# CH X : Variables aléatoires discrètes

## I. Espérance d'une variable aléatoire discrète

### I.1. Définition

**Définition**

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

Soit $X$ une v.a. discrète sur $(\Omega, \mathscr{A})$.

*1)* Si $X$ est finie : et que $X(\Omega) = \{x_1, \dots, x_n\}$.
Dans ce cas, on appelle **espérance** de la v.a. $X$ et on note $\mathbb{E}(X)$ :

$$\boxed{\mathbb{E}(X) = \sum_{i=1}^{n} x_i \, \mathbb{P}(\{X = x_i\})}$$

*2)* Si $X$ est infinie : et que $X(\Omega) = \{x_i \mid i \in \mathbb{N}\}$.
Dans ce cas, **si la série** $\sum x_n \, \mathbb{P}([X = x_n])$ **est absolument convergente**, on appelle **espérance** de la v.a. $X$ et on note $\mathbb{E}(X)$ :

$$\boxed{\mathbb{E}(X) = \sum_{i=0}^{+\infty} x_i \, \mathbb{P}(\{X = x_i\})}$$

*3)* De manière générale, si on écrit $X(\Omega) = \{x_i \mid i \in I\}$ (avec $I \subset \mathbb{N}$) alors, sous réserve d'existence :

$$\boxed{\mathbb{E}(X) = \sum_{x \in X(\Omega)} x \, \mathbb{P}(\{X = x\}) = \sum_{i \in I} x_i \, \mathbb{P}(\{X = x_i\})}$$

L'espérance est une moyenne pondérée : on somme chaque valeur possible pour $X$ pondérée par la probabilité que $X$ prenne cette valeur.

**Remarque**

L'espérance peut être pensée comme une généralisation de la notion de moyenne. Illustrons ce point avec l'exemple suivant.

- Expérience aléatoire : on effectue 1 lancer d'un dé à 6 faces.
- Univers : $\Omega = [\![1, 6]\!]$.
- Notons $X$ la v.a.r. égale au résultat du lancer.
- $X(\Omega) = [\![1, 6]\!]$ donc $X$ est donc une v.a.r. discrète finie.

  *1)* Dans un premier temps, munissons $\Omega$ de la probabilité uniforme $\mathbb{P}_1$.
  La v.a.r. $X$ étant finie, elle admet une espérance donnée par :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{x \in X(\Omega)} x \, \mathbb{P}_1(\{X = x\}) = \sum_{k=1}^{6} k \, \mathbb{P}_1([X = k]) \\ &= \frac{1}{6} \sum_{k=1}^{6} k = \frac{1}{\cancel{6}} \frac{\cancel{6}(6+1)}{2} = 3,5 \end{aligned}$$

  Le réel $3,5$ est la moyenne des résultats du lancer d'un dé équilibré (probabilité $\mathbb{P}_1$ prise en compte).

  *2)* On munit maintenant $\Omega$ de la probabilité $\mathbb{P}_2$ telle que :

$$\mathbb{P}(\{4\}) = \mathbb{P}(\{5\}) = \mathbb{P}(\{6\}) = \frac{1}{3}$$

  La v.a.r. $X$ étant finie, elle admet une espérance donnée par :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{x \in X(\Omega)} x \, \mathbb{P}_2(\{X = x\}) = \sum_{k=1}^{6} k \, \mathbb{P}_2([X = k]) \\ &= \sum_{k=4}^{6} k \, \mathbb{P}_2([X = k]) = \frac{1}{3} \sum_{k=4}^{6} k = \frac{1}{\cancel{3}} \frac{\cancel{3}(4+6)}{2} = 5 \end{aligned}$$

  Le réel 5 est la moyenne des résultats du lancer dans le cas de notre dé truqué (probabilité $\mathbb{P}_1$ prise en compte).

- Une v.a. discrète `FINIE` admet `TOUJOURS` une espérance.
- Une v.a. discrète `INFINIE` n'admet pas nécessairement une espérance. L'hypothèse de `CONVERGENCE ABSOLUE` est fondamentale pour la bonne définition de la notion d'espérance.

### À quoi sert l'hypothèse de convergence absolue ?

Il faut bien comprendre que la notation $\sum\limits_{x \in X(\Omega)} x \, \mathbb{P}(\{X = x\})$ ne permet pas de savoir dans quel ordre sont sommés les éléments.
Il n'y paraît rien, mais c'est primordial comme l'énonce le résultat suivant.

**Théorème 1.** `(CULTURE)`

*Soit $\sum u_n$ une série semi-convergente, c'est-à-dire $\left\{\begin{array}{l} \sum u_n \text{ converge} \\ \sum |u_n| \text{ diverge} \end{array}\right.$.*
*Pour tout $\alpha \in \mathbb{R} \cup \{-\infty, +\infty\}$, il existe une permutation $\sigma$ de $\mathbb{N}$ telle que :*

$$\sum_{k=0}^{n} u_{\sigma(k)} \underset{n \to +\infty}{\longrightarrow} \alpha$$

*Démonstration.*
On ne prétend pas ici démontrer le résultat mais donner un schéma de comment il se démontre dans le cas où $\alpha$ est un réel positif (par exemple).

- On démontre tout d'abord que la suite $(u_n)$ contient une infinité de termes positifs et une infinité de termes négatifs. Si ce n'était pas le cas, la suite $(u_n)$ serait de signe constant à partir d'un certain rang et la convergence de la série $\sum u_n$ serait alors équivalente à sa convergence absolue.
- Remarquons : $\max(0, u_n) = \dfrac{u_n + |\, u_n \,|}{2}$ et donc : $|\, u_n \,| = 2 \max(0, u_n) - u_n$.
  Ainsi, la série $\sum \max(0, u_n)$ est divergente (si ce n'était pas le cas, $\sum |\, u_n \,|$ serait convergente). Cette série correspond à la série $\sum u_n$ où l'on aurait remplacé tous les termes négatifs par 0. En terme de somme, elle peut être assimilée à la série des termes positifs de la suite $(u_n)$ (celle construite en sommant dans l'ordre les termes positifs de $(u_n)$).
  Or, la série $\sum \max(0, u_n)$ est :
  × à termes positifs,
  × divergente.
  La suite de ses sommes partielles tend donc vers $+\infty$. Ainsi, en sommant (dans l'ordre) suffisamment de termes positifs de $(u_n)$, on peut créer une quantité plus grande que n'importe quel réel positif fixé à l'avance.
- On peut considérer de la même manière la série $\sum \min(0, u_n)$, série qu'on peut assimiler à celle des termes négatifs de la suite $(u_n)$.
  Comme $\min(0, u_n) = \dfrac{u_n - |\, u_n \,|}{2}$, la série $\sum \min(0, u_n)$ est elle aussi divergente et la suite de ses sommes partielles tend donc vers $-\infty$. Ainsi, en sommant (dans l'ordre) suffisamment de termes négatifs de $(u_n)$, on peut donc créer une quantité aussi grande, dans les négatifs, que souhaitée.
- On construit alors la somme $\sum\limits_{k=0}^{n} u_{\sigma(k)}$ par le principe suivant :
  ***1)*** on ajoute les éléments positifs non encore utilisés de la suite $(u_n)$ (dans l'ordre des indices) tant que la somme obtenue ne dépasse pas $\alpha$.
  *(par l'étude précédente, on sait qu'en sommant des termes positifs on finira par dépasser $\alpha$)*
  ***2)*** on ajoute les éléments négatifs non encore utilisés de la suite $(u_n)$ (dans l'ordre des indices) tant que la somme obtenue dépasse strictement $\alpha$.
  *(par l'étude précédente, on sait qu'en sommant des termes négatifs on finira par repasser en dessous de $\alpha$)*
  On définit alors $\sigma(i)$ comme étant l'indice du $i^{\text{ème}}$ élément ajouté.
- Pour conclure quant à la qualité de cette stratégie, il faut comprendre que la somme ainsi construite se rapproche de $\alpha$. Cela se produit car le dépassement de $\alpha$ (par au-dessus ou en dessous) est de plus en plus faible. En effet, chaque dépassement (passage du mode ***1)*** à ***2)*** ou inversement) provient de l'ajout d'un terme de la suite $(u_n)$. Or cette suite tend vers 0 car la série $\sum u_n$ converge. Ainsi les dépassements sont de plus en plus faibles.

□

**Remarque**

- Cet énoncé est appelé théorème de réarrangement de Riemann.
- Il signifie qu'étant donnée une série semi-convergente, on peut, en modifiant l'ordre de sommation, créer une série qui converge vers n'importe quel $\alpha \in \mathbb{R}$ choisit à l'avance (on peut même prendre $\alpha = +\infty$ ou $\alpha = -\infty$).
- Ce résultat est contraire à l'intuition : on ne s'attend pas à ce que la somme de tous les termes $u_i$ dépende de l'ordre de sommation.
- Ceci ne se produit que sous l'hypothèse de semi-convergence.
  L'hypothèse de convergence absolue est celle qui permet de parler de convergence d'une série indépendamment de l'ordre de sommation choisi.
- Si la suite $(u_n)$ est à termes positifs, alors convergence de $\sum\ u_n$ et convergence absolue de $\sum\ u_n$ sont des propriétés équivalentes. Cela amène à s'intéresser de plus près aux sommations de termes positifs.

## I.2. Familles sommables

### I.2.a) L'ensemble $[0,+\infty]$

- On prolonge l'ensemble $[0,+\infty[$ en lui ajoutant $+\infty$. L'objectif de cet ajout est de simplifier la manipulation de sommes d'éléments positifs.
- On prolonge alors les opérations $+$ et $\times$ ainsi que la relation d'ordre sur $[0,+\infty]$ par les relations habituelles, à savoir :
  × $\forall x \in\ ]0,+\infty],\ x\ \times\ (+\infty) = (+\infty)\ \times x = +\infty$
  × $\forall x \in\ [0,+\infty],\ x\ +\ (+\infty) = (+\infty)\ +\ x = +\infty$
- On rappelle que toute partie $A$ de $\mathbb{R}$ non vide et majorée admet une borne supérieure (un élément réel qui est le plus petit des majorants de $A$). Si l'on considère une partie non vide $A$ de $[0,+\infty]$, alors cette partie est toujours majorée par $+\infty$. En ce sens, elle admet une forcément une borne supérieure dans $[0,+\infty]$. Cette borne supérieure étendue coïncide avec la notion de borne supérieure abordée en première année pour peu que la partie $A$ soit majorée.

### I.2.b) Notion de famille sommable

**Définition**

Soit $I$ un ensemble au plus dénombrable.
Soit $(x_i)_{i\in I}$ une famille d'éléments de $[0,+\infty]$.

- On appelle somme de la famille $(x_i)_{i\in I}$, qu'on note $\sum\limits_{i\in I} x_i$, la borne supérieure (dans $[0,+\infty]$) de l'ensemble des sommes $\sum\limits_{i\in F} u_i$ lorsque $F$ décrit l'ensemble des parties finies de $I$.
- En d'autres termes :

$$\boxed{\sum_{i\in I} x_i \ = \ \sup \left\{ \sum_{i\in F} x_i \mid F \in \mathscr{P}_f(I) \right\}}$$

- La famille $(x_i)_{i\in I}$ est dite **sommable** si $\sum\limits_{i\in I} x_i < +\infty$.

