PROBABILITES
PC*1
2025 - 2026

Fabrice Monfront
Lycée du Parc

Chapitre 1

Le modèle de Kolmogorov

1.1 Introduction

1.1.1 Un problème étudié en Sup

On jette une pièce équilibrée n fois.
Quelle est la probabilité d'avoir obtenu un nombre pair de fois Pile ?

Cette question peut être traitée de deux manières : sans ou avec variables aléatoires.
Si on n'utilise pas de variable aléatoire, on prend comme univers $\Omega = \{0;1\}^n$ (où on code Pile par 1 et Face par 0). Les évènements élémentaires sont les n-uplets de 0 et de 1.
L'évènement dont on cherche la probabilité, comme tout évènement, est un ensemble d'évènements élémentaires ie une partie de Ω. Ici, il s'agit de l'ensemble des n-uplets de 0 et de 1 avec un nombre pair de 1.
Il y a ici équiprobabilité des évènements élémentaires et la probabilité cherchée est le nombre de cas favorables divisé par le nombre total de cas.
Le nombre total de cas, c'est à dire le cardinal de l'univers, est égal à 2^n.
La question posée se ramène donc à un problème de dénombrement : combien y a-t-il de n-uplets de 0 et de 1 constitué d'un nombre pair de 1.
La réponse est $Np = \displaystyle\sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k}$.

Pour donner une expression plus simple de cette somme, on introduit $Ni = \displaystyle\sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{2k+1}$.

D'après la formule du binôme :

$$\begin{aligned} Np + Ni &= \sum_{k=0}^{n} \binom{n}{k} = 2^n \\ Np - Ni &= \sum_{k=0}^{n} \binom{n}{k} (-1)^k = (1-1)^n = 0 \end{aligned}$$

Donc $Np = 2^{n-1}$ et la probabilité cherchée vaut $\dfrac{1}{2}$.

Si on utilise les variables alétoires, on note pour $i \in [\![1;n]\!]$, X_i le résultat du i-ème lancer (1 pour Pile et 0 pour Face), sans se préoccuper d'expliciter l'univers.
Les variables aléatoires $X_1, \ldots, X_n$ sont mutuellement indépendantes et suivent toutes la loi de

Bernoulli de paramètre $\frac{1}{2}$.
Le nombre de Pile obtenu est $\sum_{k=1}^{n} X_k$. C'est une variable aléatoire qui suit la loi binomiale de paramètres n et $\frac{1}{2}$.
On déduit alors des propriétés de la loi binomiale que la probabilité cherchée est :
$\sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} \left(\frac{1}{2}\right)^{2k} \left(1-\frac{1}{2}\right)^{n-2k} = \frac{Np}{2^n}$, ce qui nous ramène au calcul précédent.

1.1.2 Un problème qu'on peut étudier en Spé mais pas en Sup

On jette une pièce équilibrée jusqu'à obtenir Pile.
Quelle est la probabilité qu'on ait lancé la pièce un nombre pair de fois ?

Ce problème peut être traité sans ou avec variables aléatoires, même si il est recommandé de le faire avec des variables aléatoires.
Sans variable aléatoire, on définit l'univers. Il est naturel de prendre comme univers $\Omega = \{\alpha; \omega_1; \omega_2, \ldots\}$ où ω_k est le cas où on obtient Pile pour la première fois au k-ème lancer et α le cas où on n'obtient jamais Pile.
Contrairement aux situations étudiées en Sup, Ω est infini.
Sur un univers infini, il ne peut y avoir équiprobabilité, car on aurait $1 = \infty \times p$ avec $p > 0$.
La probabilité de ω_k vaut $\frac{1}{2^k}$ ($k-1$ fois Face puis Pile) mais comment le justifier rigoureusement ?
Il est préférable d'utiliser des variables aléatoires. On note toujours X_i le résultat du i-ème lancer (1 pour Pile et 0 pour Face), sans se préoccuper d'expliciter l'univers mais cette fois i décrit $\mathbb{N}^*$.
Les variables aléatoires $X_i, i \in \mathbb{N}^*$ sont mutuellement indépendantes et suivent toutes la loi de Bernoulli de paramètre $\frac{1}{2}$.
Le rang d'obtention du premier Pile est une nouvelle variable aléatoire et on cherche la probabilité qu'elle soit paire.
Pour être plus rigoureux :
$T = \min\left(\{n \in \mathbb{N}^* \text{ tq } X_n = 1\}\right)$ si $\{n \in \mathbb{N}^* \text{ tq } X_n = 1\} \neq \emptyset$, $+\infty$ sinon.
On cherche $P(T \in 2\mathbb{N})$.
L'évènement $(T \in 2\mathbb{N})$ est la réunion des évènements $(T = 2k)$: $(T \in 2\mathbb{N}) = \bigcup_{k \in \mathbb{N}^*} (T = 2k)$.
Contrairement aux situations vues en Sup, la réunion est infinie. On a donc besoin d'une propriété relative aux réunions infinies d'évènements.
$(T = 2k) = \left(\bigcap_{l=1}^{2k-1} (X_l = 0)\right) \bigcap (X_{2k} = 1)$
On en déduit par indépendance que $P(T = 2k) = \frac{1}{2^{2k}} = \frac{1}{4^k}$.
La probabilité cherchée est donc $\sum_{k=1}^{+\infty} \frac{1}{4^k} = \frac{1}{4}\frac{1}{1-1/4} = \frac{1}{3}$.

1.2 Univers, évènements

1.2.1 Tribu sur un ensemble

Définition
Si Ω est un ensemble, on appelle **tribu** sur Ω toute partie $\mathcal{A}$ de l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω telle que :

i $\Omega \in \mathcal{A}$

ii pour tout $A \in \mathcal{A}$, $\overline{A} = \Omega \setminus A \in \mathcal{A}$

iii pour toute famille finie ou dénombrable [1] $(A_i)_{i\in I}$ d'éléments de $\mathcal{A}$, la réunion $\bigcup_{i\in I} A_i$ appartient à $\mathcal{A}$

Exemple
$\mathcal{P}(\Omega)$ est une tribu sur Ω. C'est la seule qui soit intéressante si Ω est fini ou dénombrable [2].
Par contre si Ω est infini, la question se complique considérablement mais je cite le programme :

La notion de tribu n'appelle aucun autre développement que sa définition.

1.2.2 Espace probabilisable : définition

On appelle espace probabilisable tout couple $(\Omega, \mathcal{A})$ où Ω est un ensemble et $\mathcal{A}$ une tribu sur Ω.
Ω est appelé l'univers et les éléments de $\mathcal{A}$ évènements.
La définition d'une tribu signifie donc que :

i Ω est un évènement, dit évènement certain

ii la non réalisation d'un évènement est encore un évènement

iii la réalisation d'au moins un des évènements d'un ensemble fini ou dénombrable d'évènements est encore un évènement.
Le programme signale la traduction de l'évènement $\bigcup_{n=0}^{+\infty} A_n$ par :
$\exists n \in \mathbb{N}$ tq A_n est réalisé

1.2.3 Espace probabilisable : propriétés

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

— Pour toute famille au plus dénombrable $(A_i)_{i\in I}$ d'éléments de $\mathcal{A}$, l'intersection $\bigcap_{i\in I} A_i$ appartient à $\mathcal{A}$ En d'autres termes, si on considère une famille au plus dénombrable d'évènements alors leur réalisation simultanée est un évènement.

Preuve
D'après la définition d'une tribu :
$\forall i \in I\ B_i = \overline{A_i} \in \mathcal{A}$
Toujours d'après la définition d'une tribu, $\bigcup_{i\in I} B_i$ appartient à $\mathcal{A}$.

1. Cf cette notion en appendice
2. On peut facilement démontrer dans ce cas que $\mathcal{P}(\Omega)$ est la seule tribu qui contienne tous les évènements élémentaires ie tous les singletons

Enfin $\bigcap_{i\in I} A_i = \overline{\bigcup_{i\in I} B_i} \in \mathcal{A}$

Le programme signale la traduction de l'évènement $\bigcap_{n=0}^{+\infty} A_n$ par :

$\forall n \in \mathbb{N}\ A_n$ est réalisé

— $\emptyset \in \mathcal{A}$.
En effet $\emptyset = \overline{\Omega}$
$\emptyset$ est appelé "évènement impossible" et Ω "évènement certain".

— Soient A et B appartenant à $\mathcal{A}$.
Alors $A \setminus B \in \mathcal{A}$.
En effet, $A \setminus B = A \cap \overline{B}$
En d'autres termes, si A et B sont deux évènements, "A est réalisé mais pas B" est aussi un évènement.

1.2.4 Un exemple de parallèle entre le vocabulaire probabiliste et le vocabulaire ensembliste

On effectue une suite infinie de lancers d'un dé. Pour tout $i \in \mathbb{N}^*$, on note :
$A_i = \{$Obtention de l'as au $i^{\text{ième}}$ lancer$\}$

1. Définir par une phrase ne comportant aucun vocabulaire mathématique chacun des évènements :
   — $E_1 = \bigcap_{i=4}^{+\infty} A_i$
   — $E_2 = \left(\bigcap_{i=1}^{3} \overline{A_i}\right) \cap \left(\bigcap_{i=4}^{+\infty} A_i\right)$
   — $E_3 = \bigcup_{i>3} A_i$
2. Ecrire à l'aide des A_i, l'évènement "on obtient au moins une fois l'as au-delà du $n^{\text{ième}}$ lancer".
3. On pose $C_n = \bigcup_{i>n} A_i$. Montrer que la suite $(C_n)_{n\in\mathbb{N}^*}$ est décroissante (ie pour tout $n \geq 1$, C_{n+1} est inclus dans C_n).
   Caractériser par une phrase ne comportant aucun vocabulaire mathématique l'évènement $C = \bigcap_{n\geq 1} C_n$.
4. Ecrire à l'aide des A_i les évènements :
   — $B_n = \{$On n'obtient plus que des as à partir du $n^{\text{ième}}$ lancer$\}$
   — $B = \{$On n'obtient plus que des as à partir d'un certain lancer$\}$

Correction

1. — E_1 est l'évènement : "on obtient l'as à chaque lancer à partir du quatrième"
   — E_2 est l'évènement : "on n'obtient l'as à aucun des trois premiers lancers puis on obtient l'as à chaque lancer"
   — E_3 est l'évènement : "on obtient l'as au moins une fois à partir du quatrième lancer (inclus)"
2. Il s'agit de $\bigcup_{i>n} A_i$.

3. $C_n = \bigcup_{i>n} A_i = A_{n+1} \bigcup \left(\bigcup_{i>n+1} A_i \right) = A_{n+1} \cup C_{n+1}$
   Donc C_{n+1} est inclus dans C_n

   C est l'évènement : "on obtient une infinité de fois l'as" :

$$\begin{aligned} \omega \in C &\iff \forall n \in \mathbb{N}^* \ \omega \in C_n \\ &\iff \forall n \in \mathbb{N}^* \ \exists i > n \text{ tq } \omega \in A_i \\ &\iff I(\omega) \text{ n'est pas majoré où on note } I(\omega) = \{n \in \mathbb{N}^* \text{ tq } \omega \in A_n\} \\ &\iff I(\omega) \text{ est infini car } I(\omega) \subset \mathbb{N} \end{aligned}$$

4. $B_n = \bigcap_{i=n}^{+\infty} A_i$ et $B = \bigcup_{n=1}^{+\infty} B_n = \bigcup_{n=1}^{+\infty} \left(\bigcap_{i=n}^{+\infty} A_i \right)$

1.3 Variables aléatoires discrètes

1.3.1 Définitions

Soit $(\Omega, \mathcal{A})$ un espace probabilisable. On appelle variable aléatoire discrète sur $(\Omega, \mathcal{A})$ toute application X de Ω dans un ensemble E (quelconque) telle que :

i $X(\Omega)$, l'image de X, est au plus dénombrable.

ii l'image réciproque de tout élément de $X(\Omega)$ appartient à $\mathcal{A}$.
En d'autres termes, pour tout $x \in X(\Omega)$, $X^{-1}(\{x\}) = \{\omega \in \Omega \text{ tq } X(\omega) = x\}$ est un évènement.

Lorsque $E = \mathbb{R}$, X est dite réelle.

1.3.2 Notations

Soit $(\Omega, \mathcal{A})$ un espace probabilisable et X une variable aléatoire discrète sur $(\Omega, \mathcal{A})$ à valeurs dans un ensemble E quelconque.
Si $x \notin X(\Omega)$ alors $X^{-1}(\{x\}) = \emptyset$ est un évènement donc en fait :
Pour tout $x \in E$, $X^{-1}(\{x\}) = \{\omega \in \Omega \text{ tq } X(\omega) = x\}$ est un évènement qu'on note $(X = x)$ ou $\{X = x\}$.

Pour tout $U \subset E$, $X^{-1}(U) = \{\omega \in \Omega \text{ tq } X(\omega) \in U\} = \bigcup_{u \in U \cap X(\Omega)} X^{-1}(\{u\})$ est un évènement comme réunion au plus dénombrable d'évènements. On le note $(X \in U)$.
Dans le cas particulier où $E = \mathbb{R}$, on note pour $x \in \mathbb{R}$, $(X \geq x)$ l'évènement $X \in [x; +\infty[$.
On définit de même les notations $X \leq x$, $X < x$ et $X > x$.

1.3.3 Exemple

Soit $(\Omega, \mathcal{A})$ un espace probabilisable et $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles discrètes sur $(\Omega, \mathcal{A})$.
Soit $T = \begin{cases} \min\left(\{n \in \mathbb{N}^* \text{ tq } X_n = 0\}\right) \text{ si } \{n \in \mathbb{N}^* \text{ tq } X_n = 0\} \neq \emptyset \\ +\infty \text{ si } \{n \in \mathbb{N}^* \text{ tq } X_n = 0\} = \emptyset \end{cases}$
T est une variable aléatoire sur $(\Omega, \mathcal{A})$:

$T(\Omega)$ est inclus dans $\mathbb{N}^* \cap \{\infty\}$ qui est dénombrable donc $T(\Omega)$ est au plus dénombrable.
$T^{-1}(\{+\infty\}) = \bigcap_{n=1}^{+\infty} (X_n \neq 0)$ est un évènement comme intersection dénombrable d'évènements.
Pour tout $n \in \mathbb{N}^*$, $T^{-1}(\{n\}) = \left(\bigcap_{k=1}^{n-1} (X_k \neq 0)\right) \cap (X_n = 0)$ est un évènement.

1.3.4 Image d'une variable aléatoire par une fonction

Soit $(\Omega, \mathcal{A})$ un espace probabilisable et X une variable aléatoire discrète sur $(\Omega, \mathcal{A})$.
Soit f une fonction définie sur $X(\Omega)$.
$f(X)$ est une variable aléatoire.

Démonstration
On note $Y = f(X) \begin{cases} \Omega \to F \\ \omega \mapsto f(X(\omega)) \end{cases}$ où F est l'ensemble d'arrivée de f.
$Y(\Omega) = f(X(\Omega))$ est au plus dénombrable car $X(\Omega)$ est au plus dénombrable.
Soit $y \in Y(\Omega)$.
Soit B l'ensemble des antécédents de y par f : $B = \{x \in X(\Omega) \text{ tq } f(x) = y\}$.
B est au plus dénombrable car B est inclus dans $X(\Omega)$ qui est au plus dénombrable.
$Y^{-1}(\{y\}) = \bigcup_{x \in B} X^{-1}(\{x\})$ est donc un évènement.

1.3.5 Couples de variables aléatoires

Soit $(\Omega, \mathcal{A})$ un espace probabilisable.
Soient $X : \Omega \to E$ et $Y : \Omega \to F$ deux applications.
Soit $C \begin{cases} \Omega \to E \times F \\ \omega \mapsto (X(\omega), Y(\omega)) \end{cases}$.
C est une variable aléatoire discrète si, et seulement si, X et Y sont des variables aléatoires discrètes.
On dit alors que C est un couple de variables aléatoires discrètes.

Démonstration

— On suppose que X et Y sont des variables aléatoires discrètes.
$X(\Omega)$ et $Y(\Omega)$ sont donc au plus dénombrables.
$X(\Omega) \times Y(\Omega)$ est alors au plus dénombrable.
Attention : $C(\Omega)$ peut-être strictement contenu dans $X(\Omega) \times Y(\Omega)$, par exemple si $X = Y$: $C(\Omega) = \{(x, x), x \in X(\Omega)\}$ alors que $X(\Omega) \times Y(\Omega) = \{(x, y), x \in X(\Omega), y \in Y(\Omega)\}$.
Soit $c \in C(\Omega)$.
$\exists (x, y) \in X(\Omega) \times Y(\Omega)$ tq $c = (x, y)$

$$\begin{aligned} C^{-1}(\{c\}) &= \{\omega \in \Omega \text{ tq } X(\omega) = x \text{ et } Y(\omega) = y\} \\ &= X^{-1}(\{x\}) \cap Y^{-1}(\{y\}) \text{ est un évènement} \end{aligned}$$

Réciproquement, on suppose que C est une variable aléatoire discrète.
Soit $\pi \begin{cases} E \times F \to E \\ (x, y) \mapsto x \end{cases}$
$X = \pi(C)$ donc X est une variable aléatoire.

On montre de même que Y est une variable aléatoire.

Plus généralement si $X_1, \dots, X_n$ sont n applications définies sur Ω alors :
($\forall i \in [\![1; n]\!]$ X_i est une variable aléatoire discrète) $\Longleftrightarrow$ $(X_1, \dots, X_n)$ est une variable aléatoire discrète
Il résulte alors de 1.3.4 que pour tout fonction définie sur un ensemble contenant $X_1(\Omega) \times \cdots \times X_n(\Omega)$, $f(X_1, \dots, X_n)$ est une variable aléatoire discrète.
Par exemple, $X_1 + \cdots + X_n$ est une variable aléatoire discrète.

1.4 Probabilités

1.4.1 Définitions

— Soit $(\Omega, \mathcal{A})$ un espace probabilisable.
On appelle **probabilité** sur $(\Omega, \mathcal{A})$ une application
$P : \mathcal{A} \to [0; 1]$ telle que :

i $P(\Omega) = 1$

ii Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'évènements deux à deux incompatibles (ie pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'élèments de $\mathcal{A}$ telle que $A_n \cap A_m = \emptyset$ pour tout couple d'entiers (n, m) tel que $n \neq m$)

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

Cette propriété est appelée σ-additivité.

— On appelle espace probabilisé un triplet $(\Omega, \mathcal{A}, P)$ où $(\Omega, \mathcal{A})$ est un espace probabilisable et P une probabilité sur $(\Omega, \mathcal{A})$.

1.4.2 A propos de la σ-additivité

Dans tout ce paragraphe, $(\Omega, \mathcal{A}, P)$ est un espace probabilisé.

— La suite $(A_n)_{n\in\mathbb{N}}$ avec $A_n = \emptyset$ pour tout $n \in \mathbb{N}$ est une suite d'éléments de $\mathcal{A}$.
Pour tout couple $(n, m) \in \mathbb{N}^2$ tel que $n \neq m$, on a :
$A_n \cap A_m = \emptyset \cap \emptyset = \emptyset$
Donc, par σ-additivité :
$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n)$
ou encore :
$P(\emptyset) = \sum_{n=0}^{+\infty} P(\emptyset)$

Cela n'est possible que si $P(\emptyset) = 0$: $P(\emptyset) \geq 0$ et si $p > 0$ alors $\sum_{n=0}^{+\infty} p = +\infty$.

Cette propriété ne figure pas dans le cours sur les séries. Pour les besoins des démonstrations théoriques du cours de probabilités, le programme introduit avec beaucoup de précautions pour en limiter l'usage aux probabilités la notion de famille sommable. Cette notion est développée en annexe à la fin de ce cours.

La probabilité de l'évènement impossible est nulle. On verra plus bas que la réciproque est fausse.

— Soient A et B deux évènements incompatibles.
Alors : $P(A \cup B) = P(A) + P(B)$.
Il suffit d'appliquer la σ-additivité avec $A_0 = A$, $A_1 = B$ et pour tout $n \geq 2$, $A_n = \emptyset$.
Plus généralement et par la même méthode ou par récurrence, on peut montrer :
— Soient $A_1, \ldots, A_p$ une famille finie d'évènements deux à deux incompatibles.
Alors $P\left(\bigcup_{i=1}^{p} A_i\right) = \sum_{i=1}^{p} P(A_i)$
On a donc pour toute famille finie $(A_i)_{i \in I}$ d'évènements deux à deux incompatibles :
$P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i)$.
Cette propriété reste vraie si I est dénombrable.
Soit φ une bijection de $\mathbb{N}$ sur I.
$P\left(\bigcup_{i \in I} A_i\right) = P\left(\bigcup_{n=0}^{+\infty} A_{\varphi(n)}\right) = \sum_{n=0}^{+\infty} P(A_{\varphi(n)})$ par σ-additivité.
D'où $P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i)$ d'après les propriétés des familles sommables.
— Pour tout évènement A, la probabilité de l'évènement contraire vaut $1 - P(A)$.
En d'autres termes :
$\forall A \in \mathcal{A}\ P(\overline{A}) = 1 - P(A)$
En effet A et $\overline{A}$ sont incompatibles donc :
$P(A) + P(\overline{A}) = P(A \cup \overline{A}) = P(\Omega) = 1$

— **Croissance de P**
$\forall (A, B) \in \mathcal{A}^2\ A \subset B \Longrightarrow P(A) \leq P(B)$
En d'autres termes, si la réalisation de l'évènement A entraîne celle de B alors la probabilité de A est inférieure ou égale à celle de B.
Pour la démonstration, il suffit de remarquer que B est la réunion des deux évènements incompatibles A et $B \setminus A$ ce qui entraîne :
$P(B) = P(A) + P(B \setminus A) \geq P(A)$

— $\forall (A, B) \in \mathcal{A}^2\ A \subset B \Longrightarrow P(B \setminus A) = P(B) - P(A)$
Il suffit de reprendre le calcul du point précédent.
Attention : ce résultat n'est pas valable si A n'est pas inclus dans B.

— $\forall (A, B) \in \mathcal{A}^2\ P(A \cup B) = P(A) + P(B) - P(A \cap B)$
On écrit $A \cup B = (A \setminus (A \cap B)) \cup (A \cap B) \cup (B \setminus (A \cap B))$
Les trois évènements étant (deux à deux) incompatibles, on a :

$$\begin{aligned} P(A \cup B) &= P(A \setminus (A \cap B)) + P(A \cap B) + P(B \setminus (A \cap B)) \\ &= P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

1.4.3 Continuité croissante

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements telle que pour tout $n \in \mathbb{N}$, on ait $A_n \subset A_{n+1}$. Alors :

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

Démonstration
Par croissance de P, la suite $(P(A_n))_{n\in\mathbb{N}}$ est croissante. Comme elle est majorée par 1, elle converge.

On pose $B_0 = A_0$ et pour tout $n \in \mathbb{N}^*$, $B_n = A_n \setminus A_{n-1}$.
D'après les propriétés des tribus :
$\forall n \in \mathbb{N}\ B_n \in \mathcal{A}$

— Les évènements B_n pour $n \in \mathbb{N}$ sont deux à deux disjoints.
Soit $(n, m) \in \mathbb{N}^2$ tq $n \neq m$.
On peut sans restreindre la généralité supposer $n < m$.
$B_n = A_n \setminus A_{n-1} \subset A_n \subset A_{m-1}$, valable aussi si $n = 0$.

$$\begin{aligned} B_m \cap B_n &= \left(A_m \cap \overline{A_{m-1}}\right) \cap B_n \\ &= A_m \cap \left(\overline{A_{m-1}} \cap B_n\right) \\ &= A_m \cap \emptyset \text{ car } B_n \subset A_{m-1} \\ &= \emptyset \end{aligned}$$

— $\displaystyle\bigcup_{n=0}^{+\infty} A_n = \bigcup_{n=0}^{+\infty} B_n$
$\forall n \in \mathbb{N}^*\ B_n = A_n \setminus A_{n-1} \subset A_n$
$B_0 = A_0 \subset A_0$
Donc :
$\forall n \in \mathbb{N}\ B_n \subset A_n$
D'où : $\displaystyle\bigcup_{n=0}^{+\infty} B_n \subset \bigcup_{n=0}^{+\infty} A_n$
Réciproquement, soit $\omega \in \displaystyle\bigcup_{n=0}^{+\infty} A_n$.
$\{n \in \mathbb{N} \text{ tq } \omega \in A_n\}$ est une partie non vide de $\mathbb{N}$ donc elle a un plus petit élément noté n_0.
Si $n_0 = 0$ alors $\omega \in A_0 = B_0$
Si $n_0 \neq 0$ alors $\omega \in A_{n_0}$ et $\omega \notin A_{n_0-1}$ donc $\omega \in B_{n_0}$
Donc $\omega \in \displaystyle\bigcup_{n=0}^{+\infty} B_n$ et $\displaystyle\bigcup_{n=0}^{+\infty} A_n \subset \bigcup_{n=0}^{+\infty} B_n$.

On a donc :

$$\begin{aligned}
P\left(\bigcup_{n=0}^{+\infty} A_n\right) = P\left(\bigcup_{n=0}^{+\infty} B_n\right) &= \sum_{n=0}^{+\infty} P(B_n) \\
&= P(A_0) + \sum_{n=1}^{+\infty} P(A_n \setminus A_{n-1}) \\
&= P(A_0) + \sum_{n=1}^{+\infty} P(A_n) - P(A_{n-1}) \text{ car } A_{n-1} \subset A_n \\
&= P(A_0) + \lim_{n\to+\infty} P(A_n) - P(A_0) \\
&= \lim_{n\to+\infty} P(A_n)
\end{aligned}$$

Remarque
Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements (non nécessairement monotone).
$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{N\to+\infty} P\left(\bigcup_{n=0}^{N} A_n\right)$
Preuve
Pour tout $N \in \mathbb{N}$, on note $B_N = \bigcup_{n=0}^{N} A_n$.
$\forall N \in \mathbb{N}\ B_{N+1} = B_N \cup A_{N+1}$
Donc :
$\forall N \in \mathbb{N}\ B_N \subset B_{N+1}$
De plus :
$\bigcup_{n=0}^{+\infty} A_n = \bigcup_{N=0}^{+\infty} B_N$
Faut-il détailler ce point ?
$\forall n \in \mathbb{N}\ A_n \subset B_n$
Donc $\bigcup_{n=0}^{+\infty} A_n \subset \bigcup_{N=0}^{+\infty} B_N$
Réciproquement :
$\forall N \in \mathbb{N}\ B_N = \bigcup_{n=0}^{N} A_n \subset \bigcup_{n=0}^{+\infty} A_n$
Donc : $\bigcup_{N=0}^{+\infty} B_N \subset \bigcup_{n=0}^{+\infty} A_n$
En tout cas :

$$\begin{aligned}
P\left(\bigcup_{n=0}^{+\infty} A_n\right) &= P\left(\bigcup_{N=0}^{+\infty} B_N\right) \\
&= \lim_{N\to+\infty} P(B_N) \text{ par continuité croissante} \\
&= \lim_{N\to+\infty} P\left(\bigcup_{n=0}^{N} A_n\right)
\end{aligned}$$

1.4.4 Continuité décroissante

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements telle que pour tout $n \in \mathbb{N}$, on ait $A_{n+1} \subset A_n$. Alors :

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

Preuve :
Pour tout $n \in \mathbb{N}$, soit $B_n = \overline{A_n}$.
$(B_n)_{n\in\mathbb{N}}$ est une suite d'évènements telle que pour tout $n \in \mathbb{N}$, on ait $B_n \subset B_{n+1}$.
Donc $\lim\limits_{n\to+\infty} P(B_n) = P\left(\bigcup\limits_{n=0}^{+\infty} B_n\right)$
Comme $P(B_n) = 1 - P(A_n)$, on a :
$\lim\limits_{n\to+\infty} P(A_n) = 1 - \lim\limits_{n\to+\infty} P(B_n) = 1 - P\left(\bigcup\limits_{n=0}^{+\infty} B_n\right) = P\left(\overline{\bigcup\limits_{n=0}^{+\infty} B_n}\right) = P\left(\bigcap\limits_{n=0}^{+\infty} A_n\right)$

Remarque
Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements (non nécessairement monotone).
$P\left(\bigcap\limits_{n=0}^{+\infty} A_n\right) = \lim\limits_{N\to+\infty} P\left(\bigcap\limits_{n=0}^{N} A_n\right)$
Preuve :
Pour tout $N \in \mathbb{N}$, on note $C_N = \bigcap\limits_{n=0}^{N} A_n$.
$\forall N \in \mathbb{N}\ C_{N+1} = C_N \cap A_{N+1} \subset C_N$
De plus :
$\bigcap\limits_{n=0}^{+\infty} A_n = \bigcap\limits_{N=0}^{+\infty} C_N$
Faut-il détailler ce point ?
$\forall N \in \mathbb{N}\ C_N \subset A_N$
Donc $\bigcap\limits_{N=0}^{+\infty} C_N \subset \bigcap\limits_{n=0}^{+\infty} A_n$
Réciproquement :
$\forall N \in \mathbb{N}\ \bigcap\limits_{n=0}^{+\infty} A_n \subset \bigcap\limits_{n=0}^{N} A_n = C_N$
Donc : $\bigcap\limits_{n=0}^{+\infty} A_n \subset \bigcap\limits_{N=0}^{+\infty} C_N$
En tout cas :

$$\begin{aligned}
P\left(\bigcap_{n=0}^{+\infty} A_n\right) &= P\left(\bigcap_{N=0}^{+\infty} C_N\right)\\
&= \lim_{N\to+\infty} P(C_N) \text{ par continuité décroissante}\\
&= \lim_{N\to+\infty} P\left(\bigcap_{n=0}^{N} A_n\right)
\end{aligned}$$

1.4.5 Sous-additivité

Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'évènements alors :

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

On rappelle qu'en cas de divergence de la série à termes positifs $\sum_{n\geq 0} P(A_n)$, $\sum_{n=0}^{+\infty} P(A_n) = +\infty$

Démonstration
$\forall (A,B) \in \mathcal{A}^2$ $P(A\cup B) = P(A) + P(B) - P(A\cap B) \leq P(A) + P(B)$ car $P(A\cap B) \geq 0$
On en déduit par récurrence :
$\forall n \in \mathbb{N}^*$ $\forall (A_1,\ldots,A_n) \in \mathcal{A}^n$ $P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$
Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements.
$\forall n \in \mathbb{N}$ $P\left(\bigcup_{i=0}^{n} A_i\right) \leq \sum_{i=0}^{n} P(A_i)$
On fait tendre n vers $+\infty$ en tenant compte des remarques faites après la continuité croissante :
$P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} P(A_n)$, cette somme étant finie ou non.

