P2 : Espaces probabilisés

I Espaces probabilisables, espaces probabilisés

On veut définir une probabilité sur un ensemble Ω. Si Ω est fini ou dénombrable, ce n'est pas trop compliqué. Mais lorsque ce n'est pas le cas, les difficultés augmentent singulièrement. La première difficulté est qu'on n'arrive pas en général à définir la probabilité de toutes les parties de Ω (contrairement au cas fini ou dénombrable). Un théorème d'Ulam dit que si Ω est en bijection avec $\mathbf{R}$, si P est une probabilité définie sur $\mathscr{P}(\Omega)$, alors il existe une partie dénombrable D de Ω telle que $P(D) = 1$ et donc $P(\Omega \setminus D) = 0$.
On ne pourra donc attribuer une probabilité qu'à certaines parties de Ω, ce sont ces parties qui seront appelées « évènements ». Pour que P ait les propriétés qu'on attend d'une probabilité, ces événements doivent former une « tribu ».

I.1 Tribu sur un ensemble

a. Définition

Soit Ω un ensemble.

On appelle tribu sur Ω toute partie $\mathscr{A}$ de $\mathscr{P}(\Omega)$ vérifiant

1. $\emptyset \in \mathscr{A}$
2. $A \in \mathscr{A} \implies \overline{A} \in \mathscr{A}$
3. Si $(A_n)_{n\in\mathbf{N}}$ est une suite d'éléments de $\mathscr{A}$, $\displaystyle\bigcup_{n=0}^{+\infty} A_n \in \mathscr{A}$

La propriété 2. est la stabilité par passage au complémentaire, la propriété 3. est la stabilité par réunion dénombrable. Vu 2., on peut remplacer 1. par $\Omega \in \mathscr{A}$.

b. Exemples

$\mathscr{P}(\Omega)$ est une tribu sur Ω (dite parfois discrète).
$\{\emptyset, \Omega\}$ est une tribu sur Ω (dite parfois grossière).
Si A est une partie de Ω, non vide et distincte de Ω, la plus petite tribu qui contient A est :

c. Propriétés

Une tribu est stable par réunion finie ou dénombrable, par intersection finie ou dénombrable :

Si $\mathscr{A}$ est une tribu sur Ω,

1. $\Omega \in \mathscr{A}$
2. Si $(A_n)_{n\in\mathbf{N}}$ est une suite d'éléments de $\mathscr{A}$, $\displaystyle\bigcap_{n=0}^{+\infty} A_n \in \mathscr{A}$
3. $(A \in \mathscr{A}$ et $B \in \mathscr{A}) \Rightarrow (A \cup B \in \mathscr{A})$
4. $(A \in \mathscr{A}$ et $B \in \mathscr{A}) \Rightarrow (A \cap B \in \mathscr{A})$
5. Si $(A_n)_{0\leq n\leq N}$ est une famille finie d'éléments de $\mathscr{A}$, $\displaystyle\bigcup_{n=0}^{N} A_n \in \mathscr{A}$
6. Si $(A_n)_{0\leq n\leq N}$ est une famille finie d'éléments de $\mathscr{A}$, $\displaystyle\bigcap_{n=0}^{N} A_n \in \mathscr{A}$

Remarque : Une partie de $\mathscr{P}(\Omega)$ contenant $\emptyset$, stable par passage au complémentaire et par réunion finie (et donc par intersection finie) est appelée une algèbre de parties de Ω (terminologie h.p.). Une tribu est parfois appelée σ-algèbre (σ-algebra en anglais), le préfixe σ désignant le "passage au dénombrable".

Exercice : Soit Ω un ensemble. On considère l'ensemble $\mathscr{A}$ des parties de Ω qui sont finies ou dénombrables ou dont le complémentaire est fini ou dénombrable. Montrer que $\mathscr{A}$ est une tribu.

I.2 Espace probabilisable

a. Définitions

On appelle espace probabilisable tout couple $(\Omega, \mathscr{A})$ formé d'un ensemble et d'une tribu sur cet ensemble. Les éléments de $\mathscr{A}$ seront appelés les évènements.

b. Vocabulaire

Un évènement élémentaire est un singleton (en supposant que les singletons appartiennent à la tribu).
Si A est un évènement, $\overline{A}$ est l'évènement contraire.
L'événement $A \cap B$ est appelé « A et B ».
L'événement $A \cup B$ est appelé « A ou B ».
Deux évènements A et B sont dits incompatibles lorsque $A \cap B = \emptyset$.
L'évènement impossible est $\emptyset$, l'évènement certain est Ω (vocabulaires peu usités).
Il ne faut pas confondre incompatibilité et indépendance : deux évènements incompatibles sont (très) rarement indépendants. Il ne faudra pas non plus confondre « impossible » et « négligeable », voir plus loin.

I.3 Probabilité sur un espace probabilisable

a. Définition

Soit $(\Omega, \mathscr{A})$ un espace probabilisable. Une probabilité sur $(\Omega, \mathscr{A})$ est une application P définie sur $\mathscr{A}$ telle que

1. $\forall A \in \mathscr{A} \quad P(A) \in [0,1]$
2. $P(\Omega) = 1$
3. Si $(A_n)_{n\in\mathbf{N}}$ est une suite d'événements deux à deux disjoints , $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n)$,
4. Autrement dit, si $(A_n)_{n\in\mathbf{N}}$ est une suite d'événements,

$$\Big(\forall (p,q) \in \mathbf{N}^2\ (p \neq q) \Rightarrow (A_p \cap A_q = \emptyset)\Big) \Longrightarrow \left(P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n)\right)$$

La propriété 3. (ou 4. car ci-dessus 3.=4.) est appelée σ-additivité, elle devrait se lire « la série $\sum_n P(A_n)$ converge, et sa somme vaut. . . » $P\left(\bigcup_{n=0}^{+\infty} A_n\right)$. Mais d'abord :

Proposition Si P est une probabilité sur $(\Omega, \mathscr{A})$, alors $P(\emptyset) = 0$.

Proposition Si P est une probabilité sur $(\Omega, \mathscr{A})$, si A et B sont deux événements, alors

$$(A \cap B) \Longrightarrow \Big(P(A \cup B) = P(A) + P(B)\Big)$$

Démonstration : Il suffit de prendre $A_n = \emptyset$ si $n \geq 2$ dans la σ-additivité.
Donc, par récurrence sur n,

Proposition Si P est une probabilité sur $(\Omega, \mathscr{A})$, si les A_k $(1 \leq k \leq n)$ sont des événements deux-à-deux disjoints, alors

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k)$$

Cette dernière proposition montre que si $(A_n)_{n\in\mathbf{N}}$ est une suite d'événements deux à deux disjoints, pour tout N on a

$$\sum_{n=0}^{N} P(A_n) = P\left(\bigcup_{n=0}^{N} A_n\right) \leq 1$$

ce qui est cohérent avec la convergence de la série $\sum P(A_n)$, car elle est à termes réels positifs (la somme est d'ailleurs inféreure ou égale à 1, ce qui est plutôt rassurant).

b. Définition

On appelle espace probabilisé tout triplet $(\Omega, \mathscr{A}, P)$ formé d'un ensemble Ω, d'une tribu $\mathscr{A}$ sur Ω et d'une probabilité P sur $\mathscr{A}$.

