

RÉSUMÉ DE COURS : COUPLES DE VARIABLES ALÉATOIRES DISCRÈTES, VECTEURS ALÉATOIRES DISCRETS

1 Définitions et premières propriétés

Lorsque l'on s'intéresse aux phénomènes aléatoires, on peut être amené à étudier simultanément plusieurs variables aléatoires qui rendent compte d'une même expérience, et à déterminer si la valeur de l'une dépend de celles des autres. Pour ce faire, on commence par une :

Définition 1.1. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A}, P)$ est une application :

$$Z : \begin{cases} \Omega & \longrightarrow & \mathbb{R}^2 \\ \omega & \longmapsto & (X(\omega), Y(\omega)) \end{cases} ,$$

où X, Y sont 2 variables aléatoires discrètes sur $(\Omega, \mathcal{A}, P)$. Elle est notée $Z = (X, Y)$.

Les couples de variables aléatoires réelles discrètes apparaissent très souvent lors d'une expérience aléatoire. Par exemple, considérons une succession de deux lancers d'un même dé. Si X désigne le résultat du premier lancer et Y celui du deuxième, alors $Z = (X, Y)$ est un couple de variables aléatoires discrètes.

Définition 1.2. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. Le support de (X, Y) , noté $(X, Y)(\Omega)$, est l'image de Ω par (X, Y) .

A noter qu'on a l'inclusion $(X, Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$, mais qu'on n'a pas nécessairement d'égalité entre ces deux ensembles. Plus précisément, il existe des couples de variables aléatoires (X, Y) pour lesquels tous les couples de valeurs (x, y) avec $x \in X(\Omega)$ et $y \in Y(\Omega)$ ne peuvent être obtenus. A titre d'exemple, si $X = Y$, alors on voit que (X, Y) ne peut prendre le couple de valeurs (x, y) que si $x = y$! Pour des raisons de commodité, on représente souvent le support du couple (X, Y) dans le plan.

Théorème - définition 1.3. Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. Alors la famille $([X = x] \cap [Y = y])_{(x, y) \in X(\Omega) \times Y(\Omega)}$ est un système complet d'événements, qui est appelé le système complet d'événements associé au couple (X, Y) .

Dans tout ce qui suit, toutes les variables aléatoires que l'on rencontrera seront supposées définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

2 Loi conjointe et lois marginales

Une fois posée la notion de couple de variables aléatoires discrètes, on peut lui associer sa loi conjointe, c'est-à-dire une fonction qui va mesurer la possibilité ou la fréquence qu'a ce couple de variables aléatoires de prendre certains couples de valeurs. Plus précisément :

Définition 2.1. Soit (X, Y) un couple de variables aléatoires discrètes. La loi conjointe de (X, Y) est l'application :

$$F_{(X, Y)} : \begin{cases} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & P([X \leq x] \cap [Y \leq y]) \end{cases} .$$

Théorème 2.2. Soit (X, Y) un couple de variables aléatoires discrètes. Alors la loi de (X, Y) est complètement déterminée par la donnée des $P([X = x] \cap [Y = y])$ lorsque (x, y) parcourt $X(\Omega) \times Y(\Omega)$.

En d'autres termes, deux couples de variables aléatoires discrètes (X, Y) et (X', Y') pour lesquelles toutes les probabilités $P([X = x] \cap [Y = y])$ et $P([X' = x] \cap [Y' = y])$ sont égales suivent la même loi. En général, on utilisera le théorème 2.2 pour décrire la loi d'un couple de variables aléatoires discrètes (et non la définition 2.1). Plus précisément, si les éléments de $X(\Omega)$ (resp. $Y(\Omega)$) sont donnés sous la forme d'une liste $(x_i)_{i \in I}$ (resp. $(y_j)_{j \in J}$) indexée sur un ensemble I (resp. J) au plus dénombrable, la loi du couple (X, Y) sera donnée par la famille $(p_{i,j})_{(i,j) \in I \times J}$ définie par :

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

A noter que, si X et Y sont des variables aléatoires finies, on peut représenter la loi conjointe du couple (X, Y) sous la forme d'un tableau à double entrée. Les lignes de ce tableau correspondent alors aux éléments $x_1, \dots, x_n$ de $X(\Omega)$, et les colonnes aux éléments $y_1, \dots, y_m$ de $Y(\Omega)$. De plus, au croisement de la i -ème ligne et de la j -ème colonne, on indique la probabilité $p_{i,j}$:

	y_1	y_2	$\dots$	y_m
x_1	$p_{1,1}$	$p_{1,2}$	$\dots$	$p_{1,m}$
x_2	$p_{2,1}$	$p_{2,2}$	$\dots$	$p_{2,m}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	$p_{n,1}$	$p_{n,2}$	$\dots$	$p_{n,m}$

Définition 2.3. Soit (X, Y) un couple de variables aléatoires discrètes. Alors la loi de X est appelée la première loi marginale de (X, Y) , et la loi de Y la deuxième loi marginale de (X, Y) .

