8-PROBABILITES et VARIABLES ALEATOIRES

A) Probabilité sur un univers dénombrable.

1) Ensembles dénombrables.

Définitions : Soit un ensemble E.

- E est fini si et seulement si $E=\varnothing$ ou qu'il existe $n\in\mathbb{N}^*$ et une bijection de $[\![1,n]\!]$ sur E.
  On peut alors écrire $E=\{x_1,...,x_n\}$, avec $x_1,...,x_n$ distincts.
- E est dénombrable si et seulement s'il existe une bijection de $\mathbb{N}$ sur E. On peut alors écrire $E=\{x_i, i\in\mathbb{N}\}$, où les x_i sont distincts.
- E est au plus dénombrable si et seulement s'il est fini ou dénombrable.

Exemples :

- $2\mathbb{N}=\{2n, n\in\mathbb{N}\}$ est dénombrable.
- $\mathbb{Z}$ est dénombrable.

Preuve rapide :

On construit l'application bijective de $\mathbb{N}$ dans $\mathbb{Z}$ définie sur $\mathbb{N}$ par :

- Si $n=2p$, $x(n)=x_n=p$
- Si $n=2p+1$, $x(n)=x_n=-p-1$

Alors $f(0)=0$, $f(1)=-1$, $f(2)=1\ldots$

Remarque : on peut montrer que $\mathbb{R}$ n'est pas dénombrable.

Propriétés admises : soient $E, F, E_1, ..., E_n$ des ensembles dénombrables. Alors :

- $E \times F$ est dénombrable.
- $E_1 \times ... \times E_n$ est dénombrable.
- Toute partie de E est au plus dénombrable.
- Toute union dénombrable d'ensembles au plus dénombrable est dénombrable.

Explication : On a $E = \{e_n, n \in \mathbb{N}\}$ et $F = \{f_n, n \in \mathbb{N}\}$.
Alors $E \times F = \{(e_0, f_0), (e_1, f_0), (e_0, f_1), (e_2, f_0), (e_1, f_1), (e_0, f_2), ...\}$

2) <u>Familles sommables.</u>

Définition : soit I et E deux ensembles.

- Une famille $(x_i)_{i\in I}$ d'éléments de E est une application de I dans E.
- Lorsque I est fini, on dit que la famille est finie.
- Lorsque I est dénombrable, on dit que la famille est dénombrable.
- Une famille au plus dénombrable est une famille finie ou dénombrable.

Propriété admise : soit $I \subset \mathbb{N}$ et $(x_i)_{i\in I}$ une famille au plus dénombrable de $[0, +\infty]$.
On peut alors définir $\sum_{i\in I} x_i \in [0, +\infty]$ telle que :

- Si $\exists i \in I, x_i = +\infty$, alors $\sum_{i\in I} x_i = +\infty$
- Si I est dénombrable, avec $I = \{i_n, n \in \mathbb{N}\}$, et que $\forall i \in I, x_i \in \mathbb{R}_+$ alors $\sum_{i\in I} x_i = +\infty$ si et seulement si la série $\sum_{n\in\mathbb{N}} x_{i_n}$ est divergente.
- Lorsque la série $\sum_{n\in\mathbb{N}} x_{i_n}$ est convergente, on dit que la famille $(x_i)_{i\in I}$ est **sommable** et $\sum_{i\in I} x_i = \sum_{n=0}^{+\infty} x_{i_n} < +\infty$.
- Pour tout découpage en paquets de I sous la forme $I = \bigcup_{n\in J} I_n$, avec J au plus dénombrable et les $(I_n)_{n\in J}$ deux à deux disjoints, alors $\sum_{i\in I} x_i = \sum_{n\in J}\left(\sum_{i\in I_n} x_i\right)$. C'est la sommation par paquets.

Remarque : si on prend $I=\mathbb{N}^*$ et $x_n=\dfrac{(-1)^n}{n}$, la série $\sum\limits_{n\in\mathbb{N}^*} x_n$ converge, mais on ne peut pas regrouper les termes comme on veut en faisant des paquets : par exemple, $\sum\limits_{n=1}^{+\infty}\dfrac{(-1)^n}{n}=\sum\limits_{\substack{n=1\\ n\text{ pair}}}^{+\infty}\dfrac{(-1)^n}{n}+\sum\limits_{\substack{n=1\\ n\text{ impair}}}^{+\infty}\dfrac{(-1)^n}{n}$ n'a pas de sens.

En pratique :

- Quand on a des nombres **positifs,** on peut les ajouter dans l'ordre qu'on veut sans se préoccuper de savoir si la série converge ou pas.
- On peut faire tous calculs souhaités sur les éléments de la somme, toute majoration sans se préoccuper de la convergence. Si on obtient à la fin une somme finie, cela justifie la sommabilité.
- Cette manière de procéder est réservée au contexte probabiliste.

Définition : soit $(x_i)_{i\in I}$ une famille au plus dénombrable d'éléments de $\mathbb{C}$. $(x_i)_{i\in I}$ est sommable si et seulement si $(|x_i|)_{i\in I}$ est sommable.

Remarque : soit $(x_i)_{i\in I}$ une famille au plus dénombrable, que l'on suppose sommable.

- Si $\forall i\in I, x_i\in\mathbb{R}$, on pose pour $i\in I$: $x_i^+=\max(x_i,0)$ et $x_i^-=\max(-x_i,0)$.
  Alors on pose $\sum\limits_{i\in I} x_i=\sum\limits_{i\in I} x_i^+-\sum\limits_{i\in I} x_i^-$
- Si $\forall i\in I, x_i\in\mathbb{C}$, on pose $\sum\limits_{j\in I} x_i=\sum\limits_{j\in I}\mathrm{Re}(x_j)+i\sum\limits_{j\in I}\mathrm{Im}(x_j)$

Propriétés : soit $(x_i)_{i\in I}$ une famille au plus dénombrable d'éléments de $\mathbb{C}$.

- Si $I=\mathbb{N}$, $(x_i)_{i\in I}$ est sommable si et seulement si $\sum\limits_{i\in\mathbb{N}} x_i$ est absolument convergente.
- Si $(y_i)_{i\in I}$ est une famille d'éléments de $[0,+\infty]$, que $\forall i\in I, |x_i|\le y_i$, et que la famille $(y_i)_{i\in I}$ est sommable, alors $(x_i)_{i\in I}$ est sommable.

Propriétés admises des familles sommables : Soient $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ deux familles sommables de nombres complexes. Soient $\lambda,\mu\in\mathbb{C}$. Alors :

- Linéarité : $\sum\limits_{i\in I}(\lambda x_i+\mu y_i)=\lambda\sum\limits_{i\in I} x_i+\mu\sum\limits_{i\in I} y_i$
- Croissance : si $\forall i\in I, 0\le x_i\le y_i$, alors $\sum\limits_{i\in I} x_i\le\sum\limits_{i\in I} y_i$.

Proposition : Sommation par paquets. Soit $(x_i)_{i\in I}$ une famille sommable de nombres complexes. Pour tout découpage en paquets de I sous la forme $I=\bigcup\limits_{n\in J} I_n$, avec J au plus dénombrable et les $\left(I_n\right)_{n\in J}$ deux à deux disjoints, alors :

- Pour tout $n\in J$, la famille $(x_i)_{i\in I_n}$ est sommable de somme notée $\sum\limits_{i\in I_n} x_i$.
- $\sum\limits_{i\in I} x_i=\sum\limits_{n\in J}\left(\sum\limits_{i\in I_n} x_i\right)$.