**Remarque**

- Pour bien comprendre la notion de famille sommable, revenons rapidement sur la notion de séries de termes positifs. Notons $\sum\ a_n$ une telle série et $(S_n)$ la suite des sommes partielles associée. Alors $(S_n)$ est une suite croissante ($\forall n \in \mathbb{N},\ S_{n+1} - S_n = a_{n+1} \geqslant 0$). Ainsi, deux cas se présentent :
  × si la suite $(S_n)$ est majorée, alors la suite $(S_n)$ converge et on dit que la série est convergente.
  × si la suite $(S_n)$ n'est pas majorée, alors la suite $(S_n)$ diverge vers $+\infty$.

  Ce que ce point met en avant c'est que, concernant les séries de termes positifs, il existe seulement 2 modes : soit la suite des sommes partielles converge, soit elle diverge vers $+\infty$.

- En travaillant dans $[0,+\infty]$, on opère un changement de point de vue. Prenons par exemple la somme $\sum\limits_{i=1}^{+\infty} \frac{1}{i}$.

*1)* Point de vue séries
La série $\sum \frac{1}{n}$ est divergente. On ne peut donc pas pas considérer sa somme et l'écriture ci-dessus n'est donc pas autorisée.

*2)* Point de vue manipulation dans $[0,+\infty]$
Cette somme a tout son sens lorsqu'on travaille dans $[0,+\infty]$. Plus précisément, $\sum\limits_{i=1}^{+\infty} \frac{1}{i} = +\infty$. La famille $\left(\frac{1}{i}\right)_{i\in\mathbb{N}^*}$ n'est donc pas sommable.

- Illustrons le point précédent avec un autre exemple.
Remarquons : $\forall x \geqslant 0,\ \mathrm{e}^x - 1 \ \geqslant\ x$. On en déduit :

$$\forall n \in \mathbb{N}^*,\ \mathrm{e}^{\frac{1}{n}} - 1 \ \geqslant\ \frac{1}{n}$$

*1)* Point de vue séries
Comme la série $\sum \frac{1}{n}$ est divergente, alors, par théorème de comparaison des séries à termes positifs, la série $\sum \left(\mathrm{e}^{\frac{1}{n}} - 1\right)$ est divergente.

*2)* Point de vue manipulation dans $[0,+\infty]$

$$\sum_{n=1}^{+\infty} \left(\mathrm{e}^{\frac{1}{n}} - 1\right) \ \geqslant\ \sum_{n=1}^{+\infty} \frac{1}{n} = +\infty$$

Ainsi, la famille $\left(\mathrm{e}^{\frac{1}{n}} - 1\right)_{n\in\mathbb{N}^*}$ n'est pas sommable.

### I.2.c) Propriétés de manipulation

**Théorème 2.** *(sommation par paquets)*
*Soit $I$ un ensemble au plus dénombrable.*
*Soit $(I_n)_{n\in\mathbb{N}}$ une partition de l'ensemble $I$.*
*Autrement dit, $I$ s'écrit comme une réunion disjointe : $I = \bigcup\limits_{n\in\mathbb{N}} I_n$.*
*Soit $(x_i)_{i\in I}$ une famille d'éléments de $[0,+\infty]$.*

$$\boxed{\sum_{i\in I} x_i \ = \ \sum_{n=0}^{+\infty} \left(\sum_{i\in I_n} x_i\right)}$$

**Remarque**

- La sommation par paquets peut être considérée comme un réarrangement de l'ordre de sommation. Du fait du théorème de réarrangement de Riemann, un tel procédé ne peut fonctionner de manière générale pour les séries semi-convergentes.
- Pour illustrer les problèmes que la sommation par paquet peut causer, considérons la série $\sum\limits_{n\geqslant 1} \frac{(-1)^{n+1}}{n}$ et notons $u_n$ son terme général.
Considérons deux partitions différentes $(I_n)_{n\in\mathbb{N}}$ et $(J_n)_{n\in\mathbb{N}}$ définies par :

$$\forall n \in \mathbb{N}, \quad I_n = \{2n+1, 4n+2, 4n+4\} \quad \text{et} \quad J_n = \{2n+1, 2n+2\}$$

On a alors, pour tout $n \in \mathbb{N}$ :

$$\begin{aligned}
\sum_{i\in I_n} u_i &= u_{2n+1} + u_{4n+2} + u_{4n+4} \\
&= \frac{1}{2n+1} - \frac{1}{4n+2} - \frac{1}{4n+4} \\
&= \left(\frac{2}{4n+2} - \frac{1}{4n+2}\right) - \frac{1}{4n+4} \\
&= \frac{1}{4n+2} - \frac{1}{4n+4} \\
&= \frac{1}{2}\left(\frac{1}{2n+1} - \frac{1}{2n+2}\right) \ = \ \frac{1}{2} \sum_{j\in J_n} u_j
\end{aligned}$$

On en déduit :

$$\sum_{n=0}^{+\infty} \left( \sum_{i \in I_n} u_i \right) = \frac{1}{2} \sum_{n=0}^{+\infty} \left( \sum_{j \in J_n} u_j \right)$$

ce qui démontre que les sommes :

$$\sum_{n=0}^{+\infty} \left( \sum_{i \in I_n} u_i \right) \quad \text{et} \quad \sum_{n=0}^{+\infty} \left( \sum_{j \in J_n} u_j \right)$$

ne peuvent toutes les deux coïncider avec $\sum\limits_{k=1}^{+\infty} \dfrac{(-1)^{k+1}}{k}$ (ce devrait être le cas si on pouvait sommer par paquets).

**Théorème 3.** *(théorème de Fubini positif)*

*Soient $I$ et $J$ deux ensembles au plus dénombrables non vides.*

*Soit $(x_{i,j})_{(i,j) \in I \times J}$ une famille d'éléments de $[0, +\infty]$.*

- *Tout d'abord :* $\boxed{\sum\limits_{(i,j) \in I \times J} x_{i,j} = \sum\limits_{i \in I} \left( \sum\limits_{j \in J} x_{i,j} \right)}$
- *En particulier si $(a_i)_{i \in I}$ et $(b_j)_{j \in J}$ sont des familles d'éléments de $[0, +\infty]$ :*

$$\begin{aligned} \sum_{(i,j) \in I \times J} a_i \times b_j &= \sum_{i \in I} \left( \sum_{j \in J} a_i \times b_j \right) \\ &= \sum_{i \in I} \left( a_i \sum_{j \in J} b_j \right) \\ &= \left( \sum_{j \in J} b_j \right) \times \left( \sum_{i \in I} a_i \right) = \left( \sum_{i \in I} a_i \right) \times \left( \sum_{j \in J} b_j \right) \end{aligned}$$

**Remarque**

- Découpage, calcul et majoration de sommes sont autorisés dans $[0, +\infty]$. La finitude de la somme vaut alors preuve de sommabilité.
  Par exemple, comme : $\forall n \in \mathbb{N}^*$, $\dfrac{1}{n(n+1)} \leqslant \dfrac{1}{n^2}$. On peut alors écrire :
  $$\sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \leqslant \sum_{n=1}^{+\infty} \frac{1}{n^2} < +\infty$$
  Ainsi, la famille $\left( \dfrac{1}{n(n+1)} \right)$ est sommable.
- Lorsqu'on travaille dans $[0, +\infty]$, on peut aussi utiliser le théorème de Fubini ou sommer par paquets. On peut ajouter que dans $[0, +\infty]$ :
  - × il y a invariance d'une somme par permutation (on peut sommer dans n'importe quel ordre).
  - × on peut utiliser une propriété de linéarité faible. Plus précisément, si $(x_i)_{i \in I}$ et $(y_i)_{i \in I}$ sont des familles d'éléments de $[0, +\infty]$ alors :
  $$\forall (\lambda, \mu) \in [0, +\infty]^2, \ \sum_{i \in I} (\lambda \ x_i + \mu \ y_i) = \lambda \sum_{i \in I} x_i + \mu \sum_{i \in I} y_i$$

### I.2.d) Famille sommable de nombres complexes

**Définition**

Soit $I$ un ensemble au plus dénombrable.

Soit $(x_i)_{i \in I}$ une famille d'éléments de $\mathbb{C}$.

- La famille $(x_i)_{i \in I}$ est dite **sommable** si $(|x_i|)_{i \in I}$ l'est.
  Autrement dit : $\boxed{\text{La famille } (x_i)_{i \in I} \text{ est sommable} \ \Leftrightarrow \ \sum\limits_{i \in I} |\, x_i \,| < +\infty}$
- En particulier, si $I = \mathbb{N}$ :

| La famille $(x_i)_{i \in \mathbb{N}}$ est sommable | $\Leftrightarrow$ | La série $\sum x_n$ est absolument convergente |
|---|---|---|

**Remarque**

- Dans cette définition, on étend la sommabilité à des familles d'éléments de $\mathbb{C}$ (et plus seulement des familles d'éléments de $[0,+\infty]$). Évidemment, cette définition est valide quand on considère une famille $(x_i)_{i\in I}$ d'éléments dans $\mathbb{R}$ (puisque $\mathbb{R}\subset\mathbb{C}$).
- Dans $[0,+\infty]$, il n'y a que deux modes : les familles sont sommables ou sont de somme $+\infty$. Cette propriété s'exprime de manière pratique par la notation :
$$\sum_{i\in I} x_i < +\infty \quad \text{et} \quad \sum_{i\in I} x_i = +\infty$$
- Dans $\mathbb{C}$, une telle disjonction n'existe pas. On peut ne pas pouvoir sommer pour beaucoup de raisons différentes. Par exemple, la série $\sum (-1)^n$ est divergente et la suite $(S_n)$ de ses sommes partielles n'admet même pas de limite. En effet :
$$\forall n\in\mathbb{N},\ S_{2n}=1 \quad \text{et} \quad S_{2n+1}=0$$
*(si la suite $(S_n)$ admettait une limite finie $\ell$, alors $\ell$ serait la limite commune à $(S_{2n})$ et $(S_{2n+1})$)*
La bonne notation pour parler d'une famille sommable $(x_i)_{i\in I}$ d'éléments de $\mathbb{C}$ est celle qui correspond à la définition, à savoir : $\sum\limits_{i\in I} |\,x_i\,| < +\infty$.
- Notons qu'une famille finie est toujours sommable.

**Théorème 4.** *(sommabilité par domination)*
*Soit $I$ un ensemble au plus dénombrable.*
*Soit $(x_i)_{i\in I}$ une famille d'éléments de $\mathbb{C}$.*
*Soit $(y_i)_{i\in I}$ une famille d'éléments de $[0,+\infty]$.*

$$\boxed{\left.\begin{array}{l} \bullet\ \forall i\in I,\ |\,x_i\,| \leqslant y_i \\ \bullet\ \textit{La famille } (y_i)_{i\in I} \textit{ est sommable}\end{array}\right\} \Rightarrow \begin{array}{c}\textit{La famille } (x_i)_{i\in I}\\ \textit{est sommable}\end{array}}$$

**Remarque**

- Si $(x_i)_{i\in I}$ est une famille d'éléments de $\mathbb{C}$, alors :
$$\forall i\in I,\quad \mathrm{Re}(x_i) \leqslant |\,x_i\,| \quad \text{et} \quad \mathrm{Im}(x_i) \leqslant |\,x_i\,|$$
Ainsi, si $(x_i)_{i\in I}$ est sommable (ce qui signifie $\big(|\,x_i\,|\big)_{i\in I}$ sommable), alors, par domination, les familles $\big(\mathrm{Re}(x_i)\big)_{i\in I}$ et $\big(\mathrm{Im}(x_i)\big)_{i\in I}$ sont sommables.
- Dans ce cas : $\sum\limits_{k\in I} x_k = \sum\limits_{k\in I} \mathrm{Re}(x_k) + i \sum\limits_{k\in I} \mathrm{Im}(x_k)$.

