I. ENSEMBLES FINIS OU DÉNOMBRABLES

1. Ensembles finis ou dénombrables

Définition 1 :

- Un ensemble E est dit **fini** s'il existe $n \in \mathbb{N}$ tel que E soit en bijection avec $[\![1, n]\!]$.
  L'entier n est alors unique et appelé **Cardinal de** E, noté $\text{Card}E$.
- Un ensemble E est dit dénombrable s'il est en bijection avec $\mathbb{N}$.

Proposition 1

Soit A une partie de E, un ensemble fini. Alors A est un ensemble fini et $\text{Card}A \leqslant \text{Card}E$.
De plus, $\text{Card}A = \text{Card}E$ si et seulement si $A = E$.

Remarque 1 :

- L'ensemble vide $\emptyset$ est un ensemble fini, de cardinal 0. C'est le seul ensemble de cardinal 0.
- Si E est fini, on peut numéroter ses éléments et en général, on notera $E = \{x_1, \ldots, x_n\}$ où $n = \text{Card}E$.
- E est dénombrable si il existe $\varphi : \mathbb{N} \to E$ bijective. Cela signifie que l'on peut numéroter les éléments de E et on notera $E = \{x_n,\ n \in \mathbb{N}\}$.
- Soit A une partie de $\mathbb{N}$, A est finie si et seulement si A est majorée.
- On regroupera parfois "fini ou dénombrable". On pourra énoncer par exemple : toute partie de $\mathbb{N}$ est finie ou dénombrable.

Exemple 1 :

a. $\mathbb{N}$ est dénombrable, $\mathbb{N}^*$ aussi, $\{2k,\ k \in \mathbb{N}\}$, $\{2k+1,\ k \in \mathbb{N}\}$ sont des ensembles dénombrables.
On remarquera à cette occasion qu'une partie stricte d'un ensemble E dénombrable peut-être dénombrable, donc en bijection avec E, alors que ceci est impossible dans le cas des ensembles finis.

b. $\mathbb{N}^2$ et $\mathbb{Z}$ sont dénombrables.

Proposition 2

Si E et F sont deux ensembles dénombrables alors le produit cartésien $E \times F$ est un ensemble dénombrable.

2. Parties d'un ensemble

Soit E un ensemble on note $\mathscr{P}(E)$ l'ensemble des parties de E.

Proposition 3

Si E est fini, alors $\mathscr{P}(E)$ est fini et $\text{Card}(\mathscr{P}(E)) = 2^{\text{Card}E}$

Remarque 2 : Si E est dénombrable, alors $\mathscr{P}(E)$ n'est pas dénombrable !

Familles de parties d'un ensemble

Soit E un ensemble. Soit I un ensemble fini ou dénombrable.
On considère une famille $(A_i)_{i\in I}$ de parties de E. Pour $I \subset \mathbb{N}$, on dit que $(A_n)_{n\in\mathbb{N}}$ est une suite de parties de E.
De plus, (A_n) est une suite croissante si $\forall n \in \mathbb{N},\ A_n \subset A_{n+1}$. De façon analogue, on définit les suites décroissantes.

Intersection : $\bigcap_{i\in I} A_i = \{x \in E \ /\ \forall i \in I,\ x \in A_i\}$.
C'est l'ensemble des éléments qui appartiennent à tous les A_i.

Réunion : $\bigcup_{i\in I} A_i = \{x \in E \ / \ \exists i \in I,\ x \in A_i\}$.
C'est l'ensemble des éléments qui appartiennent à au moins l'un des A_i.

Si $I = \mathbb{N}$, on peut noter $\bigcap_{i=0}^{+\infty} A_i$ et $\bigcup_{i=0}^{+\infty} A_i$.

Exemple 2 :

- Dans $\mathbb{N}$, $A_n = [\![1, n]\!]$. (A_n) est une suite croissante de parties de $\mathbb{N}$.
  On a $\bigcup_{n=1}^{+\infty} A_n = \mathbb{N}^*$ et $\bigcap_{n=1}^{+\infty} A_n = \{1\}$.
- Dans $\mathbb{R}$, $A_n = [0, n]$. Complétez :
  $\bigcup_{n=1}^{+\infty} A_n =$, $\bigcap_{n=1}^{+\infty} A_n =$, $\bigcap_{n=0}^{+\infty} A_n =$.
- Dans $\mathbb{R}$, $A_n = [0, \frac{1}{n}]$, pour $n \geqslant 1$. Complétez :
  $\bigcup_{n=1}^{+\infty} A_n =$, $\bigcap_{n=1}^{+\infty} A_n =$.

II. ESPACES PROBABILISÉS

1. Probabilité sur les univers finis ou dénombrables

On suppose ici que l'univers Ω est fini ou dénombrable. Dans un premier temps, nous reprenons les idées du cas fini pour définir une probabilité sur $(\Omega, \mathscr{P}(\Omega))$.
On appelle événement tout élément de $\mathscr{P}(\Omega)$.

Définition 2 : On appelle **probabilité** sur $(\Omega, \mathscr{P}(\Omega))$ toute application $P : \mathscr{P}(\Omega) \to [0,1]$ telle que

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

b. Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements 2 à 2 incompatibles, $\mathbb{P}(\bigcup_{n=0}^{+\infty} A_n) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$.

