13 Couples de variables aléatoires

13.1 Séries doubles

On aura besoin dans la suite de séries doubles.

Définition 13.1

Soit $(u_{n,p})_{n,p\in\mathbb{N}}$ une suite double de réels. On dit que cette suite double est *sommable* si :

- Pour tout $n \in \mathbb{N}$, la série $\sum\limits_{p\in\mathbb{N}} u_{n,p}$ converge
- La série $\sum\limits_{n\in\mathbb{N}} \sum\limits_{p=0}^{\infty} u_{n,p}$ converge.

Dans ce cas, la somme de cette seconde série est appelée *somme de la série double* et est notée $\sum\limits_{(n,p)\in\mathbb{N}^2} u_{n,p}$.

Dans certain cas particulier, on peut alors intervertir les sommes :

Théorème 13.2 – de Fubini

Soit $(u_{n,p})_{n,p\in\mathbb{N}}$ une suite double de réels. On suppose que pour tous $n, p \in \mathbb{N}$, $u_{n,p} \geqslant 0$. Alors cette suite double est sommable si et seulement si une des conditions suivantes est vérifiée :

- Pour tout $n \in \mathbb{N}$, la série $\sum\limits_{p\in\mathbb{N}} u_{n,p}$ converge et la série $\sum\limits_{n\in\mathbb{N}} \sum\limits_{p=0}^{\infty} u_{n,p}$ converge
- Pour tout $p \in \mathbb{N}$, la série $\sum\limits_{n\in\mathbb{N}} u_{n,p}$ converge et la série $\sum\limits_{p\in\mathbb{N}} \sum\limits_{n=0}^{\infty} u_{n,p}$ converge

Dans ce cas, on a alors

$$\sum_{(n,p)\in\mathbb{N}^2} u_{n,p} = \sum_{n=0}^{\infty} \sum_{p=0}^{\infty} u_{n,p} = \sum_{p=0}^{\infty} \sum_{n=0}^{\infty} u_{n,p}.$$

EXEMPLE : On veut montrer que la suite $(u_{i,j})_{i,j\geqslant 2}$ définie par $u_{i,j} = \frac{1}{i^j}$ est sommable.

- Soit $i \geqslant 2$. Alors la série $\sum \frac{1}{i^j}$ converge comme série géométrique, et on a
$$\sum_{j=2}^{\infty} \frac{1}{i^j} = \frac{1}{i(i-1)}.$$
- La série $\sum \frac{1}{i(i-1)}$ converge comme série téléscopique, et on a
$$\sum_{i=2}^{\infty} \frac{1}{i(i-1)} = 1.$$

Finalement, la suite double est sommable, et
$$\sum_{i=2}^{\infty}\sum_{j=2}^{\infty} \frac{1}{i^j} = \sum_{j=2}^{\infty}\sum_{i=2}^{\infty} \frac{1}{i^j} = 1.$$

13.2 Généralités

Dans cette section, on fixe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et on considère deux variables aléatoires X et Y. On veut savoir comment caractériser le comportement de ces deux variables aléatoires conjointement.

Définition 13.3

On appelle *couple de variables aléatoires X et Y*, noté (X, Y) l'application
$$(X,Y) : \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R}^2 \\ \omega & \longmapsto & (X(\omega), Y(\omega)) \end{array}.$$

Définition 13.4

On appelle *tribu associée au couple* (X, Y) la plus petite tribu contenant tous les événements
$$((X \leqslant x) \cap (Y \leqslant Y))_{x,y\in\mathbb{R}}.$$

On la note $\mathcal{A}_{(X,Y)}$

Définition 13.5

On appelle *loi du couple aléatoire* (X, Y) la donnée de fonction $F_{(X,Y)} : \mathbb{R}^2 \to \mathbb{R}$, appelée *fonction* *de répartition conjointe*, définie par

$$\forall x, y \in \mathbb{R}, \ F_{(X,Y)}(x, y) = \mathbb{P}((X \leqslant x) \cap (Y \leqslant y)).$$

On dira donc que deux couples aléatoires suivent la même loi conjointe s'ils ont les mêmes fonctions de répartitions conjointes.

On a le théorème (admis) suivant

Théorème 13.6

Soient X_1, Y_1, X_2, Y_2 quatre variables aléatoires. Soit $g \in \mathcal{C}^0(\mathbb{R}^2, \mathbb{R})$.

Alors si les couples (X_1, Y_1) et (X_2, Y_2) ont la même loi, alors les variables aléatoires $g(X_1, Y_1)$ et $g(X_2, Y_2)$ ont la même loi.

EXEMPLE : On admet que la fonction $(x, y) \mapsto x + y$ est continue. Alors si deux couples (X_1, Y_1) et (X_2, Y_2) ont la même loi, les variables aléatoires $X_1 + Y_1$ et $X_2 + Y_2$ ont la même loi.

13.3 Couples de variables aléatoires discrètes

Dans cette section, X et Y sont deux variables aléatoires discrètes. On note $\{x_i \mid i \in I\}$ (resp. $\{y_j \mid j \in J\}$) l'image de X (resp. Y).

