CH VI : Variables aléatoires réelles discrètes

I. Généralités sur les variables aléatoires discrètes

I.1. Définition

Définition

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

Soit X une variable aléatoire réelle sur $(\Omega, \mathscr{A})$.

- La v.a.r. X est dite **discrète** si son esemble image $X(\Omega)$ est au plus dénombrable (*i.e.* si $X(\Omega)$ est un ensemble fini ou si $X(\Omega)$ est infini dénombrable).
- On dit que la v.a.r. X est **finie** si $X(\Omega)$ est fini.
  On dit que la v.a.r. X est **infinie** si $X(\Omega)$ est un ensemble infini.

Remarque

- Le caractère au plus dénombrable de $X(\Omega)$ est à l'origine des spécificités des v.a.r. discrètes. La première conséquence est que l'on peut numéroter les éléments de $X(\Omega)$ *i.e.* l'écrire :

$$X(\Omega) = \{x_i \mid i \in I\}$$

  où I est un ensemble :

  × fini (*i.e.* que I peut être mis en bijection avec $[\![1, n]\!]$ pour un $n \in \mathbb{N}^*$),

  × ou infini dénombrable (*i.e.* que I peut être mis en bijection avec $\mathbb{N}$).

- Rappelons que $\mathbb{Z}$ est un ensemble (infini) dénombrable puisque l'application φ suivante est une bijection :

$$\begin{array}{rccl} \varphi & : & \mathbb{Z} & \to & \mathbb{N} \\ & & z & \mapsto & \begin{cases} -2z & \text{si } z \text{ négatif} \\ 2z-1 & \text{si } z \text{ positif} \end{cases} \end{array}$$

  (*« il y a autant d'entiers relatifs que d'entiers naturels »*)

- De même, $\mathbb{N} \times \mathbb{N}$ et $\mathbb{Q}$ sont des ensembles (infinis) dénombrables donc peuvent (en théorie) servir à indexer l'ensemble image d'une v.a.r. discrète X.
- En pratique, les v.a.r. que nous étudions ont un ensemble image $X(\Omega)$ indexé par une partie (finie ou non) de $\mathbb{N}$ (voire de $\mathbb{Z}$). Ce qui s'écrit :

$$X(\Omega) = \{x_i \mid i \in I\} \quad \text{où} \quad I \subset \mathbb{N}$$

Il faut faire attention à ne pas confondre les deux notions suivantes.

- L'ensemble image $X(\Omega)$ est indexé par $\mathbb{N}$.
  Autrement dit : $X(\Omega) = \{x_0, x_1, x_2, x_3, \ldots\}$.
- L'ensemble image $X(\Omega)$ est à valeurs dans $\mathbb{N}$.
  Par exemple : $X(\Omega) = \{3, 5, 7, 10, 11 \ldots\}$

On peut d'ailleurs préciser que :

$$X \text{ est à valeurs dans } \mathbb{N} \quad \Rightarrow \quad X \text{ est une v.a.r. discrète}$$

$$X \text{ est à valeurs dans } \mathbb{N} \quad \cancel{\Leftarrow} \quad X \text{ est une v.a.r. discrète}$$

Par exemple, si $X(\Omega) = \{1, \sqrt{2}, \mathrm{e}^6, 23\}$, alors :

× X est une v.a.r. discrète (puisque $X(\Omega)$ est un ensemble fini),

× $X(\Omega)$ n'est pas à valeurs dans $\mathbb{N}$.

I.2. Système complet d'événements associé à une v.a.r. discrète

Théorème 1.

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

Soit X *une variable aléatoire réelle discrète sur* $(\Omega, \mathscr{A})$.

Notons $X(\Omega) = \{x_i \mid i \in I\}$ *où* $I \subseteq \mathbb{N}$.

> La famille $([X = x_i])_{i \in I}$ est un système complet d'événements, appelé **système complet d'événements associé à X.**

Démonstration.
Il y a deux propriétés à démontrer.

1) $([X = x_i])_{i \in I}$ est une famille d'événements deux à deux incompatibles :

Soit $(i, j) \in I^2$ tel que $i \neq j$ et soit $\omega \in [X = x_i] \cap [X = x_j]$.
Cela signifie que :

× $\omega \in [X = x_i]$ donc $X(\omega) = x_i$,

× $\omega \in [X = x_i]$ donc $X(\omega) = x_j$.

Par définition, $x_i \neq x_j$.
On en conclut qu'il n'existe pas d'élément $\omega \in [X = x_i] \cap [X = x_j]$.
Ainsi $[X = x_i] \cap [X = x_j] = \varnothing$.

2) $\Omega = \bigcup\limits_{i \in I} [X = x_i]$:

(⊃) $\bigcup\limits_{i \in I} [X = x_i] \subset \Omega$ puisque $\bigcup\limits_{i \in I} [X = x_i]$ est un événement (en tant qu'union dénombrable d'événements).

(⊂) Soit $\omega \in \Omega$.
Alors $X(\omega) \in X(\Omega) = \{x_i \mid i \in I\}$.
Ainsi, il existe $i \in I$ tel que $X(\omega) = x_i$ *i.e.* $\omega \in [X = x_i]$. □

I.3. Loi d'une v.a.r. discrète

Définition

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

Soit X une variable aléatoire réelle discrète sur $(\Omega, \mathscr{A})$.

- On appelle **loi de probabilité** de X et on note $\mathbb{P}_X$ l'application :

$$\mathbb{P}_X \ : \left| \begin{array}{ccccc} X(\Omega) & \to & \mathbb{R} & & \\ x & \mapsto & \mathbb{P}_X(x) & = & \mathbb{P}([X = x]) \end{array} \right.$$

- Autrement dit, la loi de X est la donnée de l'ensemble des valeurs $\mathbb{P}([X = x])$ pour x décrivant $X(\Omega)$.

Remarque
Dans le cas d'une v.a.r. discrète **finie**, on pourra représenter la loi de X sous la forme d'un tableau.

Exemple
On considère l'expérience aléatoire consistant à effectuer 4 lancers successifs d'une pièce de monnaie équilibrée.

- $\Omega = \{P, F\}^4$.
- On munit Ω de la probabilité uniforme notée $\mathbb{P}$.
($(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ est un espace probabilisé)
$\text{Card}(\Omega) = 2^4 = 16$.
- On note X la v.a.r. qui compte le nombre de P obtenu lors du lancer.
$X(\Omega) = \{0, 1, 2, 3, 4\}$.

$x \in X(\Omega)$	0	1	2	3	4
$\mathbb{P}([X = x])$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Remarque `(POLY)`

- Dans le cas où X est une v.a.r. discrète, nous avons vu que la loi de X est déterminée par la suite $(\mathbb{P}([X = x]))_{x \in X(\Omega)}$.
- Inversement, si l'on se donne une suite $(p_n)_{n \in \mathbb{N}}$ dans quel cas peut-on dire que cette suite est la loi d'une v.a.r. discrète ?

  Évidemment, il faut que : $\forall n \in \mathbb{N}$, $p_n \geqslant 0$ et $\sum\limits_{n=0}^{+\infty} p_n = 1$.
  Cette condition est même suffisante.
  Plus précisément, pour toute suite $(x_n)_{n \in \mathbb{N}}$ de réels, il existe un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ et une v.a.r. discrète X telle que $X(\Omega) \subseteq \{x_n \mid n \in \mathbb{N}\}$ et : $\forall n \in \mathbb{N}$, $\mathbb{P}([X = x_n]) = p_n$.

I.4. Fonction de répartition d'une v.a.r. discrète

Théorème 2.

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

Soit X une v.a.r. discrète sur $(\Omega, \mathscr{A})$ telle que $X(\Omega) = \{x_i \mid i \in I\}$ $(I \subseteq \mathbb{N})$.

Alors la fonction de répartition F_X est déterminée par :

$$\boxed{\forall x \in \mathbb{R}, \quad F_X(x) = \sum_{\substack{x_i \leqslant x \\ i \in I}} \mathbb{P}([X = x_i])}$$

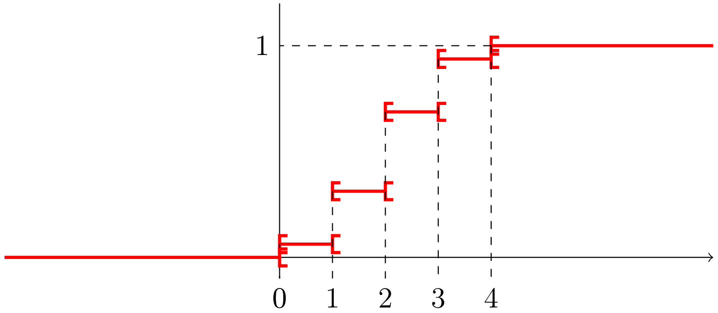
Démonstration.
On remarque tout d'abord que : $[X \leqslant x] = \bigcup\limits_{\substack{x_i \leqslant x \\ i \in I}} [X = x_i]$.

$(\subset)$ Soit $\omega \in [X \leqslant x]$. Alors $X(\omega) \leqslant x$.
Or $X(\omega)$ est un élément de l'ensemble image de X donc s'écrit sous la forme x_i pour un certain $i \in I$.

$(\supset)$ Soit $\omega \in \bigcup\limits_{\substack{x_i \leqslant x \\ i \in I}} [X = x_i]$. Alors ω est un élément de l'un des événements $[X = x_i]$ avec $i \in I$ et $x_i \leqslant x$.
Ainsi, $X(\omega) = x_i \leqslant x$ et donc $\omega \in [X \leqslant x]$.

On a affaire à une union au plus dénombrable d'événements deux à deux incompatibles. On obtient alors le résultat grâce à la σ-additivité de $\mathbb{P}$. $\square$

Exemple
On reprend l'exemple précédent. La fonction de répartition de la v.a.r. X comptant le nombre de P est donnée par le graphe suivant.



Rappelons que $X(\Omega) = \{0, 1, 2, 3, 4\}$.
Expliquons comment obtenir ce graphique :

× si $x < 0$: $[X \leqslant x] = \varnothing$.
Ainsi : $F_X(x) = \mathbb{P}([X \leqslant x]) = 0$.

× si $x \in [0, 1[$: $[X \leqslant x] = [X = 0]$.
Ainsi : $F_X(x) = \mathbb{P}([X \leqslant x]) = \mathbb{P}([X = 0]) = \frac{1}{16}$.

× si $x \in [1, 2[$: $[X \leqslant x] = [X = 0] \cup [X = 1]$. Ainsi :

$$\begin{aligned} F_X(x) &= \mathbb{P}([X \leqslant x]) = \mathbb{P}([X = 0] \cup [X = 1]) \\ &= \mathbb{P}([X = 0]) + \mathbb{P}([X = 1]) = \frac{1}{16} + \frac{4}{16} = \frac{5}{16} \end{aligned}$$

× si $x \in [2, 3[$: $F_X(x) = \mathbb{P}([X = 0]) + \mathbb{P}([X = 1]) + \mathbb{P}([X = 2]) = \frac{9}{16}$

× ...

- On obtient une fonction constante par morceaux qui présente des sauts de discontinuité. Cette forme en escalier est **caractéristique** des fonctions de répartition des v.a.r. discrètes.
- Les contremarches (*i.e.* les sauts de discontinuité) ont pour hauteur les valeurs successives de $\mathbb{P}([X = x_i])$ (les x_i étant rangés dans l'ordre).

Théorème 3.

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

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

1) *Si X est fini i.e. $X(\Omega) = \{x_1, \dots, x_n\}$ alors :* $\sum\limits_{i=1}^{n} \mathbb{P}([X = x_i]) = 1$

2) *Si X est infini i.e. $X(\Omega) = \{x_i \mid i \in \mathbb{N}\}$ alors :* $\sum\limits_{i=0}^{+\infty} \mathbb{P}([X = x_i]) = 1$

3) *On peut résumer ces propriétés en notant :*

$$\boxed{\sum_{x \in X(\Omega)} \mathbb{P}([X = x]) = 1}$$

Démonstration.
La famille $([X = x])_{x \in X(\Omega)}$ est un système complet d'événements.
En particulier :

$$\Omega = \bigcup_{x \in X(\Omega)} [X = x]$$

En appliquant $\mathbb{P}$ de part et d'autre, on obtient :

$$\underset{\substack{\shortparallel \\ 1}}{\mathbb{P}(\Omega)} = \mathbb{P}\left(\bigcup_{x \in X(\Omega)} [X = x]\right) = \sum_{x \in X(\Omega)} \mathbb{P}([X = x])$$

La dernière égalité est valide car $([X = x])_{x \in X(\Omega)}$ est une famille d'événements deux à deux incompatibles *(c'est la σ-additivité de $\mathbb{P}$: le caractère dénombrable de $X(\Omega)$ est donc indispensable ici !).* □

I.5. Opérations sur les v.a.r. discrètes

Théorème 4.

