

# Jeu du pile ou face et problème du collectionneur de coupons

Alexandre Drewery      Victor Voisin

9 avril 2021

Lecture dirigée sous la tutelle d'Émilien Manent.

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Distance en variation totale</b>	<b>2</b>
2.1	Définition et propriétés . . . . .	2
2.2	Application à la convergence en loi . . . . .	5
<b>3</b>	<b>Approximation binomiale-Poisson</b>	<b>6</b>
3.1	Convergence en loi . . . . .	7
3.2	Inégalité de Le Cam et vitesse de convergence . . . . .	8
3.3	Exemples d'applications . . . . .	10
3.3.1	Anniversaire et Nouvel An . . . . .	10
3.3.2	Objets défectueux . . . . .	10
<b>4</b>	<b>Loi de l'arcsinus</b>	<b>11</b>
4.1	Contexte et mise en situation . . . . .	11
4.2	Marches aléatoires, dernier retour en zéro . . . . .	11
4.3	Loi discrète de l'arcsinus . . . . .	14
4.4	Convergence vers la loi de l'arcsinus . . . . .	15
4.5	Principe de réflexion, théorème du scrutin et lemme principal . . . . .	17
<b>5</b>	<b>Collectionneur de coupons</b>	<b>19</b>
5.1	Présentation du modèle aléatoire . . . . .	19
5.2	Étude de la loi de $T$ . . . . .	20
5.3	Étude asymptotique . . . . .	21
<b>6</b>	<b>Conclusion</b>	<b>24</b>

## 1 Introduction

On étudie dans ce document plusieurs expériences basées sur un enchaînement de parties d'un jeu simple (type pile ou face ou lancer de dé) et les lois de probabilité qu'elles modélisent.

Lorsque l'on répète une expérience de Bernoulli de faible probabilité de succès (type lancer de pièce déséquilibrée) un grand nombre de fois, on observe une convergence de la loi associée vers une loi de Poisson. Cette convergence et sa vitesse sont détaillées.

On peut également s'intéresser au problème des changements de leader au cours d'un long duel de pile ou face. On utilise des marches aléatoires pour modéliser ce problème et on montre que l'instant du dernier changement de leader suit une loi qu'on appelle loi de l'arcsinus.

Enfin, on se penche sur le problème du collectionneur de coupons : le nombre d'expériences uniformes nécessaire pour en observer tous les résultats possibles. On détaille les caractéristiques de la loi qui correspond à ce scénario.

Ce rapport est illustré de quelques figures réalisées en Python.

## 2 Distance en variation totale

Les résultats importants de ce rapport permettent d'établir la convergence de lois de probabilité vers d'autres lois de probabilité. On se dote pour établir ces convergences d'une distance entre lois : la distance en variation totale. Celle-ci est définie et plusieurs résultats associés sont établis dans cette partie, issue de l'étude du texte de Chafai [1].

### 2.1 Définition et propriétés

**Définition 1.** Soit  $E$  un ensemble au plus dénombrable, muni de la topologie et de la tribu discrètes  $P(E)$ . On note  $LP(E)$  l'ensemble des lois de probabilité sur  $E$ . On définit sur  $LP(E)^2$  l'application suivante :

$$d_{VT} : \begin{cases} LP(E)^2 & \rightarrow & \mathbb{R} \\ (f, g) & \mapsto & \sup_{A \subseteq E} |\mu(A) - \nu(A)| \end{cases}$$

**Théorème 1.** Dans le cadre de la définition précédente,  $d_{VT}$  définit bien une distance sur  $LP(E)$ .

*Démonstration.* Montrons que  $d_{VT}$  vérifie les axiomes de la distance.

— Pour  $\mu, \nu \in LP(E)$  :

$$d_{VT}(\mu, \nu) = \sup_{A \subseteq E} |\mu(A) - \nu(A)| = \sup_{A \subseteq E} |\nu(A) - \mu(A)| = d_{VT}(\nu, \mu)$$

— Soient  $\mu, \nu \in LP(E)$ .

On a bien  $d_{VT}(\mu, \mu) = 0$

Supposons que  $d_{VT}(\mu, \nu) = 0$ . Alors :

$$\forall x \in E, 0 \leq |\nu(A) - \mu(A)| \leq \sup_{B \subseteq E} |\nu(B) - \mu(B)| = 0$$

D'où  $\mu = \nu$ .

— Soient  $\mu, \nu, \pi \in LP(E)$ .

Alors :

$$\begin{aligned} d_{VT}(\mu, \pi) &= \sup_{A \subseteq E} |\mu(A) - \pi(A)| \\ &\leq \sup_{A \subseteq E} (|\mu(A) - \nu(A)| + |\nu(A) - \pi(A)|) \\ &\leq \sup_{B \subseteq E} (\sup_{A \subseteq E} |\mu(A) - \nu(A)| + \sup_{A \subseteq E} |\nu(A) - \pi(A)|) \\ &= \sup_{A \subseteq E} |\mu(A) - \nu(A)| + \sup_{A \subseteq E} |\nu(A) - \pi(A)| \\ &= d_{VT}(\mu, \nu) + d_{VT}(\nu, \pi) \end{aligned}$$

□

**Remarque 1.** L'espace métrique  $(LP(E), d_{VT})$  est complet.

Une notion de distance a pu être imposée au sein de l'ensemble des lois de probabilité sur  $E$ . L'idée d'une telle distance est de pouvoir estimer le plus grand éloignement possible des mesures sur une partie de  $E$ . On pourra ainsi faire des majorations uniformes, à condition qu'elles soient intéressantes et utiles. Il est par conséquent essentiel de chercher une expression générale et facile à manipuler de la distance en variation totale.

On se donne à présent un ensemble  $E \neq \emptyset$  au plus dénombrable, pour lequel on va noter  $LP(E)$  l'ensemble des lois de probabilité sur  $E$ , muni de la topologie et la tribu discrètes  $P(E)$ .

**Théorème 2.** Soient  $\mu, \nu$  des lois de probabilité sur  $E$ . Alors :

$$d_{VT}(\mu, \nu) = \frac{1}{2} \sup_{f: E \rightarrow [-1,1]} \left| \int f d\mu - \int f d\nu \right| = \frac{1}{2} \sum_{x \in E} |\mu(x) - \nu(x)|$$

où pour tout  $x \in E$ , on écrit  $\mu(x) = \mu(\{x\})$  et  $\nu(x) = \nu(\{x\})$ .

De plus, la borne supérieure  $\sup_{A \subseteq E} |\mu(A) - \nu(A)|$  est atteinte pour l'ensemble  $A_* = \{x \in E / \mu(x) \geq \nu(x)\}$  et la borne supérieure  $\sup_{f: E \rightarrow [-1,1]} \left| \int f d\mu - \int f d\nu \right|$  est atteinte pour  $f_0 = \mathbb{1}_{A_*} - \mathbb{1}_{A_*^c}$ .

*Démonstration.* On remarque que pour toute fonction  $f \in [-1, 1]^E$ ,  $f$  est  $\mu$ -intégrable sur  $E$  et  $\nu$ -intégrable sur  $E$  car  $f$  est bornée et  $\mu$  et  $\nu$  sont des mesures finies, et on a :

$$\int f d\mu = \sum_{x \in E} \mu(x) f(x)$$

$$\int f d\nu = \sum_{x \in E} \nu(x) f(x)$$

(les intégrales de  $f^+$  et  $f^-$  selon les mesures  $\mu$  et  $\nu$  sont de la même forme et s'obtiennent facilement soit en tant que fonctions étagées, soit par le Théorème de convergence monotone avec des fonctions étagées car  $E$  est au plus dénombrable)

D'où, par somme :

$$\begin{aligned} \forall f \in [-1, 1]^E, \quad \left| \int f d\mu - \int f d\nu \right| &= \left| \sum_{x \in E} (\mu(x) - \nu(x)) f(x) \right| \\ &\leq \sum_{x \in E} |\mu(x) - \nu(x)| |f(x)| \\ &\leq \sum_{x \in E} |\mu(x) - \nu(x)| \\ &\quad \text{(car } f \text{ est bornée par 1)} \end{aligned}$$

Soit  $f_0 = \mathbb{1}_{A_*} - \mathbb{1}_{A_*^c}$ . Alors :

$$\begin{aligned} \left| \int f_0 d\mu - \int f_0 d\nu \right| &= \left| \sum_{x \in A_*} (\mu(x) - \nu(x)) + \sum_{x \in A_*^c} (\mu(x) - \nu(x)) \right| \\ &= \sum_{x \in A_*} (\mu(x) - \nu(x)) + \sum_{x \in A_*^c} (\mu(x) - \nu(x)) \quad \text{(par définition de } A_*) \\ &= \sum_{x \in A} |\mu(x) - \nu(x)| \end{aligned}$$

Donc,  $\left\{ \left| \int f d\mu - \int f d\nu \right| \right\}_{f \in [-1, 1]^E}$  est une partie non vide de  $\mathbb{R}$  (car  $E \neq \emptyset$ ) qui admet un maximum atteint en  $f_0$ , *i.e.* admet une borne supérieure et :

$$\frac{1}{2} \sup_{f: E \rightarrow [-1, 1]} \left| \int f d\mu - \int f d\nu \right| = \frac{1}{2} \sum_{x \in E} |\mu(x) - \nu(x)|$$

$$\begin{aligned} \forall A \subset E, \quad |\mu(A) - \nu(A)| &= \left| \sum_{x \in A} \mu(x) - \sum_{x \in A} \nu(x) \right| \\ &= \left| \sum_{x \in A} (\mu(x) - \nu(x)) \right| \\ &\leq \sum_{x \in A} |\mu(x) - \nu(x)| \end{aligned}$$

$$\begin{aligned} \text{et } \forall A \subset E, \quad |\mu(A) - \nu(A)| &= |(1 - \mu(A^c)) - (1 - \nu(A^c))| \\ &= |\mu(A^c) - \nu(A^c)| \\ &\leq \sum_{x \in A^c} |\mu(x) - \nu(x)| \end{aligned}$$

D'où, par demi-somme :

$$\forall A \subset E, |\mu(A) - \nu(A)| \leq \frac{1}{2} \sum_{x \in E} |\mu(x) - \nu(x)|$$

Or, en reprenant l'idée ci-dessus :

$$\begin{aligned} 2 |\mu(A_*) - \nu(A_*)| &= |\mu(A_*) - \nu(A_*)| + |\mu(A_*^c) - \nu(A_*^c)| \\ &= \sum_{x \in A_*} (\mu(x) - \nu(x)) + \sum_{x \in A_*^c} (\nu(x) - \mu(x)) \quad \text{(par définition de } A_*) \\ &= \sum_{x \in E} |\mu(x) - \nu(x)| \end{aligned}$$

D'où :

$$|\mu(A_*) - \nu(A_*)| = \frac{1}{2} \sum_{x \in E} |\mu(x) - \nu(x)|$$

Donc, on a bien  $d_{VT}(\mu, \nu) = \frac{1}{2} \sum_{x \in E} |\mu(x) - \nu(x)|$  et  $\sup_{A \subseteq E} |\mu(A) - \nu(A)|$  est atteint en  $A_*$ .

□

**Remarque 2.** L'écriture de  $d_{VT}(\cdot, \cdot)$  avec les intégrales est nommée *expression variationnelle fonctionnelle* de  $d_{VT}(\cdot, \cdot)$ .

Ces expressions permettent souvent un calcul pratique de la distance en variation totale bien plus simple.

**Exemple 1.** En guise d'illustration, calculons  $d_{VT}(\mu, \nu)$  où  $\mu$  est une loi de Poisson de paramètre  $p$  et  $\nu$  une loi de Bernoulli de paramètre  $p$  :

$$\begin{aligned} d_{VT}(\mu, \nu) &= \frac{1}{2} \sum_{x \in \mathbb{N}} |\mu(x) - \nu(x)| \\ &= \frac{1}{2} \left( |\mu(0) - \nu(0)| + |\mu(1) - \nu(1)| + \sum_{x \geq 2} |\mu(x) - \nu(x)| \right) \\ &= \frac{1}{2} \left( |e^{-p} - (1-p)| + |pe^{-p} - p| + \sum_{x \geq 2} \mu(x) \right) \\ &= \frac{1}{2} (e^{-p} - 1 + p + |p||e^{-p} - 1| + (1 - \mu(0) - \mu(1))) \\ &= \frac{1}{2} (e^{-p} - 1 + p + p(1 - e^{-p}) + 1 - e^{-p} - pe^{-p}) \\ &= \frac{1}{2} (p + p(1 - e^{-p}) - pe^{-p}) \\ d_{VT}(\mu, \nu) &= p(1 - e^{-p}) \end{aligned}$$

**Théorème 3.** Pour  $\mu, \nu$  lois sur  $E$ , on a :

$$d_{VT}(\mu, \nu) = 1 - \sum_{x \in E} \min(\mu(x), \nu(x))$$

*Démonstration.* Puisque pour tous  $a, b$ ,  $\min(a, b) = \frac{1}{2}(a + b - |a - b|)$ , on a :

$$\begin{aligned} \sum_{x \in E} \min(\mu(x), \nu(x)) &= \frac{1}{2} \sum_{x \in E} \mu(x) + \nu(x) - |\mu(x) - \nu(x)| \\ &= \frac{1}{2} \left( \sum_{x \in E} \mu(x) + \sum_{x \in E} \nu(x) - \sum_{x \in E} |\mu(x) - \nu(x)| \right) \\ &= \frac{1}{2} (1 + 1 - 2d_{VT}(\mu, \nu)) \\ &= 1 - d_{VT}(\mu, \nu) \end{aligned}$$

D'où l'égalité souhaitée.

