La loi géométrique

Marc Lorenzi

10 juillet 2024

1 Loi géométrique

1.1 Le premièr succès

Soit $p \in]0,1]$. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, suivant une loi de Bernoulli de paramètre p.

Pour tout $\omega \in \Omega$, soit

$$S(\omega) = \{n \in \mathbb{N}^* : X_n(\omega) = 1\}$$

Posons $\overline{\mathbb{N}} = \mathbb{N}^* \cup \{\infty\}$. L'ensemble $\overline{\mathbb{N}}$ est dénombrable, munissons-le de la tribu $\mathcal{T} = \mathcal{P}(\overline{\mathbb{N}})$.

Soit $T : \Omega \longrightarrow \overline{\mathbb{N}}$ définie par $T(\omega) = \min S(\omega)$ si $S(\omega) \neq \emptyset$ et $T(\omega) = \infty$ si $S(\omega) = \emptyset$.

De façon informelle, $X_n(\omega)$ décrit le succès ou l'échec d'une expérience aléatoire sur ω. L'entier $T(\omega)$ est le nombre d'essais nécessaires à l'obtention d'un succès, c'est à dire le *temps* d'attente avant le premier succès de l'expérience. Dans cette section et la suivante, nous allons étudier les propriétés de T. Nous nous intéresserons ensuite aux temps d'attente avant le second succès, le troisième succès, etc.

1.2 La loi de T

Commençons par un résultat sans lequel il n'y aurait pas d'article : T est bien une variable aléatoire.

Proposition 1. *T est mesurable.*

Démonstration. Comme l'ensemble d'arrivée de T est dénombrable, il suffit de montrer que pour tout $n \in \overline{\mathbb{N}}$, $\{T = n\} \in \mathcal{F}$.

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

$$\{T = n\} = \{X_1 = 0, \ldots, X_{n-1} = 0, X_n = 1\} = \{X_n = 1\} \cap \bigcap_{k=1}^{n-1} \{X_k = 0\}$$

L'ensemble $\{T = n\}$ est une intersection (finie) d'événements de la tribu $\mathcal{F}$ donc $\{T = n\} \in \mathcal{F}$. On a aussi

$$\{T = \infty\} = \{X_1 = 0, X_2 = 0, \ldots\} = \bigcap_{k=1}^{\infty} \{X_k = 0\}$$

L'ensemble $\{T = \infty\}$ est une intersection (dénombrable) d'événements de la tribu $\mathcal{F}$ donc $\{T = \infty\} \in \mathcal{F}$. □

La fonction T est ainsi une variable aléatoire définie sur Ω, à valeurs dans $\overline{\mathbb{N}}$.

Proposition 2. *Pour tout $n \in \mathbb{N}^*$, $\mathbb{P}(T = n) = p(1-p)^{n-1}$.*

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

$$\{T = n\} = \{X_1 = 0, \dots, X_{n-1} = 0, X_n = 1\}$$

Par l'indépendance des X_k,

$$\mathbb{P}(T = n) = \mathbb{P}(X_1 = 0) \dots \mathbb{P}(X_{n-1} = 0)\mathbb{P}(X_n = 1) = (1-p)^{n-1}p$$

□

Proposition 3. $\mathbb{P}(T = \infty) = 0$.

Démonstration. Pour tout $n \in \mathbb{N}^*$,

$$\{T = \infty\} = \{X_1 = 0, X_2 = 0, \dots\} \subseteq \{X_1 = 0, \dots, X_n = 0\}$$

De là, par l'indépendance des X_k,

$$\mathbb{P}(T = \infty) \leq \mathbb{P}(X_1 = 0) \dots \mathbb{P}(X_n = 0) = (1-p)^n$$

Comme $0 < p < 1$, $0 < 1 - p < 1$ et donc $(1-p)^n$ tend vers 0 lorsque n tend vers l'infini. □

Remarque. Nous avons donc $\mathbb{P}(T \in \mathbb{N}^*) = 1$. Le temps d'attente avant le premier succès de l'expérience est presque sûrement fini.

1.3 La loi géométrique

Une définition s'impose.

Définition 1. Soit $p \in]0, 1]$. La variable aléatoire $T : \Omega \longrightarrow \overline{\mathbb{N}}$ suit une *loi géométrique* de paramètre p si pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}(T = n) = p(1-p)^{n-1}$$

On note $T \sim \mathcal{G}(p)$.

Remarque. Le cas $p = 1$ est un peu à part. On a dans ce cas $\mathbb{P}(T = 1) = 1$ et pour tout $n \geq 2$, $\mathbb{P}(T = n) = 0$. La variable aléatoire T est donc presque sûrement constante, égale à 1.

