CH XXI : Variables aléatoires finies

I. Généralités sur les variables aléatoires discrètes

I.1. Notion de variable aléatoire

Définition

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

- On dit que X est une **variable aléatoire** RÉELLE définie sur $\big(\Omega, \mathscr{P}(\Omega)\big)$ si :
  (i) X est une application de Ω dans $\mathbb{R}$ ($X : \Omega \to \mathbb{R}$)
  (ii) $\forall x \in \mathbb{R},\ \ \{X \leqslant x\} = \{\omega \in \Omega \mid X(\omega) \ \leqslant\ x\} \in \mathscr{P}(\Omega)$
- L'image de Ω par l'application X, est notée $X(\Omega)$.
  Cet ensemble image $X(\Omega)$ est, par définition, l'ensemble des valeurs que peut prendre l'application X.

⚠

Le terme de « variable aléatoire réelle » peut paraître trompeur :

- X n'est pas une variable, c'est une fonction !
- X n'a rien d'aléatoire, la notion de probabilité n'entre même pas en jeu dans sa définition !

Remarque

Dans la définition au-dessus, on considère la notion de variable aléatoire réelle. On peut aussi définir des variables aléatoires complexes ou des variables aléatoires à valeurs dans n'importe quel ensemble E. Il suffit pour cela de remplacer $\mathbb{R}$ par E dans la définition.

Notation $X(\Omega)$

Il est important de bien comprendre la notation $X(\Omega)$.

- Considérons E et F deux ensembles et $f : E \to F$ une application.
  Alors, pour tout $G \subset E$, $f(G)$ est l'image de l'ensemble G par f.
  En particulier, si $G = E$, $f(E) = \text{Im}(f)$ (image de l'application f).
- L'ensemble image $X(\Omega)$ c'est l'image de l'application X. Dans le monde des v.a.r. , on préfère la notation $X(\Omega)$ à la notation $\text{Im}(X)$.
  L'ensemble $X(\Omega)$ est bien l'ensemble des valeurs que peut prendre X.
- Profitons-en pour rappeler la définition d'image réciproque. Si $f : E \to F$ est une application et si $G \subset F$ alors :

$$f^{-1}(G) \ = \ \{x \in E \,|\, f(x) \in G\}$$

  On constate alors : $\forall x \in \mathbb{R}$, $\{X \leqslant x\} = X^{-1}\big(\ ]-\infty, x]\ \big)$.

Remarque

- Reprenons la définition. X est une variable aléatoire si :
  (i) X est une fonction,
  (ii) X est une machine à créer des événements.
  (pour tout $x \in \mathbb{R}$, $\{X \leqslant x\}$ est un événement)
- Dans ce chapitre s'opère un changement de point de vue par rapport au précédent : on s'intéressait précédemment aux événements (des ensembles), l'objet de base est maintenant celui de variable aléatoire (des applications).
- Si la définition de v.a.r. ne met pas en jeu de probabilité $\mathbb{P}$, il faut bien comprendre que les v.a.r. ne s'étudient que dans un contexte aléatoire (un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$).
  Ceci légitime donc, a posteriori, le nom de variable aléatoire.
- Dans ce contexte, étudier une v.a.r. X, c'est essentiellement étudier sa loi. Cela peut se faire à l'aide de sa fonction de répartition $F_X : \mathbb{R} \to \mathbb{R}$ (ou de sa fonction génératrice dans certains cas). On est donc ramené à l'étude de fonctions de $\mathbb{R}$ dans $\mathbb{R}$. À terme, on pourra donc être amené à utiliser des résultats issus des chapitres d'analyse sur les séries / fonctions.

Exemple

1) On s'intéresse à l'expérience aléatoire consistant à effectuer 2 lancers successifs d'un même dé à 6 faces.

- $\Omega = [\![1,6]\!] \times [\![1,6]\!]$.
- L'univers Ω étant fini, on choisit la tribu $\mathscr{P}(\Omega)$.
- On considère la v.a.r. X égale à la somme des deux résultats obtenus :

$$X \ : \left| \begin{array}{ccc} [\![1,6]\!] \times [\![1,6]\!] & \to & \mathbb{R} \\ (i \quad , \quad j) & \mapsto & i+j \end{array} \right.$$

  Cette variable aléatoire est bien à valeurs dans $\mathbb{R}$.
  On peut préciser l'image de cette v.a.r. : $X(\Omega) = [\![2,12]\!] \subseteq \mathbb{R}$.
- L'ensemble $A = \{\omega \in \Omega \mid X(\omega) \leqslant 4\}$ est bien un événement.
  On peut aussi l'écrire :
  A : « la somme des deux dés est inférieur à 4 ».
  $A = \{(1,1),(1,2),(1,3),(2,1),(2,2),(3,1)\} \quad (A \in \mathscr{P}(\Omega))$
  On notera : $A = \{X \leqslant 4\} = \{\omega \in \Omega \mid X(\omega) \leqslant 4\}$

  L'ensemble $B = \{\omega \in \Omega \mid X(\omega) > 10\}$ est aussi un événement :

$$B = \bigcup_{\substack{k \in X(\Omega) \\ k > 10}} \{X = k\}$$

  B : « la somme des deux dés est strictement supérieur à 10 »
  $B = \{(5,6),(6,6),(6,5)\} \quad (B \in \mathscr{P}(\Omega))$
  On notera : $B = \{X > 10\} = \overline{\{X \leqslant 10\}} = \{\omega \in \Omega \mid X(\omega) > 10\}$

  De même, on peut considérer les événements :

  × $\{2 < X \leqslant 7\} = \{X > 2\} \cap \{X \leqslant 7\}$,
  × $\{2 \leqslant X < 7\} = \{X \leqslant 2\} \cap \{X < 7\}$,
  × $\{2.3 \leqslant X < 7.5\} = \{3 \leqslant X \leqslant 7\}$,
  × $\left\{-32 \leqslant X < \mathrm{e}^{27}\right\} = \{2 \leqslant X \leqslant 12\}$,
  × ...

2) On s'intéresse à l'expérience aléatoire consistant à observer le résultat de 4 lancers successifs d'1 dé à 6 faces.

- $\Omega = \{\text{Pile}, \text{Face}\}^4$.
- L'univers étant fini, on choisit la tribu $\mathscr{P}(\Omega)$.
- On considère la v.a.r. X égale au nombre de Pile obtenus lors du lancer.
- L'ensemble $C = \{\omega \in \Omega \mid X(\omega) \leqslant 2\}$ est un bien un événement :
  C : « le lancer a produit, au plus, deux Pile ».
  Des lancers tels que $(\text{Pile}, \text{Pile}, \text{Face}, \text{Face})$, $(\text{Face}, \text{Face}, \text{Face}, \text{Face})$, ou $(\text{Face}, \text{Face}, \text{Pile}, \text{Face})$ réalisent cet événement.
  On notera $C = \{X \leqslant 2\}$.

  On peut aussi considérer les événements :

  × $\{X = 2\}$: « le lancer contient exactement 2 Pile ».
  × $\{1 < X \leqslant 3\}$: « le lancer contient soit 2 Pile soit 3 Pile ».

Théorème 1.

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

Soit X une variable aléatoire réelle.

$$\boxed{\text{Pour tout } I \text{ intervalle de } \mathbb{R}, \ \{\omega \in \Omega \mid X(\omega) \in I\} \in \mathscr{P}(\Omega)}$$

Plus précisément, les ensembles suivants sont des événements :

- $\{X \leqslant x\} = \{\omega \in \Omega \mid X(\omega) \leqslant x\}$ *où* $x \in \mathbb{R}$
- $\{X \geqslant x\} = \{\omega \in \Omega \mid X(\omega) \geqslant x\}$ *où* $x \in \mathbb{R}$
- $\{X < x\} = \{\omega \in \Omega \mid X(\omega) < x\}$ *où* $x \in \mathbb{R}$
- $\{X > x\} = \{\omega \in \Omega \mid X(\omega) > x\}$ *où* $x \in \mathbb{R}$
- $\{X = x\} = \{\omega \in \Omega \mid X(\omega) = x\}$ *où* $x \in \mathbb{R}$
- $[x \leqslant X \leqslant y] = \{\omega \in \Omega \mid x \leqslant X(\Omega) \leqslant y\}$ *où* $(x,y) \in \mathbb{R}^2$
- ...

Démonstration.

- $\{X \leqslant x\}$ est un événement par définition de v.a.r.
- $\{X \geqslant x\} = \bigcap\limits_{n=1}^{+\infty} \left\{X > x - \frac{1}{n}\right\}$.
  C'est la seule difficulté de la démonstration.

($\subset$) Soit $\omega \in \{X \geqslant x\}$. Ceci signifie que $X(\omega) \geqslant x$.
Soit $n \geqslant 1$. Comme $x > x - \frac{1}{n}$, on a :

$$X(\omega) \geqslant x > x - \frac{1}{n}$$

Ainsi, $\omega \in \left\{X > x - \frac{1}{n}\right\}$.

($\supset$) Soit $\omega \in \bigcap\limits_{n=1}^{+\infty} \left\{X > x - \frac{1}{n}\right\}$.
Autrement dit : $\forall n \in \mathbb{N},\ X(\omega) > x - \frac{1}{n}$.
Démontrons alors :

$$\left(\forall n \in \mathbb{N},\ X(\omega) > x - \frac{1}{n}\right) \ \Rightarrow\ X(\omega) \geqslant x$$

On procède par contraposée. Autrement dit, on démontre :

$$X(\omega) < x \ \Rightarrow\ \left(\exists n_0 \in \mathbb{N},\ X(\omega) \leqslant x - \frac{1}{n_0}\right)$$

Supposons $X(\omega) < x$. Notons $\varepsilon = x - X(\omega)$.
Comme $\frac{1}{n} \underset{n \to +\infty}{\longrightarrow} 0$, il existe un rang n_0 tel que : $\forall n \geqslant n_0,\ |\frac{1}{n} - 0| \leqslant \varepsilon$.
En particulier, on a donc : $\dfrac{1}{n_0} \leqslant \varepsilon = x - X(\omega)$.
Et ainsi : $X(\omega) \leqslant x - \frac{1}{n_0}$.

Pour les autres ensembles, on utilise les propriétés de stabilité de $\mathscr{P}(\Omega)$.

