Couples et vecteurs de variables aléatoires
Préparation à l'agrégation interne

1 Couples et vecteurs aléatoires discrets

1.1 Loi conjointe

On se donne X et Y deux variables aléatoires discrètes avec $X(\Omega) = \{x_i, i \in \mathbb{N}\}$ et $Y(\Omega) = \{y_j, j \in \mathbb{N}\}$. La **loi conjointe** du couple (X, Y) est donnée par $(X, Y)(\Omega)$ (ou par $X(\Omega)$ et $Y(\Omega)$) ainsi que par les probabilités

$$\mathbf{P}(X = x, Y = y) = \mathbf{P}\{\omega, X(\omega) = x \text{ et } Y(\omega) = y\}, \quad \text{pour tout couple } (x, y) \in (X, Y)(\Omega).$$

Remarque : On doit bien entendu avoir $\sum_{x,y} \mathbf{P}(X = x, Y = y) = 1$.

Plus généralement, si $X_1, \ldots, X_n$ sont n variables aléatoires discrètes à valeurs dans $\mathbb{N}$, la loi conjointe du vecteur $(X_1, \ldots, X_n)$ est donnée par l'ensemble-image $(X_1, \ldots, X_n)(\Omega) \subset \mathbb{N}^n$ ainsi que par les probabilités $\mathbf{P}(X_1 = i_1, \ldots, X_n = i_n)$, pour tout n-uplet $(i_1, \ldots, i_n) \in \mathbb{N}^n$.

Exemple 1 : Fixons $p \in]0, 1[$ et $\lambda > 0$ et considérons le couple de variables aléatoires (X, Y) à valeurs dans $\{0, 1\} \times \mathbb{N}$ dont la loi est donnée par :

$$\begin{aligned}
\mathbf{P}(X = 0, Y = 0) &= 1 - p \\
\mathbf{P}(X = 1, Y = k) &= pe^{-\lambda}\lambda^k/k!, \text{ pour tout } k \in \mathbb{N} \\
\mathbf{P}(X = j, Y = k) &= 0 \text{ sinon.}
\end{aligned}$$

On a bien $\sum_{i,j} \mathbf{P}(X = i, Y = j) = 1$: on a donc bien écrit la loi d'un couple aléatoire discret.

Exemple 2 : On dispose d'une urne contenant quatre jetons numérotés de 1 à 4, et on tire au sort successivement deux jetons sans remise. On note (X, Y) les résultats des deux tirages.

On a : $\mathbf{P}(X = i, Y = i) = 0$ pour tout i entre 1 et 4 et $\mathbf{P}(X = i, Y = j) = 1/12$ si $1 \leq i, j \leq 4$ et $i \neq j$.

On peut écrire les probabilités sous la forme du tableau suivant (où par exemple dans la deuxième case de la première ligne, on lit $\mathbf{P}(X = 1, Y = 2)$) :

$X \backslash Y$	1	2	3	4
1	0	1/12	1/12	1/12
2	1/12	0	1/12	1/12
3	1/12	1/12	0	1/12
4	1/12	1/12	1/12	0

Exemple 3 : Loi trinomiale. On se fixe un nombre entier n strictement positif et deux paramètres réels positifs p_x et p_y tels que $p_x + p_y \leq 1$. La loi trinomiale (n, p_x, p_y) est la loi du couple (X, Y) tel que $(X, Y)(\Omega) \in \mathbb{N}^2$ et donnée pour tout $(i, j) \in \mathbb{N}^2$ tels que $i + j \leq n$ par :

$$\mathbf{P}(X = i, Y = j) = \frac{n!}{i!j!(n - i - j)!} p_x^i p_y^j (1 - p_x - p_y)^{n-i-j},$$

et $\mathbf{P}(X = i, Y = j) = 0$ sinon.

Exercice : Montrer que l'on définit bien ainsi la loi d'un couple aléatoire.

La loi trinomiale est une extension de la loi binomiale. Imaginons en effet une expérience qui a trois issues possibles, notée x, y et z, avec comme probabilité de réalisation p_x, p_y et $p_z = 1-p_x-p_y$. Répétons n fois cette expérience (n est fixé) de façon indépendante et comptons le nombre d'apparitions de x (nombre noté X) et de y (noté Z) parmi ces n répétitions. C'est alors un exercice de dénombrement que de démontrer que le couple (X, Y) suit alors une loi trinomiale de paramètres (n, p_x, p_y).

1.2 Lois marginales

Définition 1.1 *Les (deux)* **lois marginales** *du couple* (X, Y) *sont les lois des variables aléatoires* X *et* Y*. On les obtient de la façon suivante :*

$$\begin{aligned}
\mathbf{P}(X = x) &= \sum_{y \in Y(\Omega)} \mathbf{P}(X = x, Y = y), \\
\mathbf{P}(Y = y) &= \sum_{x \in X(\Omega)} \mathbf{P}(X = x, Y = y).
\end{aligned}$$

Preuve : On a

$$\{X = x\} = \{X = x, Y \in Y(\Omega)\} = \bigcup_{y \in Y(\Omega)} \{X = x, Y = y\}.$$

Comme la réunion est dénombrable et disjointe, il vient :

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

$\square$

Plus généralement, un vecteur $(X_1, \ldots, X_n)$ à valeurs dans $\mathbb{Z}^n$ possède n lois marginales unidimensionnelles, mais également $n(n-1)$ lois marginales bidimensionnelles, et ainsi de suite. On a par exemple

$$\mathbf{P}(X_1 = x) = \sum_{(x_2, \ldots, x_n) \in \mathbb{Z}^{n-1}} \mathbf{P}(X_1 = x, X_2 = x_2, \ldots, X_n = x_n).$$

Reprenons les exemples précédents :
Exemple 1.
Déterminons la loi de X : X est à valeurs dans $\{0, 1\}$ et on a :

$$\mathbf{P}(X = 0) = \sum_{j \in \mathbb{N}} \mathbf{P}(X = 0, Y = j) = 1 - p.$$