D'après la définition, la loi de (X, Y) est entièrement déterminée par la donnée de

$$\forall i \in I, \ \forall j \in J, \ p_{i,j} = \mathbb{P}(X = x_i \cap Y = y_j).$$

Dans le cas où X et Y sont finies, on pourra donc représenter la loi conjointe sous forme d'un tableau :

<table>
<tr><td colspan="2" rowspan="2">(X, Y)</td><td colspan="4">Valeurs de Y</td></tr>
<tr><td>y_1</td><td>y_2</td><td>$\cdots$</td><td>y_p</td></tr>
<tr><td rowspan="4">Valeurs de X</td><td>x_1</td><td>$p_{1,1}$</td><td>$p_{1,2}$</td><td>$\cdots$</td><td>$p_{1,p}$</td></tr>
<tr><td>x_2</td><td>$p_{2,1}$</td><td>$p_{2,2}$</td><td>$\cdots$</td><td>$p_{2,p}$</td></tr>
<tr><td>$\vdots$</td><td>$\vdots$</td><td>$\vdots$</td><td>$\ddots$</td><td>$\vdots$</td></tr>
<tr><td>x_n</td><td>$p_{n,1}$</td><td>$p_{n,2}$</td><td>$\cdots$</td><td>$p_{n,p}$</td></tr>
</table>

EXERCICE : On lance simultanément deux dés. On note X le plus petit des deux dés, et Y le plus grand. Déterminer la loi conjointe de (X, Y).

Réciproquement :

Proposition 13.7

Une suite double $(p_{i,j})_{i\in I, j\in J}$ est la loi conjointe d'un couple aléatoire discret si et seulement si

- $\forall i \in I,\ \forall j \in J,\ p_{i,j} \geqslant 0$
- la série double $\sum\limits_{i\in I}\sum\limits_{j\in J} p_{i,j}$ converge absolument vers 1.

Démonstration. Le sens direct est facile : les $p_{i,j}$ étant des probabilités, ils sont forcément positifs. De plus, les événements $(X = x_i) \cap (Y = y_j)$ forment un système complet, et donc la somme de leurs probabilités vaut 1.

Le sens réciproque est admis. □

Définition 13.8

On appelle *lois marginales du couple* (X, Y) les lois de X et de Y.

Il est facile de trouver les lois de X et Y à partir de la loi conjointe :

Proposition 13.9

La loi de X est donnée par

$$\forall i \in I,\ \mathbb{P}(X = x_i) = \sum_{j\in J} p_{i,j}.$$

De même, la loi de Y est donnée par

$$\forall j \in J,\ \mathbb{P}(Y = y_j) = \sum_{i\in I} p_{i,j}.$$

Démonstration. Il suffit d'appliquer la formule des probabilités totales. □

NOTA : Dans le cas de lois finies, en représentant la loi du couple comme un tableau, il suffit de sommer les éléments des lignes (resp. des colonnes) pour trouver la loi de X (resp. Y).

(X,Y)		Valeurs de Y				Somme de la ligne
		y_1	y_2	$\cdots$	y_p	
Valeurs de X	x_1	$p_{1,1}$	$p_{1,2}$	$\cdots$	$p_{1,p}$	$\mathbb{P}(X=x_1)$
	x_2	$p_{2,1}$	$p_{2,2}$	$\cdots$	$p_{2,p}$	$\mathbb{P}(X=x_2)$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	x_n	$p_{n,1}$	$p_{n,2}$	$\cdots$	$p_{n,p}$	$\mathbb{P}(X=x_n)$
Somme de la colonne		$\mathbb{P}(Y=y_1)$	$\mathbb{P}(Y=y_2)$	$\cdots$	$\mathbb{P}(Y=y_p)$	1

EXERCICE : Déterminer les lois marginales de l'exercice précédent.

Il est donc facile de trouver les lois marginales à partir de la loi conjointe. En revanche, il est souvent très difficile de faire l'opération inverse : il n'y a pas, dans le cas général, de façon simple de trouver la loi conjointe connaissant les lois marginales.

En revanche, si les variables sont indépendantes :

Proposition 13.10

Si X et Y sont indépendantes, alors pour tous i, j

$$p_{i,j} = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j).$$

Définition 13.11

Soit $i \in I$ tel que $\mathbb{P}(X = x_i) \neq 0$. On appelle *loi conditionnelle de Y sachant $X = x_i$* l'application

$$\begin{array}{ccl} \{y_j \mid j \in J\} & \longrightarrow & [0,1] \\ y_j & \longmapsto & \mathbb{P}_{X=x_i}(Y = y_j) = \frac{p_{i,j}}{\mathbb{P}(X=x_i)} \end{array}$$

On a alors, en appliquant directement la formule des probabilités totales

Proposition 13.12

On suppose que tous les $\mathbb{P}(X = x_i)$ et $\mathbb{P}(Y = y_j)$ sont non nuls. Alors pour tous i, j

- $p_{i,j} = \mathbb{P}(Y = y_j)\mathbb{P}_{Y=y_j}(X = x_i) = \mathbb{P}(X = x_i)\mathbb{P}_{X=x_i}(Y = y_j)$

- $\mathbb{P}(X = x_i) = \sum_{j \in J} \mathbb{P}(Y = y_j)\mathbb{P}_{Y=y_j}(X = x_i)$
- $\mathbb{P}(Y = y_j) = \sum_{i \in I} \mathbb{P}(X = x_i)\mathbb{P}_{X=x_i}(Y = y_j)$

Comme dit précédemment, les lois marginales ne suffisent pas à déterminer la loi conjointe, mais une loi marginale et les lois conditionnelles de l'autre variables suffisent.

