

## Cours 8

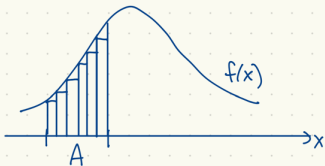
William McCausland

2020-10-27

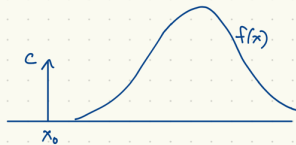
## Un peu de Chapitre 12

- ▶  $\mu, \nu, \lambda$  des mesures boréliennes sur  $\mathbb{R}$ ,  $\lambda$  Lebesgue.
- ▶  $\mu$  est absolument continue s'il existe  $f$  mesurable telle que  $\mu(A) = \int_A f(x)\lambda(dx)$ ,  $A$  borelien ( $\lambda$ -mesurable).
- ▶  $\mu$  est absolument continue par rapport à  $\nu$  s'il existe  $f$ ,  $\nu$ -mesurable, telle que  $\mu(A) = \int_A f(x)\nu(dx)$ ,  $A \in \mathcal{B}$ .
- ▶  $\mu$  est discret si  $\sum_{x \in \mathbb{R}} \mu(\{x\}) = \mu(\mathbb{R})$ .
- ▶  $\mu \ll \nu$  signifie  $\nu(A) = 0 \Rightarrow \mu(A) = 0$ .
- ▶ Théorème Radon-Nikodym :  $\mu \ll \lambda \Leftrightarrow$   
il existe  $f$ ,  $\lambda$ -mesurable, telle que  $\mu(A) = \int_A f(x)\lambda(dx)$ ,  $A$  mesurable.
- ▶  $f = \frac{d\mu}{d\lambda}$  est la dérivée Radon-Nikodym de  $\mu$  par rapport à  $\lambda$ .
- ▶  $\mu(A) = \int_A \left(\frac{d\mu}{d\lambda}\right) d\lambda$ .

# Continuité absolue



$$\mu(A) = \int f(x) \lambda(dx)$$



Problème:

$$\Pr[x_0 - \varepsilon, x_0 + \varepsilon] \sim \Pr\{\varepsilon x_0\} = c$$

$$\text{mais } \lambda(\varepsilon x_0) = 0$$

Pas de problème si  $\lambda \gg \mu$   $\lambda(A) = 0 \Rightarrow \mu(A) = 0$

$$A \in \mathcal{B}$$

## Un peu plus sur la mesurabilité

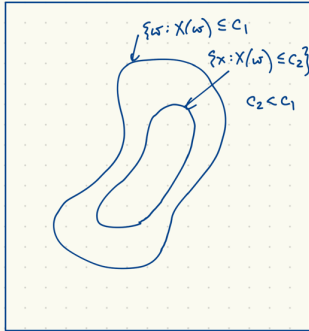
- ▶ Soit  $(\Omega, \mathcal{F}, P)$  un espace de probabilité,  $\mathcal{G} \subseteq \mathcal{F}$  une tribu.
- ▶ Soit  $X: \Omega \rightarrow \mathbb{R}$  une variable aléatoire sur  $(\Omega, \mathcal{F}, P)$ .
- ▶ Si  $X$  est  $\mathcal{G}$ -mesurable ( $\{X \leq c\} \in \mathcal{G}$ , tout  $c$  réel)
  - ▶  $\{X = c\} = \{X \leq c\} \setminus \bigcap_n \{X \leq c - 1/n\} \in \mathcal{G}$ .
- ▶ Si on sait assez pour répondre à la question suivante pour tous  $c$ , on connaît la valeur de  $X(\omega)$ .