Théorème de Fubini : soient I,J deux ensembles au plus dénombrables.

- On suppose que la famille de nombres complexes $(x_{i,j})_{(i,j)\in I\times J}$ est sommable. Alors $\sum\limits_{(i,j)\in I\times J} x_{i,j}=\sum\limits_{i\in I}\left(\sum\limits_{j\in J} x_{i,j}\right)=\sum\limits_{j\in J}\left(\sum\limits_{i\in I} x_{i,j}\right)$.
- On suppose que la famille $(x_{i,j})_{(i,j)\in I\times J}$ est constituée d'éléments de $\left[0,+\infty\right]$. Alors $\sum\limits_{(i,j)\in I\times J} x_{i,j}=\sum\limits_{i\in I}\left(\sum\limits_{j\in J} x_{i,j}\right)=\sum\limits_{j\in J}\left(\sum\limits_{i\in I} x_{i,j}\right)$. Il est ici possible que ces quantités soient égales à $+\infty$.

Produit de deux sommes : Soient I,J deux ensembles au plus dénombrables Soient $(x_i)_{i\in I}$ et $(y_j)_{j\in J}$ deux familles sommables de nombres complexes. Alors la famille $(x_i,y_j)_{(i,j)\in I\times J}$ est sommable et $\sum\limits_{(i,j)\in I\times J} x_i y_j=\left(\sum\limits_{i\in I} x_i\right)\left(\sum\limits_{j\in J} y_j\right)$.

3) Espace probabilisé.

Définitions : soit Ω un ensemble. Soit $(A_n)_{n\in\mathbb{N}}$ une famille de sous-ensembles de Ω. Alors :

- $\bigcap\limits_{n\in\mathbb{N}} A_n=\bigcap\limits_{n=0}^{+\infty} A_n$ est l'ensemble des éléments $\omega\in\Omega$ tels que $\forall n\in\mathbb{N},\omega\in A_n$.
- $\bigcup\limits_{n\in\mathbb{N}} A_n=\bigcup\limits_{n=0}^{+\infty} A_n$ est l'ensemble des éléments $\omega\in\Omega$ tels que $\exists n\in\mathbb{N},\omega\in A_n$.

Propriété : soit $n\in\mathbb{N}^*$ et soit $A,\left(B_n\right)_{n\in\mathbb{N}}$ des sous-ensembles d'un ensemble E. On a les propriétés suivantes :

- $A\cup(\bigcap\limits_{n\in\mathbb{N}} B_n)=\bigcap\limits_{n\in\mathbb{N}}(A\cup B_n)$
- $A\cap(\bigcup\limits_{n\in\mathbb{N}} B_n)=\bigcup\limits_{n\in\mathbb{N}}(A\cap B_n)$
- $\overline{\bigcap\limits_{n\in\mathbb{N}} B_n}=\bigcup\limits_{n\in\mathbb{N}}\overline{B_n}$ (ici, pour $n\in\mathbb{N}$, $\overline{B_n}=\Omega\setminus B_n$ désigne le complémentaire de B dans Ω)

Preuve du premier et du dernier :

Définition : si Ω est un ensemble, on appelle tribu sur Ω toute partie T de l'ensemble $P(\Omega)$ des parties de Ω telle que :

- $\Omega \in T$.
- Pour tout $A \in T$, $\overline{A} = \Omega \setminus A \in T$.
- Pour toute suite $(A_i)_{i\in I}$, au plus dénombrable d'éléments de T , $\bigcup_{n\in\mathbb{N}} A_n \in T$ (stabilité par union au plus dénombrable)

Lorsqu'un ensemble Ω est muni d'une tribu T , alors on dit que (Ω,T) est un espace probabilisable.

Remarques :

- En pratique, on ne s'occupe presque jamais du choix de la tribu et de l'univers : on suppose simplement qu'il existe un univers Ω et une tribu sur Ω adaptés à la description de notre expérience.
- Si T est une tribu sur Ω, et si (A_n) est une suite d'éléments de T , alors $\bigcap_{n\in\mathbb{N}} A_n \in T$. En effet, $\overline{\bigcap_{n\in\mathbb{N}} A_n} = \bigcup_{n\in\mathbb{N}} \overline{A_n} \in T$.

Définitions : soit Ω un ensemble et T une tribu sur Ω. On suppose que Ω est l'ensemble des résultats possibles d'une expérience aléatoire. Ω est appelé **univers.**

- Les **événements** sont les éléments de la tribu T.
- L'événement $\overline{A}$ est appelé **événement contraire** de A. Il est réalisé si et seulement si A ne l'est pas.
- L'événement certain est Ω. L'événement impossible est l'ensemble vide $\varnothing$.
- Deux événements A et B sont dits **incompatibles** si et seulement si $A \cap B$ est l'ensemble vide.
- Le fait que $B \subset A$ signifie que si B se produit, alors A se produit également.
- L'événement $A \cap B$ se produit lorsque A **et** B se produisent.
- L'événement $A \cup B$ se produit lorsque A **ou** B se produit (plus précisément lorsqu'au moins un des deux événements se produit).
- L'événement $\bigcap_{n\in\mathbb{N}} A_n$ se produit lorsque **pour tout** $n \in \mathbb{N}$, A_n se produit.
- L'événement $\bigcup_{n\in\mathbb{N}} A_n$ se produit lorsqu'**il existe** $n \in \mathbb{N}$ tel que A_n se produit.

Exemples (*) : soit Ω un ensemble muni d'une tribu T. Soit $(A_n)_{n\in\mathbb{N}}$ une famille d'éléments de T. Soit $\omega \in \Omega$. Traduire en langage courant les événements suivants :

1) $\omega \in \bigcap_{n\in\mathbb{N}} A_n$
2) $\omega \in \bigcup_{n\in\mathbb{N}} A_n$
3) $\omega \in \bigcap_{n\in\mathbb{N}} \left(\bigcup_{p\geq n} A_p \right)$
4) $\omega \in \bigcup_{n\in\mathbb{N}} \left(\bigcap_{p\geq n} \overline{A_p} \right)$

Exemple : On joue à Pile ou Face indéfiniment.
Pour $k \in \mathbb{N}$, on note F_k : « on obtient Face au k-ème lancer ».
Ecrire avec des ensembles les événements suivants :
F : « On a obtenu au moins une fois Face au cours de l'ensemble des lancers »
FF : « On a obtenu au moins une fois deux fois Face consécutivement »
I : « On a obtenu une infinité de fois Face »

Définitions : Soit Ω un ensemble et T une tribu sur Ω. Soit E un ensemble.

- Une **variable aléatoire discrète** à valeurs dans E est une application $X : \Omega \to E$ telle que $X(\Omega)$ est au plus dénombrable et telle que $\forall x \in X(\Omega), X^{-1}(\{x\}) \in T$.
- L'ensemble des valeurs prises par la variable aléatoire X est noté $X(\Omega) \subset E$.
- Lorsque $X(\Omega) \subset \mathbb{R}$, on dit que X est une **variable aléatoire discrète réelle.**
- Si $U \subset X(\Omega)$, l'événement $X^{-1}(U) = \{\omega \in \Omega, X(\omega) \in U\}$ est noté $(X \in U)$ ou $\{X \in U\}$. Ainsi, formellement, $(X \in U)$ est l'image réciproque de l'ensemble U par l'application X.
- L'événement $X^{-1}(\{x_k\}) = \{\omega \in \Omega, X(\omega) = x_k\}$ est noté $(X = x_k)$ ou $\{X = x_k\}$.
- Lorsque $X(\Omega) \subset \mathbb{R}$ et $x \in \mathbb{R}$, l'événement $(X \geq x)$ est $\{\omega \in \Omega, X(\omega) \geq x\}$.

