Espace probabilisé

Le calcul des probabilités est une branche de l'analyse dont le but est l'étude des phénomènes aléatoires c'est à dire des expériences dont le résultat ne peut être prédit avec certitude : par exemple, si on répète l'expérience plusieurs fois, on peut obtenir des résultats différents. Une forme d'indétermination apparaît dans l'issue de l'expérience. On peut interpréter cette forme de hasard comme notre incompétence à concevoir, expliciter et utiliser les phénomènes physiques considérés ou encore comme un manque d'information sur les conditions de l'expérience. Certains phénomènes sont peut-être par essence même sujets au hasard ...

Les exemples de telles expériences sont nombreux : on peut penser au jeu de pile ou face, au lancer d'un dé, au sexe d'un enfant à naître, au temps d'attente d'un client à la poste, à la durée de vie d'une particule radioactive ...

Pour étudier ces phénomènes, on doit tout d'abord en faire un modèle mathématique. Le modèle retenu pour décrire les expériences aléatoires est celui d'un triplet communément noté $(\Omega, \mathcal{F}, \mathbb{P})$ et appelé *espace probabilisé* ou *espace de probabilité*.

Un exemple. Il est 18 heures et vous devez impérativement vous rendre à l'autre bout de la ville pour faire une course importante. Votre petit(e) ami(e) doit passer chez vous vers 20 heures ce soir et d'autre part votre chère tante Pétunia doit vous rendre visite vers 18 heures 45.

À cause des embouteillages (on peut imaginer d'autres raisons), vous ne pouvez pas déterminez avec précision à quelle heure vous serez de retour de votre course en ville. Vous êtes alors intéressés par les trois événements suivants :

- A : retour avant 20 h.
- B : retour après 18 h. 45
- C : retour avant 20 h. 30

Les événements A et B parlent d'eux-mêmes et pour le C vous comptez sur la patience de votre ami(e).

En dehors de ces trois événements qui vous intéressent au plus haut point, on considère aussi la famille d'événements plus élémentaires formée par : « je suis de retour à la date t » où t varie continûment de 18 h. à 21 h. On appelle ces événements élémentaires des épreuves ou des réalisations. Considérons l'épreuve ω à savoir vous êtes de retour à 19 heures. Dans cette épreuve, les événements A et C sont réalisés tandis que l'événement B lui ne l'est pas.

Pour étudier ce phénomène, on est donc amené à considérer un espace fondamental Ω, l'espace de toutes les possibilités, et à l'identifier avec l'intervalle $[18, 21]$. L'épreuve ω considérée précédemment est alors identifiée à l'élément 19. L'événement A est représenté par $[18, 20]$, B par $[18.75, 21]$, C par $[18, 20.5]$. Dans cet exemple, on voit que l'appartenance « $\omega \in A$ » équivaut à « A est réalisé dans l'épreuve ω ». Cette équivalence permet de passer du vocabulaire probabiliste « A est réalisé dans l'épreuve ω » au vocabulaire ensembliste « $\omega \in A$ ».

Le triplet fondamental. Dans l'exemple précédent, on a pu expliciter complètement l'espace Ω mais ce n'est pas le cas en général ; on sait seulement qu'il fait partie d'un triplet $(\Omega, \mathcal{F}, \mathbb{P})$ possédant les propriétés suivantes :

– Ω est un ensemble non vide décrivant l'ensemble des réalisations possibles de l'expérience. Un point $\omega \in \Omega$ est une réalisation (ou épreuve) particulière. Par exemple, si on modélise le jeu de pile ou face on peut prendre comme espace Ω l'ensemble $\{0, 1\}$ ou $\{P, F\}$ ou $\{pile, face, tranche\}$, pour le lancer d'un dé on peut prendre $\Omega = \{1, 2, 3, 4, 5, 6\}$.

– $\mathcal{F}$ est une sous-classe de parties de Ω qui représente la classe des événements associés à l'expérience. Par exemple, si on observe la durée de vie d'une ampoule on peut s'intéresser aux événements suivants :

- S : la durée de vie est positive ; c'est l'événement certain.
- I : la durée de vie est strictement négative ; événement impossible.
- B : la durée de vie est supérieure à 100 heures.
- C : la durée de vie est inférieure à 200 heures.

Quelles opérations logiques peut-on effectuer sur les événements ?

