

Cours de Probabilités PC* : Chapitres 1 et 2

Lycée Bellevue

31 janvier 2026

Table des matières

1	Dénombrabilité et Sommabilité	2
1.1	Ensembles dénombrables	2
1.1.1	Equipotence	2
1.1.2	Ensembles dénombrables	2
1.1.3	Opérations sur les ensembles dénombrables	2
1.1.4	Enumeration	2
1.2	Familles sommables	3
1.2.1	Permutations	3
1.2.2	Convergence commutative	3
1.2.3	Familles sommables de réels positifs	3
1.2.4	Sommabilité d'une famille au plus dénombrables <u>de nombres complexes</u>	4
1.2.5	Premières propriétés des familles sommables	4
1.2.6	Sommation par paquets	5
1.2.7	Sommes doubles et principe de Fubini	5
1.2.8	Produit de somme	6
2	Probabilités I	7
2.1	Introduction	7
2.2	Espaces probabilisés	7
2.2.1	Espaces probabilisables	7
2.2.2	Probabilité sur un espace probabilisable	8
2.2.3	Propriétés d'une probabilité	9
2.2.4	Vocabulaire probabiliste 2	9
2.3	Probabilité conditionnelle	10
2.3.1	Présentation du concept	10
2.3.2	Formule des probabilités composées	10
2.3.3	Formule des probabilités totales	11
2.3.4	Formule de Bayes ou probabilité des causes	11
2.4	Indépendance	11
2.5	Exemples d'espaces probabilisés usuels	12
2.5.1	Ω est fini	12
2.5.2	Ω est dénombrable	12
2.6	Appendice	13
2.6.1	Rappels ensemblistes	13
2.6.2	Who is Who	13
3	Variables aléatoires discrètes	14
3.1	Généralités	14
3.2	Indépendance	14
3.3	Espérance mathématique	14
3.4	V.a.d à valeurs dans \mathbb{N}	14
3.4.1	Fonction génératrice d'une variable aléatoire à valeurs dans \mathbb{N}	14

Chapitre 1

Dénombrabilité et Sommabilité

1.1 Ensembles dénombrables

1.1.1 Equipotence

Définition 1 Deux ensembles sont équipotents s'il existe une bijection de l'un sur l'autre.

Exemple 1 \mathbb{R} et $]-\frac{\pi}{2}, \frac{\pi}{2}[$ sont équipotents via \arctan .

Remarque 1 Sur la classe des ensembles, la relation d'équipotence est une relation d'équivalence

1.1.2 Ensembles dénombrables

Définition 2 Un ensemble est dénombrable s'il est équipotent à \mathbb{N}

Remarque 2 Un ensemble dénombrable est donc infini.

Définition 3 Un ensemble est au plus dénombrable s'il est fini ou dénombrable.

Remarque 3 Ces deux notions sont invariantes par équipotence.

Exemple 2 \mathbb{N}^* , \mathbb{Z} et \mathbb{N}^2 sont dénombrables.

1.1.3 Opérations sur les ensembles dénombrables

Théorème 1 La réunion d'une famille dénombrable (donc finie a fortiori) d'ensembles dénombrables est dénombrable.

Le produit cartésien d'ensembles dénombrables est dénombrable.

1.1.4 Enumeration

Proposition 1 Soit A un ensemble. Il y a équivalence entre les assertions :

i) A est dénombrable.

ii) Il existe une suite (x_n) telle que :
$$\begin{cases} A = \{x_n, n \in \mathbb{N}\} \\ n \neq m \implies x_n \neq x_m \end{cases}$$

Dans ce cas la suite (x_n) (que l'on peut noter $n \rightarrow x_n$) est une **énumération** de A .

Corollaire 1 Toute partie d'un ensemble dénombrable est au plus dénombrable.

Toute partie infinie d'un ensemble dénombrable est dénombrable.

Toute partie infinie qui s'injecte dans un ensemble dénombrable est dénombrable.

(De cette dernière propriété et de l'exemple 2, on déduit que \mathbb{Q} est dénombrable).

Enfin et à titre culturel :

Corollaire 2 \mathbb{R} n'est pas dénombrable.

1.2 Familles sommables

1.2.1 Permutations

Définition 4 Soit A un ensemble, on appelle permutation de A toute bijection de A sur lui-même.

Remarque 4 Si A est dénombrable dont une énumération est (x_n) alors TOUTE énumération de A relève du type $(x_{\alpha(n)})$, où α est une permutation de \mathbb{N} .

1.2.2 Convergence commutative

Définition 5 Une série convergente à termes complexes $\sum u_n$ est dite **commutativement convergente** si, pour toute permutation de \mathbb{N} σ , $\sum u_{\sigma(n)}$ est encore convergente et si $\sum_{n=0}^{\infty} u_{\sigma(n)} = \sum_{n=0}^{\infty} u_n$.

Autrement dit une série est commutativement convergente si une permutation de l'ordre de ses termes ne modifie ni sa convergence ni la somme de cette série.

Proposition 2 Toute série à termes positifs convergente est commutativement convergente.

De façon plus générale et en reprenant les arguments utilisés dans la démonstration du fait qu'une série ACV est convergente, on prouve (et on admet conformément au programme) que :

Théorème 2 (Fondamental)

Une série absolument convergente (à termes complexes) est commutativement convergente.

Remarque 5 Il s'agit en fait d'une caractérisation des séries commutativement convergentes.

Voici ce qui peut se passer pour une série semi-convergente comme la série harmonique alternée.

