Couples de variables aléatoires discrètes - Covariance et indépendance

1. Loi de probabilité d'un couple de variables aléatoires discrètes

Soit (X, Y) un couple variables aléatoires discrète sur l'espace $(\Omega, \mathbb{P}(\Omega), \mathbb{P})$ tel que :

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

avec I et J deux sous-ensembles de $\mathbb{N}$

Définition.

On appelle **loi de probabilité du couple** (X, Y) (ou loi jointe ou conjointe de (X, Y)) l'ensemble des données des probabilités :

$$\mathbb{P}((X = x_i) \cap (Y = y_j)) = \mathbb{P}(X = x_i \ , \ Y = y_j)$$

lorsque $i \in I$ et $j \in J$.

☛ Lorsque ces variables sont finies, la loi du couple est souvent présentée sous forme d'un tableau à double entrée, les valeurs prises par X apparaissant par exemple en ligne et celles prises par Y en colonne.

Exemple :
On choisit deux nombres au hasard dans $\{-1, 1\}$. On note X leur somme, et Y leur produit. On cherche à déterminer la loi conjointe de (X, Y). Remarquons que l'univers est $\{-1; 1\}^2$, que X prend ses valeurs dans $\{-2, 0, 2\}$ et Y dans $\{-1, 1\}$.
La loi conjointe de (X, Y) est alors :

$\mathbb{P}(X = 2, Y = 1) = \dfrac{1}{4}$ (correspond au choix 1 et 1) $\qquad \mathbb{P}(X = 2, Y = -1) = 0 \qquad \mathbb{P}(X = 0, Y = 1) = 0$

$$\mathbb{P}(X = 0, Y = -1) = \frac{1}{2} \qquad \mathbb{P}(X = -2, Y = 1) = \frac{1}{4} \qquad \mathbb{P}(X = -2, Y = -1) = 0$$

Propriété.

Soit (X, Y) un couple de variables discrètes tel que $X(\Omega) = \{x_i, i \in I\}$ et $Y(\Omega) = \{y_j, j \in J\}$. L'ensemble des événements $(X = x_i)$ pour $i \in I$ et $(Y = y_j)$ pour $j \in J$ forment un système complet d'événements. La somme de leur probabilités vaut donc 1 :

$$\sum_{i \in I} \sum_{j \in J} \mathbb{P}((X = x_i) \cap (Y = y_j)) \ = \ 1$$

2. Lois marginales d'un couple de variables (X, Y)

Définition (Lois marginales d'un couple de variables discrètes)**.**

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et (X, Y) un couple de variables discrètes sur cet espace.
Les lois de probabilité de X et Y sont alors appelés lois marginales de (X, Y).

`Comment déterminer les lois marginales à partir de la loi conjointes de` (X, Y) `?`

Il suffit d'utiliser la formule des probabilités totales!!
En effet, si l'on cherche à déterminer $\mathbb{P}(X = x_i)$ connaissant la loi jointe de (X, Y), nous devons nous appuyer sur le fait que l'ensemble des évènements $(Y = y_j)$ avec $j \in J$ est un système complet d'évènements et ainsi que pour tout $i \in J$:

$$\mathbb{P}(X = x_i) \;=\; \mathbb{P}\left(\bigcup_{j\in J}(X = x_i) \cap (Y = y_j)\right) \;=\; \sum_{j\in J} \mathbb{P}\left((X = x_i) \cap (Y = y_i)\right)$$

☛ Lorsque les variables X et Y sont finies et que la loi du couple est donnée sous forme d'un tableau à double entrée, les lois marginales de X et Y s'obtiennent en faisant la sommes des éléments sur chaque ligne ou sur chaque colonne.

