Chapitre
Espaces probabilisés

I Résumé des techniques et formules de dénombrement

Le nombre d'éléments d'un ensemble fini E est appelé cardinal de E et noté Card(E) , $\#(E)$ ou $|E|$.

Si E et F sont des ensembles finis :

- Cardinal d'un produit. $|E \times F| =$
- Nombre d'applications de E dans F. $\left|F^E\right| =$
- Nombre d'injections de E dans F. On note $p = |E|$ et $n = |F|$:
- Nombre de bijection de E dans E ou nombre de permutation de E :
- Nombre de parties de E. $|\mathcal{P}(E)| =$

Soit E un ensemble de cardinal n et p un entier naturel non nul.

- Nombre de p-listes (ou p-uplet de E) :
- Nombre de p-listes d'éléments distincts de E :
- Nombre de parties de E à p éléments ou p-combinaisons :

✎ Méthode pratique ✎ (Méthode classiques de dénombrement)

- ▶ **Principe multiplicatif**. Lorsqu'un dénombrement de décompose en étapes successives, les possibilités offertes à chaque étape se multiplient.
- ▶ **Passage au complémentaire**. Il est parfois plus simple de dénombrer Card(A^c) que Card(A). Souvent des énoncés à base de: “au moins...”.
- ▶ **Disjonctions de cas**. On décompose l'ensemble à dénombrer en réunion d'ensemble deux à deux disjoints.
- ▶ **Reconnaître une situation de référence**. Les problèmes de dénombrement se ramènent souvent à l'une des situations suivantes pour lesquelles on connait des résultats de dénombrement:
  - listes si éléments choisis avec **ordre**, **répétition** (n^p)
  - listes d'éléments distincts si éléments choisis **avec ordre**, **sans répétition** ($\dfrac{n!}{(n-p)!}$)
  - permutation si **ordre, pas de répétition et tous les éléments sont pris** ($n!$)
  - combinaisons quand choix simultané d'objets donc **sans ordre, sans répétition**. ($\dbinom{n}{p}$).
- ▶ **Introduire une bijection.** Dans certains cas, on introduit une bijection entre l'ensemble à dénombrer et un ensemble pus facile à dénombrer.

Exemple Modèle d'urne. On extrait p boules d'une urne contenant n boules, on s'intéresse au nombre de tirages possibles. Tout dépend de la nature des tirages.

1) tirages successifs sans remise :

2) tirages successifs avec remise :

3) tirages simultanés :

Exemple

1) On tire 5 cartes dans un jeu de 32 cartes.
   - -a- Combien de mains possibles?
   - -b- Combien de mains avec exactement 2 as?
   - -c- Combien de mains avec au moins un as?
   - -d- Combien de mains avec exctement un coeur et exactement un roi?
   - -e- Combien de mains avec au moins un coeur et au moins un pique?

Quelques formules sur les coefficients binomiaux. Soient n, p deux entiers naturel.

- $\binom{n}{0} = \binom{n}{n} =$ $\qquad \binom{n}{1} = \binom{n}{n-1} =$

- Formule de symétrie. Si $p \leqslant n$,

- Formule du triangle de Pascal. Si $1 \leqslant p \leqslant n-1$,

- Formule du sélectionneur ou du capitaine. Si $1 \leqslant p \leqslant n$,

- Formule du binôme. Si $a, b \in \mathbb{C}$,

$$(a+b)^n =$$

- $\displaystyle\sum_{k=0}^{n} \binom{n}{k} =$ $\qquad \displaystyle\sum_{k=0}^{n} (-1)^k \binom{n}{k} =$

II Ensembles dénombrables

Définition (Ensemble dénombrable)

Un ensemble E est dit **dénombrable** si et seulement si, il existe une bijection $\varphi : \mathbb{N} \to E$.
Il est dit au plus dénombrable si et seulement si, il est fini ou il est dénombrable.

☞ **Explication** ☜ Un ensemble E est dénombrable si et seulement si on peut numéroter ses éléments par des entiers naturels.
En notant $x_n = \varphi(n)$, les x_n sont deux à deux distincts car φ est bijective, on peut écrire

$$E = \{x_n \mid n \in \mathbb{N}\}.$$

Exemples

1) Soit $p \in \mathbb{N}$. L'ensemble $[\![p, +\infty[\![$ est dénombrable.

2) Les ensembles des entiers pairs, des entiers impairs sont dénombrables.

Théorème

Tout ensemble de la forme $\{x_i \;/\; i \in I\}$ où I est une partie de $\mathbb{N}$ et les x_i sont deux à deux distincts est au plus dénombrable.

Propriétés ($\mathbb{Z}$)

$\mathbb{Z}$ est dénombrable.

Propriétés (Le produit cartésien d'ensembles dénombrables est dénombrable)

- **Cas de deux ensembles** Si E et F sont dénombrables alors $E \times F$ est dénombrable.
- **Cas de p ensembles** Si $E_1, \ldots, E_p$ sont dénombrables alors $E_1 \times \ldots \times E_p$ sont dénombrables.

Idée de la démonstration : repose sur le fait que $\mathbb{N} \times \mathbb{N}$ est dénombrable.

