33 Couples de variables aléatoires discrètes

ECG Maths Appliquées
Semestre 3

Dans tout ce chapitre, toutes les variables aléatoires envisagées sont des variables aléatoires discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

A. Loi d'un couple de variables aléatoires discrètes

A.1. Définition

Définition 33.1

On appelle **couple de variables aléatoires** réelles discrètes sur $(\Omega, \mathcal{A})$ toute application (X, Y) de Ω dans $\mathbb{R}^2$ telle que X et Y Soit des variables aléatoires réelles discrètes sur $(\Omega, \mathcal{A})$.

Exemple 33.1 On dispose d'une urne contenant n boules numérotées de 1 à n et on effectue dans cette urne deux tirages d'une boule, successifs et sans remise. En notant $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé permettant de modéliser cette expérience, X la variable aléatoire égale au premier numéro tiré et Y la variable aléatoire égale au deuxième numéro tiré, (X, Y) est un couple de variables aléatoires réelles discrètes sur $(\Omega, \mathcal{A})$.

Remarques

a. Comme il n'y a en général pas de confusion possible quant à l'espace probabilisable envisagé, on parle en général plus simplement de couple de variable aléatoires, sans mentionner l'espace probabilisé.

b. On a toujours :

$$(X, Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$$

En revanche, il n'y a pas toujours égalité. Ainsi, dans l'exemple précédent, on ne peut pas tirer deux fois la même boule, donc :

$$(X, Y)(\Omega) = \{(i, j) \in [\![1, n]\!]^2 \ / \ i \neq j\} \quad \text{et} \quad X(\Omega) \times Y(\Omega) = [\![1, n]\!]^2$$

c. Pour tout $(x, y) \in \mathbb{R}^2$, l'événement $[X = x] \cap [Y = y]$ est parfois aussi noté $[X = x, Y = y]$.

Définition 33.2

Soit (X, Y) un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$. Déterminer la loi de probabilité (ou **loi conjointe**) de (X, Y), c'est déterminer les probabilités $\mathbb{P}(X = x, Y = y)$ pour tout (x, y) appartenant à $X(\Omega) \times Y(\Omega)$.

Exercice 33.1 On dispose d'une urne contenant n boules numérotées de 1 à n et on effectue dans cette urne deux tirages d'une boule, successifs et sans remise.
Déterminer la loi du couple (X, Y).

Remarque Comme la probabilité $\mathbb{P}(X = x, Y = y)$ est nulle si (x, y) n'appartient pas à $(X, Y)(\Omega)$, il est aussi possible de caractériser la loi en ne donnant que les probabilités $\mathbb{P}(X = x, Y = y)$ lorsque $(x, y) \in (X, Y)(\Omega)$.

Définition 33.3

Soit (X, Y) un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$. La loi de X (respectivement de Y) est aussi appelée **loi marginale** de X (resp. de Y).

Proposition 33.4

Soit (X, Y) un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$. On a :

$$\forall x \in X(\Omega),\ \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(X = x, Y = y)$$

et :

$$\forall y \in Y(\Omega),\ \mathbb{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x, Y = y)$$

Preuve Comme X et Y sont deux variables aléatoires discrètes, $([X = x])_{x \in X(\Omega)}$ et $([Y = y])_{y \in Y(\Omega)}$ sont deux systèmes complets d'événements donc la proposition découle directement de la formule des probabilités totales.

□

A.2. Système complet d'événements associé à un couple de variables aléatoires discrètes

Proposition 33.5

Soit (X, Y) un couple de variables aléatoires discrètes. La famille $([X = x] \cap [Y = y])_{(x,y) \in (X,Y)(\Omega)}$ est un système complet d'événements, appelé **système complet d'événements associé** à (X, Y).

Preuve Comme (X, Y) est une application de Ω dans $\mathbb{R}^2$, la famille $([X = x] \cap [Y = y])_{(x,y) \in (X,Y)(\Omega)}$ est formée d'événements deux à deux incompatibles, et, par définition, on a :

$$\begin{aligned} \bigcup_{(x,y) \in (X,Y)(\Omega)} ([X = x] \cap [Y = y]) &= \bigcup_{(x,y) \in (X,Y)(\Omega)} \{\omega \in \Omega \ / \ (X, Y)(\omega) = (x, y)\} \\ &= \Omega \end{aligned}$$

ce qui prouve le résultat annoncé. □

Proposition 33.6

Soit (X, Y) un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A})$. On a :

$$\sum_{(x,y) \in X(\Omega) \times y \in Y(\Omega)} \mathbb{P}(X = x, Y = y) = 1$$

Preuve Ce résultat est une conséquence immédiate de 33.5. □

B. Fonction d'un couple de variables aléatoires discrètes

Dans ce paragraphe, (X, Y) est un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A})$.

Proposition 33.7

Si g est une application de $(X, Y)(\Omega)$ dans $\mathbb{R}$, alors l'application $Z : \omega \mapsto g(X(\omega), Y(\omega))$ est une variable aléatoire discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$, notée $g(X, Y)$.

Exemple 33.2 Si (X, Y) est un couple de variable aléatoires discrètes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$, alors $\min(X, Y)$, $\max(X, Y)$, $X+Y$, $X-Y$ et XY sont des variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$.