*Est-ce que  $\omega \in \{X \leq c\}$ ?*

- ▶ Définie  $\sigma(X) \equiv \sigma(\{X \leq c\}: c \in \mathbb{R})$ .
- ▶ Si une variable aléatoire est  $\sigma(X)$ -mesurable, c'est une fonction seulement de  $X$ —elle dépend de  $\omega$  seulement à travers  $X(\omega)$ .
  - ▶ Supposons que  $Z(\omega)$  est  $\sigma(X)$ -mesurable.
  - ▶ Pour tous  $z \in \mathbb{R}$ , il existe  $B_z \in \mathcal{B}$  tel que  $\{Z = z\} = \{X \in B_z\}$ . Alors  $Z = f(X)$ , où  $f(x) = z$  pour tous  $x \in B_z$ .

## Quelques évènements dans $\sigma(X)$

$\Omega$



# Exemple, probabilités et espérances conditionnelles

$A_1$
$A_2$
$A_3$

$$\mathcal{G}_1 = \sigma(\{A_1, A_2, A_3, \Omega\})$$

$B_1$	$B_2$	$B_3$
-------	-------	-------

$$\mathcal{G}_2 = \sigma(\{B_1, B_2, B_3, \Omega\})$$

$C_{11}$	$C_{12}$	$C_{13}$
$C_{21}$	$C_{22}$	$C_{23}$
$C_{31}$	$C_{32}$	$C_{33}$

$$\mathcal{G} = \sigma(\mathcal{G}_1 \cup \mathcal{G}_2)$$

$\Omega$
----------

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

1
1
2

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

4	5	4
---	---	---


Soit  $p_{11} = P[C_{11}]$   
 $p_{12} = P[C_{12}]$   
 $\vdots$   
 $p_{33} = P[C_{33}]$

$$P(\Lambda | \mathcal{G}_1)$$

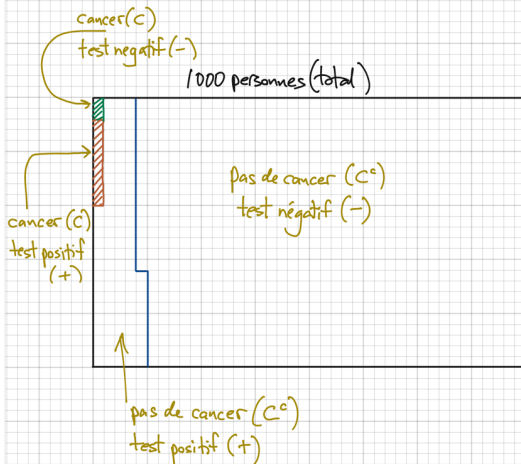
$$E[Y | \mathcal{G}_1]$$

$$P(\Lambda | X)$$

$$E[Y | X]$$

$$\Lambda = C_{13} \cup C_{21} \cup C_{22} \cup C_{32}$$

# Théorème de Bayes, événements de probabilité positive



$$\Pr[C] = 0.01$$

(10 personnes sur 1000)

$$\Pr[C^c] = 0.99$$

(990 personnes sur 1000)

$$\Pr[+|C] = 0.8$$

(8 personnes sur 10)

$$\Pr[-|C] = 0.2$$

(2 personnes sur 10)

$$\Pr[-|C^c] = 0.9$$

(891 personnes sur 990)

$$\Pr[+|C^c] = 0.1$$

(99 personnes sur 990)

$$\begin{aligned}\Pr[C|+] &= \frac{8}{8+99} = \frac{\Pr[C,+]}{\Pr[+]} \\ &= \frac{\Pr[C]\Pr[+|C]}{\Pr[+]}\end{aligned}$$

