

# Espaces probabilisés

## I. Loi de Zipf

1) D'une part les réels  $P(\{k\})$  sont tous positifs. D'autre part

$$\sum_{k=1}^{+\infty} P(\{k\}) = \sum_{k=1}^{+\infty} \frac{1}{\zeta(a)k^a} = \frac{1}{\zeta(a)} \sum_{k=1}^{+\infty} \frac{1}{k^a} = \frac{\zeta(a)}{\zeta(a)} = 1.$$

Ils définissent donc bien une probabilité sur  $\mathbb{N}^*$ .

2) Soit  $m \in \mathbb{N}^*$ . Par le calcul, on trouve

$$P(m\mathbb{N}^*) = \sum_{k=1}^{+\infty} P(\{km\}) = \sum_{k=1}^{+\infty} \frac{1}{\zeta(a)(km)^a} = \frac{1}{m^a \zeta(a)} \sum_{k=1}^{+\infty} \frac{1}{k^a} = \frac{1}{m^a}.$$

3) Soit  $(i, j) \in (\mathbb{N}^*)^2$ . D'une part, d'après le calcul précédent,  $P(i\mathbb{N}^*)P(j\mathbb{N}^*) = 1/(ij)^a$ .

D'autre part  $i\mathbb{N}^* \cap j\mathbb{N}^*$  est l'ensemble des entiers qui sont multiples à la fois de  $i$  et de  $j$ . On a donc  $i\mathbb{N}^* \cap j\mathbb{N}^* = (i \vee j)\mathbb{N}^*$  (où  $i \vee j$  désigne le p.p.c.m. de  $i$  et de  $j$ ). Comme

$$P(i\mathbb{N}^* \cap j\mathbb{N}^*) = \frac{1}{(i \vee j)^a}.$$

les événements  $i\mathbb{N}^*$  et  $j\mathbb{N}^*$  sont indépendants si et seulement si  $ij = i \vee j$ , c'est-à-dire si  $i$  et  $j$  sont premiers entre eux.

4) a) On peut écrire  $C_n = \bigcap_{i=1}^n \overline{p_i\mathbb{N}^*}$ . En reprenant la question précédente, on prouve que les  $(p_i\mathbb{N}^*)_{i \in \mathbb{N}^*}$  sont mutuellement indépendants. Il en va donc de même de leurs complémentaires, donc

$$P(C_n) = \prod_{i=1}^n (1 - P(p_i\mathbb{N}^*)) = \prod_{i=1}^n \left(1 - \frac{1}{p_i^a}\right)$$

b) L'ensemble  $\bigcap_{n \geq 1} C_n$  est l'ensemble des entiers non nuls qui ne sont divisibles par aucun des nombres premiers. Évidemment il n'y a que 1 dans ce cas, et donc  $\bigcap_{n \geq 1} C_n = \{1\}$ .

c) Comme la suite  $(C_n)_{n \in \mathbb{N}^*}$  est décroissante, on a

$$\lim_{n \rightarrow +\infty} P(C_n) = \lim_{n \rightarrow +\infty} P\left(\bigcap_{n \geq 1} C_n\right) = P(\{1\}) = \frac{1}{\zeta(a)}$$

$$\text{donc } \zeta(a) = \lim_{n \rightarrow +\infty} \prod_{i=1}^n \left(1 - \frac{1}{p_i^a}\right)^{-1}.$$

## II. Théorème de Borel-Cantelli

1) Par stabilité d'une tribu par réunion dénombrable, puis par intersection dénombrable,  $B$  est bien un évènement.

2)  $\omega \in B$  si et seulement si pour tout  $n \in \mathbb{N}$ , il existe  $k \in \mathbb{N}$  tel que  $k \geq n$  et  $\omega \in A_k$ . Ainsi  $B$  est l'ensemble des éléments qui sont dans une infinité de  $A_k$ .

3) Posons, pour tout  $p \in \mathbb{N}$ ,  $U_p = \bigcup_{k=p}^{+\infty} A_k$  et  $B_p = \bigcap_{n=0}^p \bigcup_{k=n}^{+\infty} A_k = \bigcap_{n=0}^p U_n$ .

