

Exercices corrigés

Dominique Pastor & Christophe Sintès

Version - 1 (Mai 2014)

Table des matières

1 Aléatoire et formalisme	3
2 Variables aléatoires et moments	17
3 Aléatoire multivarié	29

Introduction

Le lecteur trouvera ici les énoncés et corrigés des exercices proposés dans

["Probabilités pour l'ingénieur, des fondements aux calculs"](#)

Certains des énoncés ci-dessous ont été modifiés par rapport à ceux de l'ouvrage pour introduire quelques précisions ou, plus simplement, corriger quelques erreurs.

Nous conseillons au lecteur de consulter ce livret d'énoncés et de corrigés régulièrement car nous proposerons de nouveaux exercices. Nous envisageons notamment quelques exercices ou problèmes où les calculs seront suivis de programmations Matlab permettant de vérifier la validité des résultats trouvés par le lecteur.

Que les lecteurs intéressés n'hésitent pas à nous contacter pour nous faire part de leurs suggestions aux adresses électroniques :

Dominique.Pastor@telecom-bretagne.eu

et

Christophe.Sintes@telecom-bretagne.eu

Nous suggérons à nos éventuels correspondants de débiter le sujet de leur courriel par l'abréviation **PPI** (probabilités pour l'ingénieur), ce qui nous permettra de mieux identifier la nature de leur courriel.

Chapitre 1

Aléatoire et formalisme

EXERCICE 1.1.– [Convergences monotone et dominée]

La convergence monotone peut s'énoncer pour toute suite croissante d'applications numériques positives ou nulles, sans préciser la fonction vers laquelle cette suite converge. En effet, si $(f_k)_{k \in \mathbb{N}}$ est une suite croissante d'applications numériques mesurables positives ou nulles, alors la limite de la suite $(f_n(x))_{n \in \mathbb{N}}$ existe dans $[0, \infty]$ pour tout $x \in \mathbb{R}$. Les notions de mesurabilité et d'intégrale s'étendent sans réelle difficulté au cas des fonctions positives ou nulles à valeurs dans $[0, \infty]$. La conclusion du théorème de convergence monotone est alors inchangée : $f = \lim_n f_n$ est mesurable et :

$$\lim_k \int_{\mathbb{R}} f_k d\lambda = \int_{\mathbb{R}} f d\lambda$$

Il faut utiliser cet énoncé plus général de la convergence monotone pour répondre aux questions suivantes.

1. Soit $(g_n)_{n \in \mathbb{N}}$ une suite d'applications numériques mesurables à valeurs dans $[0, \infty[$. Montrer que $\int_{\mathbb{R}} \sum_{n=1}^{\infty} g_n(x) dx = \sum_{n=1}^{\infty} \int_{\mathbb{R}} g_n(x) dx$.
2. Soit $(f_n)_{n \in \mathbb{N}}$ une suite d'applications numériques mesurables. On suppose que $\sum_{n=1}^{\infty} \int_{\mathbb{R}} |f_n(x)| dx < \infty$. On pose $\phi(x) = \sum_{n=1}^{\infty} |f_n(x)| \in [0, \infty]$ pour tout $x \in \mathbb{R}$.
 - (a) Montrer que $\int_{\mathbb{R}} \phi(x) dx < \infty$.
 - (b) En admettant que toute application intégrable est finie presque partout, déduire de la question précédente que $\sum_{n=1}^{\infty} f_n(x)$ converge pour presque tout réel x et que $\int_{\mathbb{R}} |f(x)| dx < \infty$ avec $f(x) = \sum_{n=1}^{\infty} f_n(x)$ en tout point x

où cette série converge et $f(x) = 0$ (par exemple) en x où la série $\sum f_n$ diverge.

(c) Montrer que $\int_{\mathbb{R}} f(x) dx = \sum_{n=1}^{\infty} \int_{\mathbb{R}} f_n(x) dx$. Ce résultat est [RUD 87, Theorem 1.38, p. 29] dans le cas réel.

Solution

1) On pose $G_N = \sum_{n=1}^N g_k$ pour tout $N \in \mathbb{N}$. Pour tout $N \in \mathbb{N}$, G_N est mesurable en tant que somme finie d'applications mesurables. De plus, pour tout $N \in \mathbb{N}$, $G_N \geq 0$. Nous sommes dans les conditions de la convergence monotone et donc : $\lim_N \int_{\mathbb{R}} G_N(x) dx = \int_{\mathbb{R}} \lim_N G_N(x) dx$. D'où le résultat, car :

$$\int_{\mathbb{R}} G_N(x) dx = \sum_{n=1}^N \int_{\mathbb{R}} g_n(x) dx$$

et

$$\lim_N \int_{\mathbb{R}} G_N(x) dx = \sum_{n=1}^{\infty} \int_{\mathbb{R}} g_n(x) dx$$

2a) Par application de la question précédente, nous avons :

$$\int_{\mathbb{R}} \phi(x) dx = \sum_{n=1}^{\infty} \int_{\mathbb{R}} |f(x)| dx < \infty$$

2b) Comme $\int_{\mathbb{R}} \phi(x) dx < \infty$, ϕ est finie presque partout. Il s'ensuit que pour presque tout x , la série $\sum f_n(x)$ est absolument convergente et donc convergente. En tout point x où cette série est absolument convergente, $|f(x)| \leq \phi(x)$ et pour tout réel x où la série $\sum f_n(x)$ diverge, $f(x) = 0$. Comme ϕ est intégrable, f est elle-aussi intégrable. Il suffit même de dire que f est majorée presque partout par la fonction intégrable ϕ — sans même avoir à préciser une quelconque valeur pour f là où elle n'est pas majorée par ϕ — pour garantir que f est intégrable.

3) Nous avons $|\sum_{n=1}^N f_n| \leq \phi$ et $\lim_n \sum_{n=1}^N f_n = f$ (presque partout). Nous sommes donc dans les conditions de la convergence dominée dans un cas plus général que celui considéré dans l'énoncé du livre puisque nous avons ici une convergence presque partout au lieu d'une convergence partout. Mais cela ne change en rien les conclusions du théorème : on peut se contenter d'inégalités et d'égalités vraies presque partout dans les énoncés de la convergence montone et dominée sans que cela n'empêche d'invertir l'intégrale et la limite. Le résultat est donc une conséquence de la convergence dominée.

Le lecteur attentif le remarquera peut-être : nous n'avons en fait pas besoin de la question précédente pour garantir l'intégrabilité de f car cette intégrabilité est directement garantie par la convergence dominée !

Les 3 exercices suivants sont des adaptations d'énoncés que le lecteur trouvera dans [KHA 94].

EXERCICE 1.2.– [Application de la convergence dominée]

Soient $a > 1$, un borélien A inclus dans $[0, \infty[$ et une application numérique f intégrable sur A : $\int_A |f(x)| dx < \infty$. Montrer que $\lim_n \int_A \frac{nx f(x)}{1+n^a x^a} dx = 0$.

Indication : justifier et utiliser le fait que, pour tout $x \in [0, \infty[$, $x \leq x^a + 1$.

Solution

Si $x \in [0, 1]$, on a $x \leq 1 \leq 1 + x^a$ car $x^a \geq 0$. Si $x > 1$, $x < x^a < x^a + 1$. Donc, pour tout $x \in [0, \infty[$, $x \leq x^a + 1$. Nous déduisons de cette inégalité que $\frac{nx}{n^a x^a + 1} \leq 1$. Aussi, nous avons l'inégalité : $|\mathbb{1}_A(x) \frac{nx f(x)}{1+n^a x^a}| = \mathbb{1}_A(x) \frac{nx |f(x)|}{1+n^a x^a} \leq |f(x)|$ puisque $A \subset [0, \infty[$. Comme f est intégrable, la suite de fonctions $(f_n)_{n \in \mathbb{N}}$ avec $f_n(x) = \mathbb{1}_A(x) \frac{nx f(x)}{1+n^a x^a}$ est dominée par la fonction intégrable f . De plus, pour tout $x \in \mathbb{R}$, $\lim_n \mathbb{1}_A(x) \frac{nx f(x)}{1+n^a x^a} = 0$. D'où le résultat par application de la convergence dominée.

EXERCICE 1.3.– [Application de la convergence dominée]

Soit $a \in]0, 1[$,

1. Montrer que $e^{-x} x^{a-1}$ est intégrable sur $[0, \infty[$;
2. Montrer que $1 + x \leq e^x$ pour tout $x \in \mathbb{R}$;
3. Montrer que pour tout $x \in [0, \infty[$, $\lim_n \left(1 - \frac{x}{n}\right)^n = e^{-x}$;
4. En déduire que $\lim_n \int_0^n \left(1 - \frac{x}{n}\right)^n x^{a-1} dx = \int_0^\infty e^{-x} x^{a-1} dx$.

Solution

1) Soit $f(x) = e^{-x} x^{a-1}$ définie pour tout $x \in]0, \infty[$. Comme $f(x) \geq 0$ pour tout $x \in]0, \infty[$, la valeur de l'intégrale $\int_0^\infty f(x) dx$ existe dans $[0, \infty]$. On cherche à montrer que cette intégrale est en fait finie.

Comme $e^{-x} \leq x^{a-1}$, nous avons :

$$f(x) \mathbb{1}_{]0,1]}(x) \leq x^{a-1} \mathbb{1}_{]0,1]}(x)$$

Pour $x \geq 1$, on a $x^{a-1} \leq 1$. Nous avons donc aussi :

$$f(x) \mathbb{1}_{[1,\infty)}(x) \leq e^{-x} \mathbb{1}_{[1,\infty)}(x)$$

Il s'ensuit que :

$$\int_0^\infty f(x) dx = \int_0^1 f(x) dx + \int_1^\infty f(x) dx \leq \int_0^1 x^{a-1} dx + \int_1^\infty e^{-x} dx \quad (1.1)$$

Le second terme du membre de droite dans l'inégalité précédente est évidemment fini en raison des propriétés de l'exponentielle. On peut même préciser la valeur de

ce terme puisqu'une primitive de e^{-x} est $-e^{-x}$. On a donc $\int_1^\infty e^{-x} dx = [-e^{-x}]_1^\infty = 1$. La première intégrale du membre de droite dans l'inégalité (1.1) est elle-aussi finie. Pour le montrer, on peut utiliser la proposition 4.15 du livre. À titre d'exemple, nous allons faire ici une démonstration spécifique au cas considéré dans cet exercice, sans passer par cette proposition, afin que le lecteur s'exerce à l'emploi de la convergence monotone que nous allons employer. Soit la suite $(g_n(x))_{n \in \mathbb{N}^*}$ des applications $g_n(x) = x^{a-1} \mathbb{1}_{[1/n, 1]}(x)$ pour $x \in]0, 1]$. Pour tout $x \in]0, 1]$, cette suite est croissante et $\lim_n g_n(x) = x^{a-1} \mathbb{1}_{]0, 1]}(x)$. Par application de la convergence monotone,

$$\int_0^1 x^{a-1} dx = \lim_n \int_{1/n}^1 x^{a-1} dx \quad (1.2)$$

L'application qui associe x^{a-1} à tout $x \in [1/n, 1]$ est continue et bornée sur $[1/n, 1]$. Elle est donc intégrable sur $[1/n, 1]$. D'autre part, une primitive de x^{a-1} est $(1/a)x^a$. Cette primitive est dérivable en tout point de l'intervalle $[1/n, 1]$. Nous sommes dans les conditions du théorème fondamental du calcul intégrale (théorème 4.10 du livre). Nous obtenons donc :

$$\int_{1/n}^1 x^{a-1} dx = \left[\frac{1}{a} x^a \right]_{1/n}^1 = \frac{1}{a} \left(1 - \frac{1}{n^a} \right)$$

En reportant ce résultat dans (1.2), nous obtenons :

$$\int_0^1 x^{a-1} dx = \lim_n \frac{1}{a} \left(1 - \frac{1}{n^a} \right) = \frac{1}{a} \quad (1.3)$$

On a donc :

$$\int_0^\infty f(x) dx \leq \frac{1}{a} + 1 \quad (1.4)$$

ce qui garantit l'intégrabilité de f .

Avec un peu d'habitude, on peut aller beaucoup plus vite en passant vite sur les détails que nous venons de donner. Mais nous avons voulu donner ces détails pour montrer comment les différents résultats de la théorie s'articulent pour établir l'intégrabilité de la fonction considérée.

2) Il y a plusieurs façons de procéder. La plus simple est de faire un dessin. Si l'on veut absolument faire des calculs, une solution classique consiste à considérer la fonction $h(x) = e^x - x - 1$ définie pour tout réel x et à étudier le sens de variation de h . On a $h'(x) = e^x - 1 \geq 0$ pour $x \geq 0$. On en déduit que h est croissante sur $[0, \infty[$. cela implique que $h(x) \geq h(0)$ pour tout $x \geq 0$ et comme $h(0) = 0$, nous obtenons le résultat voulu.

3) Nous avons $\left(1 - \frac{x}{n}\right)^n = e^{n \ln(1 - \frac{x}{n})}$. Pour n assez grand, nous pouvons écrire :

$$\ln\left(1 - \frac{x}{n}\right) = -\frac{x}{n} + \frac{x}{n} \varepsilon\left(\frac{x}{n}\right)$$

avec $\lim_{t \rightarrow 0} \varepsilon(t) = 0$. On a donc : $\left(1 - \frac{x}{n}\right)^n = e^{-x + x\varepsilon(x/n)}$, d'où le résultat.

4) Soit la suite $(f_n)_{n \in \mathbb{N}^*}$ des applications f_n définies pour tout $x \in \mathbb{R}$ par :

$$f_n(x) = \left(1 - \frac{x}{n}\right)^n x^{a-1} \mathbb{1}_{[0, n]}(x)$$

Par la question 2, $1 - \frac{x}{n} \leq e^{-x/n}$ pour $x \geq 0$. Donc, pour $x \leq n$, $\left(1 - \frac{x}{n}\right)^n \leq e^{-x}$. On a donc, pour tout $n \in \mathbb{N}^*$, $f_n(x) \leq \left(1 - \frac{x}{n}\right)^n e^{-x} \mathbb{1}_{[0, n]}(x) \leq h(x)$ avec $h(x) = e^{-x} x^{a-1} \mathbb{1}_{[0, \infty[}(x)$. Comme la fonction h est intégrable en vertu de la question 1, la suite $(f_n)_{n \in \mathbb{N}^*}$ est dominée par l'application intégrable h . D'après la question 3, la suite $(f_n)_{n \in \mathbb{N}^*}$ converge simplement vers h . Nous sommes dans les conditions d'applications du théorème de la convergence dominée. D'où le résultat.

EXERCICE 1.4.- [Une autre application de la convergence dominée]

Pour tout entier naturel $n \in \mathbb{N}^*$ et tout réel x , on pose $g_n(x) = (1 + x^2/n)^{-(n+1)/2}$.

1. Pourquoi l'intégrale $c_n = \int_{\mathbb{R}} g_n(x) dx$ existe-t-elle dans $[0, \infty[$ et pourquoi peut-on écrire que : $c_n = 2 \int_0^{\infty} g_n(x) dx$?
2. On va montrer que $c_n < \infty$ pour tout entier $n \in \mathbb{N}^*$.
 - a) Montrer que pour tout réel x : $(1 + x^2/n)^{(n+1)/2} \geq 1 + x^2/2$.
 - b) Montrer que l'application $x \in \mathbb{R} \mapsto \frac{1}{1+x^2/2}$ est intégrable.
 - c) Conclure que $c_n < \infty$ pour tout $n \in \mathbb{N}^*$.
3. On veut calculer la limite de c_n lorsque n tend vers l'infini.
 - a) Montrer que $\lim_n g_n(x) = e^{-x^2/2}$;
 - b) Déterminer la limite de la suite c_n , $n \in \mathbb{N}^*$.

Solution

1) La valeur de l'intégrale $c_n = \int_{\mathbb{R}} g_n(x) dx$ existe dans $[0, \infty[$ car $g_n \geq 0$ et est mesurable. Comme g_n est paire, on a $c_n = 2 \int_0^{\infty} g_n(x) dx$.

2a) On pose $f_n(t) = (1 + \frac{t}{n})^{\frac{n+1}{2}} - 1 - \frac{t}{2}$, $t \geq 0$

$$f'_n(t) = \frac{n+1}{2n} \left(1 + \frac{t}{n}\right)^{\frac{n-1}{2}} - \frac{1}{2} = \frac{1}{2} \left(\frac{n+1}{n} \left(1 + \frac{t}{n}\right)^{\frac{n-1}{2}} - 1\right)$$

Comme $n \geq 1$ et $t \geq 0$, on a $(1 + \frac{t}{n})^{\frac{n-1}{2}} \geq 1$, ce qui implique que $\frac{n+1}{n} (1 + \frac{t}{n})^{\frac{n-1}{2}} - 1 \geq 0$. Donc $f'_n(t) \geq 0$ pour $t \in [0, \infty[$ et f_n est croissante sur $[0, \infty[$. Comme $f_n(0) = 0$, on a $f_n(t) \geq f(0) = 0$, ce qui implique le résultat.

2b) On a :

$$\int_0^{\infty} \frac{1}{1 + \frac{x^2}{2}} dx = \int_0^1 \frac{1}{1 + \frac{x^2}{2}} dx + \int_1^{\infty} \frac{1}{1 + \frac{x^2}{2}} dx$$

L'intégrale $\int_0^1 \frac{1}{1 + \frac{x^2}{2}} dx$ est finie car $x \mapsto \frac{1}{1 + \frac{x^2}{2}}$ est définie et continue sur $[0, 1]$.

Pour $x \geq 1$, $\frac{1}{1 + \frac{x^2}{2}} \geq \frac{2}{x^2}$. Or $\int_1^\infty \frac{1}{x^2} dx = \left[-\frac{1}{x} \right]_1^\infty = 1$. Donc $\int_1^\infty \frac{1}{1 + \frac{x^2}{2}} dx < \infty$. On a donc :

$$\int_0^\infty \frac{1}{1 + \frac{x^2}{2}} dx < \infty$$

Comme $x \mapsto \frac{1}{1 + \frac{x^2}{2}}$ est paire, il s'ensuit que :

$$\int_{\mathbb{R}} \frac{1}{1 + \frac{x^2}{2}} dx < \infty$$

2c) $g_n(x) = (1 + \frac{x^2}{n})^{-\frac{n+1}{2}} \leq \frac{1}{1 + \frac{x^2}{2}}$ et l'application $x \mapsto \frac{1}{1 + \frac{x^2}{2}}$ est intégrable sur $[0, \infty[$.

Donc $\int_0^\infty g_n(x) dx \leq \infty$ et c_n est fini aussi.

3a) $\ln g_n(x) = -\frac{n+1}{2} \ln \left(1 + \frac{x^2}{n} \right)$ donc $\lim_n \ln g_n(x) = -\frac{x^2}{2}$ car

$$\ln \left(1 + \frac{x^2}{n} \right) = \frac{x^2}{n} + \frac{x^2}{n} o \left(\frac{x^2}{n} \right)$$

On en déduit que $\lim_n g_n(x) = e^{-x^2/2}$ pour tout $x \in \mathbb{R}$.

3b) On sait que $g_n(x) \leq \frac{1}{1 + \frac{x^2}{2}}$, qui est intégrable sur \mathbb{R} . D'autre part, $\lim_n g_n(x) = e^{-x^2/2}$ pour tout $x \in \mathbb{R}$. Le théorème de la convergence dominée s'applique. On en déduit :

$$\lim_n c_n = \lim_n \int_{\mathbb{R}} g_n(x) dx = \int_{\mathbb{R}} \lim_n g(x) dx = \int_{\mathbb{R}} e^{-x^2/2} dx = \sqrt{2\pi}$$

La dernière égalité étant une conséquence de la condition de normalisation, on a :

$$\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$$

pour la densité de probabilité gaussienne de moyenne nulle et de variance unitaire.

EXERCICE 1.5.– $[\sin(x)/x$ n'est pas intégrable]

1. Montrer que pour tout $k \in \mathbb{N}$: $\int_{k\pi}^{(k+1)\pi} \frac{|\sin x|}{x} dx \geq \frac{2}{(k+1)\pi}$.
2. En déduire que l'application $x \rightarrow \sin(x)/x$ n'est pas intégrable.
3. Pourquoi peut-on garantir l'existence de l'intégrale $\int_{\pi}^a \frac{\sin x}{x} dx$ pour $a > \pi$?

4. À l'aide d'une intégration par parties, montrer l'égalité :

$$\int_{\pi}^a \frac{\sin x}{x} dx = \alpha \frac{\cos a}{a} + \beta \frac{1}{\pi} + \gamma \int_{\pi}^a \frac{\cos x}{x^2} dx$$

pour $a > \pi$ et où α , β et γ sont trois constantes que l'on déterminera.

5. Dédurre de ce qui précède que l'application $x \rightarrow \sin(x)/x$ est semi-intégrable dans le sens où :

$$\lim_{a \rightarrow \infty} \int_0^a \frac{\sin x}{x} dx < \infty$$

Solution

1) Faisons le changement de variable suivant $t \Leftrightarrow x - k\pi$:

$$\int_{k\pi}^{(k+1)\pi} \frac{|\sin x|}{x} dx = \int_0^{\pi} \frac{|\sin(x+k\pi)|}{x+k\pi} dx = \int_0^{\pi} \frac{|\sin(x)|}{x+k\pi} dx = \int_0^{\pi} \frac{\sin(t)}{t+k\pi} dx$$

Pour $x \leq \pi$, $\frac{1}{x+k\pi} \geq \frac{1}{(k+1)\pi}$. On a donc :

$$\int_{k\pi}^{(k+1)\pi} \frac{|\sin x|}{x} dx \geq \frac{1}{(k+1)\pi} \int_0^{\pi} \sin x dx \geq \frac{2}{(k+1)\pi}$$

2) Si $\frac{\sin x}{x}$ était intégrable, alors on aurait :

$$\int_0^{\infty} \frac{\sin x}{x} dx = \sum_{k=0}^{\infty} \int_{k\pi}^{(k+1)\pi} \frac{\sin x}{x} dx$$

par application de la convergence monotone (cf. Exercice 1.1). Comme

$$\int_{k\pi}^{(k+1)\pi} \frac{|\sin x|}{x} dx \geq \frac{2}{(k+1)\pi}$$

et que la série $\sum_{k=1}^{\infty} \frac{2}{(k+1)\pi}$ diverge, on a $\int_0^{\infty} \frac{|\sin x|}{x} dx = \infty$.