13.4 Fonctions de deux variables aléatoires

Comme on l'a vu, on peut prendre la fonction d'un couple de variable aléatoire.

Proposition 13.13

Si X et Y sont discrètes, alors pour toute fonction, $g(X, Y)$ est discrète.

On considérera donc des variables comme $X + Y$, $\max(X, Y)$, $\min(X, Y)$, *etc.*

Pour trouver la loi de $g(X, Y)$, *i.e.* pour calculer les $\mathbb{P}(g(X, Y) = z)$, il suffit de sommer les $P(X = x \cap Y = y)$ pour tous les x, y tels que $g(x, y) = z$:

$$\mathbb{P}(g(X, Y) = z) = \sum_{g(x,y)=z} \mathbb{P}(X = x \cap Y = y).$$

On peut alors appliquer le théorème de transfert

Théorème 13.14 – de transfert

La variable aléatoire $Z = g(X, Y)$ admet une espérance si et seulement si la série double $\sum g(x, y)\mathbb{P}(X = x \cap Y = y)$ converge absolument, et dans ce cas

$$\mathbb{E}(Z) = \sum_{i \in I, j \in J} g(x_i, y_j)\mathbb{P}(X = X_i \cap Y = y_j).$$

EXERCICE : En se servant du théorème de transfert, montrer la linéarité de l'espérance

$$\mathbb{E}(\lambda X + \mu Y) = \lambda\mathbb{E}(X) + \mu\mathbb{E}(Y).$$

13.4.1 Somme de deux variables indépendantes

Proposition 13.15

Si X et Y sont indépendantes, alors

$$\mathbb{P}(X+Y=z)=\sum_{\substack{i\in I, j\in J\\ x_i+y_j=z}}\mathbb{P}(X=x_i)\mathbb{P}(Y=y_j).$$

On dit alors que la loi de $X+Y$ est le *produit de convolution* des lois de X de Y.

Démonstration. On a vu comment trouver la loi de $g(X,Y)$:

$$\begin{aligned}\mathbb{P}(X+Y=z) &= \sum_{\substack{i\in I, j\in J\\ x_i+y_j=z}}\mathbb{P}(X=x_i\cap Y=y_i)\\ &= \sum_{\substack{i\in I, j\in J\\ x_i+y_j=z}}\mathbb{P}(X=x_i)\mathbb{P}(Y=y_j)\end{aligned}$$

□

On a alors certains résultats de stabilité.

Proposition 13.16

L'ensemble des lois de Poisson est stable par somme.

Plus précisément, soient $X\hookrightarrow\mathcal{P}(\lambda)$ et $Y\hookrightarrow\mathcal{P}(\mu)$, indépendantes. Alors

$$X+Y\hookrightarrow\mathcal{P}(\lambda+\mu).$$

Démonstration. On a, d'après la formule précédente

$$\begin{aligned}\mathbb{P}(X+Y=n) &= \sum_{\substack{i,j\in\mathbb{N}\\ i+j=n}}\mathbb{P}(X=i)\mathbb{P}(Y=j)\\ &= \sum_{i=0}^{n}\mathbb{P}(X=i)\mathbb{P}(Y=n-i)\end{aligned}$$

On sait que $\mathbb{P}(X=i)=\frac{\lambda^i e^{-\lambda}}{i!}$ et $\mathbb{P}(Y=n-i)=\frac{\mu^{n-i}e^{-\mu}}{(n-i)!}$. Donc

$$\mathbb{P}(X+Y=n)=\sum_{i=0}^{n}\frac{\lambda^i e^{-\lambda}}{i!}\frac{\mu^{n-i}e^{-\mu}}{(n-i)!}$$

En appliquant la formule du binôme, on obtient donc bien

$$\begin{aligned}\mathbb{P}(X+Y=n) &= \frac{e^{-(\lambda+\mu)}}{n!}\sum_{i=0}^{n}\frac{n!\lambda^i\mu^{n-i}}{i!(n-i)!}\\ &= \frac{e^{-(\lambda+\mu)}}{n!}(\lambda+\mu)^n\end{aligned}$$

□

Proposition 13.17

L'ensemble des lois binomiales de probabilité p est stable par somme.

Plus précisément, soient $X \hookrightarrow \mathcal{B}(m, p)$ et $Y \hookrightarrow \mathcal{B}(n, p)$ indépendantes. Alors

$$X + Y \hookrightarrow \mathcal{B}(m + n, p).$$

Démonstration. En revenant à la définition de loi binomiale :

- X correspond au nombre de succès dans la répétition de n épreuves de Bernoulli indépendantes de paramètre p
- Y correspond au nombre de succès dans la répétition de m épreuves de Bernoulli indépendantes de paramètre p

$X + Y$ est donc le nombre de succès dans la répétition de $n + m$ épreuves de Bernoulli indépendantes de paramètre p. □

13.4.2 Maximum de deux variables

Proposition 13.18

On a, pour tout $z \in \mathbb{R}$,

$$(\max(X, Y) \leqslant z) = (X \leqslant z) \cap (Y \leqslant z).$$

Si X et Y sont indépendantes, on a alors

$$F_{\max(X,Y)}(z) = F_X(z)F_Y(z).$$

Démonstration. Le premier point est évident par définition du max.