- $\{X = x\} = \{X \leqslant x\} \cap \{X \geqslant x\} \in \mathscr{P}(\Omega)$.
- $\{X < x\} = \{X \leqslant x\} \setminus \{X = x\} \in \mathscr{P}(\Omega)$.
- $\{x \leqslant X \leqslant y\} = \{x \leqslant X\} \cap \{X \leqslant y\} \in \mathscr{P}(\Omega)$.
- ... $\square$

I.2. Notion de variable aléatoire discrète

Définition

Soit $\big(\Omega, \mathscr{P}(\Omega)\big)$ un espace probabilisable.
Soit X une variable aléatoire réelle sur $\big(\Omega, \mathscr{P}(\Omega)\big)$.

- La v.a.r. X est dite **discrète** si son esemble image $X(\Omega)$ est au plus dénombrable, c'est-à-dire si :
  × $X(\Omega)$ est un ensemble fini,
  × `OU` $X(\Omega)$ 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 $X(\Omega)$ c'est-à-dire l'écrire :

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

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

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

On peut d'ailleurs préciser que :

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

$$X \text{ est à valeurs dans } \mathbb{N} \quad \not\Leftarrow \quad X \text{ est une v.a.r. discrète}$$

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.3. Système complet d'événements associé à une v.a.r. finie

Théorème 2.

Soit $(\Omega, \mathscr{P}(\Omega))$ *un espace probabilisable. Soit* $n \in \mathbb{N}^*$.

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

Notons $X(\Omega) = \{x_i \mid i \in I\}$ *où* $I \subseteq [\![1, n]\!]$.

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

On en déduit la propriété suivante

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

2) *Autrement dit :* $\boxed{\sum\limits_{x \in X(\Omega)} \mathbb{P}(\{X = x\}) = 1}$

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

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

($\subset$) 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.4. Loi d'une v.a.r. finie

Définition

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

Soient X et Y deux variables aléatoires (réelles) finies sur $(\Omega, \mathscr{P}(\Omega))$.

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

$$\mathbb{P}_X \ : \left| \begin{array}{ccc} X(\Omega) & \to & [0,1] \\ x & \mapsto & \mathbb{P}_X(x) \ = \ \mathbb{P}(\{X = x\}) \end{array} \right.$$

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

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

Remarque

- Dans la définition de v.a.r. de même loi, on suppose que les v.a.r. X et Y ont même ensemble image (propriété *(i)*). On peut relâcher un peu cette propriété en remplaçant $X(\Omega)$ par $\mathrm{Supp}(X) = \{x \in \mathbb{R} \mid \mathbb{P}(\{X = x\}) \neq 0\}$ (ensemble des valeurs que X prend avec probabilité non nulle) et $Y(\Omega)$ par $\mathrm{Supp}(Y) = \{y \in \mathbb{R} \mid \mathbb{P}(\{Y = y\}) \neq 0\}$ (ensemble des valeurs que X prend avec probabilité non nulle). Autrement dit, les v.a.r. X et Y ont même loi si toutes les valeurs pour lesquelles elle diffèrent sont prises avec probabilité nulle.

- Dans la définition générale de v.a.r. , on a mis en avant le fait que, pour tout $x \in \mathbb{R}$, $\{X \leqslant x\}$ est un événement. Cette propriété est primordiale pour assurer la bonne définition de la notion de fonction de répartition

$$F_X \;:\; \begin{array}{rcl} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \mathbb{P}\big(\{X \leqslant x\}\big) \end{array}$$

Pour tout $x \in \mathbb{R}$, $\{X \leqslant x\} \in \mathscr{P}(\Omega)$.
Cela permet d'assurer que la quantité $\mathbb{P}\big(\{X \leqslant x\}\big)$ est toujours bien définie

- La fonction de répartition F_X d'une v.a.r. X caractérise la loi de X. Plus précisément, donner la fonction de répartition d'une v.a.r. X, c'est donner sa loi. Dans le cadre du programme de PSI, on s'intéresse uniquement aux variable aléatoires discrètes. Pour obtenir la loi d'une v.a.r. discrète X, on peut évidemment déterminer sa fonction de répartition F_X. Mais il est plus commode de déterminer les valeurs de $\mathbb{P}\big(\{X = x\}\big)$ pour tout $x \in X(\Omega)$.
- Dans le programme officiel, X est une variable aléatoire `DISCRÈTE` si :
  (i) X est une application de Ω dans E ($X : \Omega \to E$)
  (ii) $\forall x \in E$, $\{X = x\} \in \mathscr{P}(\Omega)$

  Le programme officiel prend donc le parti de mettre en avant l'importance des événements $\{X = x\}$ dans le cas des v.a.r. discrètes.
- Il convient de noter que, si $X(\Omega)$ est fini (en fait au plus dénombrable), alors :

$$\forall x \in E, \{X = x\} \in \mathscr{P}(\Omega) \;\Leftrightarrow\; \forall x \in E, \{X \leqslant x\} \in \mathscr{P}(\Omega)$$

  La démonstration du sens ($\Rightarrow$) est faite dans le Théorème 2. Le sens direct se démontre en remarquant, pour tout $x \in E$:

$$\{X \leqslant x\} \;=\; \bigcup_{\substack{y \,\in\, X(\Omega) \\ y \,\leqslant\, x}} \{X = y\}$$

  Cette réunion est au plus dénombrable car $X(\Omega)$ l'est. Par propriété de stabilité d'une tribu par réunion au plus dénombrable, on en conclut bien que $\{X \leqslant x\}$ est un événement.

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 $\mathbb{P}$.
  ($(\Omega, \mathscr{P}(\Omega), \mathbb{P})$ est un espace probabilisé)
  $\text{Card}(\Omega) = 2^4 = 16$.
- On note X la v.a.r. qui compte le nombre de Piles obtenu lors du lancer.
  On a alors : $X(\Omega) = \{0, 1, 2, 3, 4\}$.

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

(la loi d'une v.a.r. finie peut être représentée à l'aide d'un tableau)

Remarque `(POLY)`

- Dans le cas où X est une v.a.r. finie, nous avons vu que la loi de X est déterminée par la famille $(\mathbb{P}(\{X = x\}))_{x \in X(\Omega)}$.
- Inversement, si l'on se donne une famille $(p_k)_{k \in [\![1,n]\!]}$ dans quel cas peut-on dire que cette famille est la loi d'une v.a.r. discrète ?
  Évidemment, il faut : $\forall k \in [\![1,n]\!]$, $p_k \geqslant 0$ et $\sum\limits_{k=1}^{n} p_k = 1$.
  Cette condition est même suffisante.
  Plus précisément, pour toute suite $(x_k)_{k \in [\![1,n]\!]}$ de réels, il existe un espace probabilisé $\big(\Omega, \mathscr{P}(\Omega), \mathbb{P}\big)$ et une v.a.r. discrète X telle que $X(\Omega) \subseteq \{x_k \mid k \in [\![1,n]\!]\}$ et : $\forall k \in [\![1,n]\!]$, $\mathbb{P}(\{X = x_k\}) = p_k$.

I.5. Opérations sur les v.a.r. finies

Théorème 3.

Soit $\big(\Omega, \mathscr{P}(\Omega)\big)$ *un espace probabilisable.*
Soient X *et* Y *deux v.a.r. finies sur* $\big(\Omega, \mathscr{P}(\Omega)\big)$ *et* $\lambda \in \mathbb{R}$.

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

où l'on a noté :

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

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 finies, autrement dit que leur ensemble image est finie.
Les v.a.r. X et Y étant finies, il existe $n \in \mathbb{N}^*$ tel qu'on peut noter :

× $X(\Omega) = \{x_i \mid i \in I\}$ où $I \subseteq [\![1,n]\!]$,
× $Y(\Omega) = \{y_j \mid j \in J\}$ où $J \subseteq [\![1,n]\!]$.

Ainsi : $(\lambda X)(\Omega) = \{\lambda x_i \mid i \in I\}$. Comme $I \subseteq [\![1,n]\!]$, 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 $[\![1,n]\!]^2$ donc est finie.
Ainsi, $X+Y$ et XY sont bien des v.a.r. finies. □

Remarque *(Structure de l'ensemble des v.a.r. finies)*
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. finies :

× 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. finie est un sous-espace vectoriel de $\mathcal{F}(\Omega, \mathbb{R})$.

I.6. Transformée d'une v.a.r. finie

Définition

Soit $\big(\Omega, \mathscr{P}(\Omega), \mathbb{P}\big)$ un espace probabilisé et X une v.a.r. finie sur $\big(\Omega, \mathscr{P}(\Omega)\big)$.
Soit $g : X(\Omega) \to \mathbb{R}$ une application.
On note $g(X)$ l'application composée $g \circ X$:

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

Théorème 4.

Soit $\big(\Omega, \mathscr{P}(\Omega), \mathbb{P}\big)$ *un espace probabilisé. Soit* $n \in \mathbb{N}^*$.
Soit X *une v.a.r. finie sur* $\big(\Omega, \mathscr{P}(\Omega)\big)$ *telle que* $X(\Omega) = \{x_i \mid i \in I\}$ $(I \subseteq [\![1,n]\!])$.
Soit $g : X(\Omega) \to \mathbb{R}$ *une application.*
L'application $g(X)$ *est une* **v.a.r. finie** *dont la loi est donnée par :*

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

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

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.*

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

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.

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 $I \subseteq [\![1, n]\!]$, ensemble au plus dénombrable. On récupère donc au passage que $g(X)$ est une v.a.r. finie.

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 :

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

□

Remarque

- Il faut essentiellement retenir de ce théorème que la transformée $g(X)$ d'une v.a.r. finie est une v.a.r. finie.
- La formule théorique donnant la loi de $g(X)$ n'est pas à retenir. En revanche, 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. finie.
On considère la v.a.r. finie $Y = g(X)$ 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 $y \in Y(\Omega)$.

$$\begin{array}{rcl} \mathbb{P}(\{Y = y\}) & = & \mathbb{P}(\{aX + b = y\}) \\ & = & \mathbb{P}(\{aX = y - b\}) \\ & = & \mathbb{P}\left(\left\{X = \frac{y-b}{a}\right\}\right) \end{array}$$

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

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

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 $y \in Y(\Omega)$.

$$\begin{array}{rcl} \mathbb{P}(\{Y = y\}) & = & \mathbb{P}(\{X^2 = y\}) \\ & = & \mathbb{P}(\{X = \sqrt{y}\} \cup \{X = -\sqrt{y}\}) \end{array}$$

Deux cas se présentent alors :

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

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

× 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 $y \in Y(\Omega)$.

$$\begin{array}{rcl} \mathbb{P}(\{Y = y\}) & = & \mathbb{P}(\{|X| = y\}) \\ & = & \mathbb{P}(\{X = y\} \cup \{X = -y\}) \end{array}$$

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 $y \in Y(\Omega)$.