Exemple : Soit le couple de V.A.R. finies (X, Y) dont la loi conjointe est donnée par

$X\backslash Y$	-1	1	2	7
-1	$\frac{1}{5}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$
4	$\frac{3}{10}$	$\frac{1}{6}$	0	$\frac{1}{6}$

<u>`Loi marginale de` X</u> :
Alors

$$\mathbb{P}(X = -1) = \mathbb{P}(X = -1, Y = -1)+\mathbb{P}(X = -1, Y = 1)+\mathbb{P}(X = -1, Y = 2)+\mathbb{P}(X = -1, Y = 7) = \frac{1}{5}+3\times\frac{1}{18} = \frac{11}{30}$$

Il s'agit ainsi de la somme des termes de la première ligne!! De même, on a :

$$\mathbb{P}(X = 4) = \mathbb{P}(X = 4, Y = -1)+\mathbb{P}(X = 4, Y = 1)+\mathbb{P}(X = 4, Y = 2)+\mathbb{P}(X = 4, Y = 7) = \frac{3}{10}+\frac{1}{6}+\frac{1}{6} = \frac{19}{30}$$

On a ainsi obtenue la loi marginale de X.
(On vérifie bien que $\mathbb{P}(X = -1) + \mathbb{P}(X = 4) = 1$!!)
<u>`Loi marginale de` Y</u> :
On a (en sommant les éléments de chaque colonne!!) :

$$\begin{aligned}
\mathbb{P}(Y = -1) &= \mathbb{P}(X = -1, Y = -1) + \mathbb{P}(X = 4, Y = -1) = \frac{1}{5} + \frac{3}{10} = \frac{1}{2}\\
\mathbb{P}(Y = 1) &= \mathbb{P}(X = -1, Y = 1) + \mathbb{P}(X = 4, Y = 1) = \frac{1}{18} + \frac{1}{6} = \frac{2}{9}\\
\mathbb{P}(Y = 2) &= \mathbb{P}(X = -1, Y = 2) + \mathbb{P}(X = 4, Y = 2) = \frac{1}{18}\\
\mathbb{P}(Y = 7) &= \mathbb{P}(X = -1, Y = 7) + \mathbb{P}(X = 4, Y = 7) = \frac{1}{18} + \frac{1}{6} = \frac{2}{9}
\end{aligned}$$

3. Indépendance de deux variables discrètes

Définition (Indépendance de deux V.A.R. finies)**.**

Soit (X, Y) un couple de variables discrètes tel que $X(\Omega) = \{x_i, i \in I\}$ et $Y(\Omega) = \{y_j, j \in J\}$. Soit $j_0 \in J$. On dit que X et Y sont **indépendantes** si pour tout couple $(x_i, y_j) \in X(\Omega) \times Y(\Omega)$ on a :

$$\mathbb{P}((X = x_i) \cap (Y = y_j)) = \mathbb{P}(X = x_i) \times \mathbb{P}(Y = y_j)$$

Autrement dit si et seulement si pour tout couple $(x_i, y_j) \in X(\Omega) \times Y(\Omega)$ les évènements $(X = x_i)$ et $(Y = y_j)$ sont indépendants.

☛ Dans le cas où X et Y sont deux variables indépendantes, la connaissance des lois de probabilités de X et Y permet de déterminer la loi conjointe du couple (X, Y).

Exemple - Exercice 1 : On lance un dé cubique équilibré. On note X_1 le résultat de ce lancé.
On lance un dé cubique pipé tel que le 6 a deux fois plus de chance de sortir que tout autre faces (celles-ci étant toutes équiprobables). On note X_2 le résultat du second lancé.
Les variables X_1 et X_2 sont supposées indépendantes. Compléter la loi du couple (X_1, X_2)

$X_1 \backslash X_2$	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

Exemple - Exercice 2 :
Un sac contient 4 boules numérotée de 1 à 4. On tire deux boules avec remise. On note X_1 le numéro de la première, X_2 celui de la seconde et Y le plus grand des deux!!
1. Déterminer la loi de X_1, celle de X_2 puis celle de Y.
2. Montrer que X_1 et X_2 sont indépendantes. Les variables X_1 et Y sont-elles indépendantes?

Propriété.

