{n}{2}\left(-\frac{t}{n\theta}+\mathrm{o}\left(\frac{t}{n\theta}\right)\right)\right) \to \exp\left(-\frac{t}{2\theta}\right) \text{ quand } n\to\infty.$$

On a donc $\lim_{n\to\infty} F_{Y_n}(t) = 1-\exp(-t/2\theta)$. Par conséquent $(Z_n)_{n\in\mathbb{N}}$ converge en loi vers une loi exponentielle de paramètre $1/2\theta$.

9. Soit $\varepsilon < \theta$. On a

$$\mathbf{P}(|Y_n-\theta|>\varepsilon) = 1-\mathbf{P}(\theta-\varepsilon \leq Y_n \leq \varepsilon+\theta) = 1-\int_{\theta-\varepsilon}^{\theta} f_{Y_n}(t)\,\mathrm{d}t = \left(1-\frac{\varepsilon}{\theta}\right)^{n/2} = u_n.$$

On déduit que Y_n converge en probabilité vers θ. Comme la série de terme général u_n est convergente, on déduit d'après Borel-Cantelli que Y_n converge presque surement vers θ.

Bibliographie

Barbe P. & Ledoux M. (2007). *Probabilités.* EDP Sciences.

Carbon M. (2007). Probabilités, tome 2. Polycopié de cours, ENSAI.

Foata D. & Fuchs A. (2003). *Calcul des probabilités.* Dunod, $2^{\text{ème}}$ edn..

Fromont M. (2008). Probabilités générales. Polycopié de cours, statistique 2, IES, ENSAI.

Guyader A. (2007). Probabilités et conditionnement. Polycopié de cours, http ://www.sites.univ-rennes2.fr/laboratoire-statistique/AGUYADER/index.html.

Jacod J. & Protter P. (2003). *L'essentiel en théorie des probabilités.* Cassini.

Montfort A. (1996). *Cours de probabilités.* Economica.