Si $p \neq 1$ alors pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}(T = n+1) - \mathbb{P}(T = n) = p(1-p)^n - p(1-p)^{n-1} = -p^2(1-p)^{n-1} < 0$$

La suite $(\mathbb{P}(T = n))_{n \geq 1}$ est donc strictement décroissante. De plus, elle tend clairement vers 0 lorsque n tend vers l'infini.

Remarque. Nous avons écarté de notre définition la valeur $p = 0$. Pour $p = 0$, nous obtiendrions pour tout $n \in \mathbb{N}^*$, $\mathbb{P}(T = n) = 0$. Ceci entraînerait $\mathbb{P}(T = \infty) = 1$ ou encore que T est presque sûrement infinie, ce qui est assez peu souhaitable.

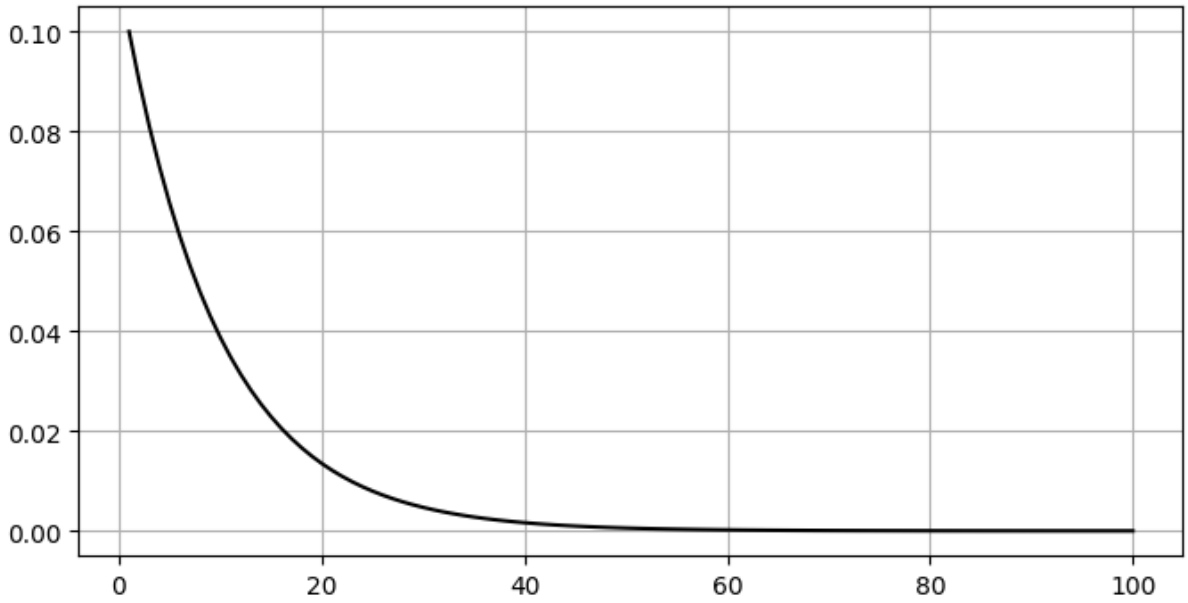

FIGURE 1 – $\mathbb{P}(T = n)$ pour $p = 1/10$

Remarque. Soit $T \sim \mathcal{G}(p)$. Que vaut $\mathbb{P}(T = \infty)$? On a

$$\Omega = \{T \in \overline{\mathbb{N}}\} = \{T = \infty\} \cup \bigcup_{n \in \mathbb{N}^*} \{T = n\}$$

Les événements de la réunion ci-dessus sont disjoints, donc

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(T = \infty) + \sum_{n \in \mathbb{N}^*} \mathbb{P}(T = n)$$

Comme $0 \leq 1 - p < 1$,

$$\begin{aligned} \sum_{n \in \mathbb{N}^*} \mathbb{P}(T = n) &= p \sum_{n \in \mathbb{N}^*} (1-p)^{n-1} \\ &= \frac{p}{1 - (1-p)} \\ &= 1 \end{aligned}$$

Ainsi, $1 = \mathbb{P}(T = \infty) + 1$ et donc $\mathbb{P}(T = \infty) = 0$.

Nous pouvons donc donner une définition quasi-équivalente de la loi géométrique pour des variables aléatoires à valeurs dans $\mathbb{N}^*$.

