

Variables discrètes

Exercices classiques, méthodes usuelles

I. Généralités sur les variables aléatoires discrètes

I.1. Notion de variable aléatoire discrète

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

Soit E un ensemble.

- On dit que X est une **variable aléatoire discrète** à valeur dans E et définie sur $(\Omega, \mathcal{A})$ si :
 - (i) X est une application de Ω dans E ($X : \Omega \rightarrow E$)
 - (ii') $\forall x \in E, \{X = x\} = \{\omega \in \Omega \mid X(\omega) = x\} = X^{-1}(\{x\}) \in \mathcal{A}$
 - (iii) L'ensemble $X(\Omega)$ (ensemble image de X) est au plus dénombrable (c'est-à-dire si $X(\Omega)$ est un ensemble fini ou si $X(\Omega)$ est infini dénombrable).
- On dit que la variable aléatoire X est **finie** si $X(\Omega)$ est fini.
On dit que la variable aléatoire X est **infinie** si $X(\Omega)$ est un ensemble infini.

Remarque

- Il est important de souligner que le programme permet de considérer des variables aléatoires à valeurs dans un ensemble E quelconque.
 - Cependant, dans les sujets, les variables aléatoires étudiées sont :
 - × majoritairement à valeurs dans $\mathbb{R}$,
 - × de temps en temps à valeurs dans $\mathbb{C}$,
 - × très rarement à valeurs dans un ensemble E qui n'est ni $\mathbb{R}$, ni $\mathbb{C}$.
 - Si les variables aléatoires réelles sont les plus étudiées aux concours, c'est parce que ce sont celles qui donnent lieu, à notre niveau, aux développements les plus importants. Typiquement, les notions de variance et d'écart-type ne sont définies que pour les variables aléatoires réelles. Par ailleurs, les lois usuelles sont toutes des lois de variables aléatoires réelles.
 - On pourrait alors se limiter à l'étude des variables aléatoires réelles. Le programme n'a pas pris ce parti car il peut arriver que l'étude d'autres variables aléatoires soit pertinent dans un exercice.
 - Profitons-en pour rappeler qu'un ensemble G est :
 - × fini s'il existe $n \in \mathbb{N}$ tel que G est en bijection avec $\llbracket 1, n \rrbracket$.
 - × (infini) dénombrable si G est en bijection avec $\mathbb{N}$,
 - × au plus dénombrable s'il existe $I \subset \mathbb{N}$ tel que G est en bijection avec I .
- Les ensembles $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{Z}$ mais aussi $\mathbb{N} \times \mathbb{N}$ ou $\mathbb{N} \times \dots \times \mathbb{N}$ sont dénombrables.

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

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

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

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

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

En particulier : $\bigcup_{i \in I} \{X = x_i\} = \Omega$. Ainsi : $\mathbb{P}\left(\bigcup_{i \in I} \{X = x_i\} \right) = \mathbb{P}(\Omega) = 1$.

Et comme les événements de la famille $\left(\{X = x_i\} \right)_{i \in I}$ sont deux à deux incompatibles :

$$\mathbb{P}\left(\bigcup_{i \in I} \{X = x_i\} \right) = \sum_{i \in I} \mathbb{P}(\{X = x_i\}) = 1$$

1) Si $X(\Omega)$ est fini alors $X(\Omega) = \{x_1, \dots, x_n\}$ alors : $\sum_{i=1}^n \mathbb{P}(\{X = x_i\}) = 1$.

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

De manière générale, si $\left(\{X = x_i\} \right)_{i \in I}$ est un système complet d'événements :

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

I.3. Loi d'une v.a.r. discrète

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

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

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

$$\mathbb{P}_X : \begin{cases} X(\Omega) & \rightarrow \mathbb{R} \\ x & \mapsto \mathbb{P}_X(x) = \mathbb{P}(\{X = x\}) \end{cases}$$

- Autrement dit, la loi de X est la donnée de l'ensemble des valeurs $\mathbb{P}(\{X = x\})$ pour x décrivant $X(\Omega)$.
- On dit que X et Y ont même loi si $\mathbb{P}_X = \mathbb{P}_Y$.
Autrement dit, X et Y ont même loi si :

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

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

Lorsque X et Y suivent la même loi, on note : $X \sim Y$.

À RETENIR

Afin de déterminer la loi d'une v.a.r. discrète, on procède toujours en deux étapes.

1) Détermination de l'ensemble image $X(\Omega)$.

2) Calcul de $\mathbb{P}(\{X = x\})$ pour tout $x \in X(\Omega)$.

I.4. Transformée d'une v.a.r. discrète

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

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

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

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

$$\begin{aligned} g(X) &: \Omega \rightarrow \mathbb{R} \\ \omega &\mapsto g(X(\omega)) \end{aligned}$$

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

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

$$2) \quad \forall y \in g(X)(\Omega), \quad \mathbb{P}(\{g(X) = y\}) = \sum_{\substack{x_i \in X(\Omega) \\ g(x_i) = y}} \mathbb{P}(\{X = x_i\})$$

- Il est à noter que si deux v.a.r. X et Y ont même loi, alors leurs transformées par la même fonction g sont aussi de même loi.

$$X \sim Y \Rightarrow g(X) \sim g(Y)$$

Remarque

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

I.5. Une propriété usuelle des variable aléatoire discrète À VALEURS ENTIÈRES

Exercice

Soit X une variable aléatoire à valeurs entières (autrement dit : $X(\Omega) \subset \mathbb{N}$).

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

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

Démonstration.

Soit $k \in \mathbb{N}^*$. Tout d'abord :

$$\begin{aligned} \{X > k - 1\} &= \{X \geq k\} && (\text{car } X \text{ est à valeurs entières}) \\ &= \{X = k\} \cup \{X > k\} \end{aligned}$$

Ainsi :

$$\mathbb{P}(\{X > k - 1\}) = \mathbb{P}(\{X = k\}) + \mathbb{P}(\{X > k\}) \quad (\text{car les événements } \{X = k\} \text{ et } \{X > k\} \text{ sont incompatibles})$$

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

II. Espérance d'une variable aléatoire discrète

II.1. Familles sommables

II.1.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 :

$$\times \quad \forall x \in]0, +\infty], x \times (+\infty) = (+\infty) \times x = +\infty$$

$$\times \quad \forall x \in [0, +\infty], x + (+\infty) = (+\infty) + x = +\infty$$
- 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.

II.1.b) Notion de famille sommable

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 $\sum_{i \in I} x_i$, la borne supérieure (dans $[0, +\infty]$) de l'ensemble des sommes $\sum_{i \in F} x_i$ lorsque F décrit l'ensemble des parties finies de I .
- En d'autres termes :

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

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

II.1.c) Propriétés de manipulation

- Le calcul de somme ne dépend pas de l'ordre de sommation

Soit I un ensemble au plus dénombrable.

Soit σ une permutation de I (c'est-à-dire une bijection de I sur lui-même).

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

$$\sum_{i \in I} x_i = \sum_{k \in I} x_{\sigma(k)}$$

- 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 : $I = \bigcup_{n \in \mathbb{N}} I_n$.

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

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

- Sommes doubles

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 :

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

× 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 : $\forall n \in \mathbb{N}^*, \frac{1}{n(n+1)} \leq \frac{1}{n^2}$. On peut alors écrire :

$$\sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \leq \sum_{n=1}^{+\infty} \frac{1}{n^2} < +\infty$$

Ainsi, la famille $\left(\frac{1}{n(n+1)} \right)$ est sommable.