Si X et Y sont indépendantes, on a alors

$$\begin{aligned} F_{\max(X,Y)}(z) &= \mathbb{P}(\max(X, Y) \leqslant z) \\ &= \mathbb{P}(X \leqslant z \cap Y \leqslant z) \\ &= F_X(z)F_Y(z) \end{aligned}$$

□

On note que dans le cas du max, c'est la fonction de répartition qui nous donne la loi de probabilité.

EXERCICE : En procédant de la même façon, trouver la loi de $\min(X, Y)$.

13.5 Covariance et corrélation linéaire

Dans cette section, on reprend les variables aléatoires discrètes de la section précédente.

Commençons par essayer de calculer l'espérance d'un produit de variables aléatoires.

Proposition 13.19

Si X, Y sont indépendantes, alors si X et Y admettent une espérance, alors XY aussi et

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

Démonstration. D'après le théorème de transfert, il suffit d'étudier la convergence et de calculer la somme double

$$\sum_{i \in I, j \in J} x_i y_j \mathbb{P}(X = x_i \cap Y = y_j).$$

On a donc

$$\begin{aligned}\left|x_i y_j \mathbb{P}(X = x_i \cap Y = y_j)\right| &= |x_i||y_j||\mathbb{P}(X = x_i \cap Y = y_j)| \\ &= |x_i \mathbb{P}(X = x_i)||y_j \mathbb{P}(Y = y_j)|\end{aligned}$$

Comme X et Y admettent une espérance, la somme converge bien, et donc XY admet une espérance. De plus

$$\begin{aligned}\sum_{i \in I, j \in J} x_i y_j \mathbb{P}(X = x_i \cap Y = y_j) &= \sum_{i \in I} \sum_{j \in J} x_i y_j \mathbb{P}(X = x_i \cap Y = y_j) \\ &= \sum_{i \in I} x_i \mathbb{P}(X = x_i) \sum_{j \in J} y_j \mathbb{P}(Y = y_j) \\ &= \mathbb{E}(X)\mathbb{E}(Y)\end{aligned}$$

□

EXERCICE : Montrer que, dans le cas où X et Y ne sont pas indépendantes mais admettent un moment d'ordre 2, alors XY admet toujours une espérance.

13.5.1 Covariance

Définition 13.20

Si X et Y admettent un moment d'ordre 2*, on appelle *covariance de X et Y* le nombre réel

$$\mathrm{Cov}(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right).$$

*. C'est équivalent à dire qu'elles possèdent une variance

NOTA : D'après l'exercice précédent, puisque X et Y ont un moment d'ordre 2, XY admet bien une espérance.

NOTA : On note que la covariance d'une variable avec elle-même est égale à la variance.

Proposition 13.21 – Formule de Hyugens

Si X et Y admettent un moment d'ordre 2, alors

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

Démonstration. On a

$$\begin{aligned}\mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right) &= \mathbb{E}\left(XY - X\mathbb{E}(Y) - Y\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y)\right) \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)\end{aligned}$$

□

La covariance est une fonction bilinéaire, symétrique et positive :

Proposition 13.22

On suppose que X, Y, Z admettent un moment d'ordre 2. Alors

- *Symétrie :* $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
- *Bilinéarité :*
  - $\forall \lambda \in \mathbb{R},\ \mathrm{Cov}(\lambda X + Y, Z) = \lambda\,\mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$
  - $\forall \lambda \in \mathbb{R},\ \mathrm{Cov}(X, \lambda Y + Z) = \lambda\,\mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$
- *Positivité :* $\mathrm{Cov}(X, X) \geqslant 0$

Démonstration. La preuve est triviale en utilisant la linéarité de l'espérance. □

Proposition 13.23

Si X et Y admettent un moment d'ordre 2 et sont indépendantes, alors $\mathrm{Cov}(X, Y) = 0$.

NOTA : Attention ! La réciproque est fausse. La nullité de la covariance n'est en aucun cas une condition nécessaire et suffisante d'indépendance. On pourra retenir le contre-exemple suivant :

Soit Z une variable qui vaut -1 et 1 avec probabilités respectives $\frac{1}{2}$ [†]. Soit X une variable aléatoire discrète quelconque admettant une espérance et une variance, indépendante de Z.

Alors X et $Y = ZX$ ne sont pas indépendantes, mais

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(X^2 Z) - \mathbb{E}(X)\mathbb{E}(ZX) = 0.$$

On peut s'en servir pour calculer, dans le cas d'indépendance, la variance d'une somme de variables aléatoires.

Théorème 13.24

Si X et Y admettent un moment d'ordre 2, alors $X+Y$ admet un moment d'ordre 2, et

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\,\mathrm{Cov}(X,Y).$$

En particulier, dans le cas où X et Y sont indépendantes, on a donc

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y).$$

Démonstration. Il suffit de calculer :

$$\begin{aligned}\mathbb{V}(X+Y) &= \mathbb{E}((X+Y)^2) - \mathbb{E}(X+Y)^2 \\ &= \mathbb{E}(X^2 + 2XY + Y^2) - \mathbb{E}(X)^2 - \mathbb{E}(Y)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 + \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 + 2(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)) \\ &= \mathbb{V}(X) + \mathbb{V}(Y) + 2\,\mathrm{Cov}(X,Y).\end{aligned}$$

□

NOTA : Dans le cas où on connaît la loi de $X+Y$, on peut s'en servir pour calculer la covariance :

$$\mathrm{Cov}(X,Y) = \frac{1}{2}(\mathbb{V}(X+Y) - \mathbb{V}(X) - \mathbb{V}(Y)).$$

Proposition 13.25 – Inégalité de Cauchy-Schwarz

Soient X et Y deux variables aléatoires admettant une variance. On a

$$|\,\mathrm{Cov}(X,Y)| \leqslant \sigma(X)\sigma(Y).$$

Démonstration. C'est une inégalité de Cauchy-Schwarz, aussi on étudie la fonction $\varphi : t \mapsto \mathbb{V}(tX - Y)$.