Définition 2. Soit $p \in]0, 1]$. La variable aléatoire $T : \Omega \longrightarrow \mathbb{N}^*$ suit une *loi géométrique* de paramètre p si pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}(T = n) = p(1-p)^{n-1}$$

Remarque. Dans la suite, nous étudierons des propriétés d'une variable aléatoire qui suit une loi géométrique. Lorsque ces propriétés ne dépendent que de la *loi* de T et pas vraiment de T, nous pourrons tout aussi bien supposer que T est le temps d'attente d'un premier succès, défini en 1.1.

1.4 Absence de mémoire

Intéressons-nous à la probabilité que le temps d'attente d'un succès soit plus grand qu'une certaine valeur.

Proposition 4. *Soit $p \in]0,1]$. Soit $T \sim \mathcal{G}(p)$. Pour tout $n \in \mathbb{N}$,*

$$\mathbb{P}(T > n) = (1-p)^n$$

Démonstration. Si $p = 1$, le résultat est immédiat. Supposons $p \neq 1$. Soit $n \in \mathbb{N}$. On a

$$\{T > n\} = \bigcup_{k=n+1}^{\infty} \{T = k\}$$

Cette union est disjointe donc

$$\begin{aligned}
\mathbb{P}(T > n) &= \sum_{k=n+1}^{\infty} \mathbb{P}(T = k) \\
&= p \sum_{k=n+1}^{\infty} (1-p)^{k-1} \\
&= p(1-p)^n \sum_{k=0}^{\infty} (1-p)^k \\
&= p(1-p)^n \frac{1}{1-(1-p)} \\
&= (1-p)^n
\end{aligned}$$

□

Proposition 5. *Soit $p \in]0,1]$. Soit $T \sim \mathcal{G}(p)$. Pour tous $m, n \in \mathbb{N}$,*

$$\mathbb{P}(T > m+n) = \mathbb{P}(T > m)\mathbb{P}(T > n)$$

Démonstration. Soient $m, n \in \mathbb{N}$. On a

$$\begin{aligned}
\mathbb{P}(T > m+n) &= (1-p)^{m+n} \\
&= (1-p)^m (1-p)^n \\
&= \mathbb{P}(T > m)\mathbb{P}(T > n)
\end{aligned}$$

□

Cette propriété traduit ce que l'on appelle *l'absence de mémoire* de la loi géométrique. En termes de probabilités conditionnelles, elle peut aussi s'écrire

$$\mathbb{P}(T > m+n | T > m) = \mathbb{P}(T > n)$$

Si j'ai réalisé une même expérience m fois de suite et que j'ai obtenu m échecs (c'est à dire $T > m$), quelle est la probabilité que je doive faire encore au moins n fois l'expérience avant d'obtenir un succès (c'est à dire $T > m+n$) ? Eh bien c'est la probabilité que je doive faire au moins n fois l'expérience avant d'obtenir un succès (c'est à dire $T > n$). Le fait de savoir que j'ai déjà échoué m fois n'apporte aucune information supplémentaire.

2 Espérance et variance

2.1 Série génératrice

Soit $T : \Omega \longrightarrow \mathbb{N}^*$ une variable aléatoire qui suit une loi géométrique de paramètre p. Considérons la série génératrice de T :

$$\mathbb{G}_T(z) = \mathbb{E}(z^T) = \sum_{k=1}^{\infty} \mathbb{P}(T = k) z^k$$

Posons $q = 1 - p$. Pour tout $n \in \mathbb{N}^*$,

$$\frac{\mathbb{P}(T = n + 1)}{\mathbb{P}(T = n)} = 1 - p = q$$

Notre série entière a donc pour rayon de convergence $R = 1/q$ (si $p = 1$, $R = \infty$). Pour tout $z \in \mathbb{C}$ tel que $|z| < R$,

$$\begin{aligned} \mathbb{G}_T(z) &= p \sum_{k=1}^{\infty} (1-p)^{k-1} z^k \\ &= pz \sum_{k=1}^{\infty} (qz)^{k-1} \\ &= \frac{pz}{1 - qz} \end{aligned}$$

2.2 Espérance

Proposition 6. *L'espérance de T est finie et*

$$\mathbb{E}(T) = \frac{1}{p}$$

Démonstration. La fonction $\mathbb{G}_T$ est indéfiniment dérivable sur le disque ouvert D de centre 0 et de rayon $R = 1/q$ et sa dérivée est obtenue en dérivant la série terme à terme. Pour tout $z \in D$,

$$\mathbb{G}'_T(z) = \frac{p}{(1 - qz)^2} = \sum_{k=1}^{\infty} k \mathbb{P}(T = k) z^{k-1}$$

