Espérance conditionnelle

I. Introduction

L'esperance conditionnelle d'une v.a. X sachant une informat. I est la valeur moyenne attendue pour X lorsque l'on connait l'information I

Exemple 1: Lancer de 2 dés

X_1 : valeur obtenue par le 1er dé

X_2 : " " " le 2 dé

$S = X_1 + X_2$

Avant de lancer le 1er dé, on ne dispose d'aucune information sur S. Donc $E(S) = E(X_1) + E(X_2) = 7$

Avant de lancer le 2e dé, on connait la valeur de X_1. Avec cette information, on attend à avoir en moyenne pour S

$E(X_2) = \frac{1}{6}(1+2+ \dots +6) =$

$\sum x_i p_i$

$$X_1 + E(X_2) = X_1 + \frac{7}{2}$$

Cette quantité s'appelle l'espérance conditionnelle de S sachant X_1. Elle est notée $E(S/X_1)$

$P(A/B) = \frac{P(A \cap B)}{P(B)}$ $\frac{P(A \cap B)}{P(B)}$

Soit Ω : univers possible

$\mathcal{F}$: ensemble de tous les evenements possibles

$P(A)$: Proba. qu'un evenement $A \in \mathcal{F}$ ait lieu

II. Conditionnement par rapport à un evenement

Soit $X : \Omega \to \mathcal{X}$ avec $\mathcal{X} = \{x_1, \dots, x_n\}$ $\mathcal{X} \subset \mathbb{R}$

(1) $E(X/B) = \sum_{x \in \mathcal{X}} x \, P(X = x / B)$, $E(X/B) \in \mathbb{R}$

Exemple 2 Lancer de dé

X : valeur prise par le dé

B : $\{X \geq 3\}$

(1): $E(X/B) = \sum_{i=1}^{6} i \, P(X = i / B) = \frac{P(X = i \cap B)}{P(B)} = \frac{P((X=i) \cap B)}{\frac{4}{6}}$

$$= \frac{3}{2} P\left((X=i)\cap B\right) = \begin{cases} \frac{3}{2}\cdot 0 = 0 & \text{si } i=1,2 \\ \frac{3}{2}\cdot\frac{1}{6} = \frac{1}{4} & \text{sinon} \end{cases}$$

d'où $E(X/B) = \frac{1}{4}(3+4+5+6) = \frac{9}{2}$

Exemple 3 (voir exemple 1)

$B = \{X_1 = 5\}$ comme les evenements $\{X_2 = i-5\}$ et $\{X_1 = 5\}$ sont independants (X_1 et X_2 sont independants)

On a : $P(S=i/X_1=5) = P(X_2 = i-5/X_1=5) = P(X_2 = i-5)$

$$= \begin{cases} 0 & \text{si } i = 1,\dots,5,12 \\ \frac{1}{6} & \text{si } i = 6,\dots,11 \end{cases}$$

$$E(S/X_1=5) = \sum_{i=2}^{12} i\,P(S=i/X_1=5) = \frac{1}{6}(6+\dots+11)$$
$$= \frac{30}{6} + \frac{21}{6} = 5 + \frac{7}{2}$$

Generalement pour $j \in \{1,\dots,6\}$ on trouve $E(S/X_1=j) = \overset{X_1}{j} + \overset{E(X_2)}{\frac{7}{2}}$

III. Conditionnement par rapport à une (ou plusieurs) v.a.

a/ Cas discret :

a-1- Conditionnement par rapport a une v.a.

Exemple 4 :

1ère étape : On lance un dé

Z : la valeur obtenue

2ème étape : On relance Z fois le dé et On multiplie entre elles les valeurs obtenues, soit X ce produit

On doit calculer $E(X/Z=5)$ si $Z(\omega) = 5$

On relance 5 fois le dé (de manière independante)

On a alors $E(X/Z=5) = E(Z_1 \dots Z_5) = E(Z_1)\dots E(Z_5)$

On trouve $E(X/Z=5) = m\dots m = m^5$ avec $m = \frac{1}{6}(1+\dots+6)$

$m = \frac{7}{2}$

Generalement si $i = 1, \dots, 6$

On a : $E(X/Z=i) = m^i$.

