LOIS CONDITIONNELLES

PRÉPARATION À L'AGRÉGATION EXTERNE DE MATHÉMATIQUES DE L'UNIVERSITÉ RENNES 1 [1]
ANNÉE 2010/2011

Pour une variable aléatoire X, implicitement définie sur l'espace probabilisé $(\Omega,\mathscr{F},\mathbb{P})$, $\mathbb{P}_X$ désigne la loi de X.

1. LA NOTION DE LOI CONDITIONNELLE

Il est facile de définir la loi conditionnelle de Y sachant X lorsque X est discrète, à valeurs dans $\mathscr{D}$. Supposons que Y est à valeurs dans $\mathbb{R}^d$. La loi conditionnelle de Y sachant X est la *famille* de lois $\{\mathbb{P}_{Y|X=x}, x\in\mathscr{D}\}$ sur $\mathbb{R}^d$, telle que si $\mathbb{P}(X=x)\neq 0$,

$$\mathbb{P}_{Y|X=x}(B)=\mathbb{P}(Y\in B|X=x)\quad \forall B\in\mathscr{B}(\mathbb{R}^d),$$

et, dans le cas contraire, $\mathbb{P}_{Y|X=x}$ est une probabilité quelconque sur $\mathbb{R}^d$. On a alors, pour $A\subset\mathscr{D}$ et $B\in\mathscr{B}(\mathbb{R}^d)$:

$$\mathbb{P}_{(X,Y)}(A\times B)=\sum_{x\in A}\mathbb{P}_{Y|X=x}(B)\mathbb{P}_X(\{x\})=\int_A \mathbb{P}_{Y|X=x}(B)\mathbb{P}_X(dx),$$

et cette relation caractérise la loi conditionnelle de Y sachant X. Si ce procédé ne s'étend pas au cas où X est quelconque, en raison de l'impossibilité de diviser par 0, il nous guide pour exporter cette notion de loi conditionnelle au cas général.

Définition 1.1 *Une application $K:\mathbb{R}^p\times\mathscr{B}(\mathbb{R}^d)\to[0,1]$ est appelée noyau de transition si elle satisfait les propriétés :*
(i) $\forall x\in\mathbb{R}^p$, $K(x,.)$ est une probabilité sur $\mathbb{R}^d$;
(ii) $\forall A\in\mathscr{B}(\mathbb{R}^d)$, l'application $K(.,A)$ est mesurable.

Si μ est une probabilité sur $\mathbb{R}^p$ et K un noyau de transition sur $\mathbb{R}^p\times\mathscr{B}(\mathbb{R}^d)$, on note $\mu.K$ l'application définie sur $\mathscr{B}(\mathbb{R}^p)\times\mathscr{B}(\mathbb{R}^d)$ par :

$$\mu.K(A\times B)=\int_A K(x,B)d\mu(x),\ A\in\mathscr{B}(\mathbb{R}^p)\text{ et }B\in\mathscr{B}(\mathbb{R}^d).$$

Cette application, qui est σ-additive sur $\mathscr{B}(\mathbb{R}^p)\times\mathscr{B}(\mathbb{R}^d)$, se prolonge en une unique probabilité sur $\mathbb{R}^p\times\mathbb{R}^d$. Le prolongement est encore noté $\mu.K$.

Définition 1.2 *Soient X et Y des v.a. à valeurs dans $\mathbb{R}^p$ et $\mathbb{R}^d$. On appelle loi conditionnelle de Y sachant X un noyau de transition K sur $\mathbb{R}^p\times\mathbb{R}^d$ tel que $\mathbb{P}_{(X,Y)}=\mathbb{P}_X.K$.*

Le *théorème de Jirina*, sur lequel nous ne nous attarderons pas, assure l'existence d'une telle loi conditionnelle, dans le cadre de la définition précédente. Le plus souvent, le noyau de transition K de la définition précédente est noté $\mathbb{P}_{Y|X=.}$, de sorte que $\mathbb{P}_{(X,Y)}=\mathbb{P}_X.\mathbb{P}_{Y|X=.}$. Noter que la loi conditionnelle est donc en réalité une *famille* de lois de probabilité.

Exercice 1.1 Soit (X,Y) un vecteur aléatoire gaussien plan centré réduit. Calculer la loi conditionnelle de X^2+Y^2 sachant $X-Y$.

Exercice 1.2 Soient X et Y des v.a.r. avec $X\sim\mathscr{U}[0,1]$, et λ la mesure de Lebesgue sur $\mathbb{R}$. On suppose que pour tout $x\in[0,1/2[$, $\mathbb{P}_{Y|X=x}=\mathbf{1}_{[1/2,1]}\lambda+(1/2)\delta_x$, et pour tout $x\in[1/2,1[$, $\mathbb{P}_{Y|X=x}=\mathscr{U}[0,1]$. Calculer la loi de Y.