I.4 Cas très simple : probabilité sur un univers fini

Si Ω est fini, on prend généralement $\mathscr{A} = \mathscr{P}(\Omega)$, et la propriété de σ-additivité est équivalente à la propriété

> 3'. Si A et B sont deux événements disjoints,
> alors $P(A \cup B) = P(A) + P(B)$

En effet, si on a 3', par récurrence on obtient que la probabilité d'une réunion finie d'événements disjoints est la somme des probabilités de ces événements ; or si l'univers est fini il n'y a qu'un nombre fini d'événements, une réunion dénombrable d'évènements s'écrit donc comme une réunion finie.
Si $\Omega = \{\omega_1, \ldots, \omega_m\}$, P est entièrement définie par la donnée des $P(\omega_i)$ $(1 \leq i \leq m)$, qui forment une famille de réels positifs dont la somme vaut 1. Et, pour toute partie A de Ω,

$$P(A) = \sum_{\omega \in A} P(\{\omega\})$$

Les probabilités des événements élémentaires déterminent donc P.

I.5 Cas simple : probabilité sur un ensemble dénombrable

Ici, on garde la propriété de σ-additivité, que l'on ne peut plus remplacer par la simple additivité (il est par ailleurs déconseillé de chercher une « pseudo-probabilité » additive et non σ-additive, mais ça existe).
Ici encore, il n'y a pas d'obstacle à prendre pour $\mathscr{A}$ la tribu « discrète », c'est-à-dire $\mathscr{P}(\Omega)$. On obtient :

Proposition Soit Ω un ensemble dénombrable. Pour toute famille $(p_\omega)_{\omega \in \Omega}$ de réels positifs sommable et de somme 1, il existe une unique probabilité sur $\mathscr{P}(\Omega)$ telle que

$$\forall \omega \in \Omega \qquad P(\{\omega\}) = p_\omega$$

Cette probabilité vérifie

$$\forall A \in \mathscr{P}(\Omega) \qquad P(A) = \sum_{\omega \in A} p_\omega$$

Démonstration C'est une conséquence des résultats sur la sommabilité.

Ici, encore une fois, P est définie de manière unique par les probabilités des singletons. Pour un univers fini ou dénombrable, la définition des tribus d'évènements n'a donc pas grand intérêt.

I.6 Un exemple d'univers non dénombrable : le Pile ou Face infini

Ici, on considère un univers non dénombrable :

$$\Omega = \{0, 1\}^{\mathbf{N}_*}$$

Il n'est pas possible de définir une probabilité modélisant l'expérience aléatoire à l'aide des probabilités des singletons : les événements élémentaires ont une probabilité nulle. On procède tout autrement. . .mais on est obligé d'admettre la plupart des résultats, qui ne sont pas tous simples. Le programme est clair : on ne doit pas vous embêter sur ces points que vous verrez en théorie de la mesure si vous continuez les mathématiques après la prépa.

Définition On appelle « évènement cylindrique élémentaire » (dénomination h.p.), toute partie X de Ω telle qu'il existe $n \in \mathbf{N}_*$ et $(\epsilon_1, \ldots, \epsilon_n) \in \{0, 1\}^n$ tels que

$$(\omega \in X) \iff ((\omega_1, \ldots, \omega_n) = (\epsilon_1, \ldots, \epsilon_n))$$

Pourquoi cyclindrique ? écrire l'équation d'un cylindre d'axe vertical dans un repère cartésien aide à comprendre le terme.

Exemples Dire quels évènements, parmi les suivants, sont cylindriques :

- Entre les lancers 1000 et 1999, on obtient exactement 350 fois Pile.
- La première séquence Pile Face précède la première séquence Pile Pile.
- Les 100 premiers lancers donnent Face.

Proposition Soit $p \in]0,1[$. Il existe une tribu $\mathscr{A}$ sur Ω et une probabilité P sur $(\Omega, \mathscr{A}, P)$ telles que

• $\mathscr{A}$ contienne les évènements cylindriques élémentaires.

• pour un tel évènement tel que décrit ci-dessus, i.e.
$X = \{\omega \in \Omega\ ;\ (\omega_1, \ldots, \omega_n) = (\epsilon_1, \ldots, \epsilon_n)\}$,

$$P(X) = p^{\sum \epsilon_i}(1-p)^{n-\sum \epsilon_i}$$

Démonstration Difficile, hors programme. Les principes en sont donnés en annexe, la lecture de cette annexe est parfaitement facultative.

Exemple d'utilisation : chercher la probabilité de l'évènement « la première séquence Pile-Face précède la première séquence Face-Face ».

II Propriétés

Dans toute cette section, $(\Omega, \mathscr{A}, P)$ est un espace probabilisé :

1. $\forall A \in \mathscr{A} \quad P(A) \in [0,1]$
2. $P(\Omega) = 1$
3. Si $(A_n)_{n\in\mathbf{N}}$ est une suite d'événements deux à deux disjoints , $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n)$

II.1 Croissance

Proposition Si A et B sont deux événements, si $A \subset B$, alors

$$P(A) \le P(B)$$

II.2 Probabilité d'une réunion d'événements

a. Réunion disjointe, finie ou dénombrable

Proposition Si $(A_n)_{n\in\mathbf{N}}$ est une suite d'événements deux à deux disjoints ,
$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n)$
Si $(A_n)_{0\le n\le N}$ est une suite d'événements deux à deux disjoints ,
$P\left(\bigcup_{n=0}^{N} A_n\right) = \sum_{n=0}^{N} P(A_n)$

b. Réunion finie

Proposition Si A et B sont deux événements, alors

$$P(A\cup B) = P(A) + P(B) - P(A\cap B)$$

Une démonstration se fait par exemple en considérant l'événement

$$A' = A \cap \overline{A\cap B}$$

Et en remarquant que

$$A\cup B = A'\cup B \qquad , \qquad A = A' \cup (A\cap B)$$

Or les seconds membres sont des unions disjointes...(faire un dessin).

On peut écrire une formule analogue pour trois,...,n événements (formule dite du crible, ou de Poincaré) ; c'est un peu moins simple.

c. Réunion croissante

Proposition Soit $(A_n)_{n\in\mathbf{N}}$ une suite croissante (pour l'inclusion) d'événements :

$$\forall n \in \mathbf{N} \qquad A_n \subset A_{n+1}$$

Alors

$$P(A_k) \xrightarrow[(k\to+\infty)]{} P\left(\bigcup_{n=0}^{+\infty} A_n\right)$$

Corollaire Soit $(A_n)_{n\in\mathbf{N}}$ une suite quelconque d'événements ; alors

$$P\left(\bigcup_{k=0}^{n} A_k\right) \xrightarrow[(n\to+\infty)]{} P\left(\bigcup_{k=0}^{+\infty} A_k\right)$$

Remarque : La réunion croissante est parfois notée

$$\bigcup_{n=0}^{+\infty}\uparrow A_n$$

mais cette notation ne figure pas dans le programme.