Exemples :

1) **Fonction indicatrice**. Soit Ω un ensemble et T une tribu sur Ω. On considère $A \subset \Omega$. La fonction indicatrice de A, notée 1_A est la variable aléatoire définie sur Ω par $\forall \omega \in A, \left| \begin{array}{l} 1_A(\omega) = 1 \text{ si } \omega \in A \\ 1_A(\omega) = 0 \text{ si } \omega \notin A \end{array} \right.$.

2) **Temps d'attente d'un premier succès :** on joue une infinité de fois à pile ou face. On considère la variable aléatoire X égale au nombre de lancers nécessaires pour obtenir le premier Pile. En adoptant la convention $X = +\infty$ lorsqu'on n'obtient jamais Pile, X est à valeurs dans $\mathbb{N}^* \cup \{+\infty\}$.

   Pour $i \in \mathbb{N}^*$, on note F_i : « on a obtenu Face au i-ème lancer ».

   Pour $k \in \mathbb{N}^*$, expliciter les événements $(X = k)$, $(X > k)$ et $(X = +\infty)$.

Définition : Soit Ω un ensemble et T une tribu sur Ω. On appelle **probabilité** sur (Ω, T) toute application $\mathrm{P} : T \to [0,1]$ telle que :

- $\mathrm{P}(\Omega) = 1$
- Pour toute famille $(A_i)_{i \in I}$, au plus dénombrable d'événements incompatibles (c'est-à-dire deux à deux disjoints), alors $\sum_{n \in \mathbb{N}} \mathrm{P}(A_n)$ converge et $\mathrm{P}(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} \mathrm{P}(A_n)$.

(Ω, T, P) est alors un espace probabilisé.

Si $A \subset \Omega$, la probabilité de A est alors notée $\mathrm{P}(A) \in [0,1]$.

Remarque : si l'univers $\Omega = \{\omega_n, n \in \mathbb{N}\}$ est dénombrable, alors il est impossible que les événements élémentaires $\{\omega_n\}$ soient équiprobables.

Propriétés (*) : soit (Ω, T, P) un espace probabilisé. Alors :

- Si $A \in T$, alors $P(\overline{A}) = 1 - P(A)$
- Si A et B sont deux événements incompatibles, alors $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$
- En général, si $A, B \in T$, $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - P\left(A \cap B\right)$
- Si $A \subset B$, $\mathrm{P}(A) \leq \mathrm{P}(B)$ (P est croissante pour l'inclusion).
- Si $A, B \in T$, on note $A \setminus B = A \cap \overline{B}$. Alors $P(A \setminus B) = P(A) - P(A \cap B)$. En particulier, lorsque $B \subset A$, $P(A \setminus B) = P(A) - P(B)$

Preuve non faite.

Exemple : soit X une variable aléatoire discrète réelles.
Montrer que $\mathrm{P}\left(\left|X-2\right| \geq 1\right) \geq \mathrm{P}(X \geq 3)$.

Proposition : continuité croissante ou décroissante. Soit (Ω, T, P) un espace probabilisé. Soit $(A_n)_{n\in\mathbb{N}}$ une famille d'éléments de T .

1) On suppose $\forall n \in \mathbb{N}, A_n \subset A_{n+1}$ (union croissante). Alors $\mathrm{P}(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to+\infty} P\left(A_n\right)$.
2) On suppose $\forall n \in \mathbb{N}, A_{n+1} \subset A_n$ (union décroissante). Alors $\mathrm{P}(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to+\infty} P\left(A_n\right)$.

Preuve du 1) : on fait un dessin.

Pour $n \in \mathbb{N}^*$, on pose $B_n = A_n \setminus A_{n-1} = A_n \cap \overline{A_{n-1}}$. On montre que les $\left(B_i\right)_{i\in\mathbb{N}^*}$ sont deux à deux disjoints et que $\bigcup_{n\in\mathbb{N}} A_n = A_0 \cup \left(\bigcup_{n\in\mathbb{N}^*} B_n\right)$. Si $\omega \in \bigcup_{n\in\mathbb{N}} A_n$, on note $p = \min\left\{k \in \mathbb{N}, \omega \in A_k\right\}$

On peut déduire le 2) du 1) à l'aide d'un passage au complémentaire.

Corollaire (*) : Soit (Ω, T, P) un espace probabilisé. Soit $(A_n)_{n\in\mathbb{N}}$ une famille d'éléments de T.

1) $\mathrm{P}(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to+\infty} P\left(\bigcup_{k=0}^{n} A_k\right)$.
2) $\mathrm{P}(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to+\infty} P\left(\bigcap_{k=0}^{n} A_k\right)$.

Preuve :

On considère $(C_n)_{n\in\mathbb{N}}$ définie pour $n \in \mathbb{N}$ par $C_n = \bigcup_{k=0}^{n} A_k$. Alors $\bigcup_{n\in\mathbb{N}} C_n = \bigcup_{n\in\mathbb{N}} A_n$.

Exemple :

Un elfe immortel prend une carte au hasard dans un jeu de 32 cartes. Il répète l'opération une infinité de fois.

a) Quelle est la probabilité de prendre à chaque fois un cœur ?
b) Quelle est la probabilité de prendre au moins une fois l'as de cœur ?
c) Quelle est la probabilité de prendre une infinité de fois l'as de cœur ?

Pour $n \in \mathbb{N}^*$, on considère les événements suivants :

A_n : « on a pris un cœur au n-ème tirage ».

B_n : « on a pris l'As de cœur au n-ème tirage ».

A : « on a pris un cœur à chaque tirage ».

B : « on a pris au moins une fois l'As de cœur »

C : « on a pris un nombre infini de fois l'As de cœur »

Pour écrire C, on pourra remarquer que si $\omega \in \Omega$, $\omega \in C \Leftrightarrow \forall n \in \mathbb{N}, \exists k \geq n, \omega \in B_k$.

Définition : soit (Ω, T, P) un espace probabilisé. Soit $A \in T$. Alors :

- On dit que A est vérifié presque sûrement (ou que A est presque sûr) si et seulement si $P(A) = 1$.
- On dit que A est négligeable si et seulement si $P(A) = 0$.

Proposition : Soit (Ω, T, P) un espace probabilisé.
Soit $(A_n)_{n\in\mathbb{N}}$ une famille d'éléments de T . Soit $N \in \mathbb{N}$. Alors :

- $P\left(\bigcup_{n=0}^{N} A_n\right) \leq \sum_{n=0}^{N} P(A_n)$
- $P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} P(A_n)$ (avec la notation $\sum_{n=0}^{+\infty} P(A_n) = +\infty$ si la série diverge).

Preuve rapide : par récurrence pour le premier et en passant l'inégalité à la limite pour le second.

Remarques :

1) La seconde propriété n'a d'intérêt que si $\sum_{n=0}^{+\infty} P(A_n) < 1$
2) Une union dénombrable d'événement négligeables est négligeable.
3) Une intersection dénombrable d'événements presque sûrs est presque sûre.