1.4.6 Evènements presque sûrs, évènements négligeables

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
On appelle évènement presque sûr tout évènement A tel que $P(A) = 1$.
On appelle évènement négligeable tout évènement A tel que $P(A) = 0$.

Exemple
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit $(A_n)_{n\in\mathbb{N}}$ une famille dénombrable d'événements presque sûrs.
Montrer que $\bigcap_{n=0}^{+\infty} A_n$ est encore un évènement presque sûr.

Démonstration

$$\begin{array}{rcl} P\left(\bigcap_{n=0}^{+\infty} A_n\right) & = & 1 - P\left(\overline{\bigcap_{n=0}^{+\infty} A_n}\right) \\ & = & 1 - P\left(\bigcup_{n=0}^{+\infty} \overline{A_n}\right) \\ 0 & \leq & P\left(\bigcup_{n=0}^{+\infty} \overline{A_n}\right) \leq \sum_{n=0}^{+\infty} P(\overline{A_n}) \text{ par sous-additivité} \\ 0 & \leq & P\left(\bigcup_{n=0}^{+\infty} \overline{A_n}\right) \leq \sum_{n=0}^{+\infty} 0 = 0 \\ P\left(\bigcup_{n=0}^{+\infty} \overline{A_n}\right) & = & 0 \\ P\left(\bigcap_{n=0}^{+\infty} A_n\right) & = & 1 \end{array}$$

1.5 Probabilités conditionnelles

1.5.1 Définition

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Pour deux évènements A et B tels que $P(B) > 0$, on appelle probabilité conditionnelle de A sachant B et on note $P_B(A)$ ou $P(A|B)$ le quotient $\dfrac{P(A \cap B)}{P(B)}$.

Remarque
Attention à la notation $P(A|B)$: $A|B$ n'est pas un évènement.

1.5.2 Probabilité P_B

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit B un évènement tel que $P(B) > 0$.
L'application $P_B \begin{cases} \mathcal{A} \to [0;1] \\ A \mapsto P_B(A) = P(A|B) \end{cases}$ est une probabilité sur $(\Omega, \mathcal{A})$.

Démonstration

- Soit $A \in \mathcal{A}$.
  $A \cap B \subset B$ donc $0 \leq P(A \cap B) \leq P(B)$.
  On en déduit $0 \leq P_B(A) \leq 1$.
- $P_B(\Omega) = \dfrac{P(\Omega \cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$.
- Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements deux à deux incompatibles.
  Les évènements $A_n \cap B$, $n \in \mathbb{N}$ sont également deux à deux incompatibles et donc :
  $P\left(\left(\bigcup_{n=0}^{+\infty} A_n\right) \cap B\right) = P\left(\bigcup_{n=0}^{+\infty} (A_n \cap B)\right) = \sum_{n=0}^{+\infty} P(A_n \cap B)$
  Divisant par $P(B)$, on obtient :

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

Remarque
P_B a donc toutes les propriétés d'une probabilité sur $(\Omega, \mathcal{A})$.
On dispose en particulier des relations suivantes, souvent utiles :
- $\forall A \in \mathcal{A}\ P_B(\overline{A}) = 1 - P_B(A)$
- $\forall (A, C) \in \mathcal{A}^2\ P_B(A \cup C) = P_B(A) + P_B(C) - P_B(A \cap C)$

1.6 Formule des probabilités composées

1.6.1 Introduction

Si A et B sont deux évènements tels que $P(B) > 0$ alors $P(A \cap B) = P(A|B) \times P(B)$.
Formellement on tourne en rond mais il existe de nombreuses situations où on sait facilement évaluer $P(A|B)$ et la formule précédente permet de calculer $P(A \cap B)$.
Examinons par exemple la situation suivante :
On considère deux urnes $\mathcal{U}_1$ et $\mathcal{U}_2$ contenant chacune initialement deux boules noires et trois boules blanches. On tire une boule de l'urne $\mathcal{U}_1$, on note sa couleur et on la met dans l'urne $\mathcal{U}_2$.
On tire alors une boule dans l'urne $\mathcal{U}_2$.
Quelle est la probabilité de tirer deux fois une boule noire ?

On note N_1 l'évènement : "la boule tirée de l'urne $\mathcal{U}_1$ est noire" et N_2 l'évènement : "la boule tirée de l'urne $\mathcal{U}_2$ est noire".
On cherche $P(N_1 \cap N_2)$.
$P(N_1 \cap N_2) = P(N_2|N_1) \times P(N_1)$.
$P(N_1) = \dfrac{2}{5}$ car $\mathcal{U}_1$ contient initialement 5 boules dont 2 noires.
$P(N_2|N_1) = \dfrac{3}{6} = \dfrac{1}{2}$ car si la première boule tirée est noire, lors du deuxième tirage $\mathcal{U}_2$ contient 6 boules dont 3 noires.
Finalement la probabilité cherchée est $\dfrac{1}{2} \times \dfrac{2}{5} = \dfrac{1}{5}$.

1.6.2 Formule des probabilités composées

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient $A_1, \ldots, A_n$ n évènements tels que $P(A_1 \cap \cdots \cap A_{n-1}) > 0$.
Alors :

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

Remarques
- Les probabilités conditionnelles écrites dans cette formule ont bien un sens :
  $\forall i \in [\![1; n-1]\!]\ P(A_1 \cap \cdots \cap A_i) > 0$
  En effet $A_1 \cap \cdots \cap A_{n-1} \subset A_1 \cap \cdots \cap A_i$ et on utilise la croissance de la probabilité.
- Cette formule se démontre par récurrence sur n.

1.6.3 Exemple

On considère une urne contenant 4 boules blanches et 3 boules noires.
On tire une à une et sans remise 3 boules de l'urne.

Quelle est la probabilité que la première boule tirée soit blanche, la seconde blanche et la troisième noire ?

Notons B_i l'évènement "la $i^{\text{ème}}$ boule tirée est blanche" et N_i l'évènement "la $i^{\text{ème}}$ boule tirée est noire".
On cherche $P(B_1 \cap B_2 \cap N_3)$.
Utilisons la formule des probabilités composées :
$P(B_1 \cap B_2 \cap N_3) = P(B_1) \times P(B_2|B_1) \times P(N_3|B_1 \cap B_2) = \dfrac{4}{7} \times \dfrac{3}{6} \times \dfrac{3}{5} = \dfrac{6}{35}$

1.7 Formule des probabilités totales

1.7.1 Système complet d'évènements

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
On appelle système complet (sous-entendu au plus dénombrable dans le cadre du programme) d'évènements tout famille $(A_i)_{i\in I}$ d'évènements telle que :

i I est un ensemble au plus dénombrable

ii $\forall (i,j) \in I^2 \ i \neq j \Rightarrow A_i \cap A_j = \emptyset$ ie les évènements de la famille sont deux à deux incompatibles.

iii $\displaystyle\bigcup_{i\in I} A_i = \Omega$

1.7.2 Formule des probabilités totales

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit B un évènement.
Si $(A_i)_{i\in I}$ est un système complet d'évènements, fini ou dénombrable, alors
$P(B) = \displaystyle\sum_{i\in I} P(B \cap A_i) = \sum_{i\in I} P(B|A_i)\, P(A_i)$

Remarque
Dans un système complet d'évènements, il est possible qu'il existe des évènements A_i de probabilité nulle.
Dans ce cas $P(B|A_i)$ n'est pas définie et la formule précédente ne peut pas être appliquée telle quelle.
En pratique pour préserver cette formule comme outil de calcul, on adopte la convention
$P(B|A_i)\, P(A_i) = 0$ lorsque $P(A_i) = 0$.

Démonstration de la formule des probabilités totales
$B = \displaystyle\bigcup_{i\in I} (B \cap A_i)$ car $\displaystyle\bigcup_{i\in I} A_i = \Omega$.
$(B \cap A_i) \cap (B \cap A_j) = B \cap (A_i \cap A_j)$ donc les $B \cap A_i$ sont deux à deux disjoints.
Donc, d'après la σ-additivité, $P(B) = \displaystyle\sum_{i\in I} P(B \cap A_i) = \sum_{i\in I} P(B|A_i)\, P(A_i)$ avec la convention ci-dessus.

1.7.3 Généralisation

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
On appelle système quasi-complet (sous-entendu au plus dénombrable dans le cadre du pro-

gramme) d'évènements tout famille $(A_i)_{i\in I}$ d'évènements telle que :

i I est un ensemble au plus dénombrable

ii $\forall (i,j) \in I^2 \; i \neq j \Rightarrow A_i \cap A_j = \emptyset$ ie les évènements de la famille sont deux à deux incompatibles.

iii $P\left(\bigcup_{i\in I} A_i\right) = 1$, ou ce qui revient au même si les deux premières hypothèses sont vérifiées : $\sum_{i\in I} P(A_i) = 1$

Soit B un évènement.

Si $(A_i)_{i\in I}$ est un système quasi-complet d'évènements, fini ou dénombrable, alors
$P(B) = \sum_{i\in I} P(B\cap A_i) = \sum_{i\in I} P(B|A_i)\,P(A_i)$

Remarque

Dans un système quasi-complet d'évènements, il est possible qu'il existe des évènements A_i de probabilité nulle.

Dans ce cas $P(B|A_i)$ n'est pas définie et la formule précédente ne peut pas être appliquée telle quelle.

En pratique pour préserver cette formule comme outil de calcul, on adopte la convention
$P(B|A_i)\,P(A_i) = 0$ lorsque $P(A_i) = 0$.

Démonstration de la formule des probabilités totales avec un système quasi-complet d'évènements

Soit $A = \Omega \setminus \left(\bigcup_{i\in I} A_i\right)$.

$P(A) = 0$

Soit $k \notin I$ et $J = \{k\} \cup I$.

Si on pose $A_k = A$ alors la famille $(A_j)_{j\in J}$ est un système complet d'évènements.

Donc :

$$\begin{array}{rcl} P(B) & = & P(B\cap A) + \sum_{i\in I} P(B\cap A_i) \\ & = & 0 + \sum_{i\in I} P(B\cap A_i) \text{ par croissance de } P \\ & = & \sum_{i\in I} P(B\cap A_i) = \sum_{i\in I} P(B|A_i)\,P(A_i) \end{array}$$

1.8 Formule de Bayes

1.8.1 Proposition

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

Soient A et B deux évènements de probabilités strictement positives.

On a : $P(A|B) = \dfrac{P(B|A)\times P(A)}{P(B)}$.

En effet, $P(A\cap B) = P(A|B)\times P(B)$
et en inversant les rôles de A et de B : $P(B\cap A) = P(B|A)\times P(A)$.

Comme $A\cap B = B\cap A$, on a : $P(A|B)\times P(B) = P(B|A)\times P(A)$ et on en déduit le résultat en divisant par $P(B) > 0$.

1.8.2 Proposition

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit $(A_i)_{i\in I}$ un système complet ou quasi complet, au plus dénombrable, d'évènements.
Soit B un évènement de probabilité non nulle.
Soit $i_0 \in I$ tel que $P(A_{i_0}) > 0$.
On a :

$$P(A_{i_0}|B) = \frac{P(B|A_{i_0})\,P(A_{i_0})}{\displaystyle\sum_{i\in I} P(B|A_i)\,P(A_i)}$$

Démonstration
D'après la formule des probabilités totales : $P(B) = \displaystyle\sum_{i\in I} P(B|A_i)\,P(A_i)$.
Il suffit donc d'appliquer le paragraphe précédent avec $A = A_{i_0}$.

Remarque
On applique souvent cette formule avec le système complet $\{A; \overline{A}\}$ où A est un évènement de probabilité différente de 0 et de 1.
On obtient, pour B évènement de probabilité non nulle :
$P(A|B) = \dfrac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|\overline{A})\,P(\overline{A})}$

Exemple

On prend un dé au hasard parmi un lot de 100 dés dont on sait que 25 sont pipés. Pour un dé pipé, la probabilité d'obtenir un 6 est de $\dfrac{1}{2}$.
On lance un dé et on obtient 6. Quelle est la probabilité que ce dé soit pipé ?
On relance alors ce dé et on obtient à nouveau 6. Quelle est la probabilité que ce dé soit pipé ?

Première question
On a, avec les notations précédentes, A : "le dé est pipé" et B : "on obtient un 6".
D'où, par application de la formule précédente, la probabilité cherchée :
$P(A|B) = \dfrac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|\overline{A})\,P(\overline{A})} = \dfrac{\frac{1}{2}\,\frac{25}{100}}{\frac{1}{2}\,\frac{25}{100} + \frac{1}{6}\,\frac{75}{100}} = \dfrac{1}{2}$

Deuxième question
On a, avec les notations précédentes, A : "le dé est pipé" et B : "on obtient deux fois 6".
D'où, par application de la formule précédente, la probabilité cherchée :
$P(A|B) = \dfrac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|\overline{A})\,P(\overline{A})} = \dfrac{\frac{1}{4}\,\frac{25}{100}}{\frac{1}{4}\,\frac{25}{100} + \frac{1}{36}\,\frac{75}{100}} = \dfrac{3}{4}$

1.9 Evènements indépendants

1.9.1 Définition

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Deux évènements A et B sont dits indépendants si, et seulement si, $P(A \cap B) = P(A) \times P(B)$.

1.9.2 Remarques

- Si $P(B) > 0$, l'indépendance de A et de B équivaut à $P(A|B) = P(A)$.
- L'indépendance est une relation symétrique entre les évènements.
- Si un des deux évènements A ou B a une probabilité nulle alors A et B sont indépendants. En effet par croissance de P, $P(A \cap B) = 0$.
- Il ne faut pas confondre indépendance et incompatibilité de deux évènements.
  Par exemple si on jette une pièce de monnaie équilibrée, $\Omega = \{P, F\}$ et $P(P) = P(F) = \dfrac{1}{2}$.
  Les deux évènements élémentaires P et F sont incompatibles mais pas indépendants :
  $P(P \cap F) = 0$ et $P(P) \times P(F) = \dfrac{1}{4}$.
  Pour un exemple d'évènements qui sont indépendants mais pas incompatibles, cf ci-dessous.
- L'indépendance n'est pas une qualité intrinsèque des évènements. Elle dépend de la probabilité considérée.
  Considérons $\Omega = [\![1; 6]\!]$, P_1 la probabilité uniforme et P_2 définie par :
  $P_2(1) = P_2(2) = \dfrac{1}{6}$, $P_2(3) = \dfrac{1}{3}$ et $P_2(4) = P_2(5) = P_2(6) = \dfrac{1}{9}$.
  Considérons enfin les évènements $A = \{1; 2\}$ et $B = \{2; 3\}$.
  On a $A \cap B = \{2\}$.
  $P_1(A) = P_1(B) = \dfrac{2}{6} = \dfrac{1}{3}$ et $P_1(A \cap B) = \dfrac{1}{6}$.
  A et B ne sont pas indépendants pour P_1.
  Par contre :
  $P_2(A) = \dfrac{1}{6} + \dfrac{1}{6} = \dfrac{1}{3}$
  $P_2(B) = \dfrac{1}{6} + \dfrac{1}{3} = \dfrac{1}{2}$
  $P_2(A \cap B) = \dfrac{1}{6} = \dfrac{1}{3} \times \dfrac{1}{2}$
  A et B sont indépendants pour P_2.

1.9.3 Indépendance et passage au complémentaire

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient A et B deux évènements indépendants.
Alors, les évènements A et $\overline{B}$ d'une part, $\overline{A}$ et B d'autre part, et enfin $\overline{A}$ et $\overline{B}$ sont indépendants.

Démonstration
$A = (A \cap \overline{B}) \cup (A \cap B)$, union disjointe donc :
$P(A \cap \overline{B}) = P(A) - P(A \cap B) = P(A) - P(A)\,P(B)$ car A et B sont indépendants.
$P(A \cap \overline{B}) = P(A)\,(1 - P(B)) = P(A)\,P(\overline{B})$.

A et B jouant des rôles symétriques, on en déduit que $\overline{A}$ et B sont indépendants.

Enfin, remplaçant A par $\overline{A}$, on a d'après le premier cas l'indépendance de $\overline{A}$ et $\overline{B}$.

1.9.4 Famille finie d'évènements mutuellement indépendants

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
On dit que n évènements $A_1, \ldots, A_n$ sont mutuellement indépendants si, et seulement si :

$$\forall I \subset [\![1; n]\!]\ I \neq \emptyset \quad P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$$

Remarques

— Pour $n = 2$, cette notion se confond avec l'indépendance définie en 1.9.1.
Si $n \geq 3$, n évènements mutuellement indépendants sont indépendants 2 à 2 :
$\forall (i, j) \in [\![1; n]\!]^2\ i \neq j \Rightarrow A_i$ et A_j sont indépendants.
(il suffit de prendre $I = \{i; j\}$ dans la définition)
mais la réciproque est fausse comme le montre l'exemple suivant :
On lance deux fois une pièce de monnaie équilibrée :
$\Omega = \{(P, P), (P, F), (F, P), (F, F)\}$ muni de la probabilité uniforme.
On considère les trois évènements suivants :
  — "A" : "le premier lancer a donné pile"
  — "B" : "le deuxième lancer a donné pile"
  — "C" : "les deux lancers ont donné le même résultat"
  — $A = \{(P, P), (P, F)\}$, $P(A) = \dfrac{1}{2}$
  — $B = \{(P, P), (F, P)\}$, $P(B) = \dfrac{1}{2}$
  — $C = \{(P, P), (F, F)\}$, $P(C) = \dfrac{1}{2}$
  — $A \cap B = \{(P, P)\}$, $P(A \cap B) = \dfrac{1}{4}$
  A et B sont indépendants.
  — $A \cap C = \{(P, P)\}$, $P(A \cap C) = \dfrac{1}{4}$
  A et C sont indépendants.
  — $B \cap C = \{(P, P)\}$, $P(B \cap C) = \dfrac{1}{4}$
  B et C sont indépendants.
  — $A \cap B \cap C = \{(P, P)\}$, $P(A \cap B \cap C) = \dfrac{1}{4}$
  A, B et C ne sont pas mutuellement indépendants.

— La définition de l'indépendance mutuelle peut-être formulée ainsi :
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit $(A_i)_{i \in I}$ une famille finie d'évènements.
Ces évènements sont dits mutuellement indépendants si, et seulement si, pour toute partie finie J de I [3], on a :
$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i)$

— Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit $(A_i)_{i \in I}$ une famille finie d'évènements mutuellement indépendants.
Soit J une partie de I.
Les évènements A_i pour i dans J et $\overline{A_i}$ pour $i \in I \setminus J$ sont également mutuellement indépendants.

3. il va de soi qu'on suppose I et J non vides

Il suffit de démontrer que si $A_1, \dots, A_n$ sont mutuellement indépendants alors $A_1, \dots, A_{n-1}$ et $\overline{A_n}$ sont mutuellement indépendants.
Dans le cas général, on passe les évènements au complémentaire un par un, peu importe qu'ils soient en dernière position ou non.
On suppose donc $A_1, \dots, A_n$ mutuellement indépendants et on considère une partie de I de $[\![1;n]\!]$ de la forme $I_1 \cup \{n\}$ (si I ne contient pas n, il n'y a rien à démontrer).
Soit $A = \bigcap\limits_{i \in I_1} A_i$.
De la défintion découle :
$P(A) = \prod\limits_{i \in I_1} P(A_i)$

$P(A \cap A_n) = P\left(\bigcap\limits_{i \in I} A_i\right) = \prod\limits_{i \in I} P(A_i)$
On en déduit $P(A \cap A_n) = P(A)P(A_n)$.
A et A_n sont donc indépendants. Comme vu plus haut, on peut en déduire que A et $\overline{A_n}$ sont indépendants :
$P\left(\left(\bigcap\limits_{i \in I_1} A_i\right) \cap \overline{A_n}\right) = P(A \cap \overline{A_n}) = P(A)P(\overline{A_n}) = \left(\prod\limits_{i \in I_1} P(A_i)\right) P(\overline{A_n})$
et on montre ainsi, en revenant à la définition, que les évènements $A_1, \dots, A_{n-1}$ et $\overline{A_n}$ sont mutuellement indépendants.

— On peut écrire la définition précédente avec I dénombrable et définir ainsi la notion de famille dénombrable d'évènements mutuellement indépendants.
Cette définition ne figure pas au programme.
— **X 2016**
On a deux dés équilibrés : un bleu et un rouge.
On note A l'évènement : "la somme des deux est égale à 9".
Trouver deux évènements relatifs au dé rouge tels que :
— $P(A \cap B \cap C) = P(A)P(B)P(C)$
— $P(A \cap B) \neq P(A)P(B)$
— $P(A \cap C) \neq P(A)P(C)$
— $P(B \cap C) \neq P(B)P(C)$
Correction
L'univers $\Omega = [\![1;6]\!]^2$, le premier terme étant le résultat du dé rouge, le second celui du dé bleu.
$A = \{(6,3);(5,4);(4,5);(3,6)\}$ et $P(A) = \dfrac{4}{36} = \dfrac{1}{9}$
On cherche d'abord ce qu'on peut dire en supposant que B et C existent.
La probabilité de tout évènement est de la forme $\dfrac{N}{36}$ où N est le nombre de cas favorables.
$0 \leq P(A \cap B \cap C) \leq P(A) = \dfrac{4}{36}$ donc $36 \times P(A \cap B \cap C) \in \{0;1;2;3;4\}$.
$P(A \cap B \cap C) = 0$ est impossible :
On aurait $P(A)P(B)P(C) = 0$ donc $P(B)$ ou $P(C) = 0$.
En contexte fini, B ou C est l'évènement impossible. A et $\emptyset$ sont indépendants.
$P(A \cap B \cap C) = P(A)$ est impossible :
On aurait $P(A)P(B)P(C) = P(A)$ donc $P(B) = P(C) = 1$.
En contexte fini, B et C sont l'évènement certain. A et Ω sont indépendants.

On essaye avec $P(A\cap B\cap C)=\dfrac{1}{36}$.
$P(B)P(C)=\dfrac{1}{36P(A)}=\dfrac{1}{4}$
B et C sont de la forme $b\times [\![1;6]\!]$ et $c\times [\![1;6]\!]$ avec b et c inclus dans $[\![1;6]\!]$.
Donc $\dfrac{1}{4}=\dfrac{9}{36}=P(B)P(C)=\dfrac{\text{Card}(b)\times 6\times \text{Card}(c)\times 6}{36\times 36}=\dfrac{\text{Card}(b)\times \text{Card}(c)}{36}$
D'où $\text{Card}(b)\times\text{Card}(c)=9$ avec $\text{Card}(b)$ et $\text{Card}(c)$ des entiers compris entre 0 et 6.
Donc $\text{Card}(b)=\text{Card}(c)=3$.
b et c ne peuvent pas être disjoints : B et C le seraient et on aurait $P(A\cap B\cap C)=0$
On tâtonne un peu et on essaie :
$b=\{1;2;3\}$ et $c=\{3;4;5\}$ (mais d'après l'ordinateur : $b=c=\{1;2;3\}$ fonctionne).
$A\cap B\cap C=\{(3,6)\}$ de probabilité $\dfrac{1}{36}$.
$P(B)=P(C)=\dfrac{3\times 6}{36}=\dfrac{1}{2}$.
On a bien $P(A\cap B\cap C)=P(A)P(B)P(C)$.
$A\cap B=\{(3,6)\}$ de probabilité $\dfrac{1}{36}$.
On a bien $P(A\cap B)\neq P(A)P(B)$.
$A\cap C=\{(3,6);(4,5);(5,4)\}$ de probabilité $\dfrac{3}{36}$.
On a bien $P(A\cap C)\neq P(A)P(C)=\dfrac{2}{36}$.
$B\cap C=\{3\}\times [\![1;6]\!]$ de probabilité $\dfrac{6}{36}$.
On a bien $P(B\cap C)\neq P(B)P(C)=\dfrac{9}{36}$.

Solution avec Python

On représente les parties de $[\![1;6]\!]$ par des listes de 0 et de 1 de longueur 6, 1 marquant l'appartenance.
Compte tenu des indices de liste en Python, l'ensemble $\{2;4\}$ est représenté par la liste `[0,1,0,1,0,0]`.
Le cardinal d'une partie est la somme des éléments de la liste la représentant.
Comme on doit passer en revue toutes les parties de $[\![1;6]\!]$, la fonction `suivant` prend en entrée une liste de longueur 6 représentant une partie de $[\![1;6]\!]$ et renvoie la liste de 0 et de 1 de longueur 6 qui suit l'entrée dans l'ordre lexicographique (sauf lorsque la liste représente $[\![1;6]\!]$ mais la fonction `suivant` n'est pas appelée avec cette entrée).

```
from time import clock

def suivant(E):
    if E[5]==0:
        E[5]=1
    else:
        i=5
        while i>=0 and E[i]==1:
            i-=1
        if i>=0:
            E[i]=1
            for j in range(i+1,6):
                E[j]=0
```

```
debut =clock()
a=[0,0,1,1,1,1]
PA=4#P(A)*36

b=[0]*6
for i in range(2**6):
    #calcul de P(B)*36
    PB=sum(x for x in b)*6
    c=[0]*6
    for j in range(2**6):
        #calcul de P(C)*36
        PC=sum(x for x in c)*6
        #calcul de P(B inter C)*36
        PBC=sum(b[i]*c[i] for i in range(6))*6
        #calcul de P(A inter B)*36
        PAB=sum(a[i]*b[i] for i in range(6))
        #calcul de P(A inter C)*36
        PAC=sum(a[i]*c[i] for i in range(6))
        #calcul de P(A inter B inter C)*36
        PABC=sum(a[i]*b[i]*c[i] for i in range(6))
        cond1=PABC*36*36==PA*PB*PC
        cond2=PAB*36!=PA*PB
        cond3=PAC*36!=PA*PC
        cond4=PBC*36!=PB*PC
        if cond1 and cond2 and cond3 and cond4:
            B=[i+1 for i in range(6) if b[i]==1]
            C=[i+1 for i in range(6) if c[i]==1]
            print(B,C)
        suivant(c)
    suivant(b)

duree=clock()-debut
print(duree)

([4, 5, 6], [1, 2, 6])
([4, 5, 6], [1, 2, 5])
([4, 5, 6], [1, 2, 4])
([3, 5, 6], [1, 2, 6])
([3, 5, 6], [1, 2, 5])
([3, 5, 6], [1, 2, 3])
([3, 4, 6], [1, 2, 6])
([3, 4, 6], [1, 2, 4])
([3, 4, 6], [1, 2, 3])
([3, 4, 5], [1, 2, 5])
([3, 4, 5], [1, 2, 4])
([3, 4, 5], [1, 2, 3])
```

```
([1, 2, 6], [4, 5, 6])
([1, 2, 6], [3, 5, 6])
([1, 2, 6], [3, 4, 6])
([1, 2, 6], [1, 2, 6])
([1, 2, 5], [4, 5, 6])
([1, 2, 5], [3, 5, 6])
([1, 2, 5], [3, 4, 5])
([1, 2, 5], [1, 2, 5])
([1, 2, 4], [4, 5, 6])
([1, 2, 4], [3, 4, 6])
([1, 2, 4], [3, 4, 5])
([1, 2, 4], [1, 2, 4])
([1, 2, 3], [3, 5, 6])
([1, 2, 3], [3, 4, 6])
([1, 2, 3], [3, 4, 5])
([1, 2, 3], [1, 2, 3])

0.064684
```

1.10 Variables aléatoires indépendantes

1.10.1 Couple de variables aléatoires indépendantes

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{A})$.
On dit que les variables aléatoires X et Y sont indépendantes, et on note $X \perp\!\!\!\perp Y$, si, et seulement si, pour toute partie A de $X(\Omega)$ et toute partie B de $Y(\Omega)$, les évènements $(X \in A)$ et $(Y \in B)$ sont indépendants.

1.10.2 Image de deux variables aléatoires par des fonctions

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{A})$.
Soient f une fonction définie sur $X(\Omega)$ et g une fonction définie sur $Y(\Omega)$.
Si X et Y sont indépendantes alors $f(X)$ et $g(Y)$ sont indépendantes.

Démonstration
Soit A une partie de $f(X(\Omega))$.
Soit C l'ensemble des antécédents des éléments de A par f :
$C = \{x \in X(\Omega) \text{ tq } f(x) \in A\}$
L'évènement $(f(X) \in A)$ est identique à l'évènement $(X \in C)$.
Soit B une partie de $g(Y(\Omega))$.
Soit D l'ensemble des antécédents des éléments de B par g :
$D = \{y \in Y(\Omega) \text{ tq } g(y) \in B\}$
L'évènement $(g(Y) \in B)$ est identique à l'évènement $(Y \in D)$.
X et Y sont indépendantes donc les évènements $(X \in C)$ et $(Y \in D)$ sont indépendants. On en déduit que les évènements $(f(X) \in A)$ et $(g(Y) \in B)$ sont indépendants.
Donc les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

1.10.3 Extension au cas de n variables aléatoires

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient $X_1, \dots, X_n$ n variables aléatoires discrètes sur $(\Omega, \mathcal{A})$.
On dit que les variables aléatoires $X_1, \dots, X_n$ sont (mutuellement) indépendantes si, et seulement si, pour tout $(A_1, \dots, A_n) \in \prod_{i=1}^{n} \mathcal{P}(X_i(\Omega))$ les évènements $(X_i \in A_i)$ sont mutuellement indépendants.
Si on s'en tient à la définition, cela signifie que :
$\forall I \subset [\![1;n]\!]\ I \neq \emptyset \quad P\left(\bigcap_{i\in I}(X_i \in A_i)\right) = \prod_{i\in I} P(X_i \in A_i)$ (où $A_i \in \mathcal{P}(X_i(\Omega))$)
et on a nécessairement :
$\forall (A_1, \dots, A_n) \in \prod_{i=1}^{n} \mathcal{P}(X_i(\Omega))\ P\left(\bigcap_{i=1}^{n}(X_i \in A_i)\right) = \prod_{i=1}^{n} P(X_i \in A_i)$
La réciproque est vraie. En effet si I est non vide et strictement contenu dans $[\![1;n]\!]$:

$$\begin{aligned}
P\left(\bigcap_{i\in I}(X_i \in A_i)\right) &= P\left(\left(\bigcap_{i\in I}(X_i \in A_i)\right) \cap \left(\bigcap_{i\in[\![1;n]\!]\setminus I}(X_i \in X_i(\Omega))\right)\right) \\
&= \left(\prod_{i\in I} P(X_i \in A_i)\right) \times \left(\prod_{i\in[\![1;n]\!]\setminus I} P(X_i \in X_i(\Omega))\right) \\
&= \left(\prod_{i\in I} P(X_i \in A_i)\right) \times \left(\prod_{i\in[\![1;n]\!]\setminus I} 1\right) \\
&= \prod_{i\in I} P(X_i \in A_i)
\end{aligned}$$

On en déduit au passage que si I est une partie non vide de $[\![1;n]\!]$ alors les variables aléatoires $X_i, i \in I$ sont indépendantes.