□

**Théorème 4** (Théorème de couplage). Pour  $\mu$  et  $\nu$  des lois sur  $E$ , on a :

$$d_{VT}(\mu, \nu) = \inf_{X \hookrightarrow \mu, Y \hookrightarrow \nu} P(X \neq Y)$$

(la borne inférieure portant sur tous les couples de variables aléatoires  $X$  et  $Y$  de lois marginales  $\mu$  et  $\nu$ )

De plus, il existe un tel couple  $(X, Y)$  pour lequel  $d_{VT}(\mu, \nu) = P(X \neq Y)$ .

*Démonstration.* Soit  $(X, Y)$  un couple de variables aléatoires sur  $E \times E$  de lois marginales  $\mu$  et  $\nu$ . Comme  $P(X = x \text{ et } Y = y) \leq \min\{\mu(x), \nu(y)\}$  pour tout  $x \in E$  :

$$1 - d_{VT}(\mu, \nu) = \sum_{x \in E} \min\{\mu(x), \nu(x)\} \geq \sum_{x \in E} P(X = x \text{ et } Y = x) = P(X = Y)$$

D'où :

$$P(X \neq Y) = 1 - P(X = Y) \geq d_{VT}(\mu, \nu)$$

Construisons un couple  $(X, Y)$  pour lequel l'égalité est atteinte. Posons  $p = 1 - d_{VT}(\mu, \nu) \in [0, 1]$ . On distingue trois cas :

— Cas où  $p = 0$  :

On prend  $(X, Y)$  avec  $X$  et  $Y$  indépendantes de lois marginales  $\mu$  et  $\nu$ . Puisque  $d_{VT}(\mu, \nu) = 1$ ,  $\mu$  et  $\nu$  ont des supports disjoints, d'où :

$$P(X = Y) = \sum_{x \in E} P(X = x \text{ et } Y = x) = \sum_{x \in E} \mu(x)\nu(x) = 0$$

— Cas où  $p = 1$  :

On a  $d_{VT}(\mu, \nu) = 0$ , ie  $\mu = \nu$ . On prend  $(X, Y)$  où  $X \hookrightarrow \mu$  et  $X = Y$ .

— Cas où  $0 < p < 1$  : Soient  $U, V, W$  des variables aléatoires de lois respectives  $\frac{1}{p} \min\{\mu, \nu\}$ ,  $\frac{1}{1-p}(\mu - \min\{\mu, \nu\})$ ,  $\frac{1}{1-p}(\nu - \min\{\mu, \nu\})$ .

Notons que par Théorème :

$$p = \sum_{x \in E} \min\{\mu(x), \nu(x)\}$$

Soit  $B$  une variable aléatoire de loi de Bernoulli de paramètre  $p$  indépendante du triplet  $(U, V, W)$ . Posons  $(X, Y) = (U, U)$  si  $B = 1$  et  $(X, Y) = (V, W)$  si  $B = 0$ .

Montrons que  $X \hookrightarrow \mu$  et  $Y \hookrightarrow \nu$ .

Soit  $(x, y) \in \text{Im}(X) \times \text{Im}(Y)$ . On a :

$$\begin{aligned} P(X = x) &= P((X = x) \cap (B = 1)) + P((X = x) \cap (B = 0)) \\ &= P_{(B=1)}(X = x) P(B = 1) + P_{(B=0)}(X = x) P(B = 0) \\ &= pP(U = x) + (1 - p)P(V = x) \\ &= \min\{\mu(x), \nu(x)\} + \mu(x) - \min\{\mu(x), \nu(x)\} \\ &= \mu(x) \end{aligned}$$

$$\begin{aligned} P(Y = y) &= P((Y = y) \cap (B = 1)) + P((Y = y) \cap (B = 0)) \\ &= P_{(B=1)}(Y = y) P(B = 1) + P_{(B=0)}(Y = y) P(B = 0) \\ &= pP(U = y) + (1 - p)P(W = y) \\ &= \min\{\mu(x), \nu(x)\} + \nu(x) - \min\{\mu(x), \nu(x)\} \\ &= \nu(x) \end{aligned}$$

Donc  $X \hookrightarrow \mu$  et  $Y \hookrightarrow \nu$ .

Puisque les lois de  $V$  et  $W$  ont des supports disjoints, on a  $P(V = W) = 0$ , donc  $P(X = Y) = P(B = 1) = p$ , ie :

$$P(X \neq Y) = 1 - p = d_{VT}(\mu, \nu)$$

□

Le théorème de couplage permet de majorer la distance entre deux lois en construisant explicitement deux variables aléatoires  $X \hookrightarrow \mu, Y \hookrightarrow \nu$  pour lesquelles on peut calculer  $P(X \neq Y)$ . On a alors  $d_{VT}(\mu, \nu) \leq P(X \neq Y)$ .

## 2.2 Application à la convergence en loi

On dispose de techniques de majoration ou de calcul de la distance entre deux lois, on peut maintenant se pencher sur le concept de convergence :

**Théorème 5.** Soient les  $(X_n)_{n \geq 1}$  des variables aléatoires sur  $E$  de lois  $(\mu_n)_{n \geq 1}$  et  $\mu$  une loi sur  $E$ . Les affirmations suivantes sont équivalentes :

1. pour toute fonction bornée  $f : E \rightarrow \mathbb{R}$ ,  $\lim_{n \rightarrow +\infty} \int f d\mu_n = \int f d\mu$  ;
2. pour tout  $x \in E$ ,  $\lim_{n \rightarrow +\infty} \mu_n(x) = \mu(x)$  ;
3.  $\lim_{n \rightarrow +\infty} d_{VT}(\mu_n, \mu) = 0$ .

Si ces affirmations sont vérifiées, on dit que  $(X_n)$  converge en loi vers  $\mu$  quand  $n \rightarrow +\infty$ .

*Démonstration.* **1. implique 2.** : s'obtient en considérant  $f = 1_{\{x\}}$ . En effet, pour tout  $x \in E$  :

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mu_n(x) &= \lim_{n \rightarrow +\infty} \int 1_{\{x\}} d\mu_n \\ &= \int 1_{\{x\}} d\mu = \mu(x) \end{aligned}$$

**2. implique 3.** : soit  $A \subset E$ . Pour tout  $n \in \mathbb{N}$ , on a :

$$d_{VT} = \frac{1}{2} \sum_{x \in E} |\mu_n(x) - \mu(x)| = \frac{1}{2} \left( \sum_{x \in A} |\mu_n(x) - \mu(x)| + \sum_{x \in A^C} |\mu_n(x) - \mu(x)| \right)$$

On cherche à majorer ces deux sommes. La seconde est sujette à la majoration suivante :

$$\begin{aligned} \sum_{x \in A^C} |\mu_n(x) - \mu(x)| &\leq \sum_{x \in A^C} \mu_n(x) + \sum_{x \in A^C} \mu(x) \\ &= \left( 1 - \sum_{x \in A} \mu_n(x) \right) + \sum_{x \in A^C} \mu(x) \\ &= \left( \sum_{x \in A} \mu(x) + \sum_{x \in A^C} \mu(x) - \sum_{x \in A} \mu_n(x) \right) + \sum_{x \in A^C} \mu(x) \\ &= \sum_{x \in A} \mu(x) - \mu_n(x) + 2 \sum_{x \in A^C} \mu(x) \\ &\leq \sum_{x \in A} |\mu_n(x) - \mu(x)| + 2 \sum_{x \in A^C} \mu(x) \end{aligned}$$

Comme  $E$  est dénombrable et que  $\mu(E) = \sum_{x \in E} \mu(x) = 1$ , pour  $\epsilon > 0$ , on peut prendre  $A$  finie telle que  $\mu(A^C) < \epsilon$ . De plus,  $A$  étant maintenant finie,  $\sum_{x \in A} |\mu_n(x) - \mu(x)|$  est une somme finie de termes qui tendent vers 0 quand  $n$  tend vers l'infini (d'après 2.). Pour tout  $\epsilon' > 0$ , on peut donc trouver  $N \in \mathbb{N}$  tel quel  $\forall n \geq N$  :

$$\sum_{x \in A} |\mu_n(x) - \mu(x)| \leq \epsilon'$$

Ainsi, pour tout  $n \geq N$  on a :

$$\sum_{x \in A} |\mu_n(x) - \mu(x)| + \sum_{x \in A^C} |\mu_n(x) - \mu(x)| \leq 2 \sum_{x \in A} |\mu_n(x) - \mu(x)| + 2 \sum_{x \in A^C} \mu(x) \leq 2\epsilon' + 2\epsilon$$

Et donc :

$$d_{VT}(\mu_n, \mu) \leq \epsilon' + \epsilon$$

Ce qui nous donne bien 3.

**3. implique 1.** : soit  $f : E \rightarrow \mathbb{R}$  bornée. Si  $f$  n'est pas à valeurs dans  $[-1, 1]$ , on s'y ramène en la divisant par  $\sup_{x \in E} |f(x)|$ .

On a alors :

$$\left| \int f d\mu_n - \int f d\mu \right| \leq \sup_{f: E \rightarrow [-1, 1]} \left| \int f d\mu_n - \int f d\mu \right| = 2d_{VT}(\mu_n, \mu)$$

Comme  $\lim_{n \rightarrow +\infty} d_{VT}(\mu_n, \mu) = 0$  d'après 3., on obtient :

$$\lim_{n \rightarrow +\infty} \left| \int f d\mu_n - \int f d\mu \right| = 0$$

Et ce pour toute  $f$  bornée. On retrouve bien 1. □

**Remarque 3.** Pour  $(\mu_n)_{n \geq 1}$  lois sur  $E$  telles que pour tout  $x \in E$ ,  $(\mu_n(x))$  converge, l'application  $\mu$  définie par  $\mu(x) := \lim_{n \rightarrow +\infty} \mu_n(x)$  n'est toujours une loi que si  $E$  est fini.

Par exemple, les lois sur  $\mathbb{N}$   $\mu_n$ , qui associent  $n^{-1}$  aux  $x < n$  engendrent  $\mu$  identiquement nulle sur  $\mathbb{N}$ .

On a maintenant tous les outils nécessaires sur la distance en variation totale et ses différentes écritures afin de légitimer au mieux la convergence de loi de probabilités.

### 3 Approximation binomiale-Poisson

Le jeu du Pile ou Face est plus souvent étudié dans le cas où on fait un grand nombre de répétitions de ce jeu, car il est plus intéressant d'étudier les résultats d'une expérience à deux issues sur une population nombreuse que de les regarder au cas pas pas (auquel cas on observerait que "pile" ou "face", ce qui ne présente pas grand intérêt). La loi de probabilité associée à cette expérience est la loi binomiale, qui présente deux paramètres qui décrivent le nombre de répétitions et la probabilité associée au succès.

Mais en se plongeant dans le calcul des probabilités des différents résultats possibles à l'issue d'un grand nombre de répétitions du jeu pile ou face, on se rend compte que le nombre d'opérations nécessaires peut rapidement s'avérer être important (pour le calcul de  $\binom{n}{k}$  il est parfois nécessaire de faire  $n$  multiplications pour aboutir à son résultat). Dans le but d'optimiser le temps de calcul, il est préférable de faire une approximation de ces expériences par une autre simulation probabiliste dont le calcul des probabilités associées à sa loi est peu coûteux, quitte à générer une erreur numérique tant qu'elle reste moindre. C'est l'objectif qu'on veut atteindre à travers l'étude du premier modèle aléatoire, qui va montrer une manière d'approcher la loi binomiale par une loi de Poisson.

Il y a deux temps dans l'ensemble des justifications de l'approximation binomiale-Poisson. On prouve en premier lieu la convergence en loi vers une loi de Poisson avant d'estimer la vitesse de cette convergence ; sont rapportés ici les résultats de notre étude du document de Chaffai [1].

### 3.1 Convergence en loi

**Théorème 6.** On se donne  $\lambda \in \mathbb{R}_+^*$  et  $(p_n)_{n \in \mathbb{N}} \in ]0, 1[^\mathbb{N}$  avec :

$$p_n \underset{n \rightarrow \infty}{\sim} \frac{\lambda}{n}$$

Soit  $(\mu_n)_{n \in \mathbb{N}}$  suite de lois binomiales de paramètres respectifs  $n$ ,  $p_n$ , et  $\mu$  une loi de Poisson de paramètre  $\lambda$ . Alors :

$$\mu_n \xrightarrow{n \rightarrow \infty} \mu$$

*Démonstration.* Soit  $p \in ]0, 1[$ . On établit d'abord la formule de la différence entre les probabilités associées à une loi de Poisson de paramètre  $np$  et les lois binomiales de paramètre de succès  $p$ .

Pour tout  $n \in \mathbb{N}$ , on se donne une variable aléatoire  $S_n$  telle que  $S_n \hookrightarrow \mathcal{B}(n, p)$ .

Pour tout  $n \in \mathbb{N}$ , pour tout  $k \in \mathbb{N}$ , on a :

$$\begin{aligned} P(S_n = k) - e^{-np} \frac{(np)^k}{k!} &= \frac{n(n-1)\dots(n-k+1)}{k!} p^k (1-p)^{n-k} - e^{-np} \frac{(np)^k}{k!} \\ &= \frac{(np)^k}{k!} \left( \left( \prod_{i=0}^{k-1} \frac{n-i}{n(1-p)} \right) (1-p)^n - e^{-np} \right) \end{aligned}$$

On se donne  $\lambda \in \mathbb{R}_+^*$  et notre suite  $(p_n)_{n \in \mathbb{N}} \in ]0, 1[^\mathbb{N}$  avec :

$$p_n \underset{n \rightarrow \infty}{\sim} \frac{\lambda}{n}$$

Cette fois, pour tout  $n \in \mathbb{N}$ ,  $S_n \hookrightarrow \mathcal{B}(n, p_n)$ .