3) Conditionnement

Définition : Soit (Ω, T, P) un espace probabilisé, et $B \in T$ un événement de probabilité non nulle. Soit $A \in T$. On appelle probabilité conditionnelle de A sachant B le nombre $\mathrm{P}_B(A) = \dfrac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$. On note également $\mathrm{P}_B(A) = \mathrm{P}(A / B)$.

Remarques :

1) Cela revient à changer d'univers et à considérer que B est vérifié.
2) Il ne faut pas confondre $\mathrm{P}(A / B)$ (on calcule la probabilité d'avoir A en supposant que B est vérifié) avec $\mathrm{P}(A \cap B)$ (on calcule la probabilité d'avoir à la fois A et B en considérant toutes les possibilités). Ainsi, $\mathrm{P}(B / B) = 1$, alors que $\mathrm{P}(B \cap B) = \mathrm{P}(B)$.
3) Si B est de probabilité non nulle, $\mathrm{P}(A \cap B) = \mathrm{P}(B).\mathrm{P}(A / B) = \mathrm{P}(B).\mathrm{P}_B(A)$

Exemple : dans un jeu de 32 cartes, on prend 2 cartes au hasard. Déterminer la probabilité d'avoir pris deux cœurs dont le roi de cœur, et la probabilité d'avoir pris deux cœurs sachant qu'on a pris le roi de cœur.
R : « on a pris le roi de cœur ».
D : « on a pris deux cœurs »

Propriété : Soit (Ω, T, P) un espace probabilisé, et $B \in T$ un événement de probabilité non nulle. Alors P_B est une probabilité sur Ω.
Par conséquent, on a, pour $A \subset \Omega$:

1) $\mathrm{P}(\overline{A} / B) = 1 - \mathrm{P}(A / B)$ (ou encore $\mathrm{P}_B(\overline{A}) = 1 - \mathrm{P}_B(A)$).
2) Pour toute suite (A_n) événements incompatibles (c'est-à-dire deux à deux disjoints), alors $\sum_{n \in \mathbb{N}} \mathrm{P}_B(A_n)$ converge et $\mathrm{P}_B(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} \mathrm{P}_B(A_n)$.
3) $\mathrm{P}(\Omega / B) = 1$ (ou encore $\mathrm{P}_B(\Omega) = 1$).

Théorème (*) : formule de probabilités composées. Soit (Ω, T, P) un espace probabilisé
Soit $n \geq 2$ et soit $A_1, .. A_n$ des événements tels que $\mathrm{P}(A_1 \cap A_2 \cap .. \cap A_{n-1}) \neq 0$. Alors :
$\forall k \leq n-1, \mathrm{P}(A_1 \cap A_2 \cap .. \cap A_k) \neq 0$ et on a
$\mathrm{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathrm{P}(A_1)\, \mathrm{P}(A_2 / A_1) ... \mathrm{P}(A_n / A_1 \cap A_2 \cap ... \cap A_{n-1})$

Exemple : un gardien de nuit doit ouvrir une porte. Il possède un jeu de n clés différentes, dont une seule ouvre la porte. Il essaie les clés au hasard, une par une, sans jamais utiliser deux fois la même. Soit $k \in [\![1, n]\!]$. Déterminer la probabilité pour que la k-ème clé soit la bonne.

Rappels : Soit (Ω, P) un espace probabilisé fini. Soit $n \in \mathbb{N}^*$ et soit $I = \{1, 2, ..., n\}$. Soit $(A_1, A_2, ..., A_n)$ une famille d'événements de Ω. Soit $B \subset \Omega$.

- **Système complet d'événements :** On dit que $(A_1, A_2, ..., A_n)$ est un **système complet d'événements** si et seulement si $(A_1, A_2, ..., A_n)$ est une partition de Ω :
  -$\forall (i,j) \in I^2, i \neq j, A_i \cap A_j = \varnothing$
  -$\Omega = \bigcup_{i \in I} A_i$
- **Formule des probabilités totales.** $\mathrm{P}(B) = \sum_{i=1}^{n} \mathrm{P}(B \cap A_i) = \sum_{i=1}^{n} \mathrm{P}(A_i)\, \mathrm{P}(B / A_i)$.
  On adopte ici la convention $\mathrm{P}(A_i)\, \mathrm{P}(B / A_i) = 0$ si $\mathrm{P}(A_i) = 0$.
  En particulier, si $A \subset \Omega$ est un événement de probabilité $p \in \left]0,1\right[$, alors :
  $\mathrm{P}(B) = \mathrm{P}(A \cap B) + \mathrm{P}(\overline{A} \cap B) = \mathrm{P}(A)\, \mathrm{P}(B / A) + \mathrm{P}(\overline{A})\, \mathrm{P}(B / \overline{A})$.

On généralise ces résultats pour un nombre dénombrable d'événements :

Définition : Soit (Ω, T, P) un espace probabilisé. Soit $\left(A_n\right)_{n \in \mathbb{N}}$ une famille d'éléments de T.

- $\left(A_n\right)_{n \in \mathbb{N}}$ est un **système complet d'événements** si et seulement si $\left(A_n\right)_{n \in \mathbb{N}}$ est une partition de Ω :
  -$\forall (i,j) \in I^2, i \neq j, A_i \cap A_j = \varnothing$
  -$\Omega = \bigcup_{n \in \mathbb{N}} A_n$
- $\left(A_n\right)_{n \in \mathbb{N}}$ est un système **quasi-complet** d'événements si et seulement si
  -$\forall (i,j) \in I^2, i \neq j, A_i \cap A_j = \varnothing$
  -$\sum_{n=0}^{+\infty} P\left(A_n\right) = 1$

Exemple (*) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire sur Ω. Alors $(X = x)_{x \in X(\Omega)}$ est un système complet d'événements.

Exemple : on lance une infinité de fois un dé à 6 faces. L'événement $(X = n)$ se produit si et seulement si on obtient 6 pour la première fois au n-ème lancer. L'événement $(X = 0)$ se produit si et seulement si on n'obtient jamais de 6. Alors $(X = n)_{n \in \mathbb{N}^*}$ est un système quasi complet d'événements.

Formule des probabilités totales (*). Soit (Ω, T, P) un espace probabilisé. Soit $\left(A_n\right)_{n\in\mathbb{N}}$ un système complet ou quasi-complet d'événements. Soit $B \in T$.
Pour $n \in \mathbb{N}$, si $P(A_n) = 0$, on adopte la convention $P(A_n)P(B / A_n) = 0$.
Alors la série $\sum P(B \cap A_n)$ converge et $P(B) = \sum_{n=0}^{+\infty} P(B \cap A_n) = \sum_{n=0}^{+\infty} P(B / A_n)P(A_n)$

Preuve : cas d'un système quasi-complet d'événements.

Remarque : la formule des probabilités totales est utile lorsque l'expérience décrite est « doublement aléatoire », c'est-à-dire qu'il y un résultat qui dépend lui-même d'un premier résultat aléatoire.

Exemple : un joueur lance une pièce équilibrée jusqu'à obtenir Pile. Si cela arrive au n-ème lancer, il prend un billet de loterie au hasard parmi n billets dont un seul est gagnant. Quelle est la probabilité que le joueur gagne ?

