Chapitre 3. Lois conjointes. Révision

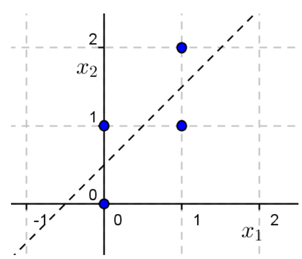
Contenu

1 Exemples introductifs

Dans une expérience aléatoire, on s'intéresse souvent à plusieurs quantités concernant les résultats de l'expérience.

Exemple 1 *On lance une pièce de monnaie 2 fois. Un ensemble fondamental pour cette expérience est* $\Omega = \{PP, PF, FP, FF\}$*, où* $P =$"*pile*" *et* $F =$"*face*"*. Soit* X_1 *le nombre de piles au* 1er *jet et* X_2 *le nombre total de piles. Dans ce cas, le couple* $X = (X_1, X_2)$ *est un* vecteur aléatoire *et* $|X(\Omega)| \leq |X_1(\Omega)| \cdot |X_2(\Omega)| = 2 \times 3 = 6$*.*

$$X(\Omega) = \{X(PP) = (1,2), X(PF) = (1,1), X(FP) = (0,1), X(FF) = (0,0)\}$$



Ou

Ω	X_1	X_2	*Probabilité*
PP	1	2	1/4
PF	1	1	1/4
FP	0	1	1/4
FF	0	0	1/4

Ou encore

X	$(0,0)$	$(0,1)$	$(1,1)$	$(1,2)$	*Total*
f_X	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

avec

X_1	0	1
P	1/2	1/2

X_2	0	1	2
P	1/4	2/4	1/4

qu'on met souvent sous forme d'un tableau croisé (ou tableau de contingence)

X_1 / X_2	0	1	f_{X_2}
0	1/4	0	1/4
1	1/4	1/4	1/2
2	0	1/4	1/4
f_{X_1}	1/2	1/2	1

Exemple 2 *On considère une urne contenant 2 boules rouges (R) , une boule blanche (B) et 2 boules vertes (V). On tire, sans remise, les boules l'une après l'autre jusqu'à l'apparition d'une rouge. Soit X le nombre de boules tirées et Y le nombre de boules blanches tirées.*

1. *Donner un ensemble fondamental décrivant cette expérience et un tableau de contingence pour le couple (X, Y).*
2. *Soit $D = X - Y$. Trouver la distribution de D.*

Réponse

1. *On peut prendre*

$$\Omega = \{R, BR, VR, BVR, VBR, VVR, BVVR, VBVR, VVBR\}.$$

De plus,

$$(X, Y)(\Omega) = \{1, 2, 3, 4\} \times \{0, 1\}.$$

Le tableau de contingence (fonction de masse conjointe) est alors

$Y \diagdown X$	1	2	3	4	f_Y
0	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{1}{15}$	0	$\frac{2}{3}$
1	0	$\frac{1}{10}$	$\frac{2}{15}$	$\frac{1}{10}$	$\frac{1}{3}$
f_X	$\frac{2}{5}$	$\frac{3}{10}$	$\frac{1}{5}$	$\frac{1}{10}$	1

2. *On voit que D est le nombre de boules tirées, sans tenir compte des boules blanches. Un calcul rapide donne*

D	1	2	3
P	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

Il est intéressant de noter que D se comporte comme s'il n'y avait pas de boules blanches dans l'urne (2 vertes et 2 rouges) car on voit que

$$\begin{aligned} P(D=1) &= P(R) = \frac{2}{4} = \frac{1}{2} \\ P(D=2) &= P(VR) = \frac{2}{4} \times \frac{2}{3} = \frac{1}{3} \\ P(D=3) &= P(VVR) = \frac{2}{4} \times \frac{1}{3} \times \frac{2}{2} = \frac{1}{6}. \end{aligned}$$

Exercice 1 *(Proposé par Hamming[1]) On choisit un nombre X entre 0 et 1. Puis on choisit un nombre réel Y_1 entre 0 et 1. Si $Y_1 > X$, on arrête, sinon on choisit un second nombre Y_2 entre 0 et 1. On continue jusqu'à obtention d'un nombre $Y_n > X$ $(n = 1, 2, \ldots)$. Si le jeu s'arrête à l'étape n, on vous donne n dollars. Quel devrait être le droit d'entrée pour que le jeu soit équitable ?*

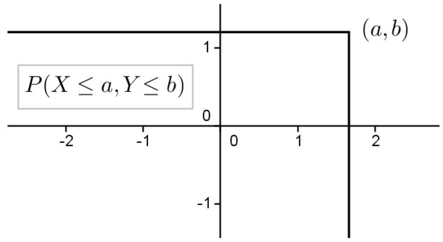
2 Fonction de répartition conjointe

Rappelons qu'une variable aléatoire n'est rien d'autre (en tout cas intuitivement) que l'association d'une valeur numérique à chaque résultat d'une expérience aléatoire. On a formalisé ceci en disant que c'est une fonction de l'ensemble Ω dans $\mathbb{R}$. Il arrive fréquemment qu'on associe plus d'une valeur numérique à une expérience aléatoire (voir les exemples introductifs). On parle dans ce cas de *vecteurs aléatoires*. La définition est similaire au cas d'une variable.

Dans le cas d'un couple (X, Y) de variables aléatoires, la fonction de répartition est définie par :

$$\begin{array}{lll} F_{(X,Y)} : \mathbb{R}^2 & \to & [0;1] \\ (x, y) & \to & F_{(X,Y)}(x, y) = P(X \leq x, Y \leq y) \end{array}$$

[1] Richard Hamming (1915-1998). Mathématicien américain, spécialisé en analyse numérique. Célèbre pour le code qui porte son nom.