A noter que la loi conjointe du couple (X, Y) permet d'en retrouver les lois marginales. En effet :

Théorème 2.4. Soit (X, Y) un couple de variables aléatoires discrètes. On pose $X(\Omega) = \{x_i\}_{i \in I}$, $Y(\Omega) = \{y_j\}_{j \in J}$ et $p_{i,j} = P([X = x_i] \cap [Y = y_j])$ pour tout $(i, j) \in I \times J$. Alors, pour tout $i \in I$ et pour tout $j \in J$, les séries $\sum_j p_{i,j}$ et $\sum_i p_{i,j}$ convergent absolument, et de plus :

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

A noter que, comme les séries en question sont absolument convergentes, l'ordre de sommation ne change ni l'absolue convergence, ni la valeur des sommes en question. A noter aussi que, si les variables aléatoires X et Y sont finies, alors la probabilité de l'événement $[X = x_i]$ s'obtient en effectuant la somme de toutes les probabilités apparaissant dans la i -ème ligne du tableau de la loi de (X, Y) , et la probabilité de $[Y = y_j]$ en effectuant la somme de toutes les probabilités apparaissant dans la j -ème colonne du même tableau. Dans tous les cas, la loi du couple (X, Y) détermine complètement les lois marginales de (X, Y) . A noter enfin que, comme la famille $([X = x_i] \cap [Y = y_j])_{(i,j) \in I \times J}$ est un système complet d'événements, la série double $\sum_{i,j} p_{i,j}$ converge absolument, et de plus :

$$\boxed{\sum_{(i,j) \in I \times J} p_{i,j} = 1.}$$

En particulier, si X et Y sont des variables aléatoires finies, cela signifie que la somme des probabilités qui apparaissent dans le tableau de la loi de (X, Y) est égale à 1. Ce résultat est très utile pour compléter ce tableau lorsqu'il y manque une probabilité ou pour vérifier la cohérence des résultats obtenus. A noter que l'encadré ci-dessus admet une réciproque (que nous admettrons), à savoir le :

Théorème 2.5. Soient I, J des ensembles au plus dénombrables. Une famille $(p_{i,j})_{(i,j) \in I \times J}$ de réels représente la loi d'un couple de variables aléatoires discrètes si et seulement si :

1. Pour tout $(i, j) \in I \times J$, on a : $p_{i,j} \geq 0$.
2. La série double $\sum_{i,j} p_{i,j}$ converge absolument et de plus : $\sum_{(i,j) \in I \times J} p_{i,j} = 1$.

A noter que, si la série double $\sum_{i,j} p_{i,j}$ est à termes positifs, il suffit d'appliquer le théorème de Fubini (en sommant par rapport à i puis par rapport à j , ou l'inverse) pour établir que cette série converge absolument, et que sa somme est égale à 1.

3 Lois conditionnelles

Définition 3.1. Soit (X, Y) un couple de variables aléatoires discrètes, et soit y un élément de $Y(\Omega)$ tel que $P([Y = y]) \neq 0$. La loi conditionnelle de X sachant $[Y = y]$ est l'application :

$$P_{[Y=y]} : \begin{cases} X(\Omega) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & P_{[Y=y]}([X = x]) \end{cases}$$

De même, si $x \in X(\Omega)$ et si $P([X = x]) \neq 0$, la loi conditionnelle de Y sachant $[X = x]$ est l'application :

$$P_{[X=x]} : \begin{cases} Y(\Omega) & \longrightarrow \mathbb{R} \\ y & \longmapsto P_{[X=x]}([Y = y]) \end{cases}$$

A noter que la condition " $P([Y = y]) \neq 0$ " (resp. " $P([X = x]) \neq 0$ ") est essentielle pour pouvoir définir la loi conditionnelle de X sachant $[Y = y]$ (resp. la loi conditionnelle de Y sachant $[X = x]$). En particulier, le théorème 2.4 entraîne que, pour tout $x \in X(\Omega)$ et tout $y \in Y(\Omega)$:

$$\begin{aligned} P([X = x]) &= \sum_{y \in Y(\Omega), P([Y=y]) \neq 0} P_{[Y=y]}([X = x])P([Y = y]) \\ P([Y = y]) &= \sum_{x \in X(\Omega), P([X=x]) \neq 0} P_{[X=x]}([Y = y])P([X = x]) \end{aligned}$$

Bien entendu, il va de soi que les sommes ci-dessus sont bien définies, c'est-à-dire que les séries en question convergent toujours absolument. L'intérêt de ces formules tient dans le fait que, si l'on connaît la première loi marginale de (X, Y) et la loi de Y sachant $[X = x]$ pour tout $x \in X(\Omega)$ tel que $P([X = x]) \neq 0$, alors on peut retrouver la deuxième loi marginale de (X, Y) .