- si A est un événement, alors sa négation est aussi un événement ;
- si A et B sont des événements alors « A ou B », « A et B » sont des événements ;
- si on a A_n une suite d'événements, « au moins un des A_n se réalise » est encore un événement.

Ces opérations logiques se traduisent au niveau ensembliste par les opérations classiques : passage au complémentaire pour la négation, union pour « ou », intersection pour « et ».

Dans les cas simples, on peut prendre pour $\mathcal{F}$ l'ensemble de toutes les parties de Ω mais dans les cas un peu plus complexes, il faudra se limiter à une sous-classe.

– Il reste à mesurer les chances qu'un événement se réalise (les chances de devenir millionnaire si on joue au Loto par exemple). Cela consiste à donner une pondération à chacun des événements. Pour obtenir cela, on peut penser à utiliser une approche empirique. Fixons un événement $A \in \mathcal{F}$. On reproduit N fois l'expérience et on compte le nombre de fois N_A où A se réalise. On évalue alors les chances que A se réalise par sa fréquence $\mathrm{Fr}(A) = N_A/N$. On peut remarquer que $0 \leq \mathrm{Fr}(A) \leq 1$ pour tout événement et que si A et B sont deux événements incompatibles i.e. $A \cap B = \emptyset$, on a $\mathrm{Fr}(A \cup B) = \mathrm{Fr}(A) + \mathrm{Fr}(B)$. Cette façon de procéder n'est pas totalement satisfaisante car le résulat dépend du nombre N d'expériences réalisées. De plus il n'est pas toujours très facile de répéter une expérience et cela peut s'avérer très couteux. C'est pourquoi nous introduirons les mesures de probabilité qui sont des applications de $\mathcal{F}$ dans $[0, 1]$ qui associe à chaque événement le nombre $\mathbb{P}(A)$ qui est la probabilité que A se réalise.

Nous voici donc un peu plus familier avec l'expression qui reviendra souvent dans ce cours : « soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé ». La connaissance explicite de ce triplet n'est importante que si l'on s'intéresse à la phase de modélisation ; dans ce cas, Ω est en général l'espace des états de l'expérience et $(E, \mathcal{E})$ remplace souvent $(\Omega, \mathcal{F})$. En dehors de cette phase, on ne spécifie pas l'espace $(\Omega, \mathcal{F}, \mathbb{P})$ qui est un espace abstrait sur lequel on n'a aucune information.

Contexte habituel. En fait, la situation la plus fréquente que nous rencontrerons est la suivante : on fait une expérience aléatoire qui est traduite par l'espace $(\Omega, \mathcal{F}, \mathbb{P})$ mais on s'intéresse à certaines conséquences de l'expérience. Par exemple, on lance deux dés et on regarde si la somme est supérieure à huit. On peut prendre pour Ω l'ensemble de tous les résultats possibles i.e. l'ensemble $\{(1,1),(1,2),\dots(6,5),(6,6)\}$; si les dés ne sont pas pipés chacun de ces couples a les mêmes chances d'être obtenu et donc on définit $\mathbb{P}$ par $\mathbb{P}\left((i,j)\right) = 1/36$. Comme on s'intéresse à la somme des points des dés, on introduit, pour tout couple $\omega = (i,j)$, $X(\omega) = i + j$. On a donc une application $X : \Omega \longrightarrow \{2,\dots,12\}$. Dans ce problème, la quantité que l'on cherche à déterminer est $\mathbb{P}\left(\{\omega \in \Omega,\ X(\omega) \in \{8,9,10,11,12\}\}\right)$ que l'on note $\mathbb{P}_X\big(\{8,9,10,11,12\}\big)$.

De façon générique, nous aurons une application définie sur Ω d'espace d'états E – nous dirons variable aléatoire – $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (E, \mathcal{E})$ et l'on cherchera à déterminer la mesure de probabilité $\mathbb{P}_X$ – la loi de X sous $\mathbb{P}$– définie par

$$\forall A \in \mathcal{E}, \qquad \mathbb{P}_X(A) = \mathbb{P}\big(\{\omega \in \Omega,\ X(\omega) \in A\}\big).$$

La fin de ce premier cours consite à donner les définitions mathématiques des objets que nous avons introduits.

