Chapitre 5

Probabilités

5.1 Couples de variables aléatoires - Variables aléatoires indépendantes

5.1.1 Premiers rappels

En première année, nous avons étudié les probabilités sur un ***univers fini*** Ω associé à une expérience aléatoire. On a défini un ***évènement*** A comme une partie de l'univers, $A \in \mathcal{P}(\Omega)$ (l'évènement A étant dit ***réalisé*** si le résultat de l'expérience appartient à A) et on a introduit le vocabulaire associé aux opérations sur les évènements : pour deux évènements A et B donnés,

- $\overline{A} = \Omega \backslash A$ (aussi noté A^c) est l'***évènement contraire*** de A
- $A \cap B$ est l'évènement A ET B
- $A \cup B$ est l'évènement A OU B
- A et B sont dits ***incompatibles*** ou disjoints si $A \cap B = \varnothing$.
- $\varnothing$ est appelé ***évènement impossible*** et Ω ***évènement certain***.

On a ensuite associé à chaque évènement A une valeur $\mathbb{P}(A)$ appelée probabilité de A (intuitivement, la probabilité d'un évènement est la fréquence d'apparition de cet évènement lorsqu'on réalise une infinité de fois, dans les mêmes conditions, l'expérience aléatoire). On a ainsi défini une application $\mathbb{P} : \mathcal{P}(\Omega) \longrightarrow \mathbb{R}_+$ vérifiant $\mathbb{P}(\Omega) = 1$ (axiome de totalité) et, pour deux évènements A_1 et A_2 disjoints (i.e. $A_1 \cap A_2 = \varnothing$), $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2)$.
Le triplet $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ a alors été appelé ***espace probabilisé fini***.
Dans ce cadre, on a défini une ***variable aléatoire réelle*** X comme étant une application $X : \Omega \longrightarrow \mathbb{R}$. On a introduit les notations

$$[X = x] = \{\omega \in \Omega,\, X(\omega) = x\} = X^{-1}(\{k\}) \qquad [X \in E] = \{\omega \in \Omega,\, X(\omega) \in E\} = X^{-1}(\{E\})$$

Remarque importante : On a toujours $\Omega = \bigcup_{x \in X(\Omega)} [X = x]$ (réunion disjointe).

La ***loi*** de la variable aléatoire X (appelée également ***probabilité image***) est l'application $\mathbb{P}_X : \mathcal{P}(X(\Omega)) \longrightarrow \mathbb{R}^+$, $E \longmapsto \mathbb{P}([X \in E])$. $\mathbb{P}_X$ est une probabilité sur $X(\Omega)$.

Comme $[X \in E] = \bigcup_{x \in E} [X = x]$ (Exercice : démontrer cette égalité par double inclusion.) et que cette réunion est disjointe, $\mathbb{P}([X \in E]) = \sum_{x \in E} \mathbb{P}([X = x])$ et donner la loi de X revient donc à donner $X(\Omega)$ et tous les $\mathbb{P}([X = x])$ pour $x \in X(\Omega)$.

5.1.2 Couples de variables aléatoires

Définition 5.1. *Soit* X *et* Y *deux variables aléatoires sur un espace probabilisé fini* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *à valeurs dans* $\mathbb{R}$.

- *On appelle* ***couple de variables aléatoires*** (X, Y) *l'application*

$$\begin{aligned}(\mathrm{X},\mathrm{Y}) : \Omega &\longrightarrow \mathbb{R}\times\mathbb{R}\\ \omega &\longmapsto (\mathrm{X}(\omega),\mathrm{Y}(\omega))\end{aligned}$$

  (X, Y) *est donc une variable aléatoire à valeurs dans* $\mathbb{R}^2$.

- *On appelle* ***loi conjointe*** *des variables aléatoires* X *et* Y *la loi du couple* (X, Y). *Cette loi est donnée par l'application*

$$\begin{aligned}(\mathrm{X},\mathrm{Y})(\Omega) &\longrightarrow [0,1]\\ (x,y) &\longmapsto \mathbb{P}([(\mathrm{X},\mathrm{Y}) = (x,y)])\end{aligned}$$

- *Les lois des variables aléatoires* X *et* Y *sont appelées* ***lois marginales*** *du couple* (X, Y).

Remarques.

- On peut aussi prendre en définition de la loi conjointe des variables aléatoires X et Y l'application $\mathrm{X}(\Omega)\times\mathrm{Y}(\Omega) \longrightarrow [0,1]$, $(x,y) \longmapsto \mathbb{P}([(\mathrm{X},\mathrm{Y}) = (x,y)])$ étant entendu qu'un couple (x,y) de $\mathrm{X}(\Omega)\times\mathrm{Y}(\Omega)$ n'est pas toujours une valeur prise par le couple (X,Y). En effet, un couple (x,y) de $\mathrm{X}(\Omega)\times\mathrm{Y}(\Omega)$ est un couple de la forme $(\mathrm{X}(\omega),\mathrm{Y}(\omega'))$ où $(\omega,\omega') \in \Omega^2$ alors qu'un couple de $(\mathrm{X},\mathrm{Y})(\Omega)$ est un couple de la forme $(\mathrm{X}(\omega),\mathrm{Y}(\omega))$ où $\omega \in \Omega$. Quelque soit la définition adoptée, l'évènement $[(\mathrm{X},\mathrm{Y}) = (x,y)]$ est l'évènement $[\mathrm{X} = x]\cap[\mathrm{Y} = y]$. Cet évènement est impossible si $(x,y) \notin (\mathrm{X},\mathrm{Y})(\Omega)$ et sa probabilité est donc nulle.
- En résumé, donner la loi conjointe des variables X et Y, c'est donner $\mathrm{X}(\Omega)$ et $\mathrm{Y}(\Omega)$ ainsi que tous les $\mathrm{P}([\mathrm{X} = x]\cap[\mathrm{Y} = y]), (x,y) \in \mathrm{X}(\Omega)\times\mathrm{Y}(\Omega)$. Lorsque cela sera possible, on présentera la loi de (X,Y) sous forme d'un tableau à double entrée.

Exemple. Une urne contient trois boules blanches, deux boules rouges et trois boules vertes indiscernables au toucher. On extrait au hasard et simultanément trois boules de l'urne et on en observe les couleurs. On note X (respectivement Y) la variable aléatoire égale au nombre de boules blanches (respectivement rouges) parmi ces trois boules. Déterminons la loi du couple (X,Y).
On introduit l'ensemble Ω des parties à trois éléments de l'ensemble des huit boules (de cardinal $\binom{8}{3} = 56$) que l'on munit de la probabilité uniforme (boules indiscernables au toucher).
La loi conjointe de X et Y est donnée par $\mathrm{X}(\Omega) = [\![0,3]\!]$, $\mathrm{Y}(\Omega) = [\![0,2]\!]$ et le tableau des $\mathbb{P}([\mathrm{X} = k]\cap[\mathrm{Y} = \ell])$:

k\\ℓ	0	1	2
0	$\frac{1}{56}$	$\frac{6}{56}$	$\frac{3}{56}$
1	$\frac{9}{56}$	$\frac{18}{56}$	$\frac{3}{56}$
2	$\frac{9}{56}$	$\frac{6}{56}$	0
3	$\frac{1}{56}$	0	0

D'une manière générale, puisque $\mathrm{X}(\Omega)$ et $\mathrm{Y}(\Omega)$ sont finis, on peut poser $\mathrm{X}(\Omega) = \{x_i, i \in \mathrm{I}\}$ et $\mathrm{Y}(\Omega) = \{y_j, j \in \mathrm{J}\}$ où I est un ensemble de la forme $[\![1,p]\!]$, $(p \in \mathbb{N}^*)$ et J est un ensemble de la forme $[\![1,q]\!]$, $(q \in \mathbb{N}^*)$. Pour $(i,j) \in \mathrm{I}\times\mathrm{J}$, on pose alors $p_{i,j} = \mathbb{P}([\mathrm{X} = x_i]\cap[\mathrm{Y} = y_j])$.