4 Indépendance de deux variables aléatoires discrètes

Définition 4.1. Deux variables aléatoires discrètes X et Y sont dites indépendantes si :

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

En d'autres termes, X et Y sont indépendantes si et seulement si, pour tout $(x, y) \in \mathbb{R}^2$, les événements $[X \leq x]$ et $[Y \leq y]$ sont indépendants. On montre alors (et nous admettrons) le :

Théorème 4.2. Soit (X, Y) un couple de variables aléatoires discrètes. Alors les conditions suivantes sont équivalentes :

1. Les variables aléatoires X et Y sont indépendantes,
2. Pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, on a : $P([X = x] \cap [Y = y]) = P([X = x])P([Y = y])$.
3. Pour tous intervalles I, J de $\mathbb{R}$, on a : $P([X \in I] \cap [Y \in J]) = P([X \in I])P([Y \in J])$.

A noter que, si X et Y sont indépendantes, alors la loi du couple (X, Y) est le produit des lois marginales de (X, Y) . A noter aussi que, si X et Y sont indépendantes et si $P([X = x]) \neq 0$ et $P([Y = y]) \neq 0$, alors $P([X = x] \cap [Y = y]) \neq 0$. En particulier, si X et Y sont des variables aléatoires finies indépendantes et si un zéro apparaît dans le tableau de la loi conjointe, alors il doit se trouver sur une ligne ou sur une colonne de zéros. Cette observation permet dans certains cas de vérifier simplement que deux variables aléatoires finies ne sont pas indépendantes à partir du tableau de leur loi conjointe.

Théorème 4.3. (Lemme des coalitions) Soit (X, Y) un couple de variables aléatoires discrètes. Si X et Y sont indépendantes, alors toute variable aléatoire fonction de X est indépendante de toute variable aléatoire fonction de Y .

5 Fonctions de deux variables aléatoires discrètes

Théorème - définition 5.1. Soit (X, Y) un couple de variables aléatoires discrètes. Alors, pour toute application $f : (X, Y)(\Omega) \longrightarrow \mathbb{R}$, l'application $Z : \Omega \longrightarrow \mathbb{R}$ définie pour tout $\omega \in \Omega$ par $Z(\omega) = f(X(\omega), Y(\omega))$ est une variable aléatoire discrète notée $f(X, Y)$.

En d'autres termes, toute fonction Z construite à partir de X, Y est une variable aléatoire discrète. C'est le cas notamment de $X + Y$, XY , $\max\{X, Y\}$, $\min\{X, Y\}$, résultat que l'on a déjà vu dans le chapitre "Variables aléatoires discrètes".

Théorème 5.2. Soient (X, Y) et (X', Y') deux couples de variables aléatoires discrètes, et soit g une fonction continue de $\mathbb{R}^2$ dans $\mathbb{R}$. Si (X, Y) et (X', Y') suivent la même loi, alors $g(X, Y)$ et $g(X', Y')$ suivent la même loi.

On verra en exercice comment déterminer les lois des variables aléatoires $\max\{X, Y\}$ et $\min\{X, Y\}$, où X et Y sont indépendantes. Dans le cas d'une somme de variables aléatoires, on dispose des résultats suivants :

Théorème 5.3. *Soit (X, Y) un couple de variables aléatoires discrètes. Alors la loi de $X + Y$ est définie pour tout $z \in (X + Y)(\Omega)$ par :*

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

Théorème 5.4. (Produit de convolution discret) *Soit (X, Y) un couple de variables aléatoires discrètes indépendantes. Alors la loi de $X + Y$ est définie pour tout $z \in (X + Y)(\Omega)$ par :*

$$P([X + Y = z]) = \sum_{x \in X(\Omega)} P([X = x])P([Y = z - x]).$$

A noter que les théorèmes 5.3 et 5.4 ne sont pas officiellement au programme, mais qu'il faut savoir les retrouver si besoin. Pour ce faire, on appliquera tout naturellement la formule des probabilités totales à l'événement $[X + Y = z]$ et au système complet d'événements $([X = x])_{x \in X(\Omega)}$. A noter enfin qu'à partir du produit de convolution discret, on obtient les deux résultats suivants :