On remarque que la suite  $(U_p)_p$  est décroissante, et donc  $B_p = U_p = \bigcup_{n=p}^{+\infty} A_n$ . La suite des  $(B_p)_p$  est donc également décroissante, et par continuité décroissante

$$P(B) = \lim_{p \rightarrow +\infty} P\left(\bigcup_{n=p}^{+\infty} A_n\right).$$

Mais

$$P\left(\bigcup_{n=p}^{+\infty} A_n\right) \leq \sum_{n=p}^{+\infty} P(A_n)$$

Et la suite des restes d'une série convergente converge vers 0, ce qui permet de conclure.

4) a) Dans ce cas, pour tout  $p \in \mathbb{N}$  il existe  $n_0 \in \mathbb{N}$  tel que  $n_0 \geq p$  et  $P(A_{n_0}) = 1$ . Alors

$$1 \geq P\left(\bigcup_{n=p}^{+\infty} A_n\right) \geq P(A_{n_0}) = 1$$

et le résultat cherché s'ensuit directement car l'égalité  $P(B) =$

$$\lim_{p \rightarrow +\infty} P\left(\bigcup_{n=p}^{+\infty} A_n\right) \text{ est toujours valable.}$$

- b) Il faut avoir une première idée : l'indépendance se traduit mieux avec des intersections qu'avec des réunions d'événements. On va donc considérer l'évènement complémentaire de la limite supérieure :

$$\overline{B} = \bigcup_{p=0}^{+\infty} \left( \bigcap_{n=p}^{+\infty} \overline{A_n} \right)$$

Par continuité croissante,

$$P(\overline{B}) = \lim_{p \rightarrow +\infty} \left( P \left( \bigcap_{n=p}^{+\infty} \overline{A_n} \right) \right)$$

- c) On a par croissance, et ensuite par indépendance des  $\overline{A_n}$

$$\forall q \geq p \quad P \left( \bigcap_{n=p}^{+\infty} \overline{A_n} \right) \leq P \left( \bigcap_{n=p}^q \overline{A_n} \right) = \prod_{n=p}^q (1 - P(A_n))$$

- d) La deuxième idée est de prendre le logarithme, l'hypothèse portant sur une série. Pour cela nous utilisons qu'au moins à partir du rang  $p$ ,  $P(A_n) < 1$ . Alors, si  $p \leq q$ ,

$$\ln \left( \prod_{n=p}^q (1 - P(A_n)) \right) = \sum_{n=p}^q \ln(1 - P(A_n))$$

La suite  $(P(A_n))$  ne converge pas vers 0 a priori, donc on ne peut pas utiliser d'équivalents. On utilise donc une comparaison basée sur

$$\forall x > -1, \quad \ln(1+x) \leq x$$

Donc

$$\ln \left( \prod_{n=p}^q (1 - P(A_n)) \right) \leq - \sum_{n=p}^q P(A_n).$$

- e) Mais  $\sum P(A_n) = +\infty$ , donc

$$\sum_{n=p}^q P(A_n) \xrightarrow{q \rightarrow +\infty} +\infty$$

et donc

$$\ln \left( \prod_{n=p}^q (1 - P(A_n)) \right) \xrightarrow{q \rightarrow +\infty} -\infty$$

et, donc,

$$P \left( \bigcap_{n=p}^{+\infty} \overline{A_n} \right) = 0$$

Donc  $P(\overline{B}) = 0$ , ce qui conclut.

### III. Une suite arithmético-géométrique

- 1) On définit l'évènement  $U$  (respectivement  $V$ ) « le premier tirage s'effectue dans l'urne  $U$  » (respectivement « le premier tirage s'effectue dans l'urne  $V$  »). La famille  $(U, V)$  est un système complet d'événements. D'après la formule des probabilités totales :

$$\begin{aligned} p_1 &= P(B_1) = P(B_1 \cap U) + P(B_1 \cap V) \\ &= P_U(B_1)P(U) + P_V(B_1)P(V). \end{aligned}$$

Compte tenu de l'énoncé :

$$p_1 = \frac{2}{6} \times \frac{1}{2} + \frac{3}{6} \times \frac{1}{2} = \frac{5}{12}.$$

- 2) La famille  $(B_n, \overline{B}_n)$  est un système complet d'événements. D'après la formule des probabilités totales :

$$\begin{aligned} p_{n+1} &= P(B_{n+1}) = P_{B_n}(B_{n+1})P(B_n) + P_{\overline{B}_n}(B_{n+1})P(\overline{B}_n) \\ &= \frac{2}{6}p_n + \frac{3}{6}(1 - p_n). \end{aligned}$$

Ainsi,

$$\forall n \in \mathbb{N}^*, p_{n+1} = -\frac{1}{6}p_n + \frac{1}{2}$$

- 3) La suite  $(p_n)_{n \geq 1}$  est une suite arithmético-géométrique. Déterminons une expression de  $p_n$  en fonction de  $n$ .