On a $\mathbb{V}(tX-Y) = t^2\mathbb{V}(X) - 2\,\mathrm{Cov}(X,Y)t + \mathbb{V}(Y)$: φ est un trinôme du second degré, qui est toujours positif. Son discriminant est donc négatif.

On en déduit donc que $4\,\mathrm{Cov}(X,Y)^2 \leqslant 4\mathbb{V}(X)\mathbb{V}(Y)$, et donc le résultat. □

†. On dit que Z suit la loi de Rademacher

13.6 Exercices

Exercice 1

Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$. On suppose que X suit une loi de Poisson de paramètre $\lambda > 0$ et que la loi de Y sachant $X = n$ est binomiale de paramètres n et $p \in]0, 1[$.

1. Déterminer la loi conjointe de (X, Y).
2. Reconnaître la loi de Y.

Réponse de l'exercice

1. Soient $x, y \in \mathbb{N}$. On a alors

$$\begin{aligned}\mathbb{P}(X = x \cap Y = y) &= \mathbb{P}_{[X=x]}(Y = y)\mathbb{P}(X = x)\\ &= \binom{x}{y} p^y (1-p)^{x-y} e^{-\lambda} \frac{\lambda^x}{x!}\end{aligned}$$

2. On peut alors trouve la loi de Y en utilisant la formule des probabilités totales :

$$\begin{aligned}\mathbb{P}(Y = y) &= \sum_{x=0}^{\infty} \mathbb{P}(X = x \cap Y = y)\\ &= \sum_{x=0}^{\infty} \binom{x}{y} p^y (1-p)^{x-y} e^{-\lambda} \frac{\lambda^x}{x!}\\ &= \sum_{x=y}^{\infty} \binom{x}{y} p^y (1-p)^{x-y} e^{-\lambda} \frac{\lambda^x}{x!}\\ &= \sum_{x=0}^{\infty} \binom{x+y}{y} p^y (1-p)^x e^{-\lambda} \frac{\lambda^{x+y}}{(x+y)!}\\ &= (p\lambda)^y e^{-\lambda} \frac{1}{y!} \sum_{x=0}^{\infty} \frac{(\lambda(1-p))^x}{x!}\\ &= (p\lambda)^y \frac{1}{y!} e^{-\lambda p}\end{aligned}$$

   Y suit alors une loi de Poisson $\mathcal{P}(\lambda p)$.

Exercice 2

Soient X et Y deux variables aléatoires réelles à valeurs dans $\mathbb{N}$ telles que pour tous $i, j \in \mathbb{N}$,

$$\mathbb{P}((X, Y) = (i, j)) = \frac{\alpha}{2^{i+1} j!}.$$

1. Déterminer la valeur de α.
2. Montrer que X et Y sont indépendantes.

Réponse de l'exercice

1. Les séries étant convergentes, on a

$$\begin{aligned}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\frac{\alpha}{2^{i+1}j!} &= \sum_{j=0}^{\infty}\frac{\alpha}{j!}\\ &= \alpha e\end{aligned}$$

   On doit donc avoir $\alpha = e$.

2. Il suffit de calculer les lois marginales de X et Y :

$$\begin{aligned}\forall i \in \mathbb{N},\ \mathbb{P}(X=i) &= \frac{1}{2^{i+1}}\\ \forall j \in \mathbb{N},\ \mathbb{P}(Y=j) &= \frac{1}{ej!}\end{aligned}$$

   Les variables sont bien indépendantes.

Exercice 3

Soit $n \in \mathbb{N}^*$, et soient X et Y deux variables aléatoires finies d'image $[\![1,n]\!]$. On suppose qu'il existe $a \in \mathbb{R}$ tel que pour tous $i, j \in [\![1,n]\!]$,

$$\mathbb{P}(X = i \cap Y = j) = aij.$$

1. Calculer la valeur de a.
2. Déterminer les lois marginales de X et Y, ainsi que leurs espérances.
3. Calculer la covariance de X et Y. Les variables X et Y sont-elles indépendantes ?
4. Calculer $\mathbb{P}(X = Y)$.
5. Déterminer la loi de la variable $M = \max(X, Y)$.