Proposition 33.8

Si g est une application de $(X, Y)(\Omega)$ dans $\mathbb{R}$ et $Z = g(X, Y)$, alors :

$$Z(\Omega) = \{g(x, y),\, (x, y) \in (X, Y)(\Omega)\}$$

et :

$$\forall z \in Z(\Omega),\ \mathbb{P}(Z = z) = \sum_{\substack{(x,y)\in(X,Y)(\Omega)\\ g(x,y)=z}} \mathbb{P}(X = x, Y = y)$$

Exercice 33.2 Dans une urne contenant n boules numérotées de 1 à n, on effectue deux tirages successifs et avec remise et on note, dans cet ordre, X et Y les variables aléatoires égales au premier et au deuxième numéros obtenus.
Déterminer la loi de $Z = X + Y$.

Théorème 33.9 ▶ Théorème de transfert

Soit g une application de $(X, Y)(\Omega)$ dans $\mathbb{R}$. Soit $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ deux suites réelles telles que (X, Y) prenne ses valeurs dans $\{(x_i, y_j),\, (i, j) \in \mathbb{N}^2\}$.
$g(X, Y)$ admet une espérance si et seulement si la somme double $\sum_{(i,j)} |g(x_i, y_j)\, \mathbb{P}(X = x_i, Y = y_j)|$ existe et, dans ce cas :

$$\mathbb{E}(g(X, Y)) = \sum_{i\in\mathbb{N}, j\in\mathbb{N}} g(x_i, y_j)\, \mathbb{P}(X = x_i, Y = y_j) = \sum_{(x,y)\in(X,Y)(\Omega)} g(x, y)\, \mathbb{P}(X = x, Y = y)$$

Exercice 33.3 On considère le couple (X, Y) défini dans l'exercice 26.2. Déterminer l'espérance de XY.

Remarques

a. Si X et Y sont des variables aléatoires discrètes finies, cette espérance existe toujours.

b. Ce résultat est essentiellement utilisé pour prouver la linéarité de l'espérance et pour étudier l'espérance de XY. À part dans ce dernier cas, il est très rare qu'elle soit utilisée aux concours.

c. En pratique, grâce au théorème de Fubini (voir chapitre 19. Séries), l'ordre de sommation n'importe pas pour le calcul si la série double est absolument convergente et on a donc, en cas d'existence de l'espérance :

$$\mathbb{E}(g(X, Y)) = \sum_{i=0}^{+\infty}\sum_{j=0}^{+\infty} g(x_i, y_j)\, \mathbb{P}(X = x_i, Y = y_j) = \sum_{j=0}^{+\infty}\sum_{i=0}^{+\infty} g(x_i, y_j)\, \mathbb{P}(X = x_i, Y = y_j)$$

Théorème 33.10

Si X et Y admettent une espérance, alors :

i. pour tout couple (a, b) de réels, $aX + bY$ admet une espérance et :

$$\mathbb{E}(aX + bY) = a\,\mathbb{E}(X) + b\,\mathbb{E}(Y)$$

ii. si $X \leqslant Y$ presque sûrement alors : $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$.

Remarque On en déduit par récurrence que :

Théorème 33.11 ▶ Linéarité de l'espérance

Si n est un entier naturel supérieur ou égal à 2, si $X_1, \ldots, X_n$ sont n variables aléatoires discrètes admettant une espérance, alors, pour tout n-uplet $(a_1, \ldots, a_n)$ de réels, $\sum\limits_{k=1}^{n} a_k X_k$ admet une espérance et :

$$\mathbb{E}\left(\sum_{k=1}^{n} a_k X_k\right) = \sum_{k=1}^{n} a_k \, \mathbb{E}(X_k)$$

C. Covariance et coefficient de corrélation linéaire

C.1. Covariance de deux variables aléatoires discrètes

Dans tout ce paragraphe, (X, Y) est un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$.

Théorème 33.12

Si X et Y admettent un moment d'ordre 2, alors XY admet une espérance.

Définition 33.13

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

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

Remarques

a. D'après 33.12, si X et Y admettent un moment d'ordre 2, XY, X et Y admettent une espérance, donc la covariance de (X, Y) est bien définie.

b. En particulier, si X admet un moment d'ordre 2, on a : $\mathrm{Cov}(X, X) = \mathbb{V}(X)$.

c. Compte tenu de la définition, on peut dire que, si $\mathrm{Cov}(X, Y) > 0$, alors, en moyenne, quand X augmente, Y augmente également, tandis que, si $\mathrm{Cov}(X, Y) < 0$, alors, en moyenne, quand X augmente, Y diminue. En revanche, s'il est possible d'interpréter le signe de la covariance, il n'est pas possible d'interpréter sa valeur. En effet, on verra plus loin que si (X, Y) admet une covariance, alors $\mathrm{Cov}(aX, Y) = a\,\mathrm{Cov}(X, Y)$, donc la valeur de la covariance est fortement influencée par les valeurs prises par chacune des deux variables aléatoires X et Y individuellement.

d. De même que pour la variance, il est rare que la définition soit utilisée pour le calcul d'une covariance. En pratique, on utilisera le plus souvent la formule suivante :

Théorème 33.14 ▶ Formule de Huygens

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

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

Preuve On a :

$$(X - \mathbb{E}(X))(Y - \mathbb{E}(Y)) = XY - \mathbb{E}(Y)X - \mathbb{E}(X)Y + \mathbb{E}(X)\,\mathbb{E}(Y)$$