Remarque 3 : La deuxième condition présuppose que la série $\sum_{n\geqslant 0} \mathbb{P}(A_n)$ converge. Dans ce cas, la somme est bien définie et on admettra qu'elle ne dépend pas de l'ordre de sommation puisque c'est une série absolument convergente dans notre cas.

Exemple 3

On lance une pièce de monnaie et on note la suite des résultats obtenus (P ou 0 pour Pile, F ou 1 pour Face). On prend ici $\Omega = \{0,1\}^{\mathbb{N}}$: cet ensemble est l'ensemble des suites de 0 et de 1, il n'est pas dénombrable. Le problème est alors de définir une probabilité sur Ω, et même de bien définir les événements.
On verra dans ce type de cas l'utilité de la notion de tribu.
Simplifions un peu ici : on lance la pièce et on s'arrête dès que l'on obtient Face.
Dans ce cas, on peut écrire $\Omega = \{1, 01, 001, \ldots, 000\ldots01, \ldots,\}$. Dans ce cas, Ω est dénombrable et on cherche à définir $\mathbb{P}(\{\omega\})$ pour chaque $\omega \in \Omega$.
On pose $\mathbb{P}(\{\underbrace{00\ldots0}_{n-1\ fois}1\}) = \dfrac{1}{2^n}$ pour traduire le fait que l'on a obtenu 0 aux $n-1$ premiers lancers (1 chance sur 2 pour chaque lancer) et 1 au dernier lancer (1 chance sur 2 également).

Exercice

On lance un dé jusqu'à obtenir un 6. On note A_k l'événement : "obtenir un premier 6 au k-ième lancer".
On peut choisir ici $\Omega = \mathbb{N}^*$, mais on ne décrit pas nécessairement l'univers Ω. On veut définir une probabilité modélisant cette expérience aléatoire.

a. Justifier que l'on peut poser $\mathbb{P}(A_k) = \mathbb{P}(\{k\}) = \dfrac{1}{6}\left(\dfrac{5}{6}\right)^{k-1}$.

b. On note A_∞ : "ne pas obtenir de 6", et on peut exprimer $A_\infty = \bigcap_{k=1}^{+\infty} \overline{A_k}$. En déduire que $\mathbb{P}(A_\infty) = 0$.

On pose donc $\mathbb{P}(\{k\}) = \dfrac{1}{6}\left(\dfrac{5}{6}\right)^{k-1}$ et on admettra que l'on définit ainsi une probabilité sur $(\Omega, \mathscr{P}(\Omega))$.

c. Quelle est la probabilité d'obtenir un 6 après un nombre pair de lancers ?

Proposition 4

Construction d'une probabilité : Soit $\Omega = \{\omega_k,\ k \in \mathbb{N}\}$ un ensemble dénombrable.
Soit $(p_k)_{k\in\mathbb{N}}$ une suite de $[0,1]$ telle que $\sum_{k\in\mathbb{N}} p_k$ converge et $\sum_{k=0}^{+\infty} p_k = 1$.
L'application $P\ :\ \mathscr{P}(\Omega) \to [0,1]$ définie par $\forall A \in \mathscr{P}(\Omega),\ \mathbb{P}(A) = \sum_{\omega_k \in A} p_k$ est une probabilité sur $(\Omega, \mathscr{P}(\Omega))$.

2. Probabilité sur un univers quelconque

Définition 3 : *Tribu sur un ensemble* Ω

Une **tribu** sur Ω est une partie $\mathscr{A}$ de $\mathscr{P}(\Omega)$ telle que

a. $\Omega \in \mathscr{A}$

b. $\forall A \in \mathscr{A}, \overline{A} \in \mathscr{A}$

c. Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathscr{A}$, on a $\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}$.

On appelle événement tout élément A de $\mathscr{A}$.

Remarque 4 :

- Pour un univers Ω fini ou dénombrable, on travaille avec $\mathscr{A} = \mathscr{P}(\Omega)$, qui est bien une tribu sur Ω.
- Les parties suivantes sont des tribus sur Ω : $\mathscr{A} = \{\emptyset, \Omega\}$, ou encore pour $A \in \mathscr{P}(\Omega)$, $\mathscr{A} = \{\emptyset, A, \overline{A}, \Omega\}$.

Définition 4 : *Espace probabilisé*

Soit Ω un ensemble et $\mathscr{A}$ une tribu sur Ω.

On appelle **probabilité sur** Ω une application $P\ :\ \mathscr{A} \to [0,1]$ telle que

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

b. Pour toute suite $(A_n)_{n\in\mathbb{N}}$ dévénements deux à deux incompatibles, $\mathbb{P}(\bigcup_{n=0}^{+\infty} A_n) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$.

$(\Omega, \mathscr{A}, \mathbb{P})$ est appelé **espace probabilisé**.

3. Propriété d'une probabilité

Proposition 5

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

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

b. $\forall A \in \mathscr{P}(\Omega),\ \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$ et $\forall A,\ B \in \mathscr{A},\ \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$

c. $\forall A,\ B \in \mathscr{P}(\Omega),\ A \subset B \implies \mathbb{P}(A) \leqslant \mathbb{P}(B)$

d. $\forall A,\ B \in \mathscr{P}(\Omega), \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Proposition 6 *Suites d'événements et probabilités*

a. *Continuité croissante*
Si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante d'événements alors $\mathbb{P}(\bigcup_{n=0}^{+\infty} A_n) = \lim_{n\to+\infty} \mathbb{P}(A_n)$

b. *Continuité décroissante*
Si $(A_n)_{n\in\mathbb{N}}$ est une suite décroissante d'événements alors $\mathbb{P}(\bigcap_{n=0}^{+\infty} A_n) = \lim_{n\to+\infty} \mathbb{P}(A_n)$

démo : dans le cas où $(A_n)_{n\in\mathbb{N}}$ est une suite croissante d'événements.