**Théorème 5.** *(linéarité pour les familles sommables)*
*Soit $I$ un ensemble au plus dénombrable.*
*Soient $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ deux familles d'éléments de $\mathbb{C}$.*
*Soit $(\lambda,\mu)\in\mathbb{C}\times\mathbb{C}$.*
*On suppose que $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ sont sommables.*
*Alors $(\lambda\, x_i+\mu\, y_i)_{i\in I}$ est une famille sommable et :*

$$\boxed{\sum_{i\in I} (\lambda\, x_i + \mu\, y_i) = \lambda \sum_{i\in I} x_i + \mu \sum_{i\in I} y_i}$$

**Remarque**

- L'idée de la notion de sommabilité est de simplifier la manipulation de sommes. Comme on l'a vu dans le cas des sommes d'éléments de $[0,+\infty]$, on peut écrire une somme infinie et justifier après coup qu'elle est finie.
- Dans le cas général des familles d'éléments de $\mathbb{C}$, **en cas de sommabilité**, les sommes se manipulent naturellement grâce aux propriétés suivante : croissance, linéarité, sommation par paquets, théorème de Fubini.
- Il ne faut pas surinterpréter ce résultat. Il ne permet pas d'écrire :
$$\begin{aligned}\sum_{n=1}^{+\infty} \big((n+1)-n\big) &= \sum_{n=1}^{+\infty} (n+1) - \sum_{n=1}^{+\infty} n \\ &= \sum_{n=2}^{+\infty} n - \sum_{n=1}^{+\infty} n \\ &= \cancel{\sum_{n=2}^{+\infty}} n - \Big(1 + \cancel{\sum_{n=2}^{+\infty}} n\Big) = -1\end{aligned}$$

Ce calcul est manifestement faux puisque :

$$\sum_{n=1}^{+\infty} 1 \cancel{=} -1$$

L'erreur provient de l'étape de linéarité. Celle-ci n'est autorisée que sous l'hypothèse de sommabilité des familles étudiées. Ce n'est pas le cas ici puisque : $\sum\limits_{n=2}^{+\infty} n = +\infty$ et $\sum\limits_{n=1}^{+\infty} n = +\infty$.

- On retiendra qu'obtenir un résultat fini à l'issue d'une manipulation de sommes `NE` permet `PAS` de justifier après coup de la sommabilité dans le cas général.
  *(plus précisément, obtenir un résultat fini justifie la sommabilité seulement lorsque toutes les manipulations s'opèrent dans* $[0, +\infty]$*)*
- De la même manière, on ne peut écrire :

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

L'étape problématique est, une nouvelle fois, la linéarité.
Celle-ci exige la sommabilité des familles étudiées ce qui n'est pas le cas puisque $\sum\limits_{n=2}^{+\infty} \frac{1}{n-1} = +\infty$ et $\sum\limits_{n=2}^{+\infty} \frac{1}{n} = +\infty$.

Remarquons quand même que le résultat obtenu (la somme initiale vaut 1) est juste ! Cela provient de l'extension de la sommation télescopique dans le cas d'une somme infinie :

$$\sum_{n=1}^{+\infty} (a_{n+1} - a_n) = -a_1 \quad \text{sous l'hypothèse } a_n \underset{n \to +\infty}{\longrightarrow} +\infty$$

## I.3. Une « nouvelle » définition de l'espérance

### I.3.a) Notion d'espérance finie

**Définition**

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé.
Soit $X$ une v.a. discrète sur $(\Omega, \mathscr{A}, \mathbb{P})$.

***1)*** Cas où $X$ est à valeurs dans $[0, +\infty]$

- Dans ce cas, on peut toujours calculer $\mathbb{E}(X)$ à l'aide de la définition suivante :

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\, \mathbb{P}\big(\{X = x\}\big)$$

  Le résultat obtenu est un élément de $[0, +\infty]$.
- On notera alors $\mathbb{E}(X) < +\infty$ pour signifer que le calcul de $E(X)$ a fourni un résultat fini.

***2)*** Cas où $X$ est à valeurs dans $\mathbb{C}$

- On dit que la variable $X$ est d'espérance finie si la famille $\Big(x\, \mathbb{P}\big(\{X = x\}\big)\Big)_{x \in X(\Omega)}$ est sommable.
- Dans ce cas, on appelle espérance de $X$ la quantité finie définie par :

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\, \mathbb{P}\big(\{X = x\}\big)$$

**Remarque**

- Comme déjà mentionné auparavant, le cas $X$ est à valeurs dans $\mathbb{C}$ permet de couvrir aussi le cas où $X$ est à valeurs réelles (mais pas toutes positives).
- Rappelons par ailleurs que si $X$ est à valeurs dans $\mathbb{C}$, la famille $\Big(x\, \mathbb{P}\big(\{X = x\}\big)\Big)_{x \in X(\Omega)}$ est aussi à valeurs dans $\mathbb{C}$. Par définition, cette famille est sommable si $\Big(|x|\, \mathbb{P}\big(\{X = x\}\big)\Big)_{x \in X(\Omega)}$ l'est.
- Cela démontre au passsage :

$$\boxed{\text{La v.a. } X \text{ est d'espérance finie} \quad \Leftrightarrow \quad \mathbb{E}\Big(|X|\Big) < +\infty}$$

#### I.3.b) Espérance finie par domination

**Théorème 6.**

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

*Soit* $X$ *une v.a. discrète à valeurs dans* $\mathbb{C}$.

*Soit* $Y$ *une v.a. discrète à valeurs dans* $\mathbb{R}$.

$$\left.\begin{array}{l} \bullet\ |X| \leqslant Y \\ \bullet\ \mathbb{E}(Y) < +\infty \end{array}\right\} \Rightarrow \begin{array}{l} \textit{La variable } X \textit{ est d'espérance finie} \\ \textit{De plus : } \mathbb{E}(|X|) \leqslant \mathbb{E}(Y) \end{array}$$

### I.4. Propriétés de l'espérance

#### I.4.a) Linéarité

**Théorème 7.**

*Soient* $X$ *et* $Y$ *deux v.a. discrètes à valeurs dans* $\mathbb{C}$.

*On suppose que* $X$ *et* $Y$ *sont d'espérances finies.*

*1)* *Soit* $\lambda \in \mathbb{C}$. *Les v.a.r.* $X+Y$ *et* $\lambda X$ *sont d'espérances finies. De plus :*

$\boxed{\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)}$ *et* $\boxed{\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X)}$

*(linéarité de l'espérance)*

*2)* *Pour tout* $(a,b) \in \mathbb{C}^2$, *la v.a.r.* $aX+b$ *est d'espérance finie. De plus :*

$\boxed{\mathbb{E}(a) = a}$ *et* $\boxed{\mathbb{E}(aX+b) = a\,\mathbb{E}(X) + b}$

**À RETENIR**

On retiendra notamment que la somme de v.a. discrètes d'espérances finies est d'espérance finie.

#### I.4.b) Positivité, croissance

**Théorème 8.**

*Soient* $X$ *et* $Y$ *deux v.a. discrètes à valeurs dans* $\mathbb{R}$.

*1)* $\boxed{X \geqslant 0 \Rightarrow \mathbb{E}(X) \geqslant 0}$

*(positivité de l'espérance)*

*2)* $\boxed{X \geqslant Y \Rightarrow \mathbb{E}(X) \geqslant \mathbb{E}(Y)}$

*(croissance de l'espérance)*

*3)* $\left.\begin{array}{l} X \geqslant 0 \\ \mathbb{E}(X) = 0 \end{array}\right\} \Rightarrow \mathbb{P}(\{X=0\}) = 1$ *(X est presque sûrement nulle)*

### I.5. Théorème de transfert

**Théorème 9.** *(théorème de transfert)*

*Soit* $X$ *une v.a. discrète à valeurs dans* $\mathbb{C}$.

*Soit* $g : X(\Omega) \to \mathbb{C}$ *une application.*

*1)* *La v.a.* $g(X)$ *est d'espérance finie* $\Leftrightarrow$ *La famille* $\Big(g(x)\ \mathbb{P}(\{X=x\})\Big)$ *est sommable*

*2)* *Dans ce cas :*

$$\mathbb{E}\big(g(X)\big) = \sum_{x \in X(\Omega)} g(x)\ \mathbb{P}\big(\{X=x\}\big)$$

**Remarque**

- De manière générale, si l'on veut connaître l'espérance d'une variable aléatoire discrète $Y$, il faut connaître sa loi, c'est-à-dire la donnée de la famille $\Big(\mathbb{P}\big(\{Y=y\}\big)\Big)_{y \in Y(\Omega)}$.
- Le théorème de transfert stipule que si $Y$ est une transformée d'une variable aléatoire $X$ alors son espérance peut être déterminée par la loi de $X$ et plus par celle de $Y$.

# II. Variance d'une variable aléatoire discrète

## II.1. Moment d'ordre 2

### Définition

Soit $X$ une v.a. discrète à valeurs dans $\mathbb{C}$.

*1)* La v.a. $X^2$ est d'espérance finie $\Leftrightarrow$ La famille $\left(x^2\ \mathbb{P}(\{X = x\})\right)$ est sommable

*2)* Dans ce cas :

$$\boxed{\mathbb{E}(X^2) = \sum_{x \in X(\Omega)} x^2\ \mathbb{P}(\{X = x\})}$$

*3)* La v.a. $X^2$ est d'espérance finie $\Rightarrow$ La v.a. $X$ est d'espérance finie

## II.2. Variance d'une v.a. discrète

### Définition

Soit $X$ une v.a. discrète.

Supposons que :

× la v.a. $X$ est d'espérance finie,

× la v.a. $\big(X - \mathbb{E}(X)\big)^2$ est d'espérance finie.

- La v.a. $X$ est de variance finie. De plus, la variance de $X$ est notée $\mathbb{V}(X)$ et définie par :

$$\boxed{\mathbb{V}(X) = \mathbb{E}\Big((X - \mathbb{E}(X))^2\Big)}$$

- Dans le cas où :
  × la variable aléatoire $X$ est à valeurs réelles,
  × la variable aléatoire $X$ est de variance finie,
  on définit l'écart-type de $X$ et on note $\sigma(X)$ la quantité : $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

**Remarque**

- Dans le programme officiel, on considère la variance n'est définie que pour les variables aléatoires **réelles** ce qui permet de ne pas avoir à faire un cas particulier pour définir l'écart-type.
- Remarquons au passage que cette notion d'écart-type est bien définie. Comme $(X - \mathbb{E}(X))^2 \geqslant 0$, alors $\mathbb{V}(X) = \mathbb{E}\Big((X - \mathbb{E}(X))^2\Big) \geqslant 0$.
  *(cf point **3)** du Théorème 8)*
- La variance est une mesure **moyenne** de l'écart existant entre $X$ et $\mathbb{E}(X)$. La variance considère un écart quadratique. Il est alors naturel d'introduire l'écart-type (la racine de la variance) pour gommer le « défaut quadratique » introduit par ce choix d'écart.
- Aurait-on pu choisir d'autres types d'écart ?
  *1)* On peut tout d'abord considérer l'écart simple.
  La mesure moyenne de l'écart simple n'est en réalité d'aucune utilité :

$$\mathbb{E}(X - \mathbb{E}(X)) \ = \ \mathbb{E}(X) - \mathbb{E}(\mathbb{E}(X)) \ = \ \mathbb{E}(X) - \mathbb{E}(X) \ = \ 0$$

  En moyenne, $X$ ne s'écarte pas de sa moyenne.