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

(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. Lois discrètes finies

II.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 prenant la valeur 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. prenant la valeur du 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. prenant la valeur numéro de la boule sortie.
Alors X suit la loi certaine d'ensemble image $X(\Omega) = \{7\}$.

II.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 $X \sim \mathcal{U}([\![a, b]\!])$ 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 prenant pour valeur 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. prenant la valeur du numéro de la boule tirée.
Alors : $X \sim \mathcal{U}([\![1, n]\!])$

2) On considère une pièce équilibrée.
L'expérience consiste en 1 lancer de cette pièce.
On note X la v.a.r. prenant la valeur 1 si on obtient Pile et 0 si on obtient Face.
Alors : $X \sim \mathcal{U}([\![0, 1]\!])$

II.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 $X \sim \mathcal{B}(p)$ 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 prenant la valeur 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. prenant la valeur 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 $X \sim \mathcal{B}(p)$.

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

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 $\mathcal{U}([\![0,1]\!])$.
- 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 $\mathcal{B}(p)$.

II.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) $\forall k \in [\![0, n]\!]$, $\boxed{\mathbb{P}(\{X = k\}) = \dbinom{n}{k}\, p^k (1-p)^{n-k}}$
- On utilisera la notation $X \sim \mathcal{B}(n, p)$ 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 de succès p.

Alors la v.a.r. prenant pour valeur 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. prenant pour valeur le 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 : $p^k \ (1-p)^{n-k}$.

On en déduit que : $\mathbb{P}(\{X = k\}) = \dbinom{n}{k} \ p^k \ (1-p)^{n-k}$.

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. prenant la valeur du nombre de boules rouges tirées.
Alors : $X \sim \mathcal{B}\left(n, \dfrac{r}{r+v}\right)$

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

II.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 $X \sim \mathcal{H}(N, n, p)$ 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. prenant pour valeur le nombre de boules blanches obtenues.
  Alors : $X \sim \mathcal{H}\left(a + b, n, \dfrac{a}{a+b}\right)$.

  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. prenant la valeur du nombre de boules blanches obtenues.
  Alors : $X \sim \mathcal{H}\left(a+b, n, \dfrac{a}{a+b}\right)$

  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 :
  $\displaystyle\sum_{k=0}^{n} \mathbb{P}(\{X = k\}) = 1$ donc $\displaystyle\sum_{k=0}^{n} \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}} = 1$.

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

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)

- $\displaystyle\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\, \mathbb{P}(\{X = x\}) = \sum_{k=0}^{n} k\frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}} = \sum_{k=1}^{n} k\frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}$.

  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 :

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

□

III. Espérance d'une variable aléatoire discrète

III.1. Définition

Définition

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

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

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

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

2) De manière générale, si on écrit $X(\Omega) = \{x_i \mid i \in I\}$ (avec $I \subset [\![1, n]\!]$) alors :

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

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).

III.2. Propriétés de l'espérance

III.2.a) Linéarité

Théorème 5.

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

On suppose que X et Y sont finies.

1) *Soit $\lambda \in \mathbb{C}$. Les v.a.r. $X+Y$ et λX sont 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 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.

III.2.b) Positivité, croissance

Théorème 6.

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

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

(positivité de l'espérance)

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

(croissance de l'espérance)

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

III.3. Théorème de transfert

Théorème 7. *(théorème de transfert)*

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

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

1) *La v.a. $g(X)$ est finie.*

2) *De plus :*

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

Remarque

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

IV. Variance d'une variable aléatoire discrète

IV.1. Moment d'ordre 2

Définition

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

1) La v.a. X^2 est finie.

2) De plus :

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

IV.2. Variance d'une v.a. discrète

Définition

Soit X une v.a. finie.

- La variance de X est notée $\mathbb{V}(X)$ et définie par :

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

- Dans le cas où :
  - × la variable aléatoire X est à valeurs réelles,
  - × la variable aléatoire X est de variance finie (ce qui est toujours le cas pour une v.a. finie),

  on définit l'écart-type de X et on note $\sigma(X)$ la quantité : $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Remarque

- Dans le programme officiel, on considère la variance n'est définie que pour les variables aléatoires **réelles** ce qui permet de ne pas avoir à faire un cas particulier pour définir l'écart-type.
- Remarquons au passage que cette notion d'écart-type est bien définie.
  Comme $(X - \mathbb{E}(X))^2 \geqslant 0$, alors $\mathbb{V}(X) = \mathbb{E}\Big((X - \mathbb{E}(X))^2\Big) \geqslant 0$.
  (cf point 3) du Théorème 6)
- La variance est une mesure **moyenne** de l'écart existant entre X et $\mathbb{E}(X)$. La variance considère un écart quadratique. Il est alors naturel d'introduire l'écart-type (la racine de la variance) pour gommer le « défaut quadratique » introduit par ce choix d'écart.
- Aurait-on pu choisir d'autres types d'écart ?

  1) On peut tout d'abord considérer l'écart simple.
  La mesure moyenne de l'écart simple n'est en réalité d'aucune utilité :

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

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

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

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

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

  On fera attention au passage :

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

IV.3. Détermination pratique de la variance

Théorème 8. *(formule de Kœnig-Huygens)*

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

$$\boxed{\text{La v.a. } X \text{ est de variance finie}}$$

De plus :

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

Démonstration.
Remarquons tout d'abord que, comme X est d'espérance finie :

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

Et ainsi :

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

Comme X^2 est d'espérance finie, alors la v.a. $\big(X - \mathbb{E}(X)\big)^2$ aussi. Ainsi la variance de X est finie.
De plus :

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

□

IV.4. Propriétés de la variance

Théorème 9.

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

L'opérateur variance vérifie les propriétés suivantes.

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

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

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

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

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

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

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

Démonstration.

1) Soit $\lambda \in \mathbb{C}$. Alors les v.a. $X+Y$ et λX sont finies, et donc de variances finies.
D'après la formule de Kœnig-Huygens (et par linéarité de l'espérance) :

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

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

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

2) Soit $(a,b) \in \mathbb{C}^2$. Par hypothèse, X est finie et donc de variance finie. La v.a. $\big((aX+b) - \mathbb{E}(aX+b)\big)^2$ est également finie et admet alors une espérance.
Par ailleurs :

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

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 6 :

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

□

⚠ Contrairement à l'espérance, l'opérateur de variance, défini à l'aide d'une élévation au carré, n'est pas linéaire.
Pour la propriété ***2)*** du Théorème 9, on parle de « quadricité » de la variance.

Remarque

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

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

IV.5. Variables centrées réduites

Définition

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

a) On rappelle que X est d'espérance finie :

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

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

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

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

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

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

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

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

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

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

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

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

Remarque

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

V. Le cas particulier des v.a. à valeurs entières

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

Théorème 10.

Soit $n \in \mathbb{N}$.

Soit X une v.a. finie à valeurs dans $[\![0, n]\!]$.

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

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

2. $\boxed{\mathbb{E}(X) = \sum\limits_{k=0}^{n-1} \mathbb{P}\big(\{X > k\}\big)}$

Démonstration.

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

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

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

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

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

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

2. On remarque :

$$\begin{aligned}\sum_{k=0}^{n-1} \mathbb{P}\big(\{X > k\}\big) &= \sum_{k=0}^{n-1} \left(\sum_{j=k+1}^{n} \mathbb{P}\big(\{X = j\}\big) \right) \quad \textit{(car } X \textit{ est à valeurs dans } [\![1, n]\!]\textit{)} \\ &= \sum_{0 \leqslant k < j \leqslant n} \mathbb{P}\big(\{X = j\}\big) \\ &= \sum_{j=1}^{n} \left(\sum_{k=0}^{j-1} \mathbb{P}\big(\{X = j\}\big) \right) \\ &= \sum_{j=1}^{n} j\, \mathbb{P}\big(\{X = j\}\big) \ = \ \mathbb{E}(X)\end{aligned}$$

Remarque

- On utilise dans le calcul la possibilité d'intervertir les symboles de sommation. On a vu plus tôt dans le cours l'interversion de sommes lorsqu'il y a indépendance des indices de sommation. Plus précisément :

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

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

Dans le cas des sommes finies

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

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

Comment retenir ces formules ?

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

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

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

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

□

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

VI. Lois associées à un couple de v.a. discrètes

VI.1. Loi d'un couple de v.a. discrètes

VI.1.a) Définition de la notion de couple de v.a. discrètes

Définition

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

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

- On dit que X est une **variable aléatoire finie à valeurs dans** E et définie sur $(\Omega, \mathscr{P}(\Omega))$ si :
  - ***(i)*** X est une application de Ω dans E ($X : \Omega \to E$).
  - ***(ii)*** $\forall x \in E, \ X^{-1}(\{x\}) = \{X = x\} = \{\omega \in \Omega \mid X(\omega) = x\} \in \mathscr{P}(\Omega)$.
  - ***(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 finies à valeurs dans** $E \times F$ et définie sur $(\Omega, \mathscr{A})$ si :
  - × Z s'écrit sous la forme $Z = (X, Y)$,
  - × X est une variable aléatoire finie à valeurs dans E,
    Y est une variable aléatoire finie à valeurs dans F.
- En particulier, un couple de variable aléatoires finies est une variable aléatoire finie à valeur dans un produit.

$$\begin{array}{ccccccc} Z & : & \Omega & \to & E & \times & F \\ & & \omega & \mapsto & \Big(X(\omega) & , & Y(\omega)\Big) \end{array}$$

Exemple

1) On lance deux fois de suite un dé à 6 faces.

- On note X la v.a. égale au plus petit des deux résultats obtenus.
- On note Y la v.a. égale au plus grand des deux résultats obtenus.

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

2) Une urne contient 3 boules blanches, 4 boules vertes et 5 boules bleues. On tire simultanément 3 boules dans l'urne.

- On note X la v.a. égale au nombre de boules blanches obtenues.
- On note Y la v.a. égale au nombre de boules vertes obtenues.

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

3) On lance indéfiniment une pièce non équilibrée pour laquelle la probabilité d'obtenir Pile vaut $\frac{3}{4}$ (et celle d'obtenir Face vaut donc $\frac{1}{4}$).
On appelle chaîne toute séquence constituée uniquement de Pile ou uniquement de Face.

- On note X la v.a. égale à longueur de la première chaîne obtenue.
- On note Y la v.a. égale à la longueur de la deuxième chaîne.

Alors, (X, Y) est un couple de v.a. discrètes.
(par exemple, si $\omega = PPPPFFP\ldots$, $X(\omega) = 4$ et $Y(\omega) = 2$)

4) On dispose d'une pièce de monnaie et d'une urne contenant en tout $2n$ boules ($n \in \mathbb{N}^*$) dont n sont blanches et n sont noires. L'expérience aléatoire consiste à lancer la pièce jusqu'à obtenir le premier pile. Puis à effectuer autant de tirages avec remise dans l'urne que de lancers effectués pour obtenir Pile.

