Chapitre 14 : couple de variables aléatoires discrètes

Dans tout le chapitre, X et Y sont des variables aléatoires discrètes.

I)Loi d'un couple, loi marginales, lois conditionnelles

Déf : on appelle couple (X,Y) l'application de Ω dans $\mathbf{R}^2$ définie par :
$\forall\omega\in\Omega,\ (X,Y)(\omega)=(X(\omega),Y(\omega))$.

Déf : on appelle loi du couple (X,Y) la donnée des couples de réels (x,y) et de la probabilité $P(X=x\cap Y=y)$ pour tout $x\in X(\Omega)$ et tout $y\in Y(\Omega)$.
La loi de X et la loi de Y sont appelées lois marginales.

Déf : pour $y\in Y(\Omega)$ avec $P(Y=y)\neq 0$, on appelle loi conditionnelle de X sachant $(Y=y)$ la donnée des couples $\big(x,P_{(Y=y)}(X=x)\big)$ où $x\in X(\Omega)$.

Déf : pour $x\in X(\Omega)$ avec $P(X=x)\neq 0$, on appelle loi conditionnelle de Y sachant $(X=x)$ la donnée des couples $\big(y,P_{(X=x)}(Y=y)\big)$ où $y\in Y(\Omega)$.

Théorème 1
1)$\forall x\in X(\Omega),\ P(X=x)=\displaystyle\sum_{y\in Y(\Omega)}P(X=x\cap Y=y)$.
2)$\forall y\in Y(\Omega),\ P(Y=y)=\displaystyle\sum_{x\in X(\Omega)}P(X=x\cap Y=y)$.

✓ Si la loi du couple est connue, on peut trouver les 2 lois marginales. Réciproque fausse.

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

<u>Exercice 1</u>
Dans une urne contenant 3 boules blanches et 2 boules rouges, on tire 3 boules une par une sans remise.
On note X et Y les variables aléatoires définies par :
$X=$nombre de boules rouges obtenues à l'issue des 3 tirages.
$Y=$rang d'obtention de la première boule blanche.
1)Déterminer la loi du couple (X,Y).
2)En déduire les lois marginales.
3)Déterminer la loi conditionnelle de Y sachant $(X=1)$.

II)Indépendance

Déf : on dit que X et Y sont indépendantes si pour tout $x\in X(\Omega)$ et tout $y\in Y(\Omega)$, on a $P(X=x\cap Y=y)=P(X=x)P(Y=y)$, c'est-à-dire si les événements $(X=x)$ et $(Y=y)$ sont indépendants.

Propriété 2
Si X et Y sont indépendantes, alors : $P(X\leq x\cap Y\leq y)=P(X\leq x)P(Y\leq y)$.

✓ On a aussi $P(X*x\cap Y**y)=P(X*x)P(Y**y)$ où * et ** sont des symboles pris parmi $<,>,=,\leq,\geq$.

Propriété 3 (lemme des coalitions)
Soient $f : \mathbf{R} \to \mathbf{R}$ et $g : \mathbf{R} \to \mathbf{R}$ des fonctions.
Si X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ aussi.

Exercice 2
Dans une urne contenant des boules blanches en proportion p et des boules noires en proportion q, on effectue une suite de tirages d'une boule avec remise jusqu'à ce qu'on obtienne 2 boules blanches au total.
On note X =rang d'apparition de la seconde boule blanche.
Soit Y définie par Y=1 si les 2 blanches ont été tirées consécutivement, Y=0 sinon.
1)Préciser $X(\Omega)$.
2)Calculer $P(X = i \cap Y = 1)$ pour $i \geq 2$.
3)En déduire la loi de Y.
4)Calculer $P(X = 2 \cap Y = 0)$. X et Y sont-elles indépendantes ?

Exercice 3
On lance une pièce de monnaie indéfiniment. On suppose que $P(pile) = p \in]0, 1[$.
On note X_1 =rang d'obtention du $1er$ pile et X_2 =rang d'obtention du $2nd$ pile.
1)Reconnaître la loi de X_1.
2)Déterminer la loi du couple (X_1, X_2).
3)En déduire la loi de X_2.
4)Reconnaître la loi de $X_2 - X_1$. En déduire $E(X_2)$.