Remarquons que $R = 1/q > 1$. Ainsi, $1 \in D$. En évaluant ci-dessus en 1,

$$\mathbb{G}'_T(1) = \frac{p}{(1 - q)^2} = \frac{1}{p} = \sum_{k=1}^{\infty} k \mathbb{P}(T = k) = \mathbb{E}(T)$$

□

2.3 Variance

En redérivant, on obtient pour tout $z \in D$

$$\mathbb{G}''_T(z) = \frac{2pq}{(1 - qz)^3} = \sum_{k=1}^{\infty} k(k-1) \mathbb{P}(T = k) z^{k-2}$$

De là,

$$\mathbb{G}_T''(z) + \mathbb{G}_T'(z) = \sum_{k=1}^{\infty}(k^2-k)\mathbb{P}(T=k)z^{k-2} + \sum_{k=1}^{\infty} k\mathbb{P}(T=k)z^{k-1}$$

En particulier,

$$\mathbb{G}_T''(1) + \mathbb{G}_T'(1) = \sum_{k=1}^{\infty}(k^2-k)\mathbb{P}(T=k) + \sum_{k=1}^{\infty} k\mathbb{P}(T=k) = \sum_{k=1}^{\infty} k^2\mathbb{P}(T=k) = \mathbb{E}(T^2)$$

Proposition 7. *La variance de T est finie et*

$$\mathbb{V}(T) = \frac{1-p}{p^2}$$

Démonstration. Par les remarques qui précèdent,

$$\mathbb{E}(T^2) = \mathbb{G}_T''(1) + \mathbb{G}_T'(1) = \frac{2pq}{(1-q)^3} + \frac{1}{p} = \frac{2q}{p^2} + \frac{1}{p}$$

De là,

$$\mathbb{V}(T) = \mathbb{E}(T^2) - \mathbb{E}(T)^2 = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q}{p^2} + \frac{p-1}{p^2} = \frac{q}{p^2}$$

□

3 Le k^{e} succès

3.1 Introduction

Reprenons la suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes définies sur Ω et suivant une loi de Bernoulli de paramètre p. Pour tout $k \geq 1$, définissons la fonction $Z_k : \Omega \longrightarrow \overline{\mathbb{N}}$ par récurrence sur k en posant $Z_1(\omega) = T(\omega)$ et pour tout $k \geq 2$,

- Si $Z_{k-1}(\omega) \neq \infty$ et l'ensemble $\{n > Z_{k-1}(\omega) : X_n(\omega) = 1\}$ est non vide,

$$Z_k(\omega) = \min\{n > Z_{k-1}(\omega) : X_n(\omega) = 1\}$$

- Sinon, $Z_k(\omega) = \infty$.

Pour tout $\omega \in \Omega$ et tout $k \geq 1$, $Z_k(\omega)$ est le temps d'attente avant le k^{e} succès de l'expérience.

Soit $\omega \in \Omega$. Supposons que pour tout $k \geq 1$, $Z_k(\omega) < \infty$. La suite $(Z_k(\omega))_{k\geq 1}$ est une suite strictement croissante d'entiers naturels non nuls. On a donc en particulier, pour tout $k \geq 1$, $Z_k(\omega) \geq k$.

Proposition 8. *Pour tout $k \geq 1$, Z_k est mesurable.*

Démonstration. La preuve se fait par récurrence sur k. Elle est analogue à celle que nous avons faite pour montrer que T est mesurable. □

3.2 La loi de Z_k

Proposition 9. *Pour tous $n, k \geq 1$,*

$$\mathbb{P}(Z_k = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}$$

Démonstration. Soient $n, k \geq 1$. On peut décomposer l'événement $\{Z_k = n\}$ comme suit.

$$\{Z_k = n\} = \bigcup_{A \in \mathcal{P}_{k-1}([\![1,n-1]\!])} E_A$$

où

$$E_A = \{X_n = 1\} \cap \bigcap_{i \in A} \{X_i = 1\} \cap \bigcap_{i \notin A} \{X_i = 0\}$$

Par l'indépendance des X_i,

$$\begin{aligned} \mathbb{P}(E_A) &= \mathbb{P}(X_n = 1) \prod_{i \in A} \mathbb{P}(X_i = 1) \prod_{i \notin A} \mathbb{P}(X_i = 0) \\ &= p^k (1-p)^{n-k} \end{aligned}$$

Les ensembles E_A sont disjoints deux à deux. De là,

$$\begin{aligned} \mathbb{P}(Z_k = n) &= \sum_{A \in \mathcal{P}_{k-1}([\![1,n-1]\!])} \mathbb{P}(E_A) \\ &= \sum_{A \in \mathcal{P}_{k-1}([\![1,n-1]\!])} p^k (1-p)^{n-k} \\ &= \binom{n-1}{k-1} p^k (1-p)^{n-k} \end{aligned}$$

