CAPES de Mathématiques **Université Joseph Fourier**
Préparation à l'écrit **Année 2007-2008**
Algèbre et probabilités

Liste des fiches

0 Quelques principes de rédaction mathématique

1 Opérations ensemblistes et dénombrement

1 bis Correction et commentaires sur la fiche 1

1 ter : Supplément à la fiche 1

2 Probabilités (1) Introduction aux espaces probabilisés

3 Probabilités (2) Variables aléatoires discrètes

4 Probabilités (3) Variables aléatoires densitables

5 Probabilités (4) Théorèmes limites

6 Probabilités (5) Applications statistiques

Fiche Probabilités : Corrigés

7 Algèbre (1) Groupes

7 Bis Algèbre (1) Groupes (supplément)

8 Algèbre (2) Groupe symétrique

9 Algèbre (3) Anneaux et corps

10 Algèbre (4) Arithmétique, anneau des entiers relatifs

11 Algèbre (5) Polynômes

12 Algèbre (6) Polynômes, racines, fractions rationnelles

13 Algèbre (7) Espaces vectoriels, dimension, dualité

14 Algèbre (8) Matrices et réduction des endomorphismes

15 Algèbre (9) Déterminants

16 Un problème d'écrit

17 Un problème de CAPES blanc

17 Bis Corrigé

CAPES de Mathématiques **Université Joseph Fourier**
Préparation à l'écrit **Année 2007-2008**
Algèbre et probabilités

Fiche 0 : Quelques principes de rédaction mathématique

« He who can, does. He who cannot, teaches. » George Bernard Shaw

Ce texte s'inspire fortement d'un texte similaire mis à la disposition des étudiants par Christophe Champetier mais les vues exprimées ici ne sont pas de sa responsabilité.

Règle 1 : Définir clairement les objets qu'on utilise

> *Tout caractère (x, t, n, a, f, i, A, F, α, ϕ, $\mathbb{N}$ ou autre) désignant un objet mathématique (élément, ensemble, fonction ou autre) doit impérativement être présenté et clairement défini avant d'être utilisé.*

Ce principe élémentaire est fondamental pour qu'une phrase, en particulier dans un raisonnement mathématique, ait un sens.

Par exemple à tout moment d'une rédaction, si on écrit « $f(x) \geqslant 0$ », ou bien « la suite réelle (a_n) est majorée par C », il faut *auparavant* avoir dit ce que sont f, x, (a_n) et C. Sinon, au mieux le raisonnement n'est pas clair, au pire il n'a pas de sens, et dans les deux cas il risque d'être interprété comme faux.

Autrement dit, on ne parle pas de quelque chose tant qu'on n'a pas dit ce que c'était. Dans un raisonnement, une variable, notée par exemple x, f, (a_n), k ou ϵ, désigne un ensemble ou un élément d'un ensemble, qui a été lui-même précédemment défini, par exemple $\mathbb{N}$, $\mathbb{R}$ ou l'ensemble des suites réelles.

Présentation d'un symbole

Un symbole représente un objet souvent présenté par un quantificateur.

Par exemple on écrira « Pour tout $x \in E$, ... » Si la suite du raisonnement est trop longue, on peut commencer par « Soit $x \in E$ [quelconque]. »

Quelques exemples :
« $\forall \epsilon > 0$, ... » signifie « Pour tout $\epsilon \in \mathbb{R}_+^*$, ... »
« Soit $x \in \mathbb{R}$ tel que $\cos x = \sin x$. »
« Soit $x \in [0, \pi]$ tel que $\cos x = \sin x$. »

Expliquons un point de détail qui va nous permettre de souligner un autre point, crucial quant à lui.

En toute rigueur, « Soit $x \in E$. » sous-entend que l'ensemble E n'est pas vide alors que « Pour tout $x \in E$, ... » est correct que E soit vide ou non. Dans la pratique, on accepte les deux. Par exemple :

« Soit E l'ensembme $E = \{x \in \mathbb{R} \,|\, x^2 + 2x + 2 = 0\}$. Soit $x \in E$. Alors $(x+1)^2 + 1 = 0$ donc $(x+1)^2 = -1$. Or on sait que le carré de tout nombre réel est un nombre réel positif ou nul. Comme -1 est strictement négatif, c'est absurde donc $E = \emptyset$. »

Exercice : Expliquer pourquoi la phrase « Soit $x \in E$. » est indispensable dans la rédaction ci-dessus.

Par conséquent, sauf dans les cas de raisonnement par l'absurde (précisément !), il *faut* rappeler ou démontrer que l'ensemble dans lequel on prend la variable n'est pas vide, sinon le raisonnement est vraisemblablement absurde ou faux.

Exemples de formulation :
« Soit $x \in \mathbb{R}$. » (sous-entendu x quelconque)
« Soit x un réel vérifiant la propriété P. »
« $\forall x \in \mathbb{R},\ \cos(x - (\pi/2)) = \sin x$. »
« Pour tout $x \in \mathbb{R}$, $\cos((\pi/2) - x) = \sin x$. »
« $\exists x \in \mathbb{R},\ \cos x = \sin x$. »
« $\exists x \in [0, \pi],\ \cos x = \sin x$. »
« $\forall x \in [0, \pi],\ ((\cos x = \sin x) \Rightarrow x = \pi/4)$. »
« Pour tout $x \in [0, \pi]$, si $\cos x = \sin x$, alors $x = \pi/4$. »
« $\forall \epsilon > 0,\ \exists x > 0,\ \frac{x}{x+1} > 1 - \epsilon$. »

Sans les quantificateurs qui introduisent x, ces phrases n'auraient pas de sens. Par exemple les phrases
« $(\cos x = \sin x) \Rightarrow x = \pi/4$ »
et
« $(\cos x = \sin x) \Rightarrow x = (\pi/4) + k\pi$ »
n'ont aucun sens si x et k ne sont pas présentés avant d'être utilisés.

Une cause fréquente des erreurs de raisonnement rencontrées dans les copies est le non-respect de ces règles : une variable n'est pas définie (impossible de comprendre dans quel ensemble elle varie) ou bien l'ensemble dans lequel elle varie change au milieu de la preuve (par exemple une constante devient soudain une variable quelconque) ou bien l'ensemble dans lequel elle varie est supposé implicitement non vide sans qu'on ait vérifié si c'était le cas.

Règle 2 : La manipulation des objets mathématiques obéit à des critères précis qui sont imposés par leur définition

Par exemple, un nombre complexe, la somme de deux fonctions réelles, la dérivée d'une fonction réelle, une suite numérique sont des objets qui sont précisément définis. Ainsi une fonction (notée par exemple f) est la donnée d'un ensemble de départ, d'un ensemble d'arrivée et pour chaque élément (noté par exemple x) de l'ensemble de départ d'un

unique élément (noté par exemple $f(x)$) de l'ensemble d'arrivée.

Les objets mathématiques ne sont pas des morceaux de théorèmes, ni des théorèmes que l'on imbrique les uns dans les autres pour faire des preuves, mais bien des objets en eux-mêmes qui ont été précisément décrits. Il faut donner un sens à ces objets, à l'aide d'exemples, de représentations visuelles ou autres. Mais ce sens ne permet pas d'écrire des preuves rigoureuses, une preuve étant un discours plus ou moins formel obéissant lui aussi à des règles précises qui sont celles de la logique mathématique.

Règle 3 : Ne pas hésiter à faire des phrases en français

Il est plus agréable, et souvent plus facile, de lire un raisonnement écrit en français qu'avec des symboles logiques.

Par exemple, la phrase mathématique
« $\exists x \in [0, \pi],\ \cos x = \sin x$ »
ne signifie pas qu'on a fixé $x = \pi/4$, mais seulement qu'il existe un réel dans l'intervalle $[0, \pi]$ dont le cosinus est égal au sinus. *La lettre x n'a plus aucun sens au-delà de la phrase mathématique considérée*, ici « $\exists x \in [0, \pi],\ \cos x = \sin x$. »

Si on veut appeler x un tel réel pour la suite du raisonnement, on dirait en français :
« Il existe un réel dans l'intervalle $[0, \pi]$ dont le cosinus est égal au sinus, soit x un tel réel » (ou alors soit x un réel dans l'intervalle $[0, \pi]$ dont le cosinus est égal au sinus).

On peut dire « Soit x un réel tel que $\cos x = \sin x$ (un tel x existe). On a alors $\tan x = 1$ ».

L'écriture « pseudo-mathématique » :

$$\exists x \in \mathbb{R}\ \ \cos x = \sin x \Rightarrow \tan x = 1$$

n'a aucun sens. On pourrait dire :

$$(\exists x \in \mathbb{R},\ \cos x = \sin x) \Rightarrow (\exists x \in \mathbb{R},\ \tan x = 1),$$

mais c'est très lourd.

Nota : Le français littéraire autorise parfois à placer le quantificateur à la fin de la phrase, par exemple : « $\cos(x - (\pi/2)) = \sin x$ pour tout $x \in \mathbb{R}$ ». Cela peut conduire à des ambiguités, donc c'est à éviter dans une rédaction mathématique.

Voici un exercice très instructif si vous n'êtes pas au clair sur ces questions.

Exercice : Montrer que pour toute fonction $f : \mathbb{R} \to \mathbb{R}$, on a :

$$\forall x \in \mathbb{R},\ \forall y \in \mathbb{R},\ \forall \epsilon > 0,\ \exists \eta > 0,\ (|x - y| < \eta\ \Rightarrow\ |f(x) - f(y)| < \epsilon).$$

Règle 4 : Utiliser correctement les symboles $\Rightarrow$ et $\Leftrightarrow$

Ici, « correctement » signifie en fait qu'on devrait presque toujours se passer de ces deux symboles et les remplacer par les mots *donc*, *ainsi*, *ce qui équivaut à*, ou autre. En effet

l'utilisation de $\Rightarrow$ et $\Leftrightarrow$ ne devrait s'inscrire que dans un cadre très rigoureux de syntaxe logique : ces symboles devraient se trouver entre deux propositions très clairement délimitées, par exemple placées entre parenthèses (voir l'énoncé de l'exercice qui conclut la Règle 3).

Exemple : La phrase $(\forall x \in E,\ f(x) = g(x) \Rightarrow f = g)$ est ambigüe, donc n'a aucun sens sauf convention, car elle pourrait signifier :

$$(\forall x \in E,\ f(x) = g(x)) \Rightarrow f = g,$$

ou bien :

$$\forall x \in E,\ (f(x) = g(x) \Rightarrow f = g).$$

Suivant le contexte, ces deux implications peuvent être vraies ou fausses (en général, la première est vraie et la seconde est fausse), en tout cas leurs significations sont très différentes.

Ainsi l'utilisation des symboles $\Rightarrow$ et $\Leftrightarrow$ dans les copies, en début de ligne et sans aucune référence à *quelle proposition* implique *quelle proposition*, est en général peu claire ou incorrecte.

Exercice : Étudier la véracité des différentes propositions obtenues en mettant des parenthèses aux différents endroits possibles dans la phrase suivante :

$$\forall x \in [0, \pi[,\ \forall n \in \mathbb{N}^*,\ \sin(nx) < \tfrac{1}{2}\ \Rightarrow\ x = 0.$$

La syntaxe logique est très lourde et on lui préfèrera presque toujours une rédaction en français. Exemples :

Si pour tout $x \in E$ on a $f(x) = g(x)$, *alors* $f = g$.

Soit $x \in E$. *Si* $f(x) = g(x)$, *alors* $f = g$.

Soit $x \in [0, \pi[$. *Si* pour tout $n \in \mathbb{N}^*$ on a $\sin(nx) < \frac{1}{2}$, *alors* $x = 0$.

À noter également une nuance entre *donc* et *implique* : la phrase mathématique $(P \Rightarrow Q)$ signifie que si P est vraie, alors Q est vraie, autrement dit, soit P est fausse, soit Q est vraie, ou encore ((non P) ou Q). Elle ne suppose pas a priori que P est vraie. Dans un raisonnement on affirmera souvent : P est vraie, donc Q est vraie, ce qui n'a pas la même signification.

Enfin le symbole $\Leftrightarrow$ est régulièrement utilisé de manière incorrecte : quand on l'utilise, il faut *impérativement* vérifier (mentalement) les deux implications $\Rightarrow$ et $\Leftarrow$ et les justifier si elles ne sont pas triviales toutes les deux.

Règle 5 : Essayer d'être concis

Il faut apprendre, par exemple en travaillant les démonstrations du cours et les exercices des TD, à distinguer le plus clairement possible les arguments essentiels d'une preuve, les idées importantes et nouvelles s'il y en a, dans un contexte d'arguments considérés

comme « évidents » (ou triviaux, immédiats, clairs ou autres) par l'enseignant-correcteur. Pour l'enseignant, dire qu'une affirmation est « évidente » ne signifie pas qu'elle est « intuitive », mais que sa preuve ne nécessite pas d'idée nouvelle et est souvent une simple application des définitions, résultant d'une vérification calculatoire ou d'une méthode classique.

Dans un devoir, surtout en temps limité, il est inutile de recopier un énoncé, un théorème de cours ou une définition. C'est purement et simplement une perte de temps car on peut supposer que le correcteur connaît le cours ou l'énoncé de la question !

Cela dit, le niveau de rédaction d'une copie doit se situer au niveau de compréhension de l'étudiant : il vaut mieux une copie où tout est démontré en plus de lignes que le minimum nécessaire, qu'une copie où il manque des arguments indispensables. Savoir rédiger correctement une preuve s'acquiert en travaillant le cours et en maîtrisant les notions acquises précédemment.

La rédaction d'une preuve mathématique consiste seulement à *convaincre* le lecteur de la justesse du raisonnement amenant à la conclusion. Personne n'écrit de preuve *complète*, c'est-à-dire lisible par un ordinateur à qui on aurait appris les règles de logique mathématique, et il arrive qu'on puisse convaincre beaucoup de monde avec un raisonnement faux. Pour l'étudiant, il s'agit de convaincre un correcteur qu'il a compris un raisonnement, et qu'il pourrait justifier toutes ses affirmations. Cela impose de n'oublier aucun argument. Avec de l'habitude et de l'expérience, on pourra parfois gagner du temps en donnant certains arguments sans autre justification que celle qu'ils sont « évidents » à vérifier.

A contrario, tout argument inutile est nuisible ; citer un certain nombre de résultats du cours qui vous semblent vaguement en rapport avec la question posée mais sans plus, pour « faire bon poids », est un moyen très sûr de persuader le correcteur que vous ne savez pas de quoi vous parlez.

. . . Et comme ce texte déroge déjà amplement à la Règle 5, on s'arrêtera là.

CAPES de Mathématiques **Université Joseph Fourier**
Préparation à l'écrit **Année 2007-2008**
Algèbre et probabilités

Fiche 1 : Opérations ensemblistes et dénombrement

« The ultimate goal of mathematics is to eliminate any need of intelligent thought. » Alfred N. Whitehead

Fonctions et ensembles

1) Soient E et F deux ensembles, f une application de E dans F, A une partie de E, B une partie de F, I un ensemble d'indices, $(A_i)_{i\in I}$ une famille de parties de E et $(B_i)_{i\in I}$ une famille de parties de F. Comparer les ensembles

$$f^{-1}\left(\bigcup_{i\in I} B_i\right) \text{ et } \bigcup_{i\in I} f^{-1}(B_i)\,; \qquad f^{-1}\left(\bigcap_{i\in I} B_i\right) \text{ et } \bigcap_{i\in I} f^{-1}(B_i)\,.$$

Comparer les ensembles

$$f\left(\bigcup_{i\in I} A_i\right) \text{ et } \bigcup_{i\in I} f(A_i); \qquad f\left(\bigcap_{i\in I} A_i\right) \text{ et } \bigcap_{i\in I} f(A_i).$$

Enfin, comparer les ensembles

$$f^{-1}(F\backslash B) \text{ et } E\backslash f^{-1}(B)\,; \qquad f(E\backslash A) \text{ et } F\backslash f(A).$$

2) Soit E et F deux ensembles et f une application de E dans F. Montrer que f est injective si et seulement si $f(X\cap Y) = f(X)\cap f(Y)$ pour toutes parties X et Y de E.

3) Soient f, g et h trois applications d'un ensemble E dans lui même.

a) Montrer que le fait que $f\circ g$ est bijective n'implique pas que f ou g est bijective. Préciser si le fait que $f\circ g$ est bijective implique que f est injective, que f est surjective, que g est injective ou que g est surjective.

b) Montrer que si $f\circ g$ et $h\circ f$ sont bijectives alors f, g et h sont bijectives.

c) Montrer que si E est un ensemble fini, alors f est injective si et seulement si f est surjective si et seulement si f est bijective. Donner des contrexemples quand E est infini.

4) Soit E un ensemble. Montrer qu'il n'existe pas de bijection de E dans $\mathcal{P}(E)$.

Indication : on pourra supposer que f est une bijection de E dans $\mathcal{P}(E)$ et considérer l'ensemble $A = \{x\in E \,|\, x\notin f(x)\}$.

5) Fonctions indicatrices : soit E un ensemble et $\{0,1\}^E$ l'ensemble des applications de E dans l'ensemble $\{0,1\}$. Pour toute partie A de E, on note désormais $\mathbf{1}_A \in \{0,1\}^E$ sa fonction indicatrice, définie pour tout élément x de E par

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{sinon.} \end{cases}$$

Montrer que l'application $A \mapsto \mathbf{1}_A$ est une bijection de $\mathcal{P}(E)$ sur $\{0,1\}^E$.

On identifie désormais chaque fonction $\mathbf{1}_A$ de E dans $\{0,1\}$ à la fonction de E dans $\mathbb{R}$ ou $\mathbb{Z}$ correspondante.

6) Difference symétrique : si A et B sont deux parties d'un ensemble E, leur différence symétrique $A\Delta B$ est définie par

$$A\Delta B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A).$$

a) Montrer que $\mathbf{1}_{A\Delta B} = (\mathbf{1}_A - \mathbf{1}_B)^2 = |\mathbf{1}_A - \mathbf{1}_B|$. En déduire que $\mathbf{1}_{A\Delta B} = \mathbf{1}_A + \mathbf{1}_B$ modulo 2, et que cette égalité suffit à reconstruire $A\Delta B$.

b) La différence symétrique est une opération clairement commutative ; montrer qu'elle est également associative, c'est-à-dire que pour toutes parties A, B et C de E,

$$(A\Delta B)\Delta C = A\Delta(B\Delta C).$$

c) On peut donc définir par récurrence sur $n \geqslant 2$ des ensembles

$$A_1\Delta A_2\Delta ...\Delta A_n = (A_1\Delta A_2\Delta ...\Delta A_{n-1})\Delta A_n,$$

pour toutes parties $A_1, \ldots, A_n$ de E. Préciser comment on peut interpréter cette partie de E.

Indication : pour chaque élément x de E, on pourra considérer le cardinal N_x de l'ensemble des indices $i \in \{1, \cdots, n\}$ tels que $x \in A_i$.

7) Déterminer une bijection de $\mathbb{N}$ dans $\mathbb{N} \times \mathbb{N}$. (On pourra faire un dessin.)

8) Montrer que l'ensemble $\{0,1\}^{\mathbb{N}}$ des suites de 0 et de 1 n'est pas dénombrable. (On pourra raisonner par l'absurde.)

En déduire que ni l'ensemble des parties de $\mathbb{N}$, ni l'ensemble $\mathbb{R}$ des nombres réels ne sont dénombrables. (On pourra utiliser l'exercice 5.)

9) Comptages On note désormais $|A|$ le cardinal d'un ensemble A.

a) Montrer que si A est une partie d'un ensemble fini E,

$$|A| = \sum_{x \in E} \mathbf{1}_A(x).$$

b) Si E est un ensemble fini de cardinal $n \geqslant 1$, calculer

$$\sum_{A \subset E} |A|, \quad \sum_{A,B \subset E} |A \cup B|, \quad \sum_{A,B \subset E} |A \cap B|.$$

c) Principe d'inclusion-exclusion de Moivre Soient $(A_i)_{1\leq i\leq n}$ des ensembles finis, parties d'un ensemble E. Démontrer le « principe d'inclusion-exclusion », c'est-à-dire que l'on a l'égalité

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{k=1}^{n}(-1)^{k-1}p_k, \quad \text{où} \quad p_k = \sum_{1\leq i_1<\ldots<i_k\leq n} |A_{i_1}\cap\ldots\cap A_{i_k}|.$$

On démontrera ce résultat de deux manières différentes, par récurrence et en exprimant la fonction indicatrice du complémentaire en fonction des fonctions caractéristiques des ensembles.

Montrer la double inégalité suivante :

$$\sum_{i=1}^{n}|E_i| - \sum_{1\leq i<j\leq n}|E_i\cap E_j| \leqslant \left|\bigcup_{i=1}^{n} E_i\right| \leqslant \sum_{i=1}^{n}|E_i|.$$

Coefficients binomiaux

10) a) Montrer la formule du binôme de Newton : $(x+y)^n = \displaystyle\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}$.

b) Montrer l'identité multinomiale :

$$(x_1+\ldots+x_r)^n = \sum_{n_1+\ldots+n_r=n}\frac{n!}{n_1!\cdots n_r!}\,x_1^{n_1}\cdots x_r^{n_r}.$$

On note pour $n_1+\ldots+n_r=n$, $\left(\begin{array}{c} n \\ n_1\cdots n_r \end{array}\right) = \dfrac{n!}{n_1!\cdots n_r!}$.

Si A est un ensemble à n éléments, $\left(\begin{array}{c} n \\ n_1\cdots n_r \end{array}\right)$ est le nombre de partitions de A en r parties, c'est-à-dire le nombre de r-uplets $(A_1,\ldots,A_r)$ de parties de A deux à deux disjointes telles que $|A_i| = n_i$ (et donc telles que A est la réunion des A_i).

Si $n_1+\ldots+n_r=n$, montrer que

$$\left(\begin{array}{c} n \\ n_1\cdots n_r \end{array}\right) = \sum_{k=1}^{r}\left(\begin{array}{c} n-1 \\ n_1\cdots n_k-1\cdots n_r \end{array}\right),$$

avec la convention que $\left(\begin{array}{c} n \\ n_1\cdots n_r \end{array}\right)$ vaut 0 si l'un des n_i est strictement négatif.

11) On note désormais Y^X l'ensemble des applications de X dans Y. On suppose que $|X|=n$ et $|Y|=m$.

a) Déterminer $|Y^X|$.

b) Déterminer le nombre d'applications injectives de X dans Y.

c) En utilisant le principe d'inclusion-exclusion, montrer que le nombre de surjections de X dans Y est égal à

$$m^n - \binom{m}{1}(m-1)^n + \binom{m}{2}(m-2)^n + \ldots + (-1)^{m-1}\binom{m}{m-1}.$$

d) On suppose que $X = \{1, 2, \cdots, n\}$ et $Y = \{1, 2, \cdots, p\}$.

Montrer que l'ensemble des applications croissantes de X dans Y est en bijection avec l'ensemble des n-uplets $(y_1, y_2, \ldots, y_n)$ d'éléments de Y tels que $y_1 \leq y_2 \leq \cdots \leq y_n$, qui est lui même en bijection avec l'ensemble des n-uplets $(z_1, \ldots, z_n)$ de $\{1, 2, \ldots, n+p-1\}$ tels que $z_1 < z_2 < \cdots < z_n$.

En déduire le nombre d'applications croissantes de X dans Y.

12) Calculer les sommes $s_1 = \sum_{k=0}^{n} k\binom{n}{k}$, $s_2 = \sum_{k=0}^{n} k^2\binom{n}{k}$ et $s_3 = \sum_{k=0}^{n} \frac{1}{k+1}\binom{n}{k}$.

13) Soit E un ensemble de cardinal n. Soit $\mathcal{F}$ l'ensemble des applications de $\mathcal{P}(E)$ dans $\mathbb{R}$. Soient ψ et φ les applications de $\mathcal{F}$ dans $\mathcal{F}$ définies par :

$$\varphi(f)(A) = \sum_{B \subset A} f(B), \qquad \psi(g)(A) = \sum_{B \subset A} (-1)^{|A \setminus B|} g(B).$$

Montrer que φ et ψ sont des bijections réciproques l'une de l'autre.

Indication : on pourra calculer $\sum_{B \subset A} (-1)^{|B|}$.

14) Soit n un entier positif et E un ensemble à n éléments ; calculer le nombre de couples (A, B) constitués de deux parties A et B de E vérifiant $A \subset B$. Généraliser au nombre de k-uplets $(A_1, \ldots, A_k)$ de parties de E tels que $A_1 \subset A_2 \subset \cdots \subset A_k$.

15) On pose $a_0 = 1$ et, pour tout entier $n \geq 1$, on note a_n le nombre de partitions d'un ensemble à n éléments. Démontrer la relation de récurrence :

$$a_{n+1} = \sum_{k=0}^{n} \binom{n}{k} a_{n-k}, \quad n \geqslant 0.$$

Préciser ce qu'on peut dire du nombre de relations d'équivalence sur un ensemble à n éléments.

16) Pour tout couple (n, m) d'entiers strictement positifs, on note :

$$S_{n,m} = \sum_{k=1}^{n} k^m.$$

a) En développant la somme $\sum_{k=0}^{n}(k+1)^m$ par la formule du binôme, trouver une relation entre les coefficients $S_{n,k}$ pour $1 \leqslant k \leqslant m-1$.

b) En déduire la valeur de $S_{n,1}$, $S_{n,2}$ et $S_{n,3}$.

c) Montrer que pour tout $m \geqslant 1$ il existe un polynôme S_m à coefficients rationnels tel que, pour tout entier $n \geqslant 1$, $S_{n,m} = S_m(n)$.

Expliquer pourquoi S_m est unique, préciser son degré et le coefficient de son terme de plus haut degré.

d) Donner un équivalent de $S_{n,m}$ quand n tend vers l'infini (m étant fixé). Retrouver ce résultat en interprétant $\dfrac{S_{n,m}}{n^{m+1}}$ comme une somme de Riemann.

17) Soient m, n et p des entiers naturels tels que $p \leqslant m + n$. Démontrer la formule :

$$\binom{n+m}{p} = \sum_{k=\max(0,p-m)}^{\min(n,p)} \binom{n}{k}\binom{m}{p-k},$$

a) en évaluant de deux façons le nombre de parties à p éléments de la réunion disjointe de deux ensembles, l'un à n éléments et l'autre à m éléments ;

b) en utilisant la formule du binôme.

En déduire la valeur de la somme $\displaystyle\sum_{k=0}^{n} \binom{n}{k}^2$.

18) Partitions d'entiers

Soient n et k des entiers naturels. On note G_n^k le nombre de n-uplets $(x_1, \ldots, x_n)$ d'entiers naturels tels que $x_1 + \cdots + x_n = k$.

a) Déterminer G_n^0, G_n^1 et G_n^2 en fonction de n et G_2^k en fonction de k.

b) Démontrer que $G_{n+1}^{k+1} = G_{n+1}^k + G_n^{k+1}$. On pourra classer les $(n+1)$-uplets tels que $x_1 + \cdots + x_{n+1} = k + 1$ suivant que $x_1 = 0$ ou non.

c) En déduire que

$$G_n^k = \binom{n+k-1}{k}.$$

19) Calculer le coefficient de x^{100} dans le développement en série entière de la fraction R définie par

$$R(x) = \frac{1}{(1-x)(1-x^2)(1-x^5)}.$$

En déduire le nombre de façons dont on peut constituer la somme de 100 € avec des pièces de 1 €, 2 € et 5 €.

Fiche 1 bis : Corrections et commentaires sur la fiche 1

1) La morale de cet exercice est que dans le sens « fonction réciproque », tout fonctionne, mais qu'il faut se méfier du sens direct.

Par contre, les implications suivantes sont vraies sans condition : si $A \subset A'$, alors $f(A) \subset f(A')$; et si $B \subset B'$, alors $f^{-1}(B) \subset f^{-1}(B')$.

On pourra penser aux exemples les plus simples possibles, par exemple la fonction $f : \{a, b\} \to \{a, b\}$ telle que $f(a) = f(b) = a$.

---

3) L'exemple le plus simple pour le cas où E est infini : $E = \mathbb{N}$, $f, g : \mathbb{N} \to \mathbb{N}$ avec $f(n) = n + 1$ et $g(n) = (n - 1)^+$.

---

4) La preuve de cet exercice fait partie de celles qu'on ne peut ignorer.

---

5) Si $f \in \{0, 1\}^E$, considérer $f^{-1}(\{1\})$.

---

7) Même remarque que pour 4).

---

8) Dans l'identification de $[0, 1]$ et de $\{0, 1\}^{\mathbb{N}}$, attention à la non unicité pour les nombres décimaux.

---

9) La morale de cet exercice est que les questions de cardinaux deviennent souvent beaucoup plus simples, une fois traduites en termes de sommes et de fonctions indicatrices.

---

12) La morale de cet exercice est qu'il faut souvent chercher la fonction génératrice derrière une somme, même finie.

---

15) Fixons un élément x de l'ensemble à $n + 1$ éléments, et soit $k + 1$ le cardinal de la partie qui contient x, donc $0 \leqslant k \leqslant n$. Pour spécifier le reste de cette partie, il faut choisir

k éléments parmi n. Pour spécifier le reste de la partition, il faut choisir une partition du reste de l'ensemble, qui a $n-k$ éléments. CQFD.

Les relations d'équivalence et les partitions sont en bijection.

---

16) Pour tout k, $(k+1)^m = \sum_{i=0}^{m} \binom{m}{i} k^i$, donc

$$S_{n+1,m} = S_{n+1,0} + \sum_{i=1}^{m} \binom{m}{i} S_{n,i}.$$

Comme $S_{n+1,m} - S_{n,m} = (n+1)^m$ et $S_{n+1,0} = n+1$, il vient

$$(n+1)^m - (n+1) = \sum_{i=1}^{m-1} \binom{m}{i} S_{n,i}.$$

Pour tout m, $S_{0,m} = 0$. On calcule $S_{n,0} = n$. Pour $m = 2$, $(n+1)^2 - (n+1) = 2S_{n,1}$ donc $S_{n,1} = \frac{1}{2}n(n+1)$. Pour $m = 3$, $(n+1)^3 - (n+1) = 3S_{n,1} + 3S_{n,2}$ donc $S_{n,2} = \frac{1}{6}n(n+1)(2n+1)$.

Si l'assertion « $S_{n,i} = S_i(n)$ pour un polynôme S_i » est vraie pour tout $i \leqslant m-2$, alors $mS_{n,m-1} = (n+1)^m - (n+1) - \sum_{i=1}^{m-2} \binom{m}{i} S_i(n)$ donc l'assertion est vraie pour $m-1$ et

$$mS_{m-1}(n) = (n+1)^m - (n+1) - \sum_{i=1}^{m-2} \binom{m}{i} S_i(n).$$

Le degré de S_0 est 1 et son terme de degré dominant est n. Si le degré de S_i est $i+1$ pour tout $i \leqslant m-2$, alors la somme sur $i \leqslant m-2$ dans le terme de droite de la formule ci-dessus est de degré au plus $m-1$ et le seul terme de degré m provient de $(n+1)^m$ donc le terme dominant de S_{m-1} est n^m/m.

Pour tout m fixé, on en déduit que $S_{n,m} \sim n^{m+1}/(m+1)$ quand $n \to \infty$. Il se trouve que $S_{n,m}/n^{m+1}$ est la somme de Riemann associée à la fonction $f : [0,1] \to \mathbb{R}$ définie par $x \mapsto x^m$, c'est-à-dire

$$\frac{S_{n,m}}{n^{m+1}} = \frac{1}{n} \sum_{k=1}^{n} f\left(\frac{k}{n}\right).$$

Comme f est continue, ses sommes de Riemann convergent vers son intégrale sur $[0,1]$, donc $S_{n,m}/n^{m+1}$ converge vers $1/(m+1)$.

---

17) Soient A et B deux ensembles disjoints de cardinaux respectifs n et m, et soit $C := A \cup B$. Le cardinal de C vaut $n+m$ donc il y a $\binom{n+m}{p}$ parties de C avec p éléments. Pour spécifier une telle partie D, il faut spécifier $D \cap A$ et $D \cap B$. Si le cardinal de $D \cap A$

vaut k, il y a $\binom{n}{k}$ façons de choisir $D \cap A$. Alors le cardinal de $D \cap B$ vaut $p-k$ et il y a $\binom{m}{p-k}$ façons de choisir $D \cap B$. Enfin, on voit que $0 \leqslant k \leqslant n$ puisque $D \cap A \subset A$, et que $0 \leqslant p-k \leqslant m$ puisque $D \cap B \subset B$, ce qui donne les bornes de la somme.

Une autre façon de procéder est de chercher le coefficient de x^p dans $(1+x)^{n+m}$ en écrivant cette fonction comme

$$(1+x)^n(1+x)^m = \sum_i \binom{n}{i} x^i \cdot \sum_j \binom{m}{j} x^j.$$

Le terme en x^p du produit correspond aux indices i et j tels que $i+j=p$, donc $j=p-i$ et l'égalité des coefficients donne

$$\binom{n+m}{p} = \sum_i \binom{n}{i}\binom{m}{p-i}.$$

Pour $n=m=p$, en utilisant le fait que $\binom{n}{n-i} = \binom{n}{i}$, on obtient

$$\binom{2n}{n} = \sum_i \binom{n}{i}\binom{n}{n-i} = \sum_i \binom{n}{i}^2.$$

Fiche 1 ter : Supplément à la fiche 1

20) Arbres sur un ensemble fini

Soit E un ensemble de cardinal supérieur ou égal à 2 et $\mathcal{P}_2(E)$ l'ensemble des parties de E à 2 éléments. Un **graphe** (non orienté et sans boucles) sur E est une partie G de $\mathcal{P}_2(E)$. Les éléments de E sont appelés les **sommets** du graphe et ceux de G les **arêtes** du graphe. Le **degré** d'un sommet est le nombre d'arêtes qui le contiennent. Pour tous sommets a et b et tout entier naturel k, un **chemin** de longueur k joignant a à b est un $(k+1)$-uplet $(x_i)_{0\leqslant i\leqslant k}$ tel que $\{x_{i-1}, x_i\}$ est une arête pour tout $1 \leqslant i \leqslant k$ et tel que $x_0 = a$ et $x_k = b$.

Un **cycle** est un chemin fermé, par exemple (a, b, c, a). Un graphe est **connexe** si deux sommets sont joints par au moins un chemin. Enfin, un **arbre** est un graphe connexe sans cycle.

a) Montrer qu'un arbre possède au moins 2 sommets de degré 1.

b) Soit $\mathcal{R}$ la relation sur E définie par : $x\mathcal{R}y$ si et seulement s'il existe un chemin joignant x à y. Montrer que $\mathcal{R}$ est une relation d'équivalence sur E.

On note $E_1, \ldots, E_r$ les classes d'équivalence de $\mathcal{R}$, et $(G_1, \ldots, G_r)$ les graphes associés par restriction.

Montrer que les graphes (E_i, G_i) sont connexes et que $|G| = \sum_{i=1}^{r} |G_i|$.

c) Montrer par récurrence que si G est connexe et sans cycle, alors il possède $n-1$ arêtes.

d) Montrer que si G est sans cycle et possède $n-1$ arêtes, alors il est connexe. (On pourra utiliser les deux questions précédentes).

e) Montrer que si G est connexe et possède $n-1$ arêtes, alors il est sans cycle. (On pourra raisonner par l'absurde).

f) Soit T un arbre sur $E = \{x_1, \ldots, x_n\}$. On note d_i le degré de x_i dans T.

Montrer que $\sum_{i=1}^{n} d_i = 2(n-1)$.

g) On note $T(n; d_1, \ldots, d_n)$ le nombre d'arbres sur $E = \{x_1, \ldots, x_n\}$ tels que d_i est le degré de x_i.

g1) Si $d_n = 1$, montrer que $T(n; d_1, \ldots, d_n) = \sum_i T(n-1; d_1, \ldots, d_i - 1, \ldots, d_{n-1})$, où la somme porte sur les indices $1 \leqslant i \leqslant n-1$ tels que $d_i \geqslant 2$.

g2) Montrer par récurrence que $T(n; d_1, \ldots, d_n) = \binom{n-2}{d_1 - 1, d_2 - 1, \ldots, d_n - 1}$.

g3) Montrer que le nombre $T(n; k)$ d'arbres sur E tel que $d_n = k$ vaut

$$T(n; k) = \binom{n-2}{k-1}(n-1)^{n-k-1}.$$

g4) En déduire que le nombre total d'arbres sur un ensemble à n éléments vaut n^{n-2}.

---

21) Soit $n \geqslant 1$ un entier naturel.

a) On dispose de n parenthèses ouvrantes et de n parenthèses fermantes. On peut simplifier une suite de parenthèses en effaçant récursivement toutes les paires « ( ) ». Calculer le nombre de façons d'ordonner ces $2n$ parenthèses pour qu'on puisse toutes les effacer.

b) On dispose d'un polygone convexe à $n+2$ sommets. Calculer le nombre de façons de tracer des cordes entre les sommets du polygone de telle sorte qu'elles ne s'intersectent pas à part en leurs extrémités et qu'elles définissent une partition du polygone en triangles.

Correction de l'exercice 20)

a) Soit $(x_i)_{0 \leqslant i \leqslant k}$ un chemin injectif, c'est-à-dire tel que $i \neq j$ entraîne $x_i \neq x_j$, de longueur maximale parmi tous les chemins injectifs (puisqu'un chemin injectif passe au plus une fois par chaque sommet, la longueur de tous les chemins injectifs est majorée par le cardinal de E donc il en existe un ou plusieurs dont la longueur est maximale). Montrons que le degré de x_k vaut 1. Sinon, soit $\{x_k, x_{k+1}\}$ une arête issue de x_k autre que l'arête $\{x_{k-1}, x_k\}$. Par maximalité de la longueur du chemin $(x_i)_{0 \leqslant i \leqslant k}$, le chemin $(x_i)_{0 \leqslant i \leqslant k+1}$ n'est pas injectif donc x_{k+1} est l'un des sommets x_i pour $0 \leqslant i \leqslant k$. Alors $(x_j)_{i \leqslant j \leqslant k+1}$ est un cycle du graphe, ce qui est absurde. De même, le degré de x_0 vaut 1. Il reste à montrer que x_0 et x_k sont bien distincts. Si ce n'est pas le cas, il n'existe aucun chemin injectif de longueur 2, donc aucune arête, donc le graphe n'est pas connexe puisque son cardinal vaut au moins 2.

b) Le chemin (x) de longueur 0 relie x à x, donc la relation est réflexive. Si $(x_i)_{0 \leqslant i \leqslant k}$ relie x à y, alors $(x_{k-i})_{0 \leqslant i \leqslant k}$ relie y à x, donc la relation est symétrique. Si $(x_i)_{0 \leqslant i \leqslant k}$ relie x à y et si $(y_i)_{0 \leqslant i \leqslant \ell}$ relie y à z, alors $(z_i)_{0 \leqslant i \leqslant k+\ell}$ relie x à z si on pose $z_i = x_i$ pour tout $0 \leqslant i \leqslant k$ et $z_i = y_{i-k}$ pour tout $k \leqslant i \leqslant k + \ell$. Donc la relation est transitive.

La connexité de chaque G_i est directe.

c) D'après a), il existe un sommet x de degré 1, soit $\{x, y\}$ l'arête unique à laquelle il appartient. Le graphe de sommets $E' = E \setminus \{x\}$ et d'arêtes $G' = G \setminus \{\{x, y\}\}$ est encore

un arbre donc (E', G') vérifie l'hypothèse de récurrence. Comme $|G| = |G'| + 1$ et $|E| = |E'| + 1$, (E, G) vérifie aussi l'hypothèse de récurrence. Reste à vérifier l'initialisation de la récurrence. Or, si $n = 1$, $E = \{x\}$ et $G = \emptyset$ donc $|E| = 1$ et $|G| = 0$ et ça marche. (Si on préfère partir de $n = 2$, on obtient $E = \{x, y\}$ et $G = \{\{x, y\}\}$ donc $|E| = 2$ et $|G| = 1$ et ça marche.)

d) Pour chaque composante connexe (E_i, G_i), la question c) donne $|E_i| = |G_i| + 1$. Comme $|E|$ est la somme des $|E_i|$ et $|G|$ est la somme des $|G_i|$, il vient $|E| = |G| + r$. Par conséquent, $r = 1$, c'est-à-dire que (E, G) est connexe.

e) Un cycle comporte autant de sommets distincts que d'arêtes. Si G comprend un cycle de longueur k, pour chacun des $n - k$ sommets x hors du cycle, il existe un chemin $c(x)$ issu de x et se terminant sur le cycle, et de longueur minimale. Soit $e(x)$ la première arête de ce chemin. La minimalité des longueurs des chemins fait que les arêtes $e(x)$ sont distinctes (si $e(x) = e(x')$, $c(x)$ commence par x, x' puis fait $\ell(x)$ pas et $c(x')$ issu de x' commence par x', x puis fait $\ell(x')$ pas ; supposons par exemple que $\ell(x) \leqslant \ell(x')$; alors le chemin issu de x' obtenu en suivant la fin du chemin $c(x)$ ne fait que $\ell(x)$ pas alors que $c(x')$ est de longueur $\ell(x') + 1 > \ell(x)$; c'est absurde). Donc G contient les k arêtes du cycle et au moins $n - k$ arêtes distinctes $c(x)$ pour x hors du cycle. En tout, $|G| \geqslant k + (n - k) = n$.

f) On utilise $d_i = \sum_x \mathbf{1}(x \sim x_i)$, où la somme porte sur tous les sommets x de E et où la notation $x \sim y$ signifie que $\{x, y\} \in G$. La somme sur i vaut $\sum_x \sum_i \mathbf{1}(x \sim x_i)$, que l'on peut réécrire

$$\sum_e \sum_x \sum_y \mathbf{1}(\{x, y\} = e).$$

Pour chaque arête e, il y a exactement deux *couples* (x, y) tels que $\{x, y\} = e$. La somme considérée vaut donc le double du nombre d'arêtes.

g1) Puisque $d_n = 1$, x_n appartient à une arête unique, disons $\{x_n, x_i\}$. L'arbre privé de x_n et de l'arête $\{x_n, x_i\}$ est de nouveau un arbre, cette fois sur l'ensemble de sommets $E' = E \setminus \{x_n\}$ et avec les arêtes $G' = G \setminus \{\{x_n, x_i\}\}$. La suite des degrés de ses sommets vaut $(d_1, \ldots, d_i - 1, \ldots, d_{n-1})$, CQFD.

g2) Si $n = 2$, $d_1 = d_2 = 1$ donc $T(2; 1, 1) = 1$ et le membre de droite vaut $\binom{0}{0,0} = 1$.

Si le résultat est vrai pour $n - 1$ avec $n \geqslant 3$ et si $d_n = 1$, alors,

$$T(n; d_1, \ldots, d_n) = \sum_i \binom{n-3}{d_1 - 1, d_2 - 1, \ldots, d_i - 2, \ldots, d_{n-1} - 1}.$$

Chaque terme de la somme vaut

$$(d_i - 1) \binom{n-3}{d_1 - 1, d_2 - 1, \ldots, d_i - 1, \ldots, d_{n-1} - 1},$$

donc la somme vaut $\sum_i (d_i - 1) = n - 2$, d'après f). On en déduit le résultat quand $d_n = 1$. Le cas général s'en déduit, puisque la valeur de $T(n, d_1, \ldots, d_n)$ est invariante par permutation circulaire sur les degrés $(d_i)_{1 \leqslant i \leqslant n}$.

g3) On doit faire la somme de $T(n; d_1, \ldots, d_{n-1}, k)$ sur tous les $(d_i)_{1 \leqslant i \leqslant n-1}$ tels que $\sum_{i=1}^{n-1} d_i = 2n - 2 - k$. Chaque coefficient multinomial est le coefficient de

$$x_1^{d_1 - 1} \cdots x_{n-1}^{d_{n-1} - 1} x_n^{k-1},$$

dans le développement de $(x_1 + x_2 + \cdots + x_n)^{n-2}$. En faisant la somme et en posant $x_1 = x_2 = \cdots = x_{n-1} = 1$, on obtient le coefficient de x_n^{k-1} dans

$$(1 + 1 + \cdots + x_n)^{n-2} = (n - 1 + x_n)^{n-2},$$

soit $\binom{n-2}{k-1} (n-1)^{n-k-1}$.

g4) La somme sur k correspond au développement de $(1 + x)^{n-2}$ quand $x = n - 1$, CQFD.

Correction de l'exercice 21)

a) Une suite de longueur $2n$ admissible commence par une parenthèse ouvrante. La parenthèse fermante correspondante occupe la position $2k$ avec $1 \leqslant k \leqslant n$. La suite restreinte aux positions $2 \leqslant i \leqslant 2k - 1$ est de longueur $2k - 2$ et doit être admissible, de même la suite restreinte aux positions $2k + 1 \leqslant j \leqslant 2n$ est de longueur $2n - 2k$ et doit être admissible. Les correspondances sont bijectives donc on en déduit une relation de récurrence sur le nombre a_n de façons admissibles d'ordonner $2n$ parenthèses, à savoir

$$a_n = \sum_{k=1}^{n} a_{k-1} a_{n-k}, \quad n \geqslant 1,$$

avec la convention $a_0 = 1$. Introduisons la série formelle $A(x) = \sum_{n \geqslant 0} a_n x^n$. Puisque le nombre de suites de longueur $2n$, admissibles ou non, formées de parenthèses ouvrantes et fermantes vaut 2^{2n}, on sait que $a_n \leqslant 4^n$ et que le rayon de convergence de A vaut au moins $\frac{1}{4}$. Pour tout x dans le disque de convergence de A, il vient successivement

$$A(x) = 1 + x\, A(x)^2, \qquad A(x) = (1 - \sqrt{1 - 4x})/(2x),$$

puisque l'autre racine du polynôme du second degré est de valuation -1 par rapport à l'inconnue x (donc le rayon de convergence de A vaut exactement $\frac{1}{4}$).

Il reste à utiliser le développement en série entière de la fonction $t \mapsto \sqrt{1 - t}$ pour obtenir

$$a_n = \binom{2n}{n} / (n + 1).$$

Donc a_n est le $n$ème nombre de Catalan.

Voici une autre solution qui consiste à compter des chemins dans le plan et qui repose sur un principe de symétrie bien connu. Notons $x_0 = 0$ et, pour chaque $1 \leqslant k \leqslant 2n$, x_k le nombre de parenthèses ouvrantes moins le nombre de parenthèses fermantes parmi les k premières parenthèses. Une suite $(x_k)_{0\leqslant k\leqslant 2n}$ est admissible si $x_0 = x_{2n} = 0$ et $x_k \geqslant 0$ pour tout k. Associons à une suite $(x_k)_k$ le chemin du plan $k \mapsto (k, x_k)$. Une suite admissible correspond à un chemin du plan reliant $(0, 0)$ à $(2n, 0)$ en faisant des pas de $(1, \pm 1)$ et en restant dans le demi-plan supérieur $y \geqslant 0$.

Le nombre total de chemins (sans condition de positivité) vaut $\binom{2n}{n}$. Tout chemin non admissible rencontre la droite d'équation $y = -1$. Après le premier instant où ce chemin rencontre cette droite, remplaçons-le par son symétrique par rapport à cette droite. Le résultat est un chemin du plan reliant $(0, 0)$ à $(2n, -2)$ en faisant des pas de $(1, \pm 1)$. Il y a $\binom{2n}{n+1}$ tels chemins. Le nombre de chemins admissibles est donc

$$a_n = \binom{2n}{n} - \binom{2n}{n+1} = \binom{2n}{n}(1 - n/(n+1)) = \binom{2n}{n}/(n+1).$$

b) $c_1 = 1$ puisqu'il n'y a qu'une seule décomposition, avec aucune corde, et $c_2 = 2$ puisqu'il faut choisir une des deux diagonales d'un quadrilatère.

Soit $n \geqslant 2$. On commence par numéroter les sommets de 0 à $n + 1$, par exemple dans le sens direct. Si 0 n'appartient à aucune corde, ses deux voisins 1 et $n + 1$ sont reliés par une corde. Il reste à choisir une décomposition du polygone de sommets $\{1, \ldots, n+1\}$, et il y en a c_{n-1}. Sinon, 0 appartient à au moins une corde. Soit k le plus petit entier tel que 0 et k sont reliés par une corde, donc $2 \leqslant k \leqslant n$. Il reste à choisir deux décompositions : une décomposition du polygone de sommets $\{0, k, k + 1, \ldots, n + 1\}$, et il y en a c_{n-k+1} ; et une décomposition du polygone de sommets $\{0, 1, \ldots, k\}$ sans corde issue de 0, donc une décomposition du polygone de sommets $\{1, \ldots, k\}$, et il y en a c_{k-2} . Finalement, avec la convention $c_0 = 1$ et en prenant garde aux premiers et aux derniers termes, il vient, pour tout $n \geqslant 1$,

$$c_n = c_{n-1} + \sum_{k=2}^{n} c_{k-2}c_{n-k+1} = \sum_{k=1}^{n} c_{k-1}c_{n-k}.$$

On retrouve la même récurrence qu'au a) avec les mêmes conditions initiales, donc $c_n = \binom{2n}{n}/(n + 1)$.

À ma connaissance, le problème de trouver une preuve bijective du fait que $a_n = c_n$ pour tout $n \geqslant 1$ est encore ouvert. On appelle « preuve bijective de $a_n = c_n$ » le fait d'exhiber une bijection entre l'ensemble énuméré par a_n et l'ensemble énuméré par c_n.

Fiche 2 : Probabilités (1) Introduction aux espaces probabilisés

« Le degré d'excitation qu'éprouve un joueur en faisant un pari est égal au produit du gain par la probabilité de gagner. » Pascal

Introduction

On fait une expérience dont on ne peut pas déterminer le résultat à l'avance, par exemple on jette une pièce de monnaie en l'air et on s'intéresse au côté sur lequel elle retombe (« pile » ou « face »). Que peut-on dire de plus du résultat de cette expérience ? Si on ne jette la pièce qu'une fois, pas grand'chose. Mais si on la jette un grand nombre de fois, on constate que la fréquence d'apparition de « pile » se stabilise et, pour de nombreuses pièces, que cette fréquence devient proche de 50%. On associe alors au résultat de l'expérience « obtenir pile » la valeur 50%, qu'on appellera la probabilité de l'événement « obtenir pile ».

Lier la probabilité d'un événement à la fréquence d'apparition de cet événement lors d'un grand nombre d'expériences résulte d'un théorème important : la « loi des grands nombres ». En effet, on montre que si F_n désigne la valeur observée de cette fréquence après n lancers et si la pièce est « équilibrée » en un certain sens, alors F_n tend vers 50% avec probabilité 1 quand n tend vers l'infini.

De plus, on peut quantifier ce premier résultat en étudiant les fluctuations de la suite $(F_n)_{n\geqslant 1}$ autour de 50%. Par exemple, on a lancé 1000 fois une pièce de monnaie et on a obtenu 537 piles ; est-ce normal ou doit-on en conclure que la pièce est biaisée ? Le « théorème central limite » répond à cette question en montrant que la vitesse de convergence de la suite aléatoire $(F_n)_n$ vers 50% est en $1/\sqrt{n}$. (Dans l'exemple, d'après le théorème central limite, la probabilité d'obtenir 537 piles ou plus que 537 piles vaut à peu près $0,964\%$, valeur approchée que l'on peut comparer avec la vraie valeur $1,046\%$.)

Le cours commence par définir un espace probabilisé $(\Omega, \mathcal{F}, P)$, les probabilités conditionnelles, les variables aléatoires et leurs lois, et la notion d'indépendance. Comme on l'a vu dans l'exemple précédent, on étudie souvent des phénomènes qui correspondent à la répétition d'une même expérience donc on s'interessera ensuite aux suites de variables aléatoires et on énoncera les deux théorèmes limites fondamentaux mentionnés ci-dessus que sont la loi forte des grands nombres et le théorème central limite.

1 Espaces probabilisés

1.1 Définition

Pour définir un espace probabilisé, on a besoin d'un ensemble Ω appelé **univers**, qui peut représenter l'ensemble des résultats possibles de l'expérience considérée. Une première solution est de choisir pour Ω un ensemble aussi petit que possible, donc la collection des résultats possibles de l'expérience.