Propriétés (La réunion dénombrable d'ensembles dénombrables est dénombrable)

Soit I un ensemble au plus dénombrable et $(E_i)_{i \in I}$ une famille d'ensembles dénombrables alors $\displaystyle\bigcup_{i \in I} E_i$ est dénombrable.

Propriétés (Partie d'un ensemble dénombrable)

Toute partie d'un ensemble dénombrable est dénombrable.

NB: $\mathbb{R}$ n'est pas dénombrable.

III Familles sommables

Ce paragraphe a pour but de donner des outils pour manipuler des sommes infinies rencontrées dans le cadre des probabilités. La démonstration des résultats est hors-programme, l'utilisation de ces résultats est limitée au cours de probabilités.

III.1 Familles sommables de nombres réels positifs

☞ **Explication** ☜ Soit I un ensemble dénombrable et une famille $(x_i)_{i \in I}$ de nombres réels. On cherche à donner un sens à la somme de tous les termes de cette famille, $\displaystyle\sum_{i \in I} x_i$, dans quel ordre on somme ?

Comme I est dénombrable, il existe une bijection $\varphi : \mathbb{N} \to I$. Il s'agit donc de calculer la somme des $x_{\varphi(n)}$ lorsque $n \in \mathbb{N}$, ce qui correspond à la somme de la série $\displaystyle\sum_{n \in \mathbb{N}} x_{\varphi(n)}$ (c'est-à-dire la limite de la suite des sommes partielles).

Une question cruciale se pose alors : la convergence de la série $\displaystyle\sum_{n \in \mathbb{N}} x_{\varphi(n)}$ et la valeur de la somme dépend-elle du choix de φ c'est-à-dire de la numérotation des éléments. La réponse est surprenante...

Exemple On somme les éléments de la famille $\left(\frac{(-1)^n}{n}\right)_{n\in\mathbb{N}^*}$.

On a déjà vu en sommant dans l'ordre naturel $\displaystyle\sum_{n=1}^{+\infty}\frac{(-1)^n}{n} = -\ln(2)$.

On somme désormais autrement, en sommant deux fois plus vite les termes d'indice pair que les termes d'indice impair

$$\begin{aligned}
S &= \underbrace{-1+\frac{1}{2}}_{=-\frac{1}{2}}+\frac{1}{4}\underbrace{-\frac{1}{3}+\frac{1}{6}}_{=-\frac{1}{6}}+\frac{1}{8}\underbrace{-\frac{1}{5}+\frac{1}{10}}_{=-\frac{1}{10}}+\frac{1}{12}+\cdots \;.....\\
&= -\frac{1}{2}+\frac{1}{4}-\frac{1}{6}+\frac{1}{8}-\frac{1}{10}+\frac{1}{12}+\cdots \;....\\
&= \frac{1}{2}\left(-1+\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\frac{1}{5}+\frac{1}{6}+\cdots \;....\right)\\
&= -\frac{\ln 2}{2}.
\end{aligned}$$

Le résultat est donc différent !!

Cependant dans le cas de termes positifs, cette curiosité disparaît.

Théorème (Invariance de la somme par permutation)

Soient $(u_n)_{n\in\mathbb{N}}$ une suite de réels positifs et $\phi : \mathbb{N} \to \mathbb{N}$ une application bijective.
La série $\sum u_n$ converge si et seulement si $\sum u_{\varphi(n)}$ converge. En cas de convergence, les deux séries ont la même somme.

Définition (Famille sommable de réels positifs)

Soit $(x_i)_{i\in I}$ une famille de réels positifs.

- Dans le cas de convergence, la famille $(x_i)_{i\in I}$ est dite **sommable** et la valeur $\displaystyle\sum_{n=0}^{+\infty} x_{\varphi(n)}$ est indépendante du choix de la bijection $\varphi : \mathbb{N} \to I$, cette valeur est appelée **somme de la famille** $(x_i)_{i\in I}$ et est notée $\displaystyle\sum_{i\in I} x_i$.
- S'il n'existe aucune bijection $\varphi : \mathbb{N} \to I$ telle que la série $\displaystyle\sum_{n\in\mathbb{N}} x_{\varphi(n)}$ converge. On note alors :

$$\sum_{i\in I} x_i = +\infty.$$

Remarques (Cas où x_i peut valoir $+\infty$)

On peut également considérer le cas où l'un au moins des x_i vaut $+\infty$ (les autres x_i sont toujours positifs), dans ce cas on au aussi

$$\sum_{i\in I} x_i = +\infty.$$

Par conséquent, toute famille d'éléments de $[0, +\infty]$, on peut écrire $\displaystyle\sum_{i\in I} x_i$ et le résultat appartient à $[0, +\infty]$.

La famille est sommable si et seulement si $\displaystyle\sum_{i\in I} x_i < +\infty$.