□

Remarque. Pour $k = 1$, nous retrouvons

$$\mathbb{P}(Z_1 = n) = \mathbb{P}(T = n) = p(1-p)^{n-1}$$

Soit $k \geq 1$. Pour tout $n \geq k$, on a

$$\begin{aligned} \frac{\mathbb{P}(Z_k = n+1)}{\mathbb{P}(Z_k = n)} &= \frac{\binom{n}{k-1} p^k (1-p)^{n+1-k}}{\binom{n-1}{k-1} p^k (1-p)^{n-k}} \\ &= \frac{n(1-p)}{n+1-k} \end{aligned}$$

On a donc $\mathbb{P}(Z_k = n+1) \geq \mathbb{P}(Z_k = n)$ si et seulement si

$$n(1-p) \geq n+1-k$$

c'est à dire

$$n \leq \frac{k-1}{p}$$

La suite $(\mathbb{P}(Z_k = n))_{n \geq 1}$ est donc croissante jusqu'à environ $(k-1)/p$, puis décroissante. Lorsque n tend vers l'infini,

$$\frac{\mathbb{P}(Z_k = n+1)}{\mathbb{P}(Z_k = n)} \longrightarrow 1 - p < 1$$

Il en résulte classiquement que $\mathbb{P}(Z_k = n)$ tend vers 0. La figure ci-dessous représente $\mathbb{P}(Z_k = n)$ en fonction de n pour $k \in [\![1, 10]\!]$. On a pris $p = 1/10$. Les maxima sont obtenus pour $n = 10, 20, 30, \ldots, 100$.

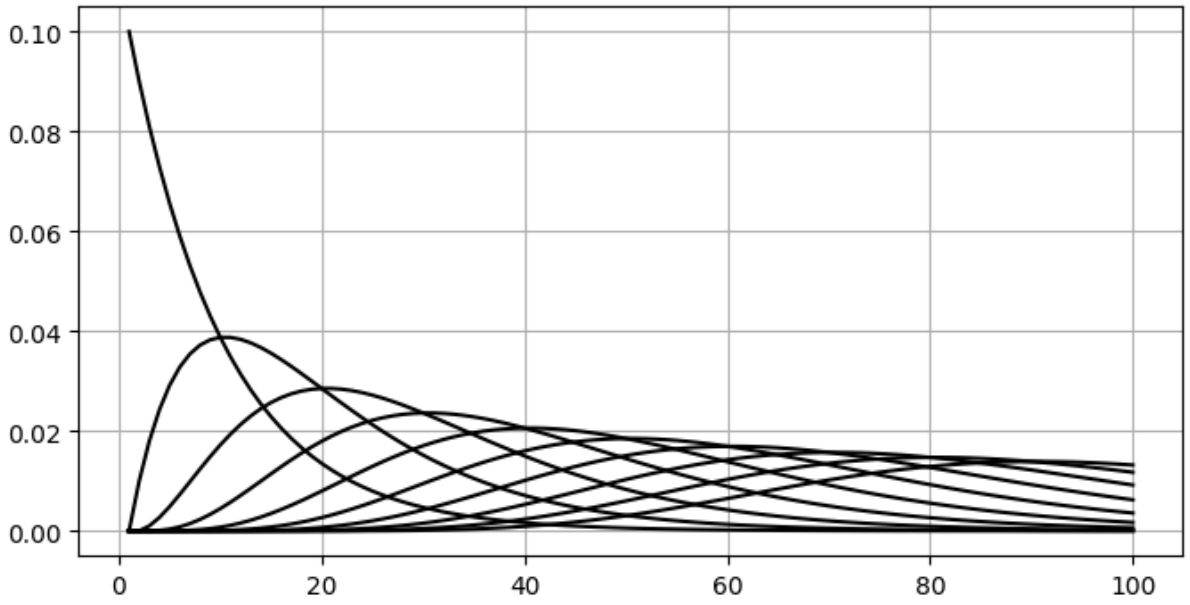

FIGURE 2 – $\mathbb{P}(Z_k = n)$ pour $k \in [\![1, 10]\!]$ et $p = 1/10$

3.3 La probabilité que Z_k soit infini

Pour déterminer complètement la loi de Z_k, il nous reste à calculer $\mathbb{P}(Z_k = \infty)$. Nous allons heureusement trouver que cette probabilité est nulle.