Quelques exemples.
- Pour le jet d'une pièce de monnaie, $\Omega = \{0, 1\}$ convient, où 0 représente « pile » et 1 représente « face » (ou vice versa !).
- Pour le jet de n pièces de monnaie ou n jets d'une seule pièce de monnaie : $\Omega = \{0, 1\}^n$ convient.
- Pour le lancer d'une fléchette sur une cible : Ω un disque du plan euclidien convient.
- Pour la durée de vie d'une ampoule électrique : $\Omega = \mathbb{R}^+$.
- Pour battre un jeu de 52 cartes : $\Omega = \mathfrak{S}_{52}$ convient, où $\mathfrak{S}_k$ désigne l'ensemble des permutations de l'ensemble $\{1, 2, \ldots, k\}$.
- Pour compter le nombre de clients dans une file d'attente : $\Omega = \mathbb{N}$ convient.
- Pour lancer une pièce de monnaie une infinité de fois : $\Omega = \{0, 1\}^{\mathbb{N}^*}$ convient, donc Ω est l'ensemble des suites à valeurs dans $\{0, 1\}$, c'est-à-dire l'ensemble des suites $(x_n)_{n\geqslant 1}$ où $x_n = 1$ si le tirage numéro n donne pile et $x_n = 0$ sinon.

Dans ces exemples, les ensembles Ω sont minimaux et on voit bien que tout ensemble plus « gros » conviendrait aussi. Par exemple, $\Omega = \{0, 1\}^{\mathbb{N}^*}$ convient pour modéliser le jet de n pièces de monnaie pour tout $n \geqslant 1$: il suffit de ne s'intéresser qu'aux n premières coordonnées.

Cette remarque suggère une deuxième solution qui consiste à ne pas se préoccuper de la forme exacte de Ω et, par contre, à spécifier ce que l'observateur peut voir ou non. Pour cela, on introduit la tribu des **événements**[1], souvent notée $\mathcal{F}$, qui correspond à « ce que l'observateur voit ». Les éléments de $\mathcal{F}$ sont donc des parties de Ω.

Quelles propriétés une telle classe d'événements doit-elle vérifier ? Si l'observateur peut voir si A est réalisé, il peut voir si A^c l'est. De même si l'observateur peut voir si A est réalisé et si B l'est, il peut voir si $A \cup B$ l'est ou non et si $A \cap B$ l'est ou non.

Enfin, on exige que si $(A_n)_{n\geqslant 0}$ est une suite d'événements, alors leur réunion l'est aussi[2].

Définition 1.1. *Une tribu (ou σ-algèbre) de E est une partie $\mathcal{F}$ de $\mathcal{P}(\Omega)$ vérifiant les trois axiomes suivants.*

1. *Ω appartient à $\mathcal{F}$.*

[1] Le lecteur attentif aura remarqué que nous utilisons dans ces notes la graphie « événement » et non pas « évènement », qui est pourtant recommandée par le service du Dictionnaire de l'Académie française dans un rapport controversé datant de 1990. À chacun de décider pour ce qui le concerne.

[2] Ce dernier point peut faire (et a fait) débat et il existe des versions « finitistes » des axiomes des probabilités qui l'excluent. Ces versions sortent du cadre du programme et nous n'en parlerons donc pas du tout.

2. *Si A appartient à $\mathcal{F}$, alors $\Omega \setminus A$ appartient à $\mathcal{F}$.*

3. *Si $(A_n)_{n\geqslant 0}$ est une suite d'éléments de $\mathcal{F}$, alors $\bigcup_{n=0}^{+\infty} A_n$ appartient à $\mathcal{F}$.*

Exemple 1.2. *Si Ω est fini, on choisira souvent $\mathcal{F} = \mathcal{P}(\Omega)$, mais pas toujours. Supposons par exemple que l'on jette un dé mais qu'on ne nous donne que la parité du résultat. La tribu associée sera $\mathcal{F} = \{\emptyset, pair, impair, \Omega\}$ et non pas une classe de 2^6 parties.*

La dernière étape de la construction consiste à mesurer chaque élément A de $\mathcal{F}$ par sa « probabilité », notée $P(A)$.

Définition 1.3. *Une probabilité sur $(\Omega, \mathcal{F})$ est une fonction $P : \mathcal{F} \to [0,1]$ telle que :*

1. *$P(\Omega) = 1$*
2. *Pour toute suite $(A_n)_{n\geqslant 0}$ d'éléments de $\mathcal{F}$ deux à deux disjoints et de réunion A,*

$$P(A) = \sum_{n=0}^{+\infty} P(A_n).$$

*Le triplet $(\Omega, \mathcal{F}, P)$ est appelé un **espace probabilisé** ou un **espace de probabilité**.*

1.2 Conséquences

Quelques propriétés faciles.

1) $P(\emptyset) = 0$;

2) Pour tout événement A, $P(\Omega \setminus A) = 1 - P(A)$;

3) Pour tous événements A et B, si $A \subset B$, $P(A) \leqslant P(B)$;

4) Pour tous événements A et B, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;

5) Si $(A_n)_{n\geqslant 0}$ est une suite croissante d'événements de réunion A, $P(A)$ est la limite de la suite croissante $(P(A_n))_{n\geqslant 0}$;

6) Si $(A_n)_{n\geqslant 0}$ est une suite décroissante d'événements d'intersection A, $P(A)$ est la limite de la suite décroissante $(P(A_n))_{n\geqslant 0}$.

Preuves 1) $\emptyset$ et $\emptyset$ sont disjoints de réunion $\emptyset$ donc $P(\emptyset) = 2P(\emptyset)$ d'où $P(\emptyset) = 0$.

2) A et A^c sont disjoints et de réunion égale à Ω donc $P(\Omega) = 1 = P(A) + P(A^c)$.

3) A et $B \setminus A$ appartiennent à $\mathcal{F}$ et sont disjoints, de réunion égale à B. On en déduit donc que $P(B) = P(A) + P(B \setminus A)$ d'où $P(B) \geqslant P(A)$.

4) A et $B \setminus A$ sont disjoints de réunion égale à $A \cup B$ d'où $P(A \cup B) = P(A) + P(B \setminus A)$. De plus $B \setminus A$ et $B \cap A$ sont disjoints de réunion égale à B d'où $P(B) = P(B \setminus A) + P(B \cap A)$. On en déduit le résultat.

5) Soit $B_n = A_{n+1} \setminus A_n$. La suite $(B_n)_{n\geqslant 0}$ est une suite d'éléments de $\mathcal{F}$ deux à deux disjoints et, pour tout $n \geqslant 0$, A_{n+1} est la réunion des B_k pour $0 \leqslant k \leqslant n$. Donc $P(A_{n+1}) = \sum_{k=0}^{n} P(B_k)$. Cette série est donc convergente et de somme égale à

$$\sum_{k=0}^{+\infty} P(B_k) = P\left(\bigcup_{k=0}^{+\infty} B_k\right) = P\left(\bigcup_{k=0}^{+\infty} A_k\right).$$

On obtient le résultat sur les suites décroissantes par passage au complémentaire.

Exemple 1.4 (Formule de Poincaré). *Soit $A_1, \ldots, A_n$ des événements. Montrer que*

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} (-1)^{k+1} s_k, \quad \textit{avec } s_k = \sum_{1\leqslant i_1<\ldots<i_k\leqslant n} P(A_{i_1} \cap \ldots \cap A_{i_k}).$$

1.3 Espaces probabilisés finis

On suppose que Ω est fini avec $\Omega = \{\omega_1, \ldots, \omega_n\}$, et que $\mathcal{F} = \mathcal{P}(\Omega)$.

Ou bien on suppose que $\mathcal{F}$ est finie donc (exercice) engendrée par une partition $(A_i)_{1\leqslant i\leqslant n}$ au sens où

$$\mathcal{F} = \{A_I \,;\, I \subset \{1, 2, \ldots, n\}\}, \quad A_I = \bigcup_{i\in I} A_i.$$

La correspondance entre les deux descriptions est $\{\omega_i\} \to A_i$. On revient désormais à la première description.

Pour i dans $\{1, 2, \ldots, n\}$, on note $p_i = P(\{\omega_i\})$, donc $0 \leqslant p_i \leqslant 1$ et $\sum_{i=1}^{n} p_i = 1$.

Réciproquement, à tout n-uplet $(p_1, \ldots, p_n)$ vérifiant ces conditions, on associe une unique probabilité P sur $(\Omega, \mathcal{F})$ donnée par $P(\{\omega_i\}) = p_i$ pour tout i dans $\{1, 2, \ldots, n\}$. Alors, pour toute partie A de Ω,

$$P(A) = \sum_{i\in I} p_i, \qquad I = \{i \in [n] \,;\, \omega_i \in A\}.$$

Premier exemple : on jette 3 dés équilibrés n fois de suite. Donc $\Omega = \left(\{1, 2, \ldots, 6\}^3\right)^n$ et $P(\{\omega\}) = 1/6^{3n}$ pour tout ω dans Ω.

La probabilité d'obtenir au moins une fois 421 vaut $1 - (1 - 1/36)^n$. Quelle est la probabilité d'obtenir au moins une fois 444 ?

Deuxième exemple : une urne contient 10 boules rouges et 5 boules noires. On tire 4 boules simultanément et on s'intéresse au nombre de boules rouges, donc on pourra choisir l'ensemble $\Omega = \{0, 1, 2, 3, 4\}$ et $\omega = i$ si on obtient i boules rouges et $4 - i$ boules noires.

Par exemple, $P(\{3\}) = \binom{5}{1}\binom{10}{3}/\binom{15}{4} = 40/91$ et la probabilité de tirer au moins une boule rouge vaut $1 - \left(\binom{5}{4}/\binom{15}{4}\right) = 272/273$.

1.4 Exemples de probabilités

1.4.1 Probabilité uniforme

Pour tout ω dans Ω de cardinal N, on pose $P(\{\omega\}) = 1/N$.

Alors, $P(A) = (\operatorname{card} A)/N$ pour toute partie A (c'est le fameux « nombre de cas favorables sur nombre de cas possibles »). Cette probabilité correspond à la notion de tirage au hasard. Tous les tirages sont équiprobables. Les calculs de telles probabilités sont donc des problèmes de dénombrement.

1.4.2 Probabilités multinomiales

Soit un ensemble fini U de cardinal N (une urne), et $\Omega = U^n$. Cela correspond à l'expérience de tirer n éléments avec replacement dans U. Le cardinal de Ω est N^n et on suppose tous les tirages équiprobables.

Fixons $U_0 \subset U$ de cardinal N_0. Pour $0 \leqslant k \leqslant n$, on s'intéresse à la probabilité p_k qu'un échantillon de taille n tiré au hasard dans U contienne exactement k éléments de U_0. Alors,

$$p_k = \binom{n}{k} p^k (1-p)^{n-k}, \qquad p = N_0/N.$$

En effet, $p_k = (\operatorname{card} A_k)/N^n$ où A_k est l'ensemble des ω tels que ω_i appartient à U_0 pour exactement k indices i. Pour une partie I donnée, de cardinal k, l'ensemble des ω tels que ω_i appartient à U_0 si et seulement si i appartient à I est de cardinal $N_0^k (N-N_0)^{n-k}$, et il y a $\binom{n}{k}$ parties I de cette sorte, donc $\operatorname{card}(A_k) = \binom{n}{k} N_0^k (N-N_0)^{n-k}$.

Par définition, on a $p_0 + p_1 + \ldots + p_n = 1$ donc $(p_0, p_1, \cdots, p_n)$ définit une probabilité sur $\{0, 1, \ldots, n\}$ appelée probabilité binomiale de paramètre (n, p).

On dit aussi que le nombre d'éléments de U_0 suit une loi binomiale de paramètre (n, p).

L'extension à plus de deux parties est la suivante : soit $U = U_1 \cup \ldots \cup U_r$ une partition de U en r parties avec $r \geqslant 2$ et avec $\operatorname{card} U_i = N_i$. Notons $p_i = N_i/N$.

Si $n_1 + \ldots + n_r = n$, la probabilité $p(n_1, \ldots, n_r)$ qu'un échantillon de taille n tiré au hasard dans U contienne exactement n_i éléments de U_i pour chaque $1 \leqslant i \leqslant r$ vaut

$$p(n_1, \ldots, n_r) = \frac{n!}{n_1!\, n_2! \ldots n_r!} p_1^{n_1} \ldots p_r^{n_r}.$$

On parle alors de probabilité multinomiale de paramètre $(n, (p_1, \ldots, p_r))$.

1.4.3 Probabilités hypergéométriques

On considère les ensembles précédents et l'expérience consistant à tirer n éléments dans U **sans replacement**, donc $0 \leqslant n \leqslant N$. On suppose tous les tirages équiprobables.

Soit $\Omega = \mathcal{P}_n(U)$ l'ensemble des parties à n éléments de U. Le cardinal de Ω est $\binom{N}{n}$.

Soit $0 \leqslant k \leqslant n$ un entier tel que $k \leqslant N_0$ et $n - k \leqslant N - N_0$. La probabilité p_k qu'une sous-population de taille n de U contienne exactement k éléments de U_0 vaut

$$p_k = \binom{N_0}{k}\binom{N - N_0}{n - k} / \binom{N}{n}.$$

En effet $p_k = \operatorname{card} A_k / \binom{N}{n}$ où A_k est l'ensemble des éléments ω de Ω (donc ω est une partie de U) tels que $\operatorname{card}(\omega \cap U_0) = k$. De tels ω s'écrivent comme $\omega = \omega_0 \cup \omega_1$ avec ω_0 dans $\mathcal{P}_k(U_0)$ et ω_1 dans $\mathcal{P}_{n-k}(U \setminus U_0)$.

Ceci définit une probabilité sur $\{0, 1, \ldots, n\}$ appelée probabilité hypergéométrique de paramètre (n, N, N_0).

On retrouve la formule $\binom{N}{n} = \displaystyle\sum_{k=0}^{n} \binom{N_0}{k}\binom{N - N_0}{n - k}$, où par convention $\binom{n}{k} = 0$ si $k > n$ ou $k < 0$.

Remarque 1.5. *Si N est grand, les deux tirages ne diffèrent pas beaucoup. En effet, si $N \to \infty$ et $N_0/N \to p$ avec $p \in]0, 1[$, alors $p_k \to \binom{n}{k} p^k (1-p)^{n-k}$. En effet,*

$$\frac{\binom{N_0}{k}\binom{N-N_0}{n-k}}{\binom{N}{n}} = \binom{n}{k} \frac{N_0 \cdots (N_0 - k + 1)(N - N_0) \cdots (N - N_0 - n + k + 1)}{N(N-1)\cdots(N-n+1)},$$

donc

$$\frac{\binom{N_0}{k}\binom{N-N_0}{n-k}}{\binom{N}{n}} \sim \binom{n}{k} \frac{N_0^k (N - N_0)^{n-k}}{N^n} \to \binom{n}{k} p^k (1-p)^{n-k}.$$

2 Probabilités conditionnelles et indépendance de deux événements

La notion de probabilité conditionnelle apparaît lorsqu'on connaît un résultat partiel de l'expérience.

Par exemple, supposons qu'une usine construise des objets de types A et B, et que ces objets peuvent ou non être de bonne qualité, propriété qu'on notera Q.

L'expérience nous permet de connaître $P(A \cap Q)$, estimé par $N_{A \cap Q}/N$ où N_C est le nombre d'objets partageant la propriété C parmi N produits.

On se restreint maintenant à la sous-population des objets de type A et on se demande quelle est la probabilité qu'un objet tiré au hasard dans cette sous-population soit de bonne qualité.

On va donc estimer ceci par $N_{A \cap Q}/N_A$. Or $N_{A \cap Q}/N_A = (N_{A \cap Q}/N)/(N_A/N)$.

On a besoin d'une nouvelle notation pour indiquer qu'on se restreint à cette sous population et on notera $P(Q|A)$, la probabilité de Q sachant A.

On a donc $P(Q|A) \approx N_{A\cap Q}/N_A \approx P(Q \cap A)/P(A)$.

Définition 2.1. *Soit A un événement de probabilité $P(A)$ non nulle. On définit une application $P(\ |A) : \mathcal{F} \to [0,1]$ par*

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

La quantité $P(B|A)$ est la probabilité conditionnelle de B sachant A.

Remarque 2.2. *La relation précédente entraîne*

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

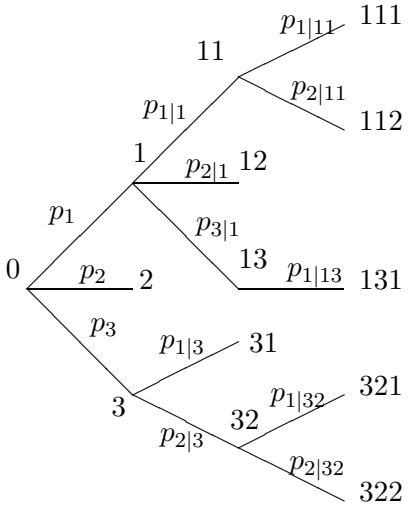
Par récurrence, on en déduit :

Proposition 2.3 (Probabilités conditionnelles en cascade).

$$P(A_1 \cap \ldots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1}),$$

à condition de donner au second membre la valeur zéro dès que l'un de ses termes est nul (même si certaines des probabilités conditionnelles y figurant ne sont alors pas définies).

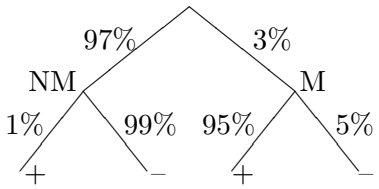
Dans certains problèmes où apparaissent les probabilités conditionnelles, on peut utiliser des arbres. Par exemple, supposons qu'on associe à un individu choisi au hasard dans une population un « comportement » $(X_1, \ldots, X_n)$, dont la longueur n peut dépendre de l'individu. On considèrera alors l'arbre représenté ici.



On a noté $p_{x_k|x_1\cdots x_{k-1}}$ la probabilité du comportement $(x_1, \ldots, x_k)$ sachant qu'on a le comportement $(x_1, \ldots, x_{k-1})$.

Considérons par exemple la situation suivante : une maladie atteint 3% d'une population. Un test de dépistage donne les résultats suivants : chez les malades, 95% des tests sont positifs ; chez les non-malades, 99% des tests sont négatifs.

On représentera ces données dans l'arbre suivant :



Les probabilités conditionnelles sont souvent utilisées pour étudier des phénomènes évoluant au cours du temps, par exemple des suites de lancers de pile ou face ou de tirages dans une urne.

Pour un exemple typique, supposons qu'une urne U contient N éléments et que l'on tire r éléments les uns après les autres de la façon suivante : quels que soient les k premiers éléments tirés, au tirage numéro $k+1$, chacun des $N-k$ éléments restants a une probabilité $1/(N-k)$ d'être choisi.

Soit Ω l'ensemble des r-uplets d'éléments distincts de U et notons X_i l'application donnant le résultat du tirage numéro i. Fixons un élément $a = (a_1, a_2, \ldots, a_r)$ de Ω et, pour $1 \leqslant k \leqslant r$, soit A_k l'événement « on a tiré l'élément a_k au tirage numéro k ». Alors,

$$P(\{a\}) = P(A_1 \cap A_2 \cap \ldots \cap A_r),$$

donc

$$P(\{a\}) = P(A_1)P(A_2|A_1)\cdots P(A_r|A_1 \cap A_2 \cap \ldots \cap A_{r-1}),$$

donc

$$P(\{a\}) = 1/(N(N-1)\cdots(N-r+1)),$$

c'est-à-dire que P est la probabilité uniforme sur Ω comme on pouvait s'en douter.

On peut calculer la probabilité de tirer k éléments dans $U_0 \subset U$ de la même façon, et on retrouve les probabilités hypergéométriques.

Proposition 2.4. *Soit A un événement de probabilité $P(A)$ non nulle. Alors $P(\ |A) : \mathcal{F} \to [0,1]$, $B \mapsto P(B|A)$, est une probabilité sur $(\Omega, \mathcal{F})$.*

La démonstration découle immédiatement de la définition.

Proposition 2.5 (Formule des probabilités totales). *Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité et $(A_i)_{i\in I}$ une partition dénombrable de Ω en événements de probabilités $P(A_i)$ strictement positives pour tout $i \in I$. Alors, pour tout événement A,*

$$P(A) = \sum_{i\in I} P(A|A_i)P(A_i).$$

Preuve : La famille $(A \cap A_i)_{i\in I}$ est une partition de A, donc,

$$P(A) = \sum_{i\in I} P(A \cap A_i) = \sum_{i\in I} P(A|A_i)P(A_i).$$

Proposition 2.6 (Formule de Bayes). *Sous les hypothèses de la proposition précédente, pour tout événement A de probabilité $P(A)$ non nulle et pour tout i,*

$$P(A_i|A) = \frac{P(A|A_i)P(A_i)}{\sum_{j\in I} P(A|A_j)P(A_j)}.$$

Preuve : Il suffit d'écrire $P(A_i|A) = \frac{P(A|A_i)P(A_i)}{P(A)}$ et d'utiliser la formule des probabilités totales.

Supposons à présent que les événements A et B sont tels que la réalisation de A n'induise rien sur la réalisation de B. Alors, $P(B|A) = P(B)$ donc $P(B \cap A) = P(A)P(B)$.

Définition 2.7. *Deux événements A et B sont indépendants si et seulement si*

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

Remarque 2.8. ***Attention*** *La notion d'indépendance dépend de la probabilité P !*

Exemple : on dispose de deux pièces de monnaie, l'une équilibrée et l'autre équilibrée ou non. On choisit une des deux pièces au hasard de manière uniforme puis on la lance deux fois. On peut donc toujours utiliser comme espace de probabilité $\Omega = \{\text{pile}, \text{face}\}^2$ et $\mathcal{F} = \mathcal{P}(\Omega)$ mais la probabilité P dépend du biais éventuel de la deuxième pièce.

Soit A l'événement « le premier jet donne face » et B l'événement « le deuxième jet donne face ». Montrer que A et B sont indépendants si et seulement si la deuxième pièce est équilibrée.

Proposition 2.9. *Soient A et B deux événements. Soient $\mathcal{A}$ et $\mathcal{B}$ les tribus engendrées respectivement par A et B, c'est à dire $\mathcal{A} = \{\emptyset, A, A^c, \Omega\}$ et $\mathcal{B} = \{\emptyset, B, B^c, \Omega\}$. Alors A et B sont indépendants si et seulement si pour tout $C \in \mathcal{A}$ et pour tout $D \in \mathcal{B}$, $P(C \cap D) = P(C)P(D)$.*

Preuve : Par exemple, $P(A \cap B^c)$ vaut

$$P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(B^c).$$

3 Indépendance d'événements

On veut généraliser la notion d'indépendance à un nombre quelconque d'événements.

Définition 3.1. *Les événements $(A_i)_{1\leqslant i\leqslant n}$ sont indépendants si et seulement si pour tout $I \subset \{1, 2, \ldots, n\}$, on a*

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

Rappelons que par convention, un produit sur un ensemble vide d'indices vaut 1 (de même qu'une somme sur un ensemble vide d'indices vaut 0), donc on peut choisir $I = \emptyset$ dans la définition précédente.

Proposition 3.2. *Considérons les propriétés suivantes :*

(i) Les événements $(A_i)_{1 \leqslant i \leqslant n}$ sont indépendants.

(ii) Pour toute famille $(B_i)_{1 \leqslant i \leqslant n}$ telle que $B_i \in \{A_i, A_i^c\}$,

$$P\left(\bigcap_{1 \leqslant i \leqslant n} B_i\right) = \prod_{1 \leqslant i \leqslant n} P(B_i).$$

(iii) Pour toute famille $(B_i)_{1 \leqslant i \leqslant n}$ telle que $B_i \in \{\emptyset, A_i, A_i^c, \Omega\}$,

$$P\left(\bigcap_{1 \leqslant i \leqslant n} B_i\right) = \prod_{1 \leqslant i \leqslant n} P(B_i).$$

Les propriétés (i), (ii) et (iii) sont équivalentes.

Preuve Tout d'abord, (iii) implique (i) et (ii) car (i) revient à se restreindre aux $B_i \in \{A_i, \Omega\}$ et (ii) revient à se restreindre aux $B_i \in \{A_i, A_i^c\}$. Ensuite (ii) implique (iii) : si on utilise une fois $\emptyset$ dans (iii), ça marche ; sinon, (iii) s'obtient en sommant 2^a égalités (ii), où a désigne le nombre de fois où (iii) utilise Ω. Reste à montrer que (i) implique (iii).

Supposons que (i) est vraie et (iii) fausse. Pour chaque famille $B = (B_i)_{1 \leqslant i \leqslant n}$ qui infirme (iii), notons $n(B)$ le nombre de fois où B utilise A_i^c et non pas A_i ou Ω (puisque dès qu'on utilise $\emptyset$, ça marche). Le cas $n(B) = 0$ est impossible car (i) est vraie, donc $n(B) \geqslant 1$. Supposons que la famille B est telle que $n(B)$ est minimal et soit i un indice tel que $B_i = A_i^c$. Alors la famille B' obtenue en remplaçant A_i^c par A_i dans la famille B est telle que $n(B') = n(B) - 1 < n(B)$ donc B' vérifie (iii). En sommant le fait que (iii) pour B est fausse et le fait que (iii) pour B' est vraie, on obtient que (iii) est fausse pour la famille B'' obtenue à partir de B en remplaçant A_i^c par Ω. Comme $n(B'') = n(B) - 1$, c'est absurde.

Remarque 3.3. ***Attention*** *Le fait que les événements $(A_i)_{i \in I}$ sont indépendants entraîne qu'ils sont indépendants deux à deux mais la réciproque est fausse.*

Exemple 3.4. *On lance une pièce deux fois de suite. Soit A l'événement « obtenir un face au premier jet », B l'événement « obtenir un face au deuxième jet », et C l'événement « obtenir deux résultats différents ».*

Alors $P(A) = P(B) = P(C) = \frac{1}{2}$ et $P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}$, donc A, B, C sont indépendants deux à deux. Mais $A \cap B \cap C = \emptyset$ donc $P(A \cap B \cap C) = 0 \neq P(A)P(B)P(C) = \frac{1}{8}$ donc (A, B, C) n'est pas une famille indépendante.

Question : Pour tout $n \geqslant 2$, trouver n événements non indépendants tels que toute collection de $n - 1$ événements parmi ces n est indépendante.

4 Exercices

Espaces de probabilité

Exercice 4.1 (Premier problème du chevalier de Méré). Quel est le plus probable : jouant avec un dé, obtenir au moins une fois 6 en 4 coups, ou bien jouant avec deux dés, obtenir au moins une fois un double 6 en 24 coups ?

Exercice 4.2 (Second problème du chevalier de Méré). Le chevalier de Méré avait posé à Pascal le problème suivant : deux joueurs jouent à un jeu de hasard en plusieurs parties ; celui qui, le premier, a gagné trois parties emporte la totalité de l'enjeu. Si les joueurs doivent arrêter le jeu alors qu'il ne manque au premier joueur, pour l'emporter, qu'une partie, et au second que deux parties, comment doit-on répartir équitablement l'enjeu ?

Exercice 4.3 (Problème des anniversaires). Quelle est la probabilité pour que n personnes prises au hasard aient toutes des jours d'anniversaire différents ?

On supposera que tous les jours de naissance sont équiprobables et on ne tiendra pas compte des années bissextiles.

Déterminer la plus petite valeur de n telle que cette probabilité soit inférieure à 50%.

Exercice 4.4 (Formule de Poincaré ; problème des rencontres). a) Montrer que si $A_1, \ldots, A_n$ sont n événements d'un même espace probabilisé et si A désigne la réunion de ces n événements, on a :

$$P(A) = \sum_{k=1}^{n} (-1)^{k+1} s_k, \quad s_k = \sum_{i_1 < i_2 < \ldots < i_k} P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}).$$

b) On tire sans remise n boules numérotées de 1 à n. Déterminer la probabilité p_n pour qu'il existe un entier k tel que la boule portant le numéro k soit tirée au tirage numéro k.

c) Déterminer la limite de p_n quand n tend vers l'infini.

Exercice 4.5 (Bridge). Calculer la probabilité qu'au bridge, chaque joueur ait un as.

Exercice 4.6 (Balles et paniers). On a r balles et n paniers numérotés de 1 à n.

On répondra aux questions dans les deux cas suivants :

(a) Les r balles sont discernables (par exemple parce qu'elles sont de couleurs différentes).

(b) Les r balles sont indiscernables.

Question 1 : Quel est le nombre de répartitions possibles (un panier peut contenir plusieurs balles) ?

Question 2 : Quelle est la probabilité p_k qu'un panier donné contienne exactement k balles. Étudier la monotonie de la suite $(p_k)_{0 \leqslant k \leqslant r}$.

Question 3 : On suppose que n et r tendent vers l'infini et que r/n tend vers λ. Montrer que chaque terme p_k admet une limite et calculer celle-ci.

Exercice 4.7 (Problème du scrutin). Lors d'un vote opposant deux candidats A et B, A obtient a voix et B obtient b voix. On suppose que $a < b$. Quelle est la probabilité pour qu'au cours du dépouillement, B ait toujours été strictement en tête ?

On pourra représenter le dépouillement par un chemin du plan constitué de segments horizontaux ou verticaux de longueur 1 joignant l'origine au point de coordonnées (a, b) et compter le nombre de chemins situés au dessus de la diagonale.

Exercice 4.8 (Milieu). Soit S l'ensemble des points de l'espace dont les coordonnées x, y et z sont entières et vérifient $0 \leqslant x \leqslant 2$, $0 \leqslant y \leqslant 3$ et $0 \leqslant z \leqslant 4$. Deux points sont choisis au hasard dans S, c'est à dire uniformément dans $\mathcal{P}_2(S)$. Quelle est la probabilité que le milieu du segment qu'ils déterminent appartiennent à S ?

Exercice 4.9 (Table). On dispose $n \geqslant 3$ personnes autour d'une table ronde. Trois personnes distinctes sont choisies au hasard. Calculer la probabilité de l'événement « au moins deux parmi les trois étaient assises l'une à coté de l'autre. »

Exercice 4.10 (Journal). Un responsable de jeux dispose de trois enveloppes de couleurs différentes. Il met de l'argent dans l'une, et du papier journal dans les deux autres. Il fait entrer un joueur et lui fait choisir une enveloppe qu'il garde fermée. Parmi les deux enveloppes restantes, il y en a toujours au moins une qui contient du papier journal. Le responsable ouvre alors une de ces deux enveloppes dont il sait qu'elle contient du papier journal, et propose au joueur de changer l'enveloppe qu'il a en main contre celle qui reste. Le joueur a t-il intérêt à changer ?

Probabilités conditionnelles et indépendance

Exercice 4.11 (Écrous). Dans une usine d'écrous, trois machines A, B et C produisent respectivement 25%, 35% et 40% du total de la production. Elles produisent respectivement 5%, 4% et 2% de pièces défectueuses. Un écrou est tiré au hasard et s'avère défectueux. Quelle est la probabilité qu'il ait été produit par la machine A ? B ? ou C ?

Exercice 4.12 (Sexe). M. Dupont a deux enfants dont une fille, quelle est la probabilité que l'autre soit un garçon ? M. Durand a deux enfants dont le plus jeune est une fille, quelle est la probabilité pour que l'autre soit un garçon ?

Exercice 4.13 (Le rouge et le noir). Une urne contient n boules noires et n boules rouges. On tire deux par deux sans remise, toutes les boules de l'urne. Quelle est la probabilité d'obtenir à chaque tirage deux boules de couleurs différentes ?

Exercice 4.14 (Urnes). On dispose de $N + 1$ urnes numérotées de 0 à N. L'urne numéro k contient k boules rouges et $N - k$ boules noires. On tire une des urnes avec équiprobabilité, puis on procède avec cette urne à une série de n tirages avec remise.

a) Calculer la probabilité d'avoir choisi l'urne numéro 1 sachant qu'on a tiré n boules rouges.

b) Calculer la probabilité de tirer n boules rouges.

c) Calculer la probabilité de tirer une boule rouge au tirage $n+1$ sachant qu'on a déjà tiré n boules rouges.

d) Déterminer les limites des probabilités précédentes quand $N \to +\infty$.

Exercice 4.15 (Un peu d'arithmétique). Pour tout entier $n \geqslant 2$ fixé, soit P_n la probabilité uniforme sur l'ensemble $\{1, 2, \ldots, n\}$. Pour tout diviseur m de n désignons par A_m le sous-ensemble de $\{1, 2, \ldots, n\}$ formé des multiples de m.

Montrer que $P(A_m) = 1/m$.

Montrer que les A_p où p parcourt les diviseurs premiers de n, sont des événements indépendants dans l'espace probabilisé $(\{1, 2, \ldots, n\}, P_n)$. En déduire que l'ensemble des entiers de $\{1, 2, \ldots, n\}$ premiers avec n a une probabilité $\prod_p \left(1 - \frac{1}{p}\right)$ où p parcourt l'ensemble des diviseurs premiers de n et en déduire le cardinal de cet ensemble. Retrouver ainsi une formule d'Euler.

Exercice 4.16 (Un peu de génétique). Les gènes (unités d'hérédité) se présentent dans les cas les plus simples en paires et sous deux formes, appelées allèles, et notées A et a. Chaque individu possède un des trois génotypes AA, aa (qui sont homozygotes) et Aa (qui est hétérozygote).

Chaque individu reçoit au hasard un gène de chacun des génotypes de ses parents, chacun des gènes de ces génotypes ayant la probabilité $\frac{1}{2}$ de passer à l'enfant. Par exemple, deux parents Aa donneront à leur enfant le génotype AA avec probabilité $\frac{1}{4}$, Aa avec probabilité $\frac{1}{2}$ et aa avec probabilité $\frac{1}{4}$.

Les génotypes des parents de la génération 0 sont supposés indépendants. La probabilité qu'un parent de la génération 0 ait AA (respectivement Aa, respectivement aa) comme génotype est notée u (resp. $2v$, resp. w), donc $u + 2v + w = 1$ et $(u, 2v, w)$ est appelé la fréquence des génotypes.

On note $p = u + v$ et $q = v + w$ la fréquence des allèles A et a.

a) Montrer qu'un individu de la première génération appartient à l'un des génotypes AA, Aa, aa avec les probabilités respectives $u_1 = p^2$, $2v_1 = 2pq$ et $w_1 = q^2$. En conclure que passé la première génération, la loi de probabilité u, v, w des génotypes des individus d'une population doit vérifier sous les hypothèses précédentes la relation de Hardy-Weinberg : $v^2 = uw$. Montrer que dans ce cas, $u = p^2$, $2v = 2pq$ et $w = q^2$.

b) On considère dans les questions suivantes que $u = p^2$, $2v = 2pq$ et $w = q^2$.

Calculer la probabilité que le génotype d'un individu de la première génération soit Aa sachant que son frère a le même génotype.

c) Calculer pour i et j dans $\mathcal{G} = \{AA, Aa, aa\}$, la probabilité $p_{i,j}$ que le génotype d'un enfant soit j sachant que le génotype de son père est i. On note P la matrice associée, indexée par $\mathcal{G} \times \mathcal{G}$.

d) Montrer que la probabilité que le génotype d'un individu de la seconde génération soit j sachant que le génotype de son grand-père paternel est i est donnée par

$$p_{i,j}^{(2)} = \sum_{k \in \mathcal{G}} p_{i,k} p_{k,j}$$

Plus généralement, montrer que la probabilité que le génotype d'un individu de la génération n soit j sachant que le génotype de son ancêtre masculin de la génération 0 était i est donnée par le coeficient (i, j) de la matrice P^n. Calculer P^n.

Exercice 4.17 (Urne de Pólya). Soient $a \geqslant 0$, $b \geqslant 0$ et $c \geqslant 0$ des entiers avec $a+b \geqslant 1$. Une urne contient a boules noires et b boules rouges. Si on tire une boule, on remet dans l'urne c boules de la couleur de la boule tirée. (Le cas du tirage avec remise simple est donnée par $c = 1$ et celui du tirage sans remise par $c = 0$).

a) Calculer la probabilité qu'au deuxième tirage, on tire une boule noire.

b) Calculer la probabilité qu'au troisième tirage, on tire une boule noire.

c) On note X_i la variable aléatoire valant 1 si on tire une boule noire au tirage numéro i et 0 sinon. Que représente la variable aléatoire $S_i = X_1 + \cdots + X_i$?

En utilisant la formule des probabilités totales montrer que

$$P(X_{i+1} = 1) = (E[S_i] + a)/(a + b + i).$$

En déduire $P(X_i = 1)$.

Exercice 4.18 (Trois). Anne, Barbara et Cléo jouent une suite de parties en suivant la règle suivante : chaque partie oppose deux joueuses et la perdante de la partie s'efface à la fin de la partie pour laisser la place aux deux autres joueuses. Les parties sont supposées indépendantes et chaque joueuse a la même probabilité à chaque fois de gagner. Le jeu s'arrête dès qu'une joueuse a gagné deux parties consécutives ; cette joueuse est alors la gagnante du jeu.

Déterminer la probabilité de chacune des trois joueuses de gagner si Anne et Barbara jouent la première partie. Pour tout entier n, calculer la probabilité que le jeu s'arrête au bout de n parties exactement.

Exercice 4.19 (QCM). Dans un QCM, il y a m réponses possibles. Un candidat a une probabilité p de connaître la réponse à une question prise au hasard parmi un ensemble fini de questions.

Sachant que le candidat a répondu juste à la question, quelle est la probabilité qu'il connaisse effectivement la réponse ? On suppose qu'un candidat qui ne connaît pas la réponse, répond au hasard et donc que les m réponses sont équiprobables.

Fiche 3 : Probabilités (2) Variables aléatoires discrètes

« Il est souvent assez difficile de faire des prédictions, surtout en ce qui concerne le futur. » Apocryphe, parfois attribué à Niels Bohr

On convient que « dénombrable » signifie « fini » ou « dénombrable infini. »

La première partie du chapitre est consacrée à des rappels, sans doute en partie indigestes, sur ce qu'il est licite et illicite de faire avec des sommes infinies ; rien de spécifiquement probabiliste donc, mais des résultats à connaître en particulier pour faire des probabilités. On introduit ensuite les variables aléatoires, leur loi, leurs moments et la propriété d'indépendance, le tout dans le cas discret.

1 Familles sommables

On rappelle que toute partie A de $\mathbb{R}$ non vide et majorée admet une borne supérieure, notée $\sup A$; c'est le plus petit des majorants de A. Si A n'est pas majorée, on note $\sup A = +\infty$.

On se donne un ensemble dénombrable I et une famille $(x_i)_{i\in I}$ de nombres réels **positifs**.

Définition 1.1. *Soit $A \subset \mathbb{R}^+$ l'ensemble des sommes finies $\sum_{i\in J} x_i$ avec J partie finie de I. On pose $\sum_{i\in I} x_i := \sup A$. Si A est borné, on dit que la famille $(x_i)_{i\in I}$ est sommable.*

La somme $\sum_{i\in I} x_i$ est donc un élément de $\mathbb{R}^+ \cup \{+\infty\}$.

Dans le cas particulier où I est vide, on pose $\sum_{i\in I} x_i := 0$, ce qui correspond au supremum de $\emptyset$ dans $\mathbb{R}^+$.

Si $(y_i)_{i\in I}$ est une seconde famille de réels positifs telle que pour tout i dans I, $x_i \leqslant y_i$, alors $\sum_{i\in I} x_i \leqslant \sum_{i\in I} y_i$. Cas particulier : si $J \subset I$, alors $\sum_{i\in J} x_i \leqslant \sum_{i\in I} x_i$.

On explique maintenant comment calculer la valeur d'une somme infinie à partir de sommes finies.

Proposition 1.2. *Soit $(J_n)_{n\geqslant 0}$ une famille croissante de parties de I, finies ou infinies, dont la réunion est égale à I. Alors*

$$\sum_{i\in I} x_i = \lim_{n\to+\infty} \sum_{i\in J_n} x_i.$$

Il suffit donc de faire le calcul pour **une** suite croissante $(J_n)_n$ de réunion I, et le résultat sera toujours le même.

Preuve : On doit montrer que $S = \ell$ où S désigne la somme sur I tout entier et ℓ la limite du membre de droite. Pour tout n, $\displaystyle\sum_{i\in J_n} x_i \leqslant S$. Comme la suite de terme général $\displaystyle\sum_{i\in J_n} x_i$ est croissante, en passant à la limite on obtient $\ell \leqslant S$.

Dans l'autre sens, soit $J \subset I$ une partie finie. À tout élément j de J on peut associer un entier $n(j)$ tel que $j \in J_{n(j)}$. En posant $N = \max_{j\in J} n(j)$, on obtient $J \subset J_N$ et

$$\sum_{i\in J} x_i \leqslant \sum_{i\in J_N} x_i \leqslant \ell,$$

ce qui termine la preuve.

Remarque 1.3. *Si $I = \mathbb{N}$, on obtient* $\displaystyle\sum_{i\in I} x_i = \sum_{n=0}^{+\infty} x_n = \lim_{n\to+\infty} \sum_{k=0}^{n} x_n.$

Désormais, $(x_i)_{i\in I}$ désigne une famille de nombres réels **pas forcément positifs**. Pour tout nombre réel x, on note x^+ sa partie positive et x^- sa partie négative, définies par

$$x^+ := \max(x, 0), \quad x^- := \max(-x, 0),$$

ou bien, ce qui est équivalent, par les deux relations

$$x = x^+ - x^-, \quad |x| = x^+ + x^-.$$

On remarquera que $x^+ \geqslant 0$ et $x^- \geqslant 0$ pour tout nombre réel x.

Définition 1.4. *La famille $(x_i)_{i\in I}$ est sommable si et seulement si la famille $(|x_i|)_{i\in I}$ d'éléments de $\mathbb{R}^+$ est sommable. Alors, comme $x_i^+ \leqslant |x_i|$ et $x_i^- \leqslant |x_i|$, les familles $(x_i^+)_{i\in I}$ et $(x_i^-)_{i\in I}$ sont également sommables. On peut donc poser*

$$\sum_{i\in I} x_i := \sum_{i\in I} x_i^+ - \sum_{i\in I} x_i^-.$$

Si les deux séries $\displaystyle\sum_{i\in I} x_i^+$ et $\displaystyle\sum_{i\in I} x_i^-$ n'étaient pas convergentes, la différence de leurs sommes n'aurait tout simplement pas de sens.

Proposition 1.5. *Soit $(x_i)_{i\in I}$ une famille sommable et $(J_n)_{n\geq 0}$ une famille croissante de parties de I, finies ou infinies, dont la réunion est égale à I. Alors*

$$\sum_{i\in I} x_i = \lim_{n\to+\infty} \sum_{i\in J_n} x_i.$$

Preuve : Comme chaque famille $(x_i)_{i\in J_n}$ est sommable,

$$\sum_{i\in J_n} x_i = \sum_{i\in J_n} x_i^+ - \sum_{i\in J_n} x_i^-,$$

et il suffit de passer à la limite en utilisant la proposition 1.2. Fin de la preuve.

Remarque 1.6. *Si $I = \mathbb{N}$, la famille $(x_n)_{n\in\mathbb{N}}$ est sommable si et seulement si la série $\sum_{n=0}^{+\infty} x_n$ est absolument convergente.*

En appliquant ce qui précède à des ensembles J_n qui sont finis, on obtient

$$\left|\sum_{i\in J_n} x_i\right| \leqslant \sum_{i\in J_n} x_i^+ + \sum_{i\in J_n} x_i^- = \sum_{i\in J_n} |x_i|.$$

D'où le résultat suivant :

Proposition 1.7. *Si la famille $(x_i)_{i\in I}$ est sommable, alors*

$$\left|\sum_{i\in I} x_i\right| \leqslant \sum_{i\in I} |x_i|.$$

Voici d'autres propriétés.

Proposition 1.8 (Linéarité). *Soit $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ deux familles sommables. Soit a et b deux nombres réels et, pour tout $i \in I$, posons $z_i := ax_i + by_i$. Alors la famille $(z_i)_{i\in I}$ est sommable et*

$$\sum_{i\in I} z_i = a\sum_{i\in I} x_i + b\sum_{i\in I} y_i.$$

Preuve : Si J est une partie finie de I,

$$\sum_{i\in J} |z_i| \leqslant |a|\sum_{i\in J} |x_i| + |b|\sum_{i\in J} |y_i| \leqslant |a|\sum_{i\in I} |x_i| + |b|\sum_{i\in I} |y_i|,$$

donc la famille $(z_i)_{i\in I}$ est sommable. Si $(J_n)_{n\geq 0}$ est une suite croissante de parties finies de I de réunion égale à I,

$$\sum_{i\in J_n} z_i = a\sum_{i\in J_n} x_i + b\sum_{i\in J_n} y_i,$$

donc on obtient le résultat par passage à la limite. Fin de la preuve.

Proposition 1.9 (Sommation par paquets). *Soit $(I_k)_{k\in K}$ une partition de I, donc les ensembles I_k sont deux à deux disjoints et leur réunion est égale à I. Soit $(x_i)_{i\in I}$ une famille sommable. Alors,*
1) Pour tout k dans K, la famille $(x_i)_{i\in I_k}$ est sommable.
2) Soit $S_k = \sum_{i\in I_k} x_i$. La famille $(S_k)_{k\in K}$ est sommable et

$$\sum_{k\in K} S_k := \sum_{i\in I} x_i.$$

Réciproquement si pour tout $k \in K$, la famille $(x_i)_{i\in I_k}$ est sommable et si la famille $(T_k)_{k\in K}$ est sommable où on a noté, pour chaque k, $T_k := \sum_{i\in I_k} |x_i|$, alors la famille $(x_i)_{i\in I}$ est sommable.

Preuve : 1) Si $J \subset I_k$ est fini,

$$\sum_{i\in J} |x_i| \leqslant \sum_{i\in I} |x_i|,$$

donc $\sum_{i\in I_k} |x_i| \leqslant \sum_{i\in I} |x_i|$, qui est fini.
2) Pour tout $k \in K$, introduisons une suite croissante $(I_{k,n})_{n\geqslant 0}$ de parties finies de I_k dont la réunion est I_k. On pose $S_{k,n} = \sum_{i\in I_{k,n}} x_i$. On sait par la proposition 1.2 que $S_{k,n}$ tend vers S_k quand n tend vers l'infini.
Pour tout $F \subset K$ fini et pour tout n,

$$\sum_{k\in F} |S_{k,n}| \leqslant \sum_{k\in F} \sum_{i\in I_{k,n}} |x_i| = \sum_{i\in F_n} |x_i| \leqslant \sum_{i\in I} |x_i|,$$

où F_n désigne la réunion des $I_{k,n}$ pour $k \in F$. En faisant tendre n vers $+\infty$,

$$\sum_{k\in F} |S_k| \leqslant \sum_{i\in I} |x_i|, \quad \sum_{k\in K} |S_k| \leqslant \sum_{i\in I} |x_i|,$$

donc la famille $(S_k)_{k\in K}$ est sommable. De plus,

$$\sum_{k\in F} S_{k,n} = \sum_{k\in F} \sum_{i\in I_{k,n}} x_i = \sum_{i\in F_n} x_i.$$

Or la suite $(F_n)_{n\geq 0}$ de parties de I est croissante et sa réunion vaut la réunion des I_k pour $k \in F$, notée I_F, donc

$$\sum_{k\in F} S_k = \sum_{i\in I_F} x_i.$$

En considérant maintenant une suite croissante de parties finies F de K, on obtient le résultat par la proposition 1.2.

Pour la réciproque, si J est une partie finie de I, $J \cap I_k$ est une partie finie de I_k donc $\sum_{i \in J \cap I_k} |x_i| \leqslant T_k$. De plus, l'ensemble K_J des indices k dans K tels que $J \cap I_k$ n'est pas vide est fini, donc $\sum_{i \in J} |x_i| = \sum_{k \in K_J} \sum_{i \in J \cap I_k} |x_i| \leqslant \sum_{k \in K} T_k$, qui est fini. Fin de la preuve.

Corollaire 1.1 (Théorème de Fubini). *Soient I et J deux ensembles dénombrables et $(x_{i,j})_{(i,j) \in I \times J}$ une famille de nombres réels.*
1) Supposons que, pour tout (i,j), $x_{i,j} \geqslant 0$. Alors

$$\sum_{(i,j) \in I \times J} x_{i,j} = \sum_{i \in I} \left(\sum_{j \in J} x_{i,j} \right) = \sum_{j \in J} \left(\sum_{i \in I} x_{i,j} \right), \tag{1}$$

avec la convention que, pour tout réel x, $x + (+\infty)$ et $(+\infty) + (+\infty)$ valent $+\infty$.
2) La famille $(x_{i,j})_{(i,j) \in I \times J}$ est sommable si et seulement si, pour tout $i \in I$, $\sum_{j \in J} |x_{i,j}|$ est fini et si $\sum_{i \in I} \left(\sum_{j \in J} |x_{i,j}| \right)$ est fini. L'égalité (1) est alors valide.

Preuve : 1) Si $\sum_{(i,j) \in I \times J} x_{i,j}$ est finie, il suffit d'appliquer la sommation par paquets à la partition $I \times J = \bigcup_{i \in I} (\{i\} \times J)$.

Si $\sum_{(i,j) \in I \times J} x_{i,j}$ est infinie, soit il existe $i \in I$ tel que $\sum_{j \in J} x_{i,j}$ est infinie et dans ce cas l'égalité est vraie, soit pour tout $i \in I$, $\sum_{j \in J} x_{i,j}$ est finie mais alors le théorème de sommation par paquets entraîne que $\sum_{i \in I} \left(\sum_{j \in J} x_{i,j} \right)$ est infinie.
2) Sommation par paquets pour la partition $I \times J = \bigcup_{i \in I} (\{i\} \times J)$. Fin de la preuve.

2 Variables aléatoires

2.1 Flèches

Définition 2.1. *Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité.*
Une variable aléatoire réelle est une application $X : \Omega \to \mathbb{R}$ telle que, pour tout nombre réel x, l'ensemble $\{\omega \in \Omega \,;\, X(\omega) \leqslant x\}$ appartient à $\mathcal{F}$.
Une variable aléatoire vectorielle est une application $X = (X_1, \ldots, X_d) : \Omega \to \mathbb{R}^d$ telle

que, pour tout i, X_i est une variable aléatoire réelle.
On dit que la variable aléatoire X est discrète quand $X(\Omega)$ est dénombrable.

Quelques remarques et notations.

Si $\mathcal{F} = \mathcal{P}(\Omega)$, toute application de Ω dans $\mathbb{R}$ ou $\mathbb{R}^d$ est une variable aléatoire.

Si Ω est dénombrable, toute variable aléatoire est discrète.

On admettra le résultat suivant :

Proposition 2.2. *Si $f : \mathbb{R}^d \to \mathbb{R}^n$ est continue et si $X : \Omega \to \mathbb{R}^d$ est une variable aléatoire, alors $f \circ X$ noté $f(X)$ est une variable aléatoire, à valeurs dans $\mathbb{R}^n$. Par exemple une somme et un produit de variables aléatoires sont des variables aléatoires.*

Pour une variable aléatoire X on notera $\{X = x\}$, $\{X \leqslant x\}$ et $\{X < x\}$ les ensembles $\{\omega \in \Omega\,;\, X(\omega) = x\}$, $\{\omega \in \Omega\,;\, X(\omega) \leqslant x\}$ et $\{\omega \in \Omega\,;\, X(\omega) < x\}$. De façon générale, on notera $\{X \in F\}$ l'ensemble $\{\omega \in \Omega\,;\, X(\omega) \in F\}$.

Quelques propriétés basées sur ces ensembles.

Proposition 2.3. *1) Soit X une application de Ω dans $\mathbb{R}$ telle que $X(\Omega)$ est dénombrable. Alors X est une variable aléatoire si et seulement si $\{X = x\} \in \mathcal{F}$ pour tout $x \in X(\Omega)$.*

2) Si X est une variable aléatoire discrète et $f : \mathbb{R}^d \to \mathbb{R}^n$ une application quelconque, alors $f(X)$ est une variable aléatoire discrète.

Pour la partie 1), on remarque que, pour tout x, $\{X = x\}$ est l'intersection dénombrable de $\{X \leqslant x\}$ et des $\{X \leqslant x - 1/n\}^c$ pour $n \geqslant 1$, et $\{X \leqslant x\}$ est la réunion dénombrable des $\{X = y\}$ pour tous les y dans $X(\Omega)$ tels que $y \leqslant x$.

On admet la partie 2).

Attention ! Dans le cas général non discret, le fait que $\{X = x\}$ soit mesurable pour tout x n'implique pas que X est une variable aléatoire.

