Chapitre 7

Couples de variables aléatoires discrètes

Compétences attendues.

- ✓ Obtenir la loi d'un couple (X, Y), ses lois marginales.
- ✓ Déterminer la loi de $X + Y$ à l'aide du produit de convolution discret.
- ✓ Déterminer la loi de $\min(X, Y)$ et $\max(X, Y)$.
- ✓ Calculer une covariance.
- ✓ Étudier l'indépendance de deux variables aléatoires discrètes.

Mathieu Mansuy - Professeur de Mathématiques en ECS2 au Lycée Louis Pergaud (Besançon)
mansuy.mathieu@hotmail.fr

1 Indépendance de deux variables aléatoires

1.1 Indépendance d'évènements

Définition.

Deux événements A et B sont dits *indépendants* lorsque $P(A \cap B) = P(A) \times P(B)$.

Remarques.

- Si $P(A) \neq 0$, A et B sont indépendants si et seulement si $P(B) = P_A(B)$.
- La plupart du temps, l'indépendance se déduit de la situation. Deux événements sont indépendants lorsque la réalisation de l'un ne donne pas d'information sur la réalisation de l'autre.

1.2 Indépendance de deux variables aléatoires

Dans cette section, X et Y sont deux variables aléatoires (discrètes ou non) d'un espace probabilisé $(\Omega, \mathscr{A}, P)$.

Définition.

On dit que les variables aléatoires X et Y sont *indépendantes* lorsque pour tout $(x, y) \in \mathbb{R}^2$, les évènements $[X \leq x]$ et $[Y \leq y]$ sont indépendants, c'est à dire :

$$P([X \leq x] \cap [Y \leq y]) = P(X \leq x)P(Y \leq y).$$

Propriété 1

Soient X et Y deux variables aléatoires (discrètes ou non) d'un espace probabilisé $(\Omega, \mathscr{A}, P)$. Alors il y a équivalence entre :

(1) X et Y sont indépendantes ;

(2) pour tous intervalles I et J de $\mathbb{R}$,

$$P([X \in I] \cap [Y \in J]) = P(X \in I)P(Y \in J)$$

(3) pour tout événement $A \in \mathscr{A}_X$ [a] et tout événement $B \in \mathscr{A}_Y$:

$$P(A \cap B) = P(A)P(B)$$

[a] Rappelons que $\mathscr{A}_X$ est la plus petite tribu contenant tous les évènements $[X \leq x]$, $x \in \mathbb{R}$. Elle représente l'information fournie par X.

Remarques.

- L'indépendance de deux variables aléatoires se déduit souvent de la situation.
- On n'aura jamais besoin de revenir aux tribus. Il faut surtout retenir que l'indépendance de X et de Y signifie que tout événement formé à partir de X est indépendant de tout événement formé à partir de Y.

Propriété 2

Si X et Y sont indépendantes alors pour toutes fonctions f et g dont les ensembles de définition contiennent respectivement $X(\Omega)$ et $Y(\Omega)$, les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

1.3 Indépendance de variables aléatoires discrètes

Propriété 3

Les variables aléatoires discrètes X et Y sont indépendantes si et seulement si pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, on a :

$$P([X = x] \cap [Y = y]) = P(X = x) \times P(Y = y).$$

Propriété 4

Soient X et Y deux variables aléatoires discrètes indépendantes d'un espace probabilisé $(\Omega, \mathscr{A}, P)$, et $B \in \mathscr{A}_Y$ de probabilité non nulle. Alors X admet une espérance si et seulement si $E(X \mid B)$ existe, et alors

$$E(X \mid B) = E(X)$$

Preuve. On fixe une numérotation du support $X(\Omega) = \{x_i,\ i \in \mathbb{N}\}$. $E(X \mid B)$ existe si et seulement si la série $\sum x_i P_B(X = x_i)$ converge absolument. Or X et Y étant indépendantes et $B \in \mathscr{A}_Y$, on a $P_B(X = x_i) = P(X = x_i)$. Ainsi $E(X \mid B)$ existe si et seulement si $\sum x_i P(X = x_i)$ converge absolument, c'est à dire si et seulement si $E(X)$ existe. De plus, on a :

$$E(X) = E(X \mid B).$$

□

Remarque. En pratique, B est de la forme $[Y \leq y]$ ou $[Y = y]$.