Comme XY, X et Y admettent une espérance, on a donc, par linéarité de l'espérance :

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

□

Exercice 33.4 On considère un couple de variables aléatoires discrètes (X,Y) à valeurs dans $\{1,2,3\}\times\{0,1,2\}$ dont la loi est caractérisée par le tableau suivant (où figurent, pour chaque couple (i,j), la probabilité $\mathbb{P}([X=i]\cap[Y=j])$) :

$i \backslash j$	0	1	2
1	$\frac{1}{18}$	$\frac{1}{9}$	$\frac{1}{9}$
2	$\frac{2}{9}$	0	$\frac{1}{9}$
3	0	$\frac{1}{6}$	$\frac{2}{9}$

Déterminer la covariance du couple (X,Y).

Exercice 33.5 On considère sur un même espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$ une variable aléatoire X prenant les valeurs $-1,0$ et 1, chacune avec probabilité $\dfrac{1}{3}$ et on considère la variable aléatoire Y prenant la valeur 1 si $X=0$ et 0 sinon.
Calculer la covariance de (X,Y).

Proposition 33.15

La covariance est une forme bilinéaire symétrique et positive sur l'espace vectoriel des variables aléatoires discrètes définies sur $(\Omega,\mathcal{A},\mathbb{P})$ et admettant un moment d'ordre 2.
Autrement dit, si X,Y et Z sont trois variables aléatoires discrètes sur $(\Omega,\mathcal{A},\mathbb{P})$ admettant un moment d'ordre 2 et si a et b sont deux réels, alors on a :

i. $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$,
ii. $\mathrm{Cov}(aX+bY,Z)=a\,\mathrm{Cov}(X,Z)+b\,\mathrm{Cov}(Y,Z)$,
iii. $\mathrm{Cov}(X,aY+bZ)=a\,\mathrm{Cov}(X,Y)+b\,\mathrm{Cov}(X,Z)$,
iv. $\mathrm{Cov}(X,X)\geqslant 0$.

Preuve On suppose que X,Y et Z admettent chacune un moment d'ordre 2.

i. On a :

$$\begin{aligned}\mathrm{Cov}(X,Y)&=\mathbb{E}(XY)-\mathbb{E}(X)\,\mathbb{E}(Y)\\&=\mathbb{E}(YX)-\mathbb{E}(Y)\,\mathbb{E}(X)\\&=\mathrm{Cov}(Y,X)\end{aligned}$$

ii. D'après la formule de Huygens et la linéarité de l'espérance, on a :

$$\begin{aligned}\mathrm{Cov}(aX+bY,Z)&=\mathbb{E}((aX+bY)Z)-\mathbb{E}(aX+bY)\,\mathbb{E}(Z)\\&=\mathbb{E}(aXZ+bYZ)-a\,\mathbb{E}(X)\,\mathbb{E}(Z)-b\,\mathbb{E}(Y)\,\mathbb{E}(Z)\\&=a\,\mathbb{E}(XZ)+b\,\mathbb{E}(YZ)-a\,\mathbb{E}(X)\,\mathbb{E}(Z)-b\,\mathbb{E}(Y)\,\mathbb{E}(Z)\\&=a\,[\mathbb{E}(XZ)-\mathbb{E}(X)\,\mathbb{E}(Z)]+b\,[\mathbb{E}(YZ)-\mathbb{E}(Y)\,\mathbb{E}(Z)]\\&=a\,\mathrm{Cov}(X,Z)+b\,\mathrm{Cov}(Y,Z)\end{aligned}$$

iii. Ce résultat est une conséquence immédiate des deux premiers.

iv. On a :

$$\begin{aligned}\mathrm{Cov}(X,X)&=\mathbb{E}(X^2)-[\mathbb{E}(X)]^2\\&=\mathbb{V}(X)\\&\geqslant 0\end{aligned}$$

□

Proposition 33.16

Soit (X, Y) un couple de variables aléatoires admettant un moment d'ordre 2. $X+Y$ admet une variance et :

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

Preuve On a :

$$(X+Y)^2 = X^2 + 2XY + Y^2$$

Comme X et Y admettent un moment d'ordre 2, XY admet une espérance d'après 33.21, donc $(X+Y)^2$ admet une espérance. On en déduit que $X+Y$ admet une variance et, d'après la formule de Huygens et par linéarité de l'espérance :

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

$\square$

Remarques

a. Avant de calculer la variance de la somme de deux variables aléatoires discrètes, ils est fondamental de se demander si les deux variables aléatoires sont indépendantes. À défaut, il faudra calculer $\mathrm{Cov}(X, Y)$.

b. On en déduit une autre formule permettant parfois de calculer la covariance :

Proposition 33.17

Si X et Y admettent une variance, alors $\mathrm{Cov}(X, Y)$ existe et :

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

Exercice 33.6 Soit p, q, r des réels positifs tels que $p+q+r=1$. On dispose d'une urne contenant des boules blanches, noires et vertes en proportion respectives p, q et r.

Soit $n \in \mathbb{N}^*$. On effectue une suite de n tirages au hasard et avec remise de boules dans cette urne et on note X (respectivement Y, Z) la variable aléatoire égale au nombre de boules blanches (respectivement noires, vertes) obtenues.