2.2 Fonctions de répartition

Définition 2.4. *Soit X une variable aléatoire réelle. On appelle fonction de répartition de X la fonction $F_X : \mathbb{R} \to [0, 1]$ définie pour tout réel x par*

$$F_X(x) = P(X \leqslant x).$$

Proposition 2.5. *1) F_X est une fonction croissante et continue à droite.*
2) La limite de F_X en $-\infty$ vaut 0.
3) La limite de F_X en $+\infty$ vaut 1.
4) Pour tout réel x, $F_X(x-) = P(X < x)$.
5) Pour tout réel x, $P(X = x) = F_X(x) - F_X(x-)$.

D'après 5), si l'on connaît la fonction de répartition d'une variable aléatoire à valeurs entières, on connaît $P(X = n)$ pour tout entier n.

Preuve de la proposition :

1) Si $x \leqslant y$, $\{X \leqslant x\} \subset \{X \leqslant y\}$. Par conséquent, $P(X \leqslant x) \leqslant P(X \leqslant y)$.

Pour le reste de la preuve, soit x un nombre réel, $(x_n)_{n\geqslant 0}$ une suite de nombres réels et $A_n := \{X \leqslant x_n\}$.

Si $(x_n)_{n\geqslant 0}$ décroît vers x, la suite $(A_n)_{n\geqslant 0}$ est décroissante et $\{X \leqslant x\}$ est l'intersection des A_n, donc $P(X \leqslant x) = \lim_n P(X \leqslant x_n)$.

2) Si $(x_n)_{n\geqslant 0}$ décroît vers $-\infty$, la suite $(A_n)_{n\geqslant 0}$ est décroissante et l'intersection des A_n est vide, donc $\lim_n P(X \leqslant x_n) = 0$.

3) Si $(x_n)_{n\geqslant 0}$ croît vers $+\infty$, la suite $(A_n)_{n\geqslant 0}$ est croissante et la réunion des A_n est Ω tout entier, donc $\lim_n P(X \leqslant x_n) = 1$.

4) Si $(x_n)_{n\geqslant 0}$ croît strictement vers x, la suite $(A_n)_{n\geqslant 0}$ est croissante et $\{X < x\}$ est la réunion des A_n, donc $P(X < x) = \lim_n P(X \leqslant x_n)$.

5) $P(X = x) = P(X \leqslant x) - P(X < x)$. Fin de la preuve.

2.3 Variables aléatoires et lois discrètes

Définition 2.6. *Soit X une variable aléatoire discrète à valeurs dans $\mathbb{R}^d$ définie sur $(\Omega, \mathcal{F}, P)$. La loi de X (sous P) est la probabilité sur $(\mathbb{R}^d, \mathcal{P}(\mathbb{R}^d))$, notée P_X et définie par la relation $P_X(B) = P(X \in B)$ pour tout $B \subset \mathbb{R}^d$.*

On note parfois aussi $X(P)$ la loi P_X de X.

Si $X(\Omega) = \{x_i\,;\, i \in I\}$ avec I dénombrable, la probabilité P_X est caractérisée par la donnée des nombres $P_X(\{x_i\}) = P(X = x_i)$.

En effet, pour $B \subset \mathbb{R}^d$, $P_X(B) = \sum_{i\in I(B)} P_X(\{x_i\})$ où $I(B)$ désigne l'ensemble des i dans I tels que x_i appartient à B, et cette somme est bien définie.

Si B est une partie de $\mathbb{R}^d$ qui contient $X(\Omega)$, alors la restriction de P_X à $(B, \mathcal{P}(B))$ est une probabilité sur $(B, \mathcal{P}(B))$.

Exemple 2.7. *On lance deux dés et on s'intéresse au plus grand nombre obtenu. On peut considérer pour cela $\Omega = \{1,2,3,4,5,6\}^2$, $\mathcal{F} = \mathcal{P}(\Omega)$ et $P(\{(i,j)\}) = 1/36$. Soit $X : \Omega \to \{1,2,3,4,5,6\}$ définie par $X(i,j) = \max(i,j)$. La loi de X est une probabilité sur $\{1,2,3,4,5,6\}$ d'ecrite par les nombres $p_i = P(X = i)$, avec*

$$p_1 = 1/36, \quad p_2 = 3/36, \quad p_3 = 5/36, \quad p_4 = 7/36, \quad p_5 = 9/36, \quad p_6 = 11/36.$$

On vérifie que $p_1 + p_2 + \cdots + p_6 = 1$.

Proposition 2.8. *Soit X une variable aléatoire discrète et $f : \mathbb{R}^d \to \mathbb{R}^n$. La loi de $Y = f(X)$ est donnée par*

$$P_Y(\{y\}) = \sum_{i \in I(y)} P_X(\{x_i\}), \qquad I(y) := \{i \in I \,;\, f(x_i) = y\}.$$

Exemple 2.9 (Suite). *On reprend l'exemple précédent et on considère la variable aléatoire Y qui vaut $Y = 1$ si $X = 1$ ou $X = 6$ et $Y = -1$ sinon. Ceci modélise le jeu suivant : on gagne 1 € si on obtient au moins un 6 ou un double 1 et on perd 1 € sinon.*

2.4 Espérance

Définition 2.10. *Soit X une variable aléatoire discrète réelle définie sur $(\Omega, \mathcal{F}, P)$ à valeurs dans $X(\Omega) = \{x_i \,;\, i \in I\}$ avec I dénombrable. Pour tout i dans I, on note $p_i = P_X(\{x_i\}) = P(X = x_i)$.*
Si la famille $(x_i p_i)_{i \in I}$ est sommable, on dit que X admet une espérance, notée $E[X]$ et définie par

$$E[X] := \sum_{i \in I} x_i p_i.$$

Quelques remarques.

Si $X(\Omega)$ est fini, X admet toujours une espérance.

Si X ne prend que des valeurs positives alors $E[X] \geqslant 0$.

Si X est une variable aléatoire constante égale à x, alors $E[X] = x$.

Si a et b sont des nombres réels et si X admet une espérance, alors $aX + b$ aussi et $E[aX + b] = aE[X] + b$. Plus géneralement, on dispose du résultat suivant.

Proposition 2.11 (Formule fondamentale). *Soit X une variable aléatoire discrète à valeurs dans $\mathbb{R}^d$ définie sur $(\Omega, \mathcal{F}, P)$. Notons $X(\Omega) = \{x_i \,;\, i \in I\}$ avec I dénombrable, l'ensemble des valeurs prises par X et, pour tout i dans I, $p_i = P_X(\{x_i\}) = P(X = x_i)$. Soit $f : \mathbb{R}^d \to \mathbb{R}$ et notons $f(X(\Omega)) = \{y_j \,;\, j \in J\}$ avec J dénombrable. Enfin, soit $Y = f(X)$.*

Si la famille $(f(x_i)p_i)_{i \in I}$ est sommable, alors Y admet une espérance et

$$E[Y] = E[f(X)] = \sum_{i \in I} f(x_i) p_i = \sum_{j \in J} y_j \, P_Y(\{y_j\}).$$

Preuve : Notons $q_j = P_Y(\{y_j\}) = \sum_{i \in I(j)} p_i$ avec $I(j) := \{i \in I \,;\, f(x_i) = y_j\}$.
Il faut d'abord montrer que la somme $\sum_{j \in J} |y_j| q_j$ converge. Pour cela, on utilise la formule

de Fubini pour la suite $(a_{i,j})_{i,j}$ définie par

$$a_{i,j} = f(x_i)p_i \text{ si } f(x_i) = y_j, \qquad a_{i,j} = 0 \text{ sinon.}$$

Alors, pour tout $i \in I$, $\sum_{j\in J} |a_{i,j}| = |f(x_i)|p_i$. Le théorème de Fubini permet de conclure.

Remarque : Si Ω est dénombrable, toute variable aléatoire X sur Ω est discrète et si X admet une espérance,

$$E[X] = \sum_{\omega\in\Omega} X(\omega)\, P(\{\omega\}).$$

Pour le voir, il suffit d'appliquer la proposition précédente à $f = X$ et à l'application identité de Ω dans Ω.

Proposition 2.12 (et définition). *Soient X et Y deux variables aléatoires réelles discrètes. On note $X(\Omega) = \{x_i\,;\, i \in I\}$ et $Y(\Omega) = \{y_j\,;\, j \in J\}$ avec I et J dénombrables. La loi de (X, Y) est donnée par $p_{i,j} = P((X, Y) = (x_i, y_j))$. Alors les lois de X et Y sont appelées les lois marginales de (X, Y) et sont données par*

$$p_i = P(X = x_i) = \sum_{j\in J} p_{i,j}, \qquad q_j = P(Y = y_j) = \sum_{i\in I} p_{i,j}.$$

Preuve : Les événements $\{Y = y_j\}$ qui sont non vides forment une partition de Ω, donc, pour tout i fixé, les événements $\{X = x_i, Y = y_j\}$ qui sont non vides forment une partition de $\{X = x_i\}$. On en déduit

$$P(X = x_i) = \sum_{j\in J} P(\{X = x_i\} \cap \{Y = y_j\}),$$

et on a la même chose pour l'autre formule. Fin de la preuve.

Exemple 2.13. *On lance deux dés. Soit X le minimum des deux nombres obtenus et Y leur maximum. On peut représenter la loi de (X, Y) dans le tableau suivant :*

$y\backslash x$	1	2	3	4	5	6	$P(Y = y)$
1	1/36	0	0	0	0	0	1/36
2	2/36	1/36	0	0	0	0	3/36
3	2/36	2/36	1/36	0	0	0	5/36
4	2/36	2/36	2/36	1/36	0	0	7/36
5	2/36	2/36	2/36	2/36	1/36	0	9/36
6	2/36	2/36	2/36	2/36	2/36	1/36	11/36
$P(X = x)$	11/36	9/36	7/36	5/36	3/36	1/36	

Proposition 2.14. *Soit X et Y deux variables aléatoires réelles discrètes admettant des espérances. Alors $X + Y$ admet une espérance et*

$$E[X + Y] = E[X] + E[Y].$$

Preuve : Gardons les notations de la proposition précédente. On utilise la formule fondamentale avec la variable aléatoire (X, Y) qui prend les valeurs (x_i, y_j) quand i décrit I et j décrit J, et avec la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $(x, y) \mapsto x + y$.

Pour tout $i \in I$, $\sum_{j\in J} |x_i|p_{i,j} = |x_i|p_i$ et la somme $\sum_{i\in I} |x_i|p_i$ converge car X admet une espérance, donc par le théorème de Fubini, la famille $(x_i p_{i,j})_{(i,j)\in I\times J}$ est sommable. Il en est de même pour la famille $(y_j p_{i,j})_{(i,j)\in I\times J}$.

Donc par linéarité la famille $((x_i + y_j)p_{i,j})_{(i,j)\in I\times J}$ est sommable et sa somme vaut

$$\sum_{(i,j)\in I\times J} x_i p_{i,j} + \sum_{(i,j)\in I\times J} y_j p_{i,j} = \sum_{i\in I} x_i p_i + \sum_{j\in J} y_j q_j = E[X] + E[Y].$$

Fin de la preuve.

Corollaire 2.1. *Soit X et Y deux variables aléatoires discrètes telles que pour tout $\omega \in \Omega$, $X(\omega) \leqslant Y(\omega)$. Si X et Y admettent des espérances alors $E[X] \leqslant E[Y]$.*

Il suffit de voir que la variable aléatoire $Y - X$ ne prend que des valeurs positives et que $E[Y - X] = E[Y] - E[X]$.

Si X et Y ne prennent que des valeurs positives, si $X \leqslant Y$ et si Y admet une espérance, alors X aussi.

2.5 Variance et covariance

Définition 2.15. *La variable aléatoire X est de carré intégrable si $E[X^2]$ est fini. On a alors*

$$E[X^2] = \sum_{i\in I} x_i^2 p_i.$$

Lemme 2.1. *Si X est de carré intégrable et si $E[X^2] = 0$, alors $X = 0$ partout.*

Preuve laissée au lecteur.

Proposition 2.16 (Inégalité de Cauchy-Schwarz). *Soit X et Y deux variables aléatoires réelles discrètes de carré intégrable. Alors la variable aléatoire XY admet une espérance et*

$$|E[XY]| \leqslant E[X^2]^{1/2} E[Y^2]^{1/2}.$$

De plus, $X + Y$ est de carré intégrable.

Preuve : Pour tout $\omega \in \Omega$, $2|X(\omega)Y(\omega)| \leqslant X^2(\omega) + Y^2(\omega)$. On en déduit par un résultat précédent que la variable aléatoire XY admet une espérance et que, pour tout réel t, la variable aléatoire $(X + tY)^2 = X^2 + 2tXY + t^2Y^2$ aussi. Comme $(X + tY)^2$ est positive partout, $E[(X + tY)^2] \geqslant 0$.

On développe et on traduit le fait que le polynome en t obtenu est toujours positif, ce qui nous donne le résultat. Fin de la preuve.

Preuve alternative : Une autre méthode consiste à considérer la variable aléatoire

$$Z := E[Y^2]^{1/2}X - E[X^2]^{1/2}Y.$$

Comme $Z^2 \leqslant 2E[Y^2]X^2 + 2E[X^2]Y^2$, Z est de carré intégrable et $E[Z^2] \geqslant 0$. En développant Z^2 et en réarrangeant les termes, on obtient

$$2E[Y^2]E[X^2] \geqslant 2E[X^2]^{1/2}E[Y^2]^{1/2}E[XY].$$

Si on peut diviser les deux membres par $2E[X^2]^{1/2}E[Y^2]^{1/2}$, on obtient

$$E[XY] \leqslant E[X^2]^{1/2}E[Y^2]^{1/2},$$

et en considérant $T := E[Y^2]^{1/2}X + E[X^2]^{1/2}Y$, on démontre l'autre inégalité.

Sinon, cela signifie que $E[X^2] = 0$ ou $E[Y^2] = 0$, donc que $X = 0$ partout ou $Y = 0$ partout, auquel cas tout fonctionne. Cette méthode fournit aussi les cas d'égalité puisque $E[Z^2] = 0$ ou $E[T^2] = 0$ implique que $Z = 0$ partout ou $T = 0$ partout, donc X et Y sont proportionnelles. Fin de la preuve alternative.

Proposition 2.17 (et définition). *Soit X une variable aléatoire discrète de carré intégrable. Alors X admet une espérance et*

$$|E[X]| \leqslant E[|X|] \leqslant E[X^2]^{1/2}.$$

On définit alors la variance de X, notée $\mathrm{var}(X)$ *ou $\sigma^2(X)$, et l'écart type de X, noté $\sigma(X)$, par*

$$\mathrm{var}(X) = E[X^2] - E[X]^2 = E[(X - E[X])^2], \quad \sigma(X) = \sqrt{\mathrm{var}(X)}.$$

Pour tout réel x, $E[(X-x)^2] = \mathrm{var}(X) + (x - E[X])^2$ donc l'espérance est aussi la valeur de x qui minimise $E[(X - x)^2]$, et la valeur minimale obtenue est la variance.

Preuve : La première inégalité est une conséquence de la proposition précédente appliquée aux variables aléatoires X et $Y = 1$. Par ailleurs,

$$E[(X - x)^2] = E[X^2] - 2xE[X] + x^2 = E[X^2] - E[X]^2 + (E[X] - x)^2,$$

ce qui termine la preuve.

Proposition 2.18 (et définition). *Soient X et Y deux variables aléatoires réelles discrètes de carré intégrable. On définit la covariance de X et Y, notée* $\mathrm{cov}(X, Y)$*, par*

$$\mathrm{cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y].$$

On a alors

$$\mathrm{var}(X + Y) = \mathrm{var}(X) + \mathrm{var}(Y) + 2\mathrm{cov}(X, Y).$$

Preuve : Pour le premier point il suffit de développer le produit et d'utiliser le résultat sur l'espérance d'une somme et le fait que si t est un réel et X une variable aléatoire discrète admettant une espérance, alors $E[tX] = tE[X]$.
Pour le second point, on développe $(X+Y)^2 = X^2 + Y^2 + 2XY$, chacun de ces trois termes admet une espérance. Fin de la preuve.

Remarque : une variance est toujours positive ou nulle mais une covariance peut être positive ou nulle ou négative.

2.6 Variables aléatoires indépendantes

Définition 2.19. *Les variables aléatoires discrètes X_k, $1 \leqslant k \leqslant n$, à valeurs dans E_1,...,E_n respectivement, sont indépendantes si et seulement si, pour tout $(x_1, \ldots, x_n)$ dans $E_1 \times \ldots \times E_n$,*

$$P(\forall 1 \leqslant k \leqslant n,\, X_k = x_k) = \prod_{k=1}^{n} P(X_k = x_k).$$

Proposition 2.20. *Les variables aléatoires discrètes $(X_1, \ldots, X_n)$ à valeurs respectivement dans E_1,...,E_n sont indépendantes si et seulement si pour toutes fonctions $f_k : E_k \to \mathbb{R}$ positives,*

$$E\left[\prod_{k=1}^{n} f_k(X_k)\right] = \prod_{k=1}^{n} E[f_k(X_k)].$$

Dans ce cas, si on se donne, pour tout k, une fonction $g_k : E_k \to \mathbb{R}$ telle que $g_k(X_k)$ est intégrable et si on pose $Z := g_1(X_1) \cdots g_n(X_n)$, alors Z est intégrable et

$$E[Z] = \prod_{k=1}^{n} E[g_k(X_k)].$$

Preuve : Supposons pour simplifier que $n = 2$ et appliquons la proposition 2.11 avec $f(x,y) = f_1(x)f_2(y)$. Notons x_i pour i dans I les valeurs prises par X_1, y_j pour j dans J les valeurs prises par X_2. Notons $p_i := P(X_1 = x_i)$, $q_j := P(X_2 = y_j)$ et $p_{i,j} := P((X_1, X_2) = (x_i, y_j))$. Par hypothèse, $p_{i,j} = p_i q_j$, donc $E[f(X_1, X_2)]$ vaut

$$\sum_{i,j} f_1(x_i) f_2(y_j) p_{i,j} = \sum_i f_1(x_i) p_i \sum_j f_2(y_j) q_j = E[f_1(X_1)] E[f_2(X_2)].$$

La réciproque s'obtient en considérant les applications $f_1 := \mathbf{1}_{\{x_i\}}$ et $f_2 = \mathbf{1}_{\{y_j\}}$.

Pour démontrer le dernier résultat, on applique ce qui précède à $|g_k|$. On en déduit que

$$E[|Z|] = \prod_{k=1}^{n} E[|g_k(X_k)|],$$

et on réutilise la proposition 2.11 de la même façon que précédemment. Fin de la preuve.

Proposition 2.21. *Soit X et Y deux variables aléatoires réelles discrètes de carré intégrable. On suppose que X et Y sont indépendantes. Alors,*

$$E[XY] = E[X]E[Y], \quad \operatorname{cov}(X,Y) = 0, \quad \operatorname{var}(X+Y) = \operatorname{var}(X) + \operatorname{var}(Y).$$

Preuve : On applique la proposition précédente.

Remarque 2.22. *Attention : la réciproque au résultat précédent est fausse, on peut trouver des variables aléatoires X et Y telles que $\operatorname{cov}(X,Y) = 0$ sans que X et Y soient indépendantes.*

Par exemple, soit X une variable aléatoire de loi $P(X = 1) = P(X = -1) = \frac{1}{2}$, Z une variable aléatoire indépendante de X et de loi $P(Z = -1) = \frac{2}{3}$ et $P(Z = 2) = \frac{1}{3}$, $Y = XZ$.

Montrer que $E[X] = E[Y] = E[XY] = 0$ donc X et Y ne sont pas corrélées, mais que $P(X = 1, Y = -2) = 0$, et en conclure que X et Y ne sont pas indépendantes.

2.7 Fonctions génératrices

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.

La distribution de X est une loi de probabilité sur $\mathbb{N}$ donc elle est entièrement déterminée par les quantités $P(X = n)$ pour n dans $\mathbb{N}$. En effet, pour tout $A \subset \mathbb{N}$,

$$P(X \in A) = \sum_{n \in A} P(X = n).$$

Donc X admet une espérance si et seulement si $\sum_n nP(X = n)$ converge, auquel cas

$$E[X] = \sum_{n \geqslant 0} nP(X = n).$$

De même, X est de carré intgrable si et seulement si $\sum_n n^2 P(X = n)$ converge, auquel cas

$$E[X^2] = \sum_{n \geqslant 0} n^2 P(X = n).$$

Si f est une application positive ou telle que $\sum_n |f(n)|P(X = n)$ converge, alors

$$E[f(X)] = \sum_{n \geqslant 0} f(n)P(X = n).$$

Définition 2.23. *Soit X une variable aléatoire à valeurs dans $\mathbb{N} \cup \{+\infty\}$. La fonction génératrice de X est la fonction $g_X : [0,1] \to [0,1]$ définie par*

$$g_X(s) := \sum_{n=0}^{+\infty} s^n P(X = n) = E[s^X].$$

Proposition 2.24. *Toute fonction génératrice vérifie les propriétés suivantes.*

1. *La fonction g_X est convexe, croissante sur $[0,1]$ et indéfiniment dérivable sur $[0,1[$.*
2. *La fonction g_X détermine la loi de X.*
3. *Si on suppose que X est à valeurs dans $\mathbb{N}$, X admet un moment d'ordre p si et seulement si g_X admet une dérivée à gauche d'ordre p en 1. Dans ce cas, pour $p=1$, $E[X] = g'_X(1-)$ et pour $p \geqslant 1$,*

$$E[X(X-1)\cdots(X-p+1)] = g_X^{(p)}(1-).$$

4. *Si X et Y sont indépendantes, alors $g_{X+Y} = g_X g_Y$.*

Preuve : La fonction g_X est donnée par une série entière de rayon de convergence au moins égal à 1, car $\sum_n P(X=n)$ converge. C'est donc une fonction indéfiniment dérivable sur $[0,1[$, et

$$g_X^{(p)}(s) = \sum_{n=0}^{+\infty} n(n-1)\cdots(n-p+1)s^{n-p}P(X=n),$$

donc

$$g_X^{(p)}(s) = E[X(X-1)\cdots(X-p+1)s^{X-p}].$$

En particulier, $g_X^{(n)}(0) = n!\,P(X=n)$. On en déduit facilement les points 1. et 2.

Démontrons le point 3. pour $p=1$. On suppose donc que X est à valeurs dans $\mathbb{N}$ et on pose $g = g_X$. Alors,

$$h(s) := \frac{g(1)-g(s)}{1-s} = \sum_{n=0}^{+\infty} \frac{1-s^n}{1-s} P(X=n).$$

Supposons que X est intégrable. On va utiliser le fait que, pour tout entier $n \geqslant 0$ et pour tout s dans $[0,1[$,

$$0 \leqslant \frac{1-s^n}{1-s} \leqslant n.$$

Par conséquent,

$$h(s) \leqslant \sum_{n=0}^{+\infty} nP(X=n) = E[X].$$

Dans l'autre sens, pour tout entier $N \geqslant 0$,

$$h(s) \geqslant \sum_{n=0}^{N} \frac{1-s^n}{1-s} P(X=n),$$

et la fraction $(1-s^n)/(1-s)$ tend vers n quand s tend vers 1. Pour N fixé, on ne manipule que des sommes finies, donc

$$\liminf_{s\to 1} h(s) \geqslant \sum_{n=0}^{N} nP(X=n).$$

Cette minoration est vraie pour tout $N \geqslant 0$, donc

$$\liminf_{s\to 1} h(s) \geqslant \sup_N \sum_{n=0}^{N} nP(X=n) = \sum_{n=0}^{+\infty} nP(X=n) = E[X],$$

donc $h(s)$ tend vers $E[X]$ quand s tend vers 1, c'est-à-dire que la dérivée à gauche en 1 de g existe et vaut $E[X]$, CQFD.

Réciproquement, supposons que g admet une dérivée à gauche en 1. Alors, pour tout $N \geqslant 0$,

$$\sum_{n=0}^{N} nP(X=n) = \lim_{s\to 1} \sum_{n=0}^{N} \frac{1-s^n}{1-s} P(X=n) \leqslant \lim_{s\to 1} \frac{1-g(s)}{1-s} = g'(1-),$$

donc X est intégrable et peut on utiliser la partie précédente de la preuve pour conclure. Finalement, g est dérivable à gauche en 1 si et seulement si $\sum_n nP(X=n)$ converge si et seulement si X est intégrable.

Le résultat général se prouve de la même manière.

Pour le point 4., pour tout s dans $[0,1]$,

$$E[s^{X+Y}] = E[s^X s^Y] = E[s^X]E[s^Y],$$

grâce à la proposition 2.20 et au fait que la fonction de $\mathbb{N}$ dans $\mathbb{R}$ définie par $x \mapsto s^x$ est bornée pour tout s dans $[0,1]$. Fin de la preuve.

Remarque 2.25. *Attention, la réciproque du point 4 est fausse. Exemple : soit (X,Y) dont la loi est donnée dans le tableau suivant :*

$x\backslash y$	*0*	*1*	*2*	$P(X=x)$
0	*1/9*	*0*	*2/9*	*1/3*
1	*2/9*	*1/9*	*0*	*1/3*
2	*0*	*2/9*	*1/9*	*1/3*
$P(Y=y)$	*1/3*	*1/3*	*1/3*	

On voit que X et Y ne sont pas indépendantes puisque $P(X=0, Y=1)$ est nul et $P(X=0)$ et $P(Y=1)$ ne le sont pas. Dans cet exemple, la loi de $X+Y$ est donnée par

k	*0*	*1*	*2*	*3*	*4*
$P(X+Y=k)$	*1/9*	*2/9*	*1/3*	*2/9*	*1/9*

On vérifie que $g_X(s) = g_Y(s) = \frac{1}{3}(1+s+s^2)$ et $g_{X+Y} = g_X g_Y$.

2.8 Exemples fondamentaux de lois discrètes

2.8.1 Loi de Dirac δ_x

Paramètre x réel (par exemple) : $X = x$ partout.

Donc $E[X] = x$, $\mathrm{var}(X) = 0$, et $g_X(s) = s^x$.

Si $\mathrm{var}(X) = 0$, on est dans ce cas.

Modèle : expérience dont le résultat est certain.

2.8.2 Loi de Bernoulli $b(p)$

Paramètre p avec $0 \leqslant p \leqslant 1$: X à valeurs dans $\{0, 1\}$ et $P(X = 1) = p = 1 - P(X = 0)$.

Donc $E[X] = p$, $\mathrm{var}(X) = p(1 - p)$, et $g_X(s) = 1 - p + ps$.

Modèle : jeu de pile ou face.

2.8.3 Loi uniforme $U(A)$

Paramètre A un ensemble fini de cardinal $n \geqslant 1$: X est à valeurs dans E et pour tout x dans A, $P(X = x) = 1/n$.

Modèle : expérience avec n résultats possibles, si on ne dispose d'aucune information supplémentaire, ou bien si la situation est entièrement symétrique.

2.8.4 Loi binomiale $B(n, p)$

Paramètres (n, p) avec $0 \leqslant p \leqslant 1$ réel et $n \geqslant 0$ entier : X est à valeurs dans $\{0, 1, \ldots, n\}$ et, pour tout $0 \leqslant k \leqslant n$, $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

Donc $g_X(s) = (1 - p + ps)^n$.

Modèle : nombre de piles parmi n résultats d'un jeu de pile ou face.

Proposition 2.26. *Soit $(X_1, \ldots, X_n)$ n variables aléatoires de Bernoulli indépendantes de paramètre p. Alors $X = X_1 + \cdots + X_n$ suit une loi binomiale de paramètre (n, p).*

Preuve : On peut utiliser les fonctions génératrices. Grâce à l'indépendance,

$$g_X(s) = \prod_k g_{X_k}(s) = (1 - p + ps)^n.$$

Comme la fonction génératrice caractérise la loi, le résultat suit.

Donc $E[X] = np$ et $\mathrm{var}(X) = np(1 - p)$.

Proposition 2.27. *Si n tend vers l'infini et p tend vers 0 de sorte que np tende vers a avec a strictement positif et fini, alors, pour tout $k \geqslant 0$,*

$$P(X_n = k) \to \frac{a^k}{k!}\mathrm{e}^{-a}.$$

2.8.5 Loi de Poisson $P(a)$

Paramètre a strictement positif : X à valeurs dans $\mathbb{N}$ et $P(X = k) = \frac{a^k}{k!}\mathrm{e}^{-a}$ pour tout $k \geqslant 0$.

Donc $g_X(s) = \mathrm{e}^{-a(1-s)}$, $E[X] = a$ et $\mathrm{var}(X) = a$.

Modèle : d'après la proposition 2.27, loi des événements rares.

2.8.6 Loi géométrique $G(p)$

Paramètre p dans $[0, 1]$: X est à valeurs dans $\mathbb{N}$ et $P(X = n) = p(1 - p)^n$ pour tout $n \geqslant 0$.

Donc $g_X(s) = p/(1 - (1 - p)s)$, $E[X] = (1 - p)/p$ et $\mathrm{var}(X) = (1 - p)/p^2$.

Modèle : instant du premier succès dans une suite d'expériences indépendantes chacune avec probabilité p de succès.

Proposition 2.28. *Soit $(X_n)_{n\geqslant 0}$ une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre p (c'est-à-dire que pour tout $n \geqslant 0$, les variables aléatoires $(X_0, \ldots, X_n)$ sont indépendantes). Soit X la variable aléatoire définie par*

$$X = \min\{n \geqslant 0 \,|\, X_n = 1\}.$$

Alors X est finie avec probabilité 1 et de loi géométrique de paramètre p.

Preuve : L'événement $\{X = n\}$ vaut l'intersection de $\{X_n = 1\}$ et des n événements $\{X_k = 0\}$ pour $0 \leqslant k \leqslant n - 1$. Par indépendance, $P(X = n) = p(1 - p)^n$.

2.9 Autres exemples de lois discrètes

2.9.1 Loi hypergéométrique $H(n, N, N_1)$

Paramètres (n, N, N_1) avec $0 \leqslant n \leqslant N$ et $0 \leqslant N_1 \leqslant N$: X est à valeurs dans $\{0, 1, \ldots, n\}$ et pour tout $0 \leqslant k \leqslant n$,

$$P(X = k) = \frac{\binom{N_1}{k}\binom{N-N_1}{n-k}}{\binom{N}{n}},$$

avec la convention que $\binom{i}{j} = 0$ si $j < 0$ ou si $j > i$.

Modèle : on effectue n tirages sans remise dans un ensemble A de cardinal N ; le nombre d'objets de $A_1 \subset A$ tirés suit cette loi si le cardinal de A est N et celui de A_1 est N_1.

Proposition 2.29. *Soit X une variable aléatoire de loi hypergéométrique de paramètres (n, N, N_1). Alors,*

$$E[X] = nN_1/N, \qquad \mathrm{var}(X) = \frac{nN_1(N-N_1)(N-n)}{N^2(N-1)}.$$

Preuve : On utilise le fait que $k\binom{N_1}{k} = N_1\binom{N_1-1}{k-1}$. Ainsi,

$$k\,P(X=k) = \frac{k\binom{N_1}{k}\binom{N-N_1}{n-k}}{\binom{N}{n}} = \frac{N_1\binom{N_1-1}{k-1}\binom{N-N_1}{(n-1)-(k-1)}}{N\binom{N-1}{n-1}/n},$$

donc

$$k\,P(X=k) = (nN_1/N)P(Y=k-1),$$

si Y suit une loi hypergéométrique de paramètres $(n-1, N-1, N_1-1)$. En sommant sur k et en utilisant le fait que la somme des $P(Y=k-1)$ vaut 1, on trouve la valeur de $E[X]$.

En appliquant le même principe à $P(Y=k-1)$, on obtient

$$k(k-1)\,P(X=k) = (nN_1/N)(k-1)P(Y=k-1),$$

et

$$(k-1)P(Y=k-1) = ((n-1)(N_1-1)/(N-1))P(Z=k-2),$$

si Z suit une loi hypergéométrique de paramètres $(n-2, N-2, N_1-2)$, donc

$$E[X(X-1)] = \frac{n(n-1)N_1(N_1-1)}{N(N-1)},$$

ce qui permet de trouver la valeur de

$$\mathrm{var}(X) = E[X(X-1)] + E[X] - E[X]^2.$$

Fin de la preuve.

Proposition 2.30. *Si N et N_1 tendent vers l'infini de sorte que N_1/N converge vers p, alors $P(X=k) \to \binom{n}{k}p^k(1-p)^{n-k}$.*

Preuve laissée en exercice.

Ceci montre qu'un tirage sans remise équivaut à un tirage avec remise si le cardinal de l'ensemble est grand et le rapport N_1/N à peu près constant.

Remarque 2.31. *Soit X_i la variable aléatoire valant $X_i = 1$ si au tirage numéro i, on obtient un objet de A_1 et $X_i = 0$ sinon. Alors $X = X_1 + \cdots + X_n$ et les variables aléatoires X_i suivent des lois de Bernoulli. Quels sont leurs paramètres ? Puisque $E[X] = nN_1/N$ et $E[X] = E[X_1] + \cdots + E[X_n]$, on voit que $P(X_i = 1) = N_1/N$ pour tout i.*

Mais attention, les variables aléatoires X_i ne sont pas indépendantes ; sinon la loi de leur somme X serait binomiale $(n, N_1/N)$.

2.9.2 Loi multinomiale $M(n; p_1, \ldots, p_r)$

Paramètres $n \geqslant 0$ entier et $(p_1, \ldots, p_r)$ avec p_k dans $[0, 1]$ et $p_1 + \cdots + p_r = 1$: alors $X = (X_1 \ldots, X_r)$ est à valeurs dans l'ensemble des r-uplets $(k_1, \ldots, k_r)$ dans $\mathbb{N}^r$ tels que $k_1 + \cdots + k_r = n$ et, pour tout tel $(k_1, \ldots, k_r)$,

$$P(X = (k_1, \ldots, k_r)) = \binom{n}{k_1, \ldots, k_r} p_1^{k_1} \ldots p_r^{k_r}, \quad \binom{n}{k_1, \ldots, k_r} = \frac{n!}{k_1! \ldots k_r!}.$$

Quand $r = 2$, (X_1, X_2) suit la loi de $(Y, n - Y)$ avec Y de loi binomiale (n, p_1). Plus généralement :

Proposition 2.32. *Si X suit la loi $M(n; p_1, \ldots, p_r)$, les lois marginales des X_i sont les lois binomiales (n, p_i) pour $1 \leqslant i \leqslant r$.*

Preuve : On peut faire le calcul directement en remarquant que $\{X_i = k\}$ est la réunion des $\{X = (k_1, \ldots, k_r)\}$ sur tous les r-uplets $(k_1, \ldots, k_r)$ tels que $k_i = k$. C'est lourd. On peut aussi utiliser les fonctions génératrices. Pour tout $(s_1, \ldots, s_r)$ dans $[0, 1]^r$,

$$E[s_1^{X_1} \cdots s_r^{X_r}] = \sum_{(k_1, \ldots, k_r)} \binom{n}{k_1, \ldots, k_r} (p_1 s_1)^{k_1} \ldots (p_r s_r)^{k_r}.$$

La somme du membre de droite vaudrait 1 si $p_1 s_1 + \cdots + p_r s_r$ valait 1 car ce serait la masse totale d'une distribution. Par homogénéité, on en déduit

$$E[s_1^{X_1} \cdots s_r^{X_r}] = (p_1 s_1 + \cdots + p_r s_r)^n.$$

Il reste à choisir $s_j = 1$ pour tout $j \neq i$ et à remarquer que la somme des p_j pour $j \neq i$ vaut $1 - p_i$ pour obtenir

$$E[s_i^{X_i}] = (1 - p_i + p_i s_i)^n.$$

C'est la fonction génératrice de la loi binomiale (n, p_i). CQFD.

Exercice : Calculer la loi de (X_1, X_2).

2.9.3 Loi binomiale négative $NB(n, p)$

Paramètres $n \geqslant 1$ et $0 \leqslant p \leqslant 1$: $P(X = k) = \binom{k-1}{n-1} p^n (1 - p)^{k-n}$ pour tout $0 \leqslant n \leqslant k$ (attention à l'ordre des deux indices).

Construction : rang du succès numéro n dans une suite d'épreuves indépendantes, chacune avec probabilité de succès p.

Donc $NB(n, p)$ est la loi de la somme de n variables aléatoires indépendantes de loi $G(p)$ et $NB(1, p) = G(p)$.

Donc $g_X(s) = p^n/(1 - (1 - p)s)^n$, $E[X] = n/p$ et $\mathrm{var}(X) = n(1 - p)/p^2$.

2.10 Exercices

2.10.1 Espérance discrète

Soit X une variable aléatoire à valeurs dans $\{0, 1, ..., n\}$. Montrer que :

$$E(X) = \sum_{k=1}^{n} P(X \geqslant k).$$

2.10.2 Loi sans mémoire

Soit T une variable aléatoire à valeurs dans $\mathbb{N}$ telle que pour tout $n \geqslant 0$ et $k \geqslant 0$:

$$P(T \geqslant n + k \,|\, T \geqslant n) = P(T \geqslant k).$$

Déterminer la loi de T.

2.10.3 Boules colorées

Une urne contient N boules dont N_1 portent le numéro 1, N_2 portent le numéro 2, ..., et N_k portent le numéro k. On fait un tirage de n boules avec remise. Soit X_i le nombre de boules tirées qui portent le numéro i, et $X = (X_1, \ldots, X_k)$.

a) Donner la loi de X.

b) Donner la loi de X_i pour $1 \leqslant i \leqslant k$.

c) Donner la loi de (X_i, X_j) pour $i \neq j$.

Facultatif : traiter les mêmes questions pour un tirage sans remise.

2.10.4 Instants

On effectue une suite infinie de lancers indépendants d'une pièce de monnaie. Le résultat du lancer numéro n est une variable aléatoire X_n qui vaut 1 si l'on a obtenu « pile », ce qui arrive avec une probabilité p, et qui vaut 0 si l'on a obtenu « face », ce qui arrive avec une probabilité $1 - p$.

On note N_1, N_2, etc., les numéros des lancers successifs où l'on a obtenu « pile », c'est-à-dire

$$N_1 = \inf\{n \geqslant 1 \,;\, X_n = 1\}, \qquad N_{k+1} = \inf\{n \geqslant N_k + 1 \,;\, X_n = 1\},\ k \geqslant 1.$$

a) Donner la loi de N_1 et montrer que N_1 est fini avec probabilité 1.

b) Donner la loi de (N_1, N_2). Montrer que les variables aléatoires N_1 et $N_2 - N_1$ sont indépendantes et de même loi.

c) Donner la loi de N_2. Donner la loi de N_1 sachant $N_2 = n_2$.

d) Donner la loi de N_k pour tout $k \geqslant 3$.

2.10.5 Lois de Bernoulli

Soit X et Y deux variables aléatoires de lois de Bernoulli, pas forcément de même loi. On suppose que $\mathrm{cov}(X, Y) = 0$. Montrer que X et Y sont indépendantes.

2.10.6 Lois binomiales

Soit n et m deux entiers strictement positifs et $0 \leqslant p \leqslant 1$. Si X suit une loi binomiale de paramètre (n, p), si Y suit une loi binomiale de paramètre (m, p) et si X et Y sont indépendantes, calculer la loi de $Z := X + Y$.
En déduire la formule suivante : pour tout $0 \leqslant k \leqslant n + m$,

$$\sum_{i=0}^{n} \binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}.$$

2.10.7 Les boîtes d'allumettes de Banach

On prétend que Stefan Banach avait toujours une boîte de N allumettes dans chacune de ses poches. Quand il voulait une allumette, il choisissait une de ses deux poches au hasard, et ses choix successifs constituaient une suite de tirages de Bernoulli avec paramètre $p = \frac{1}{2}$.

On considère le moment où Banach découvre pour la première fois qu'une boîte est vide. À ce moment, l'autre boîte contient X_N allumettes. Donner la loi de X_N.

Facultatif : montrer que $E[X_N] = (2N + 1)u_N - 1$ où $u_N = P(X_N = 0)$. Montrer en utilisant la formule de Stirling que, quand N tend vers l'infini, $E[X_N]$ est équivalent à $2\sqrt{N/\pi}$.

2.10.8 Méthode du maximum observé

Une urne contient N balles numérotées de 1 à N. On effectue n tirages avec remise. Soit X le plus grand nombre tiré lors des n tirages.

a) Donner la fonction de répartition de X.

b) Donner la loi de X.

c) Calculer $E[X]$ et donner un équivalent de $E[X]$ quand $N \to +\infty$

Remarque : Les statisticiens utilisent cette méthode du maximum observé pour estimer le paramètre inconnu N. Prolongement : comparer l'efficacité de cette méthode avec l'estimateur basé sur la moyenne des numéros observés pendant n tirages.

2.10.9 Clés

Un homme possède n clés et veut ouvrir une porte. Une seule parmi les clés dont il dispose ouvre la porte. Il essaie les clés au hasard. Trouver l'espérance et la variance du nombre d'essais nécessaires si :

a) Les clés qui ne marchent pas sont remises avec les autres.

b) Les clés qui ne marchent pas sont mises de coté.

2.10.10 Poisson(s)

Soient X et Y deux variables aléatoires indépendantes de lois de Poisson de paramètre a et b.

a) Déterminer la loi de la variable aléatoire $S = X + Y$.

b) Déterminer, pour tout couple (n, k) d'entiers naturels, la probabilité conditionnelle $P(X = k \,|\, S = n)$.

c) (Facultatif) Soit $r \geqslant 1$ un entier et X_k des variables aléatoires indépendantes de lois de Poisson de paramètres respectifs a_k. Donner la loi conditionnelle de $(X_1, ..., X_r)$ sachant $\{X_1 + ... + X_r + X_{r+1} = n\}$, pour tout $n \geqslant 0$.

2.10.11 Tirages biaisés

On dispose de n pièces numérotées de 1 à n et biaisées de façon que la probabilité d'obtenir pile avec la pièce i vaut $1/(2i + 1)$.

On lance les n pièces et on veut calculer la probabilité d'obtenir un nombre impair de piles.

Soit X_i la variable aléatoire qui vaut $X_i = 1$ si la pièce i est tombée sur pile et $X_i = 0$ sinon, et $X = X_1 + \cdots + X_n$.

a) Préciser ce que représente la variable aléatoire X.

b) Calculer la fonction génératrice φ de X.

c) Montrer que la probabilité cherchée vaut $\frac{1}{2}(\varphi(1) + \varphi(-1))$ et calculer cette valeur.

2.10.12 Lois géométriques

Soit X et Y deux variables aléatoires indépendantes de lois géométriques de paramètres respectifs a et b. Soit $Z = \min(X, Y)$ et $U = |X - Y|$.

a) Déterminer $P(X \geqslant n)$ pour tout $n \geqslant 0$. Déterminer $P(Z \geqslant n)$. Préciser la loi de Z.

b) Calculer $P(U = 0)$. Déterminer la loi de U.

c) Montrer que Z et U sont indépendantes.

2.10.13 Régression linéaire

Soit X et Y deux variables aléatoires discrètes de variances non nulles.

On pose

$$\varrho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}, \quad \overline{X} = X - E[X], \quad \overline{Y} = Y - E[Y].$$

1) Montrer que $|\mathrm{cov}(X,Y)| = |E[\overline{X}\ \overline{Y}]| \leqslant \sigma_X \sigma_Y$.

En déduire que $-1 \leqslant \varrho_{X,Y} \leqslant 1$.

2) Montrer que $|\varrho_{X,Y}| = 1$ si et seulement sil existe a non nul et b tels que $P(Y = aX + b) = 1$.

On pourra calculer $E[(\overline{Y} + t\overline{X})^2]$.

3) Préciser la valeur de $\varrho_{X,Y}$ si X et Y sont indépendantes.

4) On cherche la meilleure approximation de Y comme fonction affine de X au sens des moindres carrés, c'est-à-dire que l'on cherche les valeurs de a et b qui minimisent $E[(aX + b - Y)^2]$. Notons $\Phi(a,b) = E[(aX + b - Y)^2]$

Montrer que $\Phi(a,b) = E[(\overline{Y} - a\overline{X})^2] + (E[Y] - (aE[X] + b))^2$.

En déduire que le couple (a_0, b_0) qui minimise Φ vaut

$$a_0 = \varrho_{X,Y}\sigma_Y/\sigma_X, \qquad b_0 = E[Y] - a_0 E[X].$$

On appelle la droite d'équation $y = a_0 x + b_0$ la droite de régression linéaire de Y en X.

5) On suppose que (X,Y) suit la loi uniforme sur un ensemble de cardinal n, c'est-à-dire qu'il existe n points (x_i, y_i) dans le plan tels que $P(X = x_i, Y = y_i) = 1/n$ pour tout $1 \leqslant i \leqslant n$.

Déterminer la droite de régression linéaire de Y en X dans ce cas.

2.10.14 Loi jointe

On effectue une suite infinie de lancers indépendants d'un dé équilibré. On numérote les lancers à partir de 1. On définit X comme le numéro du premier lancer qui donne 6, et Y comme le nombre de 5 obtenus avant d'obtenir le premier 6.

Déterminer la loi du couple (X,Y).
Déterminer la loi conditionnelle de Y sachant l'événement $\{X = n\}$.
Déterminer la loi absolue de Y.

Fiche 4 : Probabilités (3) Variables aléatoires densitables

« Les questions les plus importantes de la vie ne sont en effet, pour la plupart, que des problèmes de probabilité. » Pierre-Simon Laplace

1 Description

La fonction de répartition $F_X : \mathbb{R} \to [0,1]$ d'une variable aléatoire X à valeurs réelles est définie par $F_X(x) := P(X \leqslant x)$ pour tout x réel et caractérise la loi de X. La fonction F_X est croissante, continue à droite et à valeurs dans $[0,1]$, de plus $F_X(x)$ tend vers 0 quand x tend vers $-\infty$ et $F_X(x)$ tend vers 1 quand x tend vers $+\infty$. Réciproquement, toute fonction vérifiant ces propriétés (c'est-à-dire toute fonction à valeurs dans $[0,1]$, croissante, continue à droite, de limites 0 en $-\infty$ et 1 en $+\infty$) est la fonction de répartition d'une variable aléatoire.

1.1 Densité

Définition 1.1. *On dit que la variable aléatoire $X : \Omega \to \mathbb{R}$ admet la loi de densité f si la fonction $f : \mathbb{R} \to \mathbb{R}^+$ est continue par morceaux et si, pour tout réel x,*

$$F_X(x) = P(X \leqslant x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t.$$

Si f est la densité de la loi d'une variable aléatoire, $f \geqslant 0$ et

$$\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1.$$

Si F_X est continue et admet une dérivée sauf en un nombre fini de points, alors la loi de X admet une densité donnée par la dérivée de F_X là où cette dérivée existe et par n'importe quelle valeur là où cette dérivée n'existe pas.

Remarque 1.2. *En toute généralité, la loi d'une variable aléatoire X à valeurs réelles comporte trois parties : une partie absolument continue, décrite par une densité f comme ci-dessus ; une partie atomique, correspondant à une mesure discrète et décrite dans le*

chapitre précédent par des nombres p_x ; et une troisième partie, appelée partie singulière et plus difficile à décrire, qui est en fait simplement la partie qui n'est ni atomique, ni densitable. Quand la loi comporte une partie absolument continue et une partie atomique, on utilisera la notation

$$P_X(\mathrm{d}x) = f_X(x)\,\mathrm{d}x + \sum_x p_x\,\delta_x,$$

qui signifie que, pour tout B,

$$P(X \in B) = \int_B f_X(x)\,\mathrm{d}x + \sum_{x\in B} p_x.$$

Même si on en parlera peu dans la suite, de tels « mélanges » apparaissent naturellement.

Exemple/exercice : soit X une variable aléatoire de loi uniforme sur l'intervalle $[0,4]$ (voir la définition plus bas) et $Y = \max(1, \min(X, 2))$. Décrire la loi de Y.

1.2 Espérance

On revient au cas densitable.

Définition 1.3. *Quand l'intégrale $\displaystyle\int_{-\infty}^{+\infty} |x|\,f(x)\,\mathrm{d}x$ converge, on dit que la variable aléatoire X de densité f est intégrable (ou admet une espérance). Dans ce cas, on pose*

$$E(X) := \int_{-\infty}^{+\infty} x\,f(x)\,\mathrm{d}x.$$

Si $E(X) = 0$, on dit que X est centrée.

La formule de changement de variables ci-dessous permet de calculer des espérances de fonctions de variables aléatoires densitables. C'est l'analogue de la « formule fondamentale » du chapitre 2 pour les variables aléatoires discrètes.

Proposition 1.4. *Soit X une variable aléatoire densitable de densité f_X et soit $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction. On pose $Y := \varphi(X)$. Alors la variable aléatoire Y est intégrable si et seulement si l'intégrale de la fonction $|\varphi|\,f_X$ sur $\mathbb{R}$ converge. Dans ce cas,*

$$E(Y) = \int_{-\infty}^{+\infty} \varphi(x)\,f_X(x)\,\mathrm{d}x.$$

Exemple 1.5. *Soit $Y = aX + b$ avec $a \neq 0$. Alors Y est intégrable si et seulement si X l'est et dans ce cas, $E(Y) = aE(X) + b$. Pour tout y,*

$$F_Y(y) = F_X((y-b)/a) \text{ si } a > 0, \quad F_Y(y) = 1 - F_X((y-b)/a) \text{ si } a < 0.$$

Donc Y admet une densité f_Y donnée par $f_Y(y) := f_X((y-b)/a)/|a|$.

L'espérance est un opérateur linéaire : si X et Y sont des variables aléatoires réelles intégrables et a et b des nombres réels,

$$E(aX + bY) = aE(X) + bE(Y).$$

L'espérance est un opérateur positif : si $X \geqslant 0$, alors $E(X) \geqslant 0$, donc si $X \leqslant Y$ alors $E(X) \leqslant E(Y)$.

Définition 1.6. *Soit X une variable aléatoire continue et p un réel strictement positif. On dit que X admet un moment d'ordre p si $E(|X|^p)$ est fini. Alors $E(X^p)$ existe et on l'appelle le moment d'ordre p de X.*

Le cas $p = 2$ est important et la section suivante lui est consacrée.

1.3 Second moment

Définition 1.7. *On dit que la variable aléatoire X est de carré intégrable si et seulement si X^2 est intégrable si et seulement si l'intégrale*

$$\int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x$$

est finie. Alors,

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx.$$

Voici un rappel.

Proposition 1.8 (Inégalité de Cauchy-Schwarz). *Soit f et g deux fonctions telles que f^2 et g^2 sont d'intégrale finie sur $\mathbb{R}$. Alors fg est d'intégrale finie sur $\mathbb{R}$ et*

$$\left(\int_{-\infty}^{+\infty} f(x)g(x)\mathrm{d}x\right)^2 \leqslant \int_{-\infty}^{+\infty} f(x)^2\mathrm{d}x \cdot \int_{-\infty}^{+\infty} g(x)^2\mathrm{d}x.$$

Preuve (rappel) : une méthode consiste à considérer $h := J(f)g - J(g)f$ où $J(f)$, respectivement $J(g)$, désigne la racine carrée de l'intégrale de f^2, respectivement g^2. Alors $h^2 \geqslant 0$ donc l'intégrale de h^2 est positive ou nulle, et on en déduit le résultat.

Une autre méthode, peut-être plus classique mais à mon avis moins bonne, consiste à remarquer que pour tout réel t, l'intégrale du carré de $f + tg$ est positive. Le discriminant du polynôme du second degré en t correspondant est donc négatif ou nul, ce qui fournit le résultat. **Ces deux preuves sont à connaître.**

Application. Soit f une fonction positive ou nulle. En écrivant $|x|f(x)$ comme le produit $|x|\sqrt{f(x)} \times \sqrt{f(x)}$, on obtient

$$\left(\int_{-\infty}^{+\infty} |x|f(x)\mathrm{d}x\right)^2 \leqslant \int_{-\infty}^{+\infty} x^2 f(x)\mathrm{d}x \cdot \int_{-\infty}^{+\infty} f(x)\mathrm{d}x.$$

On en déduit le résultat suivant.

Proposition 1.9 (et définition). *Si X est de carré intégrable, X admet une espérance et*

$$E[X]^2 \leqslant E[|X|]^2 \leqslant E[X^2].$$

On définit alors la variance $\mathrm{var}(X)$ *et l'écart type* $\sigma(X)$ *de X par*

$$\mathrm{var}(X) := E[(X - E[X])^2) = E[X^2] - E[X]^2, \quad \sigma(X) := \sqrt{\mathrm{var}(X)}.$$

On note souvent $\mathrm{var}(X) = \sigma^2(X)$.