Proposition 5.2.
1) Les évènements $[\mathrm{X} = x_i] \cap [\mathrm{Y} = y_j]$ *pour* $(i,j) \in \mathrm{I} \times \mathrm{J}$ *sont deux à deux incompatibles et leur réunion est égale à* Ω*. En particulier,* $\displaystyle\sum_{(i,j)\in \mathrm{I}\times\mathrm{J}} p_{i,j} = 1$
2) $\forall i \in \mathrm{I}$, $\mathbb{P}([\mathrm{X} = x_i]) = \displaystyle\sum_{j\in\mathrm{J}} p_{i,j}$ *et* $\forall j \in \mathrm{J}$, $\mathbb{P}(\mathrm{Y} = y_j) = \displaystyle\sum_{i\in\mathrm{I}} p_{i,j}$.

Démonstration :
• 1). Vérifions que $([\mathrm{X} = x_i] \cap [\mathrm{Y} = y_j])_{(i,j)\in\mathrm{I}\times\mathrm{J}}$ est un système complet d'évènements.
Soit $((i,j),(k,l)) \in (\mathrm{I}\times\mathrm{J})^2$ tel que $(i,j) \neq (k,l)$. On a ou bien $i \neq k$ et dans ce cas, $[\mathrm{X} = x_i] \cap [\mathrm{X} = x_k] = \varnothing$ puis $([\mathrm{X} = x_i] \cap [\mathrm{Y} = y_j]) \cap ([\mathrm{X} = x_k] \cap [\mathrm{Y} = y_l]) = \varnothing$, ou bien $j \neq l$ et dans ce cas, $[\mathrm{Y} = y_j] \cap [\mathrm{Y} = y_l] = \varnothing$ puis $([\mathrm{X} = x_i] \cap [\mathrm{Y} = y_j]) \cap ([\mathrm{X} = x_k] \cap [\mathrm{Y} = y_l]) = \varnothing$. Donc, les évènements $[\mathrm{X} = x_i] \cap [\mathrm{Y} = y_j], ((i,j),(k,l)) \in (\mathrm{I}\times\mathrm{J})^2$, sont deux à deux disjoints. Soit $\omega \in \Omega$. Puisque $\mathrm{X}(\omega) \in \mathrm{X}(\Omega)$ et $\mathrm{Y}(\omega) \in \mathrm{Y}(\Omega)$, on peut considérer un $(i_0, j_0) \in \mathrm{I}\times\mathrm{J}$ tel que $\mathrm{X}(\omega) = x_{i_0}$ et $\mathrm{Y}(\omega) = y_{j_0}$ et donc tel que $\omega \in [\mathrm{X} = x_{i_0}] \cap [\mathrm{Y} = y_{j_0}]$. Ceci montre que tout ω appartient à $\displaystyle\bigcup_{(i,j)\in\mathrm{I}\times\mathrm{J}} [\mathrm{X} = x_i] \cap [\mathrm{Y} = y_j]$ puis (l'inclusion réciproque étant claire) que $\Omega = \displaystyle\bigcup_{(i,j)\in\mathrm{I}\times\mathrm{J}} [\mathrm{X} = x_i] \cap [\mathrm{Y} = y_j]$. On a montré que $([\mathrm{X} = x_i] \cap [\mathrm{Y} = y_j])_{(i,j)\in\mathrm{I}\times\mathrm{J}}$ est un système complet d'évènements.

• 2). Soit $i \in \mathrm{I}$. $[\mathrm{X} = x_i] = [\mathrm{X} = x_i] \cap \left(\displaystyle\bigcup_{j\in\mathrm{J}} [\mathrm{Y} = y_j]\right) = \displaystyle\bigcup_{j\in\mathrm{J}} ([\mathrm{X} = x_i] \cap [\mathrm{Y} = y_j])$ et ces évènements sont deux à deux disjoints. Donc, $\mathbb{P}([\mathrm{X} = x_i]) = \displaystyle\sum_{j\in\mathrm{J}} \mathbb{P}([\mathrm{X} = x_i] \cap [\mathrm{Y} = y_j]) = \sum_{j\in\mathrm{J}} p_{i,j}$. De même, pour tout $j \in \mathrm{J}$, $\mathbb{P}([\mathrm{Y} = y_j]) = \displaystyle\sum_{i\in\mathrm{I}} \mathbb{P}([\mathrm{X} = x_i] \cap [\mathrm{Y} = y_j]) = \sum_{i\in\mathrm{I}} p_{i,j}$. □

Exemple. Dans l'exemple précédent, on peut alors compléter le tableau pour obtenir les lois marginales :

$k\backslash\ell$	0	1	2	$\mathbb{P}([\mathrm{X}=k])$
0	$\frac{1}{56}$	$\frac{6}{56}$	$\frac{3}{56}$	$\frac{10}{56}$
1	$\frac{9}{56}$	$\frac{18}{56}$	$\frac{3}{56}$	$\frac{30}{56}$
2	$\frac{9}{56}$	0	0	$\frac{9}{56}$
3	$\frac{1}{56}$	$\frac{6}{56}$	0	$\frac{7}{56}$
$\mathbb{P}([\mathrm{Y} = \ell])$	$\frac{20}{56}$	$\frac{30}{56}$	$\frac{6}{56}$	Somme = 1

5.1.3 Lois conditionnelles

Définition 5.3. *Soit* X *et* Y *deux variables aléatoires réelles sur un espace probabilisé fini* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$*. Soit* $y \in \mathrm{Y}(\Omega)$ *tel que* $\mathbb{P}([\mathrm{Y} = y]) \neq 0$*. La* ***loi conditionnelle*** *de X sachant* $[\mathrm{Y} = y]$ *est la loi de la variable aléatoire X dans l'espace probabilisé* $(\Omega, \mathcal{P}(\Omega), \mathbb{P}_{[\mathrm{Y}=y]})$ *c'est-à-dire :*

$$\forall x \in \mathrm{X}(\Omega),\ \mathbb{P}_{[\mathrm{Y}=y]}([\mathrm{X} = x]) = \frac{\mathbb{P}([\mathrm{X} = x] \cap [\mathrm{Y} = y])}{\mathbb{P}([\mathrm{Y} = y]}$$

On définit de même la ***loi conditionnelle*** *de Y sachant* $[\mathrm{X} = x]$ *pour un* $x \in \mathrm{X}(\Omega)$ *tel que* $\mathrm{P}([\mathrm{X} = x]) \neq 0$.