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

Soient X et Y deux v.a.r. discrètes sur $(\Omega, \mathscr{A})$ et $\lambda \in \mathbb{R}$.

$$\boxed{X + Y,\ \lambda X \text{ et } XY \text{ sont des v.a.r. discrètes.}}$$

où l'on a noté :

- $X + Y \ : \left| \begin{array}{ccc} \Omega & \to & \mathbb{R} \\ \omega & \mapsto & X(\omega) + Y(\omega) \end{array} \right.$
- $\lambda X \ : \left| \begin{array}{ccc} \Omega & \to & \mathbb{R} \\ \omega & \mapsto & \lambda X(\omega) \end{array} \right.$
- $XY \ : \left| \begin{array}{ccc} \Omega & \to & \mathbb{R} \\ \omega & \mapsto & X(\omega)\, Y(\omega) \end{array} \right.$

Démonstration.
Démontrons tout d'abord que $X + Y$, λY et XY sont des v.a.r.

(i) $X + Y$, λX et XY sont bien des fonctions de Ω dans $\mathbb{R}$.

(ii) Admis.

Il reste alors à démontrer que ces v.a.r. sont discrètes, autrement dit que leur ensemble image est au plus dénombrable.
Les v.a.r. X et Y étant discrètes, on peut noter :

× $X(\Omega) = \{x_i \mid i \in I\}$ où $I \subseteq \mathbb{N}$,

× $Y(\Omega) = \{y_j \mid j \in J\}$ où $J \subseteq \mathbb{N}$.

Ainsi : $(\lambda X)(\Omega) = \{\lambda x_i \mid i \in I\}$. Comme $I \subseteq \mathbb{N}$, la v.a.r. λX est bien discrète.
D'autre part :

$$(X + Y)(\Omega) = \{x_i + y_j \mid (i, j) \in I \times J\}$$

$$\text{et } (XY)(\Omega) = \{x_i y_j \mid (i, j) \in I \times J\}$$

Or $I \times J$ est une partie de $\mathbb{N}^2$ donc est au plus dénombrable.
Ainsi, $X + Y$ et XY sont bien des v.a.r. discrètes. □

Remarque *(Structure de l'ensemble des v.a.r. discrètes)*
On peut démontrer que l'ensemble $\mathcal{F}(\Omega, \mathbb{R})$ des fonctions de Ω dans $\mathbb{R}$ est un espace vectoriel (*cf* chapitre correspondant).
L'ensemble des v.a.r. discrètes :

× est un sous-ensemble de $\mathcal{F}(\Omega, \mathbb{R})$,

× est stable par la loi $+$ et la loi $\cdot$ (c'est l'objet du théorème précédent).

Cela permet de démontrer que l'ensemble des v.a.r. discrète est un sous-espace vectoriel de $\mathcal{F}(\Omega, \mathbb{R})$.

I.6. Transformation d'une v.a.r. discrète

Définition

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

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

On notera $g(X)$ l'application composée $g \circ X$:

$$g(X) \ : \ \begin{array}{rcl} \Omega & \to & \mathbb{R} \\ \omega & \mapsto & g(X(\omega)) \end{array}$$

Théorème 5.

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

Soit X une v.a.r. discrète sur $(\Omega, \mathscr{A})$ telle que $X(\Omega) = \{x_i \mid i \in I\}$ $(I \subseteq \mathbb{N})$.

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

L'application $g(X)$ est une **v.a.r. discrète** *dont la loi est donnée par :*

1) $g(X)(\Omega) = \{g(x_i) \mid i \in I\}$.

2) $\boxed{\forall y \in g(X)(\Omega), \quad \mathbb{P}([g(X) = y]) = \sum\limits_{\substack{x_i \in X(\Omega) \\ g(x_i) = y}} \mathbb{P}([X = x_i])}$

Démonstration.
Il faut tout d'abord démontrer que $g(X)$ est une v.a.r.
Or, par définition : $g(X) : \Omega \to \mathbb{R}$ (c'est le point ***(i)***).
Il reste à démontrer le point ***(ii)*** qui stipule que $g(X)$ est une machine à créer des événements. On ne le fait jamais en voie ECE (il manque certains notions pour le faire).

1) Par définition de $g(X)$, on a : $g(X)(\Omega) = g(X(\Omega)) = \{g(x_i) \mid i \in I\}$.
Ainsi, l'ensemble image de $g(X)$ est indéxé par $I \subseteq \mathbb{N}$, ensemble au plus dénombrable. On récupère donc au passage que $g(X)$ est une v.a.r. discrète.

2) Soit $y \in g(X)(\Omega)$. Alors :

$$\begin{array}{rcl} [g(X) = y] & = & \{\omega \in \Omega \mid g(X(\omega)) = y\} \\ & = & \{\omega \in \Omega \mid X(\omega) \in \{x_i \mid i \in I \text{ et } g(x_i) = y\}\} \\ & = & \bigcup\limits_{\substack{i \in I \\ g(x_i) = y}} [X = x_i] \end{array}$$

On obtient ainsi $[g(X) = y]$ comme union dénombrable d'événements deux à deux incompatibles. En effet :

× $\{i \in I \mid g(x_i) = y\} \subseteq I$ est au plus dénombrable,

× $[X = x_i] \cap [X = x_j] = \varnothing$ pour $i \neq j$.

On a alors :

$$\mathbb{P}([g(X) = y]) = \mathbb{P}\left(\bigcup\limits_{\substack{i \in I \\ g(x_i) = y}} [X = x_i]\right) = \sum\limits_{\substack{i \in I \\ g(x_i) = y}} \mathbb{P}([X = x_i])$$

□

Remarque

- Il faut essentiellement retenir de ce théorème que la transformée $g(X)$ d'une v.a.r. discrète est une v.a.r. discrète.
- La formule théorique donnant la loi de $g(X)$ n'est pas à retenir. Par contre, il faut savoir déterminer la loi de $g(X)$ en pratique.

Étude de la loi de $Y = g(X)$ sur quelques exemples.

Soit X une v.a.r. discrète.
On considère la v.a.r. discrète $Y = g(X)$ pour les applications g suivantes.

1) Si $g : x \mapsto ax + b$ où $a \neq 0$.

- $Y(\Omega) = g(X)(\Omega) = \{g(x) \mid x \in X(\Omega)\} = \{ax + b \mid x \in X(\Omega)\}$.
- Soit $y \in Y(\Omega)$.

$$\begin{aligned} \mathbb{P}([Y = y]) &= \mathbb{P}([aX + b = y]) \\ &= \mathbb{P}([aX = y - b]) \\ &= \mathbb{P}\left(\left[X = \tfrac{y-b}{a}\right]\right) \end{aligned}$$

(on note que comme $y \in Y(\Omega)$ alors $\frac{y-b}{a} \in X(\Omega)$)
En conclusion :

$$\boxed{\forall y \in Y(\Omega),\ \mathbb{P}([Y = y]) = \mathbb{P}\left(\left[X = \tfrac{y-b}{a}\right]\right)}$$

2) Si $g : x \mapsto x^2$.

- $Y(\Omega) = g(X)(\Omega) = \{g(x) \mid x \in X(\Omega)\} = \{x^2 \mid x \in X(\Omega)\}$.
- Soit $y \in Y(\Omega)$.

$$\begin{aligned} \mathbb{P}([Y = y]) &= \mathbb{P}(\left[X^2 = y\right]) \\ &= \mathbb{P}(\left[X = \sqrt{y}\right] \cup \left[X = -\sqrt{y}\right]) \end{aligned}$$

Deux cas se présentent alors :

× si $y \neq 0$, $\left[X = -\sqrt{y}\right]$ et $\left[X = \sqrt{y}\right]$ sont incompatibles.
Ainsi, pour tout $y \in Y(\Omega) \setminus \{0\}$:

$$\boxed{\mathbb{P}([Y = y]) = \mathbb{P}(\left[X^2 = y\right]) = \mathbb{P}([X = \sqrt{y}]) + \mathbb{P}([X = -\sqrt{y}])}$$

× si $y = 0$, on obtient $\mathbb{P}([Y = 0]) = \mathbb{P}([X = 0])$.

3) Si $g : x \mapsto |x|$.

- $Y(\Omega) = g(X)(\Omega) = \{g(x) \mid x \in X(\Omega)\} = \{|x| \mid x \in X(\Omega)\}$.
- Soit $y \in Y(\Omega)$.

$$\begin{aligned} \mathbb{P}([Y = y]) &= \mathbb{P}([|X| = y]) \\ &= \mathbb{P}([X = y] \cup [X = -y]) \end{aligned}$$

Deux cas se présentent alors :

× si $y \neq 0$, $[X = -y]$ et $[X = y]$ sont incompatibles.
Ainsi, pour tout $y \in Y(\Omega) \setminus \{0\}$:

$$\boxed{\mathbb{P}([Y = y]) = \mathbb{P}([|X| = y]) = \mathbb{P}([X = y]) + \mathbb{P}([X = -y])}$$

× si $y = 0$, on obtient $\mathbb{P}([Y = 0]) = \mathbb{P}([X = 0])$.

4) Si $g : x \mapsto \mathrm{e}^x$.

- $Y(\Omega) = g(X)(\Omega) = \{g(x) \mid x \in X(\Omega)\} = \{\mathrm{e}^x \mid x \in X(\Omega)\}$.
- Soit $y \in Y(\Omega)$.

$$\begin{aligned} \mathbb{P}([Y = y]) &= \mathbb{P}(\left[\mathrm{e}^X = y\right]) \\ &= \mathbb{P}([X = \ln(y)]) \end{aligned}$$

(il faut noter que $\ln(y)$ est bien défini puisque $y \in Y(\Omega) \subset \mathbb{R}_+^*$)
En conclusion :

$$\boxed{\forall y \in Y(\Omega),\ \mathbb{P}([Y = y]) = \mathbb{P}(\left[\mathrm{e}^X = y\right]) = \mathbb{P}([X = \ln(y)])}$$

II. Espérance d'une variable aléatoire discrète

II.1. Définition

Définition

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

Soit X une v.a.r. 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.r. 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.r. 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 Ω 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 Ω 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.r. discrète FINIE admet TOUJOURS une espérance.
- Une v.a.r. 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 6. (CULTURE)

Soit $\sum u_n$ une série semi-convergente.

i.e. $\left\{ \begin{array}{l} \sum u_n \text{ converge} \\ \sum |u_n| \text{ diverge} \end{array} \right.$

Alors, pour tout $\alpha \in \overline{\mathbb{R}}$, il existe une permutation σ de $\mathbb{N}$ telle que :

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

Remarque

- Cet énoncé est parfois nommé théorème de réarrangement.
- 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.

Exemple

On a rencontré lors du chapitre sur les séries un exemple de série semi-convergente, à savoir : $\sum \dfrac{(-1)^n}{n+1}$.

- Cette série n'est pas absolument convergente.
  En effet : $\left| \dfrac{(-1)^n}{n+1} \right| = \dfrac{1}{n+1}$ et la série $\sum \dfrac{1}{n+1}$ diverge.
- Cette série est convergente. On note ℓ sa somme (en fait $\ell = \ln(2)$).
  Ce résultat a été démontré en TD (« critère des séries alternées »).
  Pour ce faire, on démontre que les suites (S_{2n}) et (S_{2n+1}) sont adjacentes et admettent donc une limite commune qui est, par le théorème de recouvrement (à redémontrer à chaque fois), la limite de (S_n).

Le théorème de réarrangement stipule que l'on peut, en sommant les éléments dans un ordre différent, obtenir une somme qui converge vers n'importe quel réel α (éventuellement infini) fixé à l'avance (étonnant, non ?).

II.2. Propriétés de l'espérance

Théorème 7.

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé.
Soient X et Y deux v.a.r. discrètes sur $(\Omega, \mathscr{A})$.
On suppose que X et Y admettent chacune une espérance.

L'opérateur espérance vérifie les propriétés suivantes.