# Définitions de probabilité, espérance conditionnelle

- ▶ Soit  $X, Y$  deux variables aléatoires avec  $E[|Y|] < \infty$
- ▶ Soit  $A$  un évènement :  $A \in \mathcal{F}$ .
- ▶ Une variable aléatoire  $P(A|X)(\omega)$  est une *probabilité conditionnelle* de  $A$  sachant  $X$  si elle est  $\sigma(X)$ -mesurable et pour chaque  $S \in \mathcal{B}$ ,

$$E[P(A|X)1_{X \in S}] = P[A \cap \{X \in S\}].$$

- ▶ Une variable aléatoire  $E[Y|X](\omega)$  est une *espérance conditionnelle* de  $Y$  sachant  $X$  si elle est  $\sigma(X)$ -mesurable et pour chaque  $S \in \mathcal{B}$ ,

$$E[E[Y|X]1_{X \in S}] = E[Y1_{X \in S}].$$



## Trouver $P(\Lambda|\mathcal{G}_1)$

Conditions pour  $P(\Lambda|\mathcal{G}_1)$

- $\mathcal{G}_1$  mesurable ( $\mathcal{G}_1 = \sigma(A_1, A_2, A_3, \Omega)$ )

$$E[P(\Lambda|\mathcal{G}_1)1_{A_1}] = P(\Lambda \cap A_1)$$

$$E[P(\Lambda|\mathcal{G}_1)1_{A_2}] = P(\Lambda \cap A_2)$$

$$E[P(\Lambda|\mathcal{G}_1)1_{A_3}] = P(\Lambda \cap A_3)$$

Autres éléments de  $\mathcal{G}_1$  par linéarité.

## Probabilités conditionnelles de l'événement $\Lambda \in \Omega$

- Par rapport à la sous-tribu  $\mathcal{G}_1$  :

$$P(\Lambda|\mathcal{G}_1) = \begin{cases} p_{13}/P(A_1) & \omega \in A_1 \\ (p_{21} + p_{22})/P(A_2) & \omega \in A_2 \\ p_{32}/P(A_3) & \omega \in A_3 \end{cases}$$

- Par rapport à la variable aléatoire  $X$

$$P(\Lambda|X) = P(\Lambda|\sigma(X)) = \begin{cases} \frac{p_{13}+p_{21}+p_{22}}{P(A_1 \cup A_2)} & \omega \in A_1 \cup A_2 = \{X = 1\} \\ p_{32}/P(A_3) & \omega \in A_3 = \{X = 2\} \end{cases}$$

- Par rapport à la sous-tribu minimale  $\{\emptyset, \Omega\}$  :

$$P(\Lambda|\{\emptyset, \Omega\}) = P(\Lambda) = p_{13} + p_{21} + p_{22} + p_{32}, \text{ tous } \omega \in \Omega$$

- Par rapport à la sous-tribu maximal  $\mathcal{F}$  (ou par rapport à  $\mathcal{G}$ ) :

$$P(\Lambda|\mathcal{F}) = 1_\Lambda(\omega) = P(\Lambda|\mathcal{G}).$$

## Vérification de $P(\Lambda|\mathcal{G}_1)$

- ▶ À vérifier :  $E[P(\Lambda|\mathcal{G}_1)1_A] = P(\Lambda \cap A)$ ,  $A \in \mathcal{G}_1$ .

$$E[P(\Lambda|\mathcal{G}_1)1_{A_1}] = E\left[\frac{p_{13}}{P(A_1)}1_{A_1}\right] = \frac{p_{13}}{P(A_1)}E[1_{A_1}] = p_{13} = P(\Lambda \cap A_1)$$

$$\begin{aligned}E[P(\Lambda|\mathcal{G}_1)1_{A_2}] &= E\left[\frac{p_{21} + p_{23}}{P(A_2)}1_{A_2}\right] = \frac{p_{21} + p_{22}}{P(A_2)}E[1_{A_2}] \\&= p_{21} + p_{22} = P(\Lambda \cap A_2)\end{aligned}$$

$$E[P(\Lambda|\mathcal{G}_1)1_{A_3}] = E\left[\frac{p_{32}}{P(A_3)}1_{A_3}\right] = \frac{p_{32}}{P(A_3)}E[1_{A_3}] = p_{32} = P(\Lambda \cap A_3)$$