Lemme 10. *Pour tout $k \in \mathbb{N}$ et tout $z \in \mathbb{C}$ tel que $|z| < 1$,*

$$\sum_{n=k}^{\infty} \binom{n}{k} z^n = \frac{z^k}{(1-z)^{k+1}}$$

Démonstration. Considérons la série géométrique

$$F(z) = \sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$$

de rayon de convergence 1. Pour tout $k \in \mathbb{N}$ et tout $z \in \mathbb{C}$ tel que $|z| < 1$, on obtient en dérivant k fois l'égalité ci-dessus

$$\begin{aligned} F^{(k)}(z) &= \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} z^{n-k} \\ &= k! \sum_{n=k}^{\infty} \binom{n}{k} z^{n-k} \\ &= \frac{k!}{(1-z)^{k+1}} \end{aligned}$$

Ainsi,

$$\sum_{n=k}^{\infty} \binom{n}{k} z^{n-k} = \frac{1}{(1-z)^{k+1}}$$

□

Proposition 11. *Pour tout $k \geq 1$,*

$$\sum_{n=1}^{\infty} \mathbb{P}(Z_k = n) = 1$$

Démonstration. Soit $k \geq 1$. Si $p = 1$ le résultat est immédiat : pour tout $n \neq k$, $\mathbb{P}(Z_k = n) = 0$, et $\mathbb{P}(Z_k = k) = 1$. Supposons donc $p \neq 1$. On a

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{P}(Z_k = n) &= \sum_{n=1}^{\infty} \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
&= \frac{p^k}{(1-p)^{k-1}} \sum_{n=1}^{\infty} \binom{n-1}{k-1} (1-p)^{n-1} \\
&= \frac{(1-q)^k}{q^{k-1}} \sum_{n=1}^{\infty} \binom{n-1}{k-1} q^{n-1} \\
&= \frac{(1-q)^k}{q^{k-1}} \sum_{n=0}^{\infty} \binom{n}{k-1} q^n
\end{aligned}$$

où $q = 1-p$. Pour tout $n < k-1$, $\binom{n}{k-1} = 0$. On peut donc faire démarrer la somme ci-dessus à $n = k-1$. Par le lemme 10 (avec $z = q$ et $k-1$ au lieu de k), la somme est justement égale à $q^{k-1}/(1-q)^k$, d'où le résultat. □

Corollaire 12. *Pour tout $k \geq 1$, $\mathbb{P}(Z_k = \infty) = 0$.*

3.4 La série génératrice de Z_k

Soit $k \geq 1$. Nous avons vu plus haut que pour tout $n \geq k$,

$$\frac{\mathbb{P}(Z_k = n+1)}{\mathbb{P}(Z_k = n)} = \frac{n(1-p)}{n+1-k}$$

Cette quantité tend vers $q = 1-p$ lorsque n tend vers l'infini. La série génératrice de Z_k

$$\mathbb{G}_{Z_k}(z) = \sum_{n=k}^{\infty} \mathbb{P}(Z_k = n) z^n$$

a donc pour rayon de convergence $R = 1/q$ (en convenant que $R = \infty$ si $q = 0$). Rappelons-nous que la série génératrice de T a le même rayon de convergence et que pour tout $z \in \mathbb{C}$ tel que $|z| < R$,

$$\mathbb{G}_T(z) = \frac{pz}{1-qz}$$

Proposition 13. *Pour tout $k \geq 1$ et tout $z \in \mathbb{C}$ tel que $|z| < R$,*

$$\mathbb{G}_{Z_k}(z) = \mathbb{G}_T(z)^k$$

Démonstration. Soit $k \geq 1$. Soit $z \in \mathbb{C}$ tel que $|z| < R$. On a

$$\begin{aligned}
\mathbb{G}_{Z_k}(z) &= \sum_{n=k}^{\infty} \mathbb{P}(Z_k = n) z^n \\
&= \sum_{n=k}^{\infty} \binom{n-1}{k-1} p^k (1-p)^{n-k} z^n \\
&= \left(\frac{p}{1-p}\right)^k \sum_{n=k}^{\infty} \binom{n-1}{k-1} ((1-p)z)^n \\
&= \left(\frac{p}{1-p}\right)^k (1-p)z \sum_{n=k-1}^{\infty} \binom{n-1}{k-1} ((1-p)z)^n
\end{aligned}$$

Par le lemme 10,

$$\sum_{n=k-1}^{\infty} \binom{n-1}{k-1} ((1-p)z)^n = \frac{(1-p)^{k-1} z^{k-1}}{(1-(1-p)z)^k}$$