Proposition 1.10 (et définition). *1) Si X est de carré intégrable, alors $Y = aX + b$ aussi, pour tous a et b, et* $\mathrm{var}(Y) = a^2\mathrm{var}(X)$.
2) Pour toute variable aléatoire X de carré intégrable, $\mathrm{var}(X) \geqslant 0$ *et* $\mathrm{var}(X) = 0$ *si et seulement si X est constante.*
3) Pour toutes variables aléatoires X et Y de carré intégrable,

$$\mathrm{var}(X + Y) = \mathrm{var}(X) + \mathrm{var}(Y) + 2\mathrm{cov}(X, Y),$$

où la covariance $\mathrm{cov}(X, Y)$ *de X et Y est définie, si X et Y sont de carré intégrable, par la formule*

$$\mathrm{cov}(X, Y) := E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y].$$

Preuve de 1) : $E[Y] = aE[X] + b$ et $Y^2 = a^2X^2 + 2abX + b^2$ donc, toujours par linéarité, $E[Y^2]$ vaut $a^2E[X^2] + 2abE[X] + b^2$ et en effectuant la soustraction, on obtient le résultat.

1.4 Fonction caractéristique

Comme son nom l'indique, la fonction caractéristique fournit une façon de décrire entièrement la loi d'une variable aléatoire. La fonction caractéristique joue dans le cas général le rôle de la fonction génératrice dans le cas discret entier.

Définition 1.11. *On appelle fonction caractéristique de X la transformée de Fourier de la loi de X, c'est-à-dire la fonction $\varphi_X : \mathbb{R} \to \mathbb{C}$ définie pour tout réel x, par*

$$\varphi_X(x) := E[\mathrm{e}^{\mathrm{i}xX}].$$

La fonction caractéristique φ_X existe toujours puisque $|\mathrm{e}^{\mathrm{i}xX}| \leqslant 1$ est une fonction bornée de X.

Proposition 1.12. *La transformée de Fourier caractérise la loi : si les fonctions φ_X et φ_Y sont égales, alors les lois de X et de Y sont égales.*

Si $X \geqslant 0$ partout, on peut aussi considérer la transformée de Laplace λ_X de (la loi de) X, définie pour tout réel positif x, par

$$\lambda_X(x) := E[\mathrm{e}^{-xX}].$$

La transformée de Laplace λ_X existe toujours si $X \geqslant 0$ partout puisque, dans ce cas, $\mathrm{e}^{-xX} \leqslant 1$ est une fonction bornée de X.

La transformée de Laplace caractérise la loi : si X et Y sont des variables aléatoires positives et si $\lambda_X = \lambda_Y$, alors la loi de X et la loi de Y sont égales.

La notation φ_X est canonique, la notation λ_X ne l'est pas.

2 Familles de variables aléatoires

On considère maintenant un couple de variables aléatoires (X, Y). Sa fonction de répartition est définie par

$$F_{(X,Y)}(x, y) := P(X \leqslant x, Y \leqslant y).$$

Sa fonction caractéristique est définie par

$$\varphi_{(X,Y)}(x, y) := E[\mathrm{e}^{\mathrm{i}(xX+yY)}].$$

Les lois marginales sont les lois de X et de Y. Il est important de réaliser que la loi de X et la loi de Y ne déterminent pas la loi de (X, Y).

Vérifions ceci sur deux exemples.

Exercice 2.1. *Pour chaque couple de variables aléatoires (X, Y) ci-dessous, calculer la loi de X et la loi de Y.*

1) Soit $0 \leqslant t \leqslant \frac{1}{2}$. La loi de (X, Y) est

$$(\tfrac{1}{2} - t)\delta_{(0,0)} + t\delta_{(0,1)} + t\delta_{(1,0)} + (\tfrac{1}{2} - t)\delta_{(1,1)}.$$

2) Soit D le sous-ensemble du carré $[0, 1]^2$ formé des points (x, y) tels que $y \leqslant x$ ou $y \geqslant x + \frac{1}{2}$. La loi de (X, Y) est uniforme sur D.

Ensuite, décrire une loi plus simple d'un couple (X', Y') de variables aléatoires telles que X et X' ont la même loi et Y et Y' ont la même loi.

La façon la plus simple de spécifier la loi d'un couple à partir des marginales est d'imposer l'indépendance.

Définition 2.2. *On dit que deux événements A et A' sont indépendants si et seulement si $P(A \cap A') = P(A)P(A')$. On dit que deux variables aléatoires X et Y sont indépendantes si et seulement si, pour tous B et B', $\{X \in B\}$ et $\{Y \in B'\}$ sont des événements indépendants. En d'autres termes, on demande que*

$$P(X \in B, Y \in B') = P(X \in B)P(Y \in B').$$

Proposition 2.3. *Les variables aléatoires X et Y sont indépendantes :*

1. *si et seulement si, pour tous réels x et y, $F_{(X,Y)}(x, y) = F_X(x)F_Y(y)$,*

2. si et seulement si, pour toutes fonctions h et g bornées,

$$E[h(X)g(Y)] = E[h(X)]E[g(Y)],$$

3. si et seulement si, pour tous réels x et y, $\varphi_{(X,Y)}(x,y) = \varphi_X(x)\varphi_Y(y)$.

Dans le cas densitable, X et Y sont indépendantes si et seulement si, pour tout couple (x,y),

$$f_{(X,Y)}(x,y) = f_X(x)f_Y(y).$$

Dans le cas discret, X et Y sont indépendantes si et seulement si, pour tout couple (x,y),

$$P(X = x, Y = y) = P(X = x)P(Y = y).$$

Définition 2.4. *On dit qu'un vecteur $(X_k)_{1\leqslant k\leqslant n}$ de taille $n \geqslant 2$ est indépendant si et seulement si, pour tous B_k,*

$$P(\forall 1 \leqslant k \leqslant n,\, X_k \in B_k) = \prod_{k=1}^{n} P(X_k \in B_k).$$

Si le vecteur $(X_k)_{1\leqslant k\leqslant n}$ est indépendant, les variables aléatoires X_k sont indépendantes deux à deux pour $1 \leqslant k \leqslant n$. La réciproque est fausse, comme le montre l'exemple ci-dessous.

Exemple 2.5 (Important). *Soit X et Y deux variables indépendantes de Bernoulli (voir plus bas) de loi $\frac{1}{2}(\delta_1 + \delta_{-1})$, et soit $Z = XY$. Calculer la loi de (X,Y,Z) et les lois de (X,Y), (X,Z) et (Y,Z). Préciser si le triplet (X,Y,Z) est indépendant. Préciser si les couples (X,Y), (X,Z) et (Y,Z) sont indépendants.*

Si X et Y sont des variables aléatoires indépendantes et de carré intégrable, alors $\mathrm{cov}(X,Y) = 0$. La réciproque est fausse, comme le montre l'exemple ci-dessous.

Exemple 2.6. *Soit X une variable aléatoire de loi $\frac{1}{3}(\delta_{-1} + \delta_0 + \delta_1)$ et soit $Y = X^2$. Montrer que $XY = X$. En déduire que les variables aléatoires X et Y ne sont pas indépendantes mais que* $\mathrm{cov}(X,Y) = 0$.

Définition 2.7 (Coefficient de corrélation). *Toujours dans le cas de carré intégrable, on note*

$$\varrho(X,Y) := \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)\mathrm{var}(Y)}}.$$

Donc $-1 \leqslant \varrho(X,Y) \leqslant 1$. Lorsque $|\varrho(X,Y)|$ est proche de 1, on dit que les variables aléatoires sont fortement corrélées.

Cas vectoriel : loi conjointe et marginales sans changement.

Dans le cas de carré intégrable, la matrice de covariance C vaut

$$C_{i,i} := \mathrm{var}(X_i), \quad C_{i,j} := \mathrm{cov}(X_i, X_j),\ i \neq j.$$

3 Exemples de lois à densité

3.1 Loi uniforme $U(I)$

Paramètre : I intervalle d'intérieur non vide. Densité $|I|^{-1}\mathbf{1}_I$.

Si $I = (a, b)$ avec $a < b$, espérance $(a+b)/2$, variance : $(b-a)^2/12$. Fonction de répartition $F(x) = 0$ si $x \leqslant a$, $F(x) = (x-a)/(b-a)$ si $a \leqslant x \leqslant b$ et $F(x) = 1$ si $x \geqslant b$.

Si la loi de X est $U([0,1])$ (uniforme standard) et si $A : x \mapsto ax + b$, $a \neq 0$, est une application affine, alors la loi de $A(X) = aX + b$ est uniforme sur $A([0,1])$ qui vaut $[b, a+b]$ si $a > 0$ et $[a+b, b]$ si $a < 0$.

3.2 Loi exponentielle $\mathcal{E}(a)$

Paramètre : a réel strictement positif. Densité $a\,\mathrm{e}^{-ax}\,\mathbf{1}_{x\geqslant 0}$.

Espérance $1/a$, variance $1/a$. Fonction de répartition $F(x) = 0$ si $x \leqslant 0$ et $F(x) = 1-\mathrm{e}^{-ax}$ si $x \geqslant 0$.

Si la loi de X est $\mathcal{E}(1)$ (exponentielle standard) et si a est un réel positif, alors la loi de aX est $\mathcal{E}(1/a)$.

Construction : horloges. Ou $-\log U$, U uniforme standard.

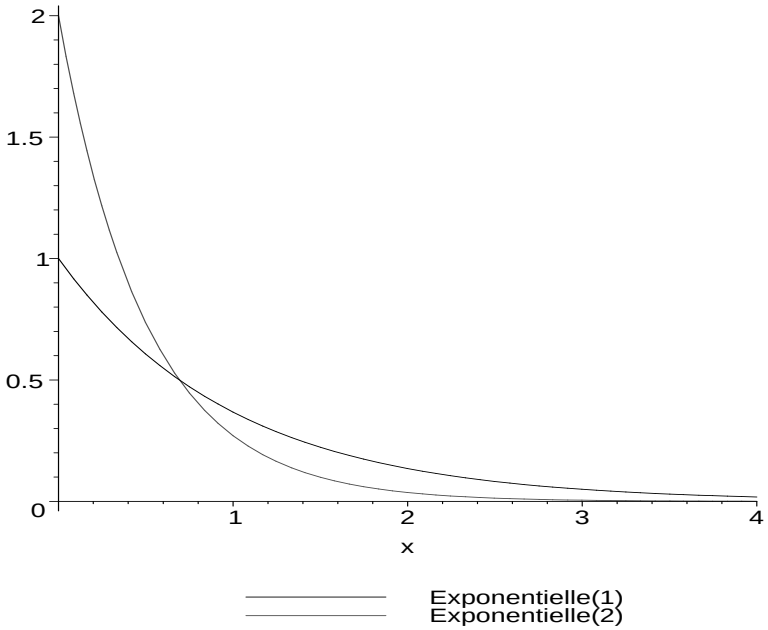


Figure 1. Densité de lois exponentielles

Les lois exponentielles sont les seules lois à valeurs dans $\mathbb{R}^+$ « sans mémoire », c'est-à-dire

telles que, pour tous réels positifs s et t,

$$P(X \geqslant s+t | X \geqslant t) = P(X \geqslant s).$$

3.3 Loi normale (ou gaussienne) $\mathcal{N}(m, v)$

Paramètres : m réel et v réel strictement positif. Densité $\mathrm{e}^{-(x-m)^2/(2v)}/\sqrt{2\pi v}$.

Espérance m, variance v.

Si la loi de X est $\mathcal{N}(0,1)$ (gaussienne standard) et si a et b sont des réels, alors la loi de $aX+b$ est $\mathcal{N}(b, a^2)$.

Par convention, $\mathcal{N}(m, 0)$ est une masse de Dirac en m.

La somme de variables aléatoires indépendantes de loi $\mathcal{N}(m_k, v_k)$ suit la loi $\mathcal{N}(m, v)$ avec

$$m := \sum_k m_k, \qquad v := \sum_k v_k.$$

Construction : loi des erreurs.

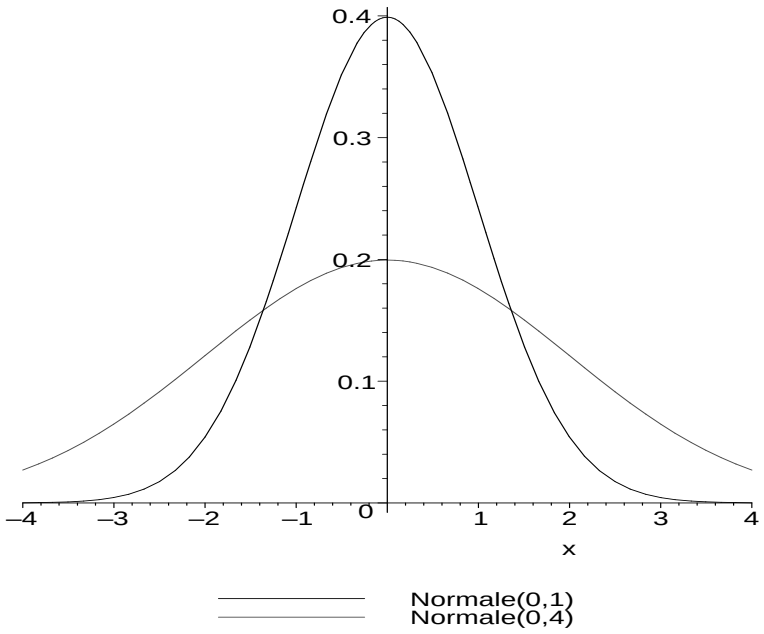


Figure 2. Densité de lois normales

3.4 Loi de Cauchy $C(m, a)$

Paramètres : a strictement positif et m réel. Densité $(a/\pi)/((x-m)^2+a^2)$.

Pas d'espérance ni de variance car les intégrales divergent. La médiane vaut m.

Si la loi de X est $C(0,1)$ (loi de Cauchy standard) et si a et b sont des réels, alors la loi de $aX+b$ est $C(b,a^2)$.

Par convention, $C(m,0)$ est une masse de Dirac en m.

La somme de variables aléatoires indépendantes de loi $C(m_k,a_k)$ suit la loi $C(m,a)$ avec $m=\sum_k m_k$ et $a=\sum_k a_k$.

Construction : la mouche. Ou bien le rapport de deux gaussiennes standard.

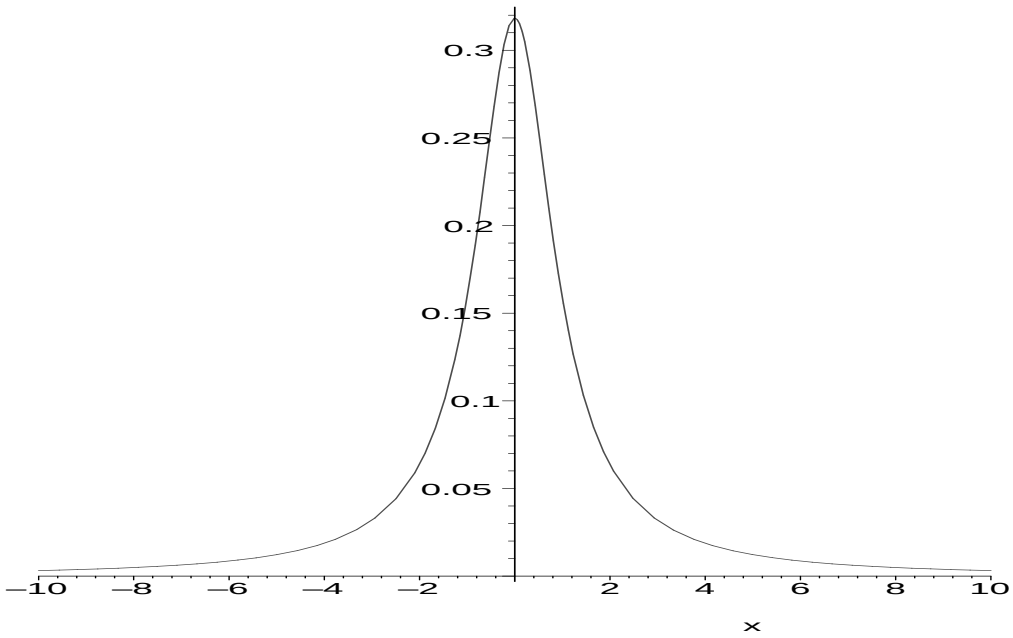


Figure 3. Densité de la loi de Cauchy

3.5 Autres lois

3.5.1 Loi gamma $G(a,\lambda)$

Paramètres : a et b strictement positifs. Densité

$$b^a\Gamma(a)^{-1}x^{a-1}\mathrm{e}^{-bx}\,\mathbf{1}_{x>0}.$$

Espérance a/b, variance a/b^2.

Construction : pour tout entier $n\geqslant 1$, la somme de n variables aléatoires indépendantes de loi $\mathcal{E}(b)$ suit la loi $G(n,b)$.

3.5.2 Loi beta $\mathcal{B}(a, b)$

Paramètres : a et b strictement positifs. Densité

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}\,\mathbf{1}_{0<x<1}.$$

Espérance $a/(a+b)$, variance $ab/((a+b)^2(a+b+1))$.

Construction : pour tous entiers $n \geqslant 1$ et $m \geqslant 1$, si on considère $n+m$ variables aléatoires indépendantes de loi $\mathcal{E}(b)$, le rapport de la somme des n premières sur la somme totale suit la loi $\mathcal{B}(n, m)$.

4 Exercices

4.1 Marginales

Soit (X, Y) un couple de variables aléatoires de densité

$$f(x, y) := \sqrt{x/y}\,\mathbf{1}_D(x, y), \qquad D := \{(x, y) \,|\, 0 < y \leqslant x \leqslant 1\}.$$

Vérifier que f est bien une densité sur $\mathbb{R}^2$.

Déterminer les lois marginales de X et Y. Préciser si les variables aléatoires X et Y sont indépendantes.

Calculer $P(Y < X/2)$.

Réaliser le couple (X, Y) à l'aide de deux variables aléatoires indépendantes à valeurs dans $[0, 1]$.

4.2 Marginales encore

Soit (X, Y) un couple de variables aléatoires de densité

$$f(x, y) := c\,\mathrm{e}^{-x} y^{-2}\,\mathbf{1}_{x>0}\,\mathbf{1}_{y>1}.$$

Calculer la constante c. Calculer $P(Y > 2, X < 1)$ et $P(XY > 1)$.

Déterminer les lois marginales de X et Y. Préciser si les variables aléatoires X et Y sont indépendantes.

4.3 Durées de vie

On fixe a strictement positif. Une usine fabrique des lampes dont la durée de vie T mesurée en heures vérifie $P(T > t) := \mathrm{e}^{-at}$ pour tout t positif.

Préciser la loi de T. Calculer la durée de vie moyenne de ces lampes et l'écart type associé.

On considère n lampes dont les durées de vie $(T_k)_{1\leqslant k\leqslant n}$ sont indépendantes et de même loi que T. On note $U := \min\{T_k \,|\, 1 \leqslant k \leqslant n\}$ le premier instant où au moins une des lampes cesse de fonctionner et $V := \max\{T_k \,|\, 1 \leqslant k \leqslant n\}$ le premier instant où toutes les lampes ont cessées de fonctionner.

Préciser les lois de U et de V.

4.4 Jacobien

Soient X et Y deux variables aléatoires indépendantes de même loi de densité la fonction f définie par $f(x) := x^{-2}\mathbf{1}_{x\geqslant 1}$. Soient $U := XY$ et $V := X/Y$.

Calculer la loi du couple (U, V) et ses lois marginales. Préciser si les variables U et V sont indépendantes.

4.5 Exponentielles

Soit X et Y deux variables indépendantes de lois exponentielles de paramètre respectif a et b strictement positifs. On pose $U := \min(X, Y)$, $V := \max(X, Y)$ et $W := V - U$.

Calculer $P(U = X)$ et montrer que U et W sont indépendantes.

4.6 Marginales enfin

Soit a un nombre réel strictement positif et (X, Y) un couple aléatoire de densité

$$f(x, y) := ce^{-ay}\,\mathbf{1}_D(x, y), \quad D := \{(x, y) \in \mathbb{R}^2 \,|\, 0 < x < y\}.$$

Préciser la valeur de c.

On pose $Z := X/Y$. Déterminer la loi du vecteur (Z, Y). Préciser si les variables Z et Y sont indépendantes. Déterminer leurs lois respectives.

Fiche 5 : Probabilités (4) Théorèmes limites

« Mathématiques : dessèchent le coeur. »
Gustave Flaubert, Dictionnaire des idées reçues

1 Théorèmes limites

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires à valeurs réelles $X_n : \Omega \to \mathbb{R}$. On rappelle que ces variables aléatoires sont indépendantes si et seulement si, pour tout $n \geqslant 1$ et tous $x_i \leqslant y_i$,

$$P(\forall 1 \leqslant i \leqslant n,\, x_i \leqslant X_i \leqslant y_i) = \prod_{i=1}^{n} P(x_i \leqslant X_i \leqslant y_i).$$

Théorème 1.1 (Loi forte des grands nombres). *Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes et de même loi admettant une espérance m. On note $S_n := X_1 + \cdots + X_n$. Alors il existe un événement A de probabilité 1 tel que pour tout $\omega \in A$, $S_n(\omega)/n \longrightarrow m$. On dit que S_n/n converge vers m presque sûrement.*

Corollaire 1.1 (Loi faible des grands nombres). *Sous les mêmes hypothèses, pour tout t strictement positif,*

$$P(nm - nt \leqslant S_n \leqslant nm + nt) \longrightarrow 1.$$

On dit que S_n/n converge vers m en probabilité.

Si on répète une même expérience un grand nombre de fois de manière identique et que l'on regarde le nombre de fois où un résultat R apparaît, la loi forte des grands nombres montre que la fréquence empirique d'apparition de R tend vers la probabilité de R quand le nombre d'expériences tend vers l'infini. Par exemple, si on lance une pièce équilibrée un grand nombre de fois, la suite des fréquences relatives des piles que l'on obtient tend avec probabilité 1 vers 50%. On peut se demander quelle est l'erreur faite quand on remplace directement la fréquence obtenue après n expériences par 50%. Le théorème 1.3 ci-dessous répond d'une certaine manière à cette question.

Nous commençons par quelques rappels sur les variables aléatoires gaussiennes. Une variable aléatoire X suit la loi $\mathcal{N}(m, v)$, donc est gaussienne (ou normale) d'espérance m et de variance v, si X admet pour densité la fonction f définie sur $\mathbb{R}$ par

$$f(x) := \frac{1}{\sqrt{2\pi v}} \mathrm{e}^{-(x-m)^2/(2v)}.$$

Quelques conséquences.

La transformée de Fourier d'une variable aléatoire X de loi $\mathcal{N}(m, v)$ vaut, en tout t réel,

$$\varphi_X(t) := E[e^{\mathrm{i}tX}] = \mathrm{e}^{\mathrm{i}mt}\mathrm{e}^{-vt^2/2}.$$

Pour tous réels a et b, si la loi de X est $\mathcal{N}(m, v)$, alors la loi de $aX+b$ est $\mathcal{N}(am+b, a^2v)$. (On rappelle que par convention $\mathcal{N}(m, 0)$ est la masse de Dirac en m.)

Si les variables aléatoires réelles X_k sont gaussiennes et indépendantes, alors pour tous nombres réels a_k, la variable aléatoire $\sum_k a_k X_k$ suit une loi gaussienne. L'hypothèse d'indépendance est importante, comme le montre le contrexemple ci-dessous.

Exercice 1.2. Soit T de loi de Bernoulli $b(p)$, donc $P(T = 1) = p$ et $P(T = 0) = 1 - p$. Soit X de loi gaussienne $\mathcal{N}(0, 1)$ et indépendante de T. Soit $Y = TX$ et $Z = X + Y$. Montrer que la loi de Y est gaussienne mais pas celle de Z.

Indication : calculer $P(Z = 0)$.

Nous sommes maintenant en mesure d'énoncer le théorème annoncé.

Théorème 1.3 (Théorème central limite). *Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes et de même loi de carré intégrable, d'espérance m et de variance v. On note $S_n = X_1 + \cdots + X_n$. Alors, pour tous x et y tels que $x \leqslant y$,*

$$P(nm + x\sqrt{nv} \leqslant S_n \leqslant nm + y\sqrt{nv}) \longrightarrow P(x \leqslant Z \leqslant y),$$

où Z désigne une variable aléatoire de loi gaussienne centrée réduite, donc

$$P(x \leqslant Z \leqslant y) = \frac{1}{\sqrt{2\pi}} \int_x^y \mathrm{e}^{-z^2/2}\mathrm{d}z.$$

2 Exemples d'application

Soit Z une variable aléatoire de loi gaussienne centrée réduite.

2.1 Binomiale

Si X_n suit une loi binomiale $B(n, p)$, alors, quand n est grand,

$$\frac{X_n - np}{\sqrt{np(1-p)}} \approx Z.$$

En effet, la loi de X_n est celle d'une somme de n variables aléatoires indépendantes de loi de Bernoulli $b(p)$, de moyenne p et de variance $p(1-p)$.

2.2 Poisson

Si X_a suit une loi de Poisson $\mathcal{P}(a)$, alors, quand a est grand,

$$\frac{X_a - a}{\sqrt{a}} \approx Z.$$

En effet, si $a = n$ est entier, la loi de X_n est celle d'une somme de n variables aléatoires indépendantes de loi de Poisson $\mathcal{P}(1)$, de moyenne 1 et de variance 1 et, si a n'est pas entier, on peut se ramener au cas entier (preuve omise).

2.3 Usage

En pratique, les statisticiens observent les règles d'usage suivantes.

1. Si $a \geqslant 10$, on peut approcher la loi de Poisson $\mathcal{P}(a)$ par la loi gaussienne $\mathcal{N}(a, a)$.
2. Si $np \geqslant 10$ et $n(1-p) \geqslant 10$, on peut approcher la loi binomiale $B(n, p)$ par la loi gaussienne $\mathcal{N}(np, np(1-p))$.
3. Si $np < 10$ et $p \leqslant \frac{1}{10}$, on peut approcher la loi binomiale $B(n, p)$ par la loi de Poisson $\mathcal{P}(np)$.
4. Si $n(1-p) < 10$ et $p \geqslant \frac{9}{10}$, on peut approcher la loi binomiale $B(n, p)$ par la loi de $n - X$ avec X de loi de Poisson $\mathcal{P}(n(1-p))$.

2.4 Fumeurs

Une proportion inconnue $0 < p < 1$ d'une certaine population est constituée de fumeurs. On utilise un tirage avec remise dans cette population pour déterminer le paramètre p. Un candidat naturel pour p est la proportion F_n de fumeurs tirés au cours de n tirages. On désire trouver p avec une erreur plus petite que $t := 0,5\%$, donc on souhaite assurer la réalisation de l'événement $A_n := \{p - t \leqslant F_n \leqslant p + t\}$. On veut déterminer combien de tirages sont nécessaires.

Tout d'abord, il faut réaliser qu'aucun nombre de tirages ne garantit que A_n est réalisé avec certitude : il est possible par exemple de ne tirer que des fumeurs pendant les n premiers tirages, auquel cas $F_n = 1$, mais ceci se produit avec probabilité p^n donc cette éventualité devient très improbable si n devient grand. On peut seulement rendre A_n très probable, donc on se contente de fixer un niveau de confiance a, par exemple $a = 95\%$, et de chercher un entier n tel que $P(A_n) \geqslant a$.

Notons $X_i = 1$ si la personne tiré au tirage numéro i est fumeur et $X_i = 0$ sinon. La suite $(X_n)_{n \geqslant 1}$ est une suite de variables aléatoires indépendantes et de même loi de Bernoulli

de paramètre p. Soit $S_n := X_1 + \cdots + X_n$ le nombre de fumeurs tirés jusqu'au temps n, et

$$T_n := \frac{S_n - np}{\sqrt{np(1-p)}}.$$

On a donc $S_n = nF_n$ et on cherche n tel que $P(A_n) \geqslant a$. Or,

$$A_n = \left\{ |T_n| < t\sqrt{n/(p(1-p))} \right\} \supset \left\{ |T_n| < 2t\sqrt{n} \right\},$$

car $\sqrt{p(1-p)} \leqslant 1/2$. Soit Z une variable aléatoire gaussienne centrée réduite et z_a un nombre réel tel que $P(|Z| \geqslant z_a) = a$.

La façon usuelle d'utiliser le théorème central limite est de décréter que $2t\sqrt{n} = z_a$ doit suffire, donc de choisir $n = z_a^2/(4t^2)$.

Si $a = 95\%$ on trouve grâce à une table de la distribution normale standard que $z_a = 1,96$. Finalement, $n = 40000$ convient.

Remarque : en toute rigueur, le théorème central limite est valide pour un niveau d'erreur fixé et quand n tend vers l'infini ; ici, on l'utilise quand n et le taux d'erreur $2t\sqrt{n}$ tendent tous les deux vers l'infini en même temps, donc en dehors de son domaine théorique de validité, comme le montre l'exercice très simple ci-dessous.

Exercice : exhiber une famille $(x_{n,k})_{n,k}$ de nombres réels tels que, pour tout k fixé, $x_{n,k} \longrightarrow 0$ quand $n \to \infty$, et pourtant $x_{n,n}$ ne tend pas vers 0 quand $n \to \infty$.

3 Exercices

3.1 Arrondis

Un ordinateur effectue les calculs avec 9 chiffres significatifs et arrondit les résultats théoriques à cette précision.

On suppose qu'il effectue 10^6 opérations élémentaires, que les erreurs d'arrondi à chaque opération sont indépendantes et de loi uniforme sur l'intervalle $[-\frac{1}{2}10^{-9}, \frac{1}{2}10^{-9}]$, et que l'erreur sur le résultat final est la somme des erreurs commises à chaque opération.

On demande d'évaluer la probabilité pour que l'erreur finale soit en valeur absolue inférieure à $\frac{1}{2}10^{-6}$.

On indique que $P(0 \leqslant Z \leqslant \sqrt{3}) \simeq 46\%$ pour une variable aléatoire Z gaussienne centrée réduite.

3.2 Preuve probabiliste de la formule de Stirling

Soit Z une variable aléatoire de loi gaussienne centrée réduite et $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes de loi de Poisson de paramètre 1. On pose $S_n = \sum_{k=1}^{n} X_k$

et $T_n = (S_n - n)/\sqrt{n}$.

a) Déterminer la loi de S_n, son espérance et sa variance.

b) Montrer que pour tout $a \leqslant b$, $P(a \leqslant T_n \leqslant b)$ tend vers $P(a \leqslant Z \leqslant b)$.

c) En déduire que $\mathrm{e}^{-n} \displaystyle\sum_{k=0}^{n} \frac{n^k}{k!} \longrightarrow \frac{1}{2}$.

d) Montrer que $\displaystyle\int_0^\infty P(T_n \geqslant x)\,\mathrm{d}x \longrightarrow \int_0^\infty P(Z \geqslant x)\,\mathrm{d}x$ et calculer cette limite.

e) Montrer que $\displaystyle\int_0^\infty P(T_n \geqslant x)\,\mathrm{d}x = \mathrm{e}^{-n} n^n \sqrt{n}/n!$.

f) En déduire la formule de Stirling $n! \sim n^n \mathrm{e}^{-n}\sqrt{2\pi n}$.

3.3 Processus de comptage

Pour tout temps $t \geqslant 0$, N_t désigne le nombre de clients qui arrivent dans un service pendant l'intervalle de temps $]0, t]$

a) Que représente, pour t et h strictement positifs, la quantité $N_{t+h} - N_t$?

On suppose qu'il existe un nombre réel a strictement positif tel que, quand $h \to 0$, pour tout t, $P(N_{t+h} - N_t = 1) = ah + o(h)$ et $P(N_{t+h} - N_t \geqslant 2) = o(h)$, et que si $t < s$, les variables aléatoires $N_s - N_t$ et N_t sont indépendantes.

b) Soit $p_n(t) = P(N_t = n)$ et $q_n(t, h) = P(N_{t+h} - N_t = n)$. On suppose que p_n est une fonction dérivable de t.

Montrer que $q_0(t, h) = 1 - ah + o(h)$, $q_1(t, h) = ah + o(h)$ et $q_n(t, h) = o(h)$ pour $n \geqslant 2$.

Montrer que $p_0(t + h) = q_0(t, h)p_0(t)$.

En déduire que $p_0'(t) = -ap_0(t)$.

c) Montrer que, pour $n \geqslant 1$, $p_n(t + h) = \displaystyle\sum_{k=0}^{n} p_k(t)q_{n-k}(t, h)$.

En déduire que $p_n'(t) = -a\,(p_n(t) + p_{n-1}(t))$.

d) Montrer que pout tout $n \geqslant 0$, $p_n(t) = \mathrm{e}^{-at}(at)^n/n!$.

e) Soit T_1 l'instant où arrive le premier client. Montrer que T_1 suit une loi exponentielle de paramètre a.

3.4 Partie entière

Soit X une variable aléatoire de loi uniforme sur $[0, 1]$, Y la partie entière de $10X$ et Z la partie entière de $100X - 10Y$.

Préciser ce que représentent Y et Z.

Trouver les lois de (Y, Z), de Y et de Z. Préciser si Y et Z sont indépendantes. Proposer une généralisation de ces résultats et en déduire une façon de simuler une variable aléatoire de loi uniforme sur $[0, 1]$ à partir d'une suite de variables aléatoires indépendantes de loi uniforme sur $\{0, 1, \ldots, 9\}$.

Déduire de tout ceci une façon de simuler une variable aléatoire de loi uniforme sur $[0, 1]$ à partir d'une suite de variables aléatoires indépendantes de loi de Bernoulli $\frac{1}{2}(\delta_0 + \delta_1)$.

Fiche 6 : Probabilités (5) Applications statistiques

« There are three kinds of lies : lies, damned lies, and statistics. » Attribué à Benjamin Disraeli par Mark Twain dans son Autobiographie

1 Estimation statistique

Dans la démarche statistique, on cherche à traiter des données, en général une série d'observations x_1, x_2, …, x_n. On suppose qu'elles correspondent à des réalisations de variables aléatoires X_1, X_2, …, X_n, et on cherche une distribution théorique du vecteur $(X_k)_{1\leqslant k\leqslant n}$ qui reflète correctement les propriétés des observations $(x_k)_{1\leqslant k\leqslant n}$.

Concrètement, les valeurs x_k sont donc connues. Par exemple, x_k représente la durée de vie du moteur de la voiture numéro k que l'on a choisi d'observer. Si le fabriquant est un grand constructeur, on ne peut pas recenser la durée de vie de toutes les voitures fabriquées donc on ne considère qu'un échantillon de taille raisonnable. Le constructeur aimerait à partir de ces valeurs améliorer la fabrication des moteurs. Il serait alors utile de connaitre la loi sous-jacente à la durée de vie d'un moteur. Cette loi est inconnue, mais à partir de l'échantillon $(x_k)_k$, on peut cependant estimer certaines valeurs, comme par exemple la durée de vie moyenne d'un moteur ; on parle alors de problèmes d'estimation et d'intervalle de confiance. Se pose parfois ensuite la question de la validité de l'estimation ; on parle alors de problème de test.

Dans toute la suite, on se restreint au cas le plus simple : **on suppose que les variables aléatoires $(X_k)_{1\leqslant k\leqslant n}$ sont indépendantes et de même loi et que cette loi appartient à une collection fixée a priori, que l'on notera $\{P_\theta\,;\,\theta\in\Theta\}$.**

Un premier exemple : les tirages avec remise (TAR). On considére une pièce de monnaie dont on ignore si elle est truquée ou non, et on veut connaitre la probabilite p de tomber sur « pile ». On lance n fois la pièce et on note X_i le résultat du lancer numéro i, avec $X_i = 1$ si on obtient « pile » (ce qui se produit avec probabilité p) et $X_i = 0$ si on obtient « face » (ce qui se produit avec probabilité $1-p$). Les variables aléatoires X_i sont indépendantes et de même loi de Bernoulli $b(p)$ donc, sur n lancers, le nombre total de « piles » $S_n = X_1 + \cdots + X_n$ suit la loi binomiale $B(n,p)$.

Un deuxième exemple : les tirages sans remise (TSR). On considère une population de taille N divisée en deux classes selon leur caractère, par exemple les personnes de rhésus

positif et celles de rhésus négatif. Il y a n_1 individus dans la classe 1 et $N - n_1$ individus dans la classe 2 mais les quantités n_1 et N sont grandes et inconnues. Pour estimer n_1 ou N ou les deux, on considère une sous-population de taille raisonnable n. On compte alors le nombre de personnes de chaque classe dans cette sous-population. Si X_k désigne la classe de l'individu obtenu au tirage numéro k et si la sous-population est choisie de façon uniforme, on voit que les variables aléatoires X_k sont de même loi mais non indépendantes. En effet, $P(X_1 = 1) = P(X_2 = 1) = p_1$ avec $p_1 := n_1/N$ mais

$$P(X_1 = X_2 = 1) = \frac{n_1(n_1 - 1)}{N(N - 1)} \neq p_1^2.$$

On considèrera donc plutôt directement le nombre $0 \leqslant X \leqslant n$ d'individus de la classe 1 parmi les n individus interrogés et on sait que la loi de X est hypergéométrique $H(n_1, n, N)$.

1.1 Estimation ponctuelle

On considère un échantillon $(X_k)_{1 \leqslant k \leqslant n}$ tel que chaque X_k est à valeurs dans un ensemble E (par exemple, $E = \mathbb{R}$) et de même loi P_θ qui dépend d'un paramètre inconnu $\theta \in \Theta$. Le but est d'estimer la valeur de θ à partir des valeurs des X_k.

Dans l'exemple TAR, le paramètre inconnu est $\theta = p$, dans l'exemple TSR, le paramètre inconnu est $\theta = (n_1, N)$.

Définition 1.1. *Un estimateur de θ est une variable aléatoire $\widehat{\theta}_n$ telle qu'il existe une fonction $F_n : E^n \to \Theta$ avec $\widehat{\theta}_n = F_n(X_1, X_2, \ldots, X_n)$; c'est donc une fonction de l'échantillon.*

Attention : un estimateur de θ ne doit pas dépendre de θ, mais seulement des observations $(X_k)_{1 \leqslant k \leqslant n}$.

Le problème est maintenant de choisir la fonction F_n de façon à estimer correctement θ.

Dans l'exemple TAR, $\widehat{\theta}_n = 10$ et $\widehat{\theta}_n = X_1 X_2$ sont des estimateurs de $\theta = p$, tous deux un peu stupides (dire pourquoi). On préfèrera utiliser l'estimateur naturel de p fourni par la fréquence empirique des succès $\overline{X}_n = S_n/n$, qui vaut le nombre de « piles » divisé par le nombre de lancers.

Pour tout θ dans Θ, on note E_θ l'espérance par rapport à la loi P_θ.

Définition 1.2. *L'estimateur $\widehat{\theta}_n$ est un estimateur consistant de θ pour la valeur θ si $\widehat{\theta}_n \longrightarrow \theta$ presque sûrement par rapport à P_θ, quand n tend vers l'infini.*

La consistance est la propriété la plus importante d'un estimateur. Ainsi, pour de grands échantillons, l'approximation de θ par $\widehat{\theta}_n$ est correcte. Par exemple, la loi forte des grands nombres affirme que, dans l'exemple TAR, $\overline{X}_n$ est un estimateur consistant de p.

Définition 1.3. *Le biais d'un estimateur $\widehat{\theta}_n$ de θ est*

$$B(\widehat{\theta}_n, \theta) := E_\theta(\widehat{\theta}_n) - \theta.$$

Le biais $B(\widehat{\theta}_n, \theta)$ est donc un nombre, qui dépend de θ et de la fonction F_n définissant l'estimateur $\widehat{\theta}_n$.

L'estimateur $\widehat{\theta}_n$ est dit sans biais si $B(\widehat{\theta}_n, \theta) = 0$ pour toute valeur de θ dans Θ, sinon l'estimateur est dit biaisé.

La valeur moyenne d'un estimateur sans biais est notre inconnue θ. Dans l'exemple TAR, $\overline{X}_n$ est un estimateur sans biais de p.

On veut maintenant comparer différents estimateurs de θ.

Définition 1.4. *Le risque quadratique de l'estimateur $\widehat{\theta}_n$ par rapport à θ est défini par*

$$R(\widehat{\theta}_n, \theta) := E_\theta\left[(\widehat{\theta}_n - \theta)^2\right].$$

Si le risque est faible, l'estimateur est souvent proche de l'inconnue θ, donc on souhaite avoir un risque le plus faible possible.

Définition 1.5. *L'estimateur $\widehat{\theta}_n$ est (quadratiquement) meilleur que l'estimateur $\widetilde{\theta}_n$ pour la valeur θ dans Θ, si $R(\widehat{\theta}_n, \theta) \leqslant R(\widetilde{\theta}_n, \theta)$.*
L'estimateur $\widehat{\theta}_n$ est (quadratiquement) uniformément meilleur que l'estimateur $\widetilde{\theta}_n$ si $\widehat{\theta}_n$ est (quadratiquement) meilleur que $\widetilde{\theta}_n$ pour tout θ dans Θ.

Attention : on ne peut pas toujours comparer uniformément deux estimateurs.

On remarque que

$$R(\widehat{\theta}_n, \theta) = \mathrm{var}_\theta(\widehat{\theta}_n) + B(\widehat{\theta}_n, \theta)^2.$$

Par conséquent, si l'estimateur est sans biais, $R(\widehat{\theta}_n, \theta) = \mathrm{var}_\theta(\widehat{\theta}_n)$. C'est aussi la raison pour laquelle on préfère souvent les estimateurs sans biais.

Attention : on pourrait donc penser utiliser $\widehat{\theta}_n - B(\widehat{\theta}_n, \theta)$ comme estimateur au lieu de $\widehat{\theta}_n$ puisque le risque quadratique est plus petit :

$$R(\widehat{\theta}_n - B(\widehat{\theta}_n, \theta), \theta) = \mathrm{var}_\theta(\widehat{\theta}_n) \leqslant \mathrm{var}_\theta(\widehat{\theta}_n) + B(\widehat{\theta}_n, \theta)^2 = R(\widehat{\theta}_n, \theta).$$

Mais $\widehat{\theta}_n - B(\widehat{\theta}_n, \theta)$ n'est pas en général un estimateur ! En effet, le terme de biais $B(\widehat{\theta}_n, \theta)$ dépend de $\widehat{\theta}_n$, ce qui est parfait puisque $\widehat{\theta}_n$ ne dépend que de l'échantillon observé, mais aussi de θ que l'on ne connaît pas.

1.2 Estimateurs des moments

Un principe général pour estimer θ dans deux situations particulières.

Premier cas : si $\theta = E[\varphi(X_1)]$, alors $\widehat{\theta}_n := \dfrac{1}{n}\sum_{k=1}^{n}\varphi(X_k)$ est un estimateur sans biais et consistant.

Deuxième cas : si $\theta = \psi(E[X_1])$ avec ψ continue, alors $\widehat{\theta}_n := \psi(\overline{X}_n)$ est un estimateur, biaisé en général, mais consistant.

1.2.1 Moyenne empirique

Si $\theta = E(X_1)$, on pourra utiliser la moyenne empirique, définie par

$$\overline{X}_n := \frac{1}{n}\sum_{k=1}^{n} X_k.$$

La moyenne empirique est un estimateur sans biais de la moyenne $E(X)$, consistant grâce à la loi des grands nombres, de risque

$$R(\overline{X}_n, \theta) = \frac{1}{n}\mathrm{var}_\theta(X_1).$$

Pour un échantillon de loi de Bernoulli $b(p)$, la loi de $n\overline{X}_n$ est binomiale $B(n,p)$. Pour un échantillon de loi normale $\mathcal{N}(\theta, v)$, la loi de $\overline{X}_n$ est normale $\mathcal{N}(\theta, v/n)$.

1.2.2 Variance et covariance empiriques

En appliquant ce qui précède aux variables aléatoires $(X_k^2)_k$, on voit qu'un estimateur consistant de $E(X_1^2)$ vaut

$$\frac{1}{n}\sum_{k=1}^{n} X_k^2.$$

Comme $\mathrm{var}(X) = E(X^2) - E(X)^2$, on obtient un estimateur consistant de $\mathrm{var}(X_1^2)$ en posant

$$\widehat{\sigma}_n^2 := \frac{1}{n}\sum_{k=1}^{n} X_k^2 - \overline{X}_n^2 = \frac{1}{n}\sum_{i=k}^{n}\left(X_k - \overline{X}_n\right)^2.$$

Mais $E(\widehat{\sigma}_n^2) = v\,\frac{n-1}{n} \neq v$ donc $\widehat{\sigma}_n^2$ est biaisé.

Proposition 1.6. *Pour tout échantillon de taille $n \geqslant 2$, indépendant et de même loi de carré intégrable et de variance $v = \mathrm{var}(X_1)$,*

$$V_n := \frac{1}{n-1}\sum_{k=1}^{n}\left(X_k - \overline{X}_n\right)^2$$

est un estimateur sans biais et consistant de v.

Pour la covariance, on considère un échantillon $(X_k, Y_k)_{1\leqslant k\leqslant n}$ indépendant et de même loi de carré intégrable dont on veut estimer la covariance

$$C = \mathrm{cov}(X_1, Y_1) = E(X_1Y_1) - E(X_1)E(Y_1).$$

Alors $\widehat{C}_n$ est un estimateur sans biais et consistant de C, si on pose

$$\widehat{C}_n := \frac{1}{n-1}\sum_{k=1}^{n} X_kY_k - \overline{X}_n\overline{Y}_n.$$

Exercice : Où doit-on mettre les parenthèses dans $\widehat{C}_n$ pour que cet estimateur soit effectivement sans biais ?

2 Estimation par intervalle de confiance

Dans la section précédente on proposait une valeur unique $\widehat{\theta}_n$ pour estimer θ. On veut maintenant proposer un ensemble $I_n \subset \Theta_n$ aussi petit que possible, tel que θ appartienne souvent à I_n.

Comme précédemment, on ne dispose pour construire I_n que des observations $(X_k)_{1\leqslant k\leqslant n}$, que l'on suppose indépendantes et de même loi P_θ pour une certaine valeur inconnue du paramètre $\theta \in \Theta$.

Définition 2.1. *Un intervalle de confiance au niveau a est un intervalle aléatoire I_n qui ne dépend que de l'échantillon $X_1,\ldots,X_n$, mais pas de θ, et tel que, pour tout θ dans Θ,*

$$P_\theta(\theta \in I_n) \geqslant a.$$

Le nombre $1 - a$ représente le taux d'erreur maximal que l'on accepte en prédisant que I_n contient θ.

Une façon de construire des intervalles de confiance consiste à considérer un estimateur $\widehat{\theta}_n$ raisonnable de θ et à trouver sa loi sous chaque P_θ.

Si $P_\theta(\widehat{\theta}_n \in [\theta - s_n, \theta + t_n]) \geqslant a$ pour tout θ dans Θ, alors $I_n = \left[\widehat{\theta}_n - t_n, \widehat{\theta}_n + s_n\right]$ est un intervalle de confiance pour θ au niveau a.

Enfin le meilleur intervalle de confiance est celui dont la longueur est la plus petite.

Définition 2.2. *Soit $(I_n)_{n\geqslant 1}$ une suite d'intervalles de confiance, donc chaque I_n ne dépend que de l'échantillon $X_1,\ldots,X_n$. Le niveau de confiance asymptotique de la suite $(I_n)_n$ vaut a si, pour tout θ dans Θ,*

$$\lim_{n\to+\infty} P_\theta(\theta \in I_n) = a.$$

2.1 Estimation de la moyenne avec une variance connue

Soit $(X_k)_{1\leqslant k\leqslant n}$ un échantillon indépendant et de même loi P_θ de carré intégrable pour une certaine valeur θ dans Θ. On pose

$$m := E_\theta(X_1), \qquad v := \mathrm{var}_\theta(X_1).$$

La moyenne empirique $\overline{X}_n$ est un estimateur sans biais et consistant de m. De plus, d'après le théorème central limite, pour tout x positif,

$$P_\theta\left(\sqrt{n}\,|\overline{X}_n - m| \leqslant x\sqrt{v}\right) \underset{n\to+\infty}{\longrightarrow} P\left(|Z| \leqslant x\right),$$

où Z désigne une variable aléatoire de loi gaussienne centrée réduite $\mathcal{N}(0,1)$. En utilisant la table de la loi normale, on choisit x_a tel que $P\left(|Z| \leqslant x_a\right) = a$. Par conséquent,

$$I_n := \left[\overline{X}_n - x_a\sqrt{v/n}, \overline{X}_n + x_a\sqrt{v/n}\right]$$

est un intervalle de confiance asymptotique de m au niveau a.

Par exemple, pour $a = 95\%$, $x_a = 1,96$ convient.

2.2 Estimation de la moyenne avec une variance inconnue

Par exemple, soit $(X_k)_k$ un échantillon indépendant et de même loi exponentielle $\mathcal{E}(1/\theta)$. Donc $m = \theta$ et $v = \theta^2$ et on cherche à estimer θ.

D'après le théorème de la limite centrale, en gardant les notations de la section précédente, pour n assez grand,

$$P_\theta\left(\sqrt{n}|\overline{X}_n - \theta| \leqslant x_a\theta\right) \approx a.$$

donc un intervalle de confiance au niveau asymptotique a, défini pour tout $n > x_a^2$, est

$$I_n := \left[\frac{\overline{X}_n}{1 + x_a/\sqrt{n}}, \frac{\overline{X}_n}{1 - x_a/\sqrt{n}}\right].$$

De manière générale, lorsque l'on ne peut pas faire autrement, on estime v par V_n et on utilise le fait que $\left(\overline{X}_n - m\right)\sqrt{n/V_n}$ suit asymptotiquement une loi gaussienne centrée réduite $\mathcal{N}(0,1)$.

3 Tests

3.1 Principe général

On s'intéresse à la répartition des enfants nouveaux-nés en garçons et filles. On dispose des résultats d'un sondage, selon lequel sur 429440 naissances, on a dénombré 221023

filles. On se demande si cette répartition entre filles et garçons est compatible avec l'hypothèse d'équiprobabilité de naissance des garçons et des filles.

On dispose donc d'un ensemble Θ de paramètres, d'une valeur particulière θ_0, et d'un échantillon $(X_k)_{1\leqslant k\leqslant n}$ de loi P_θ pour un paramètre θ inconnu. On veut pouvoir résoudre l'alternative « H_0 contre H_1 », avec

$$H_0 : \theta = \theta_0; \qquad H_1 : \theta \neq \theta_0.$$

L'hypothèse H_0 est appelée l'hypothèse nulle, et H_1 l'hypothèse alternative.

Dans la situation ci-dessus l'échantillon $(X_k)_{1\leqslant k\leqslant n}$ est indépendant et suit la loi de Bernoulli $b(\theta)$, et $\theta_0 = 50\%$.

Pour tester H_0 contre H_1, on définit une zone de rejet R ne dépendant que de $(X_k)_{1\leqslant k\leqslant n}$ et on adopte la stratégie suivante :

Si R est réalisée, on rejette l'hypothèse H_0 et on accepte $H_1 : \theta \neq \theta_0$.
Si R n'est pas réalisée, on accepte l'hypothèse $H_0 : \theta = \theta_0$.

Définition 3.1. *Le niveau de risque de première espèce de R dans le test de H_0 contre H_1 est $\alpha := \sup\{P_\theta(R)\,;|\theta \in H_0\}$.*

Le niveau de risque de première espèce mesure donc le risque de rejeter l'hypothèse H_0 alors qu'elle est réalisée. En général, on fixe un niveau de risque de l'ordre de $\alpha = 5\%$ et on calcule une zone de rejet R adaptée.

Quand $H_0 : \theta = \theta_0$, on obtient $\alpha := P_{\theta_0}(R)$.

On peut utiliser des intervalles de confiance pour calculer des zones de rejet. Supposons par exemple que I est un intervalle de confiance pour θ au niveau de confiance $1-\alpha$, c'est-à-dire que, pour tout θ dans Θ,

$$P_\theta(\theta \in I) \geqslant 1 - \alpha.$$

Alors $R = \{\theta_0 \notin I\}$ est une zone de rejet pour le test de $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$ au niveau de risque α.

L'autre erreur possible associée à une zone de rejet R donnée consiste à accepter H_0 alors que H_0 est fausse.

Définition 3.2. *Le niveau de risque de seconde espèce de R dans le test de H_0 contre H_1 est $\beta := \max\{P_\theta(R)\,;|\theta \in H_1\}$.*

La première erreur α étant en général fixée, une bonne zone de rejet minimise l'erreur de seconde espèce β.

