lorsqu'on projette cette masse sur l'axe des x.
(On accumule la densité sur l'axe verticale
$x = x_0$ pour chaque x_0.)

<u>Plus de deux variables</u>. S'il y a plusieurs, disons (x,y,z), on intègre toutes les variables, sauf une, pour les lois marginales

$$f_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y,z)\,dy\,dz \quad , \quad f_Y(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y,z)\,dx\,dz$$

$$f_Z(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y,z)\,dx\,dy.$$

Si on veut, par exemple, la loi conjointe de X et Y, on intègre en z:

$$f_{(X,Y)}(x,y) = \int_{-\infty}^{\infty} f(x,y,z)\,dz.$$

§6.4 Lois conditionnelles.

On rappelle que pour des évènements A et B avec $\mathbb{P}(B)>0$, on a défini $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$.

Si X et Y sont des variables discrètes de densité conjointe p, alors comme $X=x$ et $Y=y$ sont des évènements, on a

$$\mathbb{P}(Y=y \mid X=x) = \frac{\mathbb{P}(X=x \text{ et } Y=y)}{\mathbb{P}(X=x)} = \frac{p(x,y)}{p_X(x)}$$

pourvu que $p_X(x) \neq 0$.

où p_X est la fonction de masse marginale de X

On définit $p_{Y|X=x}(y) = \dfrac{p(x,y)}{p_X(x)}$ et on la l'appelle la <u>fonction de masse conditionnelle de Y sachant que $X = x$</u>.

Dans le cas continu, on définit la <u>densité conditionnelle de Y sachant que $X = x$</u> par

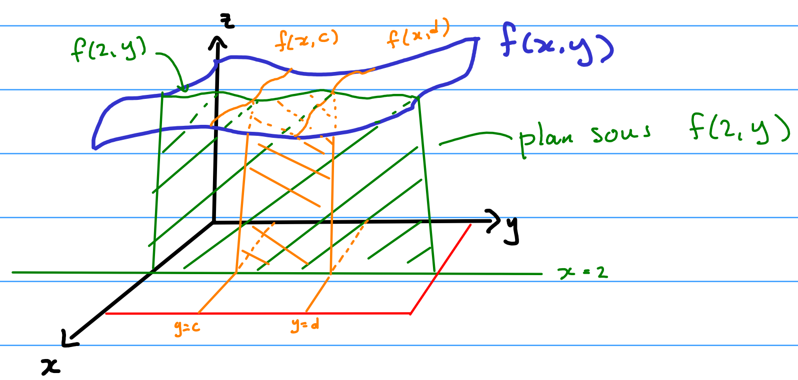
$$f_{Y|X=x}(y) = \frac{f(x,y)}{f_X(x)}$$

– fonction de densité conjointe de X et Y

– fonction de densité marginale de X

là où $f_X(x) \neq 0$.

<u>Interprétation géométrique</u>



$$P(c \leq Y \leq d \mid X = 2) = \frac{\int_c^d f(2,y)\,dy}{f_X(2)} = \frac{\int_c^d f(2,y)\,dy}{\int_{-\infty}^{\infty} f(2,y)\,dy}$$

$$= \frac{\text{aire du plan orange}}{\text{aire du plan vert}}$$

La fonction de masse conditionnelle est une fonction de masse légitime :

$$p_{Y|X=x}(y) = \frac{p(x,y)}{p_X(x)} \geq 0 \quad \text{et} \quad \sum_{y\in\mathbb{R}} p_{Y|X=x}(y) = 1$$

$$\sum_{y\in\mathbb{R}} \frac{p(x,y)}{p_X(x)} = \frac{1}{p_X(x)} \sum_{y\in\mathbb{R}} p(x,y) = \frac{1}{p_X(x)} p_X(x) = 1.$$

Ex. On lance un dé jusqu'à obtenir un 6. Soit

X = le nombre de lancers pour obtenir un premier 6

Y = le nombre de 2 parmi tous les lancers.

Alors

$$p(x,y) = p_X(x)\, p_{Y|X=x}(y)$$

$X \sim$ géométrique $(1/6)$

Si on sait que $X = x$, alors $Y \sim$ binomiale$(x-1, 1/5)$, car pour les $x-1$ premiers lancers, on sait qu'il n'y a pas de 6, donc il reste 5 valeurs possibles.

On a donc

$$p(x,y) = (p_X(x))(p_{Y|X=x}(y))$$

$$= \left(\left(1-\tfrac{1}{6}\right)^{x-1}\left(\tfrac{1}{6}\right)\right)\left(\binom{x-1}{y}\left(\tfrac{1}{5}\right)^{y}\left(1-\tfrac{1}{5}\right)^{x-y-1}\right)$$

$$= \frac{1}{6}\binom{x-1}{y}\left(\frac{5}{6}\right)^{x-1}\left(\frac{1}{5}\right)^{y}\left(\frac{4}{5}\right)^{x-y-1}.$$

si $x \geq 1$ et $0 \leq y \leq x-1$ $(x, y \in \mathbb{Z})$ et $p(x,y) = 0$ sinon.