Réponse de l'exercice

1. La somme de toutes les probabilités doit valoir 1, et on trouve $a = \frac{4}{n^2}(n+1)^2$.
2. Par la formule des probabilités totales appliquée au système complet d'événements $\{[Y=j] \mid j \in [\![1,n]\!]\}$, on a pour tout $i \in [\![1,n]\!]$:

$$\begin{aligned}\mathbb{P}(X=i) &= \sum_{j=1}^{n}\mathbb{P}(X=i\cap Y=j)\\ &= \frac{4i}{n^2(n+1)^2}\sum_{j=1}^{n} j\\ &= \frac{4i}{n^2(n+1)^2}\frac{1}{2}(n(n+1))\\ &= \frac{2i}{n(n+1)}\end{aligned}$$

Par symétrie, on a de même pour tout $j \in [\![1,n]\!]$,

$$\mathbb{P}(Y=j)=\frac{2j}{n(n+1)}.$$

L'espérance de X et Y vaut alors

$$\mathbb{E}(X)=\mathbb{E}(Y)=\sum_{i=1}^{n}\frac{2i^2}{n(n+1)}=\frac{2}{n(n+1)}\frac{1}{6}n(n+1)(2n+1)=\frac{1}{3}(2n+1).$$

3. Il est clair que X et Y sont indépendantes, et donc leur covariance est nulle.
4. On a,

$$\begin{aligned}\mathbb{P}(X=Y)&=\sum_{k=1}^{n}\mathbb{P}(X=k\cap Y=k)\\&=\frac{4}{n^2(n+1)^2}\sum_{k=1}^{n}k^2\\&=\frac{2(2n+1)}{3n(n+1)}\end{aligned}$$

5. On a par indépendance

$$\mathbb{P}(M\leqslant k)=\mathbb{P}(X\leqslant k)\mathbb{P}(Y\leqslant k)=\left(\frac{k(k+1)}{n(n+1)}\right)^2.$$

Exercice 4

Soient X une variable aléatoire définie sur un espace probabilisé fini $(\Omega\mathcal{T},\mathbb{P})$ et f une application définie sur $X(\Omega)$. À quelle condition les variables aléatoires X et $f(X)$ sont-elles indépendantes ?

Réponse de l'exercice

Supposons X et Y indépendantes. Soit x tel que $\mathbb{P}(X=x)\neq 0$. On a alors

$$\mathbb{P}_{[X=x]}(f(X)=f(x))=1=\mathbb{P}(f(X)=f(x)).$$

Ainsi, la variable $f(X)$ est une variable constante.

Réciproquement, si $f(X)$ est constante, alors les variables X et $f(X)$ sont indépendantes.

Exercice 5

Soient X, Y et Z trois variables aléatoires mutuellement indépendantes et définies sur le même espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$. On suppose que X, Y et Z suivent la loi uniforme sur $[\![1,n]\!]$.

1. a) Montrer que :

$$\forall k\in[\![2,n+1]\!],\ \mathbb{P}(X+Y=k)=\frac{k-1}{n^2}.$$

   b) Montrer que :

$$\forall k\in[\![n+2,2n]\!],\ \mathbb{P}(X+Y=k)=\frac{2n-k+1}{n^2}.$$

2. En déduire que

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

3. a) Montrer que la variable $T = n + 1 - Z$ suit la loi uniforme sur $[\![1, n]\!]$.
   b) En déduire la probabilité $\mathbb{P}(X + Y + Z = n + 1)$.

Réponse de l'exercice

1. On note que pour tout $k \in [\![2, 2n]\!]$, on a par indépendance

$$\mathbb{P}(X + Y = k) = \sum_{i=1}^{n} \mathbb{P}(X = i)\mathbb{P}(Y = k - i).$$

   a) Si $k \leqslant n + 1$, on a toujours $k - i \leqslant n$; pour avoir $k - i \in [\![1, n]\!]$, il faut donc avoir $i \leqslant k - 1$. On a donc

$$\mathbb{P}(X + Y = k) = \sum_{i=1}^{k-1} \frac{1}{n^2} = \frac{k-1}{n^2}.$$

   b) De même, si $k \geqslant n + 2$, alors $k - i$ est toujours supérieur ou égal à 1 ; pour avoir $k - i \leqslant n$, il faut donc que $i \geqslant k - n$. On a donc

$$\mathbb{P}(X + Y = k) = \sum_{i=k-n}^{n} \frac{1}{n^2} = \frac{2n-k+1}{n^2}.$$

2. On applique la formule des probabilités totales au système complet d'événements $\{[Z = k] \mid k \in [\![1, n]\!]\}$:

$$\begin{aligned}
\mathbb{P}(X + Y = Z) &= \sum_{k=1}^{n} \mathbb{P}_{[Z=k]}(X + Y = Z)\mathbb{P}(Z = k) \\
&= \sum_{k=1}^{n} \frac{k-1}{n^3} \\
&= \frac{1}{n^3}\frac{1}{2}n(n-1) \\
&= \frac{n-1}{2n^2}
\end{aligned}$$

3. a) L'univers image de T est bien $[\![1, n]\!]$, et pour tout k dans cet intervalle,

$$\mathbb{P}(T = k) = \mathbb{P}(Z = n + 1 - k) = \frac{1}{n}$$

   car $n + 1 - k \in [\![1, n]\!]$.

   b) On a donc

$$\begin{aligned}
\mathbb{P}(X + Y + Z = n + 1) &= \mathbb{P}(X + Y = T) \\
&= \frac{n-1}{2n^3}
\end{aligned}$$

   car X, Y et T sont indépendantes, et que T et Z suivent la même loi.

Exercice 6

Soit $n \in \mathbb{N}$, $n \geqslant 2$. On tire sans remise deux jetons dans une urne contenant n jetons numérotés de 1 à n. On appelle X le numéro du premier, et Y le numéro du second.