On sait que $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \ln(2)$ donc si la série harmonique alternée était commutativement convergente on aurait, en réarrangeant ses termes, $\ln(2) = \underbrace{1 - 1/2}_{-1/4} + \underbrace{1/3 - 1/6}_{-1/8} + \underbrace{1/5 - 1/10}_{-1/12} + \dots$; ce qui donnerait $\ln(2) = \frac{\ln(2)}{2}$!

Désormais les ensembles d'indices manipulés sont dénombrables ou, le contexte faisant foi, au plus dénombrables

1.2.3 Familles sommables de réels positifs

Une famille est une généralisation de la notion de suite en prenant des ensembles d'indices plus généraux que l'ensemble des entiers naturels.

On se donne un ensemble d'indices I dénombrable (le cas fini ne soulevant aucune difficulté) ainsi qu'une famille de réels positifs indexée par I et que l'on notera $(x_i)_{i \in I}$ (et plus légèrement (x_i) s'il n'y a pas d'ambiguïté). Noter donc que : $\forall i \in I, x_i \geq 0$.

Définition 6 La famille $(x_i)_{i \in I}$ est dite **sommable** s'il existe une énumération $n \rightarrow i_n$ de I pour laquelle la série à termes positifs $\sum_{n \geq 0} x_{i_n}$ soit convergente (c'est alors le cas par convergence commutative pour toute énumération de I).

Dans ce cas la somme de cette série convergente est la somme de la famille $(x_i)_{i \in I}$ et se note $\sum_{i \in I} x_i$.

On notera la cohérence de cette définition puisque la somme de la série convergente utilisée ne dépend que de I et non de telle ou telle énumération.

Définition 7 Extension de la définition de la somme d'une famille de réels positifs.

Si la famille de réels positifs $(x_i)_{i \in I}$ n'est pas sommable, on pose $\sum_{i \in I} x_i = +\infty$

Ainsi à toute famille dénombrable (et même au plus dénombrable) de réels positifs on associe sa somme qui est un élément de $[0, +\infty]$.

1.2.4 Sommabilité d'une famille au plus dénombrables de nombres complexes

I est un ensemble au plus dénombrable et, pour tout $i \in I$, $x_i \in \mathbb{C}$. Nous étendons la définition de la sommabilité de la façon suivante :

Définition 8 La famille $(x_i)_{i \in I}$ est dite sommable si I est fini ou, lorsque I est dénombrable, si la famille de réels positifs $(|x_i|)_{i \in I}$ est sommable au sens de la définition 6.

Dans ce cas la somme de cette famille, toujours notée $\sum_{i \in I} x_i$ est la somme usuelle dans le cas fini et, lorsque

I est dénombrable, il s'agit de la somme de la série absolument convergente (donc convergente) $\sum_{n \geq 0} x_{i_n}$, où

$n \rightarrow i_n$ est une énumération de I (la convergence commutative d'une série ACV nous assure bien que ce nombre ne dépend pas de l'énumération utilisée).

Notations 1 Pour signifier que $(x_i)_{i \in I}$ est sommable, on écrira (notation à savoir expliquer le cas échéant) $(x_i)_{i \in I} \in \ell^1(I)$.

Exemple 3 Si $I = \mathbb{N}$, la sommabilité de la famille (ou suite ici) de complexes (x_n) équivaut à la convergence absolue de la série $\sum_{n \geq 0} x_n$ converge et, dans ce cas, sa somme est celle de cette série.

Autrement dit $(x_n)_{n \in \mathbb{N}} \in \ell^1(\mathbb{N}) \Leftrightarrow \sum_{n \geq 0} x_n$ converge absolument. (Cas des familles de réels positifs à inclure).

1.2.5 Premières propriétés des familles sommables

On se donne I, J au plus dénombrables ainsi que $(x_i)_{i \in I}, (y_i)_{i \in I}$ des familles de nombres complexes.

Nous listons maintenant des propriétés des sommes de familles sommables qui généralisent les plus simples des sommes usuelles (i.e finies).

Proposition 3 Si $(x_i)_{i \in I}$ sommable et si $J \subset I$, alors $(x_i)_{i \in J}$ sommable.

Exercice 1 On note I l'ensemble des rationnels supérieurs à 1.

Etablir que la famille $(i^{-2})_{i \in I}$ n'est pas sommable. (Considérer la sous-famille $(x_n = 1 + 1/n)_{n \in \mathbb{N}^*}$ évidemment non sommable)

Proposition 4 (Domination)

Si $(y_i)_{i \in I} \in \ell^1(I)$ et si, pour tout $i \in I$, $|x_i| \leq |y_i|$ alors $(x_i)_{i \in I} \in \ell^1(I)$.

Par définition :

Proposition 5 a) Soient $(x_i)_{i \in I}$ sommable et σ une permutation de I :

la famille $(x_{\sigma(i)})_{i \in I}$ est sommable et $\sum_{i \in I} x_{\sigma(i)} = \sum_{i \in I} x_i$.

b) $(x_i)_{i \in I} \in \ell^1(I) \iff (|x_i|)_{i \in I} \in \ell^1(I)$.

Proposition 6 (Linéarité de la somme)

Si $(x_i)_{i \in I}$ et $(y_i)_{i \in I}$ sont dans $\ell^1(I)$ alors, pour tout $\lambda \in \mathbb{C}$, $(x_i + \lambda y_i)_{i \in I} \in \ell^1(I)$ et

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

Corollaire 3 $(x_i)_{i \in I} \in \ell^1(I) \iff (Re(x_i))_{i \in I} \in \ell^1(I)$ et $(Im(x_i))_{i \in I} \in \ell^1(I)$

Proposition 7 (Croissance de la somme)

Si $(x_i)_{i \in I}$ et $(y_i)_{i \in I}$ sont des familles de **réels** sommables et si , pour tout $i \in I$, $x_i \leq y_i$, alors $\sum_{i \in I} x_i \leq \sum_{i \in I} y_i$.