Théorème (Sommation par paquets)

Soit $(x_i)_{i\in I}$ une famille d'éléments de $[0,+\infty]$ et $I=\bigcup_{n\in\mathbb{N}} I_n$ où les I_n sont deux à deux disjoints. Alors

$$\sum_{i\in I} x_i = \sum_{n=0}^{+\infty}\left(\sum_{i\in I_n} x_i\right).$$

Remarques (Découpage fini)

Le résultat reste vrai lorsque I est découpé en un nombre fini de I_n. Par exemple, si $\mathcal{P}$ désigne l'ensemble des nombres pairs, si $\mathcal{I}$ désigne l'ensemble des nombres impairs,

$$\sum_{n\in N} x_n = \sum_{n\in\mathcal{P}} x_n + \sum_{n\in\mathcal{I}} x_n.$$

NB : attention de bien vérifier l'hypothèse $x_n \geqslant 0$.

Théorème (Fubini)

Soient I, J deux ensembles dénombrables et $(x_{ij})_{(i,j)\times I\times J}$ une famille d'éléments de $[0,+\infty]$. On a l'égalité

$$\sum_{(i,j)\times I\times J} x_{ij} = \sum_{i\in I}\left(\sum_{j\in J} x_{ij}\right) = \sum_{j\in I}\left(\sum_{i\in J} x_{ij}\right).$$

☞ **Explication** ☜ On reconnaît dans le cas où I et J sont finis la sommation sur un rectangle vue en PCSI.

✎ **En pratique** ✎ Dans le cas positif, on peut découper, calculer, majorer les sommes directement. La finitude de la somme prouve la sommabilité.

III.2 Familles sommables de nombres complexes

Définition (Famille sommable de complexes)

Soit $(x_i)_{i\in I}$ une famille dénombrable de nombres complexes. La famille est dite sommable si la famille $(|x_i|)_{i\in I}$ l'est.

Théorème-Définition (Somme d'une famille sommable)

Soit $(x_i)_{i\in I}$ une famille sommable de nombres complexes. Alors pour toute bijection $\varphi : \mathbb{N}\to I$, la série $\sum_{n\in\mathbb{N}} x_{\varphi(n)}$ est convergente et sa somme est indépendante de la bijection choisie.
Cette somme est notée $\sum_{i\in I} x_i$ et appelée **somme de la famille sommable**.

⚠ **Attention** ⚠ Contrairement aux familles de réels positifs, on n'utilise la notation $\sum_{i\in I} x_i$ que si la famille est sommable. Sinon, on ne l'utilise pas.
De plus pour les familles de réels positifs, pour lesquels l'obtention d'un résultat fini à la fin des calculs justifie a

posteriori la sommabilité de la famille, il est indispensable, dans le cas d'une famille de réels quelconques, de justifier la sommabilité en préalable à tout calcul.

Remarques (Lien avec les séries absolument convergentes)

Dans le cas où $I = \mathbb{N}$, la famille $(x_n)_{n\in\mathbb{N}}$ est sommable si et seulement si la série $\sum_{n\in\mathbb{N}} x_n$ est absolument convergente.

✎ **En pratique** ✎ Comme pour les séries, on utilise des comparaisons pour prouver qu'une famille est sommable.

Propriétés (Comparaison à l'aide d'inégalité)

Soient $(x_i)_{i\in I}$ une famille dénombrable de nombres complexes et $(y_i)_{i\in I}$ une famille de réels positifs.
On suppose que : $\forall i \in I$, $|x_i| \leqslant y_i$.
Si la famille $(y_i)_{i\in I}$ est sommable, alors $(x_i)_{i\in I}$ l'est aussi.

En cas de sommabilité, les sommes se manipulent naturellement grâce aux différentes propriétés énoncées ci-dessous.

Propriétés (Croissance)

Soient $(x_i)_{i\in I}$, $(y_i)_{i\in I}$ deux familles sommables de nombres réels.
On suppose que : $\forall i \in I$, $x_i \leqslant y_i$. Alors : $\sum_{i\in I} x_i \leqslant \sum_{i\in I} y_i$.

Propriétés (Linéarité)

Soient $(x_i)_{i\in I}$, $(y_i)_{i\in I}$ deux familles sommables de nombres complexes et $\lambda \in \mathbb{C}$.
Alors la famille $(\lambda x_i + y_i)_{i\in I}$ est sommable et

$$\sum_{i\in I}(\lambda x_i + y_i) = \lambda \sum_{i\in I} x_i + \sum_{i\in I} y_i.$$

Théorème (Sommation par paquets)

Soit $(x_i)_{i\in I}$ une famille sommable de nombres complexes et $I = \bigcup_{n\in\mathbb{N}} I_n$ où les I_n sont deux à deux à deux disjoints. Alors

$$\sum_{i\in I} x_i = \sum_{n=0}^{+\infty} \left(\sum_{i\in I_n} x_i \right).$$

NB: le résultat reste vrai lorsque I est découpé en un nombre fini de I_n.

