Probabilités, L3

Université de Picardie Jules Verne

S6, 2021

Cours 2, probabilités L3

On fixe Ω espace d'états, $\mathcal{A} \subset \mathcal{P}(\Omega)$ une tribu d'événements.

Prop. Soit $P : \mathcal{A} \to [0,1]$ une mesure de probabilité, et $(A_i)_{i\in I}$ une famille dénombrable d'événements dans $\mathcal{A}$. Alors

$$P\left(\bigcup_{i\in I} A_i\right) \leq \sum_{i\in I} P(A_i)$$

dém. (a) si I est fini on le montre par récurrence sur $\# I$

- $\# I = 1$. évident.
- si vrai pour $\# I = k$, montrons-le pour $(k+1)$

On se donne $(A_j)_{j\in J} \in \mathcal{A}^J$, $\# J = k+1$

$J = \{j_1, \dots, j_k, j_{k+1}\}$ $\qquad I := \{j_1, \dots, j_k\}$

$$P\Big(\bigcup_{j\in J} A_j\Big) = P\Big(\bigcup_{i\in I} A_i \cup A_{j_{k+1}}\Big)$$

$$= P\Big(\bigcup_{i\in I} A_i\Big) + P\big(A_{j_{k+1}}\big) - P\Big(\bigcup_{i\in I} A_i \cap A_{j_{k+1}}\Big)$$

$$\overset{HR}{\leq} \sum_{i\in I} P(A_i) + P\big(A_{j_{k+1}}\big)$$

$$= \sum_{j\in J} P(A_j)$$

(b) Supposons I dénombrable ; sans restriction, on suppose $I = \mathbb{N}^*$

Soit $B_n := \bigcup_{i=1}^{n} A_i \leadsto (B_n)_{n\in\mathbb{N}^*}$ croît vers $\bigcup_{i\in I} A_i =: B$.

Par (a) on a $P(B_n) \leq \underbrace{\sum_{i=1}^{n} P(A_i)}_{p_n}$

D'après la proposition vue précédemment (dernière prop du cours 1),

$$P(B_n) \nearrow P(B)$$

tandis que $(p_n) \nearrow$ vers $\sum_{i \in I} P(A_i)$. En passant à la limite sur n, on a.

$$P(B) = P\left(\bigcup_{i \in I} A_i\right) \leq \sum_{i \in I} P(A_i) \qquad \blacksquare$$

Soit $X : \Omega \to E$ une application à valeurs dans un ensemble E
$(E = \mathbb{R}, \mathbb{R}^k, \mathbb{N}$ etc$)$

Si $P : \mathcal{A} \to [0,1]$ est une probabilité,

pour $B \subset E$, $P_X(B) = P(X^{-1}(B))$ n'est défini que si $X^{-1}(B) \in \mathcal{A}$,
ce qui nous amène au résultat suivant.

Prop. : 1) La famille $\mathcal{E}$ des parties B de E telles que $X^{-1}(B) \in \mathcal{A}$ est une tribu de E

2) La formule $\mathcal{E} \ni B \longmapsto P_X(B) = P(X^{-1}(B))$ définit une proba. sur E, appelée la **loi** de X.

dém : Soit $\mathcal{E} = \{ B \in \mathcal{P}(E) \text{ tq. } X^{-1}(B) \in \mathcal{A} \} \subset \mathcal{P}(E)$

(A1) $X^{-1}(\emptyset) = \emptyset \in \mathcal{A}$ donc $\emptyset \in \mathcal{E}$ $\quad$ ($X^{-1}(E) = \Omega \in \mathcal{A}$ aussi)

(A2) Soit $B \in \mathcal{E}$. $\quad X^{-1}(B^c) = \{ \omega \in \Omega \text{ tq. } X(\omega) \notin B \}$

$$= \{ \omega \in \Omega \text{ tq } X(\omega) \in B \}^c$$

$$= (X^{-1}(B))^c \in \mathcal{A}$$

donc $B^c \in \mathcal{E}$

(A4) Soit $(B_k)_{k\in K}$, K dénombrable, $B_k \in \mathcal{E}$ pour tout $k \in K$.