Pour tout $i \in [\![1;n]\!]$, soit f_i une fonction définie sur $X_i(\Omega)$.
Si les variables aléatoires $X_1, \dots, X_n$ sont indépendantes alors il en est de même des variables aléatoires $f_1(X_1), \dots, f_n(X_n)$.
En effet, pour tout $i \in [\![1;n]\!]$ soit A_i une partie de $f_i(X_i(\Omega))$ et B_i l'ensemble des antécédents des éléments de A_i par f_i.
Pour tout $i \in [\![1;n]\!]$, les évènements $(f_i(X_i) \in A_i)$ et $(X_i \in B_i)$ sont identiques.
Les variables aléatoires $X_1, \dots, X_n$ sont indépendantes donc les évènements $(X_i \in B_i)$ sont indépendants.
On en déduit que les évènements $(f_i(X_i) \in A_i)$ sont indépendants.
Donc les variables aléatoires $f_1(X_1), \dots, f_n(X_n)$ sont indépendantes.

Chapitre 2

Les variables aléatoires et leurs lois

2.1 Loi d'une variable aléatoire discrète

2.1.1 Distribution de probabilités d'une variable aléatoire discrète

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, P)$.
On appelle distribution de probabilités de X la famille $(P(X = x))_{x \in X(\Omega)}$.

2.1.2 Proposition

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, P)$.
L'application $P_X \begin{cases} \mathcal{P}(X(\Omega)) \to [0;1] \\ A \mapsto P_X(A) = P(X \in A) \end{cases}$ est une probabilité sur $\mathcal{P}(X(\Omega))$.
On l'appelle loi de X.

Démonstration

- Si $A \in \mathcal{P}(X(\Omega))$, il est clair que $P_X(A) = P(X \in A) \in [0;1]$.
- $P_X(X(\Omega)) = P\left(X \in X(\Omega)\right) = P(\Omega) = 1$
- Soit $(A_n)_{n \in \mathbb{N}}$ une famille de parties deux à deux disjointes de $X(\Omega)$.
  Les évènements $(X \in A_n)$ sont deux à deux disjoints donc :
  $P_X\left(\bigcup_{n=0}^{+\infty} A_n\right) = P\left(X \in \bigcup_{n=0}^{+\infty} A_n\right) = P\left(\bigcup_{n=0}^{+\infty} (X \in A_n)\right) = \sum_{n=0}^{+\infty} P(X \in A_n) = \sum_{n=0}^{+\infty} P_X(A_n)$

Remarque
La loi P_X de X est caractérisée par sa distribution de probabilité.
En effet, si on connaît la loi de X on peut en déduire facilement sa distribution de probabilité :
$\forall x \in X(\Omega)\ P(X = x) = P\left(X \in \{x\}\right) = P_X\left(\{x\}\right)$

Récirpoquement, si on connaît la distribution de probabilité de X, on peut en déduire sa loi.
Soit A une partie de $X(\Omega)$.
$X(\Omega)$ est au plus dénombrable donc A est au plus dénombrable.
$P_X(A) = P(X \in A) = P\left(\bigcup_{a \in A} (X = a)\right) = \sum_{a \in A} P(X = a)$

2.1.3 Notations

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, P)$.
On note $X \sim Y$ lorsque X et Y suivent la même loi ie :

i $X(\Omega) = Y(\Omega)$

ii $\forall a \in X(\Omega) = Y(\Omega)\ P(X = a) = P(Y = a)$

2.1.4 Image de deux variables aléatoires de même loi par une fonction

Soit $(\Omega, \mathcal{A})$ un espace probabilisé et X, Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{A})$.
On suppose que X et Y ont la même loi.
Soit f une fonction définie sur $X(\Omega) = Y(\Omega)$.
On a déjà vu que $f(X)$ et $f(Y)$ sont des variables aléatoires.
On va montrer qu'elles ont la même loi.
Pour cela, il suffit de montrer qu'elles ont la même distribution de probabilités.

Soit $z \in f(X(\Omega)) = f(Y(\Omega))$.
Soit B l'ensemble des antécédents de z par f : $B = \{x \in X(\Omega) = Y(\Omega) \text{ tq } f(x) = z\}$.
B est au plus dénombrable car B est inclus dans $X(\Omega) = Y(\Omega)$ qui est au plus dénombrable.
$(f(X))^{-1}(z) = \bigcup\limits_{x \in B} X^{-1}(x)$ donc par σ-additivité :
$P\left(f(X) = z\right) = \sum\limits_{x \in B} P(X = x)$
$(f(Y))^{-1}(z) = \bigcup\limits_{x \in B} Y^{-1}(x)$ donc par σ-additivité :
$P\left(f(Y) = z\right) = \sum\limits_{x \in B} P(Y = x)$
X et Y ayant la même loi, donc la même distribution de probabilité, $P\left(f(X) = z\right) = P\left(f(Y) = z\right)$

2.2 Indépendance des variables aléatoires et distribution de probabilité

2.2.1 Cas de deux variables aléatoires

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, P)$.
Rappelons que cela revient à dire que (X, Y) est une variable aléatoire discrète.
$X \perp\!\!\!\perp Y \iff \forall (x, y) \in X(\Omega) \cap Y(\Omega)\ P(X = x, Y = y) = P(X = x)P(Y = y)$
où la notation $P(X = x, Y = y)$ désigne la probabilité de l'évènement

$((X, Y) = (x, y)) = (X = x) \cap (Y = y)$

Démonstration
On suppose $X \perp\!\!\!\perp Y$.
Par définition, pour tout $A \subset X(\Omega)$ et tout $B \subset Y(\Omega)$, les évènements $(X \in A)$ et $(Y \in B)$ sont indépendants.
En prenant $A = \{x\}$ et $B = \{y\}$, on a l'indépendance des évènements $(X = x)$ et $(Y = y)$ et la relation $P(X = x, Y = y) = P(X = x)P(Y = y)$

Réciproquement, on suppose :
$\forall (x,y) \in X(\Omega) \cap Y(\Omega)\ P(X=x, Y=y) = P(X=x)P(Y=y)$
Soient $A \subset X(\Omega)$ et $B \subset Y(\Omega)$.
L'évènement $(X \in A) \cap (Y \in B)$ est identique à l'évènement $\displaystyle\bigcup_{(x,y)\in A\times B} (X=x, Y=y)$.

$$\begin{aligned}
P\left((X \in A) \cap (Y \in B)\right) &= \sum_{(x,y)\in A\times B} P(X=x, Y=y) \text{ par incompatibilité} \\
&= \sum_{(x,y)\in A\times B} P(X=x)P(Y=y) \text{ par hypothèse} \\
&= \left(\sum_{x\in A} P(X=x)\right)\left(\sum_{y\in B} P(Y=y)\right) \\
&\quad \text{en appliquant sans problème Fubini puisque tout est positif} \\
&= P(X \in A)P(X \in B) \text{ par incompatibilité}
\end{aligned}$$

2.2.2 Un exemple utile de couple de variables aléatoires indépendantes

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient A et B deux évènements indépendants.
$\mathbf{1}_A$ et $\mathbf{1}_B$ sont deux v.a.r indépendantes.
Rappelons à cet effet que $\overline{A}$ et B d'une part, A et $\overline{B}$ d'autre part et $\overline{A}$ et $\overline{B}$ sont également indépendants.
On calcule alors $P(\mathbf{1}_A = 1, \mathbf{1}_B = 1)$ ainsi que les trois autres et on s'assure que la CNS de la définition est vérifiée.
Je laisse ce soin au lecteur.

2.2.3 Cas de n variables aléatoires

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient $X_1, \dots, X_n$ n variables aléatoires discrètes sur $(\Omega, \mathcal{A}, P)$.
$X_1, \dots, X_n$ indépendantes $\iff \forall (x_1, \dots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)\ P(X_1 = x_1, \dots, X_n = x_n) = \displaystyle\prod_{i=1}^{n} P(X_i = x_i)$

Démonstration
On suppose $X_1, \dots, X_n$ indépendantes.
Par définition, pour tout $(A_1, \dots, A_n) \in \mathcal{P}(X_1(\Omega)) \times \cdots \times \mathcal{P}(X_n(\Omega))$ les évènements $(X_1 \in A_1), \dots, (X_n \in A_n)$ sont indépendants.
En prenant $A = \{x_i\}$, on a l'indépendance des évènements $(X_i = x_i)$ et la relation $P(X_1 = x_1, \dots, X_n = x_n) = \displaystyle\prod_{i=1}^{n} P(X_i = x_i)$
Réciproquement, on suppose :
$\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)\ P(X_1 = x_1, \dots, X_n = x_n) = \displaystyle\prod_{i=1}^{n} P(X_i = x_i)$

Soit $(A_1, \dots, A_n) \in \mathcal{P}(X_1(\Omega)) \times \cdots \times \mathcal{P}(X_n(\Omega))$.

$$\begin{aligned}
P\left(\bigcap_{i=1}^{n}(X_i \in A_i)\right) &= P\left(\bigcup_{(x_1,\dots,x_n)\in A_1\times\cdots\times A_n}(X_1 = x_1, \dots, X_n = x_n)\right)\\
&= \sum_{(x_1,\dots,x_n)\in A_1\times\cdots\times A_n} P(X_1 = x_1, \dots, X_n = x_n) \text{ par incompatibilité}\\
&= \sum_{(x_1,\dots,x_n)\in A_1\times\cdots\times A_n} P(X_1 = x_1)\dots P(X_n = x_n) \text{ par hypothèse}\\
&= \left(\sum_{x_1\in A_1}(P(X_1 = x_1)\right)\left(\sum_{(x_2,\dots,x_n)\in A_2\times\cdots\times A_n} P(X_2 = x_2)\dots P(X_n = x_n)\right)\\
&\quad \text{par Fubini}\\
&= P(X_1 \in A_1) \times \left(\sum_{(x_2,\dots,x_n)\in A_2\times\cdots\times A_n} P(X_2 = x_2)\dots P(X_n = x_n)\right)
\end{aligned}$$

et on itère le procédé.

2.2.4 Lemme des coalitions

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient $X_1, \dots, X_n$ n variables aléatoires discrètes sur $(\Omega, \mathcal{A}, P)$.
Soit $m \in [\![1; n-1]\!]$ (ce qui sous-entend $n \geq 2$).
Soit f une fonction définie sur $X_1(\Omega) \times \cdots \times X_m(\Omega)$ et g une fonction définie sur $X_{m+1}(\Omega) \times \cdots \times X_n(\Omega)$.
Si les variables aléatoires $X_1, \dots, X_n$ sont indépendantes alors les variables aléatoires $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont indépendantes.

Démonstration
On commence par montrer que les variables aléatoires $(X_1, \dots, X_m)$ et $(X_{m+1}, \dots, X_n)$ sont indépendantes.
Pour cela, on utilise la caractérisation de 2.2.1.
Soit $(x_1, \dots, x_m) \in X_1(\Omega) \times \cdots \times X_m(\Omega)$ et $(x_{m+1}, \dots, x_n) \in X_{m+1}(\Omega) \times \cdots \times X_n(\Omega)$.

$$\begin{aligned}
& P\left((X_1, \dots, X_m) = (x_1, \dots, x_m), (X_{m+1}, \dots, X_n) = (x_{m+1}, \dots, x_n)\right)\\
=\ & P\left(X_1 = x_1, \dots, X_n = x_n\right) = P(X_1 = x_1) \times \cdots \times P(X_n = x_n)\\
=\ & \left(\prod_{i=1}^{m} P(X_i = x_i)\right) \times \left(\prod_{i=m+1}^{n} P(X_i = x_i)\right)
\end{aligned}$$

Mais les variables aléatoires $X_1, \dots, X_m$ sont mutuellement indépendantes donc :
$\displaystyle\prod_{i=1}^{m} P(X_i = x_i) = P\left(\bigcap_{i=1}^{m}(X_i = x_i)\right) = P\left((X_1, \dots, X_m) = (x_1, \dots, x_m)\right)$
On a aussi : $\displaystyle P\left((X_{m+1}, \dots, X_n) = (x_{m+1}, \dots, x_n)\right) = \prod_{i=m+1}^{n} P(X_i = x_i)$
Donc :
$P\left((X_1, \dots, X_m) = (x_1, \dots, x_m), (X_{m+1}, \dots, X_n) = (x_{m+1}, \dots, x_n)\right) = P\left((X_1, \dots, X_m) = (x_1, \dots, x_m)\right) \times P\left((X_{m+1}, \dots, X_n) = (x_{m+1}, \dots, x_n)\right)$

et les variables aléatoires $(X_1, \dots, X_m)$ et $(X_{m+1}, \dots, X_n)$ sont indépendantes.
On en déduit que les variables aléatoires $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont indépendantes.

2.2.5 Extension au cas d'un nombre quelconque de coalitions

Soient $d_1, \dots, d_p \in \mathbb{N}^*$ et $n = \displaystyle\sum_{i=1}^{p} d_i$.
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient $X_1, \dots, X_n$ n variables aléatoires discrètes sur $(\Omega, \mathcal{A}, P)$.
Soit f_1 une fonction définie sur $X_1(\Omega) \times \cdots \times X_{d_1}(\Omega)$.
Soit f_2 une fonction définie sur $X_{d_1+1}(\Omega) \times \cdots \times X_{d_1+d_2}(\Omega)$.
$\vdots$
Soit f_p une fonction définie sur $X_{d_1+\cdots+d_{p-1}+1}(\Omega) \times \cdots \times X_n(\Omega)$.
Si les variables aléatoires $X_1, \dots, X_n$ sont indépendantes alors les variables aléatoires $f_1(X_1, \dots, X_{d_1})$, $f_2(X_{d_1+1}, \dots, X_{d_1+d_2}), \dots, f_p(X_{d_1+\cdots+d_{p-1}+1}, \dots, X_n)$ sont indépendantes.

Démonstration
On raisonne par récurrence sur p.
La propriété est vraie pour $p = 2$.
Supposons la vraie pour $p - 1$.
On se place alors dans les hypothèses du théorème.
D'après le cas de deux variables aléatoires, les variables aléatoires $\Big(f_1(X_1, \dots, X_{d_1}), f_2(X_{d_1+1}, \dots, X_{d_1+d_2}), \dots, f_{p-1}(X_{d_1+\cdots+d_{p-2}+1}, \dots, X_{d_1+\cdots+d_{p-1}})\Big)$ et $f_p(X_{d_1+\cdots+d_{p-1}+1}, \dots, X_n)$ sont indépendantes.
D'après l'hypothèse de récurrence, les variables aléatoires $f_1(X_1, \dots, X_{d_1}), \dots, f_{p-1}(X_{d_1+\cdots+d_{p-2}+1}, \dots, X_{d_1+\cdots+d_{p-1}})$ sont indépendantes.
Soit $A_1 \subset f_1\left(X_1(\Omega) \times \cdots \times X_{d_1}(\Omega)\right)$.
$\vdots$
Soit $A_p \subset f_p\left(X_{d_1+\cdots+d_{p-1}+1}(\Omega) \times \cdots \times X_n(\Omega)\right)$.

$$
\begin{aligned}
& P\Big((f_1(X_1, \dots, X_{d_1}) \in A_1) \cap \cdots \cap \Big(f_p(X_{d_1+\cdots+d_{p-1}+1}, \dots, X_n) \in A_p\Big)\Big) \\
= \; & P\Big((f_1(X_1, \dots, X_{d_1}) \in A_1) \cap \cdots \cap \Big(f_{p-1}(X_{d_1+\cdots+d_{p-2}+1}, \dots, X_{d_1+\cdots+d_{p-1}}) \in A_p\Big)\Big) \\
& \times P\Big(f_p(X_{d_1+\cdots+d_{p-1}+1}, \dots, X_n) \in A_p\Big) \text{ par indépendance} \\
= \; & P\left(f_1(X_1, \dots, X_{d_1}) \in A_1\right) \times \cdots \times P\Big(f_{p-1}(X_{d_1+\cdots+d_{p-2}+1}, \dots, X_{d_1+\cdots+d_{p-1}}) \in A_p\Big) \\
& \times P\Big(f_p(X_{d_1+\cdots+d_{p-1}+1}, \dots, X_n) \in A_p\Big) \text{ par indépendance, en utilisant l'hypothèse de récurrence}
\end{aligned}
$$

2.2.6 Suites de variables aléatoires indépendantes

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires discrètes sur $(\Omega, \mathcal{A})$.
On dit que les variables aléatoires $(X_n)_{n\in\mathbb{N}}$ sont indépendantes si, et seulement si, pour toute partie finie $I = \{i_1, \dots, i_p\}$ de $\mathbb{N}$, les variables aléatoires $X_{i_1}, \dots, X_{i_p}$ sont indépendantes.
On dit que les variables aléatoires $(X_n)_{n\in\mathbb{N}}$ sont indépendantes et équidistribuées, en abrégé i.i.d si, et seulement si, elles sont indépendantes et de même loi.
Par exemple le jeu de pile ou face infini avec probabilité p de faire pile, ie la répétition indéfinie

du lancer d'une pièce ayant la probabilité p de tomber sur pile, est modélisé par une suite i.i.d de variables de Bernoulli de paramètre p (où on code pile par 1 et face par 0).

2.3 Loi géométrique

2.3.1 Introduction

On considère une pièce de monnaie ayant une probabilité p de donner pile. On la lance jusqu'à obtenir pile.
Combien de fois la lance-t-on ?
Ce nombre est aléatoire, que peut-on en dire ?
Plus généralement dans une série d'épreuves de Bernoulli (ie avec deux résultats possibles : "Succès" avec la probabilité p et "Echec" avec la probabilité $1-p$), combien faut-il faire de tentatives pour obtenir un premier succès ?

2.3.2 Temps d'attente du premier succès

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $(X_n)_{n\in\mathbb{N}^*}$ une suite i.i.d de variables aléatoires de Bernoulli de paramètre $p \in]0;1[$.
Soit $T = \begin{cases} \min\left(\{n \in \mathbb{N}^* \text{ tq } X_n = 1\}\right) \text{ si } \{n \in \mathbb{N}^* \text{ tq } X_n = 1\} \neq \emptyset \\ +\infty \text{ si } \{n \in \mathbb{N}^* \text{ tq } X_n = 1\} = \emptyset \end{cases}$
On a déjà vu que T est une variable aléatoire.
Déterminons sa loi de probabilité.
$P(T=1) = P(X_1 = 1) = p$
Pour tout $n \in [\![2; +\infty[\![$:

$$\begin{aligned} P(T=n) &= P(X_1 = 0, \dots, X_{n-1} = 0, X_n = 1) \\ &= P(X_1 = 0) \times \cdots \times P(X_{n-1} = 0) \times P(X_n = 1) \text{ par indépendance} \\ &= (1-p)^{n-1} p \end{aligned}$$

On remarquera que cette formule est valable pour $n = 1$.
On peut calculer $P(T = +\infty)$ de deux façons :

$$\begin{aligned} P(T = +\infty) &= 1 - \sum_{n=1}^{+\infty} p(1-p)^{n-1} \\ &= 1 - p\sum_{n=0}^{+\infty}(1-p)^n = 1 - \frac{p}{1-(1-p)} \\ &= 0 \end{aligned}$$

ou bien :

$$\begin{aligned} P(T = +\infty) &= P\left(\bigcap_{n=1}^{+\infty}(X_n = 0)\right) \\ &= \lim_{N\to+\infty} P\left(\bigcap_{n=1}^{N}(X_n = 0)\right) \\ &= \lim_{N\to+\infty} \prod_{n=1}^{N} P(X_n = 0) \text{ par indépendance} \\ &= \lim_{N\to+\infty} (1-p)^N = 0 \end{aligned}$$

2.3.3 Définition

Soit $p \in]0;1[$.
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, P)$.
On dit que X suit la loi géométrique de paramètre p si, et seulement si :

i $X(\Omega) = \mathbb{N}^*$

ii $\forall n \in \mathbb{N}^*$ $P(X = n) = p(1-p)^{n-1}$

On note alors : $X \sim \mathcal{G}(p)$.

2.3.4 Relation $P(X > k) = (1-p)^k$

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, P)$ qui suit la loi géométrique de paramètre p.
Pour tout $k \in \mathbb{N}$, $P(X > k) = (1-p)^k$

Deux démonstrations sont possibles :

— **Par le calcul**

$$\begin{aligned}
P(X > k) &= P\left(\bigcup_{l=k+1}^{+\infty}(X = l)\right) = \sum_{l=k+1}^{+\infty} P(X = l) \\
&= \sum_{l=k+1}^{+\infty} p(1-p)^{l-1} = \sum_{n=0}^{+\infty} p(1-p)^{k+n} \\
&= p(1-p)^k \sum_{n=0}^{+\infty}(1-p)^n = p(1-p)^k \frac{1}{1-(1-p)} \\
&= (1-p)^k
\end{aligned}$$

— **Par l'interprétation en terme de temps d'attente**

Si on reprend les notations de 2.3.2, l'évènement $(T > k)$ est l'évènement $\bigcap_{n=1}^{k}(X_n = 0)$.

Donc par indépendance, $P(T > k) = \prod_{n=1}^{k} P(X_n = 0) = (1-p)^k$

2.4 Des exemples

2.4.1 Mines 2018

On lance indéfiniment un dé équilibré.

1. Soit A_n l'événement "aucun 6 n'a été obtenu lors des n premiers lancers". Déterminer $P(A_n)$.
2. Soit F_k l'événement "le premier 6 est obtenu au k-ième lancer". Déterminer $P(F_k)$.
3. Soit K l'événement "6 n'apparaît jamais". Exprimer K à l'aide des A_n. En déduire $P(K)$.
4. Exprimer $\overline{K}$ en fonction des F_k. Retrouver la valeur de $P(K)$.
5. Soient G l'événement "6 apparaît une infinité de fois" et H l'événement "6 apparaît à tous les lancers sauf un nombre fini d'entre eux". Calculer $P(G)$ et $P(H)$.

Correction
On modélise le lancer indéfini d'un dé équilibré par un espace probabilisé $(\Omega, \mathcal{A}, P)$ sur lequel existe une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires mutuellement indépendantes qui suivent toutes la loi uniforme sur $[\![1;6]\!]$.

1.

$$\begin{aligned} P(A_n) &= P\left(\bigcap_{i=1}^{n}(X_i \neq 6)\right) \\ &= \prod_{i=1}^{n} P(X_i \neq 6) \text{ par indépendance} \\ &= \left(\frac{5}{6}\right)^n \end{aligned}$$

2. $P(F_k) = P\left(\left(\bigcap_{i=1}^{k-1}(X_i \neq 6)\right) \cap (X_k = 6)\right) = P(X_k = 6)\prod_{i=1}^{k-1} P(X_i \neq 6)$ par indépendance

   Donc $P(F_k) = \left(\frac{5}{6}\right)^{k-1}\frac{1}{6}$.

3. $K = \bigcap_{n=1}^{+\infty} A_n$ avec $A_{n+1} \subset A_n$

   Par continuité décroissante, $P(K) = \lim_{n\to+\infty} P(A_n) = 0$

4. $\overline{K} = \bigcup_{k=1}^{+\infty} F_k$ où les F_k sont deux à deux incompatibles. Donc :

   $P(\overline{K}) = \sum_{k=1}^{+\infty} P(F_k) = \sum_{k=1}^{+\infty}\left(\frac{5}{6}\right)^{k-1}\frac{1}{6} = \frac{1}{6}\frac{1}{1-5/6} = 1$

   On retrouve $P(K) = 0$.

5. $G = \bigcap_{n\geq 1}\bigcup_{i>n}$(on obtient 6 au $i^{\text{ième}}$ lancer).

   On note C_n l'évènement : $\bigcup_{i>n}$(on obtient 6 au $i^{\text{ième}}$ lancer).

   $C_{n+1} \subset C_n$.

   Par continuité décroissante, $P(G) = \lim_{n\to+\infty} P(C_n)$.

Cherchons $P(C_n)$.

$$\begin{aligned}\forall n \in \mathbb{N}^* \ P(C_n) &= 1 - P(\overline{C_n}) = 1 - P\left(\bigcap_{i=n+1}^{+\infty}(X_i \neq 6)\right) \\ &= 1 - \lim_{N \to +\infty} P\left(\bigcap_{i=n+1}^{N}(X_i \neq 6)\right) \\ &= 1 - \lim_{N \to +\infty} \prod_{i=n+1}^{N} P(X_i \neq 6) \text{ par indépendance} \\ &= 1 - \lim_{N \to +\infty} \prod_{i=n+1}^{N} \frac{5}{6} \\ &= 1 - \lim_{N \to +\infty} \left(\frac{5}{6}\right)^{N-n} \\ &= 1 - 0 = 1\end{aligned}$$

On en déduit que $P(G) = 1$.

$H = \displaystyle\bigcup_{n \geq 1} \bigcap_{i > n}$(on obtient 6 au $i^{\text{ième}}$ lancer).

On note D_n l'évènement : $\displaystyle\bigcap_{i>n}$(on obtient 6 au $i^{\text{ième}}$ lancer) $= \displaystyle\bigcap_{i>n}(X_i = 6)$.

$D_n \subset D_{n+1}$:

$D_n = \displaystyle\bigcap_{i>n}(X_i = 6) = (X_{n+1} = 6) \cap \left(\bigcap_{i>n+1}(X_i = 6)\right) = (X_{n+1} = 6) \cap D_{n+1} \subset D_{n+1}$.

Par continuité croissante, $P(H) = \displaystyle\lim_{n \to +\infty} P(D_n)$.

Cherchons $P(D_n)$.

$$\begin{aligned}\forall n \in \mathbb{N}^* \ P(D_n) &= P\left(\bigcap_{i=n+1}^{+\infty}(X_i = 6)\right) \\ &= \lim_{N \to +\infty} P\left(\bigcap_{i=n+1}^{N}(X_i = 6)\right) \\ &= \lim_{N \to +\infty} \prod_{i=n+1}^{N} P(X_i = 6) \text{ par indépendance} \\ &= \lim_{N \to +\infty} \prod_{i=n+1}^{N} \frac{1}{6} \\ &= \lim_{N \to +\infty} \left(\frac{1}{6}\right)^{N-n} \\ &= 0\end{aligned}$$

On en déduit que $P(H) = 0$.

2.4.2 Exemple 2

On dispose d'une urne contenant au départ une boule blanche et on joue indéfiniment à pile ou face avec une pièce parfaite. A chaque fois qu'on obtient face, on ajoute une boule noire au contenu de l'urne et la première fois qu'on obtient pile, on tire au hasard une boule de l'urne. Quelle est la probabilité d'obtenir une boule blanche ?

On note T le rang d'obtention du premier pile.
T est une variable aléatoire qui suit la loi géométrique de paramètre $\frac{1}{2}$. On ne prend pas la peine d'expliciter l'espace probabilisé sous-jacent.
Par la formule des probabilités totales, la probabilité de l'évènement cherché, qu'on note E est :

$$\begin{aligned} P(E) &= \sum_{n=1}^{+\infty} P(E \mid T=n)P(T=n) = \sum_{n=1}^{+\infty} \frac{1}{2}\left(1-\frac{1}{2}\right)^{n-1}\frac{1}{n} \\ &= \sum_{n=1}^{+\infty} \frac{1}{n}\left(\frac{1}{2}\right)^{n} = \ln(2) \simeq 0,69 \end{aligned}$$

2.4.3 X 2016

On imagine un jeu télévisé où les candidats doivent franchir des murs successifs. Le candidat perd s'il ne franchit pas un mur, continue sinon.
La probabilité que le candidat passe le $k^{\text{ième}}$ mur est de $\frac{1}{k}$.
On note X la variable aléatoire égale au nombre de murs franchis par le candidat.

1. Donner la loi de X.
2. Donner l'espérance, puis la variance de X (si elles ont un sens).

On trouve : $E(X) = \mathrm{e} - 1$, $E(X^2) = \mathrm{e} + 1$, $V(X) = 3\,\mathrm{e} - \mathrm{e}^2$.

Correction
Ce type d'exercice pose un problème de modélisation. La modélisation classique consiste à considérer $(\Omega, \mathcal{A}, P)$ un espace probabilisé sur lequel existe une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes, X_n suivant la loi de Bernoulli de paramètres $\frac{1}{n}$.
Les néophytes pensent à une autre modélisation :
$\Omega = \left(\bigcup_{k=1}^{+\infty}\{(1,1,\dots\ k \text{ fois } 1,1,0)\}\right) \cup \{(1,1\dots,1;\dots)\}$ (on observera que le premier mur est toujours franchi). Cette fois les épreuves successives ne sont pas indépendantes : l'existence d'une $k^{\text{ième}}$ épreuve dépend des précédentes. Une phrase comme "La probabilité que le candidat passe le $k^{\text{ième}}$ mur est de $\frac{1}{k}$" devient problématique. Il s'agit plutôt d'une probabilité conditionnelle : la probabilité de franchir le $k^{\text{ième}}$ mur sachant qu'on a franchi les $k-1$ premiers est $\frac{1}{k}$. Rappelons que l'indépendance n'est pas une propriété intrinsèque des évènements mais une propriété des probabilités.
L'utilisation de l'indépendance dans le premier cas, de la formule des probabilités composées dans le second donne :
$\forall k \in \mathbb{N}^*\ P(X \geq k) = 1 \times \frac{1}{2} \times \dots \frac{1}{k} = \frac{1}{k!}$.
Par continuité décroissante :
$P(X = +\infty) = 0$.
En toute rigueur $X(\Omega) = \mathbb{N}^* \cup \{+\infty\}$ mais $P(X = +\infty) = 0$ donc on considérera que $X(\Omega) = \mathbb{N}^*$.
$\forall k \geq 1\ P(X = k) = P(X \geq k) - P(X \geq k+1) = \frac{1}{k!} - \frac{1}{(k+1)!} = \frac{k}{(k+1)!}$

2.4.4 Exemple 3

A et B jouent au jeu suivant :
A commence à jouer. Il lance deux dés cubiques parfaits.
Si la somme des points obtenus est 6 alors A gagne sinon B lance les deux dés et gagne si la somme des points obtenus est 7.
Sinon A lance de nouveau les deux dés et ainsi de suite.
Préférez-vous être A ou B ?