1. Notion de tribu.

Définition. Soit Ω un ensemble non vide et $\mathcal{F}$ une classe de parties de Ω. On dit que $\mathcal{F}$ est une *tribu* sur Ω si

(i) $\Omega \in \mathcal{F}$;

(ii) $\mathcal{F}$ est stable par passage au complémentaire : si A appartient à $\mathcal{F}$ alors A^c appartient à $\mathcal{F}$;

(iii) $\mathcal{F}$ est stable par union dénombrable : si $(A_n)_{\mathbf{N}} \subset \mathcal{F}$ alors $\cup_{\mathbf{N}} A_n$ appartient à $\mathcal{F}$.

Si $\mathcal{F}$ est une tribu sur Ω, $(\Omega, \mathcal{F})$ est appelé espace mesurable et les éléments de $\mathcal{F}$ des événements.

Remarque(s). Si $\mathcal{F}$ est une tribu sur Ω, alors $\emptyset \in \mathcal{F}$. De plus, $\mathcal{F}$ est stable par union finie, par intersection finie et dénombrable, par différence ... Expliquons par exemple pourquoi $\mathcal{F}$ est stable par intersection dénombrable. Soit donc $(A_n)_{\mathbf{N}} \subset \mathcal{F}$. On a $\cap_n A_n = (\cup_n A_n^c)^c$. Comme, pour tout n, $A_n \in \mathcal{F}$, le point (ii) de la définition implique que A_n^c appartient à $\mathcal{F}$ pour tout n. Le point (iii) entraine donc que $\cup_n A_n^c$ est élément de $\mathcal{F}$; on peut alors appliquer à nouveau le point (ii) pour obtenir que $\cap_n A_n$ est un élément de $\mathcal{F}$.

Exemple. Soit Ω un ensemble non vide. $\mathcal{F}_1 = \{\emptyset, \Omega\}$ est une tribu appelée tribu grossière ; $\mathcal{F}_2 = \mathcal{P}(\Omega)$, la collection des parties de Ω, est une tribu ; enfin, si $A \subset \Omega$, $\mathcal{F}_3 = \{\emptyset, \Omega, A, A^c\}$ est une tribu sur Ω.

Exemple. On considère $\mathcal{A}$ l'ensemble des parties $A \subset \mathbf{N}$ vérifiant A est fini ou A^c est fini. $\mathcal{A}$ n'est pas une tribu sur $\mathbf{N}$. En effet, les deux premiers points de la définition sont satisfaits mais le troisième ne l'est pas. Pour s'en convaincre, considérons, pour $n \in \mathbf{N}$, $A_n = \{2n\}$ qui appartient à $\mathcal{A}$ puisque son cardinal est 1. Par contre, $\cup_n A_n$ n'appartient pas à $\mathcal{A}$: ni $\cup_n A_n$ ni son complémentaire ne sont de cardinal fini puisqu'il s'agit respectivement des entiers pairs et impairs.

Comme déjà dit en introduction, on aimerait travailler avec la tribu $\mathcal{P}(\Omega)$. Ceci est possible par exemple si Ω est un ensemble fini ou dénombrable. Si $\Omega = \mathbf{R}$, il n'est pas possible de mesurer toutes les parties de $\mathbf{R}$ sans aboutir à une contradiction et on se limite à une classe de parties appelée tribu borélienne.

Définition. On appelle tribu borélienne sur $\mathbf{R}$, notée $\mathcal{B}(\mathbf{R})$, la plus petite tribu, au sens de l'inclusion, contenant tous les intervalles de $\mathbf{R}$.

Cette définition appelle un commentaire : il n'est pas évident à priori que l'on puisse parler de « la plus petite tribu » contenant les intervalles. Toutefois, la définition a bien un sens car l'intersection d'une famille quelconque de tribus sur Ω est encore une tribu sur Ω. On peut donc considérer l'intersection de toutes les tribus sur $\mathbf{R}$ contenant les intervalles qui devient par construction « la plus petite ».

2. Mesure de probabilité.

Définition. Soient $(\Omega, \mathcal{F})$ un espace mesurable et $\mathbb{P}$ une application de $\mathcal{F}$ dans $[0, 1]$. $\mathbb{P}$ est une (mesure de) probabilité sur $(\Omega, \mathcal{F})$ si

(i) $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$;

(ii) pour toute suite $(A_n)_\mathbf{N}$ d'événements deux à deux disjoints, c'est à dire vérifiant $A_n \cap A_m = \emptyset$ si $n \neq m$, $\mathbb{P}(\cup_n A_n) = \sum_n \mathbb{P}(A_n)$.

