PROBABILITES - 2ème partie

Plan

Ce chapitre fait suite au chapitre L1/PROBA1.PDF, qui traite des probabilités sur un ensemble fini. Il nécessite les connaissances du chapitre sur les séries entières L2/SERIENTR.PDF, et donc également celles du chapitre sur les séries numériques L2/SERIES.PDF.

I : Probabilité sur un univers quelconque

1- Ensemble dénombrable

Nous étendons ci-dessous les notions de probabilité sur un univers fini au cas d'un ensemble Ω quelconque. Nous commençons par le cas où Ω est **dénombrable**, i.e. qui est en bijection avec $\mathbf{N}$ (on pourra lire à ce sujet l'annexe I du chapitre L1/ENSEMBLE.PDF). Cela signifie qu'on peut décrire Ω sous la forme $\{a_n, n \in \mathbf{N}\}$, les a_n étant distincts.

❑ Toute partie A d'un ensemble Ω dénombrable est finie ou dénombrable. Il suffit en effet de parcourir les éléments a_0, a_1, a_2, ..., a_n, ... de Ω et d'appeler b_0 le premier élément de la liste appartenant à A, b_1 l'élément suivant, b_2 l'élément suivant, etc... A est fini ou dénombrable suivant que la liste des b_i s'arrête ou se prolonge indéfiniment.

❑ Toute image d'un ensemble dénombrable par une application f est finie ou dénombrable. Il suffit en effet de parcourir les éléments de $f(\Omega)$: $f(a_0)$, $f(a_1)$, ..., $f(a_n)$, ... et de poser $b_0 = f(a_0)$, b_1 égal au premier $f(a_n)$ différent de b_0, b_2 le $f(a_n)$ suivant différent de b_0 et b_1, etc...

❑ La réunion de deux ensembles dénombrables est dénombrable. En effet, si ces deux ensembles sont $\{a_n, n \in \mathbb{N}\}$ et $\{b_n, n \in \mathbb{N}\}$, on peut obtenir leur réunion de la façon suivante :

i) ranger les éléments dans l'ordre a_0, b_0, a_1, b_1, ..., a_n, b_n, ...

ii) éliminer de la liste les éléments identiques (l'un des b_i peut en effet être égal à l'un des a_j) pour n'en garder qu'un exemplaire.

iii) énumérer la liste finale obtenue en désignant ses éléments par c_0, c_1, c_2, ..., c_n, ...

EXEMPLES :

❑ $\mathbb{Z}$ est dénombrable. Il est en effet la réunion de $\mathbb{N}$ et de $\{-n, n \in \mathbb{N}^*\}$. Or $\{-n, n \in \mathbb{N}^*\}$ est dénombrable au moyen de la bijection $p \in \mathbb{N} \to -p-1$.

On peut aussi donner explicitement une bijection entre $\mathbb{N}$ et $\mathbb{Z}$:

$$f : \mathbb{N} \to \mathbb{Z}$$

$$n \to \left[\begin{array}{l} \dfrac{n}{2} \text{ si } n \text{ est pair} \\ -\dfrac{n+1}{2} \text{ si } n \text{ est impair} \end{array}\right.$$

❑ Le produit de deux ensembles dénombrables est dénombrable. Montrons-le d'abord dans le cas de $\mathbb{N}^2$. Il suffit d'énumérer ses éléments dans l'ordre suivant :

$c_0 = (0, 0)$

$c_1 = (1, 0)$ $\quad c_2 = (0, 1)$

$c_3 = (2, 0)$ $\quad c_4 = (1, 1)$ $\quad c_5 = (0, 2)$

$c_6 = (3, 0)$ $\quad c_7 = (2, 1)$ $\quad c_8 = (1, 2)$ $\quad c_9 = (0, 3)$

$c_{10} = (4, 0)$ $\quad c_{11} = (3, 1)$ $\quad c_{12} = (2, 2)$ $\quad c_{13} = (1, 3)$ $\quad c_{14} = (0, 4)$

...

$c_{n(n-1)/2} = (n-1, 0)$ $\quad c_{n(n-1)/2+1} = (n-2, 1)$ $\quad$ ... $\quad c_{n(n+1)/2-1} = (0, n-1)$

...

Le cas général du produit $\{(a_n, b_m), n \in \mathbb{N}, m \in \mathbb{N}\}$ se montre exactement de la même façon en remplaçant ci-dessus les couples (n, m) par (a_n, b_m).

En particulier $\mathbb{Q}$ est dénombrable. En effet $\mathbb{Q}^+$ peut s'injecter dans $\mathbb{N}^2$ au moyen d'une application du type $\dfrac{p}{q} \to (p, q)$, où p et q sont choisis sans facteur commun. De même pour $\mathbb{Q}^-$.

A titre indicatif, voici une bijection curieuse entre $\mathbb{Q}^{+*}$ et $\mathbb{N}$. On définit la fonction f de $\mathbb{N}$ dans $\mathbb{N}^*$ de la façon suivante :

$$f(0) = 1 \text{ et } \forall\, n \geq 0,\ f(2n+1) = f(n),\ f(2n+2) = f(n) + f(n+1)$$

de sorte que les valeurs de f sont :

$$\begin{array}{l} 1, 1, 2, 1, 3, 2, 3, 1, 4, 3, 5, 2, 5, 3, 4, 1, 5, 4, 7, 3, 8\ \ldots \\ \qquad\qquad\qquad\ \ \uparrow\uparrow \qquad\qquad\qquad \uparrow\uparrow \\ \qquad\qquad\qquad\ \ n\ n{+}1 \qquad\quad 2n{+}1\ 2n{+}2 \end{array}$$

Les valeurs successives de $\dfrac{f(n)}{f(n+1)}$ sont :

$$1, \frac{1}{2}, 2, \frac{1}{3}, \frac{3}{2}, \frac{2}{3}, 3, \frac{1}{4}, \frac{4}{3}, \frac{3}{5}, \frac{5}{2}, \frac{2}{5}, \frac{5}{3}, \frac{3}{4}, 4, \frac{1}{5}, \frac{5}{4}, \frac{4}{7}, \frac{7}{3}, \frac{3}{8}\ \ldots$$

On montre[1] que tous les rationnels positifs apparaissent une fois et une seule dans cette liste, de sorte que l'application $n \to \dfrac{f(n)}{f(n+1)}$ forme une bijection de $\mathbf{N}$ dans $\mathbf{Q}^{+*}$.

La suite précédente peut également être générée par la suite récurrence suivante[2] :

$x_0 = 1$

$\forall\, n \geq 0,\ x_{n+1} = \varphi(x_n)$ avec $\varphi(x) = \dfrac{1}{1 + 2\lfloor x \rfloor - x}$, où $\lfloor x \rfloor$ désigne la partie entière de x.

2- Probabilité sur un espace dénombrable

Ω est l'ensemble des résultats possibles d'une épreuve aléatoire. La situation où Ω est dénombrable se rencontre lorsque ces résultats peuvent être en nombre infini et peuvent être énumérés. On peut également choisir un tel Ω infini dénombrable lorsque le nombre de résultats d'une expérience aléatoire est fini, mais borné par un nombre indéterminé. Les ensembles finis ou dénombrables sont appelés ensembles **discrets**.

EXEMPLES :

❑ Lancer une pièce jusqu'à ce qu'il apparaisse P(ile). Au choix, ou bien $\Omega = \{F^nP \mid n \in \mathbf{N}\}$ où F^n désigne un suite de n F(aces) ; ou bien directement $\Omega = \mathbf{N}^*$ si on considère le résultat comme le nombre de lancers. Si on envisage la possibilité que Pile n'apparaisse jamais, on ajoutera un élément supplémentaire pour cette éventualité, de sorte que $\Omega = \{F^nP \mid n \in \mathbf{N}\} \cup \{F^\infty\}$ ou bien directement $\Omega = \mathbf{N}^* \cup \{\infty\}$.

❑ Compter le nombre de voitures passant au péage de l'autoroute à Fontainebleau entre le 30 Juillet 0 h et le 1 Août 24 h. $\Omega = \mathbf{N}$.

❑ Prendre un épi de blé et compter le nombre de grains. $\Omega = \mathbf{N}$.

❑ Se promener aléatoirement sur un axe de la façon suivante :
i) On lance une pièce déterminant le sens du déplacement. Si elle tombe sur Pile, on avance d'une unité, sinon on recule d'une unité.
ii) On relance la pièce pour décider si on poursuit le mouvement. Si elle tombe sur Pile, on s'arrête. Sinon on recommence à l'étape i).

[1] N. Calkin, H. S. Wilf, *Recounting the rationals*, Amer. Math. Monthly, **107**, n°4 (avril 2000), 360-363
[2] Aimeric Malter, Dierk Schleicher, Don Zagler, *New looks at old number theory*, American Math. Monthly, **128**, n°3 (mars 2013), 243-264

Le résultat de l'épreuve est l'abscisse obtenue, lorqu'on s'arrête. $\Omega = \mathbb{Z}$. On peut montrer que l'éventualité d'un mouvement se poursuivant indéfiniment est nulle. Si on souhaite néanmoins envisager cette possibilité, on ajoutera à Ω les deux éléments $+\infty$ et $-\infty$.

DEFINITION

On appelle **événement** *une partie* A *de* Ω*. L'ensemble des événements est ici l'ensemble des parties de* Ω*, noté* $\mathcal{P}(\Omega)$.

EXEMPLES :

❑ Lancer une pièce et obtenir Pile pour la première fois entre le 15ème et le 20ème lancer. Suivant l'univers Ω choisi, ou bien $A = \{F^nP \mid 15 \le n \le 20\}$; ou bien $A = \{15, 16, 17, 18, 19, 20\}$

❑ Compter entre 10 000 et 20 000 voitures au péage de Fontainebleau.
$A = \{n \in \mathbb{N} \mid 10\,000 \le n \le 20\,000\}$.

❑ Compter plus de 30 grains de blé sur un épi. $A = \{n \in \mathbb{N} \mid n > 30\}$.

❑ Terminer la promenade aléatoire sur l'axe avec une abscisse positive ou nulle. $A = \mathbb{N}$.

L'événement A est **réalisé** si l'issue de l'épreuve appartient à A. En reprenant les exemples ci-dessus, donnons des exemples d'épreuves en indiquant respectivement si l'événement A est réalisé :

FFFFFFP	A non réalisé
16.258 voitures	A réalisé
59 grains de blé	A réalisé
abcisse = –5	A non réalisé

Ω s'appelle **événement certain**. $\varnothing$ est appelé **événement impossible**. Un **événement élémentaire** est constitué d'un singleton $\{\omega\}$, $\omega \in \Omega$. Si $A \cap B = \varnothing$, on dit que les événements A et B sont **incompatibles**. A^c (complémentaire de A dqns Ω) s'appelle l'événement **contraire** de A.

On peut définir une probabilité sur Ω à partir de la probabilité des événements élémentaires. Pour tout a_n élément de Ω, on pose $\mathbf{P}(\{a_n\}) = p_n$, où p_n est un élément de [0, 1]. Pour toute partie A de Ω, on pose :

$$\mathbf{P}(A) = \sum_{n \text{ tel que } a_n \in A} p_n = \sum_{n=0}^{\infty} p_n \mathbf{1}_A(a_n)$$

où $\mathbf{1}_A$ est la **fonction indicatrice** de la partie A, définie comme suit :

$$\mathbf{1}_A : \Omega \to \{0, 1\}$$
$$x \to \begin{cases} 1 \text{ si } x \in A \\ 0 \text{ si } x \notin A \end{cases}$$

On notera que, A pouvant contenir une infinité d'éléments, $\mathbf{P}(A)$ est défini a priori par la somme d'une série et non par une somme finie. Par ailleurs, cette somme de série à termes positifs ou nuls ne dépend pas de l'ordre dans lequel on somme ses éléments, ce qui signifie que la façon dont on a numéroté les éléments de Ω est sans importance. (Voir la notion de famille sommable dans le chapitre L2/SERIES.PDF).

La probabilité de l'événement certain doit être égale à 1. Par conséquent, les p_n doivent être de somme égale à 1 : $\sum_{n=0}^{\infty} p_n = 1$.

EXEMPLES :

❑ On lance une pièce équilibrée jusqu'à obtenir Pile. Le résultat est le nombre de lancers. On a donc comme ensemble $\Omega = \mathbb{N}^*$, ou même $\mathbb{N}^* \cup \{\infty\}$ si on envisage la possibilité de ne jamais tomber sur Pile. On pose alors, pour n variant élément de $\mathbb{N}^*$:

$$p_n = \frac{1}{2^n}$$

$$p_\infty = 0$$

La valeur de p_∞ est nécessairement nulle. En effet :

$$p_\infty = \mathbf{P}(\{\infty\}) = \mathbf{P}(\mathbb{N}^{*c}) = 1 - \mathbf{P}(\mathbb{N}^*) = 1 - \sum_{n=1}^{\infty} \frac{1}{2^n} = 1 - 1 = 0$$

La probabilité que Pile arrive au bout d'un nombre pair de coups (événement A) est :

$$\mathbf{P}(\mathrm{A}) = \sum_{n=1}^{\infty} p_{2n} = \sum_{n=1}^{\infty} \frac{1}{4^n} = \frac{1}{4} \frac{1}{1 - \frac{1}{4}} = \frac{1}{3}$$

3- Cas général

Si Ω n'est pas dénombrable (par exemple $\Omega = \mathbb{R}$ ou $\mathbb{R}^2$), on ne peut définir une probabilité sur Ω de la façon précédente. Il est même en général exclu qu'on puisse définir une probabilité sur n'importe quelle partie A de Ω tant ces parties peuvent être compliquées (comment décrire une partie générale de $\mathbb{R}$?). Pour ce faire, on limite les événements A à décrire non pas $\mathcal{P}(\Omega)$ tout entier, mais seulement une famille particulière $\mathcal{T}$ de parties de Ω. On souhaite que cette famille vérifie les propriétés suivantes :

i) $\Omega \in \mathcal{T}$

ii) Si $\mathrm{A} \in \mathcal{T}$, alors A^c aussi

iii) Si $(\mathrm{A}_n)_{n\in\mathbb{N}}$ est une famille dénombrable d'éléments de $\mathcal{T}$, alors $\bigcup_{n=0}^{\infty} \mathrm{A}_n$ aussi.

On dit que $\mathcal{T}$ est une **tribu**. On a alors également :

iv) $\emptyset \in \mathcal{T}$

v) Si (A_n) est une famille finie d'éléments de $\mathcal{T}$, alors $\bigcup \mathrm{A}_n$ aussi.

vi) Si (A_n) est une famille finie ou dénombrable d'éléments de $\mathcal{T}$, alors $\bigcap \mathrm{A}_n$ aussi.

En effet :

iv) résulte directement de ii) en prenant $\mathrm{A} = \Omega$.

v) résulte de iii) en adjoignant à la famille finie une infinité dénombrable de parties vides.

vi) résulte de iii) ou v) en appliquant ii), le complémentaire d'une réunion étant une intersection, et le complémentaire du complémentaire redonnant l'ensemble lui-même :

$$\bigcap_n A_n = (\bigcup_n A_n^c)^c$$

Bref, l'ensemble des événements est stable par les opérations suivantes : passage au complémentaire, réunion finie ou dénombrable, intersection finie ou dénombrable.

EXEMPLES :

❑ Si $\Omega = \mathbb{R}$, on prend pour $\mathcal{T}$ la tribu engendrée par les intervalles, i.e. la plus petite famille contenant les intervalles et vérifiant les propriétés ii) et iii) sans qu'on cherche davantage à décrire cette famille.

❑ Un exemple encore moins simple est donné par le lancer successif d'une pièce jusqu'à ce qu'un certain événement soit observé. En pratique, la pièce est lancée un nombre fini de fois, mais ce nombre n'est peut-être pas borné. Aussi est-il commode de prendre $\Omega = \{0, 1\}^{\mathbb{N}}$ (ensemble des suites de 0 ou 1, où 0 symbolise un lancer Pile et 1 un lancer Face), dont on peut montrer qu'il est non dénombrable. On prendra comme événement au moins tout ce qui dépend d'un nombre fini de lancers, à savoir toute partie A de Ω pour laquelle il existe un entier n et une partie $B \subset \{0, 1\}^n$ tels que $A = \{(x_k)_{k\in\mathbb{N}}, (x_k)_{k\leq n} \in B\}$. Dans ce cas, l'événement est le fait que les n premiers lancers, n étant donné, vérifient une certaine propriété caractérisée par une partie B donnée. On prend alors pour $\mathcal{T}$ la plus petite tribu contenant ces événements. On peut montrer que $\mathcal{T}$ est différent de $\mathcal{P}(\Omega)$ et qu'il est possible de définir une probabilité sur $\mathcal{T}$ correspondant à la modélisation attendue d'un lancer de pièces, mais qu'il n'est pas possible d'étendre cette probabilité à $\mathcal{P}(\Omega)$. On ne précisera pas davantage à quoi peut ressembler $\mathcal{T}$..

DEFINITION

Soit Ω *un ensemble quelconque muni d'une tribu* $\mathcal{T}$. *On dit que* **P** *est une* **probabilité** *sur* $(\Omega, \mathcal{T})$ *si* **P** *est une application de* $\mathcal{T}$ *dans* [0, 1] *telle que* :

def-i) $\mathbf{P}(\Omega) = 1$

def-ii) *Si* $(A_n)_{n\in\mathbb{N}}$ *est une famille d'éléments de* $\mathcal{T}$ *incompatibles* (*événements disjoints deux à deux*), *alors* : $\mathbf{P}(\bigcup_{n=0}^{\infty} A_n) = \sum_{n=0}^{\infty} \mathbf{P}(A_n)$

$(\Omega, \mathcal{T}, \mathbf{P})$ *s'appelle alors* **espace probabilisé**.

Dans ii), on accepte de prendre des réunions dénombrables car se limiter à des réunions finies est trop restrictif, comme le montre les propriétés v) et vi) du paragraphe suivant. Le cas infini dénombrable est également nécessaire pour retrouver la définition par somme des probabilités des événements élémentaires dans le cas où Ω est dénombrable. En effet, si on note $\Omega = \{a_n, n \in \mathbb{N}\}$, et si on pose $p_n = \mathbf{P}(\{a_n\})$, alors :

$$A = \bigcup_{n \text{ tel que } a_n \in A} \{a_n\} \qquad \text{union disjointe éventuellement infinie dénombrable}$$

donc $\mathbf{P}(A) = \sum_{n \text{ tel que } a_n \in A} \mathbf{P}(\{a_n\}) = \sum_{n \text{ tel que } a_n \in A} p_n$ comme on l'a vu dans le § I.2.

Voici un exemple de probabilité sur un ensemble non dénombrable.

EXEMPLE :

❑ On prend $\Omega = [0, 1]$. On peut montrer qu'il existe une tribu $\mathcal{T}$ contenant tous les intervalles et une probabilité $\mathbf{P}$ définie sur $\mathcal{T}$ telle que :

$$\forall (a, b) \in \Omega^2, \mathbf{P}([a, b]) = b - a$$

On remarque que, pour $a = b$, $\mathbf{P}(\{a\}) = 0$ et que $\mathbf{P}$ ne peut être définie point par point, comme dans le cas dénombrable. $\mathbf{P}$ est la loi de **probabilité uniforme** sur [0, 1].

4- Propriétés des probabilités

On retrouve ci-dessous des propriétés vues en première année dans le cas fini. D'autres sont nouvelles.

PROPOSITION :

(i) $\mathbf{P}(\emptyset) = 0$

def-ii bis) *Si* $(A_k)_{1 \le k \le n}$ *est une famille finie d'éléments de* $\mathcal{T}$ *incompatibles, alors* :

$$\mathbf{P}(\bigcup_{k=1}^{n} A_n) = \sum_{k=1}^{n} \mathbf{P}(A_n)$$

(ii) *Pour tout événement* A *et* B, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$. *En particulier* :

$$\mathbf{P}(A \cup B) \le \mathbf{P}(A) \cup \mathbf{P}(B)$$

(iii) *Pour tout événement* A, $\mathbf{P}(A^c) = 1 - \mathbf{P}(A)$

(iv) *Pour tout événement* A *et* B *tel que* A *soit inclus dans* B, *on a* :

$$\mathbf{P}(A) \le \mathbf{P}(B) \text{ et } \mathbf{P}(B - A) = \mathbf{P}(B) - \mathbf{P}(A)$$

(v) *Pour toute suite croissante* $(A_n)_{n \in \mathbf{N}}$ *d'événements* (i.e. $A_n \subset A_{n+1}$), *on a* :

$$\mathbf{P}(\bigcup_{n=0}^{\infty} A_n) = \lim_{n \to \infty} \mathbf{P}(A_n)$$

(vi) *Pour toute suite décroissante* $(A_n)_{n \in \mathbf{N}}$ *d'événements* (i.e. $A_{n+1} \subset A_n$), *on a* :

$$\mathbf{P}(\bigcap_{n=0}^{\infty} A_n) = \lim_{n \to \infty} \mathbf{P}(A_n)$$

(vii) *Pour toute suite quelconque* $(A_n)_{n \in \mathbf{N}}$ *d'événements* :

$$\mathbf{P}(\bigcup_{n=0}^{\infty} A_n) \le \sum_{n=0}^{\infty} \mathbf{P}(A_n) \qquad (\textit{propriété de sous-additivité})$$

<u>Démonstration</u> :

❑ (i) : Appliquer def-ii) avec les A_n tous égaux à $\emptyset$.

❑ def-ii bis) : Appliquer def-ii en complétant la famille finie par une infinité dénombrable d'ensembles vides.

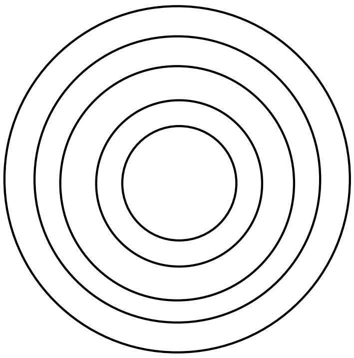
❑ (ii) : Pour tout événement A et B, incompatibles ou non, on a :

$A \cup B = B \cup (A - B)$
$A = (A \cap B) \cup (A - B)$
et les unions des membres de droites sont disjointes. D'où :
$\mathbf{P}(A \cup B) = \mathbf{P}(B) + \mathbf{P}(A - B)$
$\mathbf{P}(A) = \mathbf{P}(A \cap B) + \mathbf{P}(A - B)$
Le résultat s'en déduit en retranchant membre à membre.

❑ (iii) : $\Omega = A \cup A^c$ et cette union est disjointe. D'où :
$1 = \mathbf{P}(\Omega) = \mathbf{P}(A) + \mathbf{P}(A^c)$

❑ (iv) : Si A est inclus dans B, alors $B = A \cup (B - A)$, union disjointe, d'où :
$\mathbf{P}(B) = \mathbf{P}(A) + \mathbf{P}(B - A)$
et $\mathbf{P}(B - A)$ est positif ou nul.

❑ (v) : Posons $A = \bigcup_{n=0}^{\infty} A_n$ et $B_0 = A_0$, $B_n = A_n - A_{n-1}$ pour $n \geq 1$. Alors, on a $A = \bigcup_{n=0}^{\infty} B_n$. En effet, soit x élément de $\bigcup_{n=0}^{\infty} B_n$. Alors x est élément d'un des B_n, donc de A_n, donc est élément de A. Inversement, si x est élément de A, alors soit n le plus petit indice tel que x appartienne à A_n. x n'est donc pas dans A_{n-1}. x est alors dans B_n, donc dans leur réunion. De plus, les B_n sont deux à deux incompatibles puisque, si $k < n$, $B_k \subset A_k \subset A_{n-1}$ or $B_n \cap A_{n-1} = \varnothing$ donc $B_n \cap B_k = \varnothing$.
On montre de même que $A_n = \bigcup_{k=0}^{n} B_k$, union disjointe.
Ainsi, l'image ci-dessous peut-être vue comme représentant une suite croissante de disques A_n, ou bien un disque central B_0 entouré de couronnes disjointes successives B_n.



On a alors :

$$\mathbf{P}(A) = \mathbf{P}\Big(\bigcup_{n=0}^{\infty} B_n\Big) = \sum_{n=0}^{\infty} \mathbf{P}(B_n) \qquad \text{car l'union est disjointe}$$

$$= \lim_{n\to\infty} \sum_{k=0}^{n} \mathbf{P}(B_k) = \lim_{n\to\infty} \mathbf{P}\Big(\bigcup_{k=0}^{n} B_k\Big) = \lim_{n\to\infty} \mathbf{P}(A_n)$$

❑ (vi) : Il n'est pas difficile de voir que si $(A_n)_{n\in\mathbf{N}}$ est une suite décroissante, alors $(A_n^c)_{n\in\mathbf{N}}$ est une suite croissante. Posons $A = \bigcap_{n=0}^{\infty} A_n$. On a :
$\mathbf{P}(A) = 1 - \mathbf{P}(A^c)$

$$= 1 - \mathbf{P}((\bigcap_{n=0}^{\infty} A_n)^c)$$
$$= 1 - \mathbf{P}(\bigcup_{n=0}^{\infty} A_n^c)$$
$$= 1 - \lim_{n\to\infty} \mathbf{P}(A_n^c) \qquad \text{d'après v)}$$
$$= \lim_{n\to\infty} 1 - \mathbf{P}(A_n^c)$$
$$= \lim_{n\to\infty} \mathbf{P}(A_n)$$

❑ (vii) : Ici, la série $\sum \mathbf{P}(A_n)$ peut diverger, auquel cas $\sum_{n=0}^{\infty} \mathbf{P}(A_n) = +\infty$. La suite $B_n = \bigcup_{k=0}^{n} A_k$ est croissante, et $\bigcup_{n=0}^{\infty} B_n = \bigcup_{n=0}^{\infty} A_n$. Donc :

$$\mathbf{P}(\bigcup_{n=0}^{\infty} A_n) = \mathbf{P}(\bigcup_{n=0}^{\infty} B_n) = \lim_{n\to\infty} \mathbf{P}(B_n) = \lim_{n\to\infty} \mathbf{P}(\bigcup_{k=0}^{n} A_k)$$

Mais $\mathbf{P}(\bigcup_{k=0}^{n} A_k) \leq \mathbf{P}(\bigcup_{k=0}^{n-1} A_k) + \mathbf{P}(A_n)$ en utilisant la propriété $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$ vue en ii), et par récurrence, $\mathbf{P}(B_n) = \mathbf{P}(\bigcup_{k=0}^{n} A_k) \leq \sum_{k=0}^{n} \mathbf{P}(A_k) \leq \sum_{n=0}^{\infty} \mathbf{P}(A_n)$. Donc, en passant à la limite :

$$\mathbf{P}(\bigcup_{n=0}^{\infty} A_n) \leq \sum_{n=0}^{\infty} \mathbf{P}(A_n)$$

EXEMPLES :

❑ Soit une pièce truquée ayant la probabilité $p \in [0, 1]$ de tomber sur Pile. On considère une expérience aléatoire consistant à lancer la pièce jusqu'à obtenir Pile, où à la lancer indéfiniment s'il n'y a que des Faces qui apparaissent. Soit A_n l'événement : Pile apparaît au-delà du n-ème lancer ou bien n'apparaît jamais. A_n est aussi l'événement : les n premiers lancers sont des Faces. On a :

$$\mathbf{P}(A_n) = (1 - p)^n$$
$$A_{n+1} \subset A_n$$

L'événement $A = \bigcap_{n=0}^{\infty} A_n$ est l'événement : Pile n'apparaît jamais. C'est aussi l'événement : on tire indéfiniment des Faces. Comme la suite des événements (A_n) décroît, on a, d'après (vi) :

$$\mathbf{P}(A) = \lim_{n\to\infty} \mathbf{P}(A_n) = \lim_{n\to\infty} (1-p)^n = \begin{cases} 0 \text{ si } p > 0 \\ 1 \text{ si } p = 0 \end{cases}$$

Si $p = 0$, on ne sera guère surpris de voir que la probabilité de ne tirer que des Faces est 1. Dans le cas contraire, si $p > 0$, aussi petit soit-il, la probabilité que A se produise est nulle. Conceptuellement, A est un événement envisageable (il est différent de $\emptyset$), mais sa probabilité est nulle. On dit que A est **presque sûrement impossible**. Son événement contraire, consistant à obtenir un moment ou un autre un Pile, et qui est de probabilité 1, est dit **presque sûrement certain**.

❑ L'exemple qui suit généralise l'exemple précédent, mais relève d'une abstraction plus grande. Soit $\Omega = \{(x_n)_{n\geq 1} \mid \forall\, n, x_n = \text{P ou } x_n = \text{F}\}$ l'univers correspondant au lancer indéfiniment d'une pièce équilibrée. Montrons que la probabilité de tirer une suite stationnaire (i.e. à partir d'un certain rang on ne tire que des P ou bien on ne tire que des F) est nulle.

- Pour tout $i \geq m \geq 1$, soit $A_{im} = \{(x_n)_{n\geq 1} \mid \forall k \in [\![m, i]\!], x_k = P \text{ ou } \forall k \in [\![m, i]\!], x_k = F\}$. A_{im} est l'ensemble des suites de lancers telles que les lancers compris entre le m-ème et le i-ème sont tous des P ou bien tous des F. Les autres lancers sont quelconques. Donc le calcul de $\mathbf{P}(A_{im})$ ne fait intervenir que les lancers entre le m-ème et le i-ème et :

$$\mathbf{P}(A_{im}) = 2 \times \frac{1}{2^{i-m+1}} = \frac{1}{2^{i-m}}$$

- Pour tout $m \geq 1$, soit $A_m = \{(x_n)_{n\geq 1} \mid \forall k \geq m, x_k = P \text{ ou } \forall k \geq m, x_k = F\}$. A_m est l'ensemble des suites de lancers telles que tous les lancers à partir du m-ème sont identiques. On a :

$$A_m = \bigcap_{i=m}^{\infty} A_{im} \quad \text{intersection décroissante car } A_{i+1,m} \subset A_{im}$$

donc, d'après (vi), $\mathbf{P}(A_m) = \lim\limits_{i\to\infty} \mathbf{P}(A_{im}) = 0$

- Soit $A = \{(x_n)_{n\geq 1} \mid \exists m, \forall k \geq m, x_k = P \text{ ou } \forall k \geq m, x_k = F\}$. A est l'ensemble des suites stationnaires.

$$A = \bigcup_{m=1}^{\infty} A_m \qquad \text{union croissante car } A_m \subset A_{m+1}$$

donc, d'après (v), $\mathbf{P}(A) = \lim\limits_{m\to\infty} \mathbf{P}(A_m) = 0$

❑ Soient A et B deux événements presque sûrement certains, alors $A \cap B$ est aussi un événement presque certain. En effet :

$$1 \geq \mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$$
$$= 2 - \mathbf{P}(A \cap B) \qquad \text{puisque } \mathbf{P}(A) = \mathbf{P}(B) = 1$$

donc $\mathbf{P}(A \cap B) \geq 1$ donc $\mathbf{P}(A \cap B) = 1$

On peut aussi écrire :

$$\mathbf{P}((A \cap B)^c) = \mathbf{P}(A^c \cup B^c) \leq \mathbf{P}(A^c) + \mathbf{P}(B^c) = 0$$

5- Probabilités conditionnelles

On rappelle rapidement les notions vues en première années, en les généralisant éventuellement au cas général :

DEFINITION

Soit B *un événement de probabilité non nul, d'un espace probabilisé* $(\Omega, \mathcal{T}, \mathbf{P})$. *On définit la* **loi de probabilité sachant** B (*ou* **conditionnelle** *par rapport à* B) *par* :

$$\mathbf{P}_B : \mathcal{T} \to [0, 1]$$
$$A \to \mathbf{P}_B(A) = \frac{P(A \cap B)}{P(B)}$$

On la note aussi $\mathbf{P}(A \mid B)$

On a donc $\mathbf{P}(A \cap B) = \mathbf{P}(B) \times \mathbf{P}_B(A)$ (**formule des probabilités composées**). Dans le cas où $\mathbf{P}(B) = 0$, on peut donner à l'égalité précédente la signification $0 = 0$ même si $\mathbf{P}_B(A)$ n'est pas définie, car $\mathbf{P}(A \cap B) \leq \mathbf{P}(B)$ donc on a aussi $\mathbf{P}(A \cap B) = 0$.

$\mathbf{P}_B$ est une loi de probabilité. En effet :

i) $\mathbf{P}_B$ est définie sur $\mathcal{T}$ à valeurs dans [0, 1] car, puisque $A \cap B \subset B$, $\mathbf{P}(A \cap B) \leq \mathbf{P}(B)$.

ii) $\mathbf{P}_B(\Omega) = \dfrac{P(\Omega \cap B)}{P(B)} = 1$ car $\Omega \cap B = B$

iii) Si les $(A_n)_{n\in\mathbf{N}}$ sont deux à deux incompatibles, alors les $A_n \cap B$ le sont également et l'on a :

$$\mathbf{P}_B\left(\bigcup_{n=0}^{\infty} A_n\right) = \frac{P\left(\left(\bigcup_{n=0}^{\infty} A_n\right) \cap B\right)}{P(B)} = \frac{P\left(\bigcup_{n=0}^{\infty} (A_n \cap B)\right)}{P(B)}$$

$$= \frac{\sum_{n=0}^{\infty} P(A_n \cap B)}{P(B)} = \sum_{n=0}^{\infty} \mathbf{P}_B(A_n)$$

DEFINITION

Soit I *une famille finie ou dénombrable.* $(A_i)_{i\in I}$ *forme un* **système complet d'événements** *si et seulement si les trois conditions suivantes sont vérifiées* :

i) $\forall\, i \in I,\ A_i \neq \varnothing$

ii) $\forall\, i \neq j,\ A_i \cap A_j = \varnothing$

iii) $\Omega = \bigcup_{i\in I} A_i$

Du point de vue ensembliste, $(A_i)_{i\in I}$ forme une partition de Ω.

PROPOSITION (formule des probabilités totales)

Soit $(A_i)_{i\in I}$ *un système complet d'événements d'un espace probabilisé* $(\Omega, \mathcal{T}, \mathbf{P})$ *tels que, pour tout* i, $\mathbf{P}(A_i) > 0$, *et* B *un autre événement de cet espace. Alors* :

$$\mathbf{P}(B) = \sum_{i\in I} \mathbf{P}(A_i) \times \mathbf{P}_{A_i}(B)$$

<u>Démonstration</u> :

❑ $(A_i)_{i\in I}$ formant un système complet d'événements, on a :

$\Omega = \bigcup_{i\in I} A_i$ union disjointe finie ou dénombrable

D'où $\mathbf{P}(B) = \mathbf{P}(\Omega \cap B)$

$$= \mathbf{P}\left(\left(\bigcup_{i\in I} A_i\right) \cap B\right)$$

$$= \mathbf{P}\left(\bigcup_{i\in I} (A_i \cap B)\right) \qquad \text{cette union étant disjointe}$$

$$= \sum_{i\in I} \mathbf{P}(A_i \cap B)$$

$$= \sum_{i\in I} \mathbf{P}(A_i) \times \mathbf{P}_{A_i}(B)$$

La formule reste valide pour une famille quelconque $(A_i)_{i\in I}$ d'événements deux à deux incompatibles (toujours avec I fini ou dénombrable) à condition que $\sum_{i\in I} \mathbf{P}(A_i) = 1$. Il suffit, dans la démonstration précédente de vérifier qu'on a toujours $\mathbf{P}(\Omega \cap B) = \mathbf{P}((\bigcup_{i\in I} A_i) \cap B)$. Soit D l'événement contraire de $\bigcup_{i\in I} A_i$. On a $\Omega = \bigcup_{i\in I} A_i \cup D$ (union disjointe), donc :

$$\mathbf{P}(\Omega) = \sum_{i\in I} \mathbf{P}(A_i) + \mathbf{P}(D)$$

donc $\mathbf{P}(D) = 0$ et a fortiori $\mathbf{P}(D \cap B) = 0$, donc on a bien :

$$\mathbf{P}(\Omega \cap B) = \mathbf{P}((\bigcup_{i\in I} A_i) \cap B) + \mathbf{P}(D \cap B) = \mathbf{P}((\bigcup_{i\in I} A_i) \cap B)$$

On déduit de la formule des probabilités totales la formule suivante :

PROPOSITION (Formule de Bayes)
Soit $(A_i)_{i\in I}$ *un système complet d'événements d'un espace probabilisé* $(\Omega, \mathcal{T}, \mathbf{P})$ *tels que, pour tout i,* $\mathbf{P}(A_i) > 0$, *et* B *un autre événement de cet espace tel que* $\mathbf{P}(B) > 0$. *Alors* :

$$\mathbf{P}_B(A_i) = \frac{\mathbf{P}(B \cap A_i)}{\mathbf{P}(B)} = \frac{\mathbf{P}(A_i) \times \mathbf{P}_{A_i}(B)}{\sum \mathbf{P}(A_j) \times \mathbf{P}_{A_j}(B)}$$

EXEMPLE :

❑ Pour tout $n \geq 1$, on considère une urne U_n contenant une boule blanche et $n - 1$ boules noires. On choisit une urne au hasard, l'urne n étant choisie avec la probabilité $\frac{1}{2^n}$ (n est par exemple le nombre de lancers d'une pièce à effectuer pour obtenir Pile). L'événement A_n est réalisé si l'urne U_n est choisie. Puis on tire au hasard une boule dans l'urne choisie. L'événement B est réalisé si on tire une boule blanche. Ayant réalisé successivement ces deux choix, un expérimentateur annonce qu'il a effectivement tiré une boule blanche. On demande la probabilité qu'il ait choisi l'urne U_1.

On a ici $\mathbf{P}_{A_n}(B) = \frac{1}{n}$ pour tout n, donc :

$$\mathbf{P}(B) = \sum_{n=1}^{\infty} \mathbf{P}(A_n) \times \mathbf{P}_{A_n}(B) = \sum_{n=1}^{\infty} \frac{1}{n2^n} = -\ln(\frac{1}{2}) = \ln(2)$$

Donc $\mathbf{P}_B(A_1) = \dfrac{\mathbf{P}(A_1)\, \mathbf{P}_{A_1}(B)}{\mathbf{P}(B)} = \dfrac{1}{2} \times 1 \times \dfrac{1}{\ln(2)} = \dfrac{1}{2\ln(2)} \approx 0{,}72$

Voici ci-dessous une simulation en Python permettant de tester ce résultat :

```
# On lance une pièce jusqu'à ce qu'elle tombe sur Pile.
# Soit n le nombre de lancers.
# On tire alors une boule dans une urne contenant n boules dont
une seule
# blanche. On cherche la probabilité de tirer une boule
blanche,
# ainsi que, une boule blanche étant tirée, la probabilité
qu'elle
# provienne de la première urne.
```

```
import numpy as np

# Simulation d'un lancer de pièce. Pour tout entier n,
# np.random.randint(n) donne un nombre aléatoire entre 0 et n-1
compris
# selon une loi uniforme.
def piece():
    return(np.random.randint(2))

# On lance la pièce jusqu'à tomber sur Pile (= 1)
def lancePiece():
    n=1
    while piece()<>1:
        n+=1
    return(n)

#  Simulation  d'un  tirage  d'urne  ayant  n  boules.  On  peut
convenir que la
# boule blanche est la boule n°0
def urne(n):
    return(np.random.randint(n))

# N = nombre de tirages
# A = nombre de fois où l'urne 0 est choisie
# B = nombre de boules blanches tirées
# AB = nombre de fois où l'urne 0 est choisie et où la boule
blanche est
#      choisie
def simule(N):
    B=0
    A=0
    AB=0
    for i in range(N):
        n=lancePiece()
        m=urne(n)
        if n==1:
            A+=1
            if m==0:
                B+=1
                AB+=1
        elif m==0:
            B+=1
    pB=float(B)/N    # conversion de l'entier B en flottant car
le
                     #  comportement  de  /  diffère  suivant  les
versions de
                     # Python
    pA=float(A)/N
    pAsachantB=float(AB)/B
    print("proba de tirer la blanche = "+str(pB))
        # proba théorique : ln(2)
    print("proba de tirer l'urne 0 = "+str(pA))
        # proba théorique : 1/2
```

```
    print("proba de tirer l'urne 0 sachant qu'on a tiré la blanche = "+str(pAsachantB))
        # proba théorique : 1/(2ln(2))
# Exécuter simule(N) avec la valeur N de son choix.
```

DEFINITION

Une famille finie ou dénombrable d'événements $(A_i)_{i\in I}$ *sont* **indépendants** *si, pour toute sous-famille de p événements* $A_{i_1}, A_{i_2}, \dots, A_{i_p}$ (*les indices* $i_1, \dots, i_p$ *étant distincts*), *on a* :

$$\mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_p}) = \mathbf{P}(A_{i_1})\mathbf{P}(A_{i_2})\dots\mathbf{P}(A_{i_p})$$

Si A et B sont indépendants, on a $\mathbf{P}_B(A) = \mathbf{P}(A)$.

Un exemple typique de famille dénombrable d'événements indépendants est donné par le lancer d'une pièce indéfiniment, avec A_n l'événement "le n-ème tirage est Pile". Il en est de même pour le lancer de dés.

EXEMPLE :

❑ Voici un problème historique, proposé par Huygens dans son *De ratiociniis in ludo aleae* (1657), Proposition XIV. Nous sommes dans les débuts du développement d'une théorie des probabilités : "*Si un autre joueur et moi jettent tour à tour deux dés à condition que j'aurai gagné dès que j'aurai jeté 7 points et lui dès qu'il en aura jeté 6, tandis que je lui laisse le premier coup, trouver le rapport de ma chance à la sienne*".

Version moderne : deux joueurs lancent à tour de rôle deux dés. Le premier joueur gagne s'il obtient une somme égale à 6 et le deuxième s'il obtient une somme égale à 7, sinon le joueur suivant relance les dés. Trouver la probabilité de gain de chacun.

Il y a une probabilité de $\dfrac{5}{36}$ de tirer une somme égale à 6 avec deux dés, et une probabilité de $\dfrac{6}{36}$ de tirer une somme égale à 7 (revoir au besoin de loi de la variable aléatoire égale à la somme de deux dés dans le chapitre L1/PROBA1.PDF).