La proposition suivante est immédiate. Sa démonstration est laissée en exercice.

Proposition 5.4. *Pour tout* y *de* $\mathrm{Y}(\Omega)$ *tel que* $\mathbb{P}([\mathrm{Y} = y]) \neq 0$,

$$\forall x \in \mathrm{X}(\Omega), \mathbb{P}([\mathrm{X} = x] \cap [\mathrm{Y} = y]) = \mathbb{P}([\mathrm{Y} = y]) \times \mathbb{P}_{[\mathrm{Y}=y]}([\mathrm{X} = x]).$$

Si pour tout y de $\mathrm{Y}(\Omega), \mathbb{P}([\mathrm{Y}=y]) \neq 0$*, alors*

$$\forall x \in \mathrm{X}(\Omega), \mathbb{P}([\mathrm{X}=x]) = \sum_{y \in \mathrm{Y}(\Omega)} \mathbb{P}([\mathrm{X}=x] \cap [\mathrm{Y}=y]) = \sum_{y \in \mathrm{Y}(\Omega)} \mathbb{P}([\mathrm{Y}=y]) \times \mathbb{P}_{[\mathrm{Y}=y]}([\mathrm{X}=x])$$

5.1.4 Généralisation à un n-uplet de variables aléatoires

On généralise facilement ce qui précède à n variables aléatoires $\mathrm{X}_1, \cdots, \mathrm{X}_n$ $(n \geqslant 2)$.

Définition 5.5. *Soit* $\mathrm{X}_1, \cdots, \mathrm{X}_n$ $(n \geqslant 2)$*, des variables aléatoires réelles sur un même espace probabilisé* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.

- *L'application* $\Omega \to \mathbb{R}^n$, $\omega \mapsto (\mathrm{X}_1(\omega), ..., \mathrm{X}_n(\omega))$ *est appelée n-uplet de variables aléatoires sur l'espace* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *et se note* $(\mathrm{X}_1, ..., \mathrm{X}_n)$*. C'est une variable aléatoire sur l'espace* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *à valeurs dans* $\mathbb{R}^n$.
- *La **loi conjointe** des variables aléatoires* $\mathrm{X}_1, \cdots, \mathrm{X}_n$ *est la loi du n-uplet* $(\mathrm{X}_1, \cdots, \mathrm{X}_n)$.
- *Les **lois marginales** du n-uplet* $(\mathrm{X}_1, \cdots, \mathrm{X}_n)$ *sont les lois des variables* $\mathrm{X}_1, \cdots, \mathrm{X}_n$.

Quand on dispose de la loi du n-uplet $(\mathrm{X}_1, \cdots, \mathrm{X}_n)$, les lois marginales s'obtiennent de la façon suivante : pour $a_k \in \mathrm{X}_k(\Omega)$ donné,

$$\mathbb{P}([\mathrm{X}_k = a_k]) = \sum \mathbb{P}([\mathrm{X}_1 = x_1, \cdots, \mathrm{X}_{k-1} = x_{k-1}, \mathrm{X}_k = a_k, \mathrm{X}_{k+1} = x_{k+1}, \cdots, \mathrm{X}_n = x_n])$$

la somme portant sur tous les $(x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n)$ de $\mathrm{X}_1(\Omega) \times \cdots \times \mathrm{X}_{k-1}(\Omega) \times \mathrm{X}_{k+1}(\Omega) \times \cdots \times \mathrm{X}_n(\Omega)$.

5.1.5 Variables aléatoires indépendantes

Rappels

- n évènements $\mathrm{A}_1, \cdots, \mathrm{A}_n$ sont dits ***mutuellement indépendants*** si pour tout $k \leqslant n$, pour tout $i_1, \cdots, i_k \in \{1, \cdots, n\}$ distincts on a :
$$\mathbb{P}(\mathrm{A}_{i_1} \cap \mathrm{A}_{i_2} \cap \cdots \cap \mathrm{A}_{i_k}) = \mathbb{P}(\mathrm{A}_{i_1})\mathbb{P}(\mathrm{A}_{i_2}) \cdots \mathbb{P}(\mathrm{A}_{i_k}).$$
Par exemple, A, B et C sont trois évènements mutuellement indépendants si on a à la fois $\mathbb{P}(\mathrm{A} \cap \mathrm{B}) = \mathbb{P}(\mathrm{A})\mathbb{P}(\mathrm{B})$, $\mathbb{P}(\mathrm{A} \cap \mathrm{C}) = \mathbb{P}(\mathrm{A})\mathbb{P}(\mathrm{C}), \mathbb{P}(\mathrm{B} \cap \mathrm{C}) = \mathbb{P}(\mathrm{B})\mathbb{P}(\mathrm{C})$ mais aussi $\mathbb{P}(\mathrm{A} \cap \mathrm{B} \cap \mathrm{C}) = \mathbb{P}(\mathrm{A})\mathbb{P}(\mathrm{B})\mathbb{P}(\mathrm{C})$.
- Si des évènements sont mutuellements indépendants, alors ils sont indépendants deux à deux, mais la réciproque est fausse.
Par exemple, on jette deux dés classiques parfaitement équilibrés, l'un rouge l'autre bleu et on considère les trois évènements suivants : $\mathrm{A} = \{$le dé rouge donne un chiffre impair$\}$ $\mathrm{B} = \{$le dé bleu donne un chiffre impair$\}$ $\mathrm{C} = \{$la somme des deux dés est impaire$\}$.
A, B et C sont deux à deux indépendants mais pas mutuellement indépendants.

Définition 5.6. *Soit* X *et* Y *deux variables aléatoires sur l'espace probabilisé* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$*. Les **variables aléatoires*** X *et* Y *sont dites **indépendantes** si*

$$\forall (x, y) \in \mathrm{X}(\Omega) \times \mathrm{Y}(\Omega) \qquad \mathbb{P}([\mathrm{X}=x] \cap [\mathrm{Y}=y]) = \mathbb{P}([\mathrm{X}=x]) \times \mathbb{P}([\mathrm{Y}=y])$$

Il revient au même de dire que pour tout $(x, y) \in \mathrm{X}(\Omega) \times \mathrm{Y}(\Omega)$*, les évènements* $[\mathrm{X}=x]$ *et* $[\mathrm{Y}=y]$ *sont indépendants.*

Exemple.