Le triplet $(\Omega, \mathcal{F}, \mathbb{P})$ où $(\Omega, \mathcal{F})$ est un espace mesurable et $\mathbb{P}$ une probabilité sur cet espace s'appelle un espace probabilisé ou espace de probabilité.

Remarque(s). On peut voir facilement que la condition $\mathbb{P}(\emptyset) = 0$ est une conséquence des deux autres points. D'autre part, la deuxième propriété s'appelle la σ–additivité.

La définition conduit aux propriétés suivantes :

Proposition 1. *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Si A et B appartiennent à $\mathcal{F}$, alors*

- *additivité : si $A \cap B = \emptyset$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$;*
- *croissance : si $A \subset B$, $\mathbb{P}(A) \leq \mathbb{P}(B)$;*
- *si $A \subset B$, $\mathbb{P}(B \backslash A) = \mathbb{P}(B) - \mathbb{P}(A)$; en particulier, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$;*
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Pour obtenir l'additivité, il suffit de prendre dans la définition $A_0 = A$, $A_1 = B$ et $A_i = \emptyset$ si $i \geq 2$ et d'utiliser la σ–additivité. Pour les deux points suivants, si $A \subset B$, on écrit $B = A \cup (B \backslash A)$ et on utilise l'additivité : $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \backslash A)$. Ensuite, pour la dernière relation, on a – faites un dessin – $A \cup B = A \cup (B \backslash A)$ et $B = (B \backslash A) \cup (B \cap A)$. Par additivité,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \backslash A), \qquad \mathbb{P}(B) = \mathbb{P}(B \backslash A) + \mathbb{P}(B \cap A) \ ;$$

Finalement, on obtient

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Exemple. Soient Ω un ensemble non vide et ω un point de Ω. Pour $A \subset \Omega$, on pose $\delta_\omega(A) = 1$ si $\omega \in A$ et 0 sinon. δ_ω est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$ appelé masse de Dirac au point ω.

En effet, $\delta_\omega(\Omega) = 1$ et $\delta_\omega(\emptyset) = 0$. Seule la σ–additivité n'est pas immédiate. Soit $(A_n)_{n \in \mathbf{N}}$ une suite de parties disjointes 2 à 2. Si $\omega \notin \cup_n A_n$, alors ω n'appartient à aucun des A_n et donc $\delta_\omega(\cup A_n) = 0$, $\delta_\omega(A_n) = 0$ pour tout n. Si ω appartient à $\cup A_n$, il existe un entier p tel que $\omega \in A_p$. Pour tout $n \neq p$, $\omega \notin A_n$ puisque $A_p \cap A_n = \emptyset$. Donc, dans ce cas,

$$\delta_\omega(\cup A_n) = 1, \qquad \delta_\omega(A_p) = 1, \quad \delta_\omega(A_n) = 0, \quad \text{si } n \neq p.$$

Dans les deux cas, on a $\delta_\omega(\cup_n A_n) = \sum_{n \in \mathbf{N}} \delta_\omega(A_n)$, ce qui montre la σ–additivité.

Proposition 2. *Soient* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espace probabilisé et* $(A_n)_{n\in\mathbf{N}} \subset \mathcal{F}$. *Alors,*

- $\mathbb{P}\left(\bigcup_{n\geq 0} A_n\right) \leq \sum_{n\geq 0} \mathbb{P}(A_n)$;
- *si la suite* $(A_n)_{n\in\mathbf{N}}$ *est croissante* ($A_n \subset A_{n+1}$ *pour tout* n*),*

$$\mathbb{P}\Big(\bigcup_{n\geq 0} A_n\Big) = \lim_{n\to\infty} \mathbb{P}(A_n) = \sup_{n\in\mathbf{N}} \mathbb{P}(A_n)\,;$$

- *si la suite* $(A_n)_{n\in\mathbf{N}}$ *est décroissante* ($A_{n+1} \subset A_n$ *pour tout* n*),*

$$\mathbb{P}\Big(\bigcap_{n\geq 0} A_n\Big) = \lim_{n\to\infty} \mathbb{P}(A_n) = \inf_{n\in\mathbf{N}} \mathbb{P}(A_n).$$

