2 | Couples de variables aléatoires

Dans tout ce chapitre Ω désigne un univers **fini**. Ainsi les variables aléatoires réelles considérées ne prennent qu'un nombre **fini** de valeurs.

I – Lois de probabilités

1 – Loi d'un couple de variables aléatoires

Définition 2.1 – On appelle **couple de variables aléatoires**, tout couple (X, Y) où X et Y désignent deux variables aléatoires définies sur un même ensemble Ω *(l'univers)*.

Exemple 2.2 – Tout au long de ce chapitre, on s'appuie sur les deux exemples suivants pour illustrer les différentes notions rencontrées :

1. On lance deux dés équilibrés à six faces, l'un rouge, l'autre noir. On appelle X (resp. Y) le numéro obtenu avec le dé rouge (resp. noir).
   Comme X et Y sont des variables aléatoires, alors (X, Y) est un couple de variables aléatoires.
2. On lance les mêmes dés que dans l'exemple précédent. Cette fois, on appelle X le plus petit des deux numéros obtenus et Y le plus grand numéro obtenu (si les numéros sont égaux, X et Y prennent la valeur commune).
   Comme X et Y sont des variables aléatoires, alors (X, Y) est un couple de variables aléatoires.

Définition 2.3 – Soit (X, Y) un couple de variables aléatoires. On appelle **loi conjointe** du couple (X, Y) la donnée de toutes les probabilités $P\big([X = x] \cap [Y = y]\big)$ pour tout couple $(x, y) \in X(\Omega) \times Y(\Omega)$.

Méthode 2.4 – Déterminer la loi conjointe d'un couple de variables aléatoires

1. On détermine les supports $X(\Omega)$ et $Y(\Omega)$, ensembles des valeurs prises par X et Y.
2. On calcule les probabilités $P\big([X = x] \cap [Y = y]\big)$ pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$.

Comme pour une variable aléatoire unique, on résume souvent la loi sous la forme d'un tableau, cette fois à double entrée. La somme de toutes les probabilités est toujours égale à 1.

Exemple 2.5 – Donner la loi conjointe des couples (X, Y) dans les deux exemples précédents.

1.

2.

J'en déduis les deux tableaux suivants pour les deux lois conjointes :

1.

	$Y=1$	$Y=2$	$Y=3$	$Y=4$	$Y=5$	$Y=6$
$X=1$						
$X=2$						
$X=3$						
$X=4$						
$X=5$						
$X=6$						

2.

	$Y=1$	$Y=2$	$Y=3$	$Y=4$	$Y=5$	$Y=6$
$X=1$						
$X=2$						
$X=3$						
$X=4$						
$X=5$						
$X=6$						

Remarque 2.6 –

- On abrège souvent "loi conjointe du couple" en "loi du couple".
- On note parfois $P\big([X=x],[Y=y]\big)$ au lieu de $P\big([X=x]\cap[Y=y]\big)$, même simplement $P(X=x, Y=y)$.

2 – Lois marginales

Définition 2.7 – Soit (X,Y) un couple de variables aléatoires. La loi de X est appelée **première loi marginale** du couple et celle de Y est appelée **deuxième loi marginale** du couple.

Proposition 2.8

Soient X et Y deux variables aléatoires définies sur Ω.

- Pour tout réel $x \in X(\Omega)$,
$$P(X=x) = \sum_{y\in Y(\Omega)} P(X=x, Y=y).$$
- Pour tout réel $y \in Y(\Omega)$,
$$P(Y=y) = \sum_{x\in X(\Omega)} P(X=x, Y=y).$$

Méthode 2.9 – Déterminer les lois marginales avec la loi du couple
Une fois que l'on a déterminé la loi du couple, on peut déterminer les lois marginales.
La loi de X s'écrit par exemple

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

Lorsque la loi d'un couple (X, Y) est donnée sous la forme d'un tableau à double entrée, on obtient les lois de X et de Y en sommant les éléments d'une même ligne ou d'une même colonne, selon les cas.

