Chapitre 1

Généralités sur les vecteurs aléatoires

1.1 Introduction

La notion de vecteurs aléatoires intervient lorsqu'on s'intéresse à mesurer différentes grandeurs numériques dépendants de l'issue d'une même expérience aléatoire. Par exemple, dans l'observation d'une population quelconque on peut s'intéresser à mesurer l'âge, la taille, le poids... de chaque individu. On forme alors un vecteur d'observations aléatoires. Si on surveille une particule M évoluant dans le plan ou dans l'espace, on peut s'intéresser à sa position en fonction du temps déterminée par le vecteur de ses coordonnées aléatoires.

1.2 Rappels sur les variables aléatoires unidimensionnelles

L'espace $(\Omega, \mathcal{A}, \mathbb{P})$

On désigne par Ω l'ensemble des épreuves ou événements élémentaires w et par $\mathcal{A}$ une tribu sur Ω, c'est-à-dire une classe de parties de Ω qui vérifie les axiomes suivantes :

- (A_1) : $\varnothing \in \mathcal{A}$ et $\Omega \in \mathcal{A}$;
- (A_2) : $\mathcal{A}$ est stable par passage au complémentaire : $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$;
- (A_3) : $\mathcal{A}$ est stable par réunion et intersection dénombrables i.e. Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{A}$, alors $\bigcup_{n\in\mathbb{N}} A_n$ et $\bigcap_{n\in\mathbb{N}} A_n$ sont dans $\mathcal{A}$.

Définition 1.1. *L'ensemble des parties de Ω, $\mathcal{P}(\Omega)$ est un exemple de tribu sur Ω. Un élément A d'une tribu $\mathcal{A}$ sera appelé événement aléatoire .*

Définition 1.2. *On appelle probabilité définie sur l'espace $(\Omega, \mathcal{A})$ toute application $\mathbb{P} : \mathcal{A} \to [0, 1]$ satisfaisant les 3 axiomes suivantes :*

(i) $0 \leq \mathbb{P}(A) \leq 1, \forall A \in \mathcal{A}$.

(ii) $\mathbb{P}(\Omega) = 1$.

(iii) *Pour toute suite* $(A_i)_{i\in\mathbb{N}^*}$ *d'événements de* $\mathcal{A}$ *deux à deux disjoints (incompatibles) :*

$$\mathbb{P}\left(\bigcup_{i\in\mathbb{N}^*} A_i\right) = \sum_{i\in\mathbb{N}^*} \mathbb{P}(A_i)$$

Proposition 1.2.1. ***(Propriétés générales d'une probabilité)****. Toute probabilité* $\mathbb{P}$ *sur* $(\Omega, \mathcal{A})$ *vérifie les proprirétés suivantes :*

1. $\mathbb{P}(\varnothing) = 0$.
2. $\forall A \in \mathcal{A}, \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.
3. $\forall A \in \mathcal{A}, \forall B \in \mathcal{A}, A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$.
4. $\forall A \in \mathcal{A}, \forall B \in \mathcal{A}, \mathbb{P}\left(A \cup B\right) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Définition 1.3. *Le couple* $(\Omega, \mathcal{A})$ *est appelé espace probabilisable.*

Définition 1.4. *Le triplet* $(\Omega, \mathcal{A}, \mathbb{P})$ *est appelé espace probabilisé.*

Définition 1.5. *On appelle espace mesurable tout couple* (E, ε) *formé par un ensemble* E *et une tribu* ε *sur* E.

Définition 1.6. *Soit* $\mathcal{C}$ *un sous ensemble de* $\mathbb{P}(E)$*. La tribu engendrée par* $\mathcal{C}$ *est la plus petite tribu contenant* $\mathcal{C}$*. Cette tribu peut également être définie comme l'intersection de toutes les tribus contenant* $\mathcal{C}$*. On la note* $\sigma(\mathcal{C})$*.*

Définition 1.7. *Si* $E = \mathbb{R}$*, on appelle tribu borélienne* $\mathcal{B}(\mathbb{R})$ *la tribu* ε *engendrée par la classe des intervalles de* $\mathbb{R}$ *de la forme* $]-\infty, a[$ *pour* $a \in \mathbb{R}$.

Définition 1.8. *Soient* $(\Omega, \mathcal{A})$ *et* (E, ε) *deux espaces masurables. On dit que* f *est une application mesurable de* $(\Omega, \mathcal{A})$ *dans* (E, ε) *si l'image réciproque par* f *de tout élément de* ε *est un élément de* $\mathcal{A}$*. Autrement dit,* f *est mesurable si* $f^{-1}(\varepsilon) \subset \mathcal{A}$.

Définition 1.9. *Soit* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace de probabilité et* (E, ε) *un espace mesurable. Une variable aléatoire* X *est une application mesurable de* $(\Omega, \mathcal{A})$ *dans* (E, ε) *c'est -à- dire une application* $X : \Omega \to E$ *telle que pour tout borélien* B *de* ε *l'ensemble*

$$X^{-1}(B) = \{w \in \Omega : X(w) \in B\} \subset \mathcal{A}.$$

On distingue les cas suivants :

a. Si E est dénombrable, la variable aléatoire X est dite discrète. Dans ce cas, E est muni de la tribu $\mathcal{P}(E)$ et X vérifie

$$\forall x \in E : \{\omega \in \Omega : X = x\} \in \mathcal{A}.$$

b. Si $E = \mathbb{R}$, la variable aléatoire X est dite variable aléatoire réelle continue. En général, $\mathbb{R}$ est muni de sa tribu de Borel $\varepsilon = \mathfrak{B}(\mathbb{R})$, dans ce cas X est une variable aléatoire réelle continue de $(\Omega, \mathcal{A}, \mathbb{P})$ dans $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$.

c. Si $E = \mathbb{R}^n$, on parle de variables aléatoires réelles multidimensionnelles ou vectorielles ou encore de vecteurs aléatoires réels. En général, $\mathbb{R}^n$ est muni de sa tribu de Borel $\varepsilon = \mathfrak{B}(\mathbb{R}^n)$, dans ce cas X est un vecteur aléatoire de $(\Omega, \mathcal{A}, \mathbb{P})$ dans $(\mathbb{R}, \mathfrak{B}(\mathbb{R}^n))$. On utilisera les abréviations (v.a), (v.a.r), (v.a.d) pour désigner une variable aléatoire, une variable aléatoire réelle et une variable aléatoire discrète respectivement.

Définition 1.10. *Si X est une v.a.r et $\mathbb{P}_X : \mathfrak{B}(\mathbb{R}) \to [0,1]$ définie par $\mathbb{P}_X(B) = \mathbb{P}\left(X^{-1}(B)\right)$ une mesure de probabilité sur $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$. Cette mesure de probabilité est appelée loi de probabilité de X. On dit également que X suit la loi de probabilité $\mathbb{P}_X$ et on note*

$$X \sim \mathbb{P}_X.$$

Définition 1.11. *la tribu borélienne $\mathfrak{B}(\mathbb{R}^n)$ sur $\mathbb{R}^n$ est la tribu engendrée par les ensembles de la forme*

$$\prod_{i=1}^{n} (]-\infty, x_i[) \; où x_i \in \mathbb{R}, \forall i = 1, ..., n.$$

1.3 Couples de variables aléatoires réelles

Définition 1.12. *Un couple $Z = (X, Y)$ de variables aléatoires réelles est une application mesurable Z définie par :*

$$\begin{aligned} Z : \Omega &\to \mathbb{R}^2 \\ w &\longmapsto Z(w) = (X(w), Y(w)). \end{aligned}$$

Proposition 1.3.1. **(i)** *Etant donné un couple (X, Y) de variables aléatoires de $(\mathbb{R}^2, \mathfrak{B}(\mathbb{R}^2))$ et $B \in \mathfrak{B}(\mathbb{R}^2)$, l'ensemble $\{w \in \Omega \diagup (X(w), Y(w)) \in B\}$ est un événement.*

(ii) *Inversement, soit une application $(X, Y) : \Omega \to \mathbb{R}^2$ telle que :*

$$\forall B \in \mathfrak{B}(\mathbb{R}^2), \{w \in \Omega \diagup (X(w), Y(w)) \in B\} \in \mathcal{A}.$$

Alors (X, Y) est un couples de v.a de $(\mathbb{R}^2, \mathfrak{B}(\mathbb{R}^2))$ X et Y sont des variables aléatoires réelles.

Notation

Etant donné un couple (X, Y) de variables aléatoires de $(\mathbb{R}^2, \mathfrak{B}(\mathbb{R}^2))$, on note pour tout

$$B \in \mathfrak{B}(\mathbb{R}^2), ((X, Y) \in B) = \{w \in \Omega \diagup (X(w), Y(w)) \in B\}.$$

1.3.1 Probabilité image

Théorème 1.3.2. *Soient* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé et* (X, Y) *un couple de variables aléatoires réelles* Ω *. Alors l'application* $\mathbb{P}_{XY} : \mathfrak{B}(\mathbb{R}^2) \to [0,1]$ *définie par :*

$$\forall B \in \mathfrak{B}(\mathbb{R}^2), \mathbb{P}_{XY}(B) = \mathbb{P}\left((X,Y) \in B\right)$$

est une probabilité sur $(\mathbb{R}^2, \mathfrak{B}(\mathbb{R}^2)$ *appelée probabilité image de* $\mathbb{P}$ *par* (X, Y)*.*

Définition 1.13. *Soit un espace probabilisé* $(\Omega, \mathcal{A}, \mathbb{P})$*. On dit que deux applications de* Ω *dans* $\mathbb{R}$ *forment un couple de variables aléatoires réelles de* $\mathfrak{B}(\mathbb{R}^2)$ *si, pour tout couple* (x, y) *de réels ,* $[X \leq x]$ *et* $[Y \leq y]$ *sont des éléments de* $\mathcal{A}$*.*

Définition 1.14. *Soient* (X_1, X_2) *et* (Y_1, Y_2) *deux couples de variables aléatores réelles de* $(\mathbb{R}^2, \mathfrak{B}(\mathbb{R}^2))$*. On dit que* (X_1, X_2) *et* (Y_1, Y_2) *sont équidistribués lorsque*

$$\mathbb{P}_{X_1X_2}(B) = \mathbb{P}_{Y_1Y_2}(B), \forall B \in B(\mathbb{R}^2).$$