  *2)* On peut considérer la distance, c'est-à-dire la valeur absolue de l'écart :

$$\mathbb{E}(\ |X - \mathbb{E}(X)|\ )$$

  Cette mesure moyenne d'écart a beaucoup de sens mais possède un inconvénient : elle est plus difficile à manipuler, d'un point de vue algébrique, que la notion de variance.

  On fera attention au passage :

$$\underset{\displaystyle \sigma(X)}{\underset{\shortparallel}{\sqrt{\mathbb{E}((X - \mathbb{E}(X))^2)}}} \quad \cancel{=} \quad \mathbb{E}(\sqrt{(X - \mathbb{E}(X))^2}) \ = \ \mathbb{E}(\ |X - \mathbb{E}(X)|\ )$$

## II.3. Détermination pratique de la variance

**Théorème 10.** *(formule de Kœnig-Huygens)*

*Soit $X$ une v.a. discrète à valeurs dans $\mathbb{C}$.*

| | | |
|---|---|---|
| *La v.a. $X$ est de variance finie* | $\Leftrightarrow$ | *La v.a. $X$ est de moment d'ordre 2 fini* |

*De plus, si $X$ admet une variance, on a :*

$$\boxed{\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2}$$

*Démonstration.*
Remarquons tout d'abord, que dans le cas où $X$ est d'espérance finie :

$$\big(X - \mathbb{E}(X)\big)^2 \;=\; X^2 - 2\,\mathbb{E}(X)\,X + (\mathbb{E}(X))^2 \qquad (*)$$

Et ainsi :

$$X^2 \;=\; (X - \mathbb{E}(X))^2 + 2\,\mathbb{E}(X)\,X - (\mathbb{E}(X))^2 \qquad (**)$$

($\Rightarrow$) Supposons que $X$ est de variance finie.
Par définition de variance, $X$ et $(X - \mathbb{E}(X))^2$ sont donc d'espérances finies.
Or, d'après l'égalité $(**)$, la v.a. $X^2$ s'écrit comme la somme de v.a. d'espérances finies. Elle est donc d'espérance finie.

($\Leftarrow$) Supposons que $X$ est de moment d'ordre 2 fini.
Alors $X$ est d'espérance finie.
L'égalité $(*)$ démontre alors que la v.a. $(X - \mathbb{E}(X))^2$ est d'espérance finie car est la somme de v.a. d'espérances finies.

Supposons maintenant que $X$ est de variance finie et démontrons la formule.

$$\begin{array}{rcll}
\mathbb{V}(X) & = & \mathbb{E}((X - \mathbb{E}(X))^2) & \\
& = & \mathbb{E}(X^2 - 2\,\mathbb{E}(X)\,X + (\mathbb{E}(X))^2) & \\
& = & \mathbb{E}(X^2) - 2\mathbb{E}(\mathbb{E}(X)\,X) + \mathbb{E}((\mathbb{E}(X))^2) & \textit{(par linéarité de l'espérance)} \\
& = & \mathbb{E}(X^2) - 2\mathbb{E}(X)\,\mathbb{E}(X) + (\mathbb{E}(X))^2 & \\
& = & \mathbb{E}(X^2) - (\mathbb{E}(X))^2 &
\end{array}$$

□

## II.4. Propriétés de la variance

**Théorème 11.**

*Soit $X$ et $Y$ deux v.a. discrètes à valeurs dans $\mathbb{C}$.*

*On suppose que $X$ et $Y$ sont de variances finies.*

*L'opérateur variance vérifie les propriétés suivantes.*

***1)*** *Soit $\lambda \in \mathbb{C}$. Les v.a. $X + Y$ et $\lambda X$ admettent une variance. De plus :*

$$\boxed{\mathbb{V}(X + Y) = \mathbb{V}(X) + 2\,(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y))} + \mathbb{V}(Y)$$

$$\textit{et} \quad \boxed{\mathbb{V}(\lambda X) = \lambda^2\;\mathbb{V}(X)}$$

***2)*** *Pour tout $(a, b) \in \mathbb{C}^2$, la v.a. $aX + b$ admet est de variance finie. De plus :*

$$\boxed{\mathbb{V}(a) = 0} \quad \textit{et} \quad \boxed{\mathbb{V}(aX + b) = a^2\,\mathbb{V}(X)}$$

***3)*** *Dans le cas où $X$ est à valeurs réelles :*

$$\boxed{\mathbb{V}(X) \;\geqslant\; 0}$$

*Démonstration.*

***1)*** Soit $\lambda \in \mathbb{C}$. On admet que $X+Y$ et $\lambda X$ sont de variances finies.
D'après la formule de Kœnig-Huygens (et par linéarité de l'espérance) :

$$\begin{aligned}
& \mathbb{V}(X+Y) \\
= \ & \mathbb{E}((X+Y)^2) - (\mathbb{E}(X+Y))^2 \\
= \ & \mathbb{E}(X^2 + 2\,XY + Y^2) - (\mathbb{E}(X) + \mathbb{E}(Y))^2 \\
= \ & \mathbb{E}(X^2) + 2\,\mathbb{E}(XY) + \mathbb{E}(Y^2) - ((\mathbb{E}(X))^2 + 2\,\mathbb{E}(X)\mathbb{E}(Y) + (\mathbb{E}(Y))^2) \\
= \ & \mathbb{E}(X^2) - (\mathbb{E}(X))^2 + \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 + 2\,\mathbb{E}(XY) - 2\,\mathbb{E}(X)\mathbb{E}(Y) \\
= \ & \mathbb{V}(X) + \mathbb{V}(Y) + 2\,(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y))
\end{aligned}$$

Toujours d'après la formule de Kœnig-Huygens et linéarité de l'espérance :

$$\begin{aligned}
\mathbb{V}(\lambda X) & = \mathbb{E}((\lambda X)^2) - (\mathbb{E}(\lambda X))^2 \\
& = \mathbb{E}(\lambda^2\, X^2) - (\lambda\,\mathbb{E}(X))^2 \\
& = \lambda^2\,\mathbb{E}(X^2) - \lambda^2\,(\mathbb{E}(X))^2 \\
& = \lambda^2\,(\mathbb{E}(X^2) - (\mathbb{E}(X))^2) \\
& = \lambda^2\,\mathbb{V}(X)
\end{aligned}$$

***2)*** Soit $(a,b) \in \mathbb{C}^2$. Par hypothèse, $X$ est de variance finie. Elle est donc de moment d'ordre 2 fini et, par conséquent, est d'espérance finie.
Démontrons maintenant que la v.a. $(aX+b) - \mathbb{E}(aX+b)$ est de moment d'ordre 2 fini. Tout d'abord :

$$\begin{aligned}
((aX+b) - \mathbb{E}(aX+b))^2 & = (aX + b - (\mathbb{E}(aX) + \mathbb{E}(b)))^2 \\
& = (a\,(X - \mathbb{E}(X)) + \cancel{b} - \cancel{b})^2 \\
& = a^2\,(X - \mathbb{E}(X))^2
\end{aligned}$$

Or, par définition de la variance de $X$, la v.a. $(X - \mathbb{E}(X))^2$ est d'espérance finie. C'est donc aussi le cas de la v.a. $((aX+b) - \mathbb{E}(aX+b))^2$. En appliquant l'opérateur espérance dans l'égalité précédente, on obtient :

$$\mathbb{V}(aX+b) \ = \ \mathbb{E}((aX+b) - \mathbb{E}(aX+b))^2 \ = \ \mathbb{E}(a^2\,(X - \mathbb{E}(X))^2) \ = \ a^2\,\mathbb{V}(X)$$

Au passage, dans le cas où $a = 0$, on démontre que $\mathbb{V}(b) = 0$.

***3)*** $(X - \mathbb{E}(X))^2 \geqslant 0$ donc, d'après la propriété ***3)*** du Théorème 8 :

$$\mathbb{V}(X) \ = \ \mathbb{E}((X - \mathbb{E}(X))^2) \geqslant 0$$

□

Contrairement à l'espérance, l'opérateur de variance, défini à l'aide d'une élévation au carré, n'est pas linéaire.

**Remarque**

- La variance est une mesure moyenne de l'écart entre la v.a. et sa moyenne. Dans le cas d'une v.a. qui ne varie pas (une v.a. constante c'est-à-dire une quantité $a$), il est logique que la variance soit nulle ($\mathbb{V}(a) = 0$).
- On peut se poser la question de la réciproque de ce résultat.
  Si $X$ est une variable aléatoire réelle :

$$\boxed{\mathbb{V}(X) = 0 \quad \Rightarrow \quad \mathbb{P}\big(\{X = 0\}\big) = 1}$$

### II.5. Variables centrées réduites

**Définition**

Soit $X$ une v.a. discrète à valeurs dans $\mathbb{C}$.

***a)*** Si $X$ est d'espérance finie :

***(i)*** si $\mathbb{E}(X) = 0$, on dit que $X$ est une v.a. centrée.

***(ii)*** si $\mathbb{E}(X) \neq 0$, la v.a. $X - \mathbb{E}(X)$ est appelée v.a. centrée associée à $X$. Notons que cette v.a. est d'espérance finie car somme de v.a. d'espérances finies.
De plus, comme son nom l'indique, cette v.a. est centrée :

$$\mathbb{E}(X - \mathbb{E}(X)) \;=\; \mathbb{E}(X) - \mathbb{E}(\mathbb{E}(X)) \;=\; \mathbb{E}(X) - \mathbb{E}(X) \;=\; 0$$

***b)*** Si $X$ est une variable réelle de variance finie non nulle :

***(i)*** si $\mathbb{V}(X) = 1$, on dit que $X$ est une v.a. réduite.

***(ii)*** si $\mathbb{V}(X) \neq 1$, la v.a. $\dfrac{X}{\sigma(X)}$ (en supposant $\sigma(X) \neq 0$) est appelée v.a. réduite associée à $X$. Notons que cette v.a. est de variance finie car somme de v.a. de variances finies.
Comme son nom l'indique, cette v.a. est réduite :

$$\mathbb{V}\left(\frac{X}{\sigma(X)}\right) \;=\; \left(\frac{1}{\sigma(X)}\right)^2 \mathbb{V}(X) \;=\; \frac{1}{\mathbb{V}(X)}\,\mathbb{V}(X) \;=\; 1$$

***c)*** Si $X$ est de variance finie, la variable $X^* = \dfrac{X - \mathbb{E}(X)}{\sigma(X)}$ est appelée variable centrée réduite associée à $X$.
Remarquons déjà que $X^*$ est d'espérance et de variance finies car s'écrit sous la forme $aX + b$ avec $a = \dfrac{1}{\sigma(X)} \in \mathbb{R}$ et $b = -\dfrac{\mathbb{E}(X)}{\sigma(X)} \in \mathbb{R}$.