Par exemple, $P(X=2 \text{ et } Y=1) = \frac{1}{6}\binom{1}{1}\frac{5}{6}\times\frac{1}{5}\times\left(\frac{4}{5}\right)^{0} = \frac{1}{36}$

En effet, $[X=2 \text{ et } Y=1]$ est équivalent à « premier lancer vaut 2 et deuxième lancer vaut 2 »

Ex. On choisit un point X au hasard et de façon uniforme dans $(0,4)$, puis on choisit un point Y au hasard et de façon uniforme dans $(X,4)$. Obtenir la densité conjointe de (X,Y).

Sol.

$X \sim \text{Uniforme}(0,4)$, donc $f_X(x) = \begin{cases} 1/4, & \text{si } 0<x<4 \\ 0, & \text{sinon.} \end{cases}$

Pour tout $x \in (0,4)$, on a $f_{Y|X=x}(y) = \begin{cases} \frac{1}{4-x}, & \text{si } x<y<4 \\ 0, & \text{sinon.} \end{cases}$

On a donc

$$f(x,y) = f_X(x)\, f_{Y|X=x}(y) = \begin{cases} \frac{1}{4(4-x)}, & \text{si } 0<x<y<4 \\ 0, & \text{sinon} \end{cases}$$

La <u>fonction de répartition conditionnelle de Y sachant $X=x$</u> est

$$F_{Y|X=x}(y) = \mathbb{P}(Y \leq y \mid X=x) = \begin{cases} \int_{-\infty}^{y} f_{Y|X=x}(v)\,dv & \text{dans le cas abs. continue} \\ \sum_{v \leq y} p_{Y|X=x}(y) & \text{dans le cas discret} \end{cases}$$

L'<u>espérance conditionnelle de Y sachant $X=x$</u> est

$$\mathbb{E}[Y \mid X=x] = \begin{cases} \int_{\mathbb{R}} y\, f_{Y|X=x}(y)\,dy & \text{dans le cas abs. continue} \\ \sum_{y \in \mathbb{R}} y\, p_{Y|X=x}(y) & \text{dans le cas discret} \end{cases}$$

Ex. Soit (X,Y) avec densité conjointe $f(x,y) = \begin{cases} 2y\,e^{-y(2+x)}, & \text{si } x \geq 0 \text{ et } y \geq 0 \\ 0, & \text{sinon.} \end{cases}$

On a $f_X(x) = \dfrac{2}{(2+x)^2}$ (un exemple précédent).

et donc si $x>0$

$$f_{Y|X=x}(y) = \begin{cases} (2+x)^2 y e^{-y(2+x)}, & \text{si } y \geq 0 \\ 0, & \text{sinon} \end{cases}$$

En $x=2$, on obtient $f_{Y|X=2}(y) = \begin{cases} 16y e^{-4y}, & \text{si } y \geq 0 \\ 0, & \text{sinon} \end{cases}$

On a donc pour $y \geq 0$:

$$F_{Y|X=2}(y) = \int_{-\infty}^{y} f_{Y|X=2}(v)\,dv$$

$$= \int_0^y 16\, v e^{-4v}\,dv \qquad u = 4v, \quad du = 4dv$$

$$= \int_0^{4y} \cancel{16}\, \frac{u}{\cancel{4}}\, e^{-u}\, \frac{du}{\cancel{4}} \qquad s = u, \quad ds = du \qquad dt = e^{-u}du, \quad t = -e^{-u}$$

$$= -ue^{-u}\Big|_0^{4y} + \int_0^{4y} e^{-u}\,du$$

$$= -4y e^{-4y} - e^{-u}\Big|_0^{4y}$$

$$= 1 - 4y e^{-4y} - e^{-4y}.$$

Donc $F_{Y|X=2}(y) = \begin{cases} 1-(4y+1)e^{-4y}, & \text{si } y \geq 0 \\ 0 & \text{si } y < 0 \end{cases}$

Et $\mathbb{E}(Y|X=x) = \int_{-\infty}^{\infty} y\, f_{Y|X=x}(y)\,dy$

$$= \int_0^{\infty} (2+x)^2 y^2 e^{-y(2+x)}\,dy \qquad t = y(2+x), \quad dt = (2+x)dy$$

$$= \int_0^{\infty} \cancel{(2+x)^2}\, \frac{t^2}{\cancel{(2+x)^2}}\, e^{-t}\, \frac{dt}{2+x}$$

$$= \frac{1}{2+x}\int_0^{\infty} t^2 e^{-t}\,dt = \frac{2!}{2+x}.$$

En particulier, $\mathbb{E}(Y \mid X=2) = \dfrac{2}{2+2} = \dfrac{1}{2}$.

§6.5 Indépendance de variables aléatoires

Déf Les variables aléatoires X et Y sont indépendantes si pour tout événements $A \subseteq \mathbb{R}$, $B \subseteq \mathbb{R}$, on a

$$\mathbb{P}\left[(X \in A) \text{ et } (Y \in B)\right] = \mathbb{P}(X \in A)\, \mathbb{P}(Y \in B).$$

Théorème Si X et Y sont des variables aléatoires indépendantes et si g et h sont des fonctions continues, alors $g(X)$ et $h(Y)$ sont indépendantes.