Remarque : Cette propriété porte parfois le nom de « continuité croissante ». Hélas cette terminologie ne figure pas dans le programme officiel. On peut néanmoins penser qu'elle sera comprise.

Remarque : Lorsque $(A_n)_{n\in\mathbf{N}}$ est une suite croissante (pour l'inclusion) d'événements, on est parfois tenté de noter $\lim(A_n)$ la réunion $\bigcup\limits_{n=0}^{+\infty} A_n$. Ce n'est pas une vision stupide des choses, mais on ne le fera pas.

d. Réunion quelconque

Proposition Soit $(A_n)_{n\in\mathbf{N}}$ une suite quelconque d'événements : Alors

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} P(A_n)$$

(où, si la série $\sum P(A_n)$ diverge, on lira la formule :

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq +\infty$$

ce qui ne dit rien ; si la série converge et a une somme ≥ 1, le résultat ne dit rien non plus)

II.3 Probabilité d'une intersection décroissante

Proposition Soit $(A_n)_{n\in\mathbf{N}}$ une suite décroissante (pour l'inclusion) d'événements :

$$\forall n \in \mathbf{N} \qquad A_{n+1} \subset A_n$$

Alors

$$P(A_k) \xrightarrow[(k\to+\infty)]{} P\left(\bigcap_{n=0}^{+\infty} A_n\right)$$

Corollaire Soit $(A_n)_{n\in\mathbf{N}}$ une suite quelconque d'événements ; alors

$$P\left(\bigcap_{k=0}^{n} A_k\right) \xrightarrow[(n\to+\infty)]{} P\left(\bigcap_{k=0}^{+\infty} A_k\right)$$

Remarque : L'intersection décroissante est parfois notée

$$\bigcap_{n=0}^{+\infty}\downarrow A_n$$

Mais cette notation ne figure pas dans le programme.

Remarque : Cette propriété porte bien sûr parfois le nom de « continuité décroissante ».

III Négligeabilité

III.1 Evénements négligeables

Définition : On dit qu'un événement A est négligable lorsque $P(A) = 0$. L'événement impossible est négligeable. Un événement négligeable n'est pas en général impossible.

Proposition 1 : Une réunion finie ou dénombrable d'événements négligeables est négligeable.

Proposition 2 : Si A et B sont deux évènements, si $A \subset B$, si B est négligeable, A l'est.

Rien sur les intersections d'évènements négligeables ?
Exemple : Dans le jeu de Pile ou Face infini, l'événement « la pièce donne Pile un nombre fini de fois » est négligeable.

III.2 Evénements presque sûrs, propriété presque sûre

Définitions :

Un événement A est presque sûr, ou presque certain, lorsque $P(A) = 1$, ce qui équivaut à dire que $\overline{A}$ est négligeable.

Une propriété est dite presque sûre lorsque l'ensemble des éléments de Ω qui qui ont cette propriété est un événement presque sûr.

IV Conditionnement, indépendance

IV.1 Probabilité conditionnelle

Définition Soit $(\Omega, \mathscr{A}, P)$ un espace probabilisé, B un événement tel que $P(B) > 0$. Pour tout événement $A \in \mathscr{A}$, on définit la probabilité conditionnelle de A sachant B par

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

(*Se lit en général « P de A sachant B »*)

Proposition P_B est une probabilité sur $(\Omega, \mathscr{A})$.

Démonstration P_B est une application définie sur $\mathscr{A}$ (propriétés des tribus), à valeurs réelles positives. Et même à valeurs dans $[0, 1]$ par croissance de P.
Clairement $P_B(\Omega) = 1$.
La seule propriété qui demande un peu de travail est la σ-additivité. On peut bien sûr la montrer directement, mais il est aussi possible de montrer l'additivité d'abord (plus simple à écrire) :
Si $A \cap A' = \emptyset$, $(A \cap B) \cap \left(A' \cap B\right) = \emptyset$, et donc

$$\begin{aligned} P_B(A \cup A') &= \frac{P\left(\left(A \cup A'\right) \cap B\right)}{P(B)} \\ &= \frac{P\left((A \cap B) \cup \left(A' \cap B\right)\right)}{P(B)} \\ &= \frac{P\left(A \cap B\right) + P\left(A' \cap B\right)}{P(B)} \\ &= P_B(A) + P_B(A') \end{aligned}$$

puis, maintenant qu'on a vu comment les choses se passaient, prendre une suite (A_n) d'éléments de $\mathscr{A}$ deux à deux disjoints et faire à peu près la même chose…

IV.2 Probabilités composées

Proposition (**Formule des probabilités composées**)

Si A et B sont deux événements, si $P(B) \neq 0$, alors

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

Plus généralement, si $A_1, \dots, A_m$ sont des événements, et si $P(A_1 \cap \dots \cap A_{m-1}) \neq 0$, alors

$$P(A_1 \cap A_2 \dots \cap A_m) = P(A_1)P_{A_1}(A_2)\dots P_{A_1 \cap A_2 \cap \dots \cap A_{m-1}}(A_m)$$

ou encore

$$P\left(\bigcap_{k=1}^{m} A_k\right) = \prod_{k=1}^{m} P_{\bigcap_{j=1}^{k-1} A_j}(A_k)$$

(avec la convention habituelle $\displaystyle\bigcap_{j=1}^{0} A_j = \Omega$, Ω étant neutre pour $\cap$. On remarque que $P_\Omega = P$).

Démonstration par récurrence, bien sûr.

Exemple : le problème des clés

Dans un établissement scolaire où il y a beaucoup de serrures différentes, on vous a prêté un trousseau de n clefs pour ouvrir une porte, et vous avez oublié de demander laquelle était la bonne. De plus, elles se ressemblent suffisamment pour qu'il soit impossible de deviner à vue laquelle est la plus plausible. Vous les essayez donc l'une après l'autre. Quelle est la probabilité pour que ce soit la $k$ième clé testée qui vous ouvre la porte ($1 \leq k \leq n$) ?

Exercice Dans une urne se trouvent $n-1$ boules blanches, 1 boule rouge. On tire, sans remise, boule après boule. Quelle est la probabilité pour que la boule rouge sorte de l'urne au $k$ième tirage ?

IV.3 Probabilités totales

Définition (**Système complet d'événements**)

Soit (Ω, P) un espace probabilisé, I un ensemble fini ou dénombrable. On dit que la famille $(A_i)_{i\in I}$ d'événements est un système complet d'évènements lorsque

$$(i \neq j) \implies \left(A_i \cap A_j = \emptyset\right) \qquad \text{et} \qquad \bigcup_{i\in I} A_i = \Omega$$

(si on impose de plus les A_i non vides, ce qui se fait parfois, $\{A_i\}_{i\in I}$ est une partition de Ω).