- On note X la v.a. égale au nombre de lancers de pièce effectués.
- On note Y la v.a. égale au nombre de boules blanches obtenues lors des tirages.

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

Remarque

- On peut définir de la même manière la notion de triplet de variables aléatoires discrètes (c'est une variable aléatoire Z qui s'écrit sous la forme $Z = (X_1, X_2, X_3)$).
- De manière générale, pour tout $n \in \mathbb{N}^*$, on peut définir la notion de n-uplet de variables aléatoires discrètes. Un n-uplet de variables aléatoires discrètes est une variable aléatoire $Z = (X_1, \dots, X_n)$ où les variables aléatoires X_1, ..., X_n sont discrètes.

VI.1.b) Loi d'un couple de v.a. finies

Définition

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

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

- 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 $\mathbb{P}\big(\{X = x\} \cap \{Y = y\}\big)$ 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 :
  $$\Big(\mathbb{P}\big(\{X = x\} \cap \{Y = y\}\big)\Big)_{(x,y)\in X(\Omega)\times Y(\Omega)}$$
- On note au passage que l'événement $\{X = x\} \cap \{Y = y\}$ peut aussi s'écrire :
  - × $\{X = x, Y = y\}$.
  - × $\{(X, Y) = (x, y)\}$.

Remarque

- On peut généraliser cette définition au cas des triplets ou des n-uplets de variables aléatoires discrètes. La loi de $Z = (X_1, \dots, X_n)$ est la famille :
  $$\Big(\mathbb{P}\big(\{X_1 = x_1\} \cap \dots \cap \{X_n = x_n\}\big)\Big)_{(x_1,\dots,x_n)\in X_1(\Omega)\times\dots\times X_n(\Omega)}$$
- Dans le cas d'un couple (X, Y) constitué de deux v.a. X et Y finies, on peut représenter la loi de (X, Y) sous forme d'un tableau à double entrée.

Exemple

1) On reprend l'exemple ***1)*** qui décrit deux lancers d'un dé à 6 faces.

× $\Omega = [\![1, 6]\!]^2$. Ainsi, $\mathrm{Card}(\Omega) = \mathrm{Card}([\![1, 6]\!]^2) = 6 \times 6 = 36$.

× Ω est muni de la probabilité uniforme notée $\mathbb{P}$.
($(\Omega, \mathscr{P}(\Omega), \mathbb{P})$ est un espace probabilisé)

On a donc :
$$\forall A \in \mathscr{P}(\Omega),\ \mathbb{P}(A) = \frac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)} = \frac{\mathrm{Card}(A)}{36}$$

- Déterminons la loi de (X, Y).
  Tout d'abord : $X(\Omega) = [\![1, 6]\!]$ et $Y(\Omega) = [\![1, 6]\!]$.
- Soit $i \in [\![1, 6]\!]$ et soit $j \in [\![1, 6]\!]$.

	L'événement $\{X = i\} \cap \{Y = j\}$ est réalisé		
$\Leftrightarrow$	L'événement $\{X = i\}$ est réalisé	et	l'événement $\{Y = j\}$ est réalisé
$\Leftrightarrow$	Le plus petit des 2 résultats obtenus vaut i	et	le plus grand des 2 résultats obtenus vaut j

Trois cas se présentent.

× Si $i > j$, alors :
$$\{X = i\} \cap \{Y = j\} = \varnothing$$
En effet, le plus petit des deux résultats ne peut être strictement plus grand que le plus grand.

× Si $i < j$, alors :
$$\{X = i\} \cap \{Y = j\} = \{(i, j), (j, i)\}$$
(cet événement est réalisé par deux 2-tirages : celui où le 1^{er} résultat vaut i et le suivant j et celui où le premier résultat vaut j et le suivant i)

× Si $i = j$, alors :

$$\{X = i\} \cap \{Y = j\} = \{(i, i)\}$$

(cet événement est réalisé par un unique 2-tirages : le 1^{er} lancer donne i et le deuxième aussi)

- On en conclut :

$$\forall (i,j) \in [\![1,6]\!]^2,\ \mathbb{P}\big(\,\{X = i\} \cap \{Y = j\}\,\big) = \begin{cases} \frac{0}{36} & \text{si } i > j \\ \frac{2}{36} & \text{si } i < j \\ \frac{1}{36} & \text{si } i = j \end{cases}$$

- Ce résultat peut aussi être présenté sous forme de tableau.

$x \in X(\Omega)$ \ $y \in Y(\Omega)$	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
2	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
3	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
4	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
5	0	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$
6	0	0	0	0	0	$\frac{1}{36}$

2) On reprend maintenant le deuxième exemple.
On note $B = \{b_1, \dots, b_{12}\}$ l'ensemble de boules de l'urne.
(les boules b_1, …, b_3 sont blanches, les boules b_4, …, b_7 sont vertes et les boules b_8, …, b_{12} sont bleues)

× Ω est l'ensemble des parties à 3 éléments de l'ensemble B.
Ainsi, $\text{Card}(\Omega) = \binom{12}{3} = 220$.

× Ω est muni de la probabilité uniforme notée $\mathbb{P}$.
($(\Omega, \mathscr{P}(\Omega), \mathbb{P})$ est un espace probabilisé)

On a donc :

$$\forall A \in \mathscr{P}(\Omega),\ \mathbb{P}(A) \ = \ \frac{\text{Card}(A)}{\text{Card}(\Omega)} \ = \ \frac{\text{Card}(A)}{220}$$

- Déterminons la loi de (X, Y).
Tout d'abord : $X(\Omega) = [\![0,3]\!]$ et $Y(\Omega) = [\![0,3]\!]$.
- Soit $i \in [\![0,3]\!]$ et soit $j \in [\![0,3]\!]$.

	L'événement $\{X = i\} \cap \{Y = j\}$ est réalisé		
$\Leftrightarrow$	L'événement $\{X = i\}$ est réalisé	et	l'événement $\{Y = j\}$ est réalisé
$\Leftrightarrow$	Le 3-tirage contient i boules blanches	et	le 3-tirage contient j boules vertes

Un 3-tirage réalisant $\{X = i\} \cap \{Y = j\}$ est entièrement déterminé par :

× les numéros des i boules blanches tirées : $\binom{3}{i}$ possibilités.

× les numéros des j boules vertes tirées : $\binom{4}{j}$ possibilités.

× les numéros des boules ni blanches, ni vertes tirées : $\binom{5}{3-i-j}$ possibilités.
(ce sont forcément des boules bleues)

- On en conclut :

$$\forall (i,j) \in [\![0,3]\!]^2,\ \mathbb{P}(\{X = i\} \cap \{Y = j\}) \ = \ \frac{\binom{3}{i}\binom{4}{j}\binom{5}{3-i-j}}{\binom{12}{3}}$$

- Ce résultat peut aussi être représenté sous forme de tableau.

$x \in X(\Omega)$ \ $y \in Y(\Omega)$	0	1	2	3
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0
3	$\frac{1}{220}$	0	0	0

Remarque
De ces premières études, on peut tirer plusieurs remarques.

- Afin de déterminer la loi d'un couple (X, Y), on commence `TOUJOURS` par déterminer $X(\Omega)$ et $Y(\Omega)$.
- Il est fréquent d'avoir à réaliser une disjonction de cas lorsque l'on détermine la loi d'un couple de v.a. .
- Si l'on somme tous les éléments du tableau, on obtient 1.
  (c'est une mesure de vérification !)
  Ceci provient du fait que $\Big(\{X = x\} \cap \{Y = y\} \Big)_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}}$ est **le système complet d'événements associé au couple** (X, Y).

VI.1.c) Système complet d'événements associé à un couple de v.a. finies

Définition
Soit $(\Omega, \mathscr{P}(\Omega), \mathbb{P})$ un espace probabilisé.
Soit (X, Y) un couple de v.a. discrètes.

- On appelle **le système complet d'événements associé au couple** (X, Y) la famille :

$$\Big(\{X = x\} \cap \{Y = y\} \Big)_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}}$$

- Cette famille est un système complet d'événements. En particulier :

$$\begin{aligned} & \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} \mathbb{P}\big(\{X = x\} \cap \{Y = y\} \big) \\ = & \sum_{x \in X(\Omega)} \Big(\sum_{y \in X(\Omega)} \mathbb{P}(\{X = x\} \cap \{Y = y\}) \Big) \\ = & \sum_{y \in X(\Omega)} \Big(\sum_{x \in X(\Omega)} \mathbb{P}(\{X = x\} \cap \{Y = y\}) \Big) = 1 \end{aligned}$$

Démonstration.

1) Démontrons que cette famille est constituée d'événements 2 à 2 incompatibles. Choisissons deux événements distincts de cette famille.
Soit $(x_1, y_1) \in X(\Omega) \times Y(\Omega)$ et $(x_2, y_2) \in X(\Omega) \times Y(\Omega)$ tels que :

$$(x_1, y_1) \neq (x_2, y_2)$$

On a alors :

$$\begin{aligned} & \big(\{X = x_1\} \cap \{Y = y_1\} \big) \cap \big(\{X = x_2\} \cap \{Y = y_2\} \big) \\ = & \big(\{X = x_1\} \cap \{X = x_2\} \big) \cap \big(\{Y = y_1\} \cap \{Y = y_2\} \big) \\ = & \varnothing \end{aligned}$$

En effet, on a forcément $x_1 \neq x_2$ ou $y_1 \neq y_2$ car $(x_1, y_1) \neq (x_2, y_2)$.