Si les familles sont à termes positifs, qu'il y ait sommabilité ou pas, l'inégalité :

$\sum_{i \in I} x_i \leq \sum_{i \in I} y_i$ reste valable et peut prouver la sommabilité ou la non sommabilité d'une famille.

(La finitude de la somme majorante donne la sommabilité de la famille $(x_i)_{i \in I}$ et par contraposition....)

Corollaire 4 (Inégalité triangulaire)

Qu'il y ait sommabilité ou pas : $\sum_{i \in I} |x_i + y_i| \leq \sum_{i \in I} |x_i| + \sum_{i \in I} |y_i|$.

1.2.6 Sommation par paquets

Cette technique généralise la relation de Chasles.

Définition 9 Soient I, J ensembles au plus dénombrables, on appelle découpage (ou partage) de I toute famille $(I_j)_{j \in J}$ de parties de I telle que :

i) Les I_j sont deux à deux disjoints.

ii) $\bigcup_{j \in J} I_j = I$.

Voilà maintenant le résultat principal de cette sous-section.

Théorème 3 (Sommation par paquets)

Soient $(I_j)_{j \in J}$ un découpage de I ensemble au plus dénombrable et $(x_i)_{i \in I}$ une famille de nombres complexes **sommable**.

i) Pour chaque $j \in J$, $(x_i)_{i \in I_j}$ est sommable , en posant $S_j = \sum_{i \in I_j} x_i$, la famille $(S_j)_j$ est sommable.

ii) $\sum_{i \in I} x_i = \sum_{j \in J} S_j = \sum_{j \in J} (\sum_{i \in I_j} x_i)$.

Noter enfin que si $(x_i)_{i \in I}$ est une famille de réels positifs, l'égalité précédente est toujours vraie (qu'il y ait ou non sommabilité) et en cas de finitude du membre de droite, elle donne la sommabilité de la famille $(x_i)_{i \in I}$.

Un découpage fini et naturel de \mathbb{Z} étant $(\mathbb{N}, \mathbb{Z}_-^*)$, on a (puisque, on le rappelle, la sommabilité d'une famille indécouper par \mathbb{N} se résume à sa convergence absolue) :

Corollaire 5 (Cas d'une famille indécouper par \mathbb{Z})

Soit $(x_n)_{n \in \mathbb{Z}}$ une famille de nombres complexes.

Les assertions suivantes sont équivalentes :

i) $(x_n)_{n \in \mathbb{Z}}$ est sommable.

ii) Les séries $\sum_{n \geq 0} x_n$ et $\sum_{n \geq 1} x_{-n}$ sont absolument convergentes.

Dans ce cadre $\sum_{n \in \mathbb{Z}} x_n = \sum_{n=0}^{\infty} x_n + \sum_{n=1}^{\infty} x_{-n}$.

Exercice 2 Soit q un nombre complexe de module strictement inférieur à 1.

Etablir que la famille $(q^{|n|})_{n \in \mathbb{Z}}$ est sommable et préciser sa somme (on doit trouver $\frac{1+q}{1-q}$).

1.2.7 Sommes doubles et principe de Fubini

I et J sont des ensembles au plus dénombrables et $(x_{i,j})_{(i,j) \in I \times J}$ désigne une famille de nombres complexes. La technique qui suit est une conséquence de la sommation par paquets.

Théorème 4 (Principe de Fubini)

a) On suppose que la famille $(x_{i,j})_{(i,j) \in I \times J}$ est sommable : alors

i) Pour tout $i \in I$, la famille $(x_{i,j})_{j \in J}$ est sommable, de somme S_i et la famille $(S_i)_{i \in I}$ est sommable.

ii) Pour tout $j \in J$, la famille $(x_{i,j})_{i \in I}$ est sommable, de somme T_j et la famille $(T_j)_{j \in J}$ est sommable.

Dans ce cas (Fubini) on a : $\sum_{(i,j) \in I \times J} x_{i,j} = \sum_{i \in I} (\sum_{j \in J} x_{i,j}) = \sum_{j \in J} (\sum_{i \in I} x_{i,j})$.

Si $(x_{i,j})_{j \in J}$ est une famille de réels positifs, les égalités précédentes sont vraies qu'il y ait sommabilité ou non et, en cas de finitude ont valeur de preuve de sommabilité.

Il vient immédiatement (cf le cas famille à termes positifs et proposition 5 b)) que :

Corollaire 6 Si Pour tout $i \in I$, la famille $(x_{i,j})_{j \in J}$ est sommable et si la famille $(\sum_{j \in J} |x_{i,j}|)$ l'est aussi alors la famille $(x_{i,j})_{(i,j) \in I \times J}$ est sommable et on peut appliquer le principe de Fubini pour en calculer la somme (Même chose en échangeant les rôles de I et J)

On retiendra

donc que par ce biais on peut prouver la sommabilité d'une famille double et qu'en niant (cf TD19 Centrale 2023) ce principe (par exemple la permutation des sommations ne donne pas le même résultat) on récupère la non sommabilité■

A nouveau parce que sommabilité équivaut à convergence absolue pour des familles indexées par les entiers naturels, il vient :

Corollaire 7 (Cas d'une famille indexée par \mathbb{N}^2)

Si $(x_{n,m})_{(n,m) \in \mathbb{N}^2}$ est sommable :

i) Pour tout $n \in \mathbb{N}$, la série $\sum_{m \geq 0} x_{n,m}$ est absolument convergente, de somme S_n et la série $\sum_{n \geq 0} S_n$ est absolument convergente.

ii) Pour tout $m \in \mathbb{N}$, la série $\sum_{n \geq 0} x_{n,m}$ est absolument convergente, de somme T_m et la série $\sum_{m \geq 0} T_m$ est absolument convergente.