Théorème 5.5. (Stabilité de la loi binomiale par addition) *Soient X_1, X_2 deux variables aléatoires discrètes indépendantes, telles que $X_1 \hookrightarrow \mathcal{B}(n_1, p)$ et $X_2 \hookrightarrow \mathcal{B}(n_2, p)$. Alors $X_1 + X_2 \hookrightarrow \mathcal{B}(n_1 + n_2, p)$.*

Théorème 5.6. (Stabilité de la loi de Poisson par addition) *Soient X_1, X_2 deux variables aléatoires discrètes indépendantes, telles que $X_1 \hookrightarrow \mathcal{P}(\lambda_1)$ et $X_2 \hookrightarrow \mathcal{P}(\lambda_2)$. Alors $X_1 + X_2 \hookrightarrow \mathcal{P}(\lambda_1 + \lambda_2)$.*

En particulier, ces deux derniers résultats sont très utiles pour donner directement la loi d'une somme de variables aléatoires dans certains cas, et ce sans avoir à calculer explicitement la loi en question à l'aide du produit de convolution discret. Ils peuvent aussi servir à donner facilement l'espérance et la variance d'une somme de variables aléatoires discrètes dans certains cas, et ce sans avoir à effectuer de longs calculs.

6 Propriétés de l'espérance

Dans ce paragraphe, on se propose de rappeler quelques propriétés de l'espérance. Commençons par donner une seconde version du théorème de transfert, qui sera particulièrement utile pour calculer l'espérance d'une fonction de deux variables aléatoires discrètes (et ce en passant par des séries doubles). Plus précisément, on montre (et nous admettrons) le :

Théorème 6.1. (Transfert) *Soit (X, Y) un couple de variables aléatoires discrètes et soit f une application de $(X, Y)(\Omega)$ dans $\mathbb{R}$. Alors la variable aléatoire $f(X, Y)$ admet une espérance si et seulement si la série double $\sum_{x,y} f(x, y)P([X = x] \cap [Y = y])$ converge absolument, et dans ce cas :*

$$E(f(X, Y)) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} f(x, y)P([X = x] \cap [Y = y]).$$

En d'autres termes, si $X(\Omega) = (x_i)_{i \in I}$ et si $Y(\Omega) = (y_j)_{j \in J}$, où I et J sont au plus dénombrables, alors on a sous réserve de convergence absolue que :

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

L'intérêt du théorème de transfert vient de ce qu'il permet de calculer l'espérance de $f(X, Y)$ directement à partir de la loi conjointe de (X, Y) , et donc sans avoir à déterminer la loi de $f(X, Y)$ au préalable. A noter que, si X et Y sont des variables aléatoires finies, alors $f(X, Y)$ est aussi finie, et donc elle admet une espérance. A noter enfin que, si l'on applique le théorème de transfert à l'application $f(X, Y) = \lambda X + \mu Y$, où $\lambda, \mu \in \mathbb{R}$, alors on obtient aisément le :

Théorème 6.2. (Linéarité de l'espérance) *Soit (X, Y) un couple de variables aléatoires discrètes, et soient $\lambda, \mu \in \mathbb{R}$. Si X et Y admettent toutes deux une espérance, alors $\lambda X + \mu Y$ admet aussi une espérance, et de plus : $E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y)$.*

Théorème 6.3. (Positivité de l'espérance) Soit X une variable aléatoire discrète. Si X admet une espérance et si $X \geq 0$ presque sûrement, alors : $E(X) \geq 0$.

Théorème 6.4. (Croissance de l'espérance) Soit (X, Y) un couple de variables aléatoires discrètes. Si X et Y admettent toutes deux une espérance et si $X \leq Y$ presque sûrement, alors : $E(X) \leq E(Y)$.

A noter que la propriété de croissance se déduit presque aussitôt des propriétés de linéarité et de positivité. En effet, si $X \leq Y$ presque sûrement et si X, Y admettent toutes deux une espérance, alors la variable aléatoire $Y - X$ est positive presque sûrement et admet une espérance. Dès lors, la positivité et la linéarité de l'espérance entraîne que $E(Y - X) = E(Y) - E(X) \geq 0$, et donc $E(X) \leq E(Y)$.

Théorème 6.5. (Existence de l'espérance par domination) Soient X, Y deux variables aléatoires discrètes, telles que $0 \leq |X| \leq Y$ presque sûrement. Si Y admet une espérance, alors X admet aussi une espérance et de plus : $|E(X)| \leq E(Y)$.