- Ajoutons que lorsqu'on travaille dans $[0, +\infty]$, on peut aussi 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 :

$$\forall (\lambda, \mu) \in [0, +\infty]^2, \sum_{i \in I} (\lambda x_i + \mu y_i) = \lambda \sum_{i \in I} x_i + \mu \sum_{i \in I} y_i$$

(il est important de noter que λ et μ sont positifs, de sorte à ne pas créer un éventuel problème de différence d'infinis)

II.1.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 : La famille $(x_i)_{i \in I}$ est sommable $\Leftrightarrow \sum_{i \in I} |x_i| < +\infty$

- 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 ℓ , alors ℓ 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 : $\sum_{i \in I} |x_i| < +\infty$.

- Notons qu'une famille finie est toujours sommable.

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 $[0, +\infty]$.

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

Remarque

- Si $(x_i)_{i \in I}$ est une famille d'éléments de $\mathbb{C}$, alors :

$$\forall i \in I, \quad \operatorname{Re}(x_i) \leq |x_i| \quad \text{et} \quad \operatorname{Im}(x_i) \leq |x_i|$$

Ainsi, si $(x_i)_{i \in I}$ est sommable (ce qui signifie $(|x_i|)_{i \in I}$ sommable), alors, par domination, les familles $(\operatorname{Re}(x_i))_{i \in I}$ et $(\operatorname{Im}(x_i))_{i \in I}$ sont sommables.

- Dans ce cas : $\sum_{k \in I} x_k = \sum_{k \in I} \operatorname{Re}(x_k) + i \sum_{k \in I} \operatorname{Im}(x_k)$.

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 :

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

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.

II.2. Une « nouvelle » définition de l'espérance

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

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

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

L'espérance de X , notée $\mathbb{E}(X)$ est, par définition, la quantité de $[0, +\infty]$ définie par :

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

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

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

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

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 $\left(x \mathbb{P}(\{X = x\})\right)_{x \in X(\Omega)}$ est aussi à valeurs dans $\mathbb{C}$. Par définition, cette famille est sommable si $\left(|x| \mathbb{P}(\{X = x\})\right)_{x \in X(\Omega)}$ l'est.
- Cela démontre au passage :

La v.a. X est d'espérance finie $\Leftrightarrow$ La v.a. $|X|$ est d'espérance finie

II.3. Propriétés de l'espérance

II.3.a) Linéarité

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 λX sont d'espérances finies. De plus :

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

(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 :

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

À RETENIR

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

II.3.b) Positivité, croissance

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

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

(positivité de l'espérance)

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

(croissance de l'espérance)

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

II.4. Théorème de transfert

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

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

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

- 2) Dans ce cas :

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

III. Variance d'une variable aléatoire discrète à valeurs réelles

III.1. Moment d'ordre 2

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

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

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

III.2. Variance d'une v.a. discrète à valeurs réelles

Soit X une v.a. discrète à valeurs réelles. Supposons que :

× la v.a. X est d'espérance finie,

× la v.a. $(X - \mathbb{E}(X))^2$ 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 :

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

- 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 : $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

III.3. Détermination en pratique de la variance

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

III.4. Propriétés de la variance

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

On suppose que X et Y sont de variances finies.

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

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + 2 \left(\mathbb{E}(XY) - \mathbb{E}(X) \times \mathbb{E}(Y) \right) + \mathbb{V}(Y) \quad \text{et} \quad \mathbb{V}(\lambda X) = \lambda^2 \mathbb{V}(X)$$

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

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

III.5. Variables centrées réduites

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. $\frac{X}{\sigma(X)}$ (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 $X^* = \frac{X - \mathbb{E}(X)}{\sigma(X)}$ 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 $a = \frac{1}{\sigma(X)} \in \mathbb{R}$ et $b = -\frac{\mathbb{E}(X)}{\sigma(X)} \in \mathbb{R}$.

- 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 :

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

IV. Moment d'ordre 1 et 2 d'une variable à VALEURS ENTIÈRES

IV.1. Espérance d'une d'une variable à VALEURS ENTIÈRES

Soit X une v.a. discrètes à valeurs dans $\llbracket 0, +\infty \rrbracket$.

$$\mathbb{E}(X) = \sum_{k=0}^{+\infty} \mathbb{P}(\{X > k\}) = \sum_{k=1}^{+\infty} \mathbb{P}(\{X \geq k\})$$

Démonstration. (démonstration qui tire parti des calculs de somme dans $[0, +\infty]$)

$$\begin{aligned} \sum_{k=1}^{+\infty} \mathbb{P}(\{X \geq k\}) &= \sum_{k=1}^{+\infty} \left(\sum_{j=k}^{+\infty} \mathbb{P}(\{X = j\}) \right) && (\text{car } X \text{ est à} \\ &&& \text{valeurs entières}) \\ &= \sum_{1 \leq k \leq j \leq +\infty} \mathbb{P}(\{X = j\}) \\ &= \sum_{j=1}^{+\infty} \left(\sum_{k=1}^j \mathbb{P}(\{X = j\}) \right) \\ &= \sum_{j=1}^{+\infty} j \mathbb{P}(\{X = j\}) = \mathbb{E}(X) \end{aligned}$$

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 $\sum_{j=k}^{+\infty} \mathbb{P}(\{X = j\}) = +\infty$ alors la somme $\sum_{k=1}^{+\infty} \left(\sum_{j=k}^{+\infty} \mathbb{P}(\{X = j\}) \right)$ 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 :

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

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

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

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

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

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

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

□

IV.2. Les fonctions génératrices

IV.2.a) Définition

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

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

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

(l'expression de $G_X(t)$ est issue du théorème de transfert appliqué à la variable aléatoire $f_t(X)$ - transformée de la variable aléatoire X - où $f_t : x \mapsto t^x$)

IV.2.b) La fonction génératrice caractérise la loi

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

1. a) La série entière $\sum \mathbb{P}(\{X = n\}) t^n$ a un rayon de convergence $R \geq 1$.
b) En particulier, G_X , somme de la série entière $\sum \mathbb{P}(\{X = n\}) t^n$ est développable en série entière sur (au moins) $] -1, 1[$ et donc de classe $\mathcal{C}^\infty$ sur sur (au moins) $] -1, 1[$.

$$2. \quad \forall n \in \mathbb{N}, \mathbb{P}(\{X = n\}) = \frac{1}{n!} G_X^{(n)}(0)$$

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

IV.2.c) Continuité de la fonction génératrice d'une v.a. discrète

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

1. La série entière $\sum \mathbb{P}(\{X = n\}) t^n$ converge normalement (au moins) sur le segment $[-1, 1]$.
2. La fonction G_X est continue (au moins) sur $[-1, 1]$ et $G_X = 1$.

IV.3. Détermination des moments d'ordre 1 et 2 des variables À VALEURS ENTIÈRES

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

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

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

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

Dans ce cas :

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

V. Indépendance de variables aléatoires discrètes

V.1. Indépendance de deux variables aléatoires discrètes

V.1.a) Définition

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

Soit (X, Y) un couple de v.a. discrètes.

- Les **v.a. X et Y sont indépendantes** (pour la probabilité $\mathbb{P}$) si pour tout $A \subset X(\Omega)$ et tout $B \subset Y(\Omega)$, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants.
- De manière équivalente, les **v.a. X et Y sont indépendantes** si pour tout $x \in X(\Omega)$ et tout $y \in Y(\Omega)$, les événements $\{X = x\}$ et $\{Y = y\}$ sont indépendants, c'est-à-dire si :

$$\begin{aligned} \forall x \in X(\Omega), \forall y \in Y(\Omega), \\ \mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(\{X = x\}) \times \mathbb{P}(\{Y = y\}) \end{aligned}$$

- On note $X \perp\!\!\!\perp Y$ pour signifier que les variables X et Y sont indépendantes.