1. Déterminer la loi de X, de Y et de Z.
2. En déduire la variance de X, Y et $X+Y$.
3. Calculer la covariance du couple (X, Y).

C.2. Coefficient de corrélation linéaire

Dans tout ce paragraphe, (X, Y) est un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$.

Définition 33.18

Si X et Y admettent toutes deux une variance non nulle, on appelle **coefficient de corrélation linéaire** de (X, Y) le réel

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

Remarque On a vu précédemment que, si le signe de la covariance est interprétable, sa valeur l'est beaucoup moins car l'ordre de grandeur des valeurs prises par X et Y influe fortement sur la valeur de la covariance. C'est pour cette raison que l'on introduit la notion de coefficient de corrélation linéaire : en divisant par les écart-types de X et de Y, on normalise ainsi la covariance, ce qui permet ensuite de comparer des valeurs de manière plus significative. En effet, on peut remarquer que :

$$\rho(X,Y) = \mathrm{Cov}\left(\frac{X}{\sigma(X)}, \frac{Y}{\sigma(Y)}\right)$$

Comme les variables aléatoires $\dfrac{X}{\sigma(X)}$ et $\dfrac{Y}{\sigma(Y)}$ sont réduites, la valeur de leur covariance est moins impacté par les variations individuelles de X et de Y et il devient plus intéressant de comparer les valeurs des coefficients de corrélation linéaire de deux couples (X,Y) et (X',Y').

Proposition 33.19

Si X et Y admettent chacune une variance non nulle, alors :

i. $|\rho(X,Y)| \leqslant 1$ (inégalité de Cauchy-Schwarz),

ii. $|\rho(X,Y)| = 1 \iff \exists\,(a,b) \in \mathbb{R}\ /\ \mathbb{P}(Y = aX + b) = 1$.

Preuve

i. On considère la fonction P positive sur $\mathbb{R}$ définie par :

$$\forall \lambda \in \mathbb{R},\ P(\lambda) = \mathbb{V}(\lambda X + Y)$$

D'après 33.16 et 33.15, on a :

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

Ainsi, P est un polynôme du second degré positif sur $\mathbb{R}$, donc son discriminant $\Delta_{(X,Y)}$ est négatif ou nul, c'est-à-dire :

$$4\,[\mathrm{Cov}(X,Y)]^2 - 4\,\mathbb{V}(X)\,\mathbb{V}(Y) \leqslant 0$$

d'où :

$$[\mathrm{Cov}(X,Y)]^2 \leqslant \mathbb{V}(X)\,\mathbb{V}(Y)$$

ce qui donne le résultat attendu par croissance de la fonction $t \mapsto \sqrt{t}$ sur $\mathbb{R}^+$.

ii. On a :

$$\begin{aligned} |\rho(X,Y)| = 1 &\iff [\mathrm{Cov}(X,Y)]^2 = \mathbb{V}(X)\,\mathbb{V}(Y) \\ &\iff \Delta_{(X,Y)} = 0 \end{aligned}$$

Comme on a déjà remarqué que ce discriminant est négatif ou nul, P admet au plus une racine réelle, donc :

$$\begin{aligned} |\rho(X,Y)| = 1 &\iff \exists\, a \in \mathbb{R}\ /\ P(a) = 0 \\ &\iff \exists\, \lambda \in \mathbb{R}\ /\ \mathbb{V}(\lambda X + Y) = 0 \\ &\iff \exists\, (\lambda,\mu) \in \mathbb{R}^2\ /\ \lambda X + Y = \mu\ (p.s.) \end{aligned}$$

et donc, en notant $a = -\lambda$ et $b = \mu$:

$$\begin{aligned} |\rho(X,Y)| = 1 &\iff \exists\, (\lambda,\mu) \in \mathbb{R}^2\ /\ Y = aX + b\ (p.s.) \\ &\iff \exists\, (\lambda,\mu) \in \mathbb{R}^2\ /\ \mathbb{P}(Y = aX + b) = 1 \end{aligned}$$

□

Remarques

a. Bien retenir la preuve, aussi rencontrée en algèbre bilinéaire pour démontrer l'inégalité de Cauchy-Schwarz. Cette démonstration est parfois demandée à l'écrit comme à l'oral.

b. S'il existe un triplet (a,b,c) de réels tel que $(a,b) \neq (0,0)$ et $\mathbb{P}(aX + bY + c = 0) = 1$, on dit qu'il existe une relation presque sûrement affine entre les variables aléatoires X et Y. Cette relation sera dite affine lorsque $aX + bY + c = 0$.

Exercice 33.7 Calculer le coefficient de corrélation linéaire du couple (X,Y) étudié dans l'exercice 26.6.

D. Indépendance de variables aléatoires discrètes

D.1. Couple de variables aléatoires indépendantes

Dans ce paragraphe, (X, Y) est un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A})$.

Définition 33.20

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

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

Remarques

a. Il est assez rare que l'on soit amené à démontrer l'indépendance de deux variables aléatoires par le calcul. En général, l'indépendance sera une conséquence immédiate du modèle proposé et, dans ce cas, on peut appliquer directement le théorème précédent dans les calculs.

b. Notons qu'il est assez rare que le modèle étudié permette de justifier que deux variables aléatoires ne sont pas indépendantes : il arrive que l'on puisse penser que deux variables aléatoires ne sont pas indépendantes alors qu'elles le sont. Pour prouver que deux variables aléatoires *ne sont pas* indépendantes, il faudra en général prouver qu'il existe un couple (x, y) de réels tel que : $\mathbb{P}(X = x, Y = y) \neq \mathbb{P}(X = x)\, \mathbb{P}(Y = y)$.