Lorsque l'on connait la valeur de Z, on s'attend à avoir en moyenne pour X la valeur m^Z

$$E(X/Z) = m^Z \quad \text{V.A!}$$

Si $Z(\omega) = i \Rightarrow E(X/Z)(\omega) = E(X/Z=i)$

a-2- Conditionnement par rapport à plusieurs v.a.

$X : \Omega \to \mathcal{X} \quad \mathcal{X} = \{x_1, \dots, x_n\} \subset \mathbb{R}$

$Z : \Omega \to \mathcal{Z} \; ; \; \mathcal{Z} = \{z_1, \dots, z_n\} \subset \mathbb{R}$

Définition:

L'esperance conditionnelle de X sachant Z $(E(X/Z))$ est une v.a. définie par : $E(X/Z) : \Omega \to \mathbb{R}$

$$\omega \mapsto h(Z(\omega))$$

où $h : \mathcal{Z} \to \mathbb{R}$ définie $h(z) = E(X/Z=z) \quad \forall z \in \mathcal{Z}$

La quantité $E(X/Z=z)$ est celle definie par (1)

avec $B = \{Z = z\}$

$E(X/Z=z)$ nbr réel

$E(X/Z)$ v.a. ($E(X/Z)$ depend de ω car $Z(\omega)$ depend de ω)

Exemple 5:

Z v.a uniforme sur $\{1, \dots, n\}$

$\mathcal{E}$ v.a indep de Z t.q. $P(\mathcal{E}=1) = p$
$P(\mathcal{E}=-1) = 1-p$

On pose $X = \mathcal{E} Z$ v.a. qui prend ses valeurs dans $\{-n, \dots, -1, 1, \dots, n\}$

On a $E(X/Z=j) = \sum_{i=-n}^{n} i\, P(X=i/Z=j)$

avec $P(X=i/Z=j) = 0$ si $i \notin \{-j, j\}$ (car $X = \pm Z$)

Reste à calculer $E(X/Z=j) = -j\, P(X=-j/Z=j) + j\, P(X=j/Z=j)$

$$= -j \underbrace{P(\mathcal{E}=-1/Z=j)}_{P(\mathcal{E}=-1)} + j \underbrace{P(\mathcal{E}=1/Z=j)}_{P(\mathcal{E}=1)}$$

$= -j(1-p) + j(p) = j(2p-1)$

<u>Conclusion</u>: $E(X/Z) = (2p-1)Z$

20/10/2016

$X : \Omega \to \mathcal{X}$

$Z_1 : \Omega \to Z_1$

$\vdots$

$Z_k : \Omega \to Z_n$

$\mathcal{X} = \{x_1, \dots, x_m\} \subset \mathbb{R}$

$Z_1 = \{z_{k_1,1}, \dots, z_{k_1,n}\} \subset \mathbb{R}$

$\vdots$

$Z_n = \{z_{k_n,1}, \dots, z_{k_n,n}\} \subset \mathbb{R}$

$E(X/Z_1 = z, \dots, Z_n = z_n) = \sum_{x \in \mathcal{X}} x P(X = x / Z_1 = z_1, \dots, Z_n = z_n)$

nombre réel!

<u>Déf</u>: On appelle esperance conditionnelle de X sachant $Z_1, \dots, Z_n$ la v.a $E(X/Z_1, \dots, Z_n)$ définie par :

$$E(X/Z_1, \dots, Z_n)(\omega) = h(Z_1(\omega), \dots, Z_n(\omega))$$

où $h : Z_1 \times Z_2 \times \dots \times Z_n \to \mathbb{R}$ est la fonction à n variables définie par $h(z_1, \dots, z_n) = E(X/Z = z_1, \dots, Z = z_n)$

$$\forall \underset{Z_1}{\underset{\cap}{z_1}}, \dots, \underset{Z_n}{\underset{\cap}{z_n}}$$

