

Temps d'attente du k -ème succès, loi de PASCAL ou loi binomiale négative

Thomas CHEN

On sait que la loi géométrique modélise le temps d'attente du premier succès lorsqu'on réalise une infinité de fois la même expérience de Bernoulli de manière indépendante.

On se demande ce qu'il se passe pour le temps d'attente du k -ème succès. Appelée loi de Pascal ou loi binomiale négative, on modélise la situation avec un peu de combinatoire.

Exercice 1. On se donne une pièce qui donne face avec une probabilité $p \in]0, 1[$. On suppose les lancers indépendants.

- On note X_k variable aléatoire qui donne 1 si au k -ème lancer, on a « face » et 0 sinon.
- On note $S_n = \sum_{k=1}^n X_k$ et T_i le temps d'attente du i -ème succès avec comme convention $T_i = +\infty$ si ce i -ème succès n'est pas réalisé.
- On note $\tau_1 = T_1$ et $\tau_i = T_i - T_{i-1}$ lorsque $i \in \mathbb{N}_{\geq 2}$.

On cherche la loi des τ_i .

1. Quel est la loi de T_1 ? On considère $i \geq 2$ dans la suite.
2. Quel est la loi de T_i pour $i \in \mathbb{N}_{\geq 2}$? *Indication : on pourra exprimer $T_i = k$ grâce à S_{k-1} et X_k . La loi de T_i est appelée loi de Pascal ou loi binomiale négative.*
3. Calculer la fonction génératrice des T_i pour $i \in \mathbb{N}_{\geq 2}$.
4. En déduire que T_i a la même loi que la somme de i variable aléatoires indépendantes et identiquement distribuées suivant $\mathcal{G}(p)$.
5. *On souhaite montrer que T_i égale une telle somme ce qui est bien plus fort.*
 - (a) Quelle est la loi de τ_i ? *Indication : on pourra s'aider de T_{i-1} .*
 - (b) Montrer que les $(\tau_i)_i$ sont indépendants.
 - (c) Conclure.

Corrigé :

1. Le cours assure que c'est une loi géométrique.
2. Remarquons déjà que S_n suit une $\mathcal{B}(n, p)$ pour tout $n \in \mathbb{N}^*$. Soit $i \geq 2$. Alors $T_i(\Omega) \subset \mathbb{N}_{\geq i} \cup \{+\infty\}$. Soit donc $k \in \mathbb{N}, k \geq i$. L'événement $(T_i = k)$ égale l'événement $(S_{k-1} = i - 1) \cap (X_k = 1)$. Par ailleurs, par le lemme des coalitions, comme $S_{k-1} = f(X_1, \dots, X_{k-1})$ avec $f(x_1, \dots, x_{k-1}) = \sum_{j=1}^{k-1} x_j$, S_{k-1} et X_k sont indépendantes. On en déduit donc que

$$\begin{aligned}
 \mathbb{P}(T_i = k) &= \mathbb{P}(S_{k-1} = i - 1 \cap X_k = 1) \\
 &= \mathbb{P}(S_{k-1} = i - 1)\mathbb{P}(X_k = 1) \\
 &= \binom{k-1}{i-1} p^{i-1} (1-p)^{k-i} p \\
 &= \binom{k-1}{i-1} p^i (1-p)^{k-i}.
 \end{aligned}$$

Pour conclure que $T_i(\Omega) = \mathbb{N}_{\geq i}$, il y a deux choix. Ou bien on montre que $\mathbb{P}(T_i = +\infty) = 0$, ou bien on montre que

$$\sum_{k=i}^{+\infty} \mathbb{P}(T_i = k) = 1.$$

- Si l'événement $T_i = +\infty$ est réalisé, alors l'ensemble $\{k \in \mathbb{N} : X_k = 1\}$ est majoré. Étant une partie de \mathbb{N} , elle est soit non vide et on note k_0 son plus grand élément, soit $k_0 = 1$. Alors

$$(T_i = +\infty) \subset (\forall k \geq k_0, X_k = 0) = \bigcap_{k \geq k_0} (X_k = 0).$$

Ainsi,

$$\mathbb{P}(T_i = +\infty) \leq \prod_{k \geq k_0} \mathbb{P}(X_k = 0) = \prod_{k \geq k_0} (1 - p) = 0.$$

On obtient donc $\mathbb{P}(T_i = +\infty) = 0$.