2. PROPRIÉTÉS DES LOIS CONDITIONNELLES

Dorénavant, X et Y désignent des v.a. à valeurs dans $\mathbb{R}^p$ et $\mathbb{R}^d$ respectivement. Le théorème de type Fubini qui est énoncé ci-dessous est d'utilité constante dans la manipulation des lois conditionnelles.

1. Benoît Cadre - ENS Cachan Bretagne

Théorème 1.1 [FUBINI] *Soit* $\varphi : \mathbb{R}^p \times \mathbb{R}^d \to \mathbb{R}$ *mesurable.*

(i) Si φ *est positive, l'application* $x \mapsto \int \varphi(x,y)\mathbb{P}_{Y|X=x}(dy)$ *est* $\mathscr{B}(\mathbb{R}^p)$*-mesurable, et*

$$\int_{\mathbb{R}^p\times\mathbb{R}^d} \varphi d\mathbb{P}_{(X,Y)} = \int_{\mathbb{R}^p}\left[\int_{\mathbb{R}^d}\varphi(x,y)\mathbb{P}_{Y|X=x}(dy)\right]\mathbb{P}_X(dx).$$

(ii) Si φ *est* $\mathbb{P}_{(X,Y)}$*-intégrable, alors pour* $\mathbb{P}_X$*-p.t. x, l'application* $\varphi(x,.)$ *est* $\mathbb{P}_{Y|X=x}$*-intégrable, l'application* $x \mapsto \int \varphi(x,y)\mathbb{P}_{Y|X=x}(dy)$ *est* $\mathbb{P}_X$*-intégrable et l'égalité de (i) est encore vraie.*

APPLICATION : ESPÉRANCE CONDITIONNELLE. Fixons $\psi : \mathbb{R}^d \to \mathbb{R}$ telle que $\psi(Y) \in L^1$. Une simple application du théorème de Fubini montre que pour tout $A \in \mathscr{B}(\mathbb{R}^p)$,

$$\int_{\{X\in A\}} \psi(Y)d\mathbb{P} = \int_A\left[\int_{\mathbb{R}^d}\psi(y)\mathbb{P}_{Y|X=x}(dy)\right]\mathbb{P}_X(dx),$$

et donc que

$$\mathbb{E}[\psi(Y)|X=x] = \int_{\mathbb{R}^d}\psi(y)\mathbb{P}_{Y|X=x}(dy).$$

Par ailleurs, on peut aussi calculer $\mathbb{E}\psi(Y)$ en utilisant l'espérance conditionnelle de $\psi(Y)$ sachant X :

$$\mathbb{E}\psi(Y) = \int_{\mathbb{R}^p}\left[\int_{\mathbb{R}^d}\psi(y)\mathbb{P}_{Y|X=x}(dy)\right]\mathbb{P}_X(dx) = \int_{\mathbb{R}^p}\mathbb{E}[\psi(Y)|X=x]\mathbb{P}_X(dx).$$

On peut aussi calculer une loi conditionnelle en se ramenant à un calcul d'espérance conditionnelle (cf complément de cours : Espérance conditionnelle et introduction aux martingales). Reprenons le contexte de l'application au calcul d'espérances conditionnelles. On calcule dans un premier temps $\mathbb{P}(\psi(Y) \le y|X=x)$, puis on utilise le théorème de Dynkin pour en déduire $\mathbb{P}(\psi(Y) \in .|X=x)$. La définition 1.2 et la définition d'une espérance conditionnelle nous montrent alors que $\mathbb{P}(\psi(Y) \in .|X=.)$ est le noyau de transition recherché.

Exercice 2.1 Soient $X_1,\cdots,X_n$ i.i.d. de loi $\mathscr{U}[0,1]$ et $T = \sup(X_1,\cdots,X_n)$. Le vecteur aléatoire (X_1,T) possède-t-il une densité ? Calculer la loi conditionnelle de X_1 sachant T, puis sa limite lorsque $n \to \infty$.

Lorsque la v.a. X apparaît aussi dans Y, le calcul de la loi conditionnelle de Y sachant X peut être facilité :

Théorème 2.1 [TRANSFERT CONDITIONNEL] *Soit* φ *une fonction mesurable définie sur* $\mathbb{R}^p \times \mathbb{R}^d$*. Alors, pour tout* $x \in \mathbb{R}^p$, $\mathbb{P}_{\varphi(X,Y)|X=x} = \mathbb{P}_{\varphi(x,Y)|X=x}$*. En particulier, si X et Y sont indépendantes, on a pour tout* $x \in \mathbb{R}^p$: $\mathbb{P}_{\varphi(X,Y)|X=x} = \mathbb{P}_{\varphi(x,Y)}$.