➢ On définit une famille (B_n) d'événements 2 à 2 incompatibles en posant $B_0 = A_0$ et $\forall n \in \mathbb{N}^*,\ B_n = A_n \setminus A_{n-1}$. On a ainsi $\bigcup_{n=0}^{+\infty} B_n = \bigcup_{n=0}^{+\infty} A_n$.

➢ On utilise la définition d'une probabilité pour obtenir $\mathbb{P}(\bigcup_{n=0}^{+\infty} B_n) = \sum_{n=0}^{+\infty} \mathbb{P}(B_n)$.

➢ On vérifie que, pour $N \in \mathbb{N}$, $\sum_{n=0}^{N} \mathbb{P}(B_n) = \mathbb{P}(\bigcup_{n=0}^{N} B_n)$. Or $\bigcup_{n=0}^{N} B_n = A_N$. Donc $\sum_{n=0}^{+\infty} \mathbb{P}(B_n) = \lim_{N\to+\infty} \mathbb{P}(A_N)$.

Proposition 7

a. Soit $(A_n)_{n\in[\![1,m]\!]}$ une suite finie d'événements. On a $\mathbb{P}(\bigcup_{n=1}^{m} A_n) \leqslant \sum_{n=1}^{m} \mathbb{P}(A_n)$.

b. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements. On a $\mathbb{P}(\bigcup_{n=0}^{+\infty} A_n) \leqslant \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$.

4. Formule des probabilités totales

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

Définition 5 : Soit $(A_i)_{i\in I}$ une suite finie ou dénombrable d'événements.
On dit que $(A_i)_{i\in I}$ est un **système complet d'événements** si

- les événements A_i sont 2 à 2 incompatibles.
- la réunion $\bigcup_{i\in I} A_i$ est égale à Ω.

Remarque 5 : On dit que $(A_i)_{i\in I}$ est un **système quasi-complet d'événements** si les événements sont 2 à 2 incompatibles et $P(\bigcup_{i\in I} A_i) = 1$.
C'est-à-dire que $\bigcup_{i\in I} A_i$ n'est pas forcément Ω mais il est **quasi-certain**.

Exemple 4

- Soit A un événement, $(A, \overline{A})$ forme un système complet dévénements.
- On lance un dé équilibré et note le résultat. On note A_i : obtenir i, pour $i \in [\![1,6]\!]$.
$(A_i)_{1\leqslant i\leqslant 6}$ forme un système complet dévénements.
- On lance une pièce équilibrée, on la lance n fois et on note le résultat. On note P_i (resp. F_i) "obtenir PILE (resp. FACE) au i-ème lancer". (P_1, F_1) est un système complet dévénements mais pas (P_1, F_2)....
- On lance une pièce plusieurs fois : F_i : "obtenir FACE au i-ème lancer pour la première fois" et F_0 "ne pas obtenir FACE". $(F_i)_{i\in\mathbb{N}}$ forme un système complet dénombrable d'événements et $(F_i)_{i\in\mathbb{N}^*}$ forme un système quasi-complet dénombrable d'événements.

Théorème 1 Formule des probabilités totales

Soit $(A_n)_{n\in\mathbb{N}}$ un système complet (ou quasi-complet) d'événements.
Pour tout événement B, la série $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n \cap B)$ converge et
$\mathbb{P}(B) = \sum_{n=0}^{+\infty} \mathbb{P}(B \cap A_n) = \sum_{n=0}^{+\infty} \mathbb{P}(B|A_n)\mathbb{P}(A_n)$.

Exemple 5 : On lance un dé jusqu'à obtenir un 6.

On cherche la probabilité de n'obtenir que des nombres impairs (resp. pairs) avant le premier 6 ?

Posons, pour $k \in \mathbb{N}^*$, A_k l'événement "obtenir un 6 au kième lancer (et pas avant)" et A_0 "ne pas obtenir de 6". On a vu que $\mathbb{P}(A_0) = 0$. Notons B "obtenir seulement des nombres impairs avant le premier 6" et C "n'obtenir que des nombres pairs avant le premier 6".

Justifier que $(A_k)_{k\in\mathbb{N}^*}$ forme un système quasi-complet d'événements. Montrer que $\mathbb{P}(B) = \frac{1}{3}$. Que vaut $\mathbb{P}(C)$?

III. VARIABLES ALÉATOIRES DISCRÈTES

1. Définition

Soit $(\Omega, \mathscr{P}(\Omega), \mathbb{P})$ un espace probabilisé fini ou dénombrable.

Définition 6 : *Variable aléatoire discrète*
Une variable aléatoire X est une application définie sur Ω.
On dit que X est une **variable aléatoire réelle discrète** si $X(\Omega)$ est une partie de $\mathbb{R}$ finie ou dénombrable.

Exemple 6

On lance une pièce équilibrée indéfiniment. On note X le nombre de lancer effectué pour obtenir le premier PILE (ce lancer compris). X est une variable aléatoire réelle discrète car on a $X(\Omega) = \mathbb{N}^*$.