- Un urne contient n jetons numérotés de 1 à n. On en tire deux avec remise. Soit X et Y les variables aléatoires égales au premier et au second numéro tiré. On a :

$$\forall (i,j) \in [\![1,n]\!] \quad \mathbb{P}([\mathrm{X}=i]\cap[\mathrm{Y}=j]) = \frac{1}{n^2} = \mathbb{P}([\mathrm{X}=i])\mathbb{P}([\mathrm{Y}=j])$$

donc les variables X et Y sont indépendantes.
- On effectue la même expérience mais sans remise. On a alors $\mathbb{P}([\mathrm{X}=1]\cap[\mathrm{Y}=1])=0$ mais $\mathbb{P}([\mathrm{X}=1])\mathbb{P}([\mathrm{Y}=1]) = \dfrac{1}{n^2}$ donc les variables ne sont pas indépendantes.

Remarque. Si X et Y sont indépendantes on a aussi

$$\forall (x,y) \in \mathrm{X}(\Omega)\times\mathrm{Y}(\Omega) \qquad \mathbb{P}([\mathrm{X}\leqslant x]\cap[\mathrm{Y}\leqslant y]) = \mathbb{P}(\mathrm{X}\leqslant x)\times\mathbb{P}([\mathrm{Y}\leqslant y])$$

et, d'une manière plus générale,

$$\forall (\mathrm{A},\mathrm{B}) \in \mathcal{P}(\mathrm{X}(\Omega))\times\mathcal{P}(\mathrm{Y}(\Omega)) \qquad \mathbb{P}([\mathrm{X}\in\mathrm{A}]\cap[\mathrm{Y}\in\mathrm{B}]) = \mathbb{P}([\mathrm{X}\in\mathrm{A}])\times\mathbb{P}([\mathrm{Y}\in\mathrm{B}]).$$

Exercice 5.1. Démontrer ce résultat.

Proposition 5.7 (Fonctions de deux variables indépendantes)**.** *Soit X et Y deux variables aléatoires indépendantes sur l'espace probabilisé* $(\Omega,\mathcal{P}(\Omega),\mathbb{P})$*. Pour toute fonction f définie sur* $\mathrm{X}(\Omega)$ *et pour toute fonction g définie sur* $\mathrm{Y}(\Omega)$*, les variables aléatoires* $f(\mathrm{X})$ *et* $g(\mathrm{Y})$ *sont indépendantes.*

Démonstration :
Soit $(z,z')\in f(\mathrm{X})(\Omega)\times g(\mathrm{Y})(\Omega)$. $[(f(\mathrm{X}),g(\mathrm{Y}))=(z,z')] = [f(\mathrm{X})=z]\cap[g(\mathrm{Y})=z']$.
Or, $[f(\mathrm{X})=z] = \{\omega\in\Omega,\ f\circ\mathrm{X}(\omega)=z\} = \{\omega\in\Omega,\ \mathrm{X}(\omega)\in f^{-1}(\{z\})\} = [\mathrm{X}\in f^{-1}(\{z\})]$. De même $[g(\mathrm{Y})=z'] = [\mathrm{Y}\in g^{-1}(\{z'\})]$ donc,
$[(f(\mathrm{X}),g(\mathrm{Y}))=(z,z')] = [\mathrm{X}\in f^{-1}(\{z\})]\cap[\mathrm{Y}\in g^{-1}(\{z'\})]$. Par indépendance, on a alors

$$\begin{aligned}\mathbb{P}([(f(\mathrm{X}),g(\mathrm{Y}))=(z,z')]) &= \mathbb{P}([\mathrm{X}\in f^{-1}(\{z\})]).\mathbb{P}([\mathrm{Y}\in g^{-1}(\{z'\})]\\ &= \mathbb{P}([f(\mathrm{X})=z]).\mathbb{P}([g(\mathrm{Y})=z']\end{aligned}$$

Donc les variables $f(\mathrm{X})$ et $g(\mathrm{Y})$ sont indépendantes. □

Exemple. Si X et Y sont deux variables aléatoires indépendantes suivant respectivement la loi binomiale $\mathcal{B}(m,p)$ et la loi binomiale $\mathcal{B}(n,p)$ ($(n,m)\in(\mathbb{N}^*)^2$ et $p\in]0,1[$), alors la variable aléatoire $\mathrm{X}+\mathrm{Y}$ suit la loi binomiale $\mathcal{B}(n+m,p)$.
Démonstration : On a $\mathrm{X}(\Omega)=[\![0,m]\!]$ et $\mathrm{Y}(\Omega)=[\![0,n]\!]$ donc $(\mathrm{X}+\mathrm{Y})(\Omega)=[\![0,m+n]\!]$. Soit alors $k\in[\![0,m+n]\!]$.

$$\begin{aligned}[\mathrm{X}+\mathrm{Y}=k] &= [\mathrm{X}+\mathrm{Y}=k]\cap\Omega\\ &= [\mathrm{X}+\mathrm{Y}=k]\cap\left(\bigcup_{i=0}^{m}[\mathrm{X}=i]\right)\\ &= \bigcup_{i=0}^{m}[\mathrm{X}+\mathrm{Y}=k]\cap[\mathrm{X}=i] = \bigcup_{i=0}^{m}[\mathrm{X}=i]\cap[\mathrm{Y}=k-i]\end{aligned}$$

Cette réunion étant disjointe, par axiome de probabilité,
$\mathbb{P}([\mathrm{X}+\mathrm{Y}=k]) = \sum_{i=0}^{m}\mathbb{P}([\mathrm{X}=i]\cap[\mathrm{Y}=k-i])$ et donc, par indépendance de X et Y,

$\mathbb{P}([\mathrm{X}+\mathrm{Y}=k])=\sum_{i=0}^{m}\mathbb{P}([\mathrm{X}=i]).\mathbb{P}([\mathrm{Y}=k-i])$. En posant $\binom{n}{j}=0$ pour $j<0$ ou $j>n$, on a alors

$$\begin{aligned}\mathbb{P}([\mathrm{X}+\mathrm{Y}=k]) &=\sum_{i=0}^{m}\binom{m}{i}p^i(1-p)^{m-i}\binom{n}{k-i}p^{k-i}(1-p)^{n-k+i}\\ &=\sum_{i=0}^{m}\binom{m}{i}\binom{n}{k-i}p^k(1-p)^{m+n-k}=p^k(1-p)^{m+n-k}\sum_{i=0}^{m}\binom{m}{i}\binom{n}{k-i}\end{aligned}$$

Or, le coefficient de X^k dans le développement de $(1+\mathrm{X})^m(1+\mathrm{X})^n$ est $\sum_{i=0}^{m}\binom{m}{i}\binom{n}{k-i}$ et comme $(1+\mathrm{X})^n(1+\mathrm{X})^m=(1+\mathrm{X})^{n+m}$, on a donc $\sum_{i=0}^{m}\binom{m}{i}\binom{n}{k-i}=\binom{n+m}{k}$ (identité de Vandermonde). Finalement, $\mathbb{P}([\mathrm{X}+\mathrm{Y}=k])=\binom{n+m}{k}p^k(1-p)^{(n+m)-k}$ et donc $\mathrm{X}+\mathrm{Y}$ suit la loi $\mathcal{B}(n+m,p)$. □

Définition 5.8. *Soit* $\mathrm{X}_1,\cdots,\mathrm{X}_n$ $(n\geqslant 2)$, *des variables aléatoires sur un même espace probabilisé* $(\Omega,\mathcal{P}(\Omega),\mathbb{P})$.

