Distributions de plusieurs variables

Mathématiques Générales B
Université de Genève

Sylvain Sardy

8 mai 2008

1. Distributions conjointes

Comment généraliser les fonctions de probabilité et de densité à plus d'une variable aléatoire ?

Variables aléatoires discrètes :

Considérons 2 variables discrètes : $X =$ utilité des mathématiques et $Y =$ branche d'étude.

X $\backslash$ Y	Pharma	SdIT	Bio	Chimie	Total
Math1	5	2	16	8	31
Math2	19	4	24	12	59
Math3	2	2	6	4	14
Total	26	8	46	24	104

Tableau de contingence (2007)

X \ Y	Pharma	SdlT	Bio	Chimie	Total
Math1	0.05	0.02	0.15	0.08	0.30
Math2	0.18	0.04	0.23	0.12	0.57
Math3	0.02	0.02	0.06	0.04	0.13
Total	0.25	0.08	0.44	0.23	1

Tableau de probabilité

La probabilité conjointe est simplement donnée par un tableau de probabilités, ou

$$\mathrm{P}(\mathrm{X}=\mathrm{i},\mathrm{Y}=\mathrm{j})=\mathrm{p}_{\mathrm{ij}} \quad \text{pour tout} \quad (\mathrm{i},\mathrm{j})$$

pour deux variables. Pour trois variables, il faut définir :

$$\mathrm{P}(\mathrm{X}=\mathrm{i},\mathrm{Y}=\mathrm{j},\mathrm{Z}=\mathrm{k})=\mathrm{p}_{\mathrm{ijk}} \quad \text{pour tout} \quad (\mathrm{i},\mathrm{j},\mathrm{k}).$$

Variables aléatoires continues : deux variables aléatoires $X =$ taille et $Y =$ poids ont une **fonction de densité conjointe** si

$$\mathrm{P}((\mathrm{X},\mathrm{Y}) \in \mathrm{A}) = \int\int_{\mathrm{A}} \mathrm{f}(\mathrm{x},\mathrm{y})\ \mathrm{dx}\ \mathrm{dy},$$

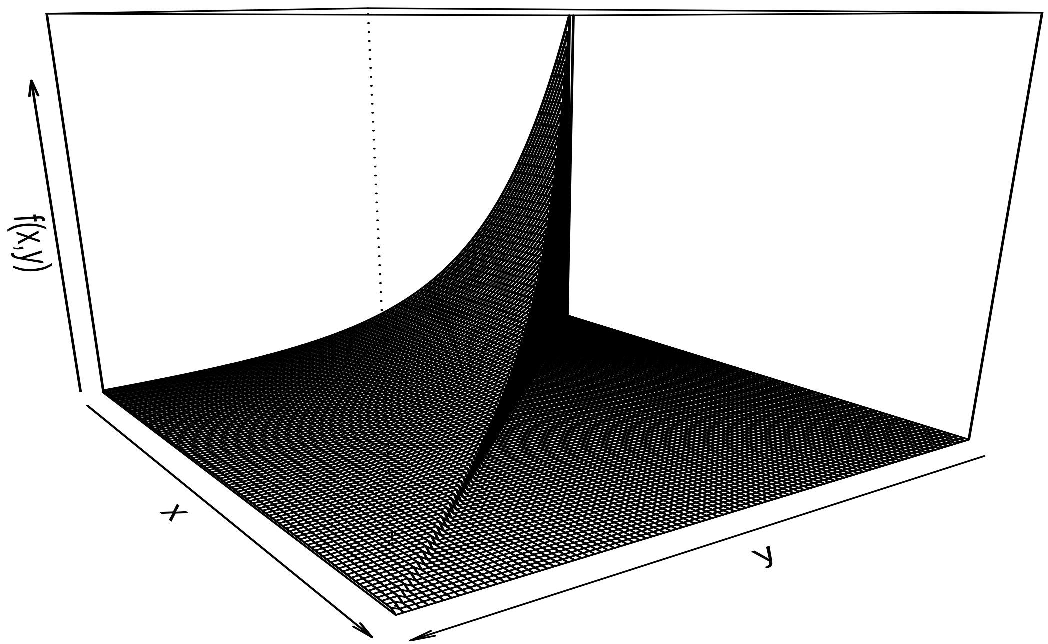
où $f(x,y) \geqslant 0$ et $\int\int f(x,y)dx\ dy = 1$.

Exemple :

$$f(x,y) = \begin{cases} \exp(-y) & 0 < x < y < \infty \\ 0 & \text{sinon} \end{cases}$$

Est-ce bien une fonction de densité ?