- On calcule :

$$\sum_{k=i}^{+\infty} \mathbb{P}(T_i = k) = \sum_{k=i}^{+\infty} \binom{k-1}{i-1} p^i (1-p)^{k-i} = p^i \sum_{k=0}^{+\infty} \binom{k+i-1}{i-1} (1-p)^k.$$

On rappelle que pour tout $x \in]-1, 1[$,

$$\begin{aligned} (1+x)^{-i} &= 1 + \sum_{n=1}^{+\infty} \frac{1}{n!} \prod_{j=0}^{n-1} (-i-j) x^n \\ &= 1 + \sum_{n=1}^{+\infty} \frac{1}{n!} (-1)^n \frac{\prod_{j=i}^{i+n-1} j \prod_{j=1}^{i-1} j}{\prod_{j=1}^{i-1} j} x^n \\ &= 1 + \sum_{n=1}^{+\infty} \frac{1}{n!} (-1)^n \frac{(n+i-1)!}{(i-1)!} x^n \\ &= 1 + \sum_{n=1}^{+\infty} \binom{n+i-1}{i-1} (-1)^n x^n. \end{aligned}$$

On note (\star) la relation :

$$\forall x \in]-1, 1[, (1+x)^{-i} = 1 + \sum_{n=1}^{+\infty} \binom{n+i-1}{i-1} (-1)^n x^n.$$

Ici, on reconnait donc

$$\sum_{k=0}^{+\infty} \binom{k+i-1}{i-1} (1-p)^k = (1 - (1-p))^{-i} = p^{-i}.$$

On obtient donc

$$\sum_{k=i}^{+\infty} \mathbb{P}(T_i = k) = p^i p^{-i} = 1.$$

3. On rappelle que $G_X(t) = \mathbb{E}[t^X] \underset{\text{transfert}}{=} \sum_{n \in X(\Omega)} t^n \mathbb{P}(X = n)$ pour tout $t \in]-R, R[$ avec $R \geq 1$. Ici, il

existe donc $R \geq 1$ tel que

$$\begin{aligned}
 \forall t \in]-R, R[, G_{T_i}(t) &= \sum_{n=i}^{+\infty} t^n \mathbb{P}(T_i = n) \\
 &= \sum_{n=i}^{+\infty} t^n \binom{n-1}{i-1} p^i (1-p)^{n-i} \\
 &= p^i t^i \sum_{n=i}^{+\infty} \binom{n-1}{i-1} [t(1-p)]^{n-i} \\
 &= (tp)^i \sum_{n=0}^{+\infty} \binom{n+i-1}{i-1} (t(1-p))^n \\
 &\stackrel{*}{=} (tp)^i (1-t(1-p))^{-i} = \left(\frac{tp}{1-t(1-p)} \right)^i
 \end{aligned}$$

la relation (*) étant établie à la question précédente.

4. Soit A_1, \dots, A_i i variables aléatoires indépendantes et identiquement distribuées avec $A_1 \hookrightarrow \mathcal{G}(p)$. Alors l'indépendance de X et Y donne $G_X G_Y = G_{X+Y}$ donc comme $G_{A_1}(t) = \frac{tp}{1-t(1-p)}$ pour $t \in]-1, 1[$, on a

$$G_{T_i}(t) = \left(\frac{tp}{1-t(1-p)} \right)^i = \prod_{j=1}^i G_{A_1}(t) = \prod_{j=1}^i G_{A_j}(t) = G_{\sum_{j=1}^i A_j}(t).$$

Comme la fonction génératrice caractérise la loi, T_i suit la même loi de $\sum_{j=1}^i A_j$.