1) *Soit $\lambda \in \mathbb{R}$. Les v.a.r. $X + Y$ et λX admettent une espérance. 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 $(a, b) \in \mathbb{R}^2$, la v.a.r. $aX + b$ admet une espérance. De plus :*

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

3) $\boxed{X \geqslant 0 \Rightarrow \mathbb{E}(X) \geqslant 0}$ *(positivité de l'espérance)*

4) $\boxed{X \geqslant Y \Rightarrow \mathbb{E}(X) \geqslant \mathbb{E}(Y)}$ *(croissance de l'espérance)*

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

Démonstration.
Soit X et Y deux v.a.r. discrètes.

1) Pour plus de simplicité, on se limite ici au cas des v.a.r. finies (cas infini analogue) et on admet la première égalité (démonstration un peu longue).
Démontrons la seconde égalité. On note $X(\Omega) = \{x_1, \dots, x_n\}$.

- Si $\lambda = 0$, la v.a.r. λX est nulle. Son espérance est $\mathbb{E}(0) = 0$.
- Si $\lambda \neq 0$, la v.a.r. λX a pour ensemble image $(\lambda X)(\Omega) = \{\lambda x_1, \dots, \lambda x_n\}$.
Étant finie, cette v.a.r. admet une espérance, qui vaut :

$$\mathbb{E}(\lambda X) = \sum_{k=1}^{n} \lambda x_i \, \mathbb{P}([\lambda X = \lambda x_i]) = \lambda \sum_{k=1}^{n} x_i \, \mathbb{P}([X = x_i])$$

2) Soit $a \in \mathbb{R}$.
Le réel a peut être vue comme une v.a.r. discrète. C'est la v.a.r. constante $X_a : \Omega \to \mathbb{R}$ qui vérifie $X_a(\omega) = a$ pour tout $\omega \in \Omega$.
Ainsi $X_a(\Omega) = \{a\}$. Comme X_a est finie, elle admet une espérance et :

$$\mathbb{E}(X_a) = \sum_{x \in X(\Omega)} x \, \mathbb{P}([X_a = x]) = a \, \mathbb{P}([X_a = a]) = a \, \mathbb{P}(\Omega) = a$$

La seconde égalité est obtenue par linéarité :

$$\mathbb{E}(aX + b) = \mathbb{E}(aX) + \mathbb{E}(b) = a\mathbb{E}(X) + b$$

3) Notons $X(\Omega) = \{x_i \mid i \in I\}$.
La condition $X \geqslant 0$ signifie que pour tout $\omega \in \Omega$, $X(\omega) \geqslant 0$.
Autrement dit, $X(\Omega) \subseteq \mathbb{R}^+$ et pour tout $i \in I$, $x_i \geqslant 0$.
L'espérance $\mathbb{E}(X)$ est alors la somme d'une série à termes positifs (son terme général est $x_i \, \mathbb{P}([X = x_i])$). Elle est donc positive.

4) Notons tout d'abord que $X - Y$ admet une espérance (d'après ***1***).
On utilise alors le résultat précédent.
$X - Y \geqslant 0$ donc $\mathbb{E}(X - Y) \geqslant 0$ et $\mathbb{E}(X - Y) = \mathbb{E}(X) - \mathbb{E}(Y)$.

5) Supposons que $X \geqslant 0$ et $\mathbb{E}(X) = 0$.
Avec les mêmes notations que précédemment, on a donc $x_i \geqslant 0$.
Supposons $x_i > 0$. Alors nécessairement $\mathbb{P}([X = x_i]) = 0$.
Si on suppose, par l'absurde, que cette probabilité est non nulle, on obtient :

$$\mathbb{E}(X) \geqslant x_i \, \mathbb{P}([X = x_i]) > 0$$

ce qui contredit l'hypothèse.
Ainsi, pour tout i tel que $x_i \neq 0$, $\mathbb{P}([X = x_i]) = 0$.
On a donc forcément $\mathbb{P}([X = 0]) = 1$. □

`À RETENIR`

On retiendra notamment que la somme de v.a.r. discrètes qui admettent une espérance, admet une espérance.

II.3. Théorème de transfert

Théorème 8. *Théorème de transfert*

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

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

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

1) Si X est finie : et que $X(\Omega) = \{x_1, \dots, x_n\}$.
Dans ce cas, la v.a.r. $g(X)$ admet une espérance de valeur :

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

2) Si X est infinie : et que $X(\Omega) = \{x_i \mid i \in \mathbb{N}\}$.

La v.a.r. $g(X)$ admet une espérance	$\Leftrightarrow$	*La série $\sum g(x_n) \ \mathbb{P}([X = x_n])$ est absolument convergente*

Dans le cas de la convergence : $\boxed{\mathbb{E}(g(X)) = \sum_{i=0}^{+\infty} g(x_i) \ \mathbb{P}([X = x_i])}$

3) De manière générale, sous réserve d'existence :

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

Démonstration.

1) $g(X)$ est une v.a.r. d'ensemble image $g(X)(\Omega) = \{g(x_1), \dots, g(x_n)\}$
(abus de notation : les $g(x_i)$ ne sont pas forcément tous différents)
$g(X)$ étant finie, elle admet une espérance.

- Si tous les $g(x_i)$ sont différents. Alors $g(X)$ a pour espérance :

$$\mathbb{E}(g(X)) \ = \ \sum_{i=1}^{n} \ g(x_i) \ \mathbb{P}([g(X) = g(x_i)]) \ = \ \sum_{i=1}^{n} \ g(x_i) \ \mathbb{P}([X = x_i])$$

- Sinon, $g(X)(\Omega) = \{y_1, \dots, y_m\}$ avec $m < n$ et pour chaque y_j, il existe au moins un $i \in I$ tel que $y_j = g(x_i)$.
Il s'agit alors de regrouper les x_i qui admettent la même image par g.
On note alors $I_j = \{i \in I \mid g(x_i) = y_j\}$. On a :

$$\mathbb{E}(g(X)) \ = \ \sum_{j=1}^{m} \ y_j \ \mathbb{P}([g(X) = y_j])$$

Il suffit alors de remarquer : $[g(X) = y_j] = \bigcup_{i \in I_j} \ [X = x_i]$.

Par σ-additivité, on a :

$$\begin{aligned} y_j \ \mathbb{P}([g(X) = y_j]) \ &= \ y_j \sum_{i \in I_j} \mathbb{P}([X = x_i]) = \sum_{i \in I_j} y_j \ \mathbb{P}([X = x_i]) \\ &= \ \sum_{i \in I_j} g(x_i) \ \mathbb{P}([X = x_i]) \end{aligned}$$

Et ainsi :

$$\mathbb{E}(g(X)) \ = \ \sum_{j=1}^{m} \sum_{i \in I_j} g(x_i) \ \mathbb{P}([X = x_i]) \ = \ \sum_{i=1}^{n} g(x_i) \ \mathbb{P}([X = x_i])$$

2) Admis. □

II.4. Moments d'ordre r

Définition

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

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

1) Si X est finie : et que $X(\Omega) = \{x_1, \dots, x_n\}$.
Dans ce cas, la v.a.r. X admet, pour tout $r \in \mathbb{N}$, un **moment d'ordre** r, noté $m_r(X)$:

$$\boxed{m_r(X) \ = \ \sum_{i=1}^{n} \ {x_i}^r \ \mathbb{P}([X = x_i])}$$

2) Si X est infinie : et que $X(\Omega) = \{x_i \mid i \in \mathbb{N}\}$ et $r \in \mathbb{N}$.
Dans ce cas, si la série $\sum \ {x_n}^r \ \mathbb{P}([X = x_n])$ est absolument convergente, on dit que la v.a.r. X admet un **moment d'ordre** r, noté $m_r(X)$:

$$\boxed{m_r(X) = \sum_{i=0}^{+\infty} \ {x_i}^r \ \mathbb{P}([X = x_i])}$$

3) De manière générale, sous réserve d'existence :

$$\boxed{m_r(X) = \sum_{x \in X(\Omega)} \ x^r \ \mathbb{P}([X = x]) = \mathbb{E}(X^r)}$$

Proposition 1.

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

Soit X une v.a.r. discrète sur $(\Omega, \mathscr{A})$ et $r \in \mathbb{N}^*$.

X *admet un moment d'ordre* r	$\Rightarrow$	X *admet un moment d'ordre* s *pour tout* $s \in [\![1, r]\!]$

On retiendra notamment que si X admet un moment d'ordre 2 alors X admet un moment d'ordre 1 i.e. une espérance.

III. Variance d'une variable aléatoire discrète

III.1. Définition

Définition

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

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

Supposons que :

× la v.a.r. X admet une espérance $\mathbb{E}(X)$,

× la v.a.r. $X - \mathbb{E}(X)$ admet un moment d'ordre 2.

- On dit alors que la v.a.r. X admet une variance $\mathbb{V}(X)$:

$$\boxed{\mathbb{V}(X) = m_2(X - \mathbb{E}(X)) = \mathbb{E}\Big((X - \mathbb{E}(X))^2\Big)}$$

- Sous ces hypothèses on appelle écart-type et on note $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Remarque

- Remarquons tout d'abord que la 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 7)*
- 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 *i.e.* 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 :

$$\underbrace{\sqrt{\mathbb{E}((X - \mathbb{E}(X))^2)}}_{= \sigma(X)} \quad \cancel{=} \quad \mathbb{E}(\sqrt{(X - \mathbb{E}(X))^2}) \ = \ \mathbb{E}(\ |X - \mathbb{E}(X)|\ )$$

- On a vu que la v.a.r. $X - \mathbb{E}(X)$ a une espérance nulle.

  Une variable qui admet une espérance nulle est dite centrée.
  Partant de la v.a.r. X pas forcément centrée, on peut toujours introduire la v.a.r. $X - \mathbb{E}(X)$ qui elle est centrée.
  On parle alors de v.a.r. centrée associée à X.

  Dans ce contexte, la variance qui n'est autre que le moment d'ordre 2 de la v.a.r. $X - \mathbb{E}(X)$ est donc le moment centré d'ordre 2 de la v.a.r. X.

III.2. Détermination pratique de la variance

Théorème 9. *Formule de Kœnig-Huygens*

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

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

La v.a.r. X *admet une variance*	$\Leftrightarrow$	*La v.a.r.* X *admet un moment d'ordre* 2

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 admet une espérance :

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

Et ainsi :

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

($\Rightarrow$) Supposons que X admet une variance.
Par définition de variance, X et $(X - \mathbb{E}(X))^2$ admettent donc une espérance.
Or, d'après l'égalité $(**)$, la v.a.r. X^2 s'écrit comme la somme de v.a.r. qui admettent une espérance. Elle admet donc une espérance.

($\Leftarrow$) Supposons que X admet un moment d'ordre 2. Alors X admet un moment d'ordre 1. Autrement dit, X admet une espérance.
L'égalité $(*)$ démontre alors que la v.a.r. $(X - \mathbb{E}(X))^2$ admet une espérance car est la somme de v.a.r. admettant une espérance.

Supposons maintenant que X admet une variance et démontrons la formule.

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

□

III.3. Propriétés de la variance

Théorème 10.
Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé.
Soit X et Y deux v.a.r. discrètes sur $(\Omega, \mathscr{A})$.
On suppose que X et Y admettent chacune une variance.

L'opérateur variance vérifie les propriétés suivantes.
1) Soit $\lambda \in \mathbb{R}$. Les v.a.r. $X+Y$ et λX admettent une variance. De plus :

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

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

2) Pour tout $(a,b) \in \mathbb{R}^2$, la v.a.r. $aX+b$ admet une variance. De plus :

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

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

Démonstration.
1) Soit $\lambda \in \mathbb{R}$. On admet que $X+Y$ et λX admettent une variance.
D'après la formule de Kœnig-Huygens (et par linéarité de l'espérance) :

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

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

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

2) Soit $(a, b \in \mathbb{R}^2)$. Par hypothèse, X admet une variance. Elle admet donc un moment d'ordre 2 et, par conséquent, un moment d'ordre 1.
Démontrons maintenant que la v.a.r. $(aX+b) - \mathbb{E}(aX+b)$ admet un moment d'ordre 2. Tout d'abord :

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

Or, par définition de la variance de X, la v.a.r. $(X - \mathbb{E}(X))^2$ admet une espérance. C'est donc aussi le cas de la v.a.r. $((aX+b) - \mathbb{E}(aX+b))^2$. 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 7 :

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

□

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.r. et sa moyenne. Dans le cas d'une v.a.r. qui ne varie pas (une v.a.r. constante *i.e.* un réel a), il est logique que la variance soit nulle ($\mathbb{V}(a) = 0$).