Théorème 6.6. (Espérance d'un produit de variables aléatoires indépendantes) Soit (X, Y) un couple de variables aléatoires discrètes indépendantes. Si X et Y admettent toutes deux une espérance, alors XY admet aussi une espérance et de plus : $E(XY) = E(X)E(Y)$.

A noter que les théorèmes 6.2 et 6.6 permettent bien souvent de simplifier les calculs d'espérance sans avoir à repasser par la définition de base ni par le théorème de transfert.

7 Covariance de deux variables aléatoires discrètes

Définition 7.1. Soit (X, Y) un couple de variables aléatoires discrètes. La covariance de (X, Y) est le réel (s'il existe) défini par :

$$\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))].$$

Théorème 7.2. (Formule de Koenig-Huygens) Soit (X, Y) un couple de variables aléatoires discrètes. Si X et Y admettent toutes deux un moment d'ordre 2, alors (X, Y) admet une covariance et de plus :

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y).$$

On dit aussi que les variables aléatoires discrètes admettent une variance (plutôt qu'un moment d'ordre 2). La formule de Koenig-Huygens permet notamment de ramener le calcul de la covariance de X, Y à celui des espérances de X, Y, XY . C'est cette formule que l'on utilise la plupart du temps dans les calculs. A noter que, d'après le théorème de transfert et sous réserve de convergence absolue, on a :

$$E(XY) = \sum_{(x,y) \in (X,Y)(\Omega)} xy P([X = x] \cap [Y = y]).$$

En d'autres termes, si $X(\Omega) = (x_i)_{i \in I}$ et si $Y(\Omega) = (y_j)_{j \in J}$, où I et J sont au plus dénombrables, alors on a sous réserve de convergence absolue que :

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

Théorème 7.3. Soient $\lambda, \lambda' \in \mathbb{R}$, et soient X, X', Y, Y' des variables aléatoires discrètes. Si X, X', Y, Y' admettent chacune une variance, alors $\lambda X + \lambda' X'$ et $\lambda Y + \lambda' Y'$ admettent une variance et de plus :

1. $\text{cov}(X, X) = V(X)$.
2. $\text{cov}(X, Y) = \text{cov}(Y, X)$ (symétrie).
3. $\text{cov}(\lambda X + \lambda' X', Y) = \lambda \text{cov}(X, Y) + \lambda' \text{cov}(X', Y)$ (linéarité à gauche).
4. $\text{cov}(X, \lambda Y + \lambda' Y') = \lambda \text{cov}(X, Y) + \lambda' \text{cov}(X, Y')$ (linéarité à droite).

Dans le langage mathématique, on résume souvent les propriétés (3) et (4) de ce théorème en disant que la covariance est bilinéaire, c'est-à-dire linéaire en les première et deuxième variables aléatoires d'entrée. Cette propriété de bilinéarité permet souvent d'alléger des calculs fastidieux.

Théorème 7.4. Soient X et Y des variables aléatoires discrètes. Si X et Y admettent chacune une variance, alors $X + Y$ admet aussi une variance, et de plus :

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

Théorème 7.5. Soient X et Y des variables aléatoires discrètes. Si X et Y admettent chacune une variance et sont indépendantes, alors : $\text{cov}(X, Y) = 0$.

Théorème 7.6. Soient X et Y des variables aléatoires discrètes. Si X et Y admettent chacune une variance et sont indépendantes, alors : $V(X + Y) = V(X) + V(Y)$.

En général, le théorème 7.5 s'utilise soit pour calculer une covariance, soit pour montrer que deux variables aléatoires ne sont pas indépendantes. En effet, considérons deux variables aléatoires discrètes X et Y admettant une variance. Si elles sont indépendantes, alors on peut conclure directement que $\text{cov}(X, Y) = 0$. Inversement, si l'on constate après calculs que $\text{cov}(X, Y) \neq 0$, alors on peut en déduire que X et Y ne sont pas indépendantes. A noter que le théorème 7.5 n'est pas une équivalence. Plus précisément, *il existe des couples de variables aléatoires discrètes de covariance nulle et non indépendantes*.

8 Coefficient de corrélation linéaire

Définition 8.1. Soient X, Y des variables aléatoires discrètes admettant des variances non nulles. Le coefficient de corrélation linéaire de X et Y est le réel :

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

Définition 8.2. Soient X et Y des variables aléatoires discrètes admettant des variances. On dit que X et Y sont non corrélées si $\text{cov}(X, Y) = 0$.

A noter que, si X et Y admettent des variances non nulles, alors X et Y sont non corrélées si et seulement si $\rho(X, Y) = 0$. En particulier, le théorème 7.5 signifie que, *si X et Y sont indépendantes et admettent des variances non nulles, alors $\rho(X, Y) = 0$* .