1.3.2 Fonction de répartition d'un couple de variables aléatoires

Notations

Soit (X, Y) un couple de variables aléatoires réelles de $(\mathbb{R}^2, \mathfrak{B}(\mathbb{R}^2))$. Etant donnés B et C dans $\mathfrak{B}(\mathbb{R}), B \times C$ est un borélien de $\mathbb{R}^2$. On note alors :

$$\mathbb{P}\left((X,Y) \in B \times C)\right) = \mathbb{P}\left(X \in B, Y \in C\right) = \mathbb{P}\left((X \in B) \cap (Y \in C)\right).$$

$$\mathbb{P}\left(X \leq x, Y \leq y\right) = \mathbb{P}\left(X \in]-\infty, x], Y \in]-\infty, y]\right).$$

$$\mathbb{P}\left(X < x, Y < y\right) = \mathbb{P}\left(X \in]-\infty, x[, Y \in]-\infty, y[\right).$$

Définition 1.15. *Soit* (X, Y) *un couple de variables aléatoires de* $(\mathbb{R}^2, \mathfrak{B}(\mathbb{R}^2))$*. On appelle fonction de répartition du couple* (X, Y)*, l'application*

$$F_{X,Y} : \mathbb{R}^2 \to [0,1]$$

définie par :

$$\forall (x,y) \in \mathbb{R}^2, F_{X,Y}(x,y) = P\left([X \leq x] \cap [Y \leq y]\right) = \mathbb{P}\left(X \leq x, Y \leq y\right).$$

On dit aussi que $F_{X,Y}$ est la loi du couple (X, Y).

Théorème 1.3.3. Soit (X, Y) un couple de variables aléatoires réelles sur l'espace $(\Omega, \mathcal{A}, \mathbb{P})$. Alors :

(i) Pour tout $(a, b) \in \mathbb{R}^2$ et tout $(h, k) \in \mathbb{R}_+^2$, on a

$$\mathbb{P}(a < X \leq a+h, b < Y \leq b+k) = F_{XY}(a+h, b+k) - F_{XY}(a, b+k) - F_{XY}(a+h, b) + F_{XY}(a, b).$$

(ii) $F_{X,Y}$ *est croissante sur* $\mathbb{R}$, et continue à droite par rapport à chacune de ses variables.

(iii) *Pour tout y réel fixé,*

$$\lim_{x\to-\infty} F_{X,Y}(x,y)=0,\ \lim_{x\to+\infty} F_{X,Y}(x,y)=F_Y(y).$$

Pour tout x réel fixé,

$$\lim_{x\to-\infty} F_{X,Y}(x,y)=0,\ \lim_{y\to+\infty} F_{X,Y}(x,y)=F_X(x).$$

(iv)

$$\lim_{x\to+\infty}(\lim_{y\to+\infty} F_{X,Y}(x,y))=\lim_{y\to+\infty}(\lim_{x\to+\infty} F_{X,Y}(x,y))=1.$$

Démonstration. (i) Pour tout $(a,b)\in\mathbb{R}^2$ et tout $(h,k)\in\mathbb{R}_+^2$, on a

$$\mathbb{P}(a<X\leq a+h, b<Y\leq b+k)=\mathbb{P}\left((X,Y)\in]a,a+h]\times]b,b+k]\right)=\mathbb{P}_{X,Y}\left(]a,a+h]\times]b,b+k]\right).$$

Par suite :

$$\begin{aligned}\mathbb{P}_{X,Y}\left(]a,a+h]\times]b,b+k]\right) &= \mathbb{P}_{X,Y}\left(]-\infty,a+h]\times]-\infty,b+k]\right)-\mathbb{P}_{X,Y}\left(]-\infty,a+h]\times]-\infty,b]\right)\\ &\quad -\mathbb{P}_{X,Y}\left(]-\infty,a]\times]-\infty,b+k]\right)+\mathbb{P}_{X,Y}\left(]-\infty,a]\times]-\infty,b]\right)\\ &= F_{XY}(a+h,b+k)-F_{XY}(a,b+k)-F_{XY}(a+h,b)+F_{XY}(a,b).\end{aligned}$$

(ii) Fixons par exemple y et soit a et x deux réels tels que $a\leq x$. $[X\leq a]\subset[X\leq x]$ *d'où,*

$$[X\leq a]\cap[Y\leq y]\subset[X\leq x]\cap[Y\leq y].$$

On en déduit que

$$\mathbb{P}\left[(X\leq a)\cap(Y\leq y)\right]\leq\mathbb{P}\left[(X\leq x)\cap(Y\leq y)\right].$$

C'est-à-dire :

$$F_{X,Y}(a,y)\leq F_{X,Y}(x,y).$$

D'autre part,

$$\begin{aligned}F_{X,Y}(x,y)-F_{X,Y}(a,y) &= \mathbb{P}\left[(X\leq x))\cap(Y\leq y)\right]-\mathbb{P}\left[(X\leq a))\cap(Y\leq y)\right]\\ &= \mathbb{P}\left[((X\leq x)-(X\leq a))\cap(Y\leq y)\right]\leq\mathbb{P}\left[(X\leq x)-(X\leq a)\right].\end{aligned}$$

Or,

$$\mathbb{P}\left[(X\leq x)-(X\leq a)\right]=\mathbb{P}\left(a<X\leq x\right)=F_X(x)-F_X(a).$$

Ainsi, $0\leq F_{X,Y}(x,y)-F_{X,Y}(a,y)\leq F_X(x)-F_X(a)$.
Par conséquent,

$$\lim_{x\to a^+}\left(F_{X,Y}(x,y)-F_{X,Y}(a,y)\right)=0.$$

$F_{X,Y}$ est continue à droite par rapport à x, pour y fixé.
On raisonne de même pour y.

(iii) Soit y réel fixé,

$$\mathbb{P}\left[(X \leq x) \cap (Y \leq y)\right] \leq \mathbb{P}(X \leq x)$$

d'où : $F_{X,Y}(x,y)) \leq F_X(x)$ *et par suite,*

$$\lim_{x \to -\infty} F_{X,Y}(x,y) = 0.$$

$$\begin{aligned} F_Y(y) - F_{X,Y}(x,y) &= \mathbb{P}(Y \leq y) - \mathbb{P}[(X \leq x) \cap (Y \leq y)] \\ &= \mathbb{P}[(Y \leq y) - (X \leq x)] \\ &= \mathbb{P}[(Y \leq y) \cap (X > x)] \leq \mathbb{P}(X > x). \end{aligned}$$

Autrement dit, $0 \leq F_Y(y) - F_{X,Y}(x,y) \leq 1 - F_X(x)$.
Par suite,

$$\lim_{x \to +\infty} [F_Y(y) - F_{X,Y}(x,y)] = 0.$$

(iv) Résulte immédiatement de (iii) par passage à la limite.

1.3.3 Indépendance de variables aléatoires

Soit (X,Y) un couple de variables aléatoires. On dit que X et Y sont indépendantes lorsque pour tout couple (A,B) de Boréliens de $\mathfrak{B}(\mathbb{R}^2)$, les événements $[X \in A]$ et $[Y \in B]$ sont indépendants. Autrement dit, les trois propriètès suivantes sont équivalentes.

Propriété 1.3.4. *1. X et Y indépendantes.*

2. $\forall (A,B) \in \mathfrak{B}(\mathbb{R}^2), \mathbb{P}\left(X \in A, Y \in B\right) = \mathbb{P}\left(X \in A\right)\mathbb{P}\left(Y \in B\right)$.
3. $\forall (A,B) \in \mathfrak{B}(\mathbb{R}^2), \mathbb{P}_{XY}\left(A \times B\right) = \mathbb{P}_X\left(A\right)\mathbb{P}_Y\left(B\right)$.

Proposition 1.3.5. *Soient X et Y deux variables aléatoires réelles. Alors X et Y sont indépendantes si et seulement si :*

$$F_{X,Y}(x,y) = F_X(x)F_Y(y).$$

1.3.4 Cas de variables aléatoires discrètes

Considérons un espace probabilisé modélisant une expérience aléatoire de deux variables aléatoires discrètes définies sur l'espace probabilisable $(\Omega, \mathcal{A})$. Chacune de ces variables prend donc ses valeurs sur un sous ensemble au plus dénombrable de $\mathbb{R}$. Nous posons donc
$X(\Omega) = \{x_i, i \in I\}$ et $Y(\Omega) = \{y_j, j \in J\}$, où I et J sont deux ensembles de $\mathbb{N}$. Les valeurs x_i sont supposées distinctes lorsque i parcourt I. De même, nous supposons que les valeurs y_j sont distinctes lorsque j parcourt J. L'ensemble des valeurs prises par le couple (X,Y) ou domaine de variation de (X,Y) est alors

$$X(\Omega) \times Y(\Omega) = \{(x_i, y_j) : i \in I, j \in J\}.$$

1.3.5 Lois conjointes d'un couple de variables aléatoires discrètes

Proposition 1.3.6. *Soit (X, Y) un couple à valeurs discrètes (v.a.d).*

(i) *Si I et J sont finis, on a*

$$\sum_{i\in I}\sum_{j\in J} P[X = x_i, Y = y_j] = 1.$$

(ii) *Pour tout événement B : $\mathbb{P}((X, Y) \in B) = \sum_{(x_i,y_j)\in B\cap X(\Omega)\times Y(\Omega)} P[X = x_i, Y = y_j]$.*

Définition 1.16. *Etant donné un couple (X, Y) de variables aléatoires discrètes (v.a.d) , on appelle loi conjoite du couple (X, Y), l'application qui à tout $(x_i, y_j) \in X(\Omega) \times Y(\Omega)$ associe*

$$\mathbb{P}_{ij} = \mathbb{P}_{XY}(x_i, y_j) = \mathbb{P}(X = x_i, Y = y_j). \tag{1.1}$$