De la même façon :

$$\mathbf{P}(X = 1) = \sum_{j \in \mathbb{N}} \mathbf{P}(X = 1, Y = j) = \sum_{j \geq 0} p e^{-\lambda} \lambda^j / j! = p.$$

La variable aléatoire X suit donc une loi de Bernoulli de paramètre p. Calculons aussi la loi de Y :

$$\mathbf{P}(Y=0) = \mathbf{P}(X=0, Y=0) + \mathbf{P}(X=1, Y=0) = 1 - p + pe^{-\lambda}.$$

Et pour tout $j \geq 1$,

$$\mathbf{P}(Y=j) = \mathbf{P}(X=0, Y=j) + \mathbf{P}(X=1, Y=j) = pe^{-\lambda}\lambda^j/j!.$$

Exemple 2 : Il suffit de sommer en colonne pour avoir la loi de X, et en ligne pour obtenir celle de Y. En pratique, on peut ajouter une colonne et une ligne au tableau pour y écrire les lois de X et Y. Et avant de conclure, on prend le soin de vérifier que la somme de cette colonne (et de cette ligne) vaut 1.
On trouve ici que X et Y suivent une loi uniforme sur $\{1, 2, 3, 4\}$.

Exemple 3 : On considère le couple (X, Y) de loi trinomiale (n, p_x, p_y). Déterminons la loi marginale de X : fixons $j \in \{0, \ldots, n\}$ et évaluons $\mathbf{P}(X=j)$.

On a

$$\begin{aligned}
\mathbf{P}(X=j) &= \sum_{k=0}^{n} \mathbf{P}(X=j, Y=k) \\
&= \sum_{k=0}^{n-j} \mathbf{P}(X=j, Y=k) + \sum_{k=n-j+1}^{n} \mathbf{P}(X=j, Y=k) \\
&= \sum_{k=0}^{n-j} \frac{n!}{j!k!(n-j-k)!} p_x^j p_y^k (1-p_x-p_y)^{n-j-k} + 0 \\
&= \frac{n!}{j!(n-j)!} p_x^j \sum_{k=0}^{n-j} \frac{(n-j)!}{k!(n-j-k)!} p_y^k (1-p_x-p_y)^{n-j-k} \\
&= \binom{n}{j} p_x^j (1-p_x)^{n-j}
\end{aligned}$$

La variable aléatoire X suit donc une loi $\mathrm{Bin}(n, p_x)$. Un calcul similaire montre que Y suit une loi binomiale $\mathrm{Bin}(n, p_y)$.

1.3 Loi de $f(X, Y)$

Problème : On dispose d'un couple de variables aléatoires discrètes (X, Y) dont on connaît la loi conjointe et on voudrait connaître la loi de la variable aléatoire $Z = f(X, Y)$, où $f : X(\Omega) \times Y(\Omega) \to \mathbb{R}$ est une fonction donnée. Par exemple, on a souvent besoin de connaître la loi de $X + Y$, ou celle de $X - Y$, ou de XY. Et déterminer la loi de X à partir de celle de (X, Y) revient à considérer la fonction $f(x, y) = x$.

Proposition 1.2 *On a $Z(\Omega) = f((X, Y)(\Omega))$ et pour tout $z \in f((X, Y)(\Omega))$, on a*

$$\mathbf{P}(Z=z) = \sum_{(x,y) \in (X,Y)(\Omega), f(x,y)=z} \mathbf{P}(X=x, Y=y).$$

Exemple : Reprenons une nouvelle fois l'exemple 1 et considérons la fonction $f(x,y) = xy$. La variable aléatoire XY est à valeurs dans $\mathbb{N}$ et on a

$$\mathbf{P}(XY=0) = \mathbf{P}(X=0, Y=0) + \mathbf{P}(X=1, Y=0) = 1-p+pe^{-\lambda}$$

et, pour tout $k \in \mathbb{N}^*$,

$$\mathbf{P}(XY=k) = \mathbf{P}(X=1, Y=k) = pe^{-\lambda}\frac{\lambda^k}{k!}.$$

Un cas particulier important. Nous considérons ici la fonction $f(x,y) = x+y$. On obtient :

$$\begin{aligned}
\mathbf{P}(X+Y=z) &= \sum_{(x,y)\in(X,Y)(\Omega), x+y=z} \mathbf{P}(X=x, Y=y) \\
&= \sum_{x\in X(\Omega)} \mathbf{P}(X=x, Y=z-x) \\
&= \sum_{y\in Y(\Omega)} \mathbf{P}(X=z-y, Y=y)
\end{aligned}$$

Plus généralement, si $\overline{X} = (X_1, \ldots, X_n)$ est un vecteur aléatoire discret et $f : \overline{X}(\Omega) \to \mathbb{R}$ est une fonction donnée, on a :

$$\mathbf{P}(f(X_1,\ldots,X_n) = z) = \sum_{x_1,\ldots,x_n, f(x_1,\ldots,x_n)=z} \mathbf{P}(X_1 = x_1, \ldots, X_n = x_n).$$

On en déduit le corollaire fondamental suivant :

Proposition 1.3 *Soient X et Y deux variables aléatoires discrètes et intégrables. Alors la variable aléatoire $Z = X+Y$ est intégrable et on a $\mathbf{E}(X+Y) = \mathbf{E}(X) + \mathbf{E}(Y)$.*

Preuve : En effet, on vient de voir que Z est une variable aléatoire discrète et que sa loi est donnée par : pour tout $z \in (X+Y)(\Omega)$,

$$\mathbf{P}(X+Y=z) = \sum_{x\in X(\Omega)} \mathbf{P}(X=x, Y=z-x)$$

On a donc

$$\mathbf{E}(|X+Y|) = \sum_{z\in(X+Y)(\Omega)} \left[|z| \sum_{x\in X(\Omega)} \mathbf{P}(X=x, Y=z-x) \right]$$

Puis

$$\mathbf{E}(|X+Y|) = \sum_{z\in(X+Y)(\Omega)} \left[\sum_{x\in X(\Omega)} |x+(z-x)| \mathbf{P}(X=x, Y=z-x) \right]$$