IV. Cadre à densité :

Deux v.a : $X : \Omega \to \mathbb{R}$

$X : \Omega \to \mathbb{R}$

possédent une densité jointe s'il existe une fct $f_{X,Y} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^+$ t.q $\forall D \subset \mathbb{R}^2$, $P((X,Y) \in D) = \int_D f_{X,Y}(x,y)\, dx\, dy$

Densités marginales

$$f_X(x) = \int_{y \in \mathbb{R}} f_{X,Y}(x,y)\, dy$$

$$f_Y(x) = \int_{x \in \mathbb{R}} f_{X,Y}(x,y)\, dx$$

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

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

Déf :

Si X et Z 2 v.a.r (avec X intégrable) ayant une densité jointe $f_{(X,Z)}$, l'espérance conditionnelle de X sachant Z est la v.a : $E(X/Z) : \Omega \to \mathbb{R}$

$$w \rightsquigarrow h(Z(w))$$

Où $h : \mathbb{R} \to \mathbb{R}$ est définie par : $h(z) = \int_{x \in \mathbb{R}} x f_X(x / Z = z)\, dz$

IV. Caractérisation et propriétés

$\sigma(Z_1, \ldots, Z_n)$ l'information donnée par n variables aléatoires $Z_1, \ldots, Z_n$ lorsqu'il sera possible de prédire la v.a d'une v.a. Y à partir de l'information $\sigma(Z_1, \ldots, Z_n)$.

On dira que Y est $\sigma(Z_1, \ldots, Z_n)$ mesurable

Déf :

On considère $n+1$ v.a

$$Y : \Omega \to \mathbb{R}$$

$$Z_1 : \Omega \to \mathbb{R}$$

$$\vdots$$

$$Z_n : \Omega \to \mathbb{R}$$

La v.a. Y est dite $\sigma(Z_1, \ldots, Z_n)$ mesurable s'il existe une fonction $h : \mathbb{R}^n \to \mathbb{R}$ (mesurable) tel que $Y(\omega) = h(Z_1(\omega), \dots, Z_n(\omega))$ pour tout $\omega \in \Omega$. On note $\mathcal{F}_n = \sigma(Z_1, \dots, Z_n)$ et on dit Y est $\mathcal{F}_n$-mesurable.

Exemple :

$n = 1$. Une v.a Y est dite $\sigma(Z)$-mesurable s'il existe une fonction h (mesurable) $h : \mathbb{R} \to \mathbb{R}$ t.q $Y = h(Z)$

$$\begin{tikzcd} \Omega \arrow[rr] \arrow[dr, "Z"'] & & \mathbb{R} \\ & \mathbb{R} \arrow[ur, "h"'] & \end{tikzcd}$$

Façon moins formelle

Y est $\sigma(Z_1, \dots, Z_n)$ mesurable lorsque "en connaissant les valeurs de $Z_1(\omega), \dots, Z_n(\omega)$, on connait la valeur de $Y(\omega)$", grâce à la relation $Y(\omega) = h(Z_1(\omega), \dots, Z_n(\omega))$

Exemple :

les v.a. Y suivantes sont $\sigma(Z_1, \dots, Z_n)$ mesurables :

a) $Y = Z_1 + \dots + Z_n$ car $Y(\omega) = h(Z_1(\omega), \dots, Z_n(\omega))$
avec $h : \mathbb{R} \times \dots \times \mathbb{R} \to \mathbb{R}$ définie par $h(z_1, \dots, z_n) = z_1 + \dots + z_n$

b) $Y = Z_1^2$ avec $h(z_1, \dots, z_n) = z_1^2$.

c) $Y = \max(Z_1, \dots, Z_n)$ avec $h(z_1, \dots, z_n) = \max(z_1, \dots, z_n)$

d) $Y = Z_1 \times \dots \times Z_{n-1}$ " " $= z_1 \times \dots \times z_{n-1}$

f) $Y = \begin{cases} 1 & \text{si } Z_1 = 1 \\ 0 & \text{sinon} \end{cases}$ avec $h(z_1, \dots, z_n) = \begin{cases} 1 & \text{si } z_n = 1 \\ 0 & \text{sinon} \end{cases}$