2) Démontrons que la réunion de ces événements est Ω.

$$\begin{aligned}
& \bigcup_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} \{X = x\} \cap \{Y = y\} \\
= \ & \bigcup_{x \in X(\Omega)} \left(\bigcup_{y \in Y(\Omega)} \{X = x\} \cap \{Y = y\} \right) \\
= \ & \bigcup_{x \in X(\Omega)} \left(\{X = x\} \cap \left(\bigcup_{y \in Y(\Omega)} \{Y = y\} \right) \right) && \textit{(car } \cap \textit{ est distributive par rapport à } \cup\textit{)} \\
= \ & \bigcup_{x \in X(\Omega)} (\{X = x\} \cap \Omega) && \textit{(car } (\{Y = y\})_{y \in Y(\Omega)} \textit{ est un sce)} \\
= \ & \bigcup_{x \in X(\Omega)} \{X = x\} \ = \ \Omega && \textit{(car } (\{X = x\})_{x \in X(\Omega)} \textit{ est un sce)}
\end{aligned}$$

On en déduit notamment :

$$\begin{aligned}
& \sum_{x \in X(\Omega)} \sum_{y \in X(\Omega)} \mathbb{P}(\{X = x\} \cap \{Y = y\}) \\
= \ & \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} \mathbb{P}(\{X = x\} \cap \{Y = y\}) \\
= \ & \mathbb{P}\Big(\bigcup_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} \{X = x\} \cap \{Y = y\} \Big) \\
= \ & \mathbb{P}(\Omega) \ = \ 1
\end{aligned}$$

$\square$

Remarque

- On souligne que cette démonstration est valable dans le cas où $X(\Omega)$ et $Y(\Omega)$ sont finis mais aussi dans le cas où ils sont infinis (dénombrables). Cela implique alors que les sommes considérées sont infinies.
- On définit de la même manière le système complet d'événements associé à un n-uplet $Z = (X_1, \ldots, X_n)$. C'est la famille :

$$\Big(\{X_1 = x_1\} \cap \ldots \cap \{X_n = x_n\} \Big)_{(x_1, \ldots, x_n) \in X_1(\Omega) \times \ldots \times X_n(\Omega)}$$

VI.2. Lois conditionnelles

Définition

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

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

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 :

$$\begin{array}{rcl} X(\Omega) & \to & \mathbb{R} \\ x & \mapsto & \mathbb{P}_{\{Y=y\}}\big(\{X = x\} \big) \ = \ \dfrac{\mathbb{P}\big(\{X = x\} \cap \{Y = y\} \big)}{\mathbb{P}\big(\{Y = y\} \big)} \end{array}$$

• 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 :

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

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 :

$$\begin{array}{rcl} Y(\Omega) & \to & \mathbb{R} \\ y & \mapsto & \mathbb{P}_{\{X=x\}}(\{Y = y\}) \ = \ \dfrac{\mathbb{P}(\{X = x\} \cap \{Y = y\})}{\mathbb{P}(\{X = x\})} \end{array}$$

• 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 :

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

Remarque

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

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

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

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

Exercice

Soient X et Y deux v.a. indépendantes.

On suppose que $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$ où $\lambda > 0$ et $\mu > 0$.

Soit $n \in \mathbb{N}$. Reconnaître la loi de X sachant $\{X + Y = n\}$.

Démonstration.

- Les v.a. X et Y étant indépendantes, $X + Y \sim \mathcal{P}(\lambda + \mu)$.
  On en déduit : $\mathbb{P}\big(\{X + Y = n\}\big) \;=\; \mathrm{e}^{-(\lambda+\mu)} \dfrac{(\lambda+\mu)^n}{n!} \neq 0$.
- Soit $k \in \mathbb{N}$. Par définition :

$$\begin{aligned}
\mathbb{P}_{\{X+Y=n\}}(\{X = k\}) &= \frac{\mathbb{P}\big(\{X = k\} \cap \{X + Y = n\}\big)}{\mathbb{P}\big(\{X + Y = n\}\big)} \\
&= \frac{\mathbb{P}\big(\{X = k\} \cap \{Y = n - k\}\big)}{\mathbb{P}\big(\{X + Y = n\}\big)} \\
&= \frac{\mathbb{P}\big(\{X = k\}\big)\, \mathbb{P}\big(\{Y = n - k\}\big)}{\mathbb{P}\big(\{X + Y = n\}\big)} \quad \textit{(car X et Y sont indépendantes)}
\end{aligned}$$

On distingue alors deux cas :

× Si $k > n$: alors $\{Y = n - k\} = \varnothing$. Et donc $\mathbb{P}_{\{X+Y=n\}}(\{X = k\}) = 0$.

× Si $k \leqslant n$:

$$\begin{aligned}
\mathbb{P}_{\{X+Y=n\}}(\{X = k\}) &= \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!}\ \mathrm{e}^{-\mu} \frac{\mu^{n-k}}{(n-k)!}\ \mathrm{e}^{\lambda+\mu} \frac{n!}{(\lambda+\mu)^n} \\
&= \frac{n!}{k!(n-k)!}\ \frac{\lambda^k\ \mu^{n-k}}{(\lambda+\mu)^k\ (\lambda+\mu)^{n-k}} \\
&= \binom{n}{k} \left(\frac{\lambda}{\lambda+\mu}\right)^k \left(1 - \frac{\lambda}{\lambda+\mu}\right)^{n-k}
\end{aligned}$$

On reconnaît la loi binomiale de paramètre $(n, \frac{\lambda}{\lambda+\mu})$. $\square$

VI.3. Lois marginales

VI.3.a) Définition

Définition

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

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

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

Remarque

- L'expression « loi marginale » n'a de sens que dans un contexte d'étude d'un couple de v.a. (X, Y). Aux concours, cette expression ne sera pas forcément utilisée, même dans les sujets sur les couples. On parlera alors tout simplement de loi de X (resp. Y) en lieu et place de $1^{\text{ère}}$ (resp. $2^{\text{ème}}$) loi marginale du couple (X, Y).
- De même, on peut définir la $i^{\text{ème}}$ loi marginale (pour $i \in [\![1, n]\!]$) d'un n-uplet de variables aléatoires $(X_1, \ldots, X_n)$.

VI.3.b) Détermination en pratique des lois marginales

Théorème 11.

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

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

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

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

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

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

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

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

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

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

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

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

Démonstration.

- Les v.a. X et Y jouent des rôles analogues.
  Ainsi, les démonstrations des points ***1)*** et ***2)*** sont similaires.
  Démontrons le point ***1)***.
- La famille $\big(\{Y = y\}\big)_{y \in Y(\Omega)}$ est un sce.
  Soit $x \in X(\Omega)$. Par la formule des probabilités totales :

$$\mathbb{P}\big(\{X = x\}\big) \; = \; \sum_{y \in Y(\Omega)} \mathbb{P}\big(\{X = x\} \cap \{Y = y\}\big)$$

- Si on sait de plus : $\forall y \in Y(\Omega), \; \mathbb{P}\big(\{Y = y\}\big) \neq 0$, alors :

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

□

Remarque

- On retiendra que pour obtenir une loi marginale à partir de la loi du couple ou d'une loi conditionnelle, il suffit d'appliquer la formule des probabilités totales.
- Cela permet de construire des énoncés du type :

1) Reconnaître la loi de X.

2) Déterminer la loi du couple (X, Y) *(si l'énoncé a un parti-pris couple).*

`OU`

Pour tout $x \in X(\Omega)$, déterminer la loi conditionnelle de Y sachant $\{X = x\}$ *(si l'énoncé a un parti-pris couple).*

3) Déterminer la loi de Y.

VII. Indépendance de variables aléatoires discrètes

VII.1. Indépendance de deux variables aléatoires discrètes

Définition

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

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

- 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** (pour la probabilité $\mathbb{P}$) 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 :

$$\boxed{\begin{array}{l} \forall x \in X(\Omega), \; \forall y \in Y(\Omega), \\ \mathbb{P}\big(\{X = x\} \cap \{Y = y\}\big) \; = \; \mathbb{P}\big(\{X = x\}\big) \times \mathbb{P}\big(\{Y = y\}\big) \end{array}}$$

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

Remarque

- Dans les énoncés, on trouvera souvent la question :
  « les v.a. X et Y sont-elles indépendantes ? ».
  Ainsi énoncée, cette question attend généralement la réponse : `NON`.
- Il s'agit alors de démontrer la négation de la propriété d'indépendance. Or :

$$\texttt{NON}(\forall x \in X(\Omega), \forall y \in Y(\Omega), \mathbb{P}\big(\{X = x\} \cap \{Y = y\}\big) = \mathbb{P}\big(\{X = x\}\big) \times \mathbb{P}\big(\{Y = y\}\big))$$
$$\Leftrightarrow \; \exists x \in X(\Omega), \exists y \in Y(\Omega), \mathbb{P}\big(\{X = x\} \cap \{Y = y\}\big) \neq \mathbb{P}\big(\{X = x\}\big) \times \mathbb{P}\big(\{Y = y\}\big)$$

On en déduit la méthodologie suivante.

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 :

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

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

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

Exemple

1) Reprenons l'expérience aléatoire consistant à tirer 3 boules dans une urne contenant 3 boules blanches, 4 vertes et 5 bleues. Alors :

$$\mathbb{P}\big(\{X = 3\} \cap \{Y = 3\}\big) \ = \ 0 \ \neq \ \underbrace{\mathbb{P}\big(\{X = 3\}\big)}_{\neq 0} \times \underbrace{\mathbb{P}\big(\{Y = 3\}\big)}_{\neq 0}$$

Ainsi, X et Y ne sont pas indépendantes.

2) Si on reprend l'exemple du lancer de pièces avec X le rang d'apparition du premier pile et Y le rang d'apparition du deuxième pile, alors :

$$\mathbb{P}\big(\{X = 2\} \cap \{Y = 1\}\big) \ = \ 0 \ \neq \ \underbrace{\mathbb{P}\big(\{X = 2\}\big)}_{\neq 0} \times \underbrace{\mathbb{P}\big(\{Y = 1\}\big)}_{\neq 0}$$

Ainsi, X et Y ne sont pas indépendantes.

Remarque

Soient X et Y sont deux v.a. finies indépendantes.

Alors, tout événement ne dépendant que de la variable X est indépendant de tout événement ne dépendant que de la variable Y.
Plus précisément, pour tout $x \in X(\Omega)$, pour tout $y \in Y(\Omega)$:

- $\mathbb{P}(\{X = x\} \cap \{Y \leqslant y\}) = \mathbb{P}\big(\{X = x\}\big) \times \mathbb{P}\big(\{Y \leqslant y\}\big)$,
- $\mathbb{P}(\{X \leqslant x\} \cap \{Y \leqslant y\}) = \mathbb{P}\big(\{X \leqslant x\}\big) \times \mathbb{P}\big(\{Y \leqslant y\}\big)$,
- $\mathbb{P}(\{X \leqslant x\} \cap \{Y > y\}) = \mathbb{P}\big(\{X \leqslant x\}\big) \times \mathbb{P}\big(\{Y > y\}\big)$,
- $\mathbb{P}(\{X > x\} \cap \{Y < y\})) = \mathbb{P}(\{X > x\} \times \mathbb{P}\big(\{Y < y\}\big)$,
- ...

Cela provient essentiellement du fait que :

$$\{Y \leqslant y\} = \bigcup_{\substack{y_j \in Y(\Omega) \\ y_j \leqslant y}} \{Y = y_j\}$$

et que $\{X = x\}$ est indépendant de tout événement $\{Y = y_j\}$.

