Additif au Chapitre 12 : Densités marginales

Théorème

Soient X et Y deux variables aléatoires continues de densité jointe $f(x, y)$. Alors les densités marginales $f_X(x)$ et $f_Y(y)$ peuvent être obtenues par les expressions suivantes :

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

et

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

Démonstration

Pour tout $A \subset \mathbb{R}$:

$$\begin{aligned}\mathbb{P}(X \in A) &= \mathbb{P}(X \in A, Y \in \mathbb{R}) \\ &= \int_A \left(\int_{\mathbb{R}} f(x,y)dy \right) dx \text{ par définition}\end{aligned}$$

On pour tout $A \subset \mathbb{R}$, on a par définition,

$$\mathbb{P}(X \in A) = \int_A f_X(x)dx$$

Ainsi par identification :

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

Généralisation

Soient $X_1, X_2, ..., X_n$ n variables aléatoires continues de densité jointe $f_X(x_1, x_2, ..., x_n)$. Alors les densités marginales $f_{X_i}(x_i)$ peuvent être obtenues par les expressions suivantes :

$$f_{X_i}(x_i) = \int ... \int_{\mathbb{R}^{n-1}} f(x_1, x_2, ..., x_n) dx_1 ... dx_{i-1} dx_{i+1} ... dx_n$$

Par exemple si $n = 3$

$$f_{X_1}(x_1) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x_1, x_2, x_3) dx_2 dx_3$$

Ainsi, si on connait la loi de distribution de X, on connait aussi les lois de ces composantes X_i, appelées lois marginales de X.

Exercice 1 Soit $F_X(x_1, x_2) = 1 - e^{-\lambda \min(x_1, x_2)}$, $x_1, x_2 \geq 0$
Quelles sont les distributions marginales de X (cad les distrib de X_1 et X_2)

$$\begin{aligned} F_{X_1}(x_1) &= \mathbb{P}(X_1 \leq x_1, X_2 \in \mathbb{R}) = \mathbb{P}(X_1 \leq x_1, X_2 \leq +\infty) \\ &= 1 - e^{-\lambda \min(x_1, +\infty)} = 1 - e^{-\lambda x_1} \end{aligned}$$

$\Rightarrow X_1 \sim \mathcal{E}(\lambda)$

De manière similaire,
$F_{X_2}(x_2) = F_X(+\infty, x_2) = 1 - e^{-\lambda \min(+\infty, x_2)} = 1 - e^{-\lambda x_2}$

$\Rightarrow X_2 \sim \mathcal{E}(\lambda)$

La réciproque est fausse : il ne suffit pas de connaître les lois marginales pour connaitre la loi du vecteur

Exercice 2 Soit $F_X(x_1, x_2) = \left(1 - e^{-\lambda x_1}\right)\left(1 - e^{-\lambda x_2}\right)$,
Alors:

$$\begin{aligned} F_{X_1}(x_1) &= F_X(x_1, +\infty) = \left(1 - e^{-\lambda x_1}\right)\left(1 - e^{-\lambda(+\infty)}\right) \\ &= 1 - e^{-\lambda x_1} \Rightarrow X_1 \sim \mathcal{E}(\lambda) \\ F_{X_2}(x_2) &= F_X(+\infty, x_2) = \left(1 - e^{-\lambda(+\infty)}\right)\left(1 - e^{-\lambda x_2}\right) \\ &= 1 - e^{-\lambda x_2} \Rightarrow X_2 \sim \mathcal{E}(\lambda) \end{aligned}$$

On voit donc que, comme dans l'exemple précédent, $X_1 \sim \mathcal{E}(\lambda)$ et $X_2 \sim \mathcal{E}(\lambda)$ mais la distribution de $X = (X_1, X_2)$ est différente dans les deux cas. Donc, on ne peut pas déduire la distrib de X de celles de X_1 et X_2

Le seul cas dans lequel ceci est possible et le cas où (toutes) les composantes sont indépendantes puisque si $X_1, ..., X_2$ sont indépendantes alors :

$$F_X(x_1, ..., x_n) = F_{X_1}(x_1)F_{X_2}(x_2)...F_{X_n}(x_n)$$

Dans l'exemple 2, X_1 et X_2 sont indépendantes alors que dans l'exemple 1 elles ne le sont pas.