Il faut commencer par déterminer la probabilité que la somme des points obtenus soit 6 ou 7.
On modélise le lancer de deux dés ainsi :
— $\Omega_1 = [\![1;6]\!]^2$
— P_1 est la probabilité uniforme.
L'évènement : "la somme des points obtenus est 6" est :
$(S=6) = \{(5,1);(4,2);(3,3);(2,4);(1,5)\}$
Sa probabilité est $\dfrac{5}{36} \simeq 0,139$.
L'évènement : "la somme des points obtenus est 7" est :
$(S=7) = \{(6,1);(5,2);(4,3);(3,4);(2,5);(1,6)\}$
Sa probabilité est $\dfrac{6}{36} = \dfrac{1}{6} \simeq 0,167$

$P(S=6) < P(S=7)$ mais c'est A qui commence.

Passons au jeu proprement dit.
On le modélise par un espace probabilisé $(\Omega, \mathcal{A}, P)$ sur lequel existe une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes telles que $X_n \sim \mathcal{B}\left(\dfrac{5}{36}\right)$ si n est impair et $X_n \sim \mathcal{B}\left(\dfrac{1}{6}\right)$ si n est pair.
$(A \text{ gagne}) = \bigcup\limits_{p=0}^{+\infty} ((X_1=0)\cap(X_2=0)\cap\cdots\cap(X_{2p-1}=0)\cap(X_{2p}=0)\cap(X_{2p+1}=1))$, union disjointe d'intersections d'évènements indépendants.
Donc :

$$\begin{aligned} P(A \text{ gagne}) &= \sum_{p=0}^{+\infty} (P(X_1=0)\times\cdots\times P(X_{2n}=0)\times P(X_{2n+1}=1)) \\ &= \sum_{p=0}^{+\infty} \left(1-\frac{5}{36}\right)^p \left(1-\frac{6}{36}\right)^p \frac{5}{36} \end{aligned}$$

$\left(1-\dfrac{5}{36}\right)\left(1-\dfrac{6}{36}\right) = \dfrac{31\times 30}{36^2} = \dfrac{155}{216}$

$$\begin{aligned} P(A \text{ gagne}) &= \frac{5}{36}\frac{1}{1-\dfrac{155}{216}} = \frac{5}{36}\frac{216 = 6\times 36}{216-155} = \frac{5\times 6}{61} \\ &= \frac{30}{61} \simeq 0,492 \end{aligned}$$

Il n'est pas a priori clair que $P(B \text{ gagne}) = 1 - P(A \text{ gagne})$: l'évènement "le jeu ne s'arrête pas" est :

$E = (X_1 = 0) \cap (X_2 = 0) \cap \cdots \cap (X_{2p-1} = 0) \cap (X_{2p} = 0) \cap \cdots \neq \emptyset$.
Néanmoins :

$$\begin{aligned}
P(E) &= P\left(\bigcap_{k=1}^{+\infty} ((X_{2k-1} = 0) \cap (X_{2k} = 0))\right) \\
&= \lim_{K \to +\infty} \left(P\left(\bigcap_{k=1}^{K} ((X_{2k-1} = 0) \cap (X_{2k} = 0))\right)\right) \\
&= \lim_{K \to +\infty} \left(\prod_{k=1}^{K} P((X_{2k-1} = 0) \cap (X_{2k} = 0))\right) \\
&\quad \text{par indépendance des variables aléatoires } (X_{2k-1}, X_{2k}) \text{ (avec le lemme des coalitions)} \\
&= \lim_{K \to +\infty} \left(\left(1 - \frac{5}{36}\right) \times \left(1 - \frac{1}{6}\right)\right)^K \\
&= 0
\end{aligned}$$

Remarques
D'autres méthodes sont possibles.

— A gagne s'il fait 6 avant que B ne fasse 7 donc $P(A \text{ gagne}) = P(X_1 \leq X_2)$ où X_1 et X_2 sont des va indépendantes, $X_1 \sim \mathcal{G}\left(\frac{5}{36}\right)$ et $X_2 \sim \mathcal{G}\left(\frac{6}{36}\right)$.

$$\begin{aligned}
P(A \text{ gagne}) &= \sum_{k=1}^{+\infty} P((X_1 \leq X_2) \cap (X_1 = k)) \\
&\quad \text{formule des probabilités totales avec le SCE } (X_1 = k)_{k \in \mathbb{N}^*} \\
&= \sum_{k=1}^{+\infty} P(k \leq X_2) \cap (X_1 = k)) \\
&= \sum_{k=1}^{+\infty} P(k \leq X_2) P(X_1 = k) \text{ par indépendance} \\
&= \sum_{k=0}^{+\infty} P(k < X_2) P(X_1 = k+1) \text{ par indépendance} \\
&= \sum_{k=0}^{+\infty} \left(1 - \frac{1}{6}\right)^k \frac{5}{36} \left(1 - \frac{5}{36}\right)^k \\
&= \frac{5}{36} \times \frac{1}{1 - 155/216} = \frac{5 \times 216}{36 \times 61} = \frac{5 \times 6 \times 36}{36 \times 61} \\
&= \frac{30}{61}
\end{aligned}$$

$P(B \text{ gagne}) = 1 - P(A \text{ gagne}) = \frac{31}{61}$
Ici cette formule est correcte car on a d'emblée négligé le cas où le jeu ne se terminerait pas en prenant X_1 et X_2 à valeurs dans $\mathbb{N}^*$ et non dans $\mathbb{N}^* \cup \{+\infty\}$.
Evidemment cette opération est légitime : on a négligé un évènement de probabilité nulle, donc indiscernable de l'évènement impossible du point de vue des probabilités.

— On note E_1 l'évènement : "on obtient 6 au premier lancer".
On note E_2 l'évènement : "on n'obtient pas 6 au premier lancer et on obtient 7 au second".

On note E_3 l'évènement : "on n'obtient pas 6 au premier lancer et on n'obtient pas 7 au second".
(E_1, E_2, E_3) est un système complet d'évènements.
On note AV l'évènement : "le joueur A est déclaré vainqueur" et BV l'évènement : "le joueur B est déclaré vainqueur".

$$\begin{aligned}
P(AV) &= P(AV|E_1)\,P(E_1) + P(AV|E_2)\,P(E_2) + P(AV|E_3)\,P(E_3) \\
&= 1 \times \frac{5}{36} + 0 \times P(E_2) + \left(1 - \frac{5}{36}\right)\left(1 - \frac{6}{36}\right) P(AV) \\
&= \frac{5}{36} + \frac{155}{216} P(AV) \\
\text{Donc : } P(AV) &= \frac{30}{61}
\end{aligned}$$

De même :

$$\begin{aligned}
P(BV) &= P(BV|E_1)\,P(E_1) + P(BV|E_2)\,P(E_2) + P(BV|E_3)\,P(E_3) \\
&= 0 \times P(E_1) + 1 \times \left(1 - \frac{5}{36}\right)\frac{6}{36} + \left(1 - \frac{5}{36}\right)\left(1 - \frac{6}{36}\right) P(BV) \\
\text{Donc : } P(BV) &= \frac{31}{61}
\end{aligned}$$

$P(AV) + P(BV) = 1$ donc le jeu se termine presque sûrement.

Remarque
On peut justifier rigoureusement que $P(AV|E_3) = P(AV)$, mais la méthode perd de son intérêt :

$$\begin{aligned}
\forall p \geq 1 \; P(A \text{ gagne après } 2p+1 \text{ coups } |E_3) &= \frac{P(E_3 \text{ et } A \text{ gagne après } 2p+1 \text{ coups})}{P(E_3)} \\
&= \frac{P(X_1 = 0, X_2 = 0, \ldots, X_{2p} = 0, X_{2p+1} = 1)}{P(X_1 = 0, X_2 = 0)} \\
&= \frac{P(X_1 = 0)P(X_2 = 0)P(X_3 = 0)\ldots P(X_{2p} = 0)P(X_{2p+1} = 1)}{P(X_1 = 0)P(X_2 = 0)} \\
&\quad \text{par indépendance} \\
&= P(X_3 = 0)\ldots P(X_{2p} = 0)P(X_{2p+1} = 1) \\
&= P(X_1 = 0)\ldots P(X_{2p-2} = 0)P(X_{2p-1} = 1) \\
&\quad \text{car } X_i \text{ et } X_{i-2} \text{ ont la même loi} \\
&= P(X_1 = 0, X_2 = 0, \ldots, X_{2p-2} = 0, X_{2p-1} = 1) \\
&= P(A \text{ gagne après } 2p-1 \text{ coups})
\end{aligned}$$

On alors :

$$\begin{aligned} P(A \text{ gagne}|E_3) &= P\left(\bigcup_{p=0}^{+\infty} A \text{ gagne après } 2p+1 \text{ coups } |E_3\right) \\ &= \sum_{p=0}^{+\infty} P(A \text{ gagne après } 2p+1 \text{ coups } |E_3) \\ &\quad \text{par incompatibilité et parce que } P(.|E_3) \text{ est une probabilité} \\ &= P(A \text{ gagne au premier coup}|E_3) + \sum_{p=1}^{+\infty} P(A \text{ gagne après } 2p-1 \text{ coups}) \\ &= 0 + P(A \text{ gagne}) = P(A \text{ gagne}) \end{aligned}$$

2.4.5 Exemple 4

A et B participent à un jeu qui consiste en une succession de parties indépendantes.
Chaque partie est gagnée par l'un des deux joueurs, A avec la probabilité p, B avec la probabilité $q = 1 - p$ ($p \in]0;1[$).
Le jeu se termine dès qu'un joueur a gagné deux parties de plus que l'autre, ce joueur étant déclaré vainqueur.
Quelle est la probabilité que A gagne ?

On modélise la succession de parties indépendantes par une suite de variables aléatoires i.i.d $(X_n)_{n\in\mathbb{N}^*}$ sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, les X_n suivant la loi de Bernoulli de paramètre p.
Après n parties, A en a gagné $S_n = \sum_{k=1}^{n} X_k$ et B $n - S_n$.
La différence de parties gagnées entre les deux joueurs est $|2S_n - n|$.
L'évènement "A gagne à la n-ième partie" s'écrit :
$(|2S_1 - 1| \neq 2) \cap (|2S_2 - 1| \neq 2) \cap \cdots \cap (|2S_{n-1} - (n-1)| \neq 2) \cap (2S_n - n = 2)$
$2S_k - k$ a la parité de k donc ne peut valoir 2 si k est impair.
Donc A ne peut gagner qu'à l'issue d'une partie de rang pair.
On note AV_{2n} l'évènement : "le joueur A est déclaré vainqueur à l'issue de la $(2n)^{\text{ième}}$ partie".
On note AV l'évènement : "le joueur A est déclaré vainqueur" : $AV = \bigcup_{n=1}^{+\infty} AV_{2n}$.
Examinons AV_{2n}.
$AV_{2n} = (|2S_2 - 2| \neq 2) \cap \cdots \cap (|2S_{2n-2} - (2n-2)| \neq 2) \cap (2S_{2n} - 2n = 2)$
ou encore :
$(|S_2 - 1| \neq 1) \cap (|S_4 - 2| \neq 1) \cap \cdots \cap (|S_{2n-2} - (n-1)| \neq 1) \cap (S_{2n} - n = 1)$
Supposons cet évènement réalisé.
$S_2 = X_1 + X_2 \neq 0, 2$ avec X_1 et $X_2 \in \{0; 1\}$ donc $(X_1, X_2) = (0, 1)$ ou $(1, 0)$ et $S_2 = 1$.
$S_4 = S_2 + X_3 + X_4 = 1 + X_3 + X_4 \neq 1, 3$ donc $X_3 + X_4 \neq 0, 2$.
On en déduit $(X_3, X_4) = (0, 1)$ ou $(1, 0)$ et $S_2 = 2$ et ainsi de suite donc :
$AV_{2n} \subset \left(\bigcap_{k=1}^{n-1} ((X_{2k-1}, X_{2k}) \in \{(0, 1), (1, 0)\})\right) \cap ((X_{2n-1}, X_{2n}) = (1, 1))$
On montre facilement l'inclusion inverse.
En d'autres termes, au cours des deux premières parties, A en a gagné une et B une (sinon AV_2 ou $\overline{AV_2}$ est réalisé). Et ainsi de suite par tranche de deux.

$$P(AV) = p^2 \sum_{n=1}^{+\infty} (2pq)^{n-1} = \frac{p^2}{1-2pq} = \frac{p^2}{1-2p+2p^2}$$

De même $P(BV) = \dfrac{q^2}{1-2pq}$.

$p^2 + q^2 = (p+q)^2 - 2pq = 1 - 2pq$ donc $P(AV) + P(BV) = 1$: le jeu se termine ps.

Remarque

Une autre méthode est possible :

$(X_1, X_2) = (1,1)$, noté $(X_1, X_2) = (1,0)$, $(X_1, X_2) = (0,1)$ et $(X_1, X_2) = (0,0)$ forment un système complet d'évènements.

$$\begin{aligned} P(AV) &= P(AV|(X_1,X_2)=(1,1))\, P(X_1=1, X_2=1) + P(AV|(X_1,X_2)=(1,0))\, P(X_1=1, X_2=0) \\ &\quad + P(AV|(X_1,X_2)=(0,1))\, P(X_1=0, X_2=1) + P(AV|(X_1,X_2)=(0,0))\, P(X_1=0, X_2=0) \\ &= 1 \times P(X_1=1, X_2=1) + P(AV)\, P(X_1=1, X_2=0) + P(AV)\, P(X_1=0, X_2=1) \\ &\quad + 0 \times P(X_1=0, X_2=0) \\ &= p^2 + 2\,p\,q\,P(AV) \end{aligned}$$

D'où : $P(AV) = \dfrac{p^2}{1 - 2\,p\,q}$

2.4.6 Loi de Poisson

— **Introduction**

Un compteur Geiger est un appareil de mesure permettant de compter le nombre de particules α (entre autres) venant frapper un capteur.

La physique du phénomène permet de faire les hypothèses suivantes :

— Le nombre *moyen* de particules détectées entre les instants t_1 et t_2 est proportionnel à $t_2 - t_1$.

— Ce nombre est variable, c'est en fait une variable aléatoire d'espérance $a \times (t_2 - t_1)$ où a est une constante. La loi de cette variable aléatoire ne dépend que de $t_2 - t_1$ (et non de t_1 ou t_2).

— Les nombres de particules détectées pendant des intervalles de temps deux à deux disjoints sont des variables aléatoires (mutuellement) indépendantes.

Il s'agit alors de déterminer la loi du nombre de particules détectées pendant un intervalle de longueur T, disons $[0;T]$ pour simplifier.

A cet effet on découpe $[0;T]$ en n intervalles [1] de longueur $\Delta t = \dfrac{T}{n}$.

Le nombre moyen de particules détectées pendant un de ces n intervalles est $a\Delta t = a\dfrac{T}{n}$.

Si Δt est très petit, la probabilité que deux particules soient détectées peut-être considérée comme négligeable. Le nombre de particules détectées suit donc une loi de Bernoulli de paramètre $a\Delta t = a\dfrac{T}{n}$.

Le nombre de particules détectées pendant $[0;T]$ suit donc une loi binomiale de paramètres n et $a\dfrac{T}{n}$:

$$P(X=k) = \binom{n}{k} \left(a\frac{T}{n}\right)^k \left(1 - a\frac{T}{n}\right)^{n-k} \xrightarrow[n \to +\infty]{} \frac{(aT)^k}{k!}\,\mathrm{e}^{-aT}$$

1. en toute rigueur ces intervalles doivent être semi-ouverts ie de la forme $[t_k; t_{k+1}[$

Démonstration
$\binom{n}{k} = \dfrac{n!}{k!(n-k)!} = \dfrac{n(n-1)\dots(n-k+1)}{k!} \underset{n\to+\infty}{\sim} \dfrac{n^k}{k!}$
$\ln\left(\left(1-\dfrac{aT}{n}\right)^{n-k}\right) = (n-k)\ln\left(1-\dfrac{aT}{n}\right) \sim n\,\dfrac{-aT}{n} = -aT$
Donc $\left(1-a\dfrac{T}{n}\right)^{n-k} \xrightarrow[n\to+\infty]{} \mathrm{e}^{-aT}$

— **Définition**
Soit $\lambda > 0$.
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, P)$.
On dit que X suit la loi de Poisson de paramètre λ si, et seulement si :

i $X(\Omega) = \mathbb{N}$

ii $\forall k \in \mathbb{N}\ P(X = k) = \mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}$

On note alors : $X \sim \mathcal{P}(\lambda)$.

— **Remarque**
La loi de Poisson est explicitement au programme.

2.5 Couples de variables aléatoires

2.5.1 Loi conjointe d'un couple de variables aléatoires

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, P)$.
On a vu en 1.3.5 que $C = (X, Y)$ est également une variable aléatoire.
On appelle loi conjointe de X et de Y la loi de C au sens de 2.1.1.
Elle est caractérisée par la distribution de probabilités de C ie la famille $(P(X = x, Y = y))_{(x,y)\in C(\Omega)}$.
Comme $C(\Omega)$ n'est pas toujours immédiat, on préfère s'intéresser à la famille $(P(X = x, Y = y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$
même si certains des évènements sont impossibles.

Plus généralement, si $X_1, \dots, X_n$ sont n variables aléatoires, $(X_1, \dots, X_n)$ est une variable aléatoire dont la loi est appelée loi conjointe des n variables aléatoires $X_1, \dots, X_n$.
Elle est caractérisée par les nombres $P(X_1 = x_1, \dots, X_n = x_n)$ pour $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$.

2.5.2 Lois marginales d'un couple de variables aléatoires

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit (X, Y) un couple de variables aléatoires discrètes.
On appelle lois marginales de (X, Y), la loi de X et la loi de Y.
Si on connaît la loi conjointe de X et de Y, on peut déterminer les lois marginales de (X, Y).
En effet, d'après la formule des probabilités totales :
$\forall x \in X(\Omega)\ P(X = x) = \displaystyle\sum_{y\in Y(\Omega)} P(X = x, Y = y)$

$\forall y \in Y(\Omega)\ P(Y = y) = \sum\limits_{x \in X(\Omega)} P(X = x, Y = y)$

Par contre, si on connaît les lois marginales on ne peut pas déterminer la loi conjointe : on ignore les liens entre X et Y.

Plus généralement, si $(X_1, \ldots, X_n)$ est une famille de variables aléatoires, les lois marginales de $(X_1, \ldots, X_n)$ sont les lois des variables aléatoires $X_1, \ldots, X_n$.
Si on connaît la loi conjointe des X_i on peut déterminer les lois marginales :
$\forall i \in [\![1;n]\!]\ \forall x_i \in X_i(\Omega)\ P(X_i = x_i) = \sum\limits_{(x_j)_{j \in [\![1;n]\!] \setminus \{i\}} \in \prod\limits_{j \in [\![1;n]\!] \setminus \{i\}} X_j(\Omega)} P\left(X_1 = x_1, \ldots, X_n = x_n\right)$
Par contre, si on connaît les lois marginales on ne peut pas déterminer la loi conjointe.

2.5.3 Loi conditionnelle de Y sachant un évènement A

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit Y une variable aléatoire sur $(\Omega, \mathcal{A}, P)$.
Soit A un évènement tel que $P(A) > 0$.
L'application $P_A \begin{cases} \mathcal{A} \to [0;1] \\ B \mapsto P_A(B) = P(B \mid A) \end{cases}$ est une probabilité sur $(\Omega, \mathcal{A})$.
Y est également une variable aléatoire sur $(\Omega, \mathcal{A}, P_A)$ (cela ne dépend que de la tribu).
On appelle loi de Y sachant A la loi de Y dans cet espace probabilisé.
C'est l'application $\begin{cases} \mathcal{P}(Y(\Omega)) \to [0;1] \\ B \mapsto P_A(Y \in B) = P(Y \in B \mid A) \end{cases}$
Un cas particulier fréquent est celui de $A = (X = x)$ où (X, Y) un couple de variables aléatoires.

Exemple
Soit $X \sim \mathcal{B}(n, p)$ et Y une variable aléatoire à valeurs dans $\mathbb{N}$ telle que la loi conditionnelle de Y sachant $X = k$ est $\mathcal{B}(k, q)$.

1. Montrer : $\forall \alpha \in [\![0;n]\!]\ \forall i \in [\![\alpha;n]\!]\ \binom{i}{\alpha}\binom{n}{i} = \binom{n}{\alpha}\binom{n-\alpha}{i-\alpha}$

2. Déterminer la loi de Y.

Correction
La situation modélisée dans l'exercice est la suivante :
On lance une pièce de monnaie n fois.
La probabilité d'obtenir face est $p \in]0;1[$.
On obtient X fois face.
On lance alors une deuxième pièce de monnaie X fois.
La probabilité d'obtenir face est $q \in]0;1[$.
On cherche la loi de Y, le nombre de faces lors du deuxième lancer.

1.
$$\begin{aligned}
\forall \alpha \in [\![0;n]\!]\ \forall i \in [\![\alpha;n]\!]\ \binom{i}{\alpha}\binom{n}{i} &= \frac{i!}{\alpha!(i-\alpha)!}\ \frac{n!}{i!(n-i)!} \\
&= \frac{n!}{\alpha!}\ \frac{1}{(n-i)!(i-\alpha)!} = \frac{n!}{\alpha!(n-\alpha)!}\ \frac{(n-\alpha)!}{(n-i)!(i-\alpha)!} \\
&= \binom{n}{\alpha}\binom{n-\alpha}{i-\alpha}
\end{aligned}$$

2. Soit $l \in \mathbb{N}$.
$P(Y=l) = \sum_{k=0}^{n} P(Y=l|X=k)P(X=k)$
$P(Y=l|X=k) = 0$ si $k < l$.
Cela se produit forcément si $l > n$ donc :
$\forall l > n\ P(Y=l) = 0$
$\forall l \in [\![0;n]\!]\ P(Y=l) = \sum_{k=l}^{n} \binom{k}{l} q^l (1-q)^{k-l} \binom{n}{k} p^k (1-p)^{n-k}$
On applique la première question avec $l = \alpha$ et $k = i$.

$$\begin{aligned}
\forall l \in [\![0;n]\!]\ P(Y=l) &= \sum_{k=l}^{n} \binom{n}{l}\binom{n-l}{k-l} q^l (1-q)^{k-l} p^k (1-p)^{n-k} \\
&= \binom{n}{l} q^l \sum_{k=l}^{n} \binom{n-l}{k-l} (1-q)^{k-l} p^k (1-p)^{n-k} \\
&= \binom{n}{l} q^l \sum_{i=0}^{n-l} \binom{n-l}{i} (1-q)^i p^{l+i} (1-p)^{n-l-i} \quad i = k-l \\
&= \binom{n}{l} q^l p^l \sum_{i=0}^{n-l} \binom{n-l}{i} \left(p(1-q)\right)^i (1-p)^{n-l-i} \\
&= \binom{n}{l} (pq)^l (p - pq + 1 - p)^{n-l}
\end{aligned}$$

Finalement $Y \sim \mathcal{B}(n, pq)$.

2.5.4 Un autre exemple de couple de variables aléatoires non indépendantes

Soit (X,Y) un couple de variables aléatoires à valeurs dans $\mathbb{N}^2$ tel que :

$$\forall (j,k) \in \mathbb{N}^2\ P(X=j, Y=k) = \frac{(j+k)\lambda^{j+k}}{\mathrm{e}j!k!}\ (\lambda > 0)$$

1. Déterminer λ.
2. Trouver les lois de X et de Y.
   X et Y sont-elles indépendantes ?
3. Les 5/2 calculeront $E(2^X)$, $E(2^{X+Y})$ et retrouveront le résultat de la question précédente.

Correction

1.

$$\begin{aligned}
1 &= \sum_{(j,k)\in\mathbb{N}^2} P(X=j,Y=k) = \sum_{j=0}^{+\infty}\sum_{k=0}^{+\infty} P(X=j,Y=k) \\
&= \sum_{j=0}^{+\infty}\left(\sum_{k=0}^{+\infty}\frac{(j+k)\lambda^{j+k}}{\mathrm{e}j!k!}\right) = \sum_{j=0}^{+\infty}\left(\frac{\lambda^j}{\mathrm{e}j!}\sum_{k=0}^{+\infty}\frac{(j+k)\lambda^k}{k!}\right)
\end{aligned}$$

Mais :

$$\begin{aligned}
\forall\lambda\in\mathbb{R}\ \sum_{k=0}^{+\infty}\frac{\lambda^k}{k!} &= \mathrm{e}^{\lambda} \\
\sum_{k=0}^{+\infty}k\frac{\lambda^k}{k!} &= \sum_{k=1}^{+\infty}k\frac{\lambda^k}{k!} = \sum_{k=1}^{+\infty}\frac{\lambda^k}{(k-1)!} = \lambda\sum_{l=0}^{+\infty}\frac{\lambda^l}{l!} \\
&= \lambda\,\mathrm{e}^{\lambda}
\end{aligned}$$

D'où :

$$\begin{aligned}
1 &= \sum_{j=0}^{+\infty}\left(\frac{\lambda^j}{\mathrm{e}j!}\left(j\,\mathrm{e}^{\lambda}+\lambda\,\mathrm{e}^{\lambda}\right)\right) = \sum_{j=0}^{+\infty}\left(\mathrm{e}^{\lambda-1}\left(j\frac{\lambda^j}{j!}+\lambda\frac{\lambda^j}{j!}\right)\right) \\
&= \mathrm{e}^{\lambda-1}\left(\lambda\,\mathrm{e}^{\lambda}+\lambda\,\mathrm{e}^{\lambda}\right) = \mathrm{e}^{\lambda-1}\,2\lambda\,\mathrm{e}^{\lambda} = 2\lambda\,\mathrm{e}^{2\lambda-1}
\end{aligned}$$

On pose $t=2\lambda$.
$2\lambda\,\mathrm{e}^{2\lambda-1}=1 \iff t\,\mathrm{e}^{t-1}=1 \iff \ln t+t-1=\varphi(t)=0$
$\forall t>0\ \varphi'(t)=\dfrac{1}{t}+1>0$
φ croît strictement de $-\infty$ à $+\infty$.
$\exists! t\in\mathbb{R}_+^*$ tq $\varphi(t)=0$
$t=1$ est racine évidente donc :
$\exists!\lambda>0$ tq $2\lambda\,\mathrm{e}^{2\lambda-1}=1$
C'est $\lambda=\dfrac{1}{2}$.

La loi de X a déjà été calculée :

$$\begin{aligned}
\forall j\in\mathbb{N}\ P(X=j) &= (j+\lambda)\frac{\lambda^j}{j!}\,\mathrm{e}^{\lambda-1} = \left(j+\frac{1}{2}\right)\frac{1}{2^j j!}\,\mathrm{e}^{-1/2} \\
&= \frac{2j+1}{2^{j+1}j!}\,\mathrm{e}^{-1/2}
\end{aligned}$$

$(j,k)\mapsto\dfrac{(j+k)\lambda^{j+k}}{\mathrm{e}j!k!}$ est symétrique en (j,k) donc :
$\forall k\in\mathbb{N}\ P(Y=k)=\dfrac{2k+1}{2^{k+1}k!}\,\mathrm{e}^{-1/2}$
$\forall (j,k)\in\mathbb{N}^2\ P(X=j)P(Y=k)=\dfrac{(2j+1)(2k+1)}{2^{j+k+2}j!k!\,\mathrm{e}}$
$\forall (j,k)\in\mathbb{N}^2\ P(X=j,Y=k)=\dfrac{j+k}{2^{j+k}j!k!\,\mathrm{e}}$
X et Y ne sont pas indépendantes. Par exemple, $P(X=0,Y=0)=0$ et $P(X=0)P(Y=0)=\dfrac{1}{4\,\mathrm{e}}$.