- ▶ Le reste par linéarité de l'espérance, additivité de probabilité

## Construction de $P(\Lambda|\mathcal{G}_1) = \frac{d\nu}{dP_0}$

- La mesure  $\nu$  :  $\nu(A) \equiv P(\Lambda \cap A)$ ,  $A \in \mathcal{G}_1$

$$\nu(A_1) = P(\Lambda \cap A_1) = p_{13}$$

$$\nu(A_2) = P(\Lambda \cap A_2) = p_{21} + p_{22}$$

$$\nu(A_3) = P(\Lambda \cap A_2) = p_{32}$$

- La mesure  $P_0$  :  $P_0(A) \equiv P(A)$ ,  $A \in \mathcal{G}_1$ .

$$P_0(A_1) = p_{11} + p_{12} + p_{13}$$

$$P_0(A_2) = p_{21} + p_{22} + p_{23}$$

$$P_0(A_3) = p_{31} + p_{32} + p_{33}$$

- Notez que  $P_0(A) = 0 \Rightarrow \nu(A) = 0$ ,  $A \in \mathcal{G}_1$  : c-à-d  $\nu \ll P_0$ .
- $P(\Lambda|\mathcal{G}_1)(\omega) \equiv \nu(A)/P_0(A)$ ,  $\omega \in A \in \mathcal{G}_1$ .

## Espérances conditionnelles de $Y$

- ▶ Par rapport à la sous-tribu  $\mathcal{G}_1$  :

$$E[Y|\mathcal{G}_1] = \begin{cases} (4(p_{11} + p_{13}) + 5p_{12})/P(A_1) & \omega \in A_1 \\ (4(p_{21} + p_{23}) + 5p_{22})/P(A_2) & \omega \in A_2 \\ (4(p_{31} + p_{33}) + 5p_{32})/P(A_3) & \omega \in A_3 \end{cases}$$

- ▶ Par rapport à la variable aléatoire  $X$

$$E[Y|X] = \begin{cases} \frac{4(p_{11}+p_{13}+p_{21}+p_{23})+5(p_{12}+p_{22})}{P(A_1 \cup A_2)} & \omega \in A_1 \cup A_2 = \{X = 1\} \\ (4(p_{31} + p_{33}) + 5p_{32})/P(A_3) & \omega \in A_3 = \{X = 2\} \end{cases}$$

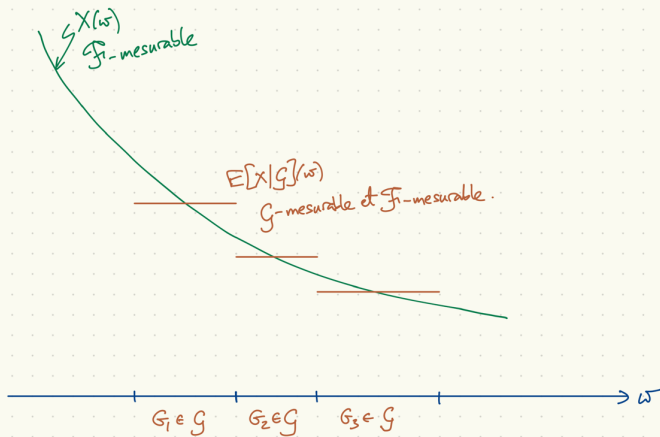
- ▶ Par rapport à la sous-tribu minimale  $\{\emptyset, \Omega\}$  :

$$E[Y|\{\emptyset, \Omega\}] = E[Y]$$

- ▶ Par rapport à la sous-tribu maximal  $\mathcal{F}$  :

$$E[Y|\mathcal{F}] = Y(\omega)$$

# Espérance conditionnelles : une illustration



## Vérification de $E[Y|\mathcal{G}_1]$

- ▶ À vérifier :  $E[E[Y|\mathcal{G}_1]1_A] = E[Y1_A]$ ,  $A \in \mathcal{G}_1$  :
- ▶ Pour  $A = A_1$  :

$$E[E[Y|\mathcal{G}_1]1_{A_1}] = \frac{4(p_{11} + p_{13}) + 5p_{12}}{P(A_1)} E[1_{A_1}] = 4(p_{11} + p_{13}) + 5p_{12}$$

$$E[Y1_{A_1}] = 4(p_{11} + p_{13}) + 5p_{12}$$

- ▶  $A = A_2$ ,  $A = A_3$  semblables
- ▶ Le reste par linéarité de l'espérance