Exercice 33.8 On effectue trois lancers successifs et indépendants d'une pièce équilibrée et on note A l'événement « on obtient au plus une fois pile » et B l'événement « on obtient au moins une fois pile et au moins une fois face ».
Montrer que les variables aléatoires $X = \mathbb{1}_A$ et $Y = \mathbb{1}_B$ sont indépendantes.

Exercice 33.9 On considère sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ une variable aléatoire X prenant les valeurs $-1, 0$ et 1, chacune avec probabilité $\dfrac{1}{3}$ et on considère la variable aléatoire Y prenant la valeur 1 si $X = 0$ et 0 sinon.
Les variables aléatoires X et Y sont-elles indépendantes ?

Théorème 33.21

Si X et Y sont indépendantes et admettent une espérance, alors XY admet une espérance et :

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

Théorème 33.22

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

$$X \text{ et } Y \text{ sont indépendantes} \Longrightarrow \operatorname{Cov}(X, Y) = 0$$

Proposition 33.23

Si X et Y admettent chacune une variance non nulle, alors :

$$X \text{ et } Y \text{ sont indépendantes} \Longrightarrow \rho(X, Y) = 0$$

Remarques

a. Attention, la réciproque est fausse en général, comme le prouvent les exercices 26.6 et 26.10.

b. Si (X, Y) est un couple de variables aléatoires discrètes admettant un moment d'ordre 2 et telles que $\rho(X, Y) = 0$ ou $\operatorname{Cov}(X, Y) = 0$, on dit que X et Y sont **non corrélées** (ce qui signifie bien que, s'il existe, le coefficient de corrélation linéaire est nul).

D.2. Familles de variables aléatoires discrètes indépendantes

Dans tout ce paragraphe, n est un entier naturel supérieur ou égal à 2 et $X_1, \dots, X_n$ sont des variables aléatoires discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Définition 33.24

On dit que $X_1, \dots, X_n$ sont mutuellement indépendantes (ou plus simplement indépendantes) si :

$$\forall (x_i)_{1 \leqslant i \leqslant n} \in \mathbb{R}^n, \ \mathbb{P}\left(\bigcap_{i=1}^{n} [X_i = x_i]\right) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i)$$

Proposition 33.25 ▶ Lemme des coalitions

Si $X_1, \dots, X_n$ sont mutuellement indépendantes et si p est un entier tel que $2 \leqslant p \leqslant n-1$, alors toute variable aléatoire fonction de $X_1, \dots, X_p$ est indépendante de toute variable aléatoire fonction de $X_{p+1}, \dots, X_n$.

E. Somme de variables aléatoires discrètes indépendantes

E.1. Formule de convolution

Théorème 33.26 ▶ Théorème de convolution

Si (X, Y) est un couple de variables aléatoires discrètes indépendantes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, alors $X + Y$ est une variable aléatoire discrète et sa loi est caractérisée par :

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

Preuve Ce résultat est une conséquence immédiate de 33.8. □

E.2. Variance de la somme de variables aléatoires discrètes indépendantes

Théorème 33.27

i. Si (X, Y) un couple de variables aléatoires indépendantes admettant une variance, $X + Y$ admet une variance et :

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

ii. Plus généralement, si n est un entier naturel supérieur ou égal à 2, si $X_1, \dots, X_n$ sont n variables aléatoires discrètes mutuellement indépendantes admettant une variance, alors, pour tout n-uplet $(a_1, \dots, a_n$ de réels :

$$\mathbb{V}\left(\sum_{k=1}^{n} a_k X_k\right) = \sum_{k=1}^{n} a_k^2 \, \mathbb{V}(X_k)$$

E.3. Stabilité pour la somme de la loi binomiale

Théorème 33.28

Soit n et m deux entiers naturels non nuls et p un réel appartenant à $]0, 1[$.
Si X et Y sont deux variables aléatoires indépendantes suivant respectivement les lois binomiales $\mathscr{B}(n, p)$ et $\mathscr{B}(m, p)$ alors $X + Y$ suit la loi binomiale $\mathscr{B}(n + m, p)$.

Preuve Comme X et Y prennent leurs valeurs respectivement dans $[\![0,n]\!]$ et $[\![0,m]\!]$, $X+Y$ prend ses valeurs dans $[\![0,n+m]\!]$ et on a, d'après 33.26 :

$$\begin{aligned}\forall k \in [\![0,n+m]\!]\,,\ \mathbb{P}(X+Y=k) &= \sum_{i=0}^{k} \mathbb{P}(X=i)\,\mathbb{P}(Y=k-i)\\ &= \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{k-i} p^{k-i} (1-p)^{m-k+i}\\ &= p^k (1-p)^{n+m-k} \sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i}\end{aligned}$$

et donc, d'après la formule de Vandermonde (*cf.* exercice 24.14 du chapitre 24. Dénombrement) :

$$\forall k \in [\![0,n+m]\!]\,,\ \mathbb{P}(X+Y=k) = \binom{n+m}{k} p^k (1-p)^{n+m-k}$$

donc $X+Y$ suit la loi binomiale $\mathscr{B}(n+m,p)$.