La démonstration de cette proposition repose sur la construction suivante : si $(A_n)_{n\in\mathbf{N}}$ est une suite de parties de $\mathcal{F}$ on construit une suite $(B_n)_{n\in\mathbf{N}} \subset \mathcal{F}$ de parties *deux à deux disjointes* telle que, pour tout n,

$$B_n \subset A_n, \qquad \bigcup_{i=0}^{n} B_i = \bigcup_{i=0}^{n} A_i, \qquad \bigcup_{i\geq 0} B_i = \bigcup_{i\geq 0} A_i.$$

Il suffit de poser pour cela $B_0 = A_0$ et $B_n = A_n \backslash \bigcup_{i=0}^{n-1} A_i$, si $n \geq 1$. On a alors,

$$\mathbb{P}\left(\bigcup_n A_n\right) = \mathbb{P}\left(\bigcup_n B_n\right) = \sum_{n\geq 0} \mathbb{P}(B_n) \leq \sum_{n\geq 0} \mathbb{P}(A_n).$$

Quand la suite $(A_n)_{n\in\mathbf{N}}$ est croissante, on a, pour tout entier $n \geq 1$, $B_n = A_n \backslash A_{n-1}$ et par suite $\mathbb{P}(B_n) = \mathbb{P}(A_n) - \mathbb{P}(A_{n-1})$. Il vient alors,

$$\mathbb{P}\left(\bigcup_n A_n\right) = \mathbb{P}\left(\bigcup_n B_n\right) = \sum_{n\geq 0} \mathbb{P}(B_n) = \lim_{p\to\infty} \sum_{n=0}^{p} \mathbb{P}(B_n) = \lim_{p\to\infty} \mathbb{P}(A_p),$$

ce qui établit la seconde assertion. Le troisième point s'obtient par passage au complémentaire.

3. Indépendance, conditionnement.

Commençons par une notion très importante pour le calcul des probabilités, la notion d'indépendance pour des événements.

3.1. Événements indépendants.

On se place sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Rappelons tout d'abord que

Définition. Deux événements A et B sont indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.

On obtient des événements indépendants lorsqu'on reproduit une expérience sans que la première expérience n'interfère avec la seconde. C'est par exemple le cas lorsque l'on joue deux fois à pile ou face. Voici un autre exemple.

Exemple. Si on lance deux fois un dé équilibré, les deux événements A « le premier dé vaut 4 » et B « le deuxième dé vaut 5 » sont indépendants puisque

$$\mathbb{P}(A \cap B) = 1/36 = 1/6 \times 1/6 = \mathbb{P}(A)\,\mathbb{P}(B).$$

Par contre, si C désigne « la somme des dés vaut 5 », A et C ne sont pas indépendants. En effet, $\mathbb{P}(A) = 1/6$, $\mathbb{P}(C) = 1/9$ et $\mathbb{P}(A \cap C) = 1/36$.

Introduisons une notion plus délicate : l'indépendance de n événements.

Définition. Soit $(A_i)_{1\leq i\leq n} \subset \mathcal{F}$ n événements. Ils sont indépendants si

$$\forall J \subset \{1,\ldots,n\}, \qquad \mathbb{P}\left(\bigcap\nolimits_{j\in J} A_j\right) = \prod\nolimits_{j\in J} \mathbb{P}(A_j).$$

Le piège est le suivant : si A, B, C sont trois événements indépendants alors (A,B), (A,C) et (B,C) sont des couples d'événements indépendants mais la réciproque est fausse.

Remarque(s). La définition 3.1 est encore valable pour une suite d'événements ; J doit alors être une partie *finie* de $\mathbf{N}$.

Exemple. On lance deux fois un dé équilibré et on s'intéresse aux événements suivants : A=« le premier résultat est pair », B=« le second résultat est pair » et C=« la somme est paire ». A, B et C sont deux à deux indépendants mais ne sont pas indépendants dans leur ensemble.

En effet, $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 1/2$, $\mathbb{P}(A\cap B) = \mathbb{P}(A\cap C) = \mathbb{P}(B\cap C) = 1/4$, mais $\mathbb{P}(A\cap B\cap C) = \mathbb{P}(A\cap B) = 1/4$.