✓ X_2 suit la loi de Pascal de paramètres 2 et p.

III)Fonction de deux variables aléatoires discrètes

Déf : soit $g : \mathbf{R}^2 \to \mathbf{R}$ une fonction de deux variables.
Par composition des applications g et (X, Y), on définit une nouvelle variable aléatoire discrète $Z = go(X, Y)$ que l'on note $Z = g(X, Y)$.

Exemple :
$Z = X + Y$, $Z = XY$, $Z = min(X, Y)$, $Z = max(X, Y)$, $Z = |X - Y|$.

Théorème 2 (thm de transfert)
Soit $g : \mathbf{R}^2 \to \mathbf{R}$. Sous réserve de convergence, on a :

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

✓ Ce thm permet de démontrer la linéarité de l'espérance.

Théorème 3
Si X et Y admettent une variance, alors XY admet une espérance donnée par :

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

Théorème 4
Soient X et Y des variables aléatoires admettant une espérance.
Si X et Y sont indépendantes, alors XY admet une espérance et on a :

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

IV)Corrélation linéaire

Déf : on appelle covariance de X et Y le nombre réel, noté $cov(X,Y)$, défini sous réserve d'existence par :

$$cov(X,Y) = E\big((X - E(X))(Y - E(Y))\big).$$

$$\checkmark\ cov(X,Y) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} \big(x - E(X)\big)\big(y - E(Y)\big)P(X = x \cap Y = y).$$

Théorème 5 (formule de Huygens)
Si X et Y admettent une variance, alors $cov(X,Y)$ existe et est donnée par :

$$cov(X,Y) = E(XY) - E(X)E(Y).$$

Propriété 4 (réciproque fausse)
Si X et Y sont indépendantes, alors $cov(X,Y) = 0$.

Propriété 5
Soient a et b des réels quelconques. On a sous réserve d'existence :

1)$cov(X,X) = V(X)$.

2)$cov(X,Y) = cov(Y,X)$ (symétrie).

3)$cov(aX_1 + bX_2, Y) = acov(X_1,Y) + bcov(X_2,Y)$ (linéarité à gauche).

4)$cov(X, aY_1 + bY_2) = acov(X,Y_1) + bcov(X,Y_2)$ (linéarité à droite).

Propriété 6
Si X et Y admettent une variance, alors $X + Y$ admet une variance donnée par :

$$V(X + Y) = V(X) + V(Y) + 2cov(X,Y).$$

✓ Si X et Y sont indépendantes, alors $V(X + Y) = V(X) + V(Y)$.

Propriété 6bis
Si X et Y admettent une variance, alors $X - Y$ admet une variance donnée par :

$$V(X - Y) = V(X) + V(Y) - 2cov(X,Y).$$

✓ Si X et Y sont indépendantes, alors $V(X - Y) = V(X) + V(Y)$.

Déf : si X et Y admettent une variance, on appelle coefficient de corrélation linéaire de X et Y, le nombre réel, noté $\rho(X,Y)$, défini par :

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

Propriété 7

1)$-1 \leq \rho(X,Y) \leq 1$.

2)$\rho(X,Y) = 1 \Longleftrightarrow \exists a > 0, b \in \mathbf{R}, P(Y = aX + b) = 1$.

3)$\rho(X,Y) = -1 \Longleftrightarrow \exists a < 0, b \in \mathbf{R}, P(Y = aX + b) = 1$.

✓ $\rho(X,Y) = \pm 1 \Leftrightarrow Y$ est une fonction presque sûrement affine de X.

Exercice 4

Soient X et Y variables aléatoires avec $X(\Omega) = \{1, 2\}$ et $Y(\Omega) = \{-1, 0\}$.

La loi du couple vaut $P(X = 1 \cap Y = -1) = \dfrac{1}{10}$, $P(X = 1 \cap Y = 0) = \dfrac{3}{10}$, $P(X = 2 \cap Y = -1) = \dfrac{2}{10}$ et $P(X = 2 \cap Y = 0) = \dfrac{4}{10}$.