En utilisant l'inégalité triangulaire, il vient :

$$\begin{aligned}
\mathbf{E}(|X+Y|) &\leq \sum_{z\in(X+Y)(\Omega)} \left[\sum_{x\in X(\Omega)} (|x|+|z-x|)\mathbf{P}(X=x, Y=z-x)\right] \\
&\leq \sum_{z\in(X+Y)(\Omega)} \left[\sum_{x\in X(\Omega)} |x|\mathbf{P}(X=x, Y=z-x)\right] + \\
&\qquad + \sum_{z\in(X+Y)(\Omega)} \left[\sum_{x\in X(\Omega)} |z-x|\mathbf{P}(X=x, Y=z-x)\right]
\end{aligned}$$

Étudions tout d'abord la première somme :

$$\begin{aligned}
\sum_{z\in(X+Y)(\Omega)} \left[\sum_{x\in X(\Omega)} |x|\mathbf{P}(X=x, Y=z-x)\right] &= \sum_{x\in X(\Omega)} \left[|x| \sum_{z\in(X+Y)(\Omega)} \mathbf{P}(X=x, Y=z-x)\right] \\
&= \sum_{x\in X(\Omega)} [|x|\mathbf{P}(X=x)] = \mathbf{E}(|X|)
\end{aligned}$$

Passons à la deuxième somme : sommer sur $x \in X(\Omega)$ ou sur x tel que $z-x \in Y(\Omega)$ ne change pas la valeur de cette somme. On a donc

$$\begin{aligned}
\sum_{z\in(X+Y)(\Omega)} \left[\sum_{x\in X(\Omega)} |z-x|\mathbf{P}(X=x, Y=z-x)\right] &= \sum_{z\in(X+Y)(\Omega)} \left[\sum_{z-x\in Y(\Omega)} |z-x|\mathbf{P}(X=x, Y=z-x)\right] \\
&= \sum_{z\in(X+Y)(\Omega)} \left[\sum_{y\in Y(\Omega)} |y|\mathbf{P}(X=z-y, Y=y)\right] \\
&= \sum_{y\in Y(\Omega)} \left[|y| \sum_{z\in(X+Y)(\Omega)} \mathbf{P}(X=z-y, Y=y)\right] \\
&= \sum_{y\in Y(\Omega)} |y|\mathbf{P}(Y=y) = \mathbf{E}(|Y|)
\end{aligned}$$

On remarque donc que si X et Y sont intégrables, $X+Y$ l'est également, et en effectuant le même calcul sans les valeurs absolues, on conclut que $\mathbf{E}(X+Y) = \mathbf{E}(X) + \mathbf{E}(Y)$.

Exemple : Supposons que (X, Y) suit une loi trinomiale (n, p_x, p_y) et calculons la loi de $X+Y$. Cette variable aléatoire est à valeurs dans $\mathbb{N}$ et on a, pour tout entier k :

$$\mathbf{P}(X+Y=k) = \sum_{j=0}^{n} \mathbf{P}(X=j, Y=k-j).$$

Pour tout $k > n$, chacun des termes de cette somme est nul donc $\mathbf{P}(X+Y=k) = 0$.

Fixons maintenant un entier $k \in \{0, \dots, n\}$. On a :

$$\begin{aligned}
\mathbf{P}(X+Y=k) &= \sum_{j=0}^{k} \mathbf{P}(X=j, Y=k-j) \\
&= \sum_{j=0}^{k} \frac{n!}{j!(k-j)!(n-k)!} p_x^j p_y^{k-j} (1-p_x-p_y)^{n-k} \\
&= \frac{n!}{(n-k)!}(1-p_x-p_y)^{n-k} \sum_{j=0}^{k} \frac{1}{j!(k-j)!} p_x^j p_y^{k-j} \\
&= \frac{n!}{k!(n-k)!}(1-p_x-p_y)^{n-k}(p_x+p_y)^k.
\end{aligned}$$

La variable aléatoire $X+Y$ suit donc une loi binomiale $\mathrm{Bin}(n, p_x+p_y)$.

1.4 Loi conditionnelle

Considérons un couple (X, Y) de variables aléatoires discrètes, dont on connaît la loi jointe et fixons y tel que $\mathbf{P}(Y=y) > 0$.

La **loi conditionnelle** de X sachant l'événement $\{Y=y\}$ est donnée par le fait que c'est une loi sur $X(\Omega)$ ainsi que par les probabilités conditionnelles $\mathbf{P}(X=x|Y=y)$ pour tout $x \in X(\Omega)$.

On vérifie aisément que

$$\sum_{x \in X(\Omega)} \mathbf{P}(X=x|Y=y) = 1$$

ce qui implique que la loi conditionnelle de X sachant $\{Y=y\}$ est la loi d'une variable aléatoire.

De plus, la formule des probabilités totales implique que

$$\mathbf{P}(X=x) = \sum_{y \in Y(\Omega)} \mathbf{P}(X=x|Y=y)\mathbf{P}(Y=y).$$

Cette définition de la loi conditionnelle s'étend à des vecteurs aléatoires : par exemple pour un triplet aléatoire (X, Y, Z), on peut étudier la loi conditionnelle de X sachant $\{Y=y\}$, la loi conditionnelle de X sachant $\{Y=y \text{ et } Z=z\}$, la loi conditionnelle du couple (X, Y) sachant $\{Z=z\}$...

Exemple 1 : La loi de Y sachant $\{X=1\}$ est donnée par $Y(\Omega) \in \mathbb{N}$ et, pour tout entier positif k, on a

$$\mathbf{P}(Y=k|X=1) = \frac{\mathbf{P}(Y=k \text{ et } X=1)}{\mathbf{P}(X=1)} = pe^{-\lambda}\frac{\lambda^k}{k!} \times \frac{1}{p} = e^{-\lambda}\frac{\lambda^k}{k!}.$$

La loi conditionnelle de Y sachant que $\{X=1\}$ est donc une loi de Poisson de paramètre λ.

Exemple 2 : La loi de X sachant $\{Y = 1\}$ est la loi uniforme sur $\{2, 3, 4\}$.