Alors $X^{-1}(B_k) \in \mathcal{A}$ pour tout $k \in K$

$$\Rightarrow \bigcap_{k\in K} X^{-1}(B_k) = X^{-1}\Big(\bigcap_{k\in K} B_k\Big) \in \mathcal{A}$$

(réunion : $\left(X^{-1}\Big(\bigcup_{k\in K} B_k\Big)\right)^c = \left(\bigcup_{k\in K} X^{-1}(B_k)\right)^c = \bigcap_{k\in K} \underbrace{X^{-1}(\underset{\in\mathcal{E}}{B_k^c})}_{\in\mathcal{A}} \in \mathcal{A}$ etc)

Remarque on peut avoir $\mathcal{E} \neq \mathcal{P}(E)$, même si $\mathcal{A} = \mathcal{P}(\Omega)$

Déf. Une famille $\mathcal{M}$ de parties d'un ensemble E est appelée **classe monotone** si

(M1) $E \in \mathcal{M}$

(M2) $A, B \in \mathcal{M}$, $A \subset B \Rightarrow B \setminus A \in \mathcal{M}$ (stable par soustraction)

(M3) $(A_n)_{n\in\mathbb{N}}$ suite $\nearrow$ d'éléments de $\mathcal{M}$ $\Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$

(stable par réunion croissante)

Pour toute famille $\mathcal{F}$ de parties de E, on note $\mathcal{M}_0(\mathcal{F})$ la classe monotone engendrée par $\mathcal{F}$ (intersection de toutes les classes monotones contenant $\mathcal{F}$).

Rem . une tribu est une classe monotone donc $\forall \mathcal{F}$ famille de parties de E,

$$\mathcal{M}_0(\mathcal{F}) \subset \sigma(\mathcal{F})$$

une classe monotone stable par intersection finie est une tribu

dém . $(M1) \Rightarrow (A1)$, et $(M2) \Rightarrow$ stabilité par passage au complémentaire

$(A^c = \Omega \setminus A$, en part $\emptyset = \Omega \setminus \Omega)$

· stabilité par union finie . $\displaystyle\bigcup_{i=1}^{n} A_i = \left(\left(\bigcup_{i=1}^{n} A_i\right)^c\right)^c$

$$= \left(\bigcap_{i=1}^{n} A_i^c\right)^c \text{ etc}$$

• stabilité par réunion dénombrable :

Soit $(A_n)_{n \in \mathbb{N}^*}$, $A_n \in \mathcal{M}$ pour tout $n \in \mathbb{N}$

⇝ $B_n := \bigcup_{i=1}^{n} A_i \in \mathcal{M}$ (stabilité par réunion finie)

$\Rightarrow (B_n)_{n \in \mathbb{N}^*}$ est croissante donc $\bigcup_{n \in \mathbb{N}^*} B_n \in \mathcal{M}$

mais $= \bigcup_{n \in \mathbb{N}^*} A_n$ $\blacksquare$

Théorème. Soit $\mathcal{F}$ famille de parties d'un ensemble E

Si $\mathcal{F}$ est stable par intersection finie, alors $\mathcal{M}(\mathcal{F}) = \sigma(\mathcal{F})$

dém : $\mathcal{M}(\mathcal{F}) \subset \sigma(\mathcal{F})$ d'après ce qu'on a vu

⇝ Il suffit de m.q. $\mathcal{M}(\mathcal{F})$ est une tribu, donc par ce qui précède, de m.q. $\mathcal{M}(\mathcal{F})$ est stable par intersection finie

Soit $\mathcal{M}_1 := \{ A \in \mathcal{M}(\mathcal{F}) \text{ tq. } \forall C \in \mathcal{F},\ A \cap C \in \mathcal{M}(\mathcal{F}) \}$

Par hypothèse sur $\mathcal{F}$, $\mathcal{F} \subset \mathcal{M}_1$

De plus, $\mathcal{M}_1$ est une classe monotone.