MÉTHODO

Démontrer que deux v.a. discrètes ne sont pas indépendantes

- Pour démontrer que X et Y ne sont pas indépendantes, il suffit de trouver $x \in X(\Omega)$ et $y \in Y(\Omega)$ tels que :

$$\mathbb{P}(\{X = x\} \cap \{Y = y\}) \neq \mathbb{P}(\{X = x\}) \times \mathbb{P}(\{Y = y\})$$

- On essaiera de trouver $x \in X(\Omega)$ et $y \in Y(\Omega)$ tels que :

$$\mathbb{P}(\{X = x\} \cap \{Y = y\}) = 0 \quad \text{et} \quad \mathbb{P}(\{X = x\}) \neq 0, \mathbb{P}(\{Y = y\}) \neq 0$$

V.1.b) Utilisation des probabilités conditionnelles pour démontrer l'indépendance

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

Soient (X, Y) un couple de v.a. discrètes.

1) Supposons : $\forall x \in X(\Omega), \mathbb{P}(\{X = x\}) \neq 0$. Alors :

$$\begin{aligned} &X \text{ et } Y \text{ sont indépendantes} \\ \Leftrightarrow &\forall x \in X(\Omega), \forall y \in Y(\Omega), \mathbb{P}_{\{X=x\}}(\{Y = y\}) = \mathbb{P}(\{Y = y\}) \end{aligned}$$

2) Supposons : $\forall y \in Y(\Omega), \mathbb{P}(\{Y = y\}) \neq 0$. Alors :

$$\begin{aligned} &X \text{ et } Y \text{ sont indépendantes} \\ \Leftrightarrow &\forall x \in X(\Omega), \forall y \in Y(\Omega), \mathbb{P}_{\{Y=y\}}(\{X = x\}) = \mathbb{P}(\{X = x\}) \end{aligned}$$

V.2. Indépendance mutuelle de variables aléatoires discrètes

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

Soient $X_1, X_2, \dots, X_n$ (avec $n \geq 2$) des v.a. discrètes.

- Les v.a. $X_1, X_2, \dots, X_n$ sont **(mutuellement) indépendantes** (pour la probabilité $\mathbb{P}$) lorsque :

$$\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega),$$

$$\mathbb{P} \left(\bigcap_{i=1}^n \{X_i = x_i\} \right) = \prod_{i=1}^n \mathbb{P}(\{X_i = x_i\})$$

- Soit $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$.
La suite $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables (mutuellement) indépendantes (pour la probabilité $\mathbb{P}$) lorsque :

Pour tout $n \geq 2$, les variables $X_1, X_2, \dots, X_n$ sont mutuellement indépendantes

ou, de manière équivalente, si pour tout J partie **finie** de $\mathbb{N}^*$ ($J \subset \mathbb{N}^*$) :

$$\forall (x_j)_{j \in J} \in \prod_{j \in J} X_j(\Omega), \mathbb{P} \left(\bigcap_{j \in J} \{X_j = x_j\} \right) = \prod_{j \in J} \mathbb{P}(\{X_j = x_j\})$$

- On dit que la suite $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires indépendantes et identiquement distribuées (i.i.d.) si :
 - × les variables de la suite $(X_n)_{n \in \mathbb{N}^*}$ sont indépendantes,
 - × les variables de la suite $(X_n)_{n \in \mathbb{N}^*}$ sont toutes de même loi.

V.3. Lemme des coalitions

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

Soit (X, Y) un couple de v.a. discrètes à valeurs dans un ensemble E .

Soit F un ensemble.

Soient $f : X(\Omega) \rightarrow F$ et $g : Y(\Omega) \rightarrow F$ deux fonctions.

- Cas de 2 v.a.

$$X \perp\!\!\!\perp Y \Rightarrow f(X) \perp\!\!\!\perp g(Y)$$

- Généralisation à n v.a.

Soient $X_1, \dots, X_n$ des v.a. discrètes.

Soient $f_1 : X_1(\Omega) \rightarrow F, \dots, f_n : X_n(\Omega) \rightarrow F$ des fonctions.

$$\begin{array}{l} \text{Les v.a. discrètes } X_1, \dots, X_n \\ \text{sont mutuellement indépendantes} \end{array} \Rightarrow \begin{array}{l} \text{Les v.a. discrètes } f_1(X_1), \dots, f_n(X_n) \\ \text{sont mutuellement indépendantes} \end{array}$$

$$\begin{array}{l} \text{Les v.a. discrètes } X_1, \dots, X_n \\ \text{sont mutuellement indépendantes} \end{array} \Rightarrow \begin{array}{l} \text{Toute v.a. s'exprimant en fonction des} \\ \text{v.a. } X_1, \dots, X_p \text{ est indépendante de} \\ \text{toute v.a. s'exprimant en fonction des} \\ \text{v.a. } X_{p+1}, \dots, X_n \text{ (pour } p \in \llbracket 2, n-1 \rrbracket) \end{array}$$

VI. Notion de couple de v.a.

VI.1. Notion de couple de v.a.

Soient E et F deux ensembles (*on considère souvent $E = \mathbb{K}$ et $F = \mathbb{K}$*).

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

- On dit que X est une **variable aléatoire discrètes à valeurs dans E** et définie sur $(\Omega, \mathcal{A})$ si :
 - (i) X est une application de Ω dans E ($X : \Omega \rightarrow E$).
 - (ii) $\forall x \in E, X^{-1}(\{x\}) = \{X = x\} = \{\omega \in \Omega \mid X(\omega) = x\} \in \mathcal{A}$.
 - (iii) L'ensemble $X(\Omega)$ est au plus dénombrable.
(*l'ensemble image $X(\Omega)$ est, par définition, l'ensemble des valeurs prise par l'application X*)
- On dit qu'une variable aléatoire Z est **couple de variables aléatoires discrètes à valeurs dans $E \times F$** et définie sur $(\Omega, \mathcal{A})$ si :
 - $\times$ Z s'écrit sous la forme $Z = (X, Y)$,
 - $\times$ X est une variable aléatoire discrète à valeurs dans E ,
 Y est une variable aléatoire discrète à valeurs dans F .
- En particulier, un couple de variable aléatoires discrètes est une variable aléatoire discrète à valeur dans un produit.

$$\begin{aligned} Z : \Omega &\rightarrow E \times F \\ \omega &\mapsto (X(\omega), Y(\omega)) \end{aligned}$$

VI.2. Loi d'un couple de v.a. discrètes, ou loi conjointe

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

Soit (X, Y) un couple de v.a. discrètes.

- On appelle **loi de probabilité du couple (X, Y)** ou **loi conjointe des v.a. X et Y** , la donnée de l'ensemble des valeurs $\mathbb{P}(\{X = x\} \cap \{Y = y\})$ pour x décrivant $X(\Omega)$ et y décrivant $Y(\Omega)$ (c'est-à-dire (x, y) décrivant $X(\Omega) \times Y(\Omega)$).
Autrement dit, la loi de X est la famille :

$$\left(\mathbb{P}(\{X = x\} \cap \{Y = y\}) \right)_{(x,y) \in X(\Omega) \times Y(\Omega)}$$

- On note au passage que l'événement $\{X = x\} \cap \{Y = y\}$ peut aussi s'écrire :
 - $\times \{X = x, Y = y\}$.
 - $\times \{(X, Y) = (x, y)\}$.

MÉTHODO

Déterminer la loi d'un couple (X, Y)

Afin de déterminer la loi du couple (X, Y) , on commence **TOUJOURS** par déterminer les ensembles image $X(\Omega)$ et $Y(\Omega)$ des v.a. X et Y .

VI.3. Lois conditionnelles

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

Soit (X, Y) un couple de v.a. discrètes.