Alors,  $(np_n)_{n \in \mathbb{N}}$  est bornée car convergente, ie  $\left( \frac{(np_n)^k}{k!} \right)_{n \in \mathbb{N}}$  est bornée pour tout  $k \in \mathbb{N}$  et :

$$\forall k \in \mathbb{N}, e^{-np_n} \xrightarrow{n \rightarrow \infty} e^{-\lambda}$$

$$\begin{aligned} \forall k \in \mathbb{N}, \left( \prod_{i=0}^{k-1} \frac{n-i}{n(1-p_n)} \right) (1-p_n)^n &\underset{n \rightarrow \infty}{\sim} (1-p_n)^{n-k} \\ &\underset{n \rightarrow \infty}{\sim} (1-p_n)^n \\ &= e^{n \ln(1-p_n)} \\ &\underset{n \rightarrow \infty}{=} e^{-np_n + o(np_n)} \\ &\underset{n \rightarrow \infty}{=} e^{-\lambda + o(1)} \xrightarrow{n \rightarrow \infty} e^{-\lambda} \end{aligned}$$

Cela prouve que si on se donne  $(p_n)_{n \in \mathbb{N}} \in [0, 1]^\mathbb{N}$  avec  $p_n \sim \frac{\lambda}{n}$  et  $(S_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires réelles telles que pour tout  $n \in \mathbb{N}$ ,  $S_n \hookrightarrow \mathcal{B}(n, p_n)$ , alors  $S_n$  tend vers une variable aléatoire  $X$  telle que  $X \hookrightarrow \mathcal{P}(\lambda)$  : il y a convergence simple.

Par le Théorème 3, on a finalement convergence en loi, et si pour tout  $n \in \mathbb{N}$  on note  $\mu_n$  la loi de  $S_n$  et si on note  $\mu$  la loi de  $X$ , alors  $d_{VT}(\mu_n, \mu) \rightarrow 0$ .  $\square$

On a maintenant une première bonne impression quant à l'approximation binomiale-Poisson sous certaines conditions sur le paramètre de succès. Reste à expliciter une vitesse de convergence convenable pour la convergence en loi associée à cette approximation. On pourra alors vraiment se convaincre qu'à partir d'un nombre suffisamment grand de répétitions d'un jeu du type succès/échec, on peut faire comme si les probabilités élémentaires sont celles obtenues par la loi de Poisson.

### 3.2 Inégalité de Le Cam et vitesse de convergence

Afin d'estimer la vitesse de convergence, il va falloir utiliser l'inégalité de Le Cam, qu'on va démontrer.

**Théorème 7.** Soit  $n \in \mathbb{N}^*$ . Soient  $X_1, \dots, X_n$  des variables aléatoires indépendantes de lois de Bernoulli respectives  $\mathcal{B}(p_1), \dots, \mathcal{B}(p_n)$ , avec  $p_1, \dots, p_n \in ]0, 1[$ .

Soit  $\mu_n$  la loi de  $S_n = \sum_{i=1}^n X_i$  et soit  $\nu_n = \mathcal{P}(\sum_{i=1}^n p_i)$  la loi de Poisson de paramètre  $\sum_{i=1}^n p_i$ , de même moyenne que  $S_n$ .

Alors, on a :

$$d_{VT}(\mu_n, \nu_n) \leq \sum_{i=1}^n p_i^2$$

*Démonstration.* La preuve est très longue. Il y a 5 grandes étapes nécessaires à l'aboutissement de l'inégalité.

1. Montrons par récurrence sur  $n \in \mathbb{N}^*$  que si  $\alpha_1, \dots, \alpha_n$  et  $\beta_1, \dots, \beta_n$  sont des lois de probabilité sur  $\mathbb{N}$ , alors :

$$d_{VT}(\alpha_1 * \dots * \alpha_n, \beta_1 * \dots * \beta_n) \leq \sum_{i=1}^n d_{VT}(\alpha_i, \beta_i) \quad (\text{Propriété } H_n)$$

Initialisation :

Avec les notations ci-dessus, pour  $n = 1$ , le résultat est évident.

Hérédité : Soit  $n \in \mathbb{N}$ , on suppose  $H_n$ .

Soient  $\alpha_1, \dots, \alpha_{n+1}$  et  $\beta_1, \dots, \beta_{n+1}$  des lois de probabilité sur  $\mathbb{N}$ .

On note  $\alpha_{0,n} = \alpha_1 * \dots * \alpha_n$  et  $\beta_{0,n} = \beta_1 * \dots * \beta_n$ .

Alors, par Théorème (de réécriture de la distance en variation totale) :

$$\begin{aligned} d_{VT}(\alpha_1 * \dots * \alpha_{n+1}, \beta_1 * \dots * \beta_{n+1}) &= \frac{1}{2} \sum_{m=0}^{+\infty} |\alpha_{0,n} * \alpha_{n+1}(m) - \beta_{0,n} * \beta_{n+1}(m)| \\ &= \frac{1}{2} \sum_{m=0}^{+\infty} \left| \sum_{k=0}^m (\alpha_{0,n}(k) \alpha_{n+1}(m-k) - \beta_{0,n}(k) \beta_{n+1}(m-k)) \right| \\ &\leq \frac{1}{2} \sum_{m=0}^{+\infty} \left| \sum_{k=0}^m (\alpha_{0,n}(k) \alpha_{n+1}(m-k) - \beta_{0,n}(k) \alpha_{n+1}(m-k)) \right| \\ &\quad + \frac{1}{2} \sum_{m=0}^{+\infty} \left| \sum_{k=0}^m (\beta_{0,n}(k) \alpha_{n+1}(m-k) - \beta_{0,n}(k) \beta_{n+1}(m-k)) \right| \\ &\leq \frac{1}{2} \sum_{m=0}^{+\infty} \sum_{k=0}^m |\alpha_{0,n}(k) - \beta_{0,n}(k)| \alpha_{n+1}(m-k) \\ &\quad + \frac{1}{2} \sum_{m=0}^{+\infty} \sum_{k=0}^m \beta_{0,n}(k) |\alpha_{n+1}(m-k) - \beta_{n+1}(m-k)| \\ &= \frac{1}{2} \left( \sum_{m=0}^{+\infty} |\alpha_{0,n}(m) - \beta_{0,n}(m)| \right) \left( \sum_{m=0}^{+\infty} \alpha_{n+1}(m) \right) \\ &\quad + \frac{1}{2} \left( \sum_{m=0}^{+\infty} |\alpha_{n+1}(m) - \beta_{n+1}(m)| \right) \left( \sum_{m=0}^{+\infty} \beta_{0,n}(m) \right) \\ &\quad (\text{produits de Cauchy}) \\ &= d_{VT}(\alpha_1 * \dots * \alpha_n, \beta_1 * \dots * \beta_n) + d_{VT}(\alpha_{n+1}, \beta_{n+1}) \\ &\leq \sum_{i=1}^{n+1} d_{VT}(\alpha_i, \beta_i) \\ &\quad (\text{hypothèse de récurrence}) \end{aligned}$$

2. Soit  $p \in ]0, 1[$ . Alors d'après l'exemple donné en sous-section 2.2, on a :

$$d_{VT}(\mathcal{P}(p), \mathcal{B}(p)) = p(1 - e^{-p}) \leq p(1 - (1 - p)) = p^2$$



(utilisation de la convexité avec la fonction exponentielle)

3. Soient  $a, b \in \mathbb{R}_+^*$ . Montrons que  $\mathcal{P}(a) * \mathcal{P}(b) = \mathcal{P}(a+b)$ .

Soient  $X_{a,b}$  et  $X_{a+b}$  des variables aléatoires réelles avec  $X_{a,b} \hookrightarrow \mathcal{P}(a) * \mathcal{P}(b)$  et  $X_{a+b} \hookrightarrow \mathcal{P}(a+b)$ .

Pour tout  $k \in \mathbb{N}$  :

$$P(X_{a,b} = k) = \sum_{j=0}^k e^{-a} \frac{a^j}{j!} e^{-b} \frac{b^{k-j}}{(k-j)!} = \frac{e^{-(a+b)}}{k!} \sum_{j=0}^k \binom{k}{j} a^j b^{k-j} = e^{-(a+b)} \frac{(a+b)^k}{k!} = P(X_{a+b} = k)$$

Donc,  $X_{a,b}$  et  $X_{a+b}$  ont la même loi.

4. Montrons par récurrence sur  $n \in \mathbb{N}^*$  que si on se donne des variables aléatoires indépendantes à valeurs dans  $\mathbb{N}$   $X_1, \dots, X_n$  de lois respectives  $\mu_1, \dots, \mu_n$ , alors :

$$\sum_{i=1}^n X_i \hookrightarrow \mu_1 * \dots * \mu_n \text{ (Propriété } H'_n \text{)}$$

Initialisation : Pour  $n = 1$ , le résultat est évident.

Hérédité : Soit  $n \in \mathbb{N}^*$ . On suppose  $H'_n$ .

Soient  $X_1, \dots, X_{n+1}$  des variables aléatoires indépendantes à valeurs dans  $\mathbb{N}$  de lois respectives  $\mu_1, \dots, \mu_{n+1}$  avec  $p_1, \dots, p_{n+1} \in ]0, 1[$ .

On a :

$$\begin{aligned} \forall k \in \mathbb{N}, P(\sum_{i=1}^{n+1} X_i = k) &= \sum_{m=0}^k P(\sum_{i=1}^n X_i = m \text{ et } X_{n+1} = k - m) \\ &= \sum_{m=0}^k P(\sum_{i=1}^n X_i = m) P(X_{n+1} = k - m) \\ &\text{(par indépendance)} \end{aligned}$$

Avec ce résultat, comme  $X_{n+1} \hookrightarrow \mu_{n+1}$  et  $\sum_{i=1}^n X_i \hookrightarrow \mu_1 * \dots * \mu_n$ , par hypothèse de récurrence, on a bien :

$$\sum_{i=1}^{n+1} X_i \hookrightarrow \mu_1 * \dots * \mu_{n+1}$$

5. Soit  $n \in \mathbb{N}^*$ . Soient  $X_1, \dots, X_n$  des variables aléatoires indépendantes de lois de Bernoulli respectives  $\mathcal{B}(p_1), \dots, \mathcal{B}(p_n)$ , avec  $p_1, \dots, p_n \in ]0, 1[$ .

Soit  $\mu_n$  la loi de  $S_n = \sum_{i=1}^n X_i$  et soit  $\nu_n = \mathcal{P}(\sum_{i=1}^n p_i)$ . Alors :

$$d_{VT}(\mu_n, \nu_n) = d_{VT}(\mathcal{B}(p_1) * \dots * \mathcal{B}(p_n), \mathcal{P}(p_1) * \dots * \mathcal{P}(p_n))$$

(par 3. et 4.)

$$\text{(par 1.)} \leq \sum_{i=1}^n d_{VT}(\mathcal{B}(p_i), \mathcal{P}(p_i))$$

$$\text{(par 2.)} \leq \sum_{i=1}^n p_i^2$$

□

En reprenant les notations données avant ce théorème, par l'inégalité de Le Cam et réécriture de la distance en variation totale (théorème 2) :

$$\forall n \in \mathbb{N}, \sum_{k=0}^{+\infty} \left| P(S_n = k) - e^{-np_n} \frac{(np_n)^k}{k!} \right| = 2d_{VT}(\mathcal{B}(n, p_n), \mathcal{P}(np_n))$$

$$\text{(par les points 3. et 4. du Théorème précédent, on a des convolutions de } n \text{ lois)} = 2d_{VT}(\mathcal{B}(p_n) * \dots * \mathcal{B}(p_n), \mathcal{P}(p_n) * \dots * \mathcal{P}(p_n))$$

$$\leq 2 \sum_{i=1}^n p_n^2 = 2np_n^2 \underset{n \rightarrow +\infty}{\sim} \frac{2\lambda^2}{n}$$

On a donc une vitesse de convergence de l'ordre de  $\frac{1}{n}$  à un facteur multiplicatif près. Ainsi, si on fait une répétition de 100 expériences du type succès échec avec une probabilité de succès de  $\frac{1}{100}$ , on obtient une distance en variation totale entre la loi binomiale sous-jacente et la loi de Poisson de paramètre 1 proche de 0,02.

**Remarque 4.** L'inégalité  $\sum_{i=1}^n p_i^2 \leq \left(\sum_{i=1}^n p_i\right) \max_{k \in \llbracket 1, n \rrbracket} p_k$  permet de retrouver la loi des petits nombres à l'aide de l'inégalité de Le Cam : si  $(X_{n,k})_{1 \leq k \leq n, n \in \mathbb{N}^*}$  est un tableau triangulaire de variables aléatoires indépendantes de lois de Bernoulli avec  $X_{n,k} \hookrightarrow \mathcal{B}(p_{n,k})$  et  $p_{n,k} \in ]0, 1[$  pour tous  $n \in \mathbb{N}$  et  $k \in \llbracket 1, n \rrbracket$  et si  $\lim_{n \rightarrow +\infty} \sum_{i=1}^n p_{n,i} = \lambda$  pour un certain  $\lambda > 0$  et  $\lim_{n \rightarrow +\infty} \max_{k \in \llbracket 1, n \rrbracket} p_{n,k} = 0$ , alors  $\left(\sum_{i=1}^n X_{n,i}\right)_{n \in \mathbb{N}^*}$  converge en loi vers  $\mathcal{P}(\lambda)$ .

On a enfin prouvé théoriquement la légitimité de l'approximation binomiale-Poisson dans le cadre restreint sur le paramètre de succès décrit dans la section 3.1, qu'on appelle loi des événements rares. Afin d'observer la puissance de cette approximation, on va s'appuyer sur des cas concrets associés à une grande répétition d'une expérience du type succès/échec et comparer la loi binomiale et la loi de Poisson.