Sous les conditions d'intégrabilité adéquates, le calcul de quantités du type $\mathbb{E}[\varphi(X,Y)|X]$ se ramène donc à un calcul de la loi conditionnelle de Y sachant X, car

$$\mathbb{E}[\varphi(X,Y)|X=x] = \mathbb{E}[\varphi(x,Y)|X=x] = \int_{\mathbb{R}^d}\varphi(x,y)\mathbb{P}_{Y|X=x}(dy).$$

Si, de plus, X et Y sont indépendantes, on a :

$$\mathbb{E}[\varphi(X,Y)|X=x] = \int_{\mathbb{R}^d}\varphi(x,y)\mathbb{P}_Y(dy).$$

Exercice 2.2 Soient $X \perp\!\!\!\perp Y$, avec $X \sim \mathscr{E}(\lambda)$ et $Y \sim \mathscr{E}(\lambda)$. Calculer la loi conditionnelle de $X+Y$ sachant X, et la loi conditionnelle de $X \wedge Y$ sachant X.

3. DEUX CAS PARTICULIERS

Lorsque le couple (X,Y) possède une densité, on a là encore une situation assez confortable. Précisons auparavant la notion de *densité conditionnelle* : si, pour $\mathbb{P}_X$-p.t. $x \in \mathbb{R}^p$, $\mathbb{P}_{Y|X=x}$ possède une densité, celle-ci est appelée *densité conditionnelle* de Y sachant $X=x$.

Théorème 2.2 [DENSITÉ CONDITIONNELLE] *Supposons que* (X,Y) *possède une densité f. Alors, pour tout x tel que* $f_X(x) > 0$*, la loi de Y sachant* $X=x$ *admet une densité, qui vaut :*

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

Preuve Soit ρ une probabilité quelconque sur $\mathbb{R}^d$. Pour tout $x \in \mathbb{R}^p$ et $B \in \mathscr{B}(\mathbb{R}^d)$, on note

$$K(x,B) = \int_B \frac{f(x,y)}{f_X(x)} dy,$$

si $f_X(x) > 0$, et $K(x,B) = \rho(B)$ si $f_X(x) = 0$. Une loi conditionnelle de Y sachant X est K, car K est un noyau de transition sur $\mathbb{R}^p \times \mathbb{R}^d$ qui vérifie $\mathbb{P}_{(X,Y)} = \mathbb{P}_X.K$. Par ailleurs, si $f_X(x) > 0$, la mesure $K(x,.)$ est à densité $f(x,.)/f_X(x)$ par rapport à la mesure de Lebesgue. •

Comme d'habitude, les vecteurs gaussiens ont un comportement très particulier.

Théorème 2.3 [CONDITIONNEMENT GAUSSIEN] *Supposons que (X,Y) est un vecteur gaussien, que Y est une v.a.r., et que le vecteur aléatoire X à valeurs dans $\mathbb{R}^p$ possède une matrice de variance inversible. On note a le vecteur défini par $a = \mathbb{V}(X)^{-1}(\mathrm{cov}(Y,X_1),\cdots,\mathrm{cov}(Y,X_p))^t$. Alors, la loi conditionnelle de Y sachant X est gaussienne, de moyenne $\mathbb{E}(Y|X) = \langle a, X - \mathbb{E}X\rangle + \mathbb{E}(Y)$, et de variance indépendante de X.*

Preuve La valeur de $\varphi(X) := \mathbb{E}(Y|X)$ a été calculée dans le complément de cours consacré aux vecteurs gaussiens. A cette occasion, on a aussi remarqué que $Y - \varphi(X) \perp\!\!\!\perp X$. On en déduit du théorème de transfert conditionnel que

$$\mathbb{P}_{Y|X=x} = \mathbb{P}_{Y-\varphi(X)+\varphi(X)|X=x} = \mathbb{P}_{Y-\varphi(X)+\varphi(x)|X=x} = \mathbb{P}_{Y-\varphi(X)+\varphi(x)}.$$

Comme φ est une fonction affine, $Y - \varphi(X)$, qui s'exprime comme une combinaison linéaire des v.a. X,Y, est donc une v.a.r. gaussienne. Enfin, on déduit de l'égalité ci-dessus que

$$\mathrm{var}(Y|X=x) = \mathrm{var}(Y - \varphi(X) + \varphi(x)) = \mathrm{var}(Y - \varphi(X)),$$

et donc que le variance de la loi conditionnelle de Y sachant X est indépendante de X. •

REFERENCES

- D. Dacunha-Castelle et M. Duflo, *Probabilités et statistiques - Tome 1 : Problèmes à temps fixe*, Masson, 1983.
- J.-Y. Ouvrard, *Probabilités 2 : maîtrise* & *agrégation*, Cassini, 2001.