2 Généralité sur les couples de variables aléatoires réelles

Dans cette section, X et Y sont deux variables aléatoires (discrètes ou non) d'un espace probabilisé $(\Omega, \mathscr{A}, P)$.

Définition.

On appelle *couple des variables aléatoires X et Y* et on note (X, Y) l'application :

$$(X, Y)\ :\ \left\{ \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R}^2 \\ \omega & \longmapsto & (X(\omega), Y(\omega)). \end{array} \right.$$

Définition.

On appelle *tribu associée au couple aléatoire* (X, Y) la tribu notée $\mathscr{A}_{(X,Y)}$ engendrée par les événements $([X \leqslant x] \cap [Y \leqslant y])_{(x,y)\in\mathbb{R}^2}$. C'est la plus petite tribu contenant tous ces évènements.

Définition.

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

$$\forall (x,y) \in \mathbb{R}^2,\ F_{(X,Y)}(x,y) = P([X \leqslant x] \cap [Y \leqslant y])$$

- La loi de X est appelée la *première loi marginale* du couple (X,Y), et la loi de Y la *seconde loi marginale* de (X,Y).

Remarque. On a $(X,Y)(\Omega) \subset X(\Omega)\times Y(\Omega)$. Ainsi on ne donnera pas en général le support $(X,Y)(\Omega)$ du couple (X,Y), mais plutôt les supports $X(\Omega)$ et $Y(\Omega)$ de X et Y.

Propriété 5

Les variables X et Y sont indépendantes si et seulement si

$$\forall (x,y) \in \mathbb{R}^2, \quad F_{(X,Y)}(x,y) = F_X(x)F_Y(y),$$

où F_X, F_Y sont les fonctions de répartition de X et Y.

Propriété 6

On suppose que :

- les couples (X_1,Y_1) et (X_2,Y_2) ont la même loi conjointe,
- g est une fonction continue[a] sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$.

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

[a] Pour la définition de la continuité d'une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$, on renvoie au **Chapitre 15. Fonctions de plusieurs variables**.

3 Couples de variables aléatoires discrètes

Dans toute la suite, X et Y sont des variables discrètes dont on fixera une numérotation de leurs supports (avec $I = [\![1,n]\!]$ ou $\mathbb{N}$ et $J = [\![1,m]\!]$ ou $\mathbb{N}$ selon que X et Y soient finies ou non) :

$$X(\Omega) = \{x_i,\ i \in I\} \quad \text{et} \quad Y(\Omega) = \{y_j,\ j \in J\}.$$

3.1 Loi d'un couple de variables discrètes

Propriété 7

La loi du couple (X,Y) est entièrement déterminée par la donnée de :

- $X(\Omega) = \{x_i,\ i \in I\}$ et $Y(\Omega) = \{y_j,\ j \in J\}$;
- $p_{i,j} = P([X = x_i] \cap [Y = y_j])$ pour tout $(i,j) \in I \times J$.

Notation. $[X = x_i] \cap [Y = y_j] = [X = x_i, Y = y_j] = [(X,Y) = (x_i, y_j)]$

Remarque. Lorsque X et Y sont **finies**, la loi du couple (X,Y) peut être représentée par un tableau à double entrée.

$X \backslash Y$	y_1	y_2	$\cdots$	y_m
x_1	$p_{1,1} = P([X = x_1] \cap [Y = y_1])$	$p_{1,2}$	$\cdots$	$p_{1,m}$
x_2	$p_{2,1}$	$p_{2,2}$	$\cdots$	$p_{2,m}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
x_n	$p_{n,1}$	$p_{n,2}$	$\cdots$	$p_{n,m}$

Exercice. Considérons l'expérience aléatoire consistant à jeter deux dés, X la variable aléatoire donnant le plus petit résultat et Y le plus grand. Déterminer la loi conjointe de (X,Y).

Propriété 8

La famille $([X = x_i] \cap [Y = y_j])_{(i,j)\in I\times J}$ est un système complet d'événements. Par conséquent, la famille $(P([X = x_i] \cap [Y = y_j]))_{(i,j)\in I\times J}$ est sommable et on a :

$$\sum_{(i,j)\in I\times J} P([X = x_i] \cap [Y = y_j]) = \sum_{(i,j)\in I\times J} p_{i,j} = 1.$$

Remarque. Lorsque la loi de (X,Y) est représentée par un tableau, la somme de toutes les probabilités du tableau vaut 1.