Exemple 7

On lance un dé non truqué n fois (avec $n \in \mathbb{N}^*$ fixé) et on note le nombre de 6 obtenus, ce qui définit une variable aléatoire notée X. Ici, $X(\Omega) = [\![0, n]\!]$, dans ce cas, $X(\Omega)$ est fini.

La situation n'est pas la même si l'énoncé ne précise pas le nombre de lancer. Par exemple, si on lance jusqu'à obtenir un 6, et X désigne le rang du 6 obtenu. On prendra $X(\Omega) = \mathbb{N}^*$.

2. Loi de probabillité

Soit X une variable aléatoire réelle discrète sur (Ω, P). On note $X(\Omega) = \{x_k,\ k \in K\}$ avec K un ensemble fini ou dénombrable d'indices.

Définition 7 :
- On appelle $((X = x_k))_{k\in K}$ un système complet d'événements associés à X.
- La loi de probabilité de X est la donnée de $X(\Omega)$ et de $\mathbb{P}(X = x)$ pour tout $x \in X(\Omega)$.

3. Les lois usuelles

Nous avons déjà revu les lois usuelles pour X une variable aléatoire fini : loi uniforme, loi de Bernoulli, loi binomiale. On présente ici deux nouvelles lois sur des univers dénombrables.

a. *Loi géométrique*

Définition 8 : Soit $p \in]0, 1[$, on note $q = 1 - p$.
On dit que la variable aléatoire X suit **la loi géométrique de paramètre** p (on note $X \hookrightarrow \mathscr{G}(p)$) si

$$X(\Omega) = \mathbb{N}^* \text{ et pour tout } k \in \mathbb{N}^*,\ \mathbb{P}(X = k) = pq^{k-1}\ .$$

Situations de référence

➢ Temps d'attente du premier succès dans une répétition infinie et indépendante d'une épreuve à deux

issues, avec succès de probabilité p.
➢ Lancers successifs et indépendants d'un dé, succès si l'on obtient 6. X égale au nombre de lancers nécessaire pour obtenir le premier 6 : $X \hookrightarrow \mathcal{G}(\frac{1}{6})$.
➢ Lancers successifs et indépendants d'une pièce, succès si l'on obtient PILE. X égale au nombre de lancers nécessaire pour obtenir le premier PILE : $X \hookrightarrow \mathcal{G}(\frac{1}{2})$.

Exercice 1

1. On suppose que la variable aléatoire X suit la loi géométrique de paramètre $p \in]0,1[$. Montrer que $\forall l, k \in \mathbb{N}^*$, $\mathbb{P}(X > k + l \mid X > k) = \mathbb{P}(X > l)$.
On dit que la loi géométrique est une loi sans mémoire.
2. Soit X une variable aléatoire prenant toutes les valeurs de $\mathbb{N}^*$. On suppose que $\mathbb{P}(X = 1) \in]0,1[$ et que $\forall l, k \in \mathbb{N}^*$, $\mathbb{P}(X > k + l \mid X > k) = \mathbb{P}(X > l)$. Montrer que X suit une loi géométrique.

b. *Loi de Poisson*

Définition 9 : Soit λ un réel strictement positif.
On dit que la variable aléatoire X suit **la loi de Poisson de paramètre** λ (on note $X \hookrightarrow \mathcal{P}(\lambda)$) si

$$X(\Omega) = \mathbb{N} \text{ et pour tout } k \in \mathbb{N},\ \mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}\ .$$

Interprétation : On montrera que la loi de Poisson est une approximation de la loi binomiale. On dit que c'est la "loi des événements rares". Elle permet de modéliser une grande variété de phénomènes : le nombre d'appels arrivant à un standard pendant un laps de temps, le nombre de petits pois dans une portion de riz cantonnais, etc. Si λ est le nombre moyen de réalisation de ce que l'on observe (durant un laps de temps, sur un domaine, etc.), alors la loi de Poisson donne la probabilité de voir se réaliser notre observation k fois sur ce laps de temps.

4. Couples de variables aléatoires

On considère deux variables aléatoires réelles discrètes X et Y sur un même espace probabilisé Ω.
On rappelle les définitions associées à la notion de couple (X, Y) de variables aléatoires.
➢ La loi de probabilité de ce couple est appelé **loi conjointe** de (X, Y). C'est la donnée de

- l'ensemble $X(\Omega) \times Y(\Omega)$ des valeurs prises par le couple (X, Y),
- $\forall (x, y) \in X(\Omega) \times Y(\Omega)$, $\mathbb{P}(X = x, Y = y)$, qui désigne la probabilité de l'événement $(X = x) \cap (Y = y)$.