- *On dit que les variables* $\mathrm{X}_1,\cdots,\mathrm{X}_n$, *sont* ***deux à deux indépendantes*** *si pour tout couple* (i,j) *tel que* $i\neq j$, *les variables* X_i *et* X_j *sont indépendantes.*
- *On dit que les variables* $\mathrm{X}_1,\cdots,\mathrm{X}_n$, *sont* ***mutuellement indépendantes*** *si pour tout* $(x_1,\cdots,x_n)\in\mathrm{X}_1(\Omega)\times\cdots\times\mathrm{X}_n(\Omega)$ *on a*

$$\mathbb{P}([\mathrm{X}_1=x_1]\cap\cdots\cap[\mathrm{X}_n=x_n])=\mathbb{P}([\mathrm{X}_1=x_1])\cdots\mathbb{P}([\mathrm{X}_n=x_n])$$

- *On dit que la suite de variables aléatoires* $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *est une* ***suite de variables aléatoires indépendantes*** *si, pour tout* $n\in\mathbb{N}$, *les variables aléatoires* $\mathrm{X}_0,\cdots,\mathrm{X}_n$ *sont mutuellement indépendantes.*

Proposition 5.9. *Soit* $\mathrm{X}_1,\cdots,\mathrm{X}_n$ $(n\geqslant 2)$, *des variables aléatoires sur un même espace probabilisé* $(\Omega,\mathcal{P}(\Omega),\mathbb{P})$. *Les variables* $\mathrm{X}_1,\cdots,\mathrm{X}_n$, *sont mutuellement indépendantes si et seulement si pour tout* $(\mathrm{E}_1,\cdots,\mathrm{E}_n)\in\mathcal{P}(\mathrm{X}_1(\Omega))\times\cdots\times\mathcal{P}(\mathrm{X}_n(\Omega))$,

$$\mathbb{P}([\mathrm{X}_1\in\mathrm{E}_1]\cap\cdots\cap[\mathrm{X}_n\in\mathrm{E}_n])=\mathbb{P}([\mathrm{X}_1\in\mathrm{E}_1])\times\cdots\times\mathbb{P}([\mathrm{X}_n\in\mathrm{E}_n])$$

Remarque. Soit $\mathrm{X}_1,\cdots,\mathrm{X}_n$ n variables aléatoires mutuellement indépendantes et soit p dans $\{2,\cdots,n-1\}$. Alors toute variable aléatoire fonction uniquement des variables $\mathrm{X}_1,\cdots,\mathrm{X}_p$ est indépendante de toute variable aléatoire fonction uniquement des variables $\mathrm{X}_{p+1},\cdots,\mathrm{X}_n$.
Par exemple, si X_1, X_2, X_3 et X_4 sont quatre variables aléatoires mutuellement indépendantes alors les variables $\mathrm{X}_1+2\mathrm{X}_3^2$ et $\mathrm{X}_2-3e^{\mathrm{X}_4}$ sont indépendantes.

5.2 Probabilités dans le cadre discret

5.2.1 Espaces probabilisés

On a rappelé la notion de probabilité sur un **univers fini** Ω associé à une expérience aléatoire. La situation se complique un peu lorsque par exemple Ω est un intervalle de $\mathbb{R}$ ou même $\mathbb{R}$ tout entier. On peut en effet montrer (on l'admettra ici) qu'il est impossible de construire une probabilité sur l'ensemble tout entier des parties de $\mathbb{R}$.
On contourne la difficulté en restreignant l'ensemble des évènements : on ne prend plus $\mathcal{P}(\mathbb{R})$ tout entier mais seulement une partie de cet ensemble. On veut toutefois continuer à faire le même type d'opérations sur les évènements. Cela conduit à introduire la notion de tribu.

Définition 5.10. *Étant donné un ensemble Ω, on appelle **tribu** de parties de Ω toute partie $\mathcal{T}$ de $\mathcal{P}(\Omega)$:*
• *contenant $\varnothing$* • *stable par passage au complémentaire* • *stable par union dénombrable*

Exemples.

- $\{\varnothing, \Omega\}$ et $\mathcal{P}(\Omega)$ sont des tribus de parties de Ω
- Si $\mathrm{A} \subset \Omega$, $\{\varnothing, \mathrm{A}, \overline{\mathrm{A}}, \Omega\}$ est une tribu de parties de Ω
- L'intersection de deux tribus de parties de Ω est encore une tribu.

Remarque. Si $\mathcal{T}$ est une tribu de parties de l'ensemble non vide Ω, on dit que $(\Omega, \mathcal{T})$ est un ***espace probabilisable***.

Définition 5.11. *Soit $(\Omega, \mathcal{T})$ un espace probabilisable. On appelle **probabilité** sur Ω toute application $\mathbb{P} : \mathcal{T} \longrightarrow \mathbb{R}$ vérifiant les trois axiomes*

(1) $\forall \mathrm{A} \in \mathcal{T} \quad \mathbb{P}(\mathrm{A}) \geqslant 0$

(2) $\mathbb{P}(\Omega) = 1$

(3) pour toute suite $\mathrm{A}_1, \mathrm{A}_2, \cdots$ *d'évènements incompatibles (i.e., pour tout $i \neq j$, $\mathrm{A}_i \cap \mathrm{A}_j = \varnothing$),*

$$\mathbb{P}\,(\bigcup_{i=1}^{\infty} \mathrm{A}_i) = \sum_{i=1}^{\infty} \mathbb{P}(\mathrm{A}_i) \qquad (\sigma\text{-additivité})$$

*On dit alors que $(\Omega, \mathcal{T}, \mathbb{P})$ est un **espace probabilisé**.*

Remarque. La définition ci-dessus contient implicitement le fait que pour toute suite $(\mathrm{A}_n)_{n\in\mathbb{N}}$ d'évènements deux à deux incompatibles, la série de terme général $\mathbb{P}(\mathrm{A}_n)$ converge.