Propriété 9

Soient I et J des ensembles finis ou dénombrables, et soit $(p_{i,j})_{(i,j)\in I\times J}$ une famille de réels satisfaisant :

(1) $\forall (i,j) \in I \times J$, $p_{i,j} \geq 0$;

(2) la famille $(p_{i,j})_{(i,j)\in I\times J}$ est sommable et sa somme vaut 1.

Alors la famille $(p_{i,j})_{(i,j)\in I\times J}$ définit une *loi conjointe de probabilité*, c'est à dire qu'il existe un couple de variables discrètes (X,Y) sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ tel que :

$$X(\Omega) \times Y(\Omega) = I \times J \quad \text{et} \quad \forall (i,j) \in I \times J,\ P([X = i] \cap [Y = j]) = p_{i,j}.$$

3.2 Lois marginales

Propriété 10

Soient (X,Y) un couple de variables aléatoire. Alors on a :

$$\forall i \in I, \quad p_i = P(X = x_i) = \sum_{j \in J} P([X = x_i] \cap [Y = y_j]) = \sum_{j \in J} p_{i,j}$$

$$\forall j \in J, \quad q_j = P(Y = y_j) = \sum_{i \in I} P([X = x_i] \cap [Y = y_j]) = \sum_{i \in I} p_{i,j}.$$

Preuve. La famille $([Y = y_j])_{j \in J}$ est un système complet d'événements. La formule des probabilités totales donne donc :

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

$\square$

Remarque. Lorsque la loi de (X,Y) est **finie**, représentée par un tableau, on obtient $P(X = x_i)$ en sommant les termes de la i-ème ligne et $P(Y = y_j)$ en sommant les termes de la j-ème colonne :

X \ Y	y_1	y_2	$\dots$	y_m	Somme par lignes $1^{\text{ère}}$ loi marginale $\downarrow$
x_1	$P([X = x_1] \cap [Y = y_1])$	$P([X = x_1] \cap [Y = y_2])$	$\dots$	$P([X = x_1] \cap [Y = y_m])$	$P([X = x_1])$
x_2	$P([X = x_2] \cap [Y = y_1])$	$P([X = x_2] \cap [Y = y_2])$	$\dots$	$P([X = x_2] \cap [Y = y_m])$	$P([X = x_2])$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_n	$P([X = x_n] \cap [Y = y_1])$	$P([X = x_n] \cap [Y = y_2])$	$\dots$	$P([X = x_n] \cap [Y = y_m])$	$P([X = x_n])$
Somme par colonnes $2^{\text{ème}}$ loi marginale $\rightarrow$	$P([Y = y_1])$	$P([Y = y_2])$	$\dots$	$P([Y = y_m])$	Somme totale $= 1$

Exercice. Déterminons les lois marginales du couple (X,Y) dans l'exercice précédent.

X \ Y	$Y = 1$	$Y = 2$	$Y = 3$	$Y = 4$	$Y = 5$	$Y = 6$	
$X = 1$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	
$X = 2$	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	
$X = 3$	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	
$X = 4$	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	
$X = 5$	0	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$	
$X = 6$	0	0	0	0	0	$\frac{1}{36}$	

Remarque. Lien entre loi conjointe et lois marginales.

- On vient de voir qu'à partir de la loi conjointe de (X,Y), on peut aisément obtenir les lois marginales de X et de Y. En revanche, il n'est en général **pas possible de retrouver la loi conjointe à partir des lois marginales** (voir le tableau ci-dessus : il n'y a pas de lien entre les p_i, q_j et les $p_{i,j}$).
- Il y a un cas particulier cependant où c'est possible : lorsque les variables X et Y sont **indépendantes**, on a :
$$p_{i,j} = P([X = x_i] \cap [Y = y_j]) = P([X = x_i])P(Y = y_j]) = p_i \times q_j.$$
Dans le cas **indépendant**, la loi conjointe se retrouve à partir des lois marginales par produit.