Propriété (formule de Bayes) : Soit (Ω, T, P) un espace probabilisé. Soit A et B sont deux événements de probabilité non nulle, on a alors : $\mathrm{P}(B / A) = \dfrac{\mathrm{P}(B)}{\mathrm{P}(A)}\mathrm{P}(A / B)$.

Remarque : on peut ensuite calculer $\mathrm{P}(A)$ en utilisant un système complet d'événements et la formule des probabilités totales.

Exemple : un virus touche 1% de la population. On met au point un test pour le dépister.
Si l'individu est malade, le test est positif dans 99% des cas.
Si l'individu n'est pas malade, le test est négatif dans 97% des cas.
Un couple va se faire tester.
Le test de la femme est négatif : quelle est la probabilité pour qu'elle soit malade ?
Le test de l'homme est positif. Quelle est la probabilité pour qu'il soit malade ?

4) <u>Evénements indépendants</u>

Définition : Deux événements A et B sont dits indépendants si et seulement si on a $\mathrm{P}(A \cap B) = \mathrm{P}(A).\mathrm{P}(B)$

Propriétés : Soit (Ω, T, P) un espace probabilisé. Soit $A, B \subset \Omega$.

1) Si $\mathrm{P}(A) \neq 0$, alors A et B sont indépendants si et seulement si $\mathrm{P}(B / A) = \mathrm{P}(B)$.
2) Si A et B sont indépendants, $\overline{A}$ et B le sont aussi.

Preuve : non faite.

Définition : Soit $n \geq 2$ et $A_1, ..., A_n \subset \Omega$.

- On dit que les événements $A_1, ..., A_n$ sont deux à deux indépendants si pour tout couple (i, j) tels que $i \neq j$, on a $\mathrm{P}(A_i \cap A_j) = \mathrm{P}(A_i)\mathrm{P}(A_j)$.
- On dit que $A_1, ..., A_n$ sont indépendants si et seulement si pour tout $k \in [\![1, n]\!]$, pour tous $1 \leq i_1 < i_2 < ... < i_k \leq n$, on a $\mathrm{P}(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_k}) = \mathrm{P}(A_{i_1})\mathrm{P}(A_{i_2})...\mathrm{P}(A_{i_k})$

Remarque : l'indépendance entraîne l'indépendance deux à deux (avec $k = 2$), mais la réciproque est fausse.

Exemple : on jette deux fois un dé à 6 faces, non pipé. On prend :
A : « on obtient un score pair au premier jet »
B : « on obtient un score pair au second jet »
C : « on obtient exactement une fois un score pair ».
On montre que ces événements sont deux à deux indépendants sans être mutuellement indépendants.

Propriété : Soient $A_1, A_2, .., A_n \subset \Omega$. On suppose $A_1, A_2, .., A_n$ indépendants. Pour $k \in [\![1, n]\!]$, on considère $B_k \in \left\{A_k, \overline{A_k}\right\}$. Alors $B_1, B_2, .., B_n$ sont indépendants.

Idée de preuve : quitte à renommer les ensembles pour considérer d'abord ceux dont on prend le complémentaire, on est amené à calculer $P\left(\overline{A_1} \cap ... \cap \overline{A_k} \cap A_{k+1} \cap ... \cap A_n\right)$ et on fait une récurrence sur k

$P\left(\overline{A_1} \cap ... \cap \overline{A_k} \cap A_{k+1} \cap ... \cap A_n\right) = P\left(\overline{A_2} \cap ... \cap \overline{A_k} \cap .. \cap A_n\right) - P\left(A_1 \cap \overline{A_2} \cap ... \cap \overline{A_k} \cap ... \cap A_n\right)$ et on peut appliquer l'hypothèse de récurrence pour $k - 1$.

B) Variable aléatoire discrète.

1) Loi d'une variable aléatoire discrète.

Propriété : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire discrète sur Ω. On note $X(\Omega)$ l'ensemble des valeurs prises par cette variable aléatoire.
On définit P_X sur $X(\Omega)$ par $\forall A \subset X(\Omega), P_X(A) = P(X \in A)$
Alors P_X est une probabilité sur $X(\Omega)$.

Preuve rapide :

Remarque : comme $X(\Omega)$ est fini ou dénombrable, si $A \subset X(\Omega)$, on a $A = \{a_n, n \in \mathbb{N}\}$.
On a alors $P_X(A) = P(X \in A) = P\left(\bigcup_{n\in\mathbb{N}} (X = a_n)\right) = \sum_{n\in\mathbb{N}} P_X(\{a_n\})$.
Ainsi, la connaissance des $P_X(\{x\}) = P(X = x)$ pour $x \in X(\Omega)$ suffit pour déterminer P_X.

On a donc la définition suivante :

Définition (*) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire discrète sur Ω. On note $X(\Omega)$ l'ensemble des valeurs prises par cette variable aléatoire.
La loi P_X de X est la probabilité définie sur $X(\Omega)$ par la donnée de $P_X(\{x\}) = P(X = x)$ pour toute valeur $x \in X(\Omega)$.

Notation : lorsque deux variables aléatoires X et Y ont la même loi (c'est-à-dire que $X(\Omega) = Y(\Omega)$ et que $p_X = p_Y$), on note $X \sim Y$.

2) Lois usuelles

Loi uniforme (*) : Soit X une variable aléatoire sur Ω. On note $X(\Omega) = E = \{x_1, x_2 ... x_n\}$. On dit qu'elle suit une loi uniforme $U(E)$ sur Ω si et seulement si $\forall\, k \in \{1, 2, ..., n\}$, $P(X = x_k) = \frac{1}{n}$. On note $X \sim U(E)$.

Loi de Bernoulli (*) : Soit X une variable aléatoire réelle sur Ω. On dit que X suit une loi de Bernoulli de paramètre p si $X(\Omega) = \{0,1\}$ et $P(X = 1) = p$. On a alors $P(X = 0) = 1 - p$. On note $X \sim B(p)$.

Exemple : lors d'une expérience aléatoire, un événement $A \subset \Omega$ a une probabilité p d'être réalisé (succès de l'expérience), et une probabilité $1 - p$ de ne pas l'être. La variable aléatoire définie par $X = 1$ si A est réalisé, et $X = 0$ sinon suit une loi de Bernoulli de paramètre p On a alors $X = 1_A$. On parle alors de variable indicatrice de Bernoulli.

Loi binomiale (*) : Soit X une variable aléatoire réelle sur Ω. Soit $p \in [0,1]$ et $n \in \mathbb{N}^*$. On dit que X suit une loi binomiale $B(n,p)$ si et seulement si $X(\Omega) = [\![0,n]\!]$ et que pour $0 \le k \le n$, il vient $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Remarque : si $X \sim B(n,p)$, on a bien $\sum_{k=0}^{n} P(X=k) = (p+1-p)^n = 1$.

Exemple : on réalise n fois une même expérience aléatoire de manière indépendante, et pour chaque expérience on a une probabilité de succès égale à p. Si on note X le nombre de succès obtenus lors de ces n expériences, alors X suit une loi binomiale $B(n,p)$.

On note A_k : "on obtient un succès au k-ème essai".

Alors $P(X=k) = \binom{n}{k} P(A_1)...P(A_k)P\left(\overline{A_{k+1}}\right)...P\left(\overline{A_n}\right)$

Définition (*) : **Loi de Poisson**. Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle discrète sur Ω. On dit que X suit une loi de Poisson de paramètre $\lambda > 0$ si et seulement si $X(\Omega) = \mathbb{N}$ et $\forall n \in \mathbb{N}, P(X=n) = \frac{\lambda^n}{n!} e^{-\lambda}$.