$E \in \mathcal{M}_1$

si $A, B \in \mathcal{M}_1$, $A \subset B$, $C \in \mathcal{F}$

$$(B \setminus A) \cap C = \underbrace{(B \cap C)}_{\in \mathcal{M}(\mathcal{F})} \setminus \underbrace{(A \cap C)}_{\in \mathcal{M}(\mathcal{F})} \in \mathcal{M}(\mathcal{F})$$

donc $B \setminus A \in \mathcal{M}_1$

- stabilité par union dénombrable croissante. pour $(A_n) \nearrow$ dans $\mathcal{M}_1$, $C \in \mathcal{F}$,

$$\Leftarrow \left(\bigcup_{n \in \mathbb{N}} A_n \right) \cap C = \bigcup_{n \in \mathbb{N}} \underbrace{(A_n \cap C)}_{\in \mathcal{M}(\mathcal{F})} \in \mathcal{M}(\mathcal{F})$$

Conséquence. $\mathcal{M}_1$ est une classe monotone contenant $\mathcal{F}$, donc $\mathcal{M}_1 = \mathcal{M}(\mathcal{F})$.

Maintenant, posons $\mathcal{M}_2 := \{A \in \mathcal{M}(\mathcal{F}) \text{ t.q. } \forall C \in \mathcal{M}(\mathcal{F}), A \cap C \in \mathcal{M}(\mathcal{F})\}$.

De même $\mathcal{M}_2$ est une classe monotone incluse dans $\mathcal{M}(\mathcal{F})$ et contenant $\mathcal{F}$ donc $\mathcal{M}_2 = \mathcal{M}(\mathcal{F})$

Par suite, $\mathcal{M}(\mathcal{F})$ est stable par intersection finie, ce qui conclut ∎

Corollaire. (unicité des mesures)

Soit Ω un ensemble muni d'une tribu $\mathcal{A}$.

Soit $\mathcal{C}$ une classe d'ensembles stable par intersection finie engendrant $\mathcal{A}$.

Si P_1 et P_2 sont deux mesures de probabilité sur $\mathcal{A}$ telles que P_1 et P_2 sont égales sur $\mathcal{C}$, alors P_1 et P_2 sont égales sur $\mathcal{A}$

dém. Soit $\mathcal{M} = \{ A \in \mathcal{A} \text{ tq. } P_1(A) = P_2(A) \}$

$\mathcal{M}$ est une classe monotone qui contient $\mathcal{C}$ donc le théorème précédent garantit que $\mathcal{M} = \mathcal{A}$. $\blacksquare$

Déf

soit Ω muni d'une tribu $\mathcal{A}$,
et E muni d'une tribu $\mathcal{E}$, $\quad (E = \mathbb{N}, \mathbb{R}, \mathbb{R}^k$ etc.)

Soit $X : \Omega \to E$ une application (variable aléatoire)

X est dite **mesurable** si $\forall B \in \mathcal{E}$, $X^{-1}(B) \in \mathcal{A}$.

ex. $E = \mathbb{R} \rightsquigarrow$ variable aléatoire réelle (v.a.r)

$E = \mathbb{R}^k, k \geq 2 \rightsquigarrow$ vecteur aléatoire

E ensemble discret, fini ou dénombrable $\rightsquigarrow$ variable aléatoire discrète

$E \subset \mathbb{N}, \mathbb{Z} \rightsquigarrow$ variable aléatoire entière

La tribu engendrée par X est la plus petite tribu $\sigma(X) \subset \mathcal{A}$ telle que $X^{-1}(B) \in \sigma(X)$ pour tout $B \in \mathcal{E}$.