Définition (**Système quasi-complet d'événements**)

Soit (Ω, P) un espace probabilisé, I un ensemble fini ou dénombrable. On dit que la famille $(A_i)_{i\in I}$ d'événements est un système quasi-complet d'évènements lorsque les A_i sont deux à deux disjoints et de réunion presque sûr : $P\left(\bigcup_{i\in I} A_i\right) = 1$.

Proposition (**Formule des probabilités totales**)

Soit $(A_i)_{i\in I}$ un système complet ou quasi-complet d'événements (I fini ou dénombrable), alors, pour tout évènement B,

$$P(B) = \sum_{i\in I} P(A_i \cap B)$$

Proposition (**Formule des probabilités totales**)

Soit $(A_i)_{i\in I}$ un système complet ou quasi-complet d'événements (I fini ou dénombrable) tels que $\forall i \in I \quad P(A_i) > 0$.

Alors, pour tout événement B,

$$P(B) = \sum_{i\in I} P(A_i \cap B) = \sum_{i\in I} P(A_i)\, P_{A_i}(B) \sum_{i\in I} P(A_i)\, P(B|A_i)$$

Remarque Si parmi les A_i certains sont négligables, si $J = \{i \in I\,;\, P(A_i) \neq 0\}$, alors on peut écrire

$$P(B) = \sum_{i\in I} P(A_i \cap B) = \sum_{i\in J} P(A_i)\, P_{A_i}(B)$$

Cette formule sert tout le temps : faire des probabilités, c'est conditionner.
C'est la première proposition qui est la plus utile.
Il faut se méfier de l'écriture fausse $P(B) = \sum_{i \in I} P(B|A_i)$. Prendre par exemple l'événement, dans un lancer Pile ou Face, pour A_1 l'évènement « la pièce tombe sur Pile », pour A_2 l'évènement « la pièce tombe sur Face », pour B l'évènement « la pièce tombe »

IV.4 Probabilité des causes (formule de Bayes)

Proposition (**Formules de Bayes**)

Soit A et B deux événements tels que $P(A) > 0$ et $P(B) > 0$. Alors

$$P_B(A) = \frac{P(A)\ P_A(B)}{P(B)}$$

Si de plus $P(A^c) > 0$ alors

$$P_B(A) = \frac{P(A)\ P_A(B)}{P(A)\ P_A(B)\ +\ P(A^c)\ P_{A^c}(B)}$$

Et plus généralement, si $(A_i)_{i\in I}$ est un système complet d'événements (I fini ou dénombrable) tels que $\forall i \in I \quad P(A_i) \neq 0$, on a

$$\forall i \in I \qquad P_B(A_i) = \frac{P(A_i)\ P_{A_i}(B)}{\sum_{k\in I} P(A_k) P_{A_k}(B)}$$

Démonstration $P(A\cap B) = P(A)P_A(B) = P(A)P(B|A)$.

$P(A\cap B) = P(B)P_B(A) = P(B)P(A|B)$.

Exemple : interprétation d'un test biologique Dans une population, 1 personne sur 10.000 est atteinte d'une maladie que l'on peut donc qualifier de relativement rare.
On dispose d'un test de détection de cette maladie. Des expérimentations ont permis de mesurer la sensibilité et la spécificité de ce test.
Sensibilité : la probabilité pour un individu malade d'être testé positif est $0,99$.
Spécificité : la probabilité pour un individu non malade d'être testé négatif est $0,999$.
On teste un individu pris au hasard dans la population. Le test est positif. Quelle est la probabilité pour qu'il soit malade ?
Cet exercice, important, fait partie de la formation en première année d'études de médecine.

Exercice On considère trois urnes X, Y et Z. Dans l'urne X, il y a 99 boules noires, une boule blanche ; dans l'urne Y, il y a 999 boules blanches, une boule noire. Dans l'urne Z, il y a 9999 boules blanches, une boule noire. On tire au hasard une première boule dans l'urne Z. Si cette première boule est blanche, on tire une deuxième boule, dans l'urne Y. Sinon, on tire la deuxième boule dans l'urne X. Quelle est la probabilité pour que la première boule soit noire sachant que la deuxième l'est ?

IV.5 Indépendance

a. Définition

Définition (pour deux événements)

On dit que deux événements A et B d'un espace probabilisé (Ω, P) sont indépendants lorsque

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

Remarquons que cela équivaut, dans le cas $P(B) > 0$, à

$$P_B(A) = P(A)$$

et donc aussi, dans le cas $P(A) > 0$, à

$$P_A(B) = P(B)$$

Définition (pour une famille quelconque d'évènements)

Soit I un ensemble quelconque (fini ou infini, dénombrable ou pas...). On dit que la famille d'événements $(A_i)_{i\in I}$ est indépendante (ou que les A_i sont mutuellement indépendants) lorsque, pour toute partie finie non vide J de I,

$$P\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} P(A_i)$$

Pour que les A_i soient mutuellement indépendants, il est donc nécessaire qu'ils le soient deux à deux. Mais cela n'est pas suffisant.
Par exemple, on lance un dé tétraédrique équilibré; les faces sont marquées $1,2,3,4$. On considère les événements suivants :
A : le dé tombe sur la face 1 ou la face 2.
B : le dé tombe sur la face 2 ou la face 3.
C : le dé tombe sur la face 1 ou la face 3.
On vérifiera que ces événements sont deux à deux indépendants, mais ne le sont pas mutuellement.

Exercice Trouver une condition nécessaire et suffisante simple pour que deux évènements disjoints (i.e. incompatibles) soient indépendants.

b. Passage au complémentaire

Proposition Soit I un ensemble quelconque. On suppose que la famille d'événements $(A_i)_{i\in I}$ est indépendante. Soit $(B_i)_{i\in I}$ telle que, pour tout $i \in I$, $B_i = A_i$ ou $B_i = \overline{A_i}$. Alors la famille $(B_i)_{i\in I}$ est aussi une famille d'événements indépendants.

Démonstration On peut supposer I fini, puisque l'indépendance d'une famille d'évènements, c'est l'indépendance de toute « sous-famille » finie de cette famille. Donc, quitte à réindexer, on va supposer $I = \{1, \ldots, n\}$, ce qui ne change rien. On veut montrer que la famille $(B_i)_{1\leq i\leq n}$ est une famille d'évènements mutuellement indépendants. Pour cela, on peut raisonner par récurrence sur le nombre de B_i égaux à $\overline{A_i}$.
Supposons $B_1 = \overline{A_1}$, $B_i = A_i$ si $2 \leq i \leq n$ (ce n'est pas restrictif, l'ordre n'a pas d'importance). Alors par additivité :

$$P(B_1 \cap A_2 \cap \ldots \cap A_n) + P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_2 \cap \ldots \cap A_n)$$

et donc, par indépendance des A_k,

$$P(B_1 \cap A_2 \cap \ldots \cap A_n) = (1 - P(A_1))\, P(A_2) \ldots P(A_n) = P(B_1) \ldots P(B_n)$$

On a donc initialisé. Mais on démontre exactement de même que la propriété est récurrente.

c. Exemple : indicatrice d'Euler

Soit $n \geq 2$, on munit $\Omega = \{1, 2, \ldots, n\}$ de la probabilité uniforme P. On note $\{p_1, \ldots, p_r\}$ l'ensemble des diviseurs premiers de n. On désigne par ϕ l'indicatrice d'Euler, on rappelle que $\phi(n)$ est le nombre d'entiers de $[\![1, n]\!]$ premiers avec n.