III.4. Variables centrées réduites

Définition

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

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

a) Si X admet une espérance :

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

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

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

(ii) si $\mathbb{V}(X) \neq 1$, la v.a.r. $\dfrac{X}{\sigma(X)}$ (en supposant $\sigma(X) \neq 0$) est appelée v.a.r. réduite associée à X. Notons que cette v.a.r. admet une variance car somme de v.a.r. admettant une variance.
Comme son nom l'indique, cette v.a.r. 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 admet une variance, la variable $X^* = \dfrac{X - \mathbb{E}(X)}{\sigma(X)}$ est appelée variable centrée réduite associée à X.
Remarquons déjà que X^* admet une espérance et une variance 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.r. X qui admet une variance non nulle, on peut associer une v.a.r. X^* centrée réduite. Cette opération de normalisation peut être utile dans des démonstrations (on démontre une propriété pour une v.a.r. centrée réduite, on l'applique à X^* et on voit ce qu'on peut en déduire sur X).

IV. Lois discrètes finies

IV.1. Loi certaine

Définition

- On dit qu'une v.a.r. X suit **une loi certaine** s'il existe $m \in \mathbb{R}$ tel que :

  a) $\boxed{X(\Omega) = \{m\}}$

  b) $\boxed{\mathbb{P}([X = m]) = 1}$

- On dira aussi que la v.a.r. X est **certaine**, égale à m.
- Si X est une v.a.r. discrète finie telle que $X(\Omega) = \{x_1, \dots, x_n\}$:
  on dit que X suit une **loi quasi-certaine** s'il existe $i \in [\![1, n]\!]$ tel que $\mathbb{P}([X = x_i]) = 1$ (dans ce cas, $\mathbb{P}([X = x_j]) = 0$ pour tout $j \neq i$).

Expérience aléatoire de référence et v.a.r. associée

- On considère une expérience qui possède une ou plusieurs issues différentes dont une issue se produit avec probabilité 1.

Alors la v.a.r. X égale à m (pour un m choisi) si l'issue particulière se produit suit une loi quasi-certaine.

Exemple

1) On considère un dé truqué dont le résultat est toujours 6.
L'expérience aléatoire consiste en un lancer de ce dé.
On note X la v.a.r. donnant le résultat du dé.
X suit une loi quasi-certaine ($X(\Omega) = [\![1, 6]\!]$ et $\mathbb{P}([X = 6]) = 1$).

2) On considère une urne contenant n boules de couleurs différentes qui sont toutes numérotées par le même chiffre 7.
L'expérience consiste à effectuer un tirage dans cette urne.
On note X la v.a.r. donnant le numéro de la boule sortie.
Alors X suit la loi certaine d'ensemble image $X(\Omega) = \{7\}$.

Remarque

- Le fait que $X(\Omega) = \{m\}$ implique que la v.a.r. X est constante égale à m.
- Au contraire, une v.a.r. X suivant une loi quasi-certaine n'est pas constante (puisqu'elle peut prendre les valeurs $x_1, \dots, x_n$).
  Elle est seulement $\mathbb{P}$-presque sûrement constante : $\mathbb{P}([X = x]) = 1$ pour un certain $x \in X(\Omega)$.

Théorème 11.

Soit X une v.a.r. discrète finie suivant une loi certaine (ou quasi-certaine). Il existe donc $m \in \mathbb{R}$ tel que $\mathbb{P}([X = m]) = 1$.

1) *Alors X admet une espérance et une variance.*

2) *De plus :* $\boxed{\mathbb{E}(X) = m}$ *et* $\boxed{\mathbb{V}(X) = 0}$.

Démonstration.
Soit X une v.a.r. discrète suivant une loi (quasi) certaine.

1) X est une v.a.r. finie.
Elle admet donc un moment à tous les ordres.
En particulier, X admet une espérance et une variance.

2) • $\mathbb{E}(X) = \sum\limits_{x \in X(\Omega)} x \, \mathbb{P}([X = x]) = m \, \mathbb{P}([X = m]) = m \times 1 = m$

• D'après la formule de Kœnig-Huygens :

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

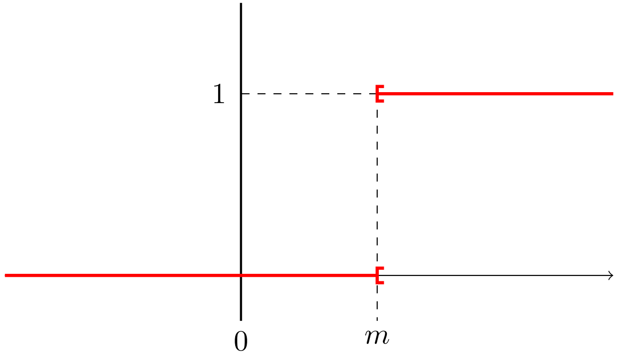
Or, par théorème de transfert :

$$\begin{aligned} \mathbb{E}(X^2) &= \sum_{x \in X(\Omega)} x^2 \, \mathbb{P}([X = x]) \\ &= m^2 \, \mathbb{P}([X = m]) = m^2 \end{aligned}$$

Ainsi : $\mathbb{V}(X) = m^2 - m^2 = 0$. □

Fonction de répartition

Si X est une v.a.r. discrète finie suivant une loi certaine (ou quasi-certaine) telle que $\mathbb{P}([X = m]) = 1$.



Proposition 2.

Soit X une v.a.r. discrète finie.

Alors on a : $\boxed{\mathbb{V}(X) = 0 \quad \Rightarrow \quad X \text{ suit une loi quasi-certaine}}$

Démonstration.

En effet, on a alors $\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = 0$ avec $(X - \mathbb{E}(X))^2 \geqslant 0$.

Ce qui prouve (point ***5)*** du Théorème 7) que : $\mathbb{P}(\left[(X - \mathbb{E}(X))^2 = 0\right]) = 1$.

Il suffit alors de remarquer que :

$$\left[(X - \mathbb{E}(X))^2 = 0\right] = [X - \mathbb{E}(X) = 0] = [X = \mathbb{E}(X)]$$

Ainsi, en notant $m = \mathbb{E}(X)$, on a $\mathbb{P}([X = m]) = 1$.

Autrement dit, $X = m$ $\mathbb{P}$-presque sûrement. □

Remarque

- Cette propriété est une équivalence : la réciproque est justifiée par le théorème précédent qui donne le calcul de $\mathbb{V}(X)$.
- Ainsi, la propriété $\mathbb{V}(X) = 0$ caractérise les v.a.r. X qui sont presque sûrement constantes et égales à leur moyenne.

IV.2. Loi uniforme

Définition

- On dit qu'une v.a.r. X suit **la loi uniforme** sur $[\![1, n]\!]$ (pour $n \in \mathbb{N}^*$) si :

  a) $\boxed{X(\Omega) = [\![1, n]\!]}$

  b) $\forall k \in [\![1, n]\!]$, $\boxed{\mathbb{P}([X = k]) = \dfrac{1}{n}}$

- Plus généralement, si $(a, b) \in \mathbb{N}^2$ et $a < b$, on dit qu'une v.a.r. X suit **la loi uniforme** sur $[\![a, b]\!]$ si :

  a) $\boxed{X(\Omega) = [\![a, b]\!]}$

  b) $\forall k \in [\![a, b]\!]$, $\boxed{\mathbb{P}([X = k]) = \dfrac{1}{b - a + 1}}$

- On utilisera la notation $X \hookrightarrow \mathcal{U}([\![a, b]\!])$ pour signifier que X suit la loi uniforme sur $[\![a, b]\!]$.

Expérience aléatoire de référence et v.a.r. associée

- On considère une expérience qui possède n issues différentes (qu'on numérote de 1 à n) qui sont équiprobables.

Alors la v.a.r. X égale à i si l'issue i est obtenue lors de l'expérience, suit la loi uniforme sur $[\![1, n]\!]$.

Exemple

1) On considère une urne contenant n boules numérotées de 1 à n.
L'expérience consiste à tirer une boule.
On note X la v.a.r. égale au numéro de la boule tirée.
Alors : $X \hookrightarrow \mathcal{U}([\![1, n]\!])$

2) On considère une pièce équilibrée.
L'expérience consiste en 1 lancer de cette pièce.
On note X la v.a.r. égale à 1 si on obtient Pile et 0 si on obtient Face.
Alors : $X \hookrightarrow \mathcal{U}(\llbracket 0,1 \rrbracket)$

Remarque

- Il s'agit bien de $\dfrac{1}{b-a+1}$ et non de $\dfrac{1}{b-a}$. En effet, il y a $b-a+1$ entiers entre a et b. Ce résultat est d'ailleurs cohérent avec le $\dfrac{1}{n}$ précédent.
- La loi uniforme sur $\mathcal{U}(\llbracket m,m \rrbracket)$ coïncide avec la loi certaine d'espérance m.

Théorème 12.

Soit X une v.a.r. discrète finie telle que $X \hookrightarrow \mathcal{U}(\llbracket 1,n \rrbracket)$ ($n \in \mathbb{N}^*$).

1) La v.a.r. X admet une espérance et une variance.

2) De plus : $\boxed{\mathbb{E}(X) = \dfrac{n+1}{2}}$ *et* $\boxed{\mathbb{V}(X) = \dfrac{n^2-1}{12}}$

Démonstration.

1) X admet une variance (et donc une espérance) car c'est une v.a.r. finie.

2) • $\mathbb{E}(X) \;=\; \sum\limits_{k=1}^{n} k\ \mathbb{P}([X=k]) = \dfrac{1}{n}\sum\limits_{k=1}^{n} k \;=\; \dfrac{1}{\cancel{n}}\ \dfrac{\cancel{n}(n+1)}{2}$

• D'après le théorème de transfert :

$$\mathbb{E}(X^2) = \sum_{k=1}^{n} k^2\ \mathbb{P}([X=k]) = \frac{1}{n}\sum_{k=1}^{n} k^2 = \frac{1}{\cancel{n}}\ \frac{\cancel{n}(n+1)(2n+1)}{6}$$

Ainsi, d'après la formule de Kœnig-Huygens :

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} \\ &= \frac{n+1}{12}(2\ (2n+1) - 3\ (n+1)) = \frac{(n+1)(n-1)}{12}\end{aligned}$$ □

Théorème 13.

Soit $(a,b) \in \mathbb{N}^2$ tel que $a < b$.
Soit X une v.a.r. discrète finie telle que $X \hookrightarrow \mathcal{U}(\llbracket a,b \rrbracket)$.
On note : $Y = X - a + 1$.

1) Alors : $Y \hookrightarrow \mathcal{U}(\llbracket 1, b-a+1 \rrbracket)$.

2) La v.a.r. X admet une espérance et une variance.

3) De plus : $\boxed{\mathbb{E}(X) = \dfrac{a+b}{2}}$ *et* $\boxed{\mathbb{V}(X) = \dfrac{(b-a)(b-a+2)}{12}}$

Démonstration.

1) On a $Y(\Omega) = \llbracket 1, b-a+1 \rrbracket$. Ainsi, pour tout $k \in \llbracket 1, b-a+1 \rrbracket$, on a :

$$\mathbb{P}([Y=k]) = \mathbb{P}([X-a+1=k]) = \mathbb{P}([X=k+a-1]) = \frac{1}{b-a+1}$$

2) X admet une variance (et donc une espérance) car c'est une v.a.r. finie.

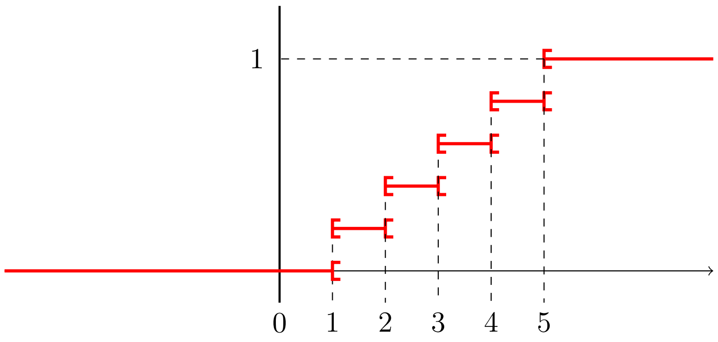
3) Comme $Y = X - a + 1$, on a $X = Y + a - 1$. On en déduit :

$$\mathbb{E}(X) = \mathbb{E}(Y) + \mathbb{E}(a-1) = \frac{(b-a+1)+1}{2} + a - 1 = \frac{a+b}{2}$$

$$\mathbb{V}(X) = \mathbb{V}(Y) = \frac{(b-a+1)^2-1}{12} = \frac{(b-a)(b-a+2)}{12}$$ □

Fonction de répartition

Si X est une v.a.r. discrète finie suivant la loi uniforme $\mathcal{U}(\llbracket 1,5 \rrbracket)$.