Proposition 5.12 (Probabilité dans le cadre discret)**.** *Soit $(\Omega, \mathcal{P}(\Omega))$ un espace probabilisable discret (c'est à dire tel que l'univers Ω est fini ou dénombrable). Se donner une probabilité sur Ω revient à se donner une suite $(p_n)_{n\in\mathbb{N}}$ de réels positifs telle que $\sum p_n$ converge et $\sum_{n=0}^{+\infty} p_n = 1$.*

Démonstration : (Incomplète.) On effectue la démonstration dans le cas particulier où $\Omega = \mathbb{N}$ mais elle est adaptable sans difficulté aux autres cas.
• Supposons que $\mathbb{P}$ soit une probabilité sur $\Omega = \mathbb{N}$.
Pour tout n de $\mathbb{N}$, posons $p_n = \mathbb{P}(\{n\})$. On définit bien ainsi une suite de réels positifs. Comme $\Omega = \bigcup_{n=0}^{\infty} \{n\}$, la σ-additivité de $\mathbb{P}$ entraîne $1 = \mathbb{P}(\Omega) = \sum_{n=0}^{+\infty} p_n$.
• Réciproquement, si une telle suite $(p_n)_{n\in\mathbb{N}}$ est donnée, on pose, pour tout A de $\mathcal{P}(\Omega)$, $\mathbb{P}(\mathrm{A}) = \sum_{n\in\mathrm{A}} p_n$. On définit ainsi une application $\mathbb{P} : \mathcal{P}(\Omega) \to \mathbb{R}_+$ qui clairement vérifie $\mathbb{P}(\Omega) = 1$.
On admet ici la σ-additivité de $\mathbb{P}$. Le résultat n'est pas difficile mais fait appel au théorème (non vu) de sommation par paquets pour les séries positives.
Finalement, $\mathbb{P}$ est bien une probabilité sur Ω. □

Exercice 5.2. Déterminer la valeur du réel $a > 0$ pour laquelle la donnée, pour $n \in \mathbb{N}^*$, des réels $p_n = \dfrac{a}{n(n+1)(n+2)}$ permet de définir une probabilité sur $\mathbb{N}^*$.

5.2.2 Propriétés élémentaires des probabilités

Tous les résultats de première année sur les probabilités s'étendent de manière naturelle à ce cadre plus général. C'est en particulier le cas de la notion de probabilité conditionnelle (formule des probabilités composées, formule des probabilités totales, formule de Bayes) mais aussi de la notion d'indépendance pour les évènements.

Définition 5.13. *Une suite* $(\mathrm{A}_n)_{n\in\mathbb{N}}$ *d'évènements est dite croissante si* $\forall n \in \mathbb{N}$, $\mathrm{A}_n \subset \mathrm{A}_{n+1}$. *Une suite* $(\mathrm{A}_n)_{n\in\mathbb{N}}$ *d'évènements est dite décroissante si* $\forall n \in \mathbb{N}$, $\mathrm{A}_n \supset \mathrm{A}_{n+1}$.

Exemple. Si $\Omega = \mathbb{N}$, la suite $(\mathrm{A}_n)_{n\in\mathbb{N}^*}$ avec $\mathrm{A}_n = \{1, \cdots, n\}$ est croissante et $\bigcup_{i=1}^{\infty} \mathrm{A}_i = \mathbb{N}^*$.

Théorème 5.14. *Si* $(\mathrm{A}_n)_{n\in\mathbb{N}}$ *est une suite croissante d'évènements , alors*

$$\lim_{n\to\infty} \mathbb{P}(\mathrm{A}_n) = \mathbb{P}\,(\bigcup_{i=0}^{\infty} \mathrm{A}_i)$$

Si $(\mathrm{A}_n)_{n\in\mathbb{N}}$ *est une suite décroissante d'évènements , alors* $\lim_{n\to\infty} \mathbb{P}(\mathrm{A}_n) = \mathbb{P}\,(\bigcap_{i=0}^{\infty} \mathrm{A}_i)$

Démonstration : Si la suite est croissante, notons $\mathrm{B}_1 = \mathrm{A}_1, \mathrm{B}_2 = \mathrm{A}_2\backslash \mathrm{A}_1, \cdots, \mathrm{B}_n = \mathrm{A}_n\backslash \mathrm{A}_{n-1}$. Les ensembles B_i sont disjoints et pour tout n

$$\bigcup_{i=1}^{n} \mathrm{B}_i = \bigcup_{i=1}^{n} \mathrm{A}_i = \mathrm{A}_n \text{ et } \bigcup_{i=1}^{\infty} \mathrm{B}_i = \bigcup_{i=1}^{\infty} \mathrm{A}_i.$$

Par conséquent

$$\begin{aligned}
\mathbb{P}\,(\bigcup_{i=0}^{\infty} \mathrm{A}_i) &= \mathbb{P}\,(\bigcup_{i=0}^{\infty} \mathrm{B}_i) \qquad \text{donc, par } \sigma\text{-additivité} \\
&= \sum_{i=0}^{\infty} \mathbb{P}(\mathrm{B}_i) = \lim_{n\to\infty} \sum_{i=0}^{n} \mathbb{P}(\mathrm{B}_i) \text{ (par définition de la somme d'une série)} \\
&= \lim_{n\to\infty} \mathbb{P}\,(\bigcup_{i=1}^{n} \mathrm{B}_i) = \lim_{n\to\infty} \mathbb{P}(\mathrm{A}_n)
\end{aligned}$$

Si la suite $(\mathrm{A}_n)_{n\geqslant 0}$ est décroissante, alors la suite $(\overline{\mathrm{A}_n})_{n\geqslant 0}$ est croissante et on obtient alors le résultat facilement en prenant le complémentaire puisque le complémentaire de $\bigcup_{i=1}^{\infty} \mathrm{A}_i$ est $\bigcap_{i=1}^{\infty} \overline{\mathrm{A}_i}$. □

Exemple. Un singe immortel tape tous les jours 100 000 caractères de façon aléatoire sur un clavier de 40 touches. On cherche la probabilité que ce singe tape, un jour donné, l'intégralité du roman *Le rouge et le noir* sans une seule faute de frappe, sachant que ce livre comporte 100 000 caractères (c'est une approximation grossière). Notons A_n l'évènement « Le singe n'a pas réussi à taper entièrement le livre avant le jour n ». La suite d'évènements (A_n) est décroissante puisque si le singe n'a pas réussi avant le jour $n+1$, c'est qu'il n'avait pas réussi avant le jour n. Posons $p = \frac{1}{40^{100\,000}}$. p est la probabilité que notre singe tape son livre un jour donné. Alors, par indépendance, $\mathbb{P}(\mathrm{A}_n) = (1-p)^n$. Le théorème de la limite monotone donne alors $\mathbb{P}\,(\bigcap_{i\in\mathbb{N}^*} \mathrm{A}_i) = \lim_{n\to\infty} \mathbb{P}(\mathrm{A}_n) = 0$ puisque $p > 0$ et donc $1 - p < 1$. Il y a donc une probabilité nulle que le singe n'arrive jamais à taper *Le rouge et le noir*. Autrement dit, le singe arrivera à taper ce livre avec probabilité de 1 : un tel évènement (de probabilité 1) est dit ***presque sûr***.

5.2.3 Variables aléatoires discrètes

Définition 5.15. *Soit* $(\Omega, \mathcal{T})$ *un espace probabilisable au plus dénombrable. Une variable aléatoire (réelle) discrète sur* Ω *est une application* X *de* Ω *dans* $\mathbb{R}$ *telle que, pour tout* x *de* $\mathbb{R}$, $\mathrm{X}^{-1}(\{x\}) \in \mathcal{T}$.

Remarque. Dans la pratique, dans ce cadre discret, on pourra encore choisir $\mathcal{T} = \mathcal{P}(\Omega)$ et donc se passer de la dernière condition de la définition.