□

Remarque On en déduit par récurrence le résultat suivant :

Théorème 33.29

Soit $n_1,\dots,n_i$ des entiers naturels non nuls, p un réel appartenant à $]0,1[$ et $X_1,\dots,X_i$ des variables aléatoires indépendantes suivant respectivement les lois binomiales $\mathscr{B}(n_1,p),\dots,\mathscr{B}(n_i,p)$.
$X_1+\cdots+X_i$ est une variable aléatoire sur $(\Omega,\mathcal{A},\mathbb{P})$ suivant la loi binomiale de paramètres $n_1+\cdots+n_i$ et p.

E.4. Somme de variables aléatoires de Bernoulli indépendantes

Théorème 33.30

Soit n un entier naturel supérieur ou égal à 2 et p un réel appartenant à $]0,1[$.
Si $X_1,\dots,X_n$ sont n variables aléatoires indépendantes, $X_1+\cdots+X_n$ suit la loi binomiale $\mathscr{B}(n,p)$.

Preuve Ce résultat est un cas particulier de la proposition 33.29. □

E.5. Stabilité pour la somme de la loi de Poisson

Théorème 33.31

Soit λ et μ deux réels strictement positifs.
Si X et Y sont deux variables aléatoires indépendantes suivant respectivement les lois de Poisson $\mathscr{P}(\lambda)$ et $\mathscr{P}(\mu)$ alors $X+Y$ suit la loi de Poisson $\mathscr{P}(\lambda+\mu)$.

Exercice 33.10 Démontrer le théorème 33.31.

Remarque On en déduit par récurrence le résultat suivant :

Théorème 33.32

Soit n un entier naturel supérieur ou égal à 2, $\lambda_1,\dots,\lambda_n$ n réels strictement positifs et $X_1,\dots,X_n$ n variables aléatoires définies sur le même espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$ et suivant respectivement les lois de Poisson $\mathscr{P}(\lambda_1),\dots,\mathscr{P}(\lambda_n)$.
$X_1+\cdots+X_n$ est une variable aléatoire sur $(\Omega,\mathcal{A},\mathbb{P})$ suivant la loi de Poisson de paramètre $\lambda_1+\cdots+\lambda_n$.

F. Correction des exercices

Correction de l'exercice 33-1

Comme on effectue deux tirages dans une urne contenant n boules numérotées de 1 à n, X et Y prennent leurs valeurs dans $[\![1,n]\!]$.
Soit alors $(i,j) \in [\![1,n]\!]^2$. Comme on effectue deux tirages au hasard et sans remise, il y a autant de couples de tirages possibles que de 2-listes d'éléments distincts de $[\![1,n]\!]$, donc $n(n-1)$ couples de tirages possibles, tous équiprobables, dont un seul réalise l'événement $[X=i]\cap[Y=j]$ si $i \neq j$ et aucun si $i=j$, donc la loi du couple (X,Y) est caractérisée par :

$$\forall (i,j) \in [\![1,n]\!]^2\,,\ \mathbb{P}([X=i]\cap[Y=j]) = \begin{cases} \dfrac{1}{n(n-1)} & \text{si } i \neq j \\ 0 & \text{si } i = j \end{cases}$$

Correction de l'exercice 33-2

Comme on effectue deux tirages dans une urne contenant n boules numérotées de 1 à n, X et Y prennent leurs valeurs dans $[\![1,n]\!]$ et (X,Y) est une couple de variables aléatoires discrètes.
Soit alors $(i,j) \in [\![1,n]\!]$. Comme on effectue deux tirages au hasard et avec remise, il y a autant de couples de tirages possibles que de 2-listes d'éléments de $[\![1,n]\!]$, donc n^2 couples de tirages possibles, tous équiprobables, dont un seul réalise l'événement $[X=i]\cap[Y=j]$ donc la loi du couple (X,Y) est caractérisée par :

$$\forall (i,j) \in [\![1,n]\!]^2\,,\ \mathbb{P}([X=i]\cap[Y=j]) = \frac{1}{n^2}$$

Comme l'application $\varphi : (x,y) \mapsto x+y$ est bien définie sur $(X,Y)(\Omega)$, Z est donc une variable aléatoire discrète qui prend ses valeurs dans $[\![2,2n]\!]$. De plus, d'après 33.8, on a :

$$\forall k \in [\![2,2n]\!]\,,\ \mathbb{P}(Z=k) = \sum_{\substack{(i,j)\in(X,Y)(\Omega)\\ i+j=k}} \mathbb{P}([X=i]\cap[Y=j])$$

De plus, on a, si $k \in [\![2,2n]\!]$:

$$\begin{aligned} \begin{cases} (i,j) \in (X,Y)(\Omega) \\ i+j=k \end{cases} &\iff \begin{cases} 1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n \\ j = k-i \end{cases} \\ &\iff \begin{cases} 1 \leqslant i \leqslant n \\ 1 \leqslant k-i \leqslant n \\ j = k-i \end{cases} \\ &\iff \begin{cases} 1 \leqslant i \leqslant n \\ k-n \leqslant i \leqslant k-1 \\ j = k-i \end{cases} \end{aligned}$$

et donc :

$$\forall k \in [\![2,2n]\!]\,,\ \mathbb{P}(Z=k) = \sum_{i=\max(1,k-n)}^{\min(n,k-1)} \mathbb{P}([X=i]\cap[Y=k-i])$$