2. $E(2^X)=\sum_{j=0}^{+\infty}2^j\dfrac{2j+1}{2^{j+1}j!}\,\mathrm{e}^{-1/2}=\dfrac{\mathrm{e}^{-1/2}}{2}\sum_{j=0}^{+\infty}\dfrac{2j+1}{j!}=\dfrac{\mathrm{e}^{-1/2}}{2}(2\,\mathrm{e}+\,\mathrm{e})=\dfrac{3}{2}\,\mathrm{e}^{1/2}$

   De même, $E(2^Y)=\dfrac{3}{2}\,\mathrm{e}^{1/2}$

   Donc $E(2^X)E(2^Y)=\dfrac{9}{4}\,\mathrm{e}$

   Pour calculer $E\left(2^{X+Y}\right)$, plusieurs méthodes sont possibles :

   — On commence par déterminer la loi de $X+Y$.
   $(X+Y)(\Omega)\subset\mathbb{N}$

$$\begin{aligned}\forall p\in\mathbb{N}\ P(X+Y=p) &= \sum_{k=0}^{p}P(X=k,Y=p-k)=\sum_{k=0}^{p}\frac{k+p-k}{2^{k+p-k}k!(p-k)!}\frac{1}{\mathrm{e}}\\ &= \frac{1}{\mathrm{e}}\frac{p}{2^p}\sum_{k=0}^{p}\frac{1}{k!(p-k)!}=\frac{1}{\mathrm{e}}\frac{p}{2^p}\frac{1}{p!}\sum_{k=0}^{p}\frac{p!}{k!(p-k)!}\\ &= \frac{1}{\mathrm{e}}\frac{p}{2^p}\frac{1}{p!}\sum_{k=0}^{p}\binom{p}{k}=\frac{1}{\mathrm{e}}\frac{p}{2^p}\frac{1}{p!}2^p=\frac{p\,\mathrm{e}^{-1}}{p!}\end{aligned}$$

   Le calcul de $E(2^{X+Y})$ est alors standard. On peut aussi remarquer que $X+Y-1\sim\mathcal{P}(1)$ et utiliser la fonction génératrice.
   $E(2^{X+Y})=2\,\mathrm{e}$.
   $E(2^{X+Y})\neq E(2^X)E(2^Y)$: on retrouve : ”X et Y ne sont pas indépendantes”

   —

$$\begin{aligned}E(2^{X+Y}) &= \sum_{j=0}^{+\infty}\sum_{k=0}^{+\infty}2^{j+k}\frac{j+k}{2^{j+k}j!k!\,\mathrm{e}}=\sum_{p=0}^{+\infty}\left(\sum_{j+k=p}2^p\frac{p}{2^pj!k!\,\mathrm{e}}\right)\\ &= \sum_{p=0}^{+\infty}\left(p\,\mathrm{e}^{-1}\sum_{k=0}^{p}\frac{1}{k!(p-k)!}\right)=\sum_{p=0}^{+\infty}\left(\frac{p\,\mathrm{e}^{-1}}{p!}\sum_{k=0}^{p}\binom{p}{k}\right)\\ &= \sum_{p=1}^{+\infty}\frac{2^p\,\mathrm{e}^{-1}}{(p-1)!}=2\,\mathrm{e}^{-1}\,\mathrm{e}^2\\ &= 2\,\mathrm{e}\end{aligned}$$

2.5.5 Un exemple surprenant de couple de variables aléatoires indépendantes

Un autostoppeur attend au péage d'une autoroute pendant une certaine période. On admet que le nombre de véhicules franchissant le péage pendant cette période est une variable aléatoire $N\sim\mathcal{P}(\lambda)$. A chaque fois qu'un véhicule franchit le péage, il lance une pièce truquée (la probabilité d'amener pile est $p\in]0;1[$).
On note P le nombre de piles obtenus et F le nombre de faces.
P et F sont indépendantes.
Démonstration

— Loi du couple (P,N) :
  $\forall(k,l)\in\mathbb{N}^2\ P(P=k,N=l)=P(P=k|N=l)\times P(N=l)$
  Il est clair que la loi de P sachant $N=l$ est $\mathcal{B}(l,p)$ donc :

si $k > l$, $P(P = k, N = l) = 0$

Si $k \in [\![0;l]\!]$ $P(P = k, N = l) = \begin{pmatrix} k \\ l \end{pmatrix} p^k (1-p)^{l-k} \, \mathrm{e}^{-\lambda} \dfrac{\lambda^l}{l!}$

— Loi de P

$$\begin{aligned}
\forall k \in \mathbb{N} \; P(P = k) &= \sum_{l=0}^{+\infty} P(P = k, N = l) = \sum_{l=k}^{+\infty} \begin{pmatrix} l \\ k \end{pmatrix} p^k (1-p)^{l-k} \, \mathrm{e}^{-\lambda} \frac{\lambda^l}{l!} \\
&= p^k \, \mathrm{e}^{-\lambda} \sum_{l=k}^{+\infty} \begin{pmatrix} l \\ k \end{pmatrix} (1-p)^{l-k} \frac{\lambda^l}{l!} \\
&= p^k \, \mathrm{e}^{-\lambda} \sum_{l=k}^{+\infty} \frac{l!}{k!(l-k)!} (1-p)^{l-k} \frac{\lambda^l}{l!} \\
&= \frac{p^k \, \mathrm{e}^{-\lambda}}{k!} \sum_{i=0}^{+\infty} \frac{1}{i!} (1-p)^i \lambda^{k+i} \quad i = l - k \\
&= \frac{p^k \, \mathrm{e}^{-\lambda}}{k!} \lambda^k \; \mathrm{e}^{\lambda(1-p)} = \frac{(\lambda p)^k}{k!} \, \mathrm{e}^{-\lambda p}
\end{aligned}$$

P suit une loi de Poisson de paramètre λp.

— Loi de F :

F suit une loi de Poisson de paramètre λq. ($q = 1 - p$)

— Indépendance de P et de F :

$$\begin{aligned}
\forall (i,j) \in \mathbb{N}^2 \; P(P = i, F = j) &= P(P = i, N = i + j) = \begin{pmatrix} i+j \\ i \end{pmatrix} p^i (1-p)^j \, \mathrm{e}^{-\lambda} \frac{\lambda^{i+j}}{(i+j)!} \\
&= \frac{(i+j)!}{i!j!} p^i q^j \, \mathrm{e}^{-\lambda} \frac{\lambda^i \lambda^j}{(i+j)!} \\
&= \frac{(p\lambda)^i}{i!} \; \frac{(q\lambda)^j}{j!} \; \mathrm{e}^{-\lambda} = \frac{(p\lambda)^i}{i!} \; \frac{(q\lambda)^j}{j!} \; \mathrm{e}^{-\lambda(p+q)} \\
&= \frac{(p\lambda)^i}{i!} \, \mathrm{e}^{-\lambda p} \; \frac{(q\lambda)^j}{j!} \; \mathrm{e}^{-\lambda q} \\
&= P(P = i) \times P(F = j)
\end{aligned}$$

Variantes

— **Centrale 2016**

Une tortue pond des oeufs. On note N la variable aléatoire comptant le nombre d'oeufs pondus. N suit une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$. Chaque oeuf a une probabilité $p \in]0;1[$ de donner naissance à une tortue.

1. Trouver la loi de D, la variable aléatoire comptant le nombre de descendants de la tortue.
2. D et N sont-elles indépendantes ?

D et N ne sont pas indépendantes : si elle l'étaient, la loi de D conditionnellement à $(N = n)$ serait indépendante de n.

Par contre D et $N - D$ sont indépendantes.

— **Mines 2016**

Un promeneur se balade en forêt pour ramasser des champignons. Il a une probabilité p

de trouver un bolet et une probabilité $q = 1 - p$ de trouver une morille.
Soient X la variable aléatoire qui compte le nombre de bolets trouvés, Y celle qui compte le nombre de morilles et N le nombre total de champignons.
On suppose que N suit une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$.

1. Déterminer la loi du couple (N, X).
2. Déterminer la loi de X.
3. X et Y sont-elles indépendantes ?

— **Exercice 1** *(CCP 2019)*
On considère un péage avec m guichets.
Soit N le nombre de voitures passant au péage pendant une heure. On suppose que $N \hookrightarrow \mathcal{P}(\lambda)$.
Soit X_k la variable aléatoire égale au nombre de voitures passant au k-ième guichet.
$P(X_k = i | N = n)$?
$P(X_k = i)$?

2.5.6 Somme de deux variables aléatoires indépendantes suivant une loi de Poisson

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient X et Y deux variables aléatoires discrètes indépendantes sur $(\Omega, \mathcal{A}, P)$.
On suppose $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$ $((\lambda, \mu) \in (\mathbb{R}_+^*)^2)$.
Alors $X + Y \sim \mathcal{P}(\lambda + \mu)$.

— Ce résultat ne figure plus au programme. On en verra une deuxième démonstration dans le cours sur les séries génératrices.

Démonstration
$X(\Omega) = Y(\Omega) = \mathbb{N}$ donc $(X + Y)(\Omega) \subset \mathbb{N}$

$$\begin{aligned}
\forall k \in \mathbb{N}\ P(X+Y=k) &= P\left(\bigcup_{l=0}^{k} ((X=l) \cap (Y=k-l))\right) \\
&= \sum_{l=0}^{k} P(X=l, Y=k-l) \text{ par incompatibilité 2 à 2} \\
&= \sum_{l=0}^{k} P(X=l)\ P(Y=k-l) \text{ par indépendance de } X \text{ et de } Y \\
&= \sum_{l=0}^{k} \frac{\lambda^l}{l!} \mathrm{e}^{-\lambda} \frac{\mu^{k-l}}{(k-l)!} \mathrm{e}^{-\mu} = \frac{\mathrm{e}^{-(\lambda+\mu)}}{k!} \sum_{l=0}^{k} \frac{k!}{l!(k-l)!} \lambda^k \mu^{k-l} \\
&= \frac{\mathrm{e}^{-(\lambda+\mu)}}{k!} \sum_{l=0}^{k} \binom{k}{l} \lambda^k \mu^{k-l} \\
&= \frac{\mathrm{e}^{-(\lambda+\mu)}}{k!} (\lambda+\mu)^k
\end{aligned}$$

— La technique ici employée pour déterminer la loi de $X + Y$ peut s'appliquer à un couple de variables aléatoires discrètes indépendantes suivant des lois autres que la loi de Poisson.Elle peut être adaptée au cas où X et Y ne sont pas indépendantes mais où on connaît la loi conjointe.

— Ce résultat peut-être généralisé, par récurrence, à la somme de n variables aléatoires indépendantes.
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient $X_1, \ldots, X_n$ n variables aléatoires discrètes mutuellement indépendantes sur $(\Omega, \mathcal{A}, P)$.
On suppose : $\forall i \in [\![1; n]\!]$ $X_i \sim \mathcal{P}(\lambda_i)$
Alors $\displaystyle\sum_{i=1}^{n} X_i \sim \mathcal{P}\left(\sum_{i=1}^{n} \lambda_i\right)$
En effet, d'après le lemme des coalitions, $\displaystyle\sum_{i=1}^{n-1} X_i$ et X_n sont indépendantes.

Chapitre 3

Moments des variables aléatoires

3.1 Espérance

3.1.1 Cas des variables aléatoires positives

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire discrète à valeurs dans $[0; +\infty]$.
L'espérance de X est définie par :
$E(X) = \sum\limits_{x \in X(\Omega)} xP(X = x)$
avec la convention $xP(X = x) = 0$ lorsque $x = +\infty$ et $P(X = +\infty) = 0$.

3.1.2 Espérance d'une variable aléatoire qui suit une loi géométrique

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire qui suit la loi géométrique de paramètre p ($p \in]0; 1[$).
$E(X) = \dfrac{1}{p}$

Ce résultat figure explicitement au programme.

Démonstration
$X(\Omega) = \mathbb{N}^*$
$E(X) = \sum\limits_{n=1}^{+\infty} nP(X = n) = p \sum\limits_{n=1}^{+\infty} n(1 - p)^{n-1}$
Or :
$\forall t \in] - 1; 1[\; \sum\limits_{n=0}^{+\infty} t^n = \dfrac{1}{1 - t} \; (R = 1)$
En dérivant, ce qui est légitime :
$\forall t \in] - 1; 1[\; \sum\limits_{n=1}^{+\infty} nt^{n-1} = \dfrac{1}{(1 - t)^2}$
Or $1 - p \in]0; 1[$ donc :
$E(X) = \dfrac{p}{(1 - (1 - p))^2} = \dfrac{p}{p^2} = \dfrac{1}{p}$

3.1.3 Espérance d'une variable aléatoire qui suit une loi de Poisson

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire qui suit la loi de Poisson de paramètre λ ($\lambda \in \mathbb{R}_+^*$).
$E(X) = \lambda$

Ce résultat figure explicitement au programme.

Démonstration
$X(\Omega) = \mathbb{N}$

$$\begin{aligned} E(X) &= \sum_{n=0}^{+\infty} nP(X=n) = \sum_{n=1}^{+\infty} nP(X=n) \\ &= \lambda\,\mathrm{e}^{-\lambda} \sum_{n=1}^{+\infty} \frac{\lambda^{n-1}}{(n-1)!} = \lambda\,\mathrm{e}^{-\lambda} \times \mathrm{e}^{\lambda} \\ &= \lambda \end{aligned}$$

3.1.4 Un exemple de variable aléatoire positive d'espérance infinie

On effectue des tirages dans une urne contenant une boule blanche et une boule noire dans les conditions suivantes :

— si on tire une boule noire on arrête.

— si on tire une boule blanche on la remet dans l'urne avec une autre boule blanche.

Soit X le rang d'obtention de la boule noire.
Calculer $P(X=n)$ et $\displaystyle\sum_{n=1}^{+\infty} P(X=n)$.
Quelle est l'espérance de X ?
Correction
$X(\Omega) = \mathbb{N}^* \cup \{+\infty\}$
On note B_n : "on tire une boule blanche au $n^{\text{ième}}$ tirage"

$$\begin{aligned} P(X=n) &= P\left(B_1 \cap \cdots \cap B_{n-1} \cap \overline{B_n}\right) \\ &= P(B_1) \times P(B_2|B_1) \times \cdots \times P(\overline{B_n}|B_1 \cap \cdots \cap B_{n-1}) \\ &= \frac{1}{2} \times \frac{2}{3} \times \cdots \times \frac{1+n-2}{2+n-2} \times \frac{1}{2+n-1} \\ &= \frac{1}{2} \times \frac{2}{3} \times \cdots \times \frac{n-1}{n} \times \frac{1}{n+1} \\ &= \frac{1}{n(n+1)} \end{aligned}$$

$\displaystyle\sum_{n=1}^{+\infty} P(X=n) = \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \lim_{n\to+\infty} \frac{1}{n} = 1$
$\displaystyle P(X=+\infty) = 1 - \sum_{n=1}^{+\infty} P(X=n) = 0$
On peut aussi utiliser la continuité décroissante :
$\displaystyle (X=+\infty) = \bigcap_{n\geq 0} (X>n)$

$(X > n) = B_1 \cap \cdots \cap B_n$ de probabilité $\dfrac{1}{n+1}$
On tire ps la boule noire (au bout d'un nombre fini de tirages)
Par contre le temps d'attente moyen est infini :
$E(X) = \displaystyle\sum_{x \in X(\Omega)} xP(X = x) = \sum_{n=1}^{+\infty} \frac{1}{n+1} = +\infty$
En effet $nP(X = n) = \dfrac{1}{n+1}$ est le terme général d'une série divergente à termes réels positifs.

3.1.5 Cas des variables aléatoires à valeurs dans $[\![0; +\infty]\!]$

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N} \cup \{+\infty\}$.
Alors $E(X) = \displaystyle\sum_{n=1}^{+\infty} P(X \geq n) = \sum_{n=0}^{+\infty} P(X > n)$.

Démonstration
$\forall n \in \mathbb{N}\ P(X \geq n) = \displaystyle\sum_{k \in [\![n;+\infty]\!]} P(X = k)$
On introduit donc la famille $(u_{n,k})_{(n,k) \in \mathbb{N}^* \times [\![0;+\infty]\!]}$ définie par :
$u_{n,k} = \begin{cases} P(X = k) \text{ si } k \geq n \\ 0 \text{ si } k < n \end{cases}$
D'après le théorème de Fubini :

$$\begin{aligned}
\sum_{n=1}^{+\infty} P(X \geq n) &= \sum_{n \in \mathbb{N}^*} \left(\sum_{k \in [\![0;+\infty]\!]} u_{n,k} \right) \\
&= \sum_{k \in [\![0;+\infty]\!]} \left(\sum_{n \in \mathbb{N}^*} u_{n,k} \right) \\
&= 0 + \sum_{k \in \mathbb{N}^*} \left(\sum_{n=1}^{k} P(X = k) \right) + \sum_{n \in \mathbb{N}^*} P(X = +\infty)
\end{aligned}$$

Si $P(X = +\infty) = 0$, cette formule devient :
$\displaystyle\sum_{n=1}^{+\infty} P(X \geq n) = \sum_{k \in \mathbb{N}^*} kP(X = k) = E(X)$
Si $P(X = +\infty) > 0$ alors cette formule devient $\displaystyle\sum_{n=1}^{+\infty} P(X \geq n) = +\infty$ et :
$E(X) = \displaystyle\sum_{n \in [\![0;+\infty]\!]} nP(X = n) = +\infty$
Il y a bien égalité.

3.1.6 Cas des variables aléatoires complexes

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire discrète à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$.
On dit que X est d'espérance finie si la famille $(xP(X = x))_{x \in X(\Omega)}$ est sommable.
Dans ce cas, la somme de cette famille est appelée espérance de X et se note $E(X)$:

$E(X) = \sum_{x \in X(\Omega)} xP(X = x)$

On dit que X est une variable aléatoire centrée si, et seulement si, X est d'espérance nulle.
En d'autres termes :

$$X \text{ centrée} \iff \begin{cases} \sum_{x \in X(\Omega)} |x| P(X = x) < +\infty \\ \sum_{x \in X(\Omega)} xP(X = x) = 0 \end{cases}$$

3.1.7 Exemple

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A})$ à valeurs dans $\mathbb{Z}$ telle que :
$\forall n \in \mathbb{Z}\ P(X = n) = \frac{a}{2^{|n|}}$

1. Déterminer a.
2. Montrer que X est d'espérance finie et calculer son espérance.

Correction

1. On détermine a en écrivant que $\sum_{n \in \mathbb{Z}} P(X = n) = 1$.
   On a donc $a\left(1 + 2\sum_{n=1}^{+\infty} \frac{1}{2^n}\right) = 1$
   (Techniquement, on utilise la sommation par paquets en écrivant $\mathbb{Z} = \mathbb{Z}_-^* \cup \{0\} \cup \mathbb{Z}_+^*$)
   Donc $a = \frac{1}{1 + \sum_{n=1}^{+\infty} \frac{1}{2^{n-1}}} = \frac{1}{3}$
2. Pour tout $n \in \mathbb{N}^*$, soit $u_n = \frac{an}{2^n} > 0$.
   $\frac{u_{n+1}}{u_n} = \frac{n+1}{n}\frac{1}{2} \xrightarrow[n \to +\infty]{} \frac{1}{2} < 1$
   Donc la série de terme général u_n converge.
   $\sum_{n \in \mathbb{Z}} |n| P(X = n) = 2\sum_{n=1}^{+\infty} u_n < +\infty$ donc X a une espérance.
   On peut alors pour calculer l'espérance utiliser la sommation par paquets de la même manière : $E(X) = \sum_{n=1}^{+\infty} \frac{-na}{2^n} + \sum_{n=1}^{+\infty} \frac{na}{2^n} = 0$
   ou utiliser la décomposition $\mathbb{Z} = \{0\} \cup \{-1; 1\} \cup \{-2; 2\} \cup \dots$:
   $E(X) = \sum_{n=1}^{+\infty} (nP(X = n) + (-n)P(X = -n)) = \sum_{n=1}^{+\infty} 0 = 0$

3.1.8 Théorème du transfert

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A})$. Soit f une application à valeurs réelles ou complexes définie sur $X(\Omega)$.
$f(X)$ est d'espérance finie si et seulement si la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ est sommable.
Dans ce cas, on a :

$$E(f(X)) = \sum_{x \in X(\Omega)} f(x)P(X = x)$$

Démonstration

$$\begin{aligned}
& f(X) \text{ est d'espérance finie} \\
\iff & \sum_{y\in f(X(\Omega))} |y|\, P(f(X)=y) < +\infty \\
\iff & \sum_{y\in f(X(\Omega))} \left(|y| \sum_{x\in X(\Omega) \text{ tq } f(x)=y} P(X=x) \right) < +\infty \\
& \text{car } (f(X)=y) = \bigcup_{x\in X(\Omega) \text{ tq } f(x)=y} (X=x) \\
\iff & \sum_{y\in f(X(\Omega))} \left(\sum_{x\in X(\Omega) \text{ tq } f(x)=y} |f(x)|\, P(X=x) \right) < +\infty \\
\iff & \sum_{x\in X(\Omega)} |f(x)P(X=x)| < \infty \text{ (sommation par paquets)} \\
\iff & \text{ la famille } (f(x)P(X=x))_{x\in X(\Omega)} \text{ est sommable}
\end{aligned}$$

On a alors :

$$\begin{aligned}
E(f(X)) &= \sum_{y\in f(X(\Omega))} yP(f(X)=y) \\
&= \sum_{y\in f(X(\Omega))} \left(y \sum_{x\in X(\Omega) \text{ tq } f(x)=y} P(X=x) \right) \\
&= \sum_{y\in f(X(\Omega))} \left(\sum_{x\in X(\Omega) \text{ tq } f(x)=y} f(x)P(X=x) \right) \\
&= \sum_{x\in X(\Omega)} f(x)P(X=x)
\end{aligned}$$

Remarque

La formule s'applique aux couples de variables aléatoires. Par exemple avec $f \begin{cases} \mathbb{C}^2 \to \mathbb{C} \\ (x,y) \mapsto x+y \end{cases}$,
on obtient pour deux variables aléatoires complexes d'espérance finie :
$E(X+Y) = \displaystyle\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} (x+y)P(X=x, Y=y)$

La formule s'étend aussi aux n-uplets de variables aléatoires et donne par exemple :
$E(X_1+\cdots+X_n) = \displaystyle\sum_{(x_1,\ldots,x_n)\in X_1(\Omega)\times\cdots\times X_n(\Omega)} (x_1+\cdots+x_n)P(X_1=x_1,\ldots,X_n=x_n)$

3.1.9 Linéarité

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient X et Y deux variables aléatoires complexes discrètes d'espérances finies.
Alors pour tous a et b dans $\mathbb{C}$, $aX+bY$ est une variable aléatoire discrète d'espérance finie et :
$E(aX+bY) = aE(X) + bE(Y)$.

Démonstration

Soit $(a,b) \in \mathbb{C}^2$.
Soit $f \begin{cases} \mathbb{C}^2 \to \mathbb{C} \\ (x,y) \mapsto ax+by \end{cases}$.

$$\begin{array}{cl} & \displaystyle\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} |ax+by|\, P(X=x, Y=y) \\ \leq & \displaystyle\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} (|a|\,|x| + |b|\,|y|)\, P(X=x, Y=y) \\ \leq & \displaystyle\sum_{x\in X(\Omega)} \left(\sum_{y\in Y(\Omega)} (|a|\,|x| + |b|\,|y|)\, P(X=x, Y=y) \right) \\ \leq & \displaystyle\sum_{x\in X(\Omega)} (|a|\,|x|\, P(X=x)) + \sum_{x\in X(\Omega)} \left(\sum_{y\in Y(\Omega)} |b|\,|y|\, P(X=x, Y=y) \right) \\ \leq & \displaystyle |a|\, E(|X|) + \sum_{y\in Y(\Omega)} \left(\sum_{x\in X(\Omega)} |b|\,|y|\, P(X=x, Y=y) \right) \\ \leq & \displaystyle |a|\, E(|X|) + \sum_{y\in Y(\Omega)} (|b|\,|y|\, P(Y=y)) \\ \leq & |a|\, E(|X|) + |b|\, E(|Y|) < +\infty \end{array}$$

D'après le théorème de transfert, $aX+bY$ est d'espérance finie.
Toujours d'après le théorème de transfert :

$$\begin{array}{rcl} E(aX+bY) & = & \displaystyle\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} (ax+by)P(X=x, Y=y) \\ & = & \displaystyle\sum_{x\in X(\Omega)} \left(\sum_{y\in Y(\Omega)} (ax+by)P(X=x, Y=y) \right) \\ & = & \displaystyle\sum_{x\in X(\Omega)} axP(X=x) + \sum_{x\in X(\Omega)} \left(\sum_{y\in Y(\Omega)} byP(X=x, Y=y) \right) \\ & = & \displaystyle aE(X) + \sum_{y\in Y(\Omega)} \left(\sum_{x\in X(\Omega)} byP(X=x, Y=y) \right) \\ & = & \displaystyle aE(X) + \sum_{y\in Y(\Omega)} byP(Y=y) \\ & = & aE(X) + bE(Y) \end{array}$$

Exemple d'application
On lance une pièce de monnaie n fois. La probabilité d'obtenir face, pour chaque lancer individuel, est $p \in]0;1[$.
Une série est une suite de lancers qui donnent le même résultat.
Par exemple dans la suite FFPFPPF, il y a cinq séries.
Quelle est l'espérance du nombre de séries ?

Correction
Le nombre de séries est égal à 1+le nombre de changements où on appelle changement le cas où

deux lancers successifs donnent des résultats différents.
Il peut être tentant de dire que le nombre de changements suit la loi binomiale de paramètres $n-1$ et r où r est la probabilité que deux lancers successifs donnent deux résultats différents mais il y a un problème d'indépendance.
Adoptons des notations plus précises.
On note X_i la variable aléatoire qui vaut 1 si on obtient face au $i^{\text{ième}}$ lancer, 0 sinon.
$(X_i)_{1\leq i\leq n}$ est une famille de variables aléatoires mutuellement indépendantes qui suivent toutes la loi de Bernoulli de paramètre p.
Pour tout i compris ente 1 et $n-1$, on pose $Y_i = 1_{X_i \neq X_{i+1}}$. En d'autres termes, Y_i vaut 1 si il y a changement entre les lancers numéros i et $i+1$, 0 sinon.
Le nombre de séries est $N = 1 + \displaystyle\sum_{i=1}^{n-1} Y_i$.
Pour tout i compris entre 1 et $n-1$, Y_i suit la loi de Bernoulli de paramètre :
$r = P(Y_i = 1) = P(X_i \neq X_{i+1}) = P(X_i = 1, X_{i+1} = 0) + P(X_i = 0, X_{i+1} = 1) = 2p(1-p)$
On a donc par linéarité de l'espérance :
$E(N) = 1 + 2(n-1)p(1-p)$
On peut observer que $E(N)$ est maximale lorsque $p = \dfrac{1}{2}$. Si p est proche de 0 ou de 1, la pièce tombe à peu près tout le temps du même côté et il y a une seule série.

Revenons sur la loi de N :

$$\begin{array}{rcl}
P(Y_1 = 1, Y_2 = 1) & = & P(X_1 = 0, X_2 = 1, X_3 = 0) + P(X_1 = 1, X_2 = 0, X_3 = 1) \\
& = & p(1-p)^2 + p^2(1-p) = p(1-p)(1-p+p) \\
& = & p(1-p) \\
P(Y_1 = 1)P(Y_2 = 1) & = & r^2 = 4p^2(1-p)^2 = 4p(1-p)P(Y_1 = 1, Y_2 = 1)
\end{array}$$

On trouve facilement :
$4p(1-p) = 1 \iff p = \dfrac{1}{2}$
Si $p \neq \dfrac{1}{2}$, Y_1 et Y_2 ne sont pas indépendantes, N ne suivra pas une loi binomiale.
Si $p = \dfrac{1}{2}$, on ne peut rien dire pour l'instant mais on va montrer que si $p = \dfrac{1}{2}$ les variables aléatoires $Y_1, \ldots, Y_{n-1}$ sont mutuellement indépendantes.
Il s'agit donc de démontrer, pour tout $n \geq 2$:
$\forall(\epsilon_1, \ldots, \epsilon_{n-1}) \in \{0;1\}^{n-1}\ P(Y_1 = \epsilon_1, \ldots, Y_{n-1} = \epsilon_{n-1}) = P(Y_1 = \epsilon_1) \times \cdots \times P(Y_n = \epsilon_n)$
On fixe $n \geq 2$ et on considère $\Phi \begin{cases} \{0;1\}^n \to \{0;1\}^n \\ (x_1, \ldots, x_n) \mapsto (x_1, z_1, \ldots, z_{n-1}) \end{cases}$ où $z_i = 1$ si $x_i \neq x_{i+1}$ et $z_i = 0$ si $x_i = x_{i+1}$ ie $z_i = \delta(x_i, x_{i+1})$ avec $\delta \begin{cases} \{0;1\}^2 \to \{0;1\} \\ (x,y) \mapsto 1 \text{ si } x \neq y \\ (x,y) \mapsto 0 \text{ si } x = y \end{cases}$.
Φ est injective :
Supposons $\Phi((a_1, \ldots, a_n)) = \Phi((b_1, \ldots, b_n))$.
On a $a_1 = b_1$: c'est la première composante de $\Phi((a_1, \ldots, a_n)) = \Phi((b_1, \ldots, b_n))$.
Supposons $(a_1, \ldots, a_k) = (b_1, \ldots, b_k)$.
Considérons z_k la $(k+1)$-ème composante de $\Phi((a_1, \ldots, a_n)) = \Phi((b_1, \ldots, b_n))$.
Si $z_k = 0$ alors $a_{k+1} = a_k$ et $b_{k+1} = b_k$ donc $a_{k+1} = b_{k+1}$.
Si $z_k = 1$ alors $a_{k+1} = 1 - a_k$ et $b_{k+1} = 1 - b_k$ donc $a_{k+1} = b_{k+1}$.

On montre donc par récurrence finie que $a_k = b_k$ pour tout k compris entre 1 et n ie $(a_1, \dots, a_n) = (b_1, \dots, b_n)$.
Φ est bien injective. Comme l'ensemble de départ et l'ensemble d'arrivée sont finis et de même cardinal, Φ est bijective.
On a alors :

$$\begin{aligned}
& \forall (\epsilon_1, \dots, \epsilon_{n-1}) \in \{0;1\}^{n-1} \\
& P(Y_1 = \epsilon_1, \dots, Y_{n-1} = \epsilon_{n-1}) \\
= \; & P(X_1 = 0, Y_1 = \epsilon_1, \dots, Y_{n-1} = \epsilon_{n-1}) + P(X_1 = 1, Y_1 = \epsilon_1, \dots, Y_{n-1} = \epsilon_{n-1}) \\
& \text{ par la formule des probabilités totales} \\
= \; & P(X_1 = 0, \delta(X_1, X_2) = \epsilon_1, \dots, \delta(X_{n-1}, X_n) = \epsilon_{n-1}) \\
& + P(X_1 = 1, \delta(X_1, X_2) = \epsilon_1, \dots, \delta(X_{n-1}, X_n) = \epsilon_{n-1}) \\
= \; & P\left(\Phi((X_1, \dots, X_n)) = (0, \epsilon_1, \dots, \epsilon_{n-1})\right) + P\left(\Phi((X_1, \dots, X_n)) = (1, \epsilon_1, \dots, \epsilon_{n-1})\right) \\
= \; & P\Big((X_1, \dots, X_n) = \Phi^{-1}((0, \epsilon_1, \dots, \epsilon_{n-1}))\Big) + P\Big((X_1, \dots, X_n) = \Phi^{-1}((1, \epsilon_1, \dots, \epsilon_{n-1}))\Big) \\
= \; & P\left((X_1, \dots, X_n) = (e_1, \dots, e_n)\right) + P\left((X_1, \dots, X_n) = (f_1, \dots, f_n)\right) \\
& \text{ où } (e_1, \dots, e_n) = \Phi^{-1}((0, \epsilon_1, \dots, \epsilon_{n-1})) \text{ et } (f_1, \dots, f_n) = \Phi^{-1}((1, \epsilon_1, \dots, \epsilon_{n-1})) \\
= \; & \prod_{i=1}^{n} P(X_i = e_i) + \prod_{i=1}^{n} P(X_i = f_i) \text{ par indépendance} \\
= \; & \prod_{i=1}^{n} \frac{1}{2} + \prod_{i=1}^{n} \frac{1}{2} \text{ car les } X_i \text{ ont toutes la même loi} \\
= \; & 2 \times \left(\frac{1}{2}\right)^n = \left(\frac{1}{2}\right)^{n-1} \\
= \; & P(Y_1 = \epsilon_1) \times \cdots \times P(Y_n = \epsilon_n) \text{ car } r = 2p(1-p) = \frac{1}{2} \text{ ici}
\end{aligned}$$

Les variables aléatoires $Y_1, \dots, Y_{n-1}$ sont mutuellement indépendantes.
Dans ce cas, $N - 1 \sim \mathcal{B}\left(n-1, \frac{1}{2}\right)$

3.1.10 Positivité de l'espérance

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire réelle discrète positive et d'espérance finie.
Alors $E(X) \geq 0$.
De plus, si $E(X) = 0$ alors $P(X = 0) = 1$.