Théorème 8.3. Soient X et Y des variables aléatoires discrètes admettant des variances non nulles. Alors, pour tous réels a, b, c, d tels que $ac \neq 0$, on a :

$$\rho(aX + b, cY + d) = \text{signe}(ac)\rho(X, Y).$$

A noter que l'expression "signe(ac)" désigne le signe du réel ac , c'est-à-dire $+$ si $ac > 0$ et $-$ si $ac < 0$. Le théorème 8.3 traduit simplement le fait que le coefficient de corrélation linéaire est (au signe près) invariant par changement d'échelle, c'est-à-dire qu'il ne change pas (au signe près) si l'on rajoute des constantes à X et à Y , ni si on les multiplie par des constantes non nulles.

Théorème 8.4. Soient X et Y des variables aléatoires discrètes admettant des variances non nulles. Alors on a : $-1 \leq \rho(X, Y) \leq 1$. De plus, $\rho(X, Y) = \pm 1$ si et seulement s'il existe des réels a, b tels que $Y = aX + b$ presque sûrement, c'est-à-dire : $P([Y = aX + b]) = 1$.

A noter que, si $\rho(X, Y) = 1$ (resp. -1), alors $Y = aX + b$ presque sûrement et $a > 0$ (resp. $a < 0$). En particulier, le coefficient de corrélation linéaire est égal à ± 1 si et seulement s'il existe une relation de dépendance linéaire entre X et Y , c'est-à-dire si Y est une fonction affine de X (presque sûrement). Géométriquement, cela signifie que les couples de valeurs de (X, Y) sont situées de façon quasi-certaine sur une droite du plan d'équation $y = ax + b$. A contrario, le fait que $\rho(X, Y) = 0$ signifie qu'il n'existe pas de relation de dépendance linéaire entre X et Y . Dès lors, si X et Y sont indépendantes, alors il n'existe pas de relation de dépendance d'aucune sorte entre X et Y , ce qui explique que $\rho(X, Y) = 0$. Par contre, si $\rho(X, Y) = 0$, alors X et Y ne sont pas nécessairement indépendantes. En effet, il peut très bien exister une relation de dépendance entre X et Y , sans qu'elle soit pour autant linéaire. De façon générale, *le coefficient de corrélation linéaire mesure l'obstruction à l'existence d'une relation de dépendance linéaire entre les variables aléatoires discrètes X et Y* .

9 Généralisation aux vecteurs aléatoires discrets

9.1 Définitions de base

Définition 9.1. Un vecteur aléatoire discret sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ est une application :

$$Z : \begin{cases} \Omega & \longrightarrow & \mathbb{R}^n \\ \omega & \longmapsto & (X_1(\omega), \dots, X_n(\omega)) \end{cases},$$

où $X_1, \dots, X_n$ sont des variables aléatoires discrètes sur $(\Omega, \mathcal{A}, P)$. On le note $Z = (X_1, \dots, X_n)$.

Comme pour les couples de variables aléatoires discrètes, on montre (et nous admettrons) le :

Théorème - définition 9.2. Soit $(X_1, \dots, X_n)$ un vecteur aléatoire discret. Alors la famille d'événements $([X_1 = x_1] \cap \dots \cap [X_n = x_n])_{(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)}$ est un système complet d'événements, qui est appelé le système complet d'événements associé à $(X_1, \dots, X_n)$.

Définition 9.3. Soit $(X_1, \dots, X_n)$ un vecteur aléatoire discret. La loi conjointe de $(X_1, \dots, X_n)$ (ou plus simplement la loi de $(X_1, \dots, X_n)$) est l'application :

$$F_{(X_1, \dots, X_n)} : \begin{cases} \mathbb{R}^n & \longrightarrow \\ (x_1, \dots, x_n) & \longmapsto P([X_1 \leq x_1] \cap \dots \cap [X_n \leq x_n]) \end{cases}.$$

De plus, les lois marginales de $(X_1, \dots, X_n)$ sont les lois des variables aléatoires discrètes $X_1, \dots, X_n$.

Théorème 9.4. Soit $(X_1, \dots, X_n)$ un vecteur aléatoire discret. Alors la loi de $(X_1, \dots, X_n)$ est complètement déterminée par la donnée des $P([X_1 = x_1] \cap \dots \cap [X_n = x_n])$ lorsque $(x_1, \dots, x_n)$ parcourt $X_1(\Omega) \times \dots \times X_n(\Omega)$.