Exemple 3 : Supposons que (X, Y) suit une loi trinomiale (n, p_x, p_y) et calculons la loi conditionnelle de X sachant $\{Y = k\}$, pour un entier $k \in \{0, \dots, n\}$. Remarquons tout d'abord que si $j < 0$ ou $j > n-k$, $\mathbf{P}(X = j|Y = k) = 0$. Fixons maintenant un entier $j \in \{0, \dots, n-k\}$. On a

$$\begin{aligned}
\mathbf{P}(X = j|Y = k) &= \frac{\mathbf{P}(X = j \text{ et } Y = k)}{\mathbf{P}(Y = k)} \\
&= \frac{n!\, p_x^j p_y^k (1 - p_x - p_y)^{n-j-k}}{j!k!(n-j-k)!} \frac{k!(n-k)!}{n!\, p_y^k (1 - p_y)^{n-k}} \\
&= \frac{(n-k)!}{j!(n-k-j)!} \left(\frac{p_x}{1 - p_y}\right)^j \left(\frac{1 - p_x - p_y}{1 - p_y}\right)^{n-k-j}.
\end{aligned}$$

La loi conditionnelle de X sachant $\{Y = k\}$ est donc une loi binomiale $(n - k, p_x/(1 - p_y))$.

1.5 Indépendance de variables aléatoires discrètes

Définition 1.4 *Deux variables aléatoires discrètes X et Y sont dites* **indépendantes** *si pour tout $x \in X(\Omega)$ et tout $y \in Y(\Omega)$, les événements $\{X = x\}$ et $\{Y = y\}$ sont indépendants, c'est-à-dire :*

$$\mathbf{P}(X = x, Y = y) = \mathbf{P}(X = x)\mathbf{P}(Y = y).$$

Si X et Y sont deux variables aléatoires discrètes indépendantes, on aura donc, pour tout $y \in Y(\Omega)$ et tout $x \in X(\Omega)$ tels que $\mathbf{P}(Y = y) > 0$ et $\mathbf{P}(X = x) > 0$, $\mathbf{P}(X = x|Y = y) = \mathbf{P}(X = x)$ et $\mathbf{P}(Y = y|X = x) = \mathbf{P}(Y = y)$.

Plus généralement, les n variables aléatoires discrètes $X_1, \dots, X_n$ sont **(mutuellement ou n à n) indépendantes** si, pour tout choix de $x_1 \in X_1(\Omega), \dots, x_n \in X_n(\Omega)$, on a

$$\mathbf{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbf{P}(X_1 = x_1) \dots \mathbf{P}(X_n = x_n).$$

Remarques :

- L'indépendance n à n entraîne l'indépendance 2 à 2 (mais la réciproque est fausse). Écrire la preuve de ce résultat pour $n = 3$.
- Des événements $A_1, \dots, A_n$ sont indépendants si et seulement si les variables aléatoires $\mathbf{1}_{A_1}, \dots, \mathbf{1}_{A_n}$ le sont.

Proposition 1.5 *Si les n variables aléatoires discrètes $X_1, \dots, X_n$ sont indépendantes alors on a $\overline{X}(\Omega) = X_1(\Omega) \times \cdots \times X_n(\Omega)$.*

Remarque : Lorsque les variables aléatoires X et Y sont indépendantes, connaître les lois marginales permet donc de connaître la loi jointe du couple (X, Y), alors que pour des variables aléatoires quelconques, cela ne suffit pas.

1.6 Espérance, matrice de covariance

Dans un souci de clarté, tous les résultats de ce paragraphe et du suivant sont énoncés pour les couples de variables aléatoires discrètes, et ils s'étendent sans peine aux vecteurs aléatoires discrets.

On considère un couple aléatoire discret (X, Y).

Définition 1.6 • *L'***espérance du couple** (X, Y) *est définie si* X *et* Y *sont intégrables et on a alors :* $\mathbf{E}(X, Y) = (\mathbf{E}(X), \mathbf{E}(Y))$.

• *Si* X *et* Y *sont deux variables aléatoires de carré intégrable, la covariance de* X *et de* Y*, ou covariance du couple* (X, Y)*, est donnée par*

$$\mathbf{cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y) = \mathbf{E}[(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))].$$

• *Si* X *et* Y *sont deux variables aléatoires de carré intégrable, la matrice de covariance du couple* (X, Y) *est la matrice*

$$C = \begin{pmatrix} \mathbf{var}\,(X) & \mathbf{cov}(X, Y) \\ \mathbf{cov}(X, Y) & \mathbf{var}\,(Y) \end{pmatrix}.$$

Plus généralement, la matrice de covariance d'un vecteur $(X_1, \ldots, X_n)$*, dont chacune des composantes est de carré intégrable, est une matrice* $n \times n$ *dont les termes diagonaux sont les variances des* X_i *et dont le terme* (i, j) *est la covariance* $\mathbf{cov}(X_i, X_j)$ *pour tout* $i \neq j$.

Remarque : Le calcul de l'espérance de X ou de Y ne fait intervenir que les lois marginales, mais nous allons voir qu'il n'est pas nécessaire d'expliciter ces lois marginales.

Proposition 1.7 *Si* (X, Y) *est un couple de variables aléatoires discrètes, pour toute fonction* $h : \mathbb{R}^2 \to \mathbb{R}$ *telle que*

$$\sum_{x \in X(\Omega), y \in Y(\Omega)} |h(x, y)|\mathbf{P}(X = x, Y = y) < \infty,$$

la variable aléatoire $h(X, Y)$ *est intégrable et on a*

$$\mathbf{E}(h(X, Y)) = \sum_{x \in X(\Omega), y \in Y(\Omega)} h(x, y)\mathbf{P}(X = x, Y = y).$$

Preuve : Notons $Z = h(X, Y)$. On a une écriture explicite de la loi de Z : cette variable aléatoire est discrète et on a, pour tout $z \in (h(X, Y))(\Omega)$,

$$\mathbf{P}(Z = z) = \sum_{(x,y) \in (X,Y)(\Omega), h(x,y) = z} \mathbf{P}(X = x, Y = y)$$