1)Déterminer les lois marginales.

2)Calculer $E(X)$, $E(Y)$ et $E(XY)$.

3)Déterminer la covariance de X et Y.

4)Déterminer le coefficient de corrélation linéaire de X et Y.

V)Annexe : somme double

Déf : soit $A = (a_{ij})_{1\leq i\leq n, 1\leq j\leq p}$ une matrice à n lignes et p colonnes.
En ajoutant les coefficients de A, on obtient une somme, appelée somme double et notée $\displaystyle\sum_{\substack{1\leq i\leq n\\1\leq j\leq p}} a_{ij}$ ou $\displaystyle\sum_{(i,j)\in I\times J} a_{ij}$ avec $I = [\![1, n]\!]$ et $J = [\![1, p]\!]$.

Introduction

Prenons $A = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 6 & 7 \end{pmatrix}$. Il y a beaucoup de façons de calculer $\displaystyle\sum_{\substack{1\leq i\leq 2\\1\leq j\leq 3}} a_{ij}$

$\displaystyle\sum_{\substack{1\leq i\leq 2\\1\leq j\leq 3}} a_{ij}$=2+7+6+3+1+5=24 ou $\displaystyle\sum_{\substack{1\leq i\leq 2\\1\leq j\leq 3}} a_{ij}$=3+5+1+2+6+7=24 etc...

Parmi toutes ces façons, il y en a deux particulières :

$$\begin{aligned}\sum_{\substack{1\leq i\leq 2\\1\leq j\leq 3}} a_{ij} &= (1 + 2 + 5) + (3 + 6 + 7)\\
&= (a_{11} + a_{12} + a_{13}) + (a_{21} + a_{22} + a_{23})\\
&= \sum_{j=1}^{3} a_{1j} + \sum_{j=1}^{3} a_{2j}\\
&= \sum_{i=1}^{2}\left(\sum_{j=1}^{3} a_{ij}\right) \text{[somme sur les lignes].}\end{aligned}$$

$$\begin{aligned}\sum_{\substack{1\leq i\leq 2\\1\leq j\leq 3}} a_{ij} &= (1 + 3) + (2 + 6) + (5 + 7)\\
&= (a_{11} + a_{21}) + (a_{12} + a_{22}) + (a_{13} + a_{23})\\
&= \sum_{i=1}^{2} a_{i1} + \sum_{i=1}^{2} a_{i2} + \sum_{i=1}^{2} a_{i3}\\
&= \sum_{j=1}^{3}\left(\sum_{i=1}^{2} a_{ij}\right) \text{[somme sur les colonnes].}\end{aligned}$$

Propriété 8 (sommation sur un rectangle)

$$\sum_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} a_{ij} = \sum_{i=1}^{n}\left(\sum_{j=1}^{p} a_{ij}\right) \text{[somme sur les lignes]}$$
$$= \sum_{j=1}^{p}\left(\sum_{i=1}^{n} a_{ij}\right) \text{[somme sur les colonnes].}$$

Propriété 9 (sommation sur un triangle – inégalité large)

$$\sum_{1\leq i\leq j\leq n} a_{ij} = \sum_{i=1}^{n}\left(\sum_{j=i}^{n} a_{ij}\right) \text{[somme sur les lignes]}$$
$$= \sum_{j=1}^{n}\left(\sum_{i=1}^{j} a_{ij}\right) \text{[somme sur les colonnes].}$$

Propriété 10 (sommation sur un triangle – inégalité stricte)

$$\sum_{1\leq i< j\leq n} a_{ij} = \sum_{i=1}^{n-1}\left(\sum_{j=i+1}^{n} a_{ij}\right) \text{[somme sur les lignes]}$$
$$= \sum_{j=2}^{n}\left(\sum_{i=1}^{j-1} a_{ij}\right) \text{[somme sur les colonnes].}$$

<u>Exercice 5</u>

1)Calculer la somme double $\displaystyle\sum_{\substack{1\leq i\leq n\\ 1\leq j\leq n}} ij^2$ en fonction de n.

2)Calculer la somme double $\displaystyle\sum_{1\leq i\leq j\leq n} ij^2$ en fonction de n.