Dans ce cas on a :
$$\sum_{(n,m) \in \mathbb{N}^2} x_{n,m} = \sum_{n=0}^{\infty} \left(\sum_{m=0}^{\infty} x_{n,m} \right) = \sum_{m=0}^{\infty} \left(\sum_{n=0}^{\infty} x_{n,m} \right)$$

1.2.8 Produit de somme

Commençons par une conséquence de Fubini (preuve détaillée en cours et à comprendre).

Proposition 8 Soient $(x_i)_{i \in I}$ et $(y_j)_{j \in J}$ deux familles sommables de nombres complexes :

a) $(x_i y_j)_{(i,j) \in I \times J} \in \ell^1(I \times J)$. b) Et
$$\sum_{(i,j) \in I \times J} x_i y_j = \left(\sum_{i \in I} x_i \right) \left(\sum_{j \in J} y_j \right).$$

Et nous pouvons désormais prouver (belle utilisation de la sommation par paquets et je vous conseille de saisir les arguments de cette démonstration, produite en classe) :

Corollaire 8 (Retour du produit de Cauchy)

Soient $\sum_{n \geq 0} a_n$, $\sum_{n \geq 0} b_n$ deux séries absolument convergentes et $\sum_{n \geq 0} c_n$ leur produit de Cauchy.

i) La série $\sum_{n \geq 0} c_n$ est absolument convergente.

ii)
$$\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n \right) \left(\sum_{n=0}^{\infty} b_n \right).$$

Chapitre 2

Probabilités I

2.1 Introduction

Les probabilités constituent désormais¹ une discipline mathématique à part entière comme la géométrie ou l'algèbre linéaire mais il fallut attendre l'approche de Kolmogorov (circa 1930) pour y parvenir. En tant que telle, elles se structurent logiquement autour d'une nervure axiomatique moins intuitive que la vision fréquentiste, utilisée au lycée pour leur introduction. C'est, malgré cela, le prix à payer pour obtenir un socle sain permettant d'engendrer des développements et applications plus substantiels.

Expliquons en quelques mots cette approche.

Une expérience aléatoire consiste en l'observation d'un phénomène théoriquement reproductible dans des conditions analogues, dont les résultats dépendent du hasard (donc non prédictibles). A ce titre un lancer de dés est une telle expérience alors que le recensement des résultats du classico Real-Barcelone n'en constitue pas une.

A une expérience aléatoire donnée on associe l'ensemble des résultats possibles de celle-ci ; il s'agit de l'univers de cette expérience. On le note traditionnellement Ω . En général (sauf éventuellement dans le cas fini) cet univers nous sera donné si son utilisation s'avère nécessaire (dans les cas concrets sa détermination relève plutôt de la statistique).

L'idée géniale de Kolmogorov consiste à définir un événement comme une partie (i.e sous-ensemble) choisie de l'univers et une probabilité comme une application, régie par des contraintes issues de l'approche fréquentiste, de l'ensemble des événements dans $[0, 1]$. Il s'affranchit ainsi de toute subjectivité, source des paradoxes en probabilité, en ne s'intéressant plus aux causes des événements.

Dans tout ce qui suit, l'univers Ω est donné.

2.2 Espaces probabilisés

2.2.1 Espaces probabilisables

On jette une pièce (non truquée) n fois de suite. Le modèle fini (i.e univers fini) de première année permet de considérer l'événement F_i : " Obtenir face au i -ième lancer", ce pour $1 \leq i \leq n$ et de déterminer la probabilité de l'événement (noté A_n) $\bigcup_{i=1}^n F_i$ (i.e "obtenir toujours face sur les n lancers"). En effet on trouve

aisément que cette probabilité vaut $\frac{1}{2^n}$.

En revanche si on décide de lancer indéfiniment cette même pièce et si on désire connaître la probabilité d'obtenir systématiquement face, ce modèle est insuffisant puisque $\bigcup_{i=1}^{\infty} F_i$ ne le catalogue pas comme événement. Tout ce que l'on peut proposer c'est un passage à la limite dans la probabilité précédente mais en quoi est-ce une probabilité ?

Il va falloir revenir sur cette notion, c'est l'objet du paragraphe suivant.

1. Poincaré doutait qu'on pût jamais les théoriser. Pour les noms propres vous pouvez vous reporter à l'appendice.

Définitions

Définition 10 On appelle **tribu** sur Ω toute partie \mathcal{A} de $\mathcal{P}(\Omega)$ telle que :

- i) $\Omega \in \mathcal{A}$.
- ii) $\forall A \in \mathcal{A}, \bar{A} \in \mathcal{A}$ (stabilité par passage au complémentaire).
- iii) Pour toute suite $(A_n)_{n \geq 0}$ d'éléments de \mathcal{A} , $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$ (stabilité par union dénombrable).

Dans ce cas, le couple (Ω, \mathcal{A}) est un **espace probabilisable** (lié à l'aléa étudié).

Remarque 6 Il s'agit de la première étape de la modélisation ; la tribu représente les parties de Ω dont on va mesurer les chances de réalisation.