Lois marginales

Théorème 1.3.7. *Soit (X, Y) un couple de variables aléatoires discrètes. Alors les lois marginales de X et Y sont :*

$$\forall x_i \in X(\Omega), \mathbb{P}_{i\bullet} = \mathbb{P}(X = x_i) = \sum_{j\in J} \mathbb{P}([X = x_i, Y = y_j]) = \sum_{j\in J} \mathbb{P}_{ij}, \tag{1.2}$$

$$\forall y_j \in Y(\Omega), \mathbb{P}_{\bullet j} = \mathbb{P}(Y = y_j) = \sum_{i\in I} \mathbb{P}([X = x_i, Y = y_j]) = \sum_{i\in I} \mathbb{P}_{ij}. \tag{1.3}$$

Nous pouvons aussi calculer la fonction de répartition du couple (X, Y) et nous obtenons :

$$\begin{aligned} F_{XY}(x, y) &= \mathbb{P}[X \le x, Y \le y] \\ &= \sum_{\substack{(i,j)\in I\times J \\ x_i\le x, y_j\le y}} \mathbb{P}_{ij} = \sum_{\substack{i\in I \\ x_i\le x}} \left(\sum_{\substack{j\in J \\ y_j\le y}} \mathbb{P}_{ij} \right) = \sum_{\substack{j\in J \\ y_j\le y}} \left(\sum_{\substack{i\in I \\ x_i\le x}} \mathbb{P}_{ij} \right). \end{aligned}$$

Remarque 1.3.1. *Les lois marginales ne permettent pas de déterminer la loi : la connaissance de la loi conjointe permet de calculer les lois marginales. Mais la réciproque est généralement fausse.*

Exemple 1.3.1. *Considérons un couple (X, Y) de variables aléatoires discrètes, chacune à valeurs dans $\{0, 1\}$. Supposons que la loi de ce couple soit :*

$$\mathbb{P}[X = 0, Y = 0] = \beta, \ \ \mathbb{P}[X = 0, Y = 1] = \frac{1}{2} - \beta$$

$$\mathbb{P}[X=1,Y=0]=\frac{1}{2}-\beta,\mathbb{P}[X=1,Y=1]=\beta.$$

où $0 \leq \beta \leq \frac{1}{2}$. Nous déterminons les marginales de X et de Y en appliquant (1.2) et (1.3) on obtient :

$$\mathbb{P}[X=0]=\mathbb{P}[X=0,Y=0]+\mathbb{P}[X=0,Y=1]=\frac{1}{2}.$$

$$\mathbb{P}[X=1]=\mathbb{P}[X=1,Y=0]+\mathbb{P}[X=1,Y=1]=\frac{1}{2}.$$

La loi de X est donc de Bernoulli de paramètre $\frac{1}{2}$. De même, on a :

$$\mathbb{P}[Y=0]=\mathbb{P}[X=0,Y=0]+\mathbb{P}[X=0,Y=1]=\frac{1}{2}.$$

$$\mathbb{P}[Y=1]=\mathbb{P}[X=0,Y=0]+P[X=1,Y=1]=\frac{1}{2}.$$

Les deux variables X et Y suivent donc des lois de Bernoulli qui ne dépendent pas de β, alors que la loi conjointe, elle, en dépend. Il est donc impossible de remonter à la loi conjointe du couple (X,Y) à partir de la seule donnée des lois marginales de X et Y.

Théorème 1.3.8. *Soit (X,Y) un couple de v.a.d. X et Y sont indépendantes si et seulement si pour tout $(x_i,y_i) \in X(\Omega)\times Y(\Omega)$, on a :*

$$\mathbb{P}[X=x_i,Y=y_j]=\mathbb{P}\left(X=x_i\right)\mathbb{P}\left(Y=y_j\right).$$

Démonstration. Supposons que les variables X et Y sont indépendantes. Classons les x_i par ordre croissant ainsi que les y_i. L'événement $[X_i=x_i]$ s'écrit encore $[X \leq x_i]\backslash[X \leq x_{i-1}]$ parceque $x_{i-1}<x_i$. De même, nous avons $[Y=y_j]=[Y\leq y_j]\backslash[Y\leq y_{j-1}]$. *Ainsi*

$$\mathbb{P}[X=x_i]=\mathbb{P}[X\leq x_i]-\mathbb{P}[X\leq x_{i-1}]=F_X(x_i)-F_X(x_{i-1}) \tag{1.4}$$

et, de la même manière,

$$\mathbb{P}[Y=y_j]=\mathbb{P}[Y\leq y_j]-\mathbb{P}[Y\leq y_{j-1}]=F_Y(y_j)-F_Y(y_{j-1}). \tag{1.5}$$

Si nous faisons le produit $\mathbb{P}[X=x_i]\times\mathbb{P}[Y=y_j]$, nous obtenons donc :

$$\begin{aligned}\mathbb{P}\left(X=x_i\right)\mathbb{P}\left(Y=y_j\right) &= F_X(x_i)F_Y(y_j)-F_X(x_i)F_Y(y_{j-1})-F_X(x_{i-1})F_Y(y_j)+F_X(x_{i-1})F_Y(y_{j-1}) &(1.6)\\ &= F_{XY}(x_i,y_j)-F_{XY}(x_i,y_{j-1})-F_{XY}(x_{i-1},y_j)+F_{XY}(x_{i-1},y_{j-1}) &(1.7)\end{aligned}$$

où la dernière égalité provient de l'indépendance de X et de Y. Par définition de la fonction de répartition et en prenant en compte que :

$$\begin{aligned}\left([X\leq x_i]\cap[Y\leq y_j]\right)\backslash\left([X\leq x_i]\cap[Y\leq y_{j-1}]\right) &= [X=x_i]\cap\left([Y\leq y_j]\backslash[Y\leq y_{j-1}]\right) &(1.8)\\ &= [X\leq x_i]\cap[Y=y_j]. &(1.9)\end{aligned}$$

Nous calculons :

$$\begin{aligned}F_{XY}(x_i,y_j)-F_{XY}(x_i,y_{j-1}) &= \mathbb{P}\left([X\le x_i]\cap[Y\le y_j]\right)-\mathbb{P}\left([X\le x_i]\cap[Y\le y_{j-1}]\right)\\ &= \mathbb{P}\left([X\le x_i]\cap[Y\le y_j]\right)\not\ \left([X\le x_i]\cap[Y\le y_{j-1}]\right)\\ &= \mathbb{P}\left([X\le x_i]\cap[Y=y_j]\right).\end{aligned}$$

De la même manière, on établit l'égalité :

$$F_{XY}(x_{i-1},y_j)-F_{XY}(x_{i-1},y_{j-1})=\mathbb{P}\left([X\le x_{i-1}]\cap[Y=y_j]\right). \tag{1.10}$$

En reportant (1.10) et (1.9) dans (1.6), il vient :

$$\begin{aligned}\mathbb{P}[X=x_i]\mathbb{P}[Y=y_j] &= \mathbb{P}\left([X\le x_i]\cap[Y=y_j]\right)-\mathbb{P}\left([X\le x_{i-1}]\cap[Y=y_j]\right)\\ &= \mathbb{P}\left([X\le x_i]\cap[Y=y_j]\right)/\left([X\le x_{i-1}]\cap[Y=y_j]\right).\end{aligned}$$

Puisque

$$\left([X\le x_i]\cap[Y=y_j]\right)\setminus\left([X\le x_{i-1}]\cap[Y=y_j]\right)=[X\le x_i]\cap[Y=y_j],$$

nous obtenons finalement :

$$\mathbb{P}\left([X=x_i]\cap[Y=y_j]\right)=\mathbb{P}\left([X=x_i]\right)\mathbb{P}\left([Y=y_j]\right).$$

Si nous calculons maintenant la fonction de répartition du couple (X,Y), nous trouvons que :

$$\begin{aligned}F_{XY}(x,y) &= \sum_{\substack{(i,j)\in I\times\jmath\\ x_i\le x,y_j\le y}}\mathbb{P}[X=x_i,Y=y_j]\\ &= \sum_{\substack{i\in I\\ x_i\le x}}\sum_{\substack{j\in J\\ y_j\le y}}\mathbb{P}[X=x_i]\mathbb{P}[Y=y_j]\\ &= \sum_{\substack{i\in I\\ x_i\le x}}\mathbb{P}[X=x_i]\sum_{\substack{j\in J\\ y_j\le y}}\mathbb{P}[Y=y_j]\\ &= F_X(x)\times F_Y(y).\end{aligned}$$

La fonction de répartition du couple (X,Y) est le produit des fonctions de répartition de X et de Y. Nous dirons donc que les variables aléaatoires X et Y sont indépendantes car les lois marginales de X et de Y ne dépendent aucunement l'une de l'autre.

1.3.6 Lois conditionnelles

Définition 1.17. *Etant donné un couple (X,Y) de v.a.d et $y_j\in Y(\Omega)$ tel que $\mathbb{P}(Y=y_j)\neq 0$. On appelle loi conditionnelle de X sachant $[Y=y_j]$, l'application définie sur $X(\Omega)$ par :*

$$\mathbb{P}\left(X=x_i/Y=y_j\right)=\frac{\mathbb{P}(X=x_i,Y=y_j)}{\mathbb{P}(Y=y_j)}. \tag{1.11}$$

De même, si $x_i \in X(\Omega)$ est tel que $\mathbb{P}(X = x_i) \neq 0$, on appelle loi conditionnelle de Y sachant $[X = x_i]$, l'application définie sur $Y(\Omega)$ par :

$$\mathbb{P}\left(Y = y_j / X = x_i\right) = \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(X = x_i)}. \tag{1.12}$$

Théorème 1.3.9. ***Théorème des probabilités composées [?].*** *Soit (X, Y) un couple de v.a.d . Pour tout $(x_i, y_j) \in X(\Omega) \times Y(\Omega)$, on a*

$$\mathbb{P}(X = x_i, Y = y_j) = \begin{cases} P(X{=}x_i/Y = y_j)\mathbb{P}(Y = y_j) & \text{si } \mathbb{P}(Y = y_j) \neq 0, \\ 0 & \text{sinon.} \end{cases} \tag{1.13}$$

Théorème 1.3.10. ***Théorème des probabilités totales [?].*** *Soit (X, Y) un couple de v.a.d . Pour tout $(x_i, y_j) \in X(\Omega) \times Y(\Omega)$, on a*

$$\mathbb{P}(X = x_i) = \sum_{j \in J} \mathbb{P}\left(X = x_i / Y = y_j\right) \mathbb{P}(Y = y_j); \mathbb{P}(Y = y_j) = \sum_{i \in I} \mathbb{P}\left(Y / X = x_i\right) \mathbb{P}(X = x_i).$$

Exercice. Un lot de n articles présente un mélange des poduits de trois usines : n_1 articles de l'usine U_1, n_2 articles de l'usine U_2 et n_3 articles de l'usine U_3. Pour les articles de l'usine U_1, la probabilité de fonctionner sans défaillance pendant un temps τ est p_1, de l'usine U_2 est p_2 et de l'usine U_3 est p_3. On tire au hasard un article, calculer la probabilité que l'article fonctionnera sans défaillance pendant un temps τ.

