Chapitre II: Espérance Conditionnelle

Probabilités II

5$^{\text{ème}}$ année DS
A.U:2022-2023

Tribus construites à partir de variables aléatoires

Soit Ω un ensemble muni d'une tribu $\mathcal{A}$. On rappelle qu'une tribu est un sous-ensemble de l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω

1. contenant $\emptyset$ et Ω.
2. stable par passage au complémentaire.
3. stable par union ou intersection dénombrable.

Remarque :
Une sous-tribu B de $\mathcal{A}$ est une tribu sur Ω tel que $B \subset \mathcal{A}$.

Rappel : Variable aléatoire

Définition

Une variable aléatoire réelle est **une application mesurable** X de $(\Omega, \mathcal{A})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, où $\mathcal{B}(\mathbb{R})$ est la tribu borélienne de $\mathbb{R}$ (c'est la plus petite tribu qui contient tous les intervalles de $\mathbb{R}$).
c-à-d $\forall B \in \mathcal{B}(\mathbb{R})$, on a :

$$\{\omega \in \Omega, X(\omega) \in B\} = \{X \in B\} \in \mathcal{A}.$$

Tribus construites à partir de variables aléatoires

Définition

Soit X une variable aléatoire sur $(\Omega, \mathcal{A})$ à valeurs dans $\mathbb{R}$. On note par $\sigma(X)$ la plus petite sous tribu $\mathcal{B}$ de $\mathcal{A}$, rendant l'application X $\mathcal{B}$-mesurable.

Cette tribu représente l'information donnée par la connaissance de X.

Exemple : Soit $X = 1_A$, avec $A \in \mathcal{A}$. la tribu engendrée par X est :

$$\sigma(X) = \{\emptyset, \Omega, A, A^c\}$$

Tribus construites à partir de variables aléatoires

- Soient $X_1, ..., X_n$, n variables aléatoires sur $(\Omega, \mathcal{A})$ à valeurs dans $\mathbb{R}$, on note :

$$\sigma(X_1, ..., X_n)$$

la plus petite sous tribu $\mathcal{B}$ de $\mathcal{A}$ rendant les applications $X_1, ..., X_n$ $\mathcal{B}$-mesurables.
De même cette tribu représente l'information donnée par la connaissance des variables aléatoires $X_1, ..., X_n$.

Filtration

Définition

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité, **une filtration** $(\mathcal{F}_n)_{n\geq 0}$ sur $(\Omega, \mathcal{A}, \mathbb{P})$ est une suite croissante de sous-tribus de $\mathcal{A}$

Proposition

Soit $(X_n, n \geq 1)$ une suite de variables aléatoires réelles $\mathcal{A}$-mesurables. On note, pour tout $n \geq 1$:

$$\mathcal{F}_n = \sigma(X_k, 1 \leq k \leq n).$$

Alors $(\mathcal{F}_n, n \geq 1)$ est une suite croissante de tribus qui porte le nom de **filtration naturelle**.

Pour un $n \geq 1$ fixé, $\mathcal{F}_n$ représente l'information disponible à l'instant n.

Filtration

Proposition

Soient X et Y deux variables aléatoires réelles, alors Y est dite $\sigma(X)$-mesurable si et seulement si il existe une fonction f de $\mathbb{R}$ dans $\mathbb{R}$, telle que :

$$\mathbb{P}(Y = f(X)) = 1$$

- Ce résultat explique l'interprétation d'une tribu comme information.
- Si Y est $\sigma(X)$-mesurable, et si on connaît la valeur de X, on peut en déduire la valeur de Y.

Espérance conditionnelle

Soit X une variable aléatoire réelle définie sur $(\Omega, \mathcal{A})$ et soit $B \subset \mathcal{A}$ une sous tribu de $\mathcal{A}$.

L'espérance conditionnelle vise à définir la notion de meilleure approximation de X par une variable aléatoire B-mesurable.