Propriétés des tribus

Voici qui résume celles-ci :

Proposition 9 Soit \mathcal{A} une tribu sur Ω , alors : i) $\emptyset \in \mathcal{A}$.

ii) Pour toute suite $(A_n)_{n \geq 0}$ d'éléments de \mathcal{A} , $\bigcap_{n \in \mathbb{N}} A_n \in \mathcal{A}$ (stabilité par intersection dénombrable).

iii) \mathcal{A} est stable par réunion et intersection finies.

iiii) Si $(A, B) \in \mathcal{A}^2$, $A \setminus B \in \mathcal{A}$.

Corollaire 9 Soient I un ensemble au plus dénombrable et $(A_i)_{i \in I}$ une famille d'éléments de \mathcal{A} alors :

$\bigcup_{i \in I} A_i \in \mathcal{A}$ et $\bigcap_{i \in I} A_i \in \mathcal{A}$

Vocabulaire probabiliste 1

Définition 11 Soit (Ω, \mathcal{A}) un espace probabilisable.

Les éléments de \mathcal{A} sont les **événements** (ou **observables**) de l'espace probabilisable.

Ω est l'événement certain.

\emptyset est l'événement impossible.

Deux événements A et B sont dits **incompatibles** s'ils sont disjoints.

Si A est un événement, \bar{A} désigne son événement contraire.

Si A et B sont des événements $A \cup B$ est aussi l'événement " A ou B " et $A \cap B$ est aussi l'événement " A et B ".

Lorsque les singletons sont des événements, ceux-ci sont nommés événements élémentaires.

Exemples

Voici les trois tribus les plus immédiates sur un univers Ω arbitraire :

Exemple 4 i) $\{\emptyset, \Omega\}$ dite tribu grossière.

ii) $\mathcal{P}(\Omega)$ (elle sera, en général, utilisée pour un univers au plus dénombrable) dite tribu pleine.

iii) $T_A = \{\emptyset, \Omega, A, \bar{A}\}$, où A est un événement ; c'est la plus petite tribu, au sens de l'inclusion contenant A , sur Ω . On dit qu'elle est engendrée par A .

2.2.2 Probabilité sur un espace probabilisable

On considère (Ω, \mathcal{A}) un espace probabilisable.

Définition 12 Une probabilité sur (Ω, \mathcal{A}) est une application \mathbb{P} , définie de \mathcal{A} et à valeurs dans \mathbb{R}_+ telle que :

i) $\mathbb{P}(\Omega) = 1$.

ii) Pour toute suite $(A_n)_{n \geq 0}$ d'éléments de \mathcal{A} , deux à deux incompatibles :

la série $\sum \mathbb{P}(A_n)$ converge et $\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n=0}^{\infty} \mathbb{P}(A_n)$.

(Cette propriété se nomme σ -additivité).

Définition 13 Si \mathbb{P} est une probabilité sur (Ω, \mathcal{A}) , le triplet $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace probabilisé (seconde phase de la modélisation).

Remarque 7 La σ -additivité se généralise comme suit pour une probabilité : si I est un ensemble au plus dénombrable et si $(A_i)_{i \in I}$ est une famille d'événements deux à deux incompatibles on a aussi :

$$\text{la famille } (\mathbb{P}(A_i))_{i \in I} \text{ est sommable et } \sum_{i \in I} \mathbb{P}(A_i) = \mathbb{P}\left(\bigcup_{i \in I} A_i\right)$$

Ainsi la σ -additivité entraîne-t-elle la notion d'additivité de la version première année d'une probabilité, il suffit de considérer I fini.

2.2.3 Propriétés d'une probabilité

On se place dans l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

On retrouve les propriétés du cas fini pour commencer.

Proposition 10 Soient A, B des événements :

o) \mathbb{P} est à valeurs dans $[0, 1]$.

i) $\mathbb{P}(\emptyset) = 0$.

ii) $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$.

iii) $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$ et $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$.

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

Voilà enfin des propriétés nouvelles :

Théorème 5 (Continuité monotone)

i) Soit (A_n) une suite croissante (pour l'inclusion) d'événements.

La suite $(\mathbb{P}(A_n))$ est convergente et $\mathbb{P}(A_n) \xrightarrow{n \rightarrow \infty} \mathbb{P}\left(\bigcup_{k \in \mathbb{N}} A_k\right)$.

ii) Soit (A_n) une suite décroissante (pour l'inclusion) d'événements.

La suite $(\mathbb{P}(A_n))$ est convergente et $\mathbb{P}(A_n) \xrightarrow{n \rightarrow \infty} \mathbb{P}\left(\bigcap_{k \in \mathbb{N}} A_k\right)$.

Une conséquence directe de ce théorème est :

Corollaire 10 Soit (A_n) une suite d'événements.

i) $\mathbb{P}\left(\bigcup_{i=0}^n A_i\right) \xrightarrow{n \rightarrow \infty} \mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right)$.

ii) $\mathbb{P}\left(\bigcap_{i=0}^n A_i\right) \xrightarrow{n \rightarrow \infty} \mathbb{P}\left(\bigcap_{i \in \mathbb{N}} A_i\right)$.

En utilisant la sous-additivité finie, la conservation des inégalités à la limite et la continuité croissante, on prouve :

Théorème 6 (Sous-additivité ou inégalité de Boole)

Soit (A_n) une suite d'événements, alors, en convenant que $\sum_{n=0}^{\infty} \mathbb{P}(A_n) = +\infty$ si la série $\sum \mathbb{P}(A_n)$ diverge,

on a : $\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n=0}^{\infty} \mathbb{P}(A_n)$.

2.2.4 Vocabulaire probabiliste 2

On travaille dans un espace probabilisé.