Exemple 2.10 – Déterminer les lois marginales de X et Y dans les deux exemples précédents.

1.

2.

Remarque 2.11 – Si l'établissement des lois marginales découle directement de la donnée de la loi conjointe, il est en revanche impossible, en général, d'obtenir la loi conjointe à partir des deux lois marginales.

3 – Lois conditionnelles

Définition 2.12 – Soit (X, Y) un couple de variables aléatoires. Pour tout $y \in Y(\Omega)$ tel que $P(Y = y) \neq 0$, on appelle loi de X **conditionnellement à l'événement** $[Y = y]$ la donnée, pour tout $x \in X(\Omega)$, de

$$P_{[Y=y]}(X = x) = \frac{P(X = x, Y = y)}{P(Y = y)}.$$

Remarque 2.13 –

- On dit aussi "loi conditionnelle de X sachant que $[Y = y]$ est réalisé", ou plus simplement "loi de X sachant $[Y = y]$".
- On définit de manière similaire la loi conditionnelle de Y sachant $[X = x]$.

Exemple 2.14 – Dans les deux exemples précédents, $P(Y=1) \neq 0$.
Déterminer alors la loi conditionnelle de X sachant $[Y=1]$ dans les deux cas.

1.

2.

Proposition 2.15 – Loi marginale et loi conditionnelle

Soit (X, Y) un couple de variables aléatoires. Si l'on connaît la loi marginale de Y, ainsi que toutes les lois conditionnelles de X sachant $[Y=y]$ pour tous les $y \in Y(\Omega)$, alors la loi de X est déterminée par

$$\forall x \in X(\Omega), \quad P(X=x) = \sum_{y \in Y(\Omega)} P(Y=y) \times P_{[Y=y]}(X=x).$$

Exemple 2.16 – J'ai calculé la loi conditionnelle de X sachant $[Y=1]$. Si je calculais les lois conditionnelles de X sachant $[Y=2]$, $[Y=3]$, etc., dans les deux exemples précédents, alors je pourrais retrouver la loi marginale de X grâce à la proposition ci-dessus.

4 – Indépendance de deux variables aléatoires

Définition 2.17 – On dit que deux variables aléatoires finies X et Y sont **indépendantes** lorsque

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad P(X=x, Y=y) = P(X=x) \times P(Y=y).$$

Remarque 2.18 – Dans le cas de deux variables aléatoires indépendantes, on peut déterminer la loi du couple (X, Y) à partir des lois de X et de Y.

Exemple 2.19 – Tester l'indépendance des variables aléatoires X et Y dans les exemples précédents.

1.

2.

Proposition 2.20

Si l'une des deux variables aléatoires X ou Y est constante, alors X et Y sont indépendantes.

II – Espérance

1 – Espérance d'une somme

Proposition 2.21

Soient X et Y deux variables aléatoires définies sur Ω. Alors

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

Exemple 2.22 – Soient X une variable aléatoire qui suit la loi uniforme sur $[\![1,9]\!]$ et Y une variable aléatoire qui suit la loi binomiale $\mathcal{B}\left(8, \frac{1}{4}\right)$. Calculer l'espérance de la variable aléatoire $Z = X + Y$.

Proposition 2.23 – Linéarité de l'espérance

Soient X et Y deux variables aléatoires définies sur Ω et a et b deux réels. Alors

$$E(aX+bY) = aE(X) + bE(Y).$$

Exemple 2.24 – Soient X une variable aléatoire qui suit la loi uniforme sur $[\![1,12]\!]$ et Y une variable aléatoire qui suit la loi binomiale $\mathcal{B}\left(7,\frac{1}{3}\right)$. Calculer l'espérance de la variable aléatoire $Z = 2X - Y$.