Théorème (Fubini)

Soient I, J deux ensembles dénombrables et $(x_{ij})_{(i,j)\times I\times J}$ une famille sommable de nombres complexes. Alors

- pour chaque $i \in I$, la famille $(x_{ij})_{j\in J}$ est sommable
- la famille $\left(\sum_{i\in I} x_{ij}\right)_{j\in J}$ est sommable, et sa somme vérifie

$$\sum_{(i,j)\times I\times J} x_{ij} = \sum_{i\in I}\left(\sum_{j\in J} x_{ij}\right) = \sum_{j\in I}\left(\sum_{i\in J} x_{ij}\right).$$

Propriétés (Produit de sommes)

Soient $(x_i)_{i\in I}$, $(y_i)_{i\in I}$ deux familles sommables de nombres complexes.
Alors la famille $(x_i y_j)_{(i,j)\in I\times J}$ est sommable et

$$\sum_{(i,j)\in I\times J} x_i y_j = \left(\sum_{i\in I} x_i\right)\left(\sum_{j\in J} y_j\right) = \left(\sum_{j\in J} y_j\right)\left(\sum_{i\in I} x_i\right).$$

IV Espaces probabilisés

IV.1 Tribus

Lorsque l'univers est infini, chercher à associer une probabilité à chaque partie de Ω est une tâche difficile, parfois impossible.
On se limite à définir la probabilité de certaines parties vérifiant des bonnes propriétés vis-à-vis des opérations effectuées sur les événements.

Définition (Tribu)

Soit Ω un ensemble non vide (univers). On appelle tribu sur Ω toute partie $\mathcal{A}$ de $\mathcal{P}(\Omega)$ telle que :

- $\Omega \in \mathcal{A}$
- $\forall A \in \mathcal{A}$, $\overline{A} \in \mathcal{A}$
- pour tout suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

$(\Omega, \mathcal{A})$ est appelé **espace probabilisable**. Les éléments de $\mathcal{A}$ sont appelés **événements**.

Exemples Tribus triviales

1) L'ensemble $\mathcal{P}(\Omega)$ est une tribu sur Ω. C'est la plus souvent utilisée, lorsque Ω est fini ou dénombrable.
2) L'ensemble $\{\emptyset, \Omega\}$ est un une tribu.

Propriétés

Soit $(\Omega, \mathcal{A})$ un espace probabilisable.

1) $\emptyset \in \mathcal{A}$
2) **Stabilité par intersection dénombrable**. Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{A}$, $\bigcap\limits_{n\in\mathbb{N}} A_n \in \mathcal{A}$.
3) **Stabilité par intersection/union finie**. Si $(A_0, \dots, A_n)$ est une suite finie d'éléments de $\mathcal{A}$, alors
$$\bigcup_{i=0}^{n} A_i \in \mathcal{A} \qquad\qquad \bigcap_{i=0}^{n} A_i \in \mathcal{A}.$$
4) **Stabilité par différence**. Si $A, B \in \mathcal{A}$ alors $A \setminus B \in \mathcal{A}$.

☞ **Explication** ☜ Une tribu est donc stable par réunion/intersection au plus dénombrable.

Remarques (Réalisation de $\bigcup\limits_{n=0}^{+\infty} A_n$ et $\bigcap\limits_{n=0}^{+\infty} A_n$)

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements.

$$\omega \in \bigcup_{n=0}^{+\infty} A_n \Leftrightarrow \exists n \in \mathbb{N} \,/\, \omega \in A_n \qquad (\omega \text{ est dans l'un des } A_n)$$

$$\omega \in \bigcap_{n=0}^{+\infty} A_n \Leftrightarrow \forall n \in \mathbb{N} \,/\, \omega \in A_n \qquad (\omega \text{ est dans tous les } A_n)$$

Remarques (Vocabulaire)

On généralise aux tribus le vocabulaire déjà vu en première année :

- A, B incompatibles si $A \cap B = \emptyset$
- événement impossible
- Ω événement certain
- $\overline{A}$ événement contraire
- ω réalise A si $\omega \in A$

IV.2 Probabilité

Définition (Probabilité)

Soit $(\Omega, \mathcal{A})$ un espace probabilisable.
Une probabilité sur $(\Omega, \mathcal{A})$ est une application $P : \mathcal{A} \to [0,1]$ vérifiant :

- $P(\Omega) = 1$
- pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux incompatibles, on a

$$P\left(\bigsqcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n) \qquad (\sigma\text{-additivité}).$$

Le triplet $(\Omega, \mathcal{A}, P)$ est appelé **espace probabilisé**.
NB : cela sous-entend implicitement que la série $\sum P(A_n)$ est convergente.

Exemple Probabilité uniforme

1) Si Ω est fini, la probabilité uniforme sur $\mathcal{P}(\Omega)$ est définie par $P(A) = \dfrac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)}$.
Cette probabilité est caractérisée par le fait que les événements élémentaires $\{w\}$ où $\omega \in \Omega$ ont tous même probabilité.
Par exemple pour le lancer d'un dé normal, $\Omega = [\![1,6]\!]$, $P(\{i\}) = \dfrac{1}{6}$.

2) Si Ω est dénombrable, $\Omega = \{\omega_n \,/\, n \in \mathbb{N}\}$ on ne peut définir de probabilité uniforme. En effet s'il existait $p \in [0,1]$ tel que : $\forall n \in \mathbb{N}$, $p_n = P(\{\omega_n\}) = p$, la série $\sum p_n$ serait grossièrement divergente et contredirait

$$1 = P(\Omega) = P\left(\bigsqcup_{n=0}^{+\infty} \{\omega_n\}\right) = \sum_{n=0}^{+\infty} p_n.$$

Exemple Cas d'un univers dénombrable Posons $\Omega = \mathbb{N}^*$. Soit P une probabilité sur Ω, telle qu'il existe $\alpha \in \mathbb{R}$ tel que : $\forall k \in \mathbb{N}^*$, $P(\{k\}) = \dfrac{\alpha}{k!}$.
Déterminer α.