Définition 14 On dit qu'un événement est négligeable si sa probabilité est nulle.

Un événement est presque sûr si sa probabilité vaut 1.

On a sans peine que (Boole pour i) et passage à l'événement contraire pour ii) :

Proposition 11 i) Toute réunion au plus dénombrable d'événements négligeables l'est également.
 ii) L'intersection au plus dénombrable d'événements presque sûrs l'est encore.

En guise de conclusion :

Exemple 5 On reconsidère l'expérience aléatoire : on jette une pièce équilibrée une infinité de fois ; on admet qu'il existe un espace probabilisé (que l'on ne précisera pas) permettant de modéliser cet aléa et on considère à nouveau les événements (on admet qu'ils en sont dans ce cadre) F_n : " les n premiers lancers donnent face ", ce pour $n \geq 1$. On rappelle que $\mathbb{P}(F_n) = 2^{-n}$.
 En utilisant la continuité décroissante, on obtient alors que $\mathbb{P}(\bigcap_{n \in \mathbb{N}^*} F_n) = 0$. Autrement dit, la probabilité d'obtenir systématiquement face est nulle soit que cet événement est négligeable, ce qui n'a rien de surprenant mais qui souligne une nuance entre impossible et négligeable.

2.3 Probabilité conditionnelle

2.3.1 Présentation du concept

Cette notion typiquement probabiliste se fonde sur le fait qu'une information supplémentaire peut modifier la vraisemblance que l'on accorde à l'événement étudié. On en prend pleinement conscience avec l'exemple très simple d'un lancer de deux dés, au sujet duquel on s'intéresse à l'événement : " obtenir une somme ≥ 10 ". Sans plus d'information, la probabilité (simple examen des possibilités (6, 6), (6, 5), (5, 6), (6, 4), (4, 6) et (5, 5) de cet événement est $\frac{1}{6}$ si on sait qu'un des jets a donné 6, cette probabilité devient $\frac{5}{11}$.

Proposition 12 Soit B un événement non négligeable, l'application définie par :

$$A \rightarrow \frac{P(A \cap B)}{P(B)}$$

est une probabilité sur (Ω, \mathcal{A}) . On la note \mathbb{P}_B ou $\mathbb{P}(\cdot | B)$.

Définition 15 Dans le contexte précédent, $\mathbb{P}_B(A) = \mathbb{P}(A | B)$ désigne, A étant un événement, la probabilité (conditionnelle) de A sachant B (réalisé).
 Si $B = \Omega$, $\mathbb{P}_B(A) = P(A)$

Exemple 6 (Classique) On admet qu'il y a équiprobabilité d'avoir pour enfant une fille ou un garçon. Quelle est la probabilité pour qu'un couple, ayant deux enfants, ait deux filles sachant tout d'abord que l'aîné est une fille, puis qu'il a au moins une fille ? (Réponse : $1/2$ et $1/3$)

2.3.2 Formule des probabilités composées

Soient A_1, \dots, A_n des événements tels que $\bigcap_{i=1}^n A_i$ soit non négligeable.

On prouve par une simple récurrence que :

Proposition 13 $\mathbb{P}(\bigcap_{i=1}^n A_i) = \mathbb{P}(A_1)\mathbb{P}(A_2 | A_1) \times \dots \times \mathbb{P}(A_n | \bigcap_{i=1}^{n-1} A_i) = \prod_{i=1}^n \mathbb{P}(A_i | \bigcap_{j=1}^{i-1} A_j)$ (en convenant que l'intersection d'un nombre nul de parties de Ω est Ω).

Remarque 8 Si la famille A_1, \dots, A_n est décroissante pour l'inclusion, cette formule devient :

$\mathbb{P}(A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2 | A_1) \times \dots \times \mathbb{P}(A_n | A_{n-1})$ puisque $\bigcap_{j=1}^{i-1} A_j = A_{j-1}$ si $j \geq 2$.

2.3.3 Formule des probabilités totales

La finitude des sommes de séries à termes positifs dans les lignes qui suivent justifie leur existence.

Définition 16 Une suite $(A_n)_n$ d'événements forme un système complet (resp. quasi-complet) d'événements si les A_n sont deux à deux incompatibles et si $\bigcup_{n \in \mathbb{N}} A_n = \Omega$ (resp. si $\bigcup_{n \in \mathbb{N}} A_n$ est presque sûr).

Noter que, dans ce cas, $\sum_{n=0}^{\infty} \mathbb{P}(A_n) = 1$.

Remarque 9 Cette notion généralise cette vue en première année où les systèmes complet ou quasi complets étaient finis. Montrons en effet que c'est un cas très particulier de la définition donnée ici. Soient $p \in \mathbb{N}$ et (A_0, \dots, A_p) un système complet d'événements (au sens de la première année); en posant $A_n = \emptyset$ pour tout $n > p$, on dispose en $(A_n)_{n \in \mathbb{N}}$ d'un système complet d'événements (au sens de la définition précédente).

Il résulte de cette remarque que le résultat qui suit contient la formule des probabilités totales de première année (somme finie).

Théorème 7 Soient $(A_n)_n$ un système quasi-complet d'événements et B un événement alors :

$$\mathbb{P}(B) = \sum_{n=0}^{\infty} \mathbb{P}(B \cap A_n) = \sum_{n=0}^{\infty} \mathbb{P}(B | A_n) \mathbb{P}(A_n), \text{ en convenant que } \mathbb{P}(B | A_n) \mathbb{P}(A_n) = 0 \text{ si } A_n \text{ est négligeable.}$$

2.3.4 Formule de Bayes ou probabilité des causes

On se donne A, B deux événements non négligeables, dont on connaît les probabilités. Nous disposons aussi de probabilité conditionnelle $\mathbb{P}(A | B)$; la question posée par Bayes se résume, pour l'essentiel, à obtenir $\mathbb{P}(B | A)$.