Théorème 12.

Soit $(\Omega, \mathscr{P}(\Omega), \mathbb{P})$ un espace probabilisé fini.
Soit (X, Y) un couple de v.a. finies.

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

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

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

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

VII.2. Indépendance mutuelle de variables aléatoires finies

Définition

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

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

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

$$\boxed{\begin{array}{l}\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega), \\ \mathbb{P}\left(\bigcap\limits_{i=1}^{n} \{X_i = x_i\}\right) = \prod\limits_{i=1}^{n} \mathbb{P}\big(\{X_i = x_i\}\big)\end{array}}$$

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

  $\forall n \geqslant 2,$ les variables X_1, X_2, ..., X_n sont mutuellement indépendantes

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

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

- 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.

Remarque

Soient $X_1, X_2, \dots, X_n$ des v.a. finies mutuellement indépendantes.
Alors toute famille de n événements dont chacun est construit à l'aide d'une v.a. X_i est une famille d'événements mutuellement indépendants.

Plus précisément, pour tout $x_1 \in X_1(\Omega)$, ..., $x_n \in X_n(\Omega)$:

× les événements $\{X_1 = x_1\}, \{X_2 \leqslant x_2\}, \dots \{X_n \leqslant x_n\}$ sont mutuellement indépendants,

× les événements $\{X_1 \leqslant x_1\}, \{X_2 > x_2\}, \dots \{X_n \leqslant x_n\}$ sont mutuellement indépendants,

× ...

Théorème 13 (Lemme des coalitions)**.**

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

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

Soit F un ensemble.

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

- *Cas de 2 v.a.*

$$\boxed{X \perp\!\!\!\perp Y \ \Rightarrow \ f(X) \perp\!\!\!\perp g(Y)}$$

- *Généralisation à n v.a.*
  Soient X_1, ..., X_n des v.a. finies.
  Soient $f_1 : X_1(\Omega) \to F$, ..., $f_n : X_n(\Omega) \to F$ des fonctions.

Les v.a. finies $X_1, \dots, X_n$ sont mutuellement indépendantes	$\Rightarrow$	*Les v.a. finies $f_1(X_1), \dots, f_n(X_n)$ sont mutuellement indépendantes*

Les v.a. finies $X_1, \dots, X_n$ sont mutuellement indépendantes	$\Rightarrow$	*Toute v.a. s'exprimant en fonction des v.a. $X_1, \dots, X_p$ est indépendante de toute v.a. s'exprimant en fonction des v.a. $X_{p+1}, \dots, X_n$ (pour $p \in [\![2, n-1]\!]$)*

Remarque

- Le lemme des coalitions ne sert qu'à une chose : démontrer que des variables aléatoires sont indépendantes. Outre le retour à la définition (méthode fastidieuse) c'est le seul outil dont on dispose pour démontrer que des variables aléatoires sont indépendantes.
- Il ne sert à rien de retenir l'énoncé de manière très formelle. Ce qu'il convient de retenir c'est que toute variable aléatoire X construite à l'aide de certaines variables aléatoires est indépendante de toute variable aléatoire Y construite à l'aide de variables aléatoires qui sont indépendantes de celles utilisées pour définir X.

Exemple

- Soient $X_1, \dots, X_5$ des v.a. finies mutuellement indépendantes. Alors :
  - × les v.a. X_1, ${X_2}^2$, $2\,X_3$, $e^{X_4} - 1$ et $|X_5|$ sont mutuellement indépendantes.
  - × les v.a. $2\,X_1 X_3 -\ X_5$ et ${X_2}^2$ sont indépendantes.
  - × les v.a. $\min(X_1, X_2)$ et $\max(X_3, X_4, X_5)$ sont indépendantes.
- Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. mutuellement indépendantes.
  Alors, pour tout $n \in \mathbb{N}^*$, $S_n = \sum\limits_{k=1}^{n} X_k$ et X_{n+1} sont indépendantes.
- Si X^2 et Y^2 ne sont pas indépendantes, alors, en procédant par l'absurde, on démontre que X et Y ne le sont pas non plus.

VIII. Opérations sur les v.a. discrètes

Dans cette section, on étudie des v.a. du type $Z = g(X, Y)$ où (X, Y) est un couple de v.a. finies et $g : X(\Omega) \times Y(\Omega) \to \mathbb{R}$.

VIII.1. Cas général : loi de $g(X, Y)$

Théorème 14.

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

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

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

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

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

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

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

Démonstration.
Il suffit de remarquer :

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

et que $(\{X = x\} \cap \{Y = y\})_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}}$ est une famille d'événements deux à deux incompatibles (car c'est un sce). □

Remarque

- Le théorème précédente stipule que la loi de $g(X,Y)$ est fournie par la loi conjointe de X et de Y.
- Si on sait de plus que X et Y sont indépendantes, les lois de X et de Y suffisent à obtenir la loi de $g(X,Y)$. L'indépendance fournit un cadre simple pour l'étude de la loi de $g(X,Y)$.

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

VIII.2.a) Cas général

Théorème 15.

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

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

- *L'application $X+Y$ est une v.a. finie.*
- *La loi de $X+Y$ est fournie par, pour tout $z \in (X+Y)(\Omega)$:*

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

Démonstration.
La famille $(\{X=x\})_{x \in X(\Omega)}$ forme un sce.
Ainsi, d'après la formule des probabilités totales :

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

□

À RETENIR

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

Remarque

- Ce qui importe le plus ici est de retenir la méthode plus que l'énoncé en lui-même. Si l'on souhaite déterminer la probabilité d'un événement construit à l'aide de deux v.a. discrètes, la bonne manière de procéder est d'utiliser la formule des probabilités totales.
- On peut illustrer le propos précédent en listant certaines cas donnant lieu à l'utilisation de la formule des probabilités totales :

$$\mathbb{P}\big(\{XY=1\}\big), \mathbb{P}\big(\{X-Y=4\}\big), \mathbb{P}\big(\{2\,X+Y=-1\}\big), \mathbb{P}\big(\{X=Y\}\big), \dots$$

En particulier :

× la loi du produit XY de deux v.a.r. finies s'obtient par cette méthode.

× il convient de remarquer : $\mathbb{P}\big(\{X=Y\}\big) = \mathbb{P}\big(\{X-Y=0\}\big)$. Déterminer la probabilité que deux v.a. soient égales n'est donc qu'un calcul de probabilité illustrant la loi de la somme.

VIII.2.b) Application : stabilité des lois usuelles

Théorème 16.

***1)* Stabilité des lois binomiales**

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

$$\left.\begin{array}{l} \bullet\ X \sim \mathcal{B}(m,p) \text{ et } Y \sim \mathcal{B}(n,p) \\ \bullet\ X \perp\!\!\!\perp Y \end{array}\right. \quad \Rightarrow \quad X+Y \sim \mathcal{B}(m+n,p)$$

***2)* Stabilité des lois de Poisson**

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

$$\left.\begin{array}{l} \bullet\ X \sim \mathcal{P}(\lambda) \text{ et } Y \sim \mathcal{P}(\mu) \\ \bullet\ X \perp\!\!\!\perp Y \end{array}\right. \quad \Rightarrow \quad X+Y \sim \mathcal{P}(\lambda+\mu)$$

Démonstration.

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

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

Or, d'après le théorème précédent :

$$\begin{aligned} \forall t \in]-\infty,+\infty[, G_{X+Y}(t) &= G_X(t) \times G_Y(t) \\ &= \big((1-p)+p\,t\big)^m \times \big((1-p)+p\,t\big)^n \\ &= \big((1-p)+p\,t\big)^{m+n} \end{aligned}$$

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 :

$$\begin{aligned} \forall t \in]-\infty,+\infty[,\ G_X(t) &= \sum_{k=0}^{+\infty} \mathbb{P}\big(\{X=k\}\big)\ t^k \\ &= \sum_{k=0}^{+\infty} \mathrm{e}^{-\lambda}\ \frac{\lambda^k}{k!}\ t^k \\ &= \mathrm{e}^{-\lambda} \sum_{k=0}^{+\infty} \frac{(\lambda\,t)^k}{k!} \\ &= \mathrm{e}^{-\lambda}\ \mathrm{e}^{\lambda\,t} \\ &= \mathrm{e}^{\lambda\,(t-1)} \end{aligned}$$

Or, d'après le théorème précédent :

$$\begin{aligned} \forall t \in]-\infty,+\infty[, G_{X+Y}(t) &= G_X(t) \times G_Y(t) \\ &= \mathrm{e}^{\lambda\,(t-1)} \times \mathrm{e}^{\mu\,(t-1)} \\ &= \mathrm{e}^{(\lambda+\mu)\,(t-1)} \end{aligned}$$

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

Exercice
Soient $X_1, \dots, X_k$ des v.a. mutuellement indépendantes.
Déterminer la loi de $X_1 + \dots + X_k$ lorsque :

(i) $X_1 \sim \mathcal{B}(p)$, …, $X_k \sim \mathcal{B}(p)$.
(ii) $X_1 \sim \mathcal{B}(n_1,p)$, …, $X_k \sim \mathcal{B}(n_k,p)$.
(iii) $X_1 \sim \mathcal{P}(\lambda_1)$, …, $X_k \sim \mathcal{P}(\lambda_k)$.

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

VIII.3.a) Cas général

Théorème 17.

Soient X et Y deux v.a. finies 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 :*

$$\begin{array}{rcccl} \min(X,Y) & : & \Omega & \to & \mathbb{R} \\ & & \omega & \mapsto & \min\big(X(\omega), Y(\omega)\big) \end{array}$$

$$\begin{array}{rcccl} \max(X,Y) & : & \Omega & \to & \mathbb{R} \\ & & \omega & \mapsto & \max\big(X(\omega), Y(\omega)\big) \end{array}$$

1) *Remarquons tout d'abord :*

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

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

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

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

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

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

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

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

Démonstration.