Exemple Cas d'un univers non dénombrable On considère un jeu de Pile ou Face infini. L'univers est $\{0,1\}^{\mathbb{N}}$ (0 pour face, 1 pour pile) (on peut démontrer que cet ensemble n'est pas dénombrable).
Pour $i \in \mathbb{N}^*$, on note P_i l'événement : "on obtient pile au i-ème lancer". Exprimer, à l'aide des P_i, les événements suivants :

1) A : "on obtient au moins un pile"

2) B : "on obtient deux piles consécutifs"

3) C : "on obtient le premier pile au k-ième lancer".

Définition

Soit A un événement d'un espace probabilisé $(\Omega, \mathcal{A})$.

- On dit que A est un événement presque sûr lorsque $P(A) = 1$.
- On dit que A est un événement négligeable lorsque $P(A) = 0$.

⚠ **Attention** ⚠ Dans le cas d'un univers non fini on peut avoir $P(A) = 0$ sans avoir $A = \emptyset$.

Exemple Retour sur le lancer de pile ou face infini avec une pièce normale. On s'intéresse à l'expérience consistant à jeter la pièce jusqu'à obtenir pile.
On pose pour $n \in \mathbb{N}^*$, D_n : "pile apparaît la première fois au n-ième tirage". Par indépendance des lancers $P(D_n) = \frac{1}{2^n}$.
On pose $D = \bigsqcup_{n \in \mathbb{N}^*} D_n$. On montre que $P(D) = 1$ et $P(\overline{D}) = 0$.

IV.3 Propriétés des probabilités

On retrouve des propriétés déjà vues pour des univers finis.

Théorème (Proprietés calculatoires)

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

1) Pour toute suite finie $(A_0, \ldots, A_n)$ d'événements deux à deux incompatibles,

$$P\left(\bigsqcup_{i=0}^{n} A_i\right) = \sum_{i=0}^{n} P(A_i).$$

2) $P(\emptyset) = 0$

3) Si A est un événement, $P(\overline{A}) = 1 - P(A)$

4) Si $A \subset B)$ sont deux événements, alors

$$P(A) \leqslant P(B) \qquad \text{et} \qquad P(B \setminus A) = P(B) - P(A).$$

5) Si A, B sont deux événements, alors

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

Puis un résultat adapté aux réunions infinies.

Théorème (Continuité monotone)

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

- **Continuité croissante**. Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements croissante pour l'inclusion, c'est-à-dire : $\forall n \in \mathbb{N}$, $A_n \subset A_{n+1}$, alors

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

- **Continuité décroissante**. Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements décroissante pour l'inclusion, c'est-à-dire : $\forall n \in \mathbb{N}$, $A_{n+1} \subset A_n$, alors

$$\lim_{n \to +\infty} P(A_n) = P\left(\bigcap_{n=0}^{+\infty} A_n\right).$$

Exemple Retour sur l'exemple du pile ou face infini.
Pour $n \in \mathbb{N}^*$, on pose A_n : "obtenir au moins un pile avant le n-ième lancer" et B_n : "ne pas obtenir de Pile pendant les n premiers lancers".
Exprimer A_n à l'aide des P_i. Monotonie des suites (A_n) et (B_n)?

Corollaire (Probabilité d'une intersection/réunion)

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements (pas nécessairement monotone), on a

$$\lim_{n\to+\infty} P\left(\bigcup_{k=0}^{n} A_k\right) = P\left(\bigcup_{k=0}^{+\infty} A_k\right) \qquad \lim_{n\to+\infty} P\left(\bigcap_{k=0}^{n} A_k\right) = P\left(\bigcap_{k=0}^{+\infty} A_k\right).$$

Corollaire (Sous-additivité)

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements (pas nécessairement monotone), on a

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

NB: la somme vaut $+\infty$ lorsque la série diverge.

V Variables aléatoires discrètes

Définition (Variable aléatoire discrète)

Soit $(\Omega, \mathcal{A})$ un espace probabilisable. Une variable aléatoire discrète sur $(\Omega, \mathcal{A})$ est une application de Ω dans un ensemble E telle que :

- l'image de Ω, $X(\Omega)$ est une partie au plus dénombrable de E
- pour tout $x \in X(\Omega)$, $X^{-1}(\{x\})$ est un événement, c'est-à-dire un élément de $\mathcal{A}$.

☞ **Explication** ☜ On rappelle que $X(\Omega) = \{X(\omega) \,/\, \omega \in \Omega\}$ est l'ensemble des valeurs prises par X, appelé **univers-image** de X.
Comme il est au plus dénombrable, ce sera souvent $\mathbb{N}$, $\mathbb{N}^*$, $\mathbb{Z}$.