Solution : Soit $A = \{$ article tiré fonctionnera sans défaillance pendant un temps $\tau\}$ et pour $i = 1, 2, 3$, $B_i = \{$article tiré appartient à l'usine $U_i\}$. On cherche à calculer $\mathbb{P}(A)$. Par la propriété des probabilités totale , on a

$$\mathbb{P}(A) = \sum_{i=1}^{3} \mathbb{P}(A \cap B_i) = \sum_{i=1}^{3} \mathbb{P}(B_i)\mathbb{P}(A/B_i), B_i \neq \varnothing, \bigcup_i B_i = \Omega, B_i \cap B_j = \varnothing, i \neq j, i = 1, 2, 3.$$

Or, pour $i = 1, 2, 3$, $\mathbb{P}(B_i) = \frac{n_i}{n}$ et $\mathbb{P}(A/B_i) = p_i$, d'où :

$$\mathbb{P}(A) = \sum_{i=1}^{3} \frac{n_i}{n} p_i = \frac{n_1 p_1 + n_2 p_2 + n_3 p_3}{n}.$$

Théorème 1.3.11. ***Théorème de Bayes [?].*** *Soit (X, Y) un couple de variables aléatoires discrètes. Alors pour tout $(x_i, y_j) \in X(\Omega) \times Y(\Omega)$, on a*

$$\mathbb{P}(X = x_i / Y = y_j) = \frac{\mathbb{P}(Y = y_j / X = x_i)\mathbb{P}(X = x_i)}{\sum_i \mathbb{P}(Y / X = x_i)\mathbb{P}(X = x_i)};\ \mathbb{P}(Y = y_j / X = x_j) = \frac{\mathbb{P}(X = x_i / Y = y_j)\mathbb{P}(Y = y_j)}{\sum_j \mathbb{P}(X = x_i / Y = y_j)\mathbb{P}(Y = y_j)} \tag{1.14}$$

Conditionnement d'une variable aléatoire discrète par un événement de probabilité non nulle

Soient $A \in \mathcal{A}$ tel que $\mathbb{P}(A) > 0$ et une variable aléatoire X et à valeurs discrètes dans $\{x_i : i \in I\}$ où I est au plus dénombrable et les valeurs x_i sont distinctes deux à deux. On a :

$$\sum_{i\in I} \mathbb{P}[X = x_i/A] = \sum_{i\in I} \mathbb{P}\left([X = x_i] \cap A\right)/\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{i\in I}(X = x_i) \cap A\right)/\mathbb{P}(A). \qquad (1.15)$$

Exemple 1.3.2. *On considère ici deux variables aléatoires discrètes X et Y dont la loi jointe est donnée dans le tableau suivant :*

$X \setminus Y$	*0*	*1*
0	$\frac{2}{8}$	$\frac{3}{8}$
1	$\frac{1}{8}$	$\frac{2}{8}$

Supposons que l'événement $[X = 0]$ est réalisé et déterminons la loi de Y . On a :

$$\mathbb{P}\left(Y = 0/X = 0\right) = \frac{\mathbb{P}\left(X = 0, Y = 0\right)}{\mathbb{P}(X = 0)} = \frac{2}{5}$$

et

$$\mathbb{P}\left(Y = 1/X = 0\right) = \frac{3}{5}.$$

On vient ainsi de définir une nouvelle variable aléatoire notée $[Y/X = 0]$ qui suit une loi de Bernoulli de paramètre $\frac{3}{5}$. De même il est facile de voir que la variable aléatoire $[Y/X = 1]$ suit une loi de Bernoulli de paramètre $\frac{2}{3}$. Ces lois sont appelées lois conditionnelles de Y sachant X.

1.3.7 Cas des V.a.r absolument continues

Définition 1.18. *Une application de $\mathbb{R}^2$ dans $\mathbb{R}$ est une densité de probabilité dans $\mathbb{R}^2$ si elle est mesurable, positive ou nulle et vérifie la condition de normalisation :*

$$\int_{\mathbb{R}^2} f(x, y)dxdy = 1. \qquad (1.16)$$

Définition 1.19. *Soit (X, Y) un couple de variables aléatoires à valeurs dans $\mathbb{R}^2$. On dit que le couple (X, Y) est absolument continu s'il existe une densité de probabilité $f_{X,Y}$ sur $\mathbb{R}^2$ telle que, pour tout couple (x, y) de réels :*

$$F_{X,Y}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(\mu, \nu)d\mu d\nu.$$

Si l'on connait la densité de probabilité $f_{X,Y}$ d'un couple (X, Y) absolument continu de variables aléatoires, on accéde donc à la fonction de répartition $F_{X,Y}$ de ce couple en intégrant la densité, variable après variable, en choisissant l'ordre d'intégration à sa guise puisque, pour tout $x, y \in \mathbb{R}$:

$$F_{X,Y}(x, y) = \int_{-\infty}^{y}\left(\int_{-\infty}^{x} f_{X,Y}(\mu, \nu)d\mu\right)d\nu = \int_{-\infty}^{x}\left(\int_{-\infty}^{y} f_{X,Y}(\mu, \nu)d\nu\right)d\mu.$$

Exemple 1.3.3. *Soit (X, Y) un couple de variables aléatoires dont la loi est donnée par la densité :*

$$f_{X,Y}(x, y) = \frac{xy}{2} 1_{0 \leq y \leq x \leq 2}.$$

Déterminons la fonction de répartition au point (x_0, y_0). Il faut distinguer plusieurs cas. Par exemple, si $0 \leq y_0 \leq x_0 \leq 2$.

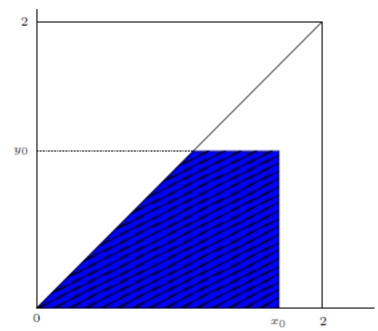

FIGURE 1.1 – Domaine d'intégration

La fonction de répartition s'obtienne en intégrant la densité sur le domaine hachuré 1.1. On peut calculer cette intégrale de deux manières différentes.

$$F_{(X,Y)}(x_0, y_0) = \int_0^{y_0} \int_y^{x_0} \frac{xy}{2} dx dy = \frac{y_0^2}{16}(2x_0^2 - y_0^2), 0 \leq y_0 \leq x_0 \leq 2$$

où

$$F_{(X,Y)}(x_0, y_0) = \int_0^{y_0} \int_0^{x_0} \frac{xy}{2} dx dy + \int_{y_0}^{x_0} \int_0^{y_0} \frac{xy}{2} dy dx = \frac{y_0}{16}(2x_0^2 - y_0^2)$$

En répétant ces calculs sur les différents ensembles, on obtient

$$F_{X,Y}(x, y) = \begin{cases} 0 & \textit{si } x < 0 \textit{ ou } y < 0 \\ \frac{y^2}{16}(2x^2 - y^2) & \textit{si} 0 \leq y \leq x \leq 2 \\ \frac{y^2}{16}(8 - y^2) & \textit{si } 0 < y \leq 2 \leq x \\ \frac{x^4}{16} & \textit{si } 0 \leq x \leq y \textit{ et } x < 2 \\ 1 & \textit{si } x \leq 2 \textit{ et } y \leq 2 \end{cases} \tag{1.17}$$

Distributions marginales

Définition 1.20. *Les fonctions de répartition marginales $F_X(x)$ et $F_Y(y)$ sont les fonctions de répartition de chacune des deux variables considérées séparément, avec :*

$$F_X(x) = \mathbb{P}(X \leq x); F_Y(y) = \mathbb{P}(Y \leq y).$$

Elles peuvent s'obtenir directement à partir de la fonction de répartition conjointe, avec

$$F_X(x) = F_{X,Y}(x, +\infty); F_Y(y) = F_{X,Y}(+\infty, y).$$

Démonstration. *Posons les événements $A \equiv (X \leq x)$ et $B \equiv (Y \leq y)$, de telle sorte que $F_{X,Y}(x,y) = \mathbb{P}(A \cap B)$. L'événement $\Omega \equiv (Y \leq \infty)$ est l'événement certain, et donc $F_{X,Y}(x,\infty) = \mathbb{P}(A \cap \Omega) = \mathbb{P}(A) = F_X(x)$. L'autre résultat se démontre de manière similaire.*

Proposition 1.3.12. *Soit (X,Y) un couple de variables aléatoires à valeurs dans $\mathbb{R}^2$ absolument continu. Soient $F_{X,Y}$ la fonction de répartition du couple (X,Y) et $f_{X,Y}$ la densité de probabilité de ce couple.*

i) $\mathbb{P}[(X,Y) \in B] = \int_B f_{X,Y}(x,y)dxdy$, *pour tout* $B \in \mathfrak{B}(\mathbb{R}^2)$.

ii) *Les variables aléatoires sont absolument continues de densité de probabilité respectives :*

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)dy \;\; et \;\; f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y)dx.$$

Ces densités de probabilité sont appelées densités marginales de X et de Y respectivement.

iii) $f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x,y) = \frac{\partial^2 F_{X,Y}}{\partial y \partial x}(x,y)$ *pour presque tout* $(x,y) \in \mathbb{R}^2$.