## Construction de $E[Y|\mathcal{G}_1]$ :

- ▶ En général,  $E[Y|\mathcal{G}_1](\omega) = E[Y^+|\mathcal{G}_1](\omega) - E[Y^-|\mathcal{G}_1](\omega)$ .
- ▶ Mêmes cas  $\infty$ ,  $-\infty$ , fini, indéfini, événement par événement
- ▶ Ici,  $Y = Y^+$ , alors  $E[Y|\mathcal{G}_1] = E[Y^+|\mathcal{G}_1] = \frac{d\rho^+}{dP_0}$ , où

$$\rho^+(A) \equiv E[Y^+1_A], \quad P_0(A) \equiv P(A), \quad A \in \mathcal{G}_1,$$

et notez que  $\rho^+ \ll P_0$  alors  $\rho^+(A) = \int_A E[Y|\mathcal{G}_1]P_0(dx)$ .

- ▶ Pour  $A = A_1$ ,
  - ▶  $\rho^+(A_1) = E[Y^+1_{A_1}] = 4(p_{11} + p_{13}) + 5p_{12}$
  - ▶  $P_0(A_1) = P(A_1) = p_{11} + p_{12} + p_{13}$
- ▶ Les cas  $A = A_2$ ,  $A = A_3$  sont semblables.
- ▶ Pour chaque  $\omega \in A \in \mathcal{G}_1$ ,  $E[Y^+|\mathcal{G}_1](\omega) = \rho^+(A)/P_0(A)$ .
- ▶ Pour  $\omega \in A_1$ ,

$$E[Y^+|\mathcal{G}_1](\omega) = \frac{4(p_{11} + p_{13}) + 5p_{12}}{p_{11} + p_{12} + p_{13}}.$$



## Exercice 13.2.3

- ▶ Soit  $\mathcal{G}_1$  et  $\mathcal{G}_2$  deux sous-tribus de  $\mathcal{F}$ .
- (a) Si  $Z$  est  $\mathcal{G}_1$ -mesurable et  $\mathcal{G}_1 \subseteq \mathcal{G}_2$ ,  $Z$  est  $\mathcal{G}_2$  mesurable :
  - ▶ Pour tous  $z \in \mathbb{R}$ ,  $\{Z \leq z\} \in \mathcal{G}_1$  alors  $\{Z \leq z\} \in \mathcal{G}_2$ .
- (b) Si  $Z$  est  $\mathcal{G}_1$ -mesurable et  $\mathcal{G}_2$ -mesurable,  $Z$  est  $(\mathcal{G}_1 \cap \mathcal{G}_2)$ -mesurable :
  - ▶ Pour tous  $z \in \mathbb{R}$ ,  $\{Z \leq z\} \in \mathcal{G}_1$  et  $\{Z \leq z\} \in \mathcal{G}_2$ , alors  $\{Z \leq z\} \in \mathcal{G}_1 \cap \mathcal{G}_2$ .

# Intersection des tribus



## Proposition 13.2.6

- Rappel, définition de la v.a.  $E[X|\mathcal{G}]$  : pour tout  $G \in \mathcal{G}$ ,

$$E[E[X|\mathcal{G}]1_G] = E[X1_G].$$

- Soit  $X, Y$  des variables aléatoires,  $X$  est  $\mathcal{G}$ -mesurable,  $E[Y] < \infty$ ,  $E[XY] < \infty$ .
- Proposition :  $E[XY|\mathcal{G}] = XE[Y|\mathcal{G}]$  avec probabilité 1.
- Preuve :
  - Soit  $G_0, G \in \mathcal{G}$ ,  $X = 1_{G_0}$ . Alors