Le premier joueur gagne au bout de $2n + 1$ lancers avec une probabilité $(\dfrac{31}{36} \times \dfrac{30}{36})^n \dfrac{5}{36}$, correspondant à une suite de tirages $xyxy\dots xy6$, où $x \neq 6$, $y \neq 7$ et où xy apparaît n fois. On obtient la probabilité précédente en supposant que les tirages sont indépendants et donc que :

$$\mathbf{P}(\{xyxy\dots xy6\}) = \mathbf{P}(\{x\})\mathbf{P}(\{y\})\mathbf{P}(\{x\})\mathbf{P}(\{y\})\dots\mathbf{P}(\{x\})\mathbf{P}(\{y\})\mathbf{P}(\{6\})$$

avec $\mathbf{P}(\{x\}) = 1 - \dfrac{5}{36} = \dfrac{31}{36}$ et $\mathbf{P}(\{y\}) = 1 - \dfrac{6}{36} = \dfrac{30}{36}$.

La probabilité de gain du premier jour est donc :

$$\sum_{n=0}^{\infty} (\frac{31}{36} \times \frac{30}{36})^n \frac{5}{36} = \frac{5}{36} \frac{1}{1 - \dfrac{31}{36} \times \dfrac{30}{36}} = \frac{180}{1296 - 930} = \frac{180}{366} = \frac{30}{61}.$$

Le deuxième joueur gagne au bout de $2n + 2$ lancers avec une probabilité $\dfrac{31}{36} (\dfrac{30}{36} \times \dfrac{31}{36})^n \dfrac{6}{36}$. Sa probabilité de gain est donc $\displaystyle\sum_{n=0}^{\infty} \frac{31}{36} (\frac{30}{36} \times \frac{31}{36})^n \frac{6}{36} = \frac{31}{36} \times \frac{6}{5} \times \frac{30}{61} = \frac{31}{61}$

On remarque que la somme des deux probabilités vaut 1, égale à la probabilité que la partie s'arrête sur le gain de l'un ou l'autre joueur. La probabilité que la partie dure indéfiniment est donc nulle.

II : Variables aléatoires

1- Définition

Soit $(\Omega, \mathcal{T}, \mathbf{P})$ un espace probabilisé et X une fonction de Ω dans $\mathbb{R}$. On se limitera dans la suite au cas où X prend un nombre fini ou dénombrable de valeurs. Soit x élément de $X(\Omega)$. On note $\{X = x\}$ l'ensemble $\{\omega \in \Omega,\ X(\omega) = x\} = X^{-1}(\{x\})$. C'est une partie de Ω. On pourra en donner la probabilité si cette partie est un élément de la tribu $\mathcal{T}$. D'où la définition :

DEFINITION

On appelle **variable aléatoire discrète** *définie sur un espace probabilisé* $(\Omega, \mathcal{T}, \mathbf{P})$ *une fonction* X *de* Ω *dans* $\mathbb{R}$ *telle que* $X(\Omega)$ *est fini ou dénombrable, et telle que l'image réciproque par* X *de tout élement de* $X(\Omega)$ *est élément de la tribu* $\mathcal{T}$.

On note $\mathbf{P}(X = x)$ ou $\mathbf{P}_X(\{x\})$ la quantité $\mathbf{P}(\{\omega \in \Omega,\ X(\omega) = x\}) = \mathbf{P}(X^{-1}(\{x\}))$, et d'une manière générale, si A est une partie de $X(\Omega)$, on note $\mathbf{P}(X \in A)$ ou $\mathbf{P}_X(A)$ la quantité :

$$\mathbf{P}(\{\omega \in \Omega,\ X(\omega) \in A\}) = \mathbf{P}(X^{-1}(A))$$

$\mathbf{P}_X$ s'appelle la **loi** de X.

EXEMPLE :

❑ Soit $\Omega = \mathbb{N}$ et $\mathbf{P}$ définie par $\mathbf{P}(\{n\}) = \dfrac{1}{2^{n+1}}$. Soit X l'application qui à n associe son chiffre des unités. Quelle est la loi de X ? X prend ses valeurs dans $\{0, 1, 2, ..., 9\}$ et, pour k dans cet ensemble, on a :

$$\mathbf{P}_X(\{k\}) = \sum_{n=0}^{\infty} \mathbf{P}(\{10n + k\}) = \sum_{n=0}^{\infty} \frac{1}{2^{10n+k+1}} = \frac{1}{2^{k+1}} \sum_{n=0}^{\infty} \frac{1}{2^{10n}} = \frac{1}{2^{k+1}} \frac{1}{1 - \dfrac{1}{2^{10}}}$$

On peut vérifier que $\displaystyle\sum_{k=0}^{9} \mathbf{P}_X(\{k\}) = 1$.

PROPOSITION

$\mathbf{P}_X$ *est une loi de probabilité sur* $X(\Omega)$ *muni de la tribu de l'ensemble de ses parties*.

Démonstration :

❑ Il est clair que $\mathbf{P}_X$ est à valeurs dans $[0, 1]$.

On a $\mathbf{P}_X(X(\Omega)) = \mathbf{P}(X^{-1}(X(\Omega)) = \mathbf{P}(\Omega) = 1$.

Par ailleurs, soit $(A_n)_{n\in\mathbb{N}}$ une famille d'événements incompatibles sur $X(\Omega)$. Alors :

$$\mathbf{P}_X\Big(\bigcup_{n=0}^{\infty} A_n\Big) = \mathbf{P}\Big(X^{-1}\Big(\bigcup_{n=0}^{\infty} A_n\Big)\Big)$$

Or $X^{-1}\Big(\displaystyle\bigcup_{n=0}^{\infty} A_n\Big) = \displaystyle\bigcup_{n=0}^{\infty} X^{-1}(A_n)$. En effet :

$$\omega \in X^{-1}(\bigcup_{n=0}^{\infty} A_n)$$
$$\Leftrightarrow \quad X(\omega) \in \bigcup_{n=0}^{\infty} A_n$$
$$\Leftrightarrow \quad \exists\, n,\ X(\omega) \in A_n$$
$$\Leftrightarrow \quad \exists\, n,\ \omega \in X^{-1}(A_n)$$
$$\Leftrightarrow \quad \omega \in \bigcup_{n=0}^{\infty} X^{-1}(A_n)$$

Donc :

$$\mathbf{P}_X(\bigcup_{n=0}^{\infty} A_n) = \mathbf{P}(\bigcup_{n=0}^{\infty} X^{-1}(A_n))$$

Les A_n étant deux à deux incompatibles, il en est de même des $X^{-1}(A_n)$. En effet, pour n différent de m :

$$\omega \in X^{-1}(A_n) \cap X^{-1}(A_m) \Leftrightarrow X(\omega) \in A_n \cap A_m = \varnothing \quad \text{impossible}$$

Donc :

$$\mathbf{P}_X(\bigcup_{n=0}^{\infty} A_n) = \sum_{n=0}^{\infty} \mathbf{P}(X^{-1}(A_n)) = \sum_{n=0}^{\infty} \mathbf{P}_X(A_n)$$

On a défini la loi de X à partir de la loi de probabilité sur Ω. Il existe une réciproque que nous admettrons. On peut se donner à priori la loi de X et en déduire une loi de probabilité **P** sur Ω de façon que X soit une variable aléatoire discrète dont la loi se déduit de **P**. Plus précisément, si X prend ses valeurs dans $\{x_n,\ n \in I\}$, I étant fini ou dénombrable et les x_n étant distincts et si (p_n) est telle que $\sum_{n \in I} p_n = 1$, alors il existe une probabilité sur Ω telle que $\mathbf{P}(X = x_n) = p_n$.

Comme nous le verrons, dans la plupart des cas, on a rarement besoin d'expliciter Ω et **P**. On raisonne directement sur $\mathbf{P}_X$.

2- Loi de Poisson, convergence de la loi binomiale vers la loi de Poisson

Considérons la situation suivante, pendant un intervalle de temps [0, T] :

- Des personnes arrivent l'une après l'autre devant un guichet. La variable aléatoire X est le nombre de personnes qui arrivent dans l'intervalle de temps [0, T].
- Un détecteur de particules élémentaires compte le nombre X de particules lors d'une mesure se déroulant pendant l'intervalle de temps [0, T].
- Un appareil est susceptible de subir de interventions ou des réparations, suite à des pannes ou des anomalies de fonctionnement. Pendant l'intervalle de temps [0, T], on compte le nombre X d'interventions.

Dans chaque cas, on s'intéresse à la loi de X. Nous dirons qu'un phénomène est observé dans l'intervalle de temps $[t_1, t_2]$ si une personne arrive au guichet pendant cet intervalle de temps, ou si une particule est détectée, ou si une anomalie de fonctionnement de l'appareil survient. Nous supposerons que la probabilité qu'un phénomène soit observé pendant l'intervalle de temps $[t_1, t_2]$ ne dépend que de la longueur $t_2 - t_1$ de l'intervalle et est indépendant de ce qui se passe avant t_1 ou après t_2. Cette hypothèse suppose une régularité des phénomènes au cours du temps (pas d'heure de pointe, pas de mémoire de la durée de vie passée dans le cas de particule radioactive, pas d'usure de la machine augmentant le risque de panne au cours du temps). Nous appellerons λ le nombre moyen de phénomènes observés pendant l'intervalle de temps [0, T].

Divisons [0, T] en n = 100 intervalles de longueur égale $[t_i, t_{i+1}]$. n est choisi assez grand de façon qu'il soit très peu probable que deux phénomènes soient observés dans le même intervalle $[t_i, t_{i+1}]$. Ainsi, dans chacun des n intervalles, ou bien on observe un phénomène, ou bien on n'observe aucun phénomène. Soit p_n la probabilité commune d'observation d'un phénomène pendant l'intervalle de temps $[t_i, t_{i+1}]$ et X_n le nombre d'événements observés pendant l'intervalle [0, T]. La loi de X_n est alors une loi binomiale $\mathcal{B}(n, p_n)$. L'hypothèse de la loi binomiale repose sur le fait que deux phénomènes ne peuvent se produire dans le même intervalle $[t_i, t_{i+1}]$, que la probabilité d'observation d'un tel phénomène ne dépend pas dudit intervalle et que la survenue de deux événements dans des intervalles distincts se fait de façon indépendante. La première hypothèse est d'autant mieux vérifiée que la longueur de l'intervalle est petite, et donc que n est grand. A la limite, quand n tend vers $+\infty$, p_n vers 0 de façon que np_n (nombre moyen d'événements observés dans le cas de X_n) tende vers λ (nombre moyen d'événements observés dans le cas de X). Nous allons en déduire une expression de $\mathbf{P}(X = k)$. Pour tout n :

$$\begin{aligned}
\mathbf{P}(X_n = k) &= \binom{n}{k} p_n^{\ k} (1 - p_n)^{n-k} \\
&= \frac{n(n-1)...(n-k+1)}{k!} p_n^{\ k} (1 - p_n)^{n-k} \\
&= \frac{1}{k!} \prod_{i=0}^{k-1} (np_n - ip_n) \times (1 - p_n)^{n-k} \\
&\sim \frac{1}{k!} \lambda^k \times (1 - p_n)^{n-k} \qquad \text{car, pour tout } i,\ np_n - ip_n \to \lambda \\
&\sim \frac{1}{k!} \lambda^k \times \exp((n-k)\ln(1 - p_n))
\end{aligned}$$

Par ailleurs :

$$\exp((n-k)\ln(1-p_n)) = \exp((n-k)(-p_n + o(p_n)) = \exp(-\lambda + o(1)) \to \exp(-\lambda)$$

A la limite, on trouve donc :

$$\mathbf{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

k peut prendre maintenant des valeurs entières positives arbitraires. On reconnaît d'ailleurs dans $\dfrac{\lambda^k}{k!}$ le terme général du développement en série entière de e^{λ}. La somme des $\mathbf{P}(X = k)$ pour k variant de 0 à $+\infty$ et donc bien 1.

La loi que nous venons de découvrir s'appelle **loi de Poisson** $\mathcal{P}(\lambda)$. Elle ne dépend que du paramètre λ. Nous avons également mis en évidence le théorème suivant de convergence de la loi binomiale vers la loi de Poisson.

PROPOSITION

Soit $\lambda > 0$ *et* (X_n) *une suite des variables aléatoires de loi binomiale* $\mathcal{B}(n, p_n)$, *où* (p_n) *est une suite telle* $\lim\limits_{n\to\infty} np_n = \lambda$. *Alors :*

$$\forall\, k \in \mathbb{N},\ \lim_{n\to\infty} \mathbf{P}(X_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

En pratique, on peut approximer la loi binomiale par la loi de Poisson dès que $p < \frac{1}{10}$, $n > 30$, $\lambda = np < 16$.

EXEMPLE :

❑ Une usine fabrique 1000 pièces dont 3 % sont défectueuses. On en tire 100. Quelle est la probabilité d'en avoir 5 défectueuses ?

Soit X le nombre de pièces défectueuses.

- Si le tirage a lieu sans remise, X suit de façon exacte une loi hypergéométrique (voir le chapitre L1/PROBA1.PDF). La probabilité cherchée est :

$$\mathbf{P}(X = 5) = \frac{\binom{30}{5}\binom{970}{95}}{\binom{1000}{100}} \approx 0{,}1024$$

- Si les tirages se font avec remise, on a une loi binomiale $\mathcal{B}(100, \frac{3}{100})$, d'où une évaluation de $\mathbf{P}(X = 5)$:

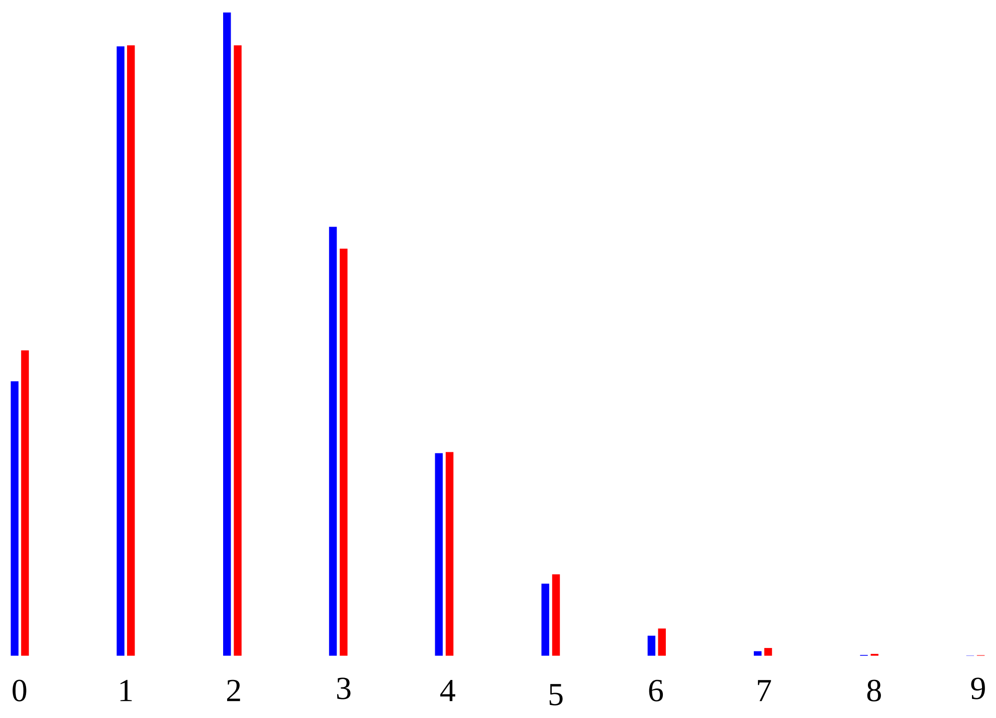
$$\mathbf{P}(X = 5) = \binom{100}{5} 0{,}03^5 \times 0{,}97^{95} \approx 0{,}1013$$

- 100 étant lui-même élevé, $p = 0.03$ étant faible, et $\lambda = np = 3$ étant raisonnable, on peut approximer la loi binomiale par la loi de Poisson $\mathcal{P}(3)$. D'où une nouvelle évaluation :

$$\mathbf{P}(X = 5) = \frac{3^5}{5!} e^{-3} \approx 0{,}1008$$

C'est évidemment cette dernière quantité qui est le plus facile à calculer. Avec deux chiffres significatifs, le résultat obtenu est correct. Les deux autres décimales n'ont été données que pour permettre d'apprécier la différence de précision entre les diverses lois.

Ci-dessous, on donne une représentation graphique en histogramme, avec $n = 20$ et $p = 0.1$, pour $0 \leq k \leq 9$. Les bâtons bleus représentent la répartition de la loi binomiale et les bâtons rouges ceux de la loi de Poisson.



3- Loi géométrique

Cette loi intervient dans les situations suivantes :

❑ On lance une pièce jusqu'à ce qu'on tombe sur Pile. On cherche la loi du nombre de lancers.

❑ On joue au loto jusqu'à ce qu'on gagne le gros lot. On cherche la loi du nombre de parties.

❑ Plus généralement, on répète la même expérience de Bernoulli (à deux issues, échec ou succès), jusqu'à obtenir un succès. On cherche la loi du nombre X d'expériences. X peut éventuellement prendre la valeur $+\infty$ si le succès ne se produit jamais.

Soit p la probabilité d'un succès et $1 - p$ celle d'un échec. Alors :

$$\mathbf{P}_X(\{k\}) = \mathbf{P}(X = k) = (1-p)^{k-1}p$$

car il y a $k - 1$ échecs avant d'avoir un succès. k varie de 1 à $+\infty$. La somme des $\mathbf{P}(X = k)$ pour k élément de $\mathbb{N}^*$ vaut 1 si p appartient à $]0, 1[$. Cela signifie aussi que $\mathbf{P}(X = \infty) = 0$, et donc qu'il est certain que le succès se produira au bout d'un temps fini (c'est encourageant pour les joueurs de loto) mais peut-être long (quelques dizaines de milliers d'années en moyenne pour le gros lot du joueur de loto !).

Deux cas dégénérés peuvent être considérés :

$p = 1$. Alors X est la variable certaine égale à 1 (succès à la première expérience).

$p = 0$. Alors X est la variable certaine $+\infty$. $\mathbf{P}(X = +\infty) = 1$ (échec certain indéfiniment).

La loi géométrique caractérise les phénomènes sans mémoire et sans usure. Calculons d'abord $\mathbf{P}(X > k)$.

$$\mathbf{P}(X > k) = \sum_{n=k+1}^{\infty} (1-p)^{n-1}p = (1-p)^k$$

Cette valeur se retrouve d'ailleurs directement en considérant que X > k signifie k échecs de suite. On en tire ensuite les probabilités conditionnelles suivantes :

$$\begin{aligned}\mathbf{P}(X > n + k \mid X > k) &= \frac{\mathbf{P}(X > n + k \text{ et } X > k)}{\mathbf{P}(X > k)} \\ &= \frac{\mathbf{P}(X > n + k)}{\mathbf{P}(X > k)} \qquad \text{car } X > n + k \text{ et } X > k \Leftrightarrow X > n + k \\ &= \frac{(1-p)^{n+k}}{(1-p)^k} = (1-p)^n \\ &= \mathbf{P}(X > n)\end{aligned}$$

Ainsi, sachant que X est supérieur à k, la loi de X – k est la même que la loi de X. Cela signifie donc que l'on peut oublier les k premières expériences et recommencer à zéro ; la loi du nombre d'expériences jusqu'au succès est inchangée. Cette règle est rarement connue, et contraire à l'intuition courante. La plupart des gens pense que le fait de lancer un dé cinq fois sans que le n° 6 sorte rend sa probabilité de sortie au coup suivant plus grande. Ou encore, le fait de jouer au loto sans gagner depuis des semaines encourage certaines personnes à persévérer en pensant que cela va bientôt être leur tour de gagner. Il n'en est rien. Elles ont autant de chance de gagner ou de perdre qu'avant, et cela indépendamment de leurs gains ou pertes passés.

Inversement, soit X une variable aléatoire sur $\mathbb{N}^*$ vérifiant :

$$\forall k > 0, \forall n > 0, \mathbf{P}(X > n + k \mid X > k) = \mathbf{P}(X > n).$$

Montrons que X suit une loi géométrique. On a :

$$\mathbf{P}(X > n) = \frac{\mathbf{P}(X > n + k \text{ et } X > k)}{\mathbf{P}(X > k)} = \frac{\mathbf{P}(X > n + k)}{\mathbf{P}(X > k)}$$

Donc $\mathbf{P}(X > n + k) = \mathbf{P}(X > n)\, \mathbf{P}(X > k)$

En particulier :

$$\mathbf{P}(X > n + 1) = \mathbf{P}(X > n)\, \mathbf{P}(X > 1)$$

Posons $p = \mathbf{P}(X = 1)$. D'où $\mathbf{P}(X > 1) = 1 - \mathbf{P}(X = 1) = 1 - p$, et $\mathbf{P}(X > n)$ est une suite géométrique de raison $1 - p$. Comme $\mathbf{P}(X > 0) = 1$, on a donc :

$$\mathbf{P}(X > n) = (1 - p)^n$$

et $$\mathbf{P}(X = n) = \mathbf{P}(X > n - 1) - \mathbf{P}(X > n) = (1 - p)^{n-1} - (1 - p)^n = (1 - p)^{n-1}p$$

Ce qui prouve que X est bien de loi géométrique.

EXEMPLE :

❑ Un exemple de phénomène sans mémoire se rencontre dans la désintégration des atomes radioactifs. Il est expérimentalement constaté que, si on dispose à l'instant initial d'un nombre N grand de radionucléides d'un certain type, il existe une constante λ telle que, au bout d'un temps t quelconque, il ne reste plus que $Ne^{-\lambda t}$ atomes non désintégrés. λ est la proportion d'atomes qui se désintègre par unité de temps. Ce comportement peut s'interpréter par la modélisation suivante. Considérons un intervalle de temps τ très court. La proportion d'atomes se désintégrant pendant la durée τ est $\lambda\tau$. Supposons donc qu'à chacun de ces intervalles de temps τ successifs, le radionucléide tire à pile ou face avec une probabilité $p = \lambda\tau$ de se désintégrer. Soit X le nombre d'intervalles de temps τ au bout duquel il se désintègre. Si les tirages sont indépendants, i.e. si le radionucléide est sans mémoire, X suit une loi géométrique de paramètre p. On a donc :

$$\mathbf{P}(X > n) = (1 - p)^n$$

Si $t = n\tau$, le nombre d'atomes qui ne se sont pas désintégrés au bout du temps t suit une loi binomiale $\mathcal{B}(N, (1-p)^n)$, en supposant que la désintégration d'un atome est indépendante de celle d'un autre, i.e. il n'y a pas de fission induite d'un atome à l'autre. Dans ce cas, l'espérance du nombre d'atomes non désintégrés au bout du temps t est :

$$N(1-p)^n = N \exp(n \ln(1-p))$$

Si on fait tendre n vers $+\infty$, de sorte que $\tau = \frac{t}{n}$ tend vers 0, ainsi que $p = \lambda\tau = \frac{\lambda t}{n}$, alors on a :

$$N(1-p)^n = N \exp(-np + no(p)) = N \exp(-np + no(\frac{1}{n})) = N \exp(-np + o(1)) \to Ne^{-\lambda t}$$

On retrouve la loi de décroissance initiale. Par ailleurs, nous verrons plus bas que l'espérance de X est $\mathbf{E}(X) = \frac{1}{p}$, de sorte que la durée de vie moyenne d'un radionucléide est $\mathbf{E}(X)\tau = \frac{\tau}{p} = \frac{1}{\lambda}$.

On rencontre, suivant les conventions adoptées, des variantes de la loi géométrique.

- On la définit parfois sur $\mathbb{N}$ par : $\mathbf{P}(X = k) = (1-p)^k p$
- Les rôles de p et $1-p$ sont parfois inversés.

Il convient donc de réfléchir à la situation proposée avant d'appliquer mécaniquement les résultats de ce paragraphe.

4- Fonction de répartition

Soit X une variable aléatoire à valeurs réelles. On appelle **fonction de répartition** de X la fonction $F_X : \mathbb{R} \to [0, 1]$ définie par : $F_X(t) = \mathbf{P}(X \le t)$.

PROPRIETES

(i) F_X *est croissante*

(ii) $\lim\limits_{t\to-\infty} F_X(t) = 0$

(iii) $\lim\limits_{t\to+\infty} F_X(t) = 1$

Démonstration :

❑ (i) Si $t \le u$, $\{\omega \in \Omega, X(\omega) \le t\} \subset \{\omega \in \Omega, X(\omega) \le u\}$ donc :

$$F_X(t) = \mathbf{P}(X \le t) \le \mathbf{P}(X \le u) = F_X(u)$$

❑ (ii) Il suffit de montrer que, pour tout suite (t_n) décroissante vers $-\infty$, on a $\lim\limits_{n\to\infty} F_X(t_n) = 0$ (caractérisation séquentielle d'une limite). Pour cela, remarquons que les événements $\{X \le t_n\}$ sont décroissants pour l'inclusion et d'intersection vide. Donc :

$$0 = \mathbf{P}(\varnothing) = \mathbf{P}(\bigcap_{n=0}^{\infty} \{X \le t_n\}) = \lim_{n\to\infty} \mathbf{P}(X \le t_n) = F_X(t_n)$$

❑ (iii) Il suffit de montrer que, pour tout suite (t_n) croissante vers $+\infty$, on a $\lim\limits_{n\to\infty} F_X(t_n) = 1$. Pour cela, remarquons que les événements $\{X \le t_n\}$ sont croissants pour l'inclusion et de réunion Ω. Donc :

$$1 = \mathbf{P}(\Omega) = \mathbf{P}(\bigcup_{n=0}^{\infty} \{X \le t_n\}) = \lim_{n\to\infty} \mathbf{P}(X \le t_n) = F_X(t_n)$$

5- Couple de variables aléatoires

La notion de couple de variables aléatoires est identique à celle vue dans le cas fini, sauf que les sommes qui interviennent peuvent être des séries, voire des séries doubles.

X et Y étant des variables aléatoires discrètes, on définit la loi conjointe du couple par la donnée des $\mathbf{P}(X = x, Y = y)$, $x \in X(\Omega)$, $y \in Y(\Omega)$. Les lois de X et de Y sont appelées **lois marginales**. On peut déterminer les lois marginales à partir de la loi conjointe en utilisant le fait que l'événement $\{X = x\}$ est la réunion disjointe et finie ou dénombrable des événements $\{X = x, Y = y\}$, $y \in Y(\Omega)$:

$$\mathbf{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbf{P}(X = x, Y = y)$$

de même :

$$\mathbf{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbf{P}(X = x, Y = y)$$

Mais la réciproque est fausse car $\{X = x, Y = y\} = \{X = x\} \cap \{Y = y\}$ et d'une façon générale, on ne peut retrouver $\mathbf{P}(A \cap B)$ à partir de $\mathbf{P}(A)$ et $\mathbf{P}(B)$, sauf si les événements sont indépendants.

Si pour tout x et tout y, les événements $\{X = x\}$ et $\{Y = y\}$ sont indépendants, on dit que les variables aléatoires discrètes X et Y sont **indépendantes**. On a dans ce cas :

$$\mathbf{P}(X = x, Y = y) = \mathbf{P}(X = x)\, \mathbf{P}(Y = y)$$

et plus généralement, pour toute partie A incluse dans $X(\Omega)$ et toute partie B incluse dans $Y(\Omega)$, on a :

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A)\, \mathbf{P}(Y \in B)$$

En effet, pour tout x, on a, en tenant compte du fait que l'événement $\{X = x, Y \in B\}$ est la réunion disjointe des événements $\{X = x, Y = y\}$ quand y parcourt B :

$$\begin{aligned}
\mathbf{P}(X = x, Y \in B) &= \sum_{y \in B} \mathbf{P}(X = x, Y = y) \\
&= \sum_{y \in B} \mathbf{P}(X = x)\, \mathbf{P}(Y = y) \\
&= \mathbf{P}(X = x) \sum_{y \in B} \mathbf{P}(Y = y) \\
&= \mathbf{P}(X = x)\, \mathbf{P}(Y \in B)
\end{aligned}$$

puis, en tenant compte du fait que l'événement $\{X \in A, Y \in B\}$ est la réunion disjointe des événements $\{X = x, Y \in B\}$ quand x parcourt A :

$$\begin{aligned}
\mathbf{P}(X \in A, Y \in B) &= \sum_{x \in A} \mathbf{P}(X = x, Y \in B) \\
&= \sum_{x \in A} \mathbf{P}(X = x)\, \mathbf{P}(Y \in B)
\end{aligned}$$

$= (\sum_{x \in A} \mathbf{P}(X = x))\, \mathbf{P}(Y \in B)$

$= \mathbf{P}(X \in A)\, \mathbf{P}(Y \in B)$

Si X et Y sont indépendantes, alors pour toute fonction f et g de $\mathbb{R}$ dans $\mathbb{R}$, $f(X)$ et $g(Y)$ sont indépendantes. En effet, Pour toute valeur u et v que peuvent prendre $f(X)$ et $g(Y)$, on a :

$$\begin{aligned}\mathbf{P}(f(X) = u, g(Y) = v) &= \mathbf{P}(X \in f^{-1}(u), Y \in g^{-1}(v)) \\ &= \mathbf{P}(X \in f^{-1}(u))\, \mathbf{P}(Y \in g^{-1}(v)) \\ &= \mathbf{P}(f(X) = u)\, \mathbf{P}(g(Y) = v)\end{aligned}$$

EXEMPLE :

❑ Pour (n, m) éléments de $\mathbb{N}^2$, soit $\mathbf{P}(X = n, Y = m) = \frac{C}{2^{n+m+1}}$. Déterminons C pour que l'on ait une loi de probabilité et voyons si X et Y sont indépendantes.

La loi de X est $\mathbf{P}(X = n) = \sum_{m=0}^{\infty} \frac{C}{2^{n+m+1}} = \frac{C}{2^n}$ et l'on a $\sum_{n=0}^{\infty} \mathbf{P}(X = n) = 1$ si et seulement si $C = \frac{1}{2}$.

De même, la loi de Y est $\mathbf{P}(Y = m) = \sum_{n=0}^{\infty} \frac{C}{2^{n+m+1}} = \frac{C}{2^m}$.

On a bien $\mathbf{P}(X = n, Y = m) = \frac{1}{2^{n+m+2}} = \frac{1}{2^{n+1}} \frac{1}{2^{m+1}} = \mathbf{P}(X = n)\, \mathbf{P}(Y = m)$

Si les variables aléatoires ne sont pas indépendantes, on aura souvent recours à des probabilités conditionnelles. En vertu de la relation $\mathbf{P}(X = x, Y = y) = \mathbf{P}(Y = y \mid X = x)\, \mathbf{P}(X = x)$, on peut reconstituer la loi du couple si on connaît la loi de X et toutes les valeurs de $\mathbf{P}(Y = y \mid X = x)$, i.e. la loi conditionnelle de Y relativement à X.

EXEMPLE :

❑ On se donne deux réels $\lambda > 0$ et $p \in]0, 1[$. Soit X de loi de Poisson de paramètre λ et Y de loi conditionnelle sachant que $X = n$ de loi binomiale $\mathcal{B}(n, p)$ et $Z = X - Y$. Quelles sont les lois de Y et Z ? Voici quelques cas où l'on rencontre cette situation :

- X est le nombre de clients d'un grand magasin durant une journée donnée. λ est le nombre moyen de clients dans le magasin, Y est le nombre de clients visitant le rayon électroménager, p la proportion de clients visitant ce rayon.
- X est le nombre de voitures arrivant à une station service durant une journée donnée, λ est le nombre moyen de voitures, Y est le nombre de clients qui achètent du diesel, Z ceux qui achètent de l'essence, p la proportion de voitures qui achète du diesel. (Cet exemple risque de devenir de plus en plus obsolète).
- X est le nombre de participants à un congrès, λ le nombre moyen de participants, Y le nombre de personnes qui s'inscrivent au repas, Z le nombre de ceux qui mangent par leur propre moyen, p la proportion de personnes qui prend un repas.

Loi conjointe de (X, Y) :

$$\mathbf{P}(X = n, Y = k) = \mathbf{P}(X = n) \times \mathbf{P}(Y = k \mid X = n)$$

$$= \frac{\lambda^n}{n!} e^{-\lambda} \binom{n}{k} p^k (1-p)^{n-k}$$

Cette quantité est nulle si $k > n$, et sinon se simplifie en $\dfrac{\lambda^n}{k!(n-k)!} e^{-\lambda} p^k (1-p)^{n-k}$. Donc :

$$\mathbf{P}(Y = k) = \sum_{n=k}^{\infty} \mathbf{P}(X = n, Y = k) = \sum_{n=k}^{\infty} \frac{\lambda^n}{k!(n-k)!} e^{-\lambda} p^k (1-p)^{n-k}$$

$$= e^{-\lambda} \frac{(\lambda p)^k}{k!} \sum_{n=k}^{\infty} \frac{\lambda^{n-k}}{(n-k)!} (1-p)^{n-k}$$

On reconnaît dans la somme le développement en série de $\exp(\lambda(1-p))$. Donc :

$$\mathbf{P}(Y = k) = e^{-\lambda} \frac{(\lambda p)^k}{k!} e^{\lambda(1-p)} = \frac{(\lambda p)^k}{k!} e^{-\lambda p}$$

qui n'est autre qu'une loi de Poisson de paramètre λp. Nous verrons plus bas que l'espérance d'une loi de Poisson de paramètre λ est $\mathbf{E}(X) = \lambda$, et donc, en particulier, $\mathbf{E}(Y) = p\lambda$, ce qui, finalement est bien naturel, puisque, dans la population X, une proportion p en moyenne forme la population Y.

Loi conjointe de (Y, Z) :

$$\mathbf{P}(Y = k, Z = n) = \mathbf{P}(X = n + k, Y = k) = \frac{\lambda^{n+k}}{(n+k)!} e^{-\lambda} \binom{n+k}{k} p^k (1-p)^n$$

$$= \frac{\lambda^{n+k}}{n!\,k!} e^{-\lambda} p^k (1-p)^n = \frac{(\lambda p)^k}{k!} e^{-\lambda p} \frac{(\lambda(1-p))^n}{n!} e^{-\lambda(1-p)}$$

En sommant sur k, on obtient $\mathbf{P}(Z = n) = \dfrac{(\lambda(1-p))^n}{n!} e^{-\lambda(1-p)}$, loi de Poisson de paramètre $\lambda(1-p)$. De plus, on constate que Y et Z sont indépendantes.

Réciproquement, on montre un peu plus bas que, si Y et Z sont deux variables aléatoires indépendantes suivant une loi de Poisson de paramètre respectif λp et $\lambda(1-p)$, alors la loi conditionnelle de Y sachant que Y + Z = n est la loi binomiale $\mathcal{B}(n, p)$.

Si l'on dispose de n variables aléatoires $X_1, ..., X_n$, on dira qu'elles sont **mutuellement indépendantes** si et seulement si pour tout $(x_1, ..., x_n)$, les événements $X = x_1, ..., X = x_n$ sont indépendants. Il ne suffit pas que les variables aléatoires soient deux à deux indépendantes pour l'être mutuellement, pas plus qu'il ne suffit pas que des événements soient deux à deux indépendants pour l'être mutuellement.

EXEMPLE :

❑ On lance n dés. Pour $1 \leq i \leq n$, X_i est le tirage du i-ème dé. Notons Y_n la parité de la somme des n dés ($Y_n = X_1 + ... + X_n$ mod 2 = 0 si la somme est paire, 1 si la somme est impaire). $(X_1, ..., X_n)$ sont indépendants, mais $(X_1, ..., X_n, Y_n)$ ne le sont pas puisque la valeur de Y_n est imposée par celle de $X_1, ..., X_n$. Cependant, pour tout $i \leq n$, (X_i, Y_n) sont indépendants. Montrons-le pour $i = n$. On a :

$\mathbf{P}(Y_n = 0, X_n = k) = \mathbf{P}(Y_{n-1} = k \text{ mod } 2, X_n = k)$

où $Y_{n-1} = X_1 + ... + X_{n-1}$ mod 2, indépendant de X_n

$= \mathbf{P}(Y_{n-1} = k \text{ mod } 2)\, \mathbf{P}(X_n = k)$

Mais, pour tout k, $\mathbf{P}(Y_{n-1} = k \text{ mod } 2) = \dfrac{1}{2} = \mathbf{P}(Y_n = 0)$ (car pour chaque dé, il y a autant de chance de tirer un nombre pair qu'un nombre impair, et il en sera de même pour la somme). Donc :

$\mathbf{P}(Y_n = 0, X_n = k) = \mathbf{P}(Y_n = 0)\, \mathbf{P}(X_n = k)$

PROPOSITION

Soit X *de loi de Poisson* $\mathcal{P}(\lambda)$ *et* Y $\mathcal{P}(\mu)$ *deux variables aléatoires indépendantes. Alors* X + Y *suit une loi de Poisson* $\mathcal{P}(\lambda + \mu)$.

Ce résultat s'interprète aussi concrètement en considérant que X est le nombre de clients qui se présenteront derrière un premier guichet pendant un intervalle de temps donné, avec une espérance λ, et que Y est le nombre de clients qui se présenteront derrière un deuxième guichet pendant le même intervalle de temps. On fusionne les deux guichets en un seul, le nombre de clients étant X + Y, avec un nombre moyen de clients de $\lambda + \mu$.

Démonstration :

❑ $\mathbf{P}(X + Y = n) = \sum_{k=0}^{n} \mathbf{P}(X = k)\, \mathbf{P}(Y = n - k)$

$$= \sum_{k=0}^{n} \frac{\lambda^k}{k!} e^{-\lambda} \frac{\mu^{n-k}}{(n-k)!} e^{-\mu}$$

$$= \frac{1}{n!} e^{-(\lambda+\mu)} \sum_{k=0}^{n} \binom{n}{k} \lambda^k \mu^{n-k}$$

On reconnaît le développement du binôme de Newton de $(\lambda + \mu)^n$. Donc :

$$\mathbf{P}(X + Y) = \frac{(\lambda + \mu)^n}{n!} e^{-(\lambda+\mu)}$$

ce qui est bien la loi cherchée.

On prouve de même, par récurrence, que la somme de n variables aléatoires indépendantes suivant des loi de Poisson de paramètre $\lambda_1, ..., \lambda_n$, suit une loi de Poisson de paramètre $\lambda_1 + ... + \lambda_n$.

EXEMPLE :

❑ Soient Y et Z deux variables aléatoires indépendantes indépendantes suivant une loi de Poisson de paramètre respectif λp et $\lambda(1 - p)$ avec $\lambda > 0$ et $p \in\ ]0, 1[$. Quelle est la loi de Y sachant que $Y + Z = n$?

D'après la proposition précédente, la loi de Y + Z est une loi de Poisson de paramètre $\lambda p + \lambda(1 - p) = \lambda$. On a, pour $0 \leq k \leq n$:

$$\mathbf{P}(Y = k \mid Y + Z = n) = \frac{\mathbf{P}(Y = k \text{ et } Y + Z = n)}{\mathbf{P}(Y + Z = n)}$$

$$= \frac{\mathbf{P}(Y = k \text{ et } Z = n - k)}{\mathbf{P}(Y + Z = n)}$$

$$= \frac{\mathbf{P}(Y = k)\mathbf{P}(Z = n - k)}{\mathbf{P}(Y + Z = n)} \qquad \text{par indépendance de Y et Z}$$

$$= \frac{e^{-\lambda p} (\lambda p)^k}{k!} \frac{e^{-\lambda(1-p)} (\lambda(1 - p))^{n-k}}{(n - k)!} \frac{n!}{e^{-\lambda} \lambda^n}$$

$$= \frac{p^k}{k!} \frac{(1 - p)^{n-k}}{(n - k)!} n! = \binom{n}{k} p^k (1 - p)^{n-k}$$

Il s'agit de la loi binomiale $\mathcal{B}(n, p)$.

6- Espérance d'une variable aléatoire

L'espérance d'une variable aléatoire discrète se calcule comme dans le cas fini, mais le nombre de valeurs prises par la variable peut être infini dénombrable. Si X est à valeurs dans $\{x_n, n \in \mathbb{N}\}$, on dit que X est d'**espérance finie** si la série $\sum x_n \mathbf{P}(X = x_n)$ est **absolument** convergente et on pose :

DEFINITION DE L'ESPERANCE

$$\mathbf{E}(X) = \sum_{n=0}^{\infty} x_n \mathbf{P}(X = x_n)$$

La raison pour laquelle on souhaite que la série $\sum x_n \mathbf{P}(X = x_n)$ soit absolument convergente est que, dans ce cas, la famille $(x_n \mathbf{P}(X = x_n))_{n\in\mathbb{N}}$ est sommable (voir la notion de famille sommable dans le chapitre L2/SERIES.PDF), et que la façon dont on a numéroté les valeurs que prend X est sans importance. On peut permuter l'ordre de numérotation et donc prendre les images dans l'ordre que l'on veut, sans changer la valeur de $\mathbf{E}(X)$.

Pour le calcul de $\mathbf{E}(f(X))$, où f est une fonction définie sur $X(\Omega)$, on dispose du théorème du transfert suivant :

THEOREME DE TRANSFERT

Si X *est une variable aléatoire discrète et* f *une fonction définie sur* $X(\Omega) = \{x_n, n \in \mathbb{N}\}$, *alors* :

(i) $f(X)$ *est une variable aléatoire discrète.*

(ii) $f(X)$ *est d'espérance finie si et seulement si* $\sum f(x_n) \mathbf{P}(X = x_n)$ *est absolument convergente et on a* $\mathbf{E}(f(X)) = \sum_{n=0}^{\infty} f(x_n) \mathbf{P}(X = x_n)$.

La démonstration utilise le théorème de sommation par paquets des familles sommables, exposé dans le chapitre L2/SERIES.PDF. Elle peut être admise si le lecteur la juge trop technique.

Démonstration :

❑ (i) : Montrons que $f(X)$ est une variable aléatoire discrète, et profitons-en pour introduire des notations que nous utiliserons dans la suite de la preuve du théorème. $X(\Omega)$ étant dénombrable (ou même fini), il en est de même de $f(X(\Omega))$. Soient y_k, $k \in K$, K fini ou dénombrable, les valeurs distinctes prises par $f(X)$. Chaque y_k est l'image par f de valeurs x_{mk} prises par X, m parcourant un ensemble fini ou dénombrable. Pour chaque k de K, notons I_k l'ensemble des indices (m, k) des x_{mk} et I la réunion (disjointe car les k sont distincts) des I_k. Les x_{mk}, $(m, k) \in I$, sont toutes les valeurs prises par X, et $\{x_{mk}, (m, k) \in I\} = X(\Omega) = \{x_n, n \in \mathbb{N}\}$.