Caractérisation : On considère $(n+1)$ v.a. $X : \Omega \to \mathbb{R}$

$Z_1 : \Omega \to \mathbb{R}$

$\vdots$

$Z_n : \Omega \to \mathbb{R}$

(avec X integrable). On note $\mathcal{F}_n = \sigma(Z_1, \dots, Z_n)$

L'espérance conditionnelle de X sachant $\mathcal{F}_n^\circ$, $E(X/\mathcal{F}_n^\circ)$ ou $E(X/Z_1, \dots, Z_n)$ est l'unique variable aléatoire vérifiant les 2 conditions :

a/ $E(X/\mathcal{F}_n^\circ)$ est $\mathcal{F}_n^\circ$-mesurable

b/ Pour toute v.a $\mathcal{F}_n^\circ$-mesurable (et bornée) Y, l'égalité

$$E\left[E(X/\mathcal{F}_n^\circ)\,Y\right] = E(XY)$$ a lieu.

Les conditions a/ et b/ traduisent que l'esperance conditionnelle $E(X/\mathcal{F}_n^\circ)$ est la meilleur approximation que l'on peut avoir de X par une v.a. $\mathcal{F}_n^\circ$-mesurable

Remarque1:

Lorsque $X, Z_1, \dots, Z_n$ prennent un nombre fini de valeurs, ou lorsque X et Z ont une densité jointe, les formules (precedentes) verifient a) et b)

Remarque2 On reformule a) par a'), b/ par b')

a') Il existe une fonction h (mesurable) t.q : $E(X/\mathcal{F}_n^\circ) = h(Z_1, \dots, Z_n)$

b') $E\left(E(X/\mathcal{F}_n^\circ)\, f(Z_1, \dots, Z_n)\right) = E\left(X f(Z_1, \dots, Z_n)\right)$

Pour toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ (mesurable et bornée)

Remarque3: "L'unicité" est à comprendre dans le sens suivant : toute v.a vérifiant a/ et b/ est indiscernable de $E(X/\mathcal{F}_n^\circ)$

Indiscernable : 2 v.a X et Y sont indiscernable s'il existe $\Omega_1 \subset \Omega$ t.q $P(\Omega_1) = 1$

$$X(\omega) = Y(\omega) \text{ pour tout } \omega \in \Omega$$

27/10/2016

② Proprietés

$X, Z_1, \dots, Z_n$ v.a à valeurs dans $\mathbb{R}$. On note $\mathcal{F}_n^\circ = \sigma(Z_1, \dots, Z_n)$

P_1. $E\left[E(X/\mathcal{F}_n^\circ)\right] = E(X)$

P_2. Si X est $\mathcal{F}_n^\circ$-mesurable, alors $E(X/\mathcal{F}_n^\circ) = X$

P_3. Si X est indépendante de $Z_1, \dots, Z_n$, alors $E(X/\mathcal{F}_n^0) = E(X)$

P_4. Si $a, b \in \mathbb{R}$, alors $E(aX + bY/\mathcal{F}_n^0) = aE(X/\mathcal{F}_n^0) + bE(Y/\mathcal{F}_n^0)$

P_5. Si Y est $\mathcal{F}_n^0$-mesurable alors $E(YX/\mathcal{F}_n^0) = YE(X/\mathcal{F}_n^0)$

P_6. Pour tout $p \in \mathbb{N}$, On a $E[E(X/\mathcal{F}_{n+p}^0)/\mathcal{F}_n^0] = E(X/\mathcal{F}_n^0)$ la plus petit des Tribu

Interpretation initiative (On rappelle que $E(X/\mathcal{F}_n^0)$ est la meilleure approximation possible de X par une v.a $\mathcal{F}_n^0$-mesurable)

P_2. Si X est $\mathcal{F}_n^0$-mesurable, la meilleure approximation $\mathcal{F}_n^0$-mesurable de X est ... X elle même!