Par définition même, on a $\mathbb{P}(B | A) = \frac{\mathbb{P}(A | B) \mathbb{P}(B)}{\mathbb{P}(A)}$ (*) puis, avec la formule des probabilités totales (*)

devient $((B, \bar{B})$ étant un système complet d'événements) $\mathbb{P}(B | A) = \frac{\mathbb{P}(A | B) \mathbb{P}(B)}{\mathbb{P}(A | B) \mathbb{P}(B) + \mathbb{P}(A | \bar{B}) \mathbb{P}(\bar{B})}$ (**).

De façon plus générale :

Théorème 8 Si (B_n) est un système complet d'événements non négligeables et si A est un événement non négligeable, alors :

$$\mathbb{P}(B_i | A) = \frac{\mathbb{P}(A | B_i) \mathbb{P}(B_i)}{\sum_{n=0}^{\infty} \mathbb{P}(A | B_n) \mathbb{P}(B_n)}.$$

Cette formule nous sera très peu utile (elle a plutôt un intérêt statistique et sert plutôt en biologie ou dans les sciences humaines) et sa démonstration est immédiate à partir de (*) et des probabilités totales. Donnons néanmoins un exemple dans lequel elle s'avère efficace.

Exemple 7 Fiabilité test ?

Un individu est pris au hasard dans une population au sein de laquelle la proportion de personnes infectées par un virus (noté V) est de 10^{-4} . On lui fait alors passer un test de détection relatif à V ; on sait par ailleurs que les probabilités d'avoir un résultat positif à ce test sont de 0,99 si l'individu est infecté, de 0,001 sinon. Sachant que le test est positif, quelle est la probabilité pour que l'individu soit réellement infecté ?

Notons B l'événement : " l'individu est infecté " et A : " le test est positif " .

Nous voulons déterminer $\mathbb{P}(B|A)$ alors que nous connaissons $\mathbb{P}(B) = 10^{-4}$ ainsi que $\mathbb{P}(A|B) = 0,99$ et $\mathbb{P}(A|\bar{B}) = 0,001$.

La formule (**) donne $\mathbb{P}(B|A) = \frac{0,99 \times 10^{-4}}{0,99 \times 10^{-4} + 0,001 \times (1 - 10^{-4})} = 0,09$.

2.4 Indépendance

Là encore, une notion purement probabiliste. Intuitivement, deux événements sont indépendants si le fait de savoir que l'un d'entre eux est réalisé n'apporte pas d'information au sujet de la réalisation de l'autre événement.

Définition 17 A, B deux événements sont indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$.

Proposition 14 Soient A, B deux événements non négligeables ; ils sont indépendants ssi :
 $\mathbb{P}(A \mid B) = \mathbb{P}(A) \Leftrightarrow \mathbb{P}(B \mid A) = \mathbb{P}(B)$.

Définition 18 Soit $(A_i)_{i \in I}$ une famille au plus dénombrable d'événements est dite (mutuellement)indépendante si , pour toute partie finie J de I , on a : $\mathbb{P}(\bigcap_{j \in J} A_j) = \prod_{j \in J} \mathbb{P}(A_j)$.

Remarque 10 L'indépendance mutuelle entraîne la deux à deux indépendance. En revanche la réciproque est fausse (cf contre-exemple de Bernstein TD 20).

Proposition 15 Si A_1, \dots, A_n sont des événements mutuellement indépendants alors $A_1, \dots, \overline{A_n}$ le sont aussi, donc $\overline{A_1}, \dots, \overline{A_n}$ forment une famille d'événements mutuellement indépendants.

2.5 Exemples d'espaces probabilisés usuels

On se donne un univers Ω au plus dénombrable.

Nous nous plaçons dans l'espace probabilisable $\mathcal{P}(\Omega)$

Dans ce contexte tout singleton est événement (élémentaire) et tout événement en est réunion au plus dénombrable, ce qui permettra (via la sigma-additivité ou l'additivité) dans un espace probabilisé associé à l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$ de définir automatiquement la probabilité d'un événement quelconque à partir de celles des événements élémentaires. Nous allons préciser davantage en revenant en premier lieu au contexte de la première année.

2.5.1 Ω est fini

Soient $n \in \mathbb{N}^*$ et $\Omega = \{\omega_1, \dots, \omega_n\}$.

Il doit être bien connu que :

Proposition 16 \mathbb{P} est une probabilité sur $\mathcal{P}(\Omega)$.

Alors, en posant pour tout $i \in \llbracket 1, n \rrbracket$, $\mathbb{P}(\{\omega_i\}) = p_i$:

$$a) \sum_{i=1}^n p_i = 1.$$

b) Soit $A \subset \Omega$. On pose $A = \{\omega_j, j \in J \subset \llbracket 1, n \rrbracket\}$.

$$\text{On a } \mathbb{P}(A) = \sum_{j \in J} p_j.$$

Inversement :

Proposition 17 Si $(q_i)_{i \in \llbracket 1, n \rrbracket}$ est un élément de \mathbb{R}_+^n tel $\sum_{i=1}^n q_i = 1$, il existe une probabilité \mathbb{P} sur $\mathcal{P}(\Omega)$ telle

que $\mathbb{P}(\{\omega_i\}) = q_i$, ce pour tout $i \in \llbracket 1, n \rrbracket$.