- Comme son nom l'indique, $X^*$ est centrée car :

$$\mathbb{E}(X^*) \;=\; a\,\mathbb{E}(X) + b \;=\; \frac{1}{\sigma(X)}\,\mathbb{E}(X) - \frac{\mathbb{E}(X)}{\sigma(X)} \;=\; 0$$

- Comme son nom l'indique, $X^*$ est réduite car :

$$\mathbb{V}(X^*) \;=\; a^2\,\mathbb{V}(X) \;=\; \left(\frac{1}{\sigma(X)}\right)^2 \mathbb{V}(X) \;=\; \frac{\mathbb{V}(X)}{(\sigma(X))^2} \;=\; \frac{\mathbb{V}(X)}{\mathbb{V}(X)} \;=\; 1$$

**Remarque**

À toute v.a. $X$ qui est de variance finie non nulle, on peut associer une v.a. $X^*$ centrée réduite. Cette opération de normalisation peut être utile dans des démonstrations (on démontre une propriété pour une v.a. centrée réduite, on l'applique à $X^*$ et on voit ce qu'on peut en déduire sur $X$).

# III. Le cas particulier des v.a. à valeurs entières

## III.1. Une propriété classique de l'espérance des variables à valeurs entières

**Théorème 12.**

*Soit $X$ une v.a. discrètes à valeurs dans $[\![0,+\infty]\!]$.*

***1. a)*** $\boxed{\forall k \in \mathbb{N}^*,\ \{X > k-1\} = \{X = k\} \cup \{X > k\}}$

***b)*** $\boxed{\forall k \in \mathbb{N}^*,\ \mathbb{P}\big(\{X = k\}\big) = \mathbb{P}\big(\{X > k-1\}\big) - \mathbb{P}\big(\{X > k\}\big)}$

***2.*** $\boxed{\mathbb{E}(X) = \sum\limits_{k=0}^{+\infty} \mathbb{P}\big(\{X > k\}\big)}$

*Démonstration.*

***1.*** Soit $k \in \mathbb{N}^*$.

***a)*** Tout d'abord, comme la variable $X$ est à valeurs entières :

$$\begin{aligned}\{X > k-1\} &= \{X \geqslant k\}\\ &= \{X = k\} \cup \{X > k\}\end{aligned}$$

***b)*** Les événements $\{X = k\}$ et $\{X > k\}$ sont incompatibles. On en déduit :

$$\mathbb{P}(\{X > k-1\}) = \mathbb{P}(\{X = k\}) + \mathbb{P}(\{X > k\})$$

et ainsi : $\mathbb{P}(\{X = k\}) = \mathbb{P}(\{X > k-1\}) - \mathbb{P}(\{X > k\})$.

***2.*** Démonstration en passant par des sommes finies

On ne considère ici que le cas où $X$ est d'espérance finie.

Autrement dit, la somme $\sum\limits_{k=1}^{+\infty} k\ \mathbb{P}\big(\{X = k\}\big)$ est finie ou encore la série $\sum n\ \mathbb{P}\big(\{X = n\}\big)$ est (absolument) convergente. Démontrons qu'il en est de même de la série $\sum \mathbb{P}\big(\{X > n\}\big)$ et que les sommes de ces deux séries sont égales. Soit $N \in \mathbb{N}^*$.

$$\begin{aligned}
& \sum_{k=1}^{N} k\,\mathbb{P}(\{X = k\})\\
= \ & \sum_{k=1}^{N} k \Big( \mathbb{P}(\{X > k-1\}) - \mathbb{P}(\{X > k\})\Big)\\
= \ & \sum_{k=1}^{N} k\,\mathbb{P}(\{X > k-1\}) - \sum_{k=1}^{N} k\,\mathbb{P}(\{X > k\})\\
= \ & \sum_{k=0}^{N-1} (k+1)\ \mathbb{P}(\{X > k\}) - \sum_{k=1}^{N} k\,\mathbb{P}(\{X > k\})\\
= \ & \sum_{k=0}^{N-1} k\ \mathbb{P}(\{X > k\}) + \sum_{k=0}^{N-1} \mathbb{P}(\{X > k\}) - \sum_{k=1}^{N} k\ \mathbb{P}(\{X > k\})\\
= \ & 0 \times \mathbb{P}(\{X > 0\}) + \cancel{\sum_{k=1}^{N-1} k\ \mathbb{P}(\{X > k\})} + \sum_{k=0}^{N-1} \mathbb{P}(\{X > k\})\\
& - \Bigg( \cancel{\sum_{k=1}^{N-1} k\ \mathbb{P}(\{X > k\}} + N\ \mathbb{P}(\{X > N\})\Bigg)\\
= \ & \sum_{k=0}^{N-1} \mathbb{P}(\{X > k\}) - N\ \mathbb{P}(\{X > N\})
\end{aligned}$$

Ce que l'on peut écrire :

$$\sum_{k=0}^{N-1} \mathbb{P}\big(\{X > k\}\big) \ = \ \sum_{k=1}^{N} k\,\mathbb{P}\big(\{X = k\}\big) - N\ \mathbb{P}\big(\{X > N\}\big) \qquad (*)$$

Or, comme $X$ est à valeurs entières, alors :

$$\{X > N\} = \bigcup_{k=N+1}^{+\infty} \{X = k\}$$

et : $N\,\mathbb{P}\big(\{X > N\}\big) \ = \ N \sum\limits_{k=N+1}^{+\infty} \mathbb{P}\big(\{X = k\}\big) \ = \ \sum\limits_{k=N+1}^{+\infty} N\,\mathbb{P}\big(\{X = k\}\big)$.

En remarquant que $k \geqslant N+1$ dès que $N \leqslant k$, on obtient :

$$\begin{aligned} N\,\mathbb{P}\big(\{X > N\}\big) &= \sum_{k=N+1}^{+\infty} N\,\mathbb{P}\big(\{X = k\}\big) \\ &\leqslant \sum_{k=N+1}^{+\infty} k\,\mathbb{P}\big(\{X = k\}\big) \end{aligned}$$

La série $\sum n\mathbb{P}\big(\{X = n\}\big)$ est convergente par hypothèse ($X$ est d'espérance finie) ce qui justifie l'utilisation de son reste d'ordre $N$ en dernière ligne.

Finalement, on a démontré :

$$0 \leqslant N\,\mathbb{P}\big(\{X > N\}\big) \leqslant \sum_{k=N+1}^{+\infty} k\,\mathbb{P}\big(\{X = k\}\big)$$

Or :

× $0 \underset{N\to+\infty}{\longrightarrow} 0$

× $\displaystyle\sum_{k=N+1}^{+\infty} k\,\mathbb{P}\big(\{X = k\}\big) \underset{N\to+\infty}{\longrightarrow} 0$ en tant que reste d'ordre $N$ d'une série convergente.

Ainsi, par théorème d'encadrement : $N\,\mathbb{P}\big(\{X > N\}\big) \underset{N\to+\infty}{\longrightarrow} 0$.

On en conclut par $(*)$ que la série $\sum \mathbb{P}\big(\{X > n\}\big)$ est convergente (car sa somme partielle s'écrit comme somme de deux quantités qui admettent une limite finie en $+\infty$). Enfin :

$$\begin{aligned} &\sum_{k=0}^{+\infty} \mathbb{P}\big(\{X > k\}\big) \\ =\ & \lim_{N\to+\infty} \sum_{k=0}^{N-1} \mathbb{P}\big(\{X > k\}\big) \\ =\ & \lim_{N\to+\infty} \sum_{k=1}^{N} k\,\mathbb{P}\big(\{X = k\}\big) - \lim_{N\to+\infty} \big(N\,\mathbb{P}\big(\{X > N\}\big)\big) \\ =\ & \sum_{k=1}^{+\infty} k\,\mathbb{P}\big(\{X = k\}\big) \end{aligned}$$

Démonstration qui tire partie des calculs de somme dans $[0, +\infty]$

$$\begin{aligned} \sum_{k=1}^{+\infty} \mathbb{P}\big(\{X \geqslant k\}\big) &= \sum_{k=1}^{+\infty} \left( \sum_{j=k}^{+\infty} \mathbb{P}\big(\{X = j\}\big) \right) \quad \textit{(car } X \textit{ est à valeurs entières)} \\ &= \sum_{1\leqslant k\leqslant j\leqslant +\infty} \mathbb{P}\big(\{X = j\}\big) \\ &= \sum_{j=1}^{+\infty} \left( \sum_{k=1}^{j} \mathbb{P}\big(\{X = j\}\big) \right) \\ &= \sum_{j=1}^{+\infty} j\,\mathbb{P}\big(\{X = j\}\big) = \mathbb{E}(X) \end{aligned}$$

**Remarque**

- Dans cette démonstration, on utilise la possibilité d'intervertir des symboles de sommation, possibilité offerte par les calculs de sommes dans $[0, +\infty]$. Chacune des sommes présentes dans ce calcul peut prendre la valeur $+\infty$. En particulier si $\displaystyle\sum_{j=k}^{+\infty} \mathbb{P}\big(\{X = j\}\big) = +\infty$ alors la somme $\displaystyle\sum_{k=1}^{+\infty} \left( \sum_{j=k}^{+\infty} \mathbb{P}\big(\{X = j\}\big) \right)$ vaut elle-même $+\infty$.
- On utilise dans le calcul la possibilité d'intervertir les symboles de sommation. On a vu plus tôt dans le cours l'interversion de sommes lorsqu'il y a indépendance des indices de sommation. Plus précisément :

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

On peut aussi procéder à l'interversion en présence de sommes qui présentent des dépendances d'indices. Rappelons les formules correspondantes.

Dans le cas des sommes finies

$$\boxed{\sum_{1\leqslant i\leqslant j\leqslant n} u_{i,j} = \sum_{i=1}^{n}\left(\sum_{j=i}^{n} u_{i,j}\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{j} u_{i,j}\right)}$$

$$\boxed{\sum_{1\leqslant i<j\leqslant n} u_{i,j} = \sum_{i=1}^{n-1}\left(\sum_{j=i+1}^{n} u_{i,j}\right) = \sum_{j=2}^{n}\left(\sum_{i=1}^{j-1} u_{i,j}\right)}$$

**Comment retenir ces formules ?**

On peut retenir la $1^{\text{ère}}$ formule en considérant l'encadrement $1 \leqslant i \leqslant j \leqslant n$.
Si on souhaite obtenir la formule de sommation suivant les lignes (c'est-à-dire commencer par une somme sur $i$), on peut procéder comme suit :

- × on supprime la variable $j$ de l'encadrement : $1 \leqslant i \leqslant \cdot \leqslant n$
  On doit donc considérer une somme : $\sum\limits_{i=1}^{n}$
- × on considère alors l'encadrement immédiat de $j$ : $i \leqslant j \leqslant n$
  On doit donc considérer une somme : $\sum\limits_{j=i}^{n}$

On retrouve alors la formule : $\sum\limits_{1\leqslant i\leqslant j\leqslant n} u_{i,j} = \sum\limits_{i=1}^{n}\left(\sum\limits_{j=i}^{n} u_{i,j}\right)$

*(on procède de même pour la formule de sommation suivant les colonnes)*

- Dans le cas des sommes infinies d'éléments de $[0, +\infty]$

$$\boxed{\sum_{1\leqslant i\leqslant j\leqslant +\infty} u_{i,j} = \sum_{i=1}^{+\infty}\left(\sum_{j=i}^{+\infty} u_{i,j}\right) = \sum_{j=1}^{+\infty}\left(\sum_{i=1}^{j} u_{i,j}\right)}$$

$$\boxed{\sum_{1\leqslant i<j\leqslant n} u_{i,j} = \sum_{i=1}^{+\infty}\left(\sum_{j=i+1}^{+\infty} u_{i,j}\right) = \sum_{j=2}^{+\infty}\left(\sum_{i=1}^{j-1} u_{i,j}\right)}$$

□

## III.2. Les fonctions génératrices

### III.2.a) Définition

**Définition**

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

- On appelle **fonction génératrice** de $X$ la somme de la série entière $\sum \mathbb{P}\big(\{X = n\}\big)\, t^n$.
- Autrement dit, la fonction génératrice de $X$ est la fonction :

$$G_X : t \mapsto \mathbb{E}\left(t^X\right) = \sum_{k=0}^{+\infty} \mathbb{P}\big(\{X = k\}\big)\, t^k$$

### III.2.b) Les fonctions génératrices des lois usuelles finies

Dans le cas où $X$ est une v.a. à valeurs entières qui suit une loi discrète finie, la série $\sum \mathbb{P}\big(\{X=n\}\big)\, t^n$ est de rayon de convergence $R=+\infty$.