IV.3. Loi de Bernoulli

Définition

- On dit qu'une v.a.r. X suit **la loi de Bernoulli** de paramètre $p \in]0,1[$ si :

  a) $\boxed{X(\Omega) = \{0,1\}}$

  b) $\boxed{\mathbb{P}([X=1]) = p}$ et $\boxed{\mathbb{P}([X=0]) = 1-p} = q$

- On utilisera la notation $X \hookrightarrow \mathcal{B}(p)$ pour signifier que X suit la loi de Bernoulli de paramètre p.

Expérience aléatoire de référence et v.a.r. associée

- On considère une expérience aléatoire possédant deux issues (qui ne sont pas forcément équiprobables). L'une de ces issues est nommée « succès » et se produit avec probabilité p ; l'autre est nommée « échec » et se produit avec probabilité $1-p$.

Alors la v.a.r. X égale à 1 en cas de succès et 0 en cas d'échec (*i.e.* calculant le nombre de succès) suit la loi de Bernoulli de paramètre p.

Exemple

1) On considère une pièce de monnaie donnant Pile avec probabilité p et Face avec probabilité $1-p$.
L'expérience consiste en 1 lancer de cette pièce de monnaie.
Ainsi : $\Omega = \{\text{Pile}, \text{Face}\}$.
On note X la v.a.r. égale à 1 si on obtient Pile et à 0 si on obtient Face.

- $X(\Omega) = \{0,1\}$.
- $\mathbb{P}([X=1]) = p$ et $\mathbb{P}([X=0]) = 1-p$.

Ainsi $X \hookrightarrow \mathcal{B}(p)$.

2) On considère une urne contenant r boules rouges et v boules vertes.
L'expérience consiste à tirer une boule.
Ainsi : $\Omega = \{b_1, \dots, b_r, b_{r+1}, \dots, b_{r+v}\}$ *(on numérote chacune des boules).*
On note X la v.a.r. égale à 1 si on tire une boule rouge et 0 si on tire une boule verte.
Alors : $X \hookrightarrow \mathcal{B}\left(\dfrac{r}{r+v}\right)$.

Remarque

- Généralement, les cas $p = 0$ et $p = 1$ sont écartés : ils correspondent à une loi quasi-certaine.
- Si $p = \dfrac{1}{2}$, la loi $\mathcal{B}\left(\dfrac{1}{2}\right)$ coïncide avec la loi uniforme $\mathcal{U}(\llbracket 0,1 \rrbracket)$.
- Si X suit une loi de Bernoulli, alors $X(\Omega) = \{0,1\}$. On en déduit que, pour tout $r \in \mathbb{N}^*$, $X^r = X$ (c'est notamment vrai pour le cas $r = 2$).
- Considérons la v.a.r. X dont la loi est définie par :

  a) $X(\Omega) = \{-1,1\}$.

  b) $\mathbb{P}([X=1]) = p$ et $\mathbb{P}([X=-1]) = 1-p$.

  Alors X **ne suit pas** une loi de Bernoulli. Déjà, $X(\Omega) \neq \{0,1\}$.
  La v.a.r. X peut-être interprétée de la manière suivante. Si on obtient Pile (succès), la banque verse 1, si on obtient Face, on verse 1 à la banque.
  Autrement dit, X est le gain dans un jeu de Pile ou Face.

  Par contre, la v.a.r. $U = \dfrac{X+1}{2}$ suit la loi de Bernoulli $\mathcal{B}(p)$.

Théorème 14.

Soit X une v.a.r. discrète finie telle que $X \hookrightarrow \mathcal{B}(p)$ ($p \in]0,1[$).

1) *Alors X admet une espérance et une variance.*

2) *De plus :* $\boxed{\mathbb{E}(X) = p}$ *et* $\boxed{\mathbb{V}(X) = p\ q} = p\ (1-p)$

Démonstration.

1) X admet une variance (et donc une espérance) car c'est une v.a.r. finie.

2) • $\mathbb{E}(X) = \sum\limits_{x \in X(\Omega)} x\ \mathbb{P}([X = x]) = 0\ \mathbb{P}([X = 0]) + 1\ \mathbb{P}([X = 1]) = p$

• D'après le théorème de transfert :

$$\mathbb{E}(X^2) = \sum_{x \in X(\Omega)} x^2\ \mathbb{P}([X = x]) = 0^2\ \mathbb{P}([X = 0]) + 1^2\ \mathbb{P}([X = 1]) = p$$

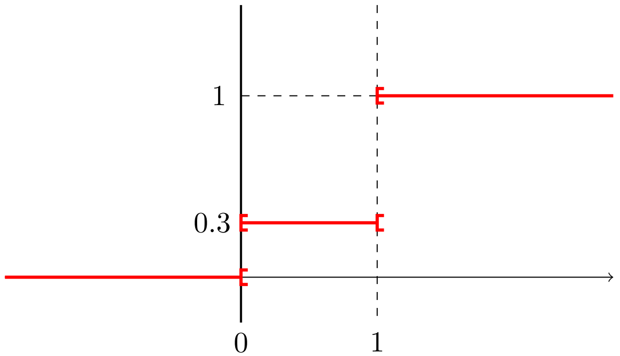
(on peut aussi remarquer que $X^2 = X$)
Ainsi, d'après le théorème de Kœnig-Huygens :

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

□

Fonction de répartition

Si X est une v.a.r. discrète finie suivant la loi de Bernoulli $\mathcal{B}(0.7)$.



IV.4. Loi binomiale

Définition

- On dit qu'une v.a.r. discrète X suit **la loi binomiale** de paramètre (n, p), où $n \in \mathbb{N}^*$ et $p \in]0,1[$ si :

  a) $\boxed{X(\Omega) = [\![0, n]\!]}$

  b) $\forall k \in [\![0, n]\!]$, $\boxed{\mathbb{P}([X = k]) = \binom{n}{k} p^k q^{n-k}}$ avec $q = 1 - p$

- On utilisera la notation $X \hookrightarrow \mathcal{B}(n, p)$ pour signifier que X suit la loi binomiale de paramètre (n, p).

Expérience aléatoire de référence et v.a.r. associée

- On considère une expérience aléatoire qui consiste en une succession de n épreuves indépendantes, chacune d'entre elles ayant deux issues : succès obtenu avec probabilité p et échec obtenu avec probabilité $q = 1 - p$.
- Autrement dit, l'expérience consiste à effectuer n épreuves de Bernoulli indépendantes (le résultat de l'une ne dépend pas du résultat des autres) et de même paramètre p.

Alors la v.a.r. donnant le nombre de succès obtenus au cours de cette expérience suit la loi binomiale de paramètre (n, p).

Exemple

1) On considère une pièce de monnaie donnant Pile avec probabilité p et Face avec probabilité $1 - p$.
L'expérience consiste en n lancers consécutifs de cette pièce de monnaie.
Ainsi : $\Omega = \{\text{Pile}, \text{Face}\}^n$.
On note X la v.a.r. égale au nombre de Pile obtenus lors de l'expérience.
Démontrons que X suit la loi binomiale de paramètre (n, p).

a) $X(\Omega) = [\![0, n]\!]$.

b) Soit $k \in [\![0, n]\!]$.

- Déterminons le nombre de n-tirages réalisant $[X = k]$.
  Un n-tirage réalisant $[X = k]$ est un n-uplet contenant k Pile. Il est entièrement déterminé par :
  × la position des k Pile : $\binom{n}{k}$ possibilités.

  Il y a donc $\binom{n}{k}$ tels tirages.
- Déterminons maintenant la probabilité d'apparition d'un tel n-tirage :
  × à chaque lancer, Pile est obtenu avec probabilité p et Face est obtenu avec probabilité $1 - p = q$.
  × or un tel n-tirage contient exactement k Pile et $n - k$ Face.

  La probabilité d'apparition d'un tel n-tirage est donc : $p^k\ q^{n-k}$.

On en déduit que : $\mathbb{P}([X = k]) = \binom{n}{k} p^k\ q^{n-k}$.

2) On considère une urne contenant r boules rouges et v boules vertes.
L'expérience consiste à tirer **successivement** n boules **avec remise**.
On note X la v.a.r. égale au nombre de boules rouges tirées.
Alors : $X \hookrightarrow \mathcal{B}\left(n, \dfrac{r}{r+v}\right)$

Cet exemple permet de retrouver la formule du binôme de Newton.
Comme $([X = k])_{k \in [\![0,n]\!]}$ est le sce associé à X :

$$\sum_{k=0}^{n} \mathbb{P}([X = k]) = 1 \quad \text{donc} \quad \sum_{k=0}^{n} \binom{n}{k} \left(\frac{r}{r+v}\right)^k \left(\frac{v}{r+v}\right)^{n-k} = 1.$$

On en conclut : $\boxed{\sum_{k=0}^{n} \binom{n}{k} r^k\ v^{n-k} = (r+v)^n}$

Remarque

- Évidemment, si une variable aléatoire suit la loi de Bernoulli de paramètre p alors $X \hookrightarrow \mathcal{B}(1,p)$.
- Dans un sujet de concours, on peut avoir à reconnaître une loi classique. On peut, dans le cas de la loi binomiale rédiger comme suit.
« La v.a.r. X compte le nombre de succès obtenus au cours d'une répétition de n épreuves de Bernoulli **indépendantes** et **de même paramètre** p. La v.a.r. X suit donc la loi binomiale de paramètre (n,p). »

Théorème 15.

Soit X une v.a.r. discrète finie telle que $X \hookrightarrow \mathcal{B}(n,p)$ $(n \in \mathbb{N}^*,\ p \in]0,1[)$.

1) *Alors X admet une espérance et une variance.*

2) *De plus :* $\boxed{\mathbb{E}(X) = n\ p}$ *et* $\boxed{\mathbb{V}(X) = n\ p\ q} = n\ p\ (1-p)$

Démonstration.

1) X admet une variance (et donc une espérance) car c'est une v.a.r. finie.

2) • $\mathbb{E}(X) = \sum\limits_{x \in X(\Omega)} x\ \mathbb{P}([X=x]) = \sum\limits_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} = \sum\limits_{k=1}^{n} k \binom{n}{k} p^k q^{n-k}$.

Or, pour $k \in [\![1,n]\!]$, on a : $k\binom{n}{k} = n\binom{n-1}{k-1}$. On en déduit :

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=1}^{n} n\binom{n-1}{k-1} p^k q^{n-k} = n \sum_{k=0}^{n-1} \binom{n-1}{k} p^{k+1} q^{n-(k+1)} \\ &= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k q^{(n-1)-k} = np\ (p+q)^{n-1} = np\end{aligned}$$

car $p+q=1$.

• $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$. On commence donc par calculer $\mathbb{E}(X^2)$.

$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{x \in X(\Omega)} x^2\ \mathbb{P}([X=x]) = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} \\ &= \sum_{k=1}^{n} k^2 \binom{n}{k} p^k q^{n-k}\end{aligned}$$

Or, pour $k \in [\![1,n]\!]$, on a : $k\binom{n}{k} = n\binom{n-1}{k-1}$.

Ainsi : $k^2\binom{n}{k} = k\ n\binom{n-1}{k-1} = n\ k\binom{n-1}{k-1}$.