Soient X et Y deux V.A.R. finies. Soit f et g deux fonctions réelles.
Si X et Y sont indépendantes alors les variables $f(X)$ et $g(Y)$ sont également indépendantes.

☛ Ainsi si X et Y indépendantes alors X^2 et Y^2 sont indépendantes!!

Propriété (Indépendance et calculs d'espérances/variances)**.**

❶ Soient X et Y deux variables aléatoires discrètes possédant toutes deux une espérance et indépendantes, alors :

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

En utilisant la propriété ci-dessus. Si f et g sont des fonctions réelles et si X et Y indépendantes alors :

$$E(f(X)g(Y)) = E(f(X))E(g(Y))$$

❷ Soient X et Y deux variables aléatoires discrètes possédant toutes deux un moment d'ordre 2 et indépendantes, alors :

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

4. Covariance d'un couple de variables aléatoires

Définition.

Soient X et Y deux variables aléatoires telles que X et Y possède (une espérance et) une variance.
On appelle covariance de X et Y, noté $Cov(X,Y)$, le nombre réel :

$$Cov(X,Y) = E\left[(X - E(X))(Y - E(Y))\right] = E(XY) - E(X)E(Y)$$

Quelques remarques importantes ...

- La covariance d'un couple de variables peut être négative...contrairement à la variance d'une variable.
- $Cov(X, X) = V(X)$.
- $Cov(X, Y) = Cov(Y, X)$.
- On a pour tout couple de variables : $\boxed{Cov(X, Y)^2 \leqslant V(X)V(Y)}$

  cf. exercice 11 TD - Coefficient de corrélation linéaire

Définition (Variables non corrélées)**.**

Soit X et Y deux variables aléatoires discrètes.
On dit que X et Y sont non-corrélées lorsque : $Cov(X\ ,\ Y)\ =\ 0$

Propriétés (Propriétés calculatoires liées à la covariance)**.**

❶ **Linéarité à droite et à gauche** :
Soient a, b deux réels et X, Y, Z trois variables aléatoires possèdant une variance

$$Cov(aX + bY, Z) = aCov(X, Z) + bCov(Y, Z)$$

$$Cov(X, aY + bZ) = aCov(X, Z) + bCov(X, Y)$$

❷ Soient X et Y deux variables aléatoires possédant une variance alors

$$V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)$$

Propriété (Variance d'une somme de variables aléatoires)**.**

Soient X_1, X_2,...,X_n variables aléatoires possédant une variance alors

$$V(X_1 + X_2 + \ldots + X_n) = \sum_{k=1}^{n} V(X_i) + 2 \sum_{1\leqslant i<j\leqslant n} Cov(X_i, X_j)$$

Exercice 3.
Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes de même loi telle que $X_1(\Omega) = \{1, 2\}$ et $\mathbb{P}(X_1 = 1) = \frac{1}{4}$ et $\mathbb{P}(X_1 = 2) = \frac{3}{4}$

1. Calculer $V\left(\sum\limits_{k=1}^{n} X_k\right)$.
2. Pour $k \in \{1, \ldots, n-1\}$. On définit la V.A.R $Y_k = X_k + X_{k+1}$.
   (a) Déterminer $Y_k(\Omega)$ puis la loi de Y_k.
   (b) Les variables Y_1 et Y_2 sont-elles indépendantes? Calculer $V(\sum_{k=1}^{n} Y_k)$.

Propriété (Indépendance et non-corrélation)**.**

La covariance d'un couple de variables indépendantes est nulle :

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

☞ D'après cette propriété, deux variables indépéndantes sont non corrélées (covariance du couple nulle). Cependant, ces notions sont bien distinctes car la réciproque n'est pas vraie : il existe des variables non corrélées et pourtant dépendantes!!

Contre-exemple : variables dépendantes mais non corrélées.
Si X et Y sont deux variables aléatoires de Bernoulli de paramètre $\frac{1}{2}$. On observe que les variables aléatoires $X + Y$ et $\mid X - Y \mid$ sont dépendantes mais non correlées.