1. Soit d un diviseur de n ($d \geq 1$), A_d l'ensemble des multiples de d dans Ω. Calculer $P(A_d)$.
2. (a) Montrer que les évènements A_{p_i} ($1 \leq i \leq r$) sont mutuellement indépendants.

   (b) En déduire que $\dfrac{\phi(n)}{n} = \displaystyle\prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right)$ et calculer $\phi(60)$

V Annexe : le lemme de Borel-Cantelli

V.1 Limites supérieure et inférieure d'une suite réelle

Ce paragraphe n'a rien à voir avec les probabilités, il permet de comprendre d'où viennent les notations.

Définition Soit (u_n) une suite bornée de réels. L'ensemble des valeurs d'adhérence de (u_n) est non vide (théorème de Bolzano-Weierstrass), borné (c'est assez évident) et fermé (exercice classique).

On appelle limite supérieure de (u_n) la plus grande valeur d'adhérence de (u_n). On la note $\limsup(u_n)$ ou $\overline{\lim}(u_n)$. On appelle limite inférieure de (u_n) la plus grande valeur d'adhérence de (u_n). On la note $\liminf(u_n)$ ou $\underline{\lim}(u_n)$.

Caractérisation On note, pour tout $n \in \mathbf{N}$,

$$v_n = \sup\left(\{u_p \ ; \ p \geq n\}\right)$$

Alors la suite (v_n) converge, et sa limite est $\overline{\lim}(u_n)$.

Proposition Une suite bornée de réels converge si et seulement si

V.2 Quelques écritures ensemblistes

Soit $(A_n)_{n\geq 0}$ une suite d'éléments d'une tribu $\mathscr{A}$ sur un univers Ω. Déterminer, en utilisant des $\cap$, des $\cup$ et éventuellement des complémentaires, un évènement B...

(i)...tel que $\omega \in \Omega$ appartienne à B si et seulement si ω appartient à tous les A_n à partir d'un certain rang.

(ii)...tel que $\omega \in \Omega$ appartienne à B si et seulement si ω appartient à une infinité de A_n.

(iii)...tel que $\omega \in \Omega$ appartienne à B si et seulement si ω appartient à au plus un nombre fini de A_n.

(iv)...tel que $\omega \in \Omega$ appartienne à B si et seulement si ω n'appartienne à aucun A_n à partir d'un certain rang.

V.3 Borel-Cantelli « facile »

Soit $(\Omega, \mathscr{A}, P)$ un espace probabilisé. Si $(A_n)_{n\in\mathbf{N}}$ est une suite d'événements, on note

$$\limsup_{n\to+\infty}(A_n) = \bigcap_{p=0}^{+\infty}\left(\bigcup_{n=p}^{+\infty} A_n\right)$$

On suppose $\sum_n P(A_n) < +\infty$. Montrer que

$$P\left(\limsup_{n\to+\infty}(A_n)\right) = 0$$

V.4 Borel-Cantelli « difficile »

On reprend les notations de Borel-Cantelli « facile », mais on change les hypothèses : on suppose $\sum_n P(A_n) = +\infty$ et les A_n indépendants. Montrer que

$$P\left(\limsup_{n\to+\infty}(A_n)\right) = 1$$

L'évènement $\limsup_{n\to+\infty}(A_n)$ *est un évènement « asymptotique », lorsque les* A_n *sont indépendants il est de probabilité* 0 *ou* 1*, c'est un exemple de vérification d'un résultat appelé la loi du* $0-1$ *de Kolmogorov.*

V.5 Le singe dactylographe

Un « singe dactylographe » tape sur une machine à écrire : à chaque frappe, il tape aléatoirement, avec même probabilité, une des 26 lettres de l'alphabet. Il ne s'arrête jamais. Montrer qu'il tapera presque sûrement une infinité de fois le mot JTKU, les Mémoires du Duc de Saint Simon, le poème « Demain dès l'aube,... », etc...(pour ces deux derniers exemples, on néglige ponctuation, espaces, accents...).

V.6 Pile ou Face : parties simultanées

Trois parties équitables de Pile ou Face, indépendantes, se déroulent simultanément. Montrer que, presque sûrement, il y a au plus un nombre fini d'instants auxquels on a égalité simultanément dans les trois parties.

VI Annexe : l'univers probabilisé du Pile-Face infini

Pour l'univers, pas de problème. On pose

$$\Omega = \{0,1\}^{\mathbf{N}_*}$$

Bien sûr, on pourrait remplacer $\{0,1\}$ par $\{P,F\}$ ou par $\{-1,1\}$. Dans la suite, on comprendra 0 comme « Pile », 1 comme « Face », on note alors que l'avantage du codage $0-1$ est que pour dénombrer les « Face » on n'a qu'à faire la somme des issues. On note p un élément arbitraire de $]0,1[$ (qui désignera la probabilité d'obtenir Face à un tirage quelconque).

Définition On appelle évènement cylindrique (ou de type fini) toute partie X de Ω telle qu'il existe $n \geq 1$ et $A \in \{0,1\}^n$, vérifiant :

$$(\omega \in X) \iff ((\omega_1,\ldots,\omega_n) \in A)$$

Proposition 1 L'ensemble $\mathscr{C}$ des évènements cylindriques est une « algèbre » :

(i) $\Omega \in \mathscr{C}$

(ii) $\left((A,B) \in \mathscr{C}^2\right) \implies (A \cup B \in \mathscr{C})$

(iii) $(A \in \mathscr{C}) \implies \left(\overline{A} \in \mathscr{C}\right)$

Remarque : $\mathscr{C}$ est stable par réunion finie et par intersection finie.

Définition On pose, si $n \geq 1$ et si $(\epsilon_1,\ldots,\epsilon_n) \in \{0,1\}^n$

$$P\left(\{\omega \in \Omega\ ;\ (\omega_1,\ldots,\omega_n) = (\epsilon_1,\ldots,\epsilon_n)\}\right) = p^{s(\epsilon_1,\ldots,\epsilon_n)}(1-p)^{n-s(\epsilon_1,\ldots,\epsilon_n)}$$

où $s(\epsilon_1,\ldots,\epsilon_n) = \sum_{k=1}^{n} \epsilon_k$.