La variable aléatoire Z est donc intégrable si la somme

$$S = \sum_{z \in (h(X,Y))(\Omega)} \left(|z| \sum_{(x,y) \in (X,Y)(\Omega), h(x,y) = z} \mathbf{P}(X = x, Y = y) \right)$$

converge. Or cette somme peut s'écrire

$$\begin{aligned} S &= \sum_{z\in(h(X,Y))(\Omega)} \left(\sum_{(x,y)\in(X,Y)(\Omega),h(x,y)=z} |h(x,y)|\mathbf{P}(X=x,Y=y) \right) \\ &= \sum_{(x,y)\in(X,Y)(\Omega)} |h(x,y)|\mathbf{P}(X=x,Y=y) \end{aligned}$$

La même procédure appliquée à la somme sans les valeurs absolues permet d'obtenir le résultat désiré pour $\mathbf{E}(h(X,Y))$. □

Application : Cette proposition permet d'écrire notamment les espérances de X, de Y ou de XY sans expliciter la loi de ces variables aléatoires. Si le couple (X,Y) est discret, on a ainsi

$$\begin{aligned} \mathbf{E}(X) &= \sum_{(x,y)\in(X,Y)(\Omega)} x\mathbf{P}(X=x,Y=y) \\ \mathbf{E}(Y) &= \sum_{(x,y)\in(X,Y)(\Omega)} y\mathbf{P}(X=x,Y=y) \\ \mathbf{E}(XY) &= \sum_{(x,y)\in(X,Y)(\Omega)} xy\mathbf{P}(X=x,Y=y) \end{aligned}$$

lorsque les séries convergent.

Revenons maintenant à la matrice de covariance :

Proposition 1.8

1. *Une matrice de covariance C est toujours une matrice symétrique et positive (i.e., pour tout $v \in \mathbb{R}^n$, $< v, Cv > \geq 0$).*
2. *Si X et Y sont deux variables aléatoires indépendantes et intégrables, on a $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$ et donc $\mathbf{cov}(X,Y) = 0$. La réciproque de ce résultat est fausse.*
3. *Si X et Y sont deux variables aléatoires indépendantes et f et g deux fonctions telles que les variables aléatoires $f(X)$ et $g(Y)$ sont intégrables, on a $\mathbf{E}(f(X)g(Y)) = \mathbf{E}(f(X))\mathbf{E}(g(Y))$. La réciproque de ce résultat est fausse.*
4. *Si les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes et de carré intégrable, alors la matrice de covariance de $(X_1, \ldots, X_n)$ est diagonale. La réciproque de ce résultat est fausse.*

Preuve : **1.** Soit $\overline{X} = (X_1, \ldots, X_n)$ un vecteur aléatoire dont chaque composante est de carré intégrable, notons C sa matrice de covariance et fixons $\overline{v} = (v_1, \ldots, v_n)$ un vecteur de $\mathbb{R}^n$. Quitte à remplacer chaque X_i par $X_i - \mathbf{E}(X_i)$, on supposera les variables aléatoires centrées (d'espérance nulle).

On a

$$\begin{aligned}
< v, Cv > &= \sum_{i,j} C_{ij} v_i v_j \\
&= \sum_{i,j} \mathbf{E}(X_i X_j) v_i v_j \\
&= \mathbf{E}\left(\sum_{i,j} v_i X_i v_j X_j \right) \\
&= \mathbf{E}\left(\left(\sum_i v_i X_i \right) \times \left(\sum_j v_j X_j \right) \right) \\
&= \mathbf{E}\left(\left(\sum_i v_i X_i \right)^2 \right)
\end{aligned}$$

On a donc bien $< v, Cv > \geq 0$.

2. Considérons deux variables aléatoires discrètes X et Y indépendantes et intégrables. Montrons que XY est intégrable et calculons son espérance. On a

$$\mathbf{E}(|XY|) = \sum_{x \in X(\Omega), y \in Y(\Omega)} |xy| \mathbf{P}(X = x, Y = y)$$

Comme X et Y sont indépendantes, $\mathbf{P}(X = x, Y = y) = \mathbf{P}(X = x)\mathbf{P}(Y = y)$ donc

$$\begin{aligned}
\mathbf{E}(|XY|) &= \sum_{x \in X(\Omega), y \in Y(\Omega)} |xy| \mathbf{P}(X = x) \mathbf{P}(Y = y) \\
&= \sum_{x \in X(\Omega)} \left(|x| \mathbf{P}(X = x) \sum_{y \in Y(\Omega)} |y| \mathbf{P}(Y = y) \right) \\
&= \left(\sum_{x \in X(\Omega)} |x| \mathbf{P}(X = x) \right) \left(\sum_{y \in Y(\Omega)} |y| \mathbf{P}(Y = y) \right) \\
&= \mathbf{E}(|X|) \mathbf{E}(|Y|).
\end{aligned}$$

On montre alors par un calcul similaire que $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$, puis on en déduit que $\mathbf{cov}(X, Y) = 0$.

3. et **4.** se déduisent aisément de **2.** $\square$

La proposition suivante, bien que très simple à prouver, est fort utile :

Proposition 1.9 *Si X et Y sont deux variables aléatoires de carré intégrable, on a*

$$\mathbf{var}\,(X + Y) = \mathbf{var}\,(X) + \mathbf{var}\,(Y) + 2\mathbf{cov}(X, Y).$$

Si de plus, X et Y sont indépendantes, on a

$$\mathbf{var}\,(X + Y) = \mathbf{var}\,(X) + \mathbf{var}\,(Y).$$

De plus, pour deux variables aléatoires de carré intégrable, il est possible d'obtenir une majoration de $\mathbf{cov}(X,Y)$ à partir des variances de X et Y :

Proposition 1.10 (Inégalité de Cauchy Schwarz) *Soient X et Y deux variables aléatoires de carré intégrable. On a*

$$|\mathbf{E}(XY)| \leq \mathbf{E}(|XY|) \leq \sqrt{\mathbf{E}(X^2)\mathbf{E}(Y^2)}$$
$$|\mathbf{cov}(X,Y)| \leq \sqrt{\mathbf{var}\,(X)\mathbf{var}\,(Y)}$$

Preuve : Le deuxième point s'obtient en appliquant le premier à $X - \mathbf{E}(X)$ et $Y - \mathbf{E}(Y)$.