L'équation  $-\frac{1}{6}x + \frac{1}{2} = x$  admet l'unique solution  $\ell = \frac{3}{7}$ . On définit alors la suite  $(u_n)_{n \geq 1}$  par  $u_n = p_n - \ell$  pour tout  $n \geq 1$ . La suite  $(u_n)_{n \geq 1}$  est géométrique de raison  $-\frac{1}{6}$  donc pour tout  $n \geq 1$ ,  $u_n = \left(-\frac{1}{6}\right)^{n-1} u_1$ . Or

$$u_1 = p_1 - \ell = \frac{5}{12} - \frac{3}{7} = -\frac{1}{84}.$$

On en déduit que

$$\forall n \geq 1, p_n = -\frac{1}{84} \left(-\frac{1}{6}\right)^{n-1} + \frac{3}{7}.$$

On vérifie que  $p_1$  est bien égal à  $\frac{5}{12}$ .

Comme la suite géométrique  $\left(\left(-\frac{1}{6}\right)^{n-1}\right)_{n \geq 1}$  converge vers 0,  $p_n \xrightarrow{n \rightarrow +\infty}$

$$\frac{3}{7}.$$

#### IV. Une chaîne de Markov

1)

$$\forall i \in \llbracket 0, 3 \rrbracket, P(x_{n+1} = i) = \sum_{j=0}^3 P_{(x_n=j)}(x_{n+1} = i) P(x_n = j)$$

Les probabilités conditionnelles sont données par l'énoncé. Ces égalités s'écrivent matriciellement  $X_{n+1} = MX_n$  avec  $M$  la matrice :

$$M = \begin{pmatrix} 1 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 1-p & 0 & 0 \\ 0 & 0 & 1-p & 1 \end{pmatrix}$$

2) Le polynôme caractéristique de  $M$  est

$$\chi_M(x) = \begin{vmatrix} x-1 & -p & 0 & 0 \\ 0 & x & -p & 0 \\ 0 & p-1 & x & 0 \\ 0 & 0 & p-1 & x-1 \end{vmatrix} = (x-1)^2 [x^2 - (p(1-p))]$$

par des développements successifs. Le spectre de  $M$  est donc

$$\text{Sp}(M) = \{1, \pm \sqrt{p(1-p)}\}$$

- Si  $p \in ]0, 1[$  alors  $M$  possède trois valeurs propres distinctes, et on observe sur la 1ère et la dernière colonne de  $M$ , que  $\dim E_M(1) \geq 2$ . La matrice  $M$  est donc diagonalisable.
- Si  $p = 0$  ou  $p = 1$  alors  $\text{Sp}(M) = \{1, 0\}$ .

Dans ce cas,  $M$  est de rang 3 donc  $\text{Ker}(f)$  est de dimension 1 alors que 0 est une valeur propre double. Donc  $M$  n'est pas diagonalisable.

- 3) Les 1ers et 4èmes vecteurs de la base canonique,  $e_1$  et  $e_4$ , forment une base de  $E_1$ . Si  $X_0 = (a, b, c, d)$ , alors  $a + b + c + d = 1$ , et  $X_n = M^n X_0 = ae_1 + \frac{1}{2^n}(be_2 + ce_3) + de_4 \xrightarrow{n \rightarrow +\infty} ae_1 + be_4$ .
- 4) Posons  $v_1 = (1, -1, -1, 1)$  et  $v_3 = (1, -3, 3, -1)$ . Alors  $(e_1, v_2, v_3, e_4)$  est une base de vecteurs propres de  $M$ . De plus  $X_0 = \frac{1}{6}[3e_1 - v_2 + v_3 + 3e_4]$ , donc  $\lim_{n \rightarrow \infty} X_n = (1/2, 0, 0, 1/2)$ .

#### V. Apparition d'un double Pile

1) Le support de  $T$  est  $T(\Omega) = \llbracket 2, +\infty \rrbracket$ .