- 1) • Pour tout $y \in Y(\Omega)$ tel que $\mathbb{P}(\{Y = y\}) \neq 0$, on appelle **loi conditionnelle de X sachant (que l'événement) $\{Y = y\}$** (est réalisé) l'application :

$$\begin{aligned} X(\Omega) &\rightarrow \mathbb{R} \\ x &\mapsto \mathbb{P}_{\{Y=y\}}(\{X = x\}) = \frac{\mathbb{P}(\{X = x\} \cap \{Y = y\})}{\mathbb{P}(\{Y = y\})} \end{aligned}$$

- Autrement dit, la loi conditionnelle de X sachant $\{Y = y\}$ est la donnée la donnée de l'ensemble des valeurs $\mathbb{P}_{\{Y=y\}}(\{X = x\})$ pour x décrivant $X(\Omega)$. Ou encore, la loi conditionnelle de X sachant l'événement $\{Y = y\}$ est la famille :

$$\left(\mathbb{P}_{\{Y=y\}}(\{X = x\}) \right)_{x \in X(\Omega)}$$

- 2) • Pour tout $x \in X(\Omega)$ tel que $\mathbb{P}(\{X = x\}) \neq 0$, on appelle **loi conditionnelle de Y sachant (que l'événement) $\{X = x\}$** (est réalisé) l'application :

$$\begin{aligned} Y(\Omega) &\rightarrow \mathbb{R} \\ y &\mapsto \mathbb{P}_{\{X=x\}}(\{Y = y\}) = \frac{\mathbb{P}(\{X = x\} \cap \{Y = y\})}{\mathbb{P}(\{X = x\})} \end{aligned}$$

- Autrement dit, la loi conditionnelle de Y sachant $\{X = x\}$ est la donnée la donnée de l'ensemble des valeurs $\mathbb{P}_{\{X=x\}}(\{Y = y\})$ pour y décrivant $Y(\Omega)$. Ou encore, la loi conditionnelle de Y sachant l'événement $\{X = x\}$ est la famille :

$$\left(\mathbb{P}_{\{X=x\}}(\{Y = y\}) \right)_{y \in Y(\Omega)}$$

MÉTHODO

Lien entre lois conditionnelles et loi conjointe via la loi d'une des v.a.

Soit $x_0 \in X(\Omega)$ tel que : $\mathbb{P}(\{X = x_0\}) \neq 0$.

Par définition : $\mathbb{P}_{\{X=x_0\}}(\{Y = y\}) = \frac{\mathbb{P}(\{X = x_0\} \cap \{Y = y\})}{\mathbb{P}(\{X = x_0\})}$

- Ainsi, si on connaît :

- × les valeurs $\mathbb{P}(\{X = x_0\} \cap \{Y = y\})$ pour tout $y \in Y(\Omega)$,
(c'est notamment le cas si on connaît la loi du couple (X, Y))
- × la valeur de $\mathbb{P}(\{X = x_0\})$,
(c'est notamment le cas si on connaît la loi de X)

alors on obtient la loi conditionnelle de Y sachant $\{X = x_0\}$.

- On peut aussi lire l'égalité dans l'autre sens. Si on connaît :

- × la valeur $\mathbb{P}(\{X = x_0\})$,
- × les valeurs $\mathbb{P}_{\{X=x_0\}}(\{Y = y\})$ pour tout $y \in Y(\Omega)$,
(c'est-à-dire la loi conditionnelle de Y sachant $\{X = x_0\}$)

alors on obtient les valeurs $\mathbb{P}(\{X = x_0\} \cap \{Y = y\})$ pour tout $y \in Y(\Omega)$.

VI.4. Lois marginales

VI.4.a) Définition

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

Soit (X, Y) un couple de v.a. discrètes.

- On appelle **1^{ère} loi marginale du couple** (X, Y) la loi de la v.a. X .
- On appelle **2^{ème} loi marginale du couple** (X, Y) la loi de la v.a. Y .

VI.4.b) Expression d'une loi marginale via la loi du couple ou via une loi conditionnelle

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

Soit (X, Y) un couple de v.a. discrètes.

1) Loi de X via la loi du couple (X, Y)

$$\forall x \in X(\Omega), \quad \mathbb{P}(\{X = x\}) = \sum_{y \in Y(\Omega)} \mathbb{P}(\{Y = y\} \cap \{X = x\})$$

Loi de X via les lois conditionnelles de X sachant $\{Y = y\}$ pour tout $y \in Y(\Omega)$:

On suppose : $\forall y \in Y(\Omega), \mathbb{P}(\{Y = y\}) \neq 0$. Alors :

$$\mathbb{P}(\{X = x\}) = \sum_{y \in Y(\Omega)} \mathbb{P}(\{Y = y\}) \times \mathbb{P}_{\{Y=y\}}(\{X = x\})$$

2) Loi de Y via la loi du couple (X, Y)

$$\forall y \in Y(\Omega), \quad \mathbb{P}(\{Y = y\}) = \sum_{x \in X(\Omega)} \mathbb{P}(\{X = x\} \cap \{Y = y\})$$

Loi de Y via les lois conditionnelles de Y sachant $\{X = x\}$ pour tout $x \in X(\Omega)$:

On suppose : $\forall x \in X(\Omega), \mathbb{P}(\{X = x\}) \neq 0$. Alors :

$$\mathbb{P}(\{Y = y\}) = \sum_{x \in X(\Omega)} \mathbb{P}(\{X = x\}) \times \mathbb{P}_{\{X=x\}}(\{Y = y\})$$

VII. Opérations sur les v.a. discrètes

VII.1. Cas général : loi de $g(X, Y)$

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

Soit (X, Y) un couple de v.a. discrètes.

Soit $g : X(\Omega) \times Y(\Omega) \rightarrow E$.

- L'application Z est une v.a. discrète.
- L'ensemble des valeurs prises par $Z = g(X, Y)$ est donnée par :

$$\begin{aligned} Z(\Omega) &= \{g(X(\omega), Y(\omega)) \mid \omega \in \Omega\} \\ &\subseteq \{g(x, y) \mid x \in X(\Omega), y \in Y(\Omega)\} \end{aligned}$$

- La loi de $Z = g(X, Y)$ est donnée par :

$$\forall z \in Z(\Omega), \mathbb{P}(\{Z = z\}) = \sum_{\substack{(x, y) \in X(\Omega) \times Y(\Omega) \\ z = g(x, y)}} \mathbb{P}(\{X = x\} \cap \{Y = y\})$$

- En pratique, on optera pour l'une des deux rédactions suivantes.

1) La famille $(\{X = x\})_{x \in X(\Omega)}$ est un sce.

Soit $z \in Z(\Omega)$. D'après la formule des probabilités totales :

$$\begin{aligned} \mathbb{P}(\{Z = z\}) &= \sum_{x \in X(\Omega)} \mathbb{P}(\{X = x\} \cap \{g(X, Y) = z\}) \\ &= \sum_{x \in X(\Omega)} \mathbb{P}(\{X = x\} \cap \{g(x, Y) = z\}) \end{aligned}$$

2) La famille $(\{Y = y\})_{y \in Y(\Omega)}$ est un sce.

Soit $z \in Z(\Omega)$. D'après la formule des probabilités totales :

$$\begin{aligned} \mathbb{P}(\{Z = z\}) &= \sum_{y \in Y(\Omega)} \mathbb{P}(\{Y = y\} \cap \{g(X, Y) = z\}) \\ &= \sum_{y \in Y(\Omega)} \mathbb{P}(\{Y = y\} \cap \{g(X, y) = z\}) \end{aligned}$$

Le choix de l'introduction du sce $(\{Y = y\})_{y \in Y(\Omega)}$ ou $(\{X = x\})_{x \in X(\Omega)}$ est guidé par les lois de X et Y . On optera toujours pour le sce le plus simple.

VII.2. Loi de la somme de deux v.a. discrètes

VII.2.a) Cas général

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

Soit (X, Y) un couple de v.a. discrètes à valeurs dans $\mathbb{C}$.

- La v.a. $X + Y$ est une v.a. discrète.
- On détermine la loi de $X + Y$ à l'aide d'une des deux rédactions suivantes.