3) $\int_{\pi}^a \frac{\sin x}{x} dx$ existe dans \mathbb{R} car $\frac{\sin x}{x}$ est définie et continue sur l'intervalle $[\pi, a]$.

4) L'intégration par partie donne $\alpha = \beta = \gamma = 1$.

5) $\lim_{a \rightarrow \infty} \frac{\cos a}{a} = 0$ et $\frac{\cos x}{x^2} \mathbb{1}_{[\pi, \infty[}(x)$ est intégrable. Donc $\lim_{a \rightarrow \infty} \int_{\pi}^a \frac{\sin x}{x} dx$ existe dans \mathbb{R} . Nous avons aussi $\int_0^{\pi} \frac{\sin x}{x} dx < \infty$. D'où le résultat.

EXERCICE 1.6.- [Fubini ne marche pas toujours]

Soit la fonction à deux variables définie par $f(x, y) = \frac{x^2 - y^2}{(x^2 + y^2)^2}$ pour $x, y \in [0, 1]$.

1. Montrer que $\frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right) = -f(x, y)$.
2. Calculer $\int_0^1 \left(\int_0^1 f(x, y) dx \right) dy$ et $\int_0^1 \left(\int_0^1 f(x, y) dy \right) dx$.
3. En déduire que f n'est pas intégrable dans \mathbb{R}^2 .

Solution

1) On a : $\frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right) = \frac{x^2 + y^2 - 2x^2}{x^2 + y^2} = -\frac{x^2 - y^2}{x^2 + y^2} = -f(x, y)$.

2) Nous pouvons écrire : $\int_0^1 f(x, y) dx = -\int_0^1 \frac{\partial}{\partial x} \left(\frac{x}{x^2 + y^2} \right) dx = -\left[\frac{x}{x^2 + y^2} \right]_0^1 = -\frac{1}{1 + y^2}$.

On en déduit :

$$\begin{aligned} \int_0^1 \left(\int_0^1 f(x, y) dx \right) dy &= -\int_0^1 \frac{1}{1 + y^2} dy \\ &= -\int_0^{\pi/4} \frac{1}{1 + \tan^2 \theta} \frac{1}{\cos^2 \theta} d\theta \\ &\quad \text{[Avec le changement de variable } y = \tan(\theta)\text{]} \\ &= -\int_0^{\pi/4} d\theta = -\frac{\pi}{4} \end{aligned}$$

D'autre part, on a :

$$\int_0^1 f(x, y) dy = \int_0^1 \frac{\partial}{\partial y} \left(\frac{y}{x^2 + y^2} \right) dy = \left[\frac{y}{x^2 + y^2} \right]_0^1 = \frac{1}{x^2 + 1}$$

Par un calcul analogue au précédent, on en déduit l'égalité :

$$\int_0^1 \left(\int_0^1 f(x, y) dy \right) dx = \int_0^1 \frac{1}{x^2 + 1} dx = \frac{\pi}{4}$$

3) f n'est pas intégrable sur \mathbb{R}^2 , car si f l'était alors le théorème de Fubini s'appliquerait et on pourrait changer l'ordre d'intégration sans changer la valeur du résultat.

EXERCICE 1.7.- [Théorème de Scheffé]

L'objet de cet exercice est de démontrer le théorème de Scheffé suivant :

Si $(f_n)_{n \in \mathbb{N}}$ est une suite de densités de probabilité et f une densité de probabilité telle que $\lim_n f_n(x) = f(x)$ pour presque tout réel x , alors $\lim_n \mathbb{F}_n(x) = \mathbb{F}(x)$ pour tout $x \in \mathbb{R}$, où \mathbb{F}_n est la fonction de répartition associée à f_n et \mathbb{F} , la fonction de répartition associée à f .

Démonstration : soit une suite $(f_n)_{n \in \mathbb{N}}$ de densités de probabilité et f est une densité de probabilité telle que $\lim_n f_n = f$ (p.p) où λ est la mesure de Lebesgue sur \mathbb{R} .

1. Montrer que $\int (f - f_n)^+ d\lambda = \int (f - f_n)^- d\lambda = \frac{1}{2} \int |f - f_n| d\lambda$;
2. Montrer que $\lim_n (f - f_n)^+ = 0$ λ -p.p. et que $0 \leq (f - f_n)^+(x) \leq f(x)$ en tout $x \in \mathbb{R}$ et en déduire la valeur de $\lim_n \int (f - f_n)^+ d\lambda$;
3. Dédire des questions précédentes que pour tout $x \in \mathbb{R}$, $\lim_{n \rightarrow \infty} F_n(x) = F(x)$.

Solution

1) f et f_n sont intégrables donc $f - f_n$ est intégrable. Etant donné que f et f_n sont des densités, alors $\int_{\mathbb{R}} f(x) dx = \int_{\mathbb{R}} f_n(x) dx$, donc $\int_{\mathbb{R}} (f(x) - f_n(x)) dx = 0$.

2) On a :

$$\int_{\mathbb{R}} (f(x) - f_n(x)) dx = \int_{\mathbb{R}} (f - f_n)^+(x) dx - \int_{\mathbb{R}} (f - f_n)^-(x) dx$$

Or $\int_{\mathbb{R}} (f(x) - f_n(x)) dx = 0$. Donc : $\int_{\mathbb{R}} (f - f_n)^+(x) dx = \int_{\mathbb{R}} (f - f_n)^-(x) dx$ et

$$\int_{\mathbb{R}} |f(x) - f_n(x)| dx = \int_{\mathbb{R}} (f - f_n)^+(x) dx + \int_{\mathbb{R}} (f - f_n)^-(x) dx = 2 \int_{\mathbb{R}} (f(x) - f_n(x))^+ dx$$

3) Comme $(f - f_n)^+ = (f - f_n) \mathbb{1}_{\{f - f_n \geq 0\}}$, on a $|f - f_n|^+ \leq |f - f_n|$. Par hypothèse, $f_n(x) \rightarrow f(x)$ pour presque tout $x \in \mathbb{R}$. Donc $\lim_n |f(x) - f_n(x)| = 0$ pour presque tout $x \in \mathbb{R}$, ce qui implique que $\lim_n (f - f_n)^+(x) = 0$ pour presque tout $x \in \mathbb{R}$.

Pour $x \in \mathbb{R}$ tel que $f(x) < f_n(x)$, on a $(f - f_n)^+(x) = 0$. Comme f est une densité, on a $f(x) \geq 0$ et donc $f(x) \geq (f - f_n)^+(x)$. Si $x \in \mathbb{R}$ est tel que $f_n(x) \leq f(x)$, on a $(f - f_n)^+(x) = f(x) - f_n(x) \leq f(x)$ car $f_n(x) \geq 0$. Au final, on obtient $(f - f_n)^+(x) \leq f(x)$ pour tout $x \in \mathbb{R}$.

Puisque $\lim_n (f - f_n)^+(x) = 0$ pour presque tout $x \in \mathbb{R}$ et que $0 \leq (f - f_n)^+ \leq f$ partout (preque partout aurait suffit), le théorème de la convergence dominée implique que $\lim_n \int_{\mathbb{R}} (f - f_n)^+ dx = 0$.

4) $F(x) - F_n(x) = \int_{-\infty}^x f(x) dx - \int_{-\infty}^x f_n(x) dx = \int_{-\infty}^x (f - f_n)(x) dx$. Cette égalité implique que $|F(x) - F_n(x)| \leq \int_{-\infty}^x |f(x) - f_n(x)| dx \leq \int_{\mathbb{R}} |f(x) - f_n(x)| dx$. D'après la question 2, on a alors : $|F(x) - F_n(x)| \leq 2 \int_{\mathbb{R}} (f(x) - f_n(x))^+ dx$. D'après la question 3,

$$\lim_n \int_{\mathbb{R}} (f(x) - f_n(x))^+ dx = 0$$

D'où le résultat.

EXERCICE 1.8.– [La réciproque du théorème de Scheffé est fausse]

Pour tout $n \in \mathbb{N}$, on définit :

$$f_n(x) = \begin{cases} 1 - \cos(2\pi nx) & \text{pour } 0 \leq x \leq 1 \\ 0 & \text{ailleurs} \end{cases}$$

1. Vérifier que f_n est une densité de probabilité.
2. Pour tout $n \in \mathbb{N}$, calculer la fonction de répartition \mathbb{F}_n associée à f_n .
3. Montrer que pour tout $x \in \mathbb{R}$, $\mathbb{F}_n(x)$ tend vers la fonction de répartition de la loi uniforme sur $[0, 1]$.
4. Pour $x \in \mathbb{R}$, est-ce que $f_n(x)$ converge lorsque n tend vers l'infini ? Conclure.

Solution

1) Soit $f_n(x) = (1 - \cos(2\pi nx))\mathbb{1}_{]0,1[}(x)$. Nous avons :

$$\int_{\mathbb{R}} f_n(x) dx = \int_0^1 (1 - \cos(2\pi nx)) dx = 1 - \int_0^1 \cos(2\pi nx) dx = 1 - \frac{1}{2\pi n} [\sin(2\pi nx)]_0^1 = 1$$

2) Pour $x \leq 0$, $\mathbb{F}_n(x) = 0$. Pour $x \in [0, 1]$,

$$\mathbb{F}_n(x) = \int_0^x (1 - \cos(2\pi nt)) dt = x - \int_0^x \cos(2\pi nt) dt = x - \frac{1}{2\pi n} \sin(2\pi nx)$$

et pour $x \geq 1$, $\mathbb{F}_n(x) = 1$.

3) On a :

$$\lim_{n \rightarrow \infty} \mathbb{F}_n(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } x \in [0, 1] \\ 1 & \text{si } x \geq 1 \end{cases}$$

D'où le résultat.

4) Lorsque n tend vers l'infini, $f_n(x)$ ne converge pas à cause de la fonction $\cos(2\pi nx)$ alors que la suite des fonctions de répartition $(\mathbb{F}_n)_{n \in \mathbb{N}}$ converge vers la fonction de répartition de la loi uniforme. Ainsi la réciproque du théorème de Scheffé n'est pas vraie.

EXERCICE 1.9.– [Expression ensembliste des événements]

Soient un espace probabilisable (Ω, \mathcal{B}) . Soit A , B , et C trois événements quelconques de \mathcal{B} . Déterminer les ensembles mesurables de \mathcal{B} qui permettent d'exprimer les propositions ou événements suivants :

1. Le résultat ω de l'expérience est dans A mais n'est ni dans B ni dans C .
2. Le résultat ω de l'expérience est dans deux et seulement deux des ensembles A , B et C .
3. Le résultat ω de l'expérience est au plus dans deux des ensembles A , B et C .

Solution

1) Si on considère l'événement A uniquement ce la signifie que B et C ne se sont pas produit donc on écrit $A \cap B^c \cap C^c$.

2) Si 2 événements au plus, cela signifie exactement que un ou deux événements se produisent. L'événement cherché est :

$$\begin{aligned} & ((A \cap B^c \cap C^c) \cup (B \cap A^c \cap C^c) \cup (C \cap B^c \cap A^c)) \quad \text{[pour un événement]} \\ & \cup \\ & ((A \cap B \cap C^c) \cup (C \cap B \cap A^c) \cup (A \cap C \cap B^c)) \quad \text{[pour deux événements]} \end{aligned}$$

3) D'après la question précédente, l'événement cherché est :

$$((A \cap B \cap C^c) \cup (C \cap B \cap A^c) \cup (A \cap C \cap B^c))$$

EXERCICE 1.10.– [Modèles probabilistes discrets et dénombrement]

On considère le jeu de bridge. Le jeu contient 52 cartes et chaque joueur reçoit 13 cartes (une main).

1. Combien de mains différentes peut recevoir un joueur ?
2. Probabilité qu'il reçoive un as exactement ?
3. Probabilité qu'il reçoive au moins un as ?
4. Probabilité qu'il reçoive au moins un as ou un roi, le "ou" n'étant pas exclusif ?
5. Probabilité qu'il lui manque au moins une couleur ?

Solution

1) Le nombre de main différentes que peut recevoir un joueur correspond au nombre de parties de cardinal 13 qui existe dans $\Omega =$ aux 52 cartes du jeu, soit C_{52}^{13} .

2) L'événement d'avoir exactement un as dans une main de 13 cartes correspond à tirer un as parmi les 4 disponibles et tirer 12 cartes parmi toutes les cartes disponibles moins celles correspondant aux 4 as. Autrement dit il existe $C_4^1 C_{48}^{12}$ mains possibles soit une probabilité de $P = C_4^1 C_{48}^{12} / C_{52}^{13}$.

3) La probabilité d'avoir un as est associé aux événements d'avoir un ou plusieurs as (c'est dire un as uniquement ou 2 as uniquement ou 3 as uniquement ou 4 as uniquement). On peut raisonner plus simplement en considérant l'événement complémentaire qui est de n'avoir aucun. Le nombre de mains associées à cet événement est C_{48}^{13} . Ainsi, la probabilité d'avoir au moins un as est $P = 1 - C_{48}^{13} / C_{52}^{13}$.

4) La probabilité d'avoir au moins un as et/ou un roi pourrait se décomposer en trois groupes d'événements : avoir au moins un as, ou au moins un roi, ou au moins un as et au moins un roi. La difficulté de cette approche concerne le fait de comptabiliser

plusieurs fois certains événements. Par exemple : avoir au moins un as n'exclut pas le fait d'avoir un roi, etc. Ainsi, il est plus simple de considérer l'événement complémentaire : avoir une main ne comptant ni roi ni as. Le nombre de mains de la sorte est le nombre de choix de 13 cartes parmi $52 - 4 - 4 = 44$. Ce nombre de mains vaut donc C_{44}^{13} . La probabilité cherchée vaut $P = 1 - C_{44}^{13}/C_{52}^{13}$.

5) La probabilité de n'avoir que 3 couleurs est simplement donnée par le choix de ne tirer que 13 cartes parmi les $53 - 13 = 39$ restantes, soit $P = C_{39}^{13}/C_{52}^{13}$.

EXERCICE 1.11.- [Le facteur très fatigué]

Dans un immeuble, un facteur est chargé de distribuer dans p boîtes le courrier constitué de N lettres, dont r_i pour la boîte i . On suppose $N > p$. Il le fait au hasard. On suppose que $N = \sum_{i=1}^p r_i$ et que le facteur extirpe une à une chaque lettre du sac.

1. Modéliser l'expérience aléatoire.
2. Probabilité que la première boîte soit correctement remplie ?
3. Probabilité qu'elle ne contienne aucune lettre des voisins ?
4. Probabilité qu'elle contienne exactement k lettres ($0 \leq k \leq N$) ?
5. Probabilité que chaque boîte contienne au moins une lettre qui lui est destinée ?

Solution

1) La distribution effectuée par le facteur est une application $f : L \rightarrow B$ de L dans B où L est l'ensemble $\{\ell_1, \ell_2, \dots, \ell_N\}$ des N lettres à distribuer et $B = \{b_1, b_2, \dots, b_p\}$ est l'ensemble des p boîtes. L'espace Ω des états ou possibles est donc l'ensemble B^L des applications de L dans B . La tribu que l'on choisit pour modéliser cette expérience aléatoire est l'ensemble des parties de Ω . La mesure de probabilité de $A \subset \Omega$ est $P(A) = \text{card}A / \text{card}\Omega$. On a $\text{card}\Omega = \text{card}B^L = \text{card}B^{\text{card}L} = p^N$.

2) La boîte 1 est correctement remplie si et seulement si l'application f est telle que $f(L_1) = b_1$ où L_1 est l'ensemble des lettres à placer dans la boîte b_1 . Le sous-ensemble de Ω concerné est donc $F_1 = \{f \in B^L : f(L_1) = b_1\}$. Cet ensemble est isomorphe à $\{b_1\}^{L_1} \times (B \setminus \{b_1\})^{(L \setminus L_1)}$, qui est le produit cartésien entre les applications de L_1 dans $\{b_1\}$ (il n'y en a qu'une) et l'ensemble des applications de $L \setminus L_1$ dans $B \setminus \{b_1\}$. Le cardinal de F_1 est alors :

$$\begin{aligned} \text{card}(\{b_1\}^{L_1}) \times \text{card}((B \setminus \{b_1\})^{(L \setminus L_1)}) &= \text{card}(\{b_1\}^{L_1}) \times \text{card}(B \setminus \{b_1\})^{\text{card}(L \setminus L_1)} \\ &= 1^{r_1} \times (p-1)^{N-r_1}. \end{aligned}$$

La probabilité cherchée est donc égale à $(p-1)^{N-r_1} / p^N$.

3) La boîte b_1 ne contient aucune lettre des voisins si l'application f prend des valeurs quelconques sur L_1 et est telle que $f(L \setminus L_1) = B \setminus \{b_1\}$. Le sous-ensemble de Ω concerné est donc

$$F_2 = \{f \in B^L : f(L_1) = B \text{ \& } f(L \setminus L_1) = B \setminus \{b_1\}\}.$$

Cet ensemble est isomorphe à $B^{L_1} \times (B \setminus \{b_1\})^{(L \setminus L_1)}$, qui est le produit cartésien entre l'ensemble des applications de L_1 dans $B \setminus \{b_1\}$ et l'ensemble des applications de $L \setminus L_1$ dans $B \setminus \{b_1\}$. Le cardinal de F_2 est alors $(p-1)^{N-r_1} p^{r_1}$: il y a p^{r_1} applications de L_1 dans B et $(p-1)^{N-r_1}$ applications de $L \setminus L_1$ dans $B \setminus \{b_1\}$. La probabilité cherchée est donc $(p-1)^{N-r_1} p^{r_1} / p^N$.

4) Il y a C_N^k sous-ensembles de k lettres dans L avec $k = 1, 2, \dots, N$. Une application qui associe k lettres à la boîte b_1 est une application qui associe tous les éléments d'un sous-ensemble A de k lettres à b_1 et dont la restriction à $L \setminus A$ est une application à valeurs dans $B \setminus \{b_1\}$. Il y a $C_N^k (p-1)^{N-k}$ applications de cette forme. La probabilité cherchée est donc $C_N^k (p-1)^{N-k} / p^N$.

5) Soit A_i , l'ensemble des applications $f : L \rightarrow B$ telles que $f(L_i) = B \setminus \{b_i\}$ où L_i est l'ensemble des lettres qui devraient être placées dans b_i . L'événement cherché est $A_1^c \cap A_2^c \cap \dots \cap A_p^c$ où A_k^c est le complémentaire de A_k dans $\Omega = B^L$. Comme $A_1^c \cap A_2^c \cap \dots \cap A_p^c = (A_1 \cup A_2 \cup \dots \cup A_p)^c$, la probabilité cherchée est $1 - \mathbb{P}(A_1 \cup A_2 \cup \dots \cup A_p)$. On applique maintenant la formule de Poincaré :

$$\mathbb{P}(A_1 \cup A_2 \cup \dots \cup A_p) = \sum_{k=1}^n (-1)^{k+1} \sum_{i_1 < i_2 < \dots < i_k} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k})$$

où $\sum_{i_1 < i_2 < \dots < i_k}$ est la somme pour tous les k -uplets (i_1, i_2, \dots, i_k) tels que $1 \leq i_1 < i_2 < \dots < i_k \leq N$. Il faut donc calculer $\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k})$. L'ensemble $A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}$ est celui de toutes les applications qui ne prennent pas leurs valeurs dans $\{b_{i_1}, b_{i_2}, \dots, b_{i_k}\}$. Cet ensemble est donc celui des applications de L dans B qui prennent leurs valeurs dans $B \setminus \{b_{i_1}, b_{i_2}, \dots, b_{i_k}\}$. Il y en a donc :

$$\text{card}\left((B \setminus \{b_{i_1}, b_{i_2}, \dots, b_{i_k}\})^L\right) = \text{card}(B \setminus \{b_{i_1}, b_{i_2}, \dots, b_{i_k}\})^{\text{card}(L)} = (p-k)^N$$

On a donc $\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = (p-k)^N / p^N = \left(1 - \frac{k}{p}\right)^N$. Comme il y a C_p^k k -uplets possibles (i_1, i_2, \dots, i_p) avec $1 \leq i_1 < i_2 < \dots < i_k \leq N$, nous avons :

$$\mathbb{P}(A_1 \cup A_2 \cup \dots \cup A_p) = \sum_{k=1}^n (-1)^{k+1} C_p^k \left(1 - \frac{k}{p}\right)^N$$

et la probabilité cherchée est :

$$1 - \sum_{k=1}^n (-1)^{k+1} C_p^k \left(1 - \frac{k}{p}\right)^N$$

Chapitre 2

Variables aléatoires et moments

EXERCICE 2.1.– [Variable aléatoire discrète et modulo]

Soient $\Omega = \mathbb{N}^*$, \mathcal{B} est l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω et $\mathbb{P}(\{k\}) = 2^{-k}$.

1. Montrer que \mathbb{P} vérifie la condition de normalisation, de sorte que $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ est un espace probabilisé.
2. Pour un entier $n \in \mathbb{N}^*$ donné, on définit la variable aléatoire X comme le reste modulo n : $X(k) = k \bmod n$, $n \in \mathbb{N}^*$. Donner la loi de X : domaine de variation, probabilité d'un singleton et fonction de répartition.

Solution

1) $\mathbb{P}(\Omega) = \mathbb{P}\left(\bigcup_{k=1}^{+\infty} \{k\}\right) = \sum_{k=1}^{+\infty} \mathbb{P}\{k\} = \sum_{k=1}^{+\infty} 2^{-k} = \frac{1}{2} \frac{1}{1-1/2} = 1$. On a $\mathbb{P}(\emptyset) = 1 - \mathbb{P}(\Omega) = 0$.