Espérance conditionnelle

Hypothèses :

On pose :

- $\mathcal{H}_{\mathcal{A}}$=$\{X$ variable aléatoire A-mesurable telle que $E(X^2) < +\infty\}$.

  C'est un espace de Hilbert : c'est un espace vectoriel complet muni d'une norme dérivant du produit scalaire $< X, Y >= E(XY)$ et $||X||^2 =< X, X >$.

- $\mathcal{H}_{\mathcal{B}}$=$\{X$ variable aléatoire B-mesurable telle que $E(X^2) < +\infty\}$.

  C'est un sous espace vectoriel fermé inclus dans $\mathcal{H}_{\mathcal{A}}$.

Espérance conditionnelle

Le théorème de projection sur un sous espace fermé dans un espace de Hilbert, nous permet d'avoir le résultat suivant.

Proposition

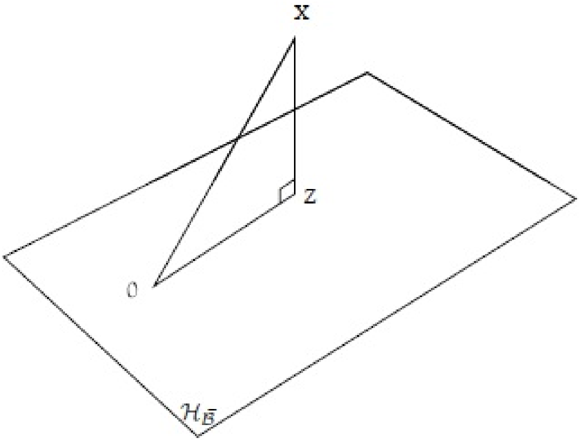
Si X est une variable aléatoire telle que $\mathbb{E}(X^2) < +\infty$, il existe une unique variable aléatoire Z, B-mesurable, telle que $\mathbb{E}(Z^2) < +\infty$ et vérifiant :

$$\mathbb{E}[(X - Z)^2] = \inf_{Y \in \mathcal{H}_{\mathcal{B}}} \mathbb{E}[(X - Y)^2].$$

- Z est la variable aléatoire la "plus proche" de X au sens de la norme de l'espace de Hilbert $\mathcal{H}_{\mathcal{B}}$.
- Z est l'espérance conditionnelle de X sachant $\mathcal{B}$ que l'on note :

$$\mathbb{E}[X|\mathcal{B}]$$

Espérance conditionnelle



Projection orthogonale sur $\mathcal{H}_{\mathcal{B}}$

Espérance conditionnelle

Proposition

Il existe une unique variable aléatoire $\mathcal{B}$-mesurable telle que $\mathbb{E}(Z^2) < +\infty$, vérifiant pout toute variable aléatoire Y, $\mathcal{B}$-mesurable et telle que $\mathbb{E}(Y^2) < +\infty$:

$$\mathbb{E}(XY) = \mathbb{E}(ZY).$$

-Cette variable aléatoire Z est la meilleure approximation, au sens de la norme $\sqrt{\mathbb{E}(X^2)}$, de X parmi les variables aléatoires $\mathcal{B}$-mesurables de carré intégrable.

Espérance conditionnelle

La définition précédente ne permet pas de traiter le cas de variables aléatoires seulement intégrables. On admet le résultat suivant :

Théorème

Soit X une variable aléatoire réelle telle que $\mathbb{E}(|X|) < +\infty$, alors il existe une unique variable aléatoire $\mathcal{B}$-mesurable Z telle que $\mathbb{E}(|Z|) < +\infty$ vérifiant $\forall Y$ $\mathcal{B}$-mesurable et **bornée** :

$$\mathbb{E}(XY) = \mathbb{E}(ZY)$$

On note la variable aléatoire Z sous la forme :

$$E(X|\mathcal{B})$$

C'est **l'espérance conditionnelle de X sachant $\mathcal{B}$.**

Espérance conditionnelle

Remarque :

Rappelons que toute variable aléatoire $\sigma(X)$-mesurable est de la forme $f(X)$. Donc on aura

$$\mathbb{E}(Y|X) = \mathbb{E}(Y|\sigma(X)) = f(X), \ \mathbb{P} - p.s$$

$\mathbb{E}(Y|X)$ est donc la meilleure approximation de Y par une fonction de X.