1) La famille $(\{X = x\})_{x \in X(\Omega)}$ est un sce.

Soit $z \in (X + Y)(\Omega)$. D'après la formule des probabilités totales :

$$\begin{aligned} \mathbb{P}(\{X + Y = z\}) &= \sum_{x \in X(\Omega)} \mathbb{P}(\{X = x\} \cap \{X + Y = z\}) \\ &= \sum_{\substack{x \in X(\Omega) \\ z - x \in Y(\Omega)}} \mathbb{P}(\{X = x\} \cap \{Y = z - x\}) \end{aligned}$$

2) La famille $(\{Y = y\})_{y \in Y(\Omega)}$ est un sce.

Soit $z \in (X + Y)(\Omega)$. D'après la formule des probabilités totales :

$$\begin{aligned} \mathbb{P}(\{X + Y = z\}) &= \sum_{y \in Y(\Omega)} \mathbb{P}(\{Y = y\} \cap \{X + Y = z\}) \\ &= \sum_{\substack{y \in Y(\Omega) \\ z - y \in X(\Omega)}} \mathbb{P}(\{Y = y\} \cap \{X = z - y\}) \end{aligned}$$

Le choix de l'introduction du sce $(\{Y = y\})_{y \in Y(\Omega)}$ ou $(\{X = x\})_{x \in X(\Omega)}$ est guidé par les lois de X et Y . On optera toujours pour le sce le plus simple.

À RETENIR

- On note que la probabilité $\mathbb{P}(\{X = x\} \cap \{Y = z - x\})$ est nulle dès que $z - x$ n'appartient pas à $Y(\Omega)$ puisqu'alors $\{Y = z - x\} = \emptyset$.
- Ceci a pour conséquence de restreindre les indices de sommation puisque :

$$\mathbb{P}(\{X = x\} \cap \{Y = z - x\}) = \mathbb{P}(\emptyset) = 0 \quad \text{dès que} \quad z - x \notin Y(\Omega)$$

MÉTHODO

Trouver les indices convenables lors de la détermination de la loi de $X + Y$

Il est fréquent, lors de l'étape de restriction des indices de la somme de tomber sur une contrainte s'exprimant par deux inégalités :

$$a \leq i \leq b \quad \text{ET} \quad c \leq i \leq d \quad \text{où} \quad (a, b, c, d) \in (\mathbb{R})^4$$

On pourra alors utiliser le résultat suivant :

$$(a \leq i \leq b \quad \text{ET} \quad c \leq i \leq d) \quad \Leftrightarrow \quad \max(a, c) \leq i \leq \min(b, d)$$

(à gauche c'est le plus grand élément qui contraint le plus, à droite c'est le plus petit élément qui contraint le plus)

VII.2.b) Cas des v.a. discrètes à valeurs entières

Soit (X, Y) un couple de v.a. discrètes à **valeurs entières**.

On suppose que les v.a. X et Y sont indépendantes.

$$\forall t \in]-1, 1[, G_{X+Y}(t) = G_X(t) \times G_Y(t)$$

Généralisation

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

On suppose que les variables aléatoires de la suite $(X_n)_{n \in \mathbb{N}^*}$ sont indépendantes.

Pour tout $n \in \mathbb{N}^*$, on note $S_n = \sum_{k=1}^n X_k$.

$$1. \quad \forall n \in \mathbb{N}^*, X_{n+1} \perp\!\!\!\perp S_n$$

$$2. \quad \forall n \in \mathbb{N}^*, G_{X_1+\dots+X_n} = \prod_{i=1}^n G_{X_i}$$

VII.2.c) Application : stabilité des lois usuelles

1) Stabilité des lois binomiales

Soit $p \in]0, 1[$ et soient $m \in \mathbb{N}^*$ et $n \in \mathbb{N}^*$.

$$\begin{aligned} & \bullet X \sim \mathcal{B}(m, p) \text{ et } Y \sim \mathcal{B}(n, p) \\ & \bullet X \perp\!\!\!\perp Y \end{aligned} \quad \Rightarrow \quad X + Y \sim \mathcal{B}(m + n, p)$$

2) Stabilité des lois de Poisson

Soient $\lambda \in \mathbb{R}_+^*$ et $\mu \in \mathbb{R}_+^*$.

$$\begin{aligned} & \bullet X \sim \mathcal{P}(\lambda) \text{ et } Y \sim \mathcal{P}(\mu) \\ & \bullet X \perp\!\!\!\perp Y \end{aligned} \quad \Rightarrow \quad X + Y \sim \mathcal{P}(\lambda + \mu)$$

Démonstration.

1) Rappelons tout d'abord que si $X \sim \mathcal{B}(m, p)$ alors :

$$\begin{aligned} \forall t \in]-\infty, +\infty[, G_X(t) &= \sum_{k=0}^m \mathbb{P}(\{X = k\}) t^k \\ &= \sum_{k=0}^m \binom{m}{k} p^k (1-p)^{m-k} t^k \\ &= \sum_{k=0}^m \binom{m}{k} (pt)^k (1-p)^{m-k} \\ &= ((1-p) + pt)^m \end{aligned}$$

Or, d'après le théorème précédent, pour tout $t \in]-\infty, +\infty[$:

$$\begin{aligned} G_{X+Y}(t) &= G_X(t) \times G_Y(t) && (\text{car } X \perp\!\!\!\perp Y) \\ &= ((1-p) + pt)^m \times ((1-p) + pt)^n \\ &= ((1-p) + pt)^{m+n} \end{aligned}$$

On reconnaît la fonction génératrice d'une v.a.r. qui suit la loi $\mathcal{B}(m+n, p)$.

Or, la fonction génératrice caractérise la loi d'une variable aléatoire.

On en déduit : $X + Y \sim \mathcal{B}(m+n, p)$.

2) Rappelons tout d'abord que si $X \sim \mathcal{P}(\lambda)$ alors :

$$\begin{aligned} \forall t \in]-\infty, +\infty[, G_X(t) &= \sum_{k=0}^{+\infty} \mathbb{P}(\{X = k\}) t^k \\ &= \sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} t^k \\ &= e^{-\lambda} \sum_{k=0}^{+\infty} \frac{(\lambda t)^k}{k!} \\ &= e^{-\lambda} e^{\lambda t} \end{aligned}$$

Or, d'après le théorème précédent, pour tout $t \in]-\infty, +\infty[$:

$$\begin{aligned} G_{X+Y}(t) &= G_X(t) \times G_Y(t) \quad (\text{car } X \perp\!\!\!\perp Y) \\ &= e^{\lambda(t-1)} \times e^{\mu(t-1)} \\ &= e^{(\lambda+\mu)(t-1)} \end{aligned}$$

On reconnaît la fonction génératrice d'une v.a.r. qui suit la loi $\mathcal{P}(\lambda + \mu)$.

Or, la fonction génératrice caractérise la loi d'une variable aléatoire.

On en déduit : $X + Y \sim \mathcal{P}(\lambda + \mu)$. □

VII.3. Loi du maximum / minimum de deux v.a. discrètes réelles indépendantes

VII.3.a) Cas général

Soient X et Y deux v.a. discrètes à valeurs réelles.

On suppose que X et Y sont **indépendantes**.