Vérifions que, pour tout y_k, $f(X)^{-1}(y_k)$ est un élément de la tribu $\mathcal{T}$ des événements de Ω. On a :

$$\begin{aligned} f(X)^{-1}(y_k) &= \{\omega \in \Omega, f(X(\omega)) = y_k\} \\ &= \{\omega \in \Omega, X(\omega) \in f^{-1}(y_k)\} \\ &= \{\omega \in \Omega, \exists\, (m, k) \in I_k, X(\omega) = x_{mk}\} \\ &= \bigcup_{(m,k)\in I_k} \{\omega \in \Omega, X(\omega) = x_{mk}\} \end{aligned}$$

$$= \bigcup_{(m,k)\in I_k} X^{-1}(\{x_{mk}\})$$

X étant une variable aléatoire discrète, chaque partie $X^{-1}(\{x_{mk}\})$ est élément de $\mathcal{T}$, et I_k étant un ensemble fini ou dénombrable, la réunion finie ou dénombrable $\bigcup_{(m,k)\in I_k} X^{-1}(\{x_{mk}\})$ est aussi un élément de $\mathcal{T}$.

❑ (ii) $\Leftarrow$: Supposons que $\sum f(x_n)\, \mathbf{P}(X = x_n)$ soit absolument convergente, et donc que la famille $(f(x_n)\, \mathbf{P}(X = x_n))_{n\in\mathbf{N}}$ soit sommable. Puisque $\{x_{mk}, (m, k) \in I\} = X(\Omega) = \{x_n, n \in \mathbf{N}\}$, l'hypothèse est identique au fait que la famille $(f(x_{mk})\, \mathbf{P}(X = x_{mk}))_{(m,k)\in I}$ est sommable. Le théorème de sommation par paquets permet alors d'écrire :

$$\sum_{n=0}^{\infty} |f(x_n)|\, \mathbf{P}(X = x_n) = \sum_{(m,k)\in I} |f(x_{mk})|\, \mathbf{P}(X = x_{mk})$$

$$= \sum_{k\in K} \sum_{(m,k)\in I_k} |f(x_{mk})|\, \mathbf{P}(X = x_{mk})$$

$$= \sum_{k\in K} \sum_{(m,k)\in I_k} |y_k|\, \mathbf{P}(X = x_{mk})$$

$$= \sum_{k\in K} |y_k| \, (\sum_{(m,k)\in I_k} \mathbf{P}(X = x_{mk}))$$

$$= \sum_{k\in K} |y_k|\, \mathbf{P}(X \in f^{-1}(y_k))$$

$$= \sum_{k\in K} |y_k|\, \mathbf{P}(f(X) = y_k)$$

ce qui prouve que $\sum_{k\in K} |y_k|\, \mathbf{P}(f(X) = y_k)$ est une quantité finie, et donc que $f(X)$ est d'espérance finie. Le même calcul, mené sans les valeurs absolue, conduit à :

$$\sum_{n=0}^{\infty} f(x_n)\, \mathbf{P}(X = x_n) = \sum_{k\in K} y_k\, \mathbf{P}(f(X) = y_k) = \mathbf{E}(f(X))$$

❑ (ii) $\Rightarrow$: Réciproquement, supposons que $f(X)$ soit d'espérance finie, et donc que $\sum_{k\in K} |y_k|\, \mathbf{P}(f(X) = y_k)$ est fini. Avec les notations précédentes, pour chaque k, la famille $(f(x_{mk})\, \mathbf{P}(X = x_{mk}))_{(m,k)\in I_k}$ est sommable puisque :

$$\sum_{(m,k)\in I_k} \left|f(x_{mk})\right| \mathbf{P}(X = x_{mk}) = \left|y_k\right| \mathbf{P}(f(X) = y_k)$$

Posons $T_k = \sum_{(m,k)\in I_k} \left|f(x_{mk})\right| \mathbf{P}(X = x_{mk}) = \left|y_k\right| \mathbf{P}(f(X) = y_k)$. Par hypothèse, la famille $(T_k)_{k\in K}$, qui n'est autre que la famille $(\left|y_k\right| \mathbf{P}(f(X) = y_k))_{k\in K}$, est sommable. Donc le théorème de sommation par paquets (dans le sens réciproque du précédent), permet de conclure que la famille $(\left|f(x_{mk})\right| \mathbf{P}(X = x_{mk}))_{(m,k)\in I}$ est sommable, et donc que $\sum f(x_n) \mathbf{P}(X = x_n)$ est absolument convergente.

EXEMPLE :

❑ Soit r un réel strictement positif. On lance une pièce équilibrée jusqu'à ce qu'elle tombe sur Pile. Si on l'a lancé n fois, on gagne r^n euros. Quelle est l'espérance de gain ?

Soit X le nombre de lancers. Pour tout $n \geq 1$, on a $\mathbf{P}(X = n) = \dfrac{1}{2^n}$. Le gain est r^X. On a :

$$\mathbf{E}(r^X) = \sum_{n=1}^{\infty} r^n \mathbf{P}(X = n) = \sum_{n=1}^{\infty} \frac{r^n}{2^n}$$

La série converge si et seulement si $r < 2$. Seulement dans ce cas, l'espérance de gain existe et vaut :

$$\mathbf{E}(r^X) = \frac{r}{2} \frac{1}{1 - \frac{r}{2}} = \frac{r}{2 - r}$$

Pour $r \geq 2$, l'espérance de gain est infinie.

La régle de transfert précédente s'applique également aux couples (X, Y) de variables aléatoires discrètes. On adapte le raisonnemment précédent en remplaçant x_n par les couples (x_n, y_m) de valeurs prises par X et par Y, et nous nous bornerons à énoncer la conclusion. Soit f une fonction définie sur $X(\Omega) \times Y(\Omega)$. $f(X, Y)$ est une variable aléatoire discrète. Elle est d'espérance finie si et seulement si la famille $(f(x_n, y_m) \mathbf{P}(X = x_n, Y = y_m))_{(n, m)\in\mathbf{N}^2}$ est sommable et :

$$\mathbf{E}(f(X, Y)) = \sum_{(n,m)\in\mathbf{N}^2} f(x_n, y_m) \mathbf{P}(X = x_n, Y = y_m)$$

On vérifie que la famille $(f(x_n, y_m) \mathbf{P}(X = x_n, Y = y_m))_{(n, m)\in\mathbf{N}^2}$ est sommable en montrant par exemple que la somme double $\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \left|f(x_n, y_m)\right| \mathbf{P}(X = x_n, Y = y_m)$ est finie.

On a également la propriété suivante :

PROPOSITION

Si Y *est d'espérance finie et si* $\left|X\right| \leq Y$, *alors* X *est d'espérance finie.*

<u>Démonstration</u> :

❑ Notons $\{x_n, n \in \mathbb{N}\}$ les valeurs prises par X et $\{y_i, i \in \mathrm{I}\}$ celles prises par Y, I étant un ensemble d'indices fini ou dénombrable. Pour tout entier n, on a :

$$|x_n|\ \mathbf{P}(\mathrm{X} = x_n) = \sum_{i\in\mathrm{I}} |x_n|\ \mathbf{P}(\mathrm{X} = x_n, \mathrm{Y} = y_i)$$

$$\leq \sum_{i\in\mathrm{I}} y_i\ \mathbf{P}(\mathrm{X} = x_n, \mathrm{Y} = y_i)$$

En effet, pour tout i, ou bien $\mathbf{P}(\mathrm{X} = x_n, \mathrm{Y} = y_i) = 0$ et le terme correspondant n'a aucune contribution dans la somme, ou bien $\mathbf{P}(\mathrm{X} = x_n, \mathrm{Y} = y_i) > 0$, mais comme $|\mathrm{X}| > \mathrm{Y}$ est impossible et est de probabilité nulle, on a nécessairement $|x_n| \leq y_i$.

On prend ensuite une somme partielle. Pour tout N :

$$\sum_{n=0}^{\mathrm{N}} |x_n|\ \mathbf{P}(\mathrm{X} = x_n) \leq \sum_{n=0}^{\mathrm{N}} \sum_{i\in\mathrm{I}} y_i\ \mathbf{P}(\mathrm{X} = x_n, \mathrm{Y} = y_i) = \sum_{i\in\mathrm{I}} \sum_{n=0}^{\mathrm{N}} y_i\ \mathbf{P}(\mathrm{X} = x_n, \mathrm{Y} = y_i)$$

$$= \sum_{i\in\mathrm{I}} y_i\ \mathbf{P}(\mathrm{X} \in \{x_0, ..., x_\mathrm{N}\}, \mathrm{Y} = y_i\}$$

$$\leq \sum_{i\in\mathrm{I}} y_i\ \mathbf{P}(\mathrm{Y} = y_i\} = \mathbf{E}(\mathrm{Y})$$

Les sommes partielles $\sum_{n=0}^{\mathrm{N}} |x_n|\ \mathbf{P}(\mathrm{X} = x_n)$ formant une suite croissante majorée, elles convergent et X a une espérance finie.

PROPRIETES

Soient X *et* Y *deux variables aléatoires ayant une espérance finie, on a* :

(i) $\mathbf{E}(\lambda\mathrm{X}) = \lambda\mathbf{E}(\mathrm{X})$

(ii) $\mathbf{E}(\mathrm{X} + \mathrm{Y}) = \mathbf{E}(\mathrm{X}) + \mathbf{E}(\mathrm{Y})$

(iii) $\mathrm{X} \geq 0 \Rightarrow \mathbf{E}(\mathrm{X}) \geq 0$

(iv) $\mathrm{X} \leq \mathrm{Y} \Rightarrow \mathbf{E}(\mathrm{X}) \leq \mathbf{E}(\mathrm{Y})$

(v) *Si* X *et* Y *sont* ***indépendantes****, alors* $\mathbf{E}(\mathrm{XY}) = \mathbf{E}(\mathrm{X})\mathbf{E}(\mathrm{Y})$.

(vi) *Si* X *est à valeurs dans* $\mathbb{N}$, $\mathbf{E}(\mathrm{X})$ *est défini si et seulement si* $\sum \mathbf{P}(\mathrm{X} \geq k)$ *converge, et dans ce cas,* $\mathbf{E}(\mathrm{X}) = \sum_{k=1}^{\infty} \mathbf{P}(\mathrm{X} \geq k)$.

Démonstration :

On note x_n les valeurs prises par X et y_m celles prises par Y.

❑ (i) : Si $\lambda = 0$, l'égalité est trivialement vérifiée. Sinon, on applique le théorème de transfert avec $f(\mathrm{X}) = \lambda\mathrm{X}$. $\sum \lambda x_n\ \mathbf{P}(\mathrm{X} = x_n)$ est absolument convergente puisque X est d'espérance finie et :

$$\mathbf{E}(\lambda\mathrm{X}) = \mathbf{E}(f(\mathrm{X})) = \sum_{n=0}^{\infty} f(x_n)\mathbf{P}(\mathrm{X} = x_n) = \sum_{n=0}^{\infty} \lambda x_n\mathbf{P}(\mathrm{X} = x_n) = \lambda \sum_{n=0}^{\infty} x_n\mathbf{P}(\mathrm{X} = x_n) = \lambda\mathbf{E}(\mathrm{X})$$

❑ (ii) : On applique le théorème de transfert en prenant $f(X, Y) = X + Y$. La famille $|f(x_n, y_m)|)\, \mathbf{P}(X = x_n, Y = y_m)_{(n,m)\in\mathbf{N}^2}$ est sommable car, pour tout (n, m) :

$$\begin{aligned} |f(x_n, y_m)|)\, \mathbf{P}(X = x_n, Y = y_m) &\leq |x_n + y_m|\, \mathbf{P}(X = x_n, Y = y_m) \\ &\leq (|x_n| + |y_m|)\, \mathbf{P}(X = x_n, Y = y_m) \\ &\leq |x_n|\, \mathbf{P}(X = x_n, Y = y_m) + |y_m|\, \mathbf{P}(X = x_n, Y = y_m) \end{aligned}$$

et que :

$$\begin{aligned} \sum_{(n,m)\in\mathbf{N}^2} |x_n|\, \mathbf{P}(X = x_n, Y = y_m) &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} |x_n|\, \mathbf{P}(X = x_n, Y = y_m) \\ &= \sum_{n=0}^{\infty} |x_n| \left(\sum_{m=0}^{\infty} \mathbf{P}(X = x_n, Y = y_m)\right) \\ &= \sum_{n=0}^{\infty} |x_n|\, \mathbf{P}(X = x_n) \end{aligned}$$

et $\displaystyle\sum_{(n,m)\in\mathbf{N}^2} |y_m|\, \mathbf{P}(X = x_n, Y = y_m) = \sum_{m=0}^{\infty} |y_m|\, \mathbf{P}(Y = y_m)$

quantités finies puisque X et Y sont d'espérance finie. On peut alors effectuer le même calcul sans les valeurs absolues :

$$\begin{aligned} \mathbf{E}(X + Y) = \mathbf{E}(f(X, Y)) &= \sum_{(n,m)\in\mathbf{N}^2} f(x_n, y_m)\mathbf{P}(X = x_n, Y = y_m) \\ &= \sum_{(n,m)\in\mathbf{N}^2} (x_n + y_m)\, \mathbf{P}(X = x_n, Y = y_m) \\ &= \sum_{(n,m)\in\mathbf{N}^2} x_n\, \mathbf{P}(X = x_n, Y = y_m) + \sum_{(n,m)\in\mathbf{N}^2} y_m\, \mathbf{P}(X = x_n, Y = y_m) \\ &= \sum_{n=0}^{\infty} x_n\, \mathbf{P}(X = x_n) + \sum_{m=0}^{\infty} y_m\, \mathbf{P}(Y = y_m) \\ &= \mathbf{E}(X) + \mathbf{E}(Y) \end{aligned}$$

❑ (iii) : évident.

❑ (iv) : Si $X \leq Y$, alors $0 \leq \mathbf{E}(Y - X) = \mathbf{E}(Y) - \mathbf{E}(X)$.

❑ (v) : On applique le théorème de transfert à la fonction $f(X, Y) = XY$. La famille $|f(x_n, y_m)|)\, \mathbf{P}(X = x_n, Y = y_m)_{(n,m)\in\mathbf{N}^2}$ est sommable car, pour tout (n, m) :

$$|f(x_n, y_m)|)\, \mathbf{P}(X = x_n, Y = y_m) = |x_n y_m|\, \mathbf{P}(X = x_n, Y = y_m)$$

$= |x_n||y_m|\ \mathbf{P}(X = x_n)\ \mathbf{P}(Y = y_m)$ d'après l'indépendance de X et Y

que, pour tout n, $(|x_n||y_m|\ \mathbf{P}(X = x_n)\ \mathbf{P}(Y = y_m))_{m\in\mathbf{N}}$ est sommable de somme T_n avec :

$$T_n = \sum_{m=0}^{\infty} |x_n||y_m|\ \mathbf{P}(X = x_n)\ \mathbf{P}(Y = y_m) = |x_n|\ \mathbf{P}(X = x_n) \sum_{m=0}^{\infty} |y_m|\ \mathbf{P}(Y = y_m)$$

et que la famille $(T_n)_{n\in\mathbf{N}}$ est sommable de somme $\left(\sum_{n=0}^{\infty} |x_n|\ \mathbf{P}(X = x_n)\right) \left(\sum_{m=0}^{\infty} |y_m|\ \mathbf{P}(Y = y_m)\right)$.

On a alors :

$$\mathbf{E}(XY) = \mathbf{E}(f(X, Y)) = \sum_{n,m} f(x_n, y_m)\mathbf{P}(X = x_n, Y = y_m)$$

$$= \sum_{n,m} x_n y_m\ \mathbf{P}(X = x_n, Y = y_m)$$

$$= \sum_{n=0}^{\infty} x_n\ \mathbf{P}(X = x_n) \times \sum_{m=0}^{\infty} y_m\ \mathbf{P}(Y = y_m)$$

$$= \mathbf{E}(X)\mathbf{E}(Y)$$

❑ (vi) : Considérons une somme partielle de $\mathbf{E}(X) = \sum_{n=0}^{\infty} n\ \mathbf{P}(X = n)$ et transformons-la :

$$\sum_{k=0}^{n} k\ \mathbf{P}(X = k) = \sum_{k=0}^{n} k\ (\mathbf{P}(X \geq k) - \mathbf{P}(X \geq k + 1))$$

$$= \sum_{k=0}^{n} k\ \mathbf{P}(X \geq k) - \sum_{k=0}^{n} k\ \mathbf{P}(X \geq k + 1)$$

$$= \sum_{k=1}^{n} k\ \mathbf{P}(X \geq k) - \sum_{k=1}^{n+1} (k - 1)\ \mathbf{P}(X \geq k) \qquad \text{en changeant d'indice}$$

$$= \sum_{k=1}^{n} (k - (k - 1))\ \mathbf{P}(X \geq k) - n\ \mathbf{P}(X \geq n + 1)$$

$$= \sum_{k=1}^{n} \mathbf{P}(X \geq k) - n\ \mathbf{P}(X \geq n + 1)$$

La relation obtenue, écrite sous la forme $\sum_{k=1}^{n} \mathbf{P}(X \geq k) = \sum_{k=0}^{n} k\ \mathbf{P}(X = k) + n\ \mathbf{P}(X \geq n + 1)$ peut aussi se visualiser sous la disposition suivante : pour tout n, disposons la somme partielle $\sum_{k=1}^{n} \mathbf{P}(X \geq k)$ en colonne :

$$
\begin{aligned}
&\ \ \mathbf{P}(X \geq 1) \\
&+ \mathbf{P}(X \geq 2) \\
&+ \mathbf{P}(X \geq 3) \\
&+ \ldots \\
&+ \mathbf{P}(X \geq n) \\
&= \mathbf{P}(X = 1) + \mathbf{P}(X = 2) + \mathbf{P}(X = 3) + \ldots + \mathbf{P}(X = n) + \mathbf{P}(X \geq n+1) \\
&+ \qquad\qquad\quad\ \mathbf{P}(X = 2) + \mathbf{P}(X = 3) + \ldots + \mathbf{P}(X = n) + \mathbf{P}(X \geq n+1) \\
&+ \qquad\qquad\qquad\qquad\qquad\ \ \mathbf{P}(X = 3) + \ldots + \mathbf{P}(X = n) + \mathbf{P}(X \geq n+1) \\
&+ \ldots \\
&+ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \mathbf{P}(X = n) + \mathbf{P}(X \geq n+1)
\end{aligned}
$$

On voit donc que :

$$\sum_{k=1}^{n} \mathbf{P}(X \geq k) = \mathbf{P}(X = 1) + 2\mathbf{P}(X = 2) + 3\mathbf{P}(X = 3) + \ldots + n\mathbf{P}(X = n) + n\mathbf{P}(X \geq n+1)$$

$$= \sum_{k=1}^{n} k\mathbf{P}(X = k) + n\mathbf{P}(X \geq n+1)$$

On laisse au lecteur le soin de vérifier que cette égalité peut aussi se prouver par récurrence.

Supposons maintenant que $\mathbf{E}(X)$ soit défini, i.e. que $\sum n\,\mathbf{P}(X = n)$ converge, alors :

$$n\,\mathbf{P}(X \geq n+1) = n \sum_{k=n+1}^{\infty} \mathbf{P}(X = k) = \sum_{k=n+1}^{\infty} n\,\mathbf{P}(X = k)$$

$$\leq \sum_{k=n+1}^{\infty} k\,\mathbf{P}(X = k) \qquad \text{car } n \leq k$$

$$\leq \mathbf{E}(X) - \sum_{k=0}^{n} k\,\mathbf{P}(X = k) \qquad \text{quantité qui tend vers 0 quand } n \text{ tend vers } +\infty$$

puisqu'on reconnaît le reste de la série définissant $\mathbf{E}(X)$

Donc $\lim\limits_{n\to\infty} n\,\mathbf{P}(X \geq n+1) = 0$.

Par conséquent, en passant à la limite dans l'égalité $\sum\limits_{k=1}^{n} \mathbf{P}(X \geq k) = \sum\limits_{k=0}^{n} k\,\mathbf{P}(X = k) + n\,\mathbf{P}(X \geq n+1)$, on obtient :

$$\sum_{k=1}^{\infty} \mathbf{P}(X \geq k) = \mathbf{E}(X)$$

Réciproquement, si $\sum \mathbf{P}(X \geq k)$ converge, alors $\mathbf{E}(X)$ existe car la suite des sommes partielles $\sum\limits_{k=0}^{n} k\,\mathbf{P}(X = k)$ est croissante avec n et majorée par $\sum\limits_{k=1}^{\infty} \mathbf{P}(X \geq k)$. En effet :

$$\sum_{k=0}^{n} k\,\mathbf{P}(X = k) = \sum_{k=1}^{n} \mathbf{P}(X \geq k) - n\,\mathbf{P}(X \geq n+1)$$

$$\leq \sum_{k=1}^{n} \mathbf{P}(X \geq k)$$

$$\leq \sum_{k=1}^{\infty} \mathbf{P}(X \geq k)$$

On est donc ramené au cas précédent, donnant l'égalité entre $\mathbf{E}(X)$ et $\sum_{k=1}^{\infty} \mathbf{P}(X \geq k)$.

❑ (vi-bis) : Une démonstration plus simple, mais utilisant les séries doubles :

$$\mathbf{E}(X) = \sum_{n=1}^{\infty} n\, \mathbf{P}(X = n)$$

$$= \sum_{n=1}^{\infty} \sum_{k=1}^{n} \mathbf{P}(X = n)$$

$$= \sum_{1 \leq k \leq n} \mathbf{P}(X = n)$$

$$= \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} \mathbf{P}(X = n)$$

$$= \sum_{k=1}^{\infty} \mathbf{P}(X \geq k)$$

La notation $\sum_{1 \leq k \leq n} \mathbf{P}(X = n)$ est légitime car on somme une famille sommable (puisqu'elle est à termes positifs et que la quantité finie initiale $\sum_{n=1}^{\infty} n\, \mathbf{P}(X = n)$ est une façon de la sommer par paquets). Donc on peut calculer cette somme double dans l'ordre que l'on veut, ce qu'on fait dans la suite du calcul en permutant les indices de sommation.

❑ (vi-ter) : On peut enfin remarquer que $X = \sum_{k=1}^{\infty} \mathbf{1}_{X \geq k}$. En effet, si $X = i$, les seules valeurs de k pour lesquelles $\mathbf{1}_{X \geq k}$ est non nul sont $k = 1, 2, ..., i$, de sorte que $\sum_{k=1}^{\infty} \mathbf{1}_{X \geq k} = i$. Ainsi, X et $\sum_{k=1}^{\infty} \mathbf{1}_{X \geq k}$ prennent simultanément les mêmes valeurs. On a alors :

$$\mathbf{E}(X) = \mathbf{E}\left(\sum_{k=1}^{\infty} \mathbf{1}_{X \geq k}\right) \overset{?}{=} \sum_{k=1}^{\infty} \mathbf{E}(\mathbf{1}_{X \geq k}) = \sum_{k=1}^{\infty} \mathbf{P}(X \geq k)$$

Il y a une difficulté à justifier l'égalité $\overset{?}{=}$ portant sur une somme infinie. Elle repose sur une variante du théorème de sommation terme à terme d'une série de fonctions tel qu'on peut le voir dans le cas d'une intégrale dans le chapitre L2/SUITESF.PDF. Le résultat est analogue ici en remplaçant calcul intégral par calcul d'espérance et nous admettrons sa validité.

On a aussi :

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

en changeant d'indice.

EXEMPLE :

❑ Le vi) s'applique a fortiori pour une variable aléatoire sur un espace fini, à valeurs entières. On dispose de n boules dans une urne, numérotées de 1 à n. On effectue des tirages ***sans remise*** jusqu'à ce que la k-ème boule tirée ait un numéro inférieur à celui de la $(k-1)$-ème. On pose alors $X = k$. Si l'on tire toutes les boules, on pose $X = n$.

a) Pour $k \in [\![0, n-1]\!]$, que vaut $\mathbf{P}(X > k)$?

b) En déduire $\lim_{n\to\infty} \mathbf{E}(X)$, où $\mathbf{E}(X)$ est l'espérance de X.

a) Soit $k < n$, et $x_1, ..., x_k$ les k premiers numéros tirés. Ces k numéros sont rangés dans un ordre quelconque. Il y a $k!$ façons de les permuter et chaque permutation a la même probabilité d'arriver. Or dire que $X > k$, c'est dire que $x_1 < x_2 < ... < x_k$. Donc $\mathbf{P}(X > k) = \dfrac{1}{k!}$. On remarque que cette formule est également valide pour $k = 1$.

b) X étant à valeurs dans $[\![1, n]\!]$, on a la somme finie $\mathbf{E}(X) = \displaystyle\sum_{k=0}^{n-1} \mathbf{P}(X > k) = \sum_{k=0}^{n-1} \frac{1}{k!} \to \sum_{k=0}^{\infty} \frac{1}{k!} = e$

Les espérances des lois usuelles sont les suivantes :

❑ *Loi de Poisson* :

$$\mathbf{E}(X) = \sum_{k=0}^{\infty} k\, \mathbf{P}(X = k) = \sum_{k=1}^{\infty} \frac{k}{k!} \lambda^k e^{-\lambda} = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \lambda^k e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \lambda^{k-1}$$

On reconnaît dans la série le développement de e^{λ}. Donc :

$$\mathbf{E}(X) = \lambda$$

❑ *Loi géométrique* :

$$\mathbf{E}(X) = \sum_{k=1}^{\infty} k\, \mathbf{P}(X = k) = \sum_{k=1}^{\infty} k\, (1-p)^{k-1}\, p$$

On reconnaît une série du type $\sum kx^{k-1}$ de somme $\dfrac{1}{(1-x)^2}$. D'où :

$$\mathbf{E}(X) = \frac{1}{p}$$

EXEMPLES :

❑ Un joueur joue au loto une grille (5 numéros sur 49) par semaine jusqu'à ce qu'il gagne le gros lot. Quelle est la durée moyenne de jeu ? Le nombre de semaines est une variable aléatoire X de loi géométrique $p = \dfrac{1}{\binom{49}{5}} = \dfrac{1}{1906884}$. Son espérance est donc de 1906884 semaines, soit plus de 35000 ans.

7- Espérance conditionnelle

Comme dans le chapitre L1/PROBA1.PDF, on peut définir l'espérance conditionnelle $\mathbf{E}(X \mid B)$ d'une variable aléatoire X sachant qu'un événement B de probabilité non nulle est réalisé. Il s'agit de l'espérance de X lorsqu'on munit Ω de la probabilité conditionnelle $\mathbf{P}_B$.

On est amené à généraliser la proposition suivante, énoncée dans le cas fini dans le susdit chapitre.

THEOREME DE L'ESPERANCE TOTALE

Soit $(A_m)_{m\in\mathbb{N}}$ *un système complet d'événements d'un espace probabilisé* $(\Omega, \mathcal{T}, \mathbf{P})$, *et* X *une variable aléatoire discrète définie sur* Ω. *Alors* :

(i) X *est d'espérance finie pour la proba* $\mathbf{P}$ *si et seulement si, pour tout m,* X *est d'espérance finie pour la probabilité conditionnelle* $\mathbf{P}_{A_m}$ *et que la série* $\sum \mathbf{P}(A_m) \times \mathbf{E}(|X| \mid A_m)$ *est convergente.*

(ii) *Dans ce cas,* $\mathbf{E}(X) = \sum_{m=0}^{\infty} \mathbf{P}(A_m) \times \mathbf{E}(X \mid A_m)$

Si X ne prend qu'un nombre fini de valeurs, les conditions (i) sont automatiquement vérifiées et (ii) s'applique sans condition.

Si les A_m sont en nombre fini, la convergence de la série $\sum \mathbf{P}(A_m) \times \mathbf{E}(|X| \mid A_m)$ est inutile puisqu'il s'agit alors d'une somme finie. Seules sont à vérifier ou bien l'existence de l'espérance de X, ou bien celle de ses espérances conditionnelles.

<u>Démonstration</u> :

On note x_n, $n \in \mathbb{N}$, les valeurs prises par X.

❑ (i) : L'équivalence énoncée repose sur les constatations suivantes :

- L'égalité suivante :

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \mathbf{P}(A_m)\, |x_n|\, \mathbf{P}_{A_m}(X = x_n) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \mathbf{P}(A_m)\, |x_n|\, \mathbf{P}_{A_m}(X = x_n) \qquad (*)$$

  est valide dès qu'un des deux membres est une quantité finie, car dans ce cas, la famille $(\mathbf{P}(A_m)\, |x_n|\, \mathbf{P}_{A_m}(X = x_n))_{(m,n)\in\mathbb{N}^2}$ est sommable, et l'autre membre est également fini, la somme des termes pouvant s'effectuer dans l'ordre que l'on veut.

- Le membre de gauche vaut :

$$\sum_{n=0}^{\infty} |x_n| \Big(\sum_{m=0}^{\infty} \mathbf{P}(A_m)\, \mathbf{P}_{A_m}(X = x_n)\Big) = \sum_{n=0}^{\infty} |x_n|\, \mathbf{P}(X = x_n) = \mathbf{E}(|X|)$$

  et est défini si et seulement si X est d'espérance finie pour la probabilité $\mathbf{P}$.

- Le membre de droite vaut :

$$\sum_{m=0}^{\infty} \mathbf{P}(\mathrm{A}_m) \left(\sum_{n=0}^{\infty} |x_n| \, \mathbf{P}_{\mathrm{A}_m}(\mathrm{X} = x_n)\right) = \sum_{m=0}^{\infty} \mathbf{P}(\mathrm{A}_m) \, \mathbf{E}(|\mathrm{X}| \mid \mathrm{A}_m)$$

et est défini si et seulement si pour tout m, X est d'espérance finie pour la probabilité conditionnelle $\mathbf{P}_{\mathrm{A}_m}$ et que la série $\sum \mathbf{P}(\mathrm{A}_m) \, \mathbf{E}(|\mathrm{X}| \mid \mathrm{A}_m)$ converge.

❑ (ii) : Dans ce cas, la famille $(\mathbf{P}(\mathrm{A}_m) \, x_n \, \mathbf{P}_{\mathrm{A}_m}(\mathrm{X} = x_n))_{(m,n)\in\mathbf{N}^2}$ est sommable et les mêmes calculs sans valeur absolue donnent :

$$\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \mathbf{P}(\mathrm{A}_m) \, x_n \, \mathbf{P}_{\mathrm{A}_m}(\mathrm{X} = x_n) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \mathbf{P}(\mathrm{A}_m) \, x_n \, \mathbf{P}_{\mathrm{A}_m}(\mathrm{X} = x_n)$$

$$\Leftrightarrow \quad \sum_{n=0}^{\infty} x_n \left(\sum_{m=0}^{\infty} \mathbf{P}(\mathrm{A}_m) \, \mathbf{P}_{\mathrm{A}_m}(\mathrm{X} = x_n)\right) = \sum_{m=0}^{\infty} \mathbf{P}(\mathrm{A}_m) \left(\sum_{n=0}^{\infty} x_n \, \mathbf{P}_{\mathrm{A}_m}(\mathrm{X} = x_n)\right)$$

$$\Leftrightarrow \quad \sum_{n=0}^{\infty} x_n \, \mathbf{P}(\mathrm{X} = x_n) = \sum_{m=0}^{\infty} \mathbf{P}(\mathrm{A}_m) \, \mathbf{E}(\mathrm{X} \mid \mathrm{A}_m)$$

$$\Leftrightarrow \quad \mathbf{E}(\mathrm{X}) = \sum_{m=0}^{\infty} \mathbf{P}(\mathrm{A}_m) \, \mathbf{E}(\mathrm{X} \mid \mathrm{A}_m)$$

EXEMPLES :
❑ On tire une pièce équilibrée jusqu'à tomber sur Pile. Soit Z le nombre de tirages pour ce faire. Pour tout $m \geq 1$, on pose $\mathrm{A}_m = \{\mathrm{Z} = m\}$. Si A_m est réalisé, on tire la pièce une nouvelle fois et on pose $\mathrm{X} = m$ si on tombe sur Pile et $\mathrm{X} = 0$ sinon.

Pour tout $m \geq 1$, $\mathbf{P}(\mathrm{A}_m) = \dfrac{1}{2^m}$.

Pour tout $m \geq 1$, X admet une espérance finie conditionnellement en A_m et :

$$\mathbf{E}(\mathrm{X} \mid \mathrm{A}_m) = \mathbf{E}(|\mathrm{X}| \mid \mathrm{A}_m) = \frac{m}{2}$$

La série $\sum \mathbf{P}(\mathrm{A}_m) \times \mathbf{E}(|\mathrm{X}| \mid \mathrm{A}_m)$, égale à $\sum \dfrac{m}{2^{m+1}}$, est convergente.

Donc X est d'espérance finie et :

$$\mathbf{E}(\mathrm{X}) = \sum_{m=1}^{\infty} \frac{m}{2^{m+1}} = 1$$

(revoir au besoin le chapitre sur les séries entières L2/SERIENTR.PDF pour le calcul de cette somme).
Voici une simulation du modèle précédent, pour MAPLE. La fonction `rand() mod 2` simule un tirage à Pile ou Face avec une probabilité $\dfrac{1}{2}$ de prendre les valeurs 0 ou 1.

```
tirage:=proc()
local Z:
Z:=1:
while rand() mod 2 <> 1 do Z:=Z+1 od:
if rand() mod 2 = 1 then Z else 0 fi:
end:
```

```
N:=2000:
s:=0:
for i to N do s:=s+tirage() od:
s*1./N;
```

❑ On modifie l'exemple précédent en posant X = m si on tombe sur Pile, et X = – m sinon, A_m étant réalisé. Le même raisonnement s'applique, mais on a maintenant $\mathbf{E}(X \mid A_m) = 0$ pour tout m, d'où $\mathbf{E}(X) = 0$.

❑ On modifie l'exemple précédent en posant $X = 2^m$ si on tombe sur Pile, et $X = -2^m$ sinon, A_m étant réalisé. Pour tout m, $\mathbf{E}(X \mid A_m) = 0$. Cependant, il est erroné de conclure que $\mathbf{E}(X) = 0$. En effet, $\mathbf{E}(|X| \mid A_m) = 2^m$, donc $\sum \mathbf{P}(A_m) \times \mathbf{E}(|X| \mid A_m)$ diverge, donc X n'admet pas d'espérance finie. On peut d'ailleurs vérifier directement que, pour tout $m \geq 1$, $\mathbf{P}(X = 2^m) = \mathbf{P}(X = -2^m) = \dfrac{1}{2^{m+1}}$ et la série $\sum 2^m\, \mathbf{P}(X = 2^m) + \sum (-2^m\, \mathbf{P}(X = -2^m))$ n'est pas absolument convergente.

8- Variance et covariance

Si X^2 est d'espérance finie, alors il en est de même de X. En effet, $\sum x_n^2\, \mathbf{P}(X = x_n)$ est une série convergente, ainsi que $\sum \mathbf{P}(X = x_n)$. Or $|x_n|\, \mathbf{P}(X = x_n) \leq \dfrac{x_n^2\, \mathbf{P}(X = x_n) + \mathbf{P}(X = x_n)}{2}$ car cette inégalité se déduit de $|x_n| \leq \dfrac{x_n^2 + 1}{2}$ ou de $x_n^2 - 2|x_n| + 1 \geq 0$ ou de $(|x_n| - 1)^2 \geq 0$ qui est vraie.

Donc $\sum |x_n|\, \mathbf{P}(X = x_n)$ converge.

On définit la **variance** de X comme suit, pour une variable aléatoire telle que X^2 soit d'espérance finie :

DEFINITION DE LA VARIANCE

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

Son **écart-type** est défini par :

$$\sigma(X) = \sqrt{\mathbf{V}(X)}$$

On remarquera que $\mathbf{V}(aX + b) = \mathbf{V}(aX) = a^2\mathbf{V}(X)$

Les variances des lois usuelles sont les suivantes :

❑ *Loi de Poisson* :

Calculons $\mathbf{E}(X(X - 1))$, plus facile à calculer que $\mathbf{E}(X^2)$.

$$\mathbf{E}(X(X-1)) = \sum_{k=0}^{\infty} k(k-1)\, \mathbf{P}(X = k) = \sum_{k=2}^{\infty} \frac{k(k-1)}{k!} \lambda^k e^{-\lambda}$$

$$= \sum_{k=2}^{\infty} \frac{1}{(k-2)!} \lambda^k e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{1}{(k-2)!} \lambda^{k-2}$$

$$= \lambda^2$$

Donc $\mathbf{E}(X^2) = \mathbf{E}(X(X-1)) + \mathbf{E}(X) = \lambda^2 + \lambda$

Donc $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \lambda$

Ces valeurs sont à rapprocher de celles de la loi binomiale :

$\mathbf{E}(X) = np$

$\mathbf{V}(X) = np(1-p)$

En effet, lorsque n tend vers l'infini, p vers 0 de sorte que np tende vers λ, la loi binomiale converge vers la loi de Poisson. On remarque qu'il y a également convergence de l'espérance et de la variance de la loi binomiale vers l'espérance et la variance de la loi de Poisson :

$np \to \lambda$

$np(1-p) \to \lambda$

Une autre façon de calculer les espérances et les variances est donnée dans le paragraphe sur les fonctions génératrices, plus loin dans ce chapitre.

❑ *Loi géométrique* :
Calculons également $\mathbf{E}(X(X-1))$, plus facile à calculer que $\mathbf{E}(X^2)$:

$$\mathbf{E}(X(X-1)) = \sum_{k=1}^{\infty} k(k-1)\,\mathbf{P}(X=k) = \sum_{k=1}^{\infty} k(k-1)\,(1-p)^{k-1}\,p$$

$$= p(1-p)\sum_{k=2}^{\infty} k(k-1)\,(1-p)^{k-2}$$

On reconnaît une série du type $\sum k(k-1)\,x^{k-2}$ de somme $\dfrac{2}{(1-x)^3}$. D'où :

$$\mathbf{E}(X(X-1)) = \frac{2(1-p)}{p^2}$$

$$\Rightarrow \qquad \mathbf{E}(X^2) = \frac{2(1-p)}{p^2} + \frac{1}{p} = \frac{2-p}{p^2}$$

Donc $\mathbf{V}(X) = \dfrac{1-p}{p^2}$

EXEMPLES :

❑ Reprenons l'exemple du loueur de loto ayant une probabilité $p = \dfrac{1}{1906884}$ de gagner le gros lot à chaque tirage. La variable aléatoire X égale au nombre de tirages jusqu'au premier gain du gros lot suit une loi géométrique de paramètre p. Son espérance est $\mathbf{E}(X) = \dfrac{1}{p} = 1906884$, sa variance est $\dfrac{1-p}{p^2}$ et son écart-type est $\sigma(X) = \dfrac{\sqrt{1-p}}{p}$, qui est quasiment égal à $\dfrac{1}{p}$. L'écart-type est une mesure de la dispersion de la variable aléatoire X autour de son espérance. La dispersion est ici importante, puisque $[\mathbf{E}(X) - \sigma(X), \mathbf{E}(X) + \sigma(X)]$ vaut $[0, 3813768]$.

❑ Dans le domaine électromagnétique, une particule chargée soumise à un champ de fréquence ν acquiert de l'énergie. En physique quantique, on suppose que l'énergie acquise E est aléatoire, mais quantifiée, i.e. elle ne peut prendre que des valeurs de la forme $nh\nu$, $n \in \mathbb{N}$, h étant une constante appelée constante de Planck. La probabilité d'acquérir l'énergie $E_n = nh\nu$ suit une distribution dite de

Boltzmann : à la température T, cette probabilité est proportionnelle à $\exp(-\frac{E_n}{kT})$, où k est une constante dite constante de Boltzmann. Pour que la somme des probabilités vaillent 1, le facteur de proportionnalité est nécessairement $\dfrac{1}{\sum_{n=0}^{\infty} \exp(-\frac{E_n}{kT})}$. Or :

$$\sum_{n=0}^{\infty} \exp(-\frac{E_n}{kT}) = \sum_{n=0}^{\infty} \exp(-\frac{nh\nu}{kT}) \qquad \text{somme d'une série géométrique}$$

$$= \frac{1}{1 - \exp(-\frac{h\nu}{kT})}$$

donc $$\frac{1}{\sum_{n=0}^{\infty} \exp(-\frac{E_n}{kT})} = 1 - \exp(-\frac{h\nu}{kT})$$

La loi de probabilité de l'énergie E est donc donnée par :

$$\forall n \in \mathbf{N},\ \mathbf{P}(E = E_n) = (1 - \exp(-\frac{h\nu}{kT}))\exp(-\frac{E_n}{kT})$$

Si on pose $p = 1 - \exp(-\frac{h\nu}{kT})$, alors $1 - p = \exp(-\frac{h\nu}{kT})$ et l'on a, en remplaçant E_n par sa valeur $nh\nu$:

$$\forall n \in \mathbf{N},\ \mathbf{P}(E = nh\nu) = p(1-p)^n$$

Dans le membre de droite, on reconnaît une variante de la loi géométrique, au détail près que n varie dans $\mathbf{N}$ et non dans $\mathbf{N}^*$. On retrouve une variable aléatoire X suivant une loi géométrique usuelle en posant $X = \frac{E}{h\nu} + 1$, de sorte que :

$$\forall n \geq 1,\ \mathbf{P}(X = n) = \mathbf{P}(E = (n-1)h\nu) = p(1-p)^{n-1}$$

D'après le calcul de l'espérance d'une loi géométrique, on a, en notant ici l'espérance entre crochet pour ne pas abuser de la lettre E :

$$\langle E \rangle = \langle h\nu X - h\nu \rangle = h\nu(\langle X \rangle - 1)$$

$$= h\nu(\frac{1}{p} - 1)$$

$$= h\nu \frac{1-p}{p}$$

$$= h\nu \frac{\exp(-\frac{h\nu}{kT})}{1 - \exp(-\frac{h\nu}{kT})}$$

$$= \frac{h\nu}{\exp(\frac{h\nu}{kT}) - 1} \qquad \text{valeur moyenne de l'énergie pour notre particule.}$$

Le résultat obtenu $\dfrac{h\nu}{\exp(\frac{h\nu}{kT}) - 1}$ est dû à Planck. On retrouve l'énergie kT d'un oscillateur élémentaire en physique classique, et plus spécialement en thermodynamique quand ν tend vers 0 ou T vers l'infini.