Exemple 3 *Soit une urne (voir l'exemple 2)*

R	B	V
2	1	2

On tire (sans remise) jusqu"à obtention d'une boule rouge. Soit $X_1 =$ *nombre de boules tirées* ($X_1(\Omega) = \{1,2,3,4\}$)*,* $X_2 =$ *nombre de boules bleues tirées* ($X_1(\Omega) = \{0,1\}$)*. Soit* $X = (X_1, X_2)$*. On a* $|X(\Omega)| \leq 4 \times 2 = 8$ *(*$X(\Omega) \subset X_1(\Omega) \times X_2(\Omega)$*).*

$$\Omega = \{R, BR, VR, BVR, VBR, VVR, BVVR, VBVR, VVBR\}$$

$$X(\Omega) = \{(1,0), (2,1), (2,0), (3,1), (3,0), (4,1)\}$$

$$|X(\Omega)| = 6$$

X_1	1	2	3	4	$\sum$
p_{X_1}	$2/5$	$3/10$	$1/5$	$1/10$	1

X_2	0	1	$\sum$
p_{X_2}	$2/3$	$1/3$	1

Ω	X_1	X_2	*Probabilité*
R	1	0	$2/5$
BR	2	1	$1/5 \times 2/4$
VR	2	0	$2/5 \times 2/4$
BVR	3	1	$2/5 \times 2/4 \times 2/3$
VBR	3	1	$2/5 \times 1/4 \times 2/3$
VVR	3	0	$2/5 \times 1/4 \times 1/3 \times 2/2$
$BVVR$	4	1	$1/5 \times 2/4 \times 1/3 \times 2/2$
$VBVR$	4	1	$2/5 \times 1/4 \times 1/3 \times 2/2$
$VVBR$	4	1	$2/5 \times 1/4 \times 1/3 \times 2/2$

Plus généralement, dans le cas de n variables aléatoires simultanées $X_1, \ldots, X_n$ (définies sur le même ensemble fondamental Ω), la fonction de répartition est définie par

$$\begin{array}{lcl} F_{(X_1,\ldots,X_n)} : \mathbb{R}^n & \to & [0;1] \\ (x_1, \ldots, x_n) & \to & F_{(X_1,\ldots,X_n)}(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n) \end{array}$$

3 Cas discret

Les variables aléatoires $X_1, \ldots, X_n$ (définies sur le même ensemble fondamental Ω) sont discrètes. Soit $\mathbf{X} = (X_1, \ldots, X_n)$ et

$$\mathbf{X}(\Omega) = \prod_{j=1}^{n} \{x_{1j}, x_{2j} \ldots\}$$

La fonction de masse conjointe est définie par

$$p_{\mathbf{X}}(x_{1j}, x_{2j} \ldots, x_{nj}) = P(X_1 = x_{1j}, \ldots, X_n = x_{nj})$$

Dans le cas $n = 2$ (deux variables simultanées X et Y), on a

$$p_{(X,Y)}(x_i, y_j) = p_{ij} = P(X = x_i, Y = y_j)$$

et

3.1 Propriétés

1. $p_{ij} \geq 0$
2. $\sum_{i,j} p_{ij} = 1$
3. $P((X,Y) \in E) = \sum_{(x_i,y_j)\in E} p_{ij}$ pour tout événement E.
4. $F_{(X,Y)}(x_i, y_j) = \sum_{\substack{x \leq xi \\ y \leq y_j}} p(x_i, y_j)$.

Exemple 4 *On lance une pièce de monnaie 2 fois (voir l'exemple 1).*

$$\Omega = \{PP, PF, FP, FF\}.$$

Soit X = nombre de piles au premier jet : $X \sim Bin\left(n = 1, p = \frac{1}{2}\right)$ et Y = nombre total de piles : $Y \sim B\left(n = 2, p = \frac{1}{2}\right)$. On a

$$(X, Y)(\Omega) = \{(0,0), (0,1), (1,1), (1,2)\}$$

La fonction de masse conjointe est :

X / Y	0	1	p_Y
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
2	0	$\frac{1}{4}$	$\frac{1}{4}$
p_X	$\frac{1}{2}$	$\frac{1}{2}$	*Total* $= 1.$

Exemple 5 *Distribution multinomiale. Généralise la distribution binomiale. Supposons qu'une expérience aléatoire peut résulter en r ($r \geq 2$) événements deux-à-deux disjoints $E_1, \ldots, E_r$. On a donc une partition de l'ensemble fondamental* :

$$\Omega = E_1 \cup \ldots \cup E_r$$

On répète cette expérience n fois ($n \geq 1$) de manière indépendante et dans des conditions identiques. Posons

$$\begin{aligned}
N_1 &= \text{ nombre de réalisations de } E_1 \\
N_2 &= \text{ nombre de réalisations de } E_2 \\
&\vdots \\
N_r &= \text{ nombre de réalisations de } E_r
\end{aligned}$$

Alors la distribution conjointe de $N = (N_1, N_2, \ldots, N_r)$ est donnée par

$$\begin{aligned}
p(N_1 = n_1, N_2 = n_2, \ldots, N_r = n_r) = p(n_1, n_2, \ldots, n_r) &= \binom{n}{n_1, n_2, \ldots, n_r} p_1^{n_1} p_2^{n_2} \ldots p_r^{n_r} \\
&= \frac{n!}{n_1! n_2! \ldots n_r!} p_1^{n_1} p_2^{n_2} \ldots p_r^{n_r}
\end{aligned}$$

où $p_j^{n_j} = P(N_j = n_j)$, $j = 1, 2, \ldots, r$. Noter que $N_j \sim Bin(n, p_j)$ sont les distributions marginales.