Soit alors $k \in [\![2,2n]\!]$. Deux cas se présentent :

◇ si $k \leqslant n$, alors $1 \geqslant k-n$ et $k-1 \leqslant n$, donc :

$$\mathbb{P}(Z=k) = \sum_{i=1}^{k-1} \frac{1}{n^2} = \frac{k-1}{n^2}$$

◇ si $k > n$, alors $k-n \geqslant 1$ et $n \leqslant k-1$, donc :

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

Finalement, la loi de Z est caractérisée par :

$$\forall k \in [\![2, 2n]\!]\,,\ \mathbb{P}(Z = k) = \begin{cases} \dfrac{k-1}{n^2} & \text{si } k \leqslant n \\ \dfrac{2n-k+1}{n^2} & \text{si } k > n \end{cases}$$

Correction de l'exercice 33-3

X et Y étant deux variables aléatoires discrètes finies, $(X, Y)(\Omega)$ est fini, donc XY admet une espérance et on a, d'après le théorème de transfert :

$$\begin{aligned} \mathbb{E}(XY) &= \sum_{(i,j)\in(X,Y)(\Omega)} ij\ \mathbb{P}([X = i] \cap [Y = j]) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{ij}{n^2} \\ &= \frac{1}{n^2} \left(\sum_{i=1}^{n} i\right)\left(\sum_{j=1}^{n} j\right) \\ &= \frac{(n+1)^2}{4} \end{aligned}$$

Correction de l'exercice 33-4

On peut déjà remarquer que X et Y sont deux variables aléatoires discrètes finies, donc X, Y et XY admettent une espérance. On commence par déterminer les lois respectives de X et Y, que l'on obtient en sommant respectivement en ligne et colonne d'après A.1 :

$i \backslash j$	0	1	2	$\mathbb{P}(X = i)$
1	$\frac{1}{18}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{5}{18}$
2	$\frac{2}{9}$	0	$\frac{1}{9}$	$\frac{1}{3}$
3	0	$\frac{1}{6}$	$\frac{2}{9}$	$\frac{7}{18}$
$\mathbb{P}(Y = j)$	$\frac{5}{18}$	$\frac{5}{18}$	$\frac{4}{9}$	

On a donc :

$$\mathbb{E}(X) = 1 \times \frac{5}{18} + 2 \times \frac{1}{3} + 3 \times \frac{7}{18} = \frac{19}{9}$$

et :

$$\mathbb{E}(Y) = 0 \times \frac{5}{18} + 1 \times \frac{5}{18} + 2 \times \frac{4}{9} = \frac{7}{6}$$

De plus, on a, d'après le théorème de transfert :

$$\begin{aligned} \mathbb{E}(XY) &= \sum_{i=1}^{3} \sum_{i=0}^{2} ij\ \mathbb{P}([X = i] \cap [Y = j]) \\ &= 1 \times 1 \times \frac{1}{9} + 1 \times 2 \times \frac{1}{9} + 2 \times 2 \times \frac{1}{9} + 3 \times 1 \times \frac{1}{6} + 3 \times 2 \times \frac{2}{9} \\ &= \frac{47}{18} \end{aligned}$$

On a donc :

$$\begin{aligned} \operatorname{Cov}(X, Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y) \\ &= \frac{47}{18} - \frac{19}{9} \times \frac{7}{6} \\ &= \frac{47 \times 3 - 19 \times 7}{9 \times 6} \\ &= \frac{4}{27} \end{aligned}$$

Correction de l'exercice 33-5

X et Y étant deux variables aléatoires discrètes finies, elles admettent un moment d'ordre 2 et $\text{Cov}(X, Y)$ existe. De plus, on a :

$$\mathbb{E}(X) = -\mathbb{P}(X = -1) + \mathbb{P}(X = 1) = 0$$

donc :

$$\mathbb{E}(X)\,\mathbb{E}(Y) = 0$$

Par ailleurs, XY est constante nulle donc on a directement :

$$\mathbb{E}(XY) = 0$$

et donc :

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y) = 0$$

Correction de l'exercice 33-6

1. ▶ X est la variable aléatoire égale au nombre de succès (tirer une boule blanche) dans une suite de n épreuves indépendantes et toutes de même probabilité de succès p (car les tirages sont faits au hasard et avec remise), donc X suit la loi binomiale $\mathcal{B}(n, p)$.
   ▶ De même Y suit la loi binomiale $\mathcal{B}(n, q)$ et Z suit la loi binomiale $\mathcal{B}(n, r)$.
2. X, Y et Z sont des variables discrètes finies donc elles admettent une variance.
   ▶ On a donc, dans un premier temps :
   $$\mathbb{V}(X) = np\,(1-p) \quad \text{et} \quad \mathbb{V}(Y) = nq\,(1-q)$$
   ▶ On peut de plus remarquer que $X + Y + Z = n$ donc on a :
   $$\mathbb{V}(X+Y) = \mathbb{V}(n-Z) = \mathbb{V}(Z) = nr\,(1-r)$$
3. (X, Y) est un couple de variables aléatoires discrètes finies, donc $\text{Cov}(X, Y)$ existe et on a :
   $$\begin{aligned}\text{Cov}(X, Y) &= \frac{1}{2}\left[\mathbb{V}(X+Y) - \mathbb{V}(X) - \mathbb{V}(Y)\right] \\ &= \frac{1}{2}\left[nr\,(1-r) - np\,(1-p) - nq\,(1-q)\right] \\ &= \frac{n\left[r\,(1-r) - p\,(1-p) - q\,(1-q)\right]}{2}\end{aligned}$$