Rappel. Étant donnée une probabilité $P : \mathcal{A} \to [0,1]$ sur $(\Omega, \mathcal{A})$ et $X : (\Omega, \mathcal{A}) \longrightarrow (E, \mathcal{E})$ une variable aléatoire,

ens tribu

la loi de X est la mesure image P_X de P par X sur $(E, \mathcal{E})$.

$$P_X(B) := P\left(X^{-1}(B)\right) = P\left(X \in B\right), \quad \forall B \in \mathcal{E}$$

$$\left(= P\left(\{\omega \in \Omega \text{ tq } X(\omega) \in B\}\right)\right)$$

Ex. $(\Omega, \mathcal{A}, P) = \left(\left\{(i,j) \in \{1,..,6\}^2\right\}, \mathcal{P}(\Omega), P\right)$

avec P = équiprobabilité sur Ω $\left(P(A) = \frac{\# A}{\# \Omega} = \frac{\# A}{36}, \forall A \in \mathcal{P}(\Omega)\right)$

$$X : (\Omega, \mathcal{P}(\Omega)) \longrightarrow (\mathbb{N}, \mathcal{P}(\mathbb{N}))$$
$$(i,j) \longmapsto i+j$$

X est mesurable. $\forall I \in \mathcal{P}(\mathbb{N})$ partie de $\mathbb{N}$,
$X^{-1}(I)$ est une partie de $\Omega = \{1,..,6\}^2$

Loi de X.

$\leadsto$ si $n \in [\![2,12]\!]$, $P_X(\{n\}) = P(\{(i,j) \in \Omega \text{ tq. } i+j=n\})$

$$= \frac{\#\{(i,j) \in \Omega \text{ tq. } i+j=n\}}{36}$$

$\leadsto$ si $n \notin [\![2,12]\!]$, $P_X(\{n\}) = 0$

Autre exemple. $(\Omega, \mathcal{A}, P) = \left([0,1], \mathcal{B}([0,1]), \lambda = \text{mesure de Lebesgue}\right)$

$$X : \Omega \longrightarrow \{0,1\}$$

$$\omega \longmapsto \begin{cases} 1 & \text{si } \omega \in \left[0,\frac{1}{2}\right] \\ 0 & \text{sinon}, \end{cases}$$

i.e. $X = \mathbb{1}_{\left[0,\frac{1}{2}\right]}$

$X^{-1}(\{0\}) = \left]\frac{1}{2},1\right] \in \mathcal{B}(\mathbb{R})$, $X^{-1}(\{1\}) = \left[0,\frac{1}{2}\right] \in \mathcal{B}(\mathbb{R})$

$$P_X(0) = P(X=0) = \lambda\left(\left]\frac{1}{2},1\right]\right) = 1-\frac{1}{2} = \frac{1}{2},$$

$$P_X(1) = P(X=1) = \lambda\left(\left[0,\frac{1}{2}\right]\right) = \frac{1}{2}-0 = \frac{1}{2}$$

Déf. (fonction de répartition)

$(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ou $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, $P : \mathcal{A} \to [0,1]$ proba.

La fonction de répartition de P est la fonction $F : \mathbb{R} \to [0,1]$,

où $F(x) = P(]-\infty, x])$ pour $x \in \mathbb{R}$

$X : (\Omega, \mathcal{A}) \to (E, \mathcal{E})$ variable aléatoire, où $E \subset \mathbb{R}$,
$P : \mathcal{A} \to [0,1]$ proba

La fonction de répartition F_X de X est la fonction de répartition de P_X, i.e.

$$F_X(x) = P_X(]-\infty, x]) \text{ pour } x \in \mathbb{R}$$

$$\left(= P(X \leq x) = P(\{\omega \in \Omega \text{ tq } X(\omega) \leq x\})\right)$$

Prop. F_X fonction de répartition de $X : (\Omega, \mathcal{A}) \to (E, \mathcal{E})$ Alors