Exemple : Distribution uniforme bivariée sur un carré, un disque, ...



Fonction de densité à deux variables.

Il est aussi possible de définir une **fonction de répartition conjointe**

$$F(x, y) = \mathrm{P}(\mathrm{X} \leqslant \mathrm{x}, \mathrm{Y} \leqslant \mathrm{y})$$

pour deux variables. Il est facile de généraliser à $n \geqslant 2$ variables.

La fonction de densité conjointe s'obtient de la fonction de répartition en différenciant

$$\frac{\partial^2 F}{\partial x \partial y} = f$$

pour $n = 2$.

Exemple : On tire deux boules sans remise d'une urne qui contient 8 Rouge, 6 Bleue et 4 Verte. Soit $X =$ le nombre de boules Rouge et $Y =$ le nombre de boules Bleue. Trouver la distribution conjointe de X et Y.

X \ Y	0	1	2
0	$\frac{6}{153}$	$\frac{24}{153}$	$\frac{15}{153}$
1	$\frac{32}{153}$	$\frac{48}{153}$	0
2	$\frac{28}{153}$	0	0

Exemple : Soit deux variables aléatoires X et Y de densité $f(x,y) = c(x+y)$ sur $[0,1]\times[0,1]$. (1) Que vaut c ? (2) Que vaut $\mathrm{P(X < 1/2)}$? (et $\mathrm{P(X \leqslant 1/2)}$?) (3) Que vaut $\mathrm{P(X+Y<1)}$?

(1)

(2) $\mathrm{P(X<1/2) = P(X<1/2, Y\in[0,1]) = \int_0^{1/2}\int_0^1 (x+y)\ dy\ dx =}$

(3) $\mathrm{P(X+Y<1) = P(X<1-Y, Y\in[0,1]) = \int_0^1\int_0^{1-y} (x+y)\ dx\ dy =}$

2. Distributions marginales

Comment trouver les **distributions marginales** de X et de Y à partir de la distribution conjointe de (X, Y) ?

Cas discret

$$\mathrm{P}(\mathrm{X}=\mathrm{x})=\sum_{\mathrm{y}} \mathrm{P}(\mathrm{X}=\mathrm{x}, \mathrm{Y}=\mathrm{y})$$

est la distribution marginale de X.

$$\mathrm{P}(\mathrm{Y}=\mathrm{y})=\sum_{\mathrm{x}} \mathrm{P}(\mathrm{X}=\mathrm{x}, \mathrm{Y}=\mathrm{y})$$

est la distribution marginale de Y.

Exemple :

$X \setminus Y$	Pharma	SdIT	Bio	Chimie	$P(X = x)$
Math1	0.05	0.02	0.15	0.08	0.30
Math2	0.18	0.04	0.23	0.12	0.57
Math3	0.02	0.02	0.06	0.04	0.13
$P(Y = y)$	0.25	0.08	0.44	0.23	1

Tableau de probabilité

Exemple :

$X \setminus Y$	0	1	2	$P(X = x)$
0	$\frac{6}{153}$	$\frac{24}{153}$	$\frac{15}{153}$	$\frac{45}{153}$
1	$\frac{32}{153}$	$\frac{48}{153}$	0	$\frac{80}{153}$
2	$\frac{28}{153}$	0	0	$\ldots$
$P(Y = y)$	$\ldots$	$\frac{72}{153}$	$\ldots$	$\ldots$

Cas continu

$$f_X(x) = \int f(x, y)dy$$

est la distribution marginale de X.

$$f_Y(y) = \int f(x, y)dx$$

est la distribution marginale de Y.

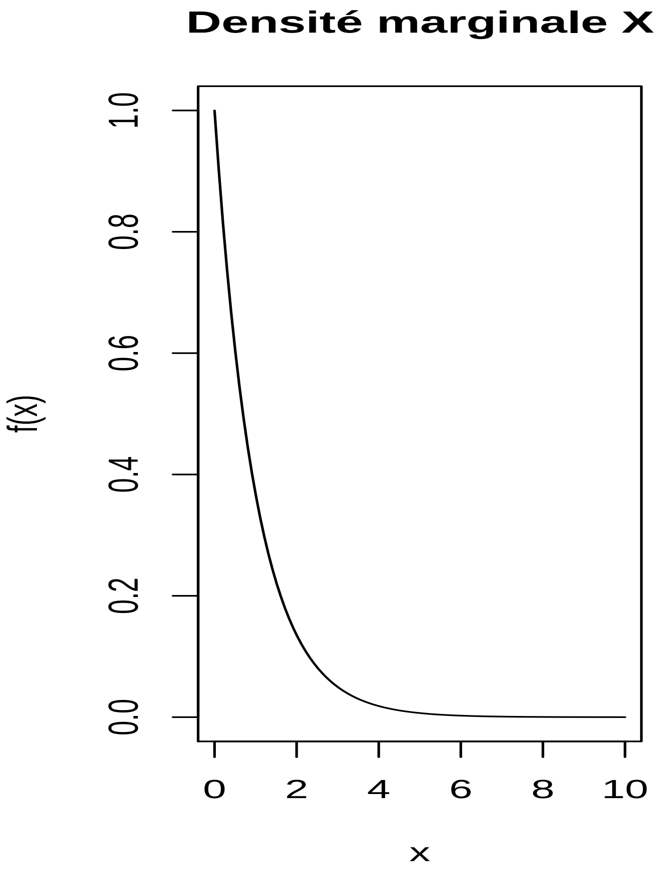
Cela définit-il bien des fonctions de densité ?