La variance de X est $\mathbf{V}(X) = \dfrac{1-p}{p^2} = \dfrac{\exp(-\frac{h\nu}{kT})}{(1-\exp(-\frac{h\nu}{kT}))^2} = \dfrac{\exp(\frac{h\nu}{kT})}{(\exp(\frac{h\nu}{kT})-1)^2}$.

Son écart-type vaut $\sigma(X) = \sqrt{\mathbf{V}(X)} = \dfrac{\exp(\frac{h\nu}{2kT})}{\exp(\frac{h\nu}{kT})-1}$ et l'écart-type de l'énergie est $h\nu \dfrac{\exp(\frac{h\nu}{2kT})}{\exp(\frac{h\nu}{kT})-1}$. Cet écart-type mesure la dispersion de l'énergie autour de l'énergie moyenne.
Si $h\nu$ est petit devant kT, elle est de l'ordre de kT.
Si $h\nu$ est grand devant kT, l'écart-type est de l'ordre de $h\nu \exp(-\frac{h\nu}{2kT})$ bien plus grand que l'énergie moyenne $h\nu \exp(-\frac{h\nu}{kT})$. A haute fréquence, les fluctuations relatives sont plus importantes qu'à basse fréquence.

On s'intéresse maintenant à la variance d'un somme :

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

Toutes les quantités intervenant ci-dessus sont définies si X et Y sont de variance finie. En effet, on sait déjà que, dans ce cas, X et Y sont d'espérance finie. Reste à montrer que $(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))$ est d'espérance finie. Posons $X' = X - \mathbf{E}(X)$ et $Y' = Y - \mathbf{E}(Y)$. Si x_i' sont les valeurs prises par X' et y_j' celles prises par Y', on a :

$$|x_i'\,y_j'|\;\mathbf{P}(X' = x_i', Y' = y_j') \leq \frac{x_i'^2 + y_j'^2}{2}\,\mathbf{P}(X' = x_i', Y' = y_j')$$

Mais $\sum_{i,j} \dfrac{x_i'^2 + y_j'^2}{2}\,\mathbf{P}(X' = x_i', Y' = y_j')$ converge. Elle vaut $\dfrac{\mathbf{E}(X'^2) + \mathbf{E}(Y'^2)}{2} = \dfrac{\mathbf{V}(X') + \mathbf{V}(Y')}{2}$

On pose alors la **covariance** de X et Y comme suit :

DEFINITION DE LA COVARIANCE

$$\text{cov}(X, Y) = \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))$$

D'où :

$$\mathbf{V}(X + Y) = \mathbf{V}(X) + \mathbf{V}(Y) + 2\,\text{cov}(X,Y)$$

La formule ci-dessus se généralise à n variables de la façon suivante :

$$\mathbf{V}(X_1 + ... + X_n) = \mathbf{V}(X_1) + ... + \mathbf{V}(X_n) + 2 \sum_{i,j} \text{cov}(X_i, X_j)$$

PROPOSITION

Variances et covariances vérifient les propriétés suivantes :

(i) $\text{cov}(X, X) = \mathbf{V}(X)$

(ii) $\text{cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$

(iii) $(X, Y) \to \text{cov}(X, Y)$ *est bilinéaire*

(iv) $|\text{cov}(X,Y)| \leq \sqrt{\mathbf{V}(X)}\,\sqrt{\mathbf{V}(Y)}$

(v) X *et* Y *indépendantes* $\Rightarrow \text{cov}(X, Y) = 0$

(vi) X *et* Y *indépendantes* $\Rightarrow \mathbf{V}(X + Y) = \mathbf{V}(X) + \mathbf{V}(Y)$

Démonstration :

❑ (i) est évident.

❑ (ii) se montre en développant cov(X,Y) :

$$\begin{aligned}
\text{cov}(X, Y) &= \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y))) \\
&= \mathbf{E}(XY - X\mathbf{E}(Y) - Y\mathbf{E}(X) + \mathbf{E}(X)\mathbf{E}(Y)) \\
&= \mathbf{E}(XY) - \mathbf{E}(X\mathbf{E}(Y)) - \mathbf{E}(Y\mathbf{E}(X)) + \mathbf{E}(\mathbf{E}(X)\mathbf{E}(Y)) \\
&= \mathbf{E}(XY) - \mathbf{E}(Y)\mathbf{E}(X) - \mathbf{E}(X)\mathbf{E}(Y) + \mathbf{E}(X)\mathbf{E}(Y) \\
&= \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)
\end{aligned}$$

❑ (iii) : La vérification est laissée au lecteur.

❑ (iv) résulte du fait qu'une forme bilinéaire symétrique positive vérifie l'inégalité de Cauchy-Schwarz (revoir la démonstration de cette inégalité dans le chapitre L1/ESPEUCL.PDF, en remplaçant le produit scalaire par la covariance et la norme euclidienne par la variance).

❑ v) : Si X et Y sont indépendantes, alors $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$. La propriété v) découle alors du (ii).

❑ vi) est clairement équivalente à v), compte tenu de la définition de la covariance.

On définit le **coefficient de corrélation** de X et Y comme étant égal à $\rho(X, Y) = \dfrac{\text{cov}(X,Y)}{\sqrt{\mathbf{V}(X)}\sqrt{\mathbf{V}(Y)}}$. Ce nombre est compris entre –1 et 1, d'après la relation iv). Il joue un rôle analogue à celui du cosinus pour deux vecteurs.

Si X et Y sont indépendantes, alors le coefficient de corrélation est nul. On dit que les variables sont **non corrélées**. On prendra garde que cette condition n'est pas équivalente à l'indépendance de X et Y (pas plus que le fait que $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$ n'entraîne l'indépendance de X et de Y).

Si $Y = aX + b$, alors $\text{cov}(X,Y) = a\mathbf{V}(X)$ et le coefficient de corrélation vaut 1 si $a > 0$ et -1 si $a < 0$.

EXEMPLE :

❑ Soient X, Y et Z trois variables aléatoires indépendantes de même loi de Poisson $\mathcal{P}(\lambda)$. Cherchons le coefficient de corrélation de X + Y et Y + Z :

$\mathbf{E}(X + Y) = \mathbf{E}(X) + \mathbf{E}(Y) = 2\lambda$

$\mathbf{E}(X + Z) = \mathbf{E}(X) + \mathbf{E}(Z) = 2\lambda$

donc $\mathbf{E}(X + Y)\mathbf{E}(Y + Z) = 4\lambda^2$

Puis $\mathbf{E}((X + Y)(Y + Z)) = \mathbf{E}(XY + XZ + Y^2 + YZ)$

$= \mathbf{E}(XY + XZ + Y^2 + YZ)$

$= \mathbf{E}(X)\mathbf{E}(Y) + \mathbf{E}(X)\mathbf{E}(Z) + \mathbf{E}(Y^2) + \mathbf{E}(Y)\mathbf{E}(Z)$

en utilisant l'indépendance des variables aléatoires

$= 4\lambda^2 + \lambda$

donc $\mathrm{cov}(X + Y, Y + Z) = \mathbf{E}((X + Y)(Y + Z)) - \mathbf{E}(X + Y)\mathbf{E}(Y + Z) = 4\lambda^2 + \lambda - 4\lambda^2 = \lambda$

Enfin $\mathbf{V}(X + Y) = \mathbf{V}(X) + \mathbf{V}(Y)$ car X et Y sont indépendantes

$= 2\lambda = \mathbf{V}(Y + Z)$

donc $\rho(X + Y, Y + Z) = \dfrac{\mathrm{cov}(X + Y, Y + Z)}{\sqrt{\mathbf{V}(X + Y)}\sqrt{\mathbf{V}(Y + Z)}} = \dfrac{1}{2}$

INEGALITE DE BIENAYME-TCHEBYCHEV

Soit X *une variable aléatoire telle que* X^2 *a une espérance finie. Soit* $\mu = \mathbf{E}(X)$ *et* $\sigma^2 = \mathbf{V}(X)$. *Alors, pour tout* $\varepsilon > 0$:

$$\mathbf{P}(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

<u>Démonstration</u> :

❑ Posons $Y = |X - \mu|$. On a $\sigma^2 = \mathbf{V}(X) = \mathbf{E}(Y^2)$. Notons $\mathbf{1}_{Y\geq\varepsilon}$ la fonction indicatrice de l'événement $\{Y \geq \varepsilon\} = \{\omega \in \Omega, Y(\omega) \geq \varepsilon\}$. Alors, on a :

$$Y^2 \geq \varepsilon^2\, \mathbf{1}_{Y\geq\varepsilon}$$

En effet :

ou bien $Y(\omega) \geq \varepsilon$ et l'inégalité ci-dessus donne bien le même résultat, puisqu'alors $\mathbf{1}_{Y\geq\varepsilon}(\omega) = 1$;

ou bien $Y(\omega) < \varepsilon$, est l'inégalité ci-dessus donne $Y(\omega)^2 \geq 0$, puisque $\mathbf{1}_{Y\geq\varepsilon}(\omega) = 0$. Dans tous les cas, l'inégalité est vérifiée. On en déduit que :

$$\sigma^2 = \mathbf{E}(Y^2) \geq \mathbf{E}(\varepsilon^2\, \mathbf{1}_{Y\geq\varepsilon}) = \varepsilon^2\, \mathbf{E}(\mathbf{1}_{Y\geq\varepsilon}) = \varepsilon^2\, \mathbf{P}(Y \geq \varepsilon)$$

D'où le résultat.

Notons que l'argument utilisé sur la variable aléatoire Y^2 se généralise à n'importe quelle variable aléatoire Z positive ou nulle sous la forme :

$\forall\, a > 0,\ \mathbf{E}(Z) \geq a\, \mathbf{P}(Z \geq a)$ (**inégalité de Markov**)

en remarquant que $Z \geq a\, \mathbf{1}_{Z\geq a}$ et en prenant l'espérance des deux membres.

EXEMPLE :

❑ Soit X une variable aléatoire telle que $\mathbf{V}(X) = 0$. Alors X est presque sûrement constante. En effet, l'inégalité de Bienaymé-Tchebychev donne, pour tout entier n :

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq \frac{1}{n}) \leq \mathbf{V}(X)n^2 = 0$$

Les événements $|X - \mathbf{E}(X)| \geq \frac{1}{n}$ sont croissants, de réunion $|X - \mathbf{E}(X)| > 0$. On a donc :

$$\mathbf{P}(|X - \mathbf{E}(X)| > 0) = \lim_{n\to\infty} \mathbf{P}(|X - \mathbf{E}(X)| \geq \frac{1}{n}) = 0$$

Donc $\mathbf{P}(|X - \mathbf{E}(X)| = 0) = 1$ et X est presque sûrement égale à la constante $\mathbf{E}(X)$.

LOI FAIBLE DES GRANDS NOMBRES

Soit (X_n) *une suite de variables aléatoires indépendantes de même loi, d'espérance* μ *et d'écart-type* σ*. Posons* $S_n = X_1 + ... + X_n$*. Alors, pour tout* $\varepsilon > 0$:

$$\lim_{n\to\infty} \mathbf{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) = 0$$

<u>Démonstration</u> :

❑ Appliquons l'inégalité de Bienaymé-Tchebychev à $Z_n = \dfrac{S_n}{n} = \dfrac{X_1 + ... + X_n}{n}$. On a $\mathbf{E}(Z_n) = \mu$ et :

$$\mathbf{V}(Z_n) = \frac{1}{n^2}\mathbf{V}(X_1 + ... + X_n) = \frac{1}{n^2} \times (\mathbf{V}(X_1) + ... + \mathbf{V}(X_n)) \text{ car les } X_i \text{ sont indépendantes}$$

$$= \frac{\sigma^2}{n}$$

Donc :

$$\mathbf{P}(|Z_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$$

quantité qui tend vers 0 quand n tend vers l'infini. Cela exprime que la probabilité que Z_n appartienne à l'intervalle $]\mu - \varepsilon, \mu + \varepsilon[$ tend vers 1 quand n tend vers l'infini, et ceci, quel que soit le nombre ε choisi.

Donnons également un exemple montrant que l'inégalité de Bienaymé-Tchebychev peut également s'appliquer à des domaines a priori sans lien avec les probabilités.

EXEMPLE :

❑ Prouver que $\displaystyle\sum_{k=1}^{\infty} \frac{x^k}{k!\sqrt{k}} \sim \frac{e^x}{\sqrt{x}}$ quand x tend vers $+\infty$.

C'est équivalent à montrer que, si on pose $S(x) = e^{-x} \displaystyle\sum_{k=1}^{\infty} \frac{\sqrt{x}}{\sqrt{k}} \frac{x^k}{k!}$, on a $\displaystyle\lim_{x\to+\infty} S(x) = 1$. Considérons une variable aléatoire X suivant une loi de Poisson de paramètre x. On constate que :

$$S(x) = \sum_{k=1}^{\infty} \frac{\sqrt{x}}{\sqrt{k}} \mathbf{P}(X = k)$$

donc $$S(x) - 1 = \sum_{k=1}^{\infty} \frac{\sqrt{x}}{\sqrt{k}} \mathbf{P}(X = k) - 1$$

$$= \sum_{k=1}^{\infty} \frac{\sqrt{x}}{\sqrt{k}} \mathbf{P}(X = k) - \sum_{k=0}^{\infty} \mathbf{P}(x = k)$$

$$= \sum_{k=1}^{\infty} \left(\frac{\sqrt{x}}{\sqrt{k}} - 1\right) \mathbf{P}(X = k) - \mathbf{P}(X = 0)$$

$$= \sum_{k=1}^{\infty} \left(\frac{\sqrt{x}}{\sqrt{k}} - 1\right) \mathbf{P}(X = k) - e^{-x}$$

Comme $\lim\limits_{x\to+\infty} e^{-x} = 0$, il suffit de montrer que $\lim\limits_{x\to\infty} \sum\limits_{k=1}^{\infty} \left(\frac{\sqrt{x}}{\sqrt{k}} - 1\right) \mathbf{P}(X = k) = 0$. Pour cela, écrivons l'inégalité de Bienaymé-Tchebychev. Pour tout a strictement positif, on a :

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq a) \leq \frac{\mathbf{V}(X)}{a^2}$$

donc $\mathbf{P}(|X - x| \geq a) \leq \frac{x}{a^2}$ puisque $\mathbf{E}(X) = x$ et $\mathbf{V}(x) = x$.

Dans la suite, on prendra $a = x^{7/8}$, un peu inférieur à x, tout en étant négligeable devant x quand x tend vers $+\infty$. En notant $m = \lfloor x - a \rfloor$ et $n = \lfloor x + a \rfloor$, on coupe la valeur absolue de la somme considérée en trois morceaux dont on va montrer que chacun tend vers 0 :

$$\left| \sum_{k=1}^{\infty} \left(\frac{\sqrt{x}}{\sqrt{k}} - 1\right) \mathbf{P}(X = k) \right| \leq \sum_{k=1}^{m} \left| \frac{\sqrt{x}}{\sqrt{k}} - 1 \right| \mathbf{P}(X = k) + \sum_{k=m+1}^{n} \left| \frac{\sqrt{x}}{\sqrt{k}} - 1 \right| \mathbf{P}(X = k) + \sum_{k=n+1}^{\infty} \left| \frac{\sqrt{x}}{\sqrt{k}} - 1 \right| \mathbf{P}(X = k)$$

Premier terme : Pour $1 \leq k \leq m \leq x - a \leq x$, on a $\left| \frac{\sqrt{x}}{\sqrt{k}} - 1 \right| = \frac{\sqrt{x}}{\sqrt{k}} - 1 \leq \sqrt{x} - 1 \leq \sqrt{x}$. Donc :

$$\begin{aligned}
\sum_{k=1}^{m} \left| \frac{\sqrt{x}}{\sqrt{k}} - 1 \right| \mathbf{P}(X = k) &\leq \sum_{k=1}^{m} \sqrt{x}\, \mathbf{P}(X = k) \\
&\leq \sqrt{x}\, \mathbf{P}(X \leq m) \\
&\leq \sqrt{x}\, \mathbf{P}(X \leq x - a) && \text{car } \{X \leq m\} \subset \{X \leq x - a\} \\
&\leq \sqrt{x}\, \mathbf{P}(|X - x| \geq a) && \text{car} \{X \leq x - a\} \subset \{|X - x| \geq a\} \\
&\leq \frac{x\sqrt{x}}{a^2} && \text{d'après l'inégalité de Bienaymé-Tchebychev} \\
&\leq \frac{1}{x^{1/4}} && \to 0 \text{ quand } x \text{ tend vers } +\infty
\end{aligned}$$

Troisième terme : de même, pour $k \geq n + 1 \geq x + a \geq x$, on a $\left| \frac{\sqrt{x}}{\sqrt{k}} - 1 \right| = 1 - \frac{\sqrt{x}}{\sqrt{k}} \leq 1$. Donc :

$$\begin{aligned}
\sum_{k=n+1}^{\infty} \left| \frac{\sqrt{x}}{\sqrt{k}} - 1 \right| \mathbf{P}(X = k) &\leq \sum_{k=n+1}^{\infty} \mathbf{P}(X = k) \\
&\leq \mathbf{P}(X \geq n + 1) \\
&\leq \mathbf{P}(X \geq x + a) \\
&\leq \mathbf{P}(|X - x| \geq a) \\
&\leq \frac{x}{a^2} \\
&\leq \frac{1}{x^{3/4}} && \to 0 \text{ quand } x \text{ tend vers } +\infty
\end{aligned}$$

Deuxième terme :

$$\sum_{k=m+1}^{n} \left| \frac{\sqrt{x}}{\sqrt{k}} - 1 \right| \mathbf{P}(X = k) \leq \operatorname*{Sup}_{k \in [\![m,n+1]\!]} \left| \frac{\sqrt{x}}{\sqrt{k}} - 1 \right| \sum_{k=m+1}^{n} \mathbf{P}(X = k)$$
$$\leq \operatorname*{Sup}_{k \in [\![m,n+1]\!]} \left| \frac{\sqrt{x}}{\sqrt{k}} - 1 \right|$$

Comme la suite $k \to \dfrac{\sqrt{x}}{\sqrt{k}} - 1$ est décroissante, positive pour $k = m$ puis négative pour $k = n + 1$, la suite $k \to \left| \dfrac{\sqrt{x}}{\sqrt{k}} - 1 \right|$ décroît, passe par un minimum, puis croît lorsque k varie de m à $n + 1$, de sorte que sa borne supérieure n'est autre que Max $\{\dfrac{\sqrt{x}}{\sqrt{m}} - 1, 1 - \dfrac{\sqrt{x}}{\sqrt{n+1}}\}$. Or :

$$0 \leq \frac{\sqrt{x}}{\sqrt{m}} - 1 \leq \frac{\sqrt{x}}{\sqrt{x-a-1}} - 1 = \frac{1}{\sqrt{x-a-1}} (\sqrt{x} - \sqrt{x-a-1})$$
$$\leq \frac{1}{\sqrt{x-a-1}} \frac{a+1}{\sqrt{x} + \sqrt{x-a-1}} \sim \frac{a}{2x} = \frac{1}{2x^{1/4}} \to 0$$

et de même pour $1 - \dfrac{\sqrt{x}}{\sqrt{n+1}}$. Donc le Max des deux termes tend aussi vers 0.

III : Fonctions génératrices

1- Définition

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. On cherche à définir une fonction dont la seule connaissance suffise à retrouver la loi. La fonction la plus naturelle est la série entière sur $\mathbb{C}$ définie par :

$$G_X(z) = \sum_{n=0}^{\infty} \mathbf{P}(X = n)\, z^n$$

puisque les coefficients d'une série entière donnée sont uniques. Cette fonction s'appelle **fonction génératrice** de X. Elle n'est autre que $\mathbf{E}(z^X)$, en vertu de la formule générale :

$$\mathbf{E}(f(X)) = \sum_{n=0}^{\infty} \mathbf{P}(X = n)\, f(n)$$

On prend ici $f(X) = z^X$. Ainsi :

$$\mathbf{E}(z^X) = \sum_{n=0}^{\infty} \mathbf{P}(X = n)\, z^n = G_X(z)$$

PROPRIETES DE LA FONCTION GENERATRICE

(i) *Le rayon* R *de convergence de la série entière est supérieur ou égal à* 1
(ii) G_X *est dérivable en* 1 *si et seulement si* X *admet une espérance, et dans ce cas :*

$$\mathbf{E}(X) = G_X'(1)$$

(iii) G_X *est deux fois dérivable en* 1 *si et seulement si* X *admet une variance, et dans ce cas :*

$$\mathbf{V}(X) = G_X''(1) + G_X'(1) - G_X'(1)^2$$

Démonstration :

❑ (i) : Pour $z = 1$, on a $G_X(1) = \sum_{n=0}^{\infty} \mathbf{P}(X = n) = 1$. Puisque la série entière définissant G_X converge en 1, le rayon de convergence R est supérieur ou égal à 1.

❑ (ii) : Si $R > 1$, G_X est indéfiniment dérivable sur $]-R, R[$, et l'on a, pour tout t élément de $]-R, R[$ (et donc en 1) :

$$G_X'(t) = \sum_{n=1}^{\infty} \mathbf{P}(X = n)\, nt^{n-1} = \mathbf{E}(X\, t^{X-1})$$

donc $G_X'(1) = \mathbf{E}(X)$

Le cas R = 1 est plus délicat :

Si X admet une espérance, le calcul précédent reste valide car $\sum n\, \mathbf{P}(X = n)$ converge, donc la série de fonctions $\sum n\, \mathbf{P}(X = n)\, t^{n-1}$ converge normalement sur [0, 1], donc G_X est C^1 sur [0, 1] et sa dérivée se calcule terme à terme (voir la notion de convergence normale d'une série de fonctions dans le chapitre L2/SUITESF.PDF).

Réciproquement, supposons que $G_X'(1)$ est défini. G_X étant une série entière de rayon 1, on sait que, pour t élément de $]-1, 1[$, $G_X'(t) = \sum_{n=1}^{\infty} n\, \mathbf{P}(X = n)\, t^{n-1}$ mais on ignore a priori si cette relation reste vraie en 1. On note cependant que G_X est continue sur [0, 1], et G_X' est croissante sur [0, 1[, donc admet une limite en 1, finie ou non. Mais si une fonction est C^1 sur un intervalle $[a, b[$, continue sur $[a, b]$ et si sa dérivée admet une limite en b finie ou non, ou bien cette limite est infinie et la fonction n'est pas dérivable en b, ou bien cette limite est finie en b et la fonction est C^1 sur $[a, b]$ (voir les conséquences du théorème de Rolle dans le chapitre L1/DERIVEE.PDF). Comme on a supposé que G_X est dérivable en 1, c'est le second cas qui s'applique et on a $\lim_{t\to 1,\, t<1} G_X'(t) = G_X'(1)$.

Or, dans le chapitre sur les séries entières L2/SERIENTR.PDF (paragraphe *comportement sur le cercle de convergence*), on a montré que, si une série entière $\sum a_n t^n$ à termes $a_n \geq 0$ et de rayon de convergence 1 est telle que sa somme $t \in]-1, 1[\rightarrow f(t) = \sum_{n=0}^{\infty} a_n t^n$ admet une limite en 1, alors $\sum a_n$ converge et $\sum_{n=0}^{\infty} a_n$ est égale à cette limite. Appliquée à $a_n = n\, \mathbf{P}(X = n)$ et $f(t) = G_X'(t)$, cette proposition permet de conclure que $\sum n\, \mathbf{P}(X = n)$ converge et que $G_X'(1) = \sum_{n=0}^{\infty} n\, \mathbf{P}(X = n) = \mathbf{E}(X)$.

❑ (iii) : La dérivation terme à terme de G_X' pour obtenir G_X'' est valide pour $R \geq 1$ en tenant un raisonnement comparable au précédent. On a alors :

$G_X''(1)$ existe

$\Leftrightarrow$ $\sum n(n-1)\, \mathbf{P}(X = n)$ converge

$\Leftrightarrow$ $\sum n^2\, \mathbf{P}(X = n)$ converge, car les termes généraux des deux séries sont de même signe et équivalents quand n tend vers l'infini

$\Leftrightarrow$ $\mathbf{E}(X^2)$ existe (et donc $\mathbf{E}(X)$ aussi)

$\Leftrightarrow$ $\mathbf{V}(X)$ existe

et l'on a alors :

$$G_X''(z) = \sum_{n=2}^{\infty} \mathbf{P}(X = n)\, n(n-1)z^{n-2} = \mathbf{E}(X(X-1)\, z^{X-2})$$

donc $G_X''(1) = \mathbf{E}(X(X-1))$

et $\mathbf{E}(X^2) = G_X''(1) + \mathbf{E}(X) = G_X''(1) + G_X'(1)$

donc $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = G_X''(1) + G_X'(1) - G_X'(1)^2$

2- Lois usuelles

Calculons la fonction génératrice des lois usuelles, et retrouvons les expressions des espérances et variances de ces lois à partir de G_X.

❑ *Loi de Bernoulli* :

$G_X(z) = 1 - p + pz$

$G_X'(z) = p$ donc $\mathbf{E}(X) = p$

$G_X''(z) = 0$ donc $\mathbf{V}(X) = p(1-p)$

❑ *Loi binomiale* $\mathcal{B}(n, p)$:

$$G_X(z) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} z^k = (pz + 1 - p)^n$$

$G_X'(z) = np(pz + 1 - p)^{n-1}$ donc $\mathbf{E}(X) = np$

$G_X''(z) = n(n-1)p^2(pz + 1 - p)^{n-2}$ donc $\mathbf{V}(X) = n(n-1)p^2 + np - n^2p^2 = np(1-p)$

❑ *Loi de Poisson* $\mathcal{P}(\lambda)$:

$$G_X(z) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} z^n = \exp(\lambda z - \lambda)$$

$G_X'(z) = \lambda\exp(\lambda z - \lambda)$ donc $\mathbf{E}(X) = \lambda$

$G_X''(z) = \lambda^2\exp(\lambda z - \lambda)$ donc $\mathbf{V}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$

❑ *Loi géométrique sur* $\mathbb{N}^*$:

$$G_X(z) = \sum_{n=1}^{\infty} (1-p)^{n-1}\, p\, z^n = \frac{pz}{1-(1-p)z}$$

Cette série n'est définie que pour $|z| < \dfrac{1}{1-p}$.

$G_X'(z) = \dfrac{p}{(1-(1-p)z)^2}$ donc $\mathbf{E}(X) = \dfrac{1}{p}$

$G_X''(z) = \dfrac{2p(1-p)}{(1-(1-p)z)^3}$ donc $\mathbf{V}(X) = \dfrac{2(1-p)}{p^2} + \dfrac{1}{p} - \dfrac{1}{p^2} = \dfrac{1-p}{p^2}$

3- Fonction génératrice d'une somme de variables aléatoires indépendantes

PROPOSITION

Soient X *et* Y *deux variables aléatoires indépendantes à valeurs dans* $\mathbb{N}$*, de fonctions génératrices* G_X *et* G_Y*. Alors la fonction génératrice* G_{X+Y} *de* X + Y *est le produit de* G_X *par* G_Y.

Démonstration :

❑ $$G_{X+Y}(z) = \sum_{n=0}^{\infty} \mathbf{P}(X + Y = n)\, z^n$$

$$= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \mathbf{P}(X = k, Y = n - k)\, z^n$$

$$= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \mathbf{P}(X = k)\, \mathbf{P}(Y = n - k)\, z^n \qquad \text{car les variables sont indépendantes}$$

On reconnaît alors l'expression d'une série produit :

$$G_{X+Y}(z) = \sum_{n=0}^{\infty} \mathbf{P}(X = n)\, z^n \times \sum_{n=0}^{\infty} \mathbf{P}(Y = n)\, z^n$$

Ainsi :

$$G_{X+Y} = G_X \times G_Y$$

❑ Une autre démonstration consiste à dire :

$$G_{X+Y}(z) = \mathbf{E}(z^{X+Y}) = \mathbf{E}(z^X z^Y) = \mathbf{E}(z^X)\, \mathbf{E}(z^Y) = G_X(z) \times G_Y(z)$$

Cette démonstration utilise les résultats suivants :
i) si X et Y sont indépendantes, alors une fonction de X (à savoir z^X) est indépendante d'une fonction de Y (à savoir z^Y).
ii) L'espérance du produit de deux variables aléatoires indépendantes est égal au produit des espérances de chacune de ces variables.

Le résultat prouvé se généralise par récurrence à la somme de n variables aléatoires indépendantes. La fonction génératrice de la somme est le produit des n fonctions génératrices.

EXEMPLES D'UTILISATION :

❑ La loi binomiale $\mathcal{B}(n, p)$ est la loi de la somme de n variables aléatoires indépendantes de Bernoulli de paramètre p. On retrouve effectivement :

$$G_{\text{binomiale}}(z) = (1 - p + pz)^n = G_{\text{Bernoulli}}(z)^n$$

❑ La somme de deux variables aléatoires indépendantes de lois binomiales $\mathcal{B}(n, p)$ et $\mathcal{B}(m, p)$ suit une loi binomiale $\mathcal{B}(n + m, p)$. Et en effet :

$$(1 - p + pz)^n\ (1 - p + pz)^m = (1 - p + pz)^{n+m}$$

❑ La somme de n variables aléatoires indépendantes de loi de Poisson de paramètres $\lambda_1, ..., \lambda_n$ suit une loi de Poisson de paramètre la somme des λ_i. En effet :

$$\exp(\lambda_1(z - 1))...\exp(\lambda_n(z - 1)) = \exp((\lambda_1 + ... + \lambda_n)(z - 1))$$

Une utilisation des séries génératrices et de la diagonalisation des matrices est donnée dans l'annexe qui suit.

Annexe : Chaînes de Markov

1- Matrice de transition d'un système

Considérons un système pouvant se trouver dans divers états possibles, numérotés de 1 à n. A intervalle de temps régulier, le système change d'état, la probabilité qu'il passe de l'état i à l'état j étant p_{ij}. Il est possible également qu'on puisse rester dans le même état i avec une certaine probabilité, ce qui signifie que p_{ii} est non nul. Parfois, parmi les n états, il existe un état (ou plusieurs) final, par exemple l'état n. Si on atteint cet état, on y reste définitivement, ce qu'on peut traduire par :

$$p_{nn} = 1 \qquad \text{et} \qquad p_{ni} = 0 \text{ si } i \neq n$$

Notons M la matrice de $\mathcal{M}_n(\mathbf{R})$ dont le terme (i, j) vaut p_{ij}. M est la **matrice de transition** du système. On remarque que la somme des coefficients de chaque ligne vaut 1, égale à la probabilité qu'on passe de l'état i à un autre état (éventuellement le même).

A l'instant k, on note $X(k)$ la variable aléatoire précisant dans quel état se trouve le système. Soit $L(k)$ le vecteur-ligne dont la i-ème composante est égale à $\mathbf{P}(X(k) = i)$, probabilité que le système à l'instant k se trouve à l'état i. $L(k)$ représente la loi de $X(k)$. A l'instant initial, le système de trouve dans l'état i suivant une certaine loi de probabilité, définissant le vecteur $L(0)$. Par exemple, si le système se trouve de manière certaine à l'état 1, alors $L(0) = (1\ 0\ \dots\ 0)$. S'il se trouve à l'instant initial dans l'état 1 ou 2 avec équiprobabilité, alors $L(0) = (\frac{1}{2}\ \ \frac{1}{2}\ \ 0\ \ \dots\ \ 0)$.

Considérons un instant k particulier. L'évènement $X(k) = i$ consiste à se trouver dans l'état i à cet instant. L'évènement $X(k + 1) = j$, consistant à se trouver dans l'état j à l'instant $k + 1$. La probabilité conditionnelle $\mathbf{P}(X(k + 1) = j \mid X(k) = i)$ n'est autre que p_{ij}. La loi des probabilité totale énonce que :

$$\mathbf{P}(X(k + 1) = j) = \sum_{i=1}^{n} \mathbf{P}(X(k + 1) = j \mid X(k) = i)\ \mathbf{P}(X(k) = i)$$

Or $\mathbf{P}(X(k) = i)$ n'est autre que la i-ème composante de $L(k)$, qu'on peut noter $L_i(k)$, et $\mathbf{P}(X(k + 1) = j)$ est la j-ème composante de $L(k + 1)$, que l'on peut noter $L_j(k + 1)$. On a donc :

$$\forall j,\ L_j(k + 1) = \sum_{i=1}^{n} p_{ij}\ L_i(k)$$

ce qui exprime exactement le fait que la ligne $L(k + 1)$ est le produit de la ligne $L(k)$ par la matrice M :

$$L(k + 1) = L(k)M$$

(Si on souhaitait manipuler les vecteurs donnant la loi de $X(k)$ en colonne, il conviendrait de prendre M^T).

Par récurrence, on en déduit que $L(k) = L(0)M^k$. On détermine explicitement $L(k)$ en parvenant à calculer M^k.

Le schéma précédent caractérise les chaînes de Markov. Dans ce type de processus, ce qui se passe en partant de l'état i est indépendant de la façon dont on est arrivé à l'état i : le futur est indépendant du passé. On peut résumer l'étude précédente de la façon suivante :

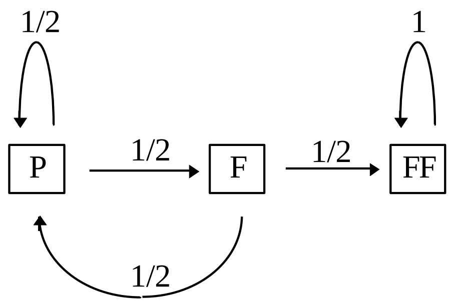
PROPOSITION

Soit une chaîne de Markov $(X(k))_{k\in\mathbf{N}}$, *où* $X(k)$ *désigne la variable aléatoire donnant l'état d'un système fini au bout du temps k. On se donne la matrice de transition* M *du système dont le terme général vaut* $p_{ij} = \mathbf{P}(X(k+1) = j \mid X(k) = i)$ *et ne dépend pas de k. Soit* $L(k)$ *le vecteur-ligne de composantes* $\mathbf{P}(X(k) = i)$. *Alors* $L(k) = L(0)M^k$ *ce qui permet de déterminer la loi de* $X(k)$ *à partir de celle de* $X(0)$.

EXEMPLE :

❑ On lance une pièce jusqu'à obtenir deux Faces successifs (FF). Il est clair que, si on tire Pile (P) avant d'avoir tiré les deux Faces, tout est à recommencer. On dispose donc des trois états suivants :

état 1	P : état initial ou état après avoir tiré un Pile
état 2	F : état suivant l'état 1 après tirage d'un Face
état 3	FF : état final, après tirage successif de deux Faces



Pour une pièce équilibrée, la matrice de transition est $M = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}$. Supposons que l'on parte de $L(0) = (1\ \ 0\ \ 0)$ (probabilité certaine d'être dans l'état initial). Les lois $L(k)$ de $X(k)$ valent successivement, en multipliant itérativement à droite par M :

$$L(0) = (1\quad 0\quad 0)$$

$$L(1) = (\frac{1}{2}\quad \frac{1}{2}\quad 0)$$

$$L(2) = (\frac{1}{2}\quad \frac{1}{4}\quad \frac{1}{4})$$

$$L(3) = (\frac{3}{8}\quad \frac{1}{4}\quad \frac{3}{8})$$

$$L(4) = (\frac{5}{16}\quad \frac{3}{16}\quad \frac{1}{2})$$

$$L(5) = (\frac{1}{4}\quad \frac{5}{32}\quad \frac{19}{32})$$

$$L(6) = (\frac{13}{64}\quad \frac{1}{8}\quad \frac{43}{64})$$

...

La somme des termes de chaque ligne doit donner 1. On peut déterminer $L(k)$ en calculant M^k en diagonalisant M. Le polynôme caractéristique de M vaut :

$$\chi_M(X) = X^3 - \frac{3X^2}{2} + \frac{X}{4} + \frac{1}{4} = (X-1)(X^2 - \frac{X}{2} - \frac{1}{4})$$

de racines 1, $\varphi = \dfrac{1+\sqrt{5}}{4}$ et $\psi = \dfrac{1-\sqrt{5}}{4}$. M est diagonalisable de matrice diagonale $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \varphi & 0 \\ 0 & 0 & \psi \end{pmatrix}$.

Une matrice de passage est donnée par $Q = \begin{pmatrix} 1 & 2\varphi & 2\psi \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$. $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ est toujours vecteur propre d'une matrice de transition, associée à la valeur propre 1, du fait que la somme des lignes de M vaut 1. Q^{-1} vaut :

$$Q^{-1} = \frac{1}{2\varphi - 2\psi} \begin{pmatrix} 0 & 0 & 2\varphi - 2\psi \\ 1 & -2\psi & -2\varphi \\ -1 & 2\varphi & 2\psi \end{pmatrix}$$

On a $M = QDQ^{-1}$ et $M^k = QD^kQ^{-1}$.

Donc : $L(k) = L(0)QD^kQ^{-1}$

$$= (1\ 0\ 0) \begin{pmatrix} 1 & 2\varphi & 2\psi \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \varphi^k & 0 \\ 0 & 0 & \psi^k \end{pmatrix} \frac{1}{2\varphi - 2\psi} \begin{pmatrix} 0 & 0 & 2\varphi - 2\psi \\ 1 & -2\psi & -2\varphi \\ -1 & 2\varphi & 2\psi \end{pmatrix}$$

$$= (1\ 2\varphi\ 2\psi) \begin{pmatrix} 1 & 0 & 0 \\ 0 & \varphi^k & 0 \\ 0 & 0 & \psi^k \end{pmatrix} \frac{1}{2\varphi - 2\psi} \begin{pmatrix} 0 & 0 & 2\varphi - 2\psi \\ 1 & -2\psi & -2\varphi \\ -1 & 2\varphi & 2\psi \end{pmatrix}$$

$$= (1\ 2\varphi^{k+1}\ 2\psi^{k+1}) \frac{1}{2\varphi - 2\psi} \begin{pmatrix} 0 & 0 & 2\varphi - 2\psi \\ 1 & -2\psi & -2\varphi \\ -1 & 2\varphi & 2\psi \end{pmatrix}$$

$$= \left(\frac{\varphi^{k+1} - \psi^{k+1}}{\varphi - \psi} \quad \frac{2(\varphi\psi^{k+1} - \varphi^{k+1}\psi)}{\varphi - \psi} \quad 1 + \frac{2(\psi^{k+2} - \varphi^{k+2})}{\varphi - \psi}\right)$$

Ces trois nombres donnent la probabilité de se trouver à l'instant k respectivement dans les états 1, 2 et 3.

Que se passe-t-il quand k tend vers l'infini ? Comme $|\varphi| < 1$ et $|\psi| < 1$, φ^k et ψ^k tendent vers 0 et $L(k)$ tend vers $(0\ \ 0\ \ 1)$ et donc la probabilité de se trouver dans l'état final FF tend vers 1.

Soit T_1 la variable aléatoire donnant l'instant k auquel on atteint l'état final FF en partant de l'état 1. On a :

$$\mathbf{P}(T_1 \leq k) = \mathbf{P}(X(k) = 3) = 1 + \frac{2(\psi^{k+2} - \varphi^{k+2})}{\varphi - \psi}$$

$$\mathbf{P}(T_1 \text{ fini}) = \mathbf{P}(\cup \{T_1 \leq k\}) = \lim_{k\to\infty} \mathbf{P}(T_1 \leq k) = 1$$

Ainsi, on est certain d'atteindre l'état final en un temps fini. Boucler indéfiniment dans le diagramme est un évènement de probabilité nulle. T_1 s'appelle le **temps d'arrêt** du processus. Cherchons son espérance :

$$\mathbf{E}(T_1) = \sum_{k=0}^{\infty} k\,\mathbf{P}(T_1 = k) = \sum_{k=0}^{\infty} \mathbf{P}(T_1 > k) = \sum_{k=0}^{\infty} (1 - \mathbf{P}(T_1 \leq k))$$

$$= \sum_{k=0}^{\infty} \frac{2(\varphi^{k+2} - \psi^{k+2})}{\varphi - \psi} = \frac{2}{\varphi - \psi}\left(\frac{\varphi^2}{1-\varphi} - \frac{\psi^2}{1-\psi}\right)$$

$$= \frac{2}{\varphi - \psi} \frac{\varphi^2 - \psi^2 - \varphi^2\psi + \varphi\psi^2}{1 - \varphi - \psi + \varphi\psi} = 2 \frac{\varphi + \psi - \varphi\psi}{1 - \varphi - \psi + \varphi\psi} \qquad \text{or } \varphi + \psi = \frac{1}{2} \text{ et } \varphi\psi = -\frac{1}{4}$$
$$= 6$$

2- Temps d'arrêt

Soit, comme précédemment, un système de Markov doté d'un état final n. Soit T_i le temps d'arrêt pour passer de l'état i à l'état final n. On a alors, pour tout $i < n$ et tout $k > 0$:

$$T_i = 1 + \sum_{j=1}^{n} X_{ij}T_j \qquad \text{où } X_{ij} = 1 \text{ si on passe de l'état } i \text{ à l'état } j \text{ au premier saut et } = 0 \text{ sinon}$$

mais $T_n = 0$

donc $T_i = 1 + \sum_{j=1}^{n-1} X_{ij}T_j$

Si $k = 1$, $\mathbf{P}(T_i = k) = \mathbf{P}(T_i = 1) = p_{in}$.
Si $k \geq 2$, on a :

$$\mathbf{P}(T_i = k) = \mathbf{P}(\sum_{j=1}^{n-1} X_{ij}T_j = k-1) = \mathbf{P}(\bigcup_{j=0}^{n-1} \{X_{ij} = 1 \cap T_j = k-1\})$$

$$= \sum_{j=1}^{n-1} \mathbf{P}(X_{ij} = 1 \cap T_j = k-1) = \sum_{j=1}^{n-1} \mathbf{P}(X_{ij} = 1)\, \mathbf{P}(T_j = k-1)$$

$$= \sum_{j=1}^{n-1} p_{ij}\, \mathbf{P}(T_j = k-1)$$

Si l'on note $T(k)$ le vecteur-colonne de composantes $\mathbf{P}(T_i = k)$, $1 \leq i \leq n-1$, on obtient la relation $T(k) = N\,T(k-1)$, où N est la matrice $(n-1) \times (n-1)$ dont le terme général est p_{ij}, $1 \leq i \leq n-1$. Il s'agit d'une sous-matrice de la matrice M déjà rencontrée, obtenue en supprimant la ligne et la colonne n, relative à l'état final.