### 3.3 Exemples d'applications

#### 3.3.1 Anniversaire et Nouvel An

On se donne une population de 1000 personnes.

Question : pour  $k \in \llbracket 0, 1000 \rrbracket$ , quelle est la probabilité que dans cette population, exactement  $k$  personnes soient nées le jour du Nouvel An ?

Pour cela, on modélise les choix aléatoires des personnes constituant notre population par 1000 variables aléatoires de loi de Bernoulli de paramètre  $p = \frac{1}{365}$ , dont la somme suit la loi binomiale de paramètres  $n = 1000$  et  $p = \frac{1}{365}$  (on exclut les personnes nées un 29 Février).

Prenons  $\lambda = \frac{1000}{365}$  pour l'approximation de Poisson.

Comparons les probabilités entre  $\mathcal{B}(1000, p)$  et  $\mathcal{P}\left(\frac{1000}{365}\right)$  :

$k \in \mathbb{N}$	0	1	2	3	4	5
binomiale	0,0643	0,1768	0,2426	0,2217	0,1518	0,0831
Poisson	0,0646	0,1770	0,2424	0,2214	0,1516	0,0831

Dans ce cadre, la loi de Poisson de paramètre  $\lambda = \frac{1000}{365}$  est une bonne approximation de la loi binomiale  $\mathcal{B}\left(1000, \frac{1}{365}\right)$ .

#### 3.3.2 Objets défectueux

Une entreprise produit des vis. Il est ainsi normal de vérifier la qualité et le caractère non défectueux des vis produits. On suppose que la probabilité qu'une vis soit défectueuse est de  $p = 0,015$ . On peut se permettre d'utiliser le schéma de Bernoulli pour tester la qualité des vis contenues dans une grosse caisse.

La probabilité pour qu'une boîte de 120 vis ne contienne aucune vis défectueuse est de  $(0,985)^{120} \simeq 0,16306$ . Celle associée à la loi de Poisson est de  $e^{-0,015 \times 120} = e^{-1,8} \simeq 0,16530$ , dont la loi est utilisée à des fins pratiques.

Posons-nous une autre question : combien de vis faut-il dans une boîte pour que la probabilité d'avoir au moins 120 vis de qualité soit supérieure ou égale à 0,08 ?

Appliquons l'approximation de Poisson pour  $n = 120 + x$  avec  $x \in \mathbb{N}$ .

Posons  $\lambda = np$ , mais  $np$  c'est presque  $120p$  pour  $x$  assez petit.

Cherchons alors le plus petit  $x \in \mathbb{N}$  tel que :

$$e^{-1,8} \left( \sum_{i=0}^x \frac{(1,8)^i}{i!} \right) \geq 0,8$$

("=  $P(X \leq x)$ " si  $X$  est la variable aléatoire qui donne le nombre de vis défectueuses).

— Pour  $x = 2$  :

$$e^{-1,8} \left( 1 + \frac{1,8}{1} + \frac{(1,8)^2}{2} \right) \simeq 0,73062 < 0,8$$

— Pour  $x = 3$  :

$$e^{-1,8} \left( 1 + \frac{1,8}{1} + \frac{(1,8)^2}{2} + \frac{(1,8)^3}{6} \right) \simeq 0,89129 > 0,8$$

Avec l'approximation de Poisson, il faudrait au moins 123 vis.

Est-ce le même résultat avec un schéma de Bernoulli ?

On reprend  $p = 0,015$ .

On cherche à calculer :

$$p_0 = \sum_{k=0}^3 \binom{123}{k} (0,015)^k (0,985)^{123-k}$$

On a  $p_0 \simeq 0,88541 \simeq 0,89129$ .

Encore une fois, on voit que l'approximation de Poisson est légitime.

## 4 Loi de l'arcsinus

L'étude précédente portait sur un paramètre de succès faible pour un nombre de répétitions important. Le comportement de la binomiale pour un grand nombre de répétitions mais avec un paramètre de succès proche de  $\frac{1}{2}$  est différent. Une symétrie dans les probabilités entre le succès et l'échec nous fait penser que la probabilité de "souvent rencontrer une égalité entre le succès et l'échec" est proche de 1, car la probabilité de  $\frac{1}{2}$  laisserait penser qu'on ne va pas s'enfoncer plus dans l'échec que dans le succès (alors qu'avec un paramètre de succès de  $\frac{3}{4}$ , on penserait plutôt qu'on va engranger beaucoup de succès). Ainsi on penserait également à un "changement fréquent de leader". L'étude des travaux de Feller [2] nous permet de lever le doute sur cette question.

### 4.1 Contexte et mise en situation

**Contexte 1.** *On va faire une analyse sur les variations d'une marche aléatoire. Selon les croyances les plus répandues, au pile ou face, aucun des deux joueurs n'aurait le dessus sur l'autre après un grand nombre d'essais.*

*On s'imagine, pour  $n \in \mathbb{N}$  assez grand, une répétition de  $2n$  jeux de pile ou face avec équiprobabilité entre pile et face. On choisit au hasard une telle situation, et on regarde le plus grand entier  $k \in \llbracket 0, n \rrbracket$  tel qu'il y ait autant de pile que de face après exactement  $k$  essais.*

*On penserait que les changements les plus fréquents de joueur en première position se font autour de  $2n$  parties. En fait ce n'est pas souvent le cas. Le prochain théorème révélera une symétrie des probabilités autour de  $2 \times \frac{n}{2}$  parties. La probabilité que la dernière égalité se fasse après exactement  $2k$  parties est la même que celle que la dernière égalité se fasse après exactement  $2(n - k)$  parties, pour  $k \in \llbracket 0, n \rrbracket$ . Cette valeur  $k$  a autant de probabilités d'être au-dessus de  $\frac{n}{2}$  que d'être en-dessous de  $\frac{n}{2}$ , et les probabilités les plus élevées sont atteintes pour  $k = 0$  et  $k = n$ . Ce résultat est assez contre-intuitif.*

**Exemples 1.** 1. *On se donne pour expérience une répétition de jeux de pile ou face pendant une année sans interruption à la fréquence d'un jeu par seconde. On s'imagine qu'on effectue simultanément un certain nombre de fois cette expérience ; alors :*

- pour une partie sur 10 en moyenne, le dernier cas d'égalité se produit dans les neuf premiers jours ;
- pour une partie sur 20 en moyenne, le dernier cas d'égalité se produit dans les 54 premières heures ;
- pour une partie sur 100 en moyenne, le dernier cas d'égalité se produit dans les 130 premières minutes.

2. *Supposons que dans une expérience d'apprentissage d'un an, un enfant soit constamment en retard, sauf peut-être la première semaine, et qu'un autre soit constamment en avance, sauf peut-être la dernière semaine. Ces enfants seraient-ils jugés égaux ?*

*On se donne 11 enfants qui vont suivre la même expérience d'apprentissage dont la progression dépend uniquement de la chance.*

*Contre toute attente, il y en aura un qui serait tout le temps le meilleur, sauf une certaine semaine, et un autre serait tout le temps le dernier, sauf une certaine semaine également.*

Les probabilités sur les valeurs de  $k$  dans  $\llbracket 0, n \rrbracket$  sont données par la loi de l'Arcsinus (qu'on va prochainement étudier).

### 4.2 Marches aléatoires, dernier retour en zéro

**Notations 1.** — Soit  $(X_n)_{n \in \mathbb{N}^*}$  une suite de variables aléatoires indépendantes et de loi de Rademacher, ie :

$$\forall n \in \mathbb{N}^*, P(X_n = -1) = P(X_n = 1) = \frac{1}{2}$$

— Pour tout  $n \in \mathbb{N}$ , on note  $S_n = \sum_{i=1}^n X_i$  (par convention,  $S_0 = 0$ ).  $(S_n)_{n \in \mathbb{N}}$  est une **marche aléatoire**.

- Pour tous  $n \in \mathbb{N}$  et  $r \in \mathbb{Z}$ , on écrira  $p_{n,r} = P(S_n = r)$ .
- Par convention, pour  $n \in \mathbb{N}$ , pour tout  $k \in \mathbb{R} \setminus \llbracket 0, n \rrbracket$ ,  $\binom{n}{k} = 0$ .
- Pour  $n \in \mathbb{N}^*$ , pour tout  $r \in \mathbb{Z}$ , on note  $N_{n,r}$  le nombre d'éléments  $(x_1, \dots, x_n)$  de  $\{-1, 1\}^n$  tels que  $\sum_{i=1}^n x_i = r$ .  
Pour  $n = 0$ , on note  $N_{0,r} = \delta_{0,r}$  pour tout  $r \in \mathbb{Z}$ .

**Théorème 8.** Pour  $n \in \mathbb{N}$  et  $r \in \mathbb{Z}$ ,  $N_{n,r} = \binom{\frac{n+r}{2}}{k_0}$ .

*Démonstration.* Pour voir clairement ce résultat, en se fixant  $n \in \mathbb{N}^*$  et  $r \in \llbracket -n, n \rrbracket$ , cela revient à chercher  $k_0 \in \llbracket 0, n \rrbracket$  tel que  $\sum_{i=1}^r x_i = k_0 \times 1 + (n - k_0) \times (-1) = r$  (car pour tout  $i \in \llbracket 1, n \rrbracket$ ,  $x_i \in \{-1, 1\}$ ). Chercher  $(x_1, \dots, x_n) \in \{-1, 1\}^n$  revient donc à choisir  $k_0$  éléments de  $\llbracket 1, n \rrbracket$  (choisir les indices  $i_1, \dots, i_{k_0}$  distincts dans  $\llbracket 1, n \rrbracket$  tels que  $x_{i_1} = \dots = x_{i_{k_0}} = 1$ ), on aura alors  $N_{n,r} = \binom{n}{k_0}$ .

Or :

$$\forall k \in \llbracket 0, n \rrbracket, k \times 1 + (n - k) \times (-1) = r \Leftrightarrow 2k - n = r \Leftrightarrow k = \frac{n+r}{2} \in \llbracket 0, n \rrbracket$$

Si  $\frac{n+r}{2} \notin \mathbb{N}$ , on a bien  $N_{n,r} = 0 = \binom{\frac{n+r}{2}}{k_0}$  et si  $\frac{n+r}{2} \in \mathbb{N}$ , l'explication ci-dessus montre bien que  $N_{n,r} = \binom{\frac{n+r}{2}}{k_0}$ .

Dans le cas  $n \in \mathbb{N}^*$  et  $r \in \mathbb{Z} \setminus \llbracket -n, n \rrbracket$ , comme  $\left\{ \sum_{i=1}^n x_i \right\}_{(x_1, \dots, x_n) \in \{-1, 1\}^n} \subset \llbracket -n, n \rrbracket$  et  $\frac{n+r}{2} \notin \llbracket 0, n \rrbracket$ , on a bien  $N_{n,r} = \binom{\frac{n+r}{2}}{k_0} = 0$ .

Il reste à regarder le cas  $n = 0$  :

$$\forall r \in \mathbb{Z}, N_{0,r} = \delta_{0,r} = \binom{0}{\frac{r}{2}}$$

□

**Théorème 9.** Pour tous  $n \in \mathbb{N}$  et  $r \in \mathbb{Z}$  :

$$p_{n,r} = \binom{\frac{n+r}{2}}{k_0} 2^{-2n}$$

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

Supposons  $n = 0$ .

Comme  $S_0 = 0$ , on a bien  $p_{0,r} = \delta_{0,r}$  pour tout  $r \in \mathbb{Z}$  et par le théorème précédent, on a finalement  $p_{0,r} = \binom{0}{\frac{r}{2}} 2^{-0}$  (de manière visuelle, pour  $2 \times 0$  parties, ie 0 parties, le seul résultat possible est d'avoir égalité puisque rien ne s'est passé).

Supposons désormais  $n$  non nul. Soit  $r \in \mathbb{Z}$ .

On compte  $2^{2n}$  situations possibles pour  $2n$  parties de pile ou face (on choisit pour chaque partie une issue parmi pile et face), toutes équiprobables (la probabilité d'avoir face, comme celle d'avoir pile, est de  $\frac{1}{2}$ , et les  $2n$  parties sont indépendantes).

Si on décide de dire que pile rapporte 1 et que face rapporte  $-1$ , alors calculer  $p_{n,r}$  revient de manière équivalente à dénombrer le nombre de situations où pile apparaît exactement  $r$  fois de plus que face (avant de diviser par le cardinal de  $\{-1, 1\}^n$ ), dont la valeur est donnée par le Théorème précédent. D'où :

$$p_{n,r} = \binom{\frac{n+r}{2}}{k_0} 2^{-2n}$$

□

**Notation 1.** Soit  $k \in \mathbb{N}$ . On note  $u_{2k} = p_{2k,0}$ . De manière visuelle,  $u_{2k}$  représente la probabilité de retour en 0 après  $2k$  déplacements unitaires depuis 0 (point de vue Marches aléatoires) ou bien la probabilité d'avoir égalité après exactement  $2k$  parties de pile ou face (point de vue du jeu).

**Rappel 1.** On rappelle la formule de Stirling :

$$n! \underset{n \rightarrow +\infty}{\sim} \left(\frac{n}{e}\right)^n \sqrt{2n\pi}$$

**Théorème 10.** (Formule de Wallis) On a :

$$\binom{2n}{n} \underset{n \rightarrow +\infty}{\sim} \frac{2^{2n}}{\sqrt{n\pi}}$$

*Démonstration.* On a, par la formule de Stirling :

$$\forall n \in \mathbb{N}, \binom{2n}{n} = \frac{(2n)!}{(n!)^2} \underset{n \rightarrow +\infty}{\sim} \frac{(2n)^{2n}}{e^{2n}} \times \sqrt{4n\pi} \times \frac{1}{2n\pi} \times \frac{e^{2n}}{n^{2n}} = \frac{2^{2n}}{\sqrt{n\pi}}$$

□

**Théorème 11.** Pour tout  $k \in \mathbb{N}$ ,  $u_{2k} = \binom{2k}{k} 2^{-2k}$ .

D'où :

$$u_{2k} \underset{k \rightarrow +\infty}{\sim} \frac{1}{\sqrt{k\pi}}$$

*Démonstration.* Pour tout  $k \in \mathbb{N}$ ,  $u_{2k} = p_{2k,0} = \binom{2k}{k} 2^{-2k}$ .