Correction de l'exercice 33-7

On peut déjà remarquer que X et Y sont des variables aléatoires discrètes finies donc elles admettent un moment d'ordre 2 et : On a donc :

$$\mathbb{E}(X^2) = 1^2 \times \frac{5}{18} + 2^2 \times \frac{1}{3} + 3^2 \times \frac{7}{18} = \frac{46}{9}$$

et :

$$\mathbb{E}(Y^2) = 0^2 \times \frac{5}{18} + 1^2 \times \frac{5}{18} + 2^2 \times \frac{4}{9} = \frac{37}{18}$$

On a donc :

$$\mathbb{V}(X) = \frac{53}{81} \quad \text{et} \quad \mathbb{V}(Y) = \frac{25}{36}$$

On a donc :

$$\begin{aligned}\rho(X, Y) &= \frac{\text{Cov}(X, Y)}{\sqrt{\mathbb{V}(X)}\,\sqrt{\mathbb{V}(Y)}} \\ &= \frac{4}{27} \times \sqrt{\frac{81}{53}} \times \sqrt{\frac{36}{25}} \\ &= \frac{24}{15\sqrt{53}}\end{aligned}$$

Correction de l'exercice 33-8

On peut déjà remarquer que X et Y sont deux variables aléatoires discrètes prenant leurs valeurs dans $\{0,1\}$ donc sont indépendantes si et seulement si $[X=1]$ et $[Y=1]$ sont indépendantes (On rappelle que, si deux événements E et F sont indépendants, alors E et $\overline{F}$ sont indépendants, $\overline{E}$ et F sont indépendants, ainsi que $\overline{E}$ et $\overline{F}$) donc si et seulement si A et B sont indépendants.
De plus, en notant N la variable aléatoire égale au nombre de « pile » obtenus, on peut remarquer, les lancers étant indépendants et la pièce équilibrée, que N suit la loi binomiale de paramètres 3 et $\dfrac{1}{2}$ et que :

$$A = [N=0] \cup [N=1]$$

et donc, les événements $[N=0]$ et $[N=1]$ étant incompatibles :

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(N=0) + \mathbb{P}(N=1)\\ &= \frac{1}{8} + 3 \times \frac{1}{8}\\ &= \frac{1}{2}\end{aligned}$$

De même, on a :

$$\begin{aligned}\mathbb{P}(B) &= \mathbb{P}(N=1) + \mathbb{P}(N=2)\\ &= \frac{3}{4}\end{aligned}$$

On a donc :

$$\mathbb{P}(X=1)\,\mathbb{P}(Y=1) = \mathbb{P}(A)\,\mathbb{P}(B) = \frac{3}{8}$$

De plus, on a :

$$[X=1] \cap [Y=1] = A \cap B = [N=1]$$

et donc :

$$\mathbb{P}([X=1] \cap [Y=1]) = \frac{3}{8} = \mathbb{P}(X=1)\,\mathbb{P}(Y=1) = \mathbb{P}(A)\,\mathbb{P}(B)$$

Finalement, les variables aléatoires X et Y sont bien indépendantes.

Correction de l'exercice 33-9

On remarque que :

$$\mathbb{P}(X=1)\,\mathbb{P}(Y=1) = \frac{1}{9} \quad \text{et} \quad \mathbb{P}([X=1] \cap [Y=1]) = \mathbb{P}([X=1] \cap [X=0]) = 0$$

donc :

$$\mathbb{P}([X=1] \cap [Y=1]) \neq \mathbb{P}(X=1)\,\mathbb{P}(Y=1)$$

donc X et Y ne sont pas indépendantes.

Correction de l'exercice 33-10

Comme X et Y prennent leurs valeurs dans $\mathbb{N}$, $X+Y$ prend ses valeurs dans $\mathbb{N}$ et on a, d'après 33.26 :

$$\begin{aligned}\forall n \in \mathbb{N},\ \mathbb{P}(X+Y=n) &= \sum_{k=0}^{+\infty} \mathbb{P}(X=k)\,\mathbb{P}(Y=n-k)\\ &= \sum_{k=0}^{n} \mathrm{e}^{-\lambda}\,\frac{\lambda^k}{k!} \times \mathrm{e}^{-\mu}\,\frac{\mu^{n-k}}{(n-k)!}\\ &= \frac{\mathrm{e}^{-(\lambda+\mu)}}{n!} \sum_{k=0}^{n} \binom{n}{k} \lambda^k \mu^{n-k}\end{aligned}$$

et d'après la formule du binôme de Newton :

$$\forall n \in \mathbb{N},\ \mathbb{P}(X+Y=n) = \mathrm{e}^{-(\lambda+\mu)}\,\frac{(\lambda+\mu)^n}{n!}$$

donc $X+Y$ suit la loi de Poisson de paramètre $\lambda+\mu$.

Sommaire