On aura donc $T(k) = N^{k-1}\,T(1)$, avec $T(1) = \begin{pmatrix} p_{1n} \\ p_{2n} \\ \dots \\ p_{n-1,n} \end{pmatrix}$ donc $T(k) = N^{k-1}\,T(1)$. Le calcul de la loi des T_i peut donc se faire également en diagonalisant la matrice N.

EXEMPLE :

❑ On reprend la matrice de transition de l'exemple précédent, on a $N = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 0 \end{pmatrix}$ et $T(1) = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix}$.
On trouve que les $T(k)$ successifs valent :

$$\begin{matrix} \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} & \begin{pmatrix} 1/4 \\ 0 \end{pmatrix} & \begin{pmatrix} 1/8 \\ 1/8 \end{pmatrix} & \begin{pmatrix} 1/8 \\ 1/16 \end{pmatrix} & \begin{pmatrix} 3/32 \\ 1/16 \end{pmatrix} & \begin{pmatrix} 5/64 \\ 3/64 \end{pmatrix} & \dots \\ T(1) & T(2) & T(3) & T(4) & T(5) & T(6) & \dots \end{matrix}$$

La loi de T_1 se trouve dans la première composante de chaque colonne, la loi de T_2 dans la deuxième composante. On retrouve comme polynôme caractéristique de N le polynôme $X^2 - \frac{X}{2} - \frac{1}{4}$

de racine φ et ψ. N est diagonalisable de matrice diagonalisable $\Delta = \begin{pmatrix} \varphi & 0 \\ 0 & \psi \end{pmatrix}$ et de matrice de passage $R = \begin{pmatrix} 2\varphi & 2\psi \\ 1 & 1 \end{pmatrix}$ et d'inverse $R^{-1} = \dfrac{1}{2\varphi - 2\psi} \begin{pmatrix} 1 & -2\psi \\ -1 & 2\varphi \end{pmatrix}$. On a $N = R\Delta R^{-1}$ donc $N^{k-1} = R\,\Delta^{k-1}\,R^{-1}$ donc :

$$
\begin{aligned}
T(k) &= R\,\Delta^{k-1}\,R^{-1}\,T(1) \\
&= \begin{pmatrix} 2\varphi & 2\psi \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi^{k-1} & 0 \\ 0 & \psi^{k-1} \end{pmatrix} \frac{1}{2\varphi - 2\psi} \begin{pmatrix} 1 & -2\psi \\ -1 & 2\varphi \end{pmatrix} \begin{pmatrix} 0 \\ 1/2 \end{pmatrix} \\
&= \frac{1}{2\varphi - 2\psi} \begin{pmatrix} 2\varphi & 2\psi \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi^{k-1} & 0 \\ 0 & \psi^{k-1} \end{pmatrix} \begin{pmatrix} -\psi \\ \varphi \end{pmatrix} \\
&= \frac{1}{2\varphi - 2\psi} \begin{pmatrix} 2\varphi & 2\psi \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -\psi\varphi^{k-1} \\ \varphi\psi^{k-1} \end{pmatrix} \\
&= \frac{1}{2\varphi - 2\psi} \begin{pmatrix} 2\varphi\psi^{k} - 2\psi\varphi^{k} \\ \varphi\psi^{k-1} - \psi\varphi^{k-1} \end{pmatrix}
\end{aligned}
$$

Ainsi :

$$\mathbf{P}(T_1 = k) = \frac{\varphi\psi^k - \psi\varphi^k}{\varphi - \psi}$$

$$\mathbf{P}(T_2 = k) = \frac{\varphi\psi^{k-1} - \psi\varphi^{k-1}}{2\varphi - 2\psi}$$

On en déduit que :

$$\mathbf{E}(T_1) = \sum_{k=0}^{\infty} k\,\frac{\varphi\psi^k - \psi\varphi^k}{\varphi - \psi} = \frac{1}{\varphi - \psi}\left(\frac{\varphi\psi}{(1-\psi)^2} - \frac{\varphi\psi}{(1-\varphi)^2}\right) = \ldots = 6 \qquad \text{valeur trouvée plus haut}$$

$$\mathbf{E}(T_2) = \sum_{k=0}^{\infty} k\,\frac{\varphi\psi^{k-1} - \psi\varphi^{k-1}}{2\varphi - 2\psi} = \frac{1}{2(\varphi - \psi)}\left(\frac{\varphi}{(1-\psi)^2} - \frac{\psi}{(1-\varphi)^2}\right) = \ldots = 4$$

On peut aussi utiliser les fonctions génératrices. Soit $g_i(z) = \sum_{k=0}^{\infty} \mathbf{P}(T_i = k)\, z^k$. n étant l'état final, on a $g_n(z) = 1$. Par ailleurs, les relations $\mathbf{P}(T_i = 0) = 0$, $\mathbf{P}(T_i = 1) = p_{in}$ et $\mathbf{P}(T_i = k) = \sum_{j=1}^{n-1} p_{ij}\, \mathbf{P}(T_j = k - 1)$ pour $i \le n - 1$ et $k \ge 2$ donne :

$$
\begin{aligned}
g_i(z) &= p_{in}\, z + \sum_{j=1}^{n-1} p_{ij} \sum_{k=2}^{\infty} \mathbf{P}(T_j = k-1)\, z^k \\
&= p_{in}\, z + \sum_{j=1}^{n-1} p_{ij} \sum_{k=1}^{\infty} \mathbf{P}(T_j = k)\, z^{k+1} \\
&= p_{in}\, z + \sum_{j=1}^{n-1} p_{ij}\, z g_j(z)
\end{aligned}
$$

$\Rightarrow \qquad g_i(z) = \sum_{j=1}^{n} p_{ij}\, z g_j(z)$

On obtient ainsi un système d'inconnues les g_i, qui, une fois résolu, permettent, en théorie, d'en tirer les lois des T_i, ainsi que leur espérance et variance.

EXEMPLE :

❑ Toujours avec l'exemple précédent, on a :

$$\begin{cases} g_1 = \dfrac{1}{2} z g_1 + \dfrac{1}{2} z g_2 \\ g_2 = \dfrac{1}{2} z g_1 + \dfrac{1}{2} z \end{cases}$$

d'où, en résolvant le système :

$$\begin{cases} g_1 = \dfrac{z^2}{4 - 2z - z^2} = \dfrac{z^2}{4} + \dfrac{z^3}{8} + \dfrac{z^4}{8} + \dfrac{3z^5}{32} + \dfrac{5z^6}{64} + \ldots \\ g_2 = \dfrac{2z - z^2}{4 - 2z - z^2} = \dfrac{z}{2} + \dfrac{z^3}{8} + \dfrac{z^4}{16} + \dfrac{z^5}{16} + \dfrac{3z^6}{64} + \ldots \end{cases}$$

Le calcul du terme général du développement en série entière demanderait de factoriser le dénominateur, de réduire la fraction obtenue en éléments simples et enfin d'en développer en série chacun des termes obtenus. Les fonctions génératrices sont plus efficaces pour calculer espérance et variance des temps d'arrêts. On trouve ainsi que :

$\mathbf{E}(T_1) = g_1'(z) = 6 \qquad \mathbf{V}(T_1) = g_1''(z) + g_1'(z) - g_1'(z)^2 = 22$

$\mathbf{E}(T_2) = g_2'(z) = 4 \qquad \mathbf{V}(T_2) = g_2''(z) + g_2'(z) - g_2'(z)^2 = 20$

On peut directement établir un système vérifié par les espérances en dérivant les relations $g_i(z) = \sum_{j=1}^{n} p_{ij}\, z g_j(z)$, ce qui donne :

$$g_i'(z) = \sum_{j=1}^{n} p_{ij}\,(g_j(z) + z g_j'(z)) \qquad \text{qu'on évalue en } z = 1$$

$$\Rightarrow \qquad \mathbf{E}(T_i) = \sum_{j=1}^{n} p_{ij}\,(1 + \mathbf{E}(T_j)) = 1 + \sum_{i=1}^{n} p_{ij}\, \mathbf{E}(T_j)$$

En moyenne, le temps d'arrêt de T_i est égal à 1 (correspondant à un saut de transition) + la moyenne des temps d'arrêt des états suivants, pondérés par la probabilité de passer de l'état *i* à l'état *j*.

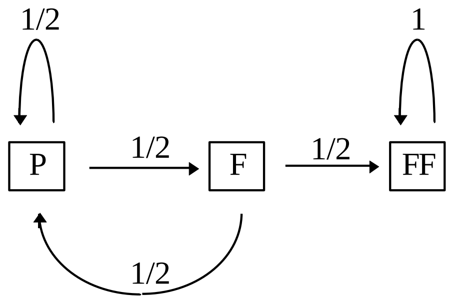
EXEMPLES :

❑ Retrouvons les valeurs déjà trouvées pour atteindre FF.

Soient $\mathbf{E}_P$ = espérance du temps d'arrêt $P \xrightarrow{*} FF$

$\mathbf{E}_F$ = espérance du temps d'arrêt $F \xrightarrow{*} FF$

où $\xrightarrow{*}$ désigne un nombre indéterminé de changement d'états



$$\begin{cases} \mathbf{E}_P = \dfrac{\mathbf{E}_P + 1}{2} + \dfrac{\mathbf{E}_F + 1}{2} \\ \mathbf{E}_F = \dfrac{\mathbf{E}_P + 1}{2} + \dfrac{1}{2} = \dfrac{\mathbf{E}_P}{2} + 1 \end{cases}$$

donc $\mathbf{E}_P = 6$ et $\mathbf{E}_F = 4$.

3- Compétition entre deux états finaux

Considérons un système possédant deux états finaux. On cherche la probabilité d'atteindre le premier état plutôt que le second. On se borne à donner un exemple.

❑ On considère une pièce ayant une probabilité p de tomber sur P et $q = 1 - p$ de tomber sur F. Sur l'axe réel, on part de 0. A chaque P, on fait un saut vers la droite et à chaque F, on fait un saut vers la gauche. Soit n un entier donné. On arrête les lancers quand on a atteint n ou $-n$. On cherche la probabilité d'atteindre n et celle d'atteindre $-n$.

Soit $\mathbf{P}(k)$ la probabilité d'atteindre n en partant de k. On a :

$\mathbf{P}(n) = 1$

$\mathbf{P}(-n) = 0$

$\forall k \in \{-(n-1), -(n-2), ..., n-2, n-1\}, \mathbf{P}(k) = p\mathbf{P}(k+1) + q\mathbf{P}(k-1)$

On obtient donc une suite récurrente linéaire d'ordre 2 dont l'équation caractérisque d'inconnue r est :

$$pr^2 - r + q = 0$$

et dont les racines sont $r = 1$ et $r = \dfrac{q}{p}$. Dans le cas où $q \neq p$, il existe λ et μ tel que, pour tout k :

$$\mathbf{P}(k) = \lambda + \mu\left(\frac{q}{p}\right)^k$$

On trouve λ et μ avec les conditions dans l'état n :

$$\begin{cases} \lambda + \mu\left(\dfrac{q}{p}\right)^n = 1 \\ \lambda + \mu\left(\dfrac{q}{p}\right)^{-n} = 0 \end{cases}$$

donc $\mu = \dfrac{1}{\left(\frac{q}{p}\right)^n - \left(\frac{q}{p}\right)^{-n}} = \dfrac{\left(\frac{q}{p}\right)^n}{\left(\frac{q}{p}\right)^{2n} - 1} = \dfrac{p^n q^n}{q^{2n} - p^{2n}}$ et $\lambda = -\dfrac{p^{2n}}{q^{2n} - p^{2n}}$

La valeur de $\mathbf{P}(0)$ est $\lambda + \mu = \dfrac{p^n}{q^n + p^n}$.

EXEMPLE :

❑ D'après Huygens[3] : "Ayant pris chacun 12 jetons, A et B jouent avec trois dés à cette condition qu'à chaque coup de 11 points, A doit donner un jeton à B et que B en doit donner un à A à chaque coup de 14 points, et que celui là gagnera qui sera le premier en possession de tous les jetons. On trouve dans ce cas que la chance de A est à celle de B comme 244 140 625 est à 282 429 536 481".
On vérifiera que la probabilité d'obtenir 11 avec trois dés est $\dfrac{27}{6^3}$ et celle d'obtenir 14 est de $\dfrac{15}{6^3}$.
Comme les autres totaux n'interviennent pas, on peut considérer les probabilités conditionnelles sachant qu'on tire 11 ou 14, celles-ci valant respectivement $\dfrac{27}{27+15} = \dfrac{9}{14} = p$ et $\dfrac{15}{27+15} = \dfrac{5}{14} = q$.
B gagne avec la probabilité $\dfrac{p^{12}}{p^{12}+q^{12}} = \dfrac{9^{12}}{9^{12}+5^{12}}$, avec $9^{12} =$ 282 429 536 481 et $5^{12} =$ 244 140 625.

Exercices

1- Enoncés sur les probabilités

Exo.1-1) a) Trois personnes A, B, C jettent une pièce équilibrée dans l'ordre A, B, C, A, B, C, ... jusqu'à ce que le côté Pile sorte. Le vainqueur est le premier qui obtient Pile. Calculer les probabilités de victoire de chacun des trois joueurs.

b) Les trois joueurs A, B, C jouent maintenant à Pile ou Face. A commence par jouer contre B, puis le vainqueur joue contre C, puis le vainqueur de cette dernière partie joue contre le perdant de la première partie. Et ainsi de suite, le vainqueur de la dernière partie jouant contre le perdant de l'avant-dernière. Le jeu s'arrête lorsqu'un joueur a gagné deux parties de suite. On note :

A_1 (respectivement B_1) l'événement "A (respectivement B) gagne la première partie".

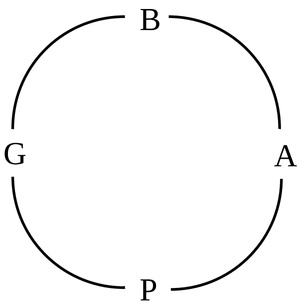
A_g (respectivement B_g, C_g) l'événement "A (respectivement B, C) gagne deux parties de suite".

Calculer les probabilités $\mathbf{P}(A_g \mid A_1)$, $\mathbf{P}(B_g \mid A_1)$, $\mathbf{P}(C_g \mid A_1)$, $\mathbf{P}(A_g \mid B_1)$, $\mathbf{P}(B_g \mid B_1)$, $\mathbf{P}(C_g \mid B_1)$, $\mathbf{P}(A_g)$, $\mathbf{P}(B_g)$, $\mathbf{P}(C_g)$.

Exo.1-2) Huygens : *De ratiociniis in ludo aleae* (1657), premier problème. "A et B jouent ensemble avec deux dés à la condition suivante : A aura gagné s'il jette 6 points, B s'il en jette 7. A fera le premier un seul coup ; ensuite B deux coups successifs ; puis de nouveau A deux coups, et ainsi de suite, jusqu'à ce que l'un ou l'autre aura gagné. On demande le rapport de la chance de A à celle de B. Réponse : comme 10355 à 12276".
Vérifier le résultat de Huygens.

Exo.1-3) Un jeu se joue de la façon suivante : on dispose quatre cases A, B, G, P en cercle. On place un jeton en A. On lance un dé et on avance le jeton du nombre de cases indiqué, dans le sens $A \to B \to G \to P \to A \to B \to$ etc. Si le jeton arrive en G, on gagne. S'il arrive en P, on perd. Sinon, on lance à nouveau le dé jusqu'à arriver en G ou P. Quelle est la probabilité de gagner à ce jeu ?

[3] Huygens, *De ratiociniis in ludo aleae* (1657), cinquième problème, http://gallica.bnf.fr/ark:/12148/bpt6k77862v/f95.item



Exo.1-4) On lance deux dés indéfiniment. A chaque tirage, on fait la somme des deux dés.

a) Quelle est la probabilité que la somme 6 sorte avant la somme 7 ?

b) Quelle est la probabilité que la somme 8 sorte avant la somme 7 ?

c) Quelle est la probabilité que les sommes 6 et 8 soient toutes deux sorties avant que le deuxième 7 ne soit sorti ?

Exo.1-5) Soit $(\Omega, \mathcal{T})$ un espace muni d'une tribu $\mathcal{T}$, et $\mathbf{P}$ une application de $\mathcal{T}$ dans $[0, 1]$ telle que :

i) $\mathbf{P}(\Omega) = 1$

ii) $\forall$ (A, B) $\in \mathcal{T} \times \mathcal{T}$, $A \cap B = \varnothing \Rightarrow \mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$

iii) Pour toute suite (A_n) décroissante d'éléments de $\mathcal{T}$ telle que $\bigcap\limits_{n=1}^{\infty} A_n = \varnothing$, $\lim\limits_{n\to\infty} \mathbf{P}(A_n) = 0$.

Montrer que $\mathbf{P}$ est une probabilité sur $(\Omega, \mathcal{T})$.

Exo.1-6) Soit $(A_n)_{n\geq 1}$ une suite d'événements d'un espace probabilisé Ω telle que :

$$\text{(i) } \forall\, n \geq 1,\ \mathbf{P}(A_n) \geq 1 - \frac{1}{2^n} \quad \text{et} \quad \text{(ii) } \exists\, k \geq 1,\ \mathbf{P}(A_k) > 1 - \frac{1}{2^k}$$

a) Montrer que $\mathbf{P}(\bigcap\limits_{n=1}^{\infty} A_n) > 0$.

b) Donner un exemple pour lequel seule la condition (i) est vérifiée et $\mathbf{P}(\bigcap\limits_{n=1}^{\infty} A_n) = 0$

Exo.1-7) Pour tout $\alpha > 1$, on pose $\zeta(\alpha) = \sum\limits_{n=1}^{\infty} \frac{1}{n^\alpha}$. Sur $\mathbb{N}^*$, on se donne une loi de probabilité $\mathbf{P}$ définie par :

$$\forall\, n \geq 1,\ \mathbf{P}(\{n\}) = \frac{1}{\zeta(\alpha)n^\alpha}$$

Pour tout entier $p \geq 1$, on considère l'événement $A_p = \{n \in \mathbb{N}^* \mid n \text{ est un multiple de } p\}$

a) Calculer $\mathbf{P}(A_p)$.

b) Vérifier que la famille d'événements $(A_p)_{p\ \text{premier}}$ forme une famille d'événements indépendants.

c) On numérote les nombres premiers par ordre croissant : $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, etc. En considérant $\lim\limits_{n\to\infty} \mathbf{P}(\bigcap\limits_{k=1}^{n} A_{p_k}{}^{c})$ et $\mathbf{P}(\{1\})$, montrer que $\zeta(\alpha) = \prod\limits_{k=1}^{\infty} \dfrac{1}{1 - \dfrac{1}{p_k{}^{\alpha}}}$.

On obtient ainsi une preuve probabiliste de la formule $\sum\limits_{n=1}^{\infty} \dfrac{1}{n^{\alpha}} = \prod\limits_{k=1}^{\infty} \dfrac{1}{1 - \dfrac{1}{p_k{}^{\alpha}}}$, dont on trouvera une autre démonstration dans le paragraphe *Famille sommable* du chapitre L2/SERIES.PDF.

Exo.1-8) Soit $\Omega = \{(x_n)_{n\geq 1} \mid \forall\ n,\ x_n = \text{P ou } x_n = \text{F}\}$ l'univers correspondant au lancer indéfiniment d'une pièce équilibrée. Calculer les probabilités des événements suivants :

a) Pour tout m et $i \geq 1$, $B_{im} = \{(x_n)_{n\geq 1} \mid \forall\ j \in [\![1, m]\!],\ \forall\ k \in [\![1, i]\!],\ x_{j+km} = x_j\}$

b) Pour tout $m \geq 1$, $B_m = \{(x_n)_{n\geq 1} \mid \forall\ j \in [\![1, m]\!],\ \forall\ k \geq 1,\ x_{j+km} = x_j\}$

c) $B = \{(x_n)_{n\geq 1} \mid \exists\ m,\ \forall\ j \in [\![1, m]\!],\ \forall\ k \geq 1,\ x_{j+km} = x_j\}$

Qu'a-t-on montré ?

Exo.1-9) Le théorème de Borel Cantelli consiste en les énoncés c) et d) de cet exercice.
Soit $(A_n)_{n\in\mathbf{N}}$ une famille d'événements d'un espace probabilisé $(\Omega, \mathbf{P})$. On pose :

$$\text{limsup}(A_n) = \bigcap_{n=0}^{\infty} \bigcup_{p=n}^{\infty} A_p \qquad \text{et} \qquad \text{liminf}(A_n) = \bigcup_{n=0}^{\infty} \bigcap_{p=n}^{\infty} A_p$$

a) Soit et ω un élément de Ω. Regrouper les propriétés ci-dessous qui ont le même sens ($\mathbf{1}_A$ désigne la fonction indicatrice de A) :

i) L'ensemble des indices n pour lesquels ω appartient à A_n est infini

ii) L'ensemble des indices n pour lesquels ω n'appartient pas à A_n est fini

iii) $\omega \in \text{liminf}(A_n)$

iv) $\omega \in \text{limsup}(A_n)$

v) $\sum\limits_{n=0}^{\infty} \mathbf{1}_{A_n}(\omega) = +\infty$

vi) $\sum\limits_{n=0}^{\infty} (1 - \mathbf{1}_{A_n})(\omega) < +\infty$

vii) $\text{Inf}_{n\in\mathbf{N}}\ \text{Sup}_{p\geq n}\ \mathbf{1}_{A_p}(\omega) = 1$

viii) $\text{Sup}_{n\in\mathbf{N}}\ \text{Inf}_{p\geq n}\ \mathbf{1}_{A_p}(\omega) = 1$

b) Comparer :

i) $\text{limsup}(A_n \cup B_n)$ et $\text{limsup}(A_n) \cup \text{limsup}(B_n)$

ii) $\text{limsup}(A_n \cap B_n)$ et $\text{limsup}(A_n) \cap \text{limsup}(B_n)$

iii) $\text{liminf}(A_n \cup B_n)$ et $\text{liminf}(A_n) \cup \text{liminf}(B_n)$

iv) $\text{liminf}(A_n \cap B_n)$ et $\text{liminf}(A_n) \cap \text{liminf}(B_n)$

c) Montrer que, si la série $\sum \mathbf{P}(A_n)$ converge, alors $\mathbf{P}(\text{limsup}(A_n)) = 0$.

d) Montrer que, si $\sum \mathbf{P}(A_n)$ diverge et si les (A_n) sont indépendantes, alors $\mathbf{P}(\limsup(A_n)) = 1$. (On pourra calculer de deux façons la quantité $\lim\limits_{n\to\infty} 1 - P(\bigcap\limits_{p=n}^{\infty} A_p{}^c)$).

e) Soit (X_n) une suite de variables aléatoires indépendantes de même loi, à valeurs dans $\mathbb{N}$. On suppose que $\mathbf{E}(X_1)$ est fini. On pose $A_n = \{X_n \geq n\}$. Montrer que la probabilité de l'événément "il existe une infinité de n tel que A_n est réalisé" est nulle.

f) Sous les mêmes hypothèses, on suppose que $\mathbf{E}(X_1) = +\infty$. Montrer que la probabilité de l'événément "il existe une infinité de n tel que A_n est réalisé" est égale à 1.

2- Enoncés sur les variables aléatoires

Exo.2-1) Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. Comparer $\mathbf{P}$(X impair) et $\mathbf{P}$(X pair) à la valeur $\frac{1}{2}$ dans les cas suivants :

a) X suit une loi géométrique sur $\mathbb{N}^*$ de paramètre $p \in]0, 1[$.

b) X suit une loi de Poisson de paramètre $\lambda > 0$.

c) X suit une loi binomiale $\mathcal{B}(n, p)$.

Exo.2-2) Deux joueurs A et B jouent plusieurs parties d'un jeu. La probabilité que A gagne une partie contre B est p. Il n'y a pas de parties nulles. L'issue d'une partie est supposée indépendantes de celles des parties précédentes. A et B jouent jusqu'à ce que l'un d'entre eux gagne deux parties (pas nécessairement de suite) de plus que son adversaire.

a) Quelle est la probabilité que ce soit A qui gagne les deux parties de plus ?

b) Quelle est la loi et l'espérance du nombre de parties jouées, que le vainqueur final soit A ou B ?

Exo.2-3) On dispose d'un jeu de $2n$ cartes numérotées 1, 1, 2, 2, ..., n, n (chaque numéro apparaît donc sur deux cartes). On bat le jeu et on pose les cartes en ligne sur la table, face cachée. On retourne une première carte puis une deuxième carte de son choix et si ces deux cartes portent le même numéro, on les retire du jeu, sinon, on les repose face cachée sur la table à leur place respective. On itère jusqu'à avoir retiré toutes les cartes du jeu.

a) Le joueur sans mémoire ne se souvient ni de la place des cartes qu'il a déjà retournées ni de leur valeur. A chaque fois, il choisit une paire au hasard et la retourne. Montrer que l'espérance du nombre de paires retournées à la fin du jeu est n^2.

b) Le joueur à mémoire partielle ne se souvient pas de la valeur des cartes mais se souvient des positions des cartes déjà retournées. A chaque fois, il retourne la même première carte ainsi qu'une autre seconde carte toujours différente jusqu'à ce que la première et la seconde carte aient même valeur. Quelle est l'espérance du nombre de paires retournées ?

c) Le joueur à mémoire parfaite se souvient aussi bien des positions des cartes retournées que de leur valeur. S'il connaît la position de deux cartes similaires, il les retourne. Sinon, il retourne une carte inconnue et s'il connaît où se trouve la deuxième carte similaire, il la retourne, sinon, il retourne une deuxième carte inconnue. Montrer que l'espérance du nombre de paires retournées est un $O(n)$.

Exo.2-4) Soit X suivant une loi de Poisson de paramètre λ. Soit α un réel strictement positif. On souhaite définir une variable aléatoire Y fonction de X et de α, mais pas de λ, et dont l'espérance est $\exp(-\alpha\lambda)$.

a) Quelle est l'espérance de X ? Vérifier que la variable aléatoire $Y = \exp(-\alpha X)$ ne répond pas à la question.

b) Déterminer une variable aléatoire Y répondant à la question.

Exo.2-5) La formule de Wald : Soient $(X_n)_{n\in\mathbb{N}}$ des variables aléatoires définies sur un espace probabilisé Ω à valeurs dans $\mathbb{N}$, indépendantes et de même loi, et U une variable aléatoire définie sur Ω à valeurs dans $\mathbb{N}$, indépendante des X_i. Soit G_X la fonction génératrice commune aux X_i, et G_U celle de U. Toutes les variables aléatoires sont supposées avoir une espérance et une variance. Soit $T = \sum_{i=1}^{U} X_i$ la variable aléatoire définie par :

$$\forall\, \omega \in \Omega,\ T(\omega) = \sum_{i=1}^{U(\omega)} X_i(\omega).$$

(Si $U(\omega) = 0$, alors $T(\omega) = 0$).

a) Montrer que la fonction génératrice G_T de T est égale à $G_U \circ G_X$. On pourra calculer $\mathbf{E}(z^T)$ au moyen de l'utilisation de la formule de l'espérance totale, en conditionnant par rapport au système complet d'événement $\{U = n\}$.

b) En déduire $\mathbf{E}(T)$ et $\mathbf{V}(T)$ en fonction de $\mathbf{E}(U)$, $\mathbf{V}(U)$, $\mathbf{E}(X)$, $\mathbf{V}(X)$.

c) On considère une pièce ayant une probabilité p élément de $]0, 1[$ de tomber sur Pile. On lance cette pièce plusieurs fois jusqu'à ce qu'elle tombe sur Pile. Soit U le nombre de lancers. Puis, on lance de nouveau la pièce U fois. Soit T le nombre de Piles apparus lors de ces U lancers. Calculer $\mathbf{E}(T)$. Donner la loi de T.

Exo.2-6) On lance une pièce plusieurs fois de suite. Soit p la probabilité de tomber sur Pile, p élément de $]0, 1[$. On note X la variable aléatoire à valeurs dans $\mathbb{N}^*$, telle que $X = n$ si et seulement si c'est à l'issue du n-ème tirage qu'on a obtenu le tirage successif de deux Piles, et pas avant. Ainsi, dans la suite de tirages suivants FPFFPFPP (F = Face, P = Pile), X prend la valeur 8.

a) Donner les valeurs de $\mathbf{P}(X = 1)$, $\mathbf{P}(X = 2)$, $\mathbf{P}(X = 3)$.

b) Montrer que, pour $n \geq 3$, $\mathbf{P}(X = n) = (1-p)\mathbf{P}(X = n-1) + p(1-p)\mathbf{P}(X = n-2)$

c) Sans calculer explicitement $\mathbf{P}(X = n)$, en déduire une relation vérifiée par $\mathbf{E}(X)$ (qu'on supposera finie), puis la valeur de $\mathbf{E}(X)$.

d) Même question dans le cas où l'on attend le tirage de trois Piles successifs.

e) Même question dans le cas où l'on attend le tirage de a Piles successifs, a étant un entier strictement positif.

f) La situation précédente s'applique au cas d'un piéton souhaitant traverser une rue. Soit t une unité de temps donnée. p est la probabilité qu'aucune voiture ne passe devant lui dans la durée de temps t qui vient. On suppose que le fait qu'une voiture passe ou pas pendant cette durée est indépendant des passages de voitures précédentes. Le piéton a besoin d'une durée $T = at$ pour traverser, a entier strictement positif donné. La variable aléatoire X de la question e) est le nombre moyen de durées t qu'il doit laisser passer pour atteindre l'autre côté de la rue, de sorte que temps moyen d'attente et de traversée du piéton est $t\mathbf{E}(X)$. Quel est le temps moyen séparant le passage de deux voitures en fonction de t et p ? Quel est le débit D du flot de circulation (nombre moyen de véhicules passant devant le piéton par seconde) ? On fait tendre a vers l'infini, t vers 0 et p vers 1 de façon que T et D soient constants. Donner la limite de $t\mathbf{E}(X)$ en fonction de D et T.

Exo.2-7) Marche aléatoire dans $\mathbb{Z}$. Soit p élément de $]0, 1[$. On part de 0, et à chaque pas, on se déplace d'une unité à droite avec une probabilité p ou d'une unité vers la gauche avec une probabilité $1 - p$. Pour tout $n \geq 0$, soit X_n la position atteinte après n pas. On définit ainsi une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires. Au cours de cette marche, on est susceptible de repasser une ou plusieurs fois par 0, voire une infinité de fois. On définit la variable aléatoire R égale à ce nombre de fois. On dit que 0 est transitoire si $\mathbf{E}(R)$ est fini (en moyenne on ne revient en 0 qu'un nombre fini de fois), et récurrent si $\mathbf{E}(R)$ est infini (on y revient une infinité de fois).

a) Exprimer R à partir des variables aléatoires $\mathbf{1}_{X_n=0}$ (fonction indicatrice de l'événement $\{X_n = 0\}$).

b) En admettant qu'on puisse permuter espérance et série de fonctions indicatrices, en déduire $\mathbf{E}(R)$ à partir des $\mathbf{P}(X_n = 0)$.

c) Donner la valeur de $\mathbf{P}(X_n = 0)$. En déduire que 0 est récurrent si et seulement si $p = \frac{1}{2}$.

Exo.2-8) La ruine du joueur : Soit a et b deux entiers strictement positifs, $b > a$. Un joueur joue à Pile ou Face avec une pièce équilibrée. Il dispose au départ de la somme de a euros. S'il tombe sur Pile, il gagne un euro, s'il tombe sur Face, il perd un euro. Il s'arrête ou bien quand il est ruiné, ou bien quand il a atteint la somme de b euros, qu'il s'est fixé à l'avance, et il continue à jouer tant qu'il n'a pas atteint l'un de ces deux cas.

a) Pour tout entier k, on note $\mathbf{P}_k$ la probabilité de finir ruiné en ayant au départ un capital de $a = k$. Donner une relation de récurrence entre $\mathbf{P}_k$, $\mathbf{P}_{k+1}$ et $\mathbf{P}_{k-1}$, pour $1 \leq k \leq b - 1$. En déduire l'expression générale de $\mathbf{P}_k$. Quelle est la probabilité d'atteindre la somme b ? Quelle est la probabilité que le jeu dure indéfiniment ?

b) Quelle est l'espérance du gain du joueur ?

c) Soit X_k la variable aléatoire égale au nombre de lancers de pièces à effectuer jusqu'à la fin de la partie, en partant d'un capital égal à k. Donner une relation de récurrence entre $\mathbf{P}(X_k = n)$, $\mathbf{P}(X_{k+1} = n - 1)$ et $\mathbf{P}(X_{k-1} = n - 1)$, $1 \leq k \leq b - 1$, $n \geq 1$. En déduire une relation de récurrence entre $\mathbf{E}(X_k)$, $\mathbf{E}(X_{k-1})$ et $\mathbf{E}(X_{k+1})$, et la valeur de $\mathbf{E}(X_a)$.

d) Refaire l'exercice en prenant p la probabilité de tomber sur Pile et $1 - p$ celle de tomber sur Face, avec $p \neq \frac{1}{2}$.

Exo.2-9) On considère une urne contenant n jetons numérotés de 1 à n. On tire les jetons un par un avec remise et on note le numéro tiré. On appelle X la variable aléatoire égale au nombre de tirages pour que tous les jetons aient été tirés au moins une fois[4]. Le but de l'exercice est de déterminer la loi de X. On note X_1 le nombre de tirage pour tirer un premier jeton (donc X_1 est la variable certaine égale à 1), X_2 le nombre de tirages qui suivent X_1 pour tirer un jeton différent du premier, X_3 le nombre de tirages qui suivent X_2 pour tirer un jeton différent des deux premiers, etc... de sorte que $X = X_1 + X_2 + ... + X_n$.

a) Quelle est la loi de X_i et sa fonction génératrice G_i ?

b) Quelle est la fonction génératrice G_X de X ?

[4] Variante : Une marque de chocolat glisse, dans chacun des paquets qu'elle fabrique, une vignette des héros de l'espace à collectionner. Il y a en tout n vignettes. Les vignettes ont toutes la même probabilité d'être sélectionnées lors de l'achat d'un paquet. X est alors le nombre de paquets de chocolat à acheter pour avoir toutes les vignettes.

c) Décomposer G_X sous la forme $\frac{n!}{n^n} z^n \sum_{k=1}^{n-1} \frac{a_k}{1 - \frac{kz}{n}}$. On donnera une expression des a_k faisant intervenir le coefficient binomial $\binom{n-1}{k}$.

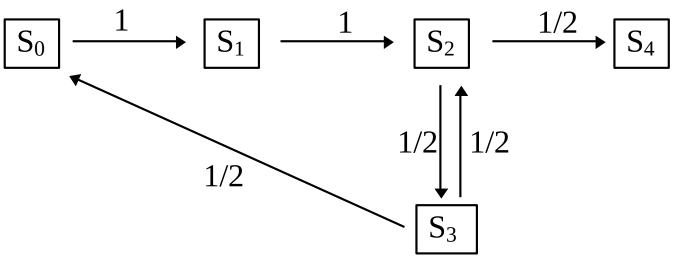
d) En déduire que, $\forall\, m \geq n$, $\mathbf{P}(X = m) = \sum_{k=1}^{n-1} (-1)^{n-1-k} \binom{n-1}{k} \frac{k^{m-1}}{n^{m-1}}$

e) Donner une expression de l'espérance de X. En donner un équivalent quand n tend vers $+\infty$.

f) Donner une expression de la variance de X. En donner un équivalent quand n tend vers $+\infty$.

Exo.2-10) Une fourmi parcourt les arêtes d'un cube. A chaque sommet, elle choisit au hasard selon une loi uniforme l'une des deux arêtes issues de ce sommet, autre que celle qu'elle vient immédiatement de parcourir. Partant d'un sommet, soit X le nombre d'arêtes qu'elle parcourt pour parvenir au sommet diamétralement opposé.

a) Justifier que le problème posé est identique au parcours de la fourmi sur le graphe suivant, où elle peut passer d'un sommet à l'autre selon la probabilité portée par l'arête orientée joignant les deux sommets.



b) Pour i variant de 0 à 3, soit T_i le nombre d'arêtes à parcourir dans le graphe précédent pour arriver au sommet S_4 en partant d'un sommet S_i (de sorte que $X = T_0$). Etablir des relations de récurrence entre les $\mathbf{P}(T_i = k)$, puis en déduire des relations entre les fonctions génératrices g_i des T_i.

c) En déduire la fonction génératrice de X, son espérance, sa variance.

d) Déterminer la loi de X.

Exo.2-11) Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires indépendantes suivant une même loi de Bernoulli de paramètre p. Pour tout n strictement positif, on pose $S_n = X_1 + ... + X_n$, nombre de succès après avoir effectué n épreuves. Soit m un entier strictement positif. On considère la variable aléatoire Y égale au plus petit entier n tel que $S_n = m$. Y est le nombre d'épreuves de Bernoulli à réaliser pour obtenir m succès.

a) Pour tout $k \geq m$, que vaut $\mathbf{P}(Y = k)$? La loi suivie par Y s'appelle **loi binomiale négative**.

b) Vérifier que $\sum_{k=m}^{\infty} \mathbf{P}(Y = k) = 1$. On pourra au préalable chercher le développement en série entière de la fonction $x \to \frac{1}{(1-x)^m}$.

c) Quelle est la fonction génératrice de la variable aléatoire Y ? Que vaut $\mathbf{E}(Y)$? $\mathbf{V}(Y)$?

d) Montrer que la somme de m variables aléatoires indépendantes de même loi géométrique de paramètre p suit la même loi que Y.

e) Réciproquement, soit $(Y_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes à valeurs dans $\mathbb{N}^*$ de même loi géométrique de paramètre p. Pour tout entier n strictement positif, on définit la variable aléatoire X_n de la façon suivante :

$$X_n = \begin{cases} 1 \text{ si et seulement si } \exists\, k \geq 1,\ Y_1 + Y_2 + ... + Y_k = n \\ 0 \text{ sinon} \end{cases}$$

Montrer que, pour tout n, X_n est une variable aléatoire de Bernoulli de paramètre p.

3- Enoncés sur les couples de variables aléatoires

Exo.3-1) On considère deux jeux dont les gains sont des variables aléatoires respectives G_1 et G_2. Un joueur tire d'abord à Pile ou Face pour savoir à quel jeu il va jouer, puis il joue à ce jeu. Soit G son gain.

a) A-t-on $\mathbf{E}(G) = \dfrac{\mathbf{E}(G_1) + \mathbf{E}(G_2)}{2}$? (où **E** désigne l'espérance)

b) A-t-on $\mathbf{V}(G) = \dfrac{\mathbf{V}(G_1) + \mathbf{V}(G_2)}{2}$? (où **V** désigne la variance)

Exo.3-2) On considère une variable aléatoire T à valeurs dans $\mathbb{N}^*$ suivant une loi géométrique de paramètre p élément de $]0, 1[$.

a) Soit X une variable aléatoire indépendante de T, à valeurs dans $\mathbb{N}$, et G_X sa fonction génératrice. Montrer que $G_X(1 - p) = \mathbf{P}(T > X)$

b) Soit X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$, telles que X, Y, et T soient des variables aléatoires indépendantes. Déduire du a) que $\mathbf{P}(T > X + Y \mid T > X) = \mathbf{P}(T > Y)$.

Cette dernière égalité est une généralisation de la propriété d'être sans mémoire des lois géométriques, énoncée dans le cours sous la forme $\mathbf{P}(T > n + k \mid T > k) = \mathbf{P}(T > n)$.

Exo.3-3) Soit X une variable aléatoire à valeurs dans $\mathbb{Z}$ telle que, pour tout n, $\mathbf{P}(X = n) = \mathbf{P}(X = -n)$, et ayant une variance finie. Soit Y une variable aléatoire indépendante de X à valeurs dans $\{-1, 1\}$, et telle que :

$\mathbf{P}(Y = 1) = p = 1 - \mathbf{P}(Y = -1)$.

a) Déterminer la loi de la variable aléatoire XY.

b) A quelle condition la covariance de X et de XY est-elle nulle ? X et XY sont-elles indépendantes ?

c) A quelle condition la covariance de $|X|$ et de XY est-elle nulle ? $|X|$ et XY sont-elles indépendantes ?

d) On suppose que $\mathbf{P}(X = 0) = 0$. Soit Z la variable aléatoire telle que Z = 1 si X > 0 et Z = –1 si X < 0. La covariance de $|X|$ et de Z est-elle nulle ? $|X|$ et Z sont-elles indépendantes ?

Exo.3-4) Soit q élément de $]0, 1[$, et X, Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$ de même loi :

$$\forall\, n \geq 0,\ \mathbf{P}(X = n) = \mathbf{P}(Y = n) = q^n(1 - q)$$

Soit $n \geq 0$. Calculer la loi de X conditionnellement à l'événement $X \leq n \leq X + Y$.

Exo.3-5) Soient $X_1, ..., X_n$ des variables aléatoires indépendantes de même loi qu'une variable aléatoire X. On pose $m = \mathbf{E}(X)$, $\sigma^2 = V(X)$, $\alpha = \mathbf{E}((X - m)^3)$, $\overline{X} = \dfrac{X_1 + ... + X_n}{n}$, $T^2 = \dfrac{1}{n}\sum_{k=1}^{n}(X_k - m)^2$,

$S^2 = \dfrac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2$.

Calculer $\mathbf{E}(\overline{X})$, $\mathbf{E}(T^2)$, $\mathbf{E}(S^2)$, $\text{cov}(\overline{X}, T^2)$, $\text{cov}(\overline{X}, S^2)$ en fonction de m, σ^2 et α.

