La loi géométrique

Christian CYRILLE

10 novembre 2020

1 La loi géométrique

1.1 Attente d'un premier événement dans un processus sans mémoire

On lance un dé jusqu'à l'obtention du nombre 6. Soit X le nombre de lancers nécessaires à la réalisation de cet événement.
$X < \Omega >= \mathbb{N}^*$ et $\forall k \in \mathbb{N}^* \quad P([X = k]) = \left(\frac{5}{6}\right)^{k-1} \frac{1}{6}$

1.2 Généralisation

> Soit p un nombre réel strictement compris entre 0 et 1.
> On dit qu'une variable aléatoire X suit une loi géométrique $\mathcal{G}(p)$ lorsque :
> $X < \Omega >= \mathbb{N}^*$ et $\forall k \in \mathbb{N}^* \quad P([X = k]) = (1 - p)^{k-1}p$
> Cette loi correspond au temps d'attente du **premier succès** dans une succession infinie d'épreuves de Bernouilli de paramètre p identiques et indépendantes.

Cette loi géométrique s'appelle aussi Loi de Pascal ou Loi Binomiale Négative avec $n = 1$

1.3 Propriétés

1.3.1 P_1

$$\sum_{k=1}^{+\infty} P([X = k]) = \sum_{k=1}^{+\infty} (1 - p)^{k-1}p = p\sum_{k=1}^{+\infty} (1 - p)^{k-1} = p\sum_{k=0}^{+\infty} (1 - p)^{k} = p\frac{1}{1 - (1 - p)} = \frac{p}{p} = 1$$

1.3.2 P_2

> Si X suit une loi géométrique $\mathcal{G}(p)$ alors $E(X) = \frac{1}{p}$ et $Var(X) = \frac{q}{p^2}$

- $E(X) = \sum_{k=1}^{+\infty} kP([X = k]) = \sum_{k=1}^{+\infty} k(1 - p)^{k-1}p = p\sum_{k=1}^{+\infty} k(1 - p)^{k-1} = \frac{p}{1 - p}\sum_{k=1}^{+\infty} k(1 - p)^{k}$
$= \frac{p}{1 - p}\frac{1 - p}{(1 - (1 - p))^2} = \frac{p}{p^2} = \frac{1}{p}$
- $E(X^2) = \sum_{k=1}^{+\infty} k^2P([X = k]) = \sum_{k=1}^{+\infty} k^2(1 - p)^{k-1}p = p\sum_{k=1}^{+\infty} k^2(1 - p)^{k-1} = \frac{p}{1 - p}\sum_{k=1}^{+\infty} k^2(1 - p)^{k}$
$= \frac{p}{1 - p}\frac{(1 - p)(1 + (1 - p))}{(1 - (1 - p))^3} = \frac{p}{1 - p}\frac{(1 - p)(2 - p)}{p^3} = \frac{2 - p}{p^2}$
$Var(X) = E(X^2) - (E(X))^2 = \frac{2 - p}{p^2} - \frac{1}{p^2} = \frac{1 - p}{p^2} = \frac{q}{p^2}$

1.3.3 P_3

> Si X suit une loi géométrique $\mathcal{G}(p)$ alors $\forall n \in \mathbb{N}^* \quad Pr([X > n] = (1-p)^n$

$Pr([X > n] = Pr([X = n+1] \cup [X = n+2] \cup [X = n+3] \cup [X = n+4] \cup \cdots) = Pr\left(\bigcup_{k=n+1}^{+\infty} [X = k]\right)$.

Comme tous ces èvènements sont dispoints deux à deux alors

$Pr([X > n] = \sum_{k=n+1}^{+\infty} Pr([X = k]) = \sum_{k=n+1}^{+\infty} p(1-p)^{k-1} = p \sum_{k=n+1}^{+\infty} (1-p)^{k-1}$.