Dès lors si $A = \{\omega_j, j \in J \subset \llbracket 1, n \rrbracket\}$, $\mathbb{P}(A) = \sum_{j \in J} q_j$.

2.5.2 Ω est dénombrable

Soit I un ensemble dénombrable et $i \rightarrow \omega_i$ une énumération de Ω .

Nous admettons la généralisation des deux propositions précédentes.

Théorème 9 (Caractérisation d'une probabilité sur un univers dénombrable via les germes de probabilité)

1) Si \mathbb{P} est une probabilité sur $\mathcal{P}(\Omega)$.

Alors, $\forall i \in I, \mathbb{P}(\{\omega_i\}) = p_i :$

a) la famille $(p_i)_{i \in I}$ est sommable, de somme 1,

b) pour tout événement $A = \{\omega_j, j \in J \subset I, \mathbb{P}(A) = \sum_{j \in J} p_j$.

2) Inversement si $(q_i)_{i \in I}$ est une famille de réels positifs, sommable et de somme 1, il existe une probabilité sur $\mathcal{P}(\Omega)$ telle que :

$\forall i \in I, \mathbb{P}(\{\omega_i\}) = q_i.$

Ces relations déterminant \mathbb{P} via le 1)b.

Dans ce cadre l'équibrosabilité s'avère impossible.

2.6 Appendice

2.6.1 Rappels ensemblistes

On se donne deux ensembles I (celui des indices, pour nous, au plus dénombrable) et E (quelconque) ainsi qu'une famille $(B_i)_{i \in I}$ de parties de E

Loi de de Morgan

Proposition 18 Dans ce contexte :

a) $\overline{\bigcup_{i \in I} B_i} = \bigcap_{i \in I} \overline{B_i}$

b) $\overline{\bigcap_{i \in I} B_i} = \bigcup_{i \in I} \overline{B_i}$

Distributivités

A est une partie de E .

Proposition 19 Dans ce contexte :

a) $A \cap \left(\bigcup_{i \in I} B_i \right) = \bigcup_{i \in I} (A \cap B_i)$ (distributivité de \cap sur \bigcup)

b) $A \cup \left(\bigcap_{i \in I} B_i \right) = \bigcap_{i \in I} (A \cup B_i)$ (distributivité de \cup sur \cap)

2.6.2 Who is Who

Mathématicien 1 Henri Poincaré (1854-1912) fut aussi phycisien, sa distraction était légendaire.

Mathématicien 2 Andrei Kolmogorov (1903-1987) mathématicien russe et père des probabilités modernes.

Mathématicien 3 Thomas Bayes (1702-1761) clergyman anglais et mathématicien à ses heures perdues, contemporain du philosophe David Hume.

Mathématicien 4 Serge Bernstein (1880-1968) mathématicien russe (en fait ukrainien), ancien élève de Supelec.

Chapitre 3

Variables aléatoires discrètes

3.1 Généralités

3.2 Indépendance

3.3 Espérance mathématique

3.4 V.a.d à valeurs dans \mathbb{N}

3.4.1 Fonction génératrice d'une variable aléatoire à valeurs dans \mathbb{N}

Dans tout ce paragraphe les variables aléatoires sont à valeurs dans \mathbb{N}

On remarquera que, sans perte de généralité et pour simplifier l'exposition, on peut prendre \mathbb{N} pour ensemble des valeurs prises par une telle v.a ; il suffit de décréter que la probabilité sur les valeurs non touchées est nulle.

Désormais X, Y, \dots sont des v.a.d dont l'image est \mathbb{N} .

Définition 19 (Série et fonction génératrices de X)

La série entière $\sum_{n \geq 0} \mathbb{P}(X = n)t^n$ est la série génératrice de X .

La fonction génératrice de X est la somme de sa série génératrice. On la note G_X .

Par définition d'une loi de probabilité, on sait que la série $\sum_{n \geq 0} \mathbb{P}(X = n)$ converge. D'où :

Proposition 20 i) La série génératrice de X possède un rayon de convergence $R_X \geq 1$.

ii) G_X est définie au moins sur $[-1, 1]$.

iii) De façon plus précise (cf cours sur les séries entières) G_X est continue sur $[-1, 1]$ au moins et de classe C^∞ sur $] -1, 1[$ au moins.

les fonctions génératrices des lois usuelles sont recensées dans le document les concernant¹

On dispose des propriétés suivantes :

Proposition 21 Soient X, Y deux variables aléatoires à valeurs dans \mathbb{N} :

1) $\forall t \in]-R_X, R_X[$, $G_X(t) = E(t^X)$.

2) $X \hookrightarrow Y \iff G_X = G_Y$. (Autrement dit, la fonction génératrice caractérise la loi de probabilité de X)

Preuve 1 1) Cf formule du transfert avec $f : x \in \mathbb{N} \rightarrow t^x$, t fixé.

2) La condition est évidemment nécessaire. Inversement si les fonctions génératrices coïncident, on a alors $\mathbb{P}(X = n) = \mathbb{P}(Y = n)$ pour tout n , ce par unicité des coefficients d'une série entière ; à partir de là les lois de probabilités respectives de X et Y sont égales soit $X \hookrightarrow Y$.

Nous allons caractériser le fait que X possède une espérance mathématique finie via G_X .

Proposition 22 $X \in L^1$ si et seulement si G_X est dérivable en 1 et, dans ce cas, $E(X) = G'_X(1)$.

La preuve sera donnée en classe

Si $R_X > 1$, ce qui prévaut pour les v.a suivant les lois usuelles, X admet une espérance mathématique