2 – Espérance d'un produit

Proposition 2.25

Soient X et Y deux variables aléatoires définies sur Ω. Alors l'espérance du produit est définie grâce à la loi conjointe par

$$E(XY) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X=x, Y=y).$$

Exemple 2.26 –

1. Un sac contient quatre boules numérotées de 1 à 4. On effectue deux tirages successifs d'une boule, avec remise. On note X_1 le numéro de la première boule, X_2 le numéro de la deuxième boule et Y le plus grand des deux numéros obtenus.
   Compléter les tableaux suivants, donnant les lois des couples (X_1, X_2) et (X_1, Y).

	$X_2=1$	$X_2=2$	$X_2=3$	$X_2=4$
$X_1=1$				
$X_1=2$				
$X_1=3$				
$X_1=4$				

	$Y=1$	$Y=2$	$Y=3$	$Y=4$
$X_1=1$				
$X_1=2$				
$X_1=3$				
$X_1=4$				

   En déduire $E(X_1X_2)$ et $E(X_1Y)$.

2. Soit (X,Y) un couple de variables aléatoires finies tel que

$$\forall (i,j) \in [\![1,2]\!]^2, \quad P(X=i, Y=j) = \begin{cases} \dfrac{i}{4} & \text{si } i=j, \\ \dfrac{1}{4} & \text{si } i<j, \\ 0 & \text{si } i>j. \end{cases}$$

Calculer $E(XY)$.

Proposition 2.27

Soient X et Y deux variables aléatoires **indépendantes** définies sur Ω. Alors

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

Exemple 2.28 – Même exemple que précédemment : un sac contient quatre boules numérotées. On effectue deux tirages successifs d'une boule, avec remise. On note X_1 le numéro de la première boule, X_2 le numéro de la deuxième boule et Y le plus grand des deux numéros obtenus.

1. Déterminer les lois marginales de X_1, X_2 et Y.

2. En déduire les valeurs de $E(X_1)$, $E(X_2)$ et $E(Y)$.

3. Les variables aléatoires X_1 et X_2 sont-elles indépendantes ? Et les variables aléatoires X_1 et Y ?

ATTENTION ! L'égalité $E(XY) = E(X)E(Y)$ peut être vérifiée sans que les variables aléatoires X et Y ne soient indépendantes.

III – Covariance, corrélation linéaire

1 – Covariance de deux variables aléatoires

Définition 2.29 – Soit (X, Y) un couple de variables aléatoires. On appelle **covariance de** X **et** Y, le réel, noté $\text{Cov}(X, Y)$, défini par

$$\text{Cov}(XY) = E\Big(\big(X - E(X)\big)\big(Y - E(Y)\big)\Big).$$

Théorème 2.30 – Formule de König-Huygens

Soit (X, Y) un couple de variables aléatoires. Alors

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

Démonstration.

□

Méthode 2.31 – Calculer directement une covariance

Pour calculer la covariance de deux variables aléatoires X et Y :

1. On calcule les trois espérances $E(X)$, $E(Y)$ et $E(XY)$ si ce n'est pas déjà fait.
2. On applique la formule de König-Huygens : $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$.

Exemple 2.32 – On reprend les deux exemples précédents.

1. Calculer $\text{Cov}(X_1, X_2)$ et $\text{Cov}(X_1, Y)$.

2. Calculer $\text{Cov}(X, Y)$.

Proposition 2.33 – Propriétés de la covariance

Soient X et Y deux variables aléatoires définies sur Ω.

- La covariance est symétrique :
$$\text{Cov}(X, Y) = \text{Cov}(Y, X).$$
- La covariance d'une variable aléatoire avec elle-même est égale à sa variance :
$$\text{Cov}(X, X) = V(X).$$
- Cas d'une variable aléatoire constante : si a est un réel, alors
$$\text{Cov}(X, a) = 0.$$

Proposition 2.34 – Linéarité à gauche et à droite de la covariance

Soient X, X_1, X_2, Y, Y_1 et Y_2 des variables aléatoires définies sur Ω. Pour tout couple $(a, b) \in \mathbb{R}^2$,

$$\mathrm{Cov}(aX_1 + bX_2, Y) = a\,\mathrm{Cov}(X_1, Y) + b\,\mathrm{Cov}(X_2, Y),$$

$$\mathrm{Cov}(X, aY_1 + bY_2) = a\,\mathrm{Cov}(X, Y_1) + b\,\mathrm{Cov}(X, Y_2).$$

Proposition 2.35

Soient X et Y deux variables aléatoires **indépendantes** définies sur Ω. Alors

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

Remarque 2.36 –

- C'est une conséquence directe de la Proposition 2.27.
- La réciproque est fausse : il se peut que $\mathrm{Cov}(X, Y) = 0$ sans que les variables aléatoires X et Y ne soient indépendantes.