Proposition 2 P s'étend de manière unique à $\mathscr{C}$ en une application que l'on notera encore P, et qui vérifie

(i) $P(\Omega) = 1$ et

(ii) $\forall (A,B) \in \mathscr{C}^2 \qquad A \cap B = \emptyset \ \Rightarrow\ P(A \cup B) = P(A) + P(B)$

Proposition 3 P a la propriété plus forte que **(ii)** suivante :

(ii') Pour toute $\forall (A_n) \in \mathscr{C}^{\mathbf{N}}$, si les A_n sont deux-à-deux disjoints et si $\bigcup_{n=0}^{+\infty} A_n \in \mathscr{C}$, alors $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n)$

Proposition 4 Il existe une unique probabilité sur la tribu $\mathscr{A}$ engendrée par $\mathscr{C}$ qui prolonge P.

Proposition 5 $\mathscr{A} \subsetneq \mathscr{P}(\Omega)$: il y a des parties de Ω qui n'ont pas de probabilité.

La proposition 3 n'est pas trop facile. La proposition 4 non plus : c'est le théorème de Caratheodory. Elle utilise la notion de tribu engendrée qui, elle, ne présente pas de difficulté : on vérifie que l'intersection des tribus contenant une partie de $\mathscr{P}(\Omega)$ donnée est une tribu. La proposition 5 est décevante, car on aimerait bien exhiber des telles parties. Or pour montrer leur existence, on a besoin du célèbre Axiome du Choix, on est donc en pleine théorie des ensembles...remarquons que c'est cela qui oblige à s'occuper de tribus : si on pouvait définir les probabilités, à chaque fois, sur $\mathscr{P}(\Omega)$, la notion de tribu d'évènements serait moins nécessaire.

VII Annexe : diverses choses

VII.1 Quelques univers finis (P1)

a. r objets numérotés dans n boîtes numérotées

On peut prendre naturellement pour Ω l'ensemble des applications de $\{x_1,\dots,x_r\}$ dans $\{b_1,\dots,b_n\}$. Dont le cardinal est n^r. On peut aussi décrire d'autres univers analogues ; par exemple, l'ensemble des suites $\left(b_{i_1},\dots,b_{i_r}\right)$ où $\forall k \in [\![1,r]\!]\ i_k \in [\![1,n]\!]$. Ce qui revient évidemment au même.

b. r objets indiscernables dans n boîtes numérotées

L'univers le plus naturel est sans doute $\Omega = \{(m_1,\dots m_n) \in \mathbf{N}^n\ ;\ m_1+\dots+m_n = r\}$ (m_k désigne le nombre d'objets dans la boîte k). Mais là aussi, plusieurs descriptions alternatives sont possibles ; par exemple, Ω' est l'ensemble des suites de 0 et de 1 comportant $n-1$ fois le chiffre 1 et r fois le chiffre 0 (les 0 représentent les objets, le contenu de la boîte 1 étant le nombre de 0 placés avant le premier 1, le contenu de la boîte k ($2 \leq k \leq n-1$) est le nombre de 0 entre le $k-1$-ième 1 et le k-ième 1, le contenu de la boîte n est le nombre de 0 après le $n-1$-ième 1.
Exemple : $n=3$, $r=7$: on codera l'événement « tous les objets dans la première boîte » par 000000011, « tous les objets dans la dernière boîte » par 110000000, « 2 objets dans la première, 4 dans la deuxième et 1 dans la troisième par 001000010.
Cardinal : un premier calcul
Cette dernière description est d'ailleurs celle qui rend le calcul du cardinal de Ω' (ou de Ω) le plus simple : il vaut

$$\binom{n+r-1}{r}$$

(nombre de manières de placer les r chiffres 0 dans la séquence, ou, ce qui revient au même, nombre de manières de placer les $n-1$ chiffres 1).
La bijection entre Ω' et Ω est facile à visualiser :

$$\underbrace{0\dots0}_{m_1 \text{ fois } 0}\ 1\ \underbrace{0\dots0}_{m_2 \text{ fois } 0}\ 1\ \underbrace{\dots\dots\dots}_{\text{des 0 et des 1}}\ 1\ \underbrace{0\dots0}_{m_n \text{ fois } 0}$$

(des m_k peuvent bien entendu être nuls).

Cardinal : un deuxième calcul
Reprenons $\Omega = \{(m_1,\ldots,m_n) \in \mathbf{N}^n \; ; \; m_1+\ldots+m_n = r\}$. L'application

$$(m_1,\ldots,m_n) \longmapsto (m_1, m_1+m_2+1, m_1+m_2+m_3+2,\ldots, m_1+\cdots+m_{n-1}+(n-2)$$

définie sur Ω, à valeurs dans l'ensemble des suites strictement croissantes de $n-1$ éléments de $[\![0, r+n-2]\!]$, est bijective, la bijection inverse étant

$$(u_1,\ldots,u_{n-1}) \longmapsto (u_1, u_2-u_1-1, u_3-u_2-1,\ldots, u_{n-1}-u_{n-2}-1, r-u_{n-1}-(n-2))$$

Or l'ensemble des suites strictement croissantes de $n-1$ éléments de $[\![0, r+n-2]\!]$ est en bijection naturelle avec l'ensemble des parties à $n-1$ éléments de $[\![0, r+n-2]\!]$ (par l'application qui à une suite strictement croissante associe son « support », c'est-à-dire l'ensemble des valeurs prises par cette suite. On retrouve donc le cardinal : $\binom{r+n-1}{n-1}$ ou $\binom{r+n-1}{r}$.

Cardinal : un troisième calcul
On peut aussi retrouver le résultat par récurrence, à partir de la première description de Ω : si $c_{n,r}$ est le cardinal de Ω, on a $c_{1,r} = 1$ pour tout $r \in \mathbf{N}$. Et ensuite,

$$c_{n,r} = \sum_{k=0}^{r} c_{n-1,r-k}$$

(on classe les éléments de Ω suivant le nombre d'objets placés dans la dernière boîte, par exemple).
Donc $c_{2,r} = r+1$ pour tout entier naturel r (facile à voir). Et pour conclure la récurrence sur n, on doit montrer la formule

$$\forall r \geq 0 \; \forall n \geq 2 \qquad \sum_{k=0}^{r} \binom{n-2+r-k}{r-k} = \binom{n+r-1}{r} = \binom{n+r-1}{n-1}$$

c'est-à-dire

$$\forall r \geq 0 \; \forall n \geq 1 \qquad \sum_{j=0}^{r} \binom{n-2+j}{j} = \binom{n+r-1}{r} = \binom{n+r-1}{n-1}$$

ce qui peut se faire par récurrence sur r : l'initialisation ($r = 0$) ne pose pas de problème, et si c'est vrai pour r,

$$\begin{aligned}\sum_{j=0}^{r+1}\binom{n-2+j}{j} &= \sum_{k=0}^{r}\binom{n-2+j}{j} + \binom{n-2+r+1}{r+1}\\ &= \binom{n-1+r}{r} + \binom{n-1+r}{r+1}\\ &= \binom{n+r}{r+1}\end{aligned}$$