Comme tout événement de \mathcal{B} peut s'écrire comme une union dénombrable d'événements élémentaires du type $\{k\}$, on a aussi :

$$\forall A \in \mathcal{B}, \quad \exists I \subset \mathbb{N}^*, \quad A = \bigcup_{k \in I} \{k\}$$
$$\mathbb{P}(A) = \sum_{k \in I} \mathbb{P}\{k\}$$

2) La variable aléatoire X est définie pour tout $k \in \Omega = \mathbb{N}^*$ par $X(k) = k \bmod n$. La variable aléatoire X prend ses valeurs dans $\{0, 1, 2, 3, \dots, n-1\}$, qui est son domaine de variation. On a :

$$\mathbb{P}(X^{-1}(\{0\})) = \mathbb{P}\left(\bigcup_{k \geq 1} \{kn\}\right) = \sum_{k \geq 1} \mathbb{P}(\{kn\}) = \sum_{k \geq 1} 2^{-kn} = \frac{1}{2^n - 1}$$

Pour $1 \leq j \leq n-1$, on a :

$$\mathbb{P}(X^{-1}(\{j\})) = \mathbb{P}\left(\bigcup_{k \geq 0} \{kn + j\}\right) = \sum_{k \geq 0} \mathbb{P}(\{kn + j\}) = \sum_{k \geq 0} 2^{-kn+j} = 2^{-j} \frac{1}{1-2^{-n}} = \frac{2^{n-j}}{2^n - 1}$$

On déduit de ces résultats l'expression de la fonction de répartition. Comme $\mathbb{P}[X < 0] = 0$, on a $\mathbb{F}_X(x) = 0$ pour $x \in]-\infty, 0[$. On a aussi $\mathbb{F}_X(0) = \mathbb{P}[X \leq 0] = \mathbb{P}[X = 0] = \frac{2^{-n}}{1-2^{-n}} = \frac{1}{2^n-1}$.

Pour $x \in [j, j+1[$, on obtient :

$$\begin{aligned} \mathbb{F}_X(x) &= \mathbb{P}[X \leq x] = \mathbb{P}_X(\{0\} \cup \{1\} \cup \dots \cup \{j\}) \\ &= \frac{1}{2^n-1} + \sum_{k=1}^j \frac{2^{n-k}}{2^n-1} = \frac{1+2^n-2^{n-j}}{2^n-1} \end{aligned}$$

EXERCICE 2.2.- [Variable aléatoire de densité triangulaire]

Soit X une variable aléatoire à valeurs dans $[-1, 1]$, dont le graphe de la densité de probabilité f_X forme un triangle isocèle (segment sur l'axe des $x =$ hypoténuse).

1. Déterminer l'expression de f_X . En déduire celle de la fonction de répartition \mathbb{F}_X .
2. Calculer l'espérance mathématique et la variance de X .
3. Calculer $\mathbb{P}[X \leq \frac{1}{4}, X^2 \leq \frac{1}{4}]$.
4. Déterminer la distribution de $Y = X^2$. Calculer la moyenne de Y . Retrouver ainsi la variance de X .

Solution

La densité de probabilité est un triangle isocèle donc f_X est paire.

1) f_X est une densité de probabilité. On utilise la condition de normalisation pour déterminer a . On a :

$$\int_{\mathbb{R}} f_X(x) dx = 1 \Rightarrow \int_{-1}^0 f_X(x) dx + \int_0^1 f_X(x) dx = 1 \Rightarrow 2 \int_0^1 f_X(x) dx = 1.$$

Or, $2 \int_0^1 f_X(x) dx$ est l'aire du triangle sous la courbe de f_X . Cette aire est égale à $a \times 1 =$

$$a. \text{ On a donc } a = 1 \text{ et } f_X(x) = \begin{cases} 1-x & \text{si } x \in [0, 1] \\ 1+x & \text{si } x \in [-1, 0] \\ 0 & \text{ailleurs} \end{cases}$$

2) On calcule la fonction de répartition de X par intégration de la densité. On a donc :

$$\begin{aligned}
 x \in]-\infty, -1[, \quad \mathbb{F}_X(x) &= 0 \\
 x \in [-1, 0], \quad \mathbb{F}_X(x) &= \int_{-\infty}^x f_X(t) dt = \int_{-1}^x (1+t) dt \\
 &= \left[\frac{(1+t)^2}{2} \right]_{-1}^x \\
 &= \frac{(1+x)^2}{2} \\
 x \in [0, 1], \quad \mathbb{F}_X(x) &= \int_{-\infty}^x f_X(t) dt \\
 &= \int_{-1}^0 f_X(t) dt + \int_0^x f_X(t) dt \\
 &= \frac{1}{2} + \left[-\frac{(1-t)^2}{2} \right]_0^x \\
 &= 1 - \frac{(1-x)^2}{2} \\
 x \leq 1, \quad \mathbb{F}_X(x) &= \int_{-\infty}^1 f_X(t) dt = \mathbb{F}_X(1) \\
 &= 1
 \end{aligned}$$

3) Comme f_X est paire :

$$\begin{aligned}
 \mathbb{E}[X] &= \int_{\mathbb{R}} x f_X(x) dx = \int_{-1}^0 x f_X(x) dx + \int_0^1 x f_X(x) dx \\
 &= -\int_0^1 x f_X(x) dx + \int_0^1 x f_X(x) dx \\
 &= 0
 \end{aligned}$$

La variance de X est $\text{Var}\{X\} = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. Par conséquent :

$$\begin{aligned} \mathbb{E}[X^2] &= \int_{\mathbb{R}} x^2 f_X(x) dx = \int_{-1}^0 x^2 f_X(x) dx + \int_0^1 x^2 f_X(x) dx \\ &= 2 \int_0^1 x^2 f_X(x) dx \\ &= 2 \int_0^1 (x^2 - x^3) dx \\ &= 2 \left(\frac{1}{3} - \frac{1}{4} \right) \\ &= \frac{1}{6} \end{aligned}$$

4)

$$\begin{aligned} \mathbb{P}\left[X \leq \frac{1}{4}, X^2 \leq \frac{1}{4}\right] &= \mathbb{P}\left[X \leq \frac{1}{4}, -\frac{1}{2} \leq X \leq \frac{1}{2}\right] \\ &= \mathbb{P}\left[-\frac{1}{2} \leq X \leq \frac{1}{4}\right] \\ &= \int_{-\frac{1}{2}}^{\frac{1}{4}} f_X(x) dx \\ &= \mathbb{F}_X(1/4) - \mathbb{F}_X(-1/2) = \left(1 - \frac{9}{32} - \frac{4}{32}\right) = \frac{19}{32} \end{aligned}$$

5) Loi de $Y = X^2$ La variable aléatoire Y est à valeurs dans $[0, 1]$. Pour tout $y \geq 0$,

$$\mathbb{P}[Y \leq y] = \mathbb{P}[X^2 \leq y] = \mathbb{P}(X \leq \sqrt{y}, X \geq -\sqrt{y}) = \mathbb{P}(X \leq \sqrt{y}) - \mathbb{P}(X \leq -\sqrt{y}).$$

On a donc $\mathbb{F}_Y(y) = \mathbb{F}_X(\sqrt{y}) - \mathbb{F}_X(-\sqrt{y})$ pour tout $y \geq 0$.

Par dérivation, il vient : $f_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}) = \frac{1}{\sqrt{y}} f_X(\sqrt{y})$. On prend maintenant en compte l'expression spécifique de f_X pour aboutir à :

$$y \in]0, 1]: \quad f_Y(y) = \frac{1 - \sqrt{y}}{\sqrt{y}}$$

$$\text{ailleurs} \quad f_Y(y) = 0$$

Avec la densité de Y on retrouve la valeur de la variance de X^2 . On a :

$$\mathbb{E}[X^2] = \mathbb{E}[Y] = \int_{\mathbb{R}} y f_Y(y) dy = \int_0^1 \sqrt{y}(1 - \sqrt{y}) dy = \left[\frac{2}{3} y^{3/2} - \frac{1}{2} y^2 \right]_0^1 = \frac{1}{6}$$

Comme $\mathbb{E}[X] = 0$, on retrouve $\text{Var}\{X\} = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] = 1/6$.

EXERCICE 2.3.– [Fonction de répartition et génération de loi]

Soit une variable aléatoire X de fonction de répartition \mathbb{F}_X supposée bijective. Montrer que $U = \mathbb{F}_X(X)$ suit une loi uniforme sur un intervalle à déterminer et que \mathbb{F}_X est la fonction de répartition de $\mathbb{F}_X^{-1}(U)$.

Solution

\mathbb{F}_X est bijective sur l'intervalle $[0, 1]$ et strictement croissante par hypothèse. Donc, si on pose $U = \mathbb{F}_X(X)$, U prend ses valeurs dans $[0, 1]$. Soit $u \in [0, 1]$.

$$\mathbb{P}[U \leq u] = \mathbb{P}[\mathbb{F}_X(X) \leq u] = \mathbb{P}[X \leq \mathbb{F}_X^{-1}(u)] = \mathbb{F}_X(\mathbb{F}_X^{-1}(u)) = u$$

Donc $U \sim \mathcal{U}[0, 1]$ (loi uniforme).

Soit alors $Y = \mathbb{F}_X^{-1}(U)$. Il est immédiat que l'on a $\mathbb{P}[Y \leq y] = \mathbb{P}[U \leq \mathbb{F}_X(y)] = \mathbb{F}_X(y)$. La variable aléatoire suit donc la loi \mathbb{F}_X . On a donc généré une variable aléatoire dont la loi est \mathbb{F}_X .

EXERCICE 2.4.– [Changement de variables en coordonnées polaires]

Soit $U \sim \mathcal{U}[0, 1[$.

1. Montrer que U et $1 - U$ suivent la même loi.
2. Quelle est la loi suivie par $V = -\ln(U)$ et $W = -\ln(1 - U)$?
3. Montrer que $R = \sqrt{-\ln(U)}$ suit une loi de Rayleigh $\mathcal{R}(\sigma)$ de paramètre σ à déterminer. On rappelle qu'une variable aléatoire suit une loi de Rayleigh $\mathcal{R}(\sigma)$ si sa densité est de la forme : $f_R(r) = (r/\sigma^2)e^{-r^2/2\sigma^2} \mathbb{1}_{]0, \infty[}(r)$.
4. Soient $X = R \cos \Phi$ et $Y = R \sin \Phi$ où $R \sim \mathcal{R}(\sigma)$ et $\Phi \sim \mathcal{U}[0, 2\pi[$ sont indépendantes. Calculer la loi conjointe de (X, Y) . En déduire une méthode de génération de deux variables aléatoires gaussiennes indépendantes.

Solution

1) On a $\mathbb{P}[1 - U \leq u] = \mathbb{P}[U \geq 1 - u] = 1 - \mathbb{F}_U(1 - u)$. Comme $U \sim \mathcal{U}[0, 1[$, $\mathbb{F}_U(u) = u$ pour $u \in]0, 1[$, $\mathbb{F}_U(u) = 0$ pour $u \leq 0$ et $\mathbb{F}_U(u) = 1$ pour $u \geq 1$, où \mathbb{F}_U est la fonction de répartition de U . On a donc $\mathbb{F}_U(1 - u) = 1 - u$ et $\mathbb{P}[1 - U \leq u] = u = \mathbb{P}[U \leq u]$, de sorte que U et $1 - U$ ont même fonction de répartition.

2) Pour $v \geq 0$, on a $\mathbb{P}[V \leq v] = \mathbb{P}[-\log U \leq v] = \mathbb{P}[U \geq e^{-v}] = 1 - e^{-v}$. Pour $v < 0$, $\mathbb{P}[V \leq v] = 0$. On dit que V suit la loi exponentielle de paramètre 1 et on écrit $V \sim \mathcal{E}(1)$. Comme U et $1 - U$ ont même loi, W suit aussi la loi exponentielle $\mathcal{E}(1)$.

3) La variable aléatoire $R = \sqrt{-\log(U)}$ a pour domaine de variation $[0, \infty[$. Pour $r < 0$,

on a $\mathbb{F}_R(r) = \mathbb{P}[R \leq r] = 0$. Pour $r \geq 0$, on a :

$$\begin{aligned}
 \mathbb{F}_R(r) &= \mathbb{P}[R \leq r] \\
 &= \mathbb{P}[\sqrt{-\log(U)} \leq r] \\
 &= \mathbb{P}[-\log(U) \leq r^2] \\
 &= \mathbb{P}[\log(U) \geq -r^2] \\
 &= \mathbb{P}[U \geq e^{-r^2}] \\
 &= 1 - e^{-r^2}
 \end{aligned}$$

On a donc :

$$\mathbb{F}_R(r) = \begin{cases} 1 - e^{-r^2} & r \geq 0 \\ 0 & \text{sinon} \end{cases}$$

Il suffit alors de dériver \mathbb{F}_R pour vérifier que $R \sim \mathcal{R}(\sigma)$ avec $\sigma = \sqrt{2}/2$.

4) Par la formule de changement de variable, on a :

$$f_{X,Y}(x, y) = \frac{1}{\sqrt{x^2 + y^2}} f_{R,\Phi}(x^2 + y^2, h(x, y)) \quad \text{pour } (x, y) \neq (0, 0)$$

avec :

$$h(x, y) = \begin{cases} \arctan(y/x) & \text{si } x > 0 \text{ et } y > 0 \\ \arctan(y/x) + 2\pi & \text{si } x > 0 \text{ et } y < 0 \\ \arctan(y/x) + \pi & \text{si } x < 0 \end{cases}$$

Puisque R et Φ sont indépendantes, on peut écrire :

$$f_{X,Y}(x, y) = \frac{1}{\sqrt{x^2 + y^2}} f_R(x^2 + y^2) f_\Phi(h(x, y)),$$

pour $(x, y) \neq 0$. Comme $\Phi \sim \mathcal{U}[0, 2\pi[$ et que $h(x, y) \in]0, 2\pi[$, $f_\Phi(h(x, y)) = \frac{1}{2\pi}$. De plus :

$$\begin{aligned}
 R \sim \mathcal{R}(\sqrt{2}/2) \quad \Rightarrow \quad f_R\left(\sqrt{x^2 + y^2}\right) &= 2\sqrt{x^2 + y^2} e^{-(x^2 + y^2)} \\
 &= 2r e^{-r^2}
 \end{aligned}$$

Au final :

$$\begin{aligned}
 \forall (x, y) \in \mathbb{R}^2, \quad f_{X,Y}(x, y) &= \frac{1}{\sqrt{x^2 + y^2}} 2\sqrt{x^2 + y^2} e^{-(x^2 + y^2)} \frac{1}{2\pi} \\
 &= \frac{e^{-(x^2 + y^2)}}{\pi}
 \end{aligned}$$

La loi marginale de X est donnée par :

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

De même, $f_Y(y) = \frac{1}{\sqrt{\pi}} e^{-y^2}$. On a donc $X \sim \mathcal{N}(0, \frac{1}{2})$ et $Y \sim \mathcal{N}(0, \frac{1}{2})$, puisque la densité d'une variable aléatoire $Z \sim \mathcal{N}(0, 1/2)$ est $\forall z \in \mathbb{R}, f_Z(z) = \frac{1}{\sqrt{\pi}} e^{-z^2}$. On remarque aussi que $f_{X,Y}(x,y) = f_X(x)f_Y(y)$. Les variables aléatoires X, Y sont donc indépendantes.

Justification de $\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$. Soit $Z \sim \mathcal{N}(0, 1/2), \forall z \in \mathbb{R}, f_Z(z) = \frac{1}{\sqrt{\pi}} e^{-z^2}$, et

$$\int_{\mathbb{R}} f_Z(z) dz = 1 \Rightarrow \int_{\mathbb{R}} e^{-z^2} dz = \sqrt{\pi}$$

Génération de variables aléatoires gaussiennes indépendantes. On peut générer $\Phi \sim \mathcal{U}[0, 2\pi[$ en prenant $\Phi = 2\pi U_1$, où $U_1 \sim \mathcal{U}[0, 1[$. On génère ensuite $U_2 \sim \mathcal{U}[0, 1[$, indépendante de U_1 . On calcule $V = -\log U_2$ puis $\sqrt{-\log U_2}$. On obtient alors $R \sim \mathcal{R}(\sqrt{2}/2)$. Les variables aléatoires R et Φ étant indépendantes, les calculs précédents montrent que $X = R \cos \Phi$ et $Y = R \sin \Phi$ sont gaussiennes centrées de variance $1/2$ et indépendantes.

EXERCICE 2.5.– [Variance d'une variable aléatoire discrète]

Soit X une variable aléatoire discrète à valeurs dans \mathbb{N}^* telle que $\mathbb{P}[X = k] = \frac{\alpha}{k!}$.

1. Calculer α .
2. Calculer $\mathbb{E}[X]$ et $\mathbb{E}[X(X-1)]$. En déduire la variance et l'écart-type de X .

Solution

1) Calcul de α . Nous devons avoir $\sum_{k>0} \mathbb{P}[X = k] = 1$ (condition de normalisation). On a alors :

$$\sum_{k>0} \mathbb{P}[X = k] = 1 \Leftrightarrow \alpha \sum_{k=1}^{+\infty} \frac{1}{k!} = 1 \Leftrightarrow \alpha [e - 1] = 1 \Leftrightarrow \alpha = \frac{1}{e - 1}$$

2) Calcul de $\mathbb{E}[X]$. Puisque X est une variable aléatoire positive ou nulle, l'espérance de X existe toujours dans $[0, \infty[$ et, par définition, nous avons :

$$\mathbb{E}[X] = \sum_{k>0} k \times \mathbb{P}[X = k] = \sum_{k=1}^{+\infty} \alpha \frac{k}{k!} = \alpha \sum_{k=0}^{+\infty} \frac{1}{k!} = \alpha$$

3) Calcul de $\mathbb{E}[X(X-1)]$. Comme $X \geq 1$, $X(X-1)$ est encore une variable aléatoire positive ou nulle dont l'espérance mathématique existe toujours. En utilisant le théorème de transfert, nous avons :

$$\mathbb{E}[X(X-1)] = \sum_{k>0} k(k-1)P[X=k] = \sum_{k=2}^{+\infty} \alpha \frac{k(k-1)}{k!} = \sum_{k=2}^{+\infty} \alpha \frac{k(k-1)}{k!} = \alpha e$$

On en déduit la variance de X en écrivant :

$$\begin{aligned} \text{Var}\{X\} &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= (\mathbb{E}[X(X-1)] + \mathbb{E}[X]) - \mathbb{E}[X]^2 \\ &= 2\alpha e - \alpha^2 e^2 \\ &= \frac{e}{e-1} \frac{2(e-1) - e}{e-1} \\ &= \frac{e}{(e-1)^2} (e-2) \end{aligned}$$

d'où l'écart-type de X égal à $\sigma_X = \sqrt{\text{Var}\{X\}} = \sqrt{\frac{e}{(e-1)^2} (e-2)}$.

EXERCICE 2.6.– [Combinatoire avec des tirages sans remise]

On dispose d'une urne remplie de N boules numérotées de 1 à N . On effectue k tirages sans remise avec $k \leq N$. Soit X le plus petit numéro obtenu.

1. Caractériser la distribution de X . En déduire l'égalité : $\sum_{i=0}^m C_{n+i}^n = C_{n+m+1}^{n+1}$.
2. Montrer que $\mathbb{E}[X] = \sum_{i=1}^{N-k+1} \mathbb{P}[X \geq i]$. En déduire que $\mathbb{E}[X] = \frac{N+1}{k+1}$.

Solution

1) Il faut commencer par modéliser l'expérience aléatoire considérée. L'espace Ω des états est l'ensemble des k -uplets $(\omega_1, \omega_2, \dots, \omega_k)$ de $\{1, 2, \dots, N\}^k$ tels que, pour tout $i, j \in \{1, 2, \dots, k\}$, $i \neq j$. On peut aussi considérer que l'espace des états est l'ensemble de tous les k -uplets $(\omega_1, \omega_2, \dots, \omega_k)$ de $\{1, 2, \dots, N\}^k$ avec $1 \leq \omega_1 < \omega_2 < \dots < \omega_k \leq N$. La tribu associée à cette expérience aléatoire sera l'ensemble des parties de Ω et la mesure de probabilité \mathbb{P} est la distribution uniforme telle que :

$$\mathbb{P}(\{(\omega_1, \omega_2, \dots, \omega_k)\}) = 1/\text{card}(\Omega)$$

avec $\text{card}\Omega = C_N^k$ pour tout k -uplet $(\omega_1, \omega_2, \dots, \omega_k) \in \Omega$. On considère alors la variable aléatoire $X : \Omega \rightarrow \mathbb{R}$ définie pour tout $(\omega_1, \omega_2, \dots, \omega_k) \in \Omega$ par :

$$X(\omega_1, \omega_2, \dots, \omega_k) = \min_{j \in \{1, 2, \dots, k\}} \omega_j$$

Cette variable aléatoire prend ses valeurs dans $\{1, 2, \dots, N - k + 1\}$ et nous avons, pour tout $j \in \{1, 2, \dots, N - k + 1\}$:

$$\mathbb{P}[X = j] = \frac{C_{N-j}^{k-1}}{C_N^k}$$

En effet, $X(\omega_1, \omega_2, \dots, \omega_k) = j$ si et seulement si une des valeurs $\omega_1, \omega_2, \dots, \omega_k$ est égale à j et toutes les autres, au nombre de $k - 1$, sont dans $\{j + 1, \dots, N\}$.

Puisque $\sum_{j=1}^{N-k+1} \mathbb{P}[X = j] = 1$ (condition de normalisation), on a : $C_N^k = \sum_{j=1}^{N-k+1} C_{N-j}^{k-1}$. Il suffit alors de poser $m = N - k$ et $n = k - 1$ pour en déduire l'égalité

$$\sum_{i=0}^m C_{n+i}^n = C_{n+m+1}^{n+1}$$

pour tous les entiers $n, m \geq 0$.