Propriété : Soit (Ω, T, P) un espace probabilisé. Soit (X_n) une suite de variables aléatoires réelles discrètes telles que pour tout $n \in \mathbb{N}^*$, X_n suive une loi binomiale de paramètres (n, p_n). On suppose $n\,p_n \underset{n\to+\infty}{\to} \lambda > 0$. Alors $\forall k \in \mathbb{N}, P(X_n = k) \underset{n\to+\infty}{\to} \frac{\lambda^k e^{-\lambda}}{k!}$.

Preuve : $P(X_n = k) = \frac{1}{k!}\left(n(n-1)...(n-k+1)\right) p_n{}^k (1-p_n)^{n-k} \underset{n\to+\infty}{\sim} \frac{(np_n)^k}{k!} e^{(n-k)\ln(1-p_n)}$

Remarque : la loi de Poisson constitue ainsi une bonne approximation de la loi binomiale lorsque n est grand et p_n petit. La loi de Poisson est « la loi des événements rares ».
Elle peut servir concrètement à modéliser le nombre de clients se présentant dans un magasin (en prenant $n = 1000000$ le nombre d'habitants d'une ville et p la probabilité que chaque habitant se rende dans le magasin un jour donné).

Définition (*) : **Loi géométrique**. Soit (Ω, T, P) un espace probabilisé. Soit $p \in]0,1[$. Soit X une variable aléatoire réelle discrète sur Ω. On dit que X suit une loi géométrique de paramètre p, notée $G(p)$ si et seulement si $X(\Omega) = \mathbb{N}^*$ et $\forall n \in \mathbb{N}^*, P(X=n) = (1-p)^{n-1} p$.

Remarques :

- On a bien $\sum_{n=1}^{+\infty} P(X=n)=1$
- Une loi géométrique sert à modéliser le rang d'un premier succès quand on répète de manière indépendante des épreuves de Bernoulli ayant une probabilité de succès égale à p.
- On a alors $\forall n \in \mathbb{N}, P(X>n)=(1-p)^n$ (n échecs lors des n premières tentatives ; se retrouve aussi avec le calcul).

Exemples :

- Matis joue aux petits chevaux. Il s'énerve car il n'obtient pas de 6. Mais quelle est la loi du nombre de lancers nécessaires pour obtenir un premier 6 ?
- $n \geq 3$ personnes jouent au jeu suivant : elles lancent simultanément une pièce équilibrée pour obtenir pile ou face. Le gagnant est celui qui obtient la face contraire de tous les autres. On note X le nombre de coups nécessaires pour obtenir un gagnant. Quelle est la loi de X ?

2) Espérance d'une variable aléatoire

Rappel : Soit X une variable aléatoire sur Ω, à valeurs réelles ou complexes. On suppose que $X(\Omega)=\{x_1, x_2 ... x_n\}$ est fini. On appelle espérance de X le nombre $E(X)=\sum_{i=1}^{n} P(X=x_i)x_i$.

Définition (*) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire discrète sur Ω, à valeurs dans $[0,+\infty]$. En adoptant la convention $x\,P(X=x)=0$ si $x=+\infty$ et $P(X=+\infty)=0$, on définit l'espérance de X par $E(X)=\sum_{x \in X(\Omega)} xP(X=x)$

Définition (*) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle ou complexe discrète sur Ω.
Alors X est d'espérance finie si et seulement si la famille $(x\,P(X=x))_{x \in X(\Omega)}$ est sommable.
Dans ce cas, l'espérance de X est définie par $E(X)=\sum_{x \in X(\Omega)} xP(X=x)$.
On dit que X est centrée si et seulement si $E(X)=0$.

Remarques :

- L'espérance est ainsi la valeur **moyenne** prise par la variable aléatoire X, pondérée par la probabilité pour que X prenne ces différentes valeurs.
- Une variable aléatoire peut prendre des valeurs finies mais être d'espérance infinie : on prend pour $n \in \mathbb{N}^*$ $P(X=n)=\dfrac{1}{n(n+1)}$.

Rappels (*) :

- Soit X une variable aléatoire à valeurs dans $\{1,2,..,n\}$. Si $X \sim U(\llbracket 1,n \rrbracket)$, alors $EX=\dfrac{(n+1)}{2}$.

- Soit X une variable aléatoire réelle sur Ω. On suppose que $X \sim B(p)$ .Alors $EX = p$
- Soit $p \in [0,1]$ et $n \in \mathbb{N}^*$. On suppose que X suit une loi binomiale $B(n,p)$. Alors $E(X) = n\,p$.

Paradoxe de Saint-Petersbourg. On vous propose le jeu suivant : il faut donner 10000 euros de mise initiale. Ensuite, on lance une pièce une infinité de fois et on gagne 2^n euros si on obtient Pile pour la première fois au n-ème lancer. Est-il rentable de jouer à ce jeu ?

Proposition (*) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle discrète sur Ω, à valeurs dans $\mathbb{N} \cup \{+\infty\}$. Alors $E(X) = \sum_{n=1}^{+\infty} P(X \geq n) = \sum_{n=0}^{+\infty} P(X > n)$

Preuve : Si $P(X = +\infty) > 0$, alors $E(X) = +\infty = \sum_{n=1}^{+\infty} P(X \geq n)$ (car $P(X \geq n) \geq P(X = +\infty)$ et cette série de nombres positifs diverge)

Sinon, $E(X) = \sum_{n=1}^{+\infty} n\,P(X = n) = \sum_{n=1}^{+\infty}\sum_{k=1}^{+\infty} a_{n,k}$, avec $a_{n,k} = P(X = n)$ si $k \leq n$ et $a_{n,k} = 0$ sinon.

Par Fubini, $E(X) = \sum_{n=1}^{+\infty}\sum_{k=1}^{+\infty} a_{n,k} = \sum_{k=1}^{+\infty}\sum_{n=k}^{+\infty} a_{n,k} = \sum_{k=1}^{+\infty}\sum_{n=k}^{+\infty} P(X = n) = \sum_{k=1}^{+\infty} P(X \geq k)$.

Proposition (*,PV) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle discrète sur Ω. suivant une loi de Poisson de paramètre $\lambda > 0$. Alors $E(X) = \lambda$

Preuve : le calcul justifie que X est d'espérance finie.

Proposition (*,PV) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle discrète sur Ω. suivant une loi géométrique de paramètre $p \in]0,1[$. Alors $E(X) = \dfrac{1}{p}$

Preuve : directement $E(X) = p\sum_{n=1}^{+\infty} n\,q^{-1} = \dfrac{p}{(1-q)^2} = \dfrac{1}{p}$, ou bien avec $E(X) = \sum_{n=0}^{+\infty} P(X > n)$.

Propriété : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire discrète sur Ω. Soit g est une application définie sur $X(\Omega)$. Alors $Y = g(X)$ est une variable aléatoire discrète sur Ω, définie par $\forall \omega \in \Omega, Y(\omega) = g(X(\omega))$.
La loi de $Y = g(X)$ est alors définie sur $g(X(\Omega))$ par :

$\forall\, y \in g(X(\Omega)), p_Y(\{y\}) = P(g(X) = y) = \sum_{x \in g^{-1}(\{y\})} P(X = x)$.