4- Solutions sur les probabilités

Sol.1-1) a) La probabilité que A gagne est la probabilité que Pile sorte la première fois à la position $1 + 3k$, $k \geq 0$. Donc :

$$\mathbf{P}(\text{A gagne}) = \sum_{k=0}^{\infty}\frac{1}{2^{1+3k}} = \frac{1}{2}\,\frac{1}{1-\frac{1}{8}} = \frac{4}{7}$$

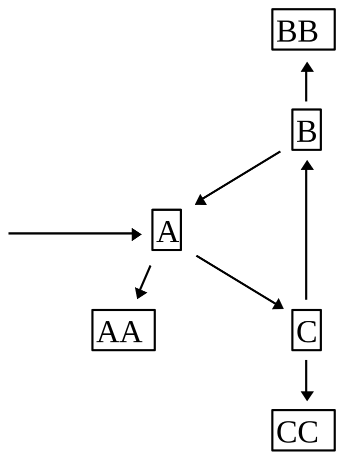
De même :

$$\mathbf{P}(\text{B gagne}) = \sum_{k=0}^{\infty}\frac{1}{2^{2+3k}} = \frac{1}{4}\,\frac{1}{1-\frac{1}{8}} = \frac{2}{7}$$

$$\mathbf{P}(\text{C gagne}) = \sum_{k=0}^{\infty}\frac{1}{2^{3+3k}} = \frac{1}{8}\,\frac{1}{1-\frac{1}{8}} = \frac{1}{7}$$

Le fait que la somme des trois probabilités soit égale à 1 montre que la probabilité que le jeu dure indéfiniment est nulle.

b) Dans le cas où A gagne la première partie contre B, le jeu se déroule selon le schéma suivant, où les lettres dans les cases désignent le vainqueur de la partie. Une case comprenant une double lettre désigne le vainqueur de deux parties successives. Chaque flèche (sauf la première qui pointe vers A et qui est supposée réalisée) a une probabilité $\frac{1}{2}$.



On calcule $\mathbf{P}(A_g \mid A_1)$ en sommant les probabilités de tous les chemins qui conduisent à la case AA, ce qui donne :

$$\mathbf{P}(A_g \mid A_1) = \sum_{n=0}^{\infty} \frac{1}{2^{1+3n}} = \frac{1}{2}\,\frac{1}{1-\frac{1}{8}} = \frac{4}{7}$$

De même :

$$\mathbf{P}(C_g \mid A_1) = \sum_{n=0}^{\infty} \frac{1}{2^{2+3n}} = \frac{1}{4}\,\frac{1}{1-\frac{1}{8}} = \frac{2}{7}$$

$$\mathbf{P}(B_g \mid A_1) = \sum_{n=0}^{\infty} \frac{1}{2^{3+3n}} = \frac{1}{8}\,\frac{1}{1-\frac{1}{8}} = \frac{1}{7}$$

Si le vainqueur de la première partie est B, on obtient les probabilités suivantes à partir des précédentes en échangeant les rôles de A et B :

$$\mathbf{P}(B_g \mid B_1) = \frac{4}{7}$$

$$\mathbf{P}(C_g \mid B_1) = \frac{2}{7}$$

$$\mathbf{P}(A_g \mid B_1) = \frac{1}{7}$$

On obtient les probabilités finales en appliquant la formule des probabilités totales :

$$\mathbf{P}(A_g) = \mathbf{P}(A_g \mid A_1) \times \mathbf{P}(A_1) + \mathbf{P}(A_g \mid B_1) \times \mathbf{P}(B_1) = \frac{4}{7} \times \frac{1}{2} + \frac{1}{7} \times \frac{1}{2} = \frac{5}{14}$$

$$\mathbf{P}(B_g) = \frac{5}{14}$$

$$\mathbf{P}(C_g) = \frac{2}{7} = \frac{4}{14}$$

Sol.1-2) La probabilité de tirer 6 points avec deux dés est $\frac{5}{36}$ et celle de tirer 7 est $\frac{6}{36}$.

A gagne au premier coup avec une probabilité $\frac{5}{36}$

A gagne au coup $4n$, $n \geq 1$, avec une probabilité $\frac{31}{36} \left(\frac{30}{36} \times \frac{30}{36} \times \frac{31}{36} \times \frac{31}{36}\right)^{n-1} \frac{30}{36} \times \frac{30}{36} \times \frac{5}{36}$

A gagne au coup $4n + 1$, $n \geq 1$, avec une probabilité $\frac{31}{36} \left(\frac{30}{36} \times \frac{30}{36} \times \frac{31}{36} \times \frac{31}{36}\right)^{n-1} \frac{30}{36} \times \frac{30}{36} \times \frac{31}{36} \times \frac{5}{36}$

La probabilité de gain de A est la somme de toutes ces probabilités, ce qui donne :

$$\frac{5}{36} + \sum_{n=1}^{\infty} \frac{31}{36} \left(\frac{25}{36} \times \frac{961}{1296}\right)^{n-1} \frac{25}{36} \times \frac{5}{36} + \sum_{n=1}^{\infty} \frac{31}{36} \left(\frac{25}{36} \times \frac{961}{1296}\right)^{n-1} \frac{25}{36} \times \frac{31}{36} \times \frac{5}{36}$$

$$= \frac{5}{36} + \frac{31}{36} \times \frac{25}{36} \times \frac{5}{36} \left(1 + \frac{31}{36}\right) \sum_{n=1}^{\infty} \left(\frac{25}{36} \times \frac{961}{1296}\right)^{n-1}$$

$$= \frac{5}{36} + \frac{31}{36} \times \frac{25}{36} \times \frac{5}{36} \left(1 + \frac{31}{36}\right) \frac{1}{1 - \frac{25}{36} \times \frac{961}{1296}} = \frac{10355}{22631}$$

B gagne au coup $4n + 2$, $n \geq 0$, avec une probabilité $\frac{31}{36} \times (\frac{30}{36} \times \frac{30}{36} \times \frac{31}{36} \times \frac{31}{36})^n \times \frac{6}{36}$

B gagne au coup $4n + 3$, $n \geq 0$, avec une probabilité $\frac{31}{36} \times (\frac{30}{36} \times \frac{30}{36} \times \frac{31}{36} \times \frac{31}{36})^n \times \frac{30}{36} \times \frac{6}{36}$

La probabilité de gain de B est la somme de toutes ces probabilités, ce qui donne :

$$\frac{31}{36} \times \frac{6}{36} \times (1 + \frac{30}{36}) \sum_{n=0}^{\infty} (\frac{25}{36} \times \frac{961}{1296})^n = \frac{31}{36} \times \frac{1}{6} \times \frac{11}{6} \times \frac{1}{1 - \frac{25}{36} \times \frac{961}{1296}} = \frac{12276}{22631}$$

On remarque que la somme des deux probabilités vaut 1. La partie s'arrête de façon presque certaine.

Sol.1-3) Soit p la probabilité de gain en partant de A et q en partant de B. En considérant la case sur laquelle on tombe au premier coup, la probabilité de gain en partant de A est égale à (formule des probabilités totales) :

probabilité de tomber en B au premier coup $\times$ probabilité de gain sachant qu'on repart de B
\+ probabilité de tomber en G au premier coup $\times$ 1
\+ probabilité de tomber en P au premier coup $\times$ 0
\+ probabilité de tomber en A au premier coup $\times$ probabilité de gain sachant qu'on repart de A

Or :

P(tomber en B au premier coup) = $\frac{2}{6}$

P(tomber en G au premier coup) = $\frac{2}{6}$

P(tomber en P au premier coup) = $\frac{1}{6}$

P(tomber en A au premier coup) = $\frac{1}{6}$

P(gain sachant qu'on part de B) = q
P(gain sachant qu'on part de A) = p

Donc la formule des probabilités totales donne $p = \frac{2q}{6} + \frac{2}{6} + \frac{p}{6}$

On applique un raisonnement comparable si on part de B, ce qui conduit au système :

$$\begin{cases} p = \frac{2q}{6} + \frac{2}{6} + \frac{p}{6} \\ q = \frac{2}{6} + \frac{p}{6} + \frac{q}{6} \end{cases}$$

$$\Leftrightarrow \begin{cases} 5p - 2q = 2 \\ -p + 5q = 2 \end{cases}$$

$$\Leftrightarrow \begin{cases} p = \frac{14}{23} \\ q = \frac{12}{23} \end{cases}$$

La probabilité de gagner est donc $p = \frac{14}{23}$.

Si on tient un raisonnement comparable pour trouver la probabilité de perdre, on trouvera $\frac{9}{23} = 1 - p$. Cela montre que la probabilité de jouer indéfiniment est nulle.

Sol.1-4) a) La probabilité que 6 sorte en un tirage est $\frac{5}{36}$. La probabilité que le 7 sorte est $\frac{6}{36}$. Comme les autres tirages n'interviennent pas, on peut les éliminer et se contenter des tirages sachant qu'on tire 6 ou 7. Les probabilités conditionnelles de tirer 6 (respectivement 7) sont alors $\frac{5}{11}$ et $\frac{6}{11}$.

La probabilité cherchée est donc $\frac{5}{11}$.

b) La probabilité de tirer un 8 est égale à celle de tirer 6, donc la réponse est identique.

c) On considère maintenant les probabilités conditionnelles sachant qu'on a tiré un 6, un 7 ou un 8. Les probabilités conditionnelles sont alors respectivement de $\frac{5}{16}, \frac{6}{16}, \frac{5}{16}$.

❑ Ou bien on tire d'abord un 6, et l'événement demandé est réalisé si on tire un 8 avant de tirer deux 7. En prenant en compte le tirage initial, cet événement correspond aux tirages 6^n8, $n \geq 1$, ou 6^n76^p8, $n \geq 1$, $p \geq 0$. Sa probabilité est :

$$\sum_{n=1}^{\infty} \left(\frac{5}{16}\right)^{n+1} + \sum_{n=1}^{\infty} \left(\frac{5}{16}\right)^{n} \frac{6}{16} \sum_{p=0}^{\infty} \left(\frac{5}{16}\right)^{p+1} = \frac{25}{256} \frac{1}{1 - \frac{5}{16}} + \frac{5}{16} \frac{1}{1 - \frac{5}{16}} \frac{6}{16} \frac{5}{16} \frac{1}{1 - \frac{5}{16}}$$

$$= \frac{25}{16 \times 11} + \frac{5}{11} \frac{6}{16} \frac{5}{11} = \frac{425}{1936}$$

❑ Ou bien on tire d'abord un 8, et la probabilité cherchée est celle de tirer un 6 avant de tirer deux 7. C'est la même que ci-dessus, en intervertissant 6 et 8.

❑ Ou bien on tire d'abord un 7. On ne doit plus alors en tirer. Les événements correspondants sont 76^n8 ou 78^n6, $n \geq 1$. La probabilité correspondante est :

$$2 \times \frac{6}{16} \sum_{n=1}^{\infty} \left(\frac{5}{16}\right)^{n+1} = 2 \times \frac{6}{16} \frac{25}{256} \frac{1}{1 - \frac{5}{16}} = 2 \times \frac{3}{8} \frac{25}{16 \times 11} = \frac{75}{704}$$

La probabilité totale est $2 \times \frac{425}{1936} + \frac{75}{704} = \frac{4225}{7744}$.

Curieusement, la probabilité de tirer 6 (ou 8) avant 7 est inférieure à $\frac{1}{2}$ (environ 0,44), mais la probabilité de tirer 6 et 8 avant le deuxième 7 est supérieure à $\frac{1}{2}$ (environ 0,55).

On peut aussi raisonner sur le diagramme suivant représentant l'état dans lequel on est :

I = état initial avant le premier lancer
7 = état où l'on a tiré un 7 avant de tirer 6 ou 8
X = état où l'on a tiré au moins un 6 ou un 8 mais pas les deux, et pas de 7
7X = état où l'on a tiré 7 et au moins un 6 ou un 8
E = échec (tirage du double 7 avant d'avoir tiré 6 et 8)
S = succès (tirage du 6 et 8 avant le deuxième 7)

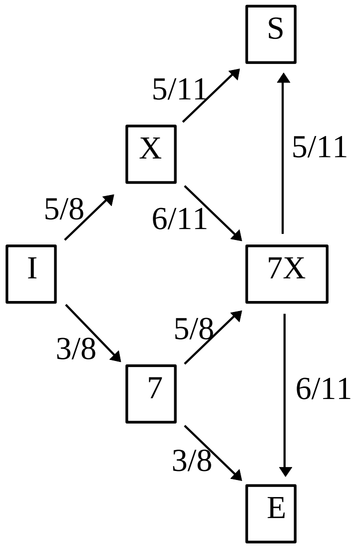
Les probabilités de transition d'un état à un autre sont indiquées sur les flèches entre les deux états. On effectue les calculs en ne tenant compte que des tirages donnant 6, 7 ou 8, les autres étant sans importance. Les probabilités de transition peuvent se calculer comme suit :

$\mathbf{P}(\mathrm{I} \to \mathrm{X}) = \mathbf{P}(6 \text{ ou } 8 \mid 6 \text{ ou } 7 \text{ ou } 8) = \dfrac{\mathrm{P}(6 \text{ ou } 8)}{\mathrm{P}(6 \text{ ou } 7 \text{ ou } 8)} = \dfrac{2 \times 5/11}{2 \times 5/11 + 6/11} = \dfrac{10}{16} = \dfrac{5}{8}$

$\mathbf{P}(\mathrm{X} \to \mathrm{S}) = \mathbf{P}(6 \mid 6 \text{ ou } 7)$ (en supposant par exemple que X = 8)

$= \dfrac{5}{11}$

etc.



On a alors, en suivant tous les chemins qui mènent de I à S :

$$\mathbf{P}(\mathrm{S}) = \frac{5}{8} \times \frac{5}{11} + \frac{5}{8} \times \frac{6}{11} \times \frac{5}{11} + \frac{3}{8} \times \frac{5}{8} \times \frac{5}{11} = \frac{4225}{7744}$$

Sol.1-5) Il s'agit de montrer que, si (A_n) est une famille d'événements disjoints, on a $\mathbf{P}(\bigcup_{n=1}^{\infty} \mathrm{A}_n) = \sum_{n=1}^{\infty} \mathbf{P}(\mathrm{A}_n)$. Pour cela, posons $\mathrm{S}_n = \bigcup_{k=1}^{n} \mathrm{A}_k$ et $\mathrm{R}_n = \bigcup_{k=n+1}^{\infty} \mathrm{A}_k$. (R_n) est une suite décroissante d'intersection vide, car si x est élément de Ω, ou bien x n'appartient à aucun A_k et donc à aucun R_n, ou bien x appartient à l'un des A_k, mais à un seul d'entre eux car les A_n sont disjoints, et alors x n'appartient pas à R_{k+1}. Donc :

$\mathbf{P}(\bigcup_{n=1}^{\infty} \mathrm{A}_n) = \mathbf{P}(\mathrm{S}_n \cup \mathrm{R}_n) = \mathbf{P}(\mathrm{S}_n) + \mathbf{P}(\mathrm{R}_n)$ car l'union finie est disjointe

$= \sum_{k=1}^{n} \mathbf{P}(\mathrm{A}_k) + \mathbf{P}(\mathrm{R}_n)$ idem

$= \sum_{k=1}^{\infty} \mathbf{P}(\mathrm{A}_k)$ en passant à la limite

Sol.1-6) a) Il vaut mieux passer au complémentaire. Posons $\mathrm{B}_n = \mathrm{A}_n^{\mathrm{c}}$:

$\forall\, n,\ \mathbf{P}(\mathrm{B}_n) \leq \dfrac{1}{2^n}$ et $\exists\, k,\ \mathbf{P}(\mathrm{B}_k) < \dfrac{1}{2^k}$

On a alors $\mathbf{P}(\bigcup_{n=1}^{\infty} B_n) \leq \sum_{n=1}^{\infty} \mathbf{P}(B_n) < \sum_{n=1}^{\infty} \frac{1}{2^n} = 1$, donc $\mathbf{P}(\bigcup_{n=1}^{\infty} B_n) < 1$.

b) On lance une pièce équilibrée indéfiniment. Soit X le nombre de lancers pour obtenir Pile. Posons $B_n = \{X = n\}$ et donc $A_n = \{X \neq n\}$. On a :

$$\mathbf{P}(B_n) = \frac{1}{2^n} \text{ et } \mathbf{P}(\bigcup_{n=1}^{\infty} B_n) = \{X \text{ fini}\} = 1.$$

$$\mathbf{P}(\bigcap_{n=1}^{\infty} A_n) = \{X \text{ infini}\} = 0$$

Sol.1-7) a) $\mathbf{P}(A_p) = \sum_{k=1}^{\infty} \mathbf{P}(\{pk\}) = \sum_{k=1}^{\infty} \frac{1}{\zeta(\alpha)p^\alpha k^\alpha} = \frac{1}{p^\alpha}$

b) Pour tout n, soient $p_1, p_2, ..., p_n$ premiers distincts.

$$\begin{aligned} A_{p_1} \cap ... \cap A_{p_n} &= \{m \in \mathbb{N}^* \mid m \text{ est un multiple de } p_1, \text{ de } p_2, ..., \text{ de } p_n\} \\ &= \{m \in \mathbb{N}^* \mid m \text{ est un multiple de } p_1...p_n\} \text{ car les } p_i \text{ sont premiers distincts} \\ &= A_{p_1...p_n} \end{aligned}$$

donc :

$$\mathbf{P}(A_{p_1} \cap ... \cap A_{p_n}) = \frac{1}{(p_1...p_n)^\alpha} = \mathbf{P}(A_{p_1}) ... \mathbf{P}(A_{p_n})$$

c) Vu l'indépendance vue en b), on a :

$$\lim_{n\to\infty} \mathbf{P}(\bigcap_{k=1}^{n} A_{p_k}{}^c) = \lim_{n\to\infty} \prod_{k=1}^{n} \mathbf{P}(A_{p_k}{}^c) = \lim_{n\to\infty} \prod_{k=1}^{n} (1 - \frac{1}{p_k{}^\alpha}) = \prod_{k=1}^{\infty} (1 - \frac{1}{p_k{}^\alpha})$$

Par ailleurs, la suite $(\bigcap_{k=1}^{n} A_{p_k}{}^c)_{n\in\mathbb{N}^*}$ est décroissante, d'intersection $\bigcap_{k=1}^{\infty} A_{p_k}{}^c$, donc :

$$\lim_{n\to\infty} \mathbf{P}(\bigcap_{k=1}^{n} A_{p_k}{}^c) = \mathbf{P}(\bigcap_{k=1}^{\infty} A_{p_k}{}^c) = \mathbf{P}((\bigcup_{k=1}^{\infty} A_{p_k})^c) = \mathbf{P}((\mathbb{N}^* - \{1\})^c) = \mathbf{P}(\{1\}) = \frac{1}{\zeta(\alpha)}$$

Sol.1-8) a) B_{im} est l'ensemble des suites de lancers qui sont périodiques de période m pendant i périodes. Les m premiers tirages sont quelconques. Les im tirages suivants sont imposés. Ceux au-delà sont quelconques.

$$\mathbf{P}(B_{im}) = \frac{1}{2^{mi}}$$

b) B_m est l'ensemble des suites périodiques de période m.

$$B_m = \bigcap_{i=1}^{\infty} B_{im} \text{ intersection décroissante car } B_{i+1,m} \subset B_{im}$$

donc $\mathbf{P}(B_m) = \lim_{i\to\infty} \mathbf{P}(B_{im}) = 0$

c) B est l'ensemble des suites périodiques.

$$B = \bigcup_{m=1}^{\infty} B_m$$

donc $\mathbf{P}(B) = \mathbf{P}(\bigcup_{m=1}^{\infty} B_m) \leq \sum_{m=1}^{\infty} \mathbf{P}(B_m) = 0$

Il y a une inégalité car l'union des B_m n'est pas disjointe.

On a prouvé qu'il y a une probabilité nulle de tirer une suite périodique.
On pourrait de même affiner le problème pour montrer qu'il y a une probabilité nulle de tirer une suite périodique à partir d'un certain rang.

Sol.1-9) a) v) $\Leftrightarrow$ i) $\Leftrightarrow \forall n, \exists p \geq n, \omega \in A_p \Leftrightarrow \forall n, \omega \in \bigcup_{p=n}^{\infty} A_p \Leftrightarrow \omega \in \bigcap_{n=0}^{\infty} \bigcup_{p=n}^{\infty} A_p \Leftrightarrow$ iv) $\Leftrightarrow$ vii)

et vi) $\Leftrightarrow$ ii) $\Leftrightarrow \exists n, \forall p \geq n, \omega \in A_p \Leftrightarrow \exists n, \omega \in \bigcap_{p=n}^{\infty} A_p \Leftrightarrow \omega \in \bigcup_{n=0}^{\infty} \bigcap_{p=n}^{\infty} A_p \Leftrightarrow$ iii) $\Leftrightarrow$ viii)

b) i) $\omega \in \mathrm{limsup}(A_n \cup B_n) \Leftrightarrow \omega$ appartient à une infinité de $A_n \cup B_n$
$\Leftrightarrow \omega$ appartient à une infinité de A_n
ou ω appartient à une infinité de B_n
$\Leftrightarrow \omega \in \mathrm{limsup}(A_n) \cup \mathrm{limsup}(B_n)$
Donc $\mathrm{limsup}(A_n \cup B_n) = \mathrm{limsup}(A_n) \cup \mathrm{limsup}(B_n)$
ii) $\omega \in \mathrm{limsup}(A_n \cap B_n) \Leftrightarrow \omega$ appartient à une infinité de $A_n \cap B_n$
$\Rightarrow \omega$ appartient à une infinité de A_n et ω appartient à une infinité de B_n
$\Leftrightarrow \omega \in \mathrm{limsup}(A_n) \cap \mathrm{limsup}(B_n)$
Donc $\mathrm{limsup}(A_n \cap B_n) \subset \mathrm{limsup}(A_n) \cap \mathrm{limsup}(B_n)$ mais la réciproque est fausse puisque, si ω appartient aux A_{2p} et pas aux A_{2p+1} et appartient aux B_{2p+1} mais pas aux B_{2p}, alors ω appartient à $\mathrm{limsup}(A_n) \cap \mathrm{limsup}(B_n)$ mais pas à $\mathrm{limsup}(A_n \cap B_n)$.
iii) $\omega \in \mathrm{liminf}(A_n \cup B_n) \Leftrightarrow \omega$ appartient à tous les $A_n \cup B_n$ sauf un nombre fini
$\Leftarrow \omega$ appartient à tous les A_n sauf un nombre fini
ou ω appartient à tous les B_n sauf un nombre fini
$\Leftrightarrow \omega \in \mathrm{liminf}(A_n) \cup \mathrm{liminf}(B_n)$.
Donc $\mathrm{liminf}(A_n) \cup \mathrm{liminf}(B_n) \subset \mathrm{liminf}(A_n \cup B_n)$.
La réciproque est fausse car si ω appartient aux A_{2p} et pas aux A_{2p+1} et appartient aux B_{2p+1} mais pas aux B_{2p}, alors ω appartient à tous les $A_n \cup B_n$ donc appartient à $\mathrm{liminf}(A_n \cup B_n)$ mais n'appartient ni à $\mathrm{liminf}(A_n)$ ni à $\mathrm{liminf}(B_n)$.
iv) $\omega \in \mathrm{liminf}(A_n \cap B_n) \Leftrightarrow \omega$ appartient à tous les $A_n \cap B_n$ sauf un nombre fini
$\Leftrightarrow \omega$ appartient à tous les A_n sauf un nombre fini
et ω appartient à tous les B_n sauf un nombre fini
$\Leftrightarrow \omega \in \mathrm{liminf}(A_n) \cap \mathrm{liminf}(B_n)$.
Donc $\mathrm{liminf}(A_n \cap B_n) = \mathrm{liminf}(A_n) \cap \mathrm{liminf}(B_n)$.
On peut aussi déduire iii) et iv) de i) et ii) puisque :

$$(\mathrm{liminf}(A_n \cup B_n))^c = \Big(\bigcup_{n=0}^{\infty} \bigcap_{p=n}^{\infty} (A_p \cup B_p)\Big)^c$$

$$= \bigcap_{n=0}^{\infty} \bigcup_{p=n}^{\infty} (A_p \cup B_p)^c$$

$$= \bigcap_{n=0}^{\infty} \bigcup_{p=n}^{\infty} (A_p^c \cap B_p^c)$$

$$= \mathrm{limsup}(A_n^c \cap B_n^c)$$

Or $\mathrm{limsup}(A_n^c \cap B_n^c) \subset \mathrm{limsup}(A_n^c) \cap \mathrm{limsup}(B_n^c)$
Donc :

$$(\mathrm{liminf}(A_n \cup B_n))^c \subset (\mathrm{liminf}(A_n))^c \cap (\mathrm{liminf}(B_n))^c = (\mathrm{liminf}(A_n) \cup \mathrm{liminf}(B_n))^c$$

Donc

$$\liminf(A_n) \cup \liminf(B_n) \subset \liminf(A_n \cup B_n)$$

On procède de même pour (iv)

c) $\mathbf{P}(\limsup(A_n)) = \mathbf{P}(\bigcap_{n=0}^{\infty} \bigcup_{p=n}^{\infty} A_p)$

$$= \lim_{n\to\infty} \mathbf{P}(\bigcup_{p=n}^{\infty} A_p) \qquad \text{car l'intersection est décroissante.}$$

$$\leq \lim_{n\to\infty} \sum_{p=n}^{\infty} \mathbf{P}(A_p) = 0$$

On peut aussi dire que $\sum_{n=0}^{\infty} \mathbf{P}(A_n) = \sum_{n=0}^{\infty} \mathbf{E}(\mathbf{1}_{A_n}) = \mathbf{E}(\sum_{n=0}^{\infty} \mathbf{1}_{A_n})$ (à condition d'appliquer un équivalent du théorème d'intégration terme à terme d'une série). Comme cette espérance est supposée finie, $\mathbf{P}(\sum_{n=0}^{\infty} \mathbf{1}_{A_n} = \infty) = 0$. On conclut en utilisant l'équivalence entre a.v) et a.iv).

d) $\mathbf{P}(\limsup(A_n)) = \lim_{n\to\infty} \mathbf{P}(\bigcup_{p=n}^{\infty} A_p)$ comme ci-dessus

$$= \lim_{n\to\infty} 1 - \mathbf{P}((\bigcup_{p=n}^{\infty} A_p)^c)$$

$$= \lim_{n\to\infty} 1 - \mathbf{P}(\bigcap_{p=n}^{\infty} A_p^c)$$

Or si les (A_n) sont indépendantes, il en est de même de leur complémentaire. On a alors :

$$\mathbf{P}(\bigcap_{p=n}^{\infty} A_p^c) = \lim_{m\to\infty} \mathbf{P}(\bigcap_{p=n}^{m} A_p^c) \text{car l'intersection est décroissante}$$

$$= \lim_{m\to\infty} \prod_{p=n}^{m} \mathbf{P}(A_p^c)$$

$$= \prod_{p=n}^{\infty} \mathbf{P}(A_p^c)$$

donc :

$$\mathbf{P}(\limsup(A_n)) = \lim_{n\to\infty} 1 - \prod_{p=n}^{\infty} \mathbf{P}(A_p^c)$$

$$= \lim_{n\to\infty} 1 - \prod_{p=n}^{\infty} (1 - \mathbf{P}(A_p))$$

Mais $\sum \mathbf{P}(A_n)$ diverge donc $\prod_{p=n}^{\infty} (1 - \mathbf{P}(A_p)) = 0$ (voir le paragraphe sur les produits infinis dans L2/SERIES.PDF. Donc :

$$\mathbf{P}(\limsup(A_n)) = 1$$

e) $\sum \mathbf{P}(A_n) = \sum \mathbf{P}(X_n \geq n) = \sum \mathbf{P}(X_1 \geq n)$ car toutes les lois sont identiques. Or cette dernière série converge, sa somme étant $\sum_{n=1}^{\infty} \mathbf{P}(X_1 \geq n) = \mathbf{E}(X_1)$. On applique c) pour conclure que $\mathbf{P}(\limsup(A_n)) = 0$, ce qui est la conclusion demandée.

f) Ici, la série diverge. En effet, pour tout K :

$$\sum_{n=1}^{K} \mathbf{P}(X_n \geq n) = \sum_{n=1}^{K} \mathbf{P}(X_1 \geq n) \qquad \text{car } X_n \text{ et } X_1 \text{ ont même loi}$$

$$= \sum_{n=1}^{K} \sum_{k=n}^{\infty} \mathbf{P}(X_1 = k)$$

$$= \sum_{1 \leq n \leq k,\ n \leq K} \mathbf{P}(X_1 = k)$$

$$= \sum_{k=0}^{K} k\, \mathbf{P}(X_1 = k) + \sum_{k=K+1}^{\infty} K\, \mathbf{P}(X_1 = k)$$

$$\geq \sum_{k=0}^{K} k\, \mathbf{P}(X = k) \to \mathbf{E}(X_1) = \infty \text{ quand K tend vers l'infini}$$

On applique d) pour conclure que $\mathbf{P}(\limsup(A_n)) = 1$.

5- Solutions sur les variables aléatoires

Sol.2-1) a) Pour une loi géométrique de paramètre p, on a $\mathbf{P}(X = n) = (1 - p)^{n-1}\, p$, et :

$$\mathbf{P}(X \text{ pair}) = \sum_{n=1}^{\infty} (1 - p)^{2n-1}\, p = \frac{p(1-p)}{1 - (1-p)^2} = \frac{1-p}{2-p} < \frac{1}{2}$$

$$\mathbf{P}(X \text{ impair}) = \sum_{n=0}^{\infty} (1 - p)^{2n}\, p = \frac{p}{1 - (1-p)^2} = \frac{1}{2-p} > \frac{1}{2}$$

b) Pour une loi de Poisson de paramètre λ :

$$\mathbf{P}(X \text{ pair}) = \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^{2n}}{(2n)!} = e^{-\lambda} \operatorname{ch}(\lambda) = \frac{1 + e^{-2\lambda}}{2} > \frac{1}{2}$$

$$\mathbf{P}(X \text{ impair}) = \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^{2n+1}}{(2n+1)!} = e^{-\lambda} \operatorname{sh}(\lambda) = \frac{1 - e^{-2\lambda}}{2} < \frac{1}{2}$$

c) Pour une loi binomiale, on a $\mathbf{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$. On sait déjà que :

$$\mathbf{P}(X \text{ pair}) + \mathbf{P}(X \text{ impair}) = 1$$

Par ailleurs :

$$\mathbf{P}(\text{X pair}) - \mathbf{P}(\text{X impair}) = \sum_{k\geq 0} \binom{n}{2k} p^{2k} (1-p)^{n-2k} - \sum_{k\geq 0} \binom{n}{2k+1} p^{2k+1} (1-p)^{n-2k-1}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (-1)^k p^k (1-p)^{n-k}$$

on reconnaît un binôme de Newton

$$= (-p + (1-p))^n$$
$$= (1-2p)^n$$

Donc, en faisant la demi-somme ou la demi-différence :

$$\mathbf{P}(\text{X pair}) = \frac{1 + (1-2p)^n}{2}$$

$$\mathbf{P}(\text{X impair}) = \frac{1 - (1-2p)^n}{2}$$

Donc :

$$\mathbf{P}(\text{X pair}) \geq \frac{1}{2} \Leftrightarrow n \text{ est pair ou } p \leq \frac{1}{2}$$

$$\mathbf{P}(\text{X impair}) > \frac{1}{2} \Leftrightarrow n \text{ est impair et } p > \frac{1}{2}$$

Sol.2-2) a) L'événement recherché est réalisé à la suite de parties ABABAB...ABAA = $(\text{AB})^n\text{AA}$, $n \geq 0$, ou bien BABA...BAA = $(\text{BA})^n\text{A}$, $n \geq 1$, où la suite de lettres désigne la séquence des vainqueurs successifs. On a :

$$\mathbf{P}((\text{AB})^n\text{AA}) = p^{n+2}(1-p)^n$$
$$\mathbf{P}((\text{BA})^n\text{A}) = p^{n+1}(1-p)^n$$

La probabilité que A gagne les deux parties de plus est :

$$\sum_{n=0}^{\infty} \mathbf{P}((\text{AB})^n\text{AA}) + \sum_{n=1}^{\infty} \mathbf{P}((\text{BA})^n\text{A}) = \frac{p^2}{1-p+p^2} + \frac{p^2(1-p)}{1-p+p^2} = \frac{p^2(2-p)}{1-p+p^2}$$

Quand p tend vers 0, elle est équivalente à $2p^2$, ou équivalente à $\mathbf{P}$(AA ou BAA). Si A est un joueur faible, ses chances de gains possibles sont concentrées sur les premiers jeux.

Quand p tend vers 1, elle tend vers 1. Plus précisément, si $p = 1 - h$ avec h tendant vers 0, on a :

$$\frac{p^2(2-p)}{1-p+p^2} = 1 - 2h^2 + o(h^2)$$

situation symétrique de la précédente, avec h probabilité que B gagne une partie. On peut d'ailleurs remarquer que, si $p = 1 - h$ avec $h \in]0, 1[$, $\frac{p^2(2-p)}{1-p+p^2} = 1 - \frac{h^2(2-h)}{1-h+h^2}$.

Si $p = \frac{1}{2}$, la probabilité de gain de deux parties de plus par A (ou B) est $\frac{1}{2}$. Plus intéressant, si $p = \frac{1}{2} + h$, elle vaut $\frac{1}{2} + \frac{5h}{3} + o(h)$ quand h tend vers 0. Pour deux joueurs quasiment de même force, le fait d'exiger le gain par deux parties de plus accentue l'écart entre les deux joueurs d'un facteur $\frac{5}{3}$.

C'est une règle analogue qui est adoptée au tennis, aussi bien pour gagner un jeu que pour gagner un match.

b) Soit X le nombre de parties jouées. Par un calcul analogue à ce qui précède :

$\forall n \geq 0$, $\mathbf{P}(X = 2n + 2) = p^{n+2}(1 - p)^n + (1 - p)^{n+2}p^n$
$\forall n \geq 1$, $\mathbf{P}(X = 2n + 1) = p^{n+1}(1 - p)^n + (1 - p)^{n+1}p^n$
On pourra vérifier que la somme de toutes ces probabilités fait 1, autrement dit, la probabilité que les parties soient menées indéfiniment est nulle.

$$\mathbf{E}(X) = \sum_{n=0}^{\infty} (2n + 2)(p^{n+2}(1 - p)^n + (1 - p)^{n+2}p^n) + \sum_{n=1}^{\infty} (2n + 1)(p^{n+1}(1 - p)^n + (1 - p)^{n+1}p^n)$$
$$= \dots = \frac{2 + p - p^2}{1 - p + p^2}$$

après un calcul assez fastidieux utilisant $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ et $\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}$ (voir L2/SERIENTR.PDF).
Il est plus simple de dire que, pour $n \geq 1$, $X > n$ est un événement correspondant aux deux suites de n gains successifs ABAB... ou BABA... alternativement, donc :
$\forall n \geq 1$, $\mathbf{P}(X > 2n) = 2p^n(1 - p)^n$
$\forall n \geq 0$, $\mathbf{P}(X > 2n + 1) = p^{n+1}(1 - p)^n + p^n(1 - p)^{n+1} = p^n(1 - p)^n$

donc $$\mathbf{E}(X) = \sum_{n=0}^{\infty} \mathbf{P}(X > n) = 1 + \sum_{n=1}^{\infty} \mathbf{P}(X > n) = 1 + \sum_{n=1}^{\infty} \mathbf{P}(X > 2n) + \sum_{n=0}^{\infty} \mathbf{P}(X > 2n + 1)$$

$$= 1 + \sum_{n=1}^{\infty} 2p^n(1 - p)^n + \sum_{n=0}^{\infty} p^n(1 - p)^n = 1 + \frac{2p(1-p)}{1 - p + p^2} + \frac{1}{1 - p + 2p^2}$$
$$= \frac{2 + p - p^2}{1 - p + p^2}$$
Si p tend vers 0 ou vers 1, $\mathbf{E}(X)$ tend vers 2, correspondant respectivement aux parties BB ou AA.

Si $p = \frac{1}{2}$, $\mathbf{E}(X) = 3 = 2 \sum_{n=2}^{\infty} \frac{n}{2^n}$.

Sol.2-3) a) Une carte étant retournée, la probabilité qu'une autre carte retournée ait la même valeur est $p = \frac{1}{2n-1}$. La loi du nombre de deuxièmes cartes à retourner pour obtenir la même valeur que la première carte suit alors une loi géométrique $\mathbf{P}(X_1 = m) = (1 - p)^{m-1}p$, $m \geq 1$. Son espérance est :
$\mathbf{E}(X_1) = \frac{1}{p} = 2n - 1$.
Il reste alors $2n - 2$ cartes sur la table. On itère le raisonnement.
Si on note X_k le nombre de paires à retourner pour trouver la k-ème paire de même valeur alors qu'il ne reste que $2(n - k + 1)$ cartes sur la table, X_k suit une loi géométrique de paramètre $\frac{1}{2(n-k+1)-1}$ de sorte que $\mathbf{E}(X_k) = 2(n - k + 1) - 1$. Remarquer que, pour la dernière paire, $k = n$ et l'on a $\mathbf{E}(X_n) = 1$, comme il se doit.
Enfin, le nombre total de paires à retourner est $X_1 + X_2 + \dots + X_n$ dont l'espérance est :

$$\mathbf{E}(X_1 + \dots + X_n) = \sum_{k=1}^{n} (2(n - k + 1) - 1) = \sum_{k=1}^{n} (2k - 1) = n^2$$

b) Une fois la même première carte retournée, on tire les $2n - 1$ autres cartes dans un certain ordre. La probabilité que la seconde carte de même valeur que la première soit celle tirée au m-ème tirage, $1 \leq m \leq 2n - 1$ est $\frac{1}{2n - 1}$. En effet, la loi donnant le rang m de cette deuxième carte est uniforme (parmi les $(2n - 1)!$ ordres de tirages possibles, il y en a $(2n - 2)!$ qui placent la carte recherchée au rang m). Donc l'espérance du nombre de tirage est $\mathbf{E}(Y_1) = \sum_{k=1}^{2n-1} \frac{m}{2n - 1} = n$.

S'il ne reste que $2k$ cartes, l'espérance pour tirer un couple sera k.

Donc l'espérance du nombre total de paires à retourner est $\sum_{k=1}^{n} k = \frac{n(n + 1)}{2}$, soit moitié moins environ que dans le cas du a).

c) On retourne chaque carte au moins une fois pour connaître sa valeur, et jamais plus de deux fois. En effet, on connaît la valeur d'une carte retournée une fois, et une fois sa valeur connue, elle ne pourra être retournée une seconde fois que pour être appariée à une autre carte de même valeur retournée juste avant. Donc le nombre de paires retournées est au plus de $2n$.

Le calcul exact de l'espérance est un problème difficile. On peut montrer[5] que l'espérance s'écrit $(3 - 2\ln(2))n + \frac{7}{8} - 2\ln(2) + o(1)$.

Sol.2-4) a) $\mathbf{E}(X) = \lambda$

$$\mathbf{E}(Y) = \mathbf{E}(\exp(-\alpha X)) = \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} e^{-\alpha n} = \exp(-\lambda + \lambda e^{-\alpha})$$

b) Si $Y = g(X)$, on souhaite que :

$$\mathbf{E}(g(X)) = \sum_{n=0}^{\infty} g(n)\, e^{-\lambda} \frac{\lambda^n}{n!} = e^{-\alpha\lambda}$$

donc $\displaystyle \sum_{n=0}^{\infty} g(n) \frac{\lambda^n}{n!} = e^{\lambda - \alpha\lambda} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} (1 - \alpha)^n$

Pour $\alpha \neq 1$, il suffit de prendre $g(n) = (1 - \alpha)^n$ pour tout n, ce qui donne $Y = g(X) = (1 - \alpha)^X$.

Pour $\alpha = 1$, on peut prendre $g(n) = \left[\begin{array}{l} 1 \text{ si } n = 0 \\ 0 \text{ sinon} \end{array}\right.$ et donc $Y = g(X) = \mathbf{1}_{X=0}$.

Sol.2-5) a) $G_T(z) = \mathbf{E}(z^T) = \sum_{n=0}^{\infty} \mathbf{E}(z^T \mid U = n)\, \mathbf{P}(U = n)$

$$= \mathbf{E}(z^T \mid U = 0)\, \mathbf{P}(U = 0) + \sum_{n=1}^{\infty} \mathbf{E}(z^{X_1 + \ldots + X_n} \mid U = n)\, \mathbf{P}(U = n)$$

[5] Daniel J. Velleman, Gregory S.Warrington, *What to expect in a game of memory*, 120:9, Amer. Math. Monthly (novembre 2013), 787-805.

$= \mathbf{P}(U = 0) + \sum_{n=1}^{\infty} \mathbf{E}(z^{X_1+...+X_n})\ \mathbf{P}(U = n)$

car $z^T = z^0 = 1$ d'espérance 1 si U = 0, et dans la série,
les variables aléatoires $X_1 + ... + X_n$ et U sont indépendantes

$= \mathbf{P}(U = 0) + \sum_{n=1}^{\infty} \mathbf{E}(z^{X_1} z^{X_2} ... z^{X_n})\ \mathbf{P}(U = n)$

$= \mathbf{P}(U = 0) + \sum_{n=1}^{\infty} \mathbf{E}(z^{X_1})\mathbf{E}(z^{X_2})...\mathbf{E}(z^{X_n})\ \mathbf{P}(U = n)$

car les variables aléatoires $X_1, ..., X_n$ sont indépendantes

$= \mathbf{P}(U = 0) + \sum_{n=1}^{\infty} \mathbf{E}(z^X)^n\ \mathbf{P}(U = n)$

où $\mathbf{E}(z^X)$ est l'espérance commune des $\mathbf{E}(z^{X_i})$

$= \mathbf{P}(U = 0) + \sum_{n=1}^{\infty} G_X(z)^n\ \mathbf{P}(U = n)$

par définition de la fonction génératrice G_X

$= \sum_{n=0}^{\infty} G_X(z)^n\ \mathbf{P}(U = n)$
$= \mathbf{E}(G_X(z)^U)$
$= G_U(G_X(z))$

b) $\mathbf{E}(T) = G_T'(1) = G_U'(G_X(1))\ G_X'(1) = G_U'(1)\ G_X'(1) = \mathbf{E}(U)\mathbf{E}(X)$
résultat qui devrait paraître naturel.
Puis :

$G_T''(z) = G_U''(G_X(z))\ G_X'^2(z) + G_U'(G_X(z))\ G_X''(z)$

donc :

$G_T''(1) = G_U''(1)\ G_X'^2(1) + G_U'(1)\ G_X''(1)$
$= (\mathbf{V}(U) - \mathbf{E}(U) + \mathbf{E}(U)^2)\ \mathbf{E}(X)^2 + \mathbf{E}(U)\ (\mathbf{V}(X) - \mathbf{E}(X) + \mathbf{E}(X)^2)$
$= (\mathbf{V}(U) + \mathbf{E}(U)^2)\ \mathbf{E}(X)^2 + \mathbf{E}(U)\ (\mathbf{V}(X) - \mathbf{E}(X))$

donc :

$\mathbf{V}(T) = G_T''(1) + G_T'(1) - G_T'(1)^2$
$= (\mathbf{V}(U) + \mathbf{E}(U)^2)\ \mathbf{E}(X)^2 + \mathbf{E}(U)\ (\mathbf{V}(X) - \mathbf{E}(X)) + \mathbf{E}(U)\ \mathbf{E}(X) - \mathbf{E}(U)^2\ \mathbf{E}(X)^2$
$= \mathbf{V}(U)\mathbf{E}(X)^2 + \mathbf{E}(U)\mathbf{V}(X)$

résultat plus difficile à interpréter que celui de l'espérance. Si U était une variable certaine égale à m, T serait la somme de m variables aléatoires indépendantes de même loi et sa variance serait $m\mathbf{V}(X)$. On peut ainsi comprendre la présence de $\mathbf{E}(U)\mathbf{V}(X)$ due à la dispersion de X. Inversement, si c'était les X_i qui étaient des variables certaines égale à x, T serait égale à xU et sa variance serait $x^2\mathbf{V}(U)$. La quantité $\mathbf{V}(U)\mathbf{E}(X)^2$ s'interprète alors comme une contribution à $\mathbf{V}(T)$ due à la dispersion de U.
c) On définit les X_i comme étant des variables aléatoires de Bernoulli indépendantes de paramètre p : $X_i = 1$ si on tire Pile au i-ème tirage lors des U lancers de pièces. On a alors $T = \sum_{i=1}^{U} X_i$ et on peut appliquer la formule de Wald :

$$\mathbf{E}(T) = \mathbf{E}(U)\mathbf{E}(X) = \frac{1}{p} \times p = 1$$

Intuitivement, on lance en moyenne la pièce $\frac{1}{p}$ fois, puis pour chacun de ces lancers, on a une probabilité p de tirer Pile, donc le nombre moyen de Piles est 1.