Exercice. Soit $p \in]0,1[$, et soit la famille $(p_{m,n})_{(m,n)\in\mathbb{N}\times\mathbb{N}}$ définie par
$$\forall (m,n) \in \mathbb{N}^2,\ p_{m,n} = \begin{cases} \dfrac{e^{-2}}{m!(n-m)!} & \text{si } n \geq m \\ 0 & \text{sinon.} \end{cases}$$
Montrer que $(p_{m,n})_{(m,n)\in\mathbb{N}\times\mathbb{N}}$ est la loi conjointe d'un couple (X,Y) de variables aléatoires discrètes, et déterminer les lois marginales. X et Y sont-elles indépendantes ?

3.3 Lois conditionnelles

Définition.

Soit (X,Y) un couple de variables aléatoires discrètes, et soit $i \in I$ tel que $P(X = x_i) \neq 0$. On appelle *loi conditionnelle de Y sachant que $[X = x_i]$* la donnée de $Y([X = x_i])$ et des probabilités
$$\forall j \in J,\ P_{[X=x_i]}(Y = y_j) = \frac{P([X = x_i] \cap [Y = y_j])}{P(X = x_i)} = \frac{p_{i,j}}{p_i}.$$
On définit de même la *loi conditionnelle de X sachant que $[Y = y_j]$*.

Propriété 11

On suppose que pour tout $(i,j) \in I \times J$, $P(X = x_i), P(Y = y_j) \neq 0$. On a pour tout $(i,j) \in I \times J$,

$$P([X = x_i] \cap [Y = y_j]) = P(Y = y_j)P_{[Y=y_j]}(X = x_i) = P(X = x_i)P_{[X=x_i]}(Y = y_j)$$

$$P(X = x_i) = \sum_{j \in J} P(Y = y_j)P_{[Y=y_j]}(X = x_i)$$

$$P(Y = y_j) = \sum_{i \in I} P(X = x_i)P_{[X=x_i]}(Y = y_j).$$

Preuve. Les premières égalités résultent de la définition des probabilités conditionnelles, les deux dernières formules résultent de la formule des probabilités totales appliquée aux systèmes complets d'événements $([Y = y_j])_{j \in J}$ et $([X = x_i])_{i \in I}$. □

Exercice. Soit N une variable aléatoire de loi $\mathscr{P}(\lambda)$ et soit X une variable aléatoire telle que la loi de X sachant $[N = n]$ est la loi $\mathscr{B}(n, \frac{1}{2})$.

a) Déterminer la loi conjointe de (N, X).

b) En déduire la loi marginale de X.

4 Variable aléatoire fonction de deux variables aléatoires discrètes

4.1 Fonction d'un couple de variables aléatoires discrètes

Propriété 12

Soit (X, Y) un couple de variables aléatoires discrètes, et $g : X(\Omega) \times Y(\Omega) \to \mathbb{R}$ une fonction. Alors $g(X, Y)$ est une variable aléatoire discrète définie sur $(\Omega, \mathscr{A})$.

Exemple. $X + Y$, XY, $\max(X, Y)$ et $\min(X, Y)$ sont des variables aléatoires discrètes sur $(\Omega, \mathscr{A}, P)$.

Remarque. Si $Z = g(X,Y)$, on a $\mathscr{A}_Z \subset \mathscr{A}_{(X,Y)}$. En effet pour tout $z \in \mathbb{R}$,

$$[Z = z] = \bigcup_{\substack{(x,y)\in X(\Omega)\times Y(\Omega),\\ g(x,y)=z}} [X = x] \cap [Y = y] \in \mathscr{A}_{(X,Y)} \quad \text{(union dénombrable d'évènements)}$$

Puisque $\mathscr{A}_Z$ est la plus petite tribu contenant les évènements $[Z = z]$, on a donc bien $\mathscr{A}_Z \subset \mathscr{A}_{(X,Y)}$.

Propriété 13 (Loi de $g(X,Y)$)

On note $Z = g(X,Y)$. Pour tout $z \in Z(\Omega)$, on a :

$$P(Z = z) = \sum_{\substack{(x,y)\in X(\Omega)\times Y(\Omega),\\ g(x,y)=z}} P([X = x] \cap [Y = y]).$$

Théorème 14 (*de transfert double* : Espérance de $g(X,Y)$)

La variable aléatoire $Z = g(X,Y)$ admet une espérance si et seulement si la famille $(g(x_i, y_j)P([X = x_i] \cap [Y = y_j]))_{(i,j)\in I\times J}$ est sommable.
Dans ce cas, on a :