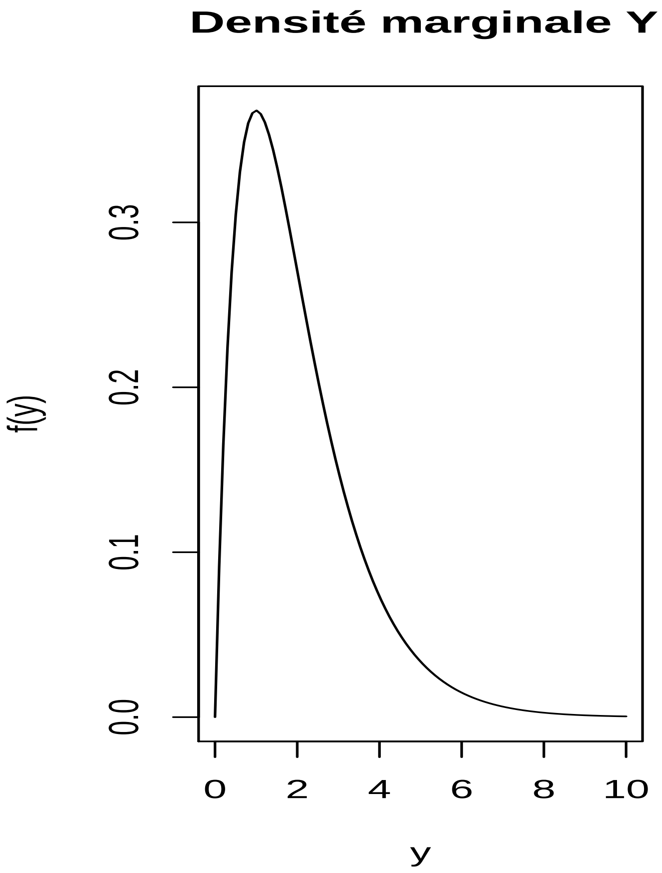
Exemple :

$$f(x,y) = \begin{cases} \exp(-y) & 0 < x < y < \infty \\ 0 & \text{sinon} \end{cases}$$

On trouve :

$$\begin{aligned} f_X(x) &= \int f(x,y)dy = \int_x^{\infty} \exp(-y)dy = \exp(-x) \\ f_Y(y) &= \end{aligned}$$





Fonctions de densité marginale.

3. Indépendance

Définition

Deux v.a. X et Y sont **indépendantes** si pour tout ensemble A et B on a

$$\mathrm{P}(\mathrm{X} \in \mathrm{A}, \mathrm{Y} \in \mathrm{B}) = \mathrm{P}(\mathrm{X} \in \mathrm{A})\mathrm{P}(\mathrm{Y} \in \mathrm{B}).$$

On peut démontrer que cette définition est équivalente à :

– Cas discret :

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

– Cas continu :

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

pour tout x, y.

Exemple :

X \ Y	0	1	2	$\mathrm{P(X = x)}$
0	$\frac{6}{153}$	$\frac{24}{153}$	$\frac{15}{153}$	$\frac{45}{153}$
1	$\frac{32}{153}$	$\frac{48}{153}$	0	$\frac{80}{153}$
2	$\frac{28}{153}$	0	0	$\frac{28}{153}$
$\mathrm{P(Y = y)}$	$\ldots$	$\frac{72}{153}$	$\ldots$	$\ldots$

Puisque

$$\mathrm{P(X = 2, Y = 2) \neq P(X = 2)P(Y = 2)},$$

on déduit que X et Y ne sont pas indépendantes.