Afin de pouvoir se débarraser du k restant, on remarque : $k = (k-1)+1$. On en déduit alors que, pour tout $k \in [\![2,n]\!]$:

$$\begin{aligned}n\ k\binom{n-1}{k-1} &= n\ (k-1)\binom{n-1}{k-1} + n\binom{n-1}{k-1} \\ &= n\ (n-1)\binom{n-2}{k-2} + n\binom{n-1}{k-1}\end{aligned}$$

On a alors :

$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{k=1}^{n} k^2\binom{n}{k} p^k q^{n-k} \\ &= \sum_{k=1}^{n} n\ (k-1)\binom{n-1}{k-1} p^k q^{n-k} + \sum_{k=1}^{n} n\binom{n-1}{k-1} p^k q^{n-k} \\ &= \sum_{k=2}^{n} n\ (k-1)\binom{n-1}{k-1} p^k q^{n-k} + np \\ &= \sum_{k=2}^{n} n\ (n-1)\binom{n-2}{k-2} p^k q^{n-k} + np \\ &= n\ (n-1) \sum_{k=0}^{n-2} \binom{n-2}{k} p^{k+2} q^{n-(k+2)} + np \\ &= n\ (n-1)\ p^2 \sum_{k=0}^{n-2} \binom{n-2}{k} p^k q^{(n-2)-k} + np \\ &= n\ (n-1)\ p^2 + np\end{aligned}$$

On en déduit :

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = n\ (n-1)\ p^2 + np\ -\ (np)^2 \\ &= \cancel{n^2p^2} - np^2 + np\ -\ \cancel{n^2p^2}\ =\ n\ (p-p^2)\ =\ n\ p\ (1-p)\end{aligned}$$

□

Remarque

- Dans le calcul de $\mathbb{E}(X^2)$, on fait apparaître une somme qui n'est autre que $\mathbb{E}(X)$ (déjà calculée).
- Afin de se focaliser sur la partie intéressante (le calcul de la première somme), on peut remarquer que :

$$\boxed{X^2\ =\ X\,(X-1) + X}$$

(qui correspond à l'idée utilisée : $n^2 = n\,(n-1) + 1$)

- Si on a déjà démontré que X admet une espérance, on obtient que :

$$X^2 \text{ admet une espérance} \Leftrightarrow X\ (X-1) \text{ admet une espérance}$$

et dans ce cas, par linéarité de l'espérance :

$$\boxed{\mathbb{E}(X^2)\ =\ \mathbb{E}(X\,(X-1)) + \mathbb{E}(X)}$$

- On aurait donc pu présenter le calcul différemment. Plus précisément en commençant par calculer $\mathbb{E}(X\,(X-1))$ et en ajoutant $\mathbb{E}(X)$.

Proposition 3.

Soit X une v.a.r. discrète finie telle que $X \hookrightarrow \mathcal{B}(n,p)$ $(n \in \mathbb{N}^*,\ p \in]0,1[)$.
Alors $n - X$ suit la loi binomiale de paramètre (n,q).

Démonstration.
Notons $Y = n - X$. On a alors :

a) $Y(\Omega) = \{n - x \mid x \in X(\Omega)\} = \{n, n-1, \dots, 1, 0\} = [\![0,n]\!]$.

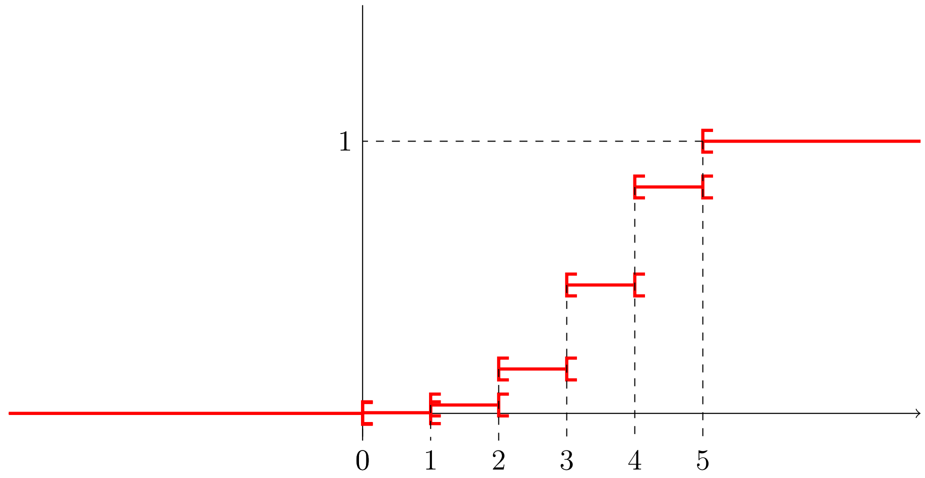
b) Soit $k \in [\![0,n]\!]$, $\mathbb{P}([Y=k]) = \mathbb{P}([n - X = k]) = \mathbb{P}([X = n-k])$

$$= \binom{n}{n-k} p^{n-k} q^{n-(n-k)} = \binom{n}{k} q^k p^{n-k}$$

□

Fonction de répartition

Si X est une v.a.r. discrète finie suivant la loi binomiale $\mathcal{B}(5,.7)$.



IV.5. Loi hypergéométrique (BONUS)

Définition

- Soit $(N, n) \in (\mathbb{N}^*)^2$ tel que $1 \leqslant n \leqslant N$ et soit $p \in]0,1[$ $(q = 1-p)$.
- On dit qu'une v.a.r. discrète X suit **la loi hypergéométrique** de paramètre (N, n, p) si :

  a) $\boxed{X(\Omega) = [\![\max(0, n-Nq), \min(n, Np)]\!]}$

  b) $\forall k \in [\![\max(0, n-Nq), \min(n, Np)]\!]$,

$$\boxed{\mathbb{P}([X=k]) = \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}}$$

- On utilisera la notation $X \hookrightarrow \mathcal{H}(N, n, p)$ pour signifier que X suit la loi hypergéométrique de paramètre (N, n, p).

Expérience aléatoire de référence et v.a.r. associée

- On considère une urne qui contient a boules blanches et b boules noires.
  L'expérience consiste à tirer **simultanément** n boules dans l'urne (on suppose $n \leqslant a+b$).
  On note X la v.a.r. égale au nombre de boules blanches obtenues.
  Alors : $X \hookrightarrow \mathcal{H}\left(a+b, n, \dfrac{a}{a+b}\right)$.
  Autrement dit, on a :

$$\forall k \in X(\Omega), \quad \mathbb{P}([X=k]) = \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}}$$

  (ici, $N = a+b$ et $p = \dfrac{a}{a+b}$ d'où $Np = a$ et $Nq = b$)

- On considère une urne qui contient a boules blanches et b boules noires.
  L'expérience consiste à tirer **successivement** et **sans remise** n boules dans l'urne (on suppose $n \leqslant a+b$).
  On note X la v.a.r. égale au nombre de boules blanches obtenues.
  Alors : $X \hookrightarrow \mathcal{H}\left(a+b, n, \dfrac{a}{a+b}\right)$

  Autrement dit, pour ces deux expériences a priori différentes, la v.a.r. donnant le nombre de boules blanches suit la même loi hypergéométrique.

Remarque

- On retrouve la formule de Vandermonde.
  Considérons l'une ou l'autre des expériences précédentes.
  On suppose de plus que $a \geqslant n$ et $b \geqslant n$ de sorte que :
  $[\![\max(0, n-Nq), \min(n, Np)]\!] = [\![0, n]\!]$.
  Comme $([X=k])_{k \in [\![0,n]\!]}$ est le système complet d'événements associé à X :
  $\sum\limits_{k=0}^{n} \mathbb{P}([X=k]) = 1$ donc $\sum\limits_{k=0}^{n} \dfrac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}} = 1$.

  On en conclut : $\boxed{\sum\limits_{k=0}^{n} \binom{a}{k}\binom{b}{n-k} = \binom{a+b}{n}}$

Théorème 16.

Soit X une v.a.r. discrète finie telle que $X \hookrightarrow \mathcal{H}(N, n, p)$ (avec $1 \leqslant n \leqslant N$, $p \in]0,1[$).

1) *X admet une espérance et une variance.*

2) *De plus :* $\boxed{\mathbb{E}(X) = n\ p}$ *et* $\boxed{\mathbb{V}(X) = n\ p\ q\ \dfrac{N-n}{N-1}}$

Démonstration.

1) X admet une variance (et donc une espérance) car c'est une v.a.r. finie.

2) Pour plus de simplicité, on se place dans le cas où $n \leqslant Np$ et $n \leqslant Np$.
On a ainsi $X(\Omega) = [\![0, n]\!]$.
(c'est le cas lorsque le nombre n de boules tirées est inférieur au nombre a de boules blanches et au nombre b de boules noires)

- $\mathbb{E}(X) = \sum\limits_{x \in X(\Omega)} x \, \mathbb{P}([X = x]) = \sum\limits_{k=0}^{n} k \dfrac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}} = \sum\limits_{k=1}^{n} k \dfrac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}$.

  Or, pour $k \in [\![1, n]\!]$, on a : $k \dbinom{Np}{k} = Np \dbinom{Np-1}{k-1}$. On en déduit :

$$\begin{aligned}
\mathbb{E}(X) &= \frac{Np}{\binom{N}{n}} \sum_{k=1}^{n} \binom{Np-1}{k-1}\binom{Nq}{n-k} \\
&= \frac{Np}{\binom{N}{n}} \sum_{k=0}^{n-1} \binom{Np-1}{k}\binom{Nq}{(n-1)-k} \\
&= \frac{Np}{\binom{N}{n}} \binom{Np-1+Nq}{n-1} = \frac{Np}{\binom{N}{n}} \binom{N-1}{n-1}
\end{aligned}$$

  car : $Np - 1 + Nq = N(p+q) - 1 = N - 1$.
  Enfin, comme : $n \dbinom{N}{n} = N \dbinom{N-1}{n-1}$, on obtient :

$$\mathbb{E}(X) = p \, \frac{N\binom{N-1}{n-1}}{\binom{N}{n}} = p \, \frac{n \cancel{\binom{N}{n}}}{\cancel{\binom{N}{n}}} = n \, p$$

- $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$. On commence donc par calculer $\mathbb{E}(X^2)$.
  Pour ce faire, on remarque que : $\mathbb{E}(X^2) = \mathbb{E}(X(X-1)) + \mathbb{E}(X)$.
  Afin de déterminer $\mathbb{E}(X^2)$, on commence par le calcul de $\mathbb{E}(X(X-1))$.

$$\begin{aligned}
\mathbb{E}(X(X-1)) &= \sum_{x \in X(\Omega)} x(x-1) \, \mathbb{P}([X = x]) \\
&= \sum_{k=0}^{n} k\,(k-1) \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}} \\
&= \sum_{k=2}^{n} k\,(k-1) \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}
\end{aligned}$$

Or, pour $k \in [\![2, n]\!]$:

$$k\,(k-1) \binom{Np}{k} = Np(k-1) \binom{Np-1}{k-1} = Np\,(Np-1) \binom{Np-2}{k-2}$$

On en déduit :

$$\begin{aligned}
\mathbb{E}(X(X-1)) &= \frac{Np\,(Np-1)}{\binom{N}{n}} \sum_{k=2}^{n} \binom{Np-2}{k-2}\binom{Nq}{n-k} \\
&= \frac{Np\,(Np-1)}{\binom{N}{n}} \sum_{k=0}^{n-2} \binom{Np-2}{k}\binom{Nq}{(n-2)-k} \\
&= \frac{Np\,(Np-1)}{\binom{N}{n}} \binom{Np-2+Nq}{n-2} \\
&= \frac{Np\,(Np-1)}{\binom{N}{n}} \binom{N-2}{n-2}
\end{aligned}$$

D'autre part, comme :

$$n\,(n-1) \binom{N}{n} = N\,(n-1) \binom{N-1}{n-1} = N\,(N-1) \binom{N-2}{n-2}$$

on obtient :

$$\mathbb{E}(X(X-1)) = \frac{\cancel{N}p\,(Np-1)}{\cancel{\binom{N}{n}}} \; \frac{n(n-1)\cancel{\binom{N}{n}}}{\cancel{N}\,(N-1)}$$

On en déduit que :

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}(X(X-1)) + \mathbb{E}(X) - (\mathbb{E}(X))^2 \\
&= \frac{np}{N-1}\,(Np-1)\,(n-1) + np - n^2\,p^2 \\
&= \frac{np}{N-1}\,((Np-1)\,(n-1) + (N-1) - np\,(N-1)) \\
&= \frac{np}{N-1}\,(\cancel{Npn} - Np - n + \cancel{1} + N - \cancel{1} - \cancel{npN} + np) \\
&= \frac{np}{N-1}\,((N-n) - p\,(N-n)) \\
&= \frac{np\,(N-n)}{N-1}\,(1-p) \\
&= n\;p\;q\;\frac{N-n}{N-1}
\end{aligned}$$

□

Remarque

- La loi hypergéométrique n'est pas au programme officiel. Ainsi, il ne peut être exigé, dans un sujet de concours, la connaissance de cette loi.
- Cependant, on peut envisager un exercice qui consisterait à démontrer tous les points précédents (loi, espérance, variance).
- Il est toutefois peu probable que le calcul de l'espérance et de la variance soit demandé dans le cas général. On verra des exercices en TD où a, b et n sont fixés.

V. Lois discrètes infinies

V.1. Loi géométrique

Définition

- On dit qu'une v.a.r. X suit **la loi géométrique** de paramètre $p \in \,]0,1[$ si :

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

  b) $\forall k \in \mathbb{N}^*$, $\boxed{\mathbb{P}([X=k]) = p\, q^{k-1}}$ avec $q = 1-p$

- On utilisera la notation $X \hookrightarrow \mathcal{G}(p)$ pour signifier que X suit la loi géométrique de paramètre p.