3.2. Probabilité conditionnelle.

Fixons $B \in \mathcal{F}$ tel que $\mathbb{P}(B) > 0$. Considérons l'application $\mathbb{P}^B$ suivante :

$$\mathbb{P}^B : \mathcal{F} \longrightarrow [0,1], \qquad A \longmapsto \mathbb{P}(A\cap B)/\mathbb{P}(B).$$

On montre que $\mathbb{P}^B$ est une probabilité sur $(\Omega,\mathcal{F})$; en effet, $\mathbb{P}^B(\Omega) = \mathbb{P}(B)/\mathbb{P}(B) = 1$, $\mathbb{P}^B(\emptyset) = 0$ et si $(A_n)_n$ est une suite d'événements deux à deux disjoints, on a

$$\mathbb{P}^B\left(\cup_n A_n\right) = \mathbb{P}(B)^{-1}\,\mathbb{P}\left(B\cap(\cup_n A_n)\right) = \mathbb{P}(B)^{-1}\,\mathbb{P}\left(\cup_n(B\cap A_n)\right) \;;$$

les événements $A_n \cap B$ sont eux-mêmes deux à deux disjoints et par suite

$$\mathbb{P}^B\left(\cup_n A_n\right) = \mathbb{P}(B)^{-1}\sum \mathbb{P}(B\cap A_n) = \sum \mathbb{P}^B(A_n).$$

Définition. Soit $B \in \mathcal{F}$ tel que $\mathbb{P}(B) > 0$. On appelle probabilité conditionnelle de A sachant B, notée $\mathbb{P}(A|B)$, la quantité $\mathbb{P}^B(A) = \mathbb{P}(A\cap B)/\mathbb{P}(B)$. Elle représente la probabilité que A se réalise sachant que B est réalisé.

On peut vérifier facilement que deux événements A et B sont indépendants si et seulement si $\mathbb{P}(A|B) = \mathbb{P}(A)$.

Exemple. Un couple possède deux enfants qui ne sont pas jumeaux. La probabilité que les deux enfants soient des filles vaut $1/4$; si on sait de plus que l'un des deux est une fille, cette probabilité vaut $(1/4)/(3/4) = 1/3$; si on sait à présent que l'aînée est une fille, la probabilité d'avoir deux filles est $(1/4)/(1/2) = 1/2$.

On peut remarquer que $\mathbb{P}(A\cap B) = \mathbb{P}(A|B)\times\mathbb{P}(B) = \mathbb{P}(B|A)\times\mathbb{P}(A)$ – la dernière égalité n'étant vraie que si $\mathbb{P}(A) > 0$. On en déduit la

Formule de Bayes. Soient $(A_i)_{i\in\mathbf{N}} \subset \mathcal{F}$ une partition de Ω telle que $\mathbb{P}(A_i) > 0$ pour tout i et $B \in \mathcal{F}$ tel que $\mathbb{P}(B) > 0$. On a alors

$$\mathbb{P}(B) = \sum_{i\geq 0} \mathbb{P}(B \cap A_i) = \sum_{i\geq 0} \mathbb{P}(B|A_i) \times \mathbb{P}(A_i),$$

et

$$\forall j \in \mathbf{N}, \qquad \mathbb{P}(A_j|B) = \mathbb{P}(B|A_j) \times \mathbb{P}(A_j) \Big/ \sum_{i\geq 0} \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

Exemple. Françoise, Bernard et Philippe vont se répartir les copies de PRBU à proportion de 50%, 25% et 25%. Françoise recevra 75% des étudiants dont elle corrige la copie, Bernard 60% et Philippe 70%. La probabilité qu'un étudiant choisi au hasard réussisse son examen s'obtient comme suit

$$\mathbb{P}(R) = \mathbb{P}(R|F)\,\mathbb{P}(F) + \mathbb{P}(R|B)\,\mathbb{P}(B) + \mathbb{P}(R|Ph)\,\mathbb{P}(Ph) = 0,7.$$

L'étudiant a réussi ; la probabilité que sa copie ait été corrigée par Bernard est

$$\mathbb{P}(B|R) = \mathbb{P}(B \cap R)/\mathbb{P}(R) = \mathbb{P}(R|B) \times \mathbb{P}(B)/\mathbb{P}(R) = 0,6 \times 0,25/0,7 \simeq 0,2143.$$