Le premier point se montre de la même façon que dans le cas classique (produit scalaire de deux vecteurs) : on étudie le polynôme

$$R(\lambda) = \mathbf{E}(X^2)\lambda^2 - 2\mathbf{E}|XY|\lambda + \mathbf{E}(Y^2).$$

Ce polynôme se factorise en

$$R(\lambda) = \mathbf{E}[(\lambda|X| - |Y|)^2].$$

Il n'admet donc pas deux racines réelles distinctes, ce qui signifie que son discriminant (réduit) est négatif :

$$(\mathbf{E}|XY|)^2 - \mathbf{E}(X^2)\mathbf{E}(Y^2) \leq 0.$$

On étudie de la même façon le cas d'égalité : le discriminant n'est nul que si le polynôme admet une racine réelle double. Dans ce cas, il existe donc λ_0 tel que

$$\mathbf{E}[(\lambda_0|X| + |Y|)^2] = 0;$$

l'espérance d'une variable aléatoire positive ne pouvant être nulle que si la variable aléatoire est (presque sûrement) nulle, on aura alors : $|Y| = -\lambda_0|X|$, presque sûrement.

Les preuves sans les valeurs absolues sont similaires. □

1.7 Fonction génératrice d'un couple

Définition 1.11 *La fonction génératrice d'un couple de variables aléatoires discrètes positives est la fonction définie sur $[0,1]^2$ par*

$$G_{(X,Y)}(s,t) = \mathbf{E}(s^X t^Y) = \sum_{x\in X(\Omega), y\in Y(\Omega)} s^x t^y \mathbf{P}(X = x, Y = y).$$

Proposition 1.12 *La fonction génératrice détermine la loi du couple (X,Y) au sens où si deux couples de variables aléatoires positives ont la même fonction génératrice, alors ils suivent la même loi.*

La fonction génératrice de (X,Y) permet de retrouver par exemple :
- la fonction génératrice de X : $G_X(s) = G_{(X,Y)}(s,1)$ et de Y,
- l'espérance de X si X est intégrable : $\mathbf{E}(X) = \frac{\partial}{\partial s}G_{(X,Y)}(1,1)$,
- l'espérance de X^2 si X est de carré intégrable : $E(X^2) = \frac{\partial^2}{\partial s^2}G_{(X,Y)}(1,1) + \frac{\partial}{\partial s}G_{(X,Y)}(1,1)$,

• l'espérance de XY si X et Y sont de carré intégrable : $\mathbf{E}(X,Y) = \dfrac{\partial^2}{\partial s\,\partial t} G_{(X,Y)}(1,1)$.

Une autre utilité importante de la fonction génératrice est de permettre de calculer simplement la loi de somme de variables aléatoires, à partir du moment où on connaît leur loi jointe. En effet, on a le résultat suivant :

Proposition 1.13 *Soient (X,Y) un couple de variables aléatoires positives dont on connaît la loi jointe. Notons $G_{X,Y}$ la fonction génératrice du couple. On a alors $G_{X+Y}(s) = G_{(X,Y)}(s,s)$.*

Preuve : On a $G_{(X,Y)}(s,s) = \mathbf{E}(s^X s^Y) = \mathbf{E}(s^{X+Y}) = G_{X+Y}(s)$.

Si X et Y sont à valeurs entières, il ne reste plus qu'à décomposer la fonction G_{X+Y} en série entière pour obtenir les probabilités $\mathbf{P}(X+Y=k)$. □

Il est facile de voir que si X et Y sont deux variables aléatoires positives indépendantes, alors on a $G_{(X,Y)}(s,t) = G_X(s)G_Y(t)$. La réciproque de ce résultat est également vraie :

Proposition 1.14 *Les variables aléatoires positives X et Y sont indépendantes si et seulement si*

$$G_{(X,Y)}(s,t) = G_X(s)G_Y(t).$$

Remarque : La fonction génératrice du vecteur aléatoire $\overline{X} = (X_1, \ldots, X_n)$ vérifiant $\overline{X}(\Omega) \subset \mathbb{N}^n$ est la fonction $G : [0,1]^n \to \mathbb{R}^+$ définie par

$$G(s_1, \ldots, s_n) = \mathbf{E}(s_1^{X_1} \times \ldots \times s_n^{X_n}).$$

Cette fonction de n variables caractérise la loi du n-uplet, permet de retrouver les lois marginales ...

1.8 Exemple d'utilisation de la fonction génératrice

Calculons la fonction génératrice de la loi trinomiale (n, p_x, p_y).
Soit (X,Y) un couple de variable aléatoire de loi trinomiale (n, p_x, p_y).
On a

$$\begin{aligned}
G_{(X,Y)}(s,t) &= \mathbf{E}(s^X t^Y) \\
&= \sum_{i,j \in \{0,\ldots,n\}, i+j \leq n} s^i t^j \mathbf{P}(X=i, Y=j) \\
&= \sum_{i=0}^{n} \sum_{j=0}^{n-i} s^i t^j \frac{n!}{i!j!(n-i-j)!} p_x^i p_y^j (1-p_x-p_y)^{n-i-j} \\
&= \sum_{i=0}^{n} \frac{n!}{i!(n-i)!} (sp_x)^i \sum_{j=0}^{n-i} \frac{(n-i)!}{j!(n-i-j)!} (tp_y)^j (1-p_x-p_y)^{n-i-j} \\
&= \sum_{i=0}^{n} \frac{n!}{i!(n-i)!} (sp_x)^i (1-p_x-(1-t)p_y)^{n-i} \\
&= (1-(1-s)p_x-(1-t)p_y)^n.
\end{aligned}$$

En prenant $t = 1$ dans le résultat ci-dessus, on trouve : $G_{(X,Y)}(s,1) = (1-(1-s)p_x)^n$; on reconnaît la fonction caractéristique d'une loi binomiale (n, p_x) ce qui prouve que X suit une loi binomiale (n, p_x).

