Chapitre 3 : Variables aléatoires sur un espace de probabilité discret
(Partie III)

Nathaël Gozlan

25 novembre

Plan

Couples de variables aléatoires indépendantes

Soit $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ un espace de probabilité discret.

Définition

Soient $X : \Omega \to E$ et $Y : \Omega \to F$ deux variables aléatoires à valeurs dans des ensembles finis ou dénombrables E et F ; on dit que X et Y sont indépendantes si

$$\mathbb{P}(X = a, Y = b) = \mathbb{P}(X = a)\mathbb{P}(Y = b), \qquad \forall a \in E, \forall b \in F.$$

Autrement dit X et Y sont indépendantes si et seulement si les événements $\{X = a\}$ et $\{Y = b\}$ sont indépendants pour tout couple $(a, b) \in E \times F$.

Intuitivement, lorsque deux variables sont indépendantes, la connaissance de la valeur prise par l'une ne permet pas d'inférer quoique ce soit sur la valeur prise par l'autre.

Couples de variables aléatoires indépendantes

Proposition

Avec les notations précédentes, il y a équivalence entre

1. X et Y sont indépendantes,
2. Pour toutes fonctions bornées $f : E \to \mathbb{R}$ et $g : F \to \mathbb{R}$,
$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$$
3. Pour tous $A \subset E$, $B \subset F$,
$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

Preuve

• Montrons que (1) $\Rightarrow$ (2). Supposons que X et Y sont indépendantes.
Posons $Z = (X, Y)$ et $h(x, y) = f(x)g(y)$, $(x, y) \in E \times F$.
La variable aléatoire Z est à valeurs dans $E \times F$. Sa loi est donnée par

$$\mathbb{P}_Z(\{(x, y)\}) = \mathbb{P}((X, Y) = (x, y)) = \mathbb{P}(X = x)\mathbb{P}(Y = y) = \mathbb{P}_X(\{x\})\mathbb{P}_Y(\{y\})$$

D'après la formule de transfert

$$\mathbb{E}[f(X)g(Y)] = \sum_{(x,y)\in E\times F} h(x, y)\mathbb{P}_Z(\{(x, y)\}) = \sum_{(x,y)\in E\times F} h(x, y)\mathbb{P}_X(\{x\})\mathbb{P}_Y(\{y\})$$

Mais,

$$\sum_{(x,y)\in E\times F} h(x, y)\mathbb{P}_X(\{x\})\mathbb{P}_Y(\{y\}) = \left(\sum_{x\in E} f(x)\mathbb{P}_X(\{x\})\right)\left(\sum_{y\in F} g(y)\mathbb{P}_Y(\{y\})\right) = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$$

• Montrons que (2) $\Rightarrow$ (3). Supposons que $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$ pour toutes fonctions f, g.
En prenant $f = \mathbf{1}_A$ et $g = \mathbf{1}_B$, on obtient

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{E}[\mathbf{1}_A(X)\mathbf{1}_B(Y)] = \mathbb{E}[\mathbf{1}_A(X)]\mathbb{E}[\mathbf{1}_B(Y)] = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

• Montrons que (3) $\Rightarrow$ (1). Supposons que $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$ pour tout $A \subset E$ et $B \subset F$.
En prenant $A = \{a\}$ et $B = \{b\}$, on trouve

$$\mathbb{P}(X = a, Y = b) = \mathbb{P}(X = a)\mathbb{P}(Y = b)$$

et donc X et Y sont indépendantes.

Couples de variables aléatoires indépendantes

Corollaire

Si X et Y sont des variables aléatoires indépendantes ayant un moment d'ordre 2 fini, alors $\mathrm{Cov}(X, Y) = 0$ et

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Démonstration.

En posant $m_X = \mathbb{E}[X]$ et $m_Y = \mathbb{E}[Y]$, on a

$$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - m_X)(Y - m_Y)] = \mathbb{E}[(X - m_X)]\mathbb{E}(Y - m_Y)] = 0$$

□

Lorsque $\mathrm{Cov}(X, Y) = 0$, on dit que X et Y sont *décorrélées.*
Attention, deux variables aléatoires décorrélées ne sont pas forcément indépendantes.