Exemple 6 *Une application de l'exemple précédent est celle des histogrammes. Un histogramme (du point de vue probabiliste) est une réalisation d'une variable aléatoire donnée et donne une approximation de la densité (ou fonction de masse) de la variable.*

4 Cas continu

On se contentera d'examiner le cas de deux variables aléatoires simultanées X et Y. La fonction de densité conjointe du couple (X, Y) est la fonction définie par

$$P((X,Y) \in E) = \int\int_E f_{(X,Y)}(u,v)\,dvdu$$

pour tout événement E. En particulier, la fonction de répartition conjointe est définie par

$$F_{(X,Y)}(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{(X,Y)}(u,v)\,dvdu$$

En particulier,

$$P(a < X \leq b, c < Y \leq d) = F_{(X,Y)}(b,d) - F_{(X,Y)}(b,c) - F_{(X,Y)}(a,d) + F_{(X,Y)}(a,c)$$

La fonction de densité conjointe est aussi liée à la fonction de répartition conjointe par la relation

$$\frac{\partial^2}{\partial x \partial y} F_{(X,Y)}(x,y) = f_{(X,Y)}(x,y) \qquad pp$$

Toute fonction $f : \mathbb{R}^2 \to \mathbb{R}$ vérifiant :

1. $f \geq 0$
2. $\int\int_{\mathbb{R}^2} f_{(X,Y)}(u,v)\,dvdu = 1$

est une fonction de densité conjointe.

Remarque 1 *On peut avoir une combinaison X continue, Y discrète ou plus (X mixte, Y mixte).*

Exemple 7 *Soit un vecteur aléatoire (X, Y) de densité conjointe*

$$f(x,y) = \begin{cases} e^{-y} & si \quad 0 < x < 1 \text{ et } y > 0 \\ 0 & ailleurs \end{cases}$$

1. *Vérifier que $f(x,y)$ est bien une fonction de densité conjointe.*
2. *Trouver la fonction de répartition conjointe $F(x,y)$.*
3. *Calculer $P\left(X < \frac{1}{2}, Y > 1\right)$.*
4. *Calculer $P(X < Y)$.*
5. *Soit $Z = X + Y$. Trouver la loi de Z.*

Réponse

1. *On a évidemment $f \geq 0$ et*

$$\int\int_{\mathbb{R}^2} f(x,y)\,dxdy = \int_0^1\int_0^\infty e^{-y}dydx = \int_0^1 1dx = 1.$$

2. *Soit $(x,y) \in \mathbb{R}^2$. Pour déterminer la fonction de répartition conjointe $F(x,y)$, on examine plusieurs cas.*

   (a) *Si $x \leq 0$ ou $y \leq 0$. Ici, évidemment, $F(x,y) = 0$ puisque $f(x,y) = 0$ dans cette région :*

$$\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,dvdu = \int_{-\infty}^{x}\int_{-\infty}^{y} 0dvdu = 0.$$

(b) *Si* $0 < x < 1$ *et* $y > 0$, *on a*

$$\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,dvdu = \int_{0}^{x}\int_{0}^{y} e^{-v}dvdu = x\left(1-e^{-y}\right).$$

(c) *Si* $x \geq 1$ *et* $y > 0$, *on a*

$$\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,dvdu = \int_{0}^{1}\int_{0}^{y} e^{-v}dvdu = 1-e^{-y}.$$

Pour nous résumer,

$$F(x,y) = \begin{cases} 0 & si \quad x \leq 0 \text{ ou } y \leq 0 \\ x\left(1-e^{-y}\right) & si \quad 0 < x < 1 \text{ et } y > 0 \\ 1-e^{-y} & si \quad x \geq 1 \text{ et } y > 0. \end{cases}$$

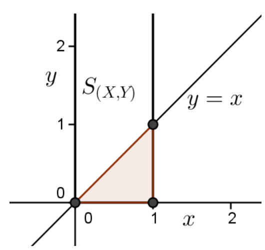
3. *On a :*

$$P\left(X < \frac{1}{2}, Y > 1\right) = \int_{-\infty}^{1/2}\int_{1}^{+\infty} f(u,v)\,dvdu$$
$$= \int_{0}^{1/2}\int_{1}^{+\infty} e^{-v}dvdu = \frac{1}{2}e^{-1}.$$

On peut aussi utiliser la fonction de répartition conjointe pour trouver ce résultat :

$$P\left(X < \frac{1}{2}, Y > 1\right) = F\left(\frac{1}{2}, +\infty\right) - F\left(\frac{1}{2}, 1\right)$$
$$= \frac{1}{2} - \frac{1}{2}\left(1-e^{-1}\right) = \frac{1}{2}e^{-1}.$$

4. *Ici, intégrons dans la région sous la bissectrice* $y = x$:



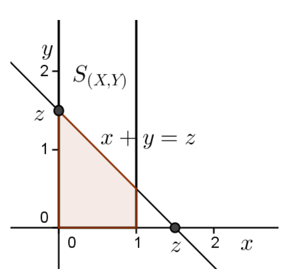
$$P(X \geq Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{u} f(u,v)\,dvdu = \int_{0}^{1}\int_{0}^{u} e^{-v}dvdu = e^{-1}$$

On tire

$$P(X < Y) = 1 - e^{-1}$$

5. *Soit* $z \in \mathbb{R}$. *La fonction de répartition de* $Z = X + Y$ *est donnée par*

$$F_Z(z) = P(X+Y \leq z) = \iint_{\{(x,y):x+y\leq z\}} f(x,y)\,dydx$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{z-x} f(u,y)\,dydx.$$



On a clairement $F_Z(z)$ *si* $z \leq 0$. *Pour* $z > 0$, *on considère deux cas.*
cas 1. $0 < z < 1$. *Alors*

$$F_Z(z) = \int_0^z \int_0^{z-x} e^{-y} dy dx = z + e^{-z} - 1.$$

cas 2. $z \geq 1$ *Alors,*

$$F_Z(z) = 1 - P(X + Y > z) = 1 - \int_0^1 \int_{z-x}^{+\infty} e^{-y} dy dx = 1 - e^{-z}(e - 1).$$