On peut aussi dire que, conditionnellement à U = n, T suit une loi binomiale $\mathcal{B}(n, p)$, donc :

$$\mathbf{E}(T \mid U = n) = np$$

donc $$\mathbf{E}(T) = \sum_{n=0}^{\infty} \mathbf{E}(T \mid U = n)\, \mathbf{P}(U = n) = \sum_{n=0}^{\infty} np\mathbf{P}(U = n) = p\, \mathbf{E}(U) = \frac{p}{p} = 1$$

Calculons maintenant la loi de T, au moyen de sa fonction génératrice :

$$G_T(z) = (G_U \circ G_X)(z) = G_U(1 - p + pz)$$

or $G_U(z) = \dfrac{pz}{1 - (1 - p)z}$ donc :

$$G_T(z) = \frac{p(1 - p + pz)}{1 - (1 - p)(1 - p + pz)} = \frac{p - p^2 + p^2 z}{1 - 1 + 2p - p^2 - p(1 - p)z}$$

$$= \frac{1 - p + pz}{2 - p - (1 - p)z} = \frac{1 - p + pz}{2 - p} \frac{1}{1 - \dfrac{(1 - p)z}{2 - p}}$$

$$= \frac{1 - p + pz}{2 - p} \sum_{n=0}^{\infty} \frac{(1 - p)^n z^n}{(2 - p)^n} = \sum_{n=0}^{\infty} \frac{(1 - p)^{n+1} z^n}{(2 - p)^{n+1}} + \sum_{n=0}^{\infty} \frac{p(1 - p)^n z^{n+1}}{(2 - p)^{n+1}}$$

$$= \sum_{n=0}^{\infty} \frac{(1 - p)^{n+1} z^n}{(2 - p)^{n+1}} + \sum_{n=1}^{\infty} \frac{p(1 - p)^{n-1} z^n}{(2 - p)^n}$$

donc $$\mathbf{P}(T = 0) = \frac{1 - p}{2 - p}$$

et $$\forall\, n \geq 1,\ \mathbf{P}(T = n) = \frac{(1 - p)^{n+1}}{(2 - p)^{n+1}} + \frac{p(1 - p)^{n-1}}{(2 - p)^n} = \frac{(1 - p)^{n+1} + (2 - p)p(1 - p)^{n-1}}{(2 - p)^{n+1}}$$

$$= \frac{(1 - p)^{n-1}}{(2 - p)^{n+1}}$$

On pourra vérifier que $\sum_{n=0}^{\infty} \mathbf{P}(T = n) = 1$.

On peut aussi calculer la loi de T par la formule des probabilités totales, en utilisant le fait que la loi de T sachant que U = m est une loi binomiale $\mathcal{B}(m, p)$.

Pour $n = 0$:

$$\mathbf{P}(T = 0) = \sum_{m=1}^{\infty} \mathbf{P}(T = 0 \mid U = m)\mathbf{P}(U = m) = \sum_{m=1}^{\infty} (1 - p)^m\, (1 - p)^{m-1} p$$

$$= p \sum_{m=1}^{\infty} (1 - p)^{2m-1}$$

$$= \frac{p(1 - p)}{1 - (1 - p)^2} = \frac{p(1 - p)}{2p - p^2} = \frac{1 - p}{2 - p}$$

Pour $m \geq 1$:

$$\mathbf{P}(T = n) = \sum_{m=n}^{\infty} \mathbf{P}(T = n \mid U = m)\mathbf{P}(U = m) = \sum_{m=n}^{\infty} \binom{m}{n} p^n(1-p)^{m-n}\,(1-p)^{m-1}p$$

$$= \sum_{m=n}^{\infty} \binom{m}{n} p^{n+1}(1-p)^{2m-n-1}$$

$$= \frac{p^{n+1}}{(1-p)^{n+1}} \sum_{m=n}^{\infty} \binom{m}{n} (1-p)^{2m}$$

$$= \frac{p^{n+1}}{(1-p)^{n+1}} \frac{(1-p)^{2n}}{(1-(1-p)^2)^{n+1}}$$

car $\forall\, k \geq 0$, $\forall\, x \in\,]-1, 1[$, $\dfrac{x^k}{(1-x)^{k+1}} = \sum_{n=k}^{\infty} \binom{n}{k} x^n$

voir L2/SERIENTR.PDF

$$= \frac{p^{n+1}(1-p)^{n-1}}{(2p-p^2)^{n+1}} = \frac{(1-p)^{n-1}}{(2-p)^{n+1}}$$

Sol.2-6) a) $\mathbf{P}(X = 1) = 0$, $\mathbf{P}(X = 2) = p^2$, $\mathbf{P}(X = 3) = (1-p)p^2$ (tirage FPP seul possible).
b) $X = n$ se produit dans deux cas :

ou bien le premier tirage est Face, puis on obtient PP au bout de $X' = n - 1$ tirages, X' étant le nombre de tirages entre le deuxième tirage et la première fois qu'on atteint PP.

ou bien le premier tirage est Pile, mais il est alors nécessairement suivi de Face, puis on obtient PP au bout de $X'' = n - 2$ tirages, X'' étant le nombre de tirages entre le troisième tirage et la première fois qu'on atteint PP.

On dispose du système complet d'événements concernant les premiers tirages {F}, {PP}, {PF} et l'on a alors :

$$\mathbf{P}(X = n) = \mathbf{P}(\{F\})\mathbf{P}(X = n \mid \{F\}) + \mathbf{P}(\{PP\})\mathbf{P}(X = n \mid \{PP\}) + \mathbf{P}(\{PF\})\mathbf{P}(X = n \mid \{PF\})$$

où :

$\mathbf{P}(X = n \mid \{PP\}) = 0$ si $n \geq 3$

$\mathbf{P}(X = n \mid \{F\}) = \mathbf{P}(X' = n - 1 \mid \{F\}))$

$= \mathbf{P}(X' = n - 1)$ du fait de l'indépendance des $n - 1$ derniers tirages par rapport au premier tirage F

$= \mathbf{P}(X = n - 1)$ car X et X' ont même loi

et de même :

$\mathbf{P}(X = n \mid \{PF\}) = \mathbf{P}(X'' = n - 2) = \mathbf{P}(X = n - 2)$

Donc :

$$\mathbf{P}(X = n) = (1-p)\,\mathbf{P}(X = n - 1) + p(1-p)\,\mathbf{P}(X = n - 2)$$

c) $\mathbf{E}(X) = \sum_{n=1}^{\infty} n\,\mathbf{P}(X = n) = 2p^2 + \sum_{n=3}^{\infty} n\,\mathbf{P}(X = n)$

$$= 2p^2 + \sum_{n=3}^{\infty} n\,((1-p)\mathbf{P}(X = n-1) + p(1-p)\mathbf{P}(X = n-2))$$

$$= 2p^2 + (1-p) \sum_{n=3}^{\infty} n\,\mathbf{P}(X = n-1) + p(1-p) \sum_{n=3}^{\infty} n\,\mathbf{P}(X = n-2)$$

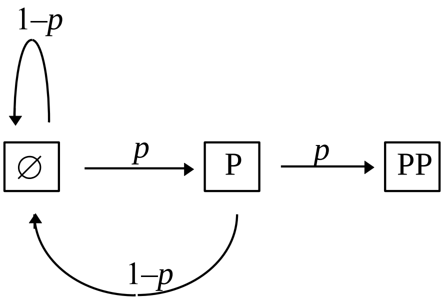
$$= 2p^2 + (1 - p) \sum_{n=2}^{\infty} (n + 1)\, \mathbf{P}(X = n) + p(1 - p) \sum_{n=1}^{\infty} (n + 2)\, \mathbf{P}(X = n)$$

$$= 2p^2 + (1 - p)(\mathbf{E}(X) + 1) + p(1 - p)(\mathbf{E}(X) + 2)$$
$$= 2p^2 + 1 - p + 2p(1 - p) + (1 - p)\mathbf{E}(X) + p(1 - p)\mathbf{E}(X)$$
$$= 1 + p + (1 - p^2)\mathbf{E}(X)$$

donc $\mathbf{E}(X) = \dfrac{1 + p}{p^2}$

REMARQUES :

❑ On peut retrouver graphiquement la valeur de cette espérance de la façon suivante. Considérons le graphe suivant :



Les cases ∅, P, PP représentent une situation dans laquelle on peut se trouver.

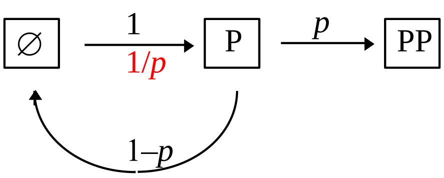
La case ∅ est la situation initiale, avant le premier lancer, ou bien une situation à l'issue de tirages sans tirages PP et dont le dernier est F.

La case P correspond à la situation correspondant à un premier tirage ayant donné un P, ou bien à la situation à l'issue de tirages sans tirages PP et dont les deux derniers sont dans l'ordre FP.

La case PP correspond à la situation finale correspondant aux tirages se terminant par les deux tirages PP pour la première fois.

Au lieu de situations, on parle aussi d'états du système. Chaque flèche représente un tirage possible de Pile ou de Face, et on indique pour chacune d'elle la probabilité de passer d'un état à un autre. Nous avons ici un exemple de chaîne de Markov, évoquée en annexe du chapitre.

Le nombre de tirages pour passer de l'état ∅ à l'état P suit une loi géométrique de paramètre p, et il est presque certain que cette transition aura lieu. Modifions le graphique pour mettre en évidence ce fait, en indiquant en rouge l'espérance $\dfrac{1}{p}$ du nombre de tirages pour effectuer cette transition, ce qui permet de supprimer la boucle $\varnothing \overset{1-p}{\rightarrow} \varnothing$.



Considérons la transition $P \overset{1-p}{\rightarrow} \varnothing$. Puisqu'elle est suivie de la transition certaine $\varnothing \overset{1}{\rightarrow} P$, on peut réduire ces deux transitions en la seule transition $P \overset{1-p}{\rightarrow} P$, avec une espérance du nombre de tirages pour y parvenir de $1 + \dfrac{1}{p}$:

$$\begin{tikzcd} \boxed{\varnothing} \arrow[r, "1", "1/p"'] & \boxed{\mathrm{P}} \arrow[r, "p"] \arrow[loop below, "1-p", "1+\frac{1}{p}"'] & \boxed{\mathrm{PP}} \end{tikzcd}$$

Enfin, le nombre de flèches à suivre pour passer de l'état P à l'état PP suit une loi géométrique de paramètre p : on boucle un certain nombre de fois la flèche P $\to$ P avant un dernier tirage P $\to$ PP. On passera donc de l'état P à l'état PP avec un nombre de flèches dont l'espérance est $\frac{1}{p}$, à savoir en moyenne $\frac{1}{p} - 1$ flèches bouclant sur l'état P avant d'appliquer une dernière flèche P $\to$ PP. Comme chaque flèche P $\to$ P correspond ici à un nombre moyen de $1 + \frac{1}{p}$ tirages, le nombre moyen de tirages pour passer de l'état P à l'état final PP est $(\frac{1}{p} - 1) \times (1 + \frac{1}{p}) + 1$. Une justification rigoureuse de ceci repose sur la formule de Wald vu dans l'exercice précédent. On a en effet $\mathrm{T} = \sum_{i=1}^{\mathrm{U}} \mathrm{Y}_i + 1$, où T est le nombre total de tirages pour passer de l'état P à l'état PP, U le nombre de fois où l'on boucle sur une transition P $\to$ P, Y_i le nombre de tirages lors de la i-ème transition P $\to$ P. Le 1 final correspond au dernier tirage P qui fait passer de l'état P à l'état PP. U + 1 suit une loi géométrique de paramètre p et chaque Y_i a une espérance égale à $\mathbf{E}(\mathrm{Y}) = 1 + \frac{1}{p}$. La formule de Wald donne :

$$\mathbf{E}(\mathrm{T}) = \mathbf{E}(\mathrm{Y})\mathbf{E}(\mathrm{U}) + 1 = (\frac{1}{p} - 1) \times (1 + \frac{1}{p}) + 1$$

$$\begin{tikzcd} \boxed{\varnothing} \arrow[r, "1", "1/p"'] & \boxed{\mathrm{P}} \arrow[r, "1", "(\frac{1}{p} - 1) \times (1 + \frac{1}{p}) + 1"'] & \boxed{\mathrm{PP}} \end{tikzcd}$$

Le nombre moyen de tirages pour passer de l'état $\varnothing$ à l'état PP est donc :

$$\mathbf{E}(\mathrm{X}) = \frac{1}{p} + (\frac{1}{p} - 1) \times (1 + \frac{1}{p}) + 1 = \frac{1}{p} + \frac{1}{p^2} = \frac{p+1}{p^2}$$

comme on l'a trouvé précédemment.

$$\begin{tikzcd} \boxed{\varnothing} \arrow[r, "1", "\frac{1+p}{p^2}"'] & \boxed{\mathrm{PP}} \end{tikzcd}$$

Pour $p = \frac{1}{2}$, $\frac{p+1}{p^2} = 6$, valeur donnée dans l'annexe sur les chaînes de Markov, pour le nombre moyen de tirages à effectuer avec une pièce équilibrée pour d'atteindre FF (au lieu de PP).

❑ On peut encore retrouver la valeur de $\mathbf{E}(\mathrm{X})$ de la façon suivante. Soit Y le nombre de tirages pour arriver au premier Pile, A l'événement "le tirage qui suit est un Pile" et B l'événement qui suit est un face". Y suit une loi géométrique de paramètre $\frac{1}{p}$. On a alors :

$$X = Y + \mathbf{1}_A + \mathbf{1}_B(1 + X') = Y + 1 + \mathbf{1}_B X'$$

où X' suit la même loi que X. En effet, si A est réalisé, on arrive directement en PP avec X = Y + 1 tirages, et si c'est B qui est réalisé, on se retrouve à l'état initial ∅ après Y + 1 tirages, et il faut X' tirages pour aller de ∅ à PP, X' ayant la même loi que X. Remarquons que, du fait de l'indépendance des tirages, X' est une variable aléatoire indépendante de $\mathbf{1}_B$. Donc :

$$\mathbf{E}(X) = \mathbf{E}(Y) + 1 + \mathbf{E}(\mathbf{1}_B X') = \mathbf{E}(Y) + 1 + \mathbf{E}(\mathbf{1}_B)\mathbf{E}(X') = \frac{1}{p} + 1 + (1-p)\mathbf{E}(X)$$

donc $\mathbf{E}(X) = \dfrac{1+p}{p^2}$

d) On a maintenant $\mathbf{P}(X = 1) = \mathbf{P}(X = 2) = 0$, $\mathbf{P}(X = 3) = p^3$, et on procèdera comme dans le b) et le c). En conditionnant par les événements correspondant aux premiers tirages F ou PF ou PPF, la formule des probabilités totales donne :

$$\forall\, n \geq 4,\ \mathbf{P}(X = n) = (1-p)\mathbf{P}(X = n-1) + p(1-p)\mathbf{P}(X = n-2) + p^2(1-p)\mathbf{P}(X = n-3)$$

donc :

$$\mathbf{E}(X) = \sum_{n=1}^{\infty} n\,\mathbf{P}(X = n) = 3p^3 + \sum_{n=4}^{\infty} n\,\mathbf{P}(X = n)$$

$$= 3p^3 + \sum_{n=4}^{\infty} n\,((1-p)\mathbf{P}(X = n-1) + p(1-p)\mathbf{P}(X = n-2) + p^2(1-p)\mathbf{P}(X = n-3))$$

$$= 3p^3 + (1-p)\sum_{n=4}^{\infty} n\,\mathbf{P}(X = n-1) + p(1-p)\sum_{n=4}^{\infty} n\,\mathbf{P}(X = n-2) + p^2(1-p)\sum_{n=4}^{\infty} n\,\mathbf{P}(X = n-3)$$

$$= 3p^3 + (1-p)\sum_{n=3}^{\infty} (n+1)\,\mathbf{P}(X = n) + p(1-p)\sum_{n=2}^{\infty} (n+2)\,\mathbf{P}(X = n) + p^2(1-p)\sum_{n=1}^{\infty} (n+3)\,\mathbf{P}(X = n)$$

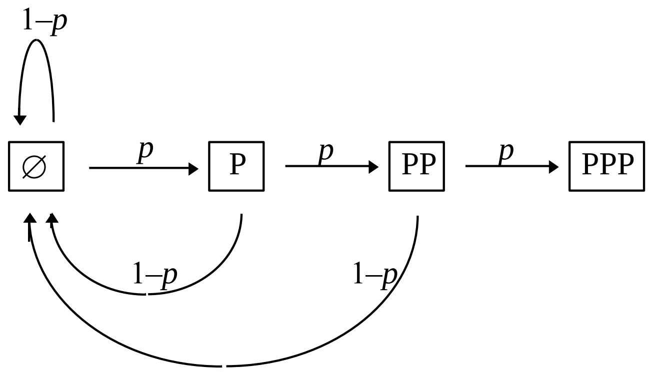
$$= 3p^3 + (1-p)(\mathbf{E}(X) + 1) + p(1-p)(\mathbf{E}(X) + 2) + p^2(1-p)(\mathbf{E}(X) + 3)$$
$$= 1 + p + p^2 + (1-p^3)\mathbf{E}(X)$$

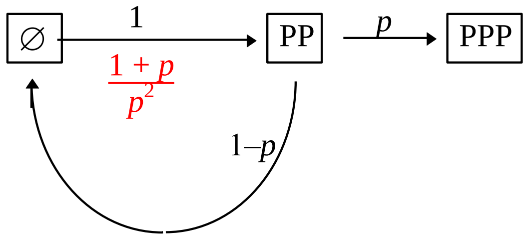
donc $\quad \mathbf{E}(X) = \dfrac{1+p+p^2}{p^3}$

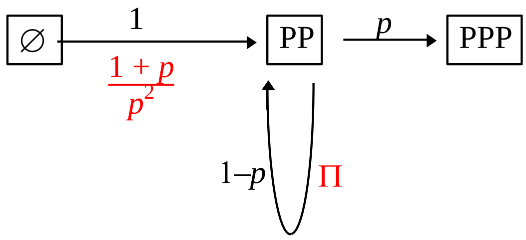
REMARQUES :

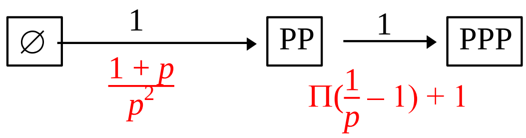
❑ Le lecteur est invité à comprendre la résolution graphique suivante, où l'on a posé :

$$\Pi = 1 + \frac{1+p}{p^2} = \frac{1+p+p^2}{p^2}$$

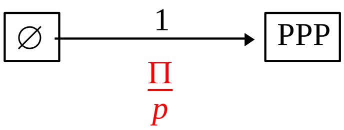








On a $\mathbf{E}(\mathrm{X}) = \dfrac{1+p}{p^2} + \Pi(\dfrac{1}{p} - 1) + 1 = \dfrac{\Pi}{p} = \dfrac{1+p+p^2}{p^3}$



❑ On a aussi :

$$\mathrm{X} = \mathrm{Y} + 1 + \mathbf{1}_{\mathrm{B}}\mathrm{X}'$$

où Y est le nombre de tirages pour passer de l'état ∅ à l'état PP pour la première fois, B l'événement "le tirage qui suit est Face", et X' une variable aléatoire de même loi que X. $\mathbf{E}(\mathrm{Y}) = \dfrac{1+p}{p^2}$ d'après le c). Donc :

$$\mathbf{E}(\mathrm{X}) = \mathbf{E}(\mathrm{Y}) + 1 + \mathbf{E}(\mathbf{1}_{\mathrm{B}}\mathrm{X}') = \mathbf{E}(\mathrm{Y}) + 1 + \mathbf{E}(\mathbf{1}_{\mathrm{B}})\mathbf{E}(\mathrm{X}')$$
$$= \frac{1+p}{p^2} + 1 + (1-p)\mathbf{E}(\mathrm{X})$$

donc $\mathbf{E}(\mathrm{X}) = \dfrac{1+p+p^2}{p^3}$

e) On a maintenant $\mathbf{P}(\mathrm{X} = 1) = \mathbf{P}(\mathrm{X} = 2) = ... = \mathbf{P}(\mathrm{X} = a-1) = 0$, $\mathbf{P}(\mathrm{X} = a) = p^a$.

Pour $n \geq a + 1$:

$$\mathbf{P}(\mathrm{X} = n) = (1-p)\mathbf{P}(\mathrm{X} = n-1) + p(1-p)\mathbf{P}(\mathrm{X} = n-2) + ... + p^{a-1}(1-p)\mathbf{P}(\mathrm{X} = n-a)$$

donc :

$$\mathbf{E}(\mathrm{X}) = \sum_{n=1}^{\infty} n\,\mathbf{P}(\mathrm{X} = n) = ap^a + \sum_{n=a+1}^{\infty} n\,\mathbf{P}(\mathrm{X} = n)$$

$$= ap^a + \sum_{n=a+1}^{\infty} n\,((1-p)\mathbf{P}(\mathrm{X} = n-1) + p(1-p)\mathbf{P}(\mathrm{X} = n-2) + ... + p^{a-1}(1-p)\mathbf{P}(\mathrm{X} = n-a))$$

$= ap^a + (1-p) \sum_{n=a+1}^{\infty} n\,\mathbf{P}(X = n-1) + p(1-p) \sum_{n=a+1}^{\infty} n\,\mathbf{P}(X = n-2) + ... + p^{a-1}(1-p) \sum_{n=a+1}^{\infty} n\,\mathbf{P}(X = n-a)$

$= ap^a + (1-p)\sum_{n=a}^{\infty} (n+1)\,\mathbf{P}(X = n) + p(1-p) \sum_{n=a-1}^{\infty} (n+2)\,\mathbf{P}(X = n) + ... + p^{a-1}(1-p) \sum_{n=1}^{\infty} (n+a)\,\mathbf{P}(X = n)$

$= ap^a + (1-p)(\mathbf{E}(X) + 1) + p(1-p)(\mathbf{E}(X) + 2) + ... + p^{a-1}(1-p)(\mathbf{E}(X) + a)$

$= 1 + p + ... + p^{a-1} + (1-p^a)\mathbf{E}(X)$

donc $\quad \mathbf{E}(X) = \dfrac{1 + p + ... + p^{a-1}}{p^a} = \dfrac{1-p^a}{p^a(1-p)}$

REMARQUES :

❑ On pourra adapter la résolution graphique du d) pour obtenir une résolution graphique du e).

❑ On peut aussi procéder par récurrence. On a :

$$X = Y + 1 + \mathbf{1}_B X'$$

où Y est le nombre de tirages pour passer de l'état ∅ à l'état P^{a-1} (suite de $a - 1$ Piles) pour la première fois, B l'événement "le tirage qui suit est Face", et X' une variable aléatoire de même loi que X. $\mathbf{E}(Y) = \dfrac{1 + p + ... + p^{a-2}}{p^{a-1}}$ d'après l'hypothèse de récurrence, et :

$$\mathbf{E}(X) = \mathbf{E}(Y) + 1 + \mathbf{E}(\mathbf{1}_B X') = \mathbf{E}(Y) + 1 + \mathbf{E}(\mathbf{1}_B)\mathbf{E}(X')$$

$$= \frac{1 + p + ... + p^{a-1}}{p^{a-1}} + (1-p)\mathbf{E}(X)$$

d'où $\quad \mathbf{E}(X) = \dfrac{1 + p + ... + p^{a-1}}{p^a}$

❑ on peut aussi chercher la fonction génératrice G de X à partir de la relation de récurrence des $\mathbf{P}(X = n)$. On trouve que :

$$G(z) = p^a z^a + \sum_{n=a+1}^{\infty} \mathbf{P}(X = n)\, z^n$$

$$= p^a z^a + \sum_{n=a+1}^{\infty} ((1-p)\mathbf{P}(X = n-1) + p(1-p)\mathbf{P}(X = n-2) + ... + p^{a-1}(1-p)\mathbf{P}(X = n-a))\, z^n$$

$$= p^a z^a + ((1-p)z + p(1-p)z^2 + ... + p^{a-1}(1-p)z^a)\, G(z)$$

$$= p^a z^a + (1-p)z\,\frac{1-(pz)^a}{1-pz}\, G(z)$$

d'où $G(z) = \dfrac{p^a z^a (1-pz)}{1 - z + p^a z^{a+1}(1-p)}$

On en déduit $\mathbf{E}(X) = G'(1) = \dfrac{1-p^a}{p^a(1-p)}$. La loi de X pourrait se trouver en développant G en série entière mais c'est compliqué.

❑ Cet exercice intervient aussi dans le nombre de touches à taper sur un digicode, sachant que le code à trouver est constitué de a lettres, que le digicode comporte m touches, qu'on tape sur chaque touche au hasard (avec donc à chaque fois une probabilité de $p = \dfrac{1}{m}$ que la touche soit bonne) et que, si on appuie sur une touche erronée, toutes les touches appuyées sont oubliées. L'espérance du nombre de touches à appuyer est $\dfrac{1-p^a}{p^a(1-p)}$ avec $p = \dfrac{1}{m}$.

f) Le nombre moyen de tirages entre un tirage F et le tirage F qui suit (le F initial étant non compté et le F final étant compté) est $\dfrac{1}{1-p}$, espérance de la loi géométrique de paramètre $1-p$. Le temps moyen de passage entre deux voitures est donc $\dfrac{t}{1-p}$. Le débit est alors :

$$\text{D} = \frac{1-p}{t} = \frac{a(1-p)}{\text{T}}$$

On peut dire aussi que, pendant une durée $\Delta = nt$, le nombre de véhicules qui passent suit une loi $\mathcal{B}(n, 1-p)$ d'espérance $n(1-p)$ et donc que le nombre moyen de véhicules par seconde est :

$$\text{D} = \frac{n(1-p)}{\Delta} = \frac{1-p}{t}$$

D'après la question e), on a :

$$t\mathbf{E}(\text{X}) = t\,\frac{1-p^a}{p^a(1-p)} = \frac{1-p^a}{\text{D}p^a}$$

or $\quad p = 1 - t\text{D} = 1 - \dfrac{\text{DT}}{a}$ donc $p^a = \left(1 - \dfrac{\text{DT}}{a}\right)^a$ qui tend vers $\exp(-\text{DT})$ quand a tend vers l'infini.

Donc $t\mathbf{E}(\text{X})$ tend vers $\dfrac{1-\exp(-\text{DT})}{\text{D}\exp(-\text{DT})} = \dfrac{\exp(\text{DT})-1}{\text{D}}$.

Si le débit D est petit, cette valeur vaut quasiment T, durée de traversée : le piéton traverse sans avoir à atteindre, puisque quasiment aucune voiture ne passe.

Le temps d'attente et de traversée augmente quasiment exponentiellement avec le débit D.

Sol.2-7) a) $\text{R} = \displaystyle\sum_{n=1}^{\infty} \mathbf{1}_{\text{X}_n=0}$ car chaque variable aléatoire $\mathbf{1}_{\text{X}_n=0}$ qui prend la valeur 1 décompte un retour en 0, et donc leur somme décompte le nombre de retours en 0.

b) Donc $\mathbf{E}(\text{R}) = \mathbf{E}\left(\displaystyle\sum_{n=1}^{\infty} \mathbf{1}_{\text{X}_n=0}\right) = \displaystyle\sum_{n=1}^{\infty} \mathbf{E}(\mathbf{1}_{\text{X}_n=0}) = \displaystyle\sum_{n=1}^{\infty} \mathbf{P}(\text{X}_n = 0)$, donc 0 est récurrent si et seulement si la série $\sum \mathbf{P}(\text{X}_n = 0)$ diverge, et transitoire si et seulement si $\sum \mathbf{P}(\text{X}_n = 0)$ converge.

c) Or $\mathbf{P}(\text{X}_n = 0) = 0$ si n est impair, et pour un indice pair, $\mathbf{P}(\text{X}_{2n} = 0) = \dbinom{2n}{n} p^n(1-p)^n$. En effet, sur $2n$ pas, il faut en effectuer n à droite, et les autres à gauche pour que l'événement $\{\text{X}_{2n} = 0\}$ soit réalisé. En utilisant la formule de Stirling $n! \sim n^n\, e^{-n} \sqrt{2\pi n}$ quand n tend vers l'infini (voir le chapitre L2/SUITESF.PDF), on a :

$$\mathbf{P}(\text{X}_{2n} = 0) \sim \frac{(2n)^{2n}\, e^{-2n} \sqrt{4\pi n}}{n^{2n}\, e^{-2n}\, 2\pi n}\, p^n(1-p)^n = \frac{(4p(1-p))^n}{\sqrt{\pi n}}$$

Remarquer que $4p(1-p)$ est un réel compris entre 0 et 1, la valeur 1 étant obtenue si et seulement si $p = \dfrac{1}{2}$, que $\sum \dfrac{1}{\sqrt{n}}$ diverge et que $\sum \dfrac{x^n}{\sqrt{n}}$ converge si $0 \le x < 1$. Donc 0 est récurrent si et seulement si $\sum \mathbf{P}(\text{X}_{2n} = 0)$ diverge, si et seulement si $p = \dfrac{1}{2}$.

Si $p > \dfrac{1}{2}$, on dérive vers la droite. Si $p < \dfrac{1}{2}$, on dérive vers la gauche. Dans ces deux cas, on ne repasse en 0 en moyenne qu'un nombre fini de fois.

COMPLEMENTS ET VARIANTES

❑ Ainsi, on a montré qu'avec une équiprobabilité de faire un pas à droite ou à gauche, l'espérance du nombre de retour en 0 est infinie. On peut aussi le voir de la façon suivante. Le nombre de retour en 0 au bout des $2n$ premiers pas est la somme partielle de la série :

$$\sum_{k=1}^{2n} \mathbf{P}(X_k = 0) = \sum_{k=1}^{n} \mathbf{P}(X_{2k} = 0) = \sum_{k=1}^{n} \frac{1}{4^k}\binom{2k}{k} \qquad \text{pour } p = \frac{1}{2}$$

et on vérifiera par récurrence sur n que cette somme vaut $\frac{2n+1}{4^n}\binom{2n}{n} - 1$, quantité équivalente à $\frac{2\sqrt{n}}{\sqrt{\pi}}$ quand n tend vers l'infini, en utilisant la formule de Stirling. Cette quantité tend vers l'infini quand n tend vers l'infini.

❑ Dans le cas général, pour $p \neq \frac{1}{2}$, on peut utiliser le développement en série entière suivant (voir les exercices du chapitre sur les séries entières L2/SERIENTR.PDF) :

$$\frac{1}{\sqrt{1-4x}} = \sum_{n=0}^{\infty} \binom{2n}{n} x^n$$

pour obtenir une expression exacte de l'espérance :

$$\mathbf{E}(R) = \sum_{n=1}^{\infty} \mathbf{P}(X_n = 0) = \sum_{n=0}^{\infty} \binom{2n}{n} p^n(1-p)^n - 1 = \frac{1}{\sqrt{1-4p(1-p)}} - 1 = \frac{1}{|2p-1|} - 1$$

❑ Si r est la probabilité d'un retour en 0 en partant de 0 au cours de la marche aléatoire, alors :

$\mathbf{P}(R = k) = r^k(1-r)$ on revient k fois en 0 puis on n'y revient plus

Si $r < 1$:

$$\sum_{k=0}^{\infty} \mathbf{P}(R = k) = \sum_{k=0}^{\infty} r^k(1-r) = 1 \text{ ce qui implique que } \mathbf{P}(R = +\infty) = 0$$

et
$$\mathbf{E}(R) = \sum_{k=1}^{\infty} kr^k(1-r) = \frac{r}{1-r} = \frac{1}{1-r} - 1$$
$$= \frac{1}{|2p-1|} - 1 \qquad \text{d'après l'expression de } \mathbf{E}(R) \text{ calculée précédemment}$$

donc $r = 1 - |2p-1|$

Si $r = 1$, $\mathbf{P}(R = k) = 0$ pour tout k, donc $\mathbf{P}(R = +\infty) = 1$ et $\mathbf{E}(R) = +\infty$.

On voit donc qu'il y a équivalence entre :

0 est récurrent
$\mathbf{E}(R) = +\infty$
la probabilité r d'un retour en 0 est égale à 1
$\mathbf{P}(R = +\infty) = 1$

Il y a équivalence entre :

0 est transitoire
$\mathbf{E}(R)$ est fini
la probabilité r d'un retour en 0 est strictement inférieure à 1
$\mathbf{P}(R = +\infty) = 0$

❑ On peut calculer la probabilité qu'il y ait un premier retour en 0 au bout de n pas de la façon suivante.
Pour tout entier $n \geq 0$, on note a_n le nombre de marches aléatoires à n pas dont la dernière étape est 0. Comme on l'a déjà vu, $a_n = 0$ si n est impair et $a_{2n} = \binom{2n}{n}$. On convient que $a_0 = 1$.
Pour tout $n \geq 1$, on note b_n le nombre de marches aléatoires à n pas se terminant en position 0, sans jamais passer par 0 avant la dernière étape. On a :

$$\forall\, n \geq 1,\ a_n = \sum_{k=1}^{n} b_k\, a_{n-k}$$

car pour pouvoir être en 0 à la n-ème étape, il faut y revenir une première fois à la k-ème étape, k pouvant prendre des valeurs de 1 à n, puis effectuer une marche aléatoire de $n - k$ pas partant de 0 et terminant en 0 (le raisonnement est valide y compris pour $k = n$ grâce à la convention $a_0 = 1$). Pour x réel, Posons $A(x) = \sum_{n=0}^{\infty} a_n x^n$ et $B(x) = \sum_{n=1}^{\infty} b_n x^n$. Comme a_n et b_n sont tous deux inférieurs à 2^n (nombre total de marches à n pas), ces deux séries entières convergent absolument si $|x| < \frac{1}{2}$. Pour un tel x, on a :

$$A(x)B(x) = \sum_{n=0}^{\infty} a_n x^n \times \sum_{n=1}^{\infty} b_n x^n$$

$$= \sum_{n=1}^{\infty} \sum_{k=1}^{n} a_{n-k} b_k\, x^n \qquad \text{en effectuant le produit de Cauchy des deux séries}$$

$$= \sum_{n=1}^{\infty} a_n x^n$$

$$= A(x) - a_0 = A(x) - 1$$

Donc $B(x) = 1 - \dfrac{1}{A(x)}$. Or :

$$A(x) = \sum_{n=0}^{\infty} a_{2n} x^{2n} \qquad \text{puisque les } a_{2n+1} \text{ sont nuls}$$

$$= \sum_{n=0}^{\infty} \binom{2n}{n} x^{2n}$$

$$= \frac{1}{\sqrt{1-4x^2}} \qquad \text{formule déjà rencontrée plus haut dans l'exercice}$$

donc $\quad B(x) = 1 - \dfrac{1}{A(x)} = 1 - \sqrt{1-4x^2}$

On vérifiera que le développement en série entière de $B(x)$ donne :

$$B(x) = \sum_{n=1}^{\infty} \frac{1 \times 3 \times \ldots \times (2n-3)}{2^n\, n!}\, 4^n\, x^{2n} = \sum_{n=1}^{\infty} \frac{1}{2n-1} \binom{2n}{n} x^{2n}.$$

Donc $b_n = 0$ si n est impair et $b_{2n} = \dfrac{1}{2n-1}\dbinom{2n}{n}$.
La probabilité de revenir en 0 pour la première fois en $2n$ pas est :

$$p_{2n} = b_{2n}p^n(1-p)^n$$

Elle est nulle si le nombre de pas est impair.
La probabilité r de revenir en 0 au moins une fois est :

$$r = \sum_{n=1}^{\infty} p_n = \sum_{n=1}^{\infty} b_{2n}p^n(1-p)^n = \mathrm{B}(\sqrt{p(1-p)})$$

qui redonne la valeur $r = 1 - |2p-1|$ déjà rencontrée. Mais on peut ici calculer l'espérance du nombre de pas pour revenir en 0 pour la première fois. Si Y est ce nombre de pas, les calculs précédents expriment que, pour tout $n \geq 1$, $\mathbf{P}(\mathrm{Y} = n) = p_n$, nul si n est impair, et :

$$\mathbf{P}(\mathrm{Y} = 2n) = p_{2n} = b_{2n}p^n(1-p)^n$$

r n'est autre que $\sum\limits_{n=1}^{\infty} \mathbf{P}(\mathrm{Y} = n) = \mathbf{P}(\mathrm{Y} < +\infty)$, ce qu'on peut aussi écrire :

$$\mathbf{P}(\mathrm{Y} = +\infty) = 1 - r = |2p-1| \qquad \text{probabilité de ne jamais revenir en 0.}$$

Si $p \neq \dfrac{1}{2}$, $\mathbf{E}(\mathrm{Y}) = +\infty$, mais curieusement, c'est aussi le cas si $p = \dfrac{1}{2}$ bien que la probabilité de revenir en 0 soit alors de 1 :

$$\mathbf{E}(\mathrm{Y}) = \sum_{n=0}^{\infty} n\mathbf{P}(\mathrm{Y} = n) = \sum_{n=1}^{\infty} 2np_{2n}$$

Or $\qquad 2np_{2n} = \dfrac{2nb_{2n}}{2^{2n}} \sim \dfrac{1}{4^n}\dfrac{(2n)^{2n}\,e^{-2n}\,2\sqrt{\pi n}}{n^{2n}\,e^{-2n}\,2\pi n} = \dfrac{1}{\sqrt{\pi n}}$ terme général d'une série divergente.

Il est intéressant d'effectuer une simulation numérique de ce cas, afin de comprendre cette apparente contradiction.
Polya a montré que 0 reste récurrent lors d'une marche aléatoire équiprobable dans $\mathbb{Z}^2$, mais devient transitoire pour les marches aléatoires dans $\mathbb{Z}^d$, $d \geq 3$.

Sol.2-8) a) Utilisons la formule des probabilités totales en considérant le système complet d'événements :

A = {on tombe sur Pile au premier lancer}

et B = {on tombe sur Face au premier lancer}

On a :

$$\mathbf{P}_k = \mathbf{P}(\text{ruine en partant de } k \mid \mathrm{A}) \times \mathbf{P}(\mathrm{A}) + \mathbf{P}(\text{ruine en partant de } k \mid \mathrm{B}) \times \mathbf{P}(\mathrm{B})$$

Mais $\mathbf{P}$(ruine en partant de k | A) = $\mathbf{P}$(ruine en partant de $k + 1$) = $\mathbf{P}_{k+1}$, puisque, si A est réalisé, on va se déplacer de k en $k + 1$.
De même, $\mathbf{P}$(ruine en partant de k | B) = $\mathbf{P}$(ruine en partant de $k + 1$) = $\mathbf{P}_{k-1}$.
Par ailleurs, $\mathbf{P}(\mathrm{A}) = \mathbf{P}(\mathrm{B}) = \dfrac{1}{2}$, donc $\mathbf{P}_k = \dfrac{1}{2}\mathbf{P}_{k+1} + \dfrac{1}{2}\mathbf{P}_{k-1}$.
Il s'agit d'une suite récurrente linéaire double dont l'équation caractéristique est $r^2 - 2r + 1 = 0$ possédant 1 comme racine double. Il existe donc λ et μ tel que, pour tout k, $\mathbf{P}_k = \lambda + \mu k$. Or $\mathbf{P}_0 = 1$ et $\mathbf{P}_b = 0$ donc $\lambda = 1$ et $\lambda + \mu b = 0$ donc $\mu = -\dfrac{1}{b}$. Ainsi :

$$\mathbf{P}_k = 1 - \frac{k}{b} = \frac{b-k}{b}$$

Symétriquement, la probabilité d'atteindre b est $\frac{k}{b}$. La somme des deux probabilités vaut 1, ce qui signifie que le jeu s'arrête presque sûrement.

b) L'espérance de gain vaut $0 \times \mathbf{P}_a + b(1 - \mathbf{P}_a) = b(1 - \frac{b-a}{b}) = a$. Le jeu est parfaitement neutre : l'espérance de gain est égale à la somme possédée initialement.

c) On peut supposer les X_k à valeurs dans $\mathbb{N}$, puisque, la partie se terminant avec une probabilité 1, $\mathbf{P}(X_k = +\infty) = 0$ pour tout k.