On définit les v.a. $\min(X, Y)$ et $\max(X, Y)$ comme suit :

$$\begin{array}{ll} \min(X, Y) : \Omega \rightarrow \mathbb{R} & \max(X, Y) : \Omega \rightarrow \mathbb{R} \\ \omega \mapsto \min(X(\omega), Y(\omega)) & \omega \mapsto \max(X(\omega), Y(\omega)) \end{array}$$

1) Remarquons tout d'abord :

$$\forall t \in \mathbb{R}, \{\min(X, Y) > t\} = \{X > t\} \cap \{Y > t\}$$

$$\forall t \in \mathbb{R}, \{\max(X, Y) \leq t\} = \{X \leq t\} \cap \{Y \leq t\}$$

2) a) La v.a. $\min(X, Y)$ est une v.a. discrète.

b) Par ailleurs, pour tout $t \in \mathbb{R}$:

$$\mathbb{P}(\{\min(X, Y) > t\}) = (1 - \mathbb{P}(\{X \leq t\})) (1 - \mathbb{P}(\{Y \leq t\}))$$

3) a) La v.a. $\max(X, Y)$ est une v.a. discrète.

b) Par ailleurs, pour tout $t \in \mathbb{R}$:

$$\mathbb{P}(\{\max(X, Y) \leq t\}) = \mathbb{P}(\{X \leq t\}) \mathbb{P}(\{Y \leq t\})$$

VII.3.b) Loi du minimum de deux v.a. indépendantes de lois géométriques

Soit X une v.a. discrète à valeurs réelles.

Soit $p \in]0, 1[$.

- Préambule

$$\begin{aligned} \times & X \text{ est une v.a.r. à valeurs entières} \\ \times & \forall k \in \mathbb{N}, \mathbb{P}(\{X > k\}) = (1 - p)^k \quad \Leftrightarrow \quad X \sim \mathcal{G}(p) \end{aligned}$$

- Loi du minimum de deux variables de lois géométriques

Soient $p_1 \in]0, 1[$ et $p_2 \in]0, 1[$.

Soient X et Y deux v.a. discrètes telles que : $X \sim \mathcal{G}(p_1)$ et $Y \sim \mathcal{G}(p_2)$.

$$1) \quad \forall x \in \mathbb{N}, \mathbb{P}(\{\min(X, Y) > k\}) = (1 - p_1)^k (1 - p_2)^k$$

2) On en déduit : $\min(X, Y) \sim \mathcal{G}(r)$ où $r = 1 - (1 - p_1)(1 - p_2)$.

VIII. Calculs d'espérance de variables aléatoires réelles ou complexes

VIII.1. Espérance de $Z = g(X, Y)$ par théorème de transfert

Soit (X, Y) un couple de v.a. discrètes à valeurs dans $\mathbb{K}$.

Soit $g : X(\Omega) \times Y(\Omega) \rightarrow \mathbb{K}$ une application.

Notons $Z = g(X, Y)$.

$$\begin{aligned} 1) \quad & \begin{aligned} & \text{La v.a. } Z \text{ est} \\ & \text{d'espérance finie} \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} & \text{La famille } \left(z \mathbb{P}(\{Z = z\}) \right) \text{ est sommable} \\ & \Leftrightarrow \quad \text{La famille } \left(g(x, y) \mathbb{P}(\{(X, Y) = (x, y)\}) \right) \text{ est sommable} \end{aligned} \end{aligned}$$

2) Dans ce cas :

$$\begin{aligned} \mathbb{E}(Z) &= \sum_{z \in Z(\Omega)} z \mathbb{P}(\{Z = z\}) \\ &= \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} g(x, y) \mathbb{P}(\{X = x\} \cap \{Y = y\}) \\ &= \sum_{x \in X(\Omega)} \left(\sum_{y \in Y(\Omega)} g(x, y) \mathbb{P}(\{X = x\} \cap \{Y = y\}) \right) \\ &= \sum_{y \in Y(\Omega)} \left(\sum_{x \in X(\Omega)} g(x, y) \mathbb{P}(\{X = x\} \cap \{Y = y\}) \right) \end{aligned}$$

VIII.2. Espérance d'une somme

Soit $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$.

Soient $X_1, \dots, X_n$ des v.a. discrètes admettant une espérance.

1) Alors $\lambda_1 \cdot X_1 + \dots + \lambda_n \cdot X_n$ admet une espérance.

2) De plus :

$$\mathbb{E} \left(\sum_{i=1}^n \lambda_i \cdot X_i \right) = \sum_{i=1}^n \lambda_i \mathbb{E}(X_i)$$

En particulier, lorsque pour tout $i \in \llbracket 1, n \rrbracket$, $\lambda_i = 1$, on obtient :

$$\mathbb{E} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \mathbb{E}(X_i)$$

MÉTHODO

Démontrer qu'une v.a. admet une espérance

On retiendra que si une v.a. s'écrit comme combinaison linéaire (resp. somme) de v.a. qui admettent une espérance, alors cette v.a. admet une espérance.

VIII.3. Espérance d'un produit

a) Formulation générale

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

Soit (X, Y) un couple de v.a. discrètes à valeurs dans $\mathbb{K}$.

On suppose que X et Y sont de moments d'ordre 2 finis.

1) Alors XY est d'espérance finie.

2) De plus :

$$\begin{aligned} \mathbb{E}(XY) &= \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xy \mathbb{P}(\{X=x\} \cap \{Y=y\}) \\ &= \sum_{x \in X(\Omega)} \left(\sum_{y \in Y(\Omega)} xy \mathbb{P}(\{X=x\} \cap \{Y=y\}) \right) \\ &= \sum_{y \in Y(\Omega)} \left(\sum_{x \in X(\Omega)} xy \mathbb{P}(\{X=x\} \cap \{Y=y\}) \right) \end{aligned}$$

b) Cas des v.a. indépendantes

Soit (X, Y) un couple de v.a. discrètes.

On suppose :
 $\times$ X et Y sont d'espérances finies.
 $\times$ X et Y sont indépendantes.

Alors XY est d'espérance finie et :

$$\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y)$$

IX. Calculs de variance et covariance de v.a. discrètes réelles

IX.1. Covariance de v.a. discrètes réelles

Soit (X, Y) un couple de v.a. discrètes réelles.

On suppose que X et Y sont de moment d'ordre 2 finis.

La **covariance** de X et Y est définie par :

$$\text{Cov}(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X)) (Y - \mathbb{E}(Y))\right)$$

IX.1.a) Calcul en pratique

Soit (X, Y) un couple de v.a. discrètes réelles.

On suppose que X et Y sont de moments d'ordre 2 finis.

Alors : $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X) \mathbb{E}(Y)$

En particulier, on a : $\text{Cov}(X, X) = \mathbb{V}(X)$

IX.1.b) Propriétés de la covariance

Soient X, Y, X_i, Y_i des v.a. discrètes.

Supposons que ces v.a. sont de moments d'ordre 2 finis.

L'opérateur de covariance vérifie les propriétés suivantes.

1) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

(propriété de symétrie)

2) $\forall (\lambda, \mu) \in \mathbb{R}^2, \text{Cov}(\lambda X_1 + \mu X_2, Y) = \lambda \text{Cov}(X_1, Y) + \mu \text{Cov}(X_2, Y)$

(linéarité à gauche)

3) $\forall (\lambda, \mu) \in \mathbb{R}^2, \text{Cov}(X, \lambda Y_1 + \mu Y_2) = \lambda \text{Cov}(X, Y_1) + \mu \text{Cov}(X, Y_2)$

(linéarité à droite)

4) $\forall a \in \mathbb{R}, \text{Cov}(X, a) = \text{Cov}(a, X) = 0$

IX.1.c) Une condition nécessaire d'indépendance

Soient X et Y deux v.a. (discrètes).

On suppose que X et Y sont de moments d'ordre 2 finis.

$$X \text{ et } Y \text{ indépendantes} \Rightarrow \text{Cov}(X, Y) = 0$$

IX.2. Variance d'une somme de v.a. discrètes réelles

IX.2.a) Cas général

Soit (X, Y) un couple de v.a. discrètes réelles.

On suppose que X et Y sont de moments d'ordre 2 finis.

(c'est-à-dire que X et Y sont de variances finies)

1) Alors $X + Y$ est de variance finie.