Pour nous résumer,

$$F_Z(z) = \begin{cases} 0 & si \quad z \leq 0 \\ z + e^{-z} - 1 & si \quad 0 < z < 1 \\ 1 - e^{-z}(e-1) & si \quad z \geq 1. \end{cases}$$

La densité de Z *est*

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \begin{cases} 0 & si \quad z \leq 0 \\ 1 - e^{-z} & si \quad 0 < z < 1 \\ e^{-z}(e-1). & si \quad z > 1. \end{cases}$$

Exemple 8 *Famille Farlie-Morgenstern (FM). Soit* $F_1(x)$ *et* $F_2(y)$ *deux fonctions de répartition (une variable) et soit une constante* $|\alpha| \leq 1$. *Alors*

$$G(x, y) = F_1(x) F_2(y) (1 + \alpha (1 - F_1(x)) (1 - F_2(y)))$$

est une fonction de répartition conjointe.

Exemple 9 *Reprenons l'exemple précédent avec* $X, Y \sim U[0; 1]$:

$$\begin{aligned} F_1(x) &= x, \qquad x \in [0; 1] \\ F_2(y) &= y, \qquad y \in [0; 1] \end{aligned}$$

Soit $(x, y) \in [0; 1]^2$. *Alors*

$$G(x, y) = xy (1 + \alpha (1 - x) (1 - y))$$

Exemple 10 *Loi binormale (ou normale bivariée). Elle est concentrée sur* $\mathbb{R}^2$ *et dépend de* 5 *paramètres* $\mu_X, \mu_Y, \sigma_X > 0, \sigma_Y > 0, -1 < \rho < 1$. *La densité conjointe est donnée par*

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)} \left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho \frac{x-\mu_X}{\sigma_X}\frac{y-\mu_Y}{\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right)\right]$$

5 Distributions marginales

5.1 Cas discret

Étant donné la fonction de masse conjointe $p_{(X,Y)}(x_i, y_j)$, les fonctions de masse marginales sont données par

$$\begin{aligned} p_X(x_i) &= \sum_j p_{(X,Y)}(x_i, y_j) \\ p_Y(y_j) &= \sum_i p_{(X,Y)}(x_i, y_j) \end{aligned}$$

Exemple 11 *Les lois marginales de la loi multinomiale sont les lois binomiales.*

5.2 Cas continu

Les fonctions de densité marginales sont données par

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

Exemple 12 *Les lois marginales de la loi multinormale sont les lois normales.*

Exemple 13 *Copule. C'est une distribution conjointe dont les lois marginales sont uniformes. Par exemple, si on considère la famille Farlie-Morgenstern où $X, Y \sim U[0;1]$. On la note $C(x,y)$. La fonction de densité conjointe est donnée par $c(x,y) = \frac{\partial^2}{\partial x \partial y} C(x,y) \geq 0$. Intérêt. Soit X, Y deux variables aléatoires continues avec fonctions de répartition respectives $F_X(x)$ et $V = F_Y(y)$. On sait que $U \sim U[0;1]$ et $V \sim U[0;1]$. Considérons la fonction de répartition conjointe ($C(u,v)$ copule donné)*

$$F(x,y) = C(F_X(x), F_Y(y))$$

On a

$$\begin{aligned} C(F_X(x), 1) &= P(U \leq F_X(x), V \leq 1) \\ &= P(U \leq F_X(x)) \\ &= F_X(x) \end{aligned}$$

De même,

$$C(1, F_Y(y)) = F_Y(y)$$

Ainsi, les lois marginales de $F(x,y)$ sont $F_X(x)$ et $F_Y(y)$.

Trouvons la densité conjointe.

$$\begin{aligned} f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y) &= \frac{\partial^2}{\partial x \partial y} C(F_X(x), F_Y(y)) \\ &= c(F_X(x), F_Y(y))\, f_X(x)\, f_Y(y) \end{aligned}$$

On voit que le terme $c(F_X(x), F_Y(y))$ capture la dépendance entre X et Y.

6 Indépendance

Les variables X et Y sont indépendantes si leur loi conjointe est égale au produit de leurs lois marginales :

$$F_{(X,Y)}(x,y) = F_X(x)\, F_Y(y)$$
$$f_{(X,Y)}(x,y) = f_X(x)\, f_Y(y)$$
$$p_{(X,Y)}(x_i, y_j) = p_X(x_i)\, p_Y(y_j)$$

On a le résultat suivant.

Théorème 1 *Les affirmations suivantes sont équivalentes :*

1. *X et Y sont indépendants.*
2. *$F_{(X,Y)}(x,y) = F_X(x)\, F_Y(y)$ pour tous x et y.*
3. *$f_{(X,Y)}(x,y) = f_X(x)\, f_Y(y)$ pour tous x et y. Même chose dans le cas discret.*

Théorème 2 *Soit X et Y deux variables aléatoires indépendantes. Alors les variables $h(X)$ et $g(Y)$ sont indépendantes pour toutes fonctions $h(x)$ et $g(y)$.*

Exemple 14 *Soit X et Y de lois uniformes sur $[0;1]^2$. Alors X et Y sont indépendantes et distribuées chacune uniformément sur $[0;1]$. Il est à noter que la forme du domaine est fondamentale.*

Exemple 15 *Prenons (X,Y) uniforme sur le quart de cercle unité. Alors X et Y sont dépendantes.*

Exemple 16 *Famille Farlie-Morgenstern. Si $\alpha = 0$, les variables sont indépendantes.*

7 Distributions conditionnelles

Étant donné la densité conjointe $f_{(X,Y)}(x,y)$, on définit les densités conditionnelles

$$f_{Y|X}(y|x) = \frac{f_{(X,Y)}(x,y)}{f_X(x)}$$

pour tout x fixé tel que $f_X(x) \neq 0$ (on considère x comme un paramètre, la variable étant y) et

$$f_{X|Y}(y|x) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)}$$

pour tout y fixé tel que $f_Y(y) \neq 0$ (on considère y comme un paramètre, la variable étant x).