Méthode 2.37 – Montrer que deux variables aléatoires ne sont pas indépendantes

Ceci est un récapitulatif des résultats à disposition pour montrer que deux variables aléatoires **ne sont pas** indépendantes. Pour rappel, si elles sont indépendantes, il n'y a aucun autre moyen que de montrer par le calcul que chaque probabilité de la loi conjointe s'obtient comme le produit des probabilités des lois marginales.

Soient X et Y deux variables aléatoires définies sur Ω.

- Si la covariance $\mathrm{Cov}(X, Y)$ est non nulle, alors les variables aléatoires X et Y ne sont pas indépendantes.
- Si les espérances ne satisfont pas l'égalité $E(XY) = E(X)E(Y)$, alors les variables aléatoires X et Y ne sont pas indépendantes.
- S'il existe un couple $(x, y) \in X(\Omega) \times Y(\Omega)$ pour qui l'égalité $P(X = x, Y = y) = P(X = x) \times P(Y = y)$ n'est pas vérifiée, alors les variables aléatoires X et Y ne sont pas indépendantes.
  Ce dernier point est souvent le plus facile à utiliser lorsqu'un couple présente une probabilité nulle dans le tableau de la loi conjointe.

2 – Variance d'une somme

Proposition 2.38

Soient X et Y deux variables aléatoires définies sur Ω. Alors

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

Méthode 2.39 – Calculer la variance d'une somme

Pour calculer la variance d'une somme de variables aléatoires, il y a deux possibilités :

- Si on connaît la loi de la somme $X + Y$, on utilise la **formule de König-Huygens** :
$$V(X + Y) = E\left((X + Y)^2\right) - E(X + Y)^2.$$
- Si on ne connaît pas la loi de la somme $X + Y$, on utilise plutôt la formule précédente :
$$V(X + Y) = V(X) + V(Y) + 2\,\mathrm{Cov}(X, Y).$$

Exemple 2.40 – On reprend les deux exemples précédents.

1. Calculer $V(X_1 + X_2)$ et $V(X_1 + Y)$.

2. Calculer $V(X + Y)$.

Remarque 2.41 – Cette formule permet également de calculer la covariance de X et Y à l'aide des trois variances $V(X)$, $V(Y)$ et $V(X+Y)$ puisque

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

Proposition 2.42

Soit (X,Y) un couple de variables aléatoires **indépendantes**. Alors

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

3 – Coefficient de corrélation linéaire

Définition 2.43 – On appelle **coefficient de corrélation linéaire** de X et Y, le réel, noté $\rho(X,Y)$, défini par

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}}.$$

Exemple 2.44 – On reprend les deux exemples de l'exemple 2.26.

1. Calculer $\rho(X_1, X_2)$ et $\rho(X_1, Y)$.

2. Calculer $\rho(X,Y)$.

Proposition 2.45

Soient X et Y deux variables aléatoires définies sur Ω. Alors

$$\left|\rho(X,Y)\right| \leqslant 1.$$

Remarque 2.46 – Le coefficient de corrélation linéaire mesure la dépendance linéaire entre deux variables aléatoires :

- Si le coefficient de corrélation linéaire est égal à 1 ou -1, X et Y sont corrélées linéairement.
  Le signe indique si les variations vont dans le même sens ou dans le sens opposé.
- Si le coefficient de corrélation linéaire est égal à 0, X et Y sont dites "non corrélées linéairement".
  Cela ne dit rien en revanche quant à une autre forme de corrélation.