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

Corollaire 15 (Linéarité de l'espérance)

Soient X et Y des variables aléatoires admettant une espérance. Alors pour tous $\lambda, \mu \in \mathbb{R}$, $\lambda X + \mu Y$ admet une espérance, et on a :

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

Preuve. Soit $g : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par $g(x,y) = \lambda x + \mu y$. D'après le théorème de transfert, $Z = g(X,Y)$ admet une espérance si et seulement si la famille des

$$g(x_i, y_j)P([X = x_i] \cap [Y = y_j]) = \lambda x_i P([X = x_i] \cap [Y = y_j]) + \mu y_j P([X = x_i] \cap [Y = y_j])$$

est sommable. Or on a :

- la série $\sum_j |x_i| P([X = x_i] \cap [Y = y_j])$ converge d'après la formule des probabilités totales, et on a :
$$\sum_{j\in J} |x_i| P([X = x_i] \cap [Y = y_j]) = |x_i| P(X = x_i).$$
- la série $\sum_i |x_i| P([X = x_i])_{i\in I}$ est convergente car X admet une espérance.

Par le théorème de Fubini, la famille $(|x_i| P([X = x_i] \cap [Y = y_j]))_{(i,j)\in I\times J}$ est sommable, et on a :

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

De même on montre que la famille $(y_j P([X = x_i] \cap [Y = y_j]))_{(i,j)\in I\times J}$ converge absolument, et sa somme est égale à $E(Y)$. On peut donc conclure que la famille $(g(x_i, y_j)P([X = x_i] \cap [Y = y_j]))_{(i,j)\in I\times J}$ est sommable comme somme de deux familles sommables. $E(Z)$ existe donc bien, et on a :

$$\begin{aligned}
E(Z) &= \sum_{(i,j)\in I\times J} g(x_i, y_j)P([X = x_i] \cap [Y = y_j]) \\
&\underset{\text{tout converge}}{=} \lambda \sum_{(i,j)\in I\times J} x_i P([X = x_i] \cap [Y = y_j]) + \mu \sum_{(i,j)\in I\times J} y_j P([X = x_i] \cap [Y = y_j]) \\
&= \lambda E(X) + \mu E(Y).
\end{aligned}$$

$\square$

4.2 Somme de deux variables indépendantes : le produit de convolution

Propriété 16 (*Produit de convolution discret*)

Soient X et Y des variables aléatoires discrètes. Alors la loi de $Z = X + Y$ est donnée par :

$$\forall z \in (X+Y)(\Omega), \quad P(X+Y = z) = \sum_{\substack{x\in X(\Omega),\\ z-x\in Y(\Omega)}} P(X = x, Y = z - x).$$

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

Preuve.

$\square$

Corollaire 17

Si X et Y sont deux variables aléatoires **indépendantes** à valeurs dans $\mathbb{N}$, alors on a :

$$\forall n \in \mathbb{N}, \quad P(X+Y = n) = \sum_{k=0}^{n} P(X = k)P(Y = n - k).$$

Preuve.

$\square$

Exercice. On lance indéfiniment une pièce de monnaie truquée de sorte que le côté pile est obtenu avec la probabilité p (avec $0 < p < 1$). On note X le rang du premier pile obtenu et Y celui du deuxième.

a) Déterminer les lois de X et $Y - X$.

b) Justifier que X et $Y - X$ sont indépendantes. En déduire la loi de Y.

4.3 Stabilité des lois binomiales et de Poisson

Propriété 18 (*Stabilité des lois de Poisson*)

Hypothèses : $X \hookrightarrow \mathscr{P}(\lambda)$ et $Y \hookrightarrow \mathscr{P}(\mu)$
X et Y sont indépendantes

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

Preuve.

$\square$

Lemme. Pour tout $n_1, n_2 \in \mathbb{N}$, pour tout $0 \leq l \leq n_1 + n_2$, on a :

$$\sum_{k=0}^{n_1} \binom{n_1}{k} \binom{n_2}{l-k} = \binom{n_1 + n_2}{l} \qquad \textit{(formule de Vandermonde)}$$

en prenant pour convention que $\displaystyle\binom{n}{k} = 0$ si $k \notin [\![0, n]\!]$.

Preuve.