Exemple 1.3.4. *Soit la fonction $f(x,y)$ suivante :*

$$f_{X,Y}(x,y) = \begin{cases} 2(x+y-2xy) & (x,y) \in [0,1] \times [0,1]; \\ 0 & \textit{ailleurs}. \end{cases} \tag{1.18}$$

1. *On demande de vérifier qu'il s'agit d'une fonction de densité de probabilié.*
2. *Calculer la probabilité de l'événement $B \equiv \left(\mid X - Y \mid \leq \frac{1}{2}\right)$.*
3. *Calculer les lois marginales de X et Y.*

1. *La fonction $f_{X,Y}(x,y)$ est une densité de proailité à condition que $f_{X,Y}(x,y) \geq 0, \forall (x,y) \in \mathbb{R}^2$ et que $\int_{\mathbb{R}^2} f(x,y)dxdy = 1$. La première condition est vérifiée, puisque $2(x+y-2xy) \geq 0$ pour tout $(x,y) \in [0,1] \times [0,1]$. Pour la seconde condition, on obtient :*

$$\int_0^1 \left(\int_0^1 2(x+y-2xy)dy \right) dx = \int_0^1 \left[2xy + y^2 - 2xy^2\right]_0^1 dx = \int_0^1 dx = 1.$$

Ce qui montre que $f_{X,Y}(x,y)$ est bien une fonction densité de probabilité.

2. *On peut calculer :*

$$\mathbb{P}(B) = \int_0^{\frac{1}{2}} \int_0^{x+\frac{1}{2}} f_{X,Y}(x,y)dydx + \int_{\frac{1}{2}}^1 \int_{x-\frac{1}{2}}^1 f_{X,Y}(x,y)dydx$$

Sur base de la symétrie de la fonction $f_{X,Y}(x,y)$, on voit cependant que $\mathbb{P}(\Omega) = \mathbb{P}(B) + 2\mathbb{P}(A) = 1$, avec :

$$\begin{aligned} \mathbb{P}(A) &= \int_{\frac{1}{2}}^1 \int_0^{x-\frac{1}{2}} f_{X,Y}(x,y)dydx = \int_{\frac{1}{2}}^1 \left[2xy + y^2 - 2xy^2\right]_0^{x-\frac{1}{2}} dx \\ &= \int_{\frac{1}{2}}^1 \left[-2x^3 + 5x^2 - \frac{5x}{2} + \frac{1}{4}\right] dx = \left[-\frac{x^4}{2} + \frac{5x^3}{3} - \frac{5x^2}{4} + \frac{x}{4}\right]_{\frac{1}{2}}^1 = \frac{17}{96} \end{aligned}$$

et l'on obtient $\mathbb{P}(B) = 1 - 2\frac{17}{96} = \frac{31}{48}$.

3. *Sur la base de la formule $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$, on obtient :*

$$f_X(x) = \int_0^1 2\left[x + y - 2xy\right] dy = \left[2xy + y^2 - 2xy\right]_0^1 = 1, \forall x \in [0,1],$$

tandis qu'elle vaut 0 ailleurs. La fonction $f(x,y)$ étant symétrique par rapport aux variables X et Y, on obtient directement $f_Y(y) = 1, \forall y \in [0,1]$ et vaut 0 ailleurs. On a donc $X \sim U_{[0,1]}$ et $Y \sim U_{[0,1]}$.

Lois conditionnelles

Définition 1.21. *Soit un couple (X,Y) absolument continu de variable aléatoires réelles, $f_{X,Y}$ la densité de probabilité de ce couple, f_X et f_Y les densités de probabilié marginales respectives de X et de Y. Pour tout $(x,y) \in \mathbb{R}^2$, on définit les densités de probabilité conditionnelles de X sachant $[Y = y]$ et de Y sachant $[X = x]$ par :*

$$f_{X/Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, f_Y(y) > 0; f_{Y/X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}, f_X(x) > 0. \tag{1.19}$$

et les fonctions de répartition conditionnelles de X sachant $[Y = y]$ et de Y sachant $[X = x]$ par :

$$F_{X/Y=y}(x) = \int_{-\infty}^{x} f_{X/Y=y}(t)dt, F_{Y/X=x}(y) = \int_{-\infty}^{y} f_{Y/X=x}(t)dt. \tag{1.20}$$

Conditionnement d'une variable aléatoire absolument continue par un événement de probabilité non nulle

Proposition 1.3.13. *Soient X une variable aléatoire réelle et un événement $A \in \mathfrak{B}(\mathbb{R})$ tel que $\mathbb{P}(A) \neq 0$. L'application $F_{X/A}$ de $\mathbb{R}$ dans $[0,1]$ qui associé à tout $x \in \mathbb{R}$ la probabilité conditionnelle $F_{X/A}(x) = \mathbb{P}[X \leq x/A]$ est une fonction de répartition appelée fonction de répartition conditionnelle de X sachant A.*

Lemme 1.3.14. *Soient un espace probabilisé $(\Omega, \mathfrak{B}(\mathbb{R}^2), \mathbb{P})$ et $A \in \mathfrak{B}(\mathbb{R}^2)$ tel que $0 < \mathbb{P}(A) < 1$ et une variable aléatoire X définie sur l'espace probabilisable $(\Omega, \mathfrak{B}(\mathbb{R}^2)$ admettant des fonctions de répartition conditionnelle $F_{X\diagup A}$ et $F_{X\diagup A^c}$ absolument continues de densité respectives $f_{X/A}$ et f_{X/A^c}. La variable aléatoire X est absolument continue de densité*

$$f_X(x) = \mathbb{P}(A) f_{X/A}(x) + \mathbb{P}(A^c) f_{X/A^c}(x). \tag{1.21}$$

Démonstration. *Par définition de la fonction de répartition conditionnelle $F_{X/A}$ de X sachant A, nous avons :*

$$F_{X/A}(x) = \frac{\mathbb{P}\left([X \leq x] \cap A\right)}{\mathbb{P}(A)}.$$

Si $F_{X/A}$ admet une densité $f_{X/A}$, alors nous avons aussi :

$$F_{X/A}(x) = \int_{-\infty}^{x} f_{X/A}(t)dt.$$

Nous déduisons de ces égalités que :

$$\mathbb{P}\left([X \leq x] \cap A\right) = \int_{-\infty}^{x} \mathbb{P}(A) f_{X/A}(t)dt.$$

Puisque $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$*, un raisonnement analogue conduit à l'égalité :*

$$\mathbb{P}\left([X \leq x] \cap A^c\right) = \int_{-\infty}^{x} \mathbb{P}(A^c) f_{X/A^c}(t)dt.$$

Puisque les ensembles $[X \leq x] \cap A$ *et* $[X \leq x] \cap A^c$ *sont disjoints et d'union égale à* $[X \leq x]$*, la fonction de répartition* F_X *de* X *s'écrit :*

$$F_X(x) = \mathbb{P}[X \leq x] = \mathbb{P}\left([X \leq x] \cap A\right) + \mathbb{P}\left([X \leq x] \cap A^c\right).$$

Il vient donc :

$$F_X(x) = \int_{-\infty}^{x} \left(\mathbb{P}(A) f_{X/A}(t) + \mathbb{P}(A^c) f_{X/A^c}(t)\right) dt = \int_{-\infty}^{x} f_X(t)dt.$$

D'où le résultat.

Exemple 1.3.5. *Soit la fonction densité* $f(x, y)$ *donnée par :*

$$f_{X,Y}(x,y) = \begin{cases} 6x & \textit{pour } 0 \leq x \leq y \leq 1; \\ 0 & \textit{ailleurs.} \end{cases}$$

Les distributions marginales sont :

$$f_X(x) = \begin{cases} \int_x^1 6x dy = 6x(1-x) & \textit{pour } x \in [0,1]; \\ 0 & \textit{ailleurs.} \end{cases}$$

$$f_Y(y) = \begin{cases} \int_0^y 6x dx = 3y^2 & \textit{pour } y \in [0,1]; \\ 0 & \textit{ailleurs.} \end{cases}$$

Les fonctions de répartition marginales :

$$F_X(x) = \begin{cases} 0 & x < 0; \\ \int_0^x 6u(1-u)du = 3x^2 - 2x^3 & \textit{pour } x \in [0,1]; \\ 1 & x > 1. \end{cases}$$

$$F_Y(y) = \begin{cases} 0 & y < 0; \\ \int_0^y 3v^2 dv = y^3 & \textit{pour } y \in [0,1]; \\ 1 & y > 1. \end{cases}$$

Les densités conditionnelles sont alors :

$$f_{X/Y=y}(x) = \begin{cases} \frac{2x}{y^2} & \textit{pour } 0 \leq x \leq y \leq 1; \\ 0 & \textit{ailleurs.} \end{cases}$$

$$f_{Y/X=x}(y) = \begin{cases} \frac{1}{1-x} & \textit{pour } 0 \leq x \leq y \leq 1; \\ 0 & \textit{ailleurs} \end{cases}$$

et les distributions conditionnelles :

$$F_{X/Y=y}(x) = \begin{cases} 0 & x < 0; \\ \int_0^x \frac{2u}{y^2} du = \frac{x^2}{y^2} & 0 \leq x \leq y \leq 1; \\ 1 & x > y. \end{cases}$$

$$F_{Y/X=x}(y) = \begin{cases} 0 & y < x; \\ \int_0^y \frac{1}{1-x} dv = \frac{y}{1-x} & 0 \leq x \leq y \leq 1; \\ 1 & y > 1. \end{cases}$$

Proposition 1.3.15. *Soient deux variables aléatoires réeeles X et Y absolument continues de densité respective f_X et f_Y. Les variables aléatoires X et Y sont indépendantes si et seulement si le couple (X, Y) est absolument continu de densité :*

$$f_{X,Y}(x, y) = f_X(x) \times f_Y(y).$$

Démonstration. *Tout d'abord, remarquons que le produit $f_{XY} = f_X \times f_Y$ est une densité de probabilité puisque :*

$$\int f(x, y) dx dy = \left(\int_{\mathbb{R}} f_X(x) dx \right) \left(\int_{\mathbb{R}} f_Y(y) dy \right) = 1.$$

Puisuqe $f_X \times f_Y$ est une densité deprobabilité du couple (X, Y), la fonction de répartition associée s'écrit :

$$\begin{aligned} F_{X,Y}(x, y) &= \int_{-\infty}^x \int_{-\infty}^y f_X(u) f_Y(v) du dv \\ &= \left(\int_{-\infty}^x f_X(u) du \right) \left(\int_{-\infty}^y f_Y(v) dv \right) \\ &= F_X(x) \times F_Y(y). \end{aligned}$$