Comme au a), on a, pour tout $n \geq 1$:

$$\begin{aligned}\mathbf{P}(X_k = n) &= \mathbf{P}(X_k = n \mid A) \times \mathbf{P}(A) + \mathbf{P}(X_k = n \mid B) \times \mathbf{P}(B)\\ &= \mathbf{P}(X_{k+1} = n-1) \times \frac{1}{2} + \mathbf{P}(X_{k-1} = n-1) \times \frac{1}{2}\end{aligned}$$

avec $\mathbf{P}(X_k = 0) = \begin{cases} 0 \text{ si } k \neq 0 \text{ et } k \neq b \\ 1 \text{ sinon} \end{cases}$.

Donc :

$$\mathbf{E}(X_k) = \sum_{n=1}^{\infty} n\, \mathbf{P}(X_k = n)$$

$$= \frac{1}{2} \sum_{n=1}^{\infty} n\, \mathbf{P}(X_{k+1} = n-1) + \frac{1}{2} \sum_{n=1}^{\infty} n\, \mathbf{P}(X_{k-1} = n-1)$$

$$= \frac{1}{2} \Big(\sum_{n=1}^{\infty} \mathbf{P}(X_{k+1} = n-1) + \sum_{n=1}^{\infty} (n-1)\, \mathbf{P}(X_{k+1} = n-1)\Big)$$

$$+ \frac{1}{2} \Big(\sum_{n=1}^{\infty} \mathbf{P}(X_{k-1} = n-1) + \sum_{n=1}^{\infty} (n-1)\, \mathbf{P}(X_{k-1} = n-1)\Big)$$

$$= \frac{1}{2} (1 + \mathbf{E}(X_{k+1})) + \frac{1}{2} (1 + \mathbf{E}(X_{k-1}) = 1 + \frac{\mathbf{E}(X_{k+1}) + \mathbf{E}(X_{k-1})}{2}$$

On a $\mathbf{E}(X_0) = \mathbf{E}(X_b) = 0$.

Les suites générales (u_k) vérifiant la relation de récurrence $u_k = 1 + \frac{u_{k-1} + u_{k+1}}{2}$ sont sommes d'une suite particulière vérifiant cette relation (par exemple la suite $u_k = -k^2$) et d'une suite générale vérifiant la relation $u_k = \frac{u_{k-1} + u_{k+1}}{2}$, ces dernières étant de la forme $\lambda + \mu k$, comme on l'a vu dans le a). Ainsi, $\mathbf{E}(X_k)$ est de la forme $\mathbf{E}(X_k) = \lambda + \mu k - k^2$. On a $\mathbf{E}(X_0) = 0 = \lambda$ et $\mathbf{E}(X_b) = 0 = \lambda + \mu b - b^2$ donc $\lambda = 0$ et $\mu = b$. Donc $\mathbf{E}(X_k) = bk - k^2$. Donc $\mathbf{E}(X_a) = a(b-a)$, produit des distances de a aux deux points terminant le jeu.

d) Posons $q = 1 - p$. La relation de récurrence vérifiée par les $\mathbf{P}_k$ devient :

$$\mathbf{P}_k = p\, \mathbf{P}_{k+1} + q\, \mathbf{P}_{k-1}$$

L'équation caractéristique devient $pr^2 - r + q = 0$, dont les racines sont 1 et $\frac{q}{p}$. Il existe λ et μ tels que, pour tout k, $\mathbf{P}_k = \lambda + \mu \frac{q^k}{p^k} = \lambda + \mu\, p^{-k}q^k$. Les valeurs $\mathbf{P}_0 = 1$ et $\mathbf{P}_b = 0$ permettent d'en déduire les expressions de λ et μ, puis l'expression finale de $\mathbf{P}_k$:

$$\mathbf{P}_k = \frac{q^b - p^{b-k}q^k}{q^b - p^b}$$

De même, la relation de récurrence portant sur les $\mathbf{E}(X_k)$ devient :

$$\mathbf{E}(X_k) = 1 + p\mathbf{E}(X_{k+1}) + q\mathbf{E}(X_{k-1})$$

Une suite particulière vérifiant la récurrence $u_k = 1 + pu_{k+1} + qu_{k-1}$ est donnée par $u_k = \frac{k}{q-p}$. La solution générale est donc de la forme $\frac{k}{q-p} + \lambda + \mu \frac{q^k}{p^k}$. Les conditions limites $\mathbf{E}(X_0) = \mathbf{E}(X_b) = 0$ permettent de trouver λ et μ, et finalement, on trouvera :

$$\mathbf{E}(X_k) = \frac{k}{q-p} - \frac{b}{q-p}\,\frac{q^k p^{b-k} - p^b}{q^b - p^b}$$

On pourra vérifier que $\lim\limits_{p\to 1/2} \mathbf{E}(X_k) = bk - k^2$, redonnant le résultat du c).

La valeur de $\mathbf{E}(X_k)$ est utile pour le joueur de roulette qui se contente de miser sur la couleur noire ou rouge de la roulette. Il y a 18 cases rouges, 18 cases noires et la case 0. On a alors $p = \frac{18}{37}$ et $q = \frac{19}{37}$.

Le joueur possède $a = 50$ euros et s'arrête de jouer s'il est ruiné ou s'il parvient à la somme de $b = 100$ euros. La probabilité d'être ruiné est $\mathbf{P}_{50} \approx 0{,}937$. Le nombre de parties jouées est en moyenne de $\mathbf{E}(X_{50}) \approx 1618$.

Si on est plus modeste avec $a = 10$ euros et $b = 20$ euros, alors $\mathbf{P}_{10} \approx 0{,}632$ et $\mathbf{E}(X_{10}) \approx 98$.

Si $p < \frac{1}{2}$ et b tend vers l'infini (cas du joueur compulsif ne parvenant pas à s'arrêter, sauf s'il est ruiné), on a à la limite $\mathbf{P}_a = 1$ (certitude de finir ruiné) et $\mathbf{E}(X_a) = \frac{a}{q-p}$.

Sol.2-9) a) X_i suit une loi géométrique de paramètre $p = \frac{n-i+1}{n}$ sur $\mathbb{N}^*$, de fonction génératrice $\frac{pz}{1-(1-p)z}$.

b) Les X_i sont indépendantes puisque les tirages successifs sont avec remise. On a donc :

$$G_X(z) = \prod_{i=1}^{n} G_i(z) = \frac{n!}{n^n}\,\frac{z^n}{1 \times (1 - \frac{z}{n}) \times (1 - \frac{2z}{n}) \times \ldots \times (1 - \frac{(n-1)z}{n})}$$

c) La somme $\displaystyle\sum_{k=1}^{n-1} \frac{a_k}{1-\frac{kz}{n}}$ est la décomposition en éléments simples de $\dfrac{1}{(1-\frac{z}{n})...(1-\frac{(n-1)z}{n})}$. Pour déterminer a_k, on multiplie les deux membres de l'égalité $\dfrac{1}{(1-\frac{z}{n})...(1-\frac{(n-1)z}{n})} = \displaystyle\sum_{k=1}^{n-1} \frac{a_k}{1-\frac{kz}{n}}$ par $1-\frac{kz}{n}$, on simplifie ce facteur au numérateur et au dénominateur du membre de gauche, puis en remplace z par $\frac{n}{k}$:

$$\begin{aligned} a_k &= \frac{1}{(1-\frac{1}{k})(1-\frac{2}{k})...(1-\frac{k-1}{k})(1-\frac{k+1}{k})...(1-\frac{n-1}{k})} \\ &= (-1)^{n-1-k} \frac{k^{n-2}}{(k-1)!\,(n-1-k)!} \\ &= (-1)^{n-1-k} \frac{k^{n-1}}{k!\,(n-1-k)!} \\ &= (-1)^{n-1-k} \frac{k^{n-1}}{(n-1)!} \binom{n-1}{k} \end{aligned}$$

Donc $\displaystyle G_X(z) = \frac{n!}{n^n} z^n \sum_{k=1}^{n-1} (-1)^{n-1-k} \frac{k^{n-1}}{(n-1)!} \binom{n-1}{k} \frac{1}{1-\frac{kz}{n}}$

d) On développe en série entière tous les $\dfrac{1}{1-\frac{kz}{n}}$, ce qui donne :

$$\begin{aligned} G_X(z) &= \frac{n!}{n^n} z^n \sum_{k=1}^{n-1} (-1)^{n-1-k} \frac{k^{n-1}}{(n-1)!} \binom{n-1}{k} \sum_{m=0}^{\infty} \frac{k^m z^m}{n^m} \\ &= \sum_{m=0}^{\infty} \sum_{k=1}^{n-1} (-1)^{n-1-k} \binom{n-1}{k} \frac{k^{m+n-1} z^{m+n}}{n^{m+n-1}} \\ &= \sum_{m=n}^{\infty} \sum_{k=1}^{n-1} (-1)^{n-1-k} \binom{n-1}{k} \frac{k^{m-1} z^m}{n^{m-1}} \end{aligned}$$

donc $\quad \displaystyle\forall\, m \geq n,\ \mathbf{P}(X = m) = \sum_{k=1}^{n-1} (-1)^{n-1-k} \binom{n-1}{k} \frac{k^{m-1}}{n^{m-1}}$

Dans le chapitre L1/DENOMBRE.PDF, on montre que le nombre de surjections d'un ensemble à n éléments dans un ensemble à q éléments est $\sigma(n, q) = \displaystyle\sum_{k=1}^{q} (-1)^{q-k} \binom{q}{k} k^n$. Donc $\mathbf{P}(X = m)$ est égal à $\dfrac{\sigma(m-1, n-1)}{n^{m-1}}$. On retrouve ce résultat directement comme suit. L'exécution de m tirages successifs est vue comme une application f de $[\![1, m]\!]$ dans $[\![1, n]\!]$, qui, à un tirage de rang k entre 1 et m

associe la valeur $f(k)$ du jeton tiré. Dire que cette application est surjective, c'est dire qu'on a procédé à m tirages successifs ayant permis de tirer tous les jetons. C'est donc dire que l'événement $X \leq m$ est réalisé. Dire que $X = m$, c'est dire que f est surjective de $[\![1, m-1]\!]$ dans $[\![1, n]\!]$ privé d'un élément r (de sorte qu'on n'a pas $X \leq m-1$) et que $f(m) = r$. Le nombre de telles applications f est $n \times \sigma(m-1, n-1)$ (n choix possibles de r multiplié par $\sigma(m-1, n-1)$ choix possibles d'une application surjective de $[\![1, m-1]\!]$ sur $[\![1, n]\!] \setminus \{r\}$). Le nombre total d'applications de $[\![1, m]\!]$ dans $[\![1, n]\!]$ est n^m, correspondant au nombre total de tirages possibles de m jetons. Tous les tirages étant équiprobables, la probabilité que $X = m$ soit réalisé est :

$$\mathbf{P}(X = m) = \frac{n \times \sigma(m-1, n-1)}{n^m} = \frac{\sigma(m-1, n-1)}{n^{m-1}} = \sum_{k=1}^{n-1} (-1)^{n-1-k} \binom{n-1}{k} \frac{k^{m-1}}{n^{m-1}}$$

e) Puisque X_k suit une loi géométrique de paramètre $p = \dfrac{n-k+1}{n}$, $\mathbf{E}(X_k) = \dfrac{n}{n-k+1}$. Donc :

$$\mathbf{E}(X) = \sum_{k=1}^{n} \mathbf{E}(X_k) = n \sum_{k=1}^{n} \frac{1}{n-k+1} = n \sum_{k=1}^{n} \frac{1}{k} \sim n \ln(n)$$

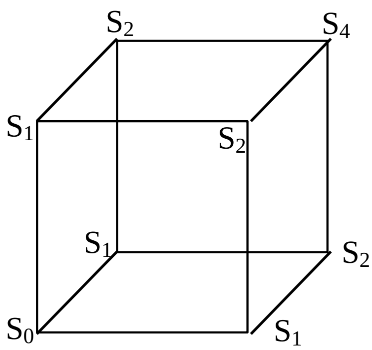
Pour l'équivalent $\displaystyle\sum_{k=1}^{n} \frac{1}{k} \sim \ln(n)$, voir le chapitre L2/SERIES.PDF

f) On a $\mathbf{V}(X_k) = \dfrac{1-p}{p^2} = \dfrac{(k-1)n}{(n-k+1)^2}$ et les X_i sont des variables aléatoires indépendantes. Donc :

$$\mathbf{V}(X) = \sum_{k=1}^{n} \mathbf{V}(X_k) = n \sum_{k=1}^{n} \frac{k-1}{(n-k+1)^2} = n \sum_{k=1}^{n} \frac{n-k}{k^2} = n^2 \sum_{k=1}^{n} \frac{1}{k^2} - n \sum_{k=1}^{n} \frac{1}{k} \sim n^2 \sum_{k=1}^{\infty} \frac{1}{k^2} = n^2 \frac{\pi^2}{6}$$

La valeur $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$ est donnée dans le chapitre L1/SERIES.PDF.

Sol.2-10) a) Le graphe indiqué correspond aux possibilité de passage d'un sommet à l'autre du cube, S_0 étant le sommet initial, S_4 le sommet diamétralement opposé, S_1 un sommet quelconque adjacent à S_0 auquel on parvient en partant de S_0, S_2 un sommet quelconque adjacent à S_4, S_3 un sommet de type S_1 auquel on parvient depuis un sommet S_2 et autre que le sommet S_1 dont on est parti pour parvenir à ce sommet S_2.



b) Pour tout $k \geq 3$:

$\mathbf{P}(T_0 = 1) = \mathbf{P}(T_0 = 2) = 0$

$\mathbf{P}(T_0 = k) = \mathbf{P}(T_1 = k-1)$

Pour $k \geq 2$:

$\mathbf{P}(T_1 = 1) = 0$

$\mathbf{P}(T_1 = k) = \mathbf{P}(T_2 = k - 1)$

Pour $k \geq 2$:

$\mathbf{P}(T_2 = 1) = \frac{1}{2}$

$\mathbf{P}(T_2 = k) = \frac{1}{2}\,\mathbf{P}(T_3 = k - 1)$

Cette dernière égalité est obtenue en appliquant la formule des probabilités totales, en conditionnant par rapport aux événement A = "on va du sommet S_2 au sommet S_3" et son complémentaire B = "on va du sommet S_2 au sommet S_4" :

$\mathbf{P}(T_2 = k) = \mathbf{P}(T_2 = k \mid A) \times \mathbf{P}(A) + \mathbf{P}(T_2 = k \mid B) \times \mathbf{P}(B)$

Mais $\mathbf{P}(T_2 = k \mid A)$ n'est autre que $\mathbf{P}(T_3 = k - 1)$ et $\mathbf{P}(T_2 = k \mid B)$ est nul puisque $k \geq 2$.

Pour $k \geq 2$:

$\mathbf{P}(T_3 = 1) = 0$

$\mathbf{P}(T_3 = k) = \frac{1}{2}\,\mathbf{P}(T_2 = k - 1) + \frac{1}{2}\,\mathbf{P}(T_0 = k - 1)$

Pour cette dernière égalité, on applique la formule des probabilités totales, en conditionnant par rapport aux événement A = "on va du sommet S_3 au sommet S_2" et son complémentaire B = "on va du sommet S_3 au sommet S_0".

Donc, en multipliant chaque $\mathbf{P}(T_i = k)$ par z^k et en sommant de $k = 1$ à l'infini, on obtient :

$$\left[\begin{array}{l} g_0(z) = z g_1(z) \\ g_1(z) = z g_2(z) \\ g_2(z) = \frac{z}{2} + \frac{z}{2} g_3(z) \\ g_3(z) = \frac{z}{2} g_2(z) + \frac{z}{2} g_0(z) \end{array}\right.$$

En effet, par exemple pour g_2 :

$$g_2(z) = \sum_{k=1}^{\infty} \mathbf{P}(T_2 = k) z^k = \frac{z}{2} + \sum_{k=2}^{\infty} \mathbf{P}(T_2 = k) z^k$$

$$= \frac{z}{2} + \frac{1}{2} \sum_{k=2}^{\infty} \mathbf{P}(T_3 = k - 1) z^k$$

$$= \frac{z}{2} + \frac{z}{2} \sum_{k=1}^{\infty} \mathbf{P}(T_3 = k) z^k$$

$$= \frac{z}{2} + \frac{z}{2} g_3(z)$$

c) On résout le système obtenu et on obtient :

$$\left[\begin{array}{l} g_0(z) = \dfrac{2z^3}{4 - z^2 - z^4} \\ g_1(z) = \dfrac{2z^2}{4 - z^2 - z^4} \\ g_2(z) = \dfrac{2z}{4 - z^2 - z^4} \\ g_3(z) = \dfrac{z^4 + z^2}{4 - z^2 - z^4} \end{array}\right.$$

A partir de g_0, on en déduit :

$\mathbf{E}(X) = \mathbf{E}(T_0) = g_0'(1) = 6$
$\mathbf{E}(X(X-1)) = g_0''(1) = 49$
$\mathbf{V}(X) = \mathbf{E}(X(X-1)) + \mathbf{E}(X) - \mathbf{E}(X)^2 = 19$

❑ On peut aussi en profiter pour tirer $\mathbf{E}(T_1) = g_1'(1) = 5$, $\mathbf{E}(T_2) = 4$, $\mathbf{E}(T_3) = 6$. On pourra réfléchir pourquoi il est assez naturel que :

$\mathbf{E}(T_0) = \mathbf{E}(T_1) + 1$
$\mathbf{E}(T_1) = \mathbf{E}(T_2) + 1$
$\mathbf{E}(T_2) = 1 + \frac{1}{2}\mathbf{E}(T_3)$
$\mathbf{E}(T_3) = \frac{1}{2}\mathbf{E}(T_2) + \frac{1}{2}\mathbf{E}(T_0) + 1$

Une démonstration convenable de la dernière égalité ci-dessus par exemple consiste à remarquer que, si on considére les événements A = "on passe du sommet 3 au sommet 2" et son complémentaire B = "on passe du sommet 3 au sommet 0", la formule de l'espérance totale donne :

$\mathbf{E}(T_3) = \mathbf{E}(T_3 \mid A) + \mathbf{E}(T_3 \mid B)$
$= \mathbf{E}(T_2 + 1 \mid A)\,\mathbf{P}(A) + \mathbf{E}(T_0 + 1 \mid B)\,\mathbf{P}(B)$

puisque, si on suppose A réalisé, $T_3 = T_2' + 1$ avec T_2' variable aléatoire de même loi que T_2. De même pour B.

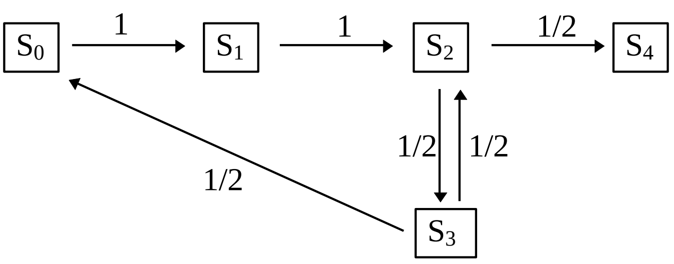
$= \frac{1}{2}(\mathbf{E}(T_2 + 1) + \mathbf{E}(T_0 + 1))$

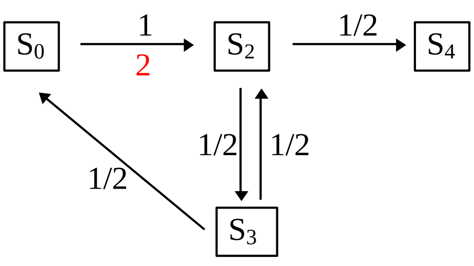
car T_2 et T_0 sont indépendants des événements A et B qui les ont précédés

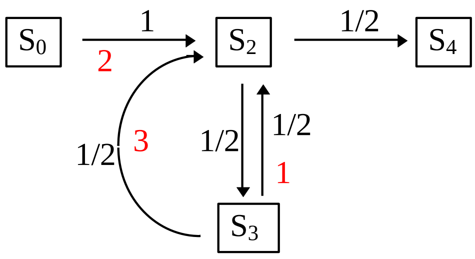
on a :

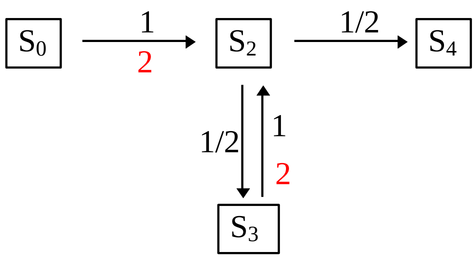
$= \frac{1}{2}\mathbf{E}(T_2) + \frac{1}{2}\mathbf{E}(T_0) + 1$

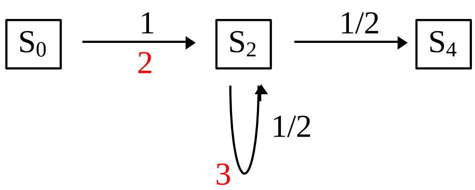
❑ On peut aussi déterminer $\mathbf{E}(T_0)$ graphiquement, comme on l'a fait dans un exercice précédent. On indique en noir la probabilité de passer d'un état à un autre, et en rouge l'espérance du nombre d'arêtes à parcourir pour changer d'état :





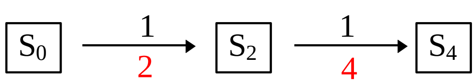


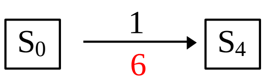




Le nombre de flèches pour passer de l'état S_2 à l'état S_4 suit une loi géométrique de paramètre $p = \frac{1}{2}$.

L'espérance de ce nombre est $\frac{1}{p} = 2$, la flèche finale correspond au parcours final par la fourmi d'un arête $S_2 \rightarrow S_4$ et la flèche précédente correspondant à un nombre moyen de 3 arêtes parcourues. On peut donc réduire le graphe précédent au graphe suivant :





et on retrouve bien $\mathbf{E}(T_0) = 6$.

d) Il suffit de développer g_0 en série entière pour trouver la loi de X. On factorise le dénominateur de g_0 pour la réduire la fraction en éléments simples :

$$g_0(z) = \frac{2z^3}{4 - z^2 - z^4} = \frac{2z^3}{\left(\frac{\sqrt{17}}{2} - \frac{1}{2} - z^2\right)\left(\frac{\sqrt{17}}{2} + \frac{1}{2} + z^2\right)}$$

$$= \frac{2z^3}{(a - z^2)(b + z^2)}$$

en posant $a = \frac{\sqrt{17} - 1}{2}$ et $b = \frac{\sqrt{17} + 1}{2}$ pour alléger les notations

$$= \frac{2z^3}{a + b}\left(\frac{1}{a - z^2} + \frac{1}{b + z^2}\right)$$

$$= \frac{2z^3}{a+b}\left(\frac{1}{a}\sum_{n=0}^{\infty}\frac{z^{2n}}{a^n} + \frac{1}{b}\sum_{n=0}^{\infty}(-1)^n\ \frac{z^{2n}}{b^n}\right)$$

$$= \frac{2}{a+b}\left(\sum_{n=0}^{\infty}\frac{1}{a^{n+1}} + (-1)^n\ \frac{1}{b^{n+1}}\right) z^{2n+3}$$

Compte tenu que $a + b = \sqrt{17}$, $\frac{1}{a} = \frac{\sqrt{17}+1}{8}$ et $b = \frac{\sqrt{17}-1}{8}$, on en tire :

$$\forall\ n \geq 0,\ \mathbf{P}(X = 2n+3) = \frac{2}{\sqrt{17}}\left(\left(\frac{1+\sqrt{17}}{8}\right)^{n+1} - \left(\frac{1-\sqrt{17}}{8}\right)^{n+1}\right)$$

Sol.2-11) a) Pour $k \geq m$:

$Y = k \Leftrightarrow X_1 + ... + X_{k-1} = m - 1$ et $X_k = 1$, donc :

$\mathbf{P}(Y = k) = \mathbf{P}(X_1 + ... + X_{k-1} = m - 1 \text{ et } X_k = 1)$

$= \mathbf{P}(X_1 + ... + X_{k-1} = m - 1)\mathbf{P}(X_k = 1)$

en vertu de l'indépendance des variables aléatoires indépendantes. $S_{k-1} = X_1 + ... + X_{k-1}$ suit une loi binomiale $\mathcal{B}(k-1, p)$ donc :

$$\mathbf{P}(Y = k) = \binom{k-1}{m-1} p^m (1-p)^{k-m}$$

On peut aussi dire que, sur les k tirages, le dernier est un succès (Pile), et qu'il faut donc choisir $m - 1$ succès parmi les $k - 1$ tirages précédents, les m succès ayant chacun une probabilité p et les $k - m$ échecs ayant une probabilité $1 - p$.

b) On utilisera la formule (voir au besoin le chapitre L2/SERIENTR.PDF) :

$$\frac{1}{(1-x)^m} = \sum_{k=m-1}^{\infty}\binom{k}{m-1} x^{k-m+1} = \sum_{k=0}^{\infty}\binom{k+m-1}{m-1} x^k$$

On a ici :

$$\sum_{k=m}^{\infty}\mathbf{P}(Y = k) = \sum_{k=m}^{\infty}\binom{k-1}{m-1} p^m (1-p)^{k-m} = p^m \sum_{k=0}^{\infty}\binom{k+m-1}{m-1}(1-p)^k = \frac{p^m}{(1-(1-p))^m} = 1.$$

Cela signifie que Y est presque sûrement fini.

c) $$G_Y(z) = \sum_{k=0}^{\infty}\mathbf{P}(Y = k)\, z^k = \sum_{k=m}^{\infty}\binom{k-1}{m-1} p^m (1-p)^{k-m} z^k$$

$$= z^m p^m \sum_{k=m-1}^{\infty}\binom{k}{m-1}(1-p)^{k-m+1} z^{k-m+1}$$

$$= \frac{p^m z^m}{(1-z(1-p))^m}$$

donc :

$$G_Y'(z) = \frac{mz^{m-1}p^m(1-z(1-p)) + m(1-p)p^m z^m}{(1-z(1-p))^{m+1}}$$

$$= \frac{mz^{m-1}p^m}{(1-z(1-p))^{m+1}}$$

$$\mathbf{E}(Y) = G_Y'(1) = \frac{mp^m}{p^{m+1}} = \frac{m}{p}$$

$$\mathbf{V}(\mathrm{Y}) = \mathrm{G_Y}''(1) + \mathrm{G_Y}'(1) - \mathrm{G_Y}'(1)^2$$
$$= \ldots = \frac{m(1-p)}{p^2}$$

d) La fonction génératrice d'une variable aléatoire géométrique est $\frac{pz}{1 - z(1-p)}$. Or la fonction génératrice d'une somme de variables aléatoires indépendantes est le produit des fonctions génératrices de chaque variable. Donc la fonction génératrice demandée est $\frac{p^m z^m}{(1 - z(1-p))^m}$. C'est celle de Y.

On retrouve alors le fait que $\mathbf{E}(\mathrm{Y}) = \frac{m}{p}$, somme des m espérances $\frac{1}{p}$ des variables géométriques, et que $\mathbf{V}(\mathrm{Y}) = \frac{m(1-p)}{p^2}$, somme des m variances $\frac{1-p}{p^2}$ des variables géométriques.

On peut dire aussi que les m variables aléatoires indépendantes Y_i dont Y est la somme sont telles que Y_1 est le nombre de lancers pour obtenir le premier succès, puis Y_2 est le nombre de lancers supplémentaires pour obtenir le deuxième succès, etc. jusqu'à Y_m qui est le nombre de lancers suivant Y_{m-1} pour obtenir le m-ème succès.

e) $\mathbf{P}(\mathrm{X}_1 = 1) = \mathbf{P}(\mathrm{Y}_1 = 1) = p$

$$\mathbf{P}(\mathrm{X}_2 = 1) = \mathbf{P}(\mathrm{Y}_1 = 2 \text{ ou } (\mathrm{Y}_1 = 1 \text{ et } \mathrm{Y}_2 = 1))$$
$$= (1-p)p + p^2 = p$$

Montrons par récurrence que $\mathbf{P}(\mathrm{X}_n = 1) = p$. Supposons que cette propriété soit vraie jusqu'au rang $n-1$. On a ensuite, en notant Y une variable aléatoire générique suivant une loi géométrique de paramètre p :

$$\mathbf{P}(\mathrm{X}_n = 1) = \mathbf{P}(\mathrm{Y}_1 = n \text{ ou } (\exists\, k \geq 2, \exists\, m \leq n-k+1, \mathrm{Y}_k = m, \mathrm{Y}_1 + \ldots + \mathrm{Y}_{k-1} = n-m))$$
$$= \mathbf{P}(\mathrm{Y} = n) + \mathbf{P}(\exists\, m \leq n-1, \exists\, k \in [\![2, n-m+1]\!], \mathrm{Y}_k = m, \mathrm{Y}_1 + \ldots + \mathrm{Y}_{k-1} = n-m)$$
$$= \mathbf{P}(\mathrm{Y} = n) + \sum_{m=1}^{n-1} \sum_{k=2}^{n-m+1} \mathbf{P}(\mathrm{Y}_k = m, \mathrm{Y}_1 + \ldots + \mathrm{Y}_{k-1} = n-m)$$
$$= \mathbf{P}(\mathrm{Y} = n) + \sum_{m=1}^{n-1} \sum_{k=2}^{n-m+1} \mathbf{P}(\mathrm{Y}_k = m)\, \mathbf{P}(\mathrm{Y}_1 + \ldots + \mathrm{Y}_{k-1} = n-m)$$
$$= \mathbf{P}(\mathrm{Y} = n) + \sum_{m=1}^{n-1} \sum_{k=2}^{n-m+1} \mathbf{P}(\mathrm{Y} = m)\, \mathbf{P}(\mathrm{Y}_1 + \ldots + \mathrm{Y}_{k-1} = n-m)$$
$$= \mathbf{P}(\mathrm{Y} = n) + \sum_{m=1}^{n-1} (\mathbf{P}(\mathrm{Y} = m) \sum_{k=2}^{n-m+1} \mathbf{P}(\mathrm{Y}_1 + \ldots + \mathrm{Y}_{k-1} = n-m))$$
$$= \mathbf{P}(\mathrm{Y} = n) + \sum_{m=1}^{n-1} (\mathbf{P}(\mathrm{Y} = m) \sum_{k=1}^{n-m} \mathbf{P}(\mathrm{Y}_1 + \ldots + \mathrm{Y}_k = n-m))$$
$$= \mathbf{P}(\mathrm{Y} = n) + \sum_{m=1}^{n-1} \mathbf{P}(\mathrm{Y} = m)\, \mathbf{P}(\mathrm{X}_{n-m} = 1)$$
$$= \mathbf{P}(\mathrm{Y} = n) + p \sum_{m=1}^{n-1} \mathbf{P}(\mathrm{Y} = m) \qquad \text{d'après l'hypothèse de récurrence}$$

$$= (1-p)^{n-1}p + p\sum_{m=1}^{n-1}(1-p)^{m-1}p$$

$$= (1-p)^{n-1}p + p(1-(1-p)^{n-1})$$

$$= p$$

On peut simuler cette situation en effectuant des expériences de Bernoulli indépendantes de probabilité p. Y_k est le nombre d'expériences à réaliser après le $k-1$ succès pour obtenir le k-ème. Mais alors X_n est exactement l'événement "obtenir un succès au n-ème lancer". Il donne donc le résultat de la n-ème expérience de Bernoulli.

6- Solutions sur les couples de variables aléatoires

Sol.3-1) a) Soit A l'événement "on choisit le jeu 1" et B son complémentaire. On a alors :

$$G = \mathbf{1}_A \times G_1 + \mathbf{1}_B \times G_2$$

G_1 et A sont indépendants, de même que G_2 et B. Donc :

$$\begin{aligned}\mathbf{E}(G) &= \mathbf{E}(\mathbf{1}_A \times G_1 + \mathbf{1}_B \times G_2)\\ &= \mathbf{E}(\mathbf{1}_A)\,\mathbf{E}(G_1) + \mathbf{E}(\mathbf{1}_B)\,\mathbf{E}(G_2)\\ &= \mathbf{P}(A)\,\mathbf{E}(G_1) + \mathbf{P}(B)\,\mathbf{E}(G_2)\\ &= \frac{\mathbf{E}(G_1) + \mathbf{E}(G_2)}{2}\end{aligned}$$

On peut aussi utiliser la formule des espérances totales :

$$\begin{aligned}\mathbf{E}(G) &= \mathbf{E}(G \mid A)\mathbf{P}(A) + \mathbf{E}(G \mid B)\mathbf{P}(B)\\ &= \mathbf{E}(G_1)\mathbf{P}(A) + \mathbf{E}(G_2)\mathbf{P}(B)\end{aligned}$$

b) De même :

$$G^2 = \mathbf{1}_A \times G_1^2 + \mathbf{1}_B \times G_2^2 \qquad \text{car } \mathbf{1}_A^2 = \mathbf{1}_A \text{ et } \mathbf{1}_A\,\mathbf{1}_B = 0$$

donc :

$$\mathbf{E}(G^2) = \frac{\mathbf{E}(G_1^2) + \mathbf{E}(G_2^2)}{2}$$

Il en résulte que :

$$\begin{aligned}\mathbf{V}(G) &= \mathbf{E}(G^2) - \mathbf{E}(G)^2\\ &= \frac{\mathbf{E}(G_1^2) + \mathbf{E}(G_2^2)}{2} - \left(\frac{\mathbf{E}(G_1) + \mathbf{E}(G_2)}{2}\right)^2\\ &= \frac{\mathbf{V}(G_1) + \mathbf{V}(G_2)}{2} + \frac{\mathbf{E}(G_1^2) + \mathbf{E}(G_2^2)}{2} - \left(\frac{\mathbf{E}(G_1) + \mathbf{E}(G_2)}{2}\right)^2\\ &= \frac{\mathbf{V}(G_1) + \mathbf{V}(G_2)}{2} + \left(\frac{\mathbf{E}(G_1) - \mathbf{E}(G_2)}{2}\right)^2\end{aligned}$$

La relation demandée est vérifiée si et seulement si $\mathbf{E}(G_1) = \mathbf{E}(G_2)$.

Sol.3-2) a) $\mathbf{P}(T > X) = \mathbf{P}(\exists\, n \in \mathbb{N},\ X = n \text{ et } T > n)$

$$= \sum_{n=0}^{\infty} \mathbf{P}(X = n \text{ et } T > n) \qquad \text{car les événements } \{X = n \text{ et } T > n\} \text{ sont disjoints}$$

$$= \sum_{n=0}^{\infty} \mathbf{P}(X = n)\mathbf{P}(T > n) \qquad \text{car X et T sont indépendantes}$$

$$= \sum_{n=0}^{\infty} \mathbf{P}(X = n)(1 - p)^n$$

$$= G_X(1 - p)$$

b) $\mathbf{P}(T > X + Y \mid T > X) = \dfrac{\mathbf{P}(T > X + Y \text{ et } T > X)}{\mathbf{P}(T > X)}$

$$= \frac{\mathbf{P}(T > X + Y)}{\mathbf{P}(T > X)}$$

$$= \frac{G_{X+Y}(1 - p)}{G_X(1 - p)}$$

$= G_Y(1 - p)$ car X et Y étant indépendantes, $G_{X+Y} = G_X\ G_Y$.

$= \mathbf{P}(T > Y)$

Sol.3-3) a) $\mathbf{P}(XY = n) = \mathbf{P}(X = n \text{ et } Y = 1) + \mathbf{P}(X = -n \text{ et } Y = -1)$

$$= p\mathbf{P}(X = n) + (1 - p)\mathbf{P}(X = -n) = \mathbf{P}(X = n)$$

donc XY a même loi que X.

b) $\mathbf{E}(X) = 0 = \mathbf{E}(XY)$, donc :

$\mathrm{cov}(X, XY) = \mathbf{E}(X^2Y) = \mathbf{E}(X^2)\mathbf{E}(Y)$ car X et Y sont indépendantes

$= (2p - 1)\mathbf{E}(X^2)$

La covariance est nulle si et seulement si $p = \frac{1}{2}$. Même dans ce cas, les deux variables ne sont pas indépendantes, puisque si on sait que $X = n$, XY ne peut prendre que les valeurs $\pm n$.

c) $\mathbf{E}(XY) = 0$ donc :

$\mathrm{cov}(|X|, XY) = \mathbf{E}(|X|XY)$

$= \sum_{n\in\mathbf{Z}} n|n|\ \mathbf{P}(X = n)p - \sum_{n\in\mathbf{Z}} n|n|\ \mathbf{P}(X = n)(1 - p) = 0$ pour tout p

Comme au b), $|X|$ et XY ne sont pas indépendantes.

d) On a $\mathbf{P}(Z = 1) = \mathbf{P}(X > 0) = \mathbf{P}(X < 0) = \mathbf{P}(Z = -1)$, donc $\mathbf{P}(Z = 1) = \mathbf{P}(Z = -1) = \frac{1}{2}$ donc $\mathbf{E}(Z) = 0$, donc :

$\mathrm{cov}(|X|, Z) = \mathbf{E}(|X|Z) = \mathbf{E}(X) = 0$

Pour $n > 0$, $\mathbf{P}(|X| = n, Z = 1) = \mathbf{P}(X = n) = \frac{1}{2}\ \mathbf{P}(|X| = n) = \mathbf{P}(|X| = n)\ \mathbf{P}(Z = 1)$, donc les deux variables aléatoires $|X|$ et Z sont indépendantes.

Sol.3-4) $\mathbf{P}(X \le n \le X + Y) = \sum_{x=0}^{n} \sum_{y=n-x}^{\infty} \mathbf{P}(X = x, Y = y)$

$$= \sum_{x=0}^{n} \sum_{y=n-x}^{\infty} (1 - q)^2\, q^x\, q^y$$

$$= \sum_{x=0}^{n} (1 - q)^2\, q^x \sum_{y=n-x}^{\infty} q^y$$

$$= \sum_{x=0}^{n} (1-q)^2 q^x \frac{q^{n-x}}{1-q}$$

$$= \sum_{x=0}^{n} (1-q)\, q^n$$

$$= (n+1)(1-q)\, q^n$$

Pour $0 \leq k \leq n$:

$$\mathbf{P}(X = k \text{ et } X \leq n \leq X + Y) = \sum_{y=n-k}^{\infty} \mathbf{P}(X = k, Y = y)$$

$$= \sum_{y=n-k}^{\infty} (1-q)^2 q^k q^y = (1-q)q^n$$

donc $\quad \mathbf{P}(X = k \mid X \leq n \leq X + Y) = \dfrac{(1-q)q^n}{(n+1)(1-q)\, q^n} = \dfrac{1}{n+1}$

Il s'agit d'une loi uniforme sur $[\![0, n]\!]$.

INTERPRETATION : On répète des expériences de Bernoulli. X est le nombre d'échecs jusqu'au premier succès (au rang X + 1) dont on suppose qu'il a lieu au plus tard au (n + 1)-ème tirage, puis Y est le nombre d'échecs jusqu'au deuxième succès (au rang X + Y + 2) qu'on suppose avoir lieu au plus tôt au (n + 2)-ème tirage. Donc, parmi les n + 1 premières expériences de Bernoulli, on suppose qu'il y a eu un et un seul succès. Celui-ci peut se produire à n'importe quel tirage de façon uniforme entre 1 et n + 1.

Sol.3-5) $\mathbf{E}(\overline{X}) = \mathbf{E}(X) = m$

$\mathbf{E}(T^2) = \mathbf{V}(X) = \sigma^2$

Si on connaît l'espérance m de X, T^2 est une variable aléatoire dont l'espérance est égale à la variance de X. Noter le facteur $\dfrac{1}{n}$ de T.

Dans la suite, pour les calculs de $\mathbf{E}(S^2)$, $\text{cov}(\overline{X}, T^2)$, $\text{cov}(\overline{X}, S^2)$, on peut raisonner sur les variables centrées $X_i - m$. En effet, chacune des ces quantités est invariante par translation des X_i. On renomme $X_i - m$ en X_i, afin d'alléger les notations. α devient $\mathbf{E}(X^3)$ et $\mathbf{E}(\overline{X})$ devient 0.

$$S^2 = \frac{1}{n-1} \Big(\sum_{k=1}^{n} X_k^2 - 2X_k\overline{X} + n\overline{X}^2\Big) = \frac{1}{n-1} \Big(\sum_{k=1}^{n} X_k^2 - 2n\overline{X}^2 + n\overline{X}^2\Big)$$

$$= \frac{1}{n-1} \Big(\sum_{k=1}^{n} X_k^2 - n\overline{X}^2\Big)$$

donc $\mathbf{E}(S^2) = \dfrac{1}{n-1} \Big(n\sigma^2 - n \times \dfrac{1}{n^2} \sum_{k=1}^{n} \mathbf{E}(X_i^2)\Big)$ (les $\mathbf{E}(X_iX_j) = \mathbf{E}(X_i)\mathbf{E}(X_j)$ sont nuls)

$$= \frac{1}{n-1} (n\sigma^2 - \sigma^2) = \sigma^2$$

Si on ignore la valeur de l'espérance m de X, S^2 est une variable aléatoire dont l'espérance est égale à la variance de X. Noter le facteur $\frac{1}{n-1}$ de S. Certaines calculatrices scientifiques donnent les deux facons d'estimer cette variance à partir d'un échantillonnage, et il convient à l'utilisateur de savoir laquelle il doit utiliser, en fonction de sa connaissance ou non de l'espérance de X.

$\mathbf{E}(X_1T^2) = \frac{1}{n}\mathbf{E}(X_1^3) = \frac{\alpha}{n}$ (les $\mathbf{E}(X_1X_i^2)$ sont nuls, pour $i > 1$)

De même, en reprenant l'expression $S^2 = \frac{1}{n-1}(\sum_{k=1}^{n} X_i^2 - n\overline{X}^2)$:

$$\mathbf{E}(X_1S^2) = \frac{1}{n-1}(\sum_{k=1}^{n}\mathbf{E}(X_1X_k^2) - n\mathbf{E}(X_1\overline{X}^2))$$

$$= \frac{1}{n-1}(\alpha - \frac{1}{n}\mathbf{E}(X_1\sum_{k=1}^{n}X_k^2)) = \frac{\alpha}{n}$$

donc $\mathbf{E}(\overline{X}T^2) = \mathbf{E}(\overline{X}S^2) = \frac{\alpha}{n} = \mathrm{cov}(\overline{X}, T^2) = \mathrm{cov}(\overline{X}, S^2)$

◆