Dans le cas discret, on a les définitions des fonctions de masse conditionnelles

$$p_{Y|X}(y_j|x_i) = \frac{p_{(X,Y)}(x_i,y_j)}{p_X(x_i)} \qquad \text{avec } p_X(x_i) \neq 0$$

et

$$p_{X|Y}(x_i|y_j) = \frac{p_{(X,Y)}(x_i,y_j)}{p_Y(y_j)} \qquad \text{avec } p_Y(y_j) \neq 0$$

Théorème 3 *$f_{Y|X}(y \mid x)$ est une densité (resp. fonction de masse) pour tout x fixé.*

Remarque 2 *La formule immédiate suivante est très utile*

$$\begin{aligned} f_{(X,Y)}(x,y) &= f_{Y|X}(y|x)\, f_X(x) = f_{X|Y}(y|x)\, f_Y(y) \\ p_{(X,Y)}(x_i,y_j) &= p_{Y|X}(y_j|x_i)\, p_X(x_i) = p_{X|Y}(x_i|y_j)\, p_Y(y_j) \end{aligned}$$

Remarque 3 *Le théorème des probabilités totales s'applique aussi bien au cas discret qu'au cas continu :*

$$f_Y(y) = \begin{cases} \int_{-\infty}^{+\infty} f_{Y|X}(y \mid x)\, f_X(x)\, dx & si \quad X \ continue \\ \\ \sum_x p_{Y|X}(y \mid x)\, p_X(x) & si \quad X \ discrète \end{cases}$$

Exemple 17 *Considérons*

$Y \setminus X$	0	1	p_Y	$p_{Y\|X}(y \mid 0)$	$p_{Y\|X}(y \mid 1)$
0	1/4	0	1/4	$\frac{1}{2} = \frac{1/4}{1/2}$	0
1	1/4	1/4	1/2	$\frac{1}{2} = \frac{1/4}{1/2}$	$\frac{1}{2}$
2	0	1/4	1/4	0	$\frac{1}{2}$
p_X	1/2	1/2			
$p_{X\|Y}(x \mid 0)$	$1 = \frac{1/4}{1/4}$	$0 = \frac{0}{1/4}$			
$p_{X\|Y}(x \mid 1)$	$\frac{1}{2}$	$\frac{1}{2}$			
$p_{X\|Y}(x \mid 2)$	0	1			

Noter que

$$E\left(Y \mid X=0\right) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}$$
$$E\left(Y \mid X=1\right) = 0 \times 0 + 1 \times \frac{1}{2} + 2 \times \frac{1}{2} = \frac{3}{2}$$

et que

$$E\left(X \mid Y=0\right) = 0 \times 1 + 1 \times 0 = 0$$
$$E\left(X \mid Y=1\right) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}$$
$$E\left(X \mid Y=2\right) = 0 \times 0 + \frac{1}{2} + 1 \times 1 = 1$$

Ce sont des espérances conditionnelles. De même, on peut chercher les variances conditionnelles.

Exemple 18 *Le nombre $N\left(t\right)$ de clients qui arrivent à un service durant une période fixe de durée t suit une loi de Poisson de paramètre λt. Le temps de service T pour un client suit une loi exponentielle de paramètre α.*

1. *Trouver la distribution du nombre de clients qui arrivent durant le temps de service d'un client donné.*
2. *Espérance et variance.*

Réponse

1. *On utilise la formule des probabilités totales*

$$\begin{aligned} P\left(N\left(T\right)=n\right) &= \int_0^\infty P\left(N\left(T\right)=n \mid T=t\right) f_T\left(t\right) dt \\ &= \int_0^\infty P\left(N\left(t\right)=n\right) f_T\left(t\right) dt \\ &= \int_0^\infty \left(\frac{\left(\lambda t\right)^n}{n!} e^{-\lambda t}\right)\left(\alpha e^{-\alpha t}\right) dt \\ &= \frac{\alpha \lambda^n}{n!} \int_0^\infty t^n e^{-\left(\lambda+\alpha\right)t} dt \\ &= \frac{\alpha \lambda^n}{n!} \frac{1}{\left(\lambda+\alpha\right)^{n+1}} \int_0^\infty u^n e^{-u} du \\ &= \frac{\alpha \lambda^n}{n!} \frac{1}{\left(\lambda+\alpha\right)^{n+1}} \Gamma\left(n+1\right) \\ &= \frac{\alpha}{\alpha+\lambda}\left(\frac{\lambda}{\alpha+\lambda}\right)^n \end{aligned}$$

pour $n=0,1,\ldots$ On reconnaît une variable géométrique décalée à gauche vers 0 (compte le nombre d'échecs avant le premier succès. Succès = instant de fin du service du client). En d'autres termes, $N\left(T\right)+1$ est géométrique de paramètre $p = \frac{\alpha}{\alpha+\lambda}$.