- Cas de la loi uniforme
  - Si $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$

$$\forall t \in ]-\infty,+\infty[,\ G_X(t) \ = \ \begin{cases} \dfrac{1}{n}\ \dfrac{t-t^{n+1}}{1-t} & si\ t \neq 1 \\ 1 & si\ t=1 \end{cases}$$

  - Si $X \sim \mathcal{U}(\llbracket a, b \rrbracket)$

$$\forall t \in ]-\infty,+\infty[,\ G_X(t) \ = \ \begin{cases} \dfrac{1}{b-a+1}\ \dfrac{t^a-t^{b+1}}{1-t} & si\ t \neq 1 \\ 1 & si\ t=1 \end{cases}$$

- Cas de la loi de Bernoulli : $X \sim \mathcal{B}(p)$

$$\forall t \in ]-\infty,+\infty[,\ G_X(t) \ = \ (1-p)+p\,t$$

- Cas de la loi binomiale : $X \sim \mathcal{B}(n,p)$

$$\forall t \in ]-\infty,+\infty[,\ G_X(t) \ = \ \big((1-p)+p\,t\big)^n$$

### III.2.c) Les fonctions génératrices des lois usuelles infinies

- Dans le cas où $X$ est une v.a. à valeurs entières qui suit une loi discrète infinie, la série $\sum \mathbb{P}\big(\{X=n\}\big)\, t^n$ est de rayon de convergence $R \geqslant 1$ (résultat à venir).
- Le rayon de convergence s'obtient par applicaiton de la règle de d'Alembert.
- Cas de la loi géométrique : $X \sim \mathcal{G}(p)$

$$\forall t \in ]-\frac{1}{1-p}, \frac{1}{1-p}[,\ G_X(t) \ = \ \frac{p\,t}{1-(1-p)\,t}$$

- Cas de la loi de Poisson : $X \sim \mathcal{P}(\lambda)$

$$\forall t \in ]-\infty,+\infty[,\ G_X(t) \ = \ e^{\lambda\,(t-1)}$$

### III.2.d) La fonction génératrice caractérise la loi

**Théorème 13.**

*Soient $X$ et $Y$ deux variables aléatoires discrètes à valeurs dans $\mathbb{N}$.*

***1.*** ***a)*** *La série entière $\sum \mathbb{P}\big(\{X=n\}\big)\, t^n$ a un rayon de convergence $R \geqslant 1$.*

***b)*** *En particulier, $G_X$, somme de la série entière $\sum \mathbb{P}\big(\{X=n\}\big)\, t^n$ est développable en série entière sur (au moins) $]-1,1[$ et donc de classe $\mathscr{C}^\infty$ sur sur (au moins) $]-1,1[$.*

***2.*** $\boxed{\forall n \in \mathbb{N},\ \mathbb{P}\big(\{X=n\}\big) \ = \ \dfrac{1}{n!}\ {G_X}^{(n)}(0)}$

***3.*** $\boxed{G_X \text{ et } G_Y \text{ coïncident sur } ]-1,1[ \quad \Leftrightarrow \quad \text{Les v.a. } X \text{ et } Y \text{ ont même loi}}$

*Démonstration.*
Dans la suite, pour tout $n \in \mathbb{N}^*$, on note :

$$f_n \; : \; \begin{array}{ccc} t & \mapsto & \mathbb{P}\big(\{X = n\}\big)\ t^n \\ \mathbb{R} & \to & \mathbb{R} \end{array}$$

***1.*** La suite $\big(|\,\mathbb{P}\big(\{X = n\}\big)\,|\ 1^n\big)$ est bornée puisque pour tout $n \in \mathbb{N}$ :

$$0 \leqslant \mathbb{P}\big(\{X = n\}\big) \leqslant 1$$

Ainsi, le rayon de convergence $R$ de la série entière $\sum \mathbb{P}\big(\{X = n\}\big)\ t^n$ vérifie $R \geqslant 1$.

***2.*** Ce résultat provient du fait que comme $G_X$ est développable en série entière sur $]-1,1[$ alors $G_X$ coïncide avec la somme de sa série de Taylor sur $]-1,1[$.

***3.*** Les fonctions $G_X$ et $G_Y$ sont développables en série entière sur $]-1,1[$ et sont donc égales à la même somme série de Taylor sur $]-1,1[$. Ainsi, elles ont les mêmes coefficients. □

### III.2.e) Continuité de la fonction génératrice d'une v.a. discrète

**Théorème 14.**

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

***1. a)*** *La série entière $\sum \mathbb{P}\big(\{X = n\}\big)\ t^n$ converge normalement (au moins) sur le segment $[-1,1]$.*

***b)*** *La fonction $G_X$ est continue (au moins) sur $[-1,1]$ et $G_X = 1$.*

*Démonstration.*

***1. a)*** Soit $n \in \mathbb{N}$.

$$\begin{array}{rcll} \forall t \in [-1,1],\ |\,\mathbb{P}\big(\{X = n\}\big)\ t^n\,| & = & \mathbb{P}\big(\{X = n\}\big)\ |\,t\,|^n & \\ & \leqslant & \mathbb{P}\big(\{X = n\}\big) & \textit{(par croissance de la fonction } t \mapsto |t|^n \textit{ sur } \mathbb{R}\textit{)} \end{array}$$

On en conclut :

$$\forall n \in \mathbb{N},\ 0 \leqslant \|f_n\|_{\infty,[-1,1]} \leqslant \mathbb{P}\big(\{X = n\}\big)$$

Or la série $\sum \mathbb{P}\big(\{X = n\}\big)$ est convergente. En effet :

$$\sum_{n=0}^{+\infty} \mathbb{P}\big(\{X = n\}\big) = \sum_{x \in X(\Omega)} \mathbb{P}\big(\{X = x\}\big) = 1$$

Par théorème de comparaison des séries à termes positifs, la série numérique $\sum \|f_n\|_{\infty,[-1,1]}$ est convergente.

***b)*** Remarquons :

× $\forall n \in \mathbb{N}$, la fonction $f_n$ est continue sur (au moins) $[-1,1]$ car polynomiale.

× $\sum f_n$ converge normalement (donc uniformément) sur $[-1,1]$.

On en conclut que la somme $G_X = \sum_{n=0}^{+\infty} f_n$ est continue sur $[-1,1]$. □

### III.2.f) Lien entre fonction génératrice et calcul de moments

**Théorème 15.**

*Soit $X$ une v.a. à valeurs dans $\mathbb{N}$ de fonction génératrice $G_X$.*

***1)*** $\boxed{X \text{ est d'espérance finie } \Leftrightarrow G_X \text{ dérivable en } 1}$

*Dans ce cas :* $\boxed{\mathbb{E}(X) = G'_X(1)}$.

***2)*** $\boxed{X \text{ est de variance finie } \Leftrightarrow G_X \text{ deux fois dérivable en } 1}$

*Dans ce cas :*

$$\boxed{\mathbb{E}\big(X\,(X-1)\big) = G''_X(1)} \quad \text{et} \quad \boxed{\mathbb{V}(X) = G''_X(1) + G'_X(1) - \big(G'_X(1)\big)^2}$$

*Démonstration.*
Pour tout $n \in \mathbb{N}$, on note : $p_n = \mathbb{P}\big(\{X = n\}\big)$.

***1)*** On procède par double implication.

($\Rightarrow$) Supposons que $X$ est d'espérance finie.
Alors la série $\sum n\,p_n$ est (absolument) convergente. On en déduit que la série $\sum n\,p_n\,t^{n-1}$ converge normalement sur $[0,1]$. La fonction $G_X$ est donc de classe $\mathscr{C}^1$ sur $[0,1]$ de dérivée, pour tout $t \in [0,1]$ :

$$G'_X(t) \;=\; \sum_{n=0}^{+\infty} n\,p_n\,t^{n-1}$$

En particulier :

$$G'_X(1) \;=\; \sum_{n=0}^{+\infty} n\,p_n \;=\; \sum_{n=0}^{+\infty} n\,\mathbb{P}\big(\{X = n\}\big) \;=\; \mathbb{E}(X)$$

($\Leftarrow$) Supposons $G_X$ dérivable en 1.
La v.a. $X$ est d'espérance finie si et seulement si la série $\sum n\,p_n$ converge absolument. Cela revient à démontrer sa convergence car c'est une série à termes positifs.

- Tout d'abord, on remarque : $\forall n \in \mathbb{N}$, $n\,p_n \geqslant 0$.
  On en déduit que la suite $\left(\sum\limits_{n=0}^{N} n\,p_n\right)_{N\in\mathbb{N}}$ est croissante.
  Il suffit donc de démontrer qu'elle est majorée pour en déduire sa convergence.
- Soit $N \in \mathbb{N}^*$. On remarque :

$$\lim_{t\to 1^-} \sum_{n=0}^{N} n\,p_n\,t^{n-1} \;=\; \sum_{n=0}^{N} n\,p_n$$

  On cherche donc à étudier $\sum n\,p_n\,t^{n-1}$.
- D'après les résultats du cours, la série dérivée $\sum n\,p_n\,t^{n-1}$ converge normalement sur tout segment inclus dans $[0,1[$.
  Ainsi, pour tout $t \in [0,1[$ :

$$G'_X(t) \;=\; \sum_{n=0}^{+\infty} n\,p_n\,t^{n-1}$$

  D'où, comme $G_X$ est dérivable en 1 :

$$\sum_{n=0}^{+\infty} n\,p_n\,t^{n-1} \underset{t\to 1^-}{\longrightarrow} G'_X(1)$$

- On obtient, pour tout $t \in [0,1[$, pour tout $N \in \mathbb{N}^*$ :

$$\begin{array}{ccc}
\displaystyle\sum_{n=0}^{N} n\,p_n\,t^{n-1} & \leqslant & \displaystyle\sum_{n=0}^{+\infty} n\,p_n\,t^{n-1} \\
{\scriptstyle t \to 1^-}\Big\downarrow & & {\scriptstyle t \to 1^-}\Big\downarrow \\
\displaystyle\sum_{n=0}^{N} n\,p_n & & G'_X(1)
\end{array}$$

Ainsi, la suite $\left(\displaystyle\sum_{n=0}^{N} n\,p_n\right)$ est :

× croissante,

× majorée par $G'_X(1)$.