Par la formule de Wallis et en utilisant l'expression qu'on vient d'obtenir, on a bien :

$$u_{2k} \underset{k \rightarrow +\infty}{\sim} \frac{1}{\sqrt{k\pi}}$$

□

**Notation 2.** Pour  $k \in \mathbb{N}$ , on note  $f_{2k}$  la probabilité que le premier retour en 0 (point de vue Marches aléatoires) se produise après exactement  $2k$  déplacements. Ainsi :

- $f_0 = 0$  (car aucun déplacement donc aucun retour);
- $\forall k \in \mathbb{N}^*, f_{2k} = P\left((S_{2k} = 0) \cap \left(\bigcap_{i=1}^{2k-1} (S_i \neq 0)\right)\right)$ .

**Théorème 12.** Pour tout  $n \in \mathbb{N}^*$ , on a :

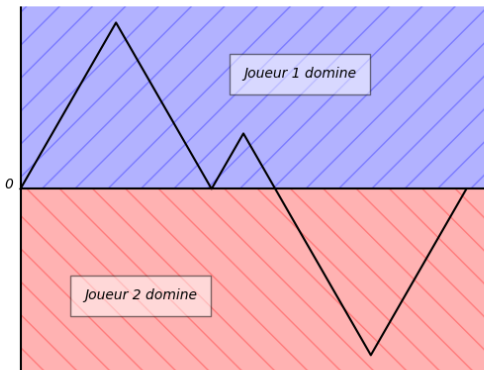
$$u_{2n} = \sum_{k=0}^{n-1} u_{2k} f_{2n-2k} = \sum_{k=1}^n f_{2k} u_{2n-2k}$$

*Démonstration.* Cette démonstration se fera de manière imagée.

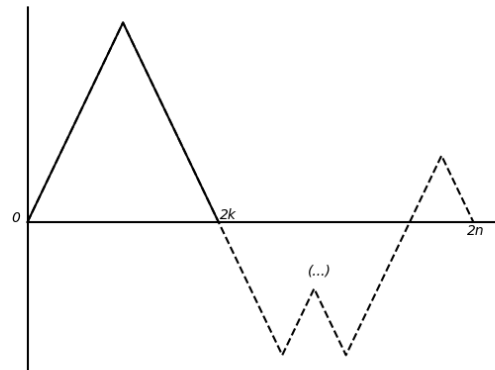
Soit  $n \in \mathbb{N}^*$ .

On sait que  $u_{2n}$  représente la probabilité pour qu'il y ait égalité entre deux joueurs au jeu du pile ou face après  $2n$  parties. Pour cela, représentons les résultats par des chemins à  $2n$  déplacements diagonaux sur un plan sur lequel on trace l'axe "égalité". Si on note  $J_1$  le joueur 1 et  $J_2$  le joueur 2, on dira qu'un déplacement se fera vers le haut si  $J_1$  gagne et vers le bas si  $J_2$  gagne.

Ainsi, grâce à l'axe "égalité", on distingue une région où  $J_1$  domine et une autre où  $J_2$  domine (figure 1a).



(a) Zones de domination des joueurs.



(b) Premier retour à zéro en  $2k$ .

FIGURE 1

Pour tout  $k \in \llbracket 1, n \rrbracket$ , on considère l'événement  $A_k$  : "le premier retour à égalité est en  $2k$ ", de probabilité  $f_{2k}$ . Alors, les événements de la famille  $(A_k)_{k \in \llbracket 1, n \rrbracket}$  sont deux-à-deux disjoints (on ne peut avoir premier retour à égalité à deux

moments différents).

De même, l'événement  $(S_{2n} = 0)$  (égalité après  $2n$  parties) est inclus dans  $\bigsqcup_{k=1}^n A_k$ ; en effet si on touche l'axe "égalité" en  $2n$ , on va forcément la retoucher pour la première fois en  $2k$  pour un certain  $k \in \llbracket 1, n \rrbracket$  (on ne peut avoir égalité après un nombre impair de parties, ie on ne touche pas l'axe "égalité" en  $2k + 1$  pour  $k \in \llbracket 0, n - 1 \rrbracket$ ).

D'où, par additivité :

$$u_{2n} = P(S_{2n} = 0) = P\left((S_{2n} = 0) \cap \left(\bigsqcup_{k=1}^n A_k\right)\right) = P\left(\bigsqcup_{k=1}^n (A_k \cap (S_{2n} = 0))\right) = \sum_{k=1}^n P((S_{2n} = 0) \cap A_k)$$

Soit  $k \in \llbracket 1, n \rrbracket$ . Être dans  $(S_{2n} = 0) \cap A_k$ , c'est avoir un premier retour à égalité en  $2k$  et un retour à égalité en  $2n$  (figure 1b).

Comme les parties de pile ou face sont indépendantes, cet événement peut être traduit par "on a un premier retour à égalité en  $2k$  et depuis  $2k$  on a un retour à égalité après  $2n - 2k$  parties" car les chemins  $0 \rightarrow 2k$  et  $2k \rightarrow 2n$  sont "indépendants" (en termes de déplacements).

D'où, en termes de probabilités :

$$P(S_{2n} = 0) = \sum_{k=1}^n P((S_{2n} = 0) \cap A_k) = \sum_{k=1}^n P(A_k)P(S_{2n-2k} = 0) = \sum_{k=1}^n f_k u_{2n-2k}$$

□

Un dernier lemme est nécessaire avant d'énoncer des résultats sur la loi du dernier retour en 0.

**Lemme 1** (Lemme principal). *Pour tout  $n \in \mathbb{N}$ , on a :*

$$u_{2n} = P(S_{2n} = 0) = P\left(\bigcap_{k=1}^{2n} S_k \neq 0\right)$$

*Autrement dit : la probabilité de revenir à 0 en  $2n$  parties est la même que la probabilité de ne jamais revenir en 0 jusqu'à la  $2n$ -ième incluse.*

Ne jamais revenir en 0 équivaut à la domination stricte d'un joueur sur l'autre. Par symétrie du problème, on en déduit que :

$$\frac{1}{2}u_{2n} = P\left(\bigcap_{k=1}^{2n} S_k > 0\right)$$

La preuve du lemme principal exploite plusieurs propositions intermédiaires, elle est détaillée en section 4.5. On admettra temporairement le résultat.

### 4.3 Loi discrète de l'arcsinus

On note  $\alpha_{2k, 2n}$  la probabilité que jusqu'à la  $2n$ -ième partie incluse, le dernier retour en 0 s'effectue après la  $2k$ -ième partie.

**Théorème 13** (Loi discrète de l'arcsinus). *Pour tout  $n \in \mathbb{N}$ , pour tout  $k \leq n$ , la probabilité  $\alpha_{2k, 2n}$  vérifie :*

$$\alpha_{2k, 2n} = u_{2k}u_{2n-2k}$$

*On nomme loi discrète de l'arcsinus d'ordre  $n$  la loi de probabilité qui affecte le poids  $\alpha_{2k, 2n}$  à  $2k$  pour  $k \leq n$ .*

*Démonstration.* On s'intéresse à l'événement suivant :

$$(S_{2k} = 0) \cap \left(\bigcap_{i=k+1}^n S_{2i} \neq 0\right)$$

Comme  $S_{2k} = 0$ , les événements  $S_{2i} \neq 0$  peuvent se réécrire comme somme des  $X_j$  pour  $j > 2k$  :

$$\begin{aligned} \alpha_{2k, 2n} &= P\left((S_{2k} = 0) \cap \left(\bigcap_{i=k+1}^n S_{2i} \neq 0\right)\right) \\ &= P\left(\left(\sum_{j=0}^{2k} X_j = 0\right) \cap \left(\bigcap_{i=k+1}^n \sum_{j=2k+1}^{2i} X_j \neq 0\right)\right) \\ &= P\left(\left(\sum_{j=0}^{2k} X_j = 0\right)\right) \times P\left(\left(\bigcap_{i=k+1}^n \sum_{j=2k+1}^{2i} X_j \neq 0\right)\right) \quad (\text{lemme des coalitions}) \end{aligned}$$

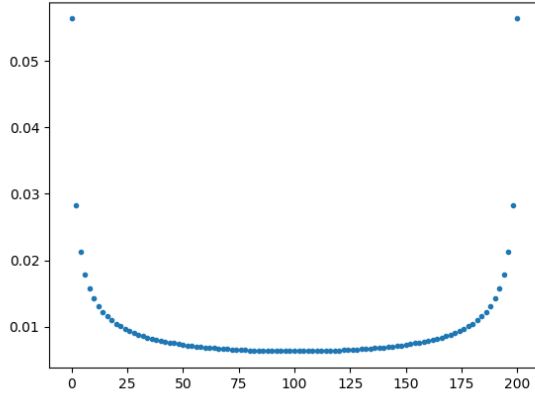
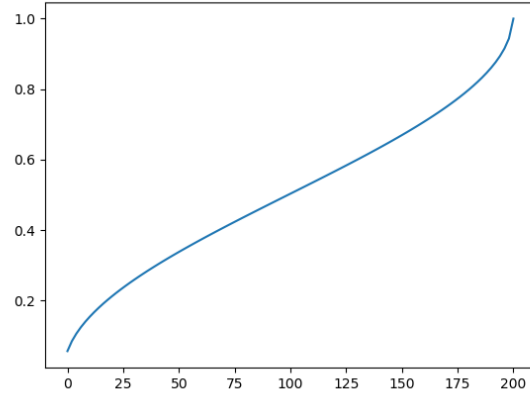
Fonction de masse de la probabilité du dernier retour à 0 ( $2n = 200$ ).Fonction de répartition de la probabilité du dernier retour à 0 ( $2n = 200$ ).

FIGURE 2 – Masse et répartition de la loi discrète de l'arcsinus d'ordre 100.

Les parties étant indépendantes et identiquement distribuées, on peut se ramener à l'étude d'une nouvelle marche partant du point  $(2k, 0)$  :

$$\begin{aligned}
 \alpha_{2k, 2n} &= u_{2k} \times P\left(\bigcap_{i=1}^{n-k} S_{2i} \neq 0\right) \\
 &= u_{2k} \times P(S_{2(n-k)} \neq 0) && \text{(lemme principal)} \\
 \alpha_{2k, 2n} &= u_{2k} u_{2n-2k} && \text{qui est bien le résultat souhaité.}
 \end{aligned}$$

□

Quelques commentaires sur ce résultat :

- puisqu'un retour à 0 se produit toujours avant la première partie, on peut toujours définir le dernier retour à 0 en  $2n$  parties. Les  $\alpha_{2k, 2n}$  pour  $k < n$  sont donc de somme égale à 1.
- on observe une symétrie dans la distribution :  $\alpha_{2k, 2n} = \alpha_{2n-2k, 2n}$ .
- la probabilité se concentre aux abords de 0 et  $2n$ , confirmant ainsi que le dernier changement de *leader* tend à se produire plutôt en début ou en fin d'expérience !

**Exemples 2.** — pour  $n = 2$ , on a  $\alpha_{0,4} = \alpha_{4,4} = \frac{3}{8}$  et  $\alpha_{2,4} = \frac{2}{8}$ .

— pour  $n = 8$ , on a la table de valeurs suivante :

$k \leq 8$	0 et 8	1 et 7	2 et 6	3 et 5	4
$\alpha_{2k, 16}$	0,1964	0,1067	0,0846	0,0769	0,0748

— la figure 2 présente la fonction de masse et la fonction de répartition pour  $n = 100$ . La figure 3 est une comparaison de ces fonctions aux fonctions de masse et de répartition d'une distribution empirique obtenue sur 4000 expériences.

#### 4.4 Convergence vers la loi de l'arcsinus

Une approximation de la loi de l'arcsinus discrète est obtenue en s'intéressant à la fonction  $f$  :

$$f : x \in ]0, 1[ \mapsto \frac{1}{\pi \sqrt{x(1-x)}}$$

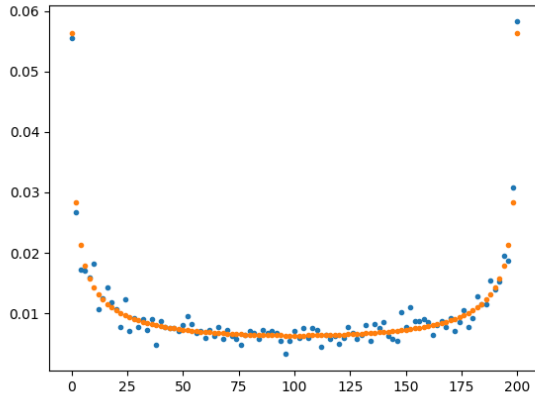
En effet, en utilisant le théorème 11, on peut montrer que lorsque  $n$  est assez grand et que  $k$  est assez proche de 0 ou de  $n$ ,

$$\alpha_{2k, 2n} = u_{2k} u_{2n-2k} \underset{n \rightarrow +\infty}{\sim} \frac{1}{\sqrt{k\pi}} \frac{1}{\sqrt{(n-k)\pi}} = \frac{1}{n} \frac{1}{\pi \sqrt{\frac{k}{n}(1-\frac{k}{n})}} = \frac{1}{n} f\left(\frac{k}{n}\right)$$

La figure 4 donne une interprétation graphique à ce résultat : pour  $k \neq 0, n$ ,  $\alpha_{2k, 2n}$  est l'aire d'un rectangle de hauteur  $f(\frac{k}{n})$  et de longueur  $\frac{1}{n}$ . On constate que la somme des  $\alpha_{2k, 2n}$  pour  $0 < k_1 \leq k \leq k_2 < n$  approxime l'aire sous la courbe entre  $\frac{k_1 - \frac{1}{2}}{n}$  et  $\frac{k_2 + \frac{1}{2}}{n}$ .