Réciproquement, supposons que X et Y soient indépendantes. Considérons alors la fonction $f = f_X \times f_Y$. Soit $(x, y) \in \mathbb{R}^2$. Nous avons :

$$\begin{aligned} F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv &= \int_{-\infty}^x \int_{-\infty}^y f_X(u) f_Y(v) du dv \\ &= \left(\int_{-\infty}^x f_X(u) du \right) \left(\int_{-\infty}^y f_Y(v) dv \right) \\ &= F_X(x) F_Y(y). \end{aligned}$$

Puisque les variables aléatoires X et Y sont supposées indépendantes,

$$F_X(x) \times F_Y(y) = F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(y, v) du dv.$$

Exemple 1.3.6. *Soit $Z = (X,Y)^t$ un couple aléatoire de loi uniforme sur $\mathcal{D} = [-1,1] \times [-1,1]$. On a alors :*

$$f_{X,Y}(x,y) = \frac{1}{4}1_{\mathcal{D}}(x,y) = \frac{1}{4}1_{[-1,1]}(x)1_{[-1,1]}(y) = f_X(x)f_Y(y)$$

avec

$$f_X(x) = \frac{1}{2}1_{[-1,1]}(x);\, f_Y(y) = \frac{1}{2}1_{[-1,1]}(y).$$

Théorème 1.3.16. ***Théorème des probabilité composées.*** *Soit (X,Y) un couple de variables aléatoires réelles (v.a.r) continu défini sur $(\mathbb{R}^2, \mathfrak{B}(\mathbb{R}^2))$. Alors pour tout $(x,y) \in \mathbb{R}^2$, on a*

$$f_{X,Y}(x,y) = f_{Y/X=x}(y)f_X(x) = f_{X/Y=y}(x)f_Y(y). \tag{1.22}$$

Théorème 1.3.17. ***Théorème des probabiités totales .*** *Soit (X,Y) un couple de v.a.r continu . Pour tout $(x,y) \in \mathbb{R}^2$, on a*

$$f_X(x) = \int_{-\infty}^{\infty} f_{X/Y}(x)f_Y(y)dy; \quad f_Y(y) = \int_{-\infty}^{\infty} f_{Y/X}(y)f_X(x)dx.$$

Théorème 1.3.18. ***Théorème de Bayes .*** *Soit (X,Y) un couple de variables aléatoires réelles. Pour tout $(x,y) \in \mathbb{R}^2$, on a*

$$f_{(X/Y=y)}(x) = \frac{f_{(Y/X=x)}(y)f_X(x)}{\int_{-\infty}^{\infty} f_{(Y/X=x)}(y)f_X(x)dx};\, f_{(Y/X=x)}(y) = \frac{f_{(X/Y=y)}(x)f_Y(y)}{\int_{-\infty}^{\infty} f_{(X/Y=y)}(x)f_Y(y)dy}.$$

1.3.8 Transformation d'un couple aléatoire

Théorème 1.3.19. Soit (X,Y) un couple de variables aléatoires discrètes et $\mathfrak{D}$ une partie contenant $X(\Omega) \times Y(\Omega)$ et $\phi : \mathfrak{D} \to \mathbb{R}$. On suppose que $Z = \phi(X,Y)$ est une variable aléatoire discrète définie sur $B \subseteq X(\Omega) \times Y(\Omega)$. Lorsque $X(\Omega)$ et $Y(\Omega)$ sont finis, on a

$$\mathbb{P}_Z(z) = \sum_{(x,y)\in B} \mathbb{P}(X = x, Y = y). \tag{1.23}$$

Théorème 1.3.20. *Soit (X,Y) un couple de variables aléatoires continues et $\mathfrak{D}$ une partie contenant $X(\Omega) \times Y(\Omega)$ et $\phi : \mathfrak{D} \to \mathbb{R}$. On suppose que $Z = \phi(X,Y)$ est une variable aléatoire continue définie sur $B \subseteq \mathfrak{B}(\mathbb{R}^2)$, alors*

$$F_Z(z) = \int_{(x,y)\in B} f(x,y)dydx. \tag{1.24}$$

Quelques cas particuliers

Quelle que soit la fonction $Z = \phi(X,Y)$, il est toujours possible de déterminer les fonctions F_Z et f_Z en utilisant le procédé général exposé ci-dessus. On remarquera cependant que, dans certaines situations, la fonction f_Z peut s'obtenir directement sans passer par le calcul de F_Z.

Cas $n°1$ **:** $Z = X + Y$ **avec** X **et** Y **sont indépendantes**

Théorème 1.3.21. *Pour une somme de deux variables aléatoires indépendantes, la fonction de masse de probabilité est la convolution des fonctions de masses de probabilité marginales pour* X *et* Y*. C'est-à-dire,*

$$\mathbb{P}_Z(z) = \left\{ \begin{array}{l} \sum_i \mathbb{P}_X(x_i)\mathbb{P}_Y(z - x_i) \\ \sum_j \mathbb{P}_X(z - y_j)\mathbb{P}_Y(y_j) \end{array} \right. ; f_Z(z) = \left\{ \begin{array}{l} \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx \\ \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy. \end{array} \right. ;$$

Démonstration. *L'événement* $Z \le z$ *est équivalent à* $X + Y \le z$*, et l'on obtient dans le cas continu*

$$F_Z(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f(x,y)dydx = \int_{-\infty}^{\infty} f_X(x)\left[\int_{-\infty}^{z-x} f_Y(y)dy\right]dx = \int_{-\infty}^{\infty} f_X(x)F_Y(z-x)dx.$$

Dans le cas continu. En dérivant , on a

$$\begin{aligned} f_Z(z) &= \frac{dF_Z(z)}{dz} = \int_{-\infty}^{\infty} f_X(x)\frac{\partial F_Y(z-x)}{\partial z}dx \\ &= \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)dx. \end{aligned}$$

La démonstration de l'autre résultat est similaire avec y *au lieu de* x*.*

Cas $n°2$ **:** $Z = \frac{X}{Y}$ **avec** X **et** Y **quelconques**

Théorème 1.3.22. *Pour un rapport de deux variables quelconques, la loi de probabilité peut s'obtenir directement à partir de la fonction de probabilité conjointe graçe aux formules :*

$$\mathbb{P}_Z(z) = \sum_j \mid y_j \mid \mathbb{P}(y_j z, y_j); f_Z(z) = \int_{-\infty}^{\infty} \mid y \mid f(yz, y)dy.$$

Cas $n°3$ **:** $Z = \max(X,Y)$ **ou** $Z = \min(X,Y)$ **avec** X **et** Y **sont indépendantes**

Théorème 1.3.23. • *Si* $Z = \max(X,Y)$ *avec* X *et* Y *sont indépendantes, alors :*

$$\begin{aligned} F_Z(z) &= F_X(z)F_Y(z). \\ f_Z(z) &= f_X(z)F_Y(z) + F_X(z)f_Y(z). \end{aligned}$$

• *Si* $Z = \min(X,Z)$ *avec* X *et* Y *sont deux variables aléatoires indépendantes, alors :*

$$1 - F_Z(z) = (1 - F_X(z))(1 - F_Y(z)). \tag{1.25}$$

$$f_Z(z) = f_X(z)(1 - F_Y(z)) + (1 - F_X(z))f_Y(z). \tag{1.26}$$

Démonstration. • *L'événement $[Z \leq z]$ est équivalent à $[(X \leq z) \cap (Y \leq z)]$, puisque tous les couples (x, y) pour lesquels $x \leq z$ et $y \leq z$ satisfont la condition $\max(x, y) \leq z$. On a donc $F_Z(z) = F_{X,Y}(z, z)$. Si X et Y sont indépendantes, on a $F_{X,Y}(z, z) = F_X(z)F_Y(z)$. Dans le continu, on obtient*

$$f_Z(z) = \frac{dF_{X,Y}(z, z)}{dz} = \frac{d[F_X(z)F_Y(z)]}{dz} = F_Y(z)\frac{dF_X(z)}{dz} + F_X(z)\frac{dF_Y(z)}{dx} = F_Y(z)f_X(z) + F_X(z)f_Y(z).$$

• *L'événement $[Z \leq z]$ est équivalent à $[(X \leq z) \cup (Y \leq z)]$, puisque tous les couples (x, y) pour lesquels $x \leq z$ ou $y \leq z$ satisfont la condition $\min(x, y) \leq z$.*
On peut donc également écrire que $(Z \leq z) = (X \leq z \cup Y \leq z)$, c'est-à-dire $(Z > z) = (X > z \cap Y > z)$. Puisque X et Y sont indépendantes, cela revient à écrire $1 - F_Z(z) = (1 - F_X(z))(1 - F_Y(z))$. En dérivant cette expression par rapport à z, on trouve la formule proposée pour $f_Z(z)$.

1.4 Vecteur aléatoire à plusieurs dimensions

Définition 1.22. *On considère un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et n variables aléatoires $X_1, X_2, ..., X_n$ définies sur Ω à valeurs dans $\mathbb{R}$. On considère l'application X de $(\Omega, \mathcal{A}, \mathbb{P})$ dans l'espace mesurable $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$ définie pour tout $w \in \Omega$ par :*

$$\begin{aligned} X : \Omega &\to \mathbb{R}^n \\ w &\to X(w) = (X_1(w), X_2(w), ..., X_n(w))^t. \end{aligned}$$

On dit que X est un vecteur aléatoire défini sur $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$

Lemme 1.4.1. *Avec les notations de la définition précédente, $X = (X_1, X_2, ..., X_n)^t$ est un vecteur aléatoire si est seulement si*

$$\{w \in \Omega, X_1(w) \leq x_1, X_2(w) \leq x_2, ..., X_n(w) \leq x_n\} \in \mathfrak{B}(\mathbb{R}^n) \text{ pour tout } n\text{−uplet } (x_1, x_2, ..., x_n)$$

de réels.