P_3. Si X est indépendante de $Z_1, \dots, Z_n$ alors ces dernières v.a n'apportent pas d'information sur X. Dans ce cas $E(X/\mathcal{F}_n^0)$ n'est pas aléatoire.

P_6. On rappelle que $E(X/\mathcal{F}_{n+p}^0)$ est la meilleure approximation de X qui soit $\mathcal{F}_{n+p}^0$-mesurable. Cette propriété traduit le fait que la meilleure approximation de cette quantité par une v.a $\mathcal{F}_n^0$-mesurable est $E(X/\mathcal{F}_n^0)$

Remarque 4: Une v.a. Y est $\mathcal{F}_n^0$-mesurable s'il existe une fonction h telle que

$Y = h(Z_1, \dots, Z_n)$. On pose $\tilde{h} : \mathbb{R}^{n+p} \longrightarrow \mathbb{R}$

$(z_1, \dots, z_{n+p}) \longrightarrow h(z_1, \dots z_n)$

On a : $Y = \tilde{h}(z_1, \dots, z_{n+p})$, donc Y est aussi $\mathcal{F}_{n+p}^0$-mesurable

conclusion:

$\mathcal{F}_n^0$-mesurable est aussi $\mathcal{F}_{n+p}^0$-mesurable

la reciproque est fausse

Conséquence:

$$E(E(X/\mathcal{F}_{n+p}^0)) = E(X/\mathcal{F}_n^0) \quad \forall p \in \mathbb{N}$$

cette dernière formule s'obtient grâce à P_2

la v.a $E(X/\mathcal{F}_n^0)$ est $\mathcal{F}_n^0$-mesurable donc $\mathcal{F}_{n+p}^0$-mesurable

d'où $E\left(E(X/\mathcal{F}_n^0)/\mathcal{F}_{n+p}^0\right) \overset{P_2}{=} E(X/\mathcal{F}_n^0)$

Exemples:

1- $S = X_1 + X_2$

On utilise la stratégie précédente :

X_1 est $\sigma(X_1)$-mesurable et X_2 est indep. de X_1 :

$$E(S/X_1) \overset{P_1}{=} E(X_1/X_1) + E(X_2/X_1)$$
$$\overset{P_2}{=} X_1 + E(X_2/X_1)$$
$$\overset{P_3}{=} X_1 + E(X_2) = X_1 + \frac{7}{2}$$

2- (Voir exemple 1)

$$R = \frac{X_2}{X_1}$$

Calcul de $E(R/X_1)$

comme $\frac{1}{X_1}$ est $\sigma(X_1)$-mesurable (car c'est une fonction de X_1), on a $E(R/X_1) \overset{P_5}{=} \frac{1}{X_1} E(X_2/X_1)$

$$\overset{P_3}{=} \frac{1}{X_1} E(X_2) = \frac{7}{2X_1}$$

3- (Voir exemple 5) $X = \varepsilon Z$

Z v.a uniforme sur $\{1, \dots, n\}$

ε indep de Z (avec $\underbrace{P(\varepsilon = 1)}_{p}$ et $\underbrace{P(\varepsilon = -1)}_{1-p}$)

On a : $E(X/Z) = E(\varepsilon Z/Z)$

$$\overset{P_5}{=} Z\, E(\varepsilon/Z)$$
$$\overset{P_3}{=} Z\, E(\varepsilon)$$
$$= (2p-1) Z$$

$E(\varepsilon) = \sum x_i p_i$

$E(\varepsilon) = 1(p) - 1(1-p)$

Propriétés :

1- Si $Y_1, \dots, Y_p$ sont indépendantes de X et de $Z_1, \dots, Z_p$ alors $E(X/Z_1, \dots, Z_n, Y_1, \dots, Y_p) = E(X/Z_1, \dots, Z_n)$