Remarque : On peut vérifier la condition de normalisation en procédant comme suit. On veut montrer que :

$$S = \sum_{j=1}^{N-k+1} \frac{C_{N-j}^{k-1}}{C_N^k} = 1$$

Par le triangle de Pascal, nous avons : $C_{N-j}^{k-1} = C_{N-j+1}^k - C_{N-j}^k$. On obtient donc :

$$S = \sum_{j=1}^{N-k+1} C_{N-j}^{k-1} = (C_N^k - C_{N-1}^k) + (C_{N-1}^k - C_{N-2}^k) + \dots + (C_{k+1}^k - C_k^k) + (C_k^{k-1}) = C_N^k$$

Il s'ensuit que $\sum_{j=1}^{N-k+1} \mathbb{P}[X = j] = \sum_{j=1}^{N-k+1} \frac{C_{N-j}^{k-1}}{C_N^k} = \frac{1}{C_N^k} (C_N^k) = 1$.

2) Montrons maintenant que $\mathbb{E}[X] = \sum_{i=1}^{N-k+1} \mathbb{P}[X \geq i]$. Tout d'abord, puisque X est à valeurs dans un ensemble fini, l'espérance de X existe. Par définition de cette espérance, on a $\mathbb{E}[X] = \sum_{j=1}^{N-k+1} j \mathbb{P}[X = j]$. On peut alors écrire cette espérance mathématique sous la forme :

$$\begin{aligned} \mathbb{E}[X] = & \mathbb{P}[X = 1] + \mathbb{P}[X = 2] + \mathbb{P}[X = 3] + \dots + \mathbb{P}[X = N - k + 1] \\ & + \mathbb{P}[X = 2] + \mathbb{P}[X = 3] + \dots + \mathbb{P}[X = N - k + 1] \\ & + \mathbb{P}[X = 3] + \dots + \mathbb{P}[X = N - k + 1] \\ & + \mathbb{P}[X = N - k + 1] \end{aligned}$$

Nous pouvons donc écrire :

$$\mathbb{E}[X] = \mathbb{P}[X \geq 1] + \mathbb{P}[X \geq 2] + \dots + \mathbb{P}[X \geq N - k + 1] = \sum_{j=1}^{N-k+1} \mathbb{P}[X \geq j].$$

D'après la question précédente, on a maintenant :

$$\mathbb{P}[X \geq j] = \sum_{i=j}^{N-k+1} \frac{C_{N-i}^{k-1}}{C_N^k} = \frac{C_{N-j+1}^k}{C_N^k}.$$

Au final :

$$\begin{aligned} \mathbb{E}[X] &= \frac{1}{C_N^k} \sum_{j=1}^{N-k+1} C_{N-j+1}^k = \frac{1}{C_N^k} \sum_{j=1}^{N-k+1} C_{N-j+2}^{k+1} - C_{N-j+1}^{k+1} \\ &= \frac{1}{C_N^k} C_{N+1}^{k+1} \\ &= \frac{(N+1)!}{(k+1)(N-k)!} \frac{(N-k)!k!}{N!} \\ &= \frac{N+1}{k+1} \end{aligned}$$

EXERCICE 2.7.– [Exemple de variable absolument continue]

Soient X une variable aléatoire absolument continue à valeurs dans \mathbb{R} et Y une variable aléatoire donnée par la relation $Y = e^X$.

1. Donner l'ensemble de valeurs décrit par Y .
2. Exprimer les relations entre \mathbb{F}_Y et \mathbb{F}_X , et entre f_Y et f_X .
3. On suppose dans la suite que X est une variable aléatoire gaussienne de moyenne μ et de variance σ^2 . Calculer la moyenne et la variance de Y .

Solution

1) La variable aléatoire X est absolument continue à valeurs dans \mathbb{R} . Elle admet une densité de probabilité $f_X(x)$. Comme la fonction exponentielle est strictement croissante, Y est une variable aléatoire absolument continue à valeur dans $]0 + \infty[$.

2) Pour $y \in]0, \infty[$, $\mathbb{P}[Y \leq y] = \mathbb{P}[e^X \leq y] = \mathbb{P}[X \leq \ln(y)]$. On a donc $\mathbb{F}_Y(y) = \mathbb{F}_X(\ln(y))$. On a donc aussi $\mathbb{F}_X(x) = \mathbb{F}_Y(e^x)$ pour tout $x \in \mathbb{R}$. De ces égalités, on en déduit $f_Y(y) = \frac{1}{y} f_X(\ln y)$ pour tout réel $y > 0$.

3) On calcule la moyenne de Y en utilisant le théorème de transfert, qui nous donne : $\mathbb{E}[Y] = \mathbb{E}[e^X] = \int_{\mathbb{R}} e^x f_X(x) dx$. En utilisant de nouveau le théorème de transfert, le moment d'ordre 2 de Y est donné par

$$\mathbb{E}[Y^2] = \mathbb{E}[e^{2X}] = \int_{\mathbb{R}} e^{2x} f_X(x) dx.$$

Les intégrales donnant les moments d'ordre 1 et 2 de Y sont de la forme $I(a) = \int_{\mathbb{R}} e^{ax} f_X(x) dx$. Puisqu'on suppose que $X \sim \mathcal{N}(\mu, \sigma^2)$, on a :

$$\begin{aligned}
 I(a) &= \int_{\mathbb{R}} e^{ax} f_X(x) dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} e^{ax} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
 &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} e^{-\frac{(x^2 - 2\mu x + \mu^2 - 2a\sigma^2 x)}{2\sigma^2}} dx \\
 &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} e^{-\frac{(x - (\mu + a\sigma^2))^2}{2\sigma^2}} e^{-\frac{\mu^2 - (\mu + a\sigma^2)^2}{2\sigma^2}} dx \\
 &= e^{\frac{a^2\sigma^4 + 2a\mu\sigma^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} e^{-\frac{t^2}{2\sigma^2}} dt \\
 &= e^{\frac{a^2\sigma^2}{2} + a\mu}
 \end{aligned}$$

On en déduit donc que $\mathbb{E}[Y] = I(1) = e^{\mu + \frac{\sigma^2}{2}}$ et que $\mathbb{E}[Y^2] = I(2) = e^{2(\mu + \sigma^2)}$. La variance de Y est alors donnée par :

$$\sigma_Y^2 = I(2) - I(1)^2 = e^{2\mu + 2\sigma^2} - e^{2\mu + \sigma^2} = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1).$$

Chapitre 3

Aléatoire multivarié

EXERCICE 3.1.– [Somme et différence de variables aléatoires]

Soit (X, Y) un couple de variables aléatoires indépendantes. On suppose que X suit une loi uniforme sur $[0, 1]$ et Y une loi exponentielle de paramètre λ sur $[0, +\infty[$ ($\lambda > 0$). On rappelle que la loi exponentielle $\mathcal{E}(\lambda)$ de paramètre λ a pour densité de probabilité $f_\lambda(x) = \lambda e^{-\lambda x}$ pour $x \geq 0$ et $f_\lambda(x) = 0$ pour $x < 0$. On construit un nouveau couple de variables aléatoires (U, V) définies par : $U = X + Y$ et $V = X - Y$.

1. Déterminer la densité de probabilité conjointe du couple (U, V) .
2. En déduire les lois marginales de U et V .
3. Calculer les matrices de covariance de $[X \ Y]^t$ et de $[U \ V]^t$.

Solution

1) On a $f_X(x) = \mathbb{1}_{[0,1]}(x)$ et $f_Y(y) = \lambda e^{-\lambda y} \mathbb{1}_{[0,+\infty[}(y)$. Aussi, les domaines de variations de U et V sont, respectivement, $[0, \infty[$ et $] -\infty, 1]$. On a $(U, V) = (X, Y)A$ avec $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. We then have $\det(A) = -2$ et son inverse est $A^{-1} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}$. On déduit alors de la formule de changement de variables dans le cas linéaire que :

$$\begin{aligned} f_{U,V}(u, v) &= |\det A|^{-1} f_{X,Y}((u, v)A^{-1}) \\ &= \frac{1}{2} f_{X,Y}\left(\frac{u+v}{2}, \frac{u-v}{2}\right) \\ &= \frac{1}{2} f_X(x(u, v)) f_Y(y(u, v)) \quad [\text{car } X \text{ et } Y \text{ sont indépendantes}] \\ &= \frac{1}{2} \lambda e^{-\lambda \frac{u-v}{2}} \mathbb{1}_{[0,1] \times [0, \infty[}\left(\frac{u+v}{2}, \frac{u-v}{2}\right) \end{aligned}$$

2) On obtient la densité de probabilité de U en écrivant :

$$f_U(u) = \int_{\mathbb{R}} f_{U,V}(u, v) dv$$

Puisque $0 \leq \frac{u+v}{2} \leq 1$ équivaut à $v \in [-u, 2-u]$ et que $\frac{u-v}{2} \geq 0$ équivaut à $v \in]-\infty, u]$, nous avons :

$$f_U(u) = \mathbb{1}_{]0, \infty[}(u) e^{-\frac{\lambda u}{2}} \int_{\mathbb{R}} \frac{1}{2} \lambda e^{\frac{\lambda v}{2}} \mathbb{1}_{[-u, 2-u] \cap]-\infty, u[\cap]-\infty, 1]}(v) dv$$

On a alors :

a) Si $0 \leq u \leq 1$, on a : $-u \leq u \leq 1 \leq 2-u$, de sorte que :

$$[-u, 2-u] \cap]-\infty, u[\cap]-\infty, 1] = [-u, u]$$

Il vient donc :

$$\begin{aligned} f_U(u) &= e^{-\frac{\lambda u}{2}} \int_{-u}^u \frac{1}{2} \lambda e^{\frac{\lambda v}{2}} dv \\ &= e^{-\frac{\lambda u}{2}} \int_{-\frac{\lambda u}{2}}^{\frac{\lambda u}{2}} e^t dt \quad [\text{Avec le changement de variable } t = \lambda v/2] \\ &= e^{-\frac{\lambda u}{2}} (e^{\frac{\lambda u}{2}} - e^{-\frac{\lambda u}{2}}) \\ &= 1 - e^{-\lambda u} \end{aligned}$$

b) Si $u > 1$, $-u \leq 2-u \leq 1 \leq u$, ce qui induit que :

$$[-u, 2-u] \cap]-\infty, u[\cap]-\infty, 1] = [-u, 2-u]$$

On obtient alors

$$\begin{aligned} f_U(u) &= e^{-\frac{\lambda u}{2}} \int_{-u}^{2-u} \frac{1}{2} \lambda e^{\frac{\lambda v}{2}} dv \\ &= e^{-\frac{\lambda u}{2}} \int_{-\frac{\lambda u}{2}}^{\frac{\lambda(2-u)}{2}} e^t dt \quad [\text{Avec le changement de variable } t = \lambda v/2] \\ &= e^{-\frac{\lambda u}{2}} (e^{\frac{2\lambda}{2}} e^{-\frac{\lambda u}{2}} - e^{-\frac{\lambda u}{2}}) \\ &= e^{-\lambda u} (e^\lambda - 1) \end{aligned}$$

En définitive, $f_U(u) = (1 - e^{-\lambda u}) \mathbb{1}_{]0, 1]}(u) + (e^{-\lambda u} (e^\lambda - 1)) \mathbb{1}_{]1, \infty[}(u)$.

La densité de probabilité de V est donnée par $f_V(v) = \int_{\mathbb{R}} f_{U,V}(u, v) du$. On procède comme précédemment et le lecteur vérifiera que :

$$f_V(v) = e^{\lambda v} (1 - e^\lambda) \mathbb{1}_{]-\infty, 0[}(v) + (1 - e^{-\lambda} e^{\lambda v}) \mathbb{1}_{]0, 1]}(v)$$

3) On a :

$$\mathbb{E}[X] = \int_0^1 x f_X(x) dx = 1/2$$

$$\mathbb{E}[Y] = \int_0^{+\infty} \lambda y e^{-\lambda y} dy = 1/\lambda$$

$$\text{Cov}\{X, Y\} = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0 \quad [\text{Puisque } X \text{ et } Y \text{ sont indépendantes.}]$$

$$\mathbb{E}[X^2] = \int_0^1 x^2 f_X(x) dx = 1/3$$

$$\mathbb{E}[Y^2] = \int_0^{+\infty} y^2 f_Y(y) dy = 2/\lambda^2$$

$$\text{Var}\{X\} = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 1/3 - 1/4 = 1/12$$

$$\text{Var}\{Y\} = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = 2/\lambda^2 - 1/\lambda^2 = 1/\lambda^2.$$

La matrice de covariance de $(X, Y)^T$ est donc :

$$K_{X,Y} = \begin{bmatrix} 1/12 & 0 \\ 0 & 1/\lambda^2 \end{bmatrix}$$

Puisque $\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = A \begin{pmatrix} X \\ Y \end{pmatrix}$, on a :

$$\begin{aligned} K_{U,V} &= \mathbb{E} \left[A \begin{pmatrix} X - \mathbb{E}(X) \\ Y - \mathbb{E}(Y) \end{pmatrix} \begin{pmatrix} X - \mathbb{E}(X) \\ Y - \mathbb{E}(Y) \end{pmatrix}^T A^T \right] \\ &= AK_{X,Y}A^T \\ &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1/12 & 0 \\ 0 & 1/\lambda^2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1/12 + 1/\lambda^2 & 1/12 - 1/\lambda^2 \\ 1/12 - 1/\lambda^2 & 1/12 + 1/\lambda^2 \end{pmatrix} \end{aligned}$$

EXERCICE 3.2.- [Suite de variables aléatoires gaussiennes et loi gamma]

Soient n variables aléatoires réelles X_1, X_2, \dots, X_n indépendantes et identiquement distribuées selon une loi normale centrée de variance unité. On définit deux nouvelles variables aléatoires :

$$M_n = \frac{1}{n} \sum_{i=1}^n X_i \text{ et } V_n = \frac{1}{n} \sum_{i=1}^n (X_i - M_n)^2$$

Soit A une matrice carrée d'ordre n , orthogonale et dont les termes de la première ligne sont tous égaux à $\frac{1}{\sqrt{n}}$. Soit le vecteur aléatoire $\vec{Y} = A\vec{X}$ où $\vec{X} = (X_1, X_2, \dots, X_n)^T$.

1. Déterminer la loi de \vec{Y} .
2. Exprimer M_n et V_n en fonction des composantes de \vec{Y} .
3. Les variables M_n et V_n sont-elles indépendantes ?
4. Donner la loi de M_n . On se propose maintenant de déterminer la loi de la variable V_n .
5. Montrer que si X est une variable aléatoire qui suit une loi normale centrée de variance unité, alors X^2 suit une loi gamma $\gamma(\frac{1}{2}, \frac{1}{2})$. On rappelle que $Y \sim \gamma(\theta, \lambda)$ si Y est absolument continue et a pour densité :

$$f_Y(y) = \frac{\lambda^\theta}{\Gamma(\theta)} y^{\theta-1} e^{-\lambda y} \mathbb{1}_{]0, +\infty[}(y)$$

$$\text{avec } \Gamma(\theta) = \int_0^{+\infty} x^{\theta-1} e^{-x} dx \text{ et } \Gamma(1/2) = \int_0^{+\infty} \frac{1}{\sqrt{x}} e^{-x} dx = 2 \int_0^{+\infty} e^{-u^2} du = \sqrt{\pi}.$$

On rappelle aussi que la fonction caractéristique d'une variable aléatoire $Y \sim \gamma(\theta, \lambda)$ est $\Phi_Y(t) = \left(\frac{\lambda}{\lambda - it}\right)^\theta$.

6. Soient deux variables aléatoires indépendantes Z_1 et Z_2 de loi respectivement égales à $\gamma(\theta_1, \lambda)$ et $\gamma(\theta_2, \lambda)$. Déterminer la loi de $Z = Z_1 + Z_2$.
7. Dédire des questions précédentes la loi de V_n .

Solution

Il faut remarquer que $\vec{Y} = A\vec{X} \Rightarrow Y_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i = \sqrt{n}M_n$. Il faut aussi noter que $\|Y\|^2 = Y^T Y = X^T A^T A X = X^T X = \|X\|^2$ car $AA^T = \text{Id}$ puisque A est orthogonale. Ces deux remarques seront très utiles dans la suite.

1) Comme X_1, X_2, \dots, X_n sont iid, on a :

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-x_i^2/2} = \frac{1}{(2\pi)^{n/2}} e^{-\|\vec{x}\|^2/2}$$

où $\|\cdot\|$ est la norme Euclidienne usuelle sur \mathbb{R}^n : si $\vec{x} = (x_1, x_2, \dots, x_n)^T$, $\|\vec{x}\| = \sum_{k=1}^n x_k^2$.

On en déduit que \vec{X} est un vecteur gaussien de vecteur moyenne nulle et de matrice de covariance identité. Puisque \vec{X} est un vecteur gaussien, $\vec{Y} = A\vec{X}$ est aussi un vecteur gaussien de moyenne nulle et de matrice de covariance $K_{\vec{Y}} = AK_{\vec{X}}A^T = \text{Id}_n$

où $K_{\vec{X}} = \text{Id}_n$ est la matrice de covariance de \vec{X} . Par conséquent, \vec{Y} a pour densité $f_{\vec{Y}}(\vec{y}) = \frac{1}{(2\pi)^{n/2}} e^{-\|\vec{y}\|^2/2}$. Comme \vec{Y} est un vecteur gaussien (ses n composantes sont mutuellement gaussiennes) et que la matrice de covariance de Y est diagonale, les n composantes de Y sont indépendantes.

2) On a :

$$M_n = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{\sqrt{n}} Y_1$$

On a aussi :

$$\begin{aligned}
 V_n &= \frac{1}{n} \sum_{i=1}^n (X_i - M_n)^2 = \frac{1}{n} \left[\sum_{i=1}^n X_i^2 - 2M_n \sum_{i=1}^n X_i + nM_n^2 \right] \\
 &= \frac{1}{n} (X^T X - 2nM_n^2 + nM_n^2) \\
 &= \frac{1}{n} (Y^T Y - Y_1^2) \quad \text{En vertu des remarques préliminaires} \\
 &= \frac{1}{n} \sum_{i=2}^n Y_i^2
 \end{aligned}$$

3) Les Y_i sont indépendantes, M_n ne fait intervenir que Y_1 et V_n ne fait intervenir que les Y_i à partir du deuxième indice ($i = 2$). Il s'ensuit que les variables aléatoires M_n et V_n sont indépendantes.

4) Puisque $M_n = \underbrace{(1/n, 1/n, \dots, 1/n)}_{m \text{ terms}}^T X$, M_n est une variable aléatoire Gaussienne en tant que transformée linéaire du vecteur gaussien X . La moyenne de M_n est nulle et la variance de M_n est égale à la somme des variances des X_i puisque ces variables aléatoires sont indépendantes donc décorrélées. On a donc $M_n \sim \mathcal{N}(0, 1/n)$.

5) Puisque X est à valeur dans \mathbb{R} , X^2 est à valeur dans $[0, \infty[$. On calcule la fonction de répartition de X^2 . On a $\mathbb{F}_{X^2}(x) = \mathbb{P}[-\sqrt{x} \leq X \leq \sqrt{x}] = \mathbb{F}_X(\sqrt{x}) - \mathbb{F}_X(-\sqrt{x})$, pour tout $x \in [0, +\infty[$. Donc, pour $x \neq 0$, la fonction de répartition de X^2 est dérivable de dérivée :

$$f_{X^2}(x) = \frac{1}{2\sqrt{x}} f_X(\sqrt{x}) + \frac{1}{2\sqrt{x}} f_X(-\sqrt{x}) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}$$

puisque $X \sim \mathcal{N}(0, 1)$. La variable aléatoire X^2 admet donc pour densité :

$$f_{X^2}(x) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}} \mathbb{1}_{]0, +\infty[}(x)$$

Ceci est bien la densité d'une loi $\gamma(1/2, 1/2)$.

6) Soient 2 variables aléatoires Z_1, Z_2 indépendantes et distribuées respectivement selon $\gamma(\theta_1, \lambda)$ et $\gamma(\theta_2, \lambda)$. La fonction caractéristique de $Z = Z_1 + Z_2$ est :

$$\begin{aligned}
 \phi_Z(t) &= \mathbb{E}[e^{itZ}] = \mathbb{E}[e^{itZ_1} e^{itZ_2}] \\
 &= \mathbb{E}[e^{itZ_1}] \mathbb{E}[e^{itZ_2}] \\
 &= \phi_{Z_1}(t) \phi_{Z_2}(t) \\
 &= \left(\frac{\lambda}{\lambda - it} \right)^{\theta_1} \left(\frac{\lambda}{\lambda - it} \right)^{\theta_2} \\
 &= \left(\frac{\lambda}{\lambda - it} \right)^{\theta_1 + \theta_2}
 \end{aligned}$$

On en déduit que $Z \sim \gamma(\theta_1 + \theta_2, \lambda)$ puisque la fonction caractéristique est unique pour une loi donnée.

7) Puisque $V_n = \frac{1}{n} \sum_{i=2}^n Y_i^2$, on a $V_n = \sum_{i=2}^n \left(\frac{Y_i}{\sqrt{n}}\right)^2$. Posons $Z_i = \frac{Y_i}{\sqrt{n}}$ avec Z_i de loi $\mathcal{N}(0, 1/n)$. Posons ensuite $U_i = Z_i^2$. Les U_i sont de loi $\gamma(1/2, 1/2)$. Or les Y_i étant indépendantes, les U_i le sont aussi. Ainsi les V_n sont $(n-1)$ variables aléatoires indépendantes de même loi $\gamma(1/2, 1/n)$. D'après ce qui précède, V_n suit une loi $\gamma\left(\frac{n-1}{2}, \frac{n}{2}\right)$.

EXERCICE 3.3.– [Probabilité et couple de variables aléatoires]

Deux personnes conviennent d'un rendez-vous entre 14h et 15h en admettant qu'aucune des deux n'attendra plus de 15 minutes. Quelle est la probabilité pour qu'elles se rencontrent? On modélise l'instant d'arrivée de chacune de ces personnes par une variable aléatoire uniformément répartie sur l'intervalle $[0, 1]$, en ayant pris 14h comme origine du temps. On suppose que ces deux variables aléatoires sont indépendantes.

Solution

Soit X , le temps d'arrivée de la première personne et Y , le temps d'arrivée de la deuxième personne. Ces deux variables sont uniformément réparties dans $[0, 1]$ et indépendantes.