Notation L'événement $X^{-1}(\{x\}) = \{\omega \in \Omega \,/\, X(\omega) = x\}$ est noté $(X = x)$ ou $\{X = x\}$.

✎ **En pratique** ✎ En général, l'univers Ω n'est pas explicité pas plus que la tribu. Il sera implicitement supposé que l'événement $(X = x)$ est élément de la tribu, l'objectif étant d'en calculer sa probabilité.

Exemples

1) Pile ou face. On peut poser la variable aléatoire X valant 0 si on obtient face, et 1 si pile.

2) Lancer d'un dé normal à 6 faces. On peut poser la variable aléatoire X, qui indique le numéro du tirage, d'univers image $[\![1, 6]\!]$.

3) Lancer de deux dés consécutifs. On peut poser la variable aléatoire X, qui indique le couple résultat, d'univers image $[\![1, 6]\!]^2$.

4) Jeu de pile ou face infini. On pose la variable aléatoire T donnant le rang du premier pile, avec pour convention $T = +\infty$ si pile n'apparaît pas. Alors T prend ses valeurs dans $\mathbb{N}^* \cup \{+\infty\}$ qui est dénombrable.

Propriétés ($X \in B$)

Soit $(\Omega, \mathcal{A})$ un espace probabilisable et X une variable aléatoire discrète à valeurs dans un ensemble E.
Pour toute partie B de E, l'ensemble $X^{-1}(B)$ est un événement, c'est-à-dire un élément de $\mathcal{A}$. Cet événement est noté $(X \in B)$.

☞ **Explication** ☜ On rappelle que $X^{-1}(B)$ est l'image réciproque de B par X c'est-à-dire l'ensemble des antécédents des éléments de B par X,

$$X^{-1}(B) = \{\omega \in \Omega \ / \ X(\omega) \in B\}.$$

Notation. Si X est à valeurs réelles et si I est un intervalle alors $X^{-1}(I)$ est un événement noté $(X \in I)$.
Si $I =]-\infty, x]$, $(X \in I)$ est noté $(X \leqslant x)$; si $I = [x, +\infty[$, $(X \in I)$ est noté $(X \geqslant x)$.

Souvent à partir d'une variable aléatoire X, on construit d'autres variables aléatoires, X^2, $|X|$...Il faut alors vérifier que ce sont des variables aléatoires.

Propriétés (Image d'une variable aléatoire)

Soit $(\Omega, \mathcal{A})$ un espace probabilisable et X une variable aléatoire discrète. On considère f une application définie sur $X(\Omega)$ à valeurs dans un ensemble E.
Alors $f \circ X$ est une variable aléatoire, que l'on note $f(X)$.

Définition (Loi d'une variable aléatoire)

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une variable aléatoire discrète.
La loi de la variable aléatoire X est la donnée de

- l'ensemble $X(\Omega)$ des valeurs prises par X
- pour tout $x \in X(\Omega)$, la probabilité $P(X = x)$ de l'événement $(X = x)$.

Exemples

1) Lancer d'un dé normal à 6 faces. On pose la variable aléatoire X, qui indique le numéro du tirage. Déterminer la loi de X.

2) On considère une urne contenant N boules numérotées de 1 à N. On tire une poignée de p boules simultanément. On définit les deux variables aléatoires réelles :

   - X le plus grand numéro tiré
   - Y le plus petit numéro tiré.

   Déterminer les lois de X et Y.

Théorème (Loi P_X de X)

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une variable aléatoire discrète.
Pour toute partie B de $X(\Omega)$, on définit : $P_X(B) = P(X \in B)$.
Alors P_X est une probabilité sur $X(\Omega)$ muni de la tribu pleine $\mathcal{P}(X(\Omega))$, P_X est appelée loi de X.
De plus si $B = \{b_i \ / \ i \in I\}$ où I est au plus dénombrable et les b_i sont deux à deux distincts :

$$P_X(B) = \sum_{i \in I} P(X = b_i) = \sum_{b \in B} P(X = b).$$

Exemple Soit X une variable aléatoire à valeurs dans N^*, tel qu'il existe $a \in \mathbb{R}$ tel que :
$\forall n \in \mathbb{N}^*$, $P(X = x) = \dfrac{a}{k(k+1)}$. Déterminer a.

Notation. Si X et Y sont deux variables aléatoires suivant la même loi, c'est-à-dire $X(\Omega) = Y(\Omega)$ et $P_X = P_Y$, on note $X \sim Y$.

Propriétés

Soient $(\Omega, \mathcal{A}, P)$ un espace probabilisé, X, Y deux variables aléatoires discrètes et f une application définie sur $X(\Omega) = Y(\Omega)$.
Si $X \sim Y$ alors $f(X) \sim f(Y)$.

VI Probabilités conditionnelles

VI.1 Définitions et premières propriétés

Dans toute cette section $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

Théorème-Définition (Probabilité conditionnelle)

Soit B un évènement tel que $P(B) > 0$.