➢ Les lois de X et de Y sont alors appelées **lois marginales** du couple (X, Y).
➢ Soit $x \in X(\Omega)$ tel que $\mathbb{P}(X = x) > 0$, on appelle **loi conditionnelle de** Y **sachant** $(X = x)$, la probabilité $\mathbb{P}_{(X=x)}$ sur $Y(\Omega)$. Celle -ci est définie par $\forall y \in Y(\Omega)$, $\mathbb{P}_{(X=x)}(Y = y) = \dfrac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(X = x)}$.
On définit de même, pour $y \in Y(\Omega)$ tel que $\mathbb{P}(Y = y) > 0$, la loi conditionnelle de X sachant $Y = y$.
On parlera aussi de la loi de X conditionnée à $Y = y$.
On a donc pour $(x, y) \in X(\Omega) \times Y(\Omega)$ tels que $\mathbb{P}(X = x) > 0$ et $\mathbb{P}(Y = y) > 0$,

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}_{(X=x)}(Y = y)\mathbb{P}(X = x) = \mathbb{P}_{(Y=y)}(X = x)\mathbb{P}(Y = y)\ .$$

➢ Soit (X, Y) un couple de variables aléatoires réelles discrètes sur Ω.
On dit que (X, Y) est un couple de variables aléatoires indépendantes si

$$\forall (x, y) \in X(\Omega) \times Y(\Omega),\ \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)\ .$$

On reprendra de même les définitions d'indépendance 2 à 2 et d'indépendance mutuellle pour une famille finie $(X_1, X_2, \ldots, X_n)$ de variables discrètes sur Ω.

Exercice 2

À la cantine, deux menus sont proposés chaque jour. Le nombre d'élèves N se présentant chaque jour suit une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$. Les élèves choisissent leur repas indépendamment les uns des autres et le menu 1 est choisi avec une probabilité p. On note X la variable égale au nombre d'élèves choisissant le menu 1. Déterminer la loi de X.

Remarque 6 :

➢ On admet l'existence d'un espace probabilisé Ω portant une suite de variables aléatoires indépendantes de lois données. Cela permet par exemple de modéliser la répétition d'une même épreuve un nombre infini de fois avec indépendance des résultats de chaque épreuve.

➢ Pour déterminer la loi de X (ou de Y) connaissant la loi conjointe du couple (X,Y), on utilise la formule des probabilités totales avec le système complet d'événements $(Y=y)_{y\in Y(\Omega)}$.

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

IV. ESPÉRANCE ET VARIANCE

1. Espérance

Définition 10 :

Soit X une variable aléatoire réelle discrète à valeurs dans $X(\Omega) = \{x_n,\ n\in I\}$, où I est une partie de $\mathbb{N}$.

On dit que X est d'**espérance finie** si la série $\sum\limits_{n\in\mathbb{N}} x_n\mathbb{P}(X=x_n)$ est absolument convergente.

Dans ce cas, l'espérance de X est la somme de la série $\mathbb{E}(X) = \sum\limits_{n=0}^{+\infty} x_n\mathbb{P}(X=x_n)$.

Remarque 7 :

a. Si I est un ensemble fini, on note $X(\Omega) = \{x_1, x_2, \dots, x_N\}$ ($N\in\mathbb{N}^*$) : X est alors une variable aléatoire finie. X est d'espérance finie et son espérance est le réel $\mathbb{E}(X) = \sum\limits_{n=1}^{N} x_n\mathbb{P}(X=x_n)$ *(c'est une somme finie).*

b. Pour une série absolument convergente, l'ordre de sommation ne change pas le résultat de la somme, donc la valeur de $\mathbb{E}(X)$ (ce n'est pas le cas pour une série semi-convergente).

c. On travaille avec des séries absolument convergentes : on se ramène donc à des séries positives. Dans ce cas (lorsque X est une variable aléatoire à valeurs positives), on peut toujours définir l'espérance en acceptant le cas où elle est infinie.

Exemple 8 ♥

- Loi uniforme : Soit $N\in\mathbb{N}^*$, si $X \hookrightarrow \mathscr{U}([\![1,N]\!])$, alors $\mathbb{E}(X) = \dfrac{N+1}{2}$.
- Loi binomiale : Soit $n\in\mathbb{N}^*$ et $p\in[0,1]$, si $X\hookrightarrow \mathscr{B}(n,p)$, alors $\mathbb{E}(X) = np$.
- Loi géométrique : Soit $p\in]0,1[$, si $X\hookrightarrow\mathscr{G}(p)$, alors $\mathbb{E}(X) = \dfrac{1}{p}$.
- Loi de Poisson : Soit $\lambda\in\mathbb{R}^*$, si $X\hookrightarrow\mathscr{P}(\lambda)$, alors $\mathbb{E}(X) = \lambda$.

Proposition 8 *Admis*

a. *Linéarité*

Soient X et Y deux variables aléatoires discrètes sur Ω, admettant des espérances finies. Soient a, b des réels. La variable aléatoire $aX+bY$ est d'espérance finie et $\mathbb{E}(aX+bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.

b. *Positivité*
Soit X un variable aléatoire discrète sur Ω ne prenant que des valeurs réelles positives, et d'espérance finie. On a alors $\mathbb{E}(X) \geqslant 0$.

c. *Croissance*
Soient X et Y deux variables aléatoires discrètes sur Ω, admettant des espérances finies.
Si $\forall \omega \in \Omega, X(\omega) \leqslant Y(\omega)$ alors $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$.

Remarque 8 : Un cas particulier de la linéarité est le suivant : pour X une variable aléatoire discrète admettant une espérance finie, pour tout $(a,b) \in \mathbb{R}^2$, la variable $aX+b$ admet une espérance finie et $\mathbb{E}(aX+b) = a\mathbb{E}(X)+b$.