$$E[XE[Y|\mathcal{G}]1_G] = E[E[Y|\mathcal{G}]1_{G \cap G_0}] = E[Y1_{G \cap G_0}] = E[XY1_G].$$

$$E[E[XY|\mathcal{G}]1_G] = E[XY1_G].$$

- $G$  est arbitraire, alors  $XE[Y|\mathcal{G}] = E[XY|\mathcal{G}]$  avec probabilité 1, tous  $X = 1_{G_0}$ .
- $G_0$  est arbitraire, alors la même chose tient pour  $X$  simple (linéarité), positive (convergence dominée), générale.

## Proposition 13.2.7 (espérances itérées)

- ▶ Définition de  $E[X|\mathcal{G}]$ :  $E[E[X|\mathcal{G}]1_G] = E[X1_G]$ , tous  $G \in \mathcal{G}$ .
- ▶ Proposition : Si  $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$ ,  $E[E[Y|\mathcal{G}_2]|\mathcal{G}_1] = E[Y|\mathcal{G}_1]$ .
- ▶ Preuve : fixez  $G \in \mathcal{G}_1 \subseteq \mathcal{G}_2$ ,

$$E[ E[E[Y|\mathcal{G}_2]|\mathcal{G}_1] 1_G ] = E[E[Y|\mathcal{G}_2]1_G] = E[Y1_G]$$

$$E[E[Y|\mathcal{G}_1]1_G] = E[Y1_G]$$

alors

$$E[E[Y|\mathcal{G}_2]|\mathcal{G}_1] = E[Y|\mathcal{G}_1] \text{ avec probabilité 1.}$$

- ▶ Cas spécial, espérance conditionnelle comme projection :

$$E[E[Y|\mathcal{G}]|\mathcal{G}] = E[Y|\mathcal{G}]$$

- ▶ Deux autres cas spéciaux :

- ▶  $E[E[X|Y]] = E[X]$  pour  $\mathcal{G}_1 = \{\emptyset, \Omega\} \subseteq \mathcal{G}_2 = \sigma(Y)$ .
- ▶  $E[E[X|Y, Z]|Z] = E[X|Z]$  pour  $\mathcal{G}_1 = \sigma(Z) \subseteq \mathcal{G}_2 = \sigma(Y, Z)$

## Loi de covariance total

La loi de covariance totale :

$$\text{Cov}[X, Y] = E[\text{Cov}[X, Y|Z]] + \text{Cov}[E[X|Z], E[Y|Z]]$$

Preuve: soit  $m_X \equiv E[X] = E[E[X|Z]]$ ,  $m_Y \equiv E[Y] = E[E[Y|Z]]$ .  
Alors

$$\begin{aligned}\text{Cov}[X, Y] &= E[E[(X - m_X)(Y - m_Y)]|Z] \\ &= E[E[(X - E[X|Z] + E[X|Z] - m_X) \\ &\quad (Y - E[Y|Z] + E[Y|Z] - m_Y)|Z]].\end{aligned}$$

Puisque  $E[(E[X|Z] - m_X)(Y - E[Y|Z])|Z] = 0$ ,

$$\begin{aligned}\text{Cov}[X, Y] &= E[\text{Cov}[X, Y|Z]] + E[(E[X|Z] - m_X)(E[Y|Z] - m_Y)] \\ &= E[\text{Cov}[X, Y|Z]] + \text{Cov}[E[X|Z], E[Y|Z]].\end{aligned}$$

## Apérçu du cours 9 (Casella et Berger)

- ▶ Statistiques exhaustives (sufficient), complètes, minimales, libres (ancillary)
- ▶ Estimation ponctuelle, méthode des moments et maximum de vraisemblance
- ▶ L'approche bayésienne et les lois a priori, conjointe et a posteriori
- ▶ Estimation ponctuelle bayésienne