Propriétés de l'espérance conditionnelle

Linéarité

$$E(\lambda X + \mu Y|\mathcal{B}) = \lambda\mathbb{E}(X|\mathcal{B}) + \mu\mathbb{E}(Y|\mathcal{B}),\ \forall(\lambda,\mu) \in \mathbb{R}^2$$

Positivité

$$\text{Si } X \geq 0, \text{ alors } \mathbb{E}(X|\mathcal{B}) \geq 0.$$

$$\text{Si } X \geq Y, \text{ alors } \mathbb{E}(X|\mathcal{B}) \geq E(Y|\mathcal{B}).$$

Mesurabilité de X

Si X est intégrable et $\mathcal{B}$-mesurable alors :

$$\mathbb{E}(X|\mathcal{B}) = X.$$

Propriétés de l'espérance conditionnelle

"Sortir ce qui est connu"

Soit Z est une v.a. $\mathcal{B}$-mesurable et bornée alors :

$$\mathbb{E}(ZX|\mathcal{B}) = Z\mathbb{E}(X|\mathcal{B})$$

L'espérance de l'espérance conditionnelle

$$\mathbb{E}(\mathbb{E}(X|\mathcal{B})) = \mathbb{E}(X)$$

Emboitement

Soit $\mathcal{C}$ une sous-tribu de $\mathcal{B}$ alors :

$$\mathbb{E}(\mathbb{E}(X|\mathcal{B})|\mathcal{C}) = \mathbb{E}(X|\mathcal{C})$$

Propriétés de l'espérance conditionnelle

Indépendance

Si X est indépendante de la tribu $\mathcal{B}$ alors :

$$\mathbb{E}(X|\mathcal{B}) = \mathbb{E}(X)$$

Contraction pour la norme L^2

$$\mathbb{E}\Big(|\ \mathbb{E}(X|\mathcal{B})\ |^2\Big) \leq \mathbb{E}(|X|^2)$$

Remarque :

On pourra déduire la contraction pour la norme L^1.
En effet d'après la positivité de l'espérance conditionnelle on a :

$$|\ \mathbb{E}(X|\mathcal{B})\ | \leq \mathbb{E}(|X||\mathcal{B})$$

Ce qui donne que :

$$\mathbb{E}\Big(\ |\ \mathbb{E}(X|\mathcal{B})\ |\ \Big) \leq \mathbb{E}(|X|)$$

Propriétés de l'espérance conditionnelle

Lemme

Soient $\mathcal{B}$ une sous-tribu de $\mathcal{A}$, X et Y deux v.a.réelles telsque X est indépendante de $\mathcal{B}$ et Y est $\mathcal{B}$-mesurable. On considère $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction mesurable bornée alors :

$$\mathbb{E}(f(X, Y)|\mathcal{B}) = \psi(Y)$$

avec

$$\psi(y) = \mathbb{E}(f(X, y)), \ \forall y \in \mathbb{R}$$

Conditionnement : Cas discret

Définition

Soit (X, Y) un couple de variable aléatoires discrètes. On appelle distribution conditionnelle de X sachant $Y = y$ la fonction

$$F_{X|Y}(x) = \mathbb{P}(X \leq x|Y = y),$$

définie pour les y tels que $\mathbb{P}(Y = y) > 0$. La distribution conditionnelle de X sachant Y est la variable

$$F_{X|Y}(x) = \mathbb{P}(X \leq x|Y).$$

On peut aussi introduire les fonctions de masse conditionnelles

$$f_{X|Y=y}(x) = \mathbb{P}(X \leq x|Y = y)$$

$$f_{X|Y}(x) = \mathbb{P}(X \leq x|Y).$$

Exercice

Une poule pond un nombre aléatoire N d'oeufs qui suit une loi de Poisson de paramètre λ. Les oeufs éclosent avec une probabilité p indépendamment les uns des autres. Quelle est la loi du nombre X de poussins ?