2. *On a $E\left(N\left(T\right)+1\right) = \frac{1}{p} = \frac{\alpha+\lambda}{\alpha} = 1 + \frac{\lambda}{\alpha}$. On en tire $E\left(N\left(T\right)\right) = \frac{\lambda}{\alpha}$. Surprenant ? La variance est donnée par*

$$VAR\left(N\left(T\right)\right) = VAR\left(N\left(T\right)+1\right) = \frac{q}{p^2} = \frac{\lambda}{\alpha}\left(1+\frac{\lambda}{\alpha}\right)$$

Exemple 19 *Loi binormale*

$$f_{(X,Y)}\left(x,y\right) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2\left(1-\rho^2\right)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\frac{x-\mu_X}{\sigma_X}\frac{x-\mu_Y}{\sigma_Y} + \left(\frac{x-\mu_Y}{\sigma_Y}\right)^2\right)\right]$$

Alors, un petit calcul montre que $Y|X = x \sim N\left(\mu_{Y|x} = \mu_Y + \rho(x - \mu_X)\frac{\sigma_Y}{\sigma_X}, \sigma^2_{Y|x} = \sigma^2_Y\left(1 - \rho^2\right)\right)$*. Noter que* $\mu_{Y|x}$ *fournit l'équation de l'estimateur linéaire de* Y *en* X *et que* $\sigma^2_{Y|x} \leq \sigma^2_Y$ *(on a égalité si* $\rho = \pm 1$ *: corrélation totale, c-à-d si on a la relation* $Y = aX + b$ *avec probabilité égale à* 1*).*

Exemple 20 *On considère* X_1 *uniforme sur* $]0;1[$ *et* X_2 *uniforme sur* $]0;X_1[$ *(on choisit un nombre* X_1 *au hasard entre* 0 *et* 1 *puis on choisit un nombre* X_2 *au hasard entre* 0 *et* X_1 *: expérience en deux étapes (le second choix dépend du premier), ou séquentielle. On demande :*

1. $f_{X_2|X_1}(x_2 \mid x_1)$.
2. *La loi conjointe.*
3. *La marginale de* X_2.
4. *L'espérance de* X_2.
5. *Généraliser. On choisit un* 3*ème nombre* X_3 *entre* 0 *et* X_2*, etc. (on choisit un nombre* X_n *entre* 0 *et* X_{n-1}*). Loi de* X_n *? Espérance et variance.*

Réponse

1. *La variable* $X_2 \mid X_1 = x_1$ *est évidemment uniforme sur* $]0;x_1[$ $(0 < x_1 < 1$*, fixé). En d'autres termes,*
$$f_{X_2|X_1}(x_2 \mid x_1) = \frac{1}{x_1} \text{ pour } 0 < x_2 < x_1 < 1.$$
2. *La densité conjointe est donnée par*
$$f_{(X_1,X_2)}(x_1, x_2) = f_{X_2|X_1}(x_2 \mid x_1) f_{X_1}(x_1) = \frac{1}{x_1} \text{ pour } 0 < x_2 < x_1 < 1.$$
3. *Soit* x_2 *tel que* $0 < x_2 < 1$*. Alors*
$$\begin{aligned} f_{X_2}(x_2) &= \int_{\mathbb{R}} f_{(X_1,X_2)}(x_1, x_2)\, dx_1 = \int_{x_2}^1 \frac{dx_1}{x_1} \\ &= -\ln x_2 \end{aligned}$$
Et
$$F_{X_2}(x_2) = \int_0^{x_2} -\ln u du = x_2 - x_2 \ln x_2.$$
4. *On a*
$$E(X_2) = -\int_0^1 x_2 \ln x_2 dx_2 = \frac{1}{4},$$
ce qui semble intuitivement clair. De même,
$$E\left(X_2^2\right) = -\int_0^1 x_2^2 \ln x_2 dx_2 = \frac{1}{9}$$
et donc
$$VAR(X_2) = \frac{1}{9} - \left(\frac{1}{4}\right)^2 = \frac{7}{144}$$
5. *Généralisation. Une petite induction donne*
$$f_{X_n}(x_n) = \frac{(-1)^{n-1}}{(n-1)!}(\ln x_n)^{n-1} \text{ pour } 0 < x_n < 1,\ n \geq 1.$$
et
$$E(X_n) = \frac{1}{2^{n+1}}.$$

Exemple 21 *Méthode du rejet. Soit donné une fonction de densité f continue sur un intervalle $]a;b[$ (nulle en dehors de l'intervalle).Soit donné une seconde fonction $M(x) \geq f(x)$ sur $]a;b[$ et définissons la densité*

$$m(x) = \frac{M(x)}{\int_a^b M(x)\,dx}$$

(on supposera l'intégrale convergente si l'intervalle d'intégration est infini). Par exemple, si $]a;b[$ est fini, on peut prendre $m(x) = \frac{1}{b-a}$ (loi uniforme). L'idée est de choisir M de façon à générer des variables aléatoires issues de la loi m.

Algorithme.

1. Générer $T \sim m$

2. Générer $U \sim [0;1]$ indépendante de T.

3. si $M(T) \times U \leq f(T)$, prendre $X = T$ (accepter T). Sinon, rejet. Recommencer (1).

Voici une preuve heuristique que X suit la loi f. On a

$$\begin{aligned}
P(x \leq X \leq x+dx) &= P(x \leq T \leq x+dx \mid accepte) \\
&= \frac{P(x \leq T \leq x+dx \text{ et } accepte)}{P(accepte)} \\
&= \frac{P(accepte \mid x \leq X \leq x+dx)\,P(x \leq X \leq x+dx)}{P(accepte)}
\end{aligned}$$

Maintenant,

$$P(accepte \mid x \leq X \leq x+dx) = P\left(U \leq \frac{f(x)}{M(x)}\right) = \frac{f(x)}{M(x)}$$

ce qui entraîne

$$P(accepte \mid x \leq X \leq x+dx)\,P(x \leq X \leq x+dx) = \frac{f(x)}{M(x)} m(x)\,dx = \frac{f(x)}{\int_a^b M(x)\,dx} dx$$

Et (probabilités totales)

$$P(accepte) = P\left(U \leq \frac{f(T)}{M(T)}\right) = \int_a^b \frac{f(t)}{M(t)} m(t)\,dt = \frac{1}{\int_a^b M(t)\,dt}$$