- Soit $t \in \mathbb{R}$.

$$\begin{array}{cll} & \mathbb{P}\big(\{\min(X,Y) > t\}\big) & \\ = & \mathbb{P}\big(\{X > t\} \cap \{Y > t\}\big) & \\ = & \mathbb{P}\big(\{X > t\}\big)\ \mathbb{P}\big(\{Y > t\}\big) & \textit{(car } X \textit{ et } Y \textit{ sont indépendantes)} \\ = & \Big(1 - \mathbb{P}\big(\{X \leqslant t\}\big)\Big)\ \Big(1 - \mathbb{P}\big(\{Y \leqslant t\}\big)\Big) & \end{array}$$

- Soit $t \in \mathbb{R}$.

$$\begin{array}{rcll} \mathbb{P}\big(\{\max(X,Y) \leqslant t\}\big) & = & \mathbb{P}\big(\{X \leqslant t\} \cap \{Y \leqslant t\}\big) & \\ & = & \mathbb{P}\big(\{X \leqslant t\}\big)\ \mathbb{P}\big(\{Y \leqslant t\}\big) & \textit{(car } X \textit{ et } Y \textit{ sont indépendantes)} \end{array}$$

□

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

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

Théorème 18. *(théorème de transfert)*

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

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

1) $\boxed{\text{La v.a. } g(Z) \text{ est finie}}$

2) De plus :

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

IX.2. Espérance d'une somme

Théorème 19. *(linéarité de l'espérance)*

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

Soient $X_1, \dots, X_n$ des v.a. discrètes d'espérances finies.

1) Alors $\lambda_1 X_1 + \ldots + \lambda_n X_n$ est d'espérance finie.

2) De plus :

$$\boxed{\mathbb{E}\Big(\sum_{i=1}^{n} \lambda_i X_i\Big) = \sum_{i=1}^{n} \lambda_i\, \mathbb{E}(X_i)}$$

En particulier, lorsque pour tout $i \in [\![1,n]\!]$, $\lambda_i = 1$, on obtient :

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

Exemple

Soient X_1, …, X_n des v.a. discrètes.

On suppose que les v.a. X_1, …, X_n sont mutuellement indépendantes.

On suppose de plus : $\forall i \in [\![1,n]\!]$, $X_i \sim \mathcal{B}(p)$.

1) $S_n = X_1 + \ldots + X_n \sim \mathcal{B}(n,p)$

(S_n prend pour valeur le nombre de succès dans une répétition de n épreuves de Bernoulli indépendantes et de même paramètre p)

2) On retrouve l'espérance d'une v.a. qui suit la loi binomiale :

$$\mathbb{E}(S_n) = \mathbb{E}(X_1 + \ldots + X_n) = \mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n) = n\,p$$

IX.3. Espérance d'un produit

IX.3.a) Formulation générale

Théorème 20.

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

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

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

1) Alors XY est finie.

2) De plus :

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

Démonstration.

1) Remarquons tout d'abord : $\big(|X| - |Y|\big)^2 \geqslant 0$. On en déduit :

$$|XY| \ \leqslant\ \frac{1}{2}\ |X|^2 + \frac{1}{2}\ |Y|^2$$

La v.a. $\frac{1}{2}\ |X|^2 + \frac{1}{2}\ |Y|^2$ est d'espérance finie en tant que combinaison linéaire de v.a. d'espérances finies ($|X|^2$ et $|Y|^2$ sont d'espérances finies car X et Y sont de moments d'ordre 2 finis).

2) Par théorème de transfert. □

IX.3.b) Cas des v.a. indépendantes

Théorème 21.

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

On suppose : × *X et Y sont indépendantes.*

Alors $X\,Y$ est d'espérance finie et :

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

⚠ Ce théorème n'énonce pas un équivalence.
Autrement dit, il existe des variables aléatoires X et Y :

- × qui ne sont pas indépendantes,
- × qui vérifient $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$.

Démonstration.

1) • Il s'agit de démontrer que la v.a. $|XY|$ est d'espérance finie. Or :

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

• Comme X et Y sont d'espérances finies, alors :

$$\mathbb{E}\big(|X|\big) < +\infty \quad \text{et} \quad \mathbb{E}\big(|Y|\big) < +\infty$$

Ainsi : $\mathbb{E}\big(|XY|\big) < +\infty$ et la v.a. XY est donc d'espérance finie.

2) On obtient le résultat souhaité par théorème de transfert avec des calculs similaires à ceux effectués en ***1)***. □

Exercice

Soient X_1, X_2 et X_3 indépendantes et de même loi $\mathcal{B}(p)$ où $p \in \,]0,1[$.

Montrer que les v.a. $Y_1 = X_1X_2$ et $Y_2 = X_2X_3$ ne sont pas indépendantes.

Démonstration.

- Comme X_1 et X_2 sont indépendantes et sont d'espérances finies, la v.a. $Y_1 = X_1\,X_2$ est d'espérance finie et :

$$\mathbb{E}(Y_1) = \mathbb{E}(X_1\,X_2) = \mathbb{E}(X_1)\,\mathbb{E}(X_2)$$

- De même, Y_2 est d'espérance finie et :

$$\mathbb{E}(Y_2) = \mathbb{E}(X_2\,X_3) = \mathbb{E}(X_2)\,\mathbb{E}(X_3)$$

- Comme $X_2 \sim \mathcal{B}(p)$, ${X_2}^2 = X_2$. Et ainsi :

$$Y_1\,Y_2 = X_1\,X_2\,X_2\,X_3 = X_1\,{X_2}^2\,X_3 = X_1\,X_2\,X_3$$

  Comme X_1, X_2, X_2 sont indépendantes, par le lemme des coalitions les v.a. $X_1\,X_2$ et X_3 sont indépendantes. Ces deux v.a. sont d'espérances finies. On en déduit que $Y_1\,Y_2$ est d'espérance finie et :

$$\begin{aligned}
\mathbb{E}(Y_1\,Y_2) &= \mathbb{E}(X_1\,X_2\,X_3) \\
&= \mathbb{E}(X_1\,X_2)\,\mathbb{E}(X_3) \\
&= \mathbb{E}(X_1)\,\mathbb{E}(X_2)\,\mathbb{E}(X_3) \;=\; p^3
\end{aligned}$$

- $$\begin{aligned}
\mathbb{E}(Y_1)\,\mathbb{E}(Y_2) &= \mathbb{E}(X_1\,X_2)\,\mathbb{E}(X_2\,X_3) \\
&= \mathbb{E}(X_1)\,\mathbb{E}(X_2)\,\mathbb{E}(X_2)\,\mathbb{E}(X_3) \;=\; p^4
\end{aligned}$$

- Or :

$$p^3 = p^4 \;\Leftrightarrow\; p^3\,(p-1) = 0 \;\Leftrightarrow\; p = 0 \;\;\texttt{OU}\;\; p = 1$$

  Ainsi, $\mathbb{E}(Y_1\,Y_2) \neq \mathbb{E}(Y_1)\,\mathbb{E}(Y_2)$ et les v.a. ne sont pas indépendantes. □

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

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

X.1.a) Définition

Définition

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

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

$$\boxed{\text{Cov}(X,Y) = \mathbb{E}\Big((X - \mathbb{E}(X))\ (Y - \mathbb{E}(Y))\Big)}$$

X.1.b) Calcul en pratique

Théorème 22. *(formule de Kœnig-Huygens)*

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

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

En particulier, on a :

$$\boxed{\text{Cov}(X,X) = \mathbb{V}(X)}$$

Démonstration.

- Par définition :

$$\begin{aligned}
& \text{Cov}(X,Y) \\
= \ & \mathbb{E}\big((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\big) \\
= \ & \mathbb{E}\big(XY - \mathbb{E}(Y)X - \mathbb{E}(X)Y + \mathbb{E}(X)\mathbb{E}(Y)\big) \\
= \ & \mathbb{E}(XY) - \mathbb{E}(Y)\mathbb{E}(X) - \mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(X)\mathbb{E}(Y) \quad \textit{(par linéarité de l'espérance)} \\
= \ & \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)
\end{aligned}$$

- $\text{Cov}(X,X) = \mathbb{E}\big((X - \mathbb{E}(X))(X - \mathbb{E}(X))\big) = \mathbb{E}\big((X - \mathbb{E}(X))^2\big) = \mathbb{V}(X)$ □

Exemple

On reprend l'exemple des v.a. Z_1 et Z_2 qui représentent respectivement le résultat de premier dé et du deuxième dé lors du lancer de deux dés à 3 faces.

$$\text{Cov}(Z_1, Z_2) \ = \ \mathbb{E}(Z_1\, Z_2) - \mathbb{E}(Z_1)\, \mathbb{E}(Z_2) \ = \ 4 - 2 \times 2 \ = \ 0$$

(c'est toujours le cas lorsque Z_1 et Z_2 sont indépendantes, puisqu'alors on a : $\mathbb{E}(Z_1 Z_2) = \mathbb{E}(Z_1)\ \mathbb{E}(Z_2)$)

X.1.c) Propriétés de la covariance

Théorème 23.

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

L'opérateur de covariance vérifie les propriétés suivantes.

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

(propriété de symétrie)

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

(linéarité à gauche)

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

(linéarité à droite)

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

Démonstration.

1) D'après la formule de Kœnig-Huygens :

$$\begin{aligned}
\text{Cov}(Y,X) &= \mathbb{E}(Y X) - \mathbb{E}(Y)\, \mathbb{E}(X) \\
&= \mathbb{E}(XY) - \mathbb{E}(X)\, \mathbb{E}(Y) \\
&= \text{Cov}(X,Y)
\end{aligned}$$

2) Soit $(\lambda, \mu) \in \mathbb{R}^2$. Comme X_1 et X_2 sont de moments d'ordre 2 finis, c'est aussi le cas de $\lambda\, X_1 + \mu\, X_2$. De plus :

$$\begin{aligned}
& \mathrm{Cov}(\lambda\, X_1 + \mu\, X_2, Y) \\
= \ & \mathbb{E}((\lambda\, X_1 + \mu\, X_2)\, Y) - \mathbb{E}(\lambda\, X_1 + \mu\, X_2)\, \mathbb{E}(Y) && \textit{(par la formule de Kœnig)} \\
= \ & \mathbb{E}(\lambda\, X_1\, Y + \mu\, X_2\, Y) - (\lambda\, \mathbb{E}(X_1) + \mu\, \mathbb{E}(X_2))\, \mathbb{E}(Y) && \textit{(par linéarité de l'espérance)} \\
= \ & \lambda\, \mathbb{E}(X_1\, Y) + \mu\, \mathbb{E}(X_2\, Y) - \lambda\, \mathbb{E}(X_1)\, \mathbb{E}(Y) - \mu\, \mathbb{E}(X_2)\, \mathbb{E}(Y) && \textit{(par linéarité de l'espérance)} \\
= \ & \lambda\, (\mathbb{E}(X_1\, Y) - \mathbb{E}(X_1)\, \mathbb{E}(Y)) + \mu\, (\mathbb{E}(X_2\, Y) - \mathbb{E}(X_2)\, \mathbb{E}(Y)) \\
= \ & \lambda\, \mathrm{Cov}(X_1, Y) + \mu\, \mathrm{Cov}(X_2, Y) && \textit{(par la formule de Kœnig)}
\end{aligned}$$

3) Soit $(\lambda, \mu) \in \mathbb{R}^2$.

$$\begin{aligned}
& \mathrm{Cov}(X, \lambda\, Y_1 + \mu\, Y_2) \\
= \ & \mathrm{Cov}(\lambda\, Y_1 + \mu\, Y_2, X) && \textit{(par symétrie)} \\
= \ & \lambda\, \mathrm{Cov}(Y_1, X) + \mu\, \mathrm{Cov}(Y_2, X) && \textit{(par le point 2))}
\end{aligned}$$

4) Soit $a \in \mathbb{R}$. $\mathrm{Cov}(X, a) = \mathbb{E}(aX) - \mathbb{E}(a)\, \mathbb{E}(X) = a\, \mathbb{E}(X) - a\, \mathbb{E}(X) = 0$.