2) De plus : $\mathbb{V}(X + Y) = \mathbb{V}(X) + 2 \operatorname{Cov}(X, Y) + \mathbb{V}(Y)$

Généralisation

Soient $X_1, \dots, X_n$ des v.a. discrètes réelles.

On suppose que les v.a. $X_1, \dots, X_n$ sont d'espérances finies.

1) Alors $X_1 + \dots + X_n$ est de variance finie.

2) De plus :

$$\begin{aligned} \mathbb{V}(X_1 + \dots + X_n) &= \sum_{\substack{1 \leq i, j \leq n \\ i=j}} \operatorname{Cov}(X_i, X_j) + \sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} \operatorname{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \mathbb{V}(X_i) + \sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} \operatorname{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \mathbb{V}(X_i) + 2 \sum_{1 \leq i < j \leq n} \operatorname{Cov}(X_i, X_j) \end{aligned}$$

IX.2.b) Variance d'une somme de v.a. indépendantes

Soit (X, Y) un couple de v.a. discrètes réelles.

On suppose que X et Y sont de moments d'ordre 2 finis.

(c'est-à-dire que X et Y sont de variances finies)

1) Alors la v.a. $X + Y$ est de variance finie. De plus :

$$X \text{ et } Y \text{ indépendantes} \Rightarrow \mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$$

2) Généralisation

On suppose :

- × $X_1, \dots, X_n$ sont de variances finies.
- × $X_1, \dots, X_n$ mutuellement indépendantes.

Alors $X_1 + \dots + X_n$ est de variance finie et :

$$\mathbb{V}(X_1 + \dots + X_n) = \mathbb{V}(X_1) + \dots + \mathbb{V}(X_n)$$

X. Inégalités de Cauchy-Schwarz pour les variables aléatoires réelles

X.1. Les semi-produit scalaires sur l'ensemble des variables aléatoires (discrètes)

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

- 1) L'ensemble des variables aléatoires à valeurs dans $\mathbb{R}$ est un sous-espace vectoriel de $\mathcal{F}(\Omega, \mathbb{R})$.

L'ensemble des variables aléatoires discrètes est un sous-espace vectoriel de l'ensemble des variables aléatoires.

L'ensemble des variables aléatoires (discrètes) d'espérance finie est un sous-espace vectoriel de l'ensemble des variables aléatoires (discrètes).

L'ensemble des variables aléatoires (discrètes) de moment d'ordre 2 fini est un sous-espace vectoriel de l'ensemble des variables aléatoires (discrètes) d'espérance finie. Notons $\mathcal{L}^2(\Omega)$ ce dernier espace vectoriel.

- 2) Les applications $\langle \cdot, \cdot \rangle_{s_1}$ et $\langle \cdot, \cdot \rangle_{s_2}$ définies par :

$$\begin{aligned} \langle \cdot, \cdot \rangle_{s_1} : (\quad X \quad, \quad Y \quad) &\mapsto \mathbb{E}(XY) \\ \mathcal{L}^2(\Omega) \times \mathcal{L}^2(\Omega) &\rightarrow \mathbb{R} \end{aligned}$$

$$\begin{aligned} \langle \cdot, \cdot \rangle_{s_2} : (\quad X \quad, \quad Y \quad) &\mapsto \text{Cov}(X, Y) \\ \mathcal{L}^2(\Omega) \times \mathcal{L}^2(\Omega) &\rightarrow \mathbb{R} \end{aligned}$$

sont :

- × bilinéaires,
- × symétriques,
- × ~~définies~~-positives.

De telles applications sont appelées des semi-produit scalaires. Elles permettent de définir des semi-normes :

$$\forall X \in \mathcal{L}^2(\Omega), \|X\|_{s_1} = \sqrt{\langle X, X \rangle_{s_1}} \quad \text{et} \quad \|X\|_{s_2} = \sqrt{\langle X, X \rangle_{s_2}}$$

Une semi-norme est une application $\mathcal{L}^2(\Omega) \rightarrow \mathbb{R}$ homogène qui vérifie les propriétés de ~~séparation~~, d'homogénéité et d'inégalité triangulaire.

X.2. Énoncés des inégalités de Cauchy-Schwarz

Énoncé dans le cas général

Soit E un $\mathbb{R}$ -espace vectoriel muni d'un semi-produit scalaire $\langle \cdot, \cdot \rangle_s$.

- $\forall (x, y) \in E^2, |\langle x, y \rangle_s| \leq \|x\|_s \|y\|_s$ *(inégalité de Cauchy-Schwarz)*

Ce qu'on peut aussi écrire :

$$\forall (x, y) \in E^2, \langle x, y \rangle_s \times \langle x, y \rangle_s \leq \langle x, x \rangle_s \times \langle y, y \rangle_s$$

- On peut caractériser le cas d'égalité. Pour tout $(x, y) \in E^2$:

$$|\langle x, y \rangle_s| = \|x\|_s \|y\|_s \Leftrightarrow \exists ! \lambda \in \mathbb{R}, \|\lambda \cdot x + y\|_s = 0$$

Inégalités de Cauchy-Schwarz probabilistes

$$1) \quad \forall (X, Y) \in \left(\mathcal{L}^2(\Omega) \right)^2, \quad \left(\mathbb{E}(XY) \right)^2 \leq \mathbb{E}(X^2) \mathbb{E}(Y^2)$$

On peut caractériser le cas d'égalité. Pour tout $(X, Y) \in \left(\mathcal{L}^2(\Omega) \right)^2$:

$$\left(\mathbb{E}(XY) \right)^2 = \mathbb{E}(X^2) \mathbb{E}(Y^2) \Leftrightarrow \text{Les v.a. } X \text{ et } Y \text{ sont presque sûrement colinéaires}$$

$$2) \quad \forall (X, Y) \in \left(\mathcal{L}^2(\Omega) \right)^2, \quad \left(\text{Cov}(X, Y) \right)^2 \leq \text{Cov}(X, X) \text{Cov}(Y, Y)$$

On peut caractériser le cas d'égalité. Pour tout $(X, Y) \in \left(\mathcal{L}^2(\Omega) \right)^2$:

$$\left(\text{Cov}(X, Y) \right)^2 = \text{Cov}(X, X) \text{Cov}(Y, Y) \Leftrightarrow \text{Une des v.a. est presque sûrement une fonction affine de l'autre}$$

À RETENIR

(détermination de la loi d'un couple de v.a. discrètes)

- La loi du couple (X, Y) n'est **JAMAIS** obtenue à l'aide de la formule des probabilités totales.
- Pour obtenir la loi du couple (X, Y) on opère **TOUJOURS** comme suit.

a) Détermination de $X(\Omega)$ et de $Y(\Omega)$.

On pourra se contenter dans cet ensemble de déterminer un sur-ensemble de ces deux ensembles image $(X(\Omega) \subset I \text{ et } Y(\Omega) \subset J)$.

b) On rédige alors comme suit.

Soit $(x, y) \in X(\Omega) \times Y(\Omega)$ (ou plutôt $(x, y) \in I \times J$).

L'événement $\{X = x\} \cap \{Y = y\}$ est réalisé

$\Leftrightarrow$ L'événement $\{X = x\}$ est réalisé et l'événement $\{Y = y\}$ est réalisé

$\Leftrightarrow \dots$ et $\dots$

Cette rédaction ne permet pas de conclure quant à la loi du couple (X, Y) .

Elle sert simplement à s'assurer de la bonne compréhension de l'expérience aléatoire décrite dans l'énoncé.

c) Il y a alors deux grandes manières de procéder.

(i) Décomposition de l'événement $\{X = x\} \cap \{Y = y\}$

On utilise en priorité cette méthode. Pour ce faire, on pensera à introduire des événements basiques liés à l'expérience :

- pour tout $k \in \mathbb{N}^*$, P_k : « on obtient Pile lors du $k^{\text{ème}}$ lancer »
- pour tout $k \in \mathbb{N}^*$, B_k : « on obtient une boule blanche lors du $k^{\text{ème}}$ tirage »
- ...