Démonstration
Comme X est positive, $E(X) = \displaystyle\sum_{x \in X(\Omega)} xP(X = x)$, cette somme étant finie ou infinie.
C'est une somme de nombres positifs donc elle est positive.
De plus si $E(X) = 0$ alors pour tout $x \in X(\Omega)$, $xP(X = x) = 0$. Donc pour tout x non nul $P(X = x) = 0$.
On en déduit $P(X \neq 0) = 0$ puis $P(X = 0) = 1$.

3.1.11 Croissance de l'espérance

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient X et Y deux variables aléatoires réelles discrètes d'espérances finies.
On suppose $X \leq Y$.
Alors $E(X) \leq E(Y)$.

— Ce résultat se démontre en combinant la linéarité et la positivité de l'espérance.
— En particulier si $X \geq a$ alors $E(X) \geq a$.
De même, si $X \leq b$ alors $E(X) \leq b$.

3.1.12 Utilisation d'une majoration

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire à valeurs complexes.
Soit Y une variable aléatoire à valeurs réelles positives.
Si $|X| \leq Y$ et si $E(Y) < +\infty$ alors X est d'espérance finie.

Démonstration

$$\begin{aligned} E(|X|) &= \sum_{x \in X(\Omega)} |x|\, P(X = x) = \sum_{x \in X(\Omega)} \left(|x| \sum_{y \in Y(\Omega)} P(X = x, Y = y) \right) \\ &= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} |x|\, P(X = x, Y = y) \end{aligned}$$

Soit $(x, y) \in X(\Omega) \times Y(\Omega)$.
Si $|x| \leq y$ alors $|x|\, P(X = x, Y = y) \leq yP(X = x, Y = y)$
Si $|x| > y$ alors $P(X = x, Y = y) = 0$ donc on a aussi $|x|\, P(X = x, Y = y) \leq yP(X = x, Y = y)$.
On en déduit :

$$\begin{aligned} E(|x|) &\leq \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} yP(X = x, Y = y) \\ &\leq \sum_{y \in Y(\Omega)} \sum_{x \in X(\Omega)} yP(X = x, Y = y) \\ &\leq \sum_{y \in Y(\Omega)} yP(Y = y) = E(Y) < +\infty \end{aligned}$$

3.1.13 Espérance du produit de deux variables aléatoires indépendantes

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient X et Y deux variables aléatoires réelles discrètes d'espérances finies.
On suppose X et Y indépendantes.
Alors XY est une variable aléatoire réelle discrète d'espérance finie et :
$E(X\,Y) = E(X)\,E(Y)$

Démonstration

$$\begin{aligned}
E(|XY|) &= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} |xy|\, P(X=x, Y=y) \\
&= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} |x|\, P(X=x)\, |y|\, P(Y=y) \text{ par indépendance} \\
&= \left(\sum_{x\in X(\Omega)} |x|\, P(X=x)\right) \times \left(\sum_{y\in Y(\Omega)} |y|\, P(Y=y)\right) \\
&= E(|X|) \times E(|Y|) < +\infty
\end{aligned}$$

Donc XY a une espérance finie.
On a alors :

$$\begin{aligned}
E(XY) &= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X=x, Y=y) \\
&= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xP(X=x)yP(Y=y) \text{ par indépendance} \\
&= \left(\sum_{x\in X(\Omega)} xP(X=x)\right) \times \left(\sum_{y\in Y(\Omega)} yP(Y=y)\right) \\
&= E(X) \times E(Y)
\end{aligned}$$

Remarque

— La réciproque est fausse : $E(X\,Y) = E(X)\,E(Y)$ n'implique pas l'indépendance de X et de Y.

  Exemple
  Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
  Soit X une v.a.r telle que $X(\Omega) = \{-2; -1; 1; 2\}$ suivant une loi uniforme.
  Soit $Y = X^2$.
  $E(X) = \frac{1}{4}(-2) + \frac{1}{4}(-1) + \frac{1}{4}1 + \frac{1}{4}2 = 0$
  $E(XY) = E(X^3) = \frac{1}{4}(-8) + \frac{1}{4}(-1) + \frac{1}{4}1 + \frac{1}{4}8 = 0$ (par le théorème de transfert)
  On a $E(XY) = E(X)\,E(Y)$
  Mais X et Y ne sont pas indépendantes :
  $P(X = 1, Y = 4) = 0$
  $P(X = 1) = \frac{1}{4}$
  $P(Y = 4) = P(X = 2 \text{ ou } X = -2) = \frac{1}{2}$

— On démontre facilement par récurrence le résultat suivant :
  Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
  Soient $X_1, \ldots, X_n$ n variables aléatoires complexes mutuellement indépendantes, toutes d'espérance finie.
  $X_1 \ldots X_n$ est alors d'espérance finie et :
  $E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E(X_i)$

3.1.14 Modèle de diffusion de Bernoulli (1769)

Une urne R contient n boules rouges.
Une urne B contient n boules bleues.
A chaque étape, une boule est tirée au sort dans chaque urne et on les échange.
Montrer que le nombre moyen de boules rouges dans l'urne R après l'étape k est $\dfrac{n}{2}\left(1+\left(1-\dfrac{2}{n}\right)^k\right)$.

— **Première méthode : point de vue macroscopique**
On note X_k le nombre de boules rouges dans R après k étapes.
Après k étapes, l'urne R contient donc X_k boules rouges et $n-X_k$ boules bleues. L'urne B contient $n-X_k$ boules rouges et X_k boules bleues.
$X_0 = n$
$\forall k \in \mathbb{N}^*\ X_k = X_{k-1} - 1_{A_k} + 1_{B_k}$
où pour tout $k \in \mathbb{N}^*$:
A_k est l'évènement : "on tire une boule rouge dans l'urne R à l'étape k"
B_k est l'évènement : "on tire une boule rouge dans l'urne B à l'étape k"
Par linéarité de l'espérance :
$\forall k \in \mathbb{N}^*\ E(X_k) = E(X_{k-1}) - E(1_{A_k}) + E(1_{B_k}) = E(X_{k-1}) - P(A_k) + P(B_k)$
Par la formule des probabilités totales :

$$\begin{aligned}
\forall k \in \mathbb{N}^*\ P(A_k) &= \sum_{i=0}^{n} P(A_k|X_{k-1}=i)\ P(X_{k-1}=i) = \sum_{i=0}^{n} \frac{i}{n}\ P(X_{k-1}=i) \\
&= \frac{1}{n}E(X_{k-1}) \\
P(B_k) &= \sum_{i=0}^{n} P(B_k|X_{k-1}=i)\ P(X_{k-1}=i) = \sum_{i=0}^{n} \frac{n-i}{n}\ P(X_{k-1}=i) \\
&= 1 - \frac{1}{n}E(X_{k-1})
\end{aligned}$$

On a donc :
$\forall k \in \mathbb{N}^*\ E(X_k) = 1 + \left(1-\dfrac{2}{n}\right)E(X_{k-1})$
On a affaire à une suite arithmético-géométrique. C'est une situation classique.
On cherche le point fixe :
$l = 1 + \left(1-\dfrac{2}{n}\right) l = 1 + l - \dfrac{l}{2n} \Longleftrightarrow l = \dfrac{n}{2}$
On a donc :

$$\begin{aligned}
E(X_k) &= 1 + \left(1-\frac{2}{n}\right)E(X_{k-1}) \\
\frac{n}{2} &= 1 + \left(1-\frac{2}{n}\right)\frac{n}{2}
\end{aligned}$$

On fait la différence :
$\forall k \in \mathbb{N}^*\ E(X_k) - \dfrac{n}{2} = \left(1-\dfrac{2}{n}\right)\left(E(X_{k-1}) - \dfrac{n}{2}\right)$
Donc :
$\forall k \in \mathbb{N}\ E(X_k) - \dfrac{n}{2} = \left(1-\dfrac{2}{n}\right)^k\left(E(X_0) - \dfrac{n}{2}\right) = \left(1-\dfrac{2}{n}\right)^k \dfrac{n}{2}$
On conclut facilement.

— **Deuxième méthode : point de vue microscopique**
Les boules rouges sont numérotées $R_1, \ldots, R_n$.
On note $Y_{k,i} = \begin{cases} 1 \text{ si } R_i \text{ est dans l'urne } R \text{ après } k \text{ étapes} \\ 0 \text{ sinon} \end{cases}$.
Le nombre de boules rouges dans l'urne R après l'étape k est $\displaystyle\sum_{i=1}^{n} Y_{k,i}$
Son espérance est :

$$\begin{aligned} \sum_{i=1}^{n} E(Y_{k,i}) &= \sum_{i=1}^{n} P(Y_{k,i} = 1) \\ &= \sum_{i=1}^{n} P(R_i \text{ est dans l'urne } R \text{ après } k \text{ étapes}) \\ &= nP(R_1 \text{ est dans l'urne } R \text{ après } k \text{ étapes}) \end{aligned}$$

Or :

$$\begin{aligned} P(R_1 \text{ est dans l'urne } R \text{ après } k \text{ étapes}) &= P(R_1 \text{ a été tirée un nombre pair de fois}) \\ &= \sum_{\substack{0 \leq l \leq k \\ l \text{ pair}}} P(R_1 \text{ a été tirée } l \text{ fois}) \\ &= \sum_{0 \leq 2m \leq k} \binom{k}{2m} \left(\frac{1}{n}\right)^{2m} \left(1 - \frac{1}{n}\right)^{k-2m} \end{aligned}$$

Le calcul de cette somme est classique : on introduit : $\displaystyle\sum_{0 \leq 2m+1 \leq k} \binom{k}{2m+1} \left(\frac{1}{n}\right)^{2m+1} \left(1 - \frac{1}{n}\right)^{k-2m-1}$.
On a :

$$\begin{aligned} &\sum_{0 \leq 2m \leq k} \binom{k}{2m} \left(\frac{1}{n}\right)^{2m} \left(1 - \frac{1}{n}\right)^{k-2m} + \sum_{0 \leq 2m+1 \leq k} \binom{k}{2m+1} \left(\frac{1}{n}\right)^{2m+1} \left(1 - \frac{1}{n}\right)^{k-2m-1} \\ = &\sum_{0 \leq m \leq k} \binom{k}{m} \left(\frac{1}{n}\right)^{m} \left(1 - \frac{1}{n}\right)^{k-m} = \left(\frac{1}{n} - 1 - \frac{1}{n}\right)^{k} = 1 \end{aligned}$$

et

$$\begin{aligned} &\sum_{0 \leq 2m \leq k} \binom{k}{2m} \left(\frac{1}{n}\right)^{2m} \left(1 - \frac{1}{n}\right)^{k-2m} - \sum_{0 \leq 2m+1 \leq k} \binom{k}{2m+1} \left(\frac{1}{n}\right)^{2m+1} \left(1 - \frac{1}{n}\right)^{k-2m-1} \\ = &\sum_{0 \leq m \leq k} \binom{k}{m} (-1)^m \left(\frac{1}{n}\right)^{m} \left(1 - \frac{1}{n}\right)^{k-m} = \left(-\frac{1}{n} + 1 - \frac{1}{n}\frac{1}{n}\right)^{k} = \left(1 - \frac{2}{n}\right)^{k} \end{aligned}$$

En sommant, on trouve :

$$2 \sum_{0 \leq 2m \leq k} \binom{k}{2m} \left(\frac{1}{n}\right)^{2m} \left(1 - \frac{1}{n}\right)^{k-2m} = 1 + \left(1 - \frac{2}{n}\right)^{k}$$

On conclut facilement.
Le calcul de la variance n'est pas facile car les $Y_{k,i}$ ne sont pas indépendantes. On a par exemple $\displaystyle\sum_{i=1}^{n} Y_{1,i} = n - 1$

3.2 Variance

3.2.1 Un résultat préliminaire

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire réelle discrète.
Si X^2 est d'espérance finie alors X est elle-même d'espérance finie.

Démonstration

— **Première méthode**
Soit $x \in \mathbb{R}$.
Si $|x| \geq 1$ alors : $|x| \leq x^2$ et $|x| \leq 1 + x^2$
Si $|x| \leq 1$, alors $|x| \leq 1 + x^2$
Donc dans tous les cas : $|X| \leq 1 + X^2$
Par hypothèse, X^2 a une espérance finie. La variable aléatoire constamment égale à 1 aussi donc par linéarité, $1 + X^2$ a une espérance finie.
Par majoration, X est d'espérance finie.

— **Deuxième méthode**
D'après le théorème de transfert, la famille $(x^2 P(X = x))_{x \in X(\Omega)}$ est sommable.
Il s'agit de prouver que la famille $(xP(X = x))_{x \in X(\Omega)}$ est sommable.
Soit J une partie finie de $X(\Omega)$.

$$\begin{aligned}
\sum_{x \in J} |xP(X = x)| &= \sum_{x \in J} \sqrt{P(X = x)}\ \sqrt{P(X = x)}\,|x| \\
&\leq \sqrt{\sum_{x \in J} P(X = x)}\ \sqrt{\sum_{x \in J} x^2 P(X = x)} \\
&\leq \sqrt{\sum_{x \in X(\Omega)} P(X = x)}\ \sqrt{\sum_{x \in X(\Omega)} x^2 P(X = x)} \\
&\leq \sqrt{1}\ \sqrt{E(X^2)} = \sqrt{E(X^2)}
\end{aligned}$$

Donc la famille $(xP(X = x))_{x \in X(\Omega)}$ est sommable.
De plus on a $|E(X)| \leq E(|X|) \leq \sqrt{E(X^2)}$.

Remarque
On peut aller plus loin :
Mines 2017
Soit X une variable aléatoire discrète possédant un moment d'ordre $n \geq 2$.
Montrer que pour tout $k \in [\![1, n-1]\!]$, X possède un moment d'ordre k.

Correction
$\forall x \in \mathbb{R}^+\ x^k \leq 1 + x^n$ en distinguant les cas $x \leq 1$ et $x \geq 1$.
On en déduit $\left|X^k\right| = |X|^k \leq 1 + |X|^n = 1 + |X^n|$
Par hypothèse, X^n a une espérance finie. La variable aléatoire constamment égale à 1 aussi donc par linéarité, $1 + |X|^n$ a une espérance finie.
Par majoration, X^k est d'espérance finie

3.2.2 Définition de la variance

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire réelle discrète telle que X^2 soit d'espérance finie.
On appelle alors variance de X le réel $V(X) = E\left((X - E(X))^2\right) = E(X^2) - E(X)^2$.

Justification de la définition
$(X - E(X))^2 = X^2 - 2E(X)X + E(X)^2$
X^2 a une espérance finie.
X a une espérance finie donc $2E(X)X$ aussi.
$E(X)^2$ est une constante.
D'après la linéarité de l'espérance, $(X - E(X))^2$ a bien une espérance.
De plus :
$E\left((X - E(X))^2\right) = E(X^2) - 2E(X)E(X) + E(X)^2E(1) = E(X^2) - E(X)^2$
On retrouve : $|E(X)| \leq \sqrt{E(X^2)}$

3.2.3 Positivité de la variance

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire réelle discrète telle que X^2 soit d'espérance finie.
On a vu que $V(X)$ est alors un nombre réel.
Ce nombre est positif.
C'est trivial sous la forme $V(X) = E\left((X - E(X))^2\right)$.

Si $V(X) = 0$ alors il existe un nombre a tel que $P(X = a) = 1$.

En effet, $E\left((X - E(X))^2\right) = 0$ donc $X = E(X)$ ps.

3.2.4 Ecart type

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire réelle discrète telle que X^2 soit d'espérance finie.
On appelle écart type de X et on note $\sigma(X)$ le réel positif $\sigma(X) = \sqrt{V(X)}$.
On dit qu'une variable aléatoire est réduite lorsque son écart type est égal à 1.

3.2.5 Variance d'une variable aléatoire qui suit une loi géométrique

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire qui suit la loi géométrique de paramètre p ($p \in]0;1[$).
$V(X) = \dfrac{1-p}{p^2}$

Ce résultat figure explicitement au programme.

Démonstration
$X(\Omega) = \mathbb{N}^* = \{n, n \geq 1\}$
Pour tout $n \in \mathbb{N}^*$, on note $u_n = n^2 p(1-p)^{n-1}$.
$\forall n \in \mathbb{N}^*\ u_n > 0$
et
$\dfrac{u_{n+1}}{u_n} = \dfrac{(n+1)^2}{n^2}(1-p) \xrightarrow[n\to+\infty]{} 1 - p < 1$

Donc $\sum u_n$ converge.
X a bien une variance.

$$\begin{aligned}
\forall t \in]-1;1[\quad \frac{1}{1-t} &= \sum_{n=0}^{+\infty} t^n \\
\frac{1}{(1-t)^2} &= \sum_{n=1}^{+\infty} n t^{n-1} \\
\frac{t}{(1-t)^2} &= t(1-t)^{-2} = \sum_{n=1}^{+\infty} n t^n \\
(1-t)^{-2} + 2t(1-t)^{-3} &= \sum_{n=1}^{+\infty} n^2 t^{n-1}
\end{aligned}$$

D'où :

$$\begin{aligned}
E(X^2) &= \sum_{n=1}^{+\infty} n^2 p(1-p)^{n-1} \\
&= p\left((1-(1-p))^{-2} + 2(1-p)\left(1-(1-p)\right)^{-3}\right) \\
&= p\left(\frac{1}{p^2} + \frac{2-2p}{p^3}\right) = \frac{p+2-2p}{p^2} = \frac{2-p}{p^2}
\end{aligned}$$

Enfin :
$V(X) = E(X^2) - E(X)^2 = \dfrac{2-p}{p^2} - \dfrac{1}{p^2} = \dfrac{1-p}{p^2}$

3.2.6 Variance d'une variable aléatoire qui suit une loi de Poisson

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire qui suit la loi de Poisson de paramètre λ ($\lambda \in \mathbb{R}_+^*$).
$V(X) = \lambda$

Ce résultat figure explicitement au programme.

Démonstration
$X(\Omega) = \mathbb{N}$
Pour tout $n \in \mathbb{N}^*$, on note $u_n = n^2 \dfrac{\lambda^n}{n!} \mathrm{e}^{-\lambda}$
$\forall n \in \mathbb{N}^* \ u_n > 0$
et
$\dfrac{u_{n+1}}{u_n} = \dfrac{(n+1)^2}{n^2} \dfrac{\lambda}{n+1} \xrightarrow[n \to +\infty]{} 0 < 1$
Donc $\sum u_n$ converge.

X a bien une variance.

$$\begin{aligned}
E(X^2) &= \sum_{n=0}^{+\infty} n^2 \frac{\lambda^n}{n!}\,\mathrm{e}^{-\lambda} = \lambda\,\mathrm{e}^{-\lambda} \sum_{n=1}^{+\infty} n^2 \frac{\lambda^{n-1}}{n!} = \lambda\,\mathrm{e}^{-\lambda} \sum_{n=1}^{+\infty} n \frac{\lambda^{n-1}}{(n-1)!} \\
&= \lambda\,\mathrm{e}^{-\lambda} \sum_{k=0}^{+\infty} (k+1) \frac{\lambda^k}{k!} = \lambda\,\mathrm{e}^{-\lambda} \left(\sum_{k=0}^{+\infty} k \frac{\lambda^k}{k!} + \mathrm{e}^{\lambda} \right) \\
&= \lambda + \lambda\,\mathrm{e}^{-\lambda} \sum_{k=1}^{+\infty} k \frac{\lambda^k}{k!} = \lambda + \lambda^2\,\mathrm{e}^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda + \lambda^2
\end{aligned}$$

Donc $V(X) = \lambda + \lambda^2 - \lambda^2 = \lambda$

3.2.7 Effet d'une homothétie ou d'une translation

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire réelle discrète telle que X^2 soit d'espérance finie.
Soient a et b deux réels.
$aX + b$ est une variable aléatoire réelle discrète dont le carré a une espérance finie et :
$V(aX + b) = a^2 V(X)$

Démonstration
$aX + b = f(X)$ est une variable aléatoire réelle discrète.
$(aX + b)^2 = a^2X^2 + 2abX + b^2$
X^2 a une espérance finie.
X a une espérance finie.
b^2 est une constante.
D'après 3.1.9, $(aX + b)^2$ a une espérance finie.
De plus :

$$\begin{aligned}
V(aX+b) &= E\left((aX+b-E(aX+b))^2\right) = E\left((aX+b-aE(X)-b)^2\right) = E\left((a(X-E(X)))^2\right) \\
&= E\left(a^2\,(X-E(X))^2\right) = a^2 E\left((X-E(X))^2\right) \\
&= a^2 V(X)
\end{aligned}$$

Exemple
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire réelle discrète telle que X^2 soit d'espérance finie et telle que $V(X) \neq 0$.
Soit $Y = \dfrac{X - E(X)}{\sigma(X)}$.
Y est une variable aléatoire réelle discrète dont le carré a une espérance finie et :
$E(Y) = 0$ et $V(Y) = 1$
En d'autres termes, Y est à la fois centrée et réduite.

3.2.8 Inégalité de Cauchy-Schwarz

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X et Y deux variables aléatoires réelle discrètes.

On suppose que X^2 et Y^2 sont d'espérances finies.
Alors XY est d'espérance finie et $E(XY)^2 \leq E(X^2)E(Y^2)$
De plus, il y a égalité si, et seulement si il existe $a \in \mathbb{R}$ tel que $P(Y = aX) = 1$ ou $b \in \mathbb{R}$ tel que $P(X = bY) = 1$

Démonstration
$\forall (x,y) \in \mathbb{R}^2 \ |xy| \leq \dfrac{x^2+y^2}{2} \leq x^2+y^2$ (classique)
Donc $|XY| \leq X^2 + Y^2$.
X^2 et Y^2 ont, par hypothèse, une espérance finie donc, par linéarité, X^2+Y^2 a une espérance finie.
On en déduit par majoration que XY a une espérance finie.

Soit $t \in \mathbb{R}$.
$(tX+Y)^2 = t^2X^2 + 2tXY + Y^2$ a une espérance finie comme combinaison linéaire de variables aléatoires d'espérance finie.
$\forall t \in \mathbb{R}\ t^2E(X^2) + 2tE(XY) + E(Y^2) = E\left((tX+Y)^2\right) \geq 0$

- **Premier cas :** $E(X^2) > 0$, ou ce qui revient au même $P(X \neq 0) > 0$
  On a affaire à un trinôme du second degré de signe constant donc :
  $\Delta = 4E(XY)^2 - 4E(X^2)E(Y^2) \leq 0$
  On en déduit $E(XY)^2 \leq E(X^2)E(Y^2)$.
- **Deuxième cas :** $X = 0$ ps
  $XY = 0$ ps donc $E(XY)^2 = 0 \leq E(X^2)E(Y^2) = 0$

Supposons qu'il y ait égalité dans l'inégalité de Cauchy-Schwarz.
Si $X = 0$ ps alors $P(X = bY) = 1$ avec $b = 0$.
Dans le cas contraire, $\Delta = 0$ donc il existe $t \in \mathbb{R}$ tel que $E((tX+Y)^2) = 0$.
Donc $P(Y = aX) = 1$ avec $a = -t$.

Supposons qu'il existe un réel a tel que $P(Y = aX) = 1$.
Si $X = 0$ ps alors comme vu ci-dessus, il y a égalité.
Sinon, $Y - aX = 0$ ps donc $a^2E(X^2) - 2aE(XY) + E(Y^2) = E((Y-aX)^2) = 0$
Donc le trinôme $t^2E(X^2) + 2tE(XY) + E(Y^2)$ a au moins une racine réelle $(-a)$ et $\Delta \geq 0$.
Comme $\Delta \leq 0$, $\Delta = 0$ et il y a égalité dans l'inégalité de Cauchy-Schwarz.

3.2.9 Covariance de deux variables aléatoires

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire réelle discrète telle que X^2 soit d'espérance finie.
Soit Y une variable aléatoire réelle discrète telle que Y^2 soit d'espérance finie.
On appelle covariance de X et de Y et on note $Cov(X,Y)$ le réel :
$Cov(X,Y) = E\left((X - E(X))(Y - E(Y))\right) = E(XY) - E(X)E(Y)$

Justification de la définition
$(X - E(X))(Y - E(Y)) = XY - E(X)Y - E(Y)X + E(X)E(Y)$
XY est une v.a.r discrète d'espérance finie.
$E(X)$ est une constante et Y une v.a.r discrète d'espérance finie.
$E(Y)$ est une constante et X une v.a.r discrète d'espérance finie.
$E(X)E(Y)$ est une constante.
Donc $(X - E(X))(Y - E(Y))$ a une espérance finie et par linéarité :

$E\left((X-E(X))(Y-E(Y))\right) = E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) = E(XY) - E(X)E(Y)$

Il résulte de ? que si X et Y sont indépendantes alors $Cov(X,Y) = 0$.
La réciproque est fausse.

3.2.10 Variance d'une somme finie de variables aléatoires

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient $X_1, \ldots, X_n$ n variables aléatoires discrètes possédant toutes une variance.
Alors $\sum_{i=1}^{n} X_i$ possède une variance et :
$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{1\leq i<j\leq n} Cov(X_i, X_j)$
En particulier, si ces variables sont deux à deux indépendantes alors :
$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i)$
Démonstration
Soit $S = \sum_{i=1}^{n} X_i$.
$S^2 = \sum_{i=1}^{n} X_i^2 + 2 \sum_{1\leq i<j\leq n} X_iX_j$.
Les X_i^2 et les X_iX_j ont toutes une espérance finie donc S^2 a une espérance finie et S a une variance.

$$\begin{aligned}
V(S) &= E(S^2) - E(S)^2 \\
&= \sum_{i=1}^{n} E(X_i^2) + 2 \sum_{1\leq i<j\leq n} E(X_iX_j) - \left(\sum_{i=1}^{n} E(X_i)\right)^2 \\
&= \sum_{i=1}^{n} E(X_i^2) + 2 \sum_{1\leq i<j\leq n} E(X_iX_j) - \sum_{i=1}^{n} E(X_i)^2 - 2 \sum_{1\leq i<j\leq n} E(X_i)E(X_j) \\
&= \sum_{i=1}^{n} \left(E(X_i^2) - E(X_i)^2\right) + 2 \sum_{1\leq i<j\leq n} \left(E(X_iX_j) - E(X_i)E(X_j)\right) \\
&= \sum_{i=1}^{n} V(X_i) + 2 \sum_{1\leq i<j\leq n} Cov(X_i, X_j)
\end{aligned}$$

Exemples

— Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires mutuellement indépendantes qui suivent toutes $\mathcal{B}(p)$.
  Pour tout $i \in \mathbb{N}^*$, soit $Y_i = X_iX_{i+1}$.
  1. Quelle est la loi de Y_i ?
  2. Soit $S_n = \sum_{i=1}^{n} Y_i$.
     Trouver l'espérance et la variance de S_n.

  Correction

1. $Y_i(\Omega) = \{0;1\}$.
$$\begin{aligned} P(Y_i = 1) &= P(X_i = 1, X_{i+1} = 1) = P(X_i = 1)P(X_{i+1} = 1) \\ &= p^2 \end{aligned}$$
$Y_i \sim \mathcal{B}(p^2)$

2. $E(S_n) = \sum_{i=1}^{n} E(Y_i) = np^2$
$V(S_n) = \sum_{i=1}^{n} V(Y_i) + 2 \sum_{1\leq i<j\leq n} Cov(Y_i, Y_j)$
$V(Y_i) = p^2(1-p^2)$ car $Y_i \sim \mathcal{B}(p^2)$.
Si $j \geq i+2$, $Cov(Y_i, Y_j) = 0$ car $Y_i = X_iX_{i+1}$ et $Y_j = X_jX_{j+1}$ sont indépendantes (cf le lemme des coalitions).
$$\begin{aligned} Cov(Y_i, Y_{i+1}) &= E(Y_iY_{i+1}) - E(Y_i)E(Y_{i+1}) = E(X_iX_{i+1}^2X_{i+2}) - p^4 \\ &= E(X_iX_{i+1}X_{i+2}) - p^4 \text{ car } X_{i+1}^2 = X_{i+1} \\ &= p^3 - p^4 \end{aligned}$$
$$\begin{aligned} V(S_n) &= np^2(1-p^2) + 2(n-1)p^3(1-p) \\ &= p^2(1-p)\left(n + np + 2(n-1)p\right) \\ &= p^2(1-p)\left((3n-2)p + n\right) \end{aligned}$$

— On considère n cartes numérotées de 1 à n.
On permute au hasard les cartes de ce jeu et on note Y la variable aléatoire égale au nombre de cartes qui occupent leur place naturelle.
On cherche l'espérance et la variance de Y.