Finalement,

$$P(x \leq X \leq x+dx) = \frac{f(x)}{\int_a^b M(x)\,dx} dx \times \int_a^b M(t)\,dt = f(x)$$

Exemple 22 *Inférence bayésienne. Soit donné une pièce de monnaie et on veut déterminer $p = P(Pile)$. On lance la pièce n fois et on examine la variable X = nombre de piles obtenus. Que nous apprend cette expérience ? Posons $\Theta = P(Pile)$. On résume notre connaissance sur Θ avant l'expérience par une densité concentrée sur $[0;1]$ (on parle de randomisation) : c'est la densité a priori. Si on ne sait rien sur Θ, on peut choisir $\Theta \sim U[0;1]$:*

$$f_\Theta(\theta) = 1 \text{ si } 0 \leq \theta \leq 1$$

Maintenant, on observe Θ pour obtenir une distribution a posteriori. Sachant θ, on a $X \mid \theta \sim Bin(n, p = \theta)$:

$$f_{X|\Theta}(x \mid \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, \qquad x = 0, 1, \ldots, n.$$

La distribution conjointe de X et Θ est alors

$$\begin{aligned} f_{(X,\Theta)}(\theta, x) &= f_{X|\Theta}(x \mid \theta) f_{\Theta}(\theta) \\ &= \binom{n}{x} \theta^{x} (1-\theta)^{n-x} \\ &= \frac{n!}{x!(n-x)!} \theta^{x} (1-\theta)^{n-x} \\ &= \frac{\Gamma(n+1)}{\Gamma(x+1)\Gamma(n+1-x)} \theta^{x} (1-\theta)^{n-x}, \qquad x = 0, 1, \ldots, n \end{aligned}$$

Intégrons par rapport à θ :

$$\begin{aligned} f_X(x) &= \frac{\Gamma(n+1)}{\Gamma(x+1)\Gamma(n+1-x)} \int_0^1 \theta^{x} (1-\theta)^{n-x} \, d\theta \\ &= \frac{\Gamma(n+1)}{\Gamma(x+1)\Gamma(n+1-x)} \frac{\Gamma(x+1)\Gamma(n+1-x)}{\Gamma(n+2)} \\ &= \frac{1}{n+1}, \qquad x = 0, 1, \ldots, n. \end{aligned}$$

On a utilisé la loi Béta pour ce résultat. Ainsi, la loi a priori de X est la loi uniforme. Maintenant, si on observe $X = x$,

$$f_{\Theta|X}(\theta \mid x) = \frac{f_{(X,\Theta)}(\theta, x)}{f_{\mathbb{X}}(x)} = \frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n+1-x)} \theta^{x} (1-\theta)^{n-x}$$

après un petit calcul. On a donc $\Theta \mid X \sim Beta(a = x+1, b = n+1-x)$. Ainsi, si par exemple on a obtenu $x = 13$ piles sur $n = 20$ jet, le maximum de la densité est donné pour la valeur approximative $p = 0.6406$.

8 Fonctions de plusieurs variables aléatoires

On se contentera (la plupart du temps) de fonctions de deux variables.

8.1 Somme de deux variables aléatoires

C'est un cas très fréquent. Si X et Y sont deux variables discrètes concentrées sur les entiers, la fonction de masse de $Z = X + Y$ est alors

$$p_Z(z) = P(X + Y = z) = \sum_{x=-\infty}^{\infty} p_{(X,Y)}(x, z - x)$$

Si, de plus, X et Y sont indépendantes, alors

$$p_Z(z) = \sum_{x=-\infty}^{\infty} p_X(x) p_Y(z - x)$$

On écrit

$$p_Z(z) = (p_X * p_Y)(z)$$

et on dit que la loi de $Z = X + Y$ est le produit de convolution des lois de X et de Y. La plupart du temps, les variables X et Y sont positives et le résulat devient :

$$p_Z(z) = (p_X * p_Y)(z) = \sum_{x=0}^{z} p_X(x) p_Y(z - x), \qquad z = 0, 1, \ldots$$

Dans le cas continu, un calcul similaire donne

$$f_Z(z) = (f_X * f_Y)(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx$$

Si, de plus, les variables sont positives,

$$f_Z(z) = \int_0^z f_X(x) f_Y(z-x)\, dx \qquad z \geq 0$$

Exemple 23 *Loi binomiale. Soient $X \sim Bin(n,p)$ et $Y \sim Bin(m,p)$ indépendantes. Si on interprète X comme une succession de n essais indépendants de Bernoulli suivi de m (Y) autres essais, alors $X+Y$ est une succession de $n+m$ essais de Bernoulli indépendants et par conséquent $X+Y \sim Bin(n+m,p)$. Ceci se vérifie par la convolution*

$$\begin{aligned}
p_{X+Y}(k) &= p_X(k) * p_Y(k) \\
&= \sum_{i=0}^{k} \binom{n}{i} p^i q^{n-i} \binom{m}{k-i} p^{k-i} q^{m-(k-i)} \\
&= p^k q^{n+m-k} \sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i} \\
&= \binom{n+m}{k} p^k q^{n+m-k}
\end{aligned}$$

pour $k = 0, 1, \ldots, n+m$.