↳ la continuité n'est pas nécessaire, il suffit que $g^{-1}(C)$ soit un événement pour tout événement $C \subseteq \mathbb{R}$

Dém. (Optionnel)

$$\mathbb{P}\left[g(X) \in A \text{ et } h(Y) \in B\right] = \mathbb{P}\left[\left(X \in g^{-1}(A)\right) \text{ et } \left(Y \in h^{-1}(B)\right)\right]$$

$$= \mathbb{P}\left[X \in g^{-1}(A)\right] \times \mathbb{P}\left[Y \in h^{-1}(B)\right]$$

$$= \mathbb{P}\left[g(X) \in A\right] \times \mathbb{P}\left[h(Y) \in B\right]. \quad \square$$

Ex. Si X et Y sont indépendantes, alors $X - \mu_X$ et $Y - \mu_Y$ le sont aussi.

Théorème Soit X, Y deux variables aléatoires.

Les énoncés suivants sont équivalents :

1. X et Y sont indépendantes
2. $F_{(X,Y)}(x,y) = F_X(x)\, F_Y(y)$, où $F_{(X,Y)}$ = fonction de répartition conjointe
   F_X, F_Y = fonctions " marginales

3. $\begin{cases} p_{(X,Y)}(x,y) = p_X(x)\, p_Y(y) & \text{dans le cas discret} \\ f_{(X,Y)}(x,y) = f_X(x)\, f_Y(y) & \text{dans le cas abs. continu} \end{cases}$

où $p_{(X,Y)} / f_{(X,Y)}$ = fonction de masse/densité conjointe

$p_X, p_Y / f_X, f_Y$ = fonctions de masse/densité marginales.

Dém. On fait seulement $1 \Rightarrow 3$:

$$\begin{aligned} p(x,y) &= \mathbb{P}[(X=x) \text{ et } (Y=y)] \\ &= \mathbb{P}[X \in \{x\} \text{ et } Y \in Y \in \{y\}] \\ &= \mathbb{P}[X \in \{x\}]\, \mathbb{P}[Y \in \{y\}] \\ &= \mathbb{P}(X=x)\, \mathbb{P}(Y=y) \\ &= p_X(x)\, p_Y(y) \qquad \square \end{aligned}$$

Ex. On lance deux dés

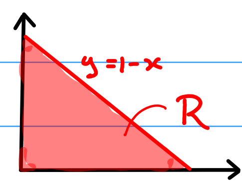
X = valeur du premier dé , Y = valeur du deuxième dé

On sait que les lois marginales sont $p_X(k) = \frac{1}{6}$ et $p_Y(k) = \frac{1}{6}$ pour $k \in \{1, \dots, 6\}$ et 0 ailleurs.

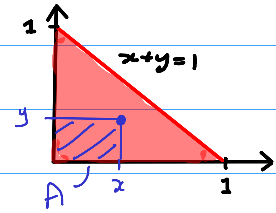
De plus, la lois conjointes est $p(x,y) = \frac{1}{36} = p_X(x)\, p_Y(y)$ pour $x, y \in \{1, \dots, 6\}$ et 0 ailleurs.

donc X et Y sont indépendants.

Ex. Soit (X,Y) un point choisi au hasard et de façon uniforme dans la région $R = \{(x,y) \in \mathbb{R}^2 \mid 0 \le x \le 1,\ 0 \le y \le 1,\ x+y \le 1\}$.



Pour $(x,y) \in R$, on a



$$F(x,y) = \mathbb{P}(X \leq x \text{ et } Y \leq y)$$
$$= \frac{\text{Aire}(A)}{\text{Aire}(R)} = \frac{xy}{1/2} = 2xy.$$

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y}(2xy) = 2 \longrightarrow f_{X,Y}(x,y) = \begin{cases} 2, & \text{si } (x,y) \in R \\ 0, & \text{sinon} \end{cases}$$

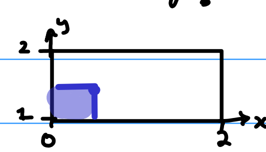
$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^{1-x} 2\,dy = 2(1-x)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^{1-y} 2\,dx = 2(1-y)$$

$$f_X(x) f_Y(y) = 4(1-x)(1-y) \neq 2 = f(x,y)$$

Ainsi, X et Y ne sont pas indépendants.

Ex. On reprend l'exemple de (X,Y) choisi au hasard et de façon uniforme, mais cette fois dans le rectangle $[0,2] \times [1,2]$.



On trouve $F(x,y) = \dfrac{x(y-1)}{2 \times 1} = \dfrac{x(y-1)}{2}$ si $x \in [0,2]$ et $y \in [1,2]$

$$\Rightarrow f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y) = \frac{1}{2}$$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_1^2 \frac{1}{2}\,dx = \frac{1}{2}$$

On a bien $f(x,y) = \frac{1}{2} = f_X(x) f_Y(y)$.

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^2 \frac{1}{2}\,dx = 1$$

$\Rightarrow X$ et Y sont indépendantes.