On a donc

$$\begin{aligned}
\mathbb{G}_{Z_k}(z) &= \left(\frac{p}{1-p}\right)^k (1-p)z \frac{(1-p)^{k-1} z^{k-1}}{(1-(1-p)z)^k} \\
&= \left(\frac{pz}{1-qz}\right)^k \\
&= \mathbb{G}_T(z)^k
\end{aligned}$$

□

4 Différences

4.1 Le temps d'attente entre deux succès

Posons $T_1 = Z_1 = T$ et pour tout $k \geq 2$, $T_k = Z_k - Z_{k-1}$, en convenant que $\infty - a = \infty$ pour tout $a \in \overline{\mathbb{N}}$. Si $k \geq 2$, T_k est le temps d'attente entre le $(k-1)^{\text{e}}$ succès et le k^{e} succès de l'expérience.

Remarquons que si $k \geq 2$,

$$\{T_k = \infty\} = \{Z_k = \infty\} \cup \{Z_{k-1} = \infty\}$$

L'événement $\{T_k = \infty\}$ est la réunion de deux événements de probabilité nulle, donc

$$\mathbb{P}(T_k = \infty) = 0$$

Remarquons aussi que pour tout $k \geq 1$,

$$Z_k = T_1 + \ldots + T_k$$

Proposition 14. *Pour tout $k \geq 1$, la fonction T_k est mesurable.*

Démonstration. C'est la différence de deux fonctions mesurables. □

4.2 La loi de T_k

Proposition 15. *Pour tout $k \geq 1$, $T_k \sim \mathcal{G}(p)$.*

Démonstration. C'est clair si $k = 1$ puisque $T_1 = T$. Soit $k \geq 2$. Soient $m, n \in \mathbb{N}^*$. On a

$$\{Z_{k-1} = m, T_k = n\} = \{Z_{k-1} = m, X_{m+1} = 0, \ldots, X_{m+n-1} = 0, X_{m+n} = 1\}$$

En reprenant les idées de la preuve de la proposition 9,

$$\{Z_{k-1} = m\} = \bigcup_{A \in \mathcal{P}_{k-2}([\![1, m-1]\!])} E_A$$

où

$$E_A = \{X_m = 1\} \cap \bigcap_{i \in A} \{X_i = 1\} \cap \bigcap_{i \notin A} \{X_i = 0\}$$

De là,

$$\{Z_{k-1} = m, T_k = n\} = \bigcup_{A \in \mathcal{P}_{k-2}([\![1, m-1]\!])} (E_A \cap \{X_{m+1} = 0, \ldots, X_{m+n-1} = 0, X_{m+n} = 1\})$$

Les X_i étant indépendantes, E_A et $\{X_{m+1} = 0, \ldots, X_{m+n-1} = 0, X_{m+n} = 1\}$ sont aussi indépendants. On a donc

$$\begin{array}{rcl} \mathbb{P}(E_A \cap \{X_{m+1} = 0, \ldots, X_{m+n-1} = 0, X_{m+n} = 1\}) & = & \\ \mathbb{P}(E_A)\mathbb{P}(X_{m+1} = 0, \ldots, X_{m+n-1} = 0, X_{m+n} = 1) & = & \mathbb{P}(E_A)p(1-p)^{n-1} \end{array}$$

Ainsi (union disjointe),

$$\begin{array}{rcl} \mathbb{P}(Z_{k-1} = m, T_k = n) & = & \displaystyle\bigcup_{A \in \mathcal{P}_{k-2}([\![1, m-1]\!])} \mathbb{P}(E_A)p(1-p)^{n-1} \\ & = & p(1-p)^{n-1} \displaystyle\bigcup_{A \in \mathcal{P}_{k-2}([\![1, m-1]\!])} \mathbb{P}(E_A) \\ & = & p(1-p)^{n-1}\mathbb{P}(Z_{k-1} = m) \end{array}$$

Remarquons aussi que $\{Z_{k-1} = \infty, T_k = n\} = \emptyset$. Les événements $\{Z_{k-1} = m\}$, $m \in \overline{\mathbb{N}}$ forment un système complet d'événements. On a donc

$$\begin{array}{rcl} \mathbb{P}(T_k = n) & = & \displaystyle\sum_{m \in \overline{\mathbb{N}}} \mathbb{P}(Z_{k-1} = m, T_k = n) \\ & = & \displaystyle\sum_{m \in \mathbb{N}^*} \mathbb{P}(Z_{k-1} = m, T_k = n) \\ & = & p(1-p)^{n-1} \displaystyle\sum_{m \in \mathbb{N}^*} \mathbb{P}(Z_{k-1} = m) \\ & = & p(1-p)^{n-1} \end{array}$$