5. (a) Soit $k \in \tau_i(\Omega) \subset \mathbb{N}_{\geq i-1} \cup \{+\infty\}$. Alors

$$\mathbb{P}(\tau_i = k) = \mathbb{P}(T_i - T_{i-1} = k).$$

On conditionne selon la valeur de T_{i-1} . Comme $(T_{i-1} = j)_{j \geq i-1}$ forme un système complet d'événements, alors

$$\mathbb{P}(T_i - T_{i-1} = k) = \sum_{j=i-1}^{+\infty} \mathbb{P}(T_i - T_{i-1} = k \mid T_{i-1} = j) \mathbb{P}(T_{i-1} = j).$$

Or on a l'égalité des événements

$$(T_i - T_{i-1} = k \mid T_{i-1} = j) = (T_i = j+k \mid T_{i-1} = j) = (X_j = 1) \cap (X_{j+k} = 1) \cap \bigcap_{p=j+1}^{j+k-1} (X_p = 0).$$

Par indépendance, on a donc

$$\mathbb{P}(T_i - T_{i-1} = k \mid T_{i-1} = j) = \mathbb{P}(X_{j+k} = 1) \prod_{p=j+1}^{j+k-1} \mathbb{P}(X_p = 0) = p(1-p)^{k-1}.$$

Ainsi,

$$\mathbb{P}(\tau_i = k) = \sum_{j=i-1}^{+\infty} \mathbb{P}(T_i - T_{i-1} = k \mid T_{i-1} = j) \mathbb{P}(T_{i-1} = j) = p(1-p)^{k-1} \underbrace{\sum_{j=i-1}^{+\infty} \mathbb{P}(T_{i-1} = j)}_{=1} = p(1-p)^{k-1}.$$

Ainsi, τ_i suit une loi géométrique de paramètre p .

(b) Soit $J \subset \mathbb{N}^*$ fini. On note $J = \{j_1, \dots, j_N\}$. Soit $k_j \in \tau_j(\Omega), j \in J$. Alors

$$\bigcap_{i=1}^N \{\tau_{j_i} = k_{j_i}\} = \bigcap_{i=0, i \notin \{k_j: j \in J\}}^K \{X_i = 0\} \cap \bigcap_{i \in k_j: j \in J} \{X_i = 1\}$$

avec $K = \sum_{j \in J} k_j$. Cet événement signifie que l'on a

- un premier succès au temps k_{j_1} ,
- on attend ensuite k_{j_2} étapes pour avoir un deuxième succès,
- ...
- on a donc $\text{card}(J)$ succès aux instants k_{j_1}, \dots, k_{j_N}

Ainsi, par indépendance,

$$\mathbb{P} \left(\bigcap_{j \in J} \tau_j = k_j \right) = \prod_{i=0, i \notin \{k_j: j \in J\}}^K \mathbb{P}(X_i = 0) \prod_{i \in k_j: j \in J} \mathbb{P}(X_i = 1) = (1-p)^{K-\text{card}(J)} p^{\text{card}(J)}.$$

Or, comme les $(\tau_j)_j$ suivent une loi géométrique de paramètres p , on a

$$\prod_{j \in J} \mathbb{P}(\tau_j = k_j) = \prod_{j \in J} p \prod_{j \in J} (1-p)^{k_j-1} = p^{\text{card}(J)} (1-p)^{\sum_{j \in J} k_j - \text{card}(J)} = p^{\text{card}(J)} (1-p)^{K-\text{card}(J)}.$$

Ainsi, les $(\tau_j)_j$ sont bien mutuellement indépendants.

(c) Puisque $\tau_i = T_i - T_{i-1}$, par télescopage,

$$\sum_{j=1}^i \tau_j = \sum_{j=2}^i \tau_j + T_1 = T_i - T_1 + T_1 = T_i$$

ce qui donne le résultat annoncé.