Pour étendre ce résultat à l'aire sous la courbe sur  $]0, 1[$  tout entier, il faut montrer que  $f$  est intégrable sur  $]0, 1[$ , d'intégrale valant la somme de tous les  $\alpha_{2k, 2n}$  pour  $0 \leq k \leq n$ , soit 1.

Masse théorique vs empirique, 4000 expériences (2n = 200).



Répartition théorique vs empirique, 4000 expériences (2n = 200).

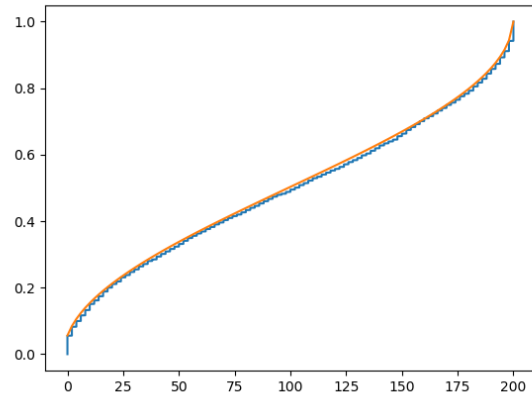


FIGURE 3 – Comparaison entre la masse et la répartition du dernier retour à zéro sur une distribution empirique (4000 expériences, en bleu) et la loi discrète de l'arcsinus d'ordre 100 (en jaune).

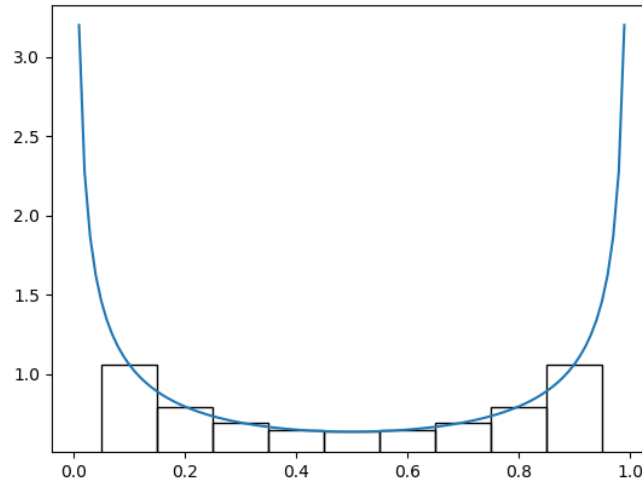


FIGURE 4 – Représentation graphique de  $f : x \mapsto \frac{1}{\pi\sqrt{x(1-x)}}$  et approximation de la loi de l'arcsinus discrète d'ordre 10.

On considère la fonction  $F : x \in ]0, 1[ \mapsto \frac{2}{\pi} \arcsin \sqrt{x}$ .  $F$  est une primitive de  $f$  sur  $]0, 1[$ , en effet :

$$\forall x \in ]0, 1[, F'(x) = \frac{2}{\pi} \frac{\frac{1}{2\sqrt{x}}}{\sqrt{1-x^2}} = \frac{1}{\pi\sqrt{x(1-x)}} = f(x)$$

Et on a :

$$\int_0^1 f(t) dt = \lim_{\epsilon \rightarrow 0} F(1-\epsilon) - F(\epsilon) = \lim_{\epsilon \rightarrow 0} \frac{2}{\pi} (\arcsin \sqrt{1-\epsilon} - \arcsin \sqrt{\epsilon}) = 1$$

Pour  $n$  assez grand, on peut ainsi assimiler la fonction de répartition de la loi de l'arcsinus discrète d'ordre  $n$  à l'aire sous la courbe de  $f$ , *i.e.* :

$$\forall x \in ]0, 1[ \sum_{0 \leq k \leq xn} \alpha_{2k, 2n} \simeq \frac{2}{\pi} \arcsin \sqrt{x}$$

Cette approximation justifie le nom de "l'arcsinus". Le prolongement par continuité de  $F$  en 0 et en 1 la rend valide même pour  $x \in [0, 1]$ . Elle est indépendante de la valeur de  $n$  permettant ainsi d'utiliser une table de valeurs et d'épargner grandement en temps de calculs. En pratique, elle est satisfaisante même pour des valeurs de  $n$  relativement faibles.



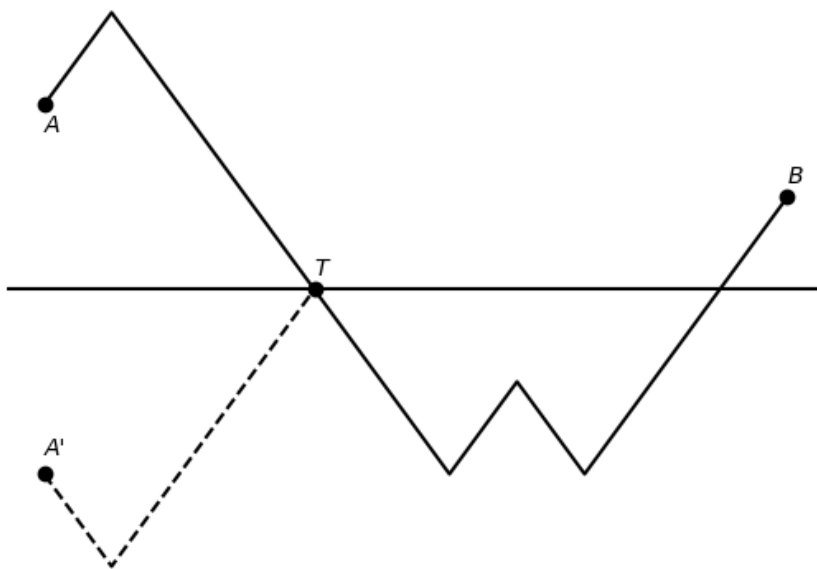


FIGURE 5 – Illustration du principe de réflexion.

## 4.5 Principe de réflexion, théorème du scrutin et lemme principal

On rappelle le lemme principal (lemme 1), évoqué et utilisé plus haut :

**Lemme 1** (Lemme principal). *Pour tout  $n \in \mathbb{N}$ , on a :*

$$u_{2n} = P(S_{2n} = 0) = P\left(\bigcap_{k=1}^{2n} S_k \neq 0\right)$$

*Autrement dit : la probabilité de revenir à 0 en  $2n$  parties est la même que la probabilité de ne jamais revenir en 0 jusqu'à la  $2n$ -ième incluse.*

La preuve de ce lemme utilise deux résultats généraux sur les marches aléatoires : le principe de réflexion et le théorème du scrutin.

**Lemme 2** (Principe de réflexion). *Pour  $0 \leq x_1 < x_2$  et  $y_1, y_2 \in \mathbb{N}^*$ , le nombre de suites de déplacements unitaires allant de  $(x_1, y_1)$  à  $(x_2, y_2)$  et intersectant l'axe d'égalité  $y = 0$  est égal au nombre de suites de déplacements unitaires allant de  $(x_1, -y_1)$  à  $(x_2, y_2)$ .*

*Démonstration.* La figure 5 illustre l'idée du principe de réflexion et de la preuve, qui va rester imagée puisqu'on voit bien sur le dessin ce qui se passe.

On va introduire une transformation bijective entre les deux ensembles de chemins.

On se donne une suite de déplacements unitaires  $\xi$  reliant  $(x_1, y_1)$  à  $(x_2, y_2)$  et intersectant l'axe  $y = 0$ . On considère  $x_1 < k_\xi < x_2$  l'abscisse du premier retour à 0 de ce chemin. On effectue sur ce chemin la transformation suivante :

- les déplacements unitaires avant l'abscisse  $k_\xi$  sont inversés (chaque  $-1$  devient un 1 et inversement) ;
- les déplacements unitaires après l'abscisse  $k_\xi$  sont conservés.

Graphiquement, on voit que ce nouveau chemin  $\xi'$  aboutit bien à  $(x_2, y_2)$  lorsque l'on place son origine en  $(x_1, -y_1)$ .

Cette transformation établit une bijection entre les deux ensembles de chemins considérés puisqu'on peut définir une transformation inverse, à savoir elle-même (pour  $\xi'$  chemin reliant  $(x_1, -y_1)$  à  $(x_2, y_2)$ ,  $k_{\xi'}$  est bien défini puisque  $-y_1 < 0$  et  $y_2 > 0$  :  $\xi'$  croise donc bien  $y = 0$  au moins une fois en partant de  $(x_1, -y_1)$ ).  $\square$

Le principe de réflexion est une astuce pratique utile pour l'étude des marches aléatoires. Elle permet notamment de montrer le théorème du scrutin, qui sert dans la preuve du lemme principal.

**Théorème 14** (Théorème du scrutin). Soit  $n > 0$  et  $y > 0$ . Il y a exactement  $\frac{y}{n}N_{n,y}$  suites de déplacement unitaires qui relient l'origine au point  $(n, y)$  et telles que  $\forall i \in \llbracket 1, n \rrbracket S_i > 0$ .

*Démonstration.* On note  $A$  l'ensemble de ces tels chemins.

Soit  $(X_1, \dots, X_n) \in \{-1, 1\}^n$  une suite de déplacements unitaires dans  $A$ . Comme  $S_1 = X_1 > 0$ , on a nécessairement  $X_1 = 1$ . On peut se limiter à l'étude des chemins reliant  $(1, 1)$  à  $(n, y)$  sans toucher l'axe  $y = 0$ .

$A$  est donc constitué de l'ensemble des chemins reliant  $(1, 1)$  à  $(n, y)$  (qu'on nommera  $B$ ) auquel on retire tous ceux qui touchent l'axe  $y = 0$  (ensemble  $C$ ) :  $A = B \setminus C$ .

En plaçant l'origine en  $(1, 1)$ , on voit qu'être un chemin de  $B$  revient à être un chemin qui relie  $(0, 0)$  à  $(n-1, y-1)$ . On a donc  $|B| = N_{n-1, y-1}$ .

On peut appliquer le principe de réflexion aux chemins de l'ensemble  $C$ , qui est donc en bijection avec l'ensemble des chemins reliant  $(1, -1)$  à  $(n, y)$ . Par un raisonnement similaire à celui pour  $|B|$ , en déplaçant l'origine en  $(1, -1)$ , on obtient  $|C| = N_{n-1, y+1}$ .

En conséquence :

$$\begin{aligned}
|A| &= |B| - |C| \\
&= N_{n-1, y-1} - N_{n-1, y+1} \\
&= \binom{p+q-1}{p-1} - \binom{p+q-1}{q-1} && \text{(en posant } p, q \text{ tels que } p+q=n, p-q=y) \\
&= \frac{(p+q-1)!}{(p-1)!q!} - \frac{(p+q-1)!}{(q-1)!p!} \\
&= (p+q)! \left( \frac{p}{p!q!(p+q)} - \frac{q}{q!p!(p+q)} \right) \\
&= \frac{(p+q)!}{p!q!} \frac{p-q}{p+q} \\
&= N_{n, y} \frac{y}{n} && \text{qui est bien le résultat voulu.}
\end{aligned}$$

□

Le théorème du scrutin établit un élégant lien de proportionnalité entre la probabilité que le joueur gagnant ait toujours mené au cours d'une expérience en  $n$  parties connaissant son score final  $y$ . Cette probabilité est  $\frac{y}{n}N_{n,y}2^{-n}$ .

Nous sommes dorénavant équipés pour démontrer le lemme principal :

*Démonstration.* On cherche à montrer :

$$u_{2n} = P(S_{2n} = 0) = P\left(\bigcap_{i=1}^{2n} S_i \neq 0\right)$$

Sachant que le problème est symétrique selon l'axe  $y = 0$  et qu'un chemin qui ne croise jamais cet axe reste toujours dans des ordonnées toutes positives ou toutes négatives, on peut se ramène à prouver :

$$\frac{1}{2}u_{2n} = P\left(\bigcap_{i=1}^{2n} S_i > 0\right)$$

On discrimine sur la valeur de  $S_{2n}$  à droite (les termes de la somme pour  $r > n$  sont nuls) :

$$P\left(\bigcap_{i=1}^{2n} S_i > 0\right) = \sum_{r=1}^{+\infty} P\left(S_{2n} = 2r \cap \bigcap_{i=1}^{2n-1} S_i > 0\right)$$

D'après l'égalité établie dans la preuve du théorème du scrutin, on a :

$$P\left(S_{2n} = 2r \cap \bigcap_{i=1}^{2n-1} S_i > 0\right) = \frac{2r}{2n} N_{2n, 2r} 2^{-2n} = 2^{-2n} (N_{2n-1, 2r-1} - N_{2n-1, 2r+1}) = \frac{1}{2} (p_{2n-1, 2-1} - p_{2n-1, 2r+1})$$

Notre somme infinie est donc télescopique, et on obtient :

$$\sum_{r=1}^{+\infty} P\left(S_{2n} = 2r \cap \bigcap_{i=1}^{2n-1} S_i > 0\right) = \frac{1}{2} \sum_{r=1}^{+\infty} p_{2n-1, 2-1} - p_{2n-1, 2r+1} = \frac{1}{2} p_{2n-1, 1}$$

Or  $p_{2n-1, 1} = u_{2n}$ . En effet :

$$\begin{aligned}
u_{2n} &= P(S_{2n} = 0) \\
&= p_{2n-1, 1} P_{(S_{2n-1}=1)}(S_{2n} = 0) + p_{2n-1, -1} P_{(S_{2n-1}=-1)}(S_{2n} = 0) && \text{(probabilités totales)} \\
&= \frac{1}{2} (p_{2n-1, 1} + p_{2n-1, -1}) \\
&= \frac{1}{2} 2p_{2n-1, 1} && \text{(par symétrie)} \\
u_{2n} &= p_{2n-1, 1}
\end{aligned}$$

D'où au final  $P\left(\bigcap_{i=1}^{2n} S_i > 0\right) = \frac{1}{2}u_{2n}$ , soit bien le résultat qu'on souhaitait.