On posera :

$$[X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n] = \{w \in \Omega : X_1(w) \leq x_1, X_2(w) \leq x_2, ..., X_n(w) \leq x_n\}.$$

Définition 1.23. ***Loi de probabilité d'un vecteur aléatoire*** *La loi de probabilité d'un vecteur aléatoire X est la probabilité image de $\mathbb{P}$ par X notée $\mathbb{P}_X$ et définie par*

$$\forall B \in \mathfrak{B}(\mathbb{R}^n), \mathbb{P}_X(B) = \mathbb{P}\left(X^{-1}(B)\right).$$

La probabilité $\mathbb{P}_X$ ainsi définie est une mesure de probabilité sur l'espace probabilisable $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n))$ appelée aussi loi conjointe de $(X_1, X_2, ..., X_n)^t$: On appelle ième loi marginale de X et on note $\mathbb{P}_{X_i}$ la loi de la ième composante $X_i (i = 1, 2, ..., n)$.

Définition 1.24. ***Fonction de répartition conjointe***. *Soit $X = (X_1, X_2, ..., X_n)^t$ un vecteur aléatoire à valeurs dans $\mathbb{R}^n$. Pour tout $x = (x_1, ..., x_n)^t \in \mathbb{R}^n$, on définit la fonction de répartition $F_X(x)$ de X en x en posant :*

$$\begin{aligned} F_X(x) &= F_X(x_1, x_2, ..., x_n) = \mathbb{P}[X_1 \le x_1, X_2 \le x_2, ..., X_n \le x_n] \\ &= \mathbb{P}\left([X_1 \le x_1] \cap [X_2 \le x_2] \cap ... \cap [X_n \le x_n]\right). \end{aligned}$$

Propriété 1.4.2. • *Si $\forall\, i = 1, ..., n$ on a $x_i = \infty$, alors $F_X(+\infty) = \mathbb{P}(\Omega) = 1$.*

- *Si $\exists i = 1, ..., n$ tel que $x_i = -\infty$, alors $F_X(x) = \mathbb{P}(\varnothing) = 0$.*

Démonstration. *L'événement $[X_i \le +\infty]$ est certain pour la variable X_i, tandis que $[X_i \le -\infty]$ est l'événement impossible. On a donc $F_X(+\infty, ..., +\infty) = \mathbb{P}(\Omega) = 1$, tandis que $F_X(x_1, ..., x_n) = \mathbb{P}(\varnothing) = 0$ si au moins l'un des x_i est égale à $-\infty$.*

Définition 1.25. ***Vecteur aléatoire discret.*** *On appelle vecteur aléatoire discret à valeurs dans $\mathbb{R}^n$ toute application de la forme :*

$$\begin{aligned} X : \Omega &\to \mathbb{R}^n \\ w &\to X(w) = (X_1(w), X_2(w), ..., X_n(w))^t. \end{aligned}$$

où $X_1, ..., X_n$ sont des variables aléatoires discrètes.

Définition 1.26. ***Loi de probabilité conjointe.*** *On appelle loi conjointe du vecteur aléatoire $X = (X_1, ..., X_n)^t$ la fonction :*

$$\mathbb{P}_X : \begin{array}{c} X_1(\Omega) \times X_2(\Omega) \times ... \times X_n(\Omega) \to [0, 1] \\ (X_1, ..., X_n) \mapsto \mathbb{P}\left(X_1 = x_1, ..., X_n = x_n\right) \end{array}$$

où $[X_1 = x_1, ..., X_n = x_n] = [X_1 = x_1] \cap ... \cap [X_n = x_n]$.

Définition 1.27. *Soit X un vecteur aléatoire à valeurs discrètes. Toute fonction discrète est une loi de probabilité conjointe à condition que les deux propriétés suivantes soient vérifiées.*

- $\mathbb{P}_X(x) \ge 0, \forall x \in \mathbb{R}^n$.
- $\sum\limits_{x} \mathbb{P}_X(x) = \sum\limits_{x_1} ... \sum\limits_{x_n} \mathbb{P}_X(x_1, ..., x_n) = 1.$

Définition 1.28. ***Probabilité d'un événement (cas discret).*** *Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace prbabilisable et $X = (X_1, X_2, ..., X_n)^t$ un vecteur aléatoire à valeurs discrètes de loi $\mathbb{P}_X$. La probabilité d'un événement $B \in \mathcal{A}$ est donnée par :*

$$\mathbb{P}(B) = \sum_{x \in B} \mathbb{P}_X(x).$$

Définition 1.29. ***Vecteur aléatoire absolument continu.*** Soit $X = (X_1, X_2, ..., X_n)^t$ un vecteur aléatoire à valeurs dans $\mathbb{R}^n$. On dit que X est absolument continu s'il existe une fonction mesurable positive ou nulle f_X, appelée densité , telle que :

$$\int_{\mathbb{R}^n} f_X(x)dx = \int_{\mathbb{R}} ... \int_{\mathbb{R}} f_X(x_1, ..., x_n)dx_1...dx_n = 1 \tag{1.27}$$

et

$$F_X(x) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} ... \int_{-\infty}^{x_n} f_X(t_1, ..., t_n)dt_1....dt_n \tag{1.28}$$

Pour tout $(x_1, x_2, ..., x_n)^t \in \mathbb{R}^n$.

Définition 1.30. ***Probabilité d'un événement (cas continu).*** *Soit $X = (X_1, X_2, ..., X_n)^t$ un vecteur aléatoire à valeurs dans $\mathbb{R}^n$ de densité f_X. La probabilité d'un événement $B \in \mathfrak{B}(\mathbb{R}^n)$ est donnée par :*

$$\mathbb{P}(B) = \int_B f_X(x)dx$$

Exemple 1.4.1. *Soit $X = (X_1, X_2)^t$ un vecteur aléatoire dont la densité est donnée par*

$$f_X(x_1, x_2) = \begin{cases} \frac{1}{2\sqrt{x_1 x_2}} & si \;\; 0 < x_1 \leq x_2 < 1 \\ 0 & sinon. \end{cases}$$

On note $A =]-2, \frac{1}{4}[\times]-1, \frac{3}{2}[$. Déterminer $\mathbb{P}(X \in A)$.

$$\mathbb{P}(X \in A) = \int_0^{\frac{1}{4}} \int_{x_1}^{1} \frac{1}{2\sqrt{x_1 x_2}} dx_2 dx_1 = \frac{3}{4}.$$

Proposition 1.4.3. Si le vecteur aléatoire X est absolument continu de fonction de répartition F_X, la densité de X est donnée par :

$$f_X(x_1, ..., x_n) = \frac{\partial^n F_X(x_1, ..., x_n)}{\partial x_1 \partial x_2 ... \partial x_n}.$$

1.5 Distributions marginales

Dans le cas d'un vecteur aléatoire comprenant n variables $(n > 1)$, on peut s'intéresser à la distribution conjointe d'un sous ensemble de m variables parmi les n variables de départ. Pour faciliter les écritures, définissons une partition du vecteur X, avec

$$X = (X_1, ...X_m, X_{m+1}, ..., X_n)^t = (X_a, X_b)^t,$$

où $X_a = (X_1, ...X_m)^t$ et $X_b = (X_{m+1}, ..., X_n)^t$ sont deux morceaux du vecteur aléatoire X, repectivement de dimension $n_a = m \; et \; n_b = n - m$.

Définition 1.31. ***Lois marginales d'un vecteur aléatoire discret.*** Soient deux vecteurs aléatoires $X_a = (X_1, ..., X_m)^t$ et $X_b = (X_{m+1}, ..., X_n)^t$ définis sur le même espace probabilisable $(\Omega, \mathcal{A})$. Si le vecteur $X = (X_1, X_2, ..., X_m, X_{m+1}, ...X_n)^t$ de dimension $m + n_b = n$ est un vecteur aléatoire discret de loi de probabilité $\mathbb{P}_X$, les vecteurs X_a et X_b sont des vecteurs aléatoires discrets et de loi de probabilité respectives :

$$\mathbb{P}_{X_a}(x_a) = \sum_{x_{m+1}} ... \sum_{x_n} \mathbb{P}_X(x) = \sum_{x_{m+1}} ... \sum_{x_n} \mathbb{P}_X(x_1, ..., x_n).$$

$$\mathbb{P}_{X_b}(x_b) = \sum_{x_1} ... \sum_{x_m} \mathbb{P}_X(x) = \sum_{x_1} ... \sum_{x_m} \mathbb{P}_X(x_1, ..., x_n).$$

avec $x_a = (x_1, ..., x_m)^t$ et $x_b = (x_{m+1}, ..., x_n)^t$.

Définition 1.32. ***Densités marginales d'un vecteur aléatoire absolument continu.*** Soient deux vecteurs aléatoires $X_a = (X_1, ..., X_m)^t$ et $X_b = (X_{m+1}, ..., X_n)^t$ définis sur le même espace probabilisable $(\Omega, \mathcal{A})$. Si le vecteur aléatoire $X = (X_1, X_2, ..., X_m, X_{m+1}, ...X_n)^t$ de dimension $m + n_b = n$ est absolument continu de densité de probabilité f_X, les vecteurs X_a et X_b sont absolument continus et de densité de probabilité respectives :

$$f_{X_a}(x_a) = \int_{\mathbb{R}^{n-m}} f_X(x)dx_b = \int_{\mathbb{R}^{n-m}} f_X(x_1, ..., x_n)dx_{m+1}...dx_n, \tag{1.29}$$

$$f_{X_b}(x_b) = \int_{\mathbb{R}^m} f_X(x)dx_a = \int_{\mathbb{R}^m} f_X(x_1, ..., x_n)dx_1...dx_m, \tag{1.30}$$

avec $x_a = (x_1, ..., x_m)^t$ *et* $x_b = (x_{m+1}, ..., x_n)^t$.

Distributions conditionnelles

Théorème 1.5.1. *Soit $X = (X_a, X_b)^t$ un vecteur aléatoire (discret ou continu), la fonction de (densité de) probabilité conditionnelle de X_a est donnée par :*

$$\mathbb{P}_{X_a/X_b=x_b}(x_a) = \frac{\mathbb{P}_X(x)}{\mathbb{P}_{X_b}(x_b)}; \quad f_{X_a/X_b=x_b}(x_a) = \frac{f_X(x)}{f_{X_b}(x_b)}, x = (x_1, ..., x_n).$$

Les fonctions $\mathbb{P}_{X_b/X_a}$ et f_{X_b/X_a} s'obtiennent de manière similaire.