Exemple 24 *Loi de Poisson. $X \sim Poi(\lambda)$, $Y \sim Poi(\lambda)$. On peut interpréter $X+Y$ comme une instance du processus de Poisson avec $t = 2$. On obtient $X+Y \curvearrowright Poi(2\lambda)$. Ceci est un cas particulier de la situation suivante.*

Exemple 25 *Loi de Poisson. $X \sim Poi(\lambda)$, $Y \sim Poi(\mu)$. On a*

$$\begin{aligned}
p_{X+Y}(k) = p_X(k) * p_Y(k) &= \sum_{i=0}^{k} \left(\frac{\lambda^i}{i!} e^{-\lambda}\right) \left(\frac{\mu^{k-i}}{(k-i)!} e^{-\mu}\right) \\
&= \frac{1}{k!} e^{-(\lambda+\mu)} \sum_{i=0}^{k} \binom{k}{i} \lambda^i \mu^{k-i} \\
&= \frac{(\lambda+\mu)^k}{k!} e^{-(\lambda+\mu)},\ k = 0.1, 2, \ldots
\end{aligned}$$

On voit que $X+Y \sim Poi(\lambda+\mu)$.

Exemple 26 *Un cas intéressant se présente quand nous avons une somme aléatoire de variables aléatoires. En d'autres termes, si X_1, X_2, … sont indépendantes de même loi (discrète ou continue) $f(x)$ et si N est une variable aléatoire discrète concentrée sur les entiers positifs et indépendante des X_i, quelle est la loi de $Z_N = X_1 + X_2 + \ldots + X_N$? La réponse est donnée en utilisant le théorème des probabilités totales (on conditionne sur N) :*

$$f_{Z_N}(x) = \sum_{k=1}^{\infty} f^{k*}(x)\, p_k$$

où on a mis $p_k = P(N = k)$, $k = 1, 2, \ldots$

Exemple 27 *Loi uniforme. Soit X et Y de lois uniformes (et indépendantes) sur l'intervalle $]0;1[$. La densité de leur somme est*

$$f_{X+Y}(z) = f_X(z) * f_Y(z) = \int_0^z f_X(x) f_Y(z-x)\,dx$$
$$= \begin{cases} 0 & si \quad z \le 0 \ ou\ z \ge 2 \\ \int_0^z dx & si \quad 0 < z < 1 \\ \int_{z-1}^1 dx & si \quad 1 \le z < 2 \end{cases}$$
$$= \begin{cases} 0 & si \quad z \le 0 \ ou\ z \ge 2 \\ z & si \quad 0 < z < 1 \\ 2-z & si \quad 1 \le z < 2 \end{cases} = \begin{cases} 1-|1-z| & si \quad 0 \le z < 2 \\ \\ 0 & ailleurs \end{cases}$$

Cette distribution est dite triangulaire.

8.2 Cas général (X et Y continues)

On a le théorème suivant (issu du cours de calcul). Considérons une transformation

$$\begin{array}{ccc} T : \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x,y) & \to & (u,v) = (g(x,y), h(x,y)) \end{array}$$

On supposera T inversible (bijective) et g, h continument dérivables. Soit la fonction réciproque de T :

$$\begin{array}{ccc} T^{-1} : \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (u,v) & \to & (x,y) = \left(\overline{g}(u,v), \overline{h}(u,v)\right) \end{array}$$

et supposons (jacobien)

$$J(x,y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} \ne 0$$

Alors la fonction de densité conjointe du couple $(U,V) = T(X,Y)$ est donnée par

$$f_{(U,V)}(u,v) = f_{(X,Y)}\left(\overline{g}(u,v), \overline{h}(u,v)\right) \left|J^{-1}\left(\overline{g}(u,v), \overline{h}(u,v)\right)\right|$$

Ce résultat peut être généralisé à plusieurs variables aléatoires.

Exercice 2 *Soit $X \sim U\,]0;1[$ et $Y \sim U\,]0;1[$ et supposons que X et Y sont indépendantes. Posons*

$$U = \sqrt{2\ln(1/X)}\cos(2\pi Y)$$
$$V = \sqrt{2\ln(1/X)}\sin(2\pi Y)$$

Montrer que $U \sim N(0;1)$, $V \sim N(0;1)$ et que ces deux variables sont indépendantes. Ceci donne donc une méthode pour simuler une variable normalement distribuée.

9 Statistiques d'ordre

Étant donné n variables aléatoires indépendantes identiquement distribuées (iid) $X_1, \ldots, X_n$ de loi (densité) $f(x)$ (et de fonction de répartition $F(x)$), on les ordonne par ordre croissant de manière à obtenir les variables

$$X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$$

On s'intéresse à la loi conjointe des $X_{(i)}$, aux lois marginales (notamment à celle de $X_{(1)}$ et de $X_{(n)}$) ainsi qu'à celle de l'étendue $R = X_{(n)} - X_{(1)}$. Il est connu (cf cours de proba) que la loi conjointe est donnée par

$$f\left(x_{(1)}, x_{(2)}, \ldots, x_{(n)}\right) = n! f\left(x_{(1)}\right) f\left(x_{(2)}\right) \ldots f\left(x_{(n)}\right)$$

avec $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$. La densité de $X_{(k)}$ est

$$f_k(x) = n\binom{n-1}{k-1} F^{k-1}(x)\left(1 - F(x)\right)^{n-k} f(x), \qquad k = 1, \ldots, n$$

En particulier,

$$f_1(x) = n\left(1 - F(x)\right)^{n-1} f(x)$$
$$f_n(x) = nF^{n-1}(x) f(x)$$

sont respectivement la densité de $X_{(1)} = \min\left(X_{(1)}, X_{(2)}, \ldots, X_{(n)}\right)$ et $X_{(n)} = \max\left(X_{(1)}, X_{(2)}, \ldots, X_{(n)}\right)$.
On montre que la loi de l'étendue $R = X_{(n)} - X_{(1)}$ est donnée par

$$f_R(x) = \int_{\mathbb{R}} f(y + x, y)\, dy$$

où

$$f(u, v) = n(n-1) f(u) f(v) \left(F(u) - F(v)\right)^{n-2}, u \leq v$$

est la densité conjointe du couple $\left(X_{(1)}, X_{(n)}\right)$.