□

## 5 Collectionneur de coupons

Le problème du collectionneur de coupons est le suivant : combien de fois faut-il lancer un dé à  $r$  faces pour voir apparaître au moins une fois chacune d'entre elles ? Le nom de ce problème est inspiré de la collection des coupons dans les paquets de céréales et il est implicitement présent dans beaucoup de situations concrètes. Nous étudions le document de Chafai [1] dans lequel il détaille quelques éléments de réponse.

### 5.1 Présentation du modèle aléatoire

Afin de modéliser cette expérience, en se fixant  $r \in \mathbb{N} \setminus \{0, 1\}$ , on considère la variable aléatoire

$$T = \min \left\{ n \in \mathbb{N}^* / \{X_i\}_{i \in \llbracket 1, n \rrbracket} = \llbracket 1, r \rrbracket \right\} = \min \left\{ n \in \mathbb{N}^* / |\{X_i\}_{i \in \llbracket 1, n \rrbracket}| = r \right\}$$

où  $(X_i)_{i \in \mathbb{N}^*}$  est une suite de variables aléatoires indépendantes et identiquement distribuées, de loi uniforme sur  $\llbracket 1, r \rrbracket$  :

$$\forall i \in \mathbb{N}^*, X_i \hookrightarrow \mathcal{U}(\llbracket 1, r \rrbracket)$$

La variable aléatoire  $T$  correspond à l'instant où on a vu apparaître pour la première fois les  $r$  faces du dé. Cet instant dépend de  $r$  mais dans un premier temps, pour simplifier, on ne prend pas en compte cette dépendance dans les notations.

On peut donner une idée de la loi de probabilité de  $T$ .

**Théorème 15.** *On a  $T \geq r$  et pour  $n \geq r$  :*

$$P(T = n) = \frac{r!}{r^n} \left\{ \begin{matrix} n-1 \\ r-1 \end{matrix} \right\}$$

où la notation avec les accolades correspond au nombre de Stirling de seconde espèce, qui désigne le nombre de manières de donner une partition de  $\llbracket 1, n-1 \rrbracket$  en  $r-1$  sous-ensembles non vides.

*Démonstration.* Soit  $n \geq r$ . Dénombrons le nombre de situations où  $T = n$ , en considérant seulement les  $n$  premières variables  $X_1, \dots, X_n$  (le reste ne nous intéresse pas car il est indépendant des  $n$  premières variables aléatoires  $X_1, \dots, X_n$  et du résultat qu'on veut obtenir :  $T = n$ ).

Donc,  $P(T = n) = \frac{|A|}{r^n}$  où  $A$  est l'ensemble des  $n$ -uplets de  $\llbracket 1, r \rrbracket^n$  qui contiennent tous les éléments de  $\llbracket 1, r \rrbracket$  et tels que pour tout  $k \in \llbracket 1, n-1 \rrbracket$ , le  $k$ -uplet issu des  $k$  premiers coefficients du  $n$ -uplet initial ne contient pas tous les éléments de  $\llbracket 1, r \rrbracket$ .

Il reste alors à dénombrer les éléments de  $A$ .

Dans  $\llbracket 1, r \rrbracket^n$ , choisir un élément de  $A$ , c'est choisir un  $n$ -uplet  $(x_1, \dots, x_n) \in \llbracket 1, r \rrbracket^n$  tel que  $|\{x_i\}_{i \in \llbracket 1, n \rrbracket}| = r$  et  $x_n \notin \{x_i\}_{i \in \llbracket 1, n-1 \rrbracket}$  (sinon  $|\{x_i\}_{i \in \llbracket 1, n-1 \rrbracket}| = |\{x_i\}_{i \in \llbracket 1, n \rrbracket}| = r$ ).

Cela revient alors à choisir un élément  $r_0$  de  $\llbracket 1, r \rrbracket$  ( $r$  choix) puis à choisir une façon de partitionner  $\llbracket 1, n-1 \rrbracket$  en  $r-1$  sous-ensembles ( $\left\{ \begin{matrix} n-1 \\ r-1 \end{matrix} \right\}$  choix) auxquels on va attribuer des numéros pour savoir quel élément de  $\llbracket 1, n-1 \rrbracket \setminus \{r_0\}$  chaque ensemble va représenter ( $(r-1)!$  choix).

D'où  $|A| = r \times \left\{ \begin{matrix} n-1 \\ r-1 \end{matrix} \right\} \times (r-1)! = r! \left\{ \begin{matrix} n-1 \\ r-1 \end{matrix} \right\}$ , ie :

$$P(T = n) = \frac{r!}{r^n} \left\{ \begin{matrix} n-1 \\ r-1 \end{matrix} \right\}$$

□

Le résultat est explicite mais ne légitime aucune intuition sur le comportement asymptotique.

De même, le calcul des nombres de Stirling nécessite un très grand nombre d'opérations élémentaires (car inclut un grand nombre de coefficients binomiaux dans l'écriture explicite du nombre de Stirling quand  $r$  est grand). Ainsi, comme pour le jeu du pile ou face, une approximation satisfaisante de la loi de probabilité  $T$  permettra de rendre les calculs plus rapides avec une erreur faible.

La seconde sous-section apportera des résultats plus parlants pour étudier le comportement asymptotique de la loi de  $T$  quand  $r$  tend vers  $+\infty$ .

## 5.2 Étude de la loi de $T$

**Lemme 3.** (Décomposition) On a  $T = \sum_{i=1}^r G_i$  où  $G_1, \dots, G_r$  sont des variables aléatoires indépendantes avec, pour tout  $i \in \llbracket 1, r \rrbracket$ , en écrivant  $\pi_i = \frac{r-i+1}{r}$  :

$$G_i \hookrightarrow \mathcal{G}(\pi_i)$$

En particulier, on a  $P(T < +\infty) = 1$  et de plus :

$$E(T) \underset{r \rightarrow \infty}{=} r(\ln r + \gamma) + o(r) ; \text{Var}(T) \underset{r \rightarrow \infty}{=} \frac{\pi^2}{6} r^2 + o(r^2)$$

où  $\gamma = \lim_{n \rightarrow \infty} \left( \left( \sum_{i=1}^n \frac{1}{i} \right) - \ln n \right) \simeq 0,577$  est la constante d'Euler.

*Démonstration.* On pose  $G_1 = 1$  et pour tout  $i \in \llbracket 2, r \rrbracket$ , on définit  $G_i$  de la façon suivante :

$$G_i = \min\{n \in \mathbb{N}^* / X_{G_1+\dots+G_{i-1}+n} \notin \{X_1, \dots, X_{G_1+\dots+G_{i-1}}\}\}$$

On a par récurrence rapide que pour tout  $i \in \llbracket 1, r \rrbracket$ ,  $|\{X_1, \dots, X_{G_1+\dots+G_i}\}| = i$  (intuitivement,  $G_1 + \dots + G_i$  correspond au premier instant où l'on obtient  $i$  coupons distincts). De même, en regardant ce résultat aux rangs  $r-1$  et  $r$  et la définition de  $G_r$ , on a  $T = G_1 + \dots + G_r$ .

Montrons par récurrence sur  $k \in \llbracket 1, r \rrbracket$  que la famille  $(G_i)_{i \in \llbracket 1, k \rrbracket}$  est indépendante et que  $G_k$  suit une loi géométrique pour un certain paramètre  $\pi_k \in [0, 1]$  (propriété  $I_k$ ).

Initialisation : La famille  $(G_1)$  est indépendante et  $G_1 \hookrightarrow \mathcal{G}(\pi_1)$  avec  $\pi_1 = 1$ .

Hérédité : Soit  $k \in \llbracket 1, r-1 \rrbracket$ , on suppose  $I_k$ . Montrons que  $(G_i)_{i \in \llbracket 1, k+1 \rrbracket}$  est indépendante.

Soit  $g_{k+1} \in \mathbb{N}^*$ . On a :

$$P(G_{k+1} = g_{k+1}) = \sum_{g_1, \dots, g_k \in \mathbb{N}^*} P\left(\bigcap_{i=1}^{k+1} (G_i = g_i)\right) = \sum_{g_1, \dots, g_k \in \mathbb{N}^*} P\left(G_{k+1} = g_{k+1} \mid \bigcap_{i=1}^k (G_i = g_i)\right) P\left(\bigcap_{i=1}^k (G_i = g_i)\right)$$

On fixe  $g_1, \dots, g_k \in \mathbb{N}^*$ . En regardant la définition de  $G_{k+1}$ , on a :

$$G_{k+1} \mid \bigcap_{i=1}^k (G_i = g_i) = \min\{n \in \mathbb{N}^* / X_{g_1+\dots+g_k+n} \notin X_1 + \dots + X_{g_1+\dots+g_k}\}$$

Comme  $|\{X_1, \dots, X_{G_1+\dots+G_k}\}| = k$  et les valeurs de  $G_1, \dots, G_k$  sont fixées dans le conditionnement de  $G_{k+1}$  précédent, on obtient finalement par lemme des coalitions que  $G_{k+1} \mid \bigcap_{i=1}^k (G_i = g_i) \hookrightarrow \mathcal{G}(\pi_{k+1})$  où  $\pi_{k+1} = \frac{r-k}{r}$ . En poursuivant le calcul précédent, on obtient finalement que :

$$P(G_{k+1} = g_{k+1}) = \sum_{g_1, \dots, g_k \in \mathbb{N}^*} \frac{k}{r} \left(1 - \frac{k}{r}\right)^{g_{k+1}} P\left(\bigcap_{i=1}^k (G_i = g_i)\right) = \frac{k}{r} \left(1 - \frac{k}{r}\right)^{g_{k+1}} \sum_{g_1, \dots, g_k \in \mathbb{N}^*} P\left(\bigcap_{i=1}^k (G_i = g_i)\right) = \frac{k}{r} \left(1 - \frac{k}{r}\right)^{g_{k+1}}$$

On a alors montré à la fois que  $G_{k+1} \hookrightarrow \mathcal{G}(\pi_{k+1})$  où  $\pi_{k+1} = \frac{r-k}{r}$  et l'indépendance de  $(G_i)_{i \in \llbracket 1, k+1 \rrbracket}$ .

Le calcul des moments de  $T$  est alors simplifié par l'indépendance de  $(G_i)_{i \in \llbracket 1, r \rrbracket}$  et donc, par linéarité de l'espérance et linéarité de la variance dans le cas indépendant, on a :

$$\begin{aligned} E(T) &= \sum_{i=1}^r E(G_i) \\ &= \sum_{i=1}^r \frac{1}{\pi_i} \\ &= r \sum_{i=1}^r \frac{1}{r-i+1} = r \sum_{j=1}^r \frac{1}{j} = r(\log r + \gamma) + o(r) \end{aligned}$$

$$\begin{aligned}
V(T) &= \sum_{i=1}^r V(G_i) \\
&= \sum_{i=1}^r \frac{1 - \pi_i}{\pi_i^2} \\
&= r \sum_{i=1}^r \frac{r - r + i - 1}{(r - i + 1)^2} = r \sum_{i=j}^{r-1} \frac{r - j}{j^2} = \frac{\pi^2}{6} r^2 - r(\log r + \gamma) + o(r^2)
\end{aligned}$$

□

**Théorème 16.** (*Queue de distribution*) Pour tout  $n \geq 1$  :

$$P(T > n) = \sum_{k=1}^r (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^n$$

*Démonstration.* Soit  $n \geq 1$ . pour tout  $i \in \llbracket 1, r \rrbracket$ , on note  $E_{n,i} = \bigcap_{k=1}^n (X_k \neq i)$ . Alors :

$$\{T > n\} = \bigcup_{i=1}^r E_{n,i}$$

(dire que  $T > n$ , cela revient à dire qu'une face du dé n'est pas apparue après les  $n$  premiers lancers)

Soit  $k \in \llbracket 1, r \rrbracket$ .

Si  $i_1, \dots, i_k \in \llbracket 1, r \rrbracket$  sont deux-à-deux distincts, alors, en notant  $R = \llbracket 1, r \rrbracket \setminus \{i_1, \dots, i_k\}$ , on a, par indépendance de  $(X_i)_{i \in \llbracket 1, n \rrbracket}$  :

$$P\left(\bigcap_{j=1}^k E_{n,i_j}\right) = P\left(\bigcap_{j=1}^n (X_j \notin \{i_1, \dots, i_k\})\right) = P\left(\bigcap_{j=1}^n X_j \in R\right) = \prod_{j=1}^n P(X_j \in R) = \left(\frac{r-k}{r}\right)^n = \left(1 - \frac{k}{r}\right)^n$$

Par la formule du crible, finalement :

$$\begin{aligned}
P(T > n) &= P\left(\bigcup_{i=1}^r E_{n,i}\right) = \sum_{k=1}^r \left( (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P\left(\bigcap_{j=1}^k E_{n,i_j}\right) \right) \\
&= \sum_{k=1}^r \left( (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \left(1 - \frac{k}{r}\right)^n \right) \\
&= \sum_{k=1}^r (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^n
\end{aligned}$$

□

À partir de ces propriétés, on va tenter de trouver un comportement asymptotique pour  $P(T > n)$  quand  $r$  tend vers  $+\infty$  et  $n$  dépend de  $r$ , car le résultat précédent ne suffit pas du fait du changement de signe entre les termes de la somme.