Exemples : Ainsi, si X est une variable aléatoire sur Ω, X^2 et e^{2X} sont des variables aléatoires.

Rappel : **théorème de transfert (*) :** Soit X une variable aléatoire sur un univers Ω. On suppose que $X(\Omega)=\{x_1,x_2...x_n\}$ est fini. Soit $f : X(\Omega)\to\mathbb{C}$.

Alors $E(f(X))=\sum_{j=1}^{n} f(x_j)P(X=x_j)=\sum_{x\in X(\Omega)} f(x)P(X=x)$

Remarque : l'intérêt est le suivant : on n'a pas besoin de la loi de $f(X)$ pour calculer son espérance. Celle de X suffit.

Théorème de transfert (*) : Soit (Ω,T,P) un espace probabilisé. Soit X une variable aléatoire discrète sur Ω. Soit $f : X(\Omega)\to\mathbb{C}$. Alors $f(X)$ est d'espérance finie si et seulement si la famille $(f(x)\,P(X=x))_{x\in X(\Omega)}$ est sommable et alors $E(f(X))=\sum_{x\in X(\Omega)} f(x)P(X=x)$.

Idée de preuve :

Sous réserve de sommabilité :
si $y\in f(X)(\Omega)$, il vient $(f(X)=y)=\bigcup_{x\in I_y}(X=x)$, avec $I_y=\{x\in X(\Omega), X(x)=y\}$.
Donc $P(f(X)=y)=\sum_{x\in I_y} P(X=x)$
$E(f(X))=\sum_{y\in f(X(\Omega))} yP(f(X)=y)=\sum_{y\in f(X(\Omega))}\sum_{x\in I_y} f(x)P(X=x)$
Par sommation par paquets, $E(f(X))=\sum_{x\in X(\Omega)} f(x)P(X=x)$.
Puis la sommabilité d'une des familles entraîne celle de l'autre.

Exemple : soit X une variable aléatoire qui suit une loi de Poisson de paramètre $\lambda>0$. Montrer que $X(X-1)$ est d'espérance finie, et la calculer.

Propriété (*) : Soit (Ω,T,P) un espace probabilisé. Soit X,Y deux variables aléatoires discrètes sur Ω, à valeurs réelles ou complexes. On suppose que X,Y sont d'espérance finie. Soit $a,b\in K$

1) Si X est presque sûrement constante, égale à $k\in\mathbb{R}$, alors $E(X)=k$.
2) $aX+bY$ est d'espérance finie et $E(aX+bY)=aE(X)+bE(Y)$ (linéarité).
3) En particulier, $E(aX+b)=a\,E(X)+b$.

Preuve : non faite. Pour la 2), on applique le théorème de transfert à $Z=(X,Y)$ à valeurs dans $X(\Omega)\times Y(\Omega)$, avec $f(Z)=f(X,Y)=aX+bY$. Il faut montrer la sommabilité pour pouvoir séparer en deux. On prouve que $\left(|a||x|P(X=x,Y=y)\right)_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est sommable en appliquant le théorème de transfert à $g(X,Y)=X$
On conclut en séparant en deux.
On peut aussi noter que $(X,Y)(\Omega)\subset X(\Omega)\times Y(\Omega)$ et que les événements qui ne sont que dans le second sont de probabilité nulle.

Exemple : on dispose de 5 bonbons et de 4 enfants. On donne chaque bonbon à un enfant choisi au hasard. Pour $k \in [\![1,4]\!]$, on note X_k la variable aléatoire égale à 1 si le k-ème enfant n'a reçu aucun bonbon, et à 0 sinon.

1) Calculer $P\left(X_k = 1\right)$
2) Combien y-a-t-il d'enfants qui n'ont reçu aucun bonbon en moyenne ?

Propriété : Soit (Ω, T, P) un espace probabilisé. Soit X, Y deux variables aléatoires discrètes sur Ω. On suppose Y à valeurs réelles et X à valeurs complexes. On suppose de plus $\forall \omega \in \Omega, |X(\omega)| \leq Y(\omega)$ et que Y est d'espérance finie. Alors X est d'espérance finie.

Preuve : non faite. Prendre $Z = (X, Y)$ et $f(X,Y) = X$.
Alors par formule de transfert $E(Y) = \sum_{(x,y)\in Z(\Omega)} yP\left(X = x, Y = y\right)$.
Pour $(x, y) \in Z\left(\Omega\right)$, on a $X(\omega) = x$ et $Y(\omega) = y$, donc $|x| \leq y$ et la famille $xP\left(X = x, Y = y\right)$ est sommable.

Propriétés : Soit (Ω, T, P) un espace probabilisé. Soit X, Y deux variables aléatoires discrètes sur Ω, à valeurs réelles.

1) L'espérance est **positive** : si $\forall \omega \in \Omega, X(\omega) \geq 0$, alors $E(X) \geq 0$
2) L'espérance est **croissante** : si $\forall \omega \in \Omega, X(\omega) \leq Y(\omega)$, et que X, Y sont d'espérance finie, alors $E(X) \leq E(Y)$.
3) Si X est à valeurs réelles positives et que $E(X) = 0$, alors $P(X = 0) = 1$.

Preuves : non faites ? Pour la 2), on applique la 1) à $Y - X$. Pour la 3), montrer que $\forall x \in X\left(\Omega\right), 0 \leq xP(X = x) \leq E(X)$ et conclure $P(X = x) = 0$ si $x \neq 0$ et conclure par union dénombrable $P(X \neq 0) = 1$.

3) <u>Variance et écart-type</u>

Dans tout ce paragraphe, les variables aléatoires sont à valeurs réelles.

Propriété : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire discrète sur Ω, à valeurs réelles. On suppose que X^2 est d'espérance finie. Alors X est d'espérance finie.

Preuve : on a pour $z \in \mathbb{C}$: $|z| \leq \frac{1}{2}\left(1 + |z|^2\right)$. Donc $|X| \leq \frac{1}{2}\left(1 + X^2\right)$.

Remarque : la réciproque n'est pas toujours vraie. Il se peut que X soit d'espérance finie, mais pas X^2. On prend pour $n \in \mathbb{N}^*$: $P(X = n) = \frac{1}{a\,n^3}$, avec $a = \sum_{n=1}^{+\infty} \frac{1}{n^3} = \zeta(3)$.

Définition (*) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle discrète sur Ω. On suppose que X^2 est d'espérance finie. On définit alors la variance de X par $V(X) = E((X - EX)^2) = E\left(X^2\right) - E\left(X\right)^2$. On dit alors que X est de variance finie.

On appelle écart-type de X le réel $\sigma(X) = \sqrt{V(X)}$.

X est réduite si et seulement si $V(X) = 1$

Explication de l'égalité :

Remarques :

- en particulier, si X^2 est d'espérance finie, $V(X) \geq 0$.
- La variance et l'écart-type mesurent la dispersion de X autour de sa moyenne. Ainsi, si X est constante, on a $V(X) = 0$.

Propriété (*) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle discrète sur Ω. On suppose que X^2 est d'espérance finie (donc que X admet une variance). Soient *a* et *b*, deux réels. Alors : $\begin{array}{l} V(aX+b) = a^2 V(X) \\ \sigma(aX+b) = |a|\sigma(X) \end{array}$.