Exemple :

$$f(x,y) = \begin{cases} \exp(-y) & 0 < x < y < \infty \\ 0 & \text{sinon} \end{cases}$$

On a trouvé :

$$\begin{aligned} f_X(x) &= \int f(x,y)dy = \int_x^\infty \exp(-y)dy = \exp(-x) \\ f_Y(y) &= y\exp(-y) \end{aligned}$$

Donc X et Y ne peuvent pas être indépendantes.

Exemple : (X,Y) a pour densité conjointe $f(x,y) = (x+y)^2 - (x-y)^2$ sur $[0,1]^2$. Les v.a. X et Y sont-elles indépendantes ?

4. Somme de deux v.a. indépendantes

Soit 2 v.a. X et Y. On s'intéresse à la distribution de leur somme $S = X + Y$.

D'une manière générale, c'est un problème difficile. En supposant que X et Y sont indépendantes, le problème est parfois simplifié.

Cas discret

$$\begin{aligned}
\mathrm{P(S = s)} &= \mathrm{P(X + Y = s)} \\
&= \sum_x \mathrm{P(X = x, Y = s - x)} \\
&= \sum_x \mathrm{P(X = x)P(Y = s - x)}.
\end{aligned}$$

Exemple : $X \sim \mathrm{Poi}(\lambda)$ et $Y \sim \mathrm{Poi}(\mu)$ sont indépendantes. Peut-on dire quelque chose de $S = X + Y$?

Puisque $\mathrm{P(X = j) = 0}$ quand $j < 0$, et $\mathrm{P(Y = k - j) = 0}$ quand $j > k$

$$
\begin{aligned}
\mathrm{P(X + Y = k)} &= \sum_{j=0}^{k} \mathrm{P(X = j)P(Y = k - j)} \\
&= \sum_{j=0}^{k} \exp(-\lambda)\frac{\lambda^j}{j!}\exp(-\mu)\frac{\mu^{k-j}}{(k-j)!} \\
&= \exp(-(\lambda + \mu))\frac{1}{k!}\sum_{j=0}^{k} C_{k,j}\lambda^j\mu^{k-j} \\
&= \exp(-(\lambda + \mu))\frac{1}{k!}\ldots
\end{aligned}
$$

Donc on peut écrire "$\mathrm{Poi}(\lambda) \overset{\mathrm{ind}}{+} \mathrm{Poi}(\mu) = \mathrm{Poi}(\lambda + \mu)$".

C'est plus l'exception que la règle de trouver une distribution simple et de même loi.

Par exemple a-t-on "$\mathrm{Bin}(n, p_1) \overset{\mathrm{ind}}{+} \mathrm{Bin}(n, p_2) = \mathrm{Bin}(n, p_1 + p_2)$" ?

Ou plutôt "$\mathrm{Bin}(n_1, p) \overset{\mathrm{ind}}{+} \mathrm{Bin}(n_2, p) = \mathrm{Bin}(n_1 + n_2, p)$" ?

Cas continu

Si $X \sim f_X$ est indépendante de $Y \sim f_Y$, alors $S = X + Y$ a pour densité

$$f_{X+Y}(s) = \int f_X(x) f_Y(s-x)dx.$$

On peut par exemple démontrer que

$$\text{"}\mathrm{N}(\mu_1, \sigma_1^2) \overset{\text{ind}}{+} \mathrm{N}(\mu_2, \sigma_2^2) = \mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)\text{"}.$$

5. Distributions conditionnelles

Cas discret

$$\mathrm{P}(\mathrm{X} = \mathrm{x} \mid \mathrm{Y} = \mathrm{y}) = \frac{\mathrm{P}(\mathrm{X} = \mathrm{x}, \mathrm{Y} = \mathrm{y})}{\mathrm{P}(\mathrm{Y} = \mathrm{y})}$$

Cas continu

$$f(x \mid Y = y) = \frac{f(x, y)}{f_Y(y)}$$

Ainsi $f(x, y) = f(x \mid Y = y) f_Y(y)$. Donc si X et Y sont indépendants, on obtient (page 14) :

$$f(x, y) = f_X(x) f_Y(y).$$

Exemple :

X \ Y	Pharma	SdIT	Bio	Chimie	$P(X = x)$
Math1	0.05	0.02	0.15	0.08	0.30
Math2	0.18	0.04	0.23	0.12	0.57
Math3	0.02	0.02	0.06	0.04	0.13
$P(Y = y)$	0.25	0.08	0.44	0.23	1

Tableau de probabilité

$$P(X = 2 \mid Y = \text{Bio}) = \frac{P(X=2, Y=\text{Bio})}{P(Y=\text{Bio})} = \frac{0.23}{0.44} = 0.52$$