Soit Z , la variable aléatoire séparant les deux instants d'arrivée. On a $Z = |X - Y|$. La variable aléatoire Z est à valeurs dans $[0, 1]$. Comme X et Y sont indépendantes, on a $f_{X,Y}(x, y) = \mathbb{1}_{[0,1]}(x) \times \mathbb{1}_{[0,1]}(y)$. Soit $z \in [0, 1]$. On calcule (cf. figure 3.1) :

$$\mathbb{P}[Z \leq z] = \mathbb{P}[|X - Y| \leq z] = \int_{\{(x,y) \in \mathbb{R}^2 : |x-y| \leq z\}} \mathbb{1}_{[0,1] \times [0,1]}(x, y) dx dy = 1 - (1-z)^2$$

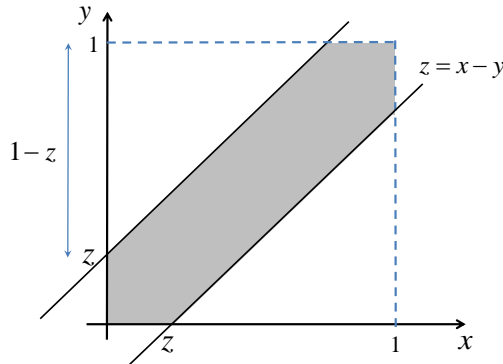


FIGURE 3.1 – Probabilité de rencontre

EXERCICE 3.4.– [File d'attente]

Soit une file d'attente à un guichet. Le nombre de clients dans la file suit une loi de Poisson de paramètre λ . La probabilité que le client puisse obtenir le service désiré est égale à p . Quelle est la loi de la variable aléatoire qui compte le nombre de requêtes satisfaites ? On supposera que le nombre de requêtes satisfaites sachant que l'on a n clients suit une loi binomiale de paramètre p .

Solution

Le nombre de clients dans la file est supposé suivre une loi de Poisson de paramètre λ . Notons N la variable aléatoire qui donne le nombre de clients dans la file. On a donc : $\mathbb{P}[N = n] = \frac{\lambda^n}{n!} e^{-\lambda}$. Soit Y , la variable aléatoire qui compte le nombre de requêtes satisfaites. On calcule la loi de Y en supposant que la loi conditionnelle de Y sachant le nombre de clients N est binomiale de sorte que pour $n \in \mathbb{N}$ donné et $k \leq n$, $\mathbb{P}[Y = k | N = n] = C_n^k p^k (1-p)^{n-k}$. L'événement " k requêtes sont satisfaites" correspond alors à l'ensemble probabilisable :

$$[Y = k] = \bigcup_{n=1}^{\infty} ([Y = k] \cap [N = n]) = \bigcup_{n=k}^{\infty} ([Y = k] \cap [N = n]).$$

En effet, $[Y = k] \cap [N = n] = \emptyset$ pour $n < k$ puisque le nombre de requêtes satisfaites ne peut pas excéder le nombre de clients. Comme les événements $[Y = k] \cap [N = n]$ sont disjoints 2 à 2, les hypothèses nous permettent d'écrire :

$$\begin{aligned} \mathbb{P}[Y = k] &= \sum_{n=k}^{\infty} \mathbb{P}([Y = k] \cap [X = n]) \\ &= \sum_{n=k}^{\infty} \mathbb{P}[Y = k | X = n] \mathbb{P}[X = n] \\ &= \sum_{n=k}^{\infty} C_n^k p^k (1-p)^{n-k} \frac{\lambda^n}{n!} e^{-\lambda} \\ &= \left(\sum_{n=k}^{\infty} \frac{(1-p)^{n-k} \lambda^{n-k}}{(n-k)!} \right) \frac{(\lambda p)^k}{k!} e^{-\lambda} \\ &= \left(\sum_{n=0}^{\infty} \frac{(1-p)^n \lambda^n}{(n)!} \right) \frac{(\lambda p)^k}{k!} e^{-\lambda} \\ &\quad \text{[En faisant le changement d'indice } n - k \leftrightarrow k\text{]} \\ &= e^{(1-p)\lambda} \frac{(\lambda p)^k}{k!} e^{-\lambda} \\ &= \frac{(\lambda p)^k}{k!} e^{-\lambda p} \end{aligned}$$

La variable aléatoire Y suit donc une loi de Poisson de paramètre λp .

EXERCICE 3.5.– [Conditionnement et inclusion]

Supposons $A \subset B$ avec $\mathbb{P}(A)\mathbb{P}(B) \neq 0$. Calculer $\mathbb{P}[B|A]$ et $\mathbb{P}[A|B]$ et interpréter les résultats obtenus. Montrer que $\mathbb{P}[A \cap B] \geq \mathbb{P}(A)\mathbb{P}(B)$ mais que cette inégalité est généralement fautive.

Solution

On a $\mathbb{P}[B|A] = \mathbb{P}[B \cap A]/\mathbb{P}(A) = \mathbb{P}(A)/\mathbb{P}(A) = 1$, car $A \cap B = A$.

Interprétation : Si $A \subset B$ et A se réalise alors B se réalise. Il est donc naturel d'avoir $\mathbb{P}[B|A] = 1$.

On a aussi $\mathbb{P}[A|B] = \mathbb{P}(B \cap A)/\mathbb{P}(B) = \mathbb{P}(A)/\mathbb{P}(B)$ car $A \subset B$ implique $\mathbb{P}(A) \leq \mathbb{P}(B)$. Comme $\mathbb{P}(B) \leq 1$, on a $\mathbb{P}(A)/\mathbb{P}(B) \geq \mathbb{P}(A)$. On obtient donc $\mathbb{P}[A|B] \geq \mathbb{P}(A)$.

Interprétation : Si $A \subset B$ et B se réalise, la probabilité que A soit réalisé aussi augmente.

Si $A \subset B$ avec $\mathbb{P}(A)\mathbb{P}(B) \neq 0$, le résultat précédent implique que $\mathbb{P}[A|B] \geq \mathbb{P}(A)$, et donc, que $\mathbb{P}[A|B] \geq \mathbb{P}(A)\mathbb{P}(B)$. En général, $\mathbb{P}[A|B] \geq \mathbb{P}(A)\mathbb{P}(B)$ est faux ! Par exemple, dans le cas où $A \cap B = \emptyset$! Mais on peut prendre aussi un exemple où $\mathbb{P}[A|B] = 0$ sans que $A \cap B$ soit vide. Ainsi, supposons que $X \sim \mathcal{U}[0, 1]$, $A = [X \in I]$ et $B = [X \in J]$ où I, J sont des intervalles inclus dans $[0, 1]$. Considérons la mesure de Lebesgue λ (longueur des intervalles). On a alors $\mathbb{P}([X \in I] \cap [X \in J]) = \mathbb{P}[A \cap B] = \lambda(I \cap J)$ et $\mathbb{P}(A)\mathbb{P}(B) = \lambda(I)\lambda(J)$. Il suffit de prendre $I = [1/4, 1]$, $J = [0, 3/4]$ pour vérifier que l'on n'a pas $\mathbb{P}[A \cap B] \geq \mathbb{P}(A)\mathbb{P}(B)$ puisque $\lambda(I \cap J) = 1/2$ et $\lambda(I) = \lambda(J) = 3/4$.

EXERCICE 3.6.– [Calcul de densité avec et sans conditionnement]

Soient un espace probabilisé $(\Omega, \mathcal{B}, \mathbb{P})$, une variable aléatoire $S : \Omega \rightarrow \{-A; +A\}$ avec $\mathbb{P}[S = A] = p \in]0, 1[$ et une variable aléatoire X absolument continue de densité f . On suppose X indépendante de S . On considère $Y = S + X$. Montrer que Y est absolument continue et calculer la densité de probabilité de Y . On fera deux types de calcul : l'un sans conditionnement, l'autre avec conditionnement. Quelle hypothèse ne faut-il pas oublier pour que le raisonnement avec conditionnement fonctionne correctement ?

Solution

On procède de manière classique en calculant $\mathbb{P}[Y \leq y]$. On a :

$$\begin{aligned} \mathbb{P}[Y \leq y] &= \mathbb{P}([Y \leq y] \cap [S = A]) + \mathbb{P}([Y \leq y] \cap [S = -A]) \\ &= \mathbb{P}([S + X \leq y] \cap [S = A]) + \mathbb{P}([S + X \leq y] \cap [S = -A]) \\ &= \mathbb{P}([A + X \leq y] \cap [S = A]) + \mathbb{P}([-A + X \leq y] \cap [S = -A]) \\ &= \mathbb{P}[A + X \leq y]p + \mathbb{P}[-A + X \leq y](1 - p) \\ &= \mathbb{P}[X \leq y - A]p + \mathbb{P}[X \leq y + A](1 - p) \end{aligned}$$

Le reste est affaire de calcul et on obtient : $f_Y(y) = pf(y - A) + (1 - p)f(y + A)$.

Avec conditionnement, on écrira :

$$\begin{aligned}\mathbb{P}[Y \leq y] &= \mathbb{P}[Y \leq y | S = A]p + \mathbb{P}[Y \leq y | S = -A](1 - p) \quad [\text{En utilisant Bayes}] \\ &= \mathbb{P}[A + X \leq y]p + \mathbb{P}[-A + X \leq y](1 - p) \quad [S \text{ et } X \text{ sont indépendants!}]\end{aligned}$$

Il est essentiel de se rendre compte que c'est parce que nous avons indépendance entre S et X que le calcul avec conditionnement fonctionne! Sans cette indépendance, on ne pourrait pas forcément écrire la deuxième égalité dans l'équation précédente. C'est pour cela que le calcul sans conditionnement est préférable car il exhibe, sans équivoque, le rôle crucial de l'hypothèse d'indépendance.

EXERCICE 3.7.- [Probabilité d'erreur et bruit additif]

Soient un espace probabilisé $(\Omega, \mathcal{B}, \mathbb{P})$, une variable aléatoire $S : \Omega \rightarrow \{-A; +A\}$ avec $\mathbb{P}[S = A] = p \in]0, 1[$ et une variable aléatoire $X \sim \mathcal{N}(0, \sigma^2)$, indépendante de S . On pose $Y = S + X$. On dit que X est le bruit d'observation et que S est le signal d'intérêt. Nous avons alors deux cas : soit $Y = A + X$, soit $Y = -A + X$. On appelle test statistique de la valeur de S , toute application \mathcal{T} définie sur \mathbb{R} et à valeurs dans $\{-A, A\}$. Lorsqu'on considère la composée $\mathcal{T}(Y)$, la valeur renvoyée par cette application sera l'estimée que nous faisons de la valeur de A . On commettra donc une erreur dans notre estimation de S si $\mathcal{T}(Y) \neq S$.

1. Calculer la probabilité d'erreur $\mathbb{P}_e(\mathcal{T}) = \mathbb{P}[\mathcal{T}(Y) \neq S]$. On fera deux calculs : l'un sans conditionnement, l'autre avec conditionnement. Quelle hypothèse ne faut-il pas oublier pour que le raisonnement avec conditionnement fonctionne correctement ?
2. On suppose que $p = 1/2$. On considère le test \mathcal{T}_λ défini pour tout $y \in \mathbb{R}$ par :

$$\mathcal{T}_\lambda(y) = \begin{cases} A & \text{si } y > \lambda \\ -A & \text{si } y \leq \lambda \end{cases}$$

Sans calcul mais en justifiant votre réponse par un argument heuristique, quelle valeur suggèreriez-vous pour λ afin de minimiser la probabilité d'erreur? Calculer ensuite la valeur de λ qui minimise effectivement la probabilité d'erreur. En déduire la valeur minimale de la probabilité d'erreur.

3. Si l'on avait défini le test $\bar{\mathcal{T}}_\lambda$ par :

$$\bar{\mathcal{T}}_\lambda(y) = \begin{cases} A & \text{si } y > \lambda, \\ -A & \text{si } y \leq \lambda \end{cases}$$

les résultats précédents auraient-ils été modifiés? Pourquoi?

Solution

1) « Première méthode. » Par définition de la probabilité d'erreur, on a :

$$\mathbb{P}_e(\mathcal{T}) = \mathbb{P}[\mathcal{T}(Y) \neq S] = \mathbb{P}([\mathcal{T}(Y) = A] \cap [S = -A]) + \mathbb{P}([\mathcal{T}(Y) = -A] \cap [S = A])$$

qu'on écrira encore $\mathbb{P}_e(\mathcal{T}) = \mathbb{P}[\mathcal{T}(Y) = A, S = -A] + \mathbb{P}[\mathcal{T}(Y) = -A, S = A]$ pour simplifier les écritures. On a donc :

$$\begin{aligned}\mathbb{P}_e(\mathcal{T}) &= \mathbb{P}[\mathcal{T}(S+X) = A, S = -A] + \mathbb{P}[\mathcal{T}(S+X) = -A, S = A] \\ &= \mathbb{P}[\mathcal{T}(-A+X) = A, S = -A] + \mathbb{P}[\mathcal{T}(A+X) = -A, S = A] \\ &= (1-p)\mathbb{P}[\mathcal{T}(-A+X) = A] + p\mathbb{P}[\mathcal{T}(A+X) = -A]\end{aligned}$$

« Deuxième méthode. » On peut écrire que :

$$\mathbb{P}_e(\mathcal{T}) = \mathbb{P}[\mathcal{T}(Y) = A|S = -A]\mathbb{P}[S = -A] + \mathbb{P}[\mathcal{T}(Y) = -A|S = A]\mathbb{P}[S = A]$$

puisque, pour tout ensemble mesurable $B \in \mathcal{B}$ tel que $\mathbb{P}(B) \neq 0$ et tout autre ensemble mesurable $C \in \mathcal{B}$: $\mathbb{P}(C) = \mathbb{P}(C|B)\mathbb{P}(B) + \mathbb{P}(C|B^c)\mathbb{P}(B^c)$ où B^c désigne le complémentaire de B dans Ω . On a donc :

$$\mathbb{P}_e(\mathcal{T}) = (1-p)\mathbb{P}[\mathcal{T}(S+X) = A|S = -A]\mathbb{P}[S = -A] + p\mathbb{P}[\mathcal{T}(S+X) = -A|S = A]$$

Puisque S et X sont indépendants, on obtient :

$$\mathbb{P}_e(\mathcal{T}) = (1-p)\mathbb{P}[\mathcal{T}(-A+X) = A] + p\mathbb{P}[\mathcal{T}(A+X) = -A]$$

c'est-à-dire, le même résultat que précédemment.

2) Cherchons \mathcal{T}_λ de la forme $\mathcal{T}_\lambda(y) = \begin{cases} A & \text{si } y \geq \lambda \\ -A & \text{si } y < \lambda \end{cases}$

- Puisque il y équiprobabilité entre les valeurs que peut prendre S et que le bruit a une densité de probabilité symétrique par rapport à l'origine, il est naturel de penser que le seuil optimal est 0, pour rester compatible de ces symétries.
- On applique le résultat de la question précédente. On a donc :

$$\begin{aligned}\mathbb{P}_e(\mathcal{T}_\lambda) &= (1/2)(\mathbb{P}[-A+X \geq \lambda] + \mathbb{P}[A+X < \lambda]) \\ &= (1/2)(\mathbb{P}[X \geq \lambda+A] + \mathbb{P}[X < \lambda-A]) \\ &= (1/2)(1 - \mathbb{F}(\lambda+A) + \mathbb{F}(\lambda-A))\end{aligned}$$

où $\mathbb{F}(y) = \mathbb{P}(X \leq y)$ est la fonction de répartition de X . La dérivée de $\mathbb{P}_e(\mathcal{T}_\lambda)$ est alors nulle si et seulement si $f(\lambda+A) = f(\lambda-A)$ où f est la densité de probabilité de S . Étant donné que $S \sim \mathcal{N}(0, \sigma^2)$, un calcul élémentaire montre que $f(\lambda+A) = f(\lambda-A)$ si et seulement si $(\lambda-A)^2 = (\lambda+A)^2$, ce qui équivaut à $\lambda = 0$. On en déduit la valeur de la probabilité d'erreur minimale que l'on peut atteindre avec un test de la forme \mathcal{T}_λ . Il s'agit de la valeur $\mathbb{P}_e^* = \mathbb{P}_e(\mathcal{T}_0)$ de $\mathbb{P}_e(\mathcal{T}_\lambda)$ pour $\lambda = 0$. On a donc : $\mathbb{P}_e^* = (1/2)(1 - \mathbb{F}(A) + \mathbb{F}(-A))$ d'après les calculs précédents. Comme $\mathbb{F}(-A) = 1 - \mathbb{F}(A)$, il vient : $\mathbb{P}_e^* = 1 - \mathbb{F}(A) = \int_A^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} dt$. Il est usuel d'exprimer cette quantité en fonction de la fonction erf ou de la fonction erfc avec : $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ et $\text{erfc}(x) = 1 - \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$. On a alors :

$$\mathbb{P}_e^* = \frac{1}{2} \text{erfc}\left(\frac{1}{\sigma\sqrt{2}}\right)$$

3) Les résultats restent les mêmes si on change l'endroit où l'on accepte l'égalité dans l'expression du test \mathcal{T}_λ , tout simplement parce que la loi de X admet une densité, de sorte que les calculs intégraux ne sont pas modifiés en changeant le domaine d'intégration sur un point.

EXERCICE 3.8.– [Calcul de probabilité par conditionnement]

Dans une population, 7% des personnes sont contaminées par un virus? On suppose disposer d'un test de dépistage ayant les propriétés suivantes : il est positif dans 99% de cas de personnes contaminées et positif dans 3% de cas de personnes non contaminées.

1. Quelle est la probabilité que ce test soit positif quand il est utilisé sur un sujet quelconque?
2. Quelle est le probabilité d'être effectivement contaminé lorsque (sachant) que le test est positif?
3. calculer la probabilité qu'un individu soit contaminé sachant que le test est négatif.

Solution

On commence par modéliser l'expérience. L'espace Ω des états est l'ensemble des individus. Ce sont eux que nous observons. On crée ensuite la variable aléatoire $S : \Omega \rightarrow \{0, 1\}$ qui, à tout individu $\omega \in \Omega$, associe son état avec $S(\omega) = 0$ si l'individu n'est pas contaminé et $S(\omega) = 1$ s'il est contaminé. Autrement dit, S nous renvoie la valeur de vérité de la proposition " ω est contaminé" avec 1 si la proposition est vraie et 0, sinon. D'après l'énoncé, on a $P[S = 1] = 0.07$ et $P[S = 0] = 1 - P[S = 1] = 0.93$.

On modélise maintenant le test de dépistage par une application $\mathcal{T} : \Omega \rightarrow \{0, 1\}$ telle que $\mathcal{T}(\omega) = 0$ si le test est négatif et $\mathcal{T}(\omega) = 1$ si le test est positif. Ce test nous renvoie donc une "estimée" de la valeur de vérité de la proposition " ω est contaminé". Autrement dit, \mathcal{T} estime la valeur de S . Dans l'idéal, on aimerait avoir $\mathcal{T}(\omega) = S(\omega)$ pour tout $\omega \in \Omega$. D'après l'énoncé, on a $P[\mathcal{T} = 1|S = 1] = 0.99$ et $P[\mathcal{T} = 1|S = 0] = 0.03$.

1) On veut calculer $\mathbb{P}[\mathcal{T} = 1] = \mathbb{P}(\{\omega \in \Omega : \mathcal{T}(\omega) = 1\})$, c'est-à-dire, la probabilité que le test soit positif sur un sujet quelconque ω . On a :

$$\begin{aligned}
 \mathbb{P}[\mathcal{T} = 1] &= \mathbb{P}([\mathcal{T} = 1] \cap [S = 1]) + \mathbb{P}([\mathcal{T} = 1] \cap [S = 0]) \\
 &= \mathbb{P}[\mathcal{T} = 1|S = 1]\mathbb{P}[S = 1] + \mathbb{P}[\mathcal{T} = 1|S = 0]\mathbb{P}[S = 0] \\
 &\quad \text{[Axiome de Bayes]} \\
 &= 0.99 \times 0.07 + 0.03 \times 0.93 \\
 &= 0.0972
 \end{aligned}$$

2) On doit calculer $\mathbb{P}[S = 1 | \mathcal{F} = 1]$. On utilise de nouveau l'axiome de Bayes pour écrire que :

$$\mathbb{P}[S = 1 | \mathcal{F} = 1] = \frac{\mathbb{P}([S = 1] \cap [\mathcal{F} = 1])}{\mathbb{P}[\mathcal{F} = 1]} = \frac{\mathbb{P}[\mathcal{F} = 1 | S = 1] \mathbb{P}[S = 1]}{\mathbb{P}[\mathcal{F} = 1]}$$

On a donc : $\mathbb{P}[S = 1 | \mathcal{F} = 1] = 0.99 \times 0.07 / 0.0972 = 0.713$.

3) Nous calculons maintenant $\mathbb{P}[S = 1 | \mathcal{F} = 0]$. Toujours par l'axiome de Bayes, nous avons :

$$\mathbb{P}[S = 1 | \mathcal{F} = 0] = \frac{\mathbb{P}([S = 1] \cap [\mathcal{F} = 0])}{\mathbb{P}[\mathcal{F} = 0]} = \frac{\mathbb{P}[\mathcal{F} = 0 | S = 1] \mathbb{P}[S = 1]}{\mathbb{P}[\mathcal{F} = 0]}$$

Comme $\mathbb{P}[\mathcal{F} = 0 | S = 1] = 1 - \mathbb{P}[\mathcal{F} = 1 | S = 1]$ et que $\mathbb{P}[\mathcal{F} = 0] = 1 - \mathbb{P}[\mathcal{F} = 1]$, nous obtenons :

$$\mathbb{P}[S = 1 | \mathcal{F} = 0] = \frac{(1 - \mathbb{P}[\mathcal{F} = 1 | S = 1]) \mathbb{P}[S = 1]}{1 - \mathbb{P}[\mathcal{F} = 1]}$$

On a donc : $\mathbb{P}[S = 1 | \mathcal{F} = 0] = \frac{(1 - 0.99) \times 0.07}{1 - 0.0972} \approx 0.0008$.