Comme pour les couples de variables aléatoires discrètes, on montre (et nous admettrons) que :

$$\sum_{(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)} P([X_1 = x_1] \cap \dots \cap [X_n = x_n]) = 1.$$

Plus précisément, la somme ci-dessus correspond à celle d'une série multiple (c'est-à-dire portant sur n indices). Tout comme pour les séries doubles à termes positifs, on montre (et nous admettrons) que cette série multiple converge absolument, et que sa somme ne dépend pas de l'ordre dans lequel la sommation est effectuée. Pour calculer une telle somme dans la pratique, on commence par additionner les termes de cette série par rapport aux termes de $X_n(\Omega)$, puis par rapport à ceux de $X_{n-1}(\Omega)$, et ainsi de suite. A noter que la loi d'un vecteur aléatoire discret permet de retrouver ses lois marginales. Plus précisément, comme pour les couples de variables aléatoires discrètes, on montre (et nous admettrons) les :

Théorème 9.5. Soit $(X_1, \dots, X_n)$ un vecteur aléatoire discret. Alors, pour tout $k \in \{1, \dots, n\}$ et tout $x_k \in X_k(\Omega)$, on a :

$$P([X_k = x_k]) = \sum_{(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \in X_1(\Omega) \times \dots \times X_{k-1}(\Omega) \times X_{k+1}(\Omega) \times \dots \times X_n(\Omega)} P([X_1 = x_1] \cap \dots \cap [X_n = x_n]).$$

9.2 Indépendance mutuelle de plusieurs variables aléatoires

Définition 9.6. Des variables aléatoires discrètes $X_1, \dots, X_n$ sont dites (mutuellement) indépendantes si, pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$, on a :

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

En d'autres termes, les variables aléatoires $X_1, \dots, X_n$ sont indépendantes si et seulement si, pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$, les événements $[X_1 \leq x_1], \dots, [X_n \leq x_n]$ sont indépendants.

Théorème 9.7. Soit $(X_1, \dots, X_n)$ un vecteur aléatoire discret. Les conditions suivantes sont équivalentes :

1. Les variables aléatoires $X_1, \dots, X_n$ sont indépendantes,
2. Pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$, on a : $P(\cap_{i=1}^n [X_i = x_i]) = \prod_{i=1}^n P([X_i = x_i])$.
3. Pour tous intervalles $I_1, \dots, I_n$ de $\mathbb{R}$, on a : $P(\cap_{i=1}^n [X_i \in I_i]) = \prod_{i=1}^n P([X_i \in I_i])$.

En d'autres termes, les variables aléatoires discrètes $X_1, \dots, X_n$ sont indépendantes si et seulement si les événements $[X_1 = x_1], \dots, [X_n = x_n]$ sont indépendants pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$. Grosso modo, cela signifie que $X_1, \dots, X_n$ ne s'influencent pas, c'est-à-dire que les valeurs prises par certaines d'entre elles ne dépendent pas des valeurs prises par les autres. En général, on utilisera surtout la condition (2) du théorème 9.7 pour vérifier que des variables aléatoires discrètes sont indépendantes.

Définition 9.8. On dit qu'une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires discrètes est une suite de variables aléatoires discrètes indépendantes si, pour toute partie finie I de $\mathbb{N}$, les variables aléatoires X_i (avec $i \in I$) sont (mutuellement) indépendantes.

9.3 Fonctions d'un vecteur aléatoire discret

Théorème 9.9. Soit $(X_1, \dots, X_n)$ un vecteur aléatoire discret sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. Alors, pour toute application $f : (X_1, \dots, X_n)(\Omega) \rightarrow \mathbb{R}$, l'application $Z : \Omega \rightarrow \mathbb{R}$ définie pour tout $\omega \in \Omega$ par $Z(\omega) = f(X_1(\omega), \dots, X_n(\omega))$ est une variable aléatoire discrète, que l'on note $f(X_1, \dots, X_n)$.

En d'autres termes, toute fonction construite à partir de $X_1, \dots, X_n$ est une variable aléatoire discrète. C'est le cas notamment de $X_1 + \dots + X_n$, $X_1 \times \dots \times X_n$, $\max\{X_1, \dots, X_n\}$, $\min\{X_1, \dots, X_n\}$, résultat que l'on a déjà vu dans le chapitre "Variables aléatoires discrètes".

Théorème 9.10. Soient $(X_1, \dots, X_n)$ et $(X'_1, \dots, X'_n)$ deux vecteurs aléatoires discrets, et soit g une fonction continue de $\mathbb{R}^n$ dans $\mathbb{R}$. Si $(X_1, \dots, X_n)$ et $(X'_1, \dots, X'_n)$ suivent la même loi, alors $g(X_1, \dots, X_n)$ et $g(X'_1, \dots, X'_n)$ suivent la même loi.