Théorème 1.5.2. ***Théorème des probabilités totales.*** *Soit $X = (X_a, X_b)^t$ un vecteur aléatoire, alors*

$$\mathbb{P}_{X_a}(x_a) = \sum_{x_b} \mathbb{P}_{X_a/X_b=x_b}(x_a)\mathbb{P}(x_b); \quad f_{X_a}(x_a) = \int_{\mathbb{R}^{n-m}} f_{X_a/X_b=x_b}(x_a)f(x_b)dx_b.$$

Théorème 1.5.3. ***Théorème des probabilités composées.*** Soit $X = (X_a, X_b)^t$ un vecteur aléatoire (discret ou continu), alors

$$\mathbb{P}(x_1, ..., x_n) = \mathbb{P}(x_1)\mathbb{P}_{X_2/X_1=x_1}\mathbb{P}_{X_3/(x_1,x_2)}...\mathbb{P}_{X_n/(x_1,...,x_{n-1})}.$$

$$f_X(x_1, ..., x_n) = f(x_1)f_{X_2/X_1=x_1}f_{X_3/(x_1,x_2)}...f_{X_n/(x_1,...,x_{n-1})}.$$

Théorème 1.5.4. ***Théorème de Bayes .*** Soit $X = (X_a, X_b)^t$ un vecteur aléatoire de $(\mathbb{R}^n, \mathfrak{B}(R^n))$, on peut écrire que :

$$\mathbb{P}_{X_a}(x_a) = \frac{\mathbb{P}_{X_a/X_b=x_b}\mathbb{P}(x_b)}{\sum_{x_b}\mathbb{P}_{X_a/X_b=x_b}\mathbb{P}_{X_b}(x_b)}; f_{X_a}(x_a) = \frac{f_{X_a/X_b=x_b}(x_a)f_{X_b}(x_b)}{\int_{\mathbb{R}^{n-m}} f_{X_a/X_b=x_b}(x_a)f_{X_b}(x_b)dx_b}.$$

1.5.1 Indépendance

Définition 1.33. Soient $X = (X_1, ..., X_m)^t$ un vecteur aléatoire de dimension m et $Y = (Y_1, ..., Y_n)^t$, un vecteur aléatoire absolument continu. On dit que X et Y sont indépendants si pour tous les ensembles mesurables $A \in \mathfrak{B}_{\mathbb{R}^m}$ et $B \in \mathfrak{B}_{\mathbb{R}^n}$,

$$\mathbb{P}[X \in A, Y \in B] = \mathbb{P}[X \in A]\,\mathbb{P}[Y \in B].$$

Avec les notations de la définition précédente, soit le vecteur $X = (x_1, ..., x_m, x_{m+1}, ..., x_n)^t$. L'indépendance de X_a et X_b implique alors que :

$$F_X(x_a, x_b) = F_{X_a}(x_a)F_{X_b}(x_b). \tag{1.31}$$

Propriété 1.5.5. Pour tous les vecteurs $x \in \mathbb{R}^m$ et $y \in \mathbb{R}^n$. Nous retrouvons donc le même type de définition que celle donnée pour l'indépendance de deux variables aléatoires. Et comme pour les variables aléatoires, la réciproque est encore vraie ; si (1.32) est vraie, alors les vecteurs aléatoires X_a et X_b sont indépendants.

Définition 1.34. ***Indépendance mutuelle***. Soient $Y_1, ..., Y_l$ des vecteurs aléatoires de dimensions respectives $n_1, ..., n_l$ où $n_1, ..., n_l$ sont l enties naturels non nuls quelconques. On dit que ces vecteurs aléatoires sont mutuellement indépendants si, pour tous les ensembles mesurables $B_1 \in \mathfrak{B}(\mathbb{R}^{n_1}), ..., B_l \in \mathfrak{B}(\mathbb{R}^{n_l})$, on a :

$$\mathbb{P}[Y_1 \in B_1, ..., Y_l \in B_l] = \prod_{i}^{l} \mathbb{P}[Y_i \in B_i].$$

Définition 1.35. *Soient $X = (X_1, ..., X_m)^t$ un vecteur aléatoire discret de dimension m et $Y = (Y_1, ..., Y_n)^t$, un vecteur aléatoire discret de dimension n. On dit que X et Y sont indépendants si pour tous les ensembles mesurables A et B*

$$\mathbb{P}[X \in A, Y \in B] = \mathbb{P}[X \in A]\,\mathbb{P}[Y \in B].$$

Avec les notations de la définition précédente, soit le vecteur $X = (x_1, ..., x_m, x_{m+1}, ..., x_n)^t$. L'indépendance de X_a et X_b implique alors que :

$$\mathbb{P}_X(X_a = x_a, X_b = x_b) = \mathbb{P}(X_a = x_a)\,\mathbb{P}(X_b = x_b). \tag{1.32}$$

Théorème 1.5.6. ***Indépendance mutuelle de n variables aléatoires [3]***. *Soit* $X = (X_1, ..., X_n)^t$ un vecteur aléatoire (discret ou continu). Les variables aléatoires $X_1, X_2, ..., X_n$ seront dites mutuellement indépendantes (ce que l'on note aussi $X_1 \perp X_2 \perp ... \perp X_n$) si et seulement si

$$\mathrm{X}_1 \perp X_2 \perp ... \perp X_n \Leftrightarrow \begin{cases} \mathrm{P}_X(x_1, ..., x_n) = \mathbb{P}_{X_1}(x_1)\mathbb{P}_{X_2}(x_2)...\mathbb{P}_{X_n}(x_n) & \text{cas discret} \\ \mathrm{f}_X(x_1, ..., x_n) = f_{X_1}(x_1)f_{X_2}(x_2)...f_{X_n}(x_n) & \text{cas continu.} \end{cases}$$

1.5.2 Loi multinomiale

Exemple 1.5.1. On rappelle que la loi binomiale permet de modéliser le nombre de succès d'une expérience à 2 issues (succès ou échec) répétée n fois de façon indépendantes. La loi multinomiale généralise la loi binomiale. Plus précisément, suite à n répétitions indépendantes d'une expérience à k issues de probabilité respectives p_k, la loi multinomiale permet de modéliser le nombre de réalisations X_i de l'issue i. Supposons que l'on répète un nombre n fixé de fois une épreuve à l'issue de laquelle un et un seul des événements $A_1, ..., A_k$ est réalisé, ces répétitions étant indépendantes les unes des autres et les probabilités des A_i étant identiques d'une répétition à l'autre. Si l'on définit le vecteur $X = (X_1, ..., X_k)^t$ où X_i est le nombre de fois que l'événement A_i s'est réalisé, alors X suit une loi multinomiale de paramètre n et $p = (p_1, ..., p_k)^t$, et on note :

$$X \sim \mathcal{M}(n, p)$$

avec $p_i = \mathbb{P}(A_i) = \mathbb{P}[X_i \in A_i]$.

Définition 1.36. On dit que le vecteur $X = (X_1, ..., X_k)^t$ suit une loi multinomiale de paramètre $(n, p_1, ..., p_k)$ si pour $x = (x_1, ..., x_k)^t \in \mathbb{N}^k$ avec $\sum\limits_{i=1}^{k} x_i = n$, on a

$$\mathbb{P}\left(X_1 = x_1, ..., X_k = x_k\right) = \frac{n!}{x_1!...x_k!} p_1^{x_1}...p_k^{x_k}.$$

Exemple 1.5.2.

a. On effectue un sondage en choisissant au hasard n individus (avec remise) dans une population partagée en k catégories et on note X_i le nombre d'individus sélectionné dans la ième catégorie. Le vecteur $X = (X_1, ..., X_k)^t$ suit une loi multinomiale.

b. On retrouve la loi multinomiale dans diverses parties de la statistique telles que la régression logistique multiclasse ou le test du Chi-Deux.

c. On s'interesse à la couleur des yeux de n individus. Sur la base d'un très grand nombre d'individus, on a pu établir de façon très fiable la probabilité de chacune des couleurs possibles, à savoir $A_1 \equiv \quad bruns$, $A_2 \equiv \quad bleu$, $A_3 \equiv \quad noisette$ et $A_4 \equiv \quad vert$, avec $\mathbb{P}(A_1) = 0.37$, $\mathbb{P}(A_2) = 0.36$, $\mathbb{P}(A_1) = 0.16$ et $\mathbb{P}(A_1) = 0.11$, la variable couleur des yeux étant donc une variables catégorique nominale. Si l'on examine n individus pris au hasard dans cette population supposée suffisamment grande et que X_i désigne le nombre d'individus appartenant à la catégorie i, alors le vecteur $X = (X_1, X_2, X_3, X_4)^t$ suit une loi multinomiale.

Remarque 1.5.1. La loi multinomiale s'intéresse donc à une seule variable catégorique, mais il n' y aucune difficulté à prendre en compte plusieurs variables de manière simultanée ; il suffit de croiser entre elles les catégories de chacunes de ces variables. Le plus souvent, les probabilités de ces catégories croisées sont alors indicées en fonction de ces variables et présentées dans un tableau multidimensionnel.

Exemple 1.5.3. Supposons que l'on ait en plus déterminer la couleur des cheveux pour ces mêmes individus. On dispose donc des deux variables catégoriques nominales "couleur des yeux " et "couleur des cheveux", avec pour les cheveux $B_1 \equiv "noir"$, $B_2 \equiv "brun"$,$B_1 \equiv "roux"$ et $B_4 \equiv "blanc"$. Les probabilités $\mathbb{P}(A_i \cap B_j)$ des différentes combinaisons possibles sont données dans le tableau suivant :

i$/j$	1	2	3	4	P(A$_i$)
1	0.12	0.20	0.04	0.01	0.37
2	0.03	0.14	0.03	0.16	0.36
3	0.03	0.09	0.02	0.02	0.16
4	0.01	0.05	0.02	0.03	0.11
P(B$_j$)	0.19	0.48	0.11	0.22	1

On peut toujours utiliser le formalisme de la loi multinomiale en considérant que chacune des 16 catégories de cette loi est obtenue en croisant une des catégories des deux variables, avec $p_{ij} = \mathbb{P}(A_i \cap B_j)$ où $i, j = 1, ..., 4$.