Ainsi, $T_k \sim \mathcal{G}(p)$. □

4.3 Indépendance

Proposition 16. *Les variables aléatoires T_k sont indépendantes.*

Démonstration. Soit $n \geq 1$. Soient $a_1, \ldots, a_n \in \mathbb{N}^*$. Soit $\omega \in \Omega$. On a pour tout $i \in [\![1, n]\!]$, $T_i(\omega) = a_i$ si et seulement si pour tout $i \in [\![1, n]\!]$, $Z_i(\omega) = a_1 + \ldots + a_i$. Ceci revient à dire que

- Pour tout $i \in [\![1, n]\!]$, $X_{a_1+\ldots+a_i}(\omega) = 1$.
- Pour tout $j \in [\![1, a_1 + \ldots + a_n]\!]$ différent de a_1, $a_1 + a_2$,..., $a_1 + \ldots + a_n$, $X_j(\omega) = 0$.

Notons

$$A = [\![1, a_1 + \ldots + a_n]\!] \setminus \{a_1, a_1 + a_2, \ldots, a_1 + \ldots + a_n\}$$

On a par ce qui précède

$$\{T_1 = a_1, \ldots, T_n = a_n\} = \{X_{a_1} = 1, X_{a_1+a_2} = 1, \ldots, X_{a_1+\ldots+a_n} = 1\} \cap \bigcap_{j \in A} \{X_j = 0\}$$

Par l'indépendance des X_k,

$$\begin{aligned} \mathbb{P}(T_1 = a_1, \ldots, T_n = a_n) &= \mathbb{P}(X_{a_1} = 1)\mathbb{P}(X_{a_1+a_2} = 1) \ldots \mathbb{P}(X_{a_1+\ldots+a_n} = 1) \\ &\quad \times \prod_{j \in A} \mathbb{P}(X_j = 0) \\ &= p^n (1-p)^{|A|} \\ &= p^n (1-p)^{a_1+\ldots+a_n-n} \end{aligned}$$

Par ailleurs,

$$\begin{aligned} \prod_{k=1}^{n} \mathbb{P}(T_k = a_k) &= \prod_{k=1}^{n} p(1-p)^{a_k-1} \\ &= p^n (1-p)^{a_1+\ldots+a_n-n} \end{aligned}$$

d'où le résultat. □

Corollaire 17. *Pour tout $k \geq 2$, T_k et $(Z_1, \ldots, Z_{k-1})$ sont indépendantes.*

Démonstration. Soit $k \geq 2$. Soient $a_1, \ldots, a_k \in \overline{\mathbb{N}}$. Deux cas sont à considérer.

Si $\mathbb{P}(Z_1 = a_1, \ldots, Z_{k-1} = a_{k-1}) > 0$, alors $a_1 < a_2 < \ldots < a_{k-1} < \infty$ et

$$\begin{aligned} \mathbb{P}(T_k = a_k | Z_1 = a_1, \ldots, Z_{k-1} = a_{k-1}) &= \mathbb{P}(T_k = a_k | T_1 = a_1, T_2 = a_2 - a_1, \\ &\quad \ldots, T_{k-1} = a_{k-1} - a_{k-2}) \\ &= \mathbb{P}(T_k = a_k) \end{aligned}$$

par l'indépendance de $T_1, \ldots, T_k$.

Si $\mathbb{P}(Z_1 = a_1, \ldots, Z_{k-1} = a_{k-1}) = 0$ alors, a fortiori,

$$\mathbb{P}(Z_1 = a_1, \ldots, Z_{k-1} = a_{k-1}, T_k = a_k) = 0$$

De là,

$$\mathbb{P}(Z_1 = a_1, \ldots, Z_{k-1} = a_{k-1}, T_k = a_k) = \mathbb{P}(Z_1 = a_1, \ldots, Z_{k-1} = a_{k-1})\mathbb{P}(T_k = a_k)$$

□

Cette propriété met en évidence le phénomène de *l'absence de mémoire*. Le temps d'attente entre le succès $k-1$ et le succès k ne dépend pas des instants où les succès $1, 2, \ldots, k-1$ ont été obtenus.

4.4 Espérance et variance de Z_k

Proposition 18. *Pour tout $k \geq 1$,*

$$\mathbb{E}(Z_k) = \frac{k}{p}$$

$$\mathbb{V}(Z_k) = \frac{k(1-p)}{p^2}$$

Démonstration. Pour tout $k \geq 1$,

$$Z_k = \sum_{i=1}^{k} T_i$$

Les T_i sont indépendantes et de même loi $\mathcal{G}(p)$, d'où le résultat. □