Exercice 1

Soit X une variable à valeurs dans $\{-1, 0, 1\}$ telle que

$$\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/3$$

et posons $Y = \mathbf{1}_{\{X=0\}}$.
Montrer que les variables X et Y sont décorrélées mains ne sont pas indépendantes.

Plan

Famille de variables aléatoires indépendantes

Soient $X_1, \ldots, X_n$ des variables aléatoires définies sur le même espace de probabilité discret $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ à valeurs dans des espaces finis ou dénombrables $E_1, \ldots, E_n$.
Le vecteur $(X_1, \ldots, X_n)$ est appelé *vecteur aléatoire*.

Définition

On dit que $(X_1, \ldots, X_n)$ est une famille de variables aléatoires indépendantes si elle vérifie les trois conditions équivalentes suivantes :

1. Pour tout $(a_1, \ldots, a_n) \in E_1 \times \cdots \times E_n$,
$$\mathbb{P}(X_1 = a_1, \ldots, X_n = a_n) = \mathbb{P}(X_1 = a_1) \cdots \mathbb{P}(X_n = a_n).$$
2. Pour toutes fonctions bornées $f_1 : E_1 \to \mathbb{R}$, ..., $f_n : E_n \to \mathbb{R}$, on a
$$\mathbb{E}[f_1(X_1) \cdots f_n(X_n)] = \mathbb{E}[f_1(X_1)] \cdots \mathbb{E}[f_n(X_n)].$$
3. Pour tous sous-ensembles $A_1 \subset E_1, \ldots, A_n \subset E_n$,
$$\mathbb{P}(X_1 \in A_1, \ldots, X_n \in A_n) = \mathbb{P}(X_1 \in A_1) \cdots \mathbb{P}(X_n \in A_n).$$

L'équivalence entre ces propriétés se montre exactement comme dans le cas de deux variables.

Famille de variables aléatoires indépendantes

Proposition

Soient $X_1, \ldots, X_n$ des variables aléatoires à valeurs réelles.

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j).$$

En particulier, si $(X_1, \ldots, X_n)$ est une famille de variables aléatoires indépendantes, alors

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i).$$

Exercice 2

Démontrer la proposition précédente.

Suite de variables aléatoires indépendantes

Définition

On dit qu'une suite $(X_i)_{i\geq 1}$ est une suite de variables aléatoires indépendantes si pour tout $n \geq 1$, $(X_1, \ldots, X_n)$ est une famille de variables aléatoires indépendantes.

On dit que des variables aléatoires $X_1, \ldots, X_n$ à valeurs dans le même espace E fini ou dénombrable sont *identiquement distribuées* si $\mathbb{P}_{X_1} = \ldots = \mathbb{P}_{X_n}$.

Quand $(X_i)_{i\geq 1}$ est une suite de variables aléatoires *indépendantes et identiquement distribuées* on parle d'une suite de variables aléatoires *i.i.d*.

Plan

Loi binomiale et nombre de succès

Théorème

Soit $(X_1, \ldots, X_n)$ une famille de variables aléatoires indépendantes telles que pour tout i, X_i suit la loi de Bernoulli $\mathcal{B}(p)$ de paramètre $p \in [0, 1]$.
Alors $S = \sum_{i=1}^{n} X_i$ suit une loi binomiale $\mathcal{B}(n, p)$.

Exercice 3

On réalise 5 lancers indépendants d'une pièce tombant sur pile avec probabilité $1/3$.
Quelle est la probabilité d'obtenir au moins 3 piles ?

Exercice 4

Retrouver l'espérance et la variance d'une variable suivant la loi binomiale $\mathcal{B}(n, p)$.

Preuve

Pour tout $k \in \{0, \ldots, n\}$,

$$\{S = k\} = \bigcup_{I \subset \{1,2,\ldots,n\}, \mathrm{Card}(I)=k} \left(\bigcap_{i \in I} \{X_i = 1\} \right) \cap \left(\bigcap_{i \in I^c} \{X_i = 0\} \right).$$