...ou par par un dénombrement des parties à r éléments d'un ensemble E à $n+r-1$ éléments $x_1,\dots,x_{n+r-1}$ effectué en classant ces parties de la manière suivante :
Les parties qui ne contiennent pas x_1 : il y en a $\binom{n-2+r}{r}$;
Les parties qui contiennent x_1 mais ne contiennent pas x_2 : il y en a $\binom{n-3+r}{r-1}$;
Les parties qui contiennent x_1 et x_2 mais ne contiennent pas x_3 : il y en a $\binom{n-4+r}{r-2}$;

...etc...

c. Allumettes de Banach

Une issue est une suite $(a_1,\dots,a_p)$ avec, pour tout i, $a_i \in \{g,d\}$, telle que
$a_p = g$, $|\{i \in [\![1,p-1]\!]\ ;\ a_i = d\}| = n$ et $|\{i \in [\![1,p-1]\!]\ ;\ a_i = g\} \le n$
ou
$a_p = d$, $|\{i \in [\![1,p-1]\!]\ ;\ a_i = g\}| = n$ et $|\{i \in [\![1,p-1]\!]\ ;\ a_i = d\} \le n$.
En remplaçant g par 0 et d par 1 (ne serait-ce que pour favoriser les dénombrements), cela donne : les suites $(a_1,\dots,a_p)$ avec, pour tout i, $a_i \in \{0,1\}$ et telles que $a_p = 1$, $\sum_{k=1}^{p-1} a_k = n$, $p \le 2n+1$ ou $a_p = 0$, $\sum_{k=1}^{p-1}(1-a_k) = n$, $p \le 2n+1$.

VII.2 Le problème des clés

Notons p_k la probabilité cherchée. On a par exemple

$$p_1 = \frac{1}{n}$$

Pour calculer p_2, on appelle A_1 l'événement « La première clé ouvre la porte », de probabilité $\frac{1}{n}$; la probabilité de l'événement A_2 « La deuxième clé ouvre la porte » **sachant** $\overline{A_1}$ est : $1/(n-1)$. Par formule des probabilités composées, on obtient donc

$$p_2 = \frac{n-1}{n} \times \frac{1}{n-1} = \frac{1}{n}$$

puis, notant A_k l'événement « on ouvre la porte au $k$ième essai », les probabilités composées donnent (en notant que $A_k = \overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_{k-1}} \cap A_k$:

$$P(A_k) = P(\overline{A_1})P_{\overline{A_1}}(\overline{A_2})P_{\overline{A_1}\cap\overline{A_2}}(\overline{A_3})\dots P_{\overline{A_{k-1}}\cap\dots\cap\overline{A_{k-1}}}(A_k)$$

Donc

$$P(A_k) = \frac{n-1}{n} \times \frac{n-2}{n-1} \times \dots \times \frac{n-(k-1)}{n-(k-2)} \times \frac{1}{n-(k-1)} = \frac{1}{n}$$

VII.3 Test biologique

On cherche la probabilité pour que l'individu soit malade sachant que le test est positif. Il est donc naturel d'introduire les événements :
A : « l'individu est malade »
B : « l'individu est testé positif »
Le problème est de calculer $P(A|B) = P_B(A)$. Or on nous donne $P_A(B) = 0,99$ et $P_{\overline{A}}(\overline{B}) = 0,999$, ou encore $P_{\overline{A}}(B) = 0.001$. On est donc typiquement dans le cadre d'application de la formule de Bayes, qui permet d'écrire (en introduisant l'autre donnée $P(A) = 0,0001$) :

$$P_B(A) = \frac{0,0001 \times 0,99}{0,0001 \times 0.99 \;+\; 0,9999 \times 0.001} \approx 0.09$$

Il y a donc moins d'une chance sur dix pour que l'individu testé soit malade. Ce résultat n'est pas très intuitif. Le "biais cognitif" qui fait qu'on a tendance à penser que l'individu testé positivement aurait plutôt neuf chances sur dix d'être malade est appelé oubli de la fréquence de base, ou négligence de la taille de l'échantillon. On comprend pourquoi ce biais cognitif fait partie du programme des études de médecine. On comprend aussi pourquoi on évite de faire des tests systématiques de dépistage des maladies très rares : on détecterait principalement de faux malades. Il vaut donc mieux limiter les tests aux éventuelles populations « à risques ».

VII.4 Indicatrice d'Euler

Comme Ω est muni de la probabilité uniforme,

$$P(A_d) = \frac{|A_d|}{|\Omega|} = \frac{\frac{n}{d}}{n} = \frac{1}{d}$$

Les p_i sont deux à deux premiers entre eux, donc, si J est une partie finie de $[\![1, r]\!]$,

$$\bigcap_{i \in J} A_{p_i} = A_{\prod_{i \in J} p_i}$$

Et donc

$$P\left(\bigcap_{i \in J} A_{p_i}\right) = \frac{1}{\prod_{i \in J} p_i} = \prod_{i \in J} P(A_{p_i})$$

ce qui donne bien l'indépendance.
Ensuite, on identifie

$$1 - \frac{1}{p_i} = P\left(\overline{A_{p_i}}\right)$$

et donc, par indépendance des $\overline{A_{p_i}}$,

$$\prod_{i=1}^{r}\left(1 - \frac{1}{p_i}\right) = P\left(\bigcap_{i=1}^{r} \overline{A_{p_i}}\right) = P\left(\overline{\bigcup_{i=1}^{r} A_{p_i}}\right)$$

Or $\left(m \in \overline{\bigcup_{i=1}^{r} A_{p_i}}\right) \iff (n \wedge m = 1)$, donc

$$P\left(\overline{\bigcup_{i=1}^{r} A_{p_i}}\right) = \frac{\phi(n)}{n}$$

ce qui donne la formule sur l'indicatrice d'Euler. Par exemple,

$$\phi(60) = 60 \times (1 - 1/2) \times (1 - 1/3) \times (1 - 1/5) = 16$$

VII.5 Le lemme de Borel-Cantelli

a. Suites réelles

La suite (v_n) est bornée, elle est d'autre part décroissante, donc elle converge. Soit α sa limite. Si $x > \alpha$, il existe n_0 tel que

$$v_{n_0} \leq \frac{x+\alpha}{2}$$

mais alors

$$\forall p \geq n_0 \qquad u_p \leq \frac{x+\alpha}{2}$$

et donc, si une suite extraite de (u_n) converge, sa limite sera $\leq \dfrac{x+\alpha}{2}$, et ne pourra donc pas être égale à x. Donc toute valeur d'adhérence de (u_n) est $\leq \alpha$. Mais , pour tout $\epsilon > 0$, l'ensemble

$$\left\{p \in \mathbf{N} \qquad u_p \in]\alpha-\epsilon, \alpha+\epsilon[\right\}$$

est infini ; sinon, à partir d'un certain rang n_0, on aurait

$$u_p \not\in]\alpha-\epsilon, \alpha+\epsilon[$$

et comme il existe de plus n_1 tel que $v_{n_1} < \alpha+\epsilon$, on aurait

$$\forall p \geq \max(n_0, n_1) \qquad u_p \leq \alpha-\epsilon$$

ce qui oblige $\alpha \leq \alpha-\epsilon$, absurde. Donc, par caractérisation, α est valeur d'adhérence de (u_n), et c'est la plus grande.
Une suite bornée de réels converge si et seulement si ses limites supérieure et inférieure sont égales, par le fait qu'une suite bornée de réels converge si et seulement si elle a une unique valeur d'adhérence.