Elle est donc convergente.

On en déduit que $X$ est d'espérance finie.

***2)*** Avec un raisonnement similaire à celui de ***1)***, on peut démontrer :

$$G_X \text{ est deux fois dérivable en } 1 \;\Leftrightarrow\; \sum n\,(n-1)\,p_n \text{ est convergente}$$

Or :

$$\sum n\,(n-1)\,p_n \text{ est convergente} \;\Leftrightarrow\; X\,(X-1) \text{ est d'espérance finie}$$

On obtient bien :

$$G_X \text{ est deux fois dérivable en } 1 \;\Leftrightarrow\; X\,(X-1) \text{ est d'espérance finie}$$

Dans ce cas :

$$G''_X(1) \;=\; \sum_{n=0}^{+\infty} n\,(n-1)\,p_n\,1^{n-2} \;=\; \sum_{n=0}^{+\infty} n\,(n-1)\,p_n \;=\; \mathbb{E}\big(X\,(X-1)\big)$$

□

**Exercice 1**

***1.*** Lorsqu'elle existe, exprimer la variance de $X$ en fonction de $G'_X(1)$ et $G''_X(1)$.

***2.*** Reprendre les exemples concernant les lois classiques.

### III.2.g) Conséquence : calcul d'espérance et de variance pour les lois usuelles

- Les fonctions génératrices étant connues, on peut retrouver les valeurs d'espérance et de variance de chacune de ces lois par un calcul de dérivée premières / seconde et à l'aide des formules présentées dans le théorème précédent.
- On place en page suivante un formulaire concernant les lois usuelles.

| | Notation et paramètres | Loi de $X$ | Fonction génératrice | $\mathbb{E}(X)$ | $\mathbb{V}(X)$ |
|---|---|---|---|---|---|
| **Loi uniforme** | $\mathcal{U}([\![1,n]\!])$ <br> $n \in \mathbb{N}^*$ | $X(\Omega) = [\![1,n]\!]$ <br> $\forall k \in [\![1,n]\!],\ \mathbb{P}\big(\{X=k\}\big) = \dfrac{1}{n}$ | $\forall t \in ]-\infty,+\infty[$, <br> $G_X(t) = \begin{cases} \dfrac{1}{n}\,\dfrac{t-t^{n+1}}{1-t} & si\ t \neq 1 \\ 1 & si\ t = 1 \end{cases}$ | $\dfrac{n+1}{2}$ | $\dfrac{(n-1)(n+1)}{12}$ |
| | $\mathcal{U}([\![a,b]\!])$ <br> $(a,b) \in \mathbb{N}^2$ <br> $b \geqslant a$ | $X(\Omega) = [\![a,b]\!]$ <br> $\forall k \in [\![a,b]\!],\ \mathbb{P}\big(\{X=k\}\big) = \dfrac{1}{b-a+1}$ | $\forall t \in ]-\infty,+\infty[$, <br> $G_X(t) = \begin{cases} \dfrac{1}{b-a+1}\,\dfrac{t^a-t^{b+1}}{1-t} & si\ t \neq 1 \\ 1 & si\ t = 1 \end{cases}$ | $\dfrac{a+b}{2}$ | $\dfrac{(b-a)(b-a+2)}{12}$ |
| **Loi de Bernoulli** | $\mathcal{B}(1,p)$ <br> $p \in ]0,1[$ | $X(\Omega) = \{0,1\}$ <br> $\mathbb{P}\big(\{X=1\}\big) = p$ et $\mathbb{P}\big(\{X=0\}\big) = 1-p$ | $\forall t \in ]-\infty,+\infty[,\ G_X(t) = (1-p)+p\,t$ | $p$ | $p\,q$ |
| **Loi binomiale** | $\mathcal{B}(n,p)$ <br> $n \in \mathbb{N}^*$, <br> $p \in ]0,1[$ | $X(\Omega) = [\![0,n]\!]$ <br> $\forall k \in [\![0,n]\!],\ \mathbb{P}\big(\{X=k\}\big) = \dbinom{n}{k} p^k\, q^{n-k}$ | $\forall t \in ]-\infty,+\infty[,\ G_X(t) = \big((1-p)+p\,t\big)^n$ | $n\,p$ | $n\,p\,q$ |
| **Loi géométrique** | $\mathcal{G}(p)$ <br> $p \in ]0,1[$ | $X(\Omega) = \mathbb{N}^*$ <br> $\forall k \in \mathbb{N}^*,\ \mathbb{P}\big(\{X=k\}\big) = p\,q^{k-1}$ | $\forall t \in ]-\frac{1}{1-p},\frac{1}{1-p}[,\ G_X(t) = \dfrac{p\,t}{1-(1-p)\,t}$ | $\dfrac{1}{p}$ | $\dfrac{q}{p^2}$ |
| **Loi de Poisson** | $\mathcal{P}(\lambda)$ <br> $\lambda > 0$ | $X(\Omega) = \mathbb{N}$ <br> $\forall k \in \mathbb{N},\ \mathbb{P}\big(\{X=k\}\big) = \mathrm{e}^{-\lambda}\,\dfrac{\lambda^k}{k!}$ | $\forall t \in ]-\infty,+\infty[,\ G_X(t) = \mathrm{e}^{\lambda\,(t-1)}$ | $\lambda$ | $\lambda$ |

# IV. Loi faible des grands nombres

## IV.1. L'inégalité de Markov

**Théorème 16** (inégalité de Markov)**.**

*Soit $X$ une v.a. qui est d'espérance finie.*

*On suppose de plus : $X(\Omega) \subset [0, +\infty[$.*
*(autrement dit, $X$ est à valeurs positives)*

$$\boxed{\forall a > 0,\ \mathbb{P}\big(\{X \geqslant a\}\big) \leqslant \frac{\mathbb{E}(X)}{a}}$$

*Démonstration.*

Démonstration adaptée au programme

Soit $a > 0$.

- Comme $X$ est une variable discrète : $\{X \geqslant a\} = \bigcup\limits_{\substack{x \in X(\Omega) \\ x \geqslant a}} \{X = x\}$.

  On en déduit :

$$\mathbb{P}\big(\{X \geqslant a\}\big) = \mathbb{P}\Big(\bigcup_{\substack{x \in X(\Omega) \\ x \geqslant a}} \{X = x\}\Big) = \sum_{\substack{x \in X(\Omega) \\ x \geqslant a}} \mathbb{P}\big(\{X = x\}\big)$$

- Comme $X$ est une variable positive, $\mathbb{E}(X) \in [0, +\infty]$ et :

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{x \in X(\Omega)} x\, \mathbb{P}\big(\{X = x\}\big) \\
&= \sum_{\substack{x \in X(\Omega) \\ x \geqslant a}} x\, \mathbb{P}\big(\{X = x\}\big) + \sum_{\substack{x \in X(\Omega) \\ x < a}} x\, \mathbb{P}\big(\{X = x\}\big) \\
&\geqslant \sum_{\substack{x \in X(\Omega) \\ x \geqslant a}} x\, \mathbb{P}\big(\{X = x\}\big) \quad \textit{(car } X \textit{ est à valeurs positives)}
\end{aligned}$$

- Finalement :

$$\sum_{\substack{x \in X(\Omega) \\ x \geqslant a}} x\, \mathbb{P}\big(\{X = x\}\big) \geqslant \sum_{\substack{x \in X(\Omega) \\ x \geqslant a}} a\, \mathbb{P}\big(\{X = x\}\big) = a\, \mathbb{P}\big(\{X \geqslant a\}\big)$$

Démonstration adaptée à toute variable aléatoire `(CULTURE)`

Soit $a > 0$. Considérons la v.a. $Y$ définie par :

$$Y : \begin{array}{ccl} \Omega & \to & \mathbb{R} \\ \omega & \mapsto & \begin{cases} a & \text{si } X(\omega) \geqslant a \\ 0 & \text{si } X(\omega) < a \end{cases} \end{array}$$

Listons quelques propriétés de $Y$ :

- $Y(\Omega) = \{0, a\}$ ($Y$ est donc une v.a. finie).
- $\{Y = a\} = \{X \geqslant a\}$ (puisque : $\forall \omega \in \Omega,\ Y(\omega) = a \Leftrightarrow X(\omega) \geqslant a$).
- $\{Y = 0\} = \{X < a\}$ (puisque : $\forall \omega \in \Omega,\ Y(\omega) = 0 \Leftrightarrow X(\omega) < a$).

*1)* Comme $Y$ est finie, $Y$ est d'espérance finie. De plus :

$$\mathbb{E}(Y) = 0 \times \underbrace{\mathbb{P}(\{Y = 0\})}_{= \mathbb{P}(\{X < a\})} + a \times \underbrace{\mathbb{P}(\{Y = a\})}_{= \mathbb{P}(\{X \geqslant a\})} = a \times \mathbb{P}(\{X \geqslant a\})$$

*2)* On remarque aussi : $Y \leqslant X$ (c'est-à-dire : $\forall \omega \in \Omega$, $Y(\omega) \leqslant X(\omega)$).
Démontrons-le. Soit $\omega \in \Omega$.
Deux cas se présentent.

× si $X(\omega) \geqslant a$ alors $Y(\omega) = a \leqslant X(\omega)$.

× si $X(\omega) < a$ alors $Y(\omega) = 0 \leqslant X(\omega)$.

*3)* Par croissance de l'espérance, on déduit des deux points précédents :

$$\underbrace{\mathbb{E}(Y)}_{= a\, \mathbb{P}(\{X \geqslant a\})} \leqslant \mathbb{E}(X)$$

On conclut en divisant de part et d'autre par $a > 0$. □

**Remarque**

- Notons que si $a \leqslant \mathbb{E}(X)$, alors l'inégalité est évidente (et donc peu intéressante). En effet, dans ce cas, on a directement : $\mathbb{P}(\{X \geqslant a\}) \leqslant 1 \leqslant \dfrac{\mathbb{E}(X)}{a}$.
Dans le cas où $a$ est « grand » par rapport à $\mathbb{E}(X)$, on obtient un résultat de bon sens : la probabilité que $X$ prenne de grandes valeurs est d'autant plus faible que celles-ci excèdent $\mathbb{E}(X)$.
- On utilise dans cette démonstration la notation $\mathbb{1}_A$. Si $A$ est un événement, $\mathbb{1}_A$ est la variable indicatrice de l'événement $A$. Plus précisément, $\mathbb{1}_A$ est définie par :
$$\begin{array}{rccl} \mathbb{1}_A & : & \Omega & \to & \mathbb{R} \\ & & \omega & \mapsto & \left\{ \begin{array}{ll} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{array} \right. \end{array}$$
Avec cette notation, on peut écrire : $Y = a \cdot \mathbb{1}_{\{X \geqslant a\}}$.
- Dans cette démonstration, on ne suppose pas que $X$ est d'espérance finie. Les v.a. en jeu étant à valeurs positives, cette précaution n'a pas lieu d'être. Dans le cas où $X$ est d'espérance infinie, l'énoncé n'a aucun intérêt : la valeur $\mathbb{P}\big(\{X \geqslant a\}\big)$ est évidemment finie puisque majorée par 1.