Or l'on sait que la série $q^n + q^{n+1} + ...$ a pour somme $S = q^n \dfrac{1}{1-q}$ donc

$\sum_{k=n+1}^{+\infty} (1-p)^{k-1} = (1-p)^n \dfrac{1}{1-(1-p))} = \dfrac{(1-p)^n}{p}$ d'où $\boxed{Pr([X > n]) = p\ \dfrac{(1-p)^n}{p} = (1-p)^n}$

1.3.4 P_4

> Si X suit une loi géométrique $\mathcal{G}(p)$ alors X est **une loi sans mémoire** c'est-à-dire que
>
> $$\forall k \in N \quad \forall n \in N \quad Pr_{[X>n]}[X > n+k] = Pr([X > k])$$

$Pr_{[X>n]}[X > n+k] = \dfrac{Pr([X > n+k] \cap [X > n])}{Pr([X > n])} = \dfrac{Pr([X > n+k])}{Pr([X > n])}$ car $[X < n+k] \subset [X > n]$.

On a donc $Pr_{[X>n]}[X > n+k] = \dfrac{(1-p)^{n+k}}{(1-p)^n} = (1-p)^k = Pr([X > k])$

1.3.5 P_5 : Corollaire

> Si X suit une loi géométrique $\mathcal{G}(p)$ et si a et b sont des entiers non nuls tels que $a < b$ alors
>
> $$Pr_{[X>a]}[X > b] = Pr([X > b-a])$$

Exemple *Source : Maths et Tiques*

> On considère que la probabilité pour qu'un couple donne naissance à un enfant gaucher est 12%.
> Sachant que ce couple a déjà un enfant droitier, quelle est la probabilité d'avoir un enfant gaucher à partir du quatrième enfant ?

Soit la variable aléatoire X correspondant au "nombre d'enfants jusqu'à la naissance du premier enfant gaucher".
X suit alors la loi géométrique $\mathcal{G}(0,12)$.
Nous cherchons en fait à déterminer $Pr_{[X>1]}([X > 3])$.
Comme une loi géométrique est aussi une loi sans mémoire alors $Pr_{[X>1]}([X > 3]) = Pr[[X > 3-1]) = Pr([X > 2]) = 1 - Pr([X \leq 2])$.
Par conséquent,

$$Pr_{[X>1]}([X > 3]) = 1 - (Pr([X = 1]) + Pr([X = 2)) = 1 - 0,12(0,88)^0 - 0,12(0,88)^1 = 0,7744 \approx 77\%$$

.

1.4 Exercices

1.4.1 Station service sur autoroute

Ti Sonson et sa femme Man Tine ont loué une caravane pour remonter la fameuse route 66 aux USA.
Pendant que Ti Sonson conduit, Man Tine vaque à ses affaires dans la caravane.
Mais Ti Sonson est distrait : quand il s'arrête chaque fois pour prendre de l'essence avec sa caravane, il y a une chance sur cinq pour qu'il reparte sans sa femme, descendue pour visiter les lieux.
Soit X la variable aléatoire égale au nombre d'étapes que Ti Sonson parcourt avec Man Tine.

1 Etablir la loi de probabilité de X puis déterminer $E(X)$

2 Déterminer la fonction de répartition de X

3 Quel est le nombre maximum d'étapes que peut comporter le voyage pour que Man Tine arrive à destination dans la voiture de Ti Sonson avec une probabilité supérieure à $0,6$?

Corrigé :

1 $X < \Omega >= \mathbb{N}^*$ et pour tout entier non nul k ,
$P((X = k]) =$ P(Ti Sonson a fait k étapes avec Man Tine) $=$ P(Ti Sonson est reparti $(k-1)$ fois avec elle et la perd à la k-ième étape) $= \left(1 - \frac{1}{5}\right)^{k-1} \frac{1}{5} = \left(\frac{4}{5}\right)^{k-1} \frac{1}{5}$
X suit la loi géométrique $\mathcal{G}(\frac{1}{5})$ donc $E(X) = \frac{1}{p} = 5$

2 Pour tout réel x, on a $F_X(x) = P([X \leq x])$. Comme $X < \Omega >= \mathbb{N}^*$ alors F_X est entièrement déterminée par la connaissance des $F_X(k)$ pour tout entier naturel non nul k

$$F_X(k) = P([X \leq k]) = P([X = 1]) + P([X = 2]) + P([X = 3]) + \cdots P([X \leq k]) = \sum_{k=1}^{n} P([X = k])$$

$$F_X(k) = \sum_{i=1}^{k} \left(\frac{4}{5}\right)^{i-1} \frac{1}{5} = \frac{1}{5} \sum_{i=1}^{k} \left(\frac{4}{5}\right)^{i-1} = \frac{1}{5}(\frac{4}{5})^0 \frac{1 - (\frac{4}{5})^k}{1 - \frac{4}{5}} = 1 - \left(\frac{4}{5}\right)^k$$