En particulier, Si $\sigma(X) > 0$, la variable aléatoire $X^* = \dfrac{X - E(X)}{\sigma(X)}$ est centrée réduite (c'est-à-dire que $E(X^*) = 0$ et $V(X^*) = 1$).

Preuve : on utilise $V(Y) = E((Y - EY)^2)$ avec $Y = aX + b$.

Rappels : soit X une variable aléatoire réelle sur Ω. Soit $p \in [0,1]$ et $n \in \mathbb{N}^*$.

- On suppose que X suit une loi de Bernoulli $B(p)$. Alors $V(X) = p(1-p)$.
- On suppose que X suit une loi binomiale $B(n, p)$. Alors $V(X) = n\,p(1-p)$.

Preuve : non faite. Avec des calculs de sommes (voir première année).

Proposition (*) : **Loi de Poisson**. Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle discrète sur Ω, qui suit une loi de Poisson de paramètre $\lambda > 0$. Alors X est de variance finie et $V(X) = \lambda$

Preuve : avec $E(X^2) = \lambda^2 + \lambda$ (le calcule justifie la sommabilité).

Proposition (*) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle discrète sur Ω. suivant une loi géométrique de paramètre $p \in \left]0,1\right[$. Alors $V(X) = \dfrac{1-p}{p^2}$.

4) Fonctions génératrices

Définition (*) : Soit (Ω, T, P) un espace probabilisé et X une variable aléatoire réelle discrète sur Ω, à valeurs dans $\mathbb{N}$. La fonction génératrice associée à cette variable aléatoire est définie en tout réel t tel que t^X est d'espérance finie et donnée par $G_X(t) = E\left(t^X\right) = \sum_{n=0}^{+\infty} P(X=n)t^n$.

Remarques :
- c'est une série entière de rayon de convergence $R \geq 1$.
- Il y a convergence normale sur $\left[-1,1\right]$, donc G_X est continue sur $\left[-1,1\right]$.

Exemples (*,PV) : Fonctions génératrices pour une loi de Poisson, une loi géométrique et une loi binomiale :

Poisson : $\forall t \in \mathbb{R}, G_X(t) = \sum_{n=0}^{+\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda} = e^{\lambda(t-1)}$.

Géométrique : si $q = 1-p$, $\forall t \in \left]-\frac{1}{q}, \frac{1}{q}\right[$, $G_X(t) = pt\sum_{n=0}^{+\infty}(qt)^n = \frac{pt}{1-qt}$

Bernoulli : $\forall t \in \mathbb{R}, G_X(t) = tp + 1 - p$

Binomiale : $\forall t \in \mathbb{R}, G_X(t) = \sum_{k=0}^{n} \binom{n}{k} t^k p^k (1-p)^{n-k} = (tp+1-p)^n$

Proposition (*) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle discrète sur Ω, à valeurs dans $\mathbb{N}$. Alors la loi de X est entièrement déterminée par sa fonction génératrice. En particulier, $\forall k \in \mathbb{N}, P(X=k) = \frac{G_X^{(k)}(0)}{k!}$

Preuve : on a une série entière de rayon de convergence $R = 1 > 0$.

Proposition (*) : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle discrète sur Ω, à valeurs dans $\mathbb{N}$. Alors X est d'espérance finie si et seulement si G_X est dérivable en 1 et dans ce cas $E(X) = G_X'(1)$.

Preuve : on fait le sens direct pour l'espérance, avec la convergence normale de la série des dérivées sur $\left[0,1\right]$. Le sens retour est admis.

Remarque : lorsque X^2 est d'espérance finie, on peut utiliser G_X pour calculer la variance : il vient $G_X'(t) = \sum_{n=0}^{+\infty} P(X=n)nt^{n-1}$ et $G_X''(t) = \sum_{n=2}^{+\infty} P(X=n)n(n-1)t^{n-2}$.
(avec $0 \leq P(X=n)n(n-1) \leq P(X=n)n^2$, il y a convergence normale et c'est vrai aussi lorsque le rayon de convergence est strictement supérieur à 1).
Donc $G_X''(1) = E\left(X(X-1)\right) = E\left(X^2\right) - E\left(X\right)$ puis $V(X) = G_X''(1) + E\left(X\right) - E\left(X\right)^2$.
donc $V(X) = G_X''(1) + G_X'(1) - G_X'(1)^2$ (à savoir retrouver).

Exemples : retrouver espérance et variance de la loi de Poisson et espérance de la loi géométrique.

Tableau récapitulatif : (avec $q = 1-p$)

Loi	$X(\Omega)$	$P(X=k)$	$E(X)$	$V(X)$	$G_X(t)$
$B(p)$	$\{0,1\}$		p	$p(1-p)$	$pt+1-p$
$B(n,p)$	$[\![0,n]\!]$	$\binom{n}{k} p^k(1-p)^{n-k}$	np	$np(1-p)$	$(pt+1-p)^n$
$U(n)$	$[\![1,n]\!]$	$\frac{1}{n}$	$\frac{n+1}{2}$		
$P(\lambda)$	$\mathbb{N}$	$\frac{\lambda^k e^{-\lambda}}{k!}$	λ	λ	$e^{\lambda(t-1)}$
$G(p)$	$\mathbb{N}^*$	pq^{n-1}	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pt}{1-qt}$

5) <u>Inégalités de Markov et Bienaymé-Tchebychev</u>

Proposition (*,PV) : inégalité de Markov. Soit (Ω, T, P) un espace probabilisé. Soit Y une variable aléatoire réelle discrète sur Ω. On suppose Y **à valeurs positives** et d'espérance finie. Alors : $\forall a > 0, \mathrm{P}(Y \geq a) \leq \frac{E(Y)}{a}$

Preuve : on écrit $E(Y) = \sum_{x_n \geq a} x_n P(Y = x_n) + \sum_{x_n < a} x_n P(Y = x_n) \geq a \sum_{x_n \geq a} P(Y = x_n)$.

Exemple : soit X une variable aléatoire discrète réelle. Montrer $\forall t \in \mathbb{R}_+, P(X \geq 0) \leq E(e^{tX})$.

Proposition (*,PV) : inégalité de Bienaymé-Tchebychev : Soit (Ω, T, P) un espace probabilisé. Soit X une variable aléatoire réelle discrète sur Ω. On suppose qu'elle admet une variance : alors $\forall \lambda > 0, P(|X - E(X)| \geq \lambda) \leq \frac{V(X)}{\lambda^2}$

Preuve : soit $\lambda > 0$. Comme X admet une variance sur Ω, elle admet aussi une espérance. On pose $Y = (X - E(X))^2$ qui est une variable aléatoire positive et on applique l'inégalité de Markov. $P(|X - E(X)| \geq \lambda) = P\left(|X - E(X)|^2 \geq \lambda^2\right) = P\left(Y \geq \lambda^2\right)$.

Donc $P(|X - E(X)| \geq \lambda) \leq \frac{E(Y)}{\lambda^2}$, avec $E(Y) = E\left(X - E(X)\right)^2 = V(X)$.

On a donc bien $P(|X - E(X)| \geq \lambda) \leq \frac{V(X)}{\lambda^2}$

Exemple : soit X une variable aléatoire suivant une loi de Poisson de paramètre $\lambda > 0$.
Montrer que $P\left(|X - \lambda| < \lambda\right) \geq \frac{\lambda - 1}{\lambda}$ et $P\left(X \leq \frac{\lambda}{2}\right) \leq \frac{4}{\lambda}$.