EXERCICE 3.9.- [Vecteur de variables aléatoires gaussiennes décorrélées]

Soit $X \sim \mathcal{N}(0, 1)$, Z une variable aléatoire indépendante de X avec

$$\mathbb{P}[Z = 1] = \mathbb{P}[Z = -1] = \frac{1}{2}.$$

On construit la variable aléatoire $Y = ZX$.

1. Montrer que $Y \sim \mathcal{N}(0, 1)$
2. Calculer la matrice de covariance de $U = (X, Y)^T$.
3. Montrer que U n'est pas gaussien
4. Moralité?

Solution

1) On utilise la fonction caractéristique et on conditionne par rapport à Z . On a :

$$\begin{aligned} \Phi_Y(t) &= \mathbb{E}[e^{itY}] = \mathbb{E}[e^{itZX} | Z = 1] \mathbb{P}[Z = 1] + \mathbb{E}[e^{itZX} | Z = -1] \mathbb{P}[Z = -1] \\ &= (1/2) \mathbb{E}[e^{itX}] + (1/2) \mathbb{E}[e^{-itX}] \end{aligned}$$

car Z et X sont indépendantes. Il vient donc $\Phi_Y(t) = \frac{1}{2}(\Phi_X(t) + \Phi_X(-t))$. Comme $X \sim \mathcal{N}(0, 1)$, on a $\Phi_X(t) = e^{-t^2/2}$. Il s'ensuit que $\Phi_Y(t) = e^{-t^2/2}$, qui est la fonction caractéristique d'une variable aléatoire gaussienne $\mathcal{N}(0, 1)$. Donc $Y \sim \mathcal{N}(0, 1)$.

2) Soit $U = \begin{pmatrix} X \\ Y \end{pmatrix}$. On a $\mathbb{E}[X] = 0$. On a aussi :

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E}[ZX] \\ &= \mathbb{E}[XZ|Z=1]\mathbb{P}[Z=1] + \mathbb{E}[XZ|Z=-1]\mathbb{P}[Z=-1] \\ &= \frac{1}{2}(\mathbb{E}[X] - \mathbb{E}[X]) = 0 \end{aligned}$$

car l'indépendance de X et Z implique $\mathbb{E}[XZ|Z=1] = \mathbb{E}[X]$ et $\mathbb{E}[XZ|Z=-1] = \mathbb{E}[-X] = -\mathbb{E}[X]$. On déduit de ce qui précède que :

$$\mathbb{E}[U] = \begin{pmatrix} \mathbb{E}[X] \\ \mathbb{E}[Y] \end{pmatrix} = 0$$

La matrice de covariance de U est alors égale à :

$$\mathbb{E}[UU^T] = \mathbb{E}\left[\begin{pmatrix} X \\ Y \end{pmatrix}(XY^T)\right] = \begin{pmatrix} \mathbb{E}[X^2] & \mathbb{E}[XY] \\ \mathbb{E}[XY] & \mathbb{E}[Y^2] \end{pmatrix}$$

On a :

$$\begin{cases} \mathbb{E}[X^2] = 1 \\ \mathbb{E}[Y^2] = \mathbb{E}[Z^2 X^2] = \mathbb{E}[Z^2]\mathbb{E}[X^2] = \mathbb{E}[X^2] \quad [\text{En raison de l'indépendance de } X \text{ et } Z] \\ \mathbb{E}[XY] = \mathbb{E}[ZX^2] = \mathbb{E}[Z]\mathbb{E}[X^2] = 0 \quad [X \text{ et } Z \text{ sont indépendantes et } \mathbb{E}[Z] = 0] \end{cases}$$

Par conséquent,

$$\mathbb{E}[UU^T] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

3) Soit $u = \begin{pmatrix} x \\ y \end{pmatrix}$. La fonction caractéristique de $U = (X, Y)^T$ est :

$$\begin{aligned} \Phi_U(u) &= \mathbb{E}[e^{i(xX+yY)}] \\ &= \mathbb{E}[e^{i(x+yZ)X}] \\ &= \mathbb{E}[e^{i(x+yZ)X}|Z=1]\mathbb{P}[Z=1] + \mathbb{E}[e^{i(x+yZ)X}|Z=-1]\mathbb{P}[Z=-1] \\ &= \frac{1}{2}\mathbb{E}[e^{i(x+y)X}] + \frac{1}{2}\mathbb{E}[e^{i(x-y)X}] \\ &= \frac{1}{2}(\Phi_X(x+y) + \Phi_X(x-y)) \end{aligned}$$

car l'indépendance de X et de Z implique que $\mathbb{E}[e^{i(x+yZ)X}|Z=1] = \mathbb{E}[e^{i(x+y)X}]$ et que $\mathbb{E}[e^{i(x+yZ)X}|Z=-1] = \mathbb{E}[e^{i(x-y)X}]$. On a donc :

$$\Phi_U(u) = \frac{1}{2} \left(e^{-\frac{(x+y)^2}{2}} + e^{-\frac{(x-y)^2}{2}} \right) = e^{-\frac{(x^2+y^2)}{2}} \cosh(xy)$$

Si U était gaussien, $\Phi_U(u)$ devrait être de la forme $e^{iu^T m} e^{(1/2)u^T K u}$ avec $m = \mathbb{E}[u]$ et $K = \mathbb{E}[(U - m)(U - m)^T]$. Comme $m = 0$ et $K = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\Phi_U(u)$ devrait être égale à $e^{-(x^2+y^2)/2}$ pour que U soit une variable aléatoire gaussienne. Le terme $\cosh(xy)$ est donc en trop dans l'expression exacte de Φ_U et nous concluons que U n'est pas gaussienne.

4) Deux conclusions :

- 2 variables aléatoires gaussiennes décorréelées ne sont pas forcément indépendantes puisqu'ici, X et Y sont décorréelées mais pas indépendantes;
- 2 variables aléatoires gaussiennes décorréelées ne forment pas forcément un vecteur aléatoire gaussien.

EXERCICE 3.10.- [Moments d'une variable gaussienne]

Soit X une v.a. gaussienne de moyenne μ et d'écart-type $\sigma > 0$.

a) Montrer que l'on a

$$\int_{\mathbb{R}} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}, \quad (3.1)$$

avec $\alpha = \frac{1}{2\sigma^2}$.

b) En déduire les moments centrés $\mathbb{E}[(X - \mu)^k]$ d'ordre $k \in \mathbb{N}^*$ de X , en dérivant l'égalité (3.1) par rapport à α .

Solution

a) Comme $X \sim \mathcal{N}(\mu, \sigma^2)$, $X - \mu \sim \mathcal{N}(0, \sigma^2)$. On a donc :

$$\int_{\mathbb{R}} f_X(x) dx = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sqrt{\frac{\alpha}{\pi}} \int_{\mathbb{R}} e^{-\alpha x^2} dx = 1$$

avec $\alpha = \frac{1}{2\sigma^2}$. D'où l'égalité (3.1).

a) En dérivant k fois le membre de droite dans l'égalité (3.1) par rapport à α , on obtient :

$$\left(\sqrt{\frac{\pi}{\alpha}} \right)^{(k)} = (-1)^k \frac{1.3.5 \dots (2k-1)}{2^k} \sqrt{\frac{\pi}{\alpha^{2k+1}}}$$

Étant donné que $\alpha > 0$, il existe α_0 tel que $0 < \alpha_0 < \alpha$. Soit l'application g définie pour tout $x \in \mathbb{R}$ et tout $\alpha \in]\alpha_0, \infty[$ par $g(x, \alpha) = e^{-\alpha x^2}$. Cette application vérifie les hypothèses de la proposition ???. On peut donc dériver sous le signe somme et on obtient :

$$\left(\int_{\mathbb{R}} e^{-\alpha x^2} dx \right)^{(k)} = (-1)^k \int_{\mathbb{R}} x^{2k} e^{-\alpha x^2} dx$$

Nous avons donc :

$$(-1)^k \int_{\mathbb{R}} x^{2k} e^{-\alpha x^2} dx = (-1)^k \frac{1.3.5 \dots (2k-1)}{2^k} \sqrt{\frac{\pi}{\alpha^{2k+1}}}$$

Par conséquent, nous avons :

$$\int_{\mathbb{R}} x^{2k} e^{-\alpha x^2} dx = \frac{1.3.5\dots(2k-1)}{2^k} \sqrt{\frac{\pi}{\alpha^{2k+1}}}$$

Comme les moments impairs de X sont nuls car la densité de X est paire, nous obtenons donc :

$$\begin{aligned} \forall p \in \mathbb{N}, \quad \mathbb{E}[X^{2p+1}] &= 0 \\ \mathbb{E}[X^{2p}] &= \frac{1.3.5\dots(2k-1)}{2^k} \sqrt{\frac{\pi}{\alpha^{2k+1}}} \end{aligned}$$

Nous constatons que les moments centrés d'une variable aléatoire gaussienne ne dépendent que de sa variance.

EXERCICE 3.11.- [Transformation linéaire d'un vecteur gaussien]

Soit $X = (X_1, X_2)^T$, un vecteur aléatoire réel centré, à deux dimensions, de loi Gaussienne et de matrice de covariance

$$K_X = \begin{pmatrix} 3 & \rho\sqrt{3} \\ \rho\sqrt{3} & 1 \end{pmatrix}$$

avec $|\rho| < 1$. On définit un nouveau vecteur aléatoire $(Y_1, Y_2)^T$ avec :

$$\begin{cases} Y_1 &= \frac{X_1}{\sqrt{3}} - X_2 \\ Y_2 &= \frac{X_1}{\sqrt{3}} + X_2 \end{cases}$$

1. Calculer la variance $\text{Var}\{X_1\}$ de X_1 et la covariance $\text{Cov}\{X_1, X_2\}$ de (X_1, X_2) ?
2. Dans quel cas, les deux variables X_1 et X_2 sont-elles indépendantes ?
3. Calculer la covariance $\text{Cov}\{Y_1, Y_2\}$ du couple (Y_1, Y_2) . Calculer les variances des variables aléatoires Y_1 et Y_2 .
4. Les deux variables Y_1 et Y_2 sont-elles indépendantes ? Justifier clairement votre réponse.

Solution

1) Il suffit de lire la matrice de covariance. On a alors $\text{Var}\{X_1\} = 3$ et $\text{Cov}\{X_1, X_2\} = \mathbb{E}[(X_1 - \mathbb{E}(X_1))(X_2 - \mathbb{E}(X_2))] = \rho\sqrt{3}$.

2) X_1, X_2 indépendantes si et seulement si $\text{Cov}\{X_1, X_2\} = 0$ car le vecteur X est gaussien. Il y a donc indépendance si et seulement si $\rho = 0$ (coefficient de corrélation).

3) $\mathbb{E}(Y_1) = \mathbb{E}(Y_2) = 0$. On calcule maintenant :

$$\begin{aligned}
 \text{Cov}\{Y_1, Y_2\} &= \mathbb{E}[Y_1 Y_2] \\
 &= \mathbb{E}\left[\left(\frac{X_1}{\sqrt{3}} - X_2\right)\left(\frac{X_1}{\sqrt{3}} + X_2\right)\right] \\
 &= \frac{1}{3}\mathbb{E}[X_1^2] + \frac{1}{\sqrt{3}}\mathbb{E}[X_1 X_2] - \frac{1}{\sqrt{3}}\mathbb{E}[X_1 X_2] - \mathbb{E}[X_2^2] \\
 &= \frac{1}{3}\mathbb{E}[X_1^2] - \mathbb{E}[X_2^2] = 1 - 1 = 0 \\
 &\quad [\text{Car } \mathbb{E}[X_1^2] = \text{Var}\{X_1\} \text{ puisque } \mathbb{E}[X_1] = 0; \text{ idem pour } X_2].
 \end{aligned}$$

On calcule maintenant les variances de Y_1 et de Y_2 . On obtient :

$$\begin{aligned}
 \text{Var}\{Y_1\} &= \mathbb{E}[Y_1^2] = \mathbb{E}\left[\left(\frac{X_1}{\sqrt{3}} - X_2\right)^2\right] \\
 &= \mathbb{E}\left[\frac{X_1^2}{3} + X_2^2 - \frac{2}{\sqrt{3}}X_1 X_2\right] \\
 &= 1 + 1 - \frac{2}{\sqrt{3}}\rho\sqrt{3} = 2(1 - \rho)
 \end{aligned}$$

et

$$\begin{aligned}
 \text{Var}\{Y_2\} &= \mathbb{E}[Y_2^2] = \mathbb{E}\left[\left(\frac{X_1}{\sqrt{3}} + X_2\right)^2\right] \\
 &= \mathbb{E}\left[\frac{X_1^2}{3} + X_2^2 + \frac{2}{\sqrt{3}}X_1 X_2\right] \\
 &= 1 + 1 + \frac{2}{\sqrt{3}}\rho\sqrt{3} = 2(1 + \rho)
 \end{aligned}$$

On peut aussi procéder plus rapidement comme suit. On a $Y = AX$ avec :

$$A = \begin{pmatrix} 1/\sqrt{3} & -1 \\ 1/\sqrt{3} & 1 \end{pmatrix}$$

Comme X est un vecteur aléatoire de loi gaussienne, la transformée linéaire $Y = AX$ est aussi un vecteur aléatoire gaussien. On a alors $Y \sim \mathcal{N}(\mathbb{E}[Y], K_Y)$ où K_Y est la matrice de covariance de Y . On a $\mathbb{E}[Y] = \mathbb{E}[AX] = A\mathbb{E}[X] = \mathbf{0}$ puisque $\mathbb{E}[X] = \mathbf{0}$. La

matrice de covariance de Y est :

$$\begin{aligned}
 K_Y &= \mathbb{E}[YY^T] = A\mathbb{E}[XX^T]A^T \\
 &= \begin{pmatrix} 1/\sqrt{3} & -1 \\ 1/\sqrt{3} & 1 \end{pmatrix} \begin{pmatrix} 3 & \rho\sqrt{3} \\ \rho\sqrt{3} & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} \\ -1 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 1/\sqrt{3} & -1 \\ 1/\sqrt{3} & 1 \end{pmatrix} \begin{pmatrix} \sqrt{3}(1-\rho) & \sqrt{3}(1+\rho) \\ \rho-1 & \rho+1 \end{pmatrix} \\
 &= \begin{pmatrix} 2(1-\rho) & 0 \\ 0 & 2(1+\rho) \end{pmatrix}
 \end{aligned}$$

On retrouve ainsi tous les résultats précédents.

4) Le vecteur aléatoire Y est obtenu par transformation linéaire du vecteur aléatoire gaussien X . Le vecteur aléatoire Y est donc lui-même gaussien. L'indépendance de ses composantes équivaut donc à la décorrélation de celles-ci. Comme $\text{Cov}\{Y_1, Y_2\} = 0$, la matrice de covariance de Y est diagonale. Il s'ensuit que ses composantes Y_1 et Y_2 sont décorrélées et donc indépendantes.

EXERCICE 3.12.- [Loi conditionnelle d'une gaussienne]

Soit $X = (X_1, X_2)^T$ un vecteur aléatoire réel à deux dimensions, de loi gaussienne $\mathcal{N}(\vec{0}, \Lambda)$.

1. Pourquoi peut-on écrire Λ sous la forme :

$$\Lambda = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \quad (3.2)$$

avec $\sigma_1^2 = \text{Var}\{X_1\}$, $\sigma_2^2 = \text{Var}\{X_2\}$? Comment s'appelle le coefficient ρ ?

2. On suppose dans la suite que $|\rho| < 1$. Donner l'expression de la densité de probabilité $f_X(x)$ du couple X .
3. Calculer la densité de probabilité conditionnelle $f_{X_2|X_1=x_1}(x_2)$ de X_2 sachant $X_1 = x_1$. Quelle est cette loi?
4. En déduire l'espérance conditionnelle, $\mathbb{E}[X_2|X_1 = x_1]$, de X_2 sachant $X_1 = x_1$.

Solution

- 1) La matrice de covariance de X est :

$$\begin{pmatrix} \text{Var}\{X_1\} & \text{Cov}\{X_1, X_2\} \\ \text{Cov}\{X_1, X_2\} & \text{Var}\{X_2\} \end{pmatrix}$$

On trouve alors l'expression de l'énoncé en posant $\rho = \text{Cov}\{X_1, X_2\} / \sqrt{\sigma_1\sigma_2}$, qui est le coefficient de corrélation entre X_1 et X_2 .

2) Si $|\rho| < 1$, la matrice de covariance de X est inversible. Le vecteur aléatoire gaussien admet donc une densité de probabilité qui est donnée par :

$$f_X(x) = f_{X_1, X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{x_1^2\sigma_2^2 + x_2^2\sigma_1^2 - 2\rho\sigma_1\sigma_2x_1x_2}{2\sigma_1^2\sigma_2^2(1-\rho^2)}}$$

3) La densité de probabilité de X_2 ne s'annule pas et nous avons :

$$f_{X_2|X_1=x_1}(x_2) = \frac{f_{X_1, X_2}(x_1, x_2)}{f_{X_1}(x_1)} = \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}} e^{-\frac{(x_2 - \rho\frac{\sigma_2}{\sigma_1}x_1)^2}{2\sigma_2^2(1-\rho^2)}}$$

La loi de X_2 conditionnellement à X_1 est donc une loi Gaussienne de moyenne $\rho\frac{\sigma_2}{\sigma_1}x_1$ et de variance $\sigma_2^2(1-\rho^2)$. On écrira $X_2|X_1 \sim \mathcal{N}(\rho\frac{\sigma_2}{\sigma_1}x_1, (1-\rho^2)\sigma_2^2)$.

4) Sans calcul, on déduit de l'expression précédente que :

$$\mathbb{E}[X_2|X_1 = x_1] = \int_{\mathbb{R}} x_2 f_{X_2|X_1=x_1}(x_2) dx_2 = \rho\frac{\sigma_2}{\sigma_1}x_1$$

EXERCICE 3.13.- [Loi exponentielle et loi d'Erlang]

La loi exponentielle $\mathcal{E}(\lambda)$ de paramètre caractéristique λ a pour densité de probabilité $f_\lambda(x) = \lambda e^{-\lambda x}$ pour $x \geq 0$ et $f_\lambda(x) = 0$ pour $x < 0$. Soit une variable aléatoire X distribuée selon la loi $\mathcal{E}(\lambda)$.

1. Donner l'expression de la fonction de répartition $F_\lambda(x) = \mathbb{P}[X \leq x]$ et de la fonction de répartition complémentaire $F_\lambda^c(x) = \mathbb{P}[X > x]$.
2. Donner l'expression de la transformée de Laplace $B^*(s) = \mathbb{E}[e^{-sX}]$ et de la fonction génératrice des moments $M(s) = \mathbb{E}[e^{sX}]$.
3. Donner l'expression de la moyenne $\mathbb{E}[X]$ et de la variance $\text{Var}\{X\}$. Que peut-on dire du coefficient de variation carré $cv^2 = \text{Var}\{X\}/\mathbb{E}[X]^2$?
4. Déduire de $M(s)$ l'expression des moments $M_i = \mathbb{E}[X^i]$. On utilisera un développement en série de l'exponentielle dans l'expression de $M(s)$ et on admettra ensuite que l'on peut échanger l'intégrale et le signe somme de la série. Retrouver l'expression des moments d'ordre 1 et 2 et de la variance de X .
5. Démontrer la propriété sans mémoire :

$$\mathbb{P}[X \geq x + u | X \geq u] = \mathbb{P}[X \geq x], \forall x \geq 0, \forall u \geq 0$$

Réciproquement, démontrer que la seule loi absolument continue vérifiant cette propriété est la loi exponentielle. Interpréter cette propriété.

6. Soit une suite X_1, X_2, \dots, X_n de variables aléatoires iid de loi $\mathcal{E}(\lambda)$. Notons $S_n = X_1 + X_2 + \dots + X_n$. Démontrer que la loi de S_n est la loi Erlang(n, λ) de densité de probabilité : $f_{n,\lambda}(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} \mathbb{1}_{]0, \infty[}(x)$;
7. Calculer la fonction de répartition de la loi Erlang(n, λ).

Solution

1) Soit \mathbb{F}_λ , la fonction de répartition de la loi exponentielle de paramètre λ . Par définition, pour tout $x \in \mathbb{R}$:

$$\mathbb{F}_\lambda(x) = \mathbb{P}[X \leq x] = \int_{-\infty}^x f_\lambda(u) du = \int_{-\infty}^x \lambda e^{-\lambda u} \mathbb{1}_{]0, \infty[}(u) du$$

On a donc :

$$\forall x \geq 0, \quad \mathbb{F}_\lambda(x) = \int_0^x \lambda e^{-\lambda u} du = [-e^{-\lambda u}]_0^x = 1 - e^{-\lambda x}$$

et pour tout $x < 0$,

$$\mathbb{F}_\lambda(x) = 0.$$

On en déduit l'expression de la fonction de répartition complémentaire :

$$\mathbb{F}_\lambda^c(x) = \mathbb{P}[X > x] = 1 - \mathbb{P}[X \leq x] = 1 - \mathbb{F}_\lambda(x).$$

D'où :

$$\mathbb{F}_\lambda^c(x) = \begin{cases} e^{-\lambda x} & \text{si } x \geq 0 \\ 1 & \text{si } x < 0. \end{cases}$$

2) On pose $B^*(s) = \mathbb{E}[e^{-sx}]$ et $M(s) = \mathbb{E}[e^{sx}]$. On a alors :

$$M(s) = \int_{\mathbb{R}} e^{sx} \lambda e^{-\lambda x} \mathbb{1}_{]0, \infty[}(x) dx = \int_0^{\infty} \lambda e^{(s-\lambda)x} dx = \frac{\lambda}{s-\lambda} \left[e^{(s-\lambda)x} \right]_0^{+\infty} = \frac{\lambda}{s-\lambda}$$

3) On calcule l'espérance de X . On a :

$$\begin{aligned} \mathbb{E}[X] &= \int_{\mathbb{R}} x f_\lambda(x) dx \\ &= \int_0^{\infty} \lambda x e^{-\lambda x} dx \\ &= \underbrace{[-x e^{-\lambda x}]_0^{\infty}}_{=0} + \int_0^{\infty} e^{-\lambda x} dx \quad [\text{Intégration par parties}] \\ &= \left[\frac{e^{-\lambda x}}{-\lambda} \right]_0^{\infty} = \frac{1}{\lambda} \end{aligned}$$

Pour calculer la variance de X , on commence par calculer le moment d'ordre 2 de X . On a :

$$\begin{aligned}\mathbb{E}[X^2] &= \int_0^{+\infty} \lambda x^2 e^{-\lambda x} dx \\ &= \underbrace{\left[-x^2 e^{-\lambda x}\right]_0^{\infty}}_{=0} + \int_0^{\infty} 2x e^{-\lambda x} dx \\ &= \frac{2}{\lambda} \mathbb{E}[X] = \frac{2}{\lambda^2}\end{aligned}$$

On en déduit la variance de X : $\text{Var}\{X\} = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{\lambda^2}$, ce qui implique que :

$$\frac{\text{Var}\{X\}}{\mathbb{E}[X]^2} = 1$$

4) En utilisant le développement en série de l'exponentielle et en admettant que l'on peut échanger l'intégrale et la somme, on a, pour tout $s < \lambda$:

$$\begin{aligned}M(s) &= \mathbb{E}[e^{sX}] = \int_{\mathbb{R}} e^{sx} f_{\lambda}(x) dx \\ &= \int_{\mathbb{R}} \sum_{k=0}^{\infty} \frac{(sx)^k}{k!} f_{\lambda}(x) dx \\ &= \sum_{k=0}^{\infty} \frac{s^k}{k!} \int_{\mathbb{R}} x^k f_{\lambda}(x) dx \\ &= \sum_{k=0}^{\infty} \frac{s^k}{k!} \mathbb{E}[X^k]\end{aligned}$$

Or,

$$M(s) = \frac{\lambda}{\lambda - s} = \frac{1}{1 - \frac{s}{\lambda}} = \sum_{k=0}^{\infty} \frac{s^k}{\lambda^k}$$

Donc :

$$\sum_{k=0}^{\infty} \frac{s^k}{\lambda^k} = \sum_{k=0}^{\infty} \mathbb{E}[X^k] \frac{s^k}{k!}$$

pour tout $s < \lambda$. Il s'ensuit que $\mathbb{E}[X^k] = \frac{k!}{\lambda^k}$. On retrouve $\mathbb{E}[X] = 1/\lambda$ et $\mathbb{E}[X^2] = 2/\lambda^2$.