**Théorème 17.**

*Soit $m \in \mathbb{N}$.*

*Soit $X$ une v.a. possédant un moment d'ordre $m$.*

$$\boxed{\forall a > 0,\ \mathbb{P}\big(\{\ |X| \geqslant a\ \}\big) \ \leqslant\ \frac{\mathbb{E}\big(|X|^m\big)}{a^m}}$$

*Démonstration.*
On utilise l'inégalité de Markov à la v.a. $|X|^m$ et au réel $a^m > 0$ :

$$\begin{array}{ccc} \mathbb{P}(\{|X|^m \geqslant a^m\}) & \leqslant & \dfrac{\mathbb{E}(|X|^m)}{a^m} \\ \shortparallel & & \\ \mathbb{P}(\{|X| \geqslant a\}) & & \end{array}$$

□

## IV.2. L'inégalité de Bienaymé-Tchebychev

**Théorème 18** (Inégalité de Bienaymé-Tchebychev)**.**
*Soit $X$ une v.a. (discrète) à valeurs réelles.*
*On suppose que $X$ est de variance finie.*
*Alors :*

$$\boxed{\forall \varepsilon > 0,\ \mathbb{P}\big(\{\ |X - \mathbb{E}(X)| \geqslant \varepsilon\ \}\big) \ \leqslant\ \frac{\mathbb{V}(X)}{\varepsilon^2}}$$

*Démonstration.*
Soit $\varepsilon > 0$. Considérons la v.a. $Y = \big(X - \mathbb{E}(X)\big)^2$.
On est dans le cadre d'application de l'inégalité de Markov :
× $Y = \big(X - \mathbb{E}(X)\big)^2$ est à valeurs positives.
× la v.a. $Y$ est d'espérance finie car $X$ est de variance finie.
En appliquant l'inégalité de Markov à la v.a. $Y$, avec $a = \varepsilon^2$ :

$$\begin{array}{ccccc} \mathbb{P}\big(\{\ Y \geqslant \varepsilon^2\ \}\big) & = & \mathbb{P}\Big(\Big\{\ \big(X - \mathbb{E}(X)\big)^2 \geqslant \varepsilon^2\ \Big\}\Big) & \leqslant & \dfrac{\mathbb{E}(Y)}{\varepsilon^2} \\ & & \shortparallel & & \shortparallel \\ & & \mathbb{P}\big(\{\ \big|X - \mathbb{E}(X)\big| \geqslant \varepsilon\ \}\big) & & \dfrac{\mathbb{E}\big((X - \mathbb{E}(X))^2\big)}{\varepsilon^2} = \dfrac{\mathbb{V}(X)}{\varepsilon^2} \end{array}$$

□

**Remarque**

- Ce résultat est une inégalité de concentration. Grâce à ce type d'inégalités, on peut mesurer / contrôler la probabilité qu'un phénomène aléatoire puisse s'écarter (on préférera le terme dévier) de la moyenne *i.e.* d'un comportement standard.
- Évidemment, il est rare de constater de grandes déviations. La question que se pose une compagnie d'assurances est de savoir si elle est prête à parier sur le fait qu'une grande déviation (*i.e.* un événement qui s'écarte fortement de la norme) ne se produira pas. Il est alors primordial dans ce cas d'obtenir des inégalités de concentration avec un majorant le plus précis possible.

- L'inégalité de Bienaymé-Tchebychev est peu précise. L'une des raisons est qu'elle s'applique à toute v.a. . On peut évidemment obtenir de meilleures majorations en tirant parti des propriétés (*i.e.* de la loi) des v.a. étudiées.

## IV.3. La loi faible des grands nombres

### IV.3.a) Énoncé et convergence en probabilité

**Théorème 19** (Loi faible des grands nombres)**.**

*Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. (discrètes) à valeurs réelles.*

*On suppose que les v.a. de la suite $(X_n)_{n\in\mathbb{N}^*}$ :*

× *sont indépendantes,*

× *admettent toutes la même espérance $m$,*

× *admettent toutes la même variance $\sigma^2$.*

*On pose, pour tout $n \in \mathbb{N}^*$, $\overline{X_n} = \dfrac{1}{n} \sum\limits_{k=1}^{n} X_k$ (moyenne empirique).*

$$\boxed{\forall \varepsilon > 0,\ \lim_{n\to+\infty} \mathbb{P}\left(\left\{\ \left|\overline{X_n} - m\right| \geqslant \varepsilon\ \right\}\right) = 0}$$

*Démonstration.*

- Rappelons tout d'abord l'inégalité de Bienaymé-Tchebychev.
  Pour toute v.a. $Y$ admettant une variance $\mathbb{V}(Y)$, on a :

$$\forall \lambda > 0,\ \mathbb{P}\big(\{|Y - \mathbb{E}(Y)| \geqslant \lambda\}\big) \ \leqslant\ \frac{\mathbb{V}(Y)}{\lambda^2}$$

- Soit $n \in \mathbb{N}^*$. Considérons la v.a. $Y = \overline{X_n} = \dfrac{1}{n} \sum\limits_{k=1}^{n} X_k$.

  × La v.a. $\overline{X_n}$ est d'espérance finie (resp. variance) car elle est la CL de v.a. qui admettent une espérance (resp. variance).

  × De plus :

$$\begin{aligned}
\mathbb{E}(\overline{X_n}) &= \mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) \\
&= \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}(X_k) && \textit{(par linéarité de l'espérance)} \\
&= \frac{1}{n}\sum_{k=1}^{n} m \ = \ \frac{1}{\cancel{n}}\, \cancel{n} \times m \ = \ m
\end{aligned}$$

  × Et :

$$\begin{aligned}
\mathbb{V}(\overline{X_n}) &= \mathbb{V}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) \\
&= \frac{1}{n^2}\, \mathbb{V}\left(\sum_{k=1}^{n} X_k\right) \\
&= \frac{1}{n^2}\sum_{k=1}^{n} \mathbb{V}(X_k) && \textit{(car les v.a. de la suite } (X_n) \textit{ sont indépendantes)} \\
&= \frac{1}{n^2}\sum_{k=1}^{n} \sigma^2 \ = \ \frac{1}{n^{\cancel{2}}}\, \cancel{n} \times \sigma^2 \ = \ \frac{1}{n}\,\sigma^2
\end{aligned}$$

- Soit $\varepsilon > 0$. D'après l'inégalité de Bienaymé-Tchebychev :

$$\mathbb{P}\left(\left\{\left|\overline{X_n} - m\right| \geqslant \varepsilon\right\}\right) \ \leqslant\ \frac{\sigma^2}{\varepsilon^2}\,\frac{1}{n} \ \underset{n\to+\infty}{\longrightarrow}\ 0$$

□

**Remarque**

- Il est fréquent de considérer (notamment pour les simulations informatiques) une suite $(X_n)_{n\in\mathbb{N}^*}$ de v.a. :
  - × indépendantes,
  - × **de même loi**,
  - × d'espérances finies et de variance finie.

  Ces hypothèses sont plus strictes que celles énoncés par la LfGN.
  On peut donc bien évidemment utiliser la LfGN dans ce cadre.
  *(des variables indépendantes et de même loi sont dites indépendantes et identiquement distribuées)*
- On dit qu'une suite de v.a. $(X_n)_{n\in\mathbb{N}^*}$ **converge en probabilté** vers une v.a. $X$ si :
  $$\forall \varepsilon > 0, \ \lim_{n\to+\infty} \mathbb{P}\big( \big\{ \, \big|X_n - X\big| \geqslant \varepsilon \, \big\} \big) = 0$$
  Dans ce cas, on note : $X_n \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} X$.
- Via cette définition, on comprend que la LfGN établit un résultat de convergence d'une suite de v.a. : la suite de v.a. $\big(\overline{X_n}\big)_{n\in\mathbb{N}^*}$ **converge en probabilité** vers la v.a. certaine égale à $m$.
- Il existe une loi forte des grands nombres. Sous des hypothèses plus exigeantes, on établit que la suite de v.a. $\big(\overline{X_n}\big)_{n\in\mathbb{N}^*}$ converge vers la v.a. certaine égale à $m$. On retrouve le résultat précédent, à ceci près que le mode de convergence (appelé convergence presque sûre) est un mode plus exigeant que la convergence en probabilité (la convergence presque sûre implique la convergence en probabilité).

### IV.3.b) Intuition derrière cet énoncé

**Exemple** *(classique)*

- On considère un dé à 6 faces (truqué ou non).
  - Expérience : on effectue un lancer de ce dé et on note $X$ la v.a. égale au résultat obtenu lors de ce lancer.
  - But : on souhaite, à l'aide d'expériences, savoir si le dé peut être considéré comme équilibré. Autrement dit, on souhaite connaître la loi de $X$.
    Plus particulièrement, on souhaiterait connaître $\mathbb{P}\big(\{X=6\}\big)$.
- Pour obtenir une approximation de cette probabilité $\mathbb{P}\big(\{X=6\}\big)$, il est naturel de procéder comme suit :
  × on effectue un grand nombre $N$ de tirages.
  × on compte le nombre d'apparitions de la face 6 au cours de ce $N$ tirages.
  On observe alors la fréquence d'apparition de cette face :

$$\begin{array}{c}\text{fréquence d'apparition}\\ \textbf{observée}\text{ de la face 6}\end{array} = \frac{\text{nombre de 6 obtenus en } N \text{ tirages}}{N}$$

  Il est naturel de penser que cette fréquence observée est une bonne approximation de $\mathbb{P}\big(\{X=6\}\big)$. La LfGN n'est rien d'autre qu'une validation mathématique de ce procédé. Faisons le lien plus formellement.
- Pour tout $i \in \mathbb{N}^*$, on note $X_i$ la v.a. qui donne le résultat du $i^{\text{ème}}$ lancer.
  On note par ailleurs $Z_i$ la v.a. définie par :

$$Z_i : \omega \mapsto \begin{cases} 1 & \text{si } X_i(\omega) = 6 \\ 0 & \text{sinon} \end{cases}$$

  Autrement dit : $Z_i = \mathbb{1}_{\{X_i=6\}}$. On a alors :

$$Z_i \sim \mathcal{B}(r) \qquad \text{où} \qquad r = \mathbb{P}\big(\{X_i=6\}\big)$$

  En particulier : $\mathbb{E}(Z_i) = \mathbb{E}\big(\mathbb{1}_{\{X_i=6\}}\big) = \mathbb{P}\big(\{X_i=6\}\big)$.
- La suite $(Z_n)_{n\in\mathbb{N}^*}$ est une suite de variables :
  × indépendantes (le résultat d'un lancer ne dépend pas des précédents),
  × de même espérance $p$,
  × de même variance $\sigma^2 = p\,(1-p)$.
  On en déduit, par la loi faible des grands nombres que, pour tout $\varepsilon > 0$ :

$$\lim_{n\to+\infty} \mathbb{P}\left( \left| \frac{Z_1+Z_2+\cdots+Z_n}{n} - p \right| \geqslant \varepsilon \right) = 0$$

  ce que l'on préférera noter : $\overline{Z_n} \xrightarrow[n\to+\infty]{\mathbb{P}} p$.
- La v.a. $\overline{Z_n}$ donne la fréquence d'apparition de la face 6 au cours des $n$ premiers lancers. La loi faible des grands nombres affirme que, plus $n$ est grand, plus la fréquence d'apparition du 6 au cours des $n$ lancers est proche de la fréquence théorique $(= \mathbb{P}\big(\{X=6\}\big))$ avec une forte probabilité.