La fonction génératrice de (X, Y) permet également de calculer celle de $X + Y$. On a en effet :

$$\mathbf{E}(s^{X+Y}) = \mathbf{E}(s^X s^Y) = G_{(X,Y)}(s,s).$$

On trouve ici : $G_{X+Y}(s) = (1-(1-s)(p_x+p_y))^n$; on retrouve donc que $X+Y$ suit une loi binomiale $(n, p_x + p_y)$.

1.9 Somme de deux variables aléatoires de loi de Poisson et indépendantes

Considérons deux variables aléatoires X et Y indépendantes et de loi de Poisson de paramètre respectivement $\mu_1 > 0$ et $\mu_2 > 0$, et calculons la loi de $X + Y$.

Première méthode

Commençons par calculer la fonction génératrice de la loi de X. On a

$$\begin{aligned} G_X(s) &= \mathbf{E}(s^X) \\ &= \sum_{n \geq 0} s^n \mathbf{P}(X = n) \\ &= \sum_{n \geq 0} s^n e^{-\mu_1} \frac{\mu_1^n}{n!} \\ &= e^{-\mu_1} e^{\mu_1 s} \\ &= e^{-\mu_1(1-s)} \end{aligned}$$

De la même façon, on a $G_Y(t) = e^{-\mu_2(1-t)}$. Comme X et Y sont des variables aléatoires indépendantes, on en déduit la fonction génératrice du couple (X, Y) :

$$G_{(X,Y)}(s,t) = e^{-\mu_1(1-s)} e^{-\mu_2(1-t)}.$$

La fonction génératrice de $X + Y$ est donc

$$G_{X+Y}(s) = G_{(X,Y)}(s,s) = e^{-(\mu_1+\mu_2)(1-s)}.$$

On reconnaît ici la fonction génératrice d'une loi de Poisson de paramètre $\mu_1 + \mu_2$ et on peut conclure que $X + Y$ suit une loi de Poisson de paramètre $\mu_1 + \mu_2$.

Attention : La variable aléatoire $X - Y$ ne suit pas une loi de Poisson : cette variable aléatoire prend ses valeurs dans $\mathbb{Z}$ et non dans $\mathbb{N}$ et, en conséquence, on ne peut pas calculer sa fonction génératrice.

Deuxième méthode

On a évidemment $(X+Y)(\Omega) \subset \mathbb{N}$. Fixons donc $n \in \mathbb{N}$ et calculons directement la probabilité $\mathbf{P}(X + Y = n)$.

On a

$$\begin{aligned}
\mathbf{P}(X+Y=n) &= \sum_{k\geq 0} \mathbf{P}(X=k, X+Y=n) \\
&= \sum_{k\geq 0} \mathbf{P}(X=k, Y=n-k) \\
&= \sum_{k\geq 0} \mathbf{P}(X=k)\mathbf{P}(Y=n-k) \\
&= \sum_{k=0}^{n} e^{-\mu_1}\frac{\mu_1^k}{k!}e^{-\mu_2}\frac{\mu_2^{n-k}}{(n-k)!} \\
&= e^{-(\mu_1+\mu_2)}\frac{1}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\mu_1^k\mu_2^{n-k} \\
&= e^{-(\mu_1+\mu_2)}\frac{(\mu_1+\mu_2)^n}{n!}
\end{aligned}$$

On retrouve donc que $X+Y$ suit une loi de Poisson de paramètre $\mu_1+\mu_2$.

2 Couples aléatoires à densité

2.1 Densité d'un couple

Définition 2.1 *Un couple (X,Y) de variables aléatoires réelles sera dit « à densité par rapport à la mesure de Lebesgue de $\mathbb{R}^2$ »s'il existe une fonction $f : \mathbb{R}^2 \to \mathbb{R}^+$ telle que, pour tous intervalles I et J et pour toute fonction continue bornée ou positive h on ait :*

$$\mathbf{P}(X \in I \text{ et } Y \in J) = \int_{I\times J} f(x,y)\, dx\, dy \qquad \text{et} \qquad \mathbf{E}(h(X,Y)) = \iint_{\mathbb{R}^2} h(x,y)f(x,y)\, dx\, dy.$$

Remarque :
- La densité f est toujours une fonction positive, intégrable et dont l'intégrale sur $\mathbb{R}^2$ par rapport à la mesure de Lebesgue vaut 1.
- On peut montrer que la définition de la densité est équivalente à : pour tout ouvert U de $\mathbb{R}^2$

$$\mathbf{P}((X,Y) \in U) = \iint_U f(x,y)\, dx\, dy.$$

Exemples :

1. Soit $D = \{(x,y) \in \mathbb{R}^2, x^2+y^2 \leq 1\}$ le disque unité de $\mathbb{R}^2$ et f la fonction définie sur $\mathbb{R}^2$ par : $f(x,y) = \mathbf{1}_D(x,y)/\pi$. f est bien une densité de probabilité sur $\mathbb{R}^2$: c'est la densité de la loi uniforme sur D, c'est-à-dire la loi que l'on obtient en jetant un point au hasard et uniformément sur D.
2. Soit X une variable aléatoire de loi exponentielle. La loi du couple (X,X) n'admet pas de densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$. En effet, si une telle densité f

existait, elle devrait être (presque partout) nulle en dehors de la droite $\{x = y\}$ puisque $\mathbf{P}(X \neq Y) = 0$. On aurait alors

$$1 = \int_{\mathbb{R}^2} f(x,y)\ dx\,dy = \int_{\{x=y\}\subset\mathbb{R}^2} f(x,y)\ dx\,dy.$$

Or une droite est de mesure nulle pour la mesure de Lebesgue sur $\mathbb{R}^2$, donc l'intégrale, toujours par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$, de n'importe quelle fonction sur ce domaine est nulle. On aboutit donc à une contradiction.