5) Soient $x \geq 0, u \geq 0$, on a :

$$\begin{aligned}\mathbb{P}[X \geq u] \mathbb{P}[X \geq x+u | X \geq u] &= \mathbb{P}([X \geq x+u] \cap [X \geq u]) \\ &= \mathbb{P}[X \geq x+u] \mathbb{P}[X \geq u | X \geq x+u].\end{aligned}$$

D'où : $\mathbb{P}[X \geq x + u | X \geq u] = \frac{\mathbb{P}[X \geq x + u]}{\mathbb{P}[X \geq u]} \mathbb{P}[X \geq u | X \geq x + u]$. Comme x est positif, $[X \geq x + u] \subset [X \geq u]$ et donc $\mathbb{P}[X \geq u | X \geq x + u] = 1$. On a donc :

$$\frac{\mathbb{P}[X \geq x + u]}{\mathbb{P}[X \geq u]} = \frac{\mathbb{F}_\lambda^c(x + u)}{\mathbb{F}_\lambda^c(u)} = \frac{e^{-\lambda(x+u)}}{e^{-\lambda u}}$$

On dérive de ce qui précède les égalités :

$$\mathbb{P}[X \geq x + u | X \geq u] = e^{-\lambda x} = \mathbb{F}_\lambda^c(x) = \mathbb{P}[X \geq x]$$

ce qui signifie exactement que $\mathbb{F}_\lambda^c(x + u) = \mathbb{F}_\lambda^c(x)\mathbb{F}_\lambda^c(u)$. Réciproquement, on suppose que X est une variable aléatoire absolument continue dont la fonction de répartition \mathbb{F} vérifie : $\mathbb{F}^c(x + u) = \mathbb{F}^c(x)\mathbb{F}^c(u)$ où x et u sont positifs ou nuls. On en déduit que $\ln(\mathbb{F}^c(x + u)) = \ln(\mathbb{F}^c(x)) + \ln(\mathbb{F}^c(u))$. Comme \mathbb{F}^c est continue et décroissante, la relation précédente implique l'existence d'un $\lambda > 0$ tel que : $\log(\mathbb{F}^c(x)) = -\lambda x$. Par conséquent, $\mathbb{F}^c(x) = e^{-\lambda x}$, de sorte que X suit bien une loi exponentielle.

6) On procède par récurrence. Pour $n = 2$, $S_2 = X_1 + X_2$. Comme X_1 et X_2 sont indépendantes, la densité de S_2 est la convolution des densités de X_1 et X_2 : $f_{S_2}(s) = (f_{X_1} * f_{X_2})(s) = \int_{\mathbb{R}} f_{X_1}(x)f_{X_2}(s - x)dx$. Comme les densités f_{X_1} et f_{X_2} sont toutes les deux celles de la loi exponentielle, on a, pour $x \geq 0$:

$$\begin{aligned} f_{S_2}(s) &= \int_{\mathbb{R}} f_\lambda(u)f_\lambda(x - u)du \\ &= \int_{\mathbb{R}} \lambda^2 e^{-\lambda u} e^{-\lambda(x-u)} \mathbb{1}_{[0,+\infty[}(u) \mathbb{1}_{[0,+\infty[}(x - u)du \\ &= \int_{\mathbb{R}} \lambda^2 e^{-\lambda x} \mathbb{1}_{[0,+\infty[\cap]-\infty, x]}(u)du \\ &= \lambda^2 e^{-\lambda x} \int_{\mathbb{R}} \mathbb{1}_{[0, x]}(u)du \\ &= \lambda^2 e^{-\lambda x} x \\ &= f_{2, \lambda}(s) \end{aligned}$$

Pour le rang n , $S_n = S_{n-1} + X_n$ et on suppose que la loi de S_{n-1} est :

$$f_{n-1, \lambda}(x) = \frac{\lambda^{n-1} x^{n-2} e^{-\lambda x}}{(n-2)!}$$

La densité f_{S_n} de S_n est donc la convolution de $f_{n-1, \lambda}$ par f_λ , puisque les variables

aléatoires S_{n-1} et X_n sont indépendantes. On a donc, pour tout $x \geq 0$:

$$\begin{aligned}
 f_{S_n}(x) &= (f_{n-1,\lambda} * f_\lambda)(x) \\
 &= \int_{\mathbb{R}} \frac{\lambda^{n-1} t^{n-2} e^{-\lambda t}}{(n-2)!} \lambda e^{-\lambda(x-t)} \mathbb{1}_{[0,\infty[}(t) \mathbb{1}_{[0,\infty[}(x-t) dt \\
 &= \frac{\lambda^n}{(n-2)!} \int_{\mathbb{R}} t^{n-2} e^{-\lambda x} \mathbb{1}_{[0,x]}(t) dt \\
 &= \frac{\lambda^n}{(n-2)!} e^{-\lambda x} \left[\frac{t^{n-1}}{n-1} \right]_0^x \\
 &= \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} \\
 &= f_{n,\lambda}(x).
 \end{aligned}$$

La loi de S_n est donc bien celle d'Erlang.

7) Pour tout $x < 0$, la fonction de répartition $\mathbb{F}_{n,\lambda}(x)$ de la loi Erlang(n, λ) vaut 0. Pour tout $x \geq 0$, on a :

$$\begin{aligned}
 \mathbb{F}_{n,\lambda}(x) &= \int_0^x f_{n,\lambda}(t) dt \\
 &= \int_0^x \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!} dt \\
 &= \frac{\lambda^n}{(n-1)!} \left(\left[-\frac{1}{\lambda} t^{n-1} e^{-\lambda t} \right]_0^x + \frac{1}{\lambda} \int_0^x (n-1) t^{n-2} e^{-\lambda t} dt \right) \\
 &\quad \text{[Intégration par parties]} \\
 &= \frac{\lambda^n}{(n-1)!} \left(-\frac{1}{\lambda} x^{n-1} e^{-\lambda x} \right) + \int_0^x \frac{\lambda^{n-1} t^{n-2} e^{-\lambda t}}{(n-2)!} dt \\
 &= \mathbb{F}_{n-1,\lambda}(x) - \frac{\lambda^{n-1}}{(n-1)!} x^{n-1} e^{-\lambda x}
 \end{aligned}$$

Par une récurrence immédiate, nous obtenons :

$$\mathbb{F}_{n,\lambda}(x) = \mathbb{F}_{1,\lambda}(x) - \sum_{k=1}^{n-1} \frac{\lambda^{n-1}}{(n-1)!} x^{n-1} e^{-\lambda x}$$

Etant donné que :

$$\mathbb{F}_{1,\lambda}(x) = \int_0^x f_{1,\lambda}(t) dt = \int_0^x \lambda e^{-\lambda t} dt = \left[-e^{-\lambda t} \right]_0^x = 1 - e^{-\lambda x}$$

nous concluons que :

$$\mathbb{F}_{n,\lambda}(x) = 1 - \sum_{k=0}^{n-1} \exp(-\lambda x) \frac{(\lambda x)^k}{k!} \quad \text{pour } x \geq 0$$

EXERCICE 3.14.– [Loi de Poisson]

La loi de Poisson de paramètre ou d'intensité λ est la loi de toute variable N prenant ses valeurs dans \mathbb{N} avec $\mathbb{P}[N = n] = e^{-\lambda} \frac{\lambda^n}{n!}$.

1. Quelle est la valeur de la moyenne $\mathbb{E}[N]$ d'une variable aléatoire N qui suit une loi Poisson(λ) ?
2. On observe des clients à l'entrée d'un système. Le processus d'arrivée des clients est tel que les intervalles entre les arrivées des clients sont des variables aléatoires iid de loi exponentielle $\mathcal{E}(\lambda)$. Montrer que le nombre de clients à l'entrée du système pendant un intervalle de temps de 1 (unité de temps) suit la loi de Poisson de paramètre λ .

Solution

1) Si N suit une loi de Poisson de paramètre λ , l'espérance de N est

$$\begin{aligned} \mathbb{E}[N] &= \sum_{k=0}^{\infty} k \frac{e^{-\lambda} \lambda^k}{k!} \\ &= \sum_{k=1}^{\infty} k \frac{e^{-\lambda} \lambda^k}{k!} \\ &= \sum_{k=1}^{\infty} \frac{e^{-\lambda} \lambda^k}{(k-1)!} \\ &= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= \lambda e^{-\lambda} e^{\lambda} \\ &= \lambda \end{aligned}$$

2) Soit N le nombre de clients à l'entrée du système pendant un intervalle de temps de 1 unité de temps. Soit X_0 l'instant d'arrivée du premier client et X_i le temps écoulé entre les instants d'arrivée des clients i et $i+1$ pour $i = 1, 2, \dots$. Le temps total écoulé jusqu'à l'arrivée du client n est alors $S_n = \sum_{i=1}^n X_i$. Comme les X_i sont supposées iid de loi $\text{Exp}(\lambda)$, l'exercice précédent nous dit que S_n suit une loi Erlang(n, λ).

La probabilité que le nombre de clients N soit exactement n pendant un intervalle d'une unité de temps est alors $\mathbb{P}[N = n] = \mathbb{P}[S_n \leq 1 < S_{n+1}]$. On peut alors écrire que $[S_n \leq 1 < S_{n+1}] = [S_n \leq 1] \cap [1 < S_{n+1}] = [S_n \leq 1] \cap [S_{n+1} \leq 1]^c = [S_n \leq 1] \setminus [S_{n+1} \leq 1]$. Comme $[S_{n+1} \leq 1] \subset [S_n \leq 1]$ et que, de manière générale, $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(B)$ dès que $B \subset A$, on a :

$$\begin{aligned} \mathbb{P}[S_n \leq 1 < S_{n+1}] &= \mathbb{P}([S_n \leq 1] \setminus [S_{n+1} \leq 1]) \\ &= \mathbb{P}[S_n \leq 1] - \mathbb{P}[S_{n+1} \leq 1] \\ &= \mathbb{F}_{n,\lambda}(1) - \mathbb{F}_{n+1,\lambda}(1) \end{aligned}$$

Puisque $\mathbb{F}_{n,\lambda}(x) = 1 - \sum_{k=0}^{n-1} e^{-\lambda x} \frac{(\lambda x)^k}{k!} \mathbb{1}_{]0, \infty[}(x)$ pour la loi Erlang(n, λ), on a :

$$\mathbb{P}[N = n] = \left(1 - \sum_{k=0}^{n-1} e^{-\lambda} \frac{\lambda^k}{k!}\right) - \left(1 - \sum_{k=0}^n e^{-\lambda} \frac{\lambda^k}{k!}\right) = e^{-\lambda} \frac{\lambda^n}{n!}$$

La variable aléatoire N suit donc une loi de Poisson de paramètre λ .

EXERCICE 3.15.- [Régression linéaire]

Soient X et Y deux variables aléatoires réelles. On cherche à estimer Y à partir de X par la variable aléatoire : $Y_{a,b} = aX + b$ où a et b sont des réels à déterminer. On cherche donc à minimiser l'erreur quadratique moyenne entre Y et son estimé $Y_{a,b}$, définie par $\epsilon_{a,b} = \mathbb{E}[|Y - Y_{a,b}|^2]$. Calculer les valeurs de a et b qui minimisent $\epsilon_{a,b}$ et donner alors l'expression de $Y_{a,b}$ obtenue pour ces valeurs. Donner une interprétation géométrique du résultat obtenu.

Solution

Méthode 1 : On procède par dérivation. On a :

$$\begin{aligned} \epsilon_{a,b} &= \mathbb{E}[(Y - Y_{a,b})^2] \\ &= \mathbb{E}[(Y - aX - b)^2] \\ &= \mathbb{E}[Y^2] - 2a\mathbb{E}[XY] - 2b\mathbb{E}[Y] + a^2\mathbb{E}[X^2] + 2ab\mathbb{E}[X] + b^2 \end{aligned}$$

Ceci implique que :

$$\begin{cases} \frac{\partial \epsilon_{a,b}}{\partial a} = -2\mathbb{E}[XY] + 2a\mathbb{E}[X^2] + 2b\mathbb{E}[X] \\ \frac{\partial \epsilon_{a,b}}{\partial b} = -2\mathbb{E}[Y] + 2a\mathbb{E}[X] + 2b \end{cases}$$

On atteint un extremum de $\epsilon_{a,b}$ en annulant ces deux dérivées partielles. On a alors :

$$\begin{cases} \frac{\partial \epsilon_{a,b}}{\partial a} = 0 \\ \frac{\partial \epsilon_{a,b}}{\partial b} = 0 \end{cases} \iff \begin{cases} \mathbb{E}[XY] = a\mathbb{E}[X^2] + b\mathbb{E}[X] \\ \mathbb{E}[Y] = a\mathbb{E}[X] + b \end{cases} \quad (3.3)$$

Les valeurs de a et b qui annulent ces dérivées partielles sont alors :

$$\begin{aligned}\hat{a} &= \frac{\text{Cov}\{X, Y\}}{\text{Var}\{X\}} \\ \hat{b} &= \mathbb{E}[Y] - \mathbb{E}[X] \frac{\text{Cov}\{X, Y\}}{\text{Var}\{X\}}\end{aligned}$$

Il faut alors vérifier que ces deux valeurs correspondent bien à un minimum de $\epsilon_{a,b}$. Pour cela, on peut montrer que $\sqrt{\epsilon_{a,b}}$ est une fonction convexe du couple (a, b) . Pour tout $u = (a, b)$, on posera $g(u) = \sqrt{\epsilon_{a,b}}$. On remarque alors que $g(u) = \sqrt{\mathbb{E}[(Y - u^T Z)^2]}$ avec $Z = (X, 1)^T$. Étant donné $u \in \mathbb{R}^2$, $v \in \mathbb{R}^2$ et $\lambda \in [0, 1]$, il faut montrer que $g(\lambda u + (1 - \lambda)v) \leq \lambda g(u) + (1 - \lambda)g(v)$. On a :

$$\begin{aligned}g(\lambda u + (1 - \lambda)v) &= \sqrt{\mathbb{E}[(Y - (\lambda u + (1 - \lambda)v)^T Z)^2]} \\ &= \sqrt{\mathbb{E}[(\lambda(Y - u^T Z) + (1 - \lambda)(Y - v^T Z))^2]} \\ &\leq \sqrt{\mathbb{E}[(\lambda(Y - u^T Z))^2]} + \sqrt{\mathbb{E}[(1 - \lambda)(Y - v^T Z)^2]} \\ &\leq \lambda \sqrt{\mathbb{E}[(Y - u^T Z)^2]} + (1 - \lambda) \sqrt{\mathbb{E}[(Y - v^T Z)^2]} \\ &\leq \lambda g(u) + (1 - \lambda)g(v)\end{aligned}$$

La fonction g est donc bien convexe. L'estimée $Y_{a,b}$ de Y sous la forme $Y_{a,b} = aX + b$ est alors :

$$\hat{Y} = \frac{\text{Cov}_{X,Y}}{\text{Var}\{X\}} (X - \mathbb{E}[X]) + \mathbb{E}[Y]$$

Méthode 2 : Le théorème de projection indique que $\epsilon_{a,b} = \mathbb{E}[(Y - (aX + b))^2]$ est minimal si l'erreur $Y - (aX + b)$ est orthogonale à tout vecteur de l'espace engendré par X et $\mathbb{1} : \Omega \rightarrow \mathbb{R}$ tel que $\mathbb{1}(\omega) = 1$ pour tout $\omega \in \Omega$. Cette condition d'orthogonalité conduit directement à $\mathbb{E}[(Y - aX - b)X] = 0$ et $\mathbb{E}[(Y - aX - b)\mathbb{1}] = 0$. En développant ces égalités, on retrouve celles de (3.3), ce qui est donc plus rapide que la méthode par dérivation. Cependant, il faut noter que l'application du théorème de projection n'est pas totalement licite dans l'espace $\mathcal{L}^2(\Omega, \mathcal{B}, \mathbb{P})$. En effet, dans cet espace, il existe des variables aléatoires X telles que $\mathbb{E}[X^2] = 0$ sans que $X(\omega)$ soit nul pour tous les $\omega \in \Omega$, ce qui fait que $q(X) = \sqrt{\mathbb{E}[X^2]}$ n'est pas une norme sur cet espace !. Le théorème de projection ne s'applique donc pas, théoriquement, dans $\mathcal{L}^2(\Omega, \mathcal{B}, \mathbb{P})$. Il fonctionne cependant dans un espace légèrement différent de $\mathcal{L}^2(\Omega, \mathcal{B}, \mathbb{P})$, de sorte que même si l'application du théorème de projection dans $\mathcal{L}^2(\Omega, \mathcal{B}, \mathbb{P})$ n'est pas parfaitement rigoureuse, elle fournit quand même le bon résultat.

EXERCICE 3.16.- [Fonction caractéristique et indépendance]

Soient X_1, X_2 deux variables aléatoires discrètes et indépendantes, et Y une nouvelle variable aléatoire définie par $Y = X_1 + X_2$.

1. Quelle est la fonction caractéristique $\Phi_Y(u)$ de la variable aléatoire Y ?
2. En supposant que les lois de X_1 et X_2 sont de Poisson de paramètres λ_1 et λ_2 , quelle est la loi de la variable Y ?

Solution

1) Par définition de la fonction caractéristique, on a :

$$\begin{aligned}\Phi_Y(u) &= \mathbb{E}[e^{iuY}] = \mathbb{E}[e^{iu(X_1+X_2)}] \\ &= \mathbb{E}[e^{iuY}] = \mathbb{E}[e^{iuX_1}] \mathbb{E}[e^{iuX_2}] = \mathbb{E}[e^{iuX_2}] \\ &= \Phi_{X_1}(u) \times \phi_{X_2}(u)\end{aligned}$$

en raison de l'indépendance des variables aléatoires X_1 et X_2 .

2) Si X_1 et X_2 suivent des lois de Poisson, leurs fonctions caractéristiques s'écrivent $\Phi_{X_1} = e^{\lambda_1(e^{iu}-1)}$ et $\Phi_{X_2} = e^{\lambda_2(e^{iu}-1)}$. Aussi, on a $\Phi_Y(u) = e^{(\lambda_1+\lambda_2)(e^{iu}-1)}$. Par conséquent Y est une loi de poisson de paramètre d'intensité $\lambda_1 + \lambda_2$.

EXERCICE 3.17.- [Loi de Poisson de paramètres aléatoires]

Soit une variable aléatoire X qui, conditionnellement à une variable aléatoire Λ positive ou nulle, suit une loi de Poisson d'intensité Λ . Pour tout $n \in \mathbb{N}$ et tout $\lambda \in [0, \infty[$, on a donc :

$$\mathbb{P}[X = n | \Lambda = \lambda] = e^{-\lambda} \frac{\lambda^n}{n!}$$

Déterminer la distribution de la variable aléatoire X , c'est-à-dire, les valeurs $\mathbb{P}[X = n]$ pour $n \in \mathbb{N}$, en supposant que l'intensité Λ suit une loi exponentielle de moyenne μ . On rappelle qu'une densité de Λ est alors $f_\Lambda(\lambda) = \frac{1}{\mu} e^{-\frac{\lambda}{\mu}} \mathbb{1}_{[0, \infty[}(\lambda)$.