Exemple 9 (utilisation de la linéarité)
On lance un dé indéfiniment, on note X le rang du premier 6 obtenu et Y le nombre de lancers nécessaires avant d'obtenir un 6. X suit une loi géométrique de paramètre $p = \frac{1}{6}$. On a $Y = X - 1$, et on peut en déduire que Y est d'espérance finie et que $\mathbb{E}(Y) = \mathbb{E}(X-1) = \mathbb{E}(X) - 1 = \frac{1}{p} - 1 = 5$.

Proposition 9 *Théorème du transfert (admis)*

Soit X une variable aléatoire réelle discrète. On note $X(\Omega) = \{x_n,\ n \in \mathbb{N}\}$.
Soit f une application à valeurs réelles définie sur $X(\Omega)$.
Alors $f(X)$ est d'espérance finie si et seulement si $\sum\limits_{n\in\mathbb{N}} f(x_n)\mathbb{P}(X = x_n)$ est absolument convergente.
Dans ce cas, on a $\mathbb{E}(f(X)) = \sum\limits_{n=0}^{+\infty} f(x_n)\mathbb{P}(X = x_n)$.

Exercice 3 ✍ : ➢ Soit $X \hookrightarrow \mathscr{G}(p)$, avec $p \in]0,1[$. Déterminer $\mathbb{E}(X(X-1))$. En déduire $\mathbb{E}(X^2)$

Proposition 10

Soit X une variable aléatoire réelle discrète sur Ω à valeurs dans $\mathbb{N}$. On suppose X d'espérance finie.
On a alors $\mathbb{E}(X) = \sum\limits_{n=1}^{+\infty} \mathbb{P}(X \geqslant n)$

démo Voici les étapes et les principaux arguments : on utilise, pour $N \in \mathbb{N}^*$ fixé,
➢ pour $n \in [\![1,N]\!]$, $\mathbb{P}(X \geqslant n) = \mathbb{P}(X \geqslant N+1) + \sum\limits_{k=n}^{N} \mathbb{P}(X = k)$ (Justifier ceci)
➢ et on obtient en faisant la somme, $\sum\limits_{n=1}^{N} \mathbb{P}(X \geqslant n) = N\mathbb{P}(X \geqslant N+1) + \sum\limits_{n=1}^{N}\sum\limits_{k=n}^{N} \mathbb{P}(X = k)$.
On intervertit la somme double pour se ramener à $\sum\limits_{n=1}^{N} \mathbb{P}(X \geqslant n) = N\mathbb{P}(X \geqslant N+1) + \sum\limits_{k=1}^{N} k\mathbb{P}(X = k)$ (expliquer).
➢ Il faut ensuite montrer que $\lim\limits_{N\to+\infty} N\mathbb{P}(X \geqslant N+1)$ existe et vaut 0 pour pouvoir conclure (en argumentant au passage sur l'existence des limites).
En utilisant $\mathbb{P}(X \geqslant N+1) = \sum\limits_{k=N+1}^{+\infty} \mathbb{P}(X = k)$, on majore $0 \leqslant N\mathbb{P}(X \geqslant N+1) \leqslant \sum\limits_{k=N+1}^{+\infty} k\mathbb{P}(X = k)$.
On conclut alors en utilisant le reste de la série $\sum k\mathbb{P}(X = k)$ (supposée convergente).

Proposition 11

Soient (X,Y) un couple de variables aléatoires réelles discrètes telles que X^2 et Y^2 sont d'espérances finies. On admet qu'alors XY est d'espérance finie.

a. *Inégalité de Cauchy-Schwarz* : on a $(\mathbb{E}(XY))^2 \leqslant \mathbb{E}(X^2)\mathbb{E}(Y^2)$.

b. (*Admis*) Si X et Y sont indépendantes, alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

2. Variance

Proposition 12

Soit X une variable aléatoire réelle discrète sur un espace probabilisé.
Si X^2 est d'espérance finie alors X est également d'espérance finie.

démo On traite le cas $X(\Omega) = \{x_n,\ n \in \mathbb{N}\}$. On suppose donc que X^2 est d'espérance finie.

➢ la série $\sum x_n^2 \mathbb{P}(X = x_n)$ est donc ACV (thm du transfert).

➢ Pour tout $n \in \mathbb{N}$, $|x_n| \leqslant \dfrac{1 + x_n^2}{2}$, donc $|x_n \mathbb{P}(X = x_n)| \leqslant \dfrac{1}{2}(\mathbb{P}(X = x_n) + x_n^2 \mathbb{P}(X = x_n))$.

➢ Les séries $\sum \mathbb{P}(X = x_n)$ et $\sum x_n^2 \mathbb{P}(X = x_n)$ convergent, on conclut donc par comparaison de séries à termes positifs.

Définition 11 :
Soit X une variable aléatoire réelle discrète (finie ou dénombrable) telle que X^2 admet une espérance finie.
On appelle **variance** de X le réel $V(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$.
On appelle **écart-type** de X le réel $\sigma(X) = \sqrt{V(X)}$.

Remarque 9 :

- On dira dans ce cas que X admet une variance finie.
- On sait aussi que $(X - \mathbb{E}(X))^2$ est une variable aléatoire à valeurs positives, donc son espérance est positive. Cela justifie la définition de $\sigma(X)$.

Proposition 13

Soit X une variable aléatoire réelle discrète telle que X^2 admet une espérance finie.
On a $V(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.

démo : On calcule en utilisant la linéarité de l'espérance

$\mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2\mathbb{E}(X)X + (\mathbb{E}(X))^2) = \mathbb{E}(X^2) - 2\mathbb{E}(X)\,\mathbb{E}(X) + (\mathbb{E}(X))^2 = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.