3 Pr(Man Tine arrive à destination dans la voiture de Ti Sonson) $> 0,6$

$\iff -Pr$(Man Tine arrive à destination dans la voiture de Ti Sonson) $< -0,6$

$\iff 1 - Pr$(Man Tine arrive à destination dans la voiture de Ti Sonson) $< 1 - 0,6$

$\iff Pr$(Man Tine est laissée en chemin) $< 0,4 \iff 1 - \left(\frac{4}{5}\right)^k < 0,4$
$\iff 0,6 < (0,8)^k \iff ln(0,6) < ln((0,8)k)$
$\iff ln(0,6) < k\ ln(0,8) \iff ln(0,6)\ ln(0,8) > k$ car $ln(0,8) < 0$ puisque $0 < 0,8 < 1$.
Or $\frac{ln(0,6)}{ln(0,8)} \approx 2,28$ donc le nombre maximum d'étapes est 2

1.4.2 Le Tricheur à l'as de carreau

> *"Le Tricheur à l'as de carreau"* est un tableau peint par Georges de La Tour vers 1636-1638, conservé au Musée du Louvre, et considéré comme l'un des chefs-d'œuvre du peintre et de la peinture française.

> Dans un jeu de 32 cartes, on remplace l'as de cœur par l'as de carreau.
>
> **1** On tire de ce jeu 4 cartes au hasard. Quelle est la probabilité de déceler la supercherie ?
>
> **2** On procède dans ce jeu à des tirages successifs de 4 cartes avec remise.
> Quel est le nombre minimal n de tirages à effectuer pour déceler la supercherie avec une probabilité d'au moins $0,95$?

Corrigé :

1 On décèle la supercherie lorsque parmi les 4 cartes tirées, il y a les 2 as de carreau.

- On choisit les 2 as de carreau et cela de $\binom{2}{2}$ façons c'est-à-dire 1 façon.
- On choisit les 2 autres cartes parmi les 30 autres cartes et ceci de $\binom{30}{2}$ façons
- Donc le nombre de cas favorables est $\binom{30}{2}$. Comme le nombre de cas possibles est $\binom{32}{4}$

Alors la probabilité de déceler la supercherie $\dfrac{\binom{30}{2}}{\binom{32}{4}} = \dfrac{435}{35960} = \dfrac{32}{48} \approx 0,12$

2 n est le nombre minimal de tirages à effectuer avant de déceler la supercherie. veut dire que :

- soit l'on détecte la supercherie au tirage 1 avec une probabilité p
- soit l'on détecte la supercherie au tirage 2 avec une probabilité $(1-p)p$
- soit l'on détecte la supercherie au tirage 3 avec une probabilité $(1-p)^2p$
- $\cdots$
- soit l'on détecte la supercherie au tirage n avec une probabilité :
$(1-p)^{n-1}p$

donc on doit résoudre l'inéquation :
$p + (1-p)p + (1-p)^2p + \cdots + (1-p)^{n-1}p \geq 0,95$

$\iff p(1 + (1-p) + (1-p)^2 + \cdots + (1-p)^{n-1}) \geq 0,95$

$\iff p(1\dfrac{1-(1-p)^n}{1-(1-p)}) \geq 0,95 \iff 1-(1-p)^n \geq 0,95 \iff 0,05 \geq (1-p)^n$

$\iff ln(0,05) \geq ln((1-p)^n) \iff ln(0,05) \geq n\ ln(1-p)$

$\iff \dfrac{ln(0,05)}{ln(1-p)} \leq n$ car $ln(1-p) \leq 0$ car $0 \leq 1-p \leq 1$ puisque $0 \leq p \leq 1$.