Comme pour les couples de variables aléatoires, on verra en exercice comment déterminer les lois des variables aléatoires $\max\{X_1, \dots, X_n\}$ et $\min\{X_1, \dots, X_n\}$, où $X_1, \dots, X_n$ sont indépendantes.

Théorème 9.11. (Espérance d'un produit de variables aléatoires indépendantes) Soient $X_1, \dots, X_n$ des variables aléatoires discrètes indépendantes. Si $X_1, \dots, X_n$ admettent chacune une espérance, alors le produit $X_1 \dots X_n$ admet aussi une espérance, et de plus :

$$E(X_1 \dots X_n) = E(X_1) \dots E(X_n).$$

De même, le lemme des coalitions se généralise sous la forme suivante :

Théorème 9.12. (Lemme des coalitions) Soient $X_1, \dots, X_n$ des variables aléatoires discrètes. Si les variables aléatoires $X_1, \dots, X_n$ sont indépendantes, alors toute variable aléatoire fonction de p d'entre elles est indépendante de toute variable aléatoire fonction des $(n - p)$ restantes.

9.4 Sommes de variables aléatoires discrètes

Théorème 9.13. (Linéarité de l'espérance) Soient $X_1, \dots, X_n$ des variables aléatoires discrètes, et soient $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Si $X_1, \dots, X_n$ admettent toutes une espérance, alors $\lambda_1 X_1 + \dots + \lambda_n X_n$ admet aussi une espérance, et de plus :

$$E(\lambda_1 X_1 + \dots + \lambda_n X_n) = \lambda_1 E(X_1) + \dots + \lambda_n E(X_n).$$

Théorème 9.14. Soient $X_1, \dots, X_n$ des variables aléatoires discrètes admettant toutes une variance. Alors $X_1 + \dots + X_n$ admet aussi une variance, et de plus :

$$V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j).$$

En particulier, si $X_1, \dots, X_n$ sont indépendantes, alors $\text{cov}(X_i, X_j) = 0$ pour tous i, j avec $i \neq j$, et donc :

Théorème 9.15. Soient $X_1, \dots, X_n$ des variables aléatoires discrètes indépendantes admettant toutes une variance. Alors $X_1 + \dots + X_n$ admet aussi une variance, et de plus :

$$V(X_1 + \dots + X_n) = V(X_1) + \dots + V(X_n).$$

A présent, considérons une expérience aléatoire à deux issues A et B , telles que $P(A) = p$ et $P(B) = 1 - p$. Supposons que l'on itère cette expérience n fois dans les mêmes conditions. Pour tout $i \in \{1, \dots, n\}$, on désigne par X_i la variable aléatoire égale à 1 si A est réalisé lors de la i -ème expérience, et à 0 sinon. Alors chaque variable aléatoire X_i suit la loi de Bernoulli de paramètre p , et il est facile de voir que toutes ces variables aléatoires sont indépendantes. De plus, la variable aléatoire $X = X_1 + \dots + X_n$ correspond au nombre de fois où A est réalisé, c'est-à-dire au nombre de succès de l'issue A , et donc X suit la loi binomiale de paramètres n, p . Par conséquent, on a le :

Théorème 9.16. Soit $n \in \mathbb{N}^*$ et soit $p \in]0, 1[$. Alors toute variable binomiale de paramètres n, p peut s'écrire comme une somme de variables de Bernoulli indépendantes de paramètre p .

Théorème 9.17. Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes qui suivent toutes la même loi de Bernoulli de paramètre p . Alors $X_1 + \dots + X_n \hookrightarrow \mathcal{B}(n, p)$.

Plus généralement, on peut montrer par récurrence les :

Théorème 9.18. (Stabilité de la loi binomiale par addition) Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes. Si $X_k \hookrightarrow \mathcal{B}(m_k, p)$ pour tout $k \in \{1, \dots, n\}$, alors $X_1 + \dots + X_n \hookrightarrow \mathcal{B}(m_1 + \dots + m_n, p)$.

Théorème 9.19. (Stabilité de la loi de Poisson par addition) Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes. Si $X_k \hookrightarrow \mathcal{P}(\lambda_k)$ pour tout $k \in \{1, \dots, n\}$, alors $X_1 + \dots + X_n \hookrightarrow \mathcal{P}(\lambda_1 + \dots + \lambda_n)$.

En particulier, ces trois derniers résultats permettent de donner sans calcul la loi d'une somme de variables aléatoires discrètes dans certains cas, ainsi que l'espérance et la variance d'une telle somme, et ce sans avoir à effectuer de longs calculs.