Théorème et définition 5.16 (Loi de probabilité d'une v.a. discrète)**.** *Soit* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *un espace probabilisé au plus dénombrable. Soit* X *une variable aléatoire discrète sur cet espace à valeurs dans* $\mathbb{R}$*. L'application* $\mathbb{P}_X : \mathcal{P}(X(\Omega)) \longrightarrow [0,1]$, $E \longmapsto \mathbb{P}([X \in E])$ *est une probabilité sur l'espace probabilisable* $(X(\Omega), \mathcal{P}(X(\Omega)))$*. On l'appelle la* ***loi de probabilité*** *de la variable aléatoire* X *(ou encore* ***probabilité image*** *de la variable aléatoire* X*).*

Démonstration : Il est clair que $\mathbb{P}_X$ est une application de $\mathcal{P}(X(\Omega))$ dans $[0,1]$. De plus, $\mathbb{P}_X(X(\Omega)) = \mathbb{P}(X \in X(\Omega)) = \mathbb{P}(\Omega) = 1$.

Soit enfin $(E_n)_{n\in\mathbb{N}} \in (\mathcal{P}(X(\Omega)))^{\mathbb{N}}$ une suite d'évènements deux à deux disjoints. Alors, $([X \in E_n])_{n\in\mathbb{N}}$ est une suite d'évènements deux à deux disjoints (car X est une application). Donc, par σ-additivité de $\mathbb{P}$,

$$\mathbb{P}_X\,(\bigcup_{n=0}^{+\infty} E_n) = \mathbb{P}\,([X \in \bigcup_{n=0}^{+\infty} E_n]) = \big(\bigcup_{n=0}^{+\infty} X^{-1}(E_n)\big) = \sum_{n=0}^{+\infty} \mathbb{P}(X^{-1}(E_n)) = \sum_{n=0}^{+\infty} \mathbb{P}_X(E_n)$$

On a montré que $\mathbb{P}_X$ est une probabilité sur l'espace probabilisable $(X(\Omega), \mathcal{P}(X(\Omega)))$. □

Remarque. La loi de probabilité de X est entièrement déterminée par les atomes de probabilité $\mathbb{P}_X(\{x\}) = \mathbb{P}(X \in \{x\}) = \mathbb{P}([X = x]) : \forall E \subset X(\Omega), \mathbb{P}_X(E) = \sum_{x\in E} \mathbb{P}([X = x]) = \sum_{x\in E} \mathbb{P}_X(\{x\})$.
Dans la pratique, donner la loi de probabilité d'une variable discrète X reviendra donc à donner $X(\Omega)$ et les $\mathbb{P}([X = x])$ pour tous les x de $X(\Omega)$.

Comme dans le cadre fini, on introduit alors les fonctions de variables aléatoires, les couples (respectivement n-uplets) de variables aléatoires et la notion de variables indépendantes. Les propriétés vues dans le cadre fini s'étendent de manière naturelle au cadre discret.
La notion d'espérance diffère pour sa part quelque peu...

Définition 5.17 (Espérance d'une variable aléatoire discrète)**.** *Soit* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *un espace probabilisé au plus dénombrable. Soit* X *une variable aléatoire discrète sur cet espace à valeurs dans* $\mathbb{R}$*. Notons alors* $\{x_1, \cdots, x_n, \cdots\}$ *l'image de* Ω *par* X*. On dit que la variable* X *admet une* ***espérance*** *si la série* $\sum x_k \mathbb{P}([X = x_k])$ *converge* *<u>absolument</u>.*

On définit alors son espérance par $\mathbb{E}[X] = \sum_{k=1}^{+\infty} x_k \mathbb{P}([X = x_k])$.

Remarque. Cette notion d'espérance garde toutes les propriétés de celle introduite dans le cadre fini : linéarité, croissance...

5.3 Lois discrètes classiques

On ne rappellera pas ici les lois finies classiques : la loi uniforme (équirépartie) sur un ensemble à n éléments, la loi de Bernoulli $\mathcal{B}(p)$ avec $p \in]0,1[$, loi Binomiale $\mathcal{B}(n,p)$ avec $p \in]0,1[$ et $n \in \mathbb{N}^*$...

- **Loi Géométrique $\mathcal{G}(p)$, avec** $p \in]0,1[$

On répète une épreuve de Bernoulli de paramètre p de manière indépendante et on s'arrête au premier succès. On note X le nombre de réalisations (indépendantes) nécessaires de l'épreuve de Bernoulli jusqu'à obtenir un succès. X est donc le rang du premier succès. On a alors $X(\Omega) = \mathbb{N}^*$ et, pour $k \in \mathbb{N}^*$, $[X = k] = E_1 \cap \cdots \cap E_{k-1} \cap S_k$ où E_i (resp. S_i) est l'évènement « La $i^{\text{ème}}$ épreuve a donné un échec (resp. succès).
Par indépendance, $\mathbb{P}([X = k]) = \mathbb{P}(E_1) \cdots \mathbb{P}(E_{k-1})\mathbb{P}(S_k) = p(1-p)^{k-1}$.

Définition 5.18. *On dit qu'une variable aléatoire* X*, définie sur un espace probabilisé dénombrable* $(\Omega, \mathcal{T}, \mathbb{P})$*, suit la* ***loi géométrique*** *de paramètre* p *et on écrit* $\mathrm{X} \hookrightarrow \mathcal{G}(p)$ *si* $\mathrm{X}(\Omega) = \mathbb{N}^*$ *et* $\forall k \in \mathbb{N}^* \quad \mathbb{P}([\mathrm{X} = k]) = p(1-p)^{k-1}$.

Proposition 5.19. *Si* X *est une variable aléatoire suivant une loi géométrique de paramètre* p *alors* X *a une espérance et on a* $\mathbb{E}[\mathrm{X}] = \dfrac{1}{p}$.

Démonstration : Notons tout d'abord que, d'après la règle de Riemann, la série $\sum k\mathbb{P}([\mathrm{X} = k])$ converge puisque, par croissances comparées, $k^2\big(k\mathbb{P}([\mathrm{X} = k])\big) = pk^3(1-p)^{k-1} \xrightarrow[k\to\infty]{} 0$ (puisque $0 < 1-p < 1$). X admet donc une espérance.

Posons à présent $q = 1-p$ et, pour $n \in \mathbb{N}^*$, $u_n = \displaystyle\sum_{k=1}^{n} kq^{k-1}$. Ce qui précède montre que la suite (u_n) est convergente. Notons ℓ sa limite. Soit $n \geqslant 2$.

On a alors $u_n = \displaystyle\sum_{k=1}^{n}(k-1+1)q^{k-1} = \sum_{k=1}^{n}(k-1)q^{k-1} + \sum_{k=1}^{n} q^{k-1}$. On en déduit (en posant $j = k-1$) $u_n = q\displaystyle\sum_{j=0}^{n-1} jq^{j-1} + \frac{1-q^n}{1-q}$ soit $u_n = qu_{n-1} + \dfrac{1-q^n}{1-q}$. Comme $|q| < 1$, on obtient à la limite $\ell = q\ell + \dfrac{1}{1-q}$ et donc $\ell = \dfrac{1}{(1-q)^2} = \dfrac{1}{p^2}$. Finalement, $\mathbb{E}(\mathrm{X}) = \displaystyle\sum_{n=1}^{+\infty} pu_n = \frac{1}{p}$. □

Définition 5.20. *Une variable aléatoire* X*, à valeurs dans* $\mathbb{N}$*, est dite* ***sans mémoire*** *si elle vérifie, pour tout* (m,n) *dans* $\mathbb{N}^2$*, l'identité :* $\mathbb{P}_{[\mathrm{X}>m]}([\mathrm{X} > m+n]) = \mathbb{P}([\mathrm{X} > n])$.