Et on montre de même que $V(X) = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - (\mathbb{E}(X))^2$.

Exemple 10 ♥

- Loi uniforme : Soit $N \in \mathbb{N}^*$, si $X \hookrightarrow \mathscr{U}([\![1, N]\!])$ alors $V(X) = \dfrac{N^2 - 1}{12}$.
- Loi binomiale : Soit $n \in \mathbb{N}^*$ et $p \in [0,1]$, $q = 1 - p$, si $X \hookrightarrow \mathscr{B}(n,p)$ alors $V(X) = npq$.
- Loi géométrique : Soit $p \in]0,1[$, $q = 1 - p$, si $X \hookrightarrow \mathscr{G}(p)$ alors $V(X) = \dfrac{q}{p^2}$.
- Loi de Poisson : Soit $\lambda \in \mathbb{R}^*$, si $X \hookrightarrow \mathscr{P}(\lambda)$ alors $V(X) = \lambda$.

Proposition 14

Soit X une variable aléatoire réelle discrète admettant une variance finie.

- Pour tout $(a,b) \in \mathbb{R}^2$, on a $V(aX + b) = a^2 V(X)$.
- Si $\mathbb{E}(X^2) = 0$ alors $\mathbb{P}(X = 0) = 1$.
- Si $V(X) = 0$ alors $\mathbb{P}(X = \mathbb{E}(X)) = 1$.

3. Inégalité de Markov et de Bienaymé-Tchebychev

On rappelle les énoncés des inégalités, dans le cadre d'une VA discrète.

Proposition 15 *Inégalité de Markov*

> Soit X une variable aléatoire réelle discrète à valeurs positives admettant une espérance finie.
> Soit $a > 0$, on a alors $\mathbb{P}(X \geqslant a) \leqslant \dfrac{\mathbb{E}(X)}{a}$.

Proposition 16 *Inégalité de Bienaymé-Tchebychev*

> Soit X une variable aléatoire réelle discrète telle que l'espérance de X^2 soit finie.
> On a alors, en notant $\mu = \mathbb{E}(X)$ et $\sigma = \sqrt{V(X)}$.
>
> $$\forall \varepsilon > 0,\ \mathbb{P}(|X-\mu| \geqslant \varepsilon) \leqslant \frac{\sigma^2}{\varepsilon^2}\ .$$

4. Covariance et coefficient de corrélation

On reprend ici les définitions données dans le cas des variables aléatoires finies. On suppose que X^2 et Y^2 sont d'espérance finie.

> **Définition** 12 : Soit (X, Y) un couple de variables aléatoires réelles discrètes sur un espace probabilisé Ω. On suppose que X^2 et Y^2 sont d'espérance finie.
> ➢ On appelle **covariance de** X **et** Y le réel : $\text{Cov}(X,Y) = \mathbb{E}((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.
> ➢ Si les variances de X et Y sont non nulles, on définit le coefficient de corrélation linéaire de X et Y par $\rho(X,Y) = \dfrac{\text{Cov}(X,Y)}{\sqrt{V(X)V(Y)}}$.

Proposition 17

> Soient X, Y, Z des variable aléatoire réelles discrètes, de variance finie.
> - $\text{Cov}(X,X) = V(X)$
> - $\text{Cov}(X,Y) = \text{Cov}(Y,X)$
> - pour tous $a,\ b \in \mathbb{R}$, $\text{Cov}(aX+bY, Z) = a\text{Cov}(X,Z) + b\text{Cov}(Y,Z)$
> - $V(X+Y) = V(X) + V(Y) + 2\text{Cov}(X,Y)$
> - On a $-1 \leqslant \rho(X,Y) \leqslant 1$. De plus, si $|\rho(X,Y)| = 1$ alors il existe $a,\ b \in \mathbb{R}$ tels que $Y = aX + b$ est certain.

Remarque 10 :

- On a vu que si X et Y sont indépendantes, alors $\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0$, donc $\text{Cov}(X,Y) = 0$.
  Attention, la réciproque est fausse.
- On peut donner une généralisation de la propriété sur $X + Y$:

$$\forall (a,b) \in \mathbb{R}^2,\ V(aX+bY) = a^2V(X) + b^2V(Y) + 2ab\text{Cov}(X,Y)\ .$$

- On peut aussi généraliser avec plusieurs variables aléatoires $X_1, \ldots, X_n$ dont le carré admet une espérance finie :

$$V(\sum_{i=1}^n X_i) = \sum_{i=1}^n V(X_i) + 2\sum_{1\leqslant i<j\leqslant n} \text{Cov}(X_i, X_j)\ .$$

V. SÉRIES GÉNÉRATRICES

1. Définition

Définition 13 : Soit X une variable aléatoire réelle discrète définie sur Ω telle que $X(\Omega) \subset \mathbb{N}$.
On appelle **série génératrice de** X la fonction somme de la série $\sum\limits_{n\geqslant 0} \mathbb{P}(X=n)\,t^n$, notée G_X :

$$G_X\;:\;t\mapsto \sum_{n=0}^{+\infty} \mathbb{P}(X=n)\,t^n\,.$$

Remarque 11 :

- On appelle ici (conformément au programme) série génératrice une fonction .... Vous trouverez dans d'autres ouvrages le terme de fonction génératrice. On remarquera aussi que si $X(\Omega)$ est fini, G_X est alors une fonction polynomiale (la somme est finie).
- Notons R le rayon de convergence de la série $\sum\limits_{n\geqslant 0} \mathbb{P}(X=n)\,t^n$.
  Pour tout $t \in]-R,R[$, $G_X(t)=\mathbb{E}(t^X)$ (l'espérance est ici bien définie d'après le théorème de tranfert).