$\square$

Propriété 19 (*Stabilité des lois binomiales*)

Hypothèses :
- $X \hookrightarrow \mathscr{B}(m, p)$ et $Y \hookrightarrow \mathscr{B}(n, p)$
- X et Y sont indépendantes

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

Preuve.

$\square$

4.4 Minimum et maximum de deux variables

Propriété 20

Soient X et Y deux variables aléatoires discrètes, $Z = \max(X, Y)$. Alors on a :

- $\forall z \in \mathbb{R}$, $[Z \leq z] = [X \leq z] \cap [Y \leq z]$;
- Si de plus les variables X et Y sont **indépendantes**, on a :

$$\forall z \in \mathbb{R}, \quad P(Z \leq z) = P(X \leq z) \times P(Y \leq z).$$

Preuve.

□

Méthode.

*Pour déterminer la loi de $Z = \max(X, Y)$ dans le cas où X et Y sont **indépendantes** à valeurs dans $\mathbb{N}$, il faudra donc :*

- *calculer pour tout $n \in \mathbb{N}$, $P(X \leq n) = \sum_{k=0}^{n} P(X = k)$, et de même pour $P(Y \leq n)$;*
- *on obtient $P(Z \leq n) = P(X \leq n)P(Y \leq n)$ pour tout $n \in \mathbb{N}$;*
- *conclure en utilisant que pour tout $n \in \mathbb{N}$, $P(Z = n) = P(Z \leq n) - P(Z \leq n - 1)$.*

Exercice. Soient X et Y des variables aléatoires indépendantes suivant toutes deux une loi uniforme $\mathscr{U}(\llbracket 1, n \rrbracket)$. Déterminer la loi de $Z = \max(X, Y)$.

Propriété 21

Soient X et Y deux variables aléatoires discrètes, $Z = \min(X, Y)$. On a :

- $\forall z \in \mathbb{R}$, $[Z > z] = [X > z] \cap [Y > z]$;
- Si de plus les variables X et Y sont **indépendantes**, on a :

$$\forall z \in \mathbb{R}, \quad P(Z > z) = P(X > z) \times P(Y > z).$$

Preuve. Laissée en exercice.

□

Méthode.

*Pour déterminer la loi de $Z = \min(X, Y)$ dans le cas X, Y **indépendantes** à valeurs dans $\mathbb{N}$, il faudra donc :*

- *calculer pour tout $n \in \mathbb{N}$, $P(X > n) = \sum_{k=n+1}^{+\infty} P(X = k)$, et de même pour $P(Y > n)$;*
- *on obtient $P(Z > n) = P(X > n)P(Y > n)$ pour tout $n \in \mathbb{N}$;*
- *conclure en utilisant que pour tout $n \in \mathbb{N}$, $P(Z = n) = P(Z > n - 1) - P(Z > n)$.*

5 Covariance, corrélation linéaire

5.1 Espérance d'un produit

Lemme. Soient $x, y \in \mathbb{R}$. Alors on a :

$$|xy| \leq \frac{x^2 + y^2}{2}.$$

Preuve.

□

Propriété 22

Soient X, Y des variables aléatoires discrètes **admettant une variance**. Alors XY admet une espérance et on a :

$$E(XY) = \sum_{(i,j)\in I\times J} x_i y_j P([X = x_i] \cap [Y = y_j]).$$

Preuve.

□

Propriété 23

Soient X, Y des variables aléatoires discrètes et **indépendantes**. On suppose que X et Y admettent une **espérance**. Alors XY admet une espérance et :

$$E(XY) = E(X)E(Y).$$

Preuve. On applique le théorème de transfert double. Pour cela, on étudie si la famille $(x_i y_j P([X = x_i] \cap [Y = y_j]))_{(i,j)\in I\times J}$ est sommable. Pour tout $(i, j) \in I \times J$, on a :

$$\begin{aligned}
|x_i y_j P([X = x_i] \cap [Y = y_j])| &\underbrace{=}_{X \text{ et } Y \text{ indépendantes}} |x_i y_j| P([X = x_i]) P([Y = y_j]) \\
&= (|x_i| P(X = x_i))(|y_j| P(Y = y_j)).
\end{aligned}$$

Comme X et Y admettent une espérance, les séries des $(|x_i| P(X = x_i))$ et $(|y_j| P(Y = y_j))$ convergent. La famille est donc bien sommable, et on en déduit par le théorème de transfert que $E(XY)$ existe. De plus, on a :

$$\begin{aligned}
E(XY) &= \sum_{(i,j)\in I\times J} x_i y_j P([X = x_i] \cap [Y = y_j]) \\
&= \sum_{i\in I} \sum_{j\in J} x_i P([X = x_i]) \times y_j P([Y = y_j]) \\
&= \left(\sum_{i\in I} x_i P([X = x_i])\right) \times \left(\sum_{j\in J} y_j P([Y = y_j])\right) = E(X)E(Y)
\end{aligned}$$

□

5.2 Covariance

Définition.