Pour  $k$  dans  $\mathbb{N}$ , étudions l'événement  $(T > k) = \llcorner \text{le motif PP n'est pas apparu avant le rang } k \rceil$  dont la probabilité est notée  $p_k$ . On a  $p_0 = p_1 = 1$ . La suite d'événements  $(T > k)_{k \geq 0}$  est une suite décroissante pour l'inclusion. Ainsi

$$P(T = k) = P(T > k-1) - P(T > k)$$

Notons  $X_k$  la variable aléatoire égale à 1 si au  $k$ -ème lancer on obtient pile et égale à 0 sinon. Les événements  $(X_k = 0), (X_k = 1)$  forment un système complet d'événements. On a ainsi la décomposition :

$$(T > k) = ((T > k) \cap (X_k = 1)) \sqcup ((T > k) \cap (X_k = 0)) \quad (\star)$$

Supposons  $k \geq 2$ .

Si au  $k$ -ème lancer le résultat est pile et si le motif PP n'est pas apparu avant le rang  $k$ , le  $(k-1)$ -ème lancer est nécessairement face. De plus, le motif PP n'a pas pu apparaître avant le  $(k-2)$ -ème lancer :

1	...	$k-2$	$k-1$	$k$
	...		$F$	$P$

On a donc

$$P((T > k) \cap (X_k = 1)) = P((T > k-2) \cap (X_{k-1} = 0) \cap (X_k = 1))$$

Par indépendance des lancers :

$$P((T > k) \cap (X_k = 1)) = p_{k-2} \times \frac{1}{3} \times \frac{2}{3} = \frac{2}{9} p_{k-2}$$

Si au  $k$ -ème lancer le résultat est face et si le motif PP n'est pas apparu avant le rang  $k$ , alors le motif PP n'est pas apparu avant le  $(k-1)$ -ème lancer. On a donc

$$P((T > k) \cap (X_k = 0)) = P((T > k-1) \cap (X_k = 0)) = p_{k-1} \times \frac{1}{3}.$$

D'après la réunion disjointe  $(\star)$ ,

$$p_k = \frac{2}{9}p_{k-2} + \frac{1}{3}p_{k-1}$$

de sorte que la suite  $(p_k)_{k \geq 0}$  vérifie la relation de récurrence linéaire d'ordre 2 :

$$\forall k \in \mathbb{N}, 9p_{k+2} - 3p_{k+1} - 2p_k = 0 \text{ et } p_0 = p_1 = 1$$

2) L'équation caractéristique est  $9r^2 - 3r - 2 = 0$  de racines  $-1/3$  et  $2/3$ . Il existe des constantes  $\alpha$  et  $\beta$  telles que

$$\forall k \in \mathbb{N}, p_k = \alpha \left(\frac{2}{3}\right)^k + \beta \left(-\frac{1}{3}\right)^k$$

Les données initiales  $p_0 = p_1 = 1$  donnent  $\alpha = 4/3$  et  $\beta = -1/3$ . Ainsi,

$$\forall k \geq 0, p_k = \frac{4}{3} \left(\frac{2}{3}\right)^k + \left(-\frac{1}{3}\right)^{k+1}$$

Pour  $k \geq 1$ ,

$$P(T = k) = p_{k-1} - p_k = \left(\frac{2}{3}\right)^{k+1} + \frac{4}{3} \left(-\frac{1}{3}\right)^k$$

Commentaires :

La formule donne bien la probabilité attendue  $P(T = 1) = 0$ . De plus,

$$P(T = 2) = P(X_1 = 1 \cap X_2 = 1) = \frac{4}{9}$$

ce qui est également cohérent.

Un calcul de somme donne  $\sum_{k=1}^{\infty} P(T = k) = 1$ .

3) Notons  $(T = +\infty)$  l'événement « le motif PP n'apparaît pas ». On a

$$(T = +\infty) = \bigcap_{k=0}^{+\infty} (T > k)$$

Par continuité décroissante (la famille  $(T > k)_{k \geq 0}$  est décroissante pour l'inclusion) :

$$P(T = +\infty) = P\left(\bigcap_{k=0}^{+\infty} (T > k)\right) = \lim_{k \rightarrow \infty} P(T > k) = \lim_{k \rightarrow \infty} p_k = 0$$

Ainsi, le motif PP arrive presque sûrement.