Solution

Pour tout $n \in \mathbb{N}$, on a :

$$\begin{aligned}\mathbb{P}[X = n] &= \int_{\mathbb{R}} \mathbb{P}[X = n | \Lambda = \lambda] f_\Lambda d\lambda \\ &= \int_0^\infty e^{-\lambda} \frac{\lambda^n}{n!} \frac{1}{\mu} e^{-\frac{\lambda}{\mu}} d\lambda \\ &= \int_0^\infty \frac{\lambda^n}{n!} \frac{1}{\mu} e^{-\lambda(1+1/\mu)} d\lambda\end{aligned}$$

En posant $u(\lambda) = \frac{\lambda^n}{n!}$ et $v(\lambda) = -\frac{1}{1+\mu} e^{-\lambda(1+1/\mu)}$, on effectue une intégration par par-

ties dans l'intégrale de la dernière égalité, de sorte que l'on obtient :

$$\begin{aligned}
 \mathbb{P}[X = n] &= \int_0^\infty u(\lambda) v'(\lambda) d\lambda \\
 &= [u(\lambda) v(\lambda)]_0^\infty - \int_0^\infty u'(\lambda) v(\lambda) d\lambda \\
 &= \left[\frac{\lambda^n}{n!} \frac{1}{\mu} e^{-\lambda(1+1/\mu)} \right]_0^\infty + \frac{1}{1+\mu} \int_0^\infty \frac{\lambda^{n-1}}{(n-1)!} e^{-\lambda(1+1/\mu)} d\lambda \\
 &= \frac{\mu}{\mu+1} \mathbb{P}[X = n-1]
 \end{aligned}$$

Par une récurrence immédiate, on obtient : $\mathbb{P}[X = n] = \frac{\mu^n}{(\mu+1)^n} \mathbb{P}[X = 0]$. Comme :

$$\mathbb{P}[X = 0] = \int_0^\infty \frac{1}{\mu} e^{-\lambda(1+1/\mu)} d\lambda = \frac{1}{\mu} \left[-\frac{1}{1+1/\mu} e^{-\lambda(1+1/\mu)} \right]_0^\infty = \frac{1}{\mu+1}$$

on aboutit finalement à :

$$\mathbb{P}[X = n] = \frac{\mu^n}{(1+\mu)^{n+1}}$$

EXERCICE 3.18.- [Convergences presque sûres de suites de variables aléatoires]

Soit (X_n) une suite de variables aléatoires indépendantes et identiquement distribuées telles que $\mathbb{E}[X_n] = \mu$ et $\text{Var}\{X_n\} = \sigma^2$. On considère deux variables aléatoires :

$$M_n = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{et} \quad V_n = \frac{1}{n} \sum_{i=1}^n (X_i - M_n)^2$$

1. Montrer que la suite (M_n) converge presque sûrement vers μ .
2. Calculer la moyenne de V_n .
3. Montrer que V_n converge presque sûrement vers σ^2 .

Solution

1) Puisque X_1, \dots, X_n sont n variables aléatoires iid d'espérance finie, alors la loi forte des grands nombres nous dit que $\frac{1}{n} \sum_{i=1}^n X_i$ converge presque sûrement vers $\mathbb{E}[X_1]$.

On a donc $M_n \xrightarrow[n \rightarrow \infty]{p.s.} \mu$.

2) On a :

$$\begin{aligned}
 \mathbb{E}[V_n] &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[(X_i - M_n)^2] \\
 &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i^2 + M_n^2 - 2X_i M_n] \\
 &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i^2] + \mathbb{E}[M_n^2] - \frac{2}{n} \mathbb{E}[M_n \sum_{i=1}^n X_i] \\
 &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i^2] + \mathbb{E}[M_n^2] - 2\mathbb{E}[M_n^2] \quad [\text{Car } \sum_{i=1}^n X_i = nM_n] \\
 &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i^2] - \mathbb{E}[M_n^2] \\
 &= (\sigma^2 + \mu^2) - \mathbb{E}[M_n^2] \quad [\text{Car } \mathbb{E}[X_n] = \mu \text{ et } \text{Var}\{X_n\} = \sigma^2]
 \end{aligned}$$

Par indépendance des X_i , $\text{Var}\{M_n\} = \frac{1}{n^2} \text{Var}\{\sum_{i=1}^n X_i\} = \frac{1}{n^2} \sum_{i=1}^n \text{Var}\{X_i\} = \frac{1}{n} \sigma^2$. On en déduit $\mathbb{E}[M_n^2] = \text{Var}\{M_n\} + \mathbb{E}[M_n]^2 = \frac{1}{n} \sigma^2 + \mu^2$, ce qui permet d'aboutir à :

$$\mathbb{E}(V_n) = \frac{n-1}{n} \sigma^2$$

3) On écrit que :

$$\begin{aligned}
 V_n &= \frac{1}{n} \sum_{i=1}^n (X_i - M_n)^2 = \frac{1}{n} \sum_{i=1}^n [(X_i - \mu) - (M_n - \mu)]^2 \\
 &= \frac{1}{n} \sum_{i=1}^n \left((X_i - \mu)^2 + (M_n - \mu)^2 - \frac{2}{n} (M_n - \mu) \sum_{i=1}^n (X_i - \mu) \right) \\
 &= \frac{1}{n} \sum_{i=1}^n \left((X_i - \mu)^2 + (M_n - \mu)^2 - \frac{2}{n} (M_n - \mu) \sum_{i=1}^n X_i - n\mu \right) \\
 &= \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 + (M_n - \mu)^2 - \frac{2}{n} (M_n - \mu) (nM_n - n\mu) \\
 &= \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 + (M_n - \mu)^2 - 2(M_n - \mu)^2 \\
 &= \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 - (M_n - \mu)^2
 \end{aligned}$$

Puisque $M_n \xrightarrow[n \rightarrow +\infty]{p.s.} \mu$ d'après ce qui précède, on a $(M_n - \mu)^2 \xrightarrow[n \rightarrow \infty]{p.s.} 0$. Posons $Y_i = (X_i - \mu)^2$ pour $i = 1, 2, \dots, n$. Les variables aléatoires X_i étant iid, les Y_i sont elles-aussi iid.

Comme on a $\mathbb{E}[Y_i] = \sigma^2$ puisque $\mathbb{E}[(X_i - \mu)^2] = \sigma^2$, on applique la forte des grands nombres pour aboutir à :

$$\frac{1}{n} \sum_{i=1}^n Y_i \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{E}(Y_i) = \sigma^2$$

On en déduit que V_n converge p.s. vers σ^2 quand n tend vers ∞ .

EXERCICE 3.19.- [Application non probabiliste du théorème de la limite centrale]

Soit X une variable aléatoire suivant une loi de Poisson de paramètre λ : $X \sim \mathcal{P}(\lambda)$.

1. Calculer $\mathbb{E}[X]$, $\mathbb{E}[X^2]$ et la variance de X ?
2. On suppose que X_1 et X_2 sont indépendantes et telles que $X_1 \sim \mathcal{P}(\lambda_1)$ et $X_2 \sim \mathcal{P}(\lambda_2)$. Montrer que $X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$.
3. On considère une suite X_n , $n = 0, 1, 2, \dots$, de variables aléatoires indépendantes, qui suivent des lois de Poisson de paramètre unité : pour tout $n = 0, 1, 2, \dots$, $X_n \sim \mathcal{P}(1)$. On pose $S_n = X_1 + X_2 + \dots + X_n$.

a) Quelle est la loi de S_n ? Soit \mathbb{F}_n la fonction de répartition de la variable aléatoire S_n . Montrer que $\mathbb{F}_n(n) = e^{-n} \sum_{k=0}^n \frac{n^k}{k!}$.

b) Montrer que $Z_n = \frac{S_n - n}{\sqrt{n}}$ converge en loi vers la loi normale unitaire centrée.

c) Dédurre des questions précédentes que $\lim_n e^{-n} \sum_{k=0}^n \frac{n^k}{k!} = 1/2$.

Solution

1) Soit $X \sim \mathcal{P}(\lambda)$. On a :

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k \mathbb{P}[X = k] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{\lambda} e^{-\lambda} = \lambda$$

Le moment d'ordre 2 est maintenant :

$$\begin{aligned}
 \mathbb{E}[X^2] &= \sum_{k=0}^{\infty} k^2 \mathbb{P}[X = k] \\
 &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} \\
 &= e^{-\lambda} \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!} \\
 &= e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} \\
 &= e^{-\lambda} \sum_{k=1}^{\infty} (k-1) \frac{\lambda^k}{(k-1)!} + e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} \\
 &= e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^k}{(k-2)!} + e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
 &= e^{-\lambda} \lambda^2 \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} + e^{-\lambda} \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\
 &= e^{-\lambda} \lambda^2 e^{\lambda} + e^{-\lambda} \lambda e^{\lambda} = \lambda + \lambda^2
 \end{aligned}$$

D'où $\text{Var}\{X\} = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda$.

2) On commence par calculer la fonction caractéristique d'une variable aléatoire $X \sim \mathcal{P}(\lambda)$. Par définition de la fonction caractéristique :

$$\Phi_X(t) = \mathbb{E}[e^{itx}] = \sum_{k=0}^{\infty} e^{itk} \mathbb{P}[X = k] = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}}$$

Donc $\Phi_X(t) = e^{\lambda(e^{it}-1)}$. Considérons maintenant deux variables aléatoires X_1 et X_2 telles que $X_1 \sim \mathcal{P}(\lambda_1)$ et $X_2 \sim \mathcal{P}(\lambda_2)$. La fonction caractéristique de $Y = X_1 + X_2$ est donnée par :

$$\begin{aligned}
 \Phi_Y(t) &= \Phi_{X_1}(t) \Phi_{X_2}(t) \quad [\text{car } X_1 \text{ et } X_2 \text{ sont indépendantes}] \\
 &= e^{\lambda_1(e^{it}-1)} e^{\lambda_2(e^{it}-1)} \quad [\text{Par le résultat précédent}] \\
 &= e^{(\lambda_1+\lambda_2)(e^{it}-1)}
 \end{aligned}$$

On voit donc que Φ_Y est la fonction caractéristique d'une variable aléatoire de Poisson de paramètre $\lambda_1 + \lambda_2$. On a donc $Y \sim \mathcal{P}(\lambda_1 + \lambda_2)$.

3) a) Une récurrence immédiate du résultat précédent montre que $S_n \sim \mathcal{P}(n)$.

b) Par définition de la fonction de répartition, on a $F_n(n) = \mathbb{P}[S_n \leq n]$. Donc, $F_n(n) =$

$\sum_{k=0}^n \mathbb{P}[S_n = k]$. Puisque $S_n \sim \mathcal{P}(n)$, $\mathbb{P}[S_n = k] = \frac{n\lambda^k}{k!} e^{-\lambda}$, d'où le résultat.

c) Les variables aléatoires X_1, \dots, X_n sont identiquement distribuées et indépendantes. D'après le théorème de la limite centrale, $\frac{S_n - nm}{\sqrt{n}\sigma}$ converge en loi vers $\mathcal{N}(0, 1)$ avec $m = \mathbb{E}[X_i]$ et $\sigma^2 = \text{Var}\{X_i\}$. Comme $X_1 \sim \mathcal{P}(1)$, on a $\mathbb{E}(X_1) = \text{Var}\{X_1\} = 1$ d'après la première question. On en déduit donc que $Z_n = \frac{S_n - n}{\sqrt{n}}$ converge en loi vers $\mathcal{N}(0, 1)$.

d) On a $[Z_n \geq 0] = [S_n \geq n]$. Donc $\mathbb{P}[Z_n \geq 0] = \mathbb{P}[S_n \geq n]$. Comme Z_n converge en loi vers $\mathcal{N}(0, 1)$, $\lim_{n \rightarrow \infty} \mathbb{P}[Z_n \geq 0] = \mathbb{P}[U \geq 0]$ avec $U \sim \mathcal{N}(0, 1)$ car 0 est un point de continuité de la fonction de répartition de la loi $\mathcal{N}(0, 1)$. Or $\mathbb{P}[U \geq 0] = 1/2$, ce qui permet de conclure.

EXERCICE 3.20.– [Convergence en probabilité]

Soit $(\Omega, \mathcal{B}, \mathbb{P})$ un espace probabilisé modélisant une expérience aléatoire et $A \in \mathcal{B}$. Au cours de n répétitions de la même expérience, on mesure la fréquence relative Z_n de l'événement A , c'est-à-dire le rapport entre le nombre de fois où l'événement A est réalisé sur le nombre n total de tests. On suppose que les n répétitions de l'expérience sont indépendantes.

1. Exprimer Z_n à l'aide de n variables aléatoires X_1, X_2, \dots, X_n qui seront supposées indépendantes et identiquement distribuées (iid) selon une loi de Bernoulli pour des raisons que l'on explicitera et qui caractérisent l'occurrence ou non de l'événement A . Quelle est alors la valeur du paramètre p de la loi de Bernoulli commune à toutes ces variables aléatoires X_k ?
2. Montrer que la fréquence relative Z_n converge en probabilité vers la valeur p en utilisant l'inégalité de Bienaymé-Tchebychev.

Solution

1) Pour tout $k \in \{1, 2, \dots, n\}$, soit la variable aléatoire X_k définie par $X_k = 1$ si A est réalisé lors de la k -ième réalisation de l'expérience, et $X_k = 0$ sinon. Ces variables aléatoires seront supposées iid puisque nous considérons des répétitions de la même expérience aléatoire et que ces itérations de cette même expérience aléatoire sont supposées indépendantes. Ces variables aléatoires seront donc supposées suivre une même loi de Bernoulli. Le nombre de fois où l'événement A se produit parmi les n itérations de notre expérience est alors $\sum_{k=1}^n X_k$ et la fréquence relative d'occurrence Z_n de l'événement A est donc : $Z_n = \frac{1}{n} \sum_{k=1}^n X_k$. Dans le modèle choisi, on a $\mathbb{P}[X_k = 1] = \mathbb{P}(A)$ et $\mathbb{P}[X_k = 0] = 1 - \mathbb{P}(A)$, pour tout $k = 1, 2, \dots, n$. Aussi, la valeur du paramètre p de la loi de Bernoulli commune choisie pour tous les variables aléatoires X_k est $p = \mathbb{P}(A)$.

2) On a $\mathbb{E}[X_k] = p$ et $\mathbb{E}[(X_k - \mathbb{E}[X_k])^2] = p(1 - p)$ pour tout $k = 1, 2, \dots, n$. Par consé-

quent $\mathbb{E}[Z_n] = p$. On calcule maintenant la variance de Z_n :

$$\begin{aligned} \text{Var}\{Z_n\} &= \text{Var}\left\{\frac{1}{n} \sum_{k=1}^n X_k\right\} \\ &= \frac{1}{n^2} \text{Var}\left\{\sum_{k=1}^n X_k\right\} \quad [\text{Car } \text{Var}\{aY\} = a^2 \text{Var}\{Y\} \text{ pour tout réel } a \\ &\quad \text{et toute variable aléatoire } Y] \\ &= \frac{1}{n^2} \sum_{k=1}^n \text{Var}\{X_k\} \quad [\text{Car la variance d'une somme de variables aléatoires} \\ &\quad \text{décorrélées est la somme des variances de ces variables} \\ &\quad \text{aléatoires}] \\ &= \frac{p(1-p)}{n} \end{aligned}$$

On applique maintenant l'inégalité de Bienaymé-Tchebychev. Soit $\epsilon > 0$,

$$\mathbb{P}[|Z_n - \mathbb{E}[Z_n]| > \epsilon] \leq \frac{\text{Var}\{Z_n\}}{\epsilon^2} \leq \frac{p(1-p)}{n\epsilon^2}.$$

D'où le résultat puisque $\lim_n \frac{p(1-p)}{n} = 0$.

EXERCICE 3.21.- [Application du théorème de la limite centrale à un calcul d'intervalle de confiance]

Nous nous intéressons aux résultats d'un référendum au sein d'une large population, de taille N . Chaque individu peut voter oui ou non (on néglige les abstentions). Nous cherchons à déterminer le pourcentage de oui, qu'on notera p . Nous disposons pour cela d'une sous-population de taille n tirée au hasard.

1. Si S_n désigne le nombre de oui, quelle est la loi suivie par S_n
2. Soient les variables aléatoires X_1, X_2, \dots, X_n avec $X_k = 1$ si le $k^{\text{ème}}$ individu a répondu oui et $X_k = 0$ sinon. On suppose que ces variables aléatoires sont iid (ce qui peut être discuté suivant la façon dont sont choisies les personnes interrogées) de même loi de Bernoulli $\mathcal{B}(p)$. On a donc $p = \mathbb{P}[X_i = 1]$ pour tout $i = 1, 2, \dots, n$. Une estimation de p peut être donné par S_n/n . Calculer la taille de l'échantillon n (en fixant un nombre positif ϵ et une probabilité d'erreur α qu'une méthode basée sur le théorème de la limite centrale préconise pour avoir

$$p \in \left[\frac{S_n}{n} - \epsilon, \frac{S_n}{n} + \epsilon \right]$$

avec une probabilité de se tromper au plus égale à α .

3. Analyser les limitations de la méthode.

Solution

1) S_n suit la loi binomiale $\mathcal{B}(n, p) : \mathbb{P}[S_n = k] = C_n^k p^k (1-p)^{n-k}$ où p désigne la proportion de oui, ou la probabilité qu'un individu pris au hasard vote oui.

2) Étant donné α et ε , on cherche donc les valeurs de n pour lesquelles $\mathbb{P}[|\frac{S_n}{n} - p| > \varepsilon] \leq \alpha$, ce qui équivaut à avoir $\mathbb{P}[|\frac{S_n}{n} - p| \leq \varepsilon] > 1 - \alpha$. On a $S_n = X_1 + \dots + X_n$ où les X_i sont supposés iid. On approxime alors la loi de $\frac{S_n}{n}$ à l'aide du théorème de la limite centrale. Comme $\mathbb{E}[S_n/n] = p$ et $\text{Var}\{S_n/n\} = \frac{p(1-p)}{n}$, on dérive du théorème de la limite centrale que $\frac{S_n - np}{\sqrt{n\sigma}}$ converge en loi vers la loi $\mathcal{N}(0, 1)$ avec $\sigma = \sqrt{p(1-p)}$. On sait de plus que la convergence de la suite des fonctions de répartition des variables aléatoires $\frac{S_n - np}{\sqrt{n\sigma}}$ est uniforme vers la fonction de répartition de la loi $\mathcal{N}(0, 1)$. Comme $\mathbb{P}[|\frac{S_n}{n} - p| \leq \varepsilon] = \mathbb{P}[|\frac{S_n - np}{\sqrt{n\sigma}}| \leq \frac{\varepsilon\sqrt{n}}{\sigma}]$ et que :

$$\int_{-\frac{\varepsilon\sqrt{n}}{\sigma}}^{+\frac{\varepsilon\sqrt{n}}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 2 \int_0^{+\frac{\varepsilon\sqrt{n}}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \text{erf}\left(\frac{\varepsilon\sqrt{n}}{\sqrt{2}\sigma}\right)$$

avec

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

on peut alors écrire pour tout $\eta > 0$, il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$:

$$\left| \mathbb{P}\left[|\frac{S_n}{n} - p| \leq \varepsilon\right] - \text{erf}\left(\frac{\varepsilon\sqrt{n}}{\sqrt{2}\sigma}\right) \right| \leq \eta$$

Pour avoir $p \in \left[\frac{S_n}{n} - \varepsilon, \frac{S_n}{n} + \varepsilon\right]$ avec une probabilité de se tromper au plus égale à α , il suffit que :

$$\text{erf}\left(\frac{\varepsilon\sqrt{n}}{\sqrt{2}\sigma}\right) \geq 1 - \alpha - \eta$$

ce qui équivaut à :

$$n \geq \frac{2\sigma^2}{\varepsilon^2} (\text{erf}^{-1}(1 - \alpha - \eta))^2$$

La valeur σ n'est pas connue. Mais on a $\sigma = \sqrt{p(1-p)} \leq 1/2$. Le nombre d'échantillon minimal que fournit cette méthode est donc :

$$n_{\min} = \frac{1}{2\varepsilon^2} (\text{erf}^{-1}(1 - \alpha - \eta))^2$$

3) Certes, on peut choisir η . Mais on ne connaît pas la valeur N à partir de laquelle les calculs précédents sont valables. Rien ne nous dit alors que n_{\min} est effectivement supérieur à cette valeur N .

EXERCICE 3.22.– [La convergence presque sûre n'implique pas la convergence en moyenne d'ordre quelconque et donc, la convergence en probabilité]

Soit une variable aléatoire U uniformément répartie sur $[0, 1]$: $U \sim \mathcal{U}([0, 1])$. Pour tout $n \in \mathbb{N}^*$, On pose $X_n = e^n \mathbb{1}_{[0, 1/n]}(U)$. Montrer que $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers 0 mais que $(X_n)_{n \in \mathbb{N}^*}$ ne converge pas en moyenne d'ordre q pour tout entier $q \geq 1$.

Solution

Soit un espace probabilisé $(\Omega, \mathcal{B}, \mathbb{P})$ et supposons U défini sur l'espace probabilisable (Ω, \mathcal{B}) . Soit Ω_0 , l'ensemble des $\omega \in \Omega$ tels que $U(\omega) = 0$. Comme $U \sim \mathcal{U}([0, 1])$, $\mathbb{P}[U = 0] = \mathbb{P}(\Omega_0) = 0$. Pour tout $\omega \in \Omega \setminus \Omega_0$, $\lim_n X_n(\omega) = 0$ puisque $X_n(\omega) = 0$ pour tout $n \geq 1/U(\omega)$. La suite $(X_n)_{n \in \mathbb{N}^*}$ converge donc presque sûrement vers 0 puisque $\lim_n X_n(\omega) = 0$ pour tout $\omega \in \Omega \setminus \Omega_0$ et que $\mathbb{P}(\Omega \setminus \Omega_0) = 1$.

Comme la variable aléatoire X_n est discrète et prend ses valeurs dans $\{0, e^n\}$, on a $\mathbb{E}[X_n^q] = \mathbb{P}[X_n = e^n] e^{nq}$ pour tout entier naturel $q \geq 1$. Or, $[X_n = e^n] = [U \leq 1/n]$. Donc, $\mathbb{P}[X_n = e^n] = \mathbb{P}[U \leq 1/n] = 1/n$. Il s'ensuit que $\mathbb{E}[X_n^q] = e^{nq}/n$. On a donc $\lim_n \mathbb{E}[X_n^q] = \lim_n e^{nq}/n = \infty$. La convergence presque sûre de X_n vers 0 n'entraîne pas la convergence d'ordre q et ce, quel que soit l'entier $q \geq 1$.

Bibliographie

[KHA 94] KHAC K. V., *Intégration er espaces de Lebesgue*, Hermann, 1994.

[RUD 87] RUDIN W., *Real and Complex Analysis*, 3^eédition, McGraw-Hill, New York, 1987.