Exemple :

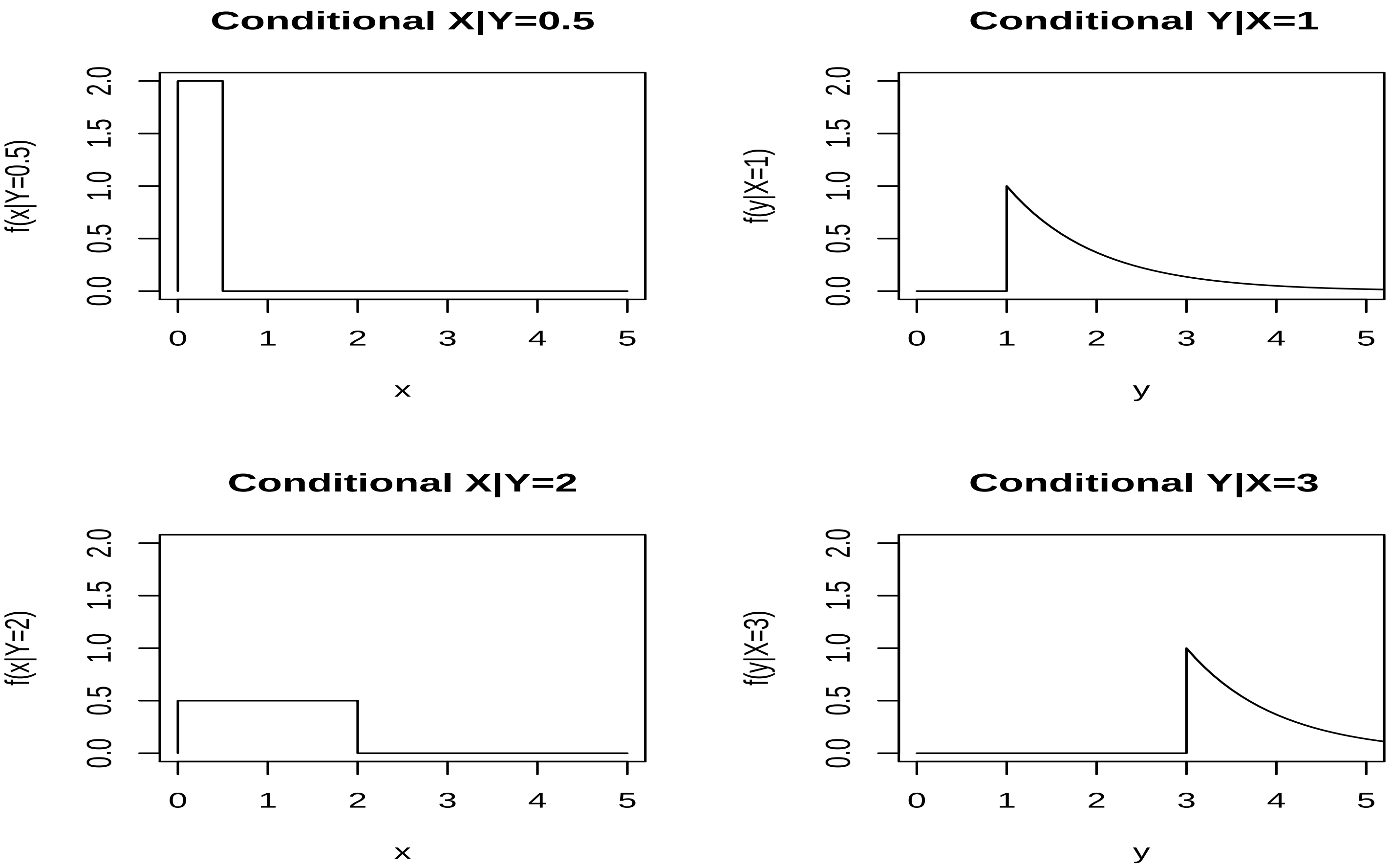
$$f(x,y) = \begin{cases} \exp(-y) & 0 < x < y < \infty \\ 0 & \text{sinon} \end{cases}$$

$f(x \mid Y = y) = \frac{f(x,y)}{f(y)} = \frac{\exp(-y)}{y\exp(-y)} = \frac{1}{y}$ pour $0 < x < y$

Donc $X \mid Y = y \sim \ldots$.

$f(y \mid X = x) = \frac{f(x,y)}{f(x)} = \frac{\exp(-y)}{\exp(-x)} = \exp(-(y-x))$ pour $y > x$.

Donc $Y - X \mid X = x \sim \text{Exp}(1)$.



Fonctions de densité conditionnelle.