I représente les instants de succès, et I^c les instants d'échec.
Par union disjointe puis indépendance et équidistribution des variables, on trouve

$$\begin{aligned}
\mathbb{P}(S = k) &= \sum_{I \subset \{1,2,\ldots,n\}, \mathrm{Card}(I)=k} \mathbb{P}\left(\bigcap_{i \in I} \{X_i = 1\} \cap \bigcap_{i \in I^c} \{X_i = 0\} \right) \\
&= \sum_{I \subset \{1,2,\ldots,n\}, \mathrm{Card}(I)=k} \mathbb{P}(X_1 = 1)^k \mathbb{P}(X_1 = 0)^{n-k} \\
&= \binom{n}{k} p^k (1-p)^{n-k}.
\end{aligned}$$

Loi géométrique et temps du premier succès

Théorème

Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires indépendantes telles que pour tout i, X_i suit une loi de Bernoulli $\mathcal{B}(p)$ de paramètre $p \in]0,1[$. Posons

$$T = \inf\{k \geq 1 : X_k = 1\},$$

avec la convention $\inf \emptyset = +\infty$.
Avec probabilité 1, $T < \infty$ et de plus T suit une loi géométrique de paramètre p.

Démonstration.

Par définition de T,

$$\{T = 1\} = \{X_1 = 1\}$$

et pour tout $k \geq 2$

$$\{T = k\} = \{X_1 = 0, \ldots, X_{k-1} = 0, X_k = 1\}.$$

On voit donc que pour tout $k \geq 2$

$$\mathbb{P}(T = k) = \mathbb{P}(X_1 = \ldots = X_{k-1} = 0, X_k = 1) = (1-p)^{k-1}p$$

et l'égalité est encore vraie pour $k = 1$. $\square$

Plan

Loi d'un couple de variables aléatoires - cas général

Définition (Loi d'un couple de variables aléatoires)

Soient X, Y deux variables aléatoires à valeurs respectivement dans E et F (finis ou dénombrables). La loi du couple (X, Y) est la mesure de probabilité $\mathbb{P}_{(X,Y)}$ sur $E \times F$ définie par

$$\mathbb{P}_{(X,Y)}(\{(a, b)\}) = \mathbb{P}(X = a, Y = b), \qquad \forall (a, b) \in E \times F.$$

Loi d'un couple de variables aléatoires - cas général

Connaissant la loi du couple (X, Y) on peut retrouver facilement la loi de X et celle de Y, comme le montre la proposition suivante :

Proposition

Avec les notations précédentes, pour tout $a \in E$,

$$\mathbb{P}_X(\{a\}) = \sum_{b \in F} \mathbb{P}_{(X,Y)}(\{(a, b)\})$$

et pour tout $b \in F$,

$$\mathbb{P}_Y(\{b\}) = \sum_{a \in E} \mathbb{P}_{(X,Y)}(\{(a, b)\}).$$

Démonstration.

Il suffit de remarquer que

$$\{X = a\} = \bigcup_{b \in F} \{X = a, Y = b\}$$

(union disjointe et dénombrable). □

Loi d'un couple de variables aléatoires - cas général

Proposition

Pour toute fonction $f : E \times F \to \mathbb{R}$ telle que $f(X, Y)$ soit intégrable, on a

$$\mathbb{E}[f(X, Y)] = \sum_{(a,b)\in E\times F} f(a, b)\mathbb{P}_{(X,Y)}(\{(a, b)\}).$$

Démonstration.

C'est la formule de transfert appliquée à $Z = (X, Y)$. $\square$

Loi d'un couple de variables aléatoires - cas général

En fait la loi du couple contient plus d'information que la simple donnée des lois de X et de Y.

Définition (Lois conditionnelles)

Avec les notations précédentes, pour tout $b \in F$ tel que $\mathbb{P}(Y = b) \neq 0$, la loi conditionnelle de X sachant $Y = b$, notée $\mathbb{P}_{X \mid Y=b}$, est définie sur E par

$$\mathbb{P}_{X|Y=b}(\{a\}) = \mathbb{P}(X = a|Y = b) = \frac{\mathbb{P}(X = a, Y = b)}{\mathbb{P}(Y = b)}, \qquad \forall a \in E.$$

La loi conditionnelle de X décrit comment la connaissance de la valeur prise par Y influe sur la valeur prise par X.

Proposition

Deux variables aléatoires X, Y sont indépendantes si et seulement si pour tout $b \in F$ tel que $\mathbb{P}(Y = b) \neq 0$, on a

$$\mathbb{P}_{X|Y=b} = \mathbb{P}_X.$$

Démonstration.

Evident. □