1) F_X est à valeurs dans $[0,1]$
2) F_X est $\nearrow$ (croissante)
3) F_X est continue à droite
4) $F_X(x) \to 0$ quand $x \to -\infty$, $F_X(x) \to 1$ quand $x \to +\infty$
5) F_X a des limites à gauche en tout point, et
$$\lim_{y \to x^-} F_X(y) = F_X(x^-) = P(X < x)$$
6) $F_X(b) - F_X(a) = P(a < X < b)$
7) $P(X = a) = F_X(a) - \lim_{x \to a^-} F_X(x) = F_X(a) - F_X(a^-)$.

dém. 1) $\forall x \in \mathbb{R}, \quad F_X(x) = P_X(]-\infty, x]) \in [0,1]$

2) $x \leq y \Rightarrow]-\infty, x] \subset]-\infty, y]$ d'où $F_X(x) = P_X(]-\infty, x])$

$$\leq P_X(]-\infty, y]) = F_X(y)$$

3) $x \in \mathbb{R}$. M.q $\displaystyle\lim_{y \to x^+} F_X(y) = F_X(x)$

Comme F_X est $\nearrow$, il suffit de montrer que $\displaystyle\lim_{n \to +\infty} F_X\left(x + \frac{1}{n}\right) = F_X(x)$.

(En effet, on a alors. $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ tq $n \geq N$

$$\Rightarrow F_X(x) \leq F_X\left(x + \frac{1}{n}\right) \leq F_X(x) + \varepsilon$$

et par croissance de F_X, $\forall\, x \le y \le x+\frac{1}{n}$,

$$F_X(x) \le F_X(y) \le F_X\left(x+\frac{1}{n}\right) \le F_X(x)+\varepsilon$$

d'où $\lim\limits_{y \to x^+} F_X(y) = F_X(x)$

On a $F_X\left(x+\frac{1}{n}\right) = P_X\Big(\underbrace{\left]-\infty, x+\frac{1}{n}\right]}_{I_n}\Big)$

(I_n) suite $\searrow$ d'intervalles donc

$$\lim_{n \to +\infty} P_X(I_n) = P_X\left(\bigcap_n I_n\right) = P_X(]-\infty, x]) = F_X(x)$$

5) $J_n :=]-\infty, x - \frac{1}{n}]$

(J_n) suite $\nearrow$ d'intervalles donc $\lim\limits_{n \to +\infty} F_X\left(x - \frac{1}{n}\right) = \lim\limits_{n \to +\infty} P_X(J_n)$

$$= P_X\left(\bigcup_n J_n\right)$$

$$= P_X\left(]-\infty, x[\right).$$

Comme F_X est croissante on en déduit que F_X admet $\underset{\leq F_X(x)}{P_X\left(]-\infty, x[\right)}$

comme limite à gauche en x

($y < x \longrightarrow F_X(y)$ est croissante et majorée par $F_X(x)$)

4) $\emptyset = \bigcap\limits_{n \in \mathbb{N}}]-\infty, -n]$ donc en raisonnant comme ci-dessus,

$$\lim_{n \to +\infty} F_X(-n) = P_X(\emptyset) = 0$$

Par croissance de F_X, on a $\lim\limits_{x \to -\infty} F_X(x) = 0$

On raisonne de même en $+\infty$

6) $F_X(b) - F_X(a) = P_X(]-\infty, b]) - P_X(]-\infty, a])$

$= P_X(]-\infty, a] \cup]a, b]) - P_X(]-\infty, a])$

$= P_X(]-\infty, a]) + P_X(]a, b]) - P_X(]-\infty, a])$

$= P_X(]a, b]) = P(X \in]a, b])$

7) exo.