3.2 Estimation du paramètre d'une loi binomiale

L'échantillon $(X_n)_{n\geqslant 1}$ suit la loi de Bernoulli de paramètre p et le paramètre p est inconnu. On fixe p_0 et veut tester $H_0 : p > p_0$ contre $H_1 : p \leqslant p_0$.

Un test de H_0 contre H_1 au niveau de risque α est associé à une zone de rejet R telle que $P_p(R) \leqslant \alpha$ pour tout $p > p_0$. On rappelle que $\overline{X}_n = \dfrac{1}{n}\sum_{i=k}^{n} X_k$ et on cherche R de la forme $R = \{\overline{X}_n < x\}$, alors

$$\alpha = \sup\{P_p(R) \,|\, p > p_0\} = P_{p_0}(R).$$

Remarque 3.3 (À omettre en première lecture). *Pour évaluer $P_{p_0}(R)$, on peut procéder comme suit. Pour tout $x < p_0$ et tout s dans $]0,1]$, l'inégalité de Crámer donne*

$$P_{p_0}(\overline{X}_n < x) \leqslant s^{nx} E_{p_0}(s^{-X_1})^n,$$

Comme $E_{p_0}(s^{-X_1}) = p_0 s^{-1} + 1 - p_0$, on peut calculer le minimum en x du membre de droite, ce qui donne après calculs,

$$P_{p_0}(\overline{X}_n < x) \leqslant \left(\frac{1-p_0}{1-x}\right)^{n(1-x)} \left(\frac{p_0}{x}\right)^{nx}.$$

On vérifie que le membre de droite est une fonction croissante de x sur l'intervalle $[0, p_0]$ et vaut 1 en $x = p_0$, donc est bien inférieur à 1 pour tout $x < p_0$ Il reste à choisir x de sorte que la valeur du membre de droite soit au plus α, ce qui est possible pour tout α supérieur à la valeur en $x = 0$, soit $\alpha \geqslant (1-p_0)^n$.

Rappelons que $\overline{X}_n$ est un estimateur consistant et sans biais de p.

3.3 Estimation de la moyenne d'une loi normale

L'échantillon $(X_n)_{n\geqslant 1}$ suit la loi normale $\mathcal{N}(m, v)$ avec v connue et m inconnue.

Rappelons que $\overline{X}_n$ est un estimateur consistant et sans biais de m et que $\overline{X}_n = m + Z_n\sqrt{v/n}$ où Z_n est asymptotiquement gaussienne centrée réduite.

Première situation On veut tester $H_0 : m > m_0$ contre $H_1 : m \leqslant m_0$.

Il s'agit d'un test unilatère. On cherche une zone de rejet $R := \{\overline{X}_n < x\}$ avec $x < m_0$. Alors,

$$\alpha = \sup\{P_m(R) \,|\, m > m_0\} = P_{m_0}\left(Z < -(m_0 - x)\sqrt{n/v}\right),$$

et il reste à consulter la table de la loi gaussienne pour choisir x. Par exemple, $\alpha = 5\%$ donne

$$x = m_0 - 1,96\sqrt{v/n}.$$

Deuxième situation On veut tester $H_0 : m \in [m_1, m_2]$ contre $H_1 : m \notin [m_1, m_2]$.

Il s'agit d'un test bilatère. On cherche une zone de rejet $R := \{\overline{X}_n \notin [x_1, x_2]\}$ avec $x_1 < m_1$ et $x_2 > m_2$. Si on suppose que $x_1 = m_1 - c$ et $x_2 = m_2 + c$ avec c positif, on montre qu'il s'agit de choisir c tel que

$$\alpha = P(Z \geqslant c\sqrt{n/v}) + P(Z \geqslant (c + m_2 - m_1)\sqrt{n/v}).$$

Les cas limites sont $m_2 = +\infty$ pour lequel on retrouve le test unilatère et $m_2 = m_1$ qui donne le test de $H_0 : m = m_1$ contre $H_1 : m \neq m_1$.

4 Exercices

4.1 Pièces

On lance une pièce équilibrée et on souhaite obtenir une proportion de « piles » entre 49% et 51% avec une probabilité au moins égale à 96%. Déterminer le nombre de jets nécessaire en utilisant l'approximation par une loi normale.

4.2 Repas

Un restaurateur peut servir 75 repas, uniquement sur réservation. En pratique, 20% des clients ayant réservé ne viennent pas. Le restaurateur souhaite pouvoir servir tous les clients qui se présentent avec une probabilité supérieure ou égale à 90%. Déterminer le nombre maximal de réservations que le restaurateur peut accepter.

4.3 Défauts

Une entreprise reçoit un lot important de pièces fabriquées en série. L'entreprise n'accepte la livraison que si la proportion p de pièces défectueuses est inférieure à 5%. Dans un échantillon de 200 pièces, on observe que 15 pièces sont défectueuses.
Décrire la conclusion d'un test de $H_0 : p \leqslant 5\%$ contre $H_1 : p > 5\%$ au niveau 1% relatif à une région de rejet $R = \{X \geqslant x\}$, où X désigne le nombre de pièces défectueuses dans l'échantillon.

4.4 Médicament

L'écart type de la teneur d'un composant dans un médicament est de 3 milligrammes. Un nouveau procédé de fabrication vise à diminuer cet écart type. Dans un échantillon de 10 unités fabriquées par le nouveau procédé, on obtient en milligrammes :

725, 722, 727, 718, 723, 731, 719, 724, 726, 726.

On supppose l'échantillon de loi $\mathcal{N}(m, v)$.
1) On suppose que m est connue et vaut 724. Donner un intervalle de confiance de niveau 95% pour la variance.
2) Dans cette question m est inconnue, tester au niveau 5% si le but recherché est atteint.

4.5 Électricité

Dans une fabrique de compteurs électriques, on vérifie le réglage des compteurs sur un échantillon de 10 compteurs. Lors d'une mesure de 100 unités, les compteurs de l'échantillon enregistrent :

983, 1002, 998, 996, 1002, 983, 994, 991, 1005, 986.

On supppose l'échantillon de loi $\mathcal{N}(m, \sigma^2)$.

Donner un intervalle de confiance au niveau 95% pour la moyenne.

Tester l'hypothèse $m = 1000$ contre $m \neq 1000$ au niveau 5%.

4.6 Confiance

Échantillon de loi de densité $h_\theta(x) = (2x/\theta^2)\mathbf{1}_{0\leqslant x\leqslant\theta}$. On suppose que n est grand et $0 < \theta \leqslant 2$. Donner un intervalle de confiance pour θ au niveau de confiance 95% basé sur $\overline{X}_n$.

4.7 Risque

Échantillon de loi de densité $f_\theta(x) = (2\theta x - \theta + 1)\mathbf{1}_{0\leqslant x\leqslant 1}$, où $-1 \leqslant \theta \leqslant 1$ est un paramètre que l'on se propose d'estimer.

Trouver a et b tels que l'estimateur $T_n = a\overline{X}_n + b$ est sans biais, pour tout θ. Calculer le risque quadratique de T_n.

Calculer la limite de $P_\theta(\sqrt{n}|T_n - \theta| \leqslant x)$ quand $n \to +\infty$, pour x réel positif. En déduire un intervalle de confiance de niveau 99% pour θ.

5 Exercices supplémentaires

Exercice 1 Un échantillon de 478 électeurs choisis aléatoirement indique que 255 d'entre eux vont voter pour A. Évaluer des intervalles de confiance à 1% et à 5% pour la proportion d'électeurs de A.

Exercice 2 On effectue un contrôle de fabrication sur des pièces dont une proportion p est défectueuse. On contrôle un lot de 200 pièces et on trouve 20 pièces défectueuses.

Donner des intervalles de confiance pour l'estimation de p, au niveau 95% puis au niveau 99%.

Exercice 3 Des appareils électriques de chauffage ont une moyenne de vie de fonctionnement de 20 000 heures avec un écart-type de 7 000 heures. À l'aide d'un changement de composant, le fabricant affirme que la durée de vie moyenne peut être accrue.

On a testé un échantillon de 127 appareils et on a observé une durée de vie moyenne de 21 000 heures. Préciser si on peut soutenir cette affirmation au risque de 5%, au risque de 1%.

Exercice 4 Une pièce jetée 660 fois tombe 312 fois sur pile, préciser si on doit penser que cette pièce est bien équilibrée, ou non.

Exercice 5 Le fabricant d'une nouvelle solution anti-rouille annonce que son produit est efficace à 90%. Dans un échantillon de 500 pièces le résultat est probant pour 420 d'entre elles. Préciser si l'affirmation du fabricant est légitime.

Exercice 6 Deux machines A et B fabriquent en série la même pièce. Lors d'une expertise de la production, on remarque que la machine A a produit 2700 pièces dont 50 sont défectueuses alors que sur les 1600 pièces produites par la machine B, 35 sont défectueuses. Préciser si on doit en conclure que la machine A est mieux réglée que la machine B, ou non.

Fiche Probabilités : Corrigés

Rédactions plus qu'abrégées, à ne pas donner aux étudiants.

Fiche 3

1. On écrit chaque $P(X \geqslant k)$ comme la somme des $P(X = i)$ de $i = k$ à $i = n$ puis on intervertit les deux sommes.

2. Donc $P(T \geqslant n + k) = P(T \geqslant n)P(T \geqslant k)$ pour tous $k \geqslant 0$ et $n \geqslant 0$. En itérant, $P(T \geqslant n) = P(T \geqslant 1)^n$ pour tout $n \geqslant 0$. En notant $a := P(T \geqslant 1)$, il vient

$$P(T = n) = P(T \geqslant n) - P(T \geqslant n + 1) = a^n(1 - a),$$

pour tout $n \geqslant 0$. Si $a = 0$, $T = 0$ partout, ce qui est une solution dégénérée. Si $a = 1$, $P(T = n) = 0$ pour tout $n \geqslant 0$, ce qui est impossible. Si $0 < a < 1$, la loi de T est géométrique de paramètre a.

3. a) Soit e_i le vecteur canonique numéro i de $\mathbb{R}^k$. Alors X suit la loi de la somme de n variables aléatoires indépendantes valant e_i avec probabilité $p_i := N_i/N$, donc la loi de X est multinomiale $(n, (p_i)_{1 \leqslant i \leqslant k})$.

b) La loi de X_i est binomiale $B(n, p_i)$.

c) La loi de (X_i, X_j) correspond aux deux premières coordonnées de la loi multinomiale $(n, (p_i, p_j, 1 - p_i - p_j))$. Donc, pour tous x et y entiers avec $0 \leqslant x, y \leqslant n$ et $x + y \leqslant n$,

$$P(X_i = x, X_j = y) = \binom{n}{x, y, n - x - y} p_i^x p_j^y (1 - p_i - p_j)^{n-x-y}.$$

4. a) $N_1 = k$ avec $k \geqslant 1$ signifie que $X_n = 0$ pour tout $1 \leqslant i \leqslant k - 1$ et que $X_k = 1$, donc $P(N_1 = k) = p(1 - p)^{k-1}$ et N_1 suit une loi géométrique décalée.

b) $N_1 = k$ et $N_2 = \ell$ avec $k \geqslant 1$ et $\ell \geqslant k + 1$ signifie que $X_n = 0$ pour tout $1 \leqslant i \leqslant k + \ell$ sauf pour $i = k$ et $i = \ell$, donc $P(N_1 = k, N_2 = \ell) = p^2(1 - p)^{\ell-2}$. Comme $N_1 = k$ et $N_2 - N_1 = \ell$ signifie que $N_1 = k$ et $N_2 = k + \ell$.

$$P(N_1 = k, N_2 - N_1 = \ell) = p^2(1 - p)^{k+\ell-2},$$

qui est le produit $\big(p(1 - p)^{k-1}\big)\ \big(p(1 - p)^{\ell-1}\big)$, cqfd.

c) En faisant la somme de $k = 1$ à $k = \ell - 1$ des $P(N_1 = k, N_2 = \ell)$, on obtient que $P(N_2 = \ell) = (\ell - 1)p^2(1-p)^{\ell-2}$, donc

$$P(N_1 = k \,|\, N_2 = \ell) = \frac{p^2(1-p)^{\ell-2}}{(\ell-1)p^2(1-p)^{\ell-2}} = \frac{1}{\ell-1}.$$

La loi de N_1 sachant $N_2 = \ell$ est uniforme sur l'ensemble $\{1, 2, \ldots, \ell - 1\}$.

d) Pour tout $\ell \geqslant k$,

$$P(N_k = \ell) = \frac{(\ell-1)\cdots(\ell-k+1)}{(k-1)!}p^k(1-p)^{\ell-k}.$$

5. La covariance de X et Y vaut

$$E(XY) - E(X)E(Y) = P(X = Y = 1) - P(X = 1)P(Y = 1),$$

donc $P(X = Y = 1) = P(X = 1)P(Y = 1)$. Les autres égalités s'en déduisent.

6. La loi de Z est binomiale $(n + m, p)$. Pour tout i,

$$P(X = i) = \binom{n}{i}p^i(1-p)^{n-i},$$

et

$$P(Y = k - i) = \binom{m}{k-i}p^{k-i}(1-p)^{m-k+i}.$$

En faisant la somme sur i, on obtient $P(Z = k)$, qui vaut par ailleurs

$$P(Z = k) = \binom{n+m}{k}p^k(1-p)^{n+m-k}.$$

En factorisant $p^k(1-p)^{n+m-k}$ des deux côtés, on obtient le résultat.

7. a) S'il reste $X_N = k$ allumettes à gauche au moment où Banach se rend compte que sa poche droite est vide, cela signifie qu'il a choisi N fois sa poche droite et $N - k$ fois sa poche gauche pendant ses $2N - k$ premiers choix, et enfin sa poche droite pour son choix $2N - k + 1$. Cela se produit avec probabilité

$$\binom{2N-k}{N}2^{-(2N-k+1)}.$$

Mais il peut aussi rester $X_N = k$ allumettes à droite, donc

$$P(X_N = k) = \binom{2N-k}{N}2^{k-2N}.$$

b) Soit $Y_N := 2N - X_N$, donc $N \leqslant Y_N \leqslant 2N$ et, pour tout $N \leqslant k \leqslant 2N$,

$$P(Y_N = k) = \binom{k}{N}2^{-k}.$$

Comme $(k+1)\binom{k}{N} = (N+1)\binom{k+1}{N+1}$,

$$E(Y_N+1) = 2(N+1)\sum_{i=N+1}^{2N+1}\binom{i}{N+1}2^{-i}.$$

Si la dernière somme allait jusqu'à $i = 2N+2$, elle correspondrait à la loi de Y_{N+1}, donc sa valeur serait 1. Il vient

$$E(Y_N+1) = 2(N+1)\left(1-\binom{2N+2}{N+1}2^{-(N+1)}\right).$$

En réordonnant et en utilisant l'égalité $(N+1)\binom{2N+2}{N+1} = 2(2N+1)\binom{2N}{N}$, on obtient le résultat. Par Stirling, $u_N \sim 1/\sqrt{\pi N}$, cqfd.

8. a)b) $X \leqslant k$ signifie que l'on a tiré une des k boules de plus petit numéro parmi les N boules possibles, et ceci n fois, donc $P(X \leqslant k) = (k/N)^n$ et, pour tout $1 \leqslant k \leqslant N$, $P(X=k) = (k^n-(k-1)^n)/N^n$.

c) D'après l'exercice 1,

$$E(X) = N - \sum_{k=1}^{N-1} P(X \leqslant k) = N - \sum_{k=1}^{N-1}(k/N)^n.$$

La dernière somme vaut N fois la somme de Riemann de la fonction $x \mapsto x^n$ sur l'intervalle $[0,1]$ donc est équivalente à $N/(n+1)$ et $E(X) \sim Nn/(n+1)$.

9. a) Premier succès dans une suite i.i.d. de piles ou face avec probabilité $1/n$ de succès, donc la loi de N est géométrique de paramètre $a := 1/n$, donc $E(N) = 1/a = n$ et $\mathrm{var}(X) = (1-a)/a^2 = n(n-1)$.

b) Continuons à essayer les clés en les tirant sans remise, même après avoir trouvé la bonne. La position de la bonne clé est aléatoire entre 1 et n et par symétrie, sa loi est uniforme sur $\{1,\dots,n\}$. Donc $E(N) = (n+1)/2$ et $\mathrm{var}(N) = (n^2-1)/12$

10. a) Poisson $a+b$.

b) Binomiale (n,p) avec $p := a/(a+b)$.

c) Multinomiale $(n,(p_k)_k)$ avec $p_k := a_k/a$ et $a := a_1+\cdots+a_{r+1}$.

11. a) X est le nombre de piles.

b) La fonction génératrice de X_i est $(2i+s)/(2i+1)$ donc

$$\varphi(s) = \prod_{i=1}^{n}\frac{2i+s}{2i+1}.$$

c) $\varphi(1) = 1$ et $\varphi(-1) = E((-1)^X) = P(X\,\text{pair}) - P(X\,\text{impair})$ donc $\varphi(-1)$ vaut $2P(X\,\text{pair})-1$, cqfd. Comme $\varphi(-1) = 1/(2n+1)$, $P(X\,\text{pair}) = n/(2n+1)$.

12. a) $P(X \geqslant n) = a^n$. $Z \geqslant n$ si $X \geqslant n$ et $Y \geqslant n$ donc $P(Z \geqslant n) = (ab)^n$ et la loi de Z est géométrique de paramètre ab : $P(Z = n) = (ab)^n(1 - ab)$.

b) $U = 0$ si $X = Y$ donc $P(U = 0)$ est la somme des $a^n(1 - a) \times b^n(1 - b)$ sur $n \geqslant 0$, soit $P(U = 0) = (1 - a)(1 - b)/(1 - ab)$. Pour $n \geqslant 1$,

$$P(X \geqslant Y + n) = \sum_{k \geqslant 0} P(Y = k)P(X \geqslant k + n) = \sum_{k \geqslant 0} (1 - b)b^k a^{k+n},$$

et cette dernière somme vaut $a^n(1 - b)/(1 - ab)$. Par symétrie $P(Y \geqslant X + n)$ vaut $b^n(1 - a)/(1 - ab)$ donc $P(U \geqslant n) = (a^n(1 - b) + b^n(1 - a))/(1 - ab)$, et finalement, pour tout $n \geqslant 1$,

$$P(U = n) = (a^n + b^n)(1 - a)(1 - b)/(1 - ab).$$

c) $Z = k$ et $U = i$ avec $i \geqslant 1$ signifie que $\{X, Y\} = \{k, k + i\}$, donc

$$P(Z = k, U = i) = (1 - a)(1 - b)(a^k b^{k+i} + a^{k+i} b^k) = P(Z = k)P(U = i).$$

Le cas $i = 0$ se traite de même.

14. Pour $0 \leqslant k < n$, $Y = k$ et $X = n$ signifie qu'on a obtenu k fois 5 et $n - k - 1$ fois ni 5 ni 6 sans savoir dans quel ordre, puis enfin un 6, donc

$$P(X = n, Y = k) = \binom{n-1}{k} 4^{n-k-1} 6^{-n}.$$

On en déduit

$$P(X = n) = 4^{n-1} 6^{-n} \sum_{k=0}^{n-1} \binom{n-1}{k} 4^{-k} = 4^{n-1} 6^{-n} (1 + 1/4)^{n-1} = 5^{n-1}/6^n.$$

Donc la loi de X sachant que $X = n$ est binomiale $(n - 1, 1/5)$.

Fiche 4

1. On intègre d'abord en y, il vient $\frac{1}{2}x$, dont l'intégrale entre 0 et 1 vaut 1.

La densité de la loi marginale de X est $\frac{1}{2}x\mathbf{1}_{0<x\leqslant 1}$ d'après le calcul précédent. En intégrant d'abord en x, on voit que la densité de la loi marginale de Y est $\frac{2}{3}(1 - y^{3/2})y^{-1/2}\mathbf{1}_{0<y\leqslant 1}$. Pas indépendance.

Soit $Z := Y/X$. Alors $0 \leqslant Z \leqslant 1$ et, pour tout $0 < z \leqslant 1$,

$$P(Z \leqslant z) = \int_0^1 \sqrt{x}\,\mathrm{d}x \int_0^{zx} \mathrm{d}y/\sqrt{y} = \int_0^1 \sqrt{x}\, 2\sqrt{zx}\,\mathrm{d}x = \sqrt{z},$$

donc la loi de Z admet la densité $1/(2\sqrt{z})$ sur $[0, 1]$. Le même calcul donne

$$P(Z \leqslant z, X \leqslant x) = \int_0^x \sqrt{t}\,\mathrm{d}t \int_0^{zt} \mathrm{d}y/\sqrt{y} = \int_0^x \sqrt{t}\, 2\sqrt{zt}\,\mathrm{d}t = x^2\sqrt{z},$$

donc X et Z sont indépendantes et conviennent.

On peut aussi utiliser deux variables aléatoires indépendantes de loi uniforme sur $[0,1]$ et poser

$$X := \sqrt{U}, \qquad Y := \sqrt{U}\, V^2.$$

2. Indépendance, $c = 1$, $P(Y \geqslant y) = 1/y$ et $P(X \geqslant x) = \mathrm{e}^{-x}$ donc la probabilité demandée $P(Y > 2, X < 1)$ vaut $\frac{1}{2}(1 - 1/\mathrm{e})$. Lois marginales de densités respectives e^{-x} sur $x \geqslant 0$ et $1/y^2$ sur $y \geqslant 1$. Comme $P(X \geqslant x) = \mathrm{e}^{-x}$,

$$P(XY > 1) = E(\mathrm{e}^{-1/Y}) = \int_1^\infty \mathrm{e}^{-1/y}\, \mathrm{d}y/y^2 = \int_0^1 \mathrm{e}^{-z}\, \mathrm{d}z = 1 - 1/\mathrm{e}.$$

3. Exponentielle a, moyenne et variance $1/a$. Pour tout $v \geqslant u \geqslant 0$,

$$P(u \leqslant U, V \leqslant v) = P(u \leqslant T_k \leqslant v, 1 \leqslant k \leqslant n).$$

En utilisant l'indépendance des T_k et $P(u \leqslant T_k \leqslant v) = \mathrm{e}^{-au} - \mathrm{e}^{-av}$, on obtient

$$P(u \leqslant U, V \leqslant v) = (\mathrm{e}^{-au} - \mathrm{e}^{-av})^n.$$

En dérivant par rapport à u puis à v, on obtient la densité du couple (U, V) sur $0 \leqslant u \leqslant v$ comme

$$n(n-1)a^2 \mathrm{e}^{-a(u+v)}(\mathrm{e}^{-au} - \mathrm{e}^{-av})^{n-2}.$$

Pas indépendance. En procédant de même (ou en faisant $v \to \infty$ ci-dessus), on voit que U est exponentielle na, et que

$$P(V \leqslant v) = (1 - \mathrm{e}^{-av})^n,$$

donc la densité de la loi de V est $na\mathrm{e}^{-av}(1 - \mathrm{e}^{-av})^{n-1}$ sur $v \geqslant 0$.

4. Jacobien $\mathrm{d}x\mathrm{d}y = \mathrm{d}u\mathrm{d}v/(2v)$ sur $uv \geqslant 1$ et $u/v \geqslant 1$ donc la densité de la loi de (U, V) est $1/(2vu^2)$ sur $1/u \leqslant v \leqslant u$. Pas d'indépendance. Loi de U de densité $(\log u)/u^2$ sur $u \geqslant 1$. Loi de V de densité $1/2$ sur $0 < v \leqslant 1$ et $1/(2v^2)$ sur $v \geqslant 1$.

5. $U = X$ signifie que $X \leqslant Y$. Comme $P(Y \geqslant y) = \mathrm{e}^{-by}$,

$$P(U = X) = E(\mathrm{e}^{-bX}) = \int_0^{+\infty} \mathrm{e}^{-bx} a \mathrm{e}^{-ax}\, \mathrm{d}x = a/(a+b).$$

Pour $u \leqslant v$, $U \leqslant u \leqslant v \leqslant V$ signifie que $X \leqslant u$ et $Y \geqslant v$ ou que $X \geqslant v$ et $Y \leqslant u$, donc

$$P(U \leqslant u, V \geqslant v) = (1 - \mathrm{e}^{-au})\mathrm{e}^{-bv} + \mathrm{e}^{-av}(1 - \mathrm{e}^{-bu}).$$

En dérivant par rapport à u puis à v, on obtient que la densité de (U, V) sur $0 \leqslant u \leqslant v$ vaut $f(u, v) := ab(\mathrm{e}^{-(au+bv)} + \mathrm{e}^{-(bu+av)})$. En posant $v = w + u$ avec $w \geqslant 0$, on obtient que la densité de (U, W) sur $u \geqslant 0$ et $w \geqslant 0$ vaut

$$ab\mathrm{e}^{-(a+b)u}(\mathrm{e}^{-bw} + \mathrm{e}^{-aw}).$$

Forme produit donc U et W sont indépendantes de lois respectives $(a+b)\mathrm{e}^{-(a+b)u}$ et $(ab/(a+b))(\mathrm{e}^{-bw}+\mathrm{e}^{-aw})$.

6. $c=a^2$. Comme $y\mathrm{d}z=\mathrm{d}x\mathrm{d}y$, la densité de (Z,Y) en tout point (z,y) tel que $0<z<1$ et $y>0$ vaut $a^2y\,\mathrm{e}^{-ay}$. Indépendance, Y de loi gamma $(2,a)$, c'est-à-dire de densité $a^2y\,\mathrm{e}^{-ay}$, et Z de loi uniforme sur $[0,1]$.

Fiche 5

1. La somme de $n=10^6$ erreurs centrées et chacune d'écart-type $t:=10^{-9}/(2\sqrt{3})$ vaut à peu près $t\sqrt{n}$ fois une gaussienne centrée réduite donc la probabilité cherchée d'une erreur totale entre $-a$ et $+a$ avec $a:=\frac{1}{2}10^{-6}$ vaut $P(|t\sqrt{n}Z|\leqslant a)$, soit $2P(0\leqslant Z\leqslant b)$ avec $b:=a/(t\sqrt{n})=\sqrt{3}$ donc la probabilité cherchée vaut à peu près 92%.

2. a) S_n est Poisson n, d'espérance et de variance n.

b) Théorème central limite pour la suite $(X_k)_k$ d'espérance et variance 1.

c) Le membre de gauche est $P(S_n\leqslant n)=P(T_n\leqslant 0)$, le membre de droite est $P(Z\leqslant 0)$, cqfd.

d) Par Bienaymé-Chebychev, $P(T_n\geqslant x)\leqslant 1/x^2$ et $P(Z\geqslant x)\leqslant 1/x^2$ pour tout $x>0$ et tout $n\geqslant 1$, donc, pour $u>0$ fixé,

$$\int_{1/u}^{+\infty}P(T_n\geqslant x)\,\mathrm{d}x\leqslant u,\quad \int_{1/u}^{+\infty}P(Z\geqslant x)\,\mathrm{d}x\leqslant u.$$

Comme intégrale d'une suite de fonctions uniformément bornées sur un intervalle borné,

$$\int_0^{1/u}P(T_n\geqslant x)\,\mathrm{d}x\to\int_0^{1/u}P(Z\geqslant x)\,\mathrm{d}x.$$

quand $n\to\infty$. Par conséquent, la limsup de la différence entre les intégrales complètes vaut au plus $2u$. Comme u est aussi petit que l'on veut, cqfd.

En échangeant l'ordre des intégrales,

$$\int_0^{+\infty}P(Z\geqslant x)\,\mathrm{d}x=\int_0^{+\infty}\frac{\mathrm{e}^{-y^2/2}}{\sqrt{2\pi}}\,\mathrm{d}y\int_0^y\mathrm{d}x=\left[-\frac{\mathrm{e}^{-y^2/2}}{\sqrt{2\pi}}\right]_0^{+\infty}=\frac{1}{\sqrt{2\pi}}.$$

e) Sur l'intervalle $k\sqrt{n}<x\leqslant(k+1)/\sqrt{n}$, $P(T_n\geqslant x)=P(S_n\geqslant n+k+1)$, donc l'intégrale recherchée vaut

$$\sum_{k\geqslant 0}P(S_n\geqslant n+k+1)/\sqrt{n}.$$

On décompose chaque $\{S_n\geqslant n+k+1\}$ en la réunion des $\{S_n=n+i\}$ pour $i\geqslant k+1$, puis on intervertit les sommes. Il vient

$$\sum_{k\geqslant 0}P(S_n\geqslant n+k+1)=\sum_{i\geqslant 1}i\,P(S_n=n+i).$$

Ensuite,

$$i\,P(S_n = n+i) = n\,(P(S_n = n+i-1) - P(S_n = n+i)),$$

donc les termes de la somme s'annulent tous sauf le terme $P(S_n = n)$ provenant de $i = 1$. Finalement,

$$\int_0^{+\infty} P(T_n \geqslant x)\,\mathrm{d}x = n\,P(S_n = n)/\sqrt{n},$$

ce qui démontre la formule. Il reste à égaler les deux limites pour obtenir l'équivalent de f).

3. a) Nombre de clients arrivant pendant l'intervalle de temps $]t, t+h]$.

b) La somme des q_n sur $n \geqslant 0$ vaut 1, d'où l'expression des $q_n(t,h)$. $N_{t+h} = 0$ signifie que $N_t = 0$ et que $N_{t+h} - N_t = 0$, deux évènements qui sont indépendants et de probabilités respectives $p_0(t)$ et $q_0(t,h)$. Quand $h \to 0$, $p_0(t+h) = p_0(t) - ahp_0(t) + o(h)$, ce qui signifie que $p_0'(t) = -ap_0(t)$ (en fait, il s'agit de la dérivée à droite). Comme $p_0(0) = 1$, il vient $p_0(t) = \mathrm{e}^{-at}$.

c) Même raisonnement : pour avoir n personnes au temps $t+h$, il faut en avoir $k \leqslant n$ au temps t et que $n-k$ arrivent pendant $]t, t+h]$. Quand $h \to 0$, ne restent que q_0 et q_1, puisque

$$p_n(t+h) = p_n(t)(1-ah) + p_{n-1}(t)ah + o(h),$$

donc $p_n'(t) = a(p_{n-1}(t) - p_n(t))$ (il y a une faute dans l'énoncé).

d) Soit $r_n(t) := p_n(t)\mathrm{e}^{at}$. Alors $r_0(t) = 1$, $r_n(0) = 0$ et $r_n'(t) = ar_{n-1}(t)$ pour tout $n \geqslant 1$. On voit que r_n est la $n$ème primitive itérée de la fonction 1, cqfd.

e) $T_1 \geqslant t$ signifie que $N_t = 0$ donc sa probabilité vaut $p_0(t) = \mathrm{e}^{-at}$, cqfd.

4. Premier et deuxième chiffres après la virgule dans l'écriture décimale de X. Indépendance, lois uniformes sur $\{0, 1, \ldots, 9\}$. Idem pour les chiffres suivants si on note $Y_1 = Y$, $Y_2 = Z$ et pour tout $n \geqslant 1$, Y_n la partie entière de

$$10^n X - 10^{n-1}Y_1 - 10^{n-2}Y_2 - \cdots - 10Y_{n-1}.$$

Alors, $X = \displaystyle\sum_{n \geqslant 1} \frac{Y_n}{10^n}$.

Fiche 6

1. On approche une variable aléatoire de loi binomiale $B(n, 1/2)$ par la variable aléatoire gaussienne $n/2 + Z\sqrt{n}/2$ donc on veut que $P(|Z/(2\sqrt{n})| \leqslant a) \geqslant 1-p$ avec $a = 1\%$ et $p = 4\%$. Si $P(Z \geqslant z) = p/2$, tout entier $n \geqslant (z/2a)^2$ convient asymptotiquement. Si $P(Z \geqslant z) = 2\%$, il faut donc de l'ordre de $2500z^2$ jets.

2. Si $n = 75$, on peut accepter de l'ordre de $(5n/4)(1 - z/\sqrt{5n})$ réservations, avec $P(Z \geqslant z) = 10\%$.

2.1. On accepte H_0 avec probabilité $P(B(200, 5\%) \geqslant 15)$, que l'on peut approcher par $P(\mathcal{N}(10,10) \geqslant 15)$, soit $P(Z \geqslant 5/\sqrt{10}) \approx P(Z \geqslant 1,6)$.

Fiche 7 : Algèbre (1) Groupes

« L'être humain n'est pas un tueur. Le groupe, si. » Konrad Lorenz

1 Rappels de cours

Une loi de composition (interne) sur un ensemble E est une application $* : E \times E \to E$. On note $x * y$ l'image de (x, y) par $*$. La loi $*$ est associative si et seulement si, pour tous x, y et z éléments de E, $(x * y) * z = x * (y * z)$. La loi $*$ est commutative si et seulement si, pour tous x et y éléments de E, $x * y = y * x$. Un élément e de E est un neutre pour la loi $*$ si et seulement si, pour tout x élément de E, $x * e = e * x = x$. Quand la loi possède un élément neutre e, un inverse d'un élément x de E est un élément y de E tel que $x * y = y * x = e$. Dans ce cas, on note souvent $y = x^{-1}$.

Un groupe $(G, *)$ est un couple formé d'un ensemble G et d'une loi de composition $*$ sur G, associative, admettant un élément neutre, et telle que tout élément possède un inverse. Si de plus, la loi $*$ est commutative, on dit que $(G, *)$ est abélien ou commutatif.

Exercice 1.1. 1) Montrer que le neutre est unique, au sens où, si e et e' sont deux éléments neutres d'un groupe $(G, *)$, alors $e = e'$.

2) Montrer que l'inverse est unique, au sens où, si x est un élément d'un groupe $(G, *)$ et si y et z sont des inverses de x, alors $y = z$.

Exercice 1.2. 1) Les ensembles $\mathbb{Z}$, $\mathbb{Z}/n\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont des groupes pour l'addition.

2) Pour tout ensemble E non vide, $\mathfrak{S}(E)$ est un groupe pour la composition.

3) Pour tout ensemble E, $\mathcal{P}(E)$ est un groupe pour la différence symétrique.

4) Pour tout ensemble E non vide et tout groupe $(G, *)$, l'ensemble G^E est un groupe pour la multiplication terme à terme définie comme suit : si φ et ψ sont des éléments de G^E, on pose, pour tout élément x de E, $(\varphi \cdot \psi)(x) = \varphi(x) * \psi(x)$.

Exercice 1.3. Montrer que, quand $(G, *) = (\mathbb{Z}/2\mathbb{Z}, +)$, l'exemple 4 de l'exercice 1.2 correspond à l'exemple 3.

Une partie H de G est stable par $*$ si, pour tous x et y éléments de H, $x * y$ appartient à H. On note alors de nouveau $*$ la restriction de $*$ à $H \times H$.

Une partie H de $(G, *)$ est un sous-groupe de $(G, *)$ si et seulement si H est stable par $*$ et si $(H, *)$ est un groupe.

Soit $A \subset G$. On note $\langle A \rangle$ l'intersection de tous les sous-groupes de $(G, *)$ contenant A. Alors la partie $\langle A \rangle$ est elle-même un sous-groupe de $(G, *)$, appelé le sous-groupe engendré par A. Comme toute intersection de sous-groupes d'un groupe est elle-même un sous-groupe, $\langle A \rangle$ est aussi le plus petit sous-groupe de $(G, *)$ contenant A.

Si $A = \{x_1, \ldots, x_n\}$ est fini, on note $\langle A \rangle = \langle x_1, \ldots, x_n \rangle$.

Soit x un élément de G et e l'élément neutre. On définit par récurrence x^n pour tout entier relatif n en posant $x^0 = e$ puis, pour tout $n \geqslant 0$,

$$x^{n+1} = x^n * x, \qquad x^{-(n+1)} = x^{-n} * x^{-1}.$$

Exercice 1.4. Montrer que $x^n * x^p = x^{n+p}$ pour tous n et p entiers relatifs.

Un groupe $(G, *)$ est cyclique s'il est engendré par un seul élément, donc s'il existe x élément de G tel que $G = \langle x \rangle$, c'est-à-dire $G = \{x^n \,;\, n \in \mathbb{Z}\}$.

Attention : cette écriture ne signifie pas que G est en bijection avec $\mathbb{Z}$.

L'ordre d'un élément x d'un groupe $(G, *)$ d'élément neutre e est

$$\inf\{n \geqslant 1 \,|\, x^n = e\}.$$

Exercice 1.5. L'ordre de x est le cardinal de $\langle x \rangle$ si $\langle x \rangle$ est fini, et $+\infty$ sinon.

Un morphisme (de groupes) d'un groupe $(G, *)$ dans un groupe $(H, \circ)$ est une application $\varphi : G \to H$ compatible avec les lois $*$ et $\circ$, c'est-à-dire telle que, pour tous x et y éléments de G, $\varphi(x * y) = \varphi(x) \circ \varphi(y)$. Si $G = H$, on dit que φ est un endomorphisme (de groupe) du groupe G.

L'image du morphisme $\varphi : G \to H$ est $\varphi(G) = \{z \in H \,;\, \exists x \in G,\, \varphi(x) = z\}$.

Le noyau du morphisme φ est $\ker(G) = \{x \in G \,;\, \varphi(x) = e_H\}$.

Exercice 1.6. L'image d'un morphisme de groupes $\varphi : (G, *) \to (H, \circ)$ est un sous-groupe de H et son noyau est un sous-groupe de G.

Un morphisme de groupes de G vers H est injectif si et seulement si son noyau est le sous-groupe trivial, donc réduit à $\{e_G\}$.

Enfin, si un morphisme de groupes φ de G vers H est bijectif, l'application inverse $\varphi^{-1} : H \to G$ est aussi un morphisme. Dans ce cas, on dit que φ est un isomorphisme, que φ^{-1} est l'isomorphisme inverse de φ et que les groupes G et H sont isomorphes.

2 Vrai ou faux

Prouver chacune des assertions suivantes ou en donner un contre-exemple.

1. $(\mathbb{N}, +)$ est un groupe abélien.
2. $(\mathbb{Q}^*, +)$ est un groupe.
3. $(\mathbb{Q}^*, \times)$ est un groupe.
4. $(\mathbb{Z}^*, +)$ est un groupe.
5. $(\mathbb{Z}^*, \times)$ est un groupe.
6. Soit $(G, \cdot)$ un groupe. Pour tout entier $n \geqslant 1$ et tous x et y éléments de G, $(x \cdot y)^n = x^n \cdot y^n$.
7. Soit $(G, \cdot)$ un groupe et x et y deux éléments de G d'ordres finis, notés respectivement $o(x)$ et $o(y)$. Alors $o(x \cdot y)$ est fini et divise le produit $o(x)o(y)$.
8. Même affirmation avec « groupe abélien » au lieu de « groupe ».
9. Soit $H \subset G$ une partie non vide d'un groupe $(G, \cdot)$. La condition suivante entraîne que H est un sous-groupe de G : pour tout h élément de H, h^{-1} appartient à H.
10. Même affirmation avec la condition : pour tous h et h' éléments de H, $h \cdot h'$ appartient à H.
11. Même affirmation avec la condition : pour tous h et h' éléments de H, $h^{-1} \cdot h'$ appartient à H.
12. Même affirmation avec la condition : pour tous h et h' éléments de H, $h \cdot h'$ et h^{-1} appartiennent à H.
13. L'ensemble $\mathbb{U} = \{z \in \mathbb{C}\,;\, |z| = 1\}$ des nombres complexes de module 1, muni du produit des nombres complexes, est un sous-groupe du groupe $\mathbb{C}^* = (\mathbb{C} \setminus \{0\}, \cdot)$.
14. La fonction exponentielle est un morphisme de groupes de $(\mathbb{R}, +)$ dans $\mathbb{R}^* = (\mathbb{R} \setminus \{0\}, \cdot)$.
15. La fonction exponentielle est un morphisme de groupes de $(\mathbb{R}, +)$ dans $\mathbb{R}_+^* = (\mathbb{R}_+ \setminus \{0\}, \cdot)$.
16. Les groupes $(\mathbb{R}, +)$ et $\mathbb{R}_+^*$ sont isomorphes.
17. Les groupes $(\mathbb{R}, +)$ et $\mathbb{R}^*$ sont isomorphes.

3 Exercices

Exercice 3.1. Soit $*$ une loi sur un ensemble E, e_g un neutre à gauche et e_d un neutre à droite. On suppose donc que, pour tout x élément de E, $e_g * x = x = x * e_d$. Prouver que $e_g = e_d$.

Exercice 3.2. Soit $E =]-\frac{\pi}{2}, \frac{\pi}{2}[$ et $* : E \times E \to E$ définie par

$$x * y = \arctan(\tan x + \tan y).$$

a) Prouver que $(E, *)$ est un groupe abélien.

b) Préciser si le groupe $(E, *)$ est isomorphe à $(\mathbb{R}, +)$.

Exercice 3.3. Soit E un ensemble, $(G, \cdot)$ un groupe, $\varphi : E \to G$ une bijection et

$$* : E \times E \to E, \qquad x * y = \varphi^{-1}(\varphi(x) \cdot \varphi(y)).$$

Prouver que $(E, *)$ est un groupe et que ce groupe est isomorphe à $(G, \cdot)$.

Exercice 3.4. Soit $G = \mathbb{R}^* \times \mathbb{R}$ et $*$ la loi sur G définie par

$$(x, y) * (u, v) = (xu, xv + y).$$

a) Prouver que $(G, *)$ est un groupe. Préciser si ce groupe est commutatif.

b) Montrer que $\mathbb{R}_+^* \times \mathbb{R}$ est un sous-groupe de G.

Exercice 3.5. a) Soit E un ensemble. Prouver que l'ensemble $\mathfrak{S}(E)$ des bijections de E sur E, muni de la loi $(s, t) \mapsto s \circ t$ où $s \circ t : E \to E$ est définie par $s \circ t(x) = s(t(x))$, est un groupe.

b) Pour tous nombres réels a et b, soit $F_{a,b} : \mathbb{R} \to \mathbb{R}$ telle que $F(t) = at + b$. Soit $\mathbb{B}$ l'ensemble des fonctions $F_{a,b}$ avec a et b réels et a non nul. Soit $\mathbb{A} \subset \mathfrak{S}(\mathbb{R})$ l'ensemble des bijections affines de $\mathbb{R}$ dans $\mathbb{R}$. On rappelle que $F : \mathbb{R} \to \mathbb{R}$ appartient à $\mathbb{A}$ si et seulement si, pour tout réel u dans $[0, 1]$ et tous réels s et t,

$$F(ut + (1 - u)s) = uF(t) + (1 - u)F(s).$$

Montrer que $\mathbb{B} = \mathbb{A}$. On pourra distinguer le cas où t appartient à $[0, 1]$, puis le cas où $t > 1$ en écrivant $1 = (1/t)t + (1 - 1/t)0$, puis le cas où $t < 0$.

En déduire que $\mathbb{A}$ est un sous-groupe de $\mathfrak{S}(\mathbb{R})$.

c) Prouver directement que $\mathbb{A}$ est un sous-groupe de $\mathfrak{S}(\mathbb{R})$.

d) Préciser si le groupe $\mathbb{A}$ est isomorphe au groupe G de l'exercice 4.

Exercice 3.6. Soit $G_1 = \mathbb{R}^* \times \mathbb{R}$ et $*$ la loi sur G_1 définie par

$$(x, y) * (u, v) = (xu, xv + yu^{-1}).$$

a) Prouver que $(G_1, *)$ est un groupe.

b) Parmi les parties suivantes, préciser lesquelles sont des sous-groupes de G_1 :

$$\mathbb{R}^* \times \{0\},\ \{-1\} \times \mathbb{R},\ \{-1, 1\} \times \mathbb{R},\ \{1\} \times \mathbb{R},\ \mathbb{Q}^* \times \mathbb{Q},\ \mathbb{Q}^* \times \mathbb{R},\ \mathbb{Q}^* \times \mathbb{Z},\ \{-1, 1\} \times \mathbb{Z}.$$

c) Pour tout t dans $\mathbb{R}$, soit $H_t = \{(x, t\,(x - x^{-1}))\,;\, x \in \mathbb{R}^*\}$. Montrer que H_t est un sous-groupe commutatif de G_1.

Exercice 3.7. Pour tout nombre complexe a, on définit $r_a : \mathbb{C} \to \mathbb{C}$ par $z \mapsto r_a(z) = az$ et $s_a : \mathbb{C} \to \mathbb{C}$ par $z \mapsto s_a(z) = a\overline{z}$. Soit

$$D = \{r_a, s_a \,|\, a \in \mathbb{C}^*\}.$$

Pour tout entier $n \geqslant 1$, soit

$$D_n = \{r_a, s_a \,|\, a \in \mathbb{C}^*,\, a^n = 1\}.$$

a) Prouver que D est un sous-groupe de $\mathfrak{S}(\mathbb{C})$ (on appelle D le groupe diédral).

b) Prouver que D_n est un sous-groupe à $2n$ éléments de D.

c) Donner les tables de composition de D_2 et D_3.

d) Prouver que D_n est commutatif si et seulement si $n = 1$ ou $n = 2$.

e) Prouver que l'application $u : D \to (\{-1, 1\}, \cdot)$ définie par $u(r_a) = +1$ et $u(s_a) = -1$ est un morphisme de groupes.

Exercice 3.8. Soit $(G, \cdot)$ un groupe. Le centre de G, noté $Z(G)$, est l'ensemble des éléments x de G qui commutent avec G, c'est-à-dire tels que, pour tout élément y de G, $x \cdot y = y \cdot x$.

a) Prouver que $Z(G)$ est un sous-groupe de G et que $Z(G) = G$ si et seulement si le groupe G est abélien.

b) Prouver que $Z(G)$ est un groupe abélien.

c) Déterminer le centre des groupes $\mathcal{A}$ de l'exercice 5, G_1 de l'exercice 6, et D, D_{2n+1} et D_{2n} de l'exercice 7.

d) Préciser si le centre d'un groupe est son plus grand sous-groupe abélien, ou non.

Exercice 3.9. a) Soit $u : \mathbb{C}^* \to \mathbb{U}$ définie par $u(z) = z/|z|$. Montrer que u est un morphisme surjectif du groupe multiplicatif $(\mathbb{C}^*, \cdot)$ dans $(\mathbb{U}, \cdot)$.

b) Construire des isomorphismes du groupe $(\mathbb{C}^*, \cdot)$ sur chacun des groupes produits $(\mathbb{U}, \cdot) \times (\mathbb{R}_+^*, \cdot)$ et $(\mathbb{U}, \cdot) \times (\mathbb{R}, +)$.

Exercice 3.10. a) Soient $(G, \cdot)$ un groupe et φ et ψ des endomorphismes du groupe G qui commutent, c'est-à-dire tels que $\varphi \circ \psi = \psi \circ \varphi$. Montrer que $\varphi(\ker \psi) \subset \ker(\psi)$.

b) Soit $(A, \cdot)$ un groupe abélien et $n \geqslant 0$ un entier. On note p_n l'application de A dans A définie par $p_n(a) = a^n$ pour tout a dans A. Montrer que p_n est un endomorphisme de A. Montrer que si ψ est un endomorphisme de A, alors $\psi \circ p_n = p_n \circ \psi$.

c) Déduire de a) et b) que, si p_n n'est pas injectif, alors il n'existe pas d'endomorphisme s de A tel que $p_n \circ s = \mathrm{Id}_A$, puis que p_n est un isomorphisme si et seulement s'il existe une application s de A dans A telle que $p_n \circ s = \mathrm{Id}_A$.

d) Pour $A = \mathbb{C}^*$, rouver que $p_{\mathbb{C}^*}^2$ est surjectif, mais qu'il n'existe pas d'endomorphisme r de $\mathbb{C}^*$ tel que $p_{\mathbb{C}^*}^2 \circ r = \mathrm{Id}_{\mathbb{C}^*}$. En d'autres termes, il n'existe pas de morphisme « racine carrée ».

Exercice 3.11. Soit $(G, \cdot)$ un groupe, e son élément neutre et A l'ensemble des éléments x de G tels que $x^2 = e$. Prouver que A est aussi l'ensemble des éléments x de G tels que $x = x^{-1}$. En déduire que l'ensemble $B = G \setminus A$ est stable par l'application $x \mapsto x^{-1}$. Montrer les assertions suivantes.

a) Si $A = G$, alors $(G, \cdot)$ est abélien.

b) Si G est fini, alors le cardinal de B est pair.

c) Si le cardinal de G est pair, alors il existe un élément x de G différent de e tel que $x^2 = e$.

Exercice 3.12. Soit $(G, \cdot)$ un groupe cyclique de générateur a et d'élément neutre e.

a) Soit $\pi : \mathbb{Z} \to G$ définie par $\pi(n) = a^n$. Montrer que π est un morphisme surjectif de groupes.

b) Prouver que π est un isomorphisme si et seulement si G est infini.

c) Soit H un sous-groupe de $(G, \cdot)$. Prouver que si G est fini ou si $H \neq \{e\}$, alors $A = \pi^{-1}(H)$ contient un élément strictement positif, puis que A est l'ensemble des multiples entiers de son plus petit élément strictement positif. En déduire les assertions suivantes.

c1) Tout sous-groupe de $(G, \cdot)$ est cyclique.

c2) Tout sous-groupe de $(G, \cdot)$ est stable par les endomorphismes de $(G, \cdot)$.

d) Donner des contre-exemples de c1) et c2) quand $(G, \cdot)$ n'est pas cyclique.

Exercice 3.13. a) Soit A une partie finie de $\mathbb{Q}$. En considérant un dénominateur commun de tous les éléments de A, prouver que le sous-groupe $\langle A \rangle$ de $(\mathbb{Q}, +)$ est cyclique.

b) Préciser si tout sous-groupe de $(\mathbb{Q}^*, \cdot)$ engendré par une partie finie est cyclique.

c) Préciser si les groupes $(\mathbb{Q}, +)$ et $(\mathbb{Q}_+ \setminus \{0\}, \times)$ sont isomorphes.

Exercice 3.14. Soit $\varphi : (G, \cdot) \to (K, \cdot)$ un morphisme de groupes et x un élément de G d'ordre fini. Prouver que l'ordre de $\varphi(x)$ dans K divise l'ordre de x dans G.

Exercice 3.15 (Supplément : exercice corrigé). Soit R un anneau et $GL_k(R)$ le groupe des matrices carrées de taille $k \geqslant 1$ à coefficients dans R et inversibles. L'application det envoie les éléments de $GL_k(R)$ dans le groupe R^* des éléments inversibles de R et le groupe $SL_k(R)$ est son noyau, donc

$$SL_k(R) := \{M \in \mathcal{M}_{k\times k}(R)\,;\, \det M = 1\}.$$

Le but de l'exercice est d'étudier certains sous-groupes $SL_k(\mathbb{Z}/n\mathbb{Z})$.

1) Déterminer $SL_2(\mathbb{Z}/2\mathbb{Z})$ et montrer que ce groupe est isomorphe à $\mathfrak{S}_3$.

2) Calculer l'ordre de $SL_2(\mathbb{Z}/n\mathbb{Z})$ quand n est un entier premier.

3) Trouver des sous-groupes de $SL_2(\mathbb{Z}/n\mathbb{Z})$ isomorphes à $\mathbb{Z}/n\mathbb{Z}$ et $(\mathbb{Z}/n\mathbb{Z})^*$.

4) Calculer l'ordre de $A := \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$ dans $G := SL_2(\mathbb{Z}/3\mathbb{Z})$.

5) D'après la question précédente, l'ordre de G est 24. Montrer que pourtant, G n'est pas isomorphe à $\mathfrak{S}_4$.

Indications de solution

1) On énumère d'abord les 6 éléments de l'ensemble $SL_2(\mathbb{Z}/2\mathbb{Z})$. Ensuite, chacun de ces éléments représente une application linéaire sur $(\mathbb{Z}/2\mathbb{Z})^2$ qui préserve $(0,0)$ et correspond à une bijection de $(\mathbb{Z}/2\mathbb{Z})^2 \setminus \{(0,0)\}$ dans lui-même.

2) On choisit d'abord la première ligne de M, qui est un vecteur non nul, disons (a,b). Soit (c,d) la deuxième ligne. Si $a \neq 0$, il reste à choisir c puis $d = (1+bc)a^{-1}$ est imposé, donc n possibilités pour b puis n possibilités pour (c,d), pour chaque a non nul. Si $a = 0$, d est libre donc n possibilités, et b doit être non nul et $c = b^{-1}$, donc $n-1$ possibilités. En tout, on obtient $n^2(n-1) + n(n-1) = n(n^2-1)$ matrices.

3) On peut associer à tout élément a de $\mathbb{Z}/n\mathbb{Z}$ la matrice $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ et à tout élément a de $(\mathbb{Z}/n\mathbb{Z})^*$ la matrice $\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$.