□

X.1.d) Une condition nécessaire d'indépendance

Théorème 24.

Soient X et Y deux v.a. (discrètes).

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

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

Démonstration.
On utilise la formule de Kœnig-Huygens :

$$\mathrm{Cov}(X, Y) = \mathbb{E}(X\, Y) - \mathbb{E}(X)\, \mathbb{E}(Y)$$

Or, comme les v.a. X et Y sont indépendantes, $\mathbb{E}(XY) = \mathbb{E}(X)\ \ E(Y)$.
On en déduit que $\mathrm{Cov}(X, Y) = 0$. □

Remarque

- Généralement c'est la contraposée de cet énoncé qui est utilisée. Elle permet de démontrer que deux v.a. X et Y ne sont pas indépendantes.

$$\boxed{\mathrm{Cov}(X, Y) \neq 0 \quad \Rightarrow \quad X \text{ et } Y \text{ ne sont pas indépendantes}}$$

Ce résultat `N'EST PAS` une équivalence. Autrement dit, il existe des variables aléatoires X et Y :

× qui ne sont pas indépendantes,

× qui vérifient $\mathrm{Cov}(X, Y) = 0$.

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

X.2.a) Cas général

Théorème 25.

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

1) Alors $X + Y$ est finie.

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

Généralisation

Soient X_1, ..., X_n des v.a. finies réelles.

1) Alors $X_1 + \ldots + X_n$ est finie.

2) De plus :

$$\begin{array}{rcl} \mathbb{V}\big(X_1 + \ldots + X_n\big) & = & \sum\limits_{i=1}^{n} \mathbb{V}(X_i) \ + \sum\limits_{1 \leqslant i,j \leqslant n} \text{Cov}(\ X_i\ ,\ X_j\ ) \\ & = & \sum\limits_{i=1}^{n} \mathbb{V}(X_i) \ + \ 2 \sum\limits_{1 \leqslant i < j \leqslant n} \text{Cov}(\ X_i\ ,\ X_j\ ) \end{array}$$

Démonstration.

1) Déjà vu précédemment.

2) On utilise la formule de Kœnig-Huygens :

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

□

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

Théorème 26.

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

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

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

2) **Généralisation**

On suppose :
- × *$X_1, \ldots, X_n$ sont finies.*
- × *$X_1, \ldots, X_n$ mutuellement indépendantes.*

Alors $X_1 + \ldots + X_n$ est finie et :

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

Démonstration.

1) Si X et Y sont indépendantes, alors $\text{Cov}(X,Y) = 0$.
Donc, d'après le théorème précédent, $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

2) Par récurrence. □

Exemple

Soient X_1, ..., X_n des v.a. finies.

On suppose que les v.a. X_1, ..., X_n sont mutuellement indépendantes.

On suppose de plus : $\forall i \in [\![1, n]\!]$, $X_i \sim \mathcal{B}(p)$.

1) $S_n = X_1 + \ldots + X_n \sim \mathcal{B}(n, p)$
(S_n prend pour valeur le nombre de succès dans une répétition de n épreuves de Bernoulli indépendantes et de même paramètre p)

2) On retrouve la variance d'une v.a. qui suit la loi binomiale :

$$\mathbb{V}(X) \ = \ \mathbb{V}(X_1 + \ldots + X_n) \ = \ \mathbb{V}(X_1) + \ldots + \mathbb{V}(X_n) = n\,p\,q$$

Bilan sur les couples : liens entre les différentes lois

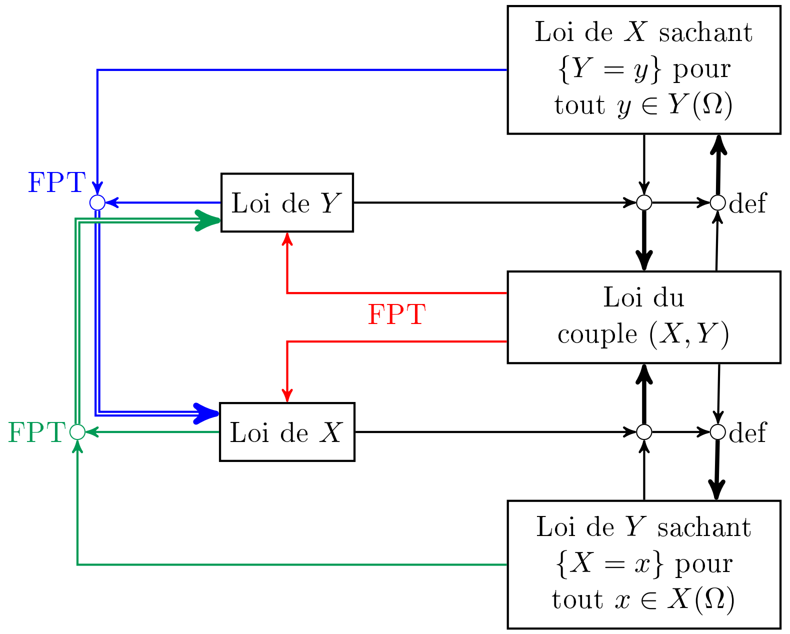
Il faut savoir :

× faire le lien entre la loi de couples et les lois conditionnelles.

× déterminer les lois marginales si on connaît la loi du couple,

× déterminer les lois marginales si on connaît les lois conditionnelles.

Les liens en ces différentes notions sont rappelés dans le schéma suivant.



On retiendra que la formule des probabilités totales est la clé pour déterminer les lois marginales. Si la détermination de la loi du couple / les lois conditionnelles constitue généralement la plus grande difficulté d'un exercice sur les couples, la détermination des lois marginales se résume simplement à une application de la formule des probabilités totales. Des difficultés calculatoires peuvent apparaître mais il n'y a aucune difficulté méthodologique.

XI. Inégalités de concentration

XI.1. L'inégalité de Markov

Théorème 27 (inégalité de Markov)**.**

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

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

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

Démonstration.
Démonstration adaptée au programme

Soit $a > 0$.

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

  On en déduit :
$$\begin{aligned} \mathbb{P}\big(\{X \geqslant a\}\big) &= \mathbb{P}\Big(\bigcup_{\substack{x \in X(\Omega) \\ x \geqslant a}} \{X = x\}\Big) \\ &= \sum_{\substack{x \in X(\Omega) \\ x \geqslant a}} \mathbb{P}\big(\{X = x\}\big) \end{aligned}$$

- Comme X est une variable positive, $\mathbb{E}(X) \in [0, +\infty]$ et :
$$\begin{aligned} \mathbb{E}(X) &= \sum_{x \in X(\Omega)} x \ \mathbb{P}\big(\{X = x\}\big) \\ &= \sum_{\substack{x \in X(\Omega) \\ x \geqslant a}} x \ \mathbb{P}\big(\{X = x\}\big) + \sum_{\substack{x \in X(\Omega) \\ x < a}} x \ \mathbb{P}\big(\{X = x\}\big) \\ &\geqslant \sum_{\substack{x \in X(\Omega) \\ x \geqslant a}} x \ \mathbb{P}\big(\{X = x\}\big) \qquad \textit{(car } X \textit{ est à valeurs positives)} \end{aligned}$$

- Finalement :

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

Démonstration adaptée à toute variable aléatoire (CULTURE)

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

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

Listons quelques propriétés de Y :

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

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

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

2) On remarque aussi : $Y \leqslant X$ (c'est-à-dire : $\forall \omega \in \Omega$, $Y(\omega) \leqslant X(\omega)$).
Démontrons-le. Soit $\omega \in \Omega$.
Deux cas se présentent.
× si $X(\omega) \geqslant a$ alors $Y(\omega) = a \leqslant X(\omega)$.
× si $X(\omega) < a$ alors $Y(\omega) = 0 \leqslant X(\omega)$.

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

$$\underset{a\, \mathbb{P}(\{X \geqslant a\})}{\underset{\shortparallel}{\mathbb{E}(Y)}} \leqslant \mathbb{E}(X)$$

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

Remarque

- Notons que si $a \leqslant \mathbb{E}(X)$, alors l'inégalité est évidente (et donc peu intéressante). En effet, dans ce cas, on a directement : $\mathbb{P}(\{X \geqslant a\}) \leqslant 1 \leqslant \dfrac{\mathbb{E}(X)}{a}$.
Dans le cas où a est « grand » par rapport à $\mathbb{E}(X)$, on obtient un résultat de bon sens : la probabilité que X prenne de grandes valeurs est d'autant plus faible que celles-ci excèdent $\mathbb{E}(X)$.
- On utilise dans cette démonstration la notation $\mathbb{1}_A$. Si A est un événement, $\mathbb{1}_A$ est la variable indicatrice de l'événement A. Plus précisément, $\mathbb{1}_A$ est définie par :

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

Avec cette notation, on peut écrire : $Y = a \cdot \mathbb{1}_{\{X \geqslant a\}}$.
- Dans cette démonstration, on ne suppose pas que X est d'espérance finie. Les v.a. en jeu étant à valeurs positives, cette précaution n'a pas lieu d'être. Dans le cas où X est d'espérance infinie, l'énoncé n'a aucun intérêt : la valeur $\mathbb{P}\big(\{X \geqslant a\}\big)$ est évidemment finie puisque majorée par 1.

Théorème 28.

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

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

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

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

$$\underset{\mathbb{P}(\{|X| \geqslant a\})}{\underset{\shortparallel}{\mathbb{P}(\{|X|^m \geqslant a^m\})}} \leqslant \frac{\mathbb{E}(|X|^m)}{a^m}$$

□

XI.2. L'inégalité de Bienaymé-Tchebychev

Théorème 29 (Inégalité de Bienaymé-Tchebychev)**.**
Soit X une v.a. (discrète) à valeurs réelles.
On suppose que X est de variance finie.
Alors :

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

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

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

□

Remarque

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