Expérience aléatoire de référence et v.a.r. associée

- On considère une expérience aléatoire qui consiste en une succession infinie d'épreuves indépendantes, chacune d'entre elles ayant deux issues : succès obtenu avec probabilité p et échec obtenu avec probabilité $q = 1-p$.
- Autrement dit, l'expérience consiste à effectuer une infinité d'épreuves de Bernoulli indépendantes (le résultat de l'une ne dépend pas du résultat des autres) et de même paramètre p.

Alors la v.a.r. donnant le rang d'apparition du premier succès obtenu lors de l'expérience suit la loi géométrique de paramètre p.

Exemple

1) On considère une pièce de monnaie déséquilibrée donnant Pile avec probabilité p et Face avec probabilité $1-p$.
L'expérience consiste en n lancers consécutifs de cette pièce de monnaie.
Ainsi : $\Omega = \{\text{Pile}, \text{Face}\}^{\mathbb{N}^*}$.
On note X la v.a.r. égale au rang d'apparition du premier Pile obtenu au cours de l'expérience.
Alors : $X \hookrightarrow \mathcal{G}(p)$.

Dans la suite, on considère les événements :

× P_i : « obtenir Pile au $i^{\text{ème}}$ lancer »,

× F_i : « obtenir Face au $i^{\text{ème}}$ lancer ».

Démontrons que X suit la loi géométrique de paramètre p.

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

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

$$[X=k] = F_1 \cap F_2 \cap \ldots \cap F_{k-1} \cap P_k$$

Les lancers étant indépendants, on obtient :

$$\begin{aligned}\mathbb{P}([X=k]) &= \mathbb{P}(F_1 \cap F_2 \cap \ldots \cap F_{k-1} \cap P_k) \\ &= \mathbb{P}(F_1) \times \mathbb{P}(F_2) \times \ldots \times \mathbb{P}(F_{k-1}) \times \mathbb{P}(P_k) \\ &= p\, q^{k-1}\end{aligned}$$

2) On considère une urne contenant r boules rouges et v boules vertes.
L'expérience consiste au tirage infini d'une boule avec remise.
On note X la v.a.r. donnant le rang d'apparition de la première boule rouge. Alors : $X \hookrightarrow \mathcal{G}\left(\dfrac{r}{r+v}\right)$

Théorème 17.

Soit X une v.a.r. discrète infinie telle que $X \hookrightarrow \mathcal{G}(p)$ ($p \in \,]0,1[$). Alors :

1) *La v.a.r. X admet une espérance et une variance.*

2) $\boxed{\mathbb{E}(X) = \dfrac{1}{p}}$ *et* $\boxed{\mathbb{V}(X) = \dfrac{q}{p^2}} = \dfrac{1-p}{p^2}$

Démonstration.

1) $X(\Omega) = \mathbb{N}^*$ et pour tout $n \in \mathbb{N}^*$, $\mathbb{P}([X = n]) = p\, q^{n-1}$.

- La v.a.r. X admet une espérance si et seulement si la série $\sum\limits_{n \geqslant 1} n\, \mathbb{P}([X = n])$ est absolument convergente. Cette série étant à termes positifs, cela revient à démontrer qu'elle est convergente.
- Soit $N \in \mathbb{N}^*$.

$$\sum_{n=1}^{N} n\, \mathbb{P}([X = n]) = \sum_{n=1}^{N} n\, p\, q^{n-1} = p \sum_{n=1}^{N} n\, q^{n-1}$$

On reconnaît la somme partielle d'une série géométrique dérivée première convergente car de raison $q \in\,]-1, 1[$. Ainsi :

$$p \sum_{n=1}^{N} n\, q^{n-1} \underset{N \to +\infty}{\longrightarrow} p\, \frac{1}{(1-q)^2} = \frac{p}{(1-(1-p))^2} = \frac{p}{p^2} = \frac{1}{p}$$

- On en conclut que la série $\sum\limits_{n \geqslant 1} n\, q^{n-1}$ est convergente.
Ainsi X admet une espérance donnée par :

$$\mathbb{E}(X) = \sum_{n=1}^{+\infty} n\, p\, q^{n-1} = \frac{1}{p}$$

2) Démontrons que X admet un moment d'ordre 2 (et donc une variance).

- Comme $X^2 = X(X-1) + X$:

$$X^2 \text{ admet une espérance } \Leftrightarrow X(X-1) \text{ admet une espérance}$$

Notons que $X(X-1)$ est une variable aléatoire discrète dont l'ensemble image est $(X(X-1))(\Omega) = \{n\,(n-1) \mid n \in \mathbb{N}^*\}$.
D'après le théorème de transfert, $X(X-1)$ admet une espérance si et seulement si la série $\sum\limits_{n \geqslant 1} n(n-1)\, \mathbb{P}([X = n])$ est absolument convergente. Cette série étant à termes positifs, cela revient à démontrer qu'elle est convergente.

- Soit $N \in \mathbb{N}^*$.

$$\begin{aligned} \sum_{n=1}^{N} n(n-1)\, \mathbb{P}([X = n]) &= \sum_{n=1}^{N} n(n-1)\, p\, q^{n-1} \\ &= \sum_{n=2}^{N} n(n-1)\, p\, q^{n-1} \\ &= p\, q \sum_{n=2}^{N} n(n-1)\, q^{n-2} \end{aligned}$$

On reconnaît la somme partielle d'une série géométrique dérivée deuxième convergente car de raison $q \in\,]-1, 1[$. Ainsi :

$$p\, q \sum_{n=2}^{N} n(n-1)\, q^{n-2} \underset{N \to +\infty}{\longrightarrow} p\, q\, \frac{2}{(1-q)^3} = p\, q\, \frac{2}{p^3} = \frac{2q}{p^2}$$

- On en conclut que la série $\sum\limits_{n \geqslant 1} n(n-1)\, p\, q^{n-1}$ converge.
Ainsi, $X(X-1)$ admet une espérance.
Ce qui démontre que X^2 admet une espérance, donnée par :

$$\begin{aligned} \mathbb{E}(X^2) &= \mathbb{E}(X(X-1)) + \mathbb{E}(X) \\ &= \frac{2q}{p^2} + \frac{1}{p} = \frac{1}{p^2}(2q + p) \end{aligned}$$

- On en déduit, par la formule de Kœnig-Huygens :

$$\begin{aligned} \mathbb{V}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \\ &= \frac{1}{p^2}\, (2\, q + p) - \left(\frac{1}{p}\right)^2 \\ &= \frac{1}{p^2}\, (2\, q + p - 1) \\ &= \frac{1}{p^2}\, (2\,(1-p) + p - 1) \\ &= \frac{1}{p^2}\, (2 - 2\, p + p - 1) = \frac{1}{p^2}\, (1-p) = \frac{q}{p^2} \end{aligned}$$

$\square$

Proposition 4.

Soit X une v.a.r. discrète infinie telle que $X \hookrightarrow \mathcal{G}(p)$ ($p \in]0,1[$).

Alors pour tout $(k,\ell) \in \mathbb{N}^2$:

1) $\boxed{\mathbb{P}([X > k]) \ = \ q^k}$

2) $\boxed{\mathbb{P}([X > k+\ell]) \ = \ \mathbb{P}([X > k]) \times \mathbb{P}([X > \ell])}$

Démonstration.

Soit $(k,\ell) \in \mathbb{N}^2$.

1) • On remarque tout d'abord :

$$\mathbb{P}([X > k]) \ = \ 1 - \mathbb{P}(\overline{[X > k]}) \ = \ 1 - \mathbb{P}([X \leqslant k])$$

• Par ailleurs, comme $X(\Omega) = \mathbb{N}^*$: $[X \leqslant k] \ = \ \bigcup\limits_{i=1}^{k} [X = i]$

On en déduit :

$$\begin{aligned} \mathbb{P}([X \leqslant k]) &= \mathbb{P}\left(\bigcup_{i=1}^{k} [X = i]\right) \\ &= \sum_{i=1}^{k} \mathbb{P}([X = i]) \qquad \textit{(par incompatibilité)} \\ &= \sum_{i=1}^{k} p\, q^{i-1} \ = \ p \sum_{i=0}^{k-1} q^i \\ &= p\, \frac{1-q^k}{1-q} \ = \ 1 - q^k \end{aligned}$$

Enfin : $\mathbb{P}([X > k]) \ = \ 1 - \mathbb{P}([X \leqslant k]) \ = \ 1 - (1-q^k) \ = \ q^k$.

2) D'après le point précédent :

$$\mathbb{P}([X > k+\ell]) \ = \ q^{k+\ell} \ = \ q^k \times q^\ell \ = \ \mathbb{P}([X > k]) \times \mathbb{P}([X > \ell]) \qquad \square$$

Remarque

Comme $[X > k+\ell] \subseteq [X > k]$, on a :

$$\mathbb{P}_{[X>k]}([X > k+\ell]) = \frac{\mathbb{P}([X > k+\ell])}{\mathbb{P}([X > k])} = \frac{q^{k+\ell}}{q^k} = q^\ell = \mathbb{P}([X > \ell])$$

On dit alors que la loi géométrique est à perte de mémoire (la propriété $X > k$ est oubliée, seul le délai est retenu) ou encore que la loi géométrique est **sans mémoire**.

Exercice

Soit X une v.a.r. telle que : $X \hookrightarrow \mathcal{G}(p)$ où $p \in]0,1[$.

1) Démontrer que, pour tout $k \in \mathbb{N}^*$:

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

2) En déduire que : $\mathbb{E}(X) = \sum\limits_{k=0}^{+\infty} \mathbb{P}([X > k])$.

3) Retrouver la valeur de $\mathbb{E}(X)$ à l'aide de cette formule.

Démonstration.

1) Soit $k \in \mathbb{N}^*$. La v.a.r. X est à valeurs entières. On en déduit :

$$[X > k-1] = [X = k] \cup [X > k]$$

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) La v.a.r. X admet une espérance car elle suit une loi géométrique.
Par définition $\mathbb{E}(X) = \sum\limits_{k=1}^{+\infty} k \ \mathbb{P}([X = k])$. Soit $N \in \mathbb{N}^*$.
Considérons la somme partielle $\sum\limits_{k=1}^{N} k \ \mathbb{P}([X = k])$ correspondante.

$$\begin{aligned}
& \sum_{k=1}^{N} k \, \mathbb{P}([X = k]) \\
= \ & \sum_{k=1}^{N} k \left(\ \mathbb{P}([X > k-1]) - \ \mathbb{P}([X > k]) \right) \\
= \ & \sum_{k=1}^{N} k \, \mathbb{P}([X > k-1]) - \sum_{k=1}^{N} k \, \mathbb{P}([X > k]) \\
= \ & \sum_{k=0}^{N-1} (k+1) \ \mathbb{P}([X > k]) - \sum_{k=1}^{N} k \, \mathbb{P}([X > k]) \\
= \ & \sum_{k=0}^{N-1} k \ \mathbb{P}([X > k]) + \sum_{k=0}^{N-1} \mathbb{P}([X > k]) - \sum_{k=1}^{N} k \ \mathbb{P}([X > k]) \\
= \ & 0 \times \mathbb{P}([X > 0]) + \cancel{\sum_{k=1}^{N-1} k \ \mathbb{P}([X > k])} + \sum_{k=0}^{N-1} \mathbb{P}([X > k]) \\
& - \left(\cancel{\sum_{k=1}^{N-1} k \ \mathbb{P}([X > k]} + N \ \mathbb{P}([X > N]) \right) \\
= \ & \sum_{k=0}^{N-1} \mathbb{P}([X > k]) - N \ \mathbb{P}([X > N]) \\
= \ & -N \ q^N \ + \sum_{k=0}^{N-1} \mathbb{P}([X > k])
\end{aligned}$$

$$\boxed{\sum_{k=1}^{N} k \, \mathbb{P}([X = k]) = -N \ q^N \ + \sum_{k=0}^{N-1} \mathbb{P}([X > k])}$$

Ainsi : $\sum\limits_{k=0}^{N-1} \mathbb{P}([X > k]) \ = \ \sum\limits_{k=1}^{N} k \ \mathbb{P}([X = k]) + N \ q^N$.
Les membres droits de l'égalité admettent une limite. Plus précisément :

$$\lim_{N \to +\infty} \sum_{k=1}^{N} k \ \mathbb{P}([X = k]) \ = \ \mathbb{E}(X) \quad \text{et} \quad \lim_{N \to +\infty} N \ q^N = 0$$

On en déduit que le membre de gauche admet une limite et :

$$\lim_{N \to +\infty} \sum_{k=0}^{N-1} \mathbb{P}([X > k]) \ = \ \sum_{k=0}^{+\infty} \mathbb{P}([X > k]) \ = \ \mathbb{E}(X)$$

Remarque
La démonstration consiste à démontrer un résultat de type somme télescopique.
On pouvait faire apparaître directement une telle somme en remarquant :

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

Ainsi, par sommation et linéarité :

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

3) D'après ce qui précède : $\mathbb{E}(X) = \sum\limits_{k=0}^{+\infty} \mathbb{P}([X > k]) = \sum\limits_{k=0}^{+\infty} q^k = \dfrac{1}{1-q} = \dfrac{1}{p}$. $\square$

Remarque

- Dans un contexte où X est un variable aléatoire mesurant une durée de vie (durée de vie d'une cellule, durée de fonctionnement d'un composant électronique, nombre de cycle de charge/décharge autorisé par une batterie), on introduit souvent la fonction :

$$S : t \mapsto \mathbb{P}([X > t]) = 1 - \mathbb{P}([X \leqslant t]) = 1 - F_X(t)$$

  Dans ce cas, $S(t) = \mathbb{P}([X > t])$ représente la probabilité que l'objet (ou l'individu) considéré soit encore en vie après une durée t.