4) L'ordre de A est 6 car $A^2 = \begin{pmatrix} 0 & 2 \\ 1 & 2 \end{pmatrix}$, $A^3 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ et $A^6 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

5) Le groupe G possède un élément d'ordre 6. Soit s un élément de $\mathfrak{S}_4$. Ou bien s est un cycle de longueur 2 ou s se décompose en un produit de deux cycles de longueur 2 disjoints, alors s est d'ordre 2, ou bien s est un cycle d'ordre 3 ou 4, alors s est d'ordre 3 ou 4. En tout cas, aucun élément de $\mathfrak{S}_4$ n'est d'ordre 6 donc les deux groupes ne sont pas isomorphes.

Fiche 8 : Algèbre (2) Groupe symétrique

« La symétrie, c'est l'ennui. » Victor Hugo, Les Misérables

1 Rappels de cours

Une permutation d'un ensemble E est une bijection de E dans E. Le groupe $\mathfrak{S}(E)$ est l'ensemble des permutations de E muni de la composition des applications (notée $\circ$ ou $\cdot$). Si $n \geqslant 1$ est un entier, on note $\mathbb{N}_n = \{1, 2, \ldots, n\}$ et $\mathfrak{S}_n$ le groupe $\mathfrak{S}(\mathbb{N}_n)$ des permutations de n objets, qu'on appelle le groupe symétrique sur n éléments.

Une transposition est un élément de $\mathfrak{S}_n$ qui laisse fixe tous les points de $\mathbb{N}_n$ sauf exactement deux d'entre eux. Pour toute transposition t, il existe donc deux entiers i et j distincts dans $\mathbb{N}_n$ tels que $t(j) = i$, $t(i) = j$, et, pour tout k dans $\mathbb{N}_n \setminus \{i, j\}$, $t(k) = k$.

Les transpositions engendrent $\mathfrak{S}_n$. Le cardinal de $\mathfrak{S}_n$ vaut $n!$. Il existe un unique morphisme $\varepsilon : \mathfrak{S}_n \to \{-1, 1\}$, qu'on appelle la signature, tel que pour toute transposition t, $\varepsilon(t) = -1$.

Exercice 1.1. Montrer que, pour toute permutation s, $\varepsilon(s) = (-1)^{i(s)}$, où $i(s)$ désigne le nombre d'inversions de s, c'est-à-dire le cardinal de l'ensemble des couples (i, j) d'entiers i et j entre 1 et n tels que $i < j$ et $s(i) > s(j)$.

Le groupe alterné est le noyau du morphisme signature ε.

Pour $\ell \geqslant 2$, un cycle de longueur ℓ de $\mathfrak{S}_n$ est une permutation s dans $\mathfrak{S}_n$ telle qu'il existe une partie $S \subset \mathbb{N}_n$ de cardinal ℓ, appelée le support du cycle s, avec $S = \{i_j \,;\, 1 \leqslant j \leqslant \ell\}$, telle que $s(i_j) = i_{j+1}$ pour tout $1 \leqslant j \leqslant \ell - 1$, $s(i_\ell) = i_1$, et $s(i) = i$ pour tout i dans $\mathbb{N}_n \setminus S$.

Toute permutation s'écrit de manière unique, à l'ordre des facteurs près, comme un produit de cycles de supports disjoints.

Si H est un sous-groupe d'un groupe $(G, \cdot)$, la relation de congruence à droite, définie par

$$\text{« } x \sim y \text{ si et seulement si } x^{-1}y \text{ appartient à } H \text{ »},$$

est une relation d'équivalence sur G. Les classes d'équivalence de cette relation sont les parties $xH = \{xh \,;\, h \in H\}$ pour x dans G. On appelle xH la classe à droite de x modulo H. L'ensemble de ces classes (ou ensemble quotient) est noté G/H.

Théorème 1.2 (Théorème de Lagrange). *Si le groupe G est fini alors son cardinal est le produit du cardinal de H et du cardinal de G/H.*

Corollaire 1.1. *Le cardinal d'un sous-groupe d'un groupe fini divise celui du groupe.*

Définition 1.3. *Un sous-groupe H d'un groupe $(G,\cdot)$ est distingué, ou normal, si pour tout élément h de H et tout élément g de G, le conjugué ghg^{-1} de h par g appartient à H. On note cette propriété $H \triangleleft G$.*

Si $H \triangleleft G$, il existe une unique structure de groupe sur G/H telle que l'application quotient $\varrho : G \to G/H$ définie par $\varrho(x) = xH$ pour tout élément x de G, est un morphisme. On appelle cette structure le groupe quotient. En d'autres termes, dans G/H, on pose $(xH)(yH) = (xy)H$ pour tous éléments x et y de G.

Une action à gauche d'un groupe $(G,\cdot)$ d'élément neutre e sur un ensemble X est une application $G \times X \to X$ notée $(g,x) \mapsto g \cdot x$, telle que pour tous g et h dans G et tout x dans X, $g \cdot (h \cdot x) = (gh) \cdot x$ et $e \cdot x = x$. On dit que G agit à gauche sur X.

Pour une action à gauche $(g,x) \mapsto g \cdot x$ donnée, le stabilisateur d'un élément x de X pour cette action est l'ensemble $G_x = \{g \in G\,;\, g \cdot x = x\}$ et son orbite sous l'effet de cette action est l'ensemble $G \cdot x = \{g \cdot x\,;\, g \in G\}$. Pour toute action de G sur X et pour tout élément x de X, G_x est un sous-groupe de $(G,\cdot)$. Pour toute action de G sur X, les orbites $G \cdot x$ pour x élément de X forment une partition de X.

La conjugaison intérieure $(g,h) \mapsto ghg^{-1}$ est une action de $(G,\cdot)$ sur G.

2 Exercices

Exercice 2.1. a) Soit $c = (i_1, \ldots, i_\ell)$ un cycle de longueur ℓ dans $\mathfrak{S}_n$ et soit s un élément de $\mathfrak{S}_n$. Établir que le conjugué scs^{-1} est le cycle de longueur ℓ qui vaut $(s(i_1), \ldots, s(i_\ell))$.

En déduire que ces deux cycles éléments de $\mathfrak{S}_n$ et de longueur ℓ ont même signature.

Désormais, on fixe $2 \leqslant \ell \leqslant n$.

b) Prouver que tout cycle élément de $\mathfrak{S}_n$ de longueur ℓ est conjugué au cycle $(1, 2 \ldots, \ell)$.

c) Déterminer la permutation $(1,2)(2,3)\cdots(i,i+1)\cdots(\ell-1,\ell)$ élément de $\mathfrak{S}_n$.

d) Prouver que tout cycle de longueur ℓ élément de $\mathfrak{S}_n$ est le produit de $\ell - 1$ transpositions. En déduire la signature d'un tel cycle.

Exercice 2.2. a) Pour tout ℓ, prouver que l'ordre d'un cycle de longueur ℓ vaut ℓ.

b) Déterminer l'ordre de $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$ et l'ordre de $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix}$. Préciser pour chacune de ces permutations s'il s'agit d'un cycle et s'il s'agit d'une puissance d'un cycle.

c) Décomposer en produit de cycles de supports disjoints la permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 6 & 4 & 5 & 8 & 7 & 9 & 11 & 10 & 1 & 12 & 3 & 2 \end{pmatrix}.$$

En déduire sa signature et son ordre.

d) Soit s élément de $\mathfrak{S}_n$ une permutation produit de k cycles de supports disjoints et de longueurs ℓ_i pour $1 \leqslant i \leqslant k$. Vérifier que $\ell_1 + \cdots + \ell_k \leqslant n$ et déterminer la signature et l'ordre de s en fonction des ℓ_i.

Exercice 2.3. Soit $(G, \cdot)$ un groupe fini et g un élément de G. Montrer que l'ordre de g divise le cardinal de G. En déduire qu'un groupe fini dont le nombre d'éléments est premier est cyclique.

Exercice 2.4. Soit p un nombre premier et $n \geqslant p$. Prouver qu'un élément d'ordre p de $\mathfrak{S}_n$ est un cycle.

Exercice 2.5. Soit $H = \langle u, v, w \rangle$ le sous-groupe de $\mathfrak{S}_9$ engendré par les éléments

$$u = (1,4,7)(2,5,8)(3,6,9),\ v = (3,6,9),\ w = (1,2,3)(4,5,6)(7,8,9).$$

a) Calculer les ordres de u, v et w.

b) Calculer wvw^{-1} et $w^{-1}vw$. En déduire la relation $u = (wv)^3$.

c) Déterminer si tout élément non trivial de H est d'ordre 3.

d) Prouver que $K = \langle v, wvw^{-1}, w^{-1}vw \rangle$ est un sous-groupe distingué de H.

e) Déterminer le nombre d'éléments de K puis celui de H.

f) Déterminer la plus grande puissance de 3 qui divise le cardinal de $\mathfrak{S}_9$.

Exercice 2.6. Soit $\mathbb{A}^+$ la partie du groupe $\mathbb{A}$ des bijections affines de $\mathbb{R}$, défini dans l'exercice 3.5 du chapitre 7, formée des bijections croissantes. Prouver que $\mathbb{A}^+$ est un sous-groupe distingué de $\mathbb{A}$. Calculer le cardinal de $\mathbb{A}/\mathbb{A}^+$, identifier le groupe quotient $\mathbb{A}/\mathbb{A}^+$ et le morphisme quotient $\mathbb{A} \to \mathbb{A}/\mathbb{A}^+$.

Exercice 2.7. a) Prouver que le centre $Z(G)$ d'un groupe $(G, \cdot)$ est un sous-groupe distingué de $(G, \cdot)$.

b) Prouver que si le groupe quotient $G/Z(G)$ est cyclique, il est trivial.

c) Déterminer si on peut remplacer « cyclique » par « abélien » dans l'assertion b). On pourra considérer la loi sur $\mathbb{Z}^3$ définie par

$$(u, v, w) * (x, y, z) = (u + x, v + y, w + uy + z).$$

Exercice 2.8. Soit $(G, \cdot)$ un groupe et H un sous-groupe de $(G, \cdot)$. On considère sur G la relation « $x \approx y$ si et seulement si $xy^{-1} \in H$ ».

a) Établir que $\approx$ est une relation d'équivalence et que ses classes d'équivalence sont les ensembles $Hy = \{hy\,;\, h \in H\}$ pour y élément de G.

b) Prouver que H est distingué si et seulement si, pour tout x élément de G, $xH = Hx$.

c) Pour deux parties A et B de G, soit $AB = \{ab\,;\, a \in A, b \in B\}$.

c1) Vérifier que si C est une troisième partie de G alors

$$(AB)C = A(BC).$$

c2) En déduire que pour tous x et y éléments de G, $(xH)(yH) = (xy)H$ si et seulement si H est distingué.

d) Conclure de c2) qu'il existe une structure de groupe sur G/H telle que l'application quotient $\varrho : G \to G/H$ est un morphisme si et seulement si H est distingué.

Exercice 2.9. On reprend les groupes G et G_1 des exercices 3.4 et 3.6 du chapitre 7.

a) Prouver que $(x, y) \mapsto (x^2, y\,x^{-1})$ est un morphisme surjectif de G_1 sur G.

b) Déduire du fait que $Z(G_1) = \{(-1, 0), (1, 0)\}$ que $G_1/Z(G_1)$ est isomorphe à G.

Exercice 2.10. a) Soit $(A, +)$ un groupe abélien dont on note 0 le neutre. Établir que l'ensemble de ses éléments d'ordre fini est un sous-groupe de $(A, +)$. On le note $T(A)$ et on l'appelle le groupe de torsion de A.

b) Prouver que $\dot{0} = T(A)$ est le seul élément d'ordre fini du groupe quotient $A/T(A)$.

c) Déterminer le sous-groupe de torsion du groupe quotient $(\mathbb{R}/\mathbb{Z}, +)$. Préciser si ce sous-groupe est fini ou non.

d) Déterminer les éléments d'ordre fini du groupe $\mathbb{A}$ des bijections affines de $\mathbb{R}$. Préciser si ces éléments forment un sous-groupe.

e) Prouver que l'application $\varphi : (\mathbb{R}, +) \to (\mathbb{C}^*, \cdot)$ définie par

$$t \mapsto \varphi(t) = \mathrm{e}^{2\mathrm{i}\pi t} = \cos(2\pi t) + \mathrm{i}\sin(2\pi t),$$

est un morphisme de groupes et déterminer son noyau et son image.

f) Montrer que $(\mathbb{R}/\mathbb{Z}, +)$ est isomorphe au groupe $(\mathbb{U}, \times)$ des nombres complexes de module 1.

g) Prouver que toute partie finie $\{\xi_1, \ldots, \xi_n\} \subset T(\mathbb{U})$ du groupe de torsion de $\mathbb{U}$ est incluse dans un sous-groupe cyclique, c'est-à-dire qu'il existe un élément ξ de $\mathbb{U}$ tel que $\{\xi_1, \ldots, \xi_n\} \subset \langle\xi\rangle$. Préciser si un tel élément ξ est nécessairement de torsion ou non.

Exercice 2.11. Soit $(G, \cdot)$ un groupe et X un ensemble. Prouver que l'application de $G \times X$ vers X définie par $(g, x) \mapsto g \cdot x$ est une action de $(G, \cdot)$ sur X si et seulement si, pour tout élément g de G, l'application $s_g : X \to X$ définie par $s_g(x) = g \cdot x$ est bijective, et si $g \mapsto s_g$ est un morphisme de $(G, \cdot)$ dans $(\mathfrak{S}(X), \circ)$.

Exercice 2.12. Soit H un sous-groupe d'un groupe $(G, \cdot)$. Pour tout élément g de G, on note $\lambda(g)$ l'application qui à toute partie K de G associe $\lambda(g)(K) = gK$.

a) Prouver que λ induit un morphisme de $(G, \cdot)$ sur $(\mathfrak{S}(G/H), \circ)$, encore noté λ et défini par $\lambda(g)(g'H) = (gg')H$. On appelle λ le morphisme de translation à gauche sur G/H.

b) Prouver que si le cardinal de G/H vaut 2, alors $\lambda(h)(H) = H$ pour tout élément h de H. En déduire que pour tout élément g de G, $\lambda(h)(gH) = gH$, puis que H est distingué.

Exercice 2.13. Soit $(G, \cdot)$ un groupe fini non réduit à un élément et $p \geqslant 2$ le plus petit nombre premier divisant le cardinal de G. Soit H un sous-groupe de $(G, \cdot)$ d'indice p. On suppose donc que le cardinal de G/H vaut p, et on voit que $\mathfrak{S}(G/H)$ est isomorphe à $\mathfrak{S}_p$.

a) Prouver que l'image $\lambda(G)$ du morphisme λ de translation à gauche sur G/H défini dans l'exercice 2.12 est un sous-groupe d'ordre p de $\mathfrak{S}(G/H)$.

b) En déduire que si g est un élément de $G \setminus \ker(\lambda)$, $\lambda(g)$ est un cycle de longueur p dans $\mathfrak{S}(G/H)$.

On pourra utiliser l'exercice 2.4.

c) En conclure que $H = \ker(\lambda)$, donc que H est distingué dans G.

Indications

Exercice 2.5

Il s'agit d'un 3-Sylow maximal.

Exercice 2.13

a) L'ordre de $\lambda(G)$ divise l'ordre de S_p et l'ordre de G, donc leur pgcd, qui vaut p. Donc l'ordre de $\lambda(G)$ vaut 1 ou p. Si $g \in H$, $\lambda(g) = \mathrm{Id}$. Si $g \in G \setminus H$, $\lambda(g)H = gH \neq H$ donc $\lambda(g) \neq \mathrm{Id}$. Comme $\lambda(G)$ possède au moins 2 éléments, donc $\lambda(G)$ est d'ordre p.

Fiche 9 : Algèbre (3) Anneaux et corps

« Le corps est le tombeau de l'âme. » Platon, Cratyle

1 Rappels de cours

Un anneau $(A, +, \times)$ est un ensemble A muni de deux lois de composition $+$, souvent appelée l'addition de A, et $\times$, souvent appelée la multiplication de A, tel que (1) $(A, +)$ est un groupe commutatif, (2) la multiplication est associative, (3) la multiplication possède un élément neutre, et (4) la multiplication est distributive par rapport à l'addition, c'est-à-dire que, pour tous a, b et c éléments de A,

$$a \times (b + c) = (a \times b) + (a \times c), \qquad (a + b) \times c = (a \times c) + (b \times c).$$

On appelle le neutre de $+$ le zéro de A et on le note souvent 0. On appelle le neutre de $\times$ le un de A et on le note souvent 1. L'anneau A est commutatif si de plus sa multiplication est commutative.

Une partie $B \subset A$ d'un anneau $(A, +, \times)$ est un sous-anneau de A si B est stable par les deux opérations $+$ et $\times$, si le un de A appartient à B et si $(B, +)$ est un sous-groupe de $(A, +)$.

Un corps (commutatif) est un anneau (commutatif) A tel que $1 \neq 0$ et tel que tout élément non nul est inversible pour la multiplication.

Un anneau A est intègre si A est commutatif et si pour tout a et b éléments de A, $a \times b = 0$ implique que $a = 0$ ou $b = 0$.

Un morphisme d'anneaux est une application $\varphi : A \to A'$ entre deux anneaux A et A', qui préserve les lois et l'élément neutre multiplicatif, c'est-à-dire telle que, pour tous a et b éléments de A,

$$\varphi(a + b) = \varphi(a) + \varphi(b), \quad \varphi(a \times b) = \varphi(a) \times \varphi(b), \quad \varphi(1_A) = 1_{A'}.$$

Un idéal (bilatère) I d'un anneau $(A, +, \times)$ est un sous-groupe de A tel que, pour tout x élément de I et tout a élément de A, ax et xa sont des éléments de I.

Le noyau $\ker(\varphi)$ (pour l'addition) d'un morphisme d'anneaux $\varphi : A \to A'$ est un idéal de A.

Si I est un idéal d'un anneau A, le groupe abélien quotient A/I possède une unique structure d'anneau telle que l'application quotient $\pi : A \to A/I$ est un morphisme d'anneaux.

Si A est un anneau, il existe un unique morphisme d'anneaux c_A de $\mathbb{Z}$ dans A. Si c_A est injectif, $\ker(c_A) = \{0\}$ et on pose $n_A = 0$. Sinon, $\ker(c_A)$ est un sous-groupe de $\mathbb{Z}$ non réduit à $\{0\}$ donc il existe un entier $n_A \geqslant 1$ tel que $\ker(c_A) = n_A\mathbb{Z}$ (c'est le plus petit entier strictement positif qui appartient à $\ker(c_A)$). Si $A = K$ est de plus un corps, alors $n_K = 0$ ou n_K est un nombre premier. On appelle n_K la caractéristique du corps K.

Si A est un anneau commutatif et si x est un élément de A, l'ensemble xA des xa avec a élément de A est un idéal de A, qu'on appelle idéal principal engendré par a et que l'on note $I(a)$. Un anneau est principal s'il est commutatif intègre et si tous ses idéaux sont principaux.

Si B est une partie de A, on note $I(B)$ l'idéal engendré par B : c'est le plus petit idéal contenant B, et aussi l'intersection de tous les idéaux contenant B. Si B est une partie finie avec $B = \{x_1, \ldots, x_n\}$, on note $I(B) = I(x_1, \ldots, x_n)$.

2 Vrai ou faux

Démontrer ou donner un contre-exemple des assertions suivantes.

1. Soit A un anneau. L'application nulle $\varphi : A \to A$ telle que $\varphi(a) = 0$ pour tout a dans A est un endomorphisme d'anneau.
2. $3\,\mathbb{Z} = \{3\,n\,;\, n \in \mathbb{Z}\}$ est un sous-anneau de $\mathbb{Z}$.
3. Si $n \geqslant 1$ est un entier, l'anneau quotient $\mathbb{Z}/n\mathbb{Z}$ est un corps de caractéristique n.
4. Soit G un groupe abélien et $A = G^G$ l'ensemble des applications de G dans G, muni des lois $f + g = [x \mapsto f(x) + g(x)]$ et $f \circ g = [x \mapsto f(g(x))]$.

   a) $(A, +, \circ)$ est un anneau.

   b) L'ensemble $B \subset A$ des endomorphismes de groupe de G est stable par $+$ et $\circ$, et $(B, +, \circ)$ est un anneau commutatif.
5. L'anneau $\mathbb{Z}/6\mathbb{Z}$ des entiers modulo 6 est un anneau intègre.

3 Exercices de cours

1. Soit A un anneau. Montrer que, pour tout a dans A, $0 \times a = 0 = a \times 0$ et $-a = (-1) \times a = a \times (-1)$.

2. Soit A un anneau. Montrer que l'ensemble C des éléments c de A tels que, pour tout élément a de A, $a \times c = c \times a$, est un sous-anneau de A.

3. Soit A un annneau commutatif et a un élément de A idempotent, c'est-à-dire tel que $a \times a = a$.

a) Prouver que $+$ et $\times$ induisent sur l'idéal principal $I(a) = aA$ une structure d'anneau.

b) Prouver que aA est un sous-anneau de A si et seulement si $a = 1$.

c) Donner des exemples d'anneaux commutatifs A possédant des éléments idempotents autres que 1 et 0.

4. Montrer que l'anneau $\mathbb{Z}/n\mathbb{Z}$ des entiers modulo un entier naturel $n \geqslant 1$ est un corps si et seulement si n est un nombre premier, c'est-à-dire que $n \geqslant 2$ et qu'il n'existe pas d'entiers p et q différents de 1 et n tels que $n = pq$.

5. Soit I un idéal d'un anneau A et $\pi : A \to A/I$ le morphisme quotient. Montrer que J est un idéal de A/I si et seulement si $\pi^{-1}(J)$ est un idéal de A contenant I. En déduire, dans le cas où A est commutatif, que l'anneau quotient A/I est un corps si et seulement si I est un idéal strict maximal de A, c'est-à-dire que $I \neq A$ et que, pour tout idéal J de A avec $I \subset J$, soit $J = I$ soit $J = A$.

6. Soit A un anneau et A^* l'ensemble des éléments de A inversibles pour la multiplication, donc un élément a de A appartient à A^* si et seulement s'il existe un élément b dans A tel que $a \times b = b \times a = 1$.

a) Prouver que A^* est stable par la multiplication et que $(A^*, \times)$ est un groupe. On appelle $(A^*, \times)$ le groupe des unités de A.

b) Déterminer le groupe des unités des anneaux $\mathbb{Z}$ et $\mathbb{Q}$. Si A est un anneau, soit $M_n(A)$ l'anneau des matrices $n \times n$ à coefficients dans A. Déterminer le groupe des unités des anneaux $M_2(\mathbb{R})$ et $M_2(\mathbb{Z})$.

7. Soit A un anneau commutatif, a et b des éléments de A, et $n \geqslant 0$ un entier. Montrer la formule du binôme

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k \cdot b^{n-k}.$$

Rappeler pourquoi cette formule n'est plus vraie dans un anneau non commutatif.

4 Exercices

1. Soit $(A, +, \times)$ un ensemble muni de deux lois $+$ et $\times$. On suppose que la loi $\times$ est distributive par rapport à la loi $+$, que la loi $\times$ possède un élément neutre 1, et que $(A, +)$ est un groupe.

Prouver que + est commutative. Indication : On pourra considérer, pour a et b dans A, l'élément $(a+b)\times(1+1)$.

2. Soit A un anneau tel que tout élément de A non nul possède un inverse à gauche : pour tout élément non nul a de A, il existe un élément b de A tel que $ba=1$. Prouver qu'un inverse à gauche de a est aussi un inverse à droite, c'est-à-dire qu'on a aussi $ab=1$.

3. Soit A un anneau et a, b et c des éléments de A tels que $c(1-ba)=1$. Calculer $(1+acb)(1-ab)$. En déduire que si $1-ba$ admet un inverse à gauche, $1-ab$ aussi.

4. Soit A un anneau idempotent, c'est-à-dire tel que tout élément a de A vérifie $a^2=a$.

a) Montrer que pour tout élément a de A, $a+a=0$. En déduire que A est commutatif.

b) Montrer que pour tout a, b et c éléments de A, $(a+b)c=0$ si et seulement si $a(b+1)c=0$ et $(a+1)bc=0$.

c) Soit X un ensemble et $\mathcal{P}(X)$ l'ensemble des parties de X. Montrer que $(\mathcal{P}(X),\triangle,\cap)$ est un anneau idempotent.

5. Soit p un nombre premier et $\mathbb{Z}_p$ l'ensemble des nombres rationnels que l'on peut écrire comme une fraction dont le dénominateur n'est pas divisible par p.

a) Prouver que $\mathbb{Z}_p$ est un sous-anneau de $\mathbb{Q}$.

b) Prouver que pour tout nombre rationnel x, soit x appartient à $\mathbb{Z}_p$, soit $x\neq 0$ et x^{-1} appartient à $\mathbb{Z}_p$.

c) Soit A un sous-anneau de $\mathbb{Q}$ contenant $\mathbb{Z}_p$. Prouver que les propriétés suivantes sont équivalentes.

1. $A\neq\mathbb{Z}_p$.
2. Il existe un entier $n\geqslant 1$ tel que $1/p^n$ appartient à A.
3. Pour tout entier $n\geqslant 1$, $1/p^n$ appartient à A.
4. $A=\mathbb{Q}$.

d) Etablir les équivalences des trois premières conditions de c) quand, au lieu de supposer que le sous-anneau A de $\mathbb{Q}$ contient $\mathbb{Z}_p$, on suppose que $\mathbb{Z}_p$ ne contient pas A. Donner un exemple avec $A\neq\mathbb{Q}$.

6. Un anneau non principal a) Prouver que l'équation $10\,y^2=x^2$ n'a pas de solution (x,y) dans $\mathbb{Z}^2$ à part $x=y=0$.

b) Déterminer l'ensemble des carrés modulo 10 : un élément y de $\mathbb{Z}/10\mathbb{Z}$ est un carré modulo 10 s'il existe un élément x de $\mathbb{Z}/10\mathbb{Z}$ tel que $y=x^2$.

c) Prouver qu'il n'existe pas de couple (x, y) dans $\mathbb{Z}^2$ tel que $10y^2 = x^2 + 3$ ou $10y^2 = x^2 - 3$.

d) Soit $v = \sqrt{10}$ et A l'ensemble des $x + yv$ pour x et y éléments de $\mathbb{Z}$. Prouver que A est un sous-anneau de $\mathbb{R}$ et que pour tout élément a de A, les entiers x et y tels que $a = x + yv$ sont uniques. On note souvent $A = \mathbb{Z}[\sqrt{10}]$.

e) Soit $c : A \to A$ définie par $c(x + yv) = x - yv$, pour tous x et y entiers. Montrer que c est un endomorphisme d'anneau et que les seuls points fixes de c sont les éléments de $\mathbb{Z}$.

f) Expliciter $ac(a)$ en fonction des « coordonnées » (x, y) de $a = x + yv$. En déduire qu'il n'existe pas d'élément a de A tel que $|ac(a)| = 3$.

g) Soit $n : A \to A$ définie par $n(a) = ac(a)$. Montrer que n est à valeurs dans $\mathbb{Z}$ et vérifie les propriétés suivantes : pour tous a et b éléments de A, $n(ab) = n(a)n(b)$; pour tout a élément de A, $n(a) = 0$ si et seulement si $a = 0$.

h) Soit I l'ensemble des $3a + (2 + v)b$ pour a et b éléments de A. Montrer que I est un idéal de A contenant 3 et $2 + v$ et déduire de ce qui précède que cet idéal n'est pas principal.

7. Soit $\mathbb{Z}[\mathrm{i}]$ l'ensemble des nombres complexes de la forme $a + \mathrm{i}b$ avec a et b éléments de $\mathbb{Z}$.

a) Montrer que $\mathbb{Z}[\mathrm{i}]$ est un sous-anneau de $\mathbb{C}$.

b) Soit z un élément de $\mathbb{Z}[\mathrm{i}]$. Montrer que le conjugué $\overline{z}$ de z appartient à $\mathbb{Z}[i]$ et que $|z|^2$ appartient à $\mathbb{N}$.

c) Soit z un élément de $\mathbb{Z}[\mathrm{i}]$. Montrer que z appartient au groupe $\mathbb{Z}[\mathrm{i}]^*$ des unités de $\mathbb{Z}[\mathrm{i}]$ si et seulement si $|z| = 1$.

d) Déterminer le groupe $\mathbb{Z}[\mathrm{i}]^*$.

e) Montrer que pour tout z élément de $\mathbb{C}$, il existe un élément z_0 de $\mathbb{Z}[\mathrm{i}]$ tel que $|z - z_0|^2 \leqslant \frac{1}{2}$.

f) Prouver que pour tous z_0 et z_1 éléments de $\mathbb{Z}[\mathrm{i}]$ avec $z_1 \neq 0$, il existe des éléments a_0 et a_1 de $\mathbb{Z}[\mathrm{i}]$ tels que $z_0 = a_0 z_1 + a_1$ avec $|a_1| < |z_1|$.

g) Montrer que $\mathbb{Z}[\mathrm{i}]$ est un anneau principal.

8. Soit K un corps de caractéristique non nulle p. Montrer que, pour tous x et y éléments de K et tout entier $n \geqslant 0$,

$$(x + y)^{p^n} = x^{p^n} + y^{p^n}.$$

Fiche 10 : Algèbre (4) Arithmétique, anneau des entiers relatifs

« One Ring to rule them all,
One Ring to find them,
One Ring to bring them all,
And in the darkness bind them. »
JRR Tolkien, The Lord of the Rings

1 Rappels de cours

L'anneau $\mathbb{Z}$ est intègre et tout sous-groupe additif de $\mathbb{Z}$ est un idéal.

Division euclidienne. Soit a et b des entiers relatifs avec $b \neq 0$. Il existe un unique élément (q, r) de $\mathbb{Z} \times \mathbb{N}$ tel que $a = bq + r$ et $0 \leqslant r < |b|$. Les entiers q et r sont respectivement le quotient et le reste de la division euclidienne de l'entier a par l'entier non nul b.

L'anneau $\mathbb{Z}$ est ainsi euclidien donc principal : tout idéal I non nul contient des éléments positifs et, si on note n_I le plus petit élément strictement positif de I, on obtient $I = n_I\mathbb{Z}$. Pour déterminer le générateur n_I de l'idéal $I(A)$ engendré par une partie A de cardinal $n \geqslant 3$, on se ramène au cas $n = 2$ en remarquant que $I(A \cup B) = I(I(A) \cup B)$. Pour cela, on utilise l'algorithme d'Euclide.

Algorithme d'Euclide. Soit a et b des entiers avec $b \neq 0$. Si la division euclidienne de a par b est $a = bq + r$ avec $0 \leqslant r < |b|$, alors les diviseurs communs de a et b sont ceux de b et r, donc $I(a, b) = I(b, r)$: soit $r = 0$ et tout diviseur de b divise a, soit b et r sont deux entiers strictement positifs avec $r < |b|$ et les diviseurs communs de a et b sont ceux de b et r. On construit ainsi une suite $b_0 = a$, $b_1 = |b| > b_2 > \cdots b_n > 0$ telle que, pour tout $2 \leqslant i \leqslant n$, b_i, est le reste de la division de b_{i-2} par b_{i-1}, et telle que b_n divise tous les b_i.

Plus grand commun diviseur. Le pgcd d'une partie A de $\mathbb{Z}$ non vide et non réduite à $\{0\}$ est l'unique entier naturel $d \geqslant 1$ tel que $I(A) = d\mathbb{Z}$.

Un entier d est le pgcd des x_k pour $1 \leqslant k \leqslant n$ si et seulement si d divise chaque x_k et s'exprime par une identité de Bezout, c'est-à-dire qu'il existe des entiers m_k tels que $d = \sum_{k=1}^{n} m_k x_k$.

Deux entiers x et y sont premiers entre eux, ou bien x est premier avec y, si 1 est le pgcd de $\{x, y\}$.

Lemme de Gauss Si d est premier avec x et divise le produit xy, alors d divise y.

Plus petit commun multiple. Le ppcm d'un ensemble fini d'entiers A est l'unique entier naturel $m \geqslant 1$ tel que $\bigcap_{x \in A} I(x) = I(m) = m\mathbb{Z}$.

Pour tous x et y entiers relatifs, $xy = \mathrm{pgcd}(x, y)\mathrm{ppcm}(x, y)$.

Un nombre premier est un entier naturel $p \geqslant 2$ dont l'ensemble des diviseurs positifs est $\{1, p\}$. L'ensemble des nombres premiers est infini.

Théorème fondamental de l'arithmétique Tout entier naturel non nul n s'écrit, de manière unique à l'ordre près, comme un produit de nombres premiers. Donc

$$n = p_1 \cdots p_m = \prod_p p^{v_p(n)},$$

où le produit porte sur les nombres premiers p, chaque $v_p(n)$ appartient à $\mathbb{N}$ et l'ensemble des nombres premiers p tels que $v_p(n) \neq 0$ est fini.

Dans la première factorisation ci-dessus, il y a éventuellement des répétitions $p_i = p_j$ pour $i \neq j$, mais pas dans la seconde. Si l'ensemble des nombres premiers p tels que $v_p(n) \neq 0$ vaut $\{p_1, \ldots, p_k\}$ avec $p_i < p_{i+1}$, la factorisation réduite de l'entier n est $n = p_1^{n_1} \cdots p_k^{n_k}$ et cette décomposition est maintenant unique.

Corollaire Si m et n sont des entiers positifs non nuls,

$$\mathrm{pgcd}(m, n) = \prod_{p \in \mathcal{P}} p^{\min(v_p(m), v_p(n))}, \quad \mathrm{ppcm}(m, n) = \prod_{p \in \mathcal{P}} p^{\max(v_p(m), v_p(n))}.$$

Soit $n \geqslant 1$ un entier naturel. L'anneau $\mathbb{Z}/n\mathbb{Z}$ des entiers modulo n possède n éléments. C'est un corps si et seulement si n est premier. Plus généralement, la classe $\overline{m}$ d'un entier m dans $\mathbb{Z}/n\mathbb{Z}$ est inversible si et seulement si m est premier avec n.

2 Vrai ou faux

1. Soient m et n des entiers relatifs. Les idéaux principaux $I(m)$, $I(n)$ et $I(mn)$ vérifient :

a) $I(m) \cap I(n) \subset I(mn)$; b) $I(mn) \subset I(m) \cap I(n)$; c) $I(mn) = I(m) \cap I(n)$.

2. Le reste de la division de -56 par 12 est -8.

3. L'entier 1457632916 est divisible par 4 mais ne l'est pas par 8.

4. Si n est un entier sans facteur carré, c'est-à-dire si $n = p_1 \cdots p_k$ où les p_i sont des nombres premiers deux à deux distincts, alors $\sqrt{n}$ est irrationnel.

5. Le pgcd de 585 et 286 est 13.

6. Les entiers 728145362718 et 782145326718 sont divisibles par 9.

7. Les entiers 33333333, 14327143271 et 34103025 sont divisibles par 11.

8. Les nombres 101, 103 et 107 sont premiers.

3 Exercices de cours

1. Soit m et n deux nombres premiers entre eux.

a) Pour tout entier $k \geqslant 1$, soit $f_k : \mathbb{Z} \to \mathbb{Z}/k\mathbb{Z}$ défini par $f_k(a) = a$ modulo k. Déterminer le noyau de $f = (f_m, f_n) : \mathbb{Z} \to (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z})$.

b) En déduire que pour tout couple d'entiers (x, y), il existe un entier a tel que $a = x$ modulo m et $a = y$ modulo n (lemme chinois).

2. a) Soit p un nombre premier et $n \in \mathbb{N}$. Calculer le nombre d'éléments inversibles de l'anneau $\mathbb{Z}/p^n\mathbb{Z}$.

b) Soit n un entier positif et $n = p_1^{n_1} \cdots p_k^{n_k}$ sa factorisation réduite. Déduire du a) de l'exercice 1 que $\mathbb{Z}/n\mathbb{Z}$ est isomorphe à l'annneau produit

$$\mathbb{Z}/p_1^{n_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_k^{n_k}\mathbb{Z}.$$

c) Prouver que, avec les notations de b), le nombre $\varphi(n)$ des inversibles de $\mathbb{Z}/n\mathbb{Z}$ vaut (formule d'Euler)

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right),$$

où le produit porte sur les nombres premiers p.

3. a) En utilisant l'algorithme d'Euclide, déterminer le pgcd de 525 et 231.

b) En remontant les calculs, trouver des entiers x et y tels que

$$x525 + y231 = \text{pgcd}(525, 231).$$

4. a) Soit $p_1, \ldots, p_n$ des entiers positifs avec $p_i \leqslant 2^{2^{i-1}}$, et $P = p_1 \cdots p_n + 1$. Prouver que P est premier avec chaque p_i et vérifie $1 \leqslant P \leqslant 2^{2^n}$.

b) En déduire que l'ensemble des nombres premiers est infini, que le $n$ème nombre premier p_n est majoré par $2^{2^{n-1}}$ et que le nombre $\pi(x)$ de nombres premiers $p \leqslant x$ majorés par x vérifie

$$\pi(x) \geqslant \frac{\log(\log x) - \log(\log 2)}{\log 2}.$$

Addendum En fait $\pi(x)$ est équivalent à $x/\log x$ quand $x \to \infty$ et le $n$ème nombre premier p_n est équivalent à $n \log n$. Des formules approchées encore plus précises sont obtenues en remplaçant le logarithme usuel par

$$\mathrm{Li}(x) = \int_1^x \frac{\mathrm{d}x}{\log x}.$$

4 Exercices

1. a) Prouver qu'il existe une infinité de nombres premiers de la forme $4n+3$.

b) Prouver qu'il existe une infinité de nombres premiers de la forme $6n+5$.

2. a) Résoudre les équations $637x + 595y = 91$ et $637x + 595y = 143$ pour x et y entiers relatifs.

b) Préciser si l'équation $442x = 317$ modulo 495 admet une solution x entier relatif.

c) Déterminer les entiers relatifs (u, v) tels que $442u + 495v = 1$.

d) En déduire les solutions de l'équation de b).

3. a) Résoudre l'équation $x^2 + y^2 = 0$ pour x et y éléments de $\mathbb{Z}/3\mathbb{Z}$.

b) En déduire que l'ensemble des entiers relatifs (x, y, z) tels que $x^2 + y^2 - 3z^2 = 0$ vaut $\{(0, 0, 0)\}$.

c) Préciser si le cercle de centre $(0, 0)$ et de rayon $\sqrt{3}$ comprend des points de coordonnées rationnelles.

4. a) Prouver que, pour tout entier n impair, n^2 est congru à 1 modulo 8.

b) En déduire que l'ensemble des entiers relatifs (x, y, z, t) tels que $x^2 + y^2 + z^2 - 7t^2 = 0$ vaut $\{(0, 0, 0, 0)\}$.

5. a) Soit $1 \leqslant a < b$ des entiers. Prouver qu'il existe un entier unique $n_1 \geqslant 2$ tel que $\dfrac{1}{n_1} \leqslant \dfrac{a}{b} < \dfrac{1}{n_1 - 1}$ et que si $a_1 = an_1 - b$ et $b_1 = bn_1$, alors $a_1 < a$ et $0 \leqslant \dfrac{a}{b} - \dfrac{1}{n_1} = \dfrac{a_1}{b_1} < \dfrac{1}{n_1}$.

b) Soit r un nombre rationnel strictement positif. Prouver qu'il existe $k \geqslant 1$ et des entiers $n_i \geqslant 1$ pour $1 \leqslant i \leqslant k$ avec $n_i < n_{i+1}$ et

$$r = \frac{1}{n_1} + \cdots + \frac{1}{n_k}.$$

On pourra d'abord utiliser a) pour traiter le cas où r appartient à $]0, 1[$.

6. Soit p un nombre premier et $n \geqslant 1$ un entier naturel. Soit p^{a_n} la plus grande puissance de p divisant $(p^n)!$, c'est-à-dire $(p^n)! = p^{a_n} q$ où q est premier avec p.

a) Dans le cas $p = 3$, calculer a_2.

b) En général établir que $a_0 = 0$ et que $a_{n+1} = p\, a_n + 1$ pour tout $n \geqslant 0$.

7. Soit $n \geqslant 1$ un entier naturel. Prouver qu'il existe un multiple de n de la forme $10^m - 1$ si et seulement si n est premier avec 10. On pourra considérer l'élément 10 de $\mathbb{Z}/n\mathbb{Z}$.

En déduire que :

a) 2004 possède un multiple dont l'écriture décimale ne comporte que le chiffre 4 ;

b) 2004 ne possède aucun multiple dont l'écriture décimale ne comporte que le chiffre 6.

9. a) Soit p un nombre premier et a un entier naturel non divisible par p. Prouver qu'il existe un entier naturel $k \geqslant 1$ vérifiant la propriété suivante : pour tout entier n, $a^n = 1$ modulo p si et seulement si k divise n. Montrer que k divise $p - 1$.

b) Prouver que 16 est le plus petit entier positif k tel que $5^k = 1$ modulo 17. On pourra d'abord vérifier que $5^8 = -1$ modulo 17.

c) Prouver que pour tout entier b non divisible par 17, il existe un entier positif n tel que $5^n = b$ modulo 17, et trouver tous les entiers naturels n tels que $5^n = 3$ modulo 17.

d) Dans le cas général de a), soit b un entier naturel non divisible par p. Prouver qu'il existe un entier naturel $\ell \geqslant 1$ vérifiant la propriété suivante : pour tout entier n, $a^n = b$ modulo p si et seulement si $n = \ell$ modulo k.

Supplément

Soit $M = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$ une matrice 2×2 à cœfficients entiers. Pour tout entier n, soit $E_1(n) = \begin{pmatrix} 1 & -n \\ 0 & 1 \end{pmatrix}$ et $E_2(n) = \begin{pmatrix} 1 & 0 \\ -n & 1 \end{pmatrix}$. Calculer $E_1(n)\, M$ et $E_2(n) M$.

Soit a et b des entiers non nuls. On définit un entier $K \geqslant 0$ et des suites $(a_k)_k$, $(b_k)_k$ et $(M_k)_k$, à valeurs dans $\mathbb{Z}$, $\mathbb{Z}$ et $M_2(\mathbb{Z})$ et indexées par $0 \leqslant k \leqslant K$, comme suit.

On pose $a_0 = a$, $b_0 = b$ et $M_0 =$ la matrice identité puis, pour tout $k \geqslant 0$, on applique les étapes suivantes.

Si $b_{2k} \neq 0$, on pose $a_{2k+1} = a_{2k} - b_{2k}q_{2k}$ avec $0 \leqslant a_{2k+1} \leqslant |b_{2k}| - 1$. Donc a_{2k+1} est le reste de la division de a_{2k} par b_{2k} et q_{2k} est le quotient. On pose $b_{2k+1} = b_{2k}$ et $M_{2k+1} = E_1(q_{2k})M_{2k}$.

Si $a_{2k+1} \neq 0$, on pose $a_{2k+2} = a_{2k+1}$, $b_{2k+2} = b_{2k+1} - a_{2k+1}q_{2k+1}$ avec $0 \leqslant b_{2k+2} \leqslant a_{2k+1} - 1$. Donc b_{2k+2} est le reste de la division de b_{2k+1} par a_{2k+1} et q_{2k+1} est le quotient. On pose $M_{2k+2} = E_2(q_{2k+1})\, M_{2k+1}$.

Si $b_{2k} = 0$ ou $a_{2k+1} = 0$, on a terminé et on pose $K = 2k$ ou $K = 2k + 1$ selon les cas.

Montrer que le pgcd de a et b vaut

$$\text{pgcd}(a, b) = a_K = x_K a + y_K b \quad \text{si } K \text{ est pair,}$$

et

$$\text{pgcd}(a, b) = b_K = z_K a + t_K b \quad \text{si } K \text{ est impair.}$$

Dans le cas $a = 525$, $b = 231$, calculer K et les suites $(a_k)_k$, $(b_k)_k$ et $(M_k)_k$, en écrivant chaque matrice M_k comme $M_k = \begin{pmatrix} x_k & y_k \\ z_k & t_k \end{pmatrix}$. Calculer $M_K \begin{pmatrix} a \\ b \end{pmatrix}$.

Supplément

a) Soit x et y deux entiers strictement positifs et premiers entre eux. Prouver qu'il existe un entier $k \geqslant 1$ tel que $x^k = 1$ modulo y.

b) Soient a et r deux entiers avec $a > r \geqslant 2$. La suite arithmétique de premier terme a et raison r vaut $(a + nr)_{n\in\mathbb{N}}$. Montrer que cette suite contient une infinité de termes ayant tous les mêmes diviseurs premiers. On pourra traiter d'abord avec a) le cas où a et r sont premiers entre eux.

Note En 2004, Ben Green et Terence Tao ont montré qu'il existe des suites arithmétiques en nombres premiers arbitrairement longues. La plus longue suite connnue est constituée de 23 termes. (Tao a reçu à 31 ans une des médailles Fields décernées au Congrès de Madrid en 2006.)

Fiche 11 : Algèbre (5) Polynômes

« Ma cohabitation passionnée avec les mathématiques m'a laissé un amour fou pour les bonnes définitions, sans lesquelles il n'y a que des à-peu-près. » Gustave Flaubert, Vie de Henry Brulard

1 Rappels de cours

On se donne un sous-corps K du corps $\mathbb{C}$ des nombres complexes et on note $K[X]$ la K-algèbre des polynômes en une indéterminée X sur K.

L'indéterminée X est un élément de $K[X]$, qui engendre $K[X]$ en tant que K-algèbre, ainsi tout élément P de $K[X]$ s'écrit comme une somme finie

$$P = a_0 + a_1 X + \cdots + a_n X^n = \sum_{i=0}^{n} a_i X^i, \qquad a_i \in K.$$

De plus, X engendre $K[X]$ librement, c'est-à-dire que $(X^i)_{i\in\mathbb{N}}$ est une base du K-espace vectoriel $K[X]$.

De façon équivalente, l'écriture de P comme une série $P = \sum_{i\geqslant 0} a_i X^i$ avec $a_i = 0$ pour i suffisamment grand est unique, c'est-à-dire que si $Q = \sum_{i\geqslant 0} b_i X^i$ avec $b_i = 0$ pour i suffisamment grand, alors $P = Q$ si et seulement si $a_i = b_i$ pour tout $i \geqslant 0$. Un morphisme naturel de K dans $K[X]$ est $a \mapsto aX^0$.

Les lois de l'algèbre $K[X]$ sont les suivantes : pour tout élément a de K et tous$P$ et Q éléments de $K[X]$ définis ci-dessus, on pose

$$aP = \sum_{i\geqslant 0} (aa_i) X^i, \quad P + Q = \sum_{i\geqslant 0} (a_i + b_i) X^i,$$

et

$$PQ = \sum_{i\geqslant 0} c_i X^i, \quad \text{avec } c_i = \sum_{k=0}^{i} a_k\, b_{i-k}.$$

L'algèbre $K[X]$ est une K-algèbre libre sur l'élément X : cela signifie que pour toute K-algèbre A et tout élément α de A, il existe un unique morphisme de K-algèbre $E_\alpha : K[X] \to A$, dit morphisme d'évaluation, tel que $E_\alpha(X) = \alpha$. Le morphisme E_α est donné par

$$P = \sum_{i \geqslant 0} a_i X^i \ \mapsto \ E_\alpha(P) = \sum_{i \geqslant 0} a_i \alpha^i.$$

On note aussi $E_\alpha(P) = P(\alpha)$. Dans le cas particulier où $A = K[X]$ et $\alpha = X$, l'unicité de E_α montre que E_X est l'application identité sur $K[X]$. C'est pourquoi on note souvent $P(X)$ le polynôme P.

Si $A = K$, la fonction polynomiale associée à P est l'application $F_P : K \to K$ telle que, pour tout x dans K, $F_P(x) = E_x(P)$. On note $F_P = P$.

Le degré du polynôme P vaut $\deg(P) = -\infty$ si $P = 0$, et, si $P \neq 0$,

$$\deg(P) = \max\{i \in \mathbb{N} \,;\, a_i \neq 0\}.$$

Pour tous P et Q dans $K[X]$,

$$\deg(P+Q) \leqslant \max\{\deg(P), \deg(Q)\}, \qquad \deg(PQ) = \deg(P) + \deg(Q),$$

avec la convention $(-\infty) + d = -\infty$ pour tout d. Le coefficient dominant de P vaut 0 si $P = 0$ et a_d si $\deg(P) = d \geqslant 0$. Le terme dominant de P vaut 0 si $P = 0$ et $a_d x^d$ si $\deg(P) = d \geqslant 0$. Un polynôme est unitaire si son coefficient dominant vaut 1.

L'anneau $K[X]$ est euclidien : cela signifie que pour tous polynômes U et V avec $V \neq 0$, il existe un unique couple de polynômes (Q, R) tel que

$$U = VQ + R, \qquad \deg(R) \leqslant \deg(V) - 1.$$

2 Vrai ou faux

1. L'anneau $K[X]$ des polynômes à coefficients dans un corps K est intègre.

2. Les inversibles de $K[X]$ sont les polynômes de degré 0.

3. Soit $d \geqslant 0$ un entier. L'ensemble des polynômes à coefficients dans K de degré inférieur ou égal à d est :

a) un sous-groupe du groupe additif $K[X]$.

b) un sous-K-espace vectoriel de $K[X]$ de dimension d sur K.

c) un idéal de $K[X]$.

4. Soit K et L deux sous-corps de $\mathbb{C}$ avec $K \subset L$. Alors $K[X] \subset L[X]$ et :

a) un polynôme de degré 1 est irréductible dans $K[X]$.

b) un polynôme irréductible dans $K[X]$ est de degré 1.

c) si un polynôme de $K[X]$ est irréductible dans $L[X]$, il est irréductible dans $K[X]$.

d) si un polynôme de $K[X]$ est irréductible dans $K[X]$, il est irréductible dans $L[X]$.

5. L'ensemble $U \subset K[X]$ des polynômes unitaires à coefficients dans un corps K est : a) stable par multiplication ; b) un idéal de $K[X]$.

6. a) Tout polynôme est multiple d'un unique polynôme unitaire.

b) Tout polynôme irréductible est multiple d'un unique polynôme unitaire.

3 Exercices de cours

1. Soit a un nombre. Effectuer les divisions euclidiennes de $X^4 - 5X^2 + 6$ par $X - a$ et par $(X - a)^2$.

2. Effectuer la division euclidienne de $X^6 - X^4 - X^2 + 1$ par $X^3 - 1$.

3. Déterminer un pgcd des polynômes $P(X) = X^4 + 2X^2 - X + 5$ et $Q(X) = X^4 + 3X^2 + X + 1$.

4. Soit P un polynôme de degré 3 sur un corps K tel que pour tout x dans K, $P(x) \neq 0$. Prouver que P est irréductible dans $K[X]$.

5. Soit $K \subset L \subset \mathbb{C}$ deux sous-corps de $\mathbb{C}$. Soit P et Q des éléments de $K[X]$ et D un élément de $L[X]$ tels que $P \neq 0$ et $P = QD$. Prouver que D appartient à $K[X]$, au sens où tous les coefficients de D appartiennent à K.

6. Soit P un élément de $\mathbb{R}[X]$ et $\alpha = a + \mathrm{i}b$ un nombre complexe tel que $b \neq 0$ et $P(\alpha) = 0$. Prouver que $Q(X) = X^2 - 2aX + a^2 + b^2$ divise P dans $\mathbb{R}[X]$.

4 Exercices

1. À l'aide de la formule du binôme et de la formule donnant le produit de deux polynômes, calculer, pour tous entiers positifs m et n, les deux membres de l'égalité

$$(X+1)^m(X+1)^n = (X+1)^{m+n}.$$

En choisissant $m = n$, en déduire la valeur de la somme $\displaystyle\sum_{k=0}^{n} \binom{n}{k}^2$.

Donner une démonstration alternative de la formule obtenue en dénombrant les parties à n éléments d'un ensemble X réunion disjointe de deux ensembles A et B de cardinal n.

2. a) Soit P et Q des éléments de $K[X]$ tels que $P^2 - XQ^2 = 0$. En considérant les degrés de P^2 et XQ^2, prouver que $P = Q = 0$.

b) On suppose que dans K l'équation $a^2 + b^2 = 0$ n'admet que la solution triviale $a = b = 0$. Prouver que si les éléments P, Q et R de $K[X]$ sont tels que $P^2 - XQ^2 + R^2 = 0$, alors $P = Q = R = 0$.

c) Donner des exemples de sous-corps K de $\mathbb{C}$ satisfaisant l'hypothèse de b).