b. Quelques écritures ensemblistes

Respectivement, on prendra $B = \displaystyle\bigcup_{p=0}^{+\infty}\left(\bigcap_{n=p}^{+\infty} A_n\right)$, $B = \displaystyle\bigcap_{p=0}^{+\infty}\left(\bigcup_{n=p}^{+\infty} A_n\right)$, $B = \displaystyle\bigcup_{p=0}^{+\infty}\left(\bigcap_{n=p}^{+\infty} \overline{A_n}\right)$, encore le même pour finir ! Les réunions sont croissantes, les intersections sont décroissantes, ce qui permet d'indexer à partir de 1, 2 ou plus au lieu de 0.

c. Borel-Cantelli « facile »

Par continuité décroissante,

$$P(\limsup_{n\to+\infty}(A_n) = \lim_{p\to+\infty}\left(P\left(\bigcup_{n=p}^{+\infty} A_n\right)\right)$$

Mais

$$P\left(\bigcup_{n=p}^{+\infty} A_n\right) \leq \sum_{n=p}^{+\infty} P(A_n)$$

Et la suite des restes d'une série convergente converge vers 0, ce qui permet de conclure.

d. Borel-Cantelli « difficile »

Il faut avoir une première idée : l'indépendance se traduit mieux avec des intersections qu'avec des réunions d'événements. On va donc considérer l'évènement complémentaire de la limite supérieure :

$$B = \bigcup_{p=0}^{+\infty}\left(\bigcap_{n=p}^{+\infty}\overline{A_n}\right)$$

Par continuité croissante,

$$P(B) = \lim_{p\to+\infty}\left(P\left(\bigcap_{n=p}^{+\infty}\overline{A_n}\right)\right)$$

Or la famille d'événements $\left(\overline{A_n}\right)_{n\geq 0}$ est une famille d'événements indépendants. On a envie d'écrire, pour utiliser cette indépendance :

$$P\left(\bigcap_{n=p}^{+\infty}\overline{A_n}\right) = \prod_{n=p}^{+\infty} P(\overline{A_n})$$

mais l'étude des produits infinis n'est pas au programme, on va donc simplement s'intéresser aux produits partiels :

$$\begin{aligned} P\left(\bigcap_{n=p}^{+\infty} \overline{A_n}\right) &= \lim_{q\to+\infty}\left(P\left(\bigcap_{n=p}^{q} \overline{A_n}\right)\right) \\ &= \lim_{q\to+\infty}\left(\prod_{n=p}^{q} P(\overline{A_n})\right) \\ &= \lim_{q\to+\infty}\left(\prod_{n=p}^{q} (1-P(A_n))\right) \end{aligned}$$

On peut aussi se contenter, par croissance, de

$$\forall q \geq p \qquad P\left(\bigcap_{n=p}^{+\infty} \overline{A_n}\right) \leq P\left(\bigcap_{n=p}^{q} \overline{A_n}\right) = \prod_{n=p}^{q} (1-P(A_n))$$

La deuxième idée est de prendre le logarithme, l'hypothèse portant sur une série. Mais il faut pour cela supposer qu'au moins à partir d'un certain rang, $P(A_n) < 1$. On fait donc l'hypothèse suivante :

$$\forall n \geq n_0 \qquad P(A_n) < 1$$

Alors, si $n_0 \leq p \leq q$,

$$\ln\left(\prod_{n=p}^{q} (1-P(A_n))\right) = \sum_{n=p}^{q} \ln(1-P(A_n))$$

La suite $(P(A_n))$ ne converge pas vers 0 a priori, donc on ne peut pas utiliser d'équivalents. On utilise donc une comparaison basée sur

$$\forall x > -1 \qquad \ln(1+x) \leq x$$

(ce n'est pas une inégalité du cours, il faut donc savoir la démontrer, en étudiant la fonction $x \longmapsto x - \ln(1+x)$ ou en remarquant que la fonction $x \ln(1+x)$ est concave sur $]-1,+\infty[$, donc au-dessous de sa tangente au point d'abscisse 0). Donc

$$\ln\left(\prod_{n=p}^{q} (1-P(A_n))\right) \leq -\sum_{n=p}^{q} P(A_n)$$

Mais $\sum P(A_n) = +\infty$, donc

$$\sum_{n=p}^{q} P(A_n) \xrightarrow[q\to+\infty]{} +\infty$$

et donc

$$\ln\left(\prod_{n=p}^{q} (1 - P(A_n))\right) \xrightarrow[q\to+\infty]{} -\infty$$

et, donc,

$$P\left(\bigcap_{n=p}^{+\infty} \overline{A_n}\right) = 0$$

Donc $P(B) = 0$ ce qui conclut.
Reste le cas où il y aurait une infinité de A_n tels que $P(A_n) = 1$. Mais dans ce cas, pour tout $p \in \mathbf{N}$,

$$P\left(\bigcup_{n=p}^{+\infty} A_n\right) = 1$$

et le résultat cherché s'ensuit directement.

e. Le singe dactylographe

Soit A_n l'évènement : « le singe frappe J, T, K, U aux frappes $4n+1$, $4n+2$, $4n+3$, $4n+4$. Les A_n sont indépendants, et $P(A_n) = \dfrac{1}{26^4}$, donc $\sum P(A_n) = +\infty$, il suffit alors d"appliquer Borel-Cantelli « difficile » : presque sûrement, une issue appartient à une infinité de A_n. Ou encore : presque sûrement, une infinité de A_n se produisent. A fortiori, on a le résultat (mais on peut préciser ce « a fortiori » en écrivant ce qu'on veut comme réunion de 4 évènements du type A_n). Souvent, on « résout » le « paradoxe » (qui n'en est d'ailleurs pas vraiment un) en cherchant l'espérance du nombre de frappes pour obtenir le premier A_n : 26^4 frappes (voir espérance d'une loi géométrique), ce qui est beaucoup. Pour le poème « Demain dès l'aube... », c'est un peu plus long : ce texte comporte très approximativement 500 lettres. L'espérance est alors 26^{500}. Pas loin de $26^{498}/3$ pages seront utilisées (en moyenne) avant d'avoir ce poème. Ce qui ferait plus de $10^{-3} \times 26^{498}$ kg de papier. C'est beaucoup. Sans compter qu'il faudrait garder un peu de matière dans l'univers pour l'encre, la nourriture du dactylographe...

f. Pile ou Face simultané

L'égalité aux trois tables après $2n$ parties a une probabilité équivalente à $(\pi n)^{-3/2}$ (probabilité d'équirépartition, équivalent obtenu avec Stirling par exemple). Qui est le terme général d'une série convergente. Il n'y a donc qu'à appliquer Borel-Cantelli « facile ».

Table des matières