- Pour tout événement A de $\mathcal{P}(\Omega)$ on définit **la probabilité conditionnelle de A sachant B** par :

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \qquad \text{noté aussi } P(A \,|\, B).$$

- L'application $\begin{array}{rccc} P_B : & \mathcal{P}(\Omega) & \to & [0,1] \\ & A & \mapsto & P_B(A) \end{array}$ est bien une probabilité sur Ω.

Remarques

- On ne sait pas définir $P_B(A)$ lorsque $P(B) = 0$, mais par convention, $P(B)P_B(A) = 0$ lorsque $P(B) = 0$ et donc l'égalité $P(A \cap B) = P(B)P_B(A)$ reste tout de même vraie dans ce cas.
- ⚠ **Attention** ⚠ Ne pas confondre la compréhension $P(A \cap B)$ et $P_B(A)$. Pour $P(A \cap B)$ on confronte les cas favorables à $A \cap B$ à tous les cas favorables à Ω, pour $P_B(A)$ on confronte les cas favorables à $A \cap B$ à tous les cas favorables à B.
- $P_B(A)$ est parfois noté $P(A \,|\, B)$, nous n'utiliserons pas cette notation, un peu ambigüe car $A \,|\, B$ n'est pas un événement.
- P_B étant une probabilité, elle vérifie toutes les propriétés calculatoires d'une probabilité : probabilité d'une union, du contraire, croissance.

VI.2 Probabilités composées

Théorème (Formule des probabilités composées)

Soient $n \geqslant 2$, $(A_i)_{1 \leqslant i \leqslant n}$ une famille d'événements telle que $P(A_1 \cap \cdots A_{n-1}) > 0$,

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \times P_{A_1}(A_2) \times \cdots \times P_{A_1 \cap \cdots \cap A_{n-1}}(A_n).$$

Remarques

- Dans le cas $n = 2$, on retrouve la réécriture de définition de la probabilité conditionnelle :
$$P(A_1 \cap A_2) = P(A_1)P_{A_1}(A_2).$$
- Comme $P(A_1 \cap \cdots A_{n-1}) > 0$, alors $P(A_1) > 0$, $P(A_1 \cap A_2) > 0$ donc toutes les probabilités conditionnelles sont bien définies.

Exemple Un Professeur de Mathématiques a décidé de faire l'effort de ne plus s'énerver.
S'il ne s'énerve pas un jour donné, alors la probabilité qu'il ne s'énerve pas le lendemain est 0.3.
En revanche, s'il s'énerve un jour donné, alors la probabilité qu'il ne s'énerve pas le lendemain est 0.9.
Le premier jour, jour de la rentrée, le professeur n'est pas énervé.
Pour $n \in \mathbb{N}^*$, on pose l'évènement E_n: "le professeur est énervé le n-ième jour" et $p_n = P(E_n)$.
Quelle est la probabilité que le professeur soit énervé les trois jours consécutifs: 2ème, 3ème et 4ème jour?

Exemple On considère une urne dans laquelle se trouve une boule rouge et une boule verte. On tire une boule de cette urne.
Si elle est verte on la remet avec une autre boule verte et on recommence; si elle est rouge le jeu s'arrête.

1) Soit $k \in \mathbb{N}^*$. Quelle est la probabilité que les tirages s'arrêtent au k-ème tirage?
2) Quelle est la probabilité que les tirage s'arrêtent?

VI.3 Formule des probabilités totales

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

Soit $(A_i)_{i\in I}$ une famille d'événements au plus dénombrable.

- $(A_i)_{i\in I}$ est un système complet d'événements si les A_i sont deux à deux incompatibles et si la réunion est Ω c'est-à-dire :
$$\forall (i,j) \in I^2, \quad i \neq j \Rightarrow A_i \cap A_j = \emptyset \qquad\qquad \bigsqcup_{i\in I} A_i = \Omega.$$
- $(A_i)_{i\in I}$ est un système quasi-complet d'événements si les A_i sont deux à deux incompatibles et si leur réunion est un événement presque sûr c'est-à-dire :
$$\forall (i,j) \in I^2, \quad i \neq j \Rightarrow A_i \cap A_j = \emptyset \qquad\qquad P\left(\bigsqcup_{i\in I} A_i\right) = 1 \quad \left(\text{ou } \sum_{i\in I} P(A_i) = 1\ \right).$$

NB: "complet" $\Rightarrow$ "quasi-complet"

Théorème (Formule des probabilités totales)

Soient $(A_i)_{1\leqslant i\leqslant n}$ un système quasi-complet ou complet d'évènements de probabilité non nulle et un évènement B :
$$P(B) = \sum_{i\in I} P(B \cap A_i) = \sum_{i\in I}^{n} P_{A_i}(B)P(A_i).$$
si $P(A_i) = 0$.
NB : la formule reste vraie si des A_i sont de proba non nulle avec la convention déjà vue : $P_{A_i}(B)P(A_i) = 0$

Exemple Lancer de deux pièces Deux pièces de monnaie ont pour probabilités respectives de tomber sur pile: 0.4, 0.7. On choisit une pièce au hasard.

1) Quelle est la probabilité d'obtenir pile?
2) On lance deux fois cette pièce préalablement choisie, on obtient deux piles. Quelle est la probabilité d'avoir choisi la première pièce.