2.2 Loi marginale

Les lois marginales du couple (X, Y) sont les lois des variables aléatoires X et Y. On peut vérifier aisément la proposition suivante :

Proposition 2.2 *Soit (X, Y) un couple de variables aléatoires de densité (conjointe) $f : \mathbb{R}^2 \to \mathbb{R}^+$. Alors les deux variables aléatoires X et Y ont chacune une densité, notée respectivement f_X et f_Y données par :*

$$f_X(x) = \int_{y\in\mathbb{R}} f(x,y)\ dy \qquad et \qquad f_Y(y) = \int_{x\in\mathbb{R}} f(x,y)\ dx.$$

Preuve : En effet, soit $t \in \mathbb{R}$ et calculons la fonction de répartition de X. On a

$$\begin{aligned}\mathbf{P}(X \le t) &= \mathbf{P}(X \in]-\infty, t] \text{ et } Y \in \mathbb{R}) \\ &= \int_{]-\infty,t]\times\mathbb{R}} f(x,y)\ dx\,dy \\ &= \int_{-\infty}^{t} \left(\int_{y\in\mathbb{R}} f(x,y)dy\right)\ dx\end{aligned}$$

□

Ce principe s'applique aux vecteurs à densité, à ceci près que les lois marginales d'un triplet sont les lois de chacune des variables qui le constituent ainsi que les lois des couples de variables.

Exercice : Calculer les lois marginales du couple (X, Y) de loi uniforme sur le disque D de centre 0 et de rayon 1.

2.3 Indépendance

Rappelons la définition de l'indépendance de variables aléatoires :

Définition 2.3 *Les variables aléatoires X et Y sont* **indépendantes** *si, pour tous intervalles I et J, on a*

$$\mathbf{P}(X \in I \text{ et } Y \in J) = \mathbf{P}(X \in I)\mathbf{P}(Y \in J).$$

Plus généralement, n variables aléatoires $X_1, \dots, X_n$ sont **(mutuellement) indépendantes** *si, pour tous intervalles $I_1, \dots, I_n$, on a*

$$\mathbf{P}\left(\bigcap_{k\le n}\{X_k \in I_k\}\right) = \prod_{k\le n}\mathbf{P}\{X_k \in I_k\}.$$

Le lien entre densités et indépendance est clair :

Proposition 2.4 *Deux variables aléatoires* (X, Y) *de densité respectivement* f *et* g *sont indépendantes si et seulement si la loi du couple admet une densité et que cette densité est la fonction* $(x, y) \to f(x)g(y)$.

2.4 Espérance, covariance

Définition 2.5 *On définit, comme pour les couples discrets l'espérance d'un couple de variables aléatoires intégrables comme étant le couple des espérances* $\mathbf{E}(X, Y) = (\mathbf{E}(X), \mathbf{E}(Y))$, *et il est facile de vérifier que*

$$\mathbf{E}(X) = \int_{\mathbb{R}^2} xf(x, y)\ dx\ dy \qquad et \qquad \mathbf{E}(Y) = \int_{\mathbb{R}^2} yf(x, y)\ dx\ dy,$$

lorsque ces intégrales sont absolument convergentes.

Si les variables aléatoires sont de carré intégrable, on définit également la matrice de covariance du couple (X, Y) *par*

$$C = \begin{pmatrix} \mathbf{var}\,(X) & \mathbf{cov}(X, Y) \\ \mathbf{cov}(X, Y) & \mathbf{var}\,(Y) \end{pmatrix}$$

Les termes non-diagonaux de cette matrice seront

$$\mathbf{cov}(X, Y) = \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y))) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$$

où

$$\mathbf{E}(XY) = \int_{\mathbb{R}^2} xyf(x, y)\ dx\,dy.$$

La matrice de covariance est, comme dans le cas discret, une matrice symétrique et positive (au sens des formes bilinéaires).

Ces définitions s'étendent naturellement au cas des vecteurs aléatoires à densité, en remplaçant les intégrales doubles par des intégrales multiples.

Remarque : Pour identifier la densité d'un couple, on utilise habituellement une fonction test $h : \mathbb{R}^2 \to \mathbb{R}$ continue et bornée et on essaie d'écrire $\mathbf{E}(h(X, Y))$ sous la forme

$$\mathbf{E}(h(X, Y))) = \int_{\mathbb{R}^2} h(x, y)f(x, y)\ dx, dy.$$

La fonction $f : \mathbb{R}^2 \to \mathbb{R}^+$ ci-dessus, si elle existe, sera la densité du couple (X, Y).

Les propriétés vues dans le cas discret restent vraies pour les couples à densité : notamment, si le couple est formé de variables aléatoires indépendantes et de carré intégrable, $\mathbf{cov}(X, Y) = 0$ et $\mathbf{var}\,(X + Y) = \mathbf{var}\,(X) + \mathbf{var}\,(Y)$, et, tout comme dans le cas discret, la covariance de deux variables aléatoires dont le couple admet une densité, peut être nulle sans que les variables soient indépendantes : reprendre par exemple l'exemple du couple de loi uniforme sur le disque de centre 0 et de rayon 1.

2.5 Loi de la somme

On peut vérifier que si le couple (X,Y) a une densité, la variable $X+Y$ aura également une densité, et on peut expliciter cette densité :

Proposition 2.6 – *Si le couple (X,Y) admet pour densité la fonction f, alors la densité de la variable aléatoire $X+Y$ est la fonction g définie par*

$$g(z) = \int_{\mathbb{R}} f(x, z-x)\ dx = \int_{\mathbb{R}} f(z-y, y)\ dy.$$

– *Si X et Y sont deux variables aléatoires indépendantes de densité f_X et f_Y, la densité, notée g, de $X+Y$ est le produit de convolution de f_X et f_Y :*

$$g(z) = \int_{\mathbb{R}} f_X(x) f_Y(z-x)\ dx = \int_{\mathbb{R}} f_X(z-y) f_Y(y)\ dy.$$

Preuve : Soit $h : \mathbb{R} \to \mathbb{R}$ une fonction continue par morceaux et bornée. On a

$$\begin{aligned} \mathbf{E}h(X+Y) &= \int_{\mathbb{R}^2} h(x+y) f(x,y)\ dx, dy \\ &= \int_{\mathbb{R}} h(z) \int_{\mathbb{R}} f(x, z-x)\ dx\ dz \end{aligned}$$

en posant $x = x$ et $z = x+y$. Si on avait effectué le changement de variables $z = x+y$ et y, on aurait obtenu l'autre expression annoncée. $\square$