1. Déterminer les lois de X et Y.
2. Les variables X et Y sont-elles indépendantes ?
3. Calculer la covariance de X et Y.
4. On note $Z = |X - Y|$. Calculer la loi de Z, puis son espérance.

Réponse de l'exercice

1. Les variables X et Y suivent une loi uniforme sur $[\![1, n]\!]$.
2. On a $\mathbb{P}(X = 1 \cap Y = 1) = 0 \neq \mathbb{P}(X = 1)\mathbb{P}(Y = 1)$. Les variables ne sont donc pas indépendantes.
3. On a pour tous $i, j \in [\![1, n]\!]$,

$$\mathbb{P}(X = i \cap Y = j) = \begin{cases} \dfrac{1}{n(n-1)} & \text{si } i \neq j \\ 0 & \text{sinon} \end{cases}.$$

Ainsi,

$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{i=1}^{n}\sum_{j=1}^{n} ij\mathbb{P}(X = i \cap Y = j) \\
&= \frac{1}{n(n-1)} \sum_{i \neq j} ij \\
&= \frac{1}{n(n-1)} \left(\sum_{i=1}^{n}\sum_{j=1}^{n} ij - \sum_{i=1}^{n} i^2 \right) \\
&= \frac{1}{n(n-1)} \left(\frac{n^2(n+1)^n}{4} - \frac{n(n+1)(2n+1)}{6} \right) \\
&= \frac{(n+1)(3n+2)}{12}
\end{aligned}$$

On trouve alors

$$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{-n-1}{12}.$$

4. On a $\mathrm{Im}(Z) = [\![1, n-1]\!]$. On a alors pour k dans cet ensemble

$$\begin{aligned}
\mathbb{P}(Z = k) &= \sum_{|i-j|=k} \mathbb{P}(X = i \cap Y = j) \\
&= \sum_{i=1}^{n-k} \mathbb{P}(X = i \cap Y = i + k) + \sum_{i=k+1}^{n} \mathbb{P}(X = i \cap Y = i - k) \\
&= \frac{2(n-k)}{n(n-1)}
\end{aligned}$$

On en déduit l'espérance de Z :

$$\begin{aligned}
\mathbb{E}(Z) &= \sum_{k=1}^{n-1} k \frac{2(n-k)}{n(n-1)} \\
&= \frac{n+1}{3}
\end{aligned}$$

Exercice 7

Soient X et Y deux variables de Bernoulli de paramètres respectifs p et q.

1. Montrer que $|\operatorname{Cov}(X,Y)| \leqslant \frac{1}{4}$.
2. Montrer qu'elles sont indépendantes si et seulement si leur covariance est nulle.

Réponse de l'exercice

1. On sait par l'inégalité de Cauchy-Schwarz que $|\operatorname{Cov}(X,Y)| \leqslant \sqrt{\mathbb{V}(X)\mathbb{V}(Y)}$. Or pour des variables de Bernoulli, on a
$$\mathbb{V}(X) = p(1-p) \quad \text{et} \quad \mathbb{V}(Y) = q(1-q).$$
La fonction $x \in [0,1] \mapsto x(1-x)$ atteint son maximum en $\frac{1}{2}$, qui vaut $\frac{1}{4}$, et on a bien le résultat voulu.
2. On note que pour des variables de Bernoulli, on a
$$\operatorname{Cov}(X,Y) = \mathbb{P}(X=1 \cap Y=1) - \mathbb{P}(X=1)\mathbb{P}(Y=1).$$
Ainsi, si la covariance est nulle, les événements $[X=1]$ et $[Y=1]$ sont bien indépendants. Les trois autres cas s'en déduisent par complémentaire.

Exercice 8

Dans cet exercice, n est un entier naturel non nul.

Soit $X_1, \dots, X_n$ n variables à densité, indépendantes et de même fonction de répartition F, définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Pour tout $\omega \in \Omega$, on ordonne les valeurs $X_1(\omega), \dots, X_n(\omega)$ et, pour tout $k \in [\![1,n]\!]$, on note $Y_k(\omega)$ la k-ème plus petite valeur. On a donc $Y_1(\omega) \leqslant Y_2(\omega) \leqslant \cdots \leqslant Y_n(\omega)$.

En particulier, on a $Y_1 = \min(X_1, \dots, X_n)$, $Y_n = \max(X_1, \dots, X_n)$.

1. Dans cette question uniquement, on suppose que les variables $X_1, \dots, X_n$ suivent la même loi exponentielle de paramètre $\lambda > 0$.
   a) Calculer $\mathbb{P}(Y_1 > x)$ pour tout réel x positif et en déduire la fonction de répartition de Y_1. Reconnaître une loi usuelle dont on donnera l'espérance et la variance.
   b) Montrer que si U est une variable qui suit la loi uniforme sur $]0,1]$ alors $\frac{-1}{\lambda}\ln(U)$ suit la loi exponentielle de paramètre $\lambda > 0$.
   c) Écrire un programme qui, pour un $n \in \mathbb{N}$ et un $i \in [\![1,n]\!]$ donnés, permet de simuler la variable aléatoire Y_i lorsque les variables $X_1, \dots, X_n$ suivent indépendamment la même loi exponentielle de paramètre $\lambda > 0$. On pourra pour cela utiliser l'instruction `B=sorted(A)` qui fournit un tableau `B` contenant les valeurs du tableau `A` rangées dans l'ordre croissant.