Exemple Retour sur l'exemple du professeur énervé

1) Que vaut p_1? Calculer p_2.
2) Soit $n \in \mathbb{N}^*$. Exprimer p_{n+1} en fonction de p_n. En déduire une expression de p_n en fonction de n.

✎ Méthode pratique ✎ (Obtenir le terme général d'une suite arithmético géométrique)

Soit $(a, b) \in \mathbb{R}^2$, $a \neq 1$ et $(u_n)_{n\in\mathbb{N}}$ une suite vérifiant : $\forall n \in \mathbb{N},\ u_{n+1} = au_n + b\ (*)$.

- On détermine le point fixe α, vérifiant l'équation $\alpha = a\alpha + b\ (**)$.
- On soustrait les relations $(*)$ et $(**)$: $\forall n \in \mathbb{N},\ u_{n+1} - \alpha = a(u_n - \alpha)$.
  La suite $(u_n - \alpha)$ est donc une suite géométrique.
- Découle alors le terme général : $\forall n \in \mathbb{N},\ u_n - \alpha = a^n(u_0 - \alpha)$.

VI.4 Formule de Bayes

Théorème (Formules de Bayes)

1) **Cas de deux évènements**. Soient A et B deux évènements de probabilité non nulle, alors :

$$P_B(A) = \frac{P(A) \times P_A(B)}{P(B)}. \qquad P_B(A)\text{: } \textit{probabilité que } A \textit{ soit la cause de } B$$

2) **Cas de n évènements**. Soit $(A_i)_{i\in I}$ un système quasi-complet d'événéments et B un évènement tel que $P(B) > 0$, alors:

$$\forall j \in I, \qquad P_B(A_j) = \frac{P(A_j)P_{A_j}(B)}{P(B)} = \frac{P(A_j)P_{A_j}(B)}{\displaystyle\sum_{i\in I} P(A_i) \times P_{A_i}(B)}.$$

Exemple On reprend l'exemple du lancer des deux pièces.
En lançant deux fois cette pièce préalablement choisie, on obtient deux piles. Quelle est la probabilité d'avoir choisi la première pièce.

VII Indépendance

Dans toute cette section $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

Intuitivement, deux évènements A et B sont indépendants lorsque la réalisation de l'un n'influe pas sur la réalisation de l'autre, dans ce cas:

$$P_B(A) = P(A) \qquad \text{et} \qquad P_A(B) = P(B)$$

Ce qui donne $P(A \cap B) = P(A)P(B)$. D'où la définition:

Définition (Indépendance de deux évènements)

Soit (Ω, P) un espace probabilisé fini.
Deux évènements A et B sont dit **indépendants** pour la probabilité P si $P(A \cap B) = P(A)P(B)$.

Remarques

1) ⚠ **Attention** ⚠ Ne pas confondre incompatibilité et indépendance. Si A et B sont indépendants et de proba non nulle $P(A \cap B) = P(A)P(B) \neq 0$ donc $A \cap B \neq \emptyset$.
L'incompatibilité est une notion ensembliste, elle ne dépend que des ensembles, des événements ici.
L'indépendance est une notion probabiliste, qui dépend de la probabilité.

2) Si $P(A) = 0$ alors A et B sont indépendants pour tout B.
Car $P(A \cap B) \leqslant P(A) = 0$ donc $P(A \cap B) = 0$.

Définition (Indépendance mutuelle)

Soit $(A_i)_{1 \leqslant i \leqslant n}$ une famille d'évènements. Les événéments sont dit mutuellement indépendants ou indépendants si :

$$\forall J \subset [\![1, n]\!], \quad P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j).$$

Remarques

- Des événements mutuellement indépendants sont aussi 2 à 2 indépendants.
La réciproque est fausse pour $n \geqslant 3$.
Contre-exemple: on lance deux dés parfaits. Soient les événements, A: “premier dé pair”, B: “second dé pair”, C: “la somme est paire”.
- ⚠ **Attention** ⚠ Deux événements peuvent apparaître intuitivement indépendants alors qu'ils ne le sont pas. Par conséquent, ou bien une hypothèse d'indépendance est clairement émise dans l'énoncé, ou bien il faut effectuer le calcul pour justifier l'indépendance.
- Si l'on souhaite calculer la probabilité d'une intersection :
  - si une hypothèse d'indépendance est connue, on utilise la formule ci-dessus
  - sinon on utilise la formule des probabilités composées.

Propriétés

Soit $(A_i)_{1 \leqslant i \leqslant n}$ une famille d'évènements mutuellements indépendants.

1) Toute sous-famille $(A_{i_1}, \ldots, A_{i_k})$ est formée d'évènements mutuellements indépendants.

2) Si $B_i = A_i$ ou $B_i = \overline{A_i}$ alors $(B_i)_{1 \leqslant i \leqslant n}$ est une famille d'évènements mutuellements indépendants.

3)
   - Si C est la réunion ou l'intersection d'événements de $\{A_1, \ldots, A_p\}$ ou de leur contraire.
   - Si D est la réunion ou l'intersection d'événements de $\{A_{p+1}, \ldots, A_n\}$ ou de leur contraire.

   Alors C et D sont indépendants.