Il est à noter que l'événement $\{X = x\} \cap \{Y = y\}$ peut aussi être décomposé à l'aide d'événements dépendant d'autres v.a.

(ii) Méthode par dénombrement

Cette méthode peut être utilisée lorsque :

- l'expérience consiste à effectuer un nombre fini n d'épreuves.
(en réalité, on peut aussi étendre cette méthode aux expériences qui procèdent à un nombre infini d'épreuves pour peu que la réalisation de l'événement $\{X = x\} \cap \{Y = y\}$ ne dépende que d'un nombre fini (notons-le n) d'épreuves de l'expérience, les résultats des épreuves ultérieures étant libres.)
- il y a équiprobabilité de toutes les issues de l'expérience (tous les n -tirages ou n -lancers ont même probabilité d'apparaître).

On rédige alors comme suit.

L'événement $\{X = x\} \cap \{Y = y\}$ est réalisé par tous les n -tirages (ou n -lancers suivant l'expérience) qui contiennent ...

Un tel n -lancer est entièrement déterminé par :

- le(s) numéro(s) apparaissant en position ... : ... possibilités.
- le(s) numéro(s) apparaissant en position ... : ... possibilités.
- ...

Le contexte oriente naturellement vers l'une ou l'autre de ces deux méthodes.

- Il est enfin à noter que la loi d'un couple (X, Y) peut aussi être obtenue à l'aide :
 - × de la loi de X et des lois conditionnelles de Y sachant l'événement $\{X = x\}$ pour tout $x \in X(\Omega)$.
Il suffit alors simplement d'écrire que pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$:

$$\mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(\{X = x\}) \times \mathbb{P}_{\{X=x\}}(\{Y = y\}) \quad (*)$$

- × de la loi de Y et des lois conditionnelles de X sachant l'événement $\{Y = y\}$ pour tout $y \in Y(\Omega)$.
Il suffit alors simplement d'écrire que pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$:

$$\mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(\{Y = y\} \cap \{X = x\}) = \mathbb{P}(\{Y = y\}) \times \mathbb{P}_{\{Y=y\}}(\{X = x\}) \quad (**)$$

Il y a peu de chance que cette méthode permette de déterminer la loi du couple (X, Y) si l'énoncé n'est pas écrit dans un contexte de probabilités conditionnelles. Les égalités $(*)$ et $(**)$ ont un intérêt limité : déterminer une loi conditionnelle est aussi difficile que déterminer une loi de couple. Écrire $(*)$ ou $(**)$ n'a donc d'intérêt que si l'énoncé amène naturellement à travailler sur les lois conditionnelles.

XI. Loi faible des grands nombres

XI.1. L'inégalité de Markov

Soit X une v.a. (discrètes) à valeurs réelles.

On suppose que X est à valeurs positives ($X(\Omega) \subset [0, +\infty[$).

$$\forall a > 0, \mathbb{P}(\{X \geq a\}) \leq \frac{\mathbb{E}(X)}{a}$$

Démonstration.

Démonstration adaptée au programme

Soit $a > 0$.

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

On en déduit :

$$\begin{aligned} \mathbb{P}(\{X \geq a\}) &= \mathbb{P}\left(\bigcup_{\substack{x \in X(\Omega) \\ x \geq a}} \{X = x\}\right) \\ &= \sum_{\substack{x \in X(\Omega) \\ x \geq a}} \mathbb{P}(\{X = x\}) \end{aligned}$$

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

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

- Finalement :

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

□

Remarque

- 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 :

$$\begin{aligned} \mathbb{1}_A &: \Omega \rightarrow \mathbb{R} \\ \omega &\mapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases} \end{aligned}$$

Avec cette notation, on peut écrire : $Y = a \cdot \mathbb{1}_{\{X \geq a\}}$.

- 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 $\mathbb{P}(\{X \geq a\})$ est évidemment finie puisque majorée par 1.

XI.2. L'inégalité de Bienaymé-Tchebychev

Soit X une v.a. (discrète) à valeurs réelles.

On suppose que X est de variance finie.

$$\forall \varepsilon > 0, \mathbb{P}(\{ |X - \mathbb{E}(X)| \geq \varepsilon \}) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}$$

Démonstration.

Soit $\varepsilon > 0$. Considérons la v.a.r. $Y = (X - \mathbb{E}(X))^2$.

On est dans le cadre d'application de l'inégalité de Markov puisque $Y = (X - \mathbb{E}(X))^2$ est à valeurs positives.

En appliquant l'inégalité de Markov à la v.a.r. Y , avec $a = \varepsilon^2$:

$$\begin{aligned} \mathbb{P}(\{ Y \geq \varepsilon^2 \}) &= \mathbb{P}(\{ (X - \mathbb{E}(X))^2 \geq \varepsilon^2 \}) \leq \frac{\mathbb{E}(Y)}{\varepsilon^2} \\ &\quad \parallel \quad \parallel \\ &\mathbb{P}(\{ |X - \mathbb{E}(X)| \geq \varepsilon \}) \quad \frac{\mathbb{E}((X - \mathbb{E}(X))^2)}{\varepsilon^2} = \frac{\mathbb{V}(X)}{\varepsilon^2} \end{aligned}$$

L'inégalité de gauche est obtenue par stricte croissance de la fonction racine sur $[0, +\infty[$. □

XI.3. La loi faible des grands nombres

XI.3.a) Convergence en probabilité

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

Soit X une v.a.

- On dit que la suite de v.a.r. $(X_n)_{n \in \mathbb{N}^*}$ **converge en probabilité** vers X si :

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P}(\{ |X_n - X| \geq \varepsilon \}) = 0$$

Dans ce cas, on note : $X_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} X$.

XI.3.b) Loi faible des grands nombres

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

- × indépendantes,
- × ~~de même loi,~~
- × toutes de même espérance finie m ,
- × toutes de même variance finie σ^2 .

On pose, pour tout $n \in \mathbb{N}^*$, $\overline{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ (moyenne empirique).

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P}(\{ |\overline{X}_n - m| \geq \varepsilon \}) = 0$$

Autrement dit : $\overline{X}_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} m$.

(la suite de v.a.r. $(\overline{X}_n)_{n \in \mathbb{N}^*}$ **converge en probabilité** vers la v.a.r. certaine égale à m)

Démonstration.

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

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

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

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

× De plus :

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

× Et :

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

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

$$0 \leq \mathbb{P}(\{|\overline{X_n} - m| \geq \varepsilon\}) \leq \frac{\sigma^2}{\varepsilon^2} \frac{1}{n}$$

Or :

$$\begin{aligned} &\times 0 \xrightarrow{n \rightarrow +\infty} 0 \\ &\times \frac{\sigma^2}{\varepsilon^2} \frac{1}{n} \xrightarrow{n \rightarrow +\infty} 0 \end{aligned}$$

Ainsi, par théorème d'encadrement : $\lim_{n \rightarrow +\infty} \mathbb{P}(\{|\overline{X_n} - m| \geq \varepsilon\}) = 0$. □

Remarque

- Le programme officiel préconise l'utilisation de la loi faible des grands nombres dans le cas où $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a. (discrètes) à valeurs réelles :
 - × indépendantes,
 - × **de même loi**,
 - × de variance finie σ^2 .

Évidemment, si les variables de la suite (X_n) sont de même loi et sont de variances finies, alors elles ont toutes même espérance et même variance.

- Ce cadre est plus restrictif (les hypothèses sont plus exigeantes) mais est très classique en probabilité.
- Des variables aléatoires de même loi sont dites identiquement distribuées (i.i.d.). La loi faible des grands nombres peut donc s'appliquer à une suite de variables i.i.d de variance finie.