Proposition 5.21. *Une variable aléatoire* X *à valeurs dans* $\mathbb{N}^*$ *est sans mémoire si et seulement si sa loi est géométrique.*

Démonstration : • Soit X une variable aléatoire sans mémoire à valeurs dans $\mathbb{N}^*$. Pour $n \in \mathbb{N}$, notons $u_n = \mathbb{P}([\mathrm{X} > n])$. On a alors :

$$u_{n+1} = \mathbb{P}([\mathrm{X} > n+1]) = \mathbb{P}([\mathrm{X} > n+1] \cap [\mathrm{X} > n]) = \mathbb{P}_{[\mathrm{X}>n]}([\mathrm{X} > n+1])\mathbb{P}([\mathrm{X} > n])$$

soit $u_{n+1} = \mathbb{P}([\mathrm{X} > 1])\mathbb{P}([\mathrm{X} > n]) = u_1u_n$. En posant $q = u_1$, on constate que la suite (u_n) est géométrique de raison q.
Enfin, pour $n \in \mathbb{N}^*$, $[\mathrm{X} = n] = [\mathrm{X} > n-1] \cap \overline{[\mathrm{X} > n]}$ donc $\mathbb{P}([\mathrm{X} = n]) = u_{n-1} - u_n = q^{n-1} - q^n$ soit finalement $\mathbb{P}([\mathrm{X} = n]) = p.q^{n-1}$ où $p = 1-q$. X suit la loi géométrique de paramètre p.
• Réciproquement, supposons que X suive la loi géométrique de paramètre p et soit $(m,n) \in \mathbb{N}^2$.

$$\mathbb{P}_{[\mathrm{X}>m]}([\mathrm{X} > m+n]) = \frac{\mathbb{P}([\mathrm{X} > m+n] \cap [\mathrm{X} > m])}{\mathbb{P}([\mathrm{X} > m])} = \frac{\mathbb{P}([\mathrm{X} > m+n])}{\mathbb{P}([\mathrm{X} > m])}$$

Or, $[\mathrm{X} > m+n] = \displaystyle\bigcup_{k=m+n+1}^{+\infty}[\mathrm{X} = k]$ (car X est à valeurs entières) donc

$$\mathbb{P}([\mathrm{X} + \mathrm{Y} > m+n]) = \sum_{k=m+n+1}^{+\infty} pq^{k-1} = pq^{m+n}\frac{1}{1-q} = q^{m+n}$$

De même, $\mathbb{P}([\mathrm{X} > m]) = q^m$. On en déduit alors

$$\mathbb{P}_{[\mathrm{X}>m]}([\mathrm{X} > m+n]) = \frac{q^{m+n}}{q^m} = q^n = \mathbb{P}([\mathrm{X} > n]$$

X est donc bien sans mémoire. □

• Loi de Poisson, $\mathcal{P}(\lambda)$, avec $\lambda > 0$

Cette loi est souvent utilisée pour dénombrer des évènements « rares », comme des pannes, des sinistres, des clients entrant dans une boutique pendant un intervalle de temps donné.

Définition 5.22. *On dit qu'une variable aléatoire* X*, définie sur un espace probabilisé dénombrable* $(\Omega, \mathcal{T}, \mathbb{P})$*, suit la **loi de Poisson** de paramètre* λ *et on écrit* $\mathrm{X} \hookrightarrow \mathcal{P}(\lambda)$ *si* $\mathrm{X}(\Omega) = \mathbb{N}$ *et* $\forall k \in \mathbb{N} \quad \mathbb{P}([\mathrm{X} = k]) = \dfrac{\lambda^k}{k!}e^{-\lambda}$.

Proposition 5.23. *Si* X *est une variable aléatoire suivant une loi de Poisson de paramètre* λ *alors* X *a une espérance et on a* $\mathbb{E}[\mathrm{X}] = \lambda$.

Démonstration : On a déjà vu que, pour tout réel x, $e^x = \displaystyle\sum_{k=0}^{+\infty} \frac{x^k}{k!}$. Notons, pour $n \in \mathbb{N}$, $u_n = \displaystyle\sum_{k=0}^{n} k\mathbb{P}([\mathrm{X} = k])$. On a alors $u_n = \displaystyle\sum_{k=0}^{n} k\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{n} \lambda\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}$. On en déduit (en posant $j = k-1$) $u_n = \lambda e^{-\lambda}\displaystyle\sum_{j=0}^{n-1}\frac{\lambda^j}{j!}$. Il est alors clair que la suite $(u_n)_{n\in\mathbb{N}}$ converge (donc X a une espérance) et à la limite $\mathbb{E}(\mathrm{X}) = \lambda e^{-\lambda}e^{\lambda} = \lambda$. □

Théorème 5.24 (Approximation de la loi binomiale par la loi de Poisson)**.** *Soit* λ *un réel strictement positif et* $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires discrètes telles que* X_n *suit la loi binomiale de paramètres* n *et* $\dfrac{\lambda}{n}$*. Alors* $(\mathrm{X}_n)_{n\in\mathbb{N}}$ *converge en loi vers une variable aléatoire* X *qui suit une loi de Poisson de paramètre* λ*. Autrement dit,*

$$\lim_{n\to+\infty} \mathrm{P}([\mathrm{X}_n = k]) = \frac{\lambda^k}{k!}e^{-\lambda}$$

Démonstration : On a, pour $k \in [\![0, n]\!]$,

$$\begin{aligned}\mathrm{P}([\mathrm{X}_n = k]) &= \frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}\\ &= \frac{\lambda^k}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k}e^{(n-k)\ell n\,(1-\frac{\lambda}{n})}\end{aligned}$$

Or, d'une part $n(n-1)\cdots(n-k+1) \underset{n\to+\infty}{\sim} n^k$ et d'autre part $\ell n\,(1-\lambda/n) \underset{n\to+\infty}{\sim} -\dfrac{\lambda}{n}$ (car $\dfrac{\lambda}{n} \xrightarrow[n\to+\infty]{} 0$. Donc on en déduit :

$$\lim_{n\to+\infty} \mathrm{P}([\mathrm{X}_n = k]) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

□

Remarques.

- Dans la pratique, on considère souvent que, lorsque $n \geqslant 30$ et $p \leqslant 0,1$, on peut approcher la loi $\mathscr{B}(n,p)$ par la loi $\mathcal{P}(np)$.
- On dit que la loi de Poisson est la loi des évènements rares (elle approche le tirage de n boules avec remise dans une urne contenant des boules blanche en proportion égale à p qui est faible).