  Dans le cas d'un phénomène à durée de vie continue (durée de vie d'une cellule), la modélisation s'appuiera sur une v.a.r. X à densité.
  Dans le cas d'un phénomène à durée de vie discrète (nombre de cycles d'une batterie), la modélisation s'appuiera sur une v.a.r. X discrète.

- Considérons la propriété d'absence de mémoire dans ce contexte.

$$\mathbb{P}_{[X>k]}([X > k + \ell]) = \mathbb{P}([X > \ell])$$

  Considérons que X compte la durée de fonctionnement d'un composant avant une panne. Alors cette propriété signifie que la durée de vie restante d'un objet est indépendante de la durée de vie écoulée de l'objet (période durant laquelle il a fonctionné sans tomber en panne). Autrement dit, il n'y a pas de vieillissement ou encore d'usure du composant électronique considéré. C'est un cas assez fréquent en réalité : on peut considérer que les diodes, transistors, résistances, condensateurs sont sans usure puisque leur usure ne débute que bien après la fin de vie de l'objet dans lequel ils sont installés. C'est pourquoi la durée de vie d'un composant est souvent modélisée par une v.a.r. qui suit loi exponentielle qui est, elle aussi, sans mémoire (c'est même une propriété qui caractérise la loi exponentielle).

- Il est fréquent de tomber sur une étude de la fonction de survie aux concours (cadre de l'analyse de fiabilité qui mesure la probabilité d'occurrence d'une défaillance d'un système).

V.2. Loi de Poisson

Définition

- On dit qu'une v.a.r. X suit **la loi de Poisson** de paramètre $\lambda > 0$ si :

  a) $\boxed{X(\Omega) = \mathbb{N}}$

  b) $\forall k \in \mathbb{N}$, $\boxed{\mathbb{P}([X = k]) = \mathrm{e}^{-\lambda} \dfrac{\lambda^k}{k!}}$

- On utilisera la notation $X \hookrightarrow \mathcal{P}(\lambda)$ pour signifier que X suit la loi de Poisson de paramètre λ.

Exemple

Le `BO` suggère d'introduire la loi de Poisson comme loi limite.
On considère $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. telle que, pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{B}\left(n, \frac{\lambda}{n}\right)$ où $\lambda > 0$. On a alors, pour tout $k \leqslant n$:

$$\begin{aligned}
\mathbb{P}([X_n = k]) &= \binom{n}{k} \times \left(\frac{\lambda}{n}\right)^k \times \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n!}{k!(n-k)!} \, \frac{\lambda^k}{n^k} \, \left(1 - \frac{\lambda}{n}\right)^n \, \left(1 - \frac{\lambda}{n}\right)^{-k} \\
&= \frac{\lambda^k}{k!} \, \frac{\frac{n!}{(n-k)!}}{n^k} \, \left(1 - \frac{\lambda}{n}\right)^n \, \left(1 - \frac{\lambda}{n}\right)^{-k}
\end{aligned}$$

Intéressons-nous aux termes de ce produit.

- $\dfrac{\frac{n!}{(n-k)!}}{n^k} = \dfrac{n(n-1)\dots(n-(k-1))}{n^k} \underset{n\to+\infty}{\sim} \dfrac{n^k}{n^k} = 1$
  En effet, le numérateur apparaît comme produit de k éléments qui sont tous équivalents, lorsque n tend vers $+\infty$, à n.

- Notons $u_n = \left(1 - \dfrac{\lambda}{n}\right)^n$. Alors : $u_n = \exp\left(n \ln\left(1 - \dfrac{\lambda}{n}\right)\right)$.
  Or :
  $$\ln(u_n) = n \ln\left(1 - \frac{\lambda}{n}\right) \underset{n\to+\infty}{\sim} \cancel{n} \frac{-\lambda}{\cancel{n}} = -\lambda$$
  (c'est une instance de la propriété : $\lim\limits_{x\to 0} \dfrac{\ln(1+x)}{x} = 1$*)*
  D'où : $\ln(u_n) \to -\lambda$ et $u_n = \mathrm{e}^{\ln(u_n)} \underset{n\to+\infty}{\longrightarrow} \mathrm{e}^{-\lambda}$.

- Enfin, comme $1 - \dfrac{\lambda}{n} \underset{n\to+\infty}{\longrightarrow} 1$, on a $\left(1 - \dfrac{\lambda}{n}\right)^{-k} = \dfrac{1}{\left(1 - \frac{\lambda}{n}\right)^k} \underset{n\to+\infty}{\longrightarrow} \dfrac{1}{1} = 1$

On en déduit que :

$$\boxed{\mathbb{P}([X_n = k]) \underset{n\to+\infty}{\longrightarrow} \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} = \mathbb{P}([X = k])}$$

où X est une v.a.r. telle que $X \hookrightarrow \mathcal{P}(\lambda)$.
La loi de Poisson apparaît comme la limite de lois binomiales $\mathcal{B}\left(n, \dfrac{\lambda}{n}\right)$.
Ainsi, si n grand (et donc $\frac{\lambda}{n}$ proche de 0) la loi $\mathcal{P}(\lambda)$ est une bonne approximation de la loi $\mathcal{B}\left(n, \dfrac{\lambda}{n}\right)$.

Théorème 18.

Soit X une v.a.r. telle que $X \hookrightarrow \mathcal{P}(\lambda)$ $(\lambda > 0)$. Alors :

1) *La v.a.r. admet une espérance et une variance.*

2) $\boxed{\mathbb{E}(X) = \lambda}$ *et* $\boxed{\mathbb{V}(X) = \lambda}$

Démonstration.

1) $X(\Omega) = \mathbb{N}$ et pour tout $n \in \mathbb{N}$, $\mathbb{P}([X = n]) = \mathrm{e}^{-\lambda}\, \dfrac{\lambda^n}{n!}$.

La v.a.r. X admet une espérance si et seulement si la série $\sum\limits_{n \geqslant 1} n\, \mathbb{P}([X = n])$ est absolument convergente. Cette série étant à termes positifs, cela revient à démontrer qu'elle est convergente.

Soit $N \in \mathbb{N}$. Notons alors : $S_N = \sum\limits_{n=0}^{N} n\ \mathbb{P}([X = n]) = \sum\limits_{n=1}^{N} n\ \mathrm{e}^{-\lambda}\ \dfrac{\lambda^n}{n!}$.

$$\begin{aligned} S_N &= \mathrm{e}^{-\lambda} \sum_{n=1}^{N} n\ \frac{\lambda^n}{n!} = \mathrm{e}^{-\lambda} \sum_{n=1}^{N} \frac{\lambda^n}{(n-1)!} \\ &= \mathrm{e}^{-\lambda}\, \lambda \sum_{n=1}^{N} \frac{\lambda^{n-1}}{(n-1)!} = \mathrm{e}^{-\lambda}\, \lambda \sum_{n=0}^{N-1} \frac{\lambda^n}{n!} \end{aligned}$$

Or : $\sum\limits_{n=0}^{N-1} \dfrac{\lambda^n}{n!} \underset{N \to +\infty}{\longrightarrow} \mathrm{e}^{\lambda}$ par convergence de la série exponentielle.

On en déduit que X admet une espérance qui vaut :

$$\mathbb{E}(X) = \lim_{N \to +\infty} S_N = \mathrm{e}^{-\lambda}\, \lambda\, \mathrm{e}^{\lambda} = \lambda$$

2) Démontrons que X admet un moment d'ordre 2 (et donc une variance). Comme $X^2 = X(X-1) + X$:

$$X^2 \text{ admet une espérance} \Leftrightarrow X(X-1) \text{ admet une espérance}$$

Or : $(X(X-1))(\Omega) = \{n\,(n-1) \mid n \in \mathbb{N}\}$.

Ainsi, X admet une variance si $X(X-1)$ admet une espérance si et seulement si la série $\sum\limits_{n \geqslant 0} n(n-1)\, \mathbb{P}([X = n])$ est absolument convergente.

Cette série étant à termes positifs, cela revient à démontrer qu'elle est convergente.

Soit $N \in \mathbb{N}$. Notons alors :

$$T_N = \sum_{n=0}^{N} n(n-1)\ \mathrm{e}^{-\lambda}\ \frac{\lambda^n}{n!} = \sum_{n=2}^{N} n(n-1)\ \mathrm{e}^{-\lambda}\ \frac{\lambda^n}{n!}$$

$$\begin{aligned} T_N &= \mathrm{e}^{-\lambda} \sum_{n=2}^{N} n(n-1)\ \frac{\lambda^n}{n!} = \mathrm{e}^{-\lambda} \sum_{n=2}^{N} \frac{\lambda^n}{(n-2)!} \\ &= \mathrm{e}^{-\lambda}\, \lambda^2 \sum_{n=2}^{N} \frac{\lambda^{n-2}}{(n-2)!} = \mathrm{e}^{-\lambda}\, \lambda^2 \sum_{n=0}^{N-2} \frac{\lambda^n}{n!} \end{aligned}$$

Or : $\sum\limits_{n=0}^{N-2} \dfrac{\lambda^n}{n!} \underset{N \to +\infty}{\longrightarrow} \mathrm{e}^{\lambda}$ par convergence de la série exponentielle.

On en déduit que $X(X-1)$ admet une espérance qui vaut :

$$\mathbb{E}(X(X-1)) = \lim_{N \to +\infty} T_N = \mathrm{e}^{-\lambda}\, \lambda^2\, \mathrm{e}^{\lambda} = \lambda^2$$

Ainsi, X admet une variance et d'après la formule de Kœnig-Huygens :

$$\begin{aligned} \mathbb{V}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \\ &= \mathbb{E}(X(X-1)) + \mathbb{E}(X) - (\mathbb{E}(X))^2 \\ &= \cancel{\lambda^2} + \lambda - \cancel{\lambda^2} = \lambda \end{aligned}$$

□

Exemple

- La loi de Poisson est généralement utilisée comme loi de v.a.r. consistant à calculer le nombre d'événements d'un certain type se produisant sur un laps de temps donné.
  Cette modélisation est valide si :
  1) les événements se produisant sont indépendants,
  2) la probabilité d'apparition du phénomène dans un laps de temps donné T ne dépend que de cette durée T.

  Par exemple :

  a) On considère une autoroute pour laquelle il y a en moyenne 1.8 accidents par semaine. On note X la v.a.r. donnant le nombre d'accidents par semaine sur cette autoroute.
  On peut considérer $X \hookrightarrow \mathcal{P}(1.8)$.
  (notez que λ n'est pas forcément entier)

  b) On considère un livre contenant des fautes d'impression. On sait de plus qu'il y a en moyenne 0.8 faute d'impression par page.
  On note X le nombre de fautes d'impression sur une page de ce livre.
  On peut considérer $X \hookrightarrow \mathcal{P}(0.8)$.

  c) On considère un serveur téléphonique qui reçoit en moyenne trois appels toutes les minutes.
  On note X le nombre d'appels reçus par le serveur en une minute.
  On peut considérer $X \hookrightarrow \mathcal{P}(3)$.