3. a) Prouver que $X^3 = 1$ dans l'anneau $K[X]/(X^2 + X + 1)$.

En déduire les valeurs de l'entier naturel m pour lesquelles le polynôme $X^2 + X + 1$ divise le polynôme $(X+1)^m - X^m - 1$.

b) Prouver que pour tout entier naturel m, le polynôme $X^2 - X + 1$ divise $(X-1)^{m+2} + X^{2m+1}$.

4. Si a et b sont des éléments d'un anneau commutatif et $i \geqslant 1$ un entier, on rappelle l'identité remarquable

$$a^i - b^i = (a-b)\sum_{j=0}^{i-1} a^j b^{i-j-1}.$$

a) Soit a un élément de K et P un élément de $K[X]$ tels que $P(a) = 0$. Montrer que $X - a$ divise P.

a') Donner une autre preuve de a) en utilisant la division euclidienne.

b) Soit P un élément de $K[X]$. Montrer que $P - X$ divise $P \circ P - X$, c'est-à-dire qu'il existe Q dans $K[X]$ tel que $P(P(X)) - X = Q(X)(P(X) - X)$.

c) Soit $k \geqslant 1$ et $P = \displaystyle\sum_{i=1}^{k} X^{n_i}$ où les entiers $n_i \geqslant 0$ vérifient les congruences $n_i \equiv i - 1$ modulo k. Prouver que $(X-1)P$ est divisible par $X^k - 1$. En déduire que P est divisible

par $Q = \sum_{i=1}^{k} X^{i-1}$.

5. Soit Δ la transformation de $K[X]$ définie par

$$\Delta(P)(X) = P(X+1) - P(X).$$

On pose $\Delta^0 = \mathrm{Id}_{K[X]}$ puis, pour tout $n \geqslant 0$, $\Delta^{n+1} = \Delta \circ \Delta^n$.

Résoudre la question a) ou la question a'), puis la question b).

a) a1) Prouver que si $\deg(P) = d$, alors $\Delta^{d+1}(P) = 0$ et $\Delta^d(P) \neq 0$.

a2) On pose $C_0(X) = \binom{X}{0} = 1$ puis, pour tout nombre entier $d \geqslant 1$,

$$C_d(X) = \binom{X}{d} = \frac{1}{d!} X(X-1)\cdots(X-d+1).$$

Calculer $\Delta(C_d(X))$ pour tout $d \geqslant 0$. En déduire que, si $\deg(P) = d \geqslant 0$,

$$P(X) = \sum_{i=0}^{d} \Delta^i(P)(0)\, C_i(X).$$

a3) Déduire de a2) que si P est un polynôme de degré $d \geqslant 0$ à valeurs entières sur les entiers naturels, c'est-à-dire si $P(\mathbb{N}) \subset \mathbb{Z}$ (on rappelle que $\mathbb{N} \subset K$ puisque K est un sous-corps de $\mathbb{C}$), alors le polynôme P s'écrit de manière unique comme une combinaison linéaire à coefficients dans $\mathbb{Z}$ des polynômes binomiaux $C_i(X)$ pour $1 \leqslant i \leqslant d$.

a') Montrer que l'application $\Delta : K_d[X] \to K_{d-1}[X]$ est linéaire et calculer son image et son noyau.

b) Prouver que pour tout polynôme Q de degré au plus $d-1$, il existe un unique polynôme P de degré au plus d tel que $\Delta(P) = Q$ et $P(0) = 0$.

En déduire une expression pour les sommes $Q(0) + Q(1) + \cdots + Q(n)$ pour tout entier $n \geqslant 0$, puis la valeur des sommes $1^2 + 2^2 + \cdots + n^2$ et $1^3 + 2^3 + \cdots + n^3$.

6. Soit P un polynôme de $\mathbb{R}[X]$ tel que $P(t) \geqslant 0$ pour tout nombre réel t. Prouver qu'il existe deux polynômes A et B de $\mathbb{R}[X]$ tels que $P = A^2 + B^2$.

Indication : on pourra déduire de la décomposition de P en facteurs irréductibles dans $\mathbb{C}[X]$ qu'il existe un polynôme Q de $\mathbb{C}[X]$ tel que $P = Q\overline{Q}$, où on note

$$Q = \sum_{i=0}^{m} b_i X^i, \quad b_i \in \mathbb{C}, \quad \overline{Q} = \sum_{i=0}^{m} \overline{b}_i X^i.$$

7. Soient K un corps et $a_0, \ldots, a_n$ des éléments de K deux à deux distincts.

a) Prouver que si $b_0, \ldots, b_n$ sont des éléments de K, il existe un unique polynôme P de $K[X]$ de degré $\deg(P) \leqslant n$ tel que pour tout $0 \leqslant i \leqslant n$, $P(a_i) = b_i$.

Montrer les formules d'interpolation de Lagrange, qui affirment que P vaut

$$P = \sum_{i=0}^{n} b_i \frac{P_i(X)}{P_i(a_i)}, \quad P_i(X) = \prod_{j \neq i} (X - a_j).$$

b) Soit Q un polynôme de degré $\deg(Q) \leqslant n-1$ En considérant le coefficient du terme de degré n du polynôme P, fourni par la question a), tel que $\deg(P) \leqslant n$ et, pour tout $0 \leqslant i \leqslant n$, $P(a_i) = Q(a_i)$, montrer que

$$\frac{Q(a_0)}{P_0(a_0)} + \frac{Q(a_1)}{P_1(a_1)} + \cdots + \frac{Q(a_n)}{P_n(a_n)} = 0.$$

8. Calculer le pgcd de P et Q pour :

a) $P = X^2 + aX + b$ et $Q = X^2 + pX + q$ avec $p \neq a$;

b) $P = aX^2 + bX + c$ et $Q = 2X + a$ avec $a \neq 0$;

c) $P = X^3 + pX + q$ et $Q = 3X^2 + p$ avec $p \neq 0$.

En déduire qu'il existe des polynômes A et B dont les coefficients sont des expressions polynômiales à coefficients entiers en (a, b, p, q) dans le cas a), en (a, b, c) dans le cas b) et en (p, q) dans le cas c), tels que $AP + BQ$ soit égal à $(b-q)^2 + p(b-q)(p-a) + q(p-a)^2$ dans le cas a), $b^2 - 4ac$ dans le cas b), et $4p^3 - 27q^2$ dans le cas c).

Fiche 12 : Algèbre (6) Polynômes, racines et fractions rationnelles

« Un homme est comme une fraction dont le numérateur est ce qu'il est et le dénominateur ce qu'il pense de lui-même. Plus le dénominateur est grand, plus la fraction est petite. » Léon Tolstoï

1 Exercices de cours

On se donne un polynôme P de $\mathbb{C}[X]$ de degré $n \geqslant 1$ et sa factorisation dans $\mathbb{C}[X]$, comme

$$P(X) = \sum_{i=0}^{n} a_i X^i = a_n \prod_{j=1}^{n} (X - x_j).$$

1. Relations entre coefficients et racines a) Prouver que pour tout nombre complexe t,

$$\frac{P(X) - P(t)}{X - t} = \sum_{k=1}^{n} \left(\sum_{i=k}^{n} a_i \, t^{i-k} \right) X^{k-1}.$$

b) Soit $S_0 = n$ et, pour tout entier $m \geqslant 1$, S_m la $m$ème somme de Newton en les racines du polynôme P, définie comme $S_m = \sum_{j=1}^{n} (x_j)^m$. Montrer que

$$P'(X) = \sum_{j=1}^{n} \frac{P(X)}{X - x_j} = \sum_{j=1}^{n} \frac{P(X) - P(x_j)}{X - x_j}.$$

En déduire que les sommes de Newton S_m se calculent par les relations de récurrence, dites relations de Newton, suivantes : si $m \leqslant n - 1$,

$$\sum_{i=m+1}^{n} a_i S_{i-m} + (n - m) a_m = 0,$$

et si $m \geqslant n$,

$$\sum_{i=0}^{n} a_{n-i} S_{m-i} = 0.$$

2. Fonctions symétriques élémentaires Pour tout $1 \leqslant k \leqslant n$, on note

$$\sigma_k = \sum x_{j_1} \cdots x_{j_k},$$

où la somme porte sur tous les $k$uplets $(j_i)_{1 \leqslant i \leqslant k}$ d'entiers tels que

$$1 \leqslant j_1 < j_2 < \cdots < j_k \leqslant n.$$

Donc σ_k est la $k$ème fonction symétrique élémentaire en les racines du polynôme P.

a) Expliciter σ_1 et σ_n dans le cas général, et tous les σ_k si $n \leqslant 4$.

b) Préciser de combien de monômes en les inconnues $(x_j)_j$ on aurait besoin pour expliciter σ_k.

c) Vérifier les relations entre coefficients et racines affirmant que, pour tout $1 \leqslant k \leqslant n$,

$$a_n \sigma_k = (-1)^k a_{n-k}.$$

3. Dérivation des polynômes Le polynôme P' dérivé de P est

$$P' = \sum_{i=1}^{n} i a_i X^{i-1} = \sum_{j=0}^{d-1} (j+1) a_{j+1} X^j.$$

Pour tout entier $k \geqslant 0$, le $k$ème polynôme dérivée de P est le polynôme $P^{(k)}$ défini par récurrence par $P^{(0)} = P$, puis $P^{(k+1)} = (P^{(k)})'$ pour tout $k \geqslant 0$. Par exemple, $P^{(1)} = P'$.

a) Vérifier que $P^{(k)}(0) = k! a_k$ pour tout $k \geqslant 0$ et établir les formules de Leibniz selon lesquelles $(PQ)' = P'Q + PQ'$ et plus généralement, pour tout entier $k \geqslant 1$,

$$(PQ)^{(k)} = \sum_{i=0}^{k} \binom{k}{i} P^{(i)} Q^{(k-i)}.$$

b) Prouver que pour tout élément u de K, le morphisme de K-algèbre

$$E_{X-u} : K[X] \to K[X], X \mapsto X - u$$

d'évaluation en $X - u$ est un isomorphisme. Préciser l'isomorphisme inverse.

c) Déduire de b) l'existence de coefficients a_i éléments de $\mathbb{C}$ tels que

$$P(X) = \sum_{i=0}^{n} a_i (X-u)^i,$$

puis la formule de Taylor pour P, selon laquelle

$$P(X) = \sum_{i=0}^{n} \frac{P^{(i)}(u)}{i!} (X-u)^i.$$

d) Pour tout entier $k \geqslant 0$, on dit que u est une racine d'ordre k de P si et seulement s'il existe un polynôme Q tel que $Q(u) \neq 0$ et $P = (X-u)^k Q$. Déduire de c) que u est une racine d'ordre k de P si et seulement si

$$P(u) = P'(u) = \cdots = P^{(k-1)}(u) = 0, \qquad P^{(k)}(u) \neq 0.$$

Redémontrer ce résultat, par récurrence sur $k \geqslant 0$, en utilisant la première formule de Leibniz.

2 Exercices

1. Soient p et A deux nombres réels strictement positifs. Déterminer le ploynôme $X^2 + bX + c$ dont les deux racines sont la longueur et la largeur d'un rectangle de périmètre p et d'aire A. En déduire une relation entre le périmètre et l'aire d'un rectangle.

2. a) Déterminer le polynôme $X^2 + bX + c$ dont les racines valent $\mathrm{e}^{\mathrm{i}\pi/3}$ et $\mathrm{e}^{-\mathrm{i}\pi/3}$.

b) Prouver que trois points a, b et c du plan complexe $\mathbb{C}$ sont les sommets d'un triangle équilatéral si et seulement si $(a-c)^2 - (a-c)(b-c) + (b-c)^2 = 0$ si et seulement si $a^2 + b^2 + c^2 = ab + bc + ca$.

c) Déterminer tous les polynômes de degré trois dont les racines sont les sommets d'un triangle équilatéral.

3. a) Déterminer tous les polynômes de degré 4 dont les racines dans $\mathbb{C}$ sont les quatre sommets d'un carré. On pourra traiter d'abord le cas où le carré est centré en 0, puis se ramener à ce cas.

b) Plus généralement déterminer tous les polynômes de degré n dont les racines dans $\mathbb{C}$ sont les sommets d'un polygone régulier à n cotés.

c) Déduire de b) pour le cas $n = 3$ une solution alternative à 2. c).

4. En utilisant les relations de Newton, donner les polynômes de degré 3 dont les racines x, y et z dans $\mathbb{C}$ sont les solutions du système d'équations algébriques

$$x^2 + y^2 + z^2 = 2, \quad x^3 + y^3 + z^3 = 2, \quad x^4 + y^4 + z^4 = 2.$$

5. Soient $n \geqslant m \geqslant 0$ deux entiers et P et Q des polynômes à coefficients complexes de degrés $\deg(P) = m$ et $\deg(Q) = n$, donc

$$P(X) = \sum_{i=0}^{m} a_i X^i = a_m \prod_{j=1}^{m} (X - x_j), \qquad Q(X) = \sum_{i=0}^{n} b_i X^i = b_n \prod_{k=1}^{n} (X - y_k).$$

On suppose que pour tout entier $\ell \geqslant 1$, $\sum_{j=1}^{m}(x_j)^\ell = \sum_{k=1}^{n}(y_k)^\ell$. Prouver que $Q(X) = (b_n/a_m)X^{n-m}P(X)$.

6. Prouver que si $n > m \geqslant 1$ sont deux entiers positifs et a et b deux nombres complexes, le trinôme $X^n + aX^m + b$ n'admet une racine au moins triple que dans deux cas : ou bien $b = 0$ et $m \geqslant 3$, ou bien $a = b = 0$ et $n \geqslant 3$.

7. Soit $a = (a_k)_{0\leqslant k\leqslant n}$ une suite de nombres réels. Le nombre de changements de signe de la suite a est le nombre $V(a)$ de couples d'entiers (i, j) tels que $0 \leqslant i < j \leqslant n$, $a_i a_j < 0$, et $a_k = 0$ pour tout $i < k < j$.

a) Soit $P(X) = a_0 + a_1X + \cdots + a_nX^n$ un polynôme à coefficients réels. Prouver par récurrence sur le degré de P que le nombre de ses racines strictement positives, comptées avec leur multiplicité, est au plus égal à $V(a)$.

b) En déduire que si un polynôme à coefficients réels possède exactement k coefficients non nuls, alors il possède au plus $k-1$ racines positives et au plus $2k-1$ racines réelles distinctes.

8. Soit P un polynôme $P(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$ à coefficients complexes.

a) Prouver que si α est une racine de P, $|\alpha| < 1 + \max\{|a_i|\,;\, 0 \leqslant i \leqslant n-1\}$.

b) On suppose à présent que P est un polynôme à coefficients réels et que α est une racine réelle de P.

b1) Si $a_i \geqslant 0$ pour tout i, on pose $m = -1$ et $B = 0$. Sinon, on pose $m = \max\{i\,;\, a_i < 0\}$ et $B = \max\{-a_i\,;\, a_i < 0\}$. Prouver que

$$\alpha < 1 + B^{1/(n-m)}.$$

On pourra supposer que $\alpha > 1$ et remarquer qu'alors $0 \geqslant \alpha^n - B\,(\alpha^m + \cdots + 1)$.

b2) On note r le nombre des indices i tels que $a_i < 0$. Prouver que

$$\alpha \leqslant \max\{(r|a_i|)^{1/(n-i)}\,;\, a_i < 0\}.$$

9. Soit α une racine réelle d'un polynôme P de $\mathbb{Z}[X]$ de degré $d \geqslant 2$ et irréductible dans $\mathbb{Q}[X]$.

a) Prouver que pour tout nombre rationnel p/q avec p et q entiers et $q \geqslant 1$, $|P(p/q)| \geqslant 1/q^d$.

b) En déduire qu'il existe une constante $A > 0$ telle que pour tout nombre rationnel p/q, $|\alpha - (p/q)| \geqslant A/q^d$.

Indication : on pourra montrer que la valeur $A = \min(c, \delta)$ convient, où c et δ sont strictement positifs et tels que, si $|t - \alpha| \leqslant \delta$, alors $|P'(t)| \leqslant 1/c$.

Application : le nombre $\sum_n 10^{-n!}$ est transcendant.

10. Soient P_1/Q_1 et P_2/Q_2 deux fractions rationnelles à coefficients complexes dont les fonctions rationnelles associées coïncident sur un ensemble infini $X \subset \{Q_1 \neq 0\} \cap \{Q_2 \neq 0\}$ inclus dans l'intersection de leurs domaines de définition. Prouver que ces deux fractions rationnelles sont égales dans le corps $\mathbb{C}(X)$ des fractions rationnelles.

Préciser si $P_1 = P_2$ et $Q_1 = Q_2$ ou pas forcément.

11. Décomposer en éléments simples sur $\mathbb{C}$ puis sur $\mathbb{R}$ les fractions rationnelles suivantes :

$$R_1 = \frac{X^2+1}{X(X^2-1)}, \quad R_2 = \frac{2}{(X-1)(X-2)(X-3)},$$

$$R_3 = \frac{X^5 - X^3 - X^2}{X^2-1}, \quad R_4 = \frac{4X^3}{(X^2+1)^2},$$

et

$$R_5 = \frac{X^6 - X^2 + 1}{(X-1)^2}, \quad R_6 = \frac{3X^2+3}{X^3 - 3X - 2}, \quad R_7 = \frac{X^5}{(X^4-1)^2}.$$

12. Soit α un nombre réel. Décomposer en éléments simples sur $\mathbb{C}$ la fraction rationnelle

$$R(x) = \frac{1}{1 - 2\cos(\alpha)x + x^2}.$$

En déduire, pour tout entier naturel n, la dérivée $n$ème de la fonction R.

Fiche 13 : Algèbre (7) Espaces vectoriels, dimension, dualité

« Nous autres comédiens sommes les vecteurs de l'imaginaire des réalisateurs. » Nicole Kidman, Studio magazine

1 Vrai ou faux

1. On note $\{f(x,y,z) = 0\}$ la partie de $\mathbb{R}^3$ formée des triplets (x,y,z) tels que $f(x,y,z) = 0$. Déterminer si les parties suivantes de $\mathbb{R}^3$ sont des sous-espaces vectoriels : $\{2x+3y=0\}$, $\{2x+3y=4\}$, $\{x^2+y^2=0\}$, $\{x^3+x=0\}$, $\{x^2-y^2+z^2=0\}$, $\{x^2-6xy+9y^2-z^2=0\}$, et $\{x^2-6xy+9y^2+z^2=0\}$.

2. Soit $a=(2,1,-3)$, $b=(3,2,-5)$ et $c=(1,1,-2)$.

a) Les vecteurs a, b et c forment une base de $\mathbb{R}^3$.

b) Le vecteur $x=(5,2,-7)$ est une combinaison linéaire de a, b et c.

2 Exercices de cours

1. Soient M et N deux sous-espaces vectoriels d'un espace vectoriel E de dimension finie. On note $M+N$ le sous-espace vectoriel de E engendré par M et N. En considérant l'application linéaire $f : M+N \to E$ définie par $f(m,n) = m+n$, établir la relation

$$\dim(M) + \dim(N) = \dim(M+N) + \dim(M \cap N).$$

2. a) Soit E et F des espaces vectoriels et $f : E \to F$ une application linéaire. Prouver qu'il existe une application linéaire $g : \mathrm{Im}(f) \to E$ telle que $f \circ g$ est l'application identité de $\mathrm{Im}(f)$ dans F.

b) En déduire que E est isomorphe à la somme directe $\ker(f) \oplus \mathrm{Im}(f)$ puis que, si E est de dimension finie, $\ker(f)$ et $\mathrm{Im}(f)$ sont de dimension finie et vérifient

$$\dim(E) = \dim(\ker(f)) + \dim(\mathrm{Im}(f)).$$

c) Soit f un endomorphisme d'un espace vectoriel E de dimension finie. Prouver que f est injectif si et seulement si f est surjectif. Préciser si ce résultat subsiste si on ne suppose plus que la dimension de E est finie.

4. Soit k un corps et E un espace vectoriel sur k. On note E^* le dual de E, ensemble des formes linéaires sur E, c'est-à-dire des applications k-linéaires de E dans k.

a) Vérifier que si E est un k-espace vectoriel à gauche, E^* est un k-espace vectoriel à droite pour la loi externe $E^* \times k \to E^*$, $(\varphi, \lambda) \mapsto [x \mapsto \varphi(x)\lambda]$.

a') Vérifier que si E est un k-espace vectoriel à droite, E^* est un k-espace vectoriel à gauche pour la loi externe $k \times E^* \to E^*$, $(\lambda, \psi) \mapsto [x \mapsto \lambda\psi(x)]$.

b) Prouver que l'application $J : E \to (E^*)^*$ définie par $J(x) = [\varphi \mapsto \varphi(x)]$ est linéaire et injective.

c) Si E est de dimension finie n et muni d'une base $(e_i)_{1 \leqslant i \leqslant n}$, on note e_i^* la forme linéaire élément de E^* définie par $e_i^*(e_i) = 1$ et $e_i^*(e_j) = 0$ si $i \neq j$. Prouver que $(e_i^*)_{1 \leqslant i \leqslant n}$ est une base de E^*.

En déduire que si E est de dimension finie alors $\dim(E) = \dim(E^*)$ et l'application canonique J définie en b) est un isomorphisme.

d) Soit $E = \mathbb{R}[X]$ muni de la base $(e_n)_{n \in \mathbb{N}}$ avec $e_n = X^n$. Prouver que si φ appartient au sous-espace de E^* engendré par les e_n^* alors il existe un entier $N = N_\varphi$ tel que pour tout polynôme P divisible par X^N, $\varphi(P) = 0$.

Donner un exemple de forme linéaire φ élément de E^* qui n'a pas cette propriété.

En déduire que les formes linéaires $(e_n^*)_{n \in \mathbb{N}}$ ne forment pas une base de E^* et que l'application canonique J définie en b) n'est pas surjective.

5. Soit E et F des espaces vectoriels de dimensions finies n et m munis de bases $(e_i)_{1 \leqslant i \leqslant n}$ et $(f_j)_{1 \leqslant j \leqslant m}$ et $f : E \to F$ une application linéaire.

a) Soit A la matrice de f dans les bases $(e_i)_i$ et $(f_j)_j$. Prouver que l'application $f^* : F^* \to E^*$ définie par $\psi \mapsto \psi \circ f$ est linéaire, et que sa matrice dans les bases $(f_j^*)_j$ et $(e_i^*)_i$ est la transposée A^t de la matrice A.

b) Soit A des matrices sur k de tailles respectives $n \times m$ et $m \times p$. Montrer que les transposées A^t et B^t, de tailles respectives $m \times n$ et $p \times m$, vérifient la relation $B^t A^t = (AB)^t$.

Retrouver cette relation en explicitant les matrices A^t, B^t, $B^t A^t$ et $(AB)^t$.

3 Exercices

1. Prouver que les vecteurs

$$a = (1, 2, -1, -2),\ b = (2, 3, 0, -1),\ c = (1, 2, 1, 3),\ d = (1, 3, 1, 0),$$

forment une base de $\mathbb{R}^4$ et donner la décomposition dans cette base du vecteur $x = (7, 13, -1, -2)$.

b) Montrer que les vecteurs a, b et c ci-dessous forment une base de $\mathbb{R}^3$ et trouver les coordonnées du vecteur x par rapport à cette base, dans les deux cas suivants :

b1) $a = (1, 1, 1)$, $b = (1, 1, 2)$, $c = (1, 2, 3)$, $x = (6, 9, 14)$;
b2) $a = (2, 1, -3)$, $b = (3, 2, -5)$, $c = (1, -1, 1)$, $x = (6, 2, -7)$.

2. a) Soit e_i le $i$ème vecteur de la base canonique de $\mathbb{R}^{n+1}$, dont la $i$ème coordonnée vaut 1 et toutes les autres 0. Prouver que les vecteurs

$$f_0 = e_1 - e_{n+1}, \quad f_i = e_{i+1} - e_i, \ 1 \leqslant i \leqslant n,$$

engendrent le sous-espace E de $\mathbb{R}^{n+1}$ d'équation $x_1 + \cdots + x_{n+1} = 0$.

b) Prouver que toute partie à n éléments de $\{f_0, \ldots, f_n\}$ est une base de E.

c) Donner l'expression de f_0 dans la base $(f_i)_{1 \leqslant i \leqslant n}$.

En déduire la matrice dans la base $(f_i)_{1 \leqslant i \leqslant n}$ de la restriction à E de la permutation circulaire des coordonnées $(x_i)_{1 \leqslant i \leqslant n+1} \mapsto (x_{n+1}, x_1, \ldots, x_n)$.

d) Donner un suplémentaire de E dans $\mathbb{R}^{n+1}$.

3. a) Pour $v = (z_1, \ldots, z_n)$ dans $\mathbb{C}^n$, on note $\|v\| = \sum_{i=1}^{n} |z_i|$. Prouver que l'application $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}_+$ est une norme sur $\mathbb{C}^n$, c'est-à-dire que l'on a :

1. $\|v\| = 0$ si et seulement si $v = 0$;
2. pour tous v et w dans $\mathbb{C}^n$ et α dans $\mathbb{C}$, $\|\alpha v\| = |\alpha| \, \|v\|$;
3. pour tous v et w dans $\mathbb{C}^n$, $\|v + w\| \leqslant \|v\| + \|w\|$.

b) Soit $1 \leqslant m \leqslant n$. Pour tout $1 \leqslant i \leqslant m$, soit $v_i = (z_{i,k})_{1 \leqslant k \leqslant n}$ un vecteur de $\mathbb{C}^n$ tel que

$$|z_{i,i}| > \sum_{j=1, j \neq i}^{n} |z_{i,j}|$$

Soit $\lambda_1 v_1 + \cdots \lambda_m v_m = 0$ une relation linéaire entre les v_i. En notant e_i le $i$ème vecteur de la base canonique, vérifier l'égalité

$$\left\| \sum_{i=1}^{m} \lambda_i (v_i - z_{i,i} e_i) \right\| = \left\| \sum_{i=1}^{m} \lambda_i z_{i,i} e_i \right\|.$$

En déduire que les vecteurs $(v_i)_{1 \leqslant i \leqslant m}$ sont linéairement indépendants.

4. (Suite de l'exercice 3) a) Si $m \leqslant n - 1$, prouver qu'il existe des vecteurs $(v_i)_{m+1 \leqslant i \leqslant n}$ avec, pour tout $m + 1 \leqslant i \leqslant n$,

$$v_i = (z_{i,k})_{1 \leqslant k \leqslant n}, \qquad |z_{i,i}| > \sum_{j=1, j \neq i}^{n} |z_{i,j}|.$$

b) On suppose que $m = n$, on note $w_j = (z_{k,j})_{1 \leqslant k \leqslant n}$, et on suppose que $\sum_{j=1}^{n} \mu_j w_j = 0$. Prouver que tous les coefficients μ_j sont nuls.

On pourra considérer un indice $1 \leqslant j_0 \leqslant n$ tel que $|\mu_{j_0}| \geqslant |\mu_j|$ pour tout $1 \leqslant j \leqslant n$, puis la $j_0$ème coordonnée de la somme des $\mu_j w_j$.

c) Déduire de a) et b) et de l'exercice de cours 2 une autre solution de l'exercice 3.

5. Pour tout nombre réel α, on note $f_\alpha : \mathbb{R} \to \mathbb{R}$ la fonction définie par $f_\alpha(t) = \mathrm{e}^{\alpha t}$.

a) Prouver que la famille $(f_\alpha)_{\alpha \in \mathbb{R}}$ est libre dans l'espace vectoriel $\mathcal{F}(\mathbb{R}, \mathbb{R})$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

b) Soient $\alpha_1 < \alpha_2 < \cdots < \alpha_n$ des nombres réels et $(a_1, \ldots, a_n)$ un élément non nul de $\overset{n}{R}$. Prouver que l'ensemble des nombres réels $x > 0$ tels que $\sum_{i=1}^{n} a_i x^{\alpha_i} = 0$ comporte strictement moins de n éléments.

En déduire que pour toute partie infinie $T \subset \mathbb{R}$, la famille $(f_\alpha^T)_{\alpha \in \mathbb{R}}$ des restrictions f_α^T à T des fonctions f_α est libre dans l'espace vectoriel $\mathcal{F}(T, \mathbb{R})$ des applications de T dans $\mathbb{R}$.

6. Pour $0 \leqslant i \leqslant n$, soit $f_i : E_i \to E_{i+1}$ une application linéaire. On suppose que $E_0 = \{0\} = E_{n+1}$ et que, pour tout $1 \leqslant i \leqslant n$, $\ker(f_i) = f_{i-1}(E_{i-1})$. Prouver que si les E_i sont de dimension finie alors

$$\sum_{i=1}^{n} (-1)^i \dim(E_i) = 0.$$

7. Pour tout nombre réel a, on note $g_a : \mathbb{R} \to \mathbb{R}$ l'application $g_a(x) = |x - a|$.

a) Soit A une partie finie de $\mathbb{R}$ et $a_1 < a_2 < \cdots < a_n$ les éléments de A. On note E l'espace vectoriel des fonctions dérivables sur $\mathbb{R} \setminus A$ et F l'espace vectoriel des fonctions sur $\mathbb{R} \setminus A$. En considérant l'image de g_{a_n} par l'opérateur de dérivation $D : E \to F$, défini par $D(f) = f'$, et, pour $1 \leqslant i \leqslant n-1$, l'image de $g_{a_{i+1}} - g_{a_i}$, montrer que les fonctions $(g_{a_i})_{1 \leqslant i \leqslant n}$ sont linéairement indépendantes dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

b) Prouver que la famille $(g_a)_{a \in \mathbb{R}}$ est une famille libre dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

8. Soit $E = \mathbb{R}[X]$ l'espace vectoriel des polynômes à coefficients réels. Déterminer l'image et le noyau de l'endomorphisme de dérivation $D : E \to E$ défini par $D(P) = P'$.

9. Soit E un espace vectoriel de dimension finie et f un endomorphisme de E. Prouver que $\ker(f) = f(E)$ si et seulement si $\dim(E) = 2\dim(f(E))$ et $f^2 = 0$.

10. Soit E un espace vectoriel et f un endomorphisme de E tel que, pour tout x dans E, il existe un entier $n \geqslant 1$ tel que $f^n(x) = 0$. Prouver que $h = \mathrm{Id}_E - f$ est injective. Pour tout élément x de E, soit $n \geqslant 0$ tel que $f^n(x) = 0$ et

$$g(x) = x + \sum_{i=1}^{n-1} f^i(x).$$

Montrer que cette définition de $g(x)$ ne dépend pas du choix de n et que l'application g ainsi définie est linéaire. Calculer $g \circ h$ et $h \circ g$.

11. La trace d'une matrice carrée $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ de $M_{n,n}(k)$ est

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{i,i}.$$

Prouver que si A et B sont des matrices de $M_{n,m}(k)$ et $M_{m,n}(k)$, alors les produits AB et BA sont bien définis et que $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

b) En déduire que si f est un endomorphisme d'un espace vectoriel de dimension finie muni d'une base $(e_i)_i$, la trace de la matrice de f dans la base $(e_i)_i$ ne dépend pas du choix de la base.

12. a) Pour tous nombres complexes a et b, on note $h_{a,b} = \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix}$. Prouver que l'ensemble $H = \{h_{a,b}\,;\, a, b \in \mathbb{C}\}$ est une sous-algèbre réelle de l'algèbre $M_{2,2}(\mathbb{C})$ des matrices carrées 2×2 à coefficients complexes vue comme algèbre sur $\mathbb{R}$.

On pose $u = h_{\mathrm{i},0}$, $v = h_{0,1}$ et $w = h_{0,\mathrm{i}}$. Calculer, pour tous x et y dans l'ensemble $\{u, v, w\}$, les produits xy. Préciser si l'algèbre H est commutative.

b) Prouver que l'application $\sigma : H \to H$ définie par $\sigma(h_{a,b}) = h_{\overline{a},-b}$ est un morphisme additif tel que, pour tous h et h' éléments de H, $\sigma(hh') = \sigma(h')\sigma(h)$.

c) Prouver qu'il existe une application $n : H \to \mathbb{R}_+$ telle que pour tout élément h de H, $n(h) = 0$ si et seulement si $h = 0$ et $h\sigma(h) = \sigma(h)h = n(h)I_2$. Déduire de cette dernière relation et de b) que n est un morphisme multiplicatif, c'est-à-dire que $n(I_2) = 1$ et, pour tous h et h' éléments de H, $n(hh') = n(h)n(h')$.

d) Déduire de c) que si $h \neq 0$ alors $h^{-1} = n(h)^{-1}\sigma(h)$ est un inverse de h pour la multiplication de H, c'est-à-dire que $hh^{-1} = h^{-1}h = I_2$.

Ainsi H est un « corps non commutatif ». En admettant que la théorie de la dimension a lieu dans les espaces vectoriels sur les corps non commutatifs, vérifier que les questions b), c) et d) de l'exercice de cours 2 sont également vraies quand $K = H$.

e) On considère la matrice $A = \begin{pmatrix} 1 & u \\ v & v\,u \end{pmatrix}$. Montrer que $r_g(A) = c_d(A) = 1$ et $r_d(A) = c_g(A) = 2$.

Fiche 14 : Algèbre (8) Matrices et réduction des endomorphismes

« The Matrix is everywhere. It is all around us. Even now, in this very room. » Morpheus, The Matrix

1 Matrices

Soit K un sous-corps de $\mathbb{C}$.

1. Soit $E = M_{2,2}(K)$ et $\sigma : E \to E$ définie par

$$\sigma\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

a) Prouver que σ est un anti-automorphisme involutif d'algèbre, c'est-à-dire que $\sigma^2 = \mathrm{Id}_E$, $\sigma(I_2) = I_2$ et, pour tout élément x de K et tous éléments A et B de E,

$$\sigma(A+B) = \sigma(A) + \sigma(B), \quad \sigma(xA) = x\sigma(A), \quad \sigma(AB) = \sigma(B)\sigma(A).$$

b) Déterminer l'ensemble des points fixes de σ et vérifier que pour toute matrice A dans E, les matrices $A + \sigma(A)$ et $\sigma(A)A$ sont fixés par σ.

c) Déduire de ce qui précède une preuve des identités

$$(ax+bz)(cy+dt) - (ay+bt)(cx+dz) = (ad-bc)(xt-yz),$$

et

$$A^2 - (a+d)A + (ad-bc)I_2 = 0 \quad \text{avec } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

2. Soit U l'ensemble des matrices $(a_{i,j})_{1 \leqslant i,j \leqslant n}$ de $M_{n,n}(K)$ telles que, pour tous $1 \leqslant i < j \leqslant n$, $a_{i,j} = 0$, et, pour tous $1 \leqslant i \leqslant n$, $a_{i,i} = 1$. Soit T l'ensemble des matrices $(a_{i,j})_{1 \leqslant i,j \leqslant n}$ de $M_{n,n}(K)$ telles que, pour tous $1 \leqslant i < j \leqslant n$, $a_{i,j} = 0$, et, pour tous $1 \leqslant i \leqslant n$, $a_{i,i} \neq 0$.

Prouver que U et T sont des sous-groupes de $GL_n(K)$ et que U est un sous-groupe distingué de T.

3. Soient A et B deux matrices $n \times n$ telles que $AB = A + B$. Calculer $(I_n - A)(I_n - B)$ en déduire que A et B commutent.

4. a) Soit A une matrice de $M_{n,n}(K)$ telle que pour toute matrice B de $M_{n,n}(K)$, $AB = BA$. Prouver qu'il existe un élément x de K tel que $A = xI_n$.

b) Soit A et B des matrices de $M_{n,n}(K)$. Calculer $\mathrm{Tr}(AB)$. En déduire que si f est une forme linéaire sur $M_{n,n}(K)$, il existe une unique matrice A telle que pour toute matrice X de $M_{n,n}(K)$, $f(X) = \mathrm{Tr}(AX)$.

c) Soit f une forme linéaire sur $M_{n,n}(K)$ telle que pour toutes matrices X et Y de $M_{n,n}(K)$, $f(XY) = f(YX)$. Prouver qu'il existe un élément x de K tel que $f = x\mathrm{Tr}$.

5. Soit E un espace vectoriel de dimension finie sur K et f et g des endomorphismes de E. Montrer que

$$\mathrm{rang}(f) + \mathrm{rang}(g) - \dim(E) \leqslant \mathrm{rang}(f \circ g) \leqslant \min(\mathrm{rang}(f), \mathrm{rang}(g)).$$

6. Soit $n \geqslant 0$ un entier et E l'espace vectoriel des polynômes de degré au plus n.

a) Prouver qu'il existe un endomorphisme f de E tel que, pour tout P dans E,

$$\frac{\mathrm{d}}{\mathrm{d}x}(P(x)\,\mathrm{e}^{-x}) = f(P)(x)\,\mathrm{e}^{-x}.$$

b) Déterminer le noyau et l'image de f.

c) Donner la matrice de f dans la base $(X^i)_{0 \leqslant i \leqslant n}$ de E.

7. Soit $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Calculer A^n pour tout $n \in \mathbb{N}$.

8. Pour $1 \leqslant i, j \leqslant n$ on pose $a_{i,j} = 1$ si $j = i + 1$ et $a_{i,j} = 0$ si $j \neq i + 1$.

a) Écrire sous forme de tableau la matrice $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ et, pour tout $1 \leqslant k \leqslant n$, calculer A^k.

b) Soit x dans K, calculer l'inverse de la matrice $M(x) = (m_{i,j}(x))_{1 \leqslant i,j \leqslant n}$ définie par $m_{i,j}(x) = x^{j-i}$ si $j \geqslant i$ et $m_{i,j}(x) = 0$ si $j < i$.

9. Soit $n \geqslant 2$ et $M = (m_{i,j})_{1 \leqslant i,j \leqslant n}$ définie par $m_{i,i} = 0$ et $m_{i,j} = 1/(n-1)$ pour tous $i \neq j$.

Calculer M^2. En déduire que M est inversible et calculer M^{-1}.

10. On rappelle que K désigne un sous-corps de $\mathbb{C}$ et on se donne r vecteurs colonnes $(v_i)_{1\leqslant i\leqslant r}$, donc chaque v_i est un élément de $M_{n,1}(K)$.

a) On suppose que les vecteurs $(v_i)_{1\leqslant i\leqslant r}$ sont K-linéairement indépendants. On note M la matrice de $M_{n,r}(K)$ dont la $i$ème colonne vaut v_i. Prouver qu'il existe une matrice N appartenant à $M_{r,n}(K)$ telle que $NM = I_r$. En déduire que, vus comme vecteurs dans $M_{n,1}(\mathbb{C})$, les vecteurs $(v_i)_{1\leqslant i\leqslant r}$ sont $\mathbb{C}$-linéairement indépendants.

b) Établir l'égalité des rangs sur K et sur $\mathbb{C}$ de la famille $(v_i)_{1\leqslant i\leqslant r}$.

11. Soit $AX = B$ un système linéaire à coefficients dans un sous-corps K de $\mathbb{C}$, donc A et B sont des éléments de $M_{m,n}(K)$ et $M_{m,1}(K)$ donnés et l'élément X de taille $n \times 1$ est l'inconnue. On note $\mathcal{S}(\mathbb{C})$ l'ensemble des solutions du système dans $M_{n,1}(\mathbb{C})$, et $\mathcal{S}(K) = \mathcal{S}(\mathbb{C}) \cap M_{n,1}(K)$ l'ensemble des solutions du système dans $M_{n,1}(K)$.

Montrer que $\dim_K \mathcal{S}(K) = \dim_{\mathbb{C}} \mathcal{S}(\mathbb{C})$.

12. a) Prouver que si X est une matrice $n \times n$ à coefficients dans K, $X^tX = I_n$ si et seulement si $XX^t = I_n$.

b) Prouver que l'ensemble des matrices X de $M_{n,n}(K)$ telles que $XX^t = I_n$ est inclus dans $GL_n(K)$ et forme un sous-groupe de $GL_n(K)$.

2 Réduction des endomorphismes

1. Soit E un espace vectoriel de dimension finie n sur un corps K et f un endomorphisme de E.

a) Soit v un vecteur non nul de E. Soit E_v le sous-espace vectoriel de E engendré par v et par tous les $f^k(v)$ pour $k \geqslant 1$. Soit d la dimension de E_v. Montrer qu'il existe des coefficients a_i tels que

$$a_0 v + a_1 f(v) + \cdots + a_{d-1} f^{d-1}(v) + f^d(v) = 0.$$

b) On note $P_v = X^d + a_{d-1}X^{d-1} + \cdots + a_0$.

Prouver que $P_v(f) = f^d + a_{d-1}f^{d-1} + \cdots + a_0\mathrm{Id}_E$ induit sur E_v l'endomorphisme nul.

c) Prouver que f induit un endomorphisme de l'espace quotient E/E_v.

d) Prouver qu'il existe une base $(e_i)_{1\leqslant i\leqslant n}$ de E telle que $e_1 = v$ et, pour tout $2 \leqslant i \leqslant d$, $e_i = f^{i-1}(v)$. Écrire la matrice de f dans cette base et en déduire une autre solution de c).

e) Déduire de ce qui précède, en raisonnant par récurrence sur la dimension de E, qu'il existe un polynôme P de $K[X]$ de degré au plus $\dim(E)$ tel que $P(f) = 0$.

2. a) Soit $C = (c_{i,j})_{i,j}$ la matrice de taille $n \times n$ telle que $c_{i,i-1} = 1$ pour tout $2 \leqslant i \leqslant n$, $c_{i,n} = a_i$ pour tout $1 \leqslant i \leqslant n$, et $c_{i,j} = 0$ pour tous les autres couples (i, j). Calculer un polynôme P de $K[X]$ de degré n tel $P(C) = 0$.

b) Soit E un espace vectoriel de dimension n et f et g deux endomorphismes de E. On suppose qu'il existe un polynôme de degré n et des vecteurs v et w de E tels que $P(f) = P(g) = 0$ et tels que $(v, f(v), \ldots, f^{n-1}(v))$ et $(w, g(w), \ldots, g^{n-1}(w))$ sont tous deux linéairements indépendants.

Montrer que f et g sont conjugués, c'est-à-dire qu'il existe un isomorphisme $\varphi : E \to E$ tel que $f = \varphi^{-1} \circ g \circ \varphi$.

c) Soit f et g les endomorphismes de $\mathbb{C}^3$ dont les matrices dans la base canonique sont $M = \begin{pmatrix} 3 & 0 & 1 \\ -1 & 2 & -1 \\ -2 & 0 & 0 \end{pmatrix}$ et $N = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. Calculer les polynômes caractéristiques de f et g. Préciser si f et g sont conjugués ou non.

3. Soit f un endomorphisme d'un espace vectoriel E sur un corps K et P un polynôme de $K[X]$ tel que $P(f) = 0$. Un élément a de K est une valeur propre de f s'il existe un vecteur v non nul de E tel que $f(v) = av$.

Prouver que, si a est une valeur propre de f, $P(a) = 0$.

En déduire que, si E est de dimension finie n, les valeurs propres de f sont en nombre au plus n.

4. Soit E un espace vectoriel et f et g deux endomorphismes de E tels que $f \circ g = g \circ f$. Montrer que $\ker(g)$ et $\mathrm{Im}(g)$ sont stables par f, c'est-à-dire que $f(\ker(g)) \subset \ker(g)$ et $f(\mathrm{Im}(g)) \subset \mathrm{Im}(g)$.

5. a) **Théorème des noyaux** Soit A et B deux polynômes de $K[X]$ premiers entre eux et $P = AB$, E un espace vectoriel de dimension finie sur K, et f un endomorphisme de E. Montrer que

$$\ker(P(f)) = \ker(A(f)) \oplus \ker(B(f)).$$

b) En déduire qu'un endomorphisme f d'un espace vectoriel de dimension finie est diagonalisable si et seulement s'il existe un polynôme P de $K[X]$, scindé et sans racines multiples, tel que $P(f) = 0$.

On rappelle que P est scindé si P est un multiple d'un produit de monômes $X - a_i$ avec a_i dans K.

6. Soit E un espace vectoriel et f et g deux endomorphismes de E tels que $f \circ g = g \circ f$.

a) Montrer que les espaces propres de f sont stables par g, d'abord directement, puis en utilisant l'exercice 4.

a') Même question pour les espaces $E_i(a,f)$ définis, pour a dans K et i dans $\mathbb{N}$, par $E_i(a,f) = \ker((a\mathrm{Id}_E - f)^i)$.

b) En déduire que si E est de dimension finie et si f et g sont diagonalisables, alors f et g sont diagonalisables dans la même base.

c) Application : résoudre l'équation $X^2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$ dans $M_{3,3}(\mathbb{C})$.

c') résoudre l'équation $X^2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ dans $M_{3,3}(\mathbb{C})$.

7. Soit E un espace vectoriel de dimension finie sur $\mathbb{C}$ et f un endomorphisme de E tel que $f \circ f = -\mathrm{Id}_E$. Prouver que f est diagonalisable.

8. Soit $\mathcal{M}$ l'ensemble des matrices $(a_{i,j})_{i,j}$ de $M_{n,n}(\mathbb{R})$ telles que pour tout $1 \leqslant i,j \leqslant n$, $0 \leqslant a_{i,j} \leqslant 1$, et, pour tout $1 \leqslant i \leqslant n$, $\sum_{j=1}^{n} a_{i,j} = 1$.

Prouver que si M et N appartiennent à $\mathcal{M}$, MN aussi, et que tout M dans $\mathcal{M}$ admet 1 pour valeur propre.

Fiche 15 : Algèbre (9) Déterminants

« Faire des mathématiques, c'est donner le même nom à des choses différentes. » Henri Poincaré

1 Rappels de cours

Soit V un espace vectoriel sur un corps K. Une forme n-linéaire sur V est une application $f : V^n \to K$ telle que, pour tout $1 \leqslant i \leqslant n$ et pour tout choix de vecteurs $(v_j)_{j\neq i}$ de V, l'application

$$v \mapsto f\,(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n),$$

est linéaire. La forme f est alternée si $f(v_1, \ldots, v_n) = 0$ dès que $v_i = v_j$ pour au moins une paire d'indices $j \neq i$. En ce cas, pour tout élément s de $\mathfrak{S}_n$,

$$f(v_{s(1)}, \ldots, v_{s(n)}) = \varepsilon(s) f(v_1, \ldots, v_n),$$

où $\varepsilon(s)$ désigne la signature de la permutation s. On note $\mathfrak{A}_n(V)$ l'espace vectoriel des formes n-linéaires alternées sur V.

Si $V = K^n$, le déterminant sur K^n est la forme det définie sur $\mathfrak{A}_n(K^n)$ comme suit : pour tout choix de n vecteurs $v_i = (v_{i,j})_{1\leqslant j\leqslant n}$ de K^n,

$$\det(v_1, \ldots, v_n) = \sum_{s\in\mathfrak{S}_n} \varepsilon(s) v_{1,s(1)} \cdots v_{n,s(n)}.$$

Une définition équivalente est que det est l'unique forme n-linéaire alternée telle que $\det(e_1, \ldots, e_n) = 1$, où $(e_i)_{1\leqslant i\leqslant n}$ désigne la base canonique de K^n.

Pour tout espace vectoriel V de dimension finie d et pour tout $n \geqslant 1$, la dimension de $\mathfrak{A}_n(V)$ vaut 0 si $n \geqslant d+1$ et $\binom{d}{n}$ si $n \leqslant d$. En particulier cette dimension vaut 1 si $n = d$, donc $\mathfrak{A}_d(V)$ est engendré par la forme déterminant dans une base quelconque et les formes déterminants dans deux bases différentes sont proportionnelles.

Tout morphisme d'espaces vectoriels $f : V \to W$ induit, pour tout $n \geqslant 1$, un endomorphisme $\mathfrak{A}_n(f) : \mathfrak{A}_n(W) \to \mathfrak{A}_n(V)$, défini par

$$\mathfrak{A}_n(f)(\alpha)(v_1, \ldots, v_n) = \alpha(f(v_1), \ldots, f(v_n)).$$

Dans le cas où $V = W$, le déterminant d'un endomorphisme f d'un espace vectoriel V de dimension d est l'unique élément $\det f$ de K tel que pour tout α dans $\mathfrak{A}_d(V)$,

$$\mathfrak{A}_d(f)(\alpha) = (\det f) \cdot \alpha.$$

Si $V = K^d$ et si M est la matrice de f dans la base canonique,

$$\det f = \det(C_1, \ldots, C_d),$$

où C_i désigne la $i$ème colonne de la matrice M. On note ce nombre $\det M$.

Une conséquence est que $\det f \circ g = \det f \cdot \det g$ pour tous endomorphismes f et g. En particulier, $\det MN = \det M \cdot \det N$ pour tout $n \geqslant 1$ et toutes matrices M et N de taille $n \times n$.

Soit $M = (m_{i,j})_{1 \leqslant i,j \leqslant n}$ une matrice $n \times n$. Pour tous $1 \leqslant i, j \leqslant n$, soit $M_{i,j}$ la matrice obtenue à partir de M en effaçant la $i$ème ligne et la $j$ème colonne. Donc $M_{i,j}$ est de taille $(n-1) \times (n-1)$ et le coefficient (a, b) de $M_{i,j}$ est $m_{a,b}$ si $a < i$ et $b < j$, $m_{a,b+1}$ si $a < i$ et $b \geqslant j$, $m_{a+1,b}$ si $a \geqslant i$ et $b < j$, et $m_{a+1,b+1}$ si $a \geqslant i$ et $b \geqslant j$. Pour tous $1 \leqslant i, j \leqslant n$, le cofacteur (i, j) de M est

$$C_{i,j}(M) = (-1)^{i+j} \det M_{i,j}.$$

La matrice des cofacteurs de M est $C(M) = (C_{i,j}(M))_{1 \leqslant i,j \leqslant n}$. On a

$$MC(M)^t = C(M)^t M = \det M \, I_n.$$

Le développement du déterminant de M suivant la ligne i ou suivant la colonne i donne le $i$ème terme diagonal de cette relation. Quant aux termes non diagonaux, le terme (i, j) avec $i \neq j$ s'obtient en développant le déterminant de la matrice obtenue en remplaçant dans M la colonne j par la colonne i et en développant le déterminant obtenu, qui est nul puisque deux de ses colonnes sont égales, par rapport à la ligne i.

2 Exercices de cours

1. Soit V un espace vectoriel, $(v_i)_{1 \leqslant i \leqslant n}$ une famille de vecteurs linéairement dépendants et f une forme de $\mathfrak{A}_n(V)$. Prouver que $f(v_1, \ldots, v_n) = 0$.

2. Prouver que la forme det définie par une somme sur $\mathfrak{S}_n$ dans les rappels de cours appartient à $\mathfrak{A}_n(K^n)$.

3. Prouver que $\varepsilon(s^{-1}) = \varepsilon(s)$ pour tout élément s de $\mathfrak{S}_n$. En déduire que, pour toute matrice M de taille $n \times n$, $\det M^t = \det M$.

4. Soit M une matrice $n \times n$. On considère la matrice $B = XI_n - M$, donc B appartient à l'algèbre $M_{n,n}(K[X])$ des matrices $n \times n$ à coefficients dans l'anneau des polynômes $K[X]$. On considère le polynôme caractéristique de M, défini par

$$\chi_M(X) = \det B = \det(XI_n - M).$$

Donc $\chi_M(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$ pour certains coefficients a_i dans K. Enfin, soit C la matrice des cofacteurs de B.

a) Prouver qu'il existe des matrices A_i éléments de $M_{n,n}(K)$, telles que

$$C^t = A_{n-1}X^{n-1} + \cdots + A_0.$$

b) Utiliser l'identité $C^t B = \det(B)\, I_n$ dans $M_{n,n}(K[X])$ pour exprimer les coefficients a_i de $\chi_M(X)$ en fonction de M et des matrices A_i.

c) En déduire une « preuve matricielle » du théorème de Cayley-Hamilton, affirmant que $\chi_M(M) = 0_n$, avec

$$\chi_M(M) = M^n + a_{n-1}M^{n-1} + \cdots + a_0 I_n.$$

3 Exercices

1. a) Soit $M = \begin{pmatrix} 0_n & N \\ I_n & 0_n \end{pmatrix}$ dans $M_{2n,2n}(K)$. Déterminer $\det M$ en fonction de $\det N$. On pourra consisidérer la matrice $\begin{pmatrix} 0_n & I_n \\ I_n & 0_n \end{pmatrix}$.

b) Soient A, B, C et D quatre matrices de $M_{n,n}(K)$ telles que l'une d'entre elles est nulle. Dans les quatre cas, calculer le déterminant de $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ dans $M_{2n,2n}(K)$.

c) Avec les notations de a), déterminer le polynôme caractéristique de la matrice M en fonction de celui de la matrice N.