Soient X et Y des variables aléatoires discrètes admettant une variance (ou de manière équivalente un moment d'ordre 2). On appelle *covariance de X et Y* le réel :

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

On dit que X et Y sont *non corrélées* si $Cov(X,Y) = 0$.

Remarque. $\mathrm{Cov}(X,Y)$ existe bien car $X - E(X)$ et $Y - E(Y)$ admettent toutes deux une variance, et donc leur produit admet une espérance.

Théorème 24 (*Formule de Huygens*)

Si X et Y admettent toutes deux une variance, alors on a :

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

Preuve.

□

Propriété 25 (*Covariance avec une variable aléatoire constante*)

Si X admet une variance et $Y = a$ $p.s.$ alors $\mathrm{Cov}(X,Y) = 0$.
En particulier, $\mathrm{Cov}(X,a) = 0$.

Preuve. Si $Y = a$ presque surement, alors $Y - E(Y) = Y - a = 0$ presque surement. Ainsi $(X - E(X))(Y - E(Y)) = 0$ presque surement et on a :

$$\mathrm{Cov}(X,Y) = E((X - E(X))(Y - E(Y))) = 0.$$

□

Propriété 26

On suppose que X, Y, Z, T admettent toutes une variance. On a alors :

- *Symétrie :* $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.
- *Bilinéarité* (i.e. linéaire par rapport à chacune des variables) :
  - *linéarité à gauche* : $\forall \lambda, \mu \in \mathbb{R}$, $\mathrm{Cov}(\lambda X + \mu Y, Z) = \lambda \mathrm{Cov}(X,Z) + \mu \mathrm{Cov}(Y,Z)$;
  - *linéarité à droite* : $\forall \lambda, \mu \in \mathbb{R}$, $\mathrm{Cov}(X, \lambda Z + \mu T) = \lambda \mathrm{Cov}(X,Z) + \mu \mathrm{Cov}(X,T)$;
- *Positivité :* pour toute variable aléatoire discrète X admettant une variance, on a :

$$\mathrm{Cov}(X,X) = V(X) \geq 0.$$

Preuve.

□

Propriété 27

Soient X et Y admettant toutes deux une variance.
Si X et Y sont **indépendantes**, **alors** $\text{Cov}(X, Y) = 0$.

Preuve.

□

Mise en garde.

Ainsi si $\text{Cov}(X, Y) \neq 0$, alors X et Y ne sont pas indépendantes. Mais attention, **la réciproque est fausse** : on peut avoir $\text{Cov}(X, Y) = 0$ avec X et Y non indépendantes

Exercice. Soient X et Y deux variables aléatoires indépendantes, suivant une même loi géométrique $\mathscr{G}(p)$.On pose $S = X + Y$ et $U = X - Y$.

a) Sans déterminer les lois de S et U, montrer que $\text{Cov}(S, U) = 0$.

b) Les variables S et U sont-elles indépendantes ?

Théorème 28

Soit (X, Y) un couple de variables aléatoires discrètes admettant chacune une variance.

- La variable aléatoire $X + Y$ admet une variance et on a :

$$V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y)$$

- Dans le cas où X et Y sont **indépendantes**, on a :

$$V(X + Y) = V(X) + V(Y)$$

Preuve.

□

Mise en garde.

La variance n'est pas linéaire : en général, on n'a pas $V(X+Y) = V(X) + V(Y)$.

Remarque. On retiendra en particulier que $\mathrm{Cov}(X,Y) = \dfrac{1}{2}(V(X+Y) - V(X) - V(Y))$, qui permet de calculer la covariance à partir des variances.