On retourne maintenant au cas général.

2. Exprimer la fonction de répartition de Y_n à l'aide de F.

3. Les variables Y_1 et Y_n sont-elles indépendantes ?

4. On souhaite maintenant obtenir la fonction de répartition de Y_i, pour n'importe quel i dans $[\![1, n]\!]$. On fixe donc i dans $[\![1, n]\!]$ et x dans $\mathbb{R}$ et on cherche à calculer $\mathbb{P}(Y_i \leqslant x)$. C'est la probabilité qu'au moins i variables parmi $X_1, \ldots, X_n$ soient inférieures ou égales à x.

   a) Pour tout $k \in [\![1, n]\!]$, on note Z_k la variable telle que $Z_k(\omega) = 1$ si $X_k(\omega) \leqslant x$ et $Z_k(\omega) = 0$ sinon. Reconnaître la loi de Z_k (on exprimera le(s) paramètre(s) à l'aide de $F(x)$).

   b) On note $S = \sum_{k=1}^{n} Z_k$. Que représente S ? Reconnaître sa loi.

   c) Montrer que $\mathbb{P}(Y_i \leqslant x) = \mathbb{P}(S \geqslant i)$ et en déduire l'expression de $\mathbb{P}(Y_i \leqslant x)$ sous la forme d'une somme que l'on ne cherchera pas à simplifier.

Réponse de l'exercice

1. a) Soit $x \in \mathbb{R}_+$. On sait qu'un minimum est supérieur à x si et seulement si toutes les valeurs sont supérieures à x, et donc

$$\begin{aligned}\mathbb{P}(Y_1 > x) &= \mathbb{P}\left(\bigcap_{i=1}^{n} X_i > x\right)\\ &= \prod_{i=1}^{n} \mathbb{P}(X_i > x) \quad \text{par indépendance}\\ &= e^{-\lambda n x}\end{aligned}$$

   De plus, il est clair que $\mathbb{P}(Y_1 > x) = 1$ pour tout $x < 0$, donc Y_1 suit une loi exponentielle de paramètre λn.

   b) Soit donc U une variable aléatoire de loi uniforme sur $]0, 1]$. On pose $Z = \frac{-1}{\lambda} \ln(U)$.
   Alors Z est une variable presque sûrement positive, donc pour tout $x < 0$, $\mathbb{P}(Z \leqslant x) = 0$.
   Soit alors $x \in \mathbb{R}_+$. On a alors

$$\begin{aligned}\mathbb{P}(Z \leqslant x) &= \mathbb{P}(U \leqslant 1 - e^{-\lambda x})\\ &= 1 - e^{-\lambda x}\end{aligned}$$

   car $1 - e^{-\lambda x} \in]0, 1]$.
   Ainsi, Z suit bien une loi exponentielle de paramètre λ.

   c)
```python
import random as rd
from math import log

def expo(l):
    """
    Simule une loi exponentielle de parametre l
    """
    u = rd.random()
    return (-1/l)*log(u)

def Y(n,i,l):
    X = [expo(l) for _ in range(n)]
```

```
    Xsorted = sorted(X)
    return Xsorted[i-1] #attention au decalage d'indice

```

2. Soit $x \in \mathbb{R}$. On a alors

$$\begin{aligned}
\mathbb{P}(Y_n \leqslant x) &= \mathbb{P}\left(\bigcap_{i=1}^{n} X_i \leqslant x\right) \\
&= \prod_{i=1}^{n} \mathbb{P}(X_i \leqslant n) \quad \text{par indépendance} \\
&= F(x)^n
\end{aligned}$$

3. De même qu'en 1a, on peut trouver :

$$\forall x \in \mathbb{R},\ \mathbb{P}(Y_1 > x) = (1 - F(x))^n.$$

On sait que F est continue sur $\mathbb{R}$, tend vers 0 en $-\infty$ et vers 1 en $+\infty$, et donc on peut trouver $x \in \mathbb{R}$ tel que $F(x) = \frac{1}{2}$.
On a alors

$$\mathbb{P}(Y_1 > x \cap Y_n \leqslant x) = 0 \neq \mathbb{P}(Y_1 > x)\mathbb{P}(Y_n \leqslant x).$$

Les variables Y_1 et Y_n ne sont donc pas indépendantes.

4. a) L'univers image de Z_k est $\{0, 1\}$, et donc Z_k suit une loi de Bernoulli. On a $\mathbb{P}(Z_k = 1) = \mathbb{P}(X_k \leqslant x) = F(x)$, et donc Z_k suit une loi de Bernoulli de paramètre $F(x)$.

   b) La variable S représente donc le nombre de variables parmi $X_1, \dots, x_n$ qui sont inférieures ou égales à x.

   De plus, les X_k étant indépendantes, les Z_k le sont aussi, et donc S suit une loi binomiale de paramètres n et $F(x)$.

   c) D'après l'énoncé et la question précédente, on a donc bien $\mathbb{P}(Y_i \leqslant x) = \mathbb{P}(S \geqslant i)$, et on a donc

$$\mathbb{P}(Y_i \leqslant x) = \sum_{k=i}^{n} \mathbb{P}(S = k) = \sum_{k=i}^{n} \binom{n}{k} F(x)^k (1 - F(x))^{n-k}.$$