$\blacksquare$

Prop. une fonction de répartition présente au plus un nombre dénombrable de discontinuités

Théorème. une fonction caractérise complètement la probabilité, i.e,

Si P_1 et P_2 sont 2 proba sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ de fonctions de répartition F_1, F_2

Alors $P_1 = P_2 \iff F_1 = F_2$

dém. $\boxed{\Rightarrow}$ clair

$\boxed{\Leftarrow}$ $\forall \,]a, b] \subset \mathbb{R}$ intervalle,

$$P_1(]a,b]) = F_1(b) - F_1(a) = F_2(b) - F_2(a) = P_2(]a,b])$$

La classe des intervalles $]a, b]$, $a, b \in \mathbb{R}$ est stable par intersection finie et engendre $\mathcal{B}(\mathbb{R})$.

Par le corollaire vu à la p 10, on déduit que $P_1 = P_2$ ∎

Corollaire Soit X, Y deux variables aléatoires définies sur $(\Omega, \mathcal{A}, P)$, resp. $(\Omega', \mathcal{A}', P')$, avec X, Y à valeurs réelles ou entières

Si $F_X = F_Y$ alors X et Y ont même loi

dém. on applique le résultat précédent pour $P_1 = P_X$, $P_2 = P_Y$ ∎

3 Indépendance et Conditionnement

Déf. $(\Omega, \mathcal{A}, P)$ espace probabilisé

a) 2 événements $A, B \in \mathcal{A}$ sont dits **indépendants** si

$$P(A \cap B) = P(A)\, P(B)$$

b) Les événements $(A_i)_{i \in I}$, $A_i \in \mathcal{A}$ pour tout $i \in I$,

sont dits **indépendants** si $\forall J \subset I$ finie,

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j).$$

Prop. Si A et B sont indépendants, (A, B^c), (A^c, B), (A^c, B^c) sont indépendants

Exemples. a) on tire 4 fois un dé

Si A_i, $i=1,..,4$, est un événement qui ne dépend que du i-ème tirage, alors $(A_i)_{i=1,..,4}$ sont indépendants.

b) $\Omega = [\![1,..,4]\!]$, $\mathcal{A} = \mathcal{P}(\Omega)$, P proba uniforme

Alors $A = \{1,2\}$, $B = \{1,3\}$, $C = \{2,3\}$ sont 2 à 2 indépendants $\left(P(A \cap B) = P(\{1\}) = \frac{1}{4} = P(A)\,P(B) \text{ etc} \right)$

mais A, B, C ne sont pas (mutuellement) indépendants

$$\left(P(\underbrace{A \cap B \cap C}_{\emptyset}) = 0 \text{ mais } P(A)\,P(B)\,P(C) = \left(\frac{1}{2}\right)^3 = \frac{1}{8} \right).$$

Déf.. A, B 2 événements $P(A) > 0$. La **probabilité conditionnelle** de B sachant A est

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

Prop. Supposons $P(A) > 0$

a) A, B indépendants $\iff P(B|A) = P(B)$

b) $P(\cdot|B) : \mathcal{A} \to [0,1]$, $B \mapsto P(B|A)$ est une proba sur $\mathcal{A}$, appelée proba conditionnelle sachant A

dém. a) $P(B|A) = \dfrac{P(A \cap B)}{P(A)}$ donc $P(B|A) = P(B)$

$\iff P(A \cap B) = P(A)\,P(B)$

b) • Si $B \in \mathcal{A}$, $A \cap B \subset A$ et $A \cap B \in \mathcal{A}$ donc $0 \leq P(A \cap B) \leq P(A)$

et $P(B|A) \in [0,1]$

(P1) • $P(\Omega | A) = \dfrac{P(\Omega \cap A)}{P(A)} = \dfrac{P(A)}{P(A)} = 1$

(P3) • soit $(B_n)_n \in \mathcal{A}^{\mathbb{N}}$ tq $\forall m \neq n \in \mathbb{N}$, $B_m \cap B_n = \emptyset$