Correction
On note $X_k = \begin{cases} 1 \text{ si la } k^{\text{ième}} \text{ carte est à sa place} \\ 0 \text{ sinon} \end{cases}$.
$X_k \sim \mathcal{B}(p)$ avec $p = P(X_k = 1) = \dfrac{(n-1)!}{n!} = \dfrac{1}{n}$ (indépendant de k, commenter, symétrie des rôles).
$Y = \sum_{k=1}^{n} X_k$ donc $E(Y) = \sum_{k=1}^{n} E(X_k) = n\,\dfrac{1}{n} = 1$
$V(Y) = V\left(\sum_{k=1}^{n} X_k\right)$ mais quid de l'indépendance ?
Pour $k \neq l$, $X_kX_l \begin{cases} 1 \text{ si la } k^{\text{ième}} \text{ carte et la } l^{\text{ième}} \text{ sont à leurs places} \\ 0 \text{ sinon} \end{cases}$.
$X_kX_l \sim \mathcal{B}(r)$ avec $r = P(X_k = 1, X_l = 1) = \dfrac{(n-2)!}{n!} = \dfrac{1}{n(n-1)}$
$$\begin{aligned} Cov(X_k, X_l) &= E(X_kX_l) - E(X_k)E(X_l) = \frac{1}{n(n-1)} - \frac{1}{n^2} = \frac{1}{n}\left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= \frac{1}{n(n-1)^2} \end{aligned}$$

Cette covariance est positive : si la $k^{\text{ième}}$ est à sa place, la $l^{\text{ième}}$ a plus de chances de l'être, cf $n = 2$ ou 3.
D'où :

$$\begin{aligned}
V(Y) &= \sum_{k=1}^{n} V(X_k) + 2 \sum_{1 \le k < l \le n} Cov(X_k, X_l) \\
&= np(1-p) + 2\,\frac{n(n-1)}{2}\,\frac{1}{n^2(n-1)} \\
&= n\,\frac{1}{n}\left(1 - \frac{1}{n}\right) + \frac{1}{n} = 1 - \frac{1}{n} + \frac{1}{n} \\
&= 1
\end{aligned}$$

Tout cela sans calcul de la loi.

Chapitre 4

Fonctions génératrices d'une variable aléatoire à valeurs dans $\mathbb{N}$

4.1 Définitions et exemples

4.1.1 Définition

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.
On appelle fonction génératrice de la variable aléatoire X la fonction G_X définie par

$$G_X(t) = E\left(t^X\right) = \sum_{n=0}^{+\infty} P(X=n)t^n$$

Remarques

— La fonction G_X est la somme d'une série entière.
$\forall n \in \mathbb{N}\ \forall t \in [-1;1]\ \ |P(X=n)t^n| \leq P(X=n)$ indépendant de t et terme général d'une série convergente.
On en déduit que la série entière converge normalement sur $[-1;1]$.
Il en résulte :
— La série entière $\sum\limits_{n\geq 0} P(X=n)t^n$ a un rayon de convergence supérieur ou égal à 1.
— La fonction G_X est continue sur $[-1;1]$.
Notons qu'on a systématiquement $G_X(0) = P(X=0)$ et $G_X(1) = 1$.

— La loi d'une variable aléatoire X à valeurs dans $\mathbb{N}$ est caractérisée par sa fonction génératrice G_X.
En effet si on connaît G_X, qui est la somme d'une série entière, on peut récupérer les nombres $P(X=n)$ pour $n \in \mathbb{N}$ qui sont ses coefficients.
Rappelons qu'on a :
$\forall n \in \mathbb{N}\ P(X=n) = \dfrac{G_X^{(n)}(0)}{n!}$.
En particulier si deux variables aléatoires à valeurs dans $\mathbb{N}$ ont la même fonction génératrice, elles ont la même loi.

4.1.2 Exemples

Je cite le programme.

Les étudiants doivent savoir calculer rapidement la fonction génératrice d'une variable aléatoire de Bernoulli, binomiale, géométrique, de Poisson.

— **Variable aléatoire constante ou presque sûrement constante**
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire telle que :
$\exists n_0 \in \mathbb{N}$ tq $P(X = n_0) = 1$
Alors G_X est définie sur $\mathbb{R}$ et :
$\forall t \in \mathbb{R}\ G_X(t) = t^{n_0}$

— **Variable aléatoire suivant une loi de Bernoulli**
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire telle que :
$P(X = 1) = p \in]0; 1[$ et $P(X = 0) = 1 - p$.
G_X est définie sur $\mathbb{R}$ et :
$\forall t \in \mathbb{R}\ G_X(t) = 1 - p + p\, t$

— **Variable aléatoire suivant une loi binomiale**
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire qui suit la loi binomiale $\mathcal{B}(n, p)$ ($n \in \mathbb{N}^*$, $p \in]0; 1[$)
G_X est définie sur $\mathbb{R}$ et :
$\forall t \in \mathbb{R}\ G_X(t) = (1 - p + p\, t)^n$

Démonstration
$G_X(t) = \displaystyle\sum_{k=0}^{n} P(X = k)t^k$ car $X(\Omega) = [\![0; n]\!]$.
G_X est donc bien définie sur $\mathbb{R}$ comme toute fonction polynômiale.

$$\begin{aligned} \forall t \in \mathbb{R}\ G_X(t) &= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} t^k = \sum_{k=0}^{n} \binom{n}{k} (pt)^k (1-p)^{n-k} \\ &= (pt + 1 - p)^n \end{aligned}$$

On verra une autre preuve plus loin.

— **Variable aléatoire suivant une loi géométrique**
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire qui suit la loi géométrique $\mathcal{G}(p)$ ($p \in]0; 1[$).
Le domaine de définition de G_X est $\left]-\dfrac{1}{1-p}; \dfrac{1}{1-p}\right[$ et :
$\forall t \in \left]-\dfrac{1}{1-p}; \dfrac{1}{1-p}\right[\ G_X(t) = \dfrac{p\, t}{1 - t\,(1-p)}$

Démonstration
G_X est la somme de la série entière $\displaystyle\sum_{n \geq 1} p(1-p)^{n-1} t^n$ ou encore $\displaystyle\sum_{n \geq 1} pt\,((1-p)t)^{n-1}$.

Donc le domaine de définition de G_X est $\left]-\dfrac{1}{1-p};\dfrac{1}{1-p}\right[$ et :
$\forall t \in \left]-\dfrac{1}{1-p};\dfrac{1}{1-p}\right[\ G_X(t) = \dfrac{p\,t}{1-t\,(1-p)}$

— **Variable aléatoire suivant une loi de Poisson**
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire qui suit la loi de Poisson $\mathcal{P}(\lambda)$ $(\lambda \in \mathbb{R})$.
G_X est définie sur $\mathbb{R}$ et :
$\forall t \in \mathbb{R}\ G_X(t) = \mathrm{e}^{\lambda(t-1)}$

Démonstration
G_X est la somme de la série entière $\displaystyle\sum_{n\geq 0} \frac{\lambda^n}{n!}\,\mathrm{e}^{-\lambda} t^n$ ou encore $\displaystyle\sum_{n\geq 0} \frac{(\lambda t)^n}{n!}\,\mathrm{e}^{-\lambda}$.
Donc le domaine de définition de G_X est $\mathbb{R}$ et :
$\forall t \in \mathbb{R}\ G_X(t) = \mathrm{e}^{\lambda(t-1)}$

4.2 Fonction génératrice et moments d'une variable aléatoire

4.2.1 Fonction génératrice et espérance d'une variable aléatoire

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.
X admet une espérance $E(X)$ si, et seulement si, G_X est dérivable en 1.
Si tel est le cas, on a $E(X) = G'_X(1)$.

Démonstration
Il s'agit de prouver :
La famille $(nP(X=n))_{n\in\mathbb{N}}$ est sommable $\iff G_X$ est dérivable en 1
Comme il s'agit famille de nombres positifs indéxée par $\mathbb{N}$, cela revient à prouver :
$\displaystyle\sum_{n\geq 0} nP(X=n)$ converge $\iff G_X$ est dérivable en 1

— On suppose que $\displaystyle\sum_{n\geq 0} nP(X=n)$ converge.
  — **Premier cas :** $R > 1$
  G_X est $\mathcal{C}^\infty$ sur $]-R;R[$, dérivable terme à terme donc G_X est dérivable en 1 et :
  $G'_X(1) = \displaystyle\sum_{n=1}^{+\infty} nP(X=n)1^{n-1} = \sum_{n=0}^{+\infty} nP(X=n) = E(X)$
  — **Deuxième cas :** $R = 1$
  Pour tout $n \in \mathbb{N}$, soit $f_n \begin{cases} [0;1] \to \mathbb{R} \\ t \mapsto P(X=n)t^n \end{cases}$.
    — Pour tout $n \in \mathbb{N}$, f_n est $\mathcal{C}^1$ sur $[0;1]$.
    — La série de fonctions $\displaystyle\sum_{n\geq 0} f_n$ converge simplement sur $[0;1]$.
    — $\forall n \in \mathbb{N}^*\ \forall t \in [0;1]\ f'_n(t) = nt^{n-1}P(X=n)$ (et $f_0 = 0$)
    $\forall n \in \mathbb{N}^*\ \|f'_n\|_\infty = \displaystyle\sup_{t\in[0;1]} |f'_n(t)| = nP(X=n)$

Donc la série de fonctions $\sum_{n\geq 0} f_n'$ converge normalement sur $[0;1]$.

Donc G_X est $\mathcal{C}^1$ sur $[0;1]$ et

$\forall t \in [0;1]\ G_X'(t) = \sum_{n=1}^{+\infty} nP(X=n)t^{n-1}$

En particulier, G_X est dérivable en 1 (G_X n'est pas définie à droite en 1) et :

$G_X'(1) = \sum_{n=1}^{+\infty} nP(X=n) = \sum_{n=0}^{+\infty} nP(X=n) = E(X)$

— La démonstration de la réciproque n'est pas exigible.
On suppose que la série $\sum_{n\geq 0} nP(X=n)$ diverge.

$R = R_{CV}\left(\sum_{n\geq 0} P(X=n)t^n\right) = R_{CV}\left(\sum_{n\geq 1} nP(X=n)t^{n-1}\right) \leq 1$

Comme on sait que ce rayon de convergence est supérieur ou égal à 1, on a : $R = 1$.

$$\begin{aligned} \forall t \in [0;1[\ G_X'(t) &= \sum_{n=1}^{+\infty} nP(X=n)t^{n-1} \\ G_X''(t) &= \sum_{n=2}^{+\infty} n(n-1)P(X=n)t^{n-2} \geq 0 \end{aligned}$$

Donc G_X' est croissante sur $[0;1[$.

Donc $G_X'(t) \underset{\substack{t\to 1\\ t<1}}{\longrightarrow} l = \sup_{t\in[0;1[} G_X'(t) \in [0;+\infty]$

Supposons $l < +\infty$:

$\forall N \in \mathbb{N}\ \forall t \in [0;1[\ \sum_{n=1}^{N} \left(nP(X=n)t^{n-1} \geq 0\right) \leq \sum_{n=1}^{+\infty} nP(X=n)t^{n-1}$

et :

$\forall t \in [0;1[\ \sum_{n=1}^{+\infty} nP(X=n)t^{n-1} \leq l$

Donc :

$\forall N \in \mathbb{N}\ \forall t \in [0;1[\ \sum_{n=1}^{N} nP(X=n)t^{n-1} \leq l$

On fait tendre t vers 1 à N fixé :

$\forall N \in \mathbb{N}\ \sum_{n=1}^{N} nP(X=n) \leq l < +\infty$

La suite des sommes partielles de la série à termes positifs $\sum_{n\geq 0} nP(X=n)$ est majorée.

Donc $\sum_{n\geq 0} nP(X=n)$ converge : absurde.

Donc $G_X'(t) \underset{\substack{t\to 1\\ t<1}}{\longrightarrow} +\infty$.

G_X étant continue sur $[0;1]$ on peut invoquer le théorème de la limite de la dérivée pour obtenir :

$\dfrac{G_X(t) - G_X(1)}{t-1} \underset{\substack{t\to 1\\ t<1}}{\longrightarrow} +\infty$.

Donc G_X n'est pas dérivable en 1.

Par contrapposition :

$$G_X \text{ dérivable en } 1 \Longrightarrow \sum_{n\geq 0} nP(X=n) \text{ converge}$$

4.2.2 Fonction génératrice et variance d'une variable aléatoire

Je cite le programme.

Utilisation de G_X pour calculer $E(X)$ et $V(X)$.

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.

On a sans problème :

$$\begin{array}{rcl}\forall t \in]-1;1[\ G'_X(t) & = & \displaystyle\sum_{n=1}^{+\infty} nP(X=n)t^{n-1} \\ G''_X(t) & = & \displaystyle\sum_{n=2}^{+\infty} n(n-1)P(X=n)t^{n-2} \geq 0\end{array}$$

Supposons que X^2 a une espérance finie (c'est la condition de définition de $V(X)$ vue plus haut).
X a alors une espérance finie (donc G_X est dérivable en 1).
Donc $\displaystyle\sum_{n\geq 0} nP(X=n)$ et $\displaystyle\sum_{n\geq 0} n^2P(X=n)$ convergent.
Donc $\displaystyle\sum_{n\geq 0} n(n-1)P(X=n)$ converge.
On a alors comme précédemment, G_X est deux fois dérivable en 1 et :

$$\begin{array}{rcl}G''_X(1) & = & \displaystyle\sum_{n=0}^{+\infty} n(n-1)P(X=n) \\ & = & E(X(X-1)) \text{ par le théorème de transfert} \\ & = & E(X^2-X) = E(X^2)-E(X)\end{array}$$

Mais $V(X) = E(X^2) - E(X)^2$ donc :

$$\begin{array}{rcl}G''_X(1) & = & V(X) + E(X)^2 - E(X) \\ & = & V(X) + G'_X(1)^2 - G'_X(1)\end{array}$$

et :

$$V(X) = G''_X(1) - G'_X(1)^2 + G'_X(1)$$

4.2.3 Exemple

Considérons le programme suivant :

```
def recherche_maximum(t):
    maxi=t[0]
    for v in t:
        if v>maxi:
            maxi=v
    return(maxi)
```

Combien d'affectations effectue-t-il ? En d'autres termes, combien de fois s'éxécute `maxi=v` ?

Nous allons modéliser la situation de la manière suivante :
On recherche le maximum de n nombres deux à deux distincts.
Il y a $n!$ permutations de ces entiers, on les suppose équiprobables.
Dans ce cadre, le nombre d'affectations est une variable aléatoire. Il s'agit d'en déterminer l'espérance.

Soit $n \geq 1$.
Soit $\mathcal{E}_n = [\![0; n[\![$.
Soit Ω_n l'ensemble des bijections de $\mathcal{E}_n$ sur lui-même.
(Bien sûr, les t_i à trier ne sont pas forcément compris entre 0 et $n-1$. Il faudrait introduire $\mathcal{T}_n = \{x_0; \dots; x_{n-1}\}$ avec $x_i < x_{i+1}$ et une bijection φ de $\mathcal{E}_n$ sur $\mathcal{T}_n$, par exemple $i \mapsto x_i$. Ω_n serait $\{\varphi\sigma\varphi^{-1}, \sigma \in \sigma([\![0; n-1]\!])\}$ mais il s'agit d'un modèle)
Soit $A_n \begin{cases} \Omega_n \to [\![1; n]\!] \\ \omega \in \Omega_n \mapsto \text{ nombre d'affectations dans le tri de } \omega(0)\omega(1)\dots\omega(n-1) \end{cases}$.
$A_n(\omega) = \mathrm{Card}\left(\left\{i \in [\![1; n[\![\text{ tq } \omega(i) > \max\limits_{0 \leq j \leq i-1} \omega(j)\right\}\right) + 1$ (qui correspond à l'affectation initiale).
A_n est une variable aléatoire et on cherche son espérance.
$A_n(\Omega_n) = [\![1; n]\!]$.
On note $p_{n,k} = P(A_n = k)$ avec par convention $p_{n,0} = 0$ et :
$\forall k \geq n+1 \ p_{n,k} = 0$
On note $X_n \begin{cases} \Omega_n \to [\![0; n-1]\!] \\ \omega \mapsto \omega(n-1) \end{cases}$.
$\omega \in (X_n = j) \iff \omega(n-1) = j$ et ω réalise une bijection entre $[\![0; n-2]\!]$ et $[\![0; n[\![\backslash\{j\}$
$\mathrm{Card}\left((X_n = j)\right) = (n-1)!$
$P(X_n = j) = \dfrac{(n-1)!}{\mathrm{Card}(\Omega_n)} = \dfrac{1}{n}$

$$\begin{aligned}
\forall n \geq 2 \ \forall k \in \mathbb{N}^* \ p_{n,k} &= P(A_n = k) \\
&= \sum_{l=0}^{n-1} P(A_n = k | X_n = l) P(X_n = l) \text{ probabilités totales} \\
&= \frac{1}{n} \sum_{l=0}^{n-1} P(A_n = k | X_n = l) \\
&= \frac{1}{n} \sum_{l=0}^{n-2} P(A_n = k | X_n = l) \text{ où il n'y a pas d'affectation pour le dernier} \\
&\quad + \frac{1}{n} P(A_n = k | X_n = n-1) \text{ où il y a une affectation pour le dernier} \\
&= \frac{1}{n} \sum_{l=0}^{n-2} P(A_{n-1} = k) + \frac{1}{n} P(A_{n-1} = k-1) \\
&= \frac{n-1}{n} p_{n-1,k} + \frac{1}{n} p_{n-1,k-1}
\end{aligned}$$

Soit G_{A_n} la fonction génératrice de A_n.

$$\begin{aligned}
\forall t \in \mathbb{R}\ G_{A_n}(t) &= \sum_{k=1}^{+\infty} p_{n,k} t^k \quad p_{n,0} = 0 \text{ et la somme est en fait finie} \\
&= \sum_{k=1}^{+\infty} \left(\left(1 - \frac{1}{n}\right) p_{n-1,k} + \frac{1}{n} p_{n-1,k-1} \right) t^k \\
&= \left(1 - \frac{1}{n}\right) G_{A_{n-1}}(t) + \frac{1}{n} \sum_{k=1}^{+\infty} p_{n-1,k-1} t^k \\
&= \left(1 - \frac{1}{n}\right) G_{A_{n-1}}(t) + \frac{t}{n} \sum_{l=0}^{+\infty} p_{n-1,l} t^l \\
&= \left(1 - \frac{1}{n} + \frac{t}{n}\right) G_{A_{n-1}}(t) = \frac{t+n-1}{n} G_{A_{n-1}}(t)
\end{aligned}$$

D'où :
$\forall n \geq 2\ \forall t \in \mathbb{R}\ G_{A_n}(t) = \prod_{k=1}^{n} \frac{t+k-1}{k}\ (G_{A_1}(t) = t) = \frac{1}{n!} \prod_{k=0}^{n-1} (t+k)$
La loi de A_n est "connue" : $P(A_n = l)$ est le coefficient de t^l dans $G_{A_n}(t)$.

$\forall t > 0\ \ln(G_{A_n}(t)) = \sum_{k=0}^{n-1} \ln(t+k) - \ln(n!)$
On dérive :
$\forall t > 0\ \frac{G'_{A_n}(t)}{G_{A_n}(t)} = \sum_{k=0}^{n-1} \frac{1}{t+k}$
D'où :
$\frac{G'_{A_n}(1)}{G_{A_n}(1) = 1} = \sum_{k=0}^{n-1} \frac{1}{1+k} = \sum_{k=1}^{n} \frac{1}{k}$
Donc :

$$E(A_n) = \sum_{k=1}^{n} \frac{1}{k} \sim \ln(n)$$

4.3 Fonctions génératrices et sommes de variables aléatoires indépendantes

4.3.1 Fonction génératrice de la somme de deux variables aléatoires indépendantes

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient X et Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$.
On a : $G_{X+Y} = G_X \times G_Y$
Plus précisément si on note R_X le rayon de convergence de la fonction génératrice de X et R_Y celui de Y, on a :
$\forall t \in]-\min(R_X, R_Y); \min(R_X, R_Y)[\ G_{X+Y}(t) = G_X(t) \times G_Y(t)$
En particulier, on a dans tous les cas :
$\forall t \in]-1; 1[\ G_{X+Y}(t) = G_X(t) \times G_Y(t)$
Cette formule est en fait vraie dans tous les cas sur $[-1; 1]$.

Démonstration

— **Première preuve**

$$\begin{aligned}
& \forall t \in]-\min(R_X, R_Y); \min(R_X, R_Y)[\ G_X(t) \times G_Y(t) \\
= & \left(\sum_{n=0}^{+\infty} P(X=n)t^n\right) \times \left(\sum_{n=0}^{+\infty} P(Y=n)t^n\right) \\
= & \sum_{n=0}^{+\infty} \left(\sum_{p=0}^{n} P(X=p)P(Y=n-p)\right) t^n \\
& \text{ cf le cours sur les séries entières} \\
= & \sum_{n=0}^{+\infty} P(X+Y=n)t^n \\
= & G_{X+Y}(t)
\end{aligned}$$

— **Deuxième preuve**

Soit $t \in]-\min(R_X, R_Y); \min(R_X, R_Y)[$.
X et Y sont indépendantes donc t^X et t^Y aussi
Comme t^X et t^Y ont une espérance finie :

$$\begin{aligned}
G_{X+Y}(t) &= E(t^{X+Y}) = E(t^X\ t^Y) = E(t^X)\ E(t^Y) \\
&= G_X(t) \times G_Y(t)
\end{aligned}$$

Enfin pour l'extension à $[-1; 1]$, on utilise la continuité des séries génératrices sur $[-1; 1]$.

4.3.2 Exemple d'application

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient X et Y deux variables aléatoires discrètes indépendantes sur $(\Omega, \mathcal{A}, P)$.
On suppose $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$ $((\lambda, \mu) \in (\mathbb{R}_+^*)^2)$.
Alors $X + Y \sim \mathcal{P}(\lambda + \mu)$.

Démonstration

$$\begin{aligned}
\forall t \in \mathbb{R}\ G_{X+Y}(t) &= G_X(t) \times G_Y(t) = \mathrm{e}^{\lambda(t-1)}\ \mathrm{e}^{\mu(t-1)} = \mathrm{e}^{\lambda(t-1)+\mu(t-1)} \\
&= \mathrm{e}^{(\lambda+\mu)(t-1)} \\
&= G_Z(t) \text{ si } Z \sim \mathcal{P}(\lambda + \mu)
\end{aligned}$$

Donc $X + Y \sim \mathcal{P}(\lambda + \mu)$
(Rappelons que la fonction génératrice caractérise la loi)

4.3.3 Fonction génératrice de la somme de plusieurs variables aléatoires indépendantes

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soient $X_1, \ldots, X_n$ n variables aléatoires indépendantes à valeurs dans $\mathbb{N}$.
On a $G_{X_1+\cdots+X_n} = \displaystyle\prod_{i=1}^{n} G_{X_i}$

Plus précisément si on note R_{X_i} le rayon de convergence de la série génératrice de X_i, on a :
$\forall t \in]-\min\limits_{1\leq i\leq n}(R_{X_i}); \min\limits_{1\leq i\leq n}(R_{X_i})[\ G_{X_1+\cdots+X_n}(t) = \prod\limits_{i=1}^{n} G_{X_i}(t)$
En particulier, on a dans tous les cas :
$\forall t \in [-1;1]\ G_{X_1+\cdots+X_n}(t) = \prod\limits_{i=1}^{n} G_{X_i}(t)$

Ce résultat se déduit par récurrence immédiate et utilisation du lemme des coalitions de 4.3.1.

4.3.4 Exemple d'application

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire qui suit la loi binomiale $\mathcal{B}(n,p)$ ($n \in \mathbb{N}$, $p \in]0;1[$)
G_X est définie sur $\mathbb{R}$ et :
$\forall t \in \mathbb{R}\ G_X(t) = (1-p+p\,t)^n$

Démonstration
Il n'est pas forcément possible d'écrire $X = Y_1+\cdots+Y_n$ avec $Y_1, \ldots, Y_n$ indépendantes et suivant la loi de Bernoulli de paramètre p.
Par contre, il est possible de construire $(\Omega', \mathcal{A}', P')$ un espace probabilisé sur lequel existent $Y_1, \ldots, Y_n$ n variables aléatoires mutuellement indépendantes suivant la loi de Bernoulli de paramètre p.
$Y_1 + \cdots + Y_n \sim \mathcal{B}(n,p)$
Donc $G_X = G_{Y_1+\ldots Y_n} = G_{Y_1}^n$

Chapitre 5

Inégalités probabilistes

5.1 Inégalité de Markov

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire réelle discrète d'espérance finie.
On suppose en outre :

i $X(\Omega) \subset \mathbb{R}_+$

ii $m = E(X) > 0$

Alors :

$$\forall \lambda \in \mathbb{R}_+^* \; P(X \geq \lambda\, m) \leq \frac{1}{\lambda}$$

Démonstration
Soit $Y = 1_{X \geq \lambda m}$.
$Y \sim \mathcal{B}(p)$ avec $p = P(X \geq \lambda m)$.
$\forall \omega \in \Omega \; \lambda m Y(\omega) = \begin{cases} 0 \leq X(\omega) \text{ si } \omega \notin (X \geq \lambda m) \\ \lambda m \leq X(\omega) \text{ si } \omega \in (X \geq \lambda m) \end{cases}$
Donc :
$\forall \omega \in \Omega \; \lambda m Y(\omega) \leq X(\omega)$
Donc $\lambda m Y \leq X$

En pratique, on rédige plus rapidement :

$$\begin{aligned} \lambda m Y &= \lambda m 1_{X \geq \lambda m} \\ &= \begin{cases} 0 \leq X \quad (X \geq 0) \text{ si } X < \lambda m \\ \lambda m \leq X \text{ si } X \geq \lambda m \end{cases} \\ &\leq X \text{ dans tous les cas} \end{aligned}$$

Par croissance et linéarité de l'espérance :
$\lambda m E(Y) \leq E(X)$
$\lambda m P(X \geq \lambda m) \leq m$
Après simplification par $m > 0$:
$P(X \geq \lambda m) \leq \dfrac{1}{\lambda}$ Cette inégalité n'est pas d'une grande précision :
Si $\lambda < 1$, on écrit $P(X \geq \lambda m) \leq \dfrac{1}{\lambda}$ avec $\dfrac{1}{\lambda} > 1$.

5.2 Inégalité de Bienaymé-Tchebychev

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit X une variable aléatoire réelle discrète telle que X^2 soit d'espérance finie.
On a alors, en notant m l'espérance de X et σ^2 sa variance :

$$\forall a > 0 \; P\left(|X - m| \geq a\right) \leq \frac{\sigma^2}{a^2}$$

Démonstration
$Y = (X - m)^2$ a une espérance finie (cf ?).
$E(Y) = V(X) = \sigma^2$

- **Premier cas :** $\sigma > 0$
  D'après l'inégalité de Markov :
  $\forall \lambda > 0 \; P(Y \geq \lambda\sigma^2) \leq \dfrac{1}{\lambda}$
  On prend $\lambda = \dfrac{a^2}{\sigma^2} > 0$
  $\forall a > 0 \; P\left((X - m)^2 \geq a^2\right) \leq \dfrac{\sigma^2}{a^2}$
  D'où :
  $\forall a > 0 \; P\left(|X - m| \geq a\right) \leq \dfrac{\sigma^2}{a^2}$
- **Deuxième cas :** $\sigma = 0$
  $V(X) = \sigma^2 = 0$
  $P(X = m) = 1$
  $\forall a > 0 \; P\left(|X - m| \geq a\right) = 0 \leq \dfrac{\sigma^2}{a^2}$

5.3 Remarque

Ces inégalités ont un intérêt essentiellement théorique. Elles fournissent des majorations très médiocres des probabilités considérées.
Supposons par exemple $X \sim \mathcal{B}\left(10, \dfrac{1}{2}\right)$.
$E(X) = 10 \times \dfrac{1}{2} = 5$ et $\sigma^2 = V(X) = 10 \times \dfrac{1}{2} \times \left(1 - \dfrac{1}{2}\right) = \dfrac{5}{2}$
On a donc :
$P\left(|X - 5| \geq 4\right) \leq \dfrac{5}{32} \simeq 0,156$
Mais :

$$\begin{aligned} P\left(|X - 5| \geq 4\right) &= P(X = 0) + P(X = 1) + P(X = 9) + P(X = 10) \\ &= \left(\binom{10}{0} + \binom{10}{1} + \binom{10}{9} + \binom{10}{10}\right)\left(\frac{1}{2}\right)^{10} \\ &= \frac{22}{2^{10}} = \frac{11}{512} \simeq 0,021 \end{aligned}$$

5.4 Loi faible des grands nombres

5.4.1 Théorème

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.
Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d ayant une variance finie.
On note $S_n = \sum_{k=1}^{n} X_k$, $m = E(X_1)$ et $\sigma = \sigma(X_1)$. Alors :

i $\forall \epsilon > 0\ P\left(\left|\frac{1}{n}S_n - m\right| \geq \epsilon\right) \xrightarrow[n\to+\infty]{} 0$

ii Plus précisément :
$$\forall \epsilon > 0\ P\left(\left|\frac{1}{n}S_n - m\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}$$

Démonstration
$E\left(\frac{S_n}{n}\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} m = m$
$V\left(\frac{S_n}{n}\right) = \frac{1}{n^2}V\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i)$ car les X_i sont mutuellement indépendantes mais on peut remarquer qu'il suffit qu'elles soient deux à deux indépendantes.
$V\left(\frac{S_n}{n}\right) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$
On applique alors l'inégalité de Bienaymé-Tchebychev :
$\forall \epsilon > 0\ P\left(\left|\frac{1}{n}S_n - m\right| \geq \epsilon\right) \leq \frac{\sigma^2/n}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$

5.4.2 Remarques

— Je cite le programme :

> Les étudiants doivent savoir retrouver, avec $\sigma = \sigma(X_1)$:
> $$P\left(\left|\frac{1}{n}S_n - m\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}$$
> On peut affaiblir les hypothèses : il suffir que les X_n soient deux à deux indépendantes et aient toutes la même espérance et la même variance.

5.5 Un exemple

Soit $\alpha \in \mathbb{R}_+$.
On considère la série entière $\sum_{n\geq 0} \frac{n^\alpha}{n!}x^n$.
Pour tout $n \in \mathbb{N}$, on note $a_n = \frac{n^\alpha}{n!}$.
Soit $r \in \mathbb{R}_+^*$.
$\forall n \in \mathbb{N}^*\ a_n r^n > 0$ et :
$\frac{a_{n+1}r^{n+1}}{a_n r^n} = \left(1 + \frac{1}{n}\right)^\alpha \frac{r}{n+1} \xrightarrow[n\to+\infty]{} 0 < 1$

D'après la règle de d'Alembert, la série de terme général $a_n r^n$ converge pour tout $r > 0$.
Le rayon de convergence de série entière $\displaystyle\sum_{n\geq 0} \frac{n^\alpha}{n!} x^n$ est donc infini.

On en déduit par linéarité que pour tout polynôme P la série entière $\displaystyle\sum_{n\geq 0} \frac{P(n)}{n!} x^n$ a un rayon de convergence infini.