5.3 Coefficient de corrélation linéaire

Propriété 29 (*Inégalité de Cauchy-Schwarz*)

On suppose que X et Y admettent une variance. Alors on a :

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

Preuve.

□

Propriété 30 (*Cas d'égalité dans l'inégalité de Cauchy-Schwarz*)

On suppose que X et Y admettent une variance. On a l'équivalence :

$$[\ |\mathrm{Cov}(X,Y)| = \sigma(X)\,\sigma(Y)\ ]$$
$$\Updownarrow$$
$$[X \text{ est constante p.s.}] \ \text{ ou } \ \left[\exists (a,b) \in \mathbb{R}^2,\ Y = aX + b \ p.s.\right].$$

De plus si X n'est pas constante presque sûrement, alors a est du signe de $\mathrm{Cov}(X,Y)$.

Preuve.

□

Définition.

On suppose que X et Y admettent une variance non nulle. On appelle *cœfficient de corrélation linéaire de* (X, Y), et on note $\rho_{X,Y}$, le réel :

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

Remarque. Le coefficient de corrélation linéaire est la normalisation de la covariance par le produit des écarts types. L'intérêt est que X et Y ne s'expriment pas forcément dans la même unité (par exemple si X représente la température extérieure et Y la facture de chauffage). Cette normalisation permet d'obtenir un réel indépendant des unités de mesure des observations.

Propriété 31

On suppose que X et Y admettent une variance non nulle. On a :

(1) $\rho_{X,Y} \in [-1, 1]$,

(2) $\rho_{X,Y} = 1 \Leftrightarrow \exists (a, b) \in \mathbb{R}_+^* \times \mathbb{R}$ tel que $Y = aX + b$,

(3) $\rho_{X,Y} = -1 \Leftrightarrow \exists (a, b) \in \mathbb{R}_-^* \times \mathbb{R}$ tel que $Y = aX + b$.

Preuve.

□

Remarques.

- Le coefficient de corrélation linéaire permet donc de mesurer la dépendance **linéaire** qui peut exister entre un couple de variables aléatoires (X, Y).
  - Si $\rho_{X,Y}$ est proche de 1 ou de -1, la dépendance linéaire entre X et Y est importante.
    Exemple : X = volume des ventes, Y = nombre de remises promotionnelles.
  - si $\rho_{X,Y}$ est proche de 0 les deux variables sont **linéairement indépendantes**.
    Exemple : X = production de comté, Y = nombre de décès sur les routes de Franche Comté.

Mise en garde.

Attention encore une fois de ne pas confondre « indépendance linéaire » et « indépendance » : si $\rho_{X,Y} = 0$ (ou $\text{Cov}(X,Y) = 0$), alors X et Y sont linéairement indépendantes. Mais elles ne sont pas forcément indépendantes : il peut exister une dépendance non linéaire entre les variables, par exemple $Y = e^X, \ln(X), \ldots$

- Le signe de $\rho_{X,Y}$ (et de $\text{Cov}(X,Y)$) indique le sens d'évolution d'une variable par rapport à l'autre :
  - si $\rho_{X,Y} > 0$, alors les deux variables évoluent en moyenne dans le même sens.
    Exemple : X = température extérieure, Y = consommation de crèmes glacées.
  - si $\rho_{X,Y} < 0$, les deux variables évoluent en moyenne en sens inverse.
    Exemple : X = température extérieure, Y = facture de chauffage.

Le saviez vous ?

Le statisticien et chimiste anglais William Gosset (1876 - 1937) travaillait aux brasseries Guiness à Dublin. Dans le but d'améliorer la qualité du houblon, il l'étudie de manière statistique. Cela l'amène à introduire la notion de variance sous le nom de *fluctuation*. Son compatriote Ronald Fisher (1890 - 1962) préfère la nommer *variance*, mot qui a pour sens désaccord en anglais. Le préfixe *co* vient de la préposition latine *cum* signifiant *avec*. Ainsi *covariance* correspond à « désaccord avec ».

L'entreprise Guiness avait pour règle que ses chimistes ne publient pas leurs découvertes. Gosset argua que ses travaux ne seraient d'aucune utilité pour les concurrents et obtint l'autorisation de publier mais sous un pseudonyme, Student, pour éviter les difficultés avec les autres membres de son équipe. C'est ce nom qu'on retrouve aujourd'hui encore associé à de nombreux outils statistiques.