Alors $\forall m \neq n \in \mathbb{N}$, $(A \cap B_m) \cap (A \cap B_n) = \emptyset \Rightarrow P\left(\left(\bigcup_{n \in \mathbb{N}} B_n\right) \cap A\right) = P\left(\bigcup_{n \in \mathbb{N}} (B_n \cap A)\right)$

$$= \sum_{n \in \mathbb{N}} P(B_n \cap A)$$

donc $P\left(\bigcup_{n \in \mathbb{N}} B_n \,\Big|\, A\right) = \sum_{n \in \mathbb{N}} \dfrac{P(B_n \cap A)}{P(A)} = \sum_{n \in \mathbb{N}} P(B_n | A)$ ■

Prop.. (proba. composées)

Si $(A_i)_{i=1,\dots,n}$ sont des évènements t.q. $P(A_1 \cap \dots \cap A_{n-1}) > 0$,

alors $P(A_1 \cap \dots \cap A_n) = P(A_1) P(A_2 | A_1) \dots P(A_n | A_1 \cap \dots \cap A_{n-1})$.

dém. récurrence sur $n \geq 2$

- $n = 2$. $P(A_1 \cap A_2) = P(A_1)\, P(A_2 \mid A_1)$ par définition.
- Si vrai pour $n \geq 2$ montrons le pour $n+1$. Soit $(A_i)_{i=1,\dots,n+1}$ tq $P(A_1 \cap \dots \cap A_n) > 0$.

$$P(A_1 \cap \dots \cap A_{n+1}) = \underbrace{P(A_1 \cap \dots \cap A_n)}_{P(A_1)\, P(A_2 \mid A_1) \dots P(A_n \mid A_1 \cap \dots \cap A_{n-1}) \text{ par HR.}} P(A_{n+1} \mid A_1 \cap \dots \cap A_n)$$

■

Prop. Soit $(B_i)_{i \in I}$ partition finie ou dénombrable de Ω, tq $B_i \in \mathcal{A}$ et $P(B_i) > 0$, $\forall i \in I$. Alors $\forall A \in \mathcal{A}$,

$$P(A) = \sum_{i \in I} P(A \mid B_i)\, P(B_i)$$

dém : Dans ce cas, les $(A \cap B_i)_{i \in I}$ forment une partition de A

$$\left(\bigcup_{i \in I} (A \cap B_i) = \Big(\underbrace{\bigcup_{i \in I} B_i}_{\Omega}\Big) \cap A = A \text{ et } (A \cap B_i) \cap (A \cap B_j) \subset \underbrace{(B_i \cap B_j)}_{= \emptyset} \text{ pour } i \neq j \in I \right)$$

donc $P(A) = P\left(\bigcup_{i \in I} (A \cap B_i)\right) = \sum_{i \in I} P(\underbrace{A \cap B_i}_{})$

$$= \sum_{i \in I} P(A \mid B_i)\, P(B_i)$$

$\left(\text{car } P(B_i) > 0,\ \forall i \in I\right)$ ∎

Prop. . (théorème de Bayes)

Soit $(B_i)_{i \in I}$ partition finie ou dénombrable de Ω tq $P(B_i) > 0$, $\forall i \in I$

Alors $\forall A \in \mathcal{A}$ t.q. $P(A) > 0$, on a : $\forall i \in I$,

$$P(B_i | A) = \frac{P(A|B_i)\, P(B_i)}{\sum_{j \in I} P(A|B_j)\, P(B_j)}$$

dém : Comme $P(A) > 0$, $P(B_i) > 0$, $\forall i \in I$ on a $P(B_i|A) = \frac{P(A \cap B_i)}{P(A)} = \frac{P(A|B_i)\, P(B_i)}{P(A)}$.

Mais d'après la proposition précédente, on peut réécrire le dénominateur.

$$P(A) = \sum_{j \in I} P(A|B_j)\, P(B_j). \qquad \left(\text{car } P(B_j) > 0, \forall j \in I\right)$$

$\blacksquare$