Solution

Conditionnellement à N le nombre X suit une loi de binomial de paramètre N et p.c-a-d

$$\mathbb{P}(X = k | N = n) = C_k^n p^k (1-p)^{n-k}, \quad k = 0, \dots, n$$

Par conséquent

$$\mathbb{P}(X = k, N = n) = \mathbb{P}(X = k | N = n)\mathbb{P}(N = n) = C_k^n p^k (1-p)^{n-k} e^{-\lambda} \frac{\lambda^n}{n!}.$$

Puis

$$\mathbb{P}(X = k) = \sum_{n=k}^{+\infty} \frac{\lambda^n p^k (1-p)^{n-k} e^{-\lambda}}{k!(n-k)!}$$

En faisant le changement d'indice $m = n - k$ on trouve

$$\mathbb{P}(X = k) = e^{-\lambda p} \frac{(\lambda p)^k}{k!}$$

Donc X suit une loi de Poisson de paramètre λp.

Conditionnement : Cas discret

Définition

Soit X une v.a.d à valeurs dans $\{x_1, x_2, \ldots, x_n\}$. On suppose que X est intégrable. On appelle espérance conditionnelle de X sachant $Y = y$ la quantité

$$\mathbb{E}(X|Y = y) = \sum_i x_i \mathbb{P}(X = x_i|Y = y)$$

définie pour les y tels que $\mathbb{P}(Y = y) > 0$. On appelle espérance conditionnelle de X sachant Y la v.a.

$$\mathbb{E}(X|Y) = \sum_i x_i \mathbb{P}(X = x_i|Y)$$

Conditionnement : Cas continu

Définition

Soit X, Y un couple de v.a. ayant une densité $f_{(X,Y)}$. On appelle densité conditionnelle de X sachant $Y = y$ la fonction

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

définie pour tout y telle que $f_Y(y) > 0$.

Définition

On suppose que X est intégrable. On appelle espérance conditionnelle de X sachant $Y = y$ la quantité

$$\mathbb{E}(X|Y = y) = \int_{\mathbb{R}} x f_{X|Y=y}(x) dx$$

Exemple

Soit (X, Y) ayant pour densité la fonction la fonction

$$f_{(X,Y)}(x,y) = \frac{1}{x} 1_{0<y<x<1}(x,y).$$

La densité X est

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)dy = \frac{1}{x}(\int_0^x dy)1_{0<x<1}(x) = 1_{0<x<1}(x)$$

Autrement dit X suit une loi uniforme sur $[0,1]$. La densité conditionnelle de Y sachant X est donnée par

$$f_{Y|X=x}(y) = \frac{f_{(X,Y)}(x,y)}{f_X(x)} = \frac{1_{0<y<x}(y)}{x}$$

Pour tout $y \in \mathbb{R}$ et tout $x \in]0,1[$. Autrement dit, conditionnellement à X, la variable Y est uniforme sur $[0, X]$. En particulier

$$\mathbb{E}(Y|X) = \int_{\mathbb{R}} y f_{Y|X=x}(y)dy = \int_{\mathbb{R}} \frac{y}{x} 1_{0<y<x} dy$$
$$= \frac{1}{x}\int_0^x y dy = \frac{X}{2}.$$

Exercice

Soit $(B_t)_{t \geq 0}$ un mouvement brownien et $0 < s \leq t$.

1. Calculer $E[B_t | B_s]$ et $E[B_s | B_t]$.
2. Déterminer $E[B_t^2 | B_s]$ et $E[e^{\lambda B_t} | B_s]$, $\lambda \in \mathbb{R}$