

Apprentissage statistique

Chapitre 3 :

Lucie Le Briquer

29 janvier 2019

Table des matières

1	2
2	2
3	2
4 Désintégration de la mesure	3

1

2

3

4 Désintégration de la mesure

$$f_{X|Y}(x|y) = \begin{cases} \frac{f_{(X,Y)}(x,y)}{f_Y(y)} & \text{si } f_Y(y) \neq 0 \\ f_0(x) & \text{où } f_0(y) \text{ densité fixée par rapport à } \mu_X \end{cases}$$

Théorème 1

On a $\mathbb{P}_X(dx) \otimes \mathbb{P}_Y(dy)$ -p.p. :

$$f(x, y) = f(x|y)f(y)$$

Preuve.

Soit $A \in \mathcal{X}$, $B \in \mathcal{Y}$.

$$\mathbb{P}(X \in A, Y \in B) = \int_{A \times B} f(x, y) \mathbb{1}_{\{f_Y(y) \neq 0\}} \mu_X(dx) \mu_Y(dy)$$

car $\mathbb{P}_Y(dy)$ -p.p. $f_Y(y) > 0$, en effet :

$$\mathbb{P}_Y(\{f_Y = 0\}) = \int f_Y(y) \mathbb{1}_{\{f_Y = 0\}} d\mu_Y(y) = 0$$

Conclusion $\forall a \in \mathcal{X}$, $B \in \mathcal{Y}$:

$$\mathbb{P}(X \in A, Y \in B) = \int_{A \times B} f_Y(y) f_{X|Y}(x|y) \mu_X(dx) \mu_Y(dy)$$

On conclut par unicité de la densité. ■

$$\mathbb{P}_{(X,Y)}(A) = \mathbb{E}[\mathbb{1}_A(X, Y)] = \mathbb{E}[\mathbb{1}_{A_Y}(X)]$$

où $A_y = \{x \in \mathbb{X} : (x, y) \in A\}$ est une section de A par rapport à y . On peut réécrire :

$$\begin{aligned} \mathbb{P}_{(X,Y)}(A) &= \mathbb{E}[\mathbb{E}[\mathbb{1}_{A_Y}(X)|Y]] \\ &= \mathbb{E}[Kf(Y)] \quad \text{où } f(x, y) = \mathbb{1}_{A_y}(x) \\ &= \int_{\mathbb{Y}} \mathbb{P}_Y(dy) \int_{\mathbb{X}} K(y, A_y) \end{aligned}$$

Définition 1 (désintégration)

Soit (E, \mathcal{A}) et (F, \mathcal{B}) deux espaces mesurables. Prenons :

- $T: E \rightarrow F$ mesurable
- λ mesure σ -finie sur (E, \mathcal{A})
- $\mu = \lambda_{\#T}$ mesure image de λ par T

On dit que un noyau de transition K sur $F \times \mathcal{A}$ est une désintégration de λ par rapport à T si :

1. $\forall y \in F, \text{supp}(K(y, \cdot)) \subset \{T = y\} \Leftrightarrow \forall y \in F, K(y, \{T \neq y\}) = 0$
2. $\lambda = \mu K$

Remarque. Dans le cas de deux variables (X, Y) sur $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{P}_{X|Y}$ est un noyau sur $\mathbb{Y} \times \mathcal{X}$. Alors que dans le cas de la désintégration $\lambda = \mathbb{P}_{(X,Y)}$, $T: (x, y) \mapsto y$ \mathbb{P}_Y . Ici $E = \mathbb{X} \times \mathbb{Y}$.

Ici dans le cas de la désintégration, on a un noyau K qui est défini sur $\mathbb{Y} \times (\mathcal{X} \otimes \mathcal{Y})$. Cependant on a une bijection entre les noyaux de transition sur $\mathbb{Y} \times \mathcal{X}$ et les noyaux de transitions sur $\mathbb{Y} \times (\mathcal{X} \otimes \mathcal{Y})$ qui vérifie (2) pour $T: (x, y) \mapsto y$.

Prenons P un noyau sur $\mathbb{Y} \times \mathcal{X}$. On définit simplement :

$$K(y, A \times B) = \delta_y(B)P(y, A) = P(y, (A \times B)_y)$$

pour tout $y \in \mathbb{Y}$ et $A \in \mathcal{X}$, $B \in \mathcal{Y}$.

On en déduit alors qu'il existe une unique mesure pour tout y telle que :

$$\mu_y(X) = P(y, C_y)$$

Cette mesure vit sur $\mathcal{A} \otimes \mathcal{B}$ et vérifie :

$$\mu_y(\{T \neq y\}) = 0$$

Il reste à montrer que $y \mapsto \mu_y(C)$ est mesurable. Or $\mu_y(X) = P(y, C_y)$ donc mesurable. Pour conclure on pose $K(y, C) = \mu_y(C)$.

Réciproquement, soit K un noyau sur $\mathbb{Y} \times (\mathcal{X} \otimes \mathcal{Y})$ qui vérifie $K(y, \{T \neq y\}) = 0$. On définit alors le noyau :

$$P(y, A) = K(y, A \times \{y\}) = K(y, A \times \mathbb{Y})$$

sur $\mathbb{Y} \times \mathcal{X}$.

Théorème 2

Soit (X, Y) v.a. de $(\Omega, \mathcal{F}, \mathbb{P})$. L'existence d'une loi conditionnelle régulière de $X|Y$ est équivalent à l'existence d'une désintégration de $\mathbb{P}_{(X,Y)}$ par $T: (x, y) \mapsto y$.

Preuve.

On a construit pour tout noyau $\mathbb{Y} \times \mathcal{X}$ un noyau sur $\mathbb{Y} \times (\mathcal{X} \otimes \mathcal{Y})$ qui vérifie (1). ■

Théorème 3

Soit E un espace métrique séparable complet (polonais) muni de $\mathcal{A} = \mathcal{B}(E)$ et F idem muni de $\mathcal{B} = \mathcal{B}(F)$. Soit T une application mesurable et λ mesure σ -finie sur $(\mathbb{E}, \mathcal{B}(E))$. Alors il existe K une désintégration de λ par Y si λ est de Radon.

Remarque. λ est de Radon si :

$$\lambda(A) = \sup\{\lambda(C) : C \text{ compact} \subset A\} \quad \lambda(C) < +\infty \quad \forall C \text{ compact}$$

Théorème 4

Soit $E = E_1 \times E_2$ avec (E_1, \mathcal{F}_1) espace mesurable et E_2 polonais muni de $\mathcal{F}_2 = \mathcal{B}(E_2)$. Soit λ une mesure sur $(E, \mathcal{B}(E_1 \times E_2))$ et $T: (x, y) \mapsto y$. Alors il existe une désintégration de λ par T .