### 5.3 Étude asymptotique

On peut déjà commencer par une majoration.

**Théorème 17.** (*Déviaton*) Pour tout réel  $t > 0$  :

$$P(T > rt + r \ln r) = P\left(\frac{T - r \ln r}{r} > t\right) \leq e^{-t+\epsilon_r}$$

où  $\epsilon_r = \frac{1}{r}(rt + r \ln r - \lfloor rt + r \ln r \rfloor) \in [0, \frac{1}{r}[$ .

*Démonstration.* Pour tout  $n \in \mathbb{N}^*$ , on a, par inégalité de Boole :

$$P(T > n) = P\left(\bigcup_{i=1}^r E_{n,i}\right) \leq \sum_{i=1}^r P(E_{n,i}) = \sum_{i=1}^r \left(1 - \frac{1}{r}\right)^n = r e^{n \ln(1-\frac{1}{r})} \leq r e^{-\frac{n}{r}}$$

Soit  $t > 0$ . On pose  $n = \lfloor rt + r \ln r \rfloor$ . Alors :

$$P(T > rt + r \ln r) = P(T > n) \leq r e^{-\frac{n}{r}} = r e^{\epsilon_r - t - \ln r} = r e^{\ln(\frac{1}{r})} e^{-t+\epsilon_r} = e^{-t+\epsilon_r}$$

□

**Remarque 5.** Pour  $\alpha > 0$  et  $r \in \mathbb{N}^*$  fixés, on peut choisir  $t$  assez grand pour que  $e^{-t+\epsilon_r} \leq \alpha$ , par exemple  $t = -\ln(\alpha) + \frac{1}{r}$ ; on retrouve encore le résultat  $P(T = +\infty) = 0$  (il est presque sûr de trouver tous les coupons).

Avec ce résultat, on peut désormais trouver un début de comportement asymptotique pour  $T$ .

**Théorème 18.** On a :

$$\frac{T}{r \ln r} \xrightarrow[r \rightarrow +\infty]{P} 1$$

*Démonstration.* Soit  $\epsilon > 0$ . Par l'inégalité de Markov et le lemme démontré précédemment, comme  $T$  admet un moment d'ordre 2, on a :

$$\begin{aligned} P\left(\left|\frac{T}{r \ln r} - 1\right| > \epsilon\right) &\leq \frac{E((T - r \ln r)^2)}{\epsilon^2 r^2 (\ln r)^2} = \frac{\text{Var}(T) + (E(T) - r \ln r)^2}{\epsilon^2 r^2 (\ln r)^2} \\ &\stackrel{r \rightarrow +\infty}{=} \frac{1}{\epsilon^2 r^2 (\ln r)^2} (\frac{\pi^2}{6} r^2 + o(r^2) + (r\gamma + o(r))^2) \\ &\stackrel{r \rightarrow +\infty}{=} \frac{1}{\epsilon^2 r^2 (\ln r)^2} (\frac{\pi^2}{6} r^2 + r^2 \gamma^2 + o(r^2)) \\ &\stackrel{r \rightarrow +\infty}{=} O\left(\frac{1}{(\ln r)^2}\right) \end{aligned}$$

D'où :

$$\frac{T}{r \ln r} \xrightarrow[r \rightarrow +\infty]{P} 1$$

□

Comme la famille  $\left(\frac{1}{(\ln r)^2}\right)_{r \geq 2}$  n'est pas sommable (critère sur les séries de Bertrand), on ne peut pas utiliser le lemme de Borel-Cantelli pour montrer une éventuelle convergence presque sûre.

On peut néanmoins prouver une convergence en loi vers une loi de Gumbel.

**Théorème 19.** (Fluctuations asymptotiques) On a :

$$\frac{T - r \ln r}{r} = \ln r \left( \frac{T}{r \ln r} - 1 \right) \xrightarrow[r \rightarrow +\infty]{\mathcal{L}} \text{Gumbel}$$

où la loi de Gumbel a pour fonction de répartition  $t \in \mathbb{R} \mapsto e^{-e^{-t}}$ .

**Illustration 1.** La figure 6 illustre l'approximation de la queue de distribution de la variable aléatoire  $T$  par celle de  $r \ln r + rG$  où  $G \hookrightarrow \text{Gumbel}$  (par le Théorème énoncé ci-dessus) ainsi que par une méthode de Monte Carlo avec  $N = 4000$  tirages et  $r = 20$ , en utilisant la relation  $T = \sum_{i=1}^r G_i$ .

*Démonstration.* Montrons la convergence en loi en montrant la convergence de la suite des fonctions de répartition associée en tout point de  $\mathbb{R}$  (car celle de la loi de Gumbel est continue sur  $\mathbb{R}$ ).

Soit  $t \in \mathbb{R}$ . Montrons que  $\lim_{r \rightarrow +\infty} P(T > r \ln r + rt) = S(t)$  où  $S$  est la fonction  $x \in \mathbb{R} \mapsto 1 - e^{-e^{-x}}$ .

Soit  $r \in \mathbb{N}^*$ .

Supposons  $r$  assez grand pour que  $r \ln r + rt > r$ . Soit  $n_{t,r} = \lfloor r \ln r + rt \rfloor$ . On sait que, par Théorème :

$$P(T > r \ln r + rt) = P(T > n_{t,r}) = \sum_{k=1}^r (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^{n_{t,r}}$$

Comme  $\binom{r}{k} = \frac{\prod_{i=0}^{k-1} (r-i)}{k!} \leq \frac{r^k}{k!}$  pour tout  $k \in \llbracket 1, r \rrbracket$  et  $e^{-u} \geq 1 - u$  pour tout  $u \geq 0$  (convexité de la fonction exponentielle), on a :

$$\forall k \in \llbracket 1, r \rrbracket, \binom{r}{k} \left(1 - \frac{k}{r}\right)^{n_{t,r}} \leq \frac{r^k}{k!} e^{-\frac{kn_{t,r}}{r}} \leq \frac{r^k}{k!} e^{\frac{k}{r} - (rt+r \ln r) \frac{k}{r}} \leq \frac{r^k}{k!} e^{-tk} e^{-k \ln r} = e \times \frac{e^{-tk}}{k!}$$

Aussi, en ne fixant plus  $r$ , on a :

$$\begin{aligned} \forall k \in \mathbb{N}^*, \forall r \geq k, \binom{r}{k} \left(1 - \frac{k}{r}\right)^{n_{t,r}} &= \binom{r}{k} e^{\lfloor r \ln r + rt \rfloor \ln(1 - \frac{k}{r})} \\ &\underset{r \rightarrow +\infty}{\sim} \frac{r^k}{k!} e^{\lfloor r \ln r + rt \rfloor \ln(1 - kr)} \\ &= \frac{r^k}{k!} e^{(r \ln r + rt + O(1))(-\frac{k}{r} - \frac{k^2}{2r^2} + o(\frac{1}{r^2}))} \\ &\stackrel{r \rightarrow +\infty}{=} \frac{r^k}{k!} e^{-k \ln r - kt - \frac{k^2}{2r} \ln r - \frac{k^2 t}{2r} + O(\frac{1}{r}) + o(\frac{\ln r}{r})} \\ &= \frac{r^k}{k!} r^{-k} e^{-kt} e^{o(1)} \\ &\stackrel{r \rightarrow +\infty}{=} \frac{e^{-kt}}{k!} (1 + o(1)) \xrightarrow[r \rightarrow +\infty]{} \frac{e^{-kt}}{k!} \end{aligned}$$

Répartition Gumbel vs taille de la queue empirique, 4000 expériences ( $r = 20$ ).

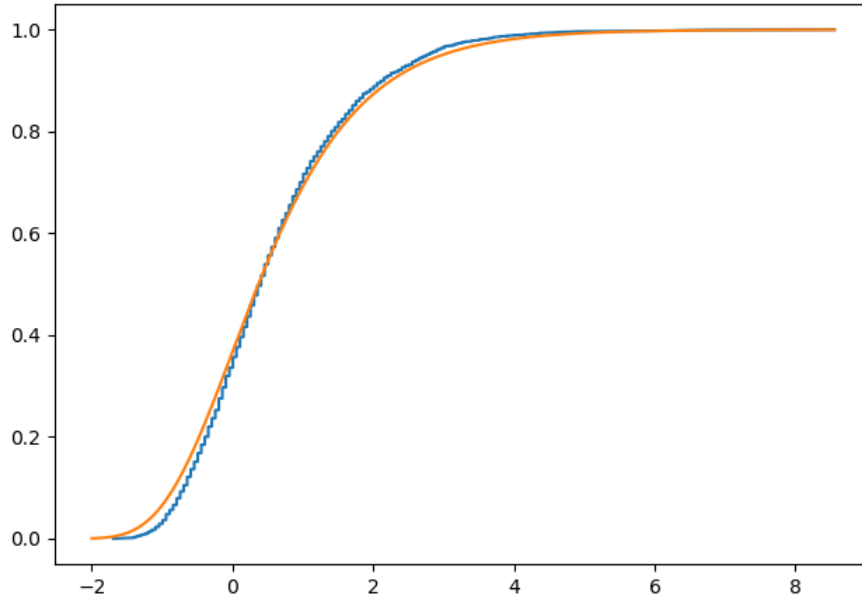


FIGURE 6 – Répartition empirique de  $\frac{1}{r} \times$  taille de la queue comparée à celle de la loi de Gumbel.

Comme  $\left(\sum e^{\frac{-kt}{k!}}\right)$  converge absolument, par Théorème de Convergence Dominée :

$$\lim_{r \rightarrow +\infty} \sum_{k=1}^r (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^{n_{t,r}} = \sum_{k=1}^{+\infty} (-1)^{k-1} \frac{e^{-kt}}{k!} = 1 - \sum_{k=0}^{+\infty} \frac{(-e^{-t})^k}{k!} = 1 - e^{-e^{-t}}$$

Ceci prouve la convergence en loi. □

**Remarque 6.** Par ce résultat et linéarité de la limite, on obtient finalement :

$$\forall a, b \in \mathbb{R}, a \leq b \Rightarrow \lim_{r \rightarrow +\infty} P(T \in [r \ln r + ra, r \ln r + rb]) = e^{-e^{-b}} - e^{-e^{-a}}$$

On peut observer que  $e^{-e^{-t}}$  tend très rapidement vers 1 quand  $t$  tend vers  $+\infty$ , on a alors une montée abrupte de 0 à 1.

Cela donne un phénomène de seuil pour  $T$ , ie  $P(T > n)$  passe soudainement de 1 environ à 0 environ pour  $n = r \ln r$  avec  $r \gg 1$ .

**Théorème 20.** (Convergence abrupte en  $n = r \ln r$ ) Pour tout réel  $c > 0$ , par Théorème :

$$P(T > r \ln r + rc) \leq e^{-c + \epsilon_r}$$

avec  $\epsilon_r \in [0, \frac{1}{r}[$

Si  $(c_r)_{r \in \mathbb{N}^*}$  est une suite de réels strictement positifs divergeant vers  $+\infty$ , alors :

$$\lim_{r \rightarrow +\infty} P(T > r \ln r - rc_r) = 1$$

*Démonstration.* On sait que  $\frac{T}{r \ln r} \xrightarrow{\mathcal{L}}_{r \rightarrow +\infty}$  Gumbel, montrons alors que  $P(T > r \ln r - rc_r) = P\left(\frac{T - r \ln r}{r} > -c_r\right) \xrightarrow{r \rightarrow +\infty} 1$ .

Soit  $A > 0$ . On sait que :

$$\exists r_0 \in \mathbb{N}^* / \forall r \geq r_0, -A > -c_r$$

D'où :

$$\forall r \geq r_0, 1 \geq P\left(\frac{T - r \ln r}{r} > -c_r\right) \geq P\left(\frac{T - r \ln r}{r} > -A\right) \geq 0$$

D'où :

$$1 \geq \liminf_{r \rightarrow +\infty} P\left(\frac{T - r \ln r}{r} > -c_r\right) \geq \liminf_{r \rightarrow +\infty} P\left(\frac{T - r \ln r}{r} > -A\right) = \lim_{r \rightarrow +\infty} P\left(\frac{T - r \ln r}{r} > -A\right) = 1 - e^{-e^{-A}}$$

(par le Théorème précédent)

Ceci étant vrai pour tout  $A > 0$ , on a, par encadrement :

$$\liminf_{r \rightarrow +\infty} P\left(\frac{T - r \ln r}{r} > -c_r\right) = 1$$

Comme  $P$  est de masse 1, on a finalement :

$$\liminf_{r \rightarrow +\infty} P\left(\frac{T - r \ln r}{r} > -c_r\right) = \limsup_{r \rightarrow +\infty} P\left(\frac{T - r \ln r}{r} > -c_r\right) = \lim_{r \rightarrow +\infty} P\left(\frac{T - r \ln r}{r} > -c_r\right) = 1$$

□

## 6 Conclusion

Finalement, en partant seulement de la loi de Bernoulli ou plus simplement du jeu du Pile ou Face, équilibré ou non, la création de nouvelles variables élémentaires telles que la loi binomiale, la loi de Rademacher ou même la loi Géométrique, il est possible de modéliser des situations d'un point de vue probabiliste et d'y établir des résultats sophistiqués et qui parfois contredisent l'intuition humaine, faisant appel à énormément d'outils de la théorie des probabilités et de l'analyse réelle.

## Références

- [1] Djalil Chafai and Florent Malrieu. *Recueil de modèles aléatoires*. 2018. Chapitre 1.
- [2] William Feller. *An introduction to probability theory and its applications. Vol. I*. Troisième édition. John Wiley & Sons Inc., New York, 1968. Chapitre 3.