Proposition 18

Soit X une variable aléatoireréelle discrète définie sur Ω telle que $X(\Omega)\subset\mathbb{N}$.
- Le rayon de la série $\sum\limits_{n\geqslant 0} \mathbb{P}(X=n)\,t^n$ est supérieur ou égal à 1 et $G_X(1)=1$
- La loi d'une variable aléatoire réelle discrète est entièrement déterminée par la donnée de G_X.

démo

Exemple 11

- Loi uniforme : Soit $N\in\mathbb{N}^*$, si $X\hookrightarrow \mathscr{U}([\![1,N]\!])$ alors $\forall\, t\in\mathbb{R},\; G_X(t)=\dfrac{t}{N}\sum\limits_{k=0}^{N-1} t^k$.
- Loi binomiale : Soit $n\in\mathbb{N}^*$ et $p\in[0,1]$, $q=1-p$, si $X\hookrightarrow\mathscr{B}(n,p)$ alors $\forall\, t\in\mathbb{R},\; G_X(t)=(pt+q)^n$.
- Loi géométrique : Soit $p\in]0,1[$, $q=1-p$, si $X\hookrightarrow\mathscr{G}(p)$ alors $\forall\, t\in]-\frac{1}{q},\frac{1}{q}[,\; G_X(t)=\dfrac{pt}{1-qt}$.
- Loi de Poisson : Soit $\lambda\in\mathbb{R}^*$, si $X\hookrightarrow\mathscr{P}(\lambda)$ alors $\forall\, t\in\mathbb{R},\; G_X(t)=e^{\lambda(t-1)}$.

2. Lien avec espérance et variance

Proposition 19

Soit X un variable aléatoire réelle discrète à valeurs dans $\mathbb{N}$.

a. X admet une espérance finie si et seulement si G_X est dérivable en 1 et dans ce cas : $G'_X(1)=\mathbb{E}(X)$.

b. X admet une variance finie si et seulement si G_X est deux fois dérivable en 1, et dans ce cas :
$G''_X(1)=\mathbb{E}(X(X-1))$ et $V(X)=G''_X(1)+G'_X(1)-(G'_X(1))^2$.

démo

3. Série génératrice d'une somme

Proposition 20

Soient X, Y deux variables aléatoires réelles discrètes à valeurs dans $\mathbb{N}$, **indépendantes**.
On a alors $G_{X+Y}=G_X\,G_Y$

démo : C'est un produit de Cauchy de deux séries entières.

VI. RÉSULTATS ASYMPTOTIQUES

1. Approximation de la loi binomiale par la loi de Poisson

Proposition 21

Soit $\lambda \in \mathbb{R}$. Si, pour tout $n \in \mathbb{N}$, la variable aléatoire X_n suit une loi binomiale $\mathscr{B}(n, p_n)$ avec (p_n) une suite d'éléments de $]0,1[$ telle que $\lim_{n\to+\infty} np_n = \lambda$, alors pour tout $k \in \mathbb{N}$, on a

$$\lim_{n\to+\infty} \mathbb{P}(X_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

Remarque 12 : On dit que la loi de Poisson donne une approximation de la loi binomiale. On utilise cette approximation lorsque n est grand et lorsque np_n n'est pas trop grand !
Dans la pratique, il semble admis que l'on peut approcher une loi binomiale par une loi de Poisson lorsque $n \geqslant 30$, que $p \leqslant 0,1$.

2. Loi faible des grands nombres

Proposition 22

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires 2 à 2 indépendantes et de même loi admettant une variance finie. On note μ leur espérance et σ leur écart-type. On pose $S_n = \sum_{k=1}^{n} X_k$.
On a pour tout $\varepsilon > 0$, $\lim_{n\to+\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geqslant \varepsilon\right) = 0$

démo : Soit $n \in \mathbb{N}^*$, on a $S_n = X_1 + \cdots + X_n$. On utilise l'inégalité de Bienaymé-Tchebychev pour $Y = \frac{S_n}{n}$.
L'espérance est linéaire, donc S_n admet une espérance finie et $\mathbb{E}(S_n) = n\mu$, et donc $\mathbb{E}(Y) = \mu$.
Chaque variable X_i admet une variance finie et ces variables X_1,...X_n sont indépendantes donc Y admet une variance finie et $V(Y) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{\sigma^2}{n}$.
Ainsi, pour tout $\varepsilon > 0$, $\mathbb{P}(|Y - \mathbb{E}(Y)| \geqslant \varepsilon) \leqslant \frac{V(Y)}{\varepsilon^2}$, c'est-à-dire $0 \leqslant \mathbb{P}(|\frac{S_n}{n} - \mu| \geqslant \varepsilon) \leqslant \frac{\sigma^2}{n\varepsilon^2}$. Or ceci fournit une majoration par une suite qui tend vers 0, quand n tend vers $+\infty$, on obtient donc la conclusion souhaitée.