Si $\alpha = p$ est un entier naturel, le calcul de la somme est facile.
On pose $H_0 = 1$ et pour tout $k \in \mathbb{N}^*$, $H_k = \displaystyle\prod_{l=0}^{k-1} (X - l)$.
La famille $(H_0, \dots, H_p)$ est une famille de polynômes de $\mathbb{R}_p[X]$.
Elle est échelonnée en degré donc elle est libre.
Elle a $p + 1 = \dim(\mathbb{R}_p[X])$ éléments donc c'est une base de $\mathbb{R}_p[X]$ et :
$\exists!(a_0, \dots, a_p) \in \mathbb{R}^{p+1}$ tq $X^p = \displaystyle\sum_{l=0}^{p} a_l H_l$.
Par comparaison des coefficients dominants, on a $a_p = 1$.

$$
\begin{aligned}
\forall x \in \mathbb{R} \ \sum_{n=0}^{+\infty} \frac{n^p}{n!} x^n &= \sum_{n=0}^{+\infty} \left(\sum_{l=0}^{p} \frac{a_l H_l(n)}{n!} x^n \right) \\
&= \sum_{l=0}^{p} \left(a_l \sum_{n=0}^{+\infty} \frac{H_l(n)}{n!} x^n \right) \text{ par simple linéarité} \\
&= \sum_{l=0}^{p} \left(a_l \sum_{n=l}^{+\infty} \frac{n(n-1)\dots(n-l+1)}{n!} x^n \right) \\
&= \sum_{l=0}^{p} \left(a_l \sum_{n=l}^{+\infty} \frac{1}{(n-l)!} x^n \right) \\
&= \sum_{l=0}^{p} \left(a_l x^l \, \mathrm{e}^x \right) \\
&= \mathrm{e}^x \sum_{l=0}^{p} a_l x^l
\end{aligned}
$$

On en déduit :

$$\forall p \in \mathbb{N} \ \sum_{n=0}^{+\infty} \frac{n^p}{n!} x^n \sim_{x\to+\infty} x^p \, \mathrm{e}^x$$

On va montrer :
$\displaystyle\sum_{n=0}^{+\infty} \frac{\sqrt{n}}{n!} x^n \sim_{x\to+\infty} \sqrt{x} \, \mathrm{e}^x$
mais le calcul de la somme n'est plus faisable.
Il s'agit de prouver :
$\displaystyle\frac{1}{\mathrm{e}^x \sqrt{x}} \sum_{n=0}^{+\infty} \frac{\sqrt{n}}{n!} x^n = \frac{1}{\sqrt{x}} \sum_{n=0}^{+\infty} \frac{\sqrt{n}}{n!} x^n \, \mathrm{e}^{-x} \xrightarrow[x\to+\infty]{} 0$
ou encore :
$\forall \epsilon > 0 \ \exists A > 0$ tq $\forall x \geq A \ \left| \displaystyle\frac{1}{\sqrt{x}} \sum_{n=0}^{+\infty} \frac{\sqrt{n}}{n!} x^n \, \mathrm{e}^{-x} - 1 \right| \leq \epsilon$

On remarque que pour $x > 0$, $\displaystyle\sum_{n=0}^{+\infty} \frac{\sqrt{n}}{n!} x^n \, \mathrm{e}^{-x} = E(\sqrt{X_x})$ où X_x est une variable aléatoire qui suit la loi de Poisson de paramètre x.
Soit $\epsilon > 0$.
Soit $\eta \in]0;1[$ à choisir plus tard en fonction de ϵ et seulement en fonction de ϵ.

$$\begin{aligned}
\forall x > 0 \ \left|\frac{1}{\sqrt{x}}\sum_{n=0}^{+\infty} \frac{\sqrt{n}}{n!} x^n \, \mathrm{e}^{-x} - 1\right| = \left|\frac{1}{\sqrt{x}}E(\sqrt{X_x}) - 1\right| = \left|E\left(\sqrt{\frac{X_x}{x}}\right) - 1\right| \\
= \quad \left|E\left(\sqrt{\frac{X_x}{x}}1_{|X_x - x| < \eta x}\right) - 1 + E\left(\sqrt{\frac{X_x}{x}}1_{|X_x - x| \geq \eta x}\right)\right| \\
\leq \quad \left|E\left(\sqrt{\frac{X_x}{x}}1_{|X_x - x| < \eta x}\right) - 1\right| + E\left(\sqrt{\frac{X_x}{x}}1_{|X_x - x| \geq \eta x}\right)
\end{aligned}$$

$\sqrt{1-\eta}1_{|X_x - x| < \eta x} \leq \sqrt{\frac{X_x}{x}}1_{|X_x - x| < \eta x} \leq \sqrt{1+\eta}1_{|X_x - x| < \eta x}$
En effet si $|X_x - x| \geq \eta x$ alors l'inégalité s'écrit $0 \leq 0 \leq 0$ et
si $|X_x - x| < \eta x$ alors $-\eta x < X_x - x < \eta x$ donc $(1-\eta)x < X_x < (1+\eta)x$ puis
$\sqrt{1-\eta} < \sqrt{\frac{X_x}{x}} < \sqrt{1+\eta}$ qui est l'inégalité voulue.
On en déduit :
$\sqrt{1-\eta}P\left(|X_x - x| < \eta x\right) \leq E\left(\sqrt{\frac{X_x}{x}}1_{|X_x - x| < \eta x}\right) \leq \sqrt{1+\eta}P\left(|X_x - x| < \eta x\right)$
Donc :
$\sqrt{1-\eta}\left(1 - P\left(|X_x - x| \geq \eta x\right)\right) \leq E\left(\sqrt{\frac{X_x}{x}}1_{|X_x - x| < \eta x}\right) \leq \sqrt{1+\eta}\left(1 - P\left(|X_x - x| \geq \eta x\right)\right)$
Donc :
$\sqrt{1-\eta} - 1 - \sqrt{1-\eta}P\left(|X_x - x| \geq \eta x\right) \leq E\left(\sqrt{\frac{X_x}{x}}1_{|X_x - x| < \eta x}\right) - 1 \leq \sqrt{1-\eta} - 1 - \sqrt{1+\eta}P\left(|X_x - x| \geq \eta x\right)$
Mais par inégalité triangulaire :
$|\sqrt{1-\eta} - 1 - \sqrt{1-\eta}P\left(|X_x - x| \geq \eta x\right)| \leq 1 - \sqrt{1-\eta} + P\left(|X_x - x| \geq \eta x\right)$
et :
$|\sqrt{1+\eta} - 1 - \sqrt{1+\eta}P\left(|X_x - x| \geq \eta x\right)| \leq \sqrt{1+\eta} - 1 + \sqrt{1+\eta}P\left(|X_x - x| \geq \eta x\right)$
On en déduit :
$\left|E\left(\sqrt{\frac{X_x}{x}}1_{|X_x - x| < \eta x}\right) - 1\right| \leq \max\left(1 - \sqrt{1-\eta}, \sqrt{1+\eta} - 1\right) + \sqrt{1+\eta}P\left(|X_x - x| \geq \eta x\right)$
Par ailleurs avec Cauchy-Schwarz :

$$\begin{aligned}
E\left(\sqrt{\frac{X_x}{x}}1_{|X_x - x| \geq \eta x}\right) \quad &\leq \quad \left(E\left(\left(\sqrt{\frac{X_x}{x}}\right)^2\right)\right)^{1/2} \left(E\left(\left(1_{|X_x - x| \geq \eta x}\right)^2\right)\right)^{1/2} \\
&\leq \quad \left(E\left(\frac{X_x}{x}\right)\right)^{1/2} \left(E\left(1_{|X_x - x| \geq \eta x}\right)\right)^{1/2} \\
&\leq \quad \sqrt{P\left(|X_x - x| \geq \eta x\right)}
\end{aligned}$$

On a donc :

$$\begin{aligned}&\forall x > 0 \ \left| \frac{1}{\sqrt{x}} \sum_{n=0}^{+\infty} \frac{\sqrt{n}}{n!} x^n \, \mathrm{e}^{-x} - 1 \right| \\ \leq \ & \max\left(1 - \sqrt{1-\eta}, \sqrt{1+\eta} - 1\right) + \sqrt{1+\eta} P\left(|X_x - x| \geq \eta x\right) + \sqrt{P\left(|X_x - x| \geq \eta x\right)}\end{aligned}$$

$\max\left(1 - \sqrt{1-\eta}, \sqrt{1+\eta} - 1\right) \xrightarrow[\eta \to 0^+]{} 0$ donc :

$\exists \eta \in]0;1[$ tq $\max\left(1 - \sqrt{1-\eta}, \sqrt{1+\eta} - 1\right) \leq \dfrac{\epsilon}{3}$

On a alors :

$\forall x > 0 \ \sqrt{1+\eta} P\left(|X_x - x| \geq \eta x\right) \leq \sqrt{1+\eta} \dfrac{V(X_x)}{\eta^2 x^2} = \sqrt{1+\eta} \dfrac{1}{\eta^2 x}$

$\sqrt{1+\eta} \dfrac{1}{\eta^2 x} \xrightarrow[x \to +\infty]{} 0$ donc :

$\exists x_1 > 0$ tq $\forall x \geq x_1 \ \sqrt{1+\eta} \dfrac{1}{\eta^2 x} \leq \dfrac{\epsilon}{3}$

Enfin :

$\sqrt{P\left(|X_x - x| \geq \eta x\right)} \leq \dfrac{1}{\eta \sqrt{x}}$

$\dfrac{1}{\eta \sqrt{x}} \xrightarrow[x \to +\infty]{} 0$ donc :

$\exists x_2 > 0$ tq $\forall x \geq x_2 \ \dfrac{1}{\eta \sqrt{x}} \leq \dfrac{\epsilon}{3}$

Finalement, si on note $x_0 = \max(x_1, x_2)$ alors :

$\forall x \geq x_0 \ \left| \dfrac{1}{\sqrt{x}} \displaystyle\sum_{n=0}^{+\infty} \dfrac{\sqrt{n}}{n!} x^n \, \mathrm{e}^{-x} - 1 \right| \leq \epsilon$

Annexe A

Ensembles dénombrables

A.1 Définition des ensembles dénombrables

Un ensemble est dit dénombrable si, et seulement si, il est en bijection avec $\mathbb{N}$ ie s'il peut être écrit en extension $\{x_i, i \in \mathbb{N}\}$ avec des x_i distincts.
Un ensemble est dit au plus dénombrable si, et seulement si, il est en bijection avec une partie I de $\mathbb{N}$ ie s'il peut être écrit en extension $\{x_i, i \in I\}$ avec I une partie de $\mathbb{N}$ et des x_i distincts.
Les ensembles au plus dénombrables sont les ensembles finis ou dénombrables.

A.2 Exemples et propriétés

A.2.1 Parties d'un ensemble dénombrable

Une partie d'un ensemble dénombrable est au plus dénombrable.

La démonstration n'est pas au programme.

A.2.2 $\mathbb{Z}$ est dénombrable

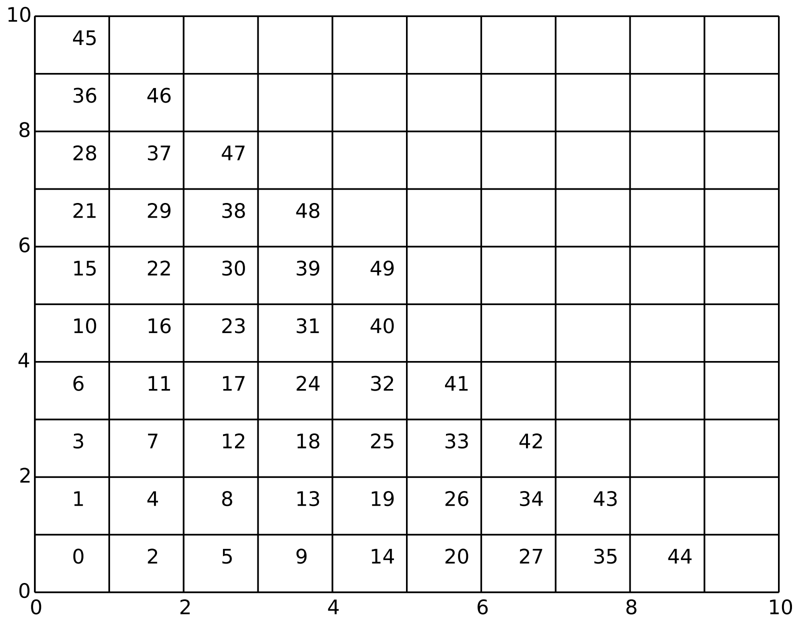
Ce résultat est explicitement au programme mais pas sa démonstration.
$\begin{pmatrix} \dots & -2 & -1 & 0 & 1 & 2 & \dots \\ \dots & 3 & 1 & 0 & 2 & 4 & \dots \end{pmatrix}$
Soit $\varphi \begin{cases} \mathbb{N} \to \mathbb{Z} \\ n \mapsto n/2 \text{ si } n \text{ est pair} \\ n \mapsto -(n+1)/2 \text{ si } n \text{ est impair} \end{cases}$
En d'autres termes :
$\forall p \in \mathbb{N}\ \varphi(2p) = p$
$\forall p \in \mathbb{N}^*\ \varphi(2p-1) = -p$
φ est bien une application de $\mathbb{N}$ dans $\mathbb{Z}$.
φ est injective :
Supposons $\varphi(n_1) = \varphi(n_2)$.
n_1 et n_2 ont la même parité : si n est pair, $\varphi(n) \geq 0$ et si n est impair, $\varphi(n) < 0$.
Si n_1 et n_2 sont pairs, alors $\dfrac{n_1}{2} = \dfrac{n_2}{2}$ et $n_1 = n_2$.
Si n_1 et n_2 sont impairs, alors $-\dfrac{n_1+1}{2} = -\dfrac{n_2+1}{2}$ et $n_1 = n_2$.

φ est surjective :
Si $n \in \mathbb{N}$, $n = \varphi(2\,n)$.
Si $n \in \mathbb{Z}_-^*$ alors $n = \varphi\left(2(-n) - 1\right)$ et $2(-n) - 1 \in \mathbb{N}$.

φ est donc une bijection de $\mathbb{N}$ sur $\mathbb{Z}$ et $\mathbb{Z}$ est dénombrable.

A.2.3 Produit cartésien d'un nombre fini d'ensembles dénombrables

Si $E_1, \ldots, E_n$ sont n ensembles dénombrables alors $E_1 \times \cdots \times E_n$ est un ensemble dénombrable.
Ce résultat est mentionné dans le programme.
Par contre, la démonstration n'est pas au programme.
Elle se fait par récurrence sur n.
La propriété est triviale pour $n = 1$ et se démontre comme suit pour $n = 2$.
On commence par montrer que $\mathbb{N}^2$ est dénombrable en explicitant une bijection de $\mathbb{N}$ sur $\mathbb{N}^2$.



(i, j) se trouve sur la même couche que $(0, i + j)$.
Sur les couches précédentes, il y a eu :
$1 + 2 + \cdots + (i + j) = \dfrac{(i + j)(i + j + 1)}{2}$ éléments.
Sur la couche de $(0, i + j)$, (i, j) est le $(i + 1)^{\text{ème}}$.
Donc $\varphi_1 \begin{cases} \mathbb{N}^2 \to \mathbb{N}^* \\ (i, j) \mapsto \dfrac{(i + j)(i + j + 1)}{2} + i + 1 \end{cases}$ et $\varphi \begin{cases} \mathbb{N}^2 \to \mathbb{N} \\ (i, j) \mapsto \dfrac{(i + j)(i + j + 1)}{2} + i \end{cases}$ sont des bijections.

Justifions le rigoureusement.
φ va bien de $\mathbb{N}^2$ dans $\mathbb{N}$: $(i + j)(i + j + 1)$ est pair.
Supposons $\varphi(i_1, j_1) = \varphi(i_2, j_2)$.

Supposons $i_1 + j_1 < i_2 + j_2$.

$$\begin{aligned} i_1 - i_2 &= \frac{(i_2+j_2)(i_2+j_2+1)}{2} - \frac{(i_1+j_1)(i_1+j_1+1)}{2} \\ &= \sum_{k=0}^{i_2+j_2} k - \sum_{k=0}^{i_1+j_1} k \\ &= \sum_{k=i_1+j_1+1}^{i_2+j_2} k \\ &\geq i_1 + j_1 + 1 \end{aligned}$$

Donc $i_1 \geq i_1 + (i_2 + j_1 \geq 0) + 1 > i_1$.
On aboutit à une absurdité.
Idem si on suppose $i_1 + j_1 > i_2 + j_2$.
Donc $i_1 + j_1 = i_2 + j_2$.
$\varphi(i_1, j_1) = \varphi(i_2, j_2)$ donne alors $i_1 = i_2$, puis $j_1 = j_2$.
φ est injective.

Soit $N \in \mathbb{N}$.
Soit $(s_n)_{n\in\mathbb{N}} = \left(\frac{n(n+1)}{2}\right)_{n\in\mathbb{N}}$.

— $s_0 = 0$
— $(s_n)_{n\in\mathbb{N}}$ est strictement croissante
— $s_n \xrightarrow[n\to+\infty]{} +\infty$

Donc :
$\exists! n_0 \in \mathbb{N}$ tq $s_{n_0} \leq N < s_{n_0+1}$
$0 \leq N - s_{n_0} < s_{n_0+1} - s_{n_0} = n_0 + 1$
On pose $i = N - s_{n_0} \in [\![0; n_0]\!]$.
$j = n_0 - i$ est bien défini.
$(i, j) \in \mathbb{N}^2$ et $\varphi(i, j) = N$.
φ est surjective.

Donc φ est bijective.

Soient ensuite E et F deux ensembles dénombrables.
Il existe e une bijection de E sur $\mathbb{N}$ et f une bijection de F sur $\mathbb{N}$.
$\begin{cases} E \times F \to \mathbb{N}^2 \\ (x, y) \mapsto (e(x), f(y)) \end{cases}$ est une bijection de $E \times F$ sur $\mathbb{N}^2$. Comme $\mathbb{N}^2$ est en bijection avec $\mathbb{N}$, $E \times F$ est en bijection avec $\mathbb{N}$.

Enfin, pour passer du rang n au rang $n+1$, on remarque que $\begin{cases} E_1 \times \cdots \times E_{n+1} \to (E_1 \times \cdots \times E_n) \times E_{n+1} \\ (x_1, \ldots, x_{n+1}) \mapsto ((x_1, \ldots, x_n), x_{n+1}) \end{cases}$ est une bijection.

Remarque
$\mathbb{Q}$ est un ensemble dénombrable.
Ce résultat n'est pas mentionné dans le programme.
Le principe de la démonstration est le suivant :

tout rationnel strictement positif r s'écrit de manière unique
$r = \dfrac{p}{q}$ où $(p,q) \in \mathbb{N}^* \times \mathbb{N}^*$ et $PGCD(p,q) = 1$, ce qui permet de mettre en bijection $\mathbb{Q}_+^*$ et une partie de $\mathbb{N}^2$.

A.2.4 Remarque

Il existe des ensembles infinis qui ne sont pas dénombrables.
Cette propriété n'est pas mentionnée dans le programme et ne peut donc a priori pas faire l'objet de questions pendant les concours. Il s'agit toutefois d'un point de culture générale et il me paraît indispensable d'en avoir entendu parler.
L'exemple le plus frappant est celui de $\mathbb{R}$: $\mathbb{R}$ est un ensemble infini non dénombrable.
En effet, il a été vu en première année que tout réel appartenant à $[0;1[$ a un, et un seul, développement décimal propre ie $x \in [0;1[$ s'écrit de manière unique $x = 0, c_1 c_2 \dots c_n \dots$ avec $c_n \in [\![0;9]\!]$ et c_n qui n'est pas égal à 9 à partir d'un certain rang.
Supposons $\mathbb{R}$ dénombrable.
Il existerait une bijection φ de $\mathbb{N}^*$ sur $[0;1[$.
Pour tout $n \in \mathbb{N}^*$, $\varphi(n) = 0, c_1^n c_2^n \dots$ (développement décimal propre de $\varphi(n)$).
Soit $(d_n)_{n\in\mathbb{N}^*}$ la suite définie par $\begin{cases} d_n = 1 \text{ si } c_n^n = 0 \\ d_n = 0 \text{ si } c_n^n \neq 0 \end{cases}$.
Le nombre $0, d_1 d_2 \dots$ est dans $[0;1[$ et il est différent de tous les $\varphi(n)$.
On aboutit donc à une contradiction.

A.2.5 Union au plus dénombrable d'ensembles dénombrables

Une union au plus dénombrable d'ensembles dénombrables est un ensemble dénombrable.
Plus précisément, soit I un ensemble non vide au plus dénombrable et $(E_i)_{i\in I}$ une famille d'ensembles dénombrables. Alors :
$\displaystyle\bigcup_{i\in I} E_i$ est un ensemble dénombrable.
La démonstration est hors programme.

Annexe B

Familles sommables

B.1 Familles sommables de réels positifs

B.1.1 Définitions

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

On dit que la famille $(x_i)_{i\in I}$ est sommable si, et seulement si, $\left\{\sum_{j\in J} x_j, J \text{ partie finie de } I\right\}$ est majoré.

Si la famille $(x_i)_{i\in I}$ est sommable, sa somme est $\sum_{i\in I} x_i = \sup_{\substack{J\subset I\\ J \text{ finie}}} \left(\sum_{j\in J} x_j\right) \in \mathbb{R}$.

Si la famille $(x_i)_{i\in I}$ n'est pas sommable, par convention sa somme est $\sum_{i\in I} x_i = +\infty$

On généralise cette notion au cas où les x_i peuvent prendre la valeur $+\infty$:
si l'un des x_i est infini, la famille $(x_i)_{i\in I}$ n'est pas sommable et par convention sa somme est $\sum_{i\in I} x_i = +\infty$

B.1.2 Lien avec les séries

— Si $(x_i)_{i\in I}$ est une famille finie de réels positifs, sa somme $\sum_{i\in I} x_i$ telle qu'on vient de la définir est la même que sa somme habituelle, somme qui peut être calculée en additionnant les éléments de la famille dans n'importe quel ordre.

— Soit $(x_n)_{n\in\mathbb{N}}$ une suite de réels positifs.
$(x_n)_{n\in\mathbb{N}}$ est sommable $\Longleftrightarrow$ la série $\sum_{n\geq 0} x_n$ converge

Dans ce cas, la somme $\sum_{n\in\mathbb{N}} x_n$ telle qu'on vient de la définir est égale à la somme $\sum_{n=0}^{+\infty} x_n$ de la série.

— Soit I un ensemble dénombrable et $(x_i)_{i\in I}$ une famille de réels positifs.
Soit φ une bijection de $\mathbb{N}$ sur I ie une énumération des éléments de I : $I = \{\varphi(0); \varphi(1); \ldots\}$.
$(x_i)_{i\in I}$ est sommable $\Longleftrightarrow$ la série $\sum_{n\geq 0} x_{\varphi(n)}$ converge

Dans ce cas : $\sum_{i\in I} x_i = \sum_{n=0}^{+\infty} x_{\varphi(n)}$
On voit donc que dans une famille sommable l'ordre de la sommation n'a pas d'importance.

B.1.3 Sommation par paquets

Soit $(x_i)_{i\in I}$ une famille au plus dénombrable d'éléments de $[0;+\infty]$.
On suppose que I se décompose en $I = \bigcup_{k\in K} I_k$ où K est au plus dénombrable et les I_k sont deux à deux disjoints.
$\sum_{i\in I} x_i = \sum_{k\in K}\left(\sum_{i\in I_k} x_i\right)$
Rappelons que toutes ces sommes peuvent prendre la valeur $+\infty$.

B.1.4 Théorème de Fubini

Soient I et J deux ensembles au plus dénombrables.
Soit $(x_{i,j})_{(i,j)\in I\times J}$ une famille de nombres réels positifs.
$\sum_{(i,j)\in I\times J} x_{i,j} = \sum_{i\in I}\left(\sum_{j\in J} x_{i,j}\right) = \sum_{j\in J}\left(\sum_{i\in I} x_{i,j}\right)$
ces sommes pouvant être égales à $+\infty$.

B.2 Familles sommables de nombres complexes

B.2.1 Définition

Soit $(x_i)_{i\in I}$ une famille au plus dénombrable de nombres complexes.
On dit que la famille $(x_i)_{i\in I}$ est sommable si, et seulement si, la famille (au plus dénombrable de réels positifs) $(|x_i|)_{i\in I}$ est sommable.

Si les x_i sont tous réels, on définit $y_i = \sup(0, x_i) \in \mathbb{R}_+$ et $z_i = -\inf(0, x_i) \in \mathbb{R}_+$ et on a $x_i = y_i - z_i$ et $|x_i| = y_i + z_i$.
$\sum_{i\in I} y_i \leq \sum_{i\in I} |x_i| < +\infty$ donc la famille $(y_i)_{i\in I}$ est sommable.
De même la famille $(z_i)_{i\in I}$ est une famille sommable de réels positifs.
$\sum_{i\in I} y_i$ et $\sum_{i\in I} z_i$ sont donc définies et on définit $\sum_{i\in I} x_i$ par $\sum_{i\in I} x_i = \sum_{i\in I} y_i - \sum_{i\in I} z_i$
Dans le cas général, on pose $y_i = \Re e(x_i)$ et $z_i = \Im m(x_i)$.
$\sum_{i\in I} |y_i| \leq \sum_{i\in I} \sqrt{y_i^2 + z_i^2} = \sum_{i\in I} |x_i| < +\infty$ donc la famille $(y_i)_{i\in I}$ est sommable.
De même la famille $(z_i)_{i\in I}$ est une famille sommable de nombres réels.
$\sum_{i\in I} y_i$ et $\sum_{i\in I} z_i$ sont donc définies et on définit $\sum_{i\in I} x_i$ par $\sum_{i\in I} x_i = \sum_{i\in I} y_i + \mathrm{i}\sum_{i\in I} z_i$

B.2.2 Lien avec les séries

— Si $(x_i)_{i\in I}$ est une famille finie de nombres complexes, c'est une famille sommable et sa somme $\sum_{i\in I} x_i$ telle qu'on vient de la définir est la même que sa somme habituelle, somme qui peut être calculée en additionnant les éléments de la famille dans n'importe quel ordre.

— Soit $(x_n)_{n\in\mathbb{N}}$ une suite de nombres complexes.
$(x_n)_{n\in\mathbb{N}}$ est sommable $\iff$ la série $\sum_{n\geq 0} x_n$ converge absolument

Dans ce cas, la somme $\sum_{n\in\mathbb{N}} x_n$ telle qu'on vient de la définir est égale à la somme $\sum_{n=0}^{+\infty} x_n$ de la série.

— Soit I un ensemble dénombrable et $(x_i)_{i\in I}$ une famille de nombres complexes.
Soit φ une bijection de $\mathbb{N}$ sur I ie une énumération des éléments de I : $I = \{\varphi(0); \varphi(1); \ldots\}$.
$(x_i)_{i\in I}$ est sommable $\iff$ la série $\sum_{n\geq 0} x_{\varphi(n)}$ converge absolument

Dans ce cas : $\sum_{i\in I} x_i = \sum_{n=0}^{+\infty} x_{\varphi(n)}$
On voit donc que dans une famille sommable l'ordre de la sommation n'a pas d'importance.

B.2.3 Propriétés

— **Croissance**
Soit $(x_i)_{i\in I}$ une famille au plus dénombrable de nombres complexes.
Soit $(y_i)_{i\in I}$ une famille de réels positifs.
On suppose :
— $\forall i \in I \ \ |x_i| \leq y_i$
— La famille $(y_i)_{i\in I}$ est sommable
Alors la famille $(x_i)_{i\in I}$ est sommable.

Soient $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ deux familles sommables de nombres réels telle que :
$\forall i \in I \ x_i \leq y_i$
On a : $\sum_{i\in I} x_i \leq \sum_{i\in I} y_i$

— **Linéarité**
Soit $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ deux familles au plus dénombrables de nombres complexes.
Si ces deux familles sont sommables alors pour tout $(\lambda, \mu) \in \mathbb{C}^2$, la famille $(\lambda x_i + \mu y_i)_{i\in I}$ est sommable et :
$\sum_{i\in I} (\lambda x_i + \mu y_i) = \lambda \sum_{i\in I} x_i + \mu \sum_{i\in I} y_i$

— **Sommation par paquets**
Soit $(x_i)_{i\in I}$ une famille au plus dénombrable de nombres complexes.
On suppose que I se décompose en $I = \bigcup_{k\in K} I_k$ où K est au plus dénombrable et les I_k sont deux à deux disjoints.
Il y a équivalence entre

(i) La famille $(x_i)_{i\in I}$ est sommable.

(ii) Pour tout $k \in K$, la famille $(x_i)_{i \in I_k}$ est sommable et la famille $\left(\sum_{i \in I_k} |x_i|\right)_{k \in K}$ est sommable, ce qui peut s'écrire $\sum_{k \in K} \left(\sum_{i \in I_k} |x_i|\right) < +\infty$.

On a alors :
$\sum_{i \in I} x_i = \sum_{k \in K} \left(\sum_{i \in I_k} x_i\right)$

— **Théorème de Fubini**
Soient I et J deux ensembles au plus dénombrables.
Soit $(x_{i,j})_{(i,j) \in I \times J}$ une famille de nombres complexes.

$$\begin{aligned} (x_{i,j})_{(i,j) \in I \times J} \text{ sommable} \quad & \Longleftrightarrow \quad \sum_{i \in I} \left(\sum_{j \in J} |x_{i,j}|\right) < +\infty \\ & \Longleftrightarrow \quad \sum_{j \in J} \left(\sum_{i \in I} |x_{i,j}|\right) < +\infty \end{aligned}$$

On a alors :
$\sum_{(i,j) \in I \times J} x_{i,j} = \sum_{i \in I} \left(\sum_{j \in J} x_{i,j}\right) = \sum_{j \in J} \left(\sum_{i \in I} x_{i,j}\right)$

— **Produit de deux sommes**
Soit $(x_i)_{i \in I}$ une famille au plus dénombrable et sommable de nombres complexes.
Soit $(y_j)_{j \in J}$ une famille au plus dénombrable et sommable de nombres complexes.
Alors la famille $(x_i y_j)_{(i,j) \in I \times J}$ est sommable et :
$\left(\sum_{i \in I} x_i\right) \times \left(\sum_{j \in J} y_j\right) = \sum_{(i,j) \in I \times J} x_i y_j$