2. a) Soient A et B des éléments de $M_{n,n}(\mathbb{R})$ tels que $A+\mathrm{i}B$ est inversible dans $M_{n,n}(\mathbb{C})$. Prouver qu'il existe un réel x tel que $A + xB$ est inversible.

b) Soient M et N des éléments de $M_{n,n}(K)$. Prouver que M et N sont semblables si et seulement si l'ensemble $C(M,N)$ des matrices X de $M_{n,n}(K)$ telles que $XM = NX$ contient une matrice inversible.

c) Déduire de b) et de l'exercice 11 de la fiche 14 que si $K \subset \mathbb{C}$ et si les matrices M et N sont semblables sur $\mathbb{C}$, M et N sont semblables sur K.

d) Application. Prouver qu'une matrice carrée à coefficients réels est diagonalisable sur $\mathbb{R}$ si et seulement si elle est diagonalisable sur $\mathbb{C}$ et si toutes ses valeurs propres sont réelles.

3. Pour tout $n \geqslant 1$ et tout élément $x = (x_i)_{1 \leqslant i \leqslant n}$ de K^n, on note $V(x)$ le déterminant de la matrice $M = (m_{i,j})_{1 \leqslant i,j \leqslant n}$ définie par $m_{i,j} = x_i^{j-1}$.

a) Calculer $V(x_1)$, $V(x_1, x_2)$ et $V(x_1, x_2, x_3)$.

b) On suppose que, pour tout $1 \leqslant k \leqslant n-1$, $V(x_1, \ldots, x_k) \neq 0$.

Prouver que $V(x_1, \ldots, x_n)$ est un polynôme en x_n dont on déterminera le degré et les racines.

En déduire une relation de récurrence entre $V(x_1, \ldots, x_k)$ et $V(x_1, \ldots, x_{k-1})$ puis la valeur de $V(x_1, \ldots, x_n)$.

c) Dans le cas général, donner la valeur de $V(x_1, \ldots, x_n)$.

4. Soit V un espace vectoriel de dimension n sur un corps K et $(a_i)_{1 \leqslant i \leqslant n}$ une base de V. Soit f une forme bilinéaire alternée sur V. On suppose que $a = 0$ est le seul vecteur de V tel que pour tout x dans V, $f(a, x) = 0$.

a) Soit $A = (f(a_i, a_j))_{1 \leqslant i,j \leqslant n}$. Prouver que A est inversible.

b) Soient a et b des vecteurs de V tels que $f(a, b) \neq 0$. On pose $f_a(x) = f(a, x)$ et $f_b(x) = f(b, x)$. Prouver que f_a et f_b sont des formes linéaires non proportionnelles et que l'ensemble des vecteurs x de V tels que

$$(*) \qquad f(a, x) = 0 = f(b, x),$$

est un sous-espace vectoriel de V de dimension $n - 2$.

c) Prouver que V est somme directe du sous-espace engendré par a et b et du sous-espace W des solutions de $(*)$. Prouver que si un vecteur z de W est tel que, pour tout vecteur y de W, $f(z, y) = 0$, alors $z = 0$.

d) En raisonnant par récurrence sur la dimension, prouver que la dimension de V est paire, soit $2p$, et qu'il existe une base de V composée des vecteurs a_i et b_i pour $1 \leqslant i \leqslant p$, telle que, pour $1 \leqslant i, j \leqslant p$, $f(a_i, a_j) = f(b_i, b_j) = 0$, pour $1 \leqslant i \neq j \leqslant n$, $f(a_i, b_j) = f(b_j, a_i) = 0$, et pour $1 \leqslant i \leqslant n$, $f(a_i, b_i) = 1 = -f(b_i, a_i)$.

e) Une matrice carrée $A = (\alpha_{i,j})_{1 \leqslant i,j \leqslant n}$ est alternée si, pour tous $1 \leqslant i, j \leqslant n$,

$$\alpha_{i,i} = 0, \qquad \alpha_{i,j} + \alpha_{j,i} = 0.$$

Prouver que si une matrice alternée est inversible, sa taille $n = 2p$ est paire et il existe une matrice U dans $GL_n(K)$ telle que $U^t A U = \begin{pmatrix} 0_p & I_p \\ -I_p & 0_p \end{pmatrix}$.

f) Prouver que pour tous x, y et z dans K, $\begin{vmatrix} 0 & x & z \\ -x & 0 & y \\ -z & -y & 0 \end{vmatrix} = 0$.

5. Calculer les déterminants

$$D_3 = \begin{vmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{vmatrix}, \qquad D_5 = \begin{vmatrix} 1 & 0 & p & q & 0 \\ 0 & 1 & 0 & p & q \\ 3 & 0 & p & 0 & 0 \\ 0 & 3 & 0 & p & 0 \\ 0 & 0 & 3 & 0 & p \end{vmatrix},$$

et

$$D_7 = \begin{vmatrix} 1 & 0 & a & b & 0 & 0 & 0 \\ 0 & 1 & 0 & a & b & 0 & 0 \\ 0 & 0 & 1 & 0 & a & b & 0 \\ 0 & 0 & 0 & 1 & 0 & a & b \\ 1 & 0 & -2a & -8b & a^2 & 0 & 0 \\ 0 & 1 & 0 & -2a & -8b & a^2 & 0 \\ 0 & 0 & 1 & 0 & -2a & -8b & a^2 \end{vmatrix}.$$

6. Soit $n \geqslant 3$ et $(x_i)_{1\leqslant i\leqslant n}$ et $(y_i)_{1\leqslant i\leqslant n}$ des éléments de K^n. Pour tous $1 \leqslant i, j \leqslant n$, on pose $a_{i,j} = 1 + x_i y_j$. Soit $A = (a_{i,j})_{1\leqslant i,j\leqslant n}$. Montrer que $\det A = 0$.

7. Pour tout $n \geqslant 1$, on note D_n le déterminant $n \times n$ suivant :

$$D_n = \begin{vmatrix} 3 & 2 & 0 & 0 & \cdots & 0 \\ 1 & 3 & 2 & 0 & \cdots & 0 \\ 0 & 1 & 3 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 3 \end{vmatrix}.$$

Pour tout $n \geqslant 1$, établir une relation de récurrence entre D_{n+2}, D_{n+1} et D_n. Déterminer toutes les suites réelles $(u_n)_{n\geqslant 1}$ qui vérifient la même relation de récurrence. En déduire la valeur de D_n pour tout $n \geqslant 1$.

8. Calculer le déterminant 8×8 ci-dessous :

$$D_8 = \begin{vmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 2 \end{vmatrix}.$$

Fiche 16 : Un problème d'écrit

Notations

n est un entier supérieur ou égal à 2.

Dans cet énoncé toutes les matrices sont à coefficients réels.

On désigne par M_n l'algèbre des matrices carrées $n \times n$.

On identifie le vecteur $x = (x_1, \ldots, x_n) = (x_i) \in \mathbb{R}^n$ à la matrice ligne $X = (x_1, \ldots, x_n) = (x_i)$.

Si A est une matrice, on note A^t la matrice transposée de A. On notera ainsi X^t la matrice colonne transposée de la matrice ligne $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Questions préliminaires (étude d'exemples dans le cas $n = 3$)

0.1 Soit $U = (1, 2, 3)$ et $V = (1, 2, 5)$.

a) Calculer $m = UV^t$ et $M = V^tU$.

b) Calculer le rang de la matrice M.

c) Calculer M^2 et exprimer le résultat en fonction de M.

0.2 Soit $M' = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ -4 & -8 & -12 \end{pmatrix}$.

a) Calculer le rang de M'.

b) Déterminer deux matrices U' et V' telles que $M' = (V')^tU'$.

On revient au cas général $n \geqslant 2$.

Première partie

Soit $U = (u_i)$ et $V = (v_i)$ deux éléments non nuls de $\mathbb{R}^n$.

On pose $m = UV^t$ et $M = V^tU$.

1.1 Expliciter le réel m et le terme général de la matrice M en fonction des coefficients des matrices lignes U et V.

1.2 Résoudre le système $MX^t = 0$ en l'inconnue $X = (x_i)$ dans $\mathbb{R}^n$.

En déduire le rang de M.

1.3 On note E l'élément de M_n dont tous les coefficients sont nuls sauf celui qui est situé sur la première ligne et la première colonne, qui vaut 1.

a) Soient R et S deux éléments de M_n. Exprimer la matrice RES à l'aide d'une colonne de R et d'une ligne de S.

b) Prouver qu'une matrice M' de M_n est de rang 1 si et seulement s'il existe des matrices inversibles R et S de M_n telles que $M' = RES$.

c) Déduire de a) et b) que si un élément M' de M_n est de rang 1 alors il existe deux matrices lignes non nulles U' et V' de $\mathbb{R}^n$ telles que $M' = (V')^t U'$.

Deuxième partie

On garde les notations de la première partie : U et V sont deux matrices lignes non nulles de $\mathbb{R}^n$, $m = UV^t$ et $M = V^t U$. Soit $N = I_n + M$.

2.1 Calculer M^2 en fonction de m et M. En déduire pour tout entier $p \geqslant 1$ la valeur de M^p en fonction de M, m et p.

On se propose de calculer N^p par trois méthodes différentes.

2.2 En appliquant la formule du binôme à $(I_n + M)^p$ exprimer N^p en fonction de I_n, M, m et p.

[Le résultat ne devra pas faire intervenir le symbole $\sum$, pour cela on distinguera les cas $m = 0$ et $m \neq 0$.]

Dans les deux questions suivantes on ne traite que le cas $m \neq 0$.

2.3 Soient les matrices $P = \dfrac{1}{m} M$ et $Q = I_n - P$.

a) Calculer P^2, PQ, QP et Q^2.

b) Exprimer N en fonction de P, Q et m.

c) En déduire la valeur de N^p en fonction de P, Q, m et p.

2.4 a) Calculer N^2 en fonction de N, I_n et m.

b) En déduire un polynôme $f(X)$ de $\mathbb{R}[X]$ du second degré tel que $f(N) = 0$.

c) Calculer le reste de la division euclidienne de X^p par $f(X)$.

d) En déduire une expression de N^p.

Troisième partie

Soit A un élément de M_n. On dira que A possède la propriété $\mathcal{P}$ si le rang de $A - I_n$ vaut 1.

3.0 Prouver que A possède la propriété $\mathcal{P}$ si et seulement s'il existe deux matrices lignes non nulles U et V de $\mathbb{R}^n$ telles que $A = I_n + V^t U$.

3.1 Soit $N = I_n + V^t U$ une matrice possédant la propriété $\mathcal{P}$ et $m = UV^t$.

a) Prouver que N est inversible si et seulement si $m \neq -1$.

b) Prouver dans ce cas que N^{-1} possède la propriété $\mathcal{P}$.

Fiche 17 : Un problème de CAPES blanc

Préambule

Dans tout le sujet, n désigne un entier fixé $\geqslant 2$.

On utilise les notations et observations suivantes.

- $[[1,n]]$ désigne l'ensemble $\{1,2,\ldots,n\}$.
- $\mathcal{M}$ désigne l'espace vectoriel des matrices carrées d'ordre n à coefficients réels.
- $\mathrm{Sp}(M)$ désigne l'ensemble des valeurs propres (réelles ou complexes) de $M \in \mathcal{M}$.
- $\mathcal{C}$ désigne l'espace vectoriel des matrices unicolonnes à n lignes à coefficients réels.
- $\mathcal{L}$ désigne l'espace vectoriel des matrices unilignes à n colonnes à coefficients réels.
- I désigne la matrice identité d'ordre n. Les colonnes de I, notées $E_1,\ldots,E_n$, forment la base canonique de $\mathcal{C}$ et ses lignes ${}^tE_1,\ldots,{}^tE_n$ forment la base canonique de $\mathcal{L}$.
- Si $M \in \mathcal{M}$, φ_M désigne l'application $\varphi_M : X \in \mathcal{C} \mapsto MX \in \mathcal{C}$ qui est linéaire de matrice M par rapport à la base canonique de $\mathcal{C}$.
- Si $M \in \mathcal{M}$, ψ_M désigne l'application $\psi_M : X \in \mathcal{L} \mapsto XM \in \mathcal{L}$ qui est linéaire de matrice ${}^tM$ par rapport à la base canonique de $\mathcal{L}$.
- U désigne le vecteur colonne $U = {}^t(1,\ldots,1)$, c'est-à-dire l'élément de $\mathcal{C}$ dont toutes les composantes sont égales à 1.
- Pour toute matrice $M = (m_{i,j})$ (carrée ou rectangulaire), $M \geqslant 0$ désigne le fait que $m_{i,j} \geqslant 0$ pour tous i et j, et $\|M\|$ désigne le nombre $\max|m_{i,j}|$, qui est le maximum de la valeur absolue des coefficients de M.
- Un vecteur ligne $V = (v_1,\ldots,v_n) \in \mathcal{L}$ est dit stochastique si $\displaystyle\sum_{j=1}^{n} v_j = 1$ et si $v_j \geqslant 0$ pour tout $j \in [[1,n]]$.
- Les espaces vectoriels $\mathcal{M}$, $\mathcal{C}$ et $\mathcal{L}$ sont considérés comme des espaces normés pour la norme $\|\cdot\|$.
- On note $\mathcal{M}^+ = \{A \in \mathcal{M} \mid A \geqslant 0\}$, $\mathcal{C}^+ = \{X \in \mathcal{C} \mid X \geqslant 0\}$ et $\mathcal{L}^+ = \{X \in \mathcal{L} \mid X \geqslant 0\}$.

Nota : La partie IV et le début de la partie V sont indépendants de la partie III.

I Matrices stochastiques

Soit $\mathcal{S}$ l'ensemble des matrices $A = (a_{i,j})$ de $\mathcal{M}$ vérifiant les conditions (1) et (2) suivantes :

$$(1)\quad A \geqslant 0. \qquad\qquad (2)\quad \forall i \in [[1,n]],\ \sum_{j=1}^{n} a_{i,j} = 1.$$

Les éléments de $\mathcal{S}$ sont des *matrices stochastiques*.

1. On considère les propriétés $(1')$, $(1'')$ et $(2')$ suivantes :

$$(1') \quad \varphi_A(\mathcal{C}^+) \subset \mathcal{C}^+. \qquad (1'') \quad \psi_A(\mathcal{L}^+) \subset \mathcal{L}^+. \qquad (2') \quad \varphi_A(U) = AU = U.$$

Montrer les équivalences : $(1) \iff (1') \iff (1'')$.
Montrer l'équivalence : $(2) \iff (2')$.

2. Montrer que pour toutes matrices A et B de $\mathcal{S}$ et tout réel $t \in [0,1]$, la matrice $tA + (1-t)B$ est dans $\mathcal{S}$ (c'est-à-dire que $\mathcal{S}$ est une partie convexe de $\mathcal{M}$).

3. Montrer que pour toute suite $(A_k)_{k\in\mathbb{N}}$ de matrices de $\mathcal{S}$ convergeant vers B dans $\mathcal{M}$, la matrice B est dans $\mathcal{S}$ (c'est-à-dire que $\mathcal{S}$ est une partie fermée de $\mathcal{M}$). Montrer que la partie $\mathcal{S}$ est compacte.

4. Soit $V = (v_1, \ldots, v_n)$ un vecteur ligne stochastique. Montrer que pour toute matrice A de $\mathcal{S}$, le vecteur ligne VA est stochastique.

5. Montrer que $\mathcal{S}$ est une partie de $\mathcal{M}$, stable pour la multiplication matricielle.

6.a. Soit $A \in \mathcal{S}$ telle que $A^tA = I$. Montrer que chaque ligne et chaque colonne de A a tous ses coefficients nuls sauf un qui vaut 1. Préciser comment agit φ_A sur la base canonique $E_1, \ldots, E_n$ de $\mathcal{C}$.

6.b. Que peut-on dire de $G = \{A \in \mathcal{S} \mid A\,{}^tA = I\}$?

II Éléments propres des matrices stochastiques

Soit $A = (a_{i,j})$ une matrice de $\mathcal{S}$.

1. Montrer que 1 est valeur propre de A.

2. Montrer que $\mathrm{Sp}(A)$ est inclus dans $\{z \in \mathbb{C} \mid |z| \leqslant 1\}$.

[Pour traiter cette question, il pourra être utile de considérer pour $\lambda \in \mathrm{Sp}(A)$, un vecteur colonne $Y = {}^t(y_1, \ldots, y_n) \neq 0$ tel que $AY = \lambda Y$ et $\mu = \max\{|y_i|, i \in [[1,n]]\}$.]

3. Soit $\lambda \in \mathrm{Sp}(A)$. Justifier le fait qu'il existe $p \in [[1,n]]$ tel que $|\lambda - a_{p,p}| \leqslant 1 - a_{p,p}$.

4. On suppose ici que $a_{i,i} > 1/2$ pour tout $i \in [[1,n]]$. Montrer que A est inversible.

5. On suppose ici que $a_{i,i} > 0$ pour tout $i \in [[1,n]]$.

5.a. Montrer qu'il existe un réel $\alpha \in]0,1]$ tel que $\mathrm{Sp}(A)$ est inclus dans l'ensemble $D_\alpha = \{z \in \mathbb{C} \,|\, |z-\alpha| \leqslant 1-\alpha\}$.
Interpréter géométriquement le résultat précédent et faire la figure correspondante.

5.b. Que peut-on dire du module des valeurs propres de A différentes de 1 ?

6. On suppose ici que $a_{i,j} > 0$ pour tous $i,j \in [[1,n]]$. Montrer que les valeurs propres (complexes) de A autres que 1, sont de module strictement inférieur à 1, et préciser le rang de la matrice $A - I$.

7. Soit $\lambda \in \mathrm{Sp}_{\mathbb{C}}(A)$ avec $|\lambda| = 1$.

7.a. Soit un vecteur colonne $Y = {}^t(y_1, \ldots, y_n) \neq 0$ tel que $AY = \lambda Y$. On note

$$\mu = \max\{|y_i| \,|\, i \in [[1,n]]\}, \qquad K = \{k \in [[1,n]] \,|\, |y_k| = \mu\}.$$

Construire une application f de K dans K telle que pour tout $k \in K$, $y_{f(k)} = \lambda y_k$.

7.b. En déduire qu'il existe un entier p compris entre 1 et n tel que $\lambda^p = 1$.

III Convergence

Soit A dans $\mathcal{S}$. On s'intéresse dans cette partie à la convergence éventuelle de la suite $(A^k)_{k\in\mathbb{N}}$. Pour tout $k \in \mathbb{N}$, on note $a_{i,j}^{(k)}$ le coefficient d'indice (i,j) de A^k pour tous $i,j \in [[1,n]]$.

1. La suite $(A^k)_{k\in\mathbb{N}}$ peut-elle converger si A possède une valeur propre λ de module 1 et différente de 1 ?

2. On suppose que la suite $(A^k)_{k\in\mathbb{N}}$ converge vers B dans $\mathcal{M}$.

2.a. Montrer que $B \in \mathcal{S}$ et $B^2 = B$.

2.b. Montrer qu'on a aussi $BA = AB = B$. Ces égalités traduisent des propriétés remarquables des colonnes et des lignes de B, lesquelles ?

2.c. Lorsque $n = 2$, donner toutes les possibilités de matrices B.

3. On suppose ici que A est diagonalisable dans l'espace des matrices à coefficients complexes et qu'elle ne possède pas de valeur propre de module 1 autre que 1. Montrer que la suite $(A^k)_{k\in\mathbb{N}}$ converge.

4. On suppose ici que $\varepsilon = \min\{a_{i,j} \mid i,j \in [[1,n]]\} > 0$. Pour tout k dans $\mathbb{N}$, et tout j dans $[[1,n]]$, on note :

$$\alpha_j^{(k)} = \min\{a_{i,j}^{(k)} \mid i \in [[1,n]]\}, \quad \beta_j^{(k)} = \max\{a_{i,j}^{(k)} \mid i \in [[1,n]]\}, \quad \delta_j^{(k)} = \beta_j^{(k)} - \alpha_j^{(k)}.$$

4.a. Montrer que pour tout k dans $\mathbb{N}$, et tout j dans $[[1,n]]$, on a :

$$\alpha_j^{(k)} \leqslant \alpha_j^{(k+1)} \leqslant \beta_j^{(k+1)} \leqslant \beta_j^{(k)}, \qquad \delta_j^{(k+1)} \leqslant (1-2\varepsilon)\delta_j^{(k)}.$$

4.b. En déduire que $(A^k)_{k\in\mathbb{N}}$ converge. Si $B = \lim_{k\to\infty} A^k$, comparer les lignes de B.

5. Dire dans chacun des cas suivants, en faisant le minimum possible de calculs, si la suite $(A^k)_{k\in\mathbb{N}}$ converge ou non :

$$(i)\ A = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}\ ; \qquad (ii)\ A = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}\ ;$$

$$(iii)\ A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}\ ; \qquad (iv)\ A = \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 1/6 & 1/2 & 1/3 \\ 1/3 & 1/6 & 1/2 \end{pmatrix}.$$

IV Chaînes de Markov

Soit une suite de variables aléatoires $(X_k)_{k\in\mathbb{N}}$ définies sur un même espace de probabilité $(\Omega, \mathcal{T}, P)$ à valeurs dans $[[1,n]]$. On dit que la suite $(X_k)_{k\in\mathbb{N}}$ est une *chaîne de Markov* (à n états) si les propriétés suivantes sont réalisées :

(CM1) Il existe une matrice $M = (p_{i,j}) \in \mathcal{M}$, appelée *matrice de transition* telle que pour tout entier $k > 0$, tous $i, j \in [[1,n]]$, on ait

$$P(X_k = j \mid X_{k-1} = i) = p_{i,j}.$$

Autrement dit, l'état X_k dépend de l'état X_{k-1} de manière invariable.

(CM2) Pour tout entier $k > 1$ et tous $i, j, i_{k-2}, \ldots, i_0 \in [[1, n]]$, on a

$$P(X_k = j \mid X_{k-1} = i, X_{k-2} = i_{k-2}, \ldots, X_0 = i_0) = P(X_k = j \mid X_{k-1} = i).$$

Autrement dit, l'état X_k ne dépend que de l'état X_{k-1}.

On considère une telle chaîne de Markov $(X_k)_{k\in\mathbb{N}}$. On note $M = (p_{i,j}) \in \mathcal{M}$ sa matrice de transition. Pour tout entier k, on note $P(X_k = i) = q_i^{(k)}$ pour tout $i \in [[1, n]]$ et Q_k le vecteur ligne $Q_k = (q_1^{(k)}, \ldots, q_n^{(k)})$ (la loi de X_k). La loi de X_0 donnée par Q_0 que l'on notera simplement $Q_0 = (q_1, \ldots, q_n)$ est appelée *la loi initiale* de la chaîne.

1. Montrer que M est dans $\mathcal{S}$ (M est stochastique) et que pour tout $k \geqslant 0$, Q_k est un vecteur ligne stochastique, c'est-à-dire $\displaystyle\sum_{i=1}^{n} q_i^{(k)} = 1$.

2. Pour tout entier $k \geqslant 0$, montrer que $Q_{k+1} = Q_k M$. En déduire l'expression de la loi de X_k donnée par Q_k à l'aide de M et de la loi initiale donnée par Q_0.

3. Montrer qu'il existe une loi initiale telle que $Q_0 = Q_0 M$. Une telle loi est dite *stationnaire*.

[On sera amené à montrer que si un vecteur $V = (v_1, \ldots, v_n)$ vérifie $VM = V$ alors $V' = (|v_1|, \ldots, |v_n|)$ vérifie aussi $V'M = V'$.]

4. On suppose que la loi initiale est stationnaire, c'est-à-dire $Q_0 = Q_0 M$. Montrer qu'il existe une matrice $N = (q_{i,j})$ de $\mathcal{S}$ telle que $q_{i,j} = P(X_k = j \mid X_{k+1} = i)$ pour tout k entier.

V Un exemple de chaîne de Markov

Une urne contient initialement deux boules rouges et deux boules bleues. On décide de faire une succession de tirages avec la règle suivante : pour tout $k \geqslant 0$, la boule tirée au tirage k est laissée de côté au tirage $k+1$ et n'est remise dans l'urne que pour effectuer le tirage $k+2$. On considère la suite de variables aléatoires $(X_k)_{k\in\mathbb{N}}$ où X_k est le nombre de boules rouges dans l'urne après le tirage k. Tous les tirages sont équiprobables sauf le tirage initial de numéro $k = 0$, pour lequel on décide que $P(X_0 = 1) = p$ pour un certain $p \in]0, 1[$ fixé.

0. Déterminer la loi de X_1, sa moyenne $E(X_1)$ et sa variance $\mathrm{var}(X_1)$.

1. Justifier que la suite $(X_k)_{k\in\mathbb{N}}$ est une chaîne de Markov dont on précisera la matrice de transition M et la loi initiale.

2. Montrer qu'il y a une seule loi stationnaire que l'on précisera.

3. En s'aidant des parties 1. Matrices stochastiques à 3. Convergence, montrer que la suite des lois des X_k converge vers la loi stationnaire.

On note ici $R_k = [X_k = 1]$ l'événement "Au tirage k, une boule rouge a été tirée". On considère la variable aléatoire N égale au numéro d'ordre du premier tirage d'une boule rouge (autrement dit N est égal au plus petit entier k tel que R_k est réalisé).

4. Exprimer pour $k \geqslant 0$ l'événement $[N = k]$ à l'aide de $R_0, \ldots, R_k$.

5. En déduire la loi de N, sa moyenne $E(N)$ et sa variance $\mathrm{var}(N)$.

CAPES de Mathématiques **Université Joseph Fourier**
Préparation à l'écrit **Année 2007-2008**
Algèbre et probabilités

Corrigé du problème de CAPES blanc

I Matrices stochastiques

1. Montrons que (1) et (1′) sont équivalentes.

Soient $C_1, \dots, C_n$ les colonnes de A. Par définition, $AE_i = C_i$ pour tout $i \in [[1, n]]$. Donc $A \geqslant 0$ si et seulement si $\varphi_A(E_i) = AE_i \geqslant 0$ pour tout $i \in [[1, n]]$. Comme $\mathbb{C}^+ = \{t_1E_1 + \cdots + t_nE_n, t_i \geqslant 0,\ \forall i \in [[1, n]]\}$, cette condition équivaut à $\varphi_A(\mathbb{C}^+) \subset \mathbb{C}^+$.

On raisonne de même sur les lignes pour obtenir la condition équivalente $\psi_A(\mathcal{L}^+) \subset \mathcal{L}^+$.

Montrons que (2) et (2′) sont équivalentes.

Par définition, $\sum_{j=1}^{n} a_{i,j} = (AU)_i$ donc la condition (2) revient à imposer que toutes les composantes de AU sont égales à 1, donc que par définition, $AU = U$.

2. Soit A et B dans $\mathfrak{S} \subset \mathcal{M}^+$ et $t \in [0, 1]$. La matrice $tA + (1-t)B$ est dans $\mathcal{M}^+$ et $(tA + (1-t)B)U = tAU + (1-t)BU = tU + (1-t)U = U$. La matrice $tA + (1-t)B$ vérifie (1) et (2′) donc elle est stochastique.

3. Les coefficients de chaque A_k sont positifs ou nuls donc leurs limites, qui sont les coefficients de B, aussi. Comme l'application $A \in \mathcal{M} \mapsto AU \in \mathbb{C}$ est continue, la suite de terme général $A_kU = U$ converge vers BU, d'où $BU = U$; donc B vérifie (1) et (2′), ce qui montre que $B \in \mathfrak{S}$. Ainsi, la partie $\mathfrak{S}$ est fermée dans $\mathcal{M}$.

De plus, pour toute matrice $A = (a_{i,j}) \in \mathcal{M}$, $0 \leqslant a_{i,j} \leqslant \sum_{k=1}^{n} a_{i,k} = 1$ pour tous i et j. La partie $\mathfrak{S}$ est donc bornée, et comme $\mathfrak{S}$ est déjà fermée, $\mathfrak{S}$ est compacte.

4. Observons que le fait que $V = (v_1, \dots, v_n) \in \mathcal{L}$ soit stochastique revient à dire que $V \in \mathcal{L}^+$ et que $\sum_{j=1}^{n} v_j = VU = 1$. Soit $A \in \mathfrak{S}$. Alors $VA = \psi_A(V) \in \mathcal{L}^+$ par la propriété (1′) et $(VA)U = V(AU) = VU$ par la propriété (2′), donc VA est stochastique.

5. Soient A et B dans $\mathfrak{S}$. Alors $\varphi_{AB} = \varphi_A \circ \varphi_B$, donc $\varphi_{AB}(\mathbb{C}^+) = \varphi_A(\varphi_B(\mathbb{C}^+)) \subset \varphi_A(\mathbb{C}^+) \subset \mathbb{C}^+$. De plus, $ABU = AU = U$ par les propriétés (1′) et (2′) satisfaites par A et B. Donc $AB \in \mathfrak{S}$.

6.a. Soit C_i la colonne numéro i de ${}^tA$. Alors $A\,{}^tA = ({}^tC_iC_j) = I$, en particulier ${}^tC_iC_i = \sum_{j=1}^{n} a_{i,j}^2 = 1$. D'après (2), $\sum_{j=1}^{n} a_{i,j} = 1$. Comme $0 \leqslant a_{i,j} \leqslant 1$ pour tous i et j, $\sum_{j=1}^{n} a_{i,j}^2 \leqslant \sum_{j=1}^{n} a_{i,j} = 1$ et l'inégalité est stricte sauf si tous les $a_{i,j}$ quand j varie de 1 à n valent 0 ou 1. On est donc dans ce cas-là et, comme la somme sur j des $a_{i,j}$ vaut 1, exactement un des $a_{i,j}$ vaut 1.

Comme $A\,{}^tA = {}^tAA = I$, $\sum_{i=1}^{n} a_{i,j}^2 = 1$. Sachant maintenant que chacun des $a_{i,j}$ vaut 0 ou 1, on a aussi que pour j fixé, tous les $a_{i,j}$ quand i varie de 1 à n sont nuls sauf l'un d'eux qui vaut 1.

Ceci revient à dire que la colonne C_j est un élément $E_{\sigma(j)}$ de la base canonique, donc $C_j = AE_{\sigma(j)} = E_{\sigma(j)}$ où σ est une application de $[[1,n]]$ dans lui même. La condition précédente sur les lignes de A indique que σ est injective et donc σ est une permutation. L'application φ_A permute donc les éléments de la base canonique. Une matrice associée à une permutation de la base canonique est appelée une *matrice de permutation*.

6.b. On vient de justifier que $G = \{A \in \mathfrak{S} \mid A\,{}^tA = I\}$ est le groupe des matrices de permutations associées aux permutations de la base canonique de $\mathbb{C}$, isomorphe au groupe symétrique $\mathfrak{S}_n$.

II Éléments propres des matrices stochastiques

1. D'après $(1')$, $\varphi_A(U) = AU = U$ donc U est vecteur propre de A pour la valeur propre 1.

2. Soit $\lambda \in \mathrm{Sp}(A)$ une valeur propre (complexe) de A. On reprend les notations de l'indication de l'énoncé et on choisit un indice p pour lequel $|y_p| = \mu$. Comme $(\lambda Y)_p = (AY)_p$,

$$(*) \qquad \lambda y_p = \sum_{j=1}^{n} a_{p,j} y_j.$$

Par l'inégalité triangulaire, $(*)$ implique

$$(**) \qquad |\lambda|\mu = |\lambda y_p| \leqslant \sum_{j=1}^{n} a_{p,j}|y_j| \leqslant \mu \sum_{j=1}^{n} a_{p,j} = \mu,$$

soit $\lambda|\mu \leqslant \mu$. Comme $Y \neq 0$, $\mu \neq 0$ donc $|\lambda| \leqslant 1$.

3. Pour le même indice p que dans la question précédente, on peut écrire $(*)$ sous la forme $(\lambda - a_{p,p})y_p = \sum_{j\neq p} a_{p,j} y_j$, donc $|\lambda - a_{p,p}|\mu \leqslant \sum_{j\neq p} a_{p,j}\mu = (1 - a_{p,p})\mu$. En simplifiant par $\mu > 0$, on en déduit l'inégalité cherchée.

4. Pour tout $\lambda \in \mathrm{Sp}(A)$, par l'inégalité triangulaire, $|\lambda| \geqslant a_{p,p} - |\lambda - a_{p,p}| \geqslant 2a_{p,p} - 1 > 0$ donc $\lambda \neq 0$. Donc 0 n'est pas valeur propre de A et A est inversible.

5.

5.a. Montrons que si $0 < \alpha \leqslant \beta \leqslant 1$ alors $D_\beta \subset D_\alpha$. (C'est évident géométriquement : les disques D_α sont tangents en 1 à l'axe vertical passant par 1 donc ils sont emboîtés les uns dans les autres et le rayon de D_α vaut α.)

En effet $z \in D_\beta$ implique que $|z-\alpha| = |z-\beta+\beta-\alpha| \leqslant |z-\beta|+|\beta-\alpha| \leqslant 1-\beta+\beta-\alpha = 1-\alpha$, donc $z \in D_\alpha$.

Posons $\alpha = \min\{a_{i,i} \,|\, i \in [[1,n]]\}$. Par hypothèse, $\alpha > 0$ et par la question 3., $\mathrm{Sp}(A) \subset D_\alpha$.

[Dessin : on savait depuis la question 2. que $\mathrm{Sp}(A)$ était une partie du disque unité, en fait $\mathrm{Sp}(A)$ est même une partie d'un disque D_α de centre α et dont le bord passe par 1, inclus strictement dans le disque unité.]

5.b. L'intersection de D_α avec le bord $\{z \in \mathbb{C} \,|\, |z| = 1\}$ du disque unité est réduite au point $\{1\}$.

En effet, tout point z dans cette intersection vérifie $1 = |z| \leqslant |z-\alpha|+\alpha \leqslant (1-\alpha)+\alpha = 1$, donc les inégalités sont des égalités. En particulier $|z| = |z-\alpha| + \alpha$ donc $|z-0| = |z-\alpha| + |\alpha - 0|$, ce qui signifie que le point α appartient au segment $[0,z] \subset \mathbb{C}$. Le seul point z du cercle unité vérifiant cette propriété est $z = 1$, donc on a terminé.

6. D'après 2. et 5.b., le module de chaque valeur propre de A autre que 1 est strictement inférieur à 1.

Montrons que le rang de $A - I$ est exactement $n-1$. En effet, sinon, il existerait un vecteur propre $Y = {}^t(y_1, \ldots, y_n)$ pour la valeur propre 1 linéairement indépendant de U. En remplaçant Y par $Y - y_n U$, on peut supposer que $y_n = 0$. L'inégalité $(**)$ du 2. avec $\lambda = 1$ s'écrit alors $\mu \leqslant \displaystyle\sum_{j=1}^{n-1} a_{p,j}|y_j| \leqslant \mu \sum_{j=1}^{n-1} a_{p,j} = (1-a_{p,n})\mu < \mu$, puisque $a_{p,n} > 0$. C'est absurde.

7.

7.a. Soit p un indice tel que $|y_p| = \mu$. L'inégalité $(**)$ du 2. avec $|\lambda| = 1$ s'écrit $\mu \leqslant \displaystyle\sum_{j=1}^{n} a_{p,j}|y_j| \leqslant \mu$. La deuxième inégalité sera stricte si $|y_j| < \mu$ pour au moins un indice j tel que $a_{p,j} \neq 0$. On utilise alors l'évidence géométrique suivante : si le barycentre G d'un système pondéré de points du cercle de centre 0 et de rayon μ est sur ce même cercle, alors tous les points du système de poids strictement positifs sont égaux à G. En voyant les y_j comme les affixes de points pondérés par les coefficients $a_{p,j}$ pour $1 \leqslant j \leqslant n$, on voit que $y_j = \lambda y_p$ pour tout j tel que $a_{p,j} \neq 0$.

À tout indice $p \in K$, on a associé un indice j tel que $y_j = \lambda y_p$. Donc $|y_j| = \mu$ et $j \in K$. Choisissons un quelconque de ces indices j et notons-le $f(p)$, alors f envoie bien K sur

K.

7.b. Soit $k \in K$ quelconque. Comme $\{k, f(k), f^2(k), \ldots, f^n(k)\}$ est une partie de K et que le cardinal de K vaut au plus n, au moins deux éléments de cette orbite sont égaux, par exemple $f^s(k) = f^t(k)$ avec $0 \leqslant s < t \leqslant n$. Soit $z = y_{f^t(k)}$ et $p = t - s$, donc $1 \leqslant p \leqslant n$ et $z = \lambda^p y_{f^s(k)} = \lambda^p z$.

Comme $Y \neq 0$, $\mu \neq 0$ donc $z \neq 0$ et on a montré que $\lambda^p = 1$.

III Convergence

1. Supposons que A possède une valeur propre $\lambda \neq 1$ telle que $|\lambda| = 1$. D'après II-7.b, il existe un entier $p > 1$ tel que $\lambda^p = 1$. Soit Y un vecteur propre de A pour la valeur propre λ. Alors $A^{kp}Y = \lambda^{kp}Y = Y$ et $A^{kp+1}Y = \lambda^{kp+1}Y = \lambda Y$. Or $\lambda \neq 1$ et $Y \neq 0$ donc $\lambda Y \neq Y$. La suite $(A^kY)_{k\in\mathbb{N}}$ possède au moins deux valeurs d'adhérence distinctes donc elle diverge. Par continuité de l'application $M \in \mathcal{M} \mapsto MY \in \mathbb{C}$, la suite $(A_k)_{k\in\mathbb{N}}$ diverge.

2.a. D'après I-3, $\mathfrak{S}$ est fermé. Comme $A^k \in \mathfrak{S}$ pour tout k, $B \in \mathfrak{S}$.

La suite $(A^{2k})_{k\in\mathbb{N}}$ est une suite extraite de $(A^k)_{k\in\mathbb{N}}$, donc elle converge aussi vers B. Par continuité de l'application $M \times N \in \mathcal{M} \times \mathcal{M} \mapsto MN \in \mathcal{M}$, cette suite converge vers B^2. Donc $B^2 = B$, ce qui signifie que B est la matrice d'un projecteur.

2.b. On a $BA = (\lim_{k\to\infty} A^k)A = \lim_{k\to\infty} A^{k+1} = B$. On procède de même pour AB.

Comme $BA = AB = B$, les colonnes et les lignes de B sont respectivement vecteurs propres de φ_A et de ψ_A pour la valeur propre 1.

2.c. Si B est inversible et $B^2 = B$ alors $B = I$. (Comme $BA = B$, $A = I$.)

Lorsque $n = 2$ une matrice stochastique non inversible B s'écrit $\begin{pmatrix} a & 1-a \\ a & 1-a \end{pmatrix}$ qui vérifie automatiquement $B^2 = B$. (C'était prévisible car alors le noyau de φ_B et la droite propre $\mathbb{R}U$ pour la valeur propre 1 sont en somme directe donc φ_B est un projecteur.)

3. On dispose donc d'une matrice inversible M et d'une matrice diagonale D à coefficients complexes telles que $A = MDM^{-1}$. D'après ce qui précède, D est la diagonale des (λ_i) et $\lambda_i = 1$ ou $|\lambda_i| < 1$. Donc $\lambda_i^k = 1$ ou $\lambda_i^k \to 0$, quand k tend vers l'infini, en tous les cas $(D^k)_{k\in\mathbb{N}}$ converge et $(A^k)_{k\in\mathbb{N}}$ aussi.

4.

4.a. Les inégalités demandées sont vraies même si $\varepsilon = 0$. En ce qui concerne les coefficients α, pour tous i et j,

$$a_{i,j}^{(k+1)} = \sum_{\ell=1}^{n} a_{i,\ell} a_{\ell,j}^{(k)} \geqslant \sum_{\ell=1}^{n} a_{i,\ell} \alpha_j^{(k)} = \alpha_j^{(k)}.$$

En considérant l'infimum sur i, pour j fixé, on obtient $\alpha_j^{(k+1)} \geqslant \alpha_j^{(k)}$. La preuve pour les

coefficients β est similaire.

En ce qui concerne les coefficients δ, on remarque d'abord que pour tout j,

$$\delta_j^{(k)} = \max\left\{a_{i,j}^{(k)} - a_{i',j}^{(k)} \,|\, i \in [[1,n]],\, i' \in [[1,n]]\right\}.$$

Pour tous i, i' et j,

$$a_{i,j}^{(k)} - a_{i',j}^{(k+1)} = \sum_{\ell=1}^{n} a_{i',\ell}(a_{i,j}^{(k)} - a_{\ell,j}^{(k)}).$$

On majore la dernière parenthèse par $\delta_j^{(k)}$ si $\ell \neq i$ et par 0 si $\ell = i$. Il vient

$$a_{i,j}^{(k)} - a_{i',j}^{(k+1)} \leqslant \delta_j^{(k)} \sum_{\ell \neq j} a_{i',\ell} = \delta_j^{(k)}(1 - a_{i',i}) \leqslant (1-\varepsilon)\delta_j^{(k)}.$$

Si on choisit les indices i et i' tels que $a_{i,j}^{(k)}$ réalise $\beta_j^{(k)}$ et $a_{i',j}^{(k+1)}$ réalise $\beta_j^{(k+1)}$, on obtient

$$\beta_j^{(k)} - \beta_j^{(k+1)} \leqslant (1-\varepsilon)\delta_j^{(k)}.$$

De même, pour tous i, i' et j,

$$a_{i,j}^{(k+1)} - a_{i',j}^{(k)} = \sum_{\ell=1}^{n} a_{i,\ell}(a_{\ell,j}^{(k)} - a_{i',j}^{(k)}) \leqslant \delta_j^{(k)} \sum_{\ell \neq i'} a_{i,\ell} \leqslant \delta_j^{(k)}(1-\varepsilon).$$

Si on choisit les indices i et i' tels que $a_{i,j}^{(k+1)}$ réalise $\alpha_j^{(k+1)}$ et $a_{i',j}^{(k)}$ réalise $\alpha_j^{(k)}$, on obtient

$$\alpha_j^{(k+1)} - \alpha_j^{(k)} \leqslant (1-\varepsilon)\delta_j^{(k)}.$$

En sommant ces deux inégalités, il vient $\delta_j^{(k+1)} - \delta_j^{(k)} \leqslant 2(1-\varepsilon)\delta_j^{(k)}$, ce qui est équivalent à l'inégalité demandée.

4.b. Désormais, $\varepsilon > 0$. De plus la somme d'une ligne de A vaut 1 et est au moins égale à $n\varepsilon$ donc $\varepsilon \leqslant 1/n < 1$ car $n \geqslant 2$. Si on note $c = 1 - 2\varepsilon$, $0 \leqslant c < 1$ et $\delta_j^{(k)} \leqslant c^k \delta_j^{(0)} = c^k$, donc $\delta_j^{(k)}$ tend vers 0 lorsque $k \to \infty$. Les suites $\alpha_j^{(k)}$ et $\beta_j^{(k)}$ sont adjacentes, donc elles convergent vers une limite commune b_j.

Pour tous i et j, $\alpha_j^{(k)} \leqslant a_{i,j}^{(k)} \leqslant \beta_j^{(k)}$, donc $a_{i,j}^{(k)}$ tend vers b_j.

Finalement, $(A^k)_{k\in\mathbb{N}}$ converge vers une matrice stochastique B dont toutes les lignes sont égales à $(b_1, \ldots, b_n)$ (qui est vecteur propre de ψ_A pour la valeur propre 1 d'après III-2.b).

5.

(i) Les valeurs propres sont 1 valeur propre simple et $\frac{1}{2}$ valeur propre double. Comme A est triangulaire, A est diagonalisable donc la question 3. montre la convergence. Si B

désigne la limite, la question 2.b. montre que les lignes de B sont des vecteurs propres de ψ_A pour la valeur propre 1. Le vecteur E_3 est tel, donc $B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

(ii) Le polynôme caractéristique est $\chi_A(X) = (X-1)(X^2 + X/2 - 1/4)$ donc les valeurs propres sont distinctes, A est diagonalisable et A^k converge vers B d'après III-1. Le vecteur $(1,1,1)$ est un vecteur propre de ψ_A pour la valeur propre 1 donc $B = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$

(iii) A est la matrice du cycle (123) donc $A^3 = I$ mais $A \neq I$. La suite $(A^k)_{k\in\mathbb{N}}$ est périodique de période 3 et non constante donc elle diverge. Comme la somme des colonnes de A vaut 1, le vecteur $(1,1,1)$ est un vecteur propre de ψ_A pour la valeur propre 1 donc la limite est la même que dans le cas (ii).

IV Chaînes de Markov

1. Pour tout i, $\sum_{j=1}^{n} p_{i,j} = \sum_{j=1}^{n} P(X_k = j \,|\, X_{k-1} = i)$. D'après la formule des causes totales, la dernière somme vaut $P(\Omega \,|\, X_{k-1} = i) = 1$ donc $M \in \mathfrak{S}$.

D'autre part $\sum_{i=1}^{n} q_i^{(k)} = \sum_{i=1}^{n} P(X_k = i) = 1$ car la variable aléatoire X_k prend ses valeurs dans $[[1,n]]$, donc Q_k est stochastique.

2. On décompose l'événement $[(X_{k+1} = j]$ selon les valeurs de X_k. D'après la formule des causes totales et la formule de Bayes,

$$P(X_{k+1} = j) = \sum_{j=1}^{n} P(X_{k+1} = j,\, X_k = i) = \sum_{j=1}^{n} P(X_k = i) P(X_{k+1} = j | X_k = i),$$

donc

$$q_j^{(k+1)} = \sum_{i=1}^{n} q_i^{(k)} p_{i,j},$$

ce qui donne bien l'égalité entre matrices $Q_{k+1} = Q_k M$. Donc la loi de X_k est décrite par le vecteur $Q_k = Q_0 M^k$.

3. La matrice de la transformation ψ_M est ${}^tM$ qui a le même polynôme caractéristique que M. Comme 1 est valeur propre de M, elle l'est aussi de ${}^tM$, et donc il existe un vecteur propre V de ψ_M pour la valeur propre 1 autrement dit $V = VM$ et $V \neq 0$.

Notons $V = (v_1, \ldots, v_n)$. Soit $V' = (|v_1|, \ldots, |v_n|)$. Montrons que $V'M = V'$. En effet,

soit $w = |v_1| + \cdots + |v_n|$. Pour tout $j \in [[1, n]]$,

$$|v_j| = \left| \sum_{i=1}^{n} v_i p_{i,j} \right| \leqslant \sum_{i=1}^{n} |v_i| p_{i,j}.$$

Sommant ces égalités sur j, il vient

$$w = \sum_{j=1}^{n} |v_j| \leqslant \sum_{i=1}^{n} |v_i| \sum_{j=1}^{n} p_{i,j} = \sum_{i=1}^{n} |v_i| = w,$$

donc ces n inégalités sont en fait des égalités, ce qui montre exactement que $V'M = V'$.

Soit $Q_0 = w^{-1}V'$. Alors Q_0 est un vecteur stochastique stable, donc une loi stationnaire.

4. Par hypothèse, Q_0 est la loi de X_k, pour tout $k \in \mathbb{N}$. Supposons que $q_i > 0$ pour tout i. Encore une formule de Bayes donne, pour tous i et j,

$$P(X_k = j \mid X_{k+1} = i) = \frac{P(X_k = j)}{P(X_{k+1} = i)} P(X_{k+1} = i \mid X_k = j) = \frac{q_j}{q_i} p_{j,i}.$$

V Un exemple de chaîne de Markov

0. Pour tout $k \geqslant 0$, X_k vaut 1 ou 2 donc $P(X_k = 2) = 1 - P(X_k = 1)$. Soit $p_k = P(X_k = 1)$, donc $p_0 = p$.

Si $X_0 = 1$, il reste 1 boule rouge et 2 bleues avant le tirage 1 donc $X_1 = 1$ avec probabilité $\frac{1}{3}$. Si $X_0 = 2$, il reste 2 boules rouges et 1 bleue avant le tirage 1 donc $X_1 = 1$ avec probabilité $\frac{2}{3}$. Au total, $p_1 = P(X_1 = 1) = p\frac{1}{3} + (1-p)\frac{2}{3} = (2-p)/3$.

Donc $E(X_1) = p_1 \times 1 + (1-p_1) \times 2 = 2 - p_1 = (4+p)/3$, $E(X_1^2) = p_1 \times 1 + (1-p_1) \times 4 = 4 - 3p_1 = 2 + p$ et la variance de X_1 vaut $E(X_1^2) - E(X_1)^2 = \frac{1}{9}(2 + p - p^2)$.

1. La loi initiale est donnée par l'énoncé et vaut $q_1 = P(X_0 = 1) = p = 1 - P(X_0 = 2)$.

La couleur de la boule hors de l'urne entre les tirages k et $k+1$ détermine le contenu de l'urne avant le tirage $k+1$ donc la loi de X_{k+1}. Or, cette couleur est rouge si $X_k = 1$ et bleue si $X_k = 2$, donc X_k détermine la loi de X_{k+1}. Ceci prouve (CM2). De plus, si $X_k = 1$, on sait que $[X_{k+1} = 1]$ avec probabilité $\frac{1}{3}$, et si $X_k = 2$, on sait que $[X_{k+1} = 1]$ avec probabilité $\frac{2}{3}$. Ceci prouve (CM1) et fournit la matrice de transition $M = \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{pmatrix}$.

2. La trace de M vaut $\frac{2}{3}$, c'est la somme des deux valeurs propres et l'une vaut 1 donc l'autre vaut $\frac{2}{3} - 1 \neq 1$, donc 1 est valeur propre simple. Il n'y a donc qu'une loi stationnaire. Comme M est symétrique, c'est la loi uniforme décrite par le vecteur $(1/2, 1/2)$.

3. Tous les coefficients de M sont strictement positifs donc d'après III-4, $(M^k)_{k\in\mathbb{N}}$ converge vers une matrice stochastique B dont toutes les lignes sont égales au vecteur (b_1, b_2), vecteur propre de ψ_M pour la valeur propre 1 d'après III-2.b. D'après la question précédente ce vecteur décrit l'unique loi stationnaire donc $B = (1/2, 1/2)$.

4. On a $[N = 0] = R_0$ et pour $k \geqslant 1$, $[N = k] = R_0^c \cap \cdots \cap R_{k-1}^c \cap R_k$.

5. $P[N = 0] = P(R_0) = p$. Pour $k \geqslant 1$, $[N = k] = [X_0 = 2, \ldots, X_{k-1} = 2, X_k = 1]$. Par la propriété de Markov,

$$P(N = k) = P(X_0 = 2)P(X_1 = 2 \,|\, X_0 = 2)^{k-1}P((X_1 = 1 \,|\, X_0 = 2).$$

Donc $P(N = k) = 2(1 - p)/3^k$ pour tout $k \geqslant 1$.

On vérifie que la somme des $P(N = k)$ sur $k \geqslant 0$ vaut 1 et on calcule l'espérance et la variance de N en utilisant les formules suivantes. Pour tout nombre complexe $|z| < 1$,

$$s(z) = \sum_{n \geqslant 0} z^n = \frac{1}{1 - z}.$$

En dérivant cette série entière, on obtient également

$$s'(z) = \sum_{n \geqslant 0} nz^{n-1} = \frac{1}{(1 - z)^2}, \qquad s''(z) = \sum_{n \geqslant 0} n(n - 1)z^{n-2} = \frac{2}{(1 - z)^3}.$$

La première formule pour $z = \frac{1}{3}$ donne $P(N \geqslant 1) = \frac{2}{3}(1 - p)s(z) = 1 - p$, comme il se doit. En utilisant la première dérivée, toujours pour $z = \frac{1}{3}$, on obtient

$$E(N) = \sum_{n \geqslant 1} (1 - p)nz^{n-1}(1 - z) = (1 - p)(1 - z)s'(z),$$

donc $E(N) = (1 - p)/(1 - z) = \frac{3}{2}(1 - p)$. Enfin, la deuxième dérivée donne

$$E(N(N - 1)) = \sum_{n \geqslant 0} (1 - p)n(n - 1)z^{n-1}(1 - z) = (1 - p)z(1 - z)s''(z),$$

donc $E(N(N - 1)) = 2(1 - p)z/(1 - z)^2 = \frac{3}{2}(1 - p)$.

La variance de N vaut $E(N(N - 1)) + E(N) - E(N)^2 = 2a - a^2$ avec $a = \frac{3}{2}(1 - p)$, donc la variance de N vaut $\frac{3}{4}(1 - p)(1 + 3p)$.