Or $\dfrac{ln(0,05)}{ln(1-p)} = \dfrac{(ln(0,05)}{ln(\dfrac{245}{248})} \approx 246,146$ donc le nombre minimal de tirages n est 247

1.4.3 Edhec 2007

On lance une pièce équilibrée (la probabilité d'obtenir "pile" et celle d'obtenir "face" étant toutes égales à $\frac{1}{2}$) et on note Z la variable aléatoire égale au rang du lancer où l'on obtient le premier "pile".
Après cette série de lancers, si Z a pris la valeur k ($k \in \mathbb{N}^*$), on remplit une urne de k boules numérotées $1, 2, \cdots, k$ puis on extrait au hasard une boule de cette urne.
On note X la variable aléatoire égale au numéro de la boule tirée après la procédure décrite ci-dessus.

1 On décide de coder l'événement "obtenir un pile" par 1 et l'événement "obtenir un face" par 0.
On rappelle que la fonction *random* renvoie , pour un argument k de type integer(où k désigne un entier supérieur ou égal à 1) un entier aléatoire compris entre 0 et $k-1$.

a. Recopier et compléter le programme suivant pour qu'il affiche la valeur prise par Z lors de la première partie de l'expérience décrite ci-dessus :

```
Program edhec2007;
  var z, hasard : integer;
begin
    randomize;
    z:=0;
    repeat
         z := ............;
         hasard := .................;
    until (hasard = 1);
    writeln(z);
end.
```

b. Quelle instruction faut-il rajouter avant la dernière ligne de ce programme pour qu'il simule l'expérience aléatoire décrite dans ce problème et affiche la valeur prise par la valeur aléatoire X ? Vous n'oublierez pas de compléter alors la partie déclarations du programme.

2 Déterminer la loi de Z. Donner alors sans justification son espérance et sa variance.

Corrigé :

1 On décide de coder l'événement "obtenir un pile" par 1 et l'événement "obtenir un face" par 0.
On rappelle que la fonction *random* renvoie , pour un argument k de type integer(où k désigne un entier supérieur ou égal à 1) un entier aléatoire compris entre 0 et $k-1$.

a. Le programme suivant affiche la valeur prise par Z lors de la première partie de l'expérience puis affiche la valeur de X :

```
Program edhec2007;
  var z, hasard : integer;
       x: integer;
begin
    randomize;
    z:=0;
    repeat
         z := z + 1;
         hasard := random(2).;
    until (hasard = 1);
    writeln(z);
    x:= 1 + random(z);
    writeln(x);
end.
```

2 Z suit la loi géometrique de paramètre $p = \dfrac{1}{2}$ et $Z < \Omega >= \mathbb{N}^*$:
$\forall k \in \mathbb{N}^*$ on a $P([Z = k] = (1-p)^{k-1}p = (\dfrac{1}{2})^{k-1}\dfrac{1}{2} = (\dfrac{1}{2})^k$.
Donc $E(Z) = \dfrac{1}{p} = 2$ et $Var(Z) = \dfrac{1-p}{p^2} = 2$.

1.5 Ecricome ECT 04

Soit n un entier naturel supérieur ou égal à 2.
On considère une urne contenant :

- une boule numérotée 1
- deux boules numérotées 2
- trois boules numérotées 3
- $\cdots$
- n boules numérotées n

1 Epreuve 1 :
On tire une boule de cette urne. On note X la variable aléatoire représentant le numéro de la boule obtenue.

a. Déterminer le nombre total de boules dans l'urne puis déterminer l'univers-image $X < \Omega >$

b. Déterminer $P([X = k])$ pour tout $k \in X < \Omega >$ puis vérifier que $P([X = n]) = \dfrac{2}{n+1}$

c. Déterminer $E(X)$

2 Epreuve 2 :
On tire maintenant 10 fois une boule avec remise dans cette urne. On note Y la variable aléatoire représentant le nombre de fois où l'on a obtenu une boule numérotée n.

a. Reconnaître la loi de Y. Justifier.

b. Donner alors les valeurs de $E(Y)$ et de $Var(Y)$

3 Epreuve 3 :
On tire maintenant une infinité de fois une boule avec remise dans cette urne. On note Z la variable aléatoire représentant le numéro du tirage où pour la première fois l'on a obtenu une boule numérotée n.

a. Reconnaître la loi de Z. Justifier.

b. Donner alors les valeurs de $E(Z)$ et de $Var(Z)$