Hugo SALOU
Dept. Informatique

Devoir maison n° 1

13 mars 2025.

Probabilités

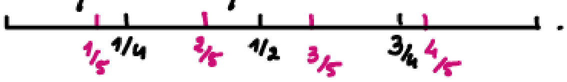
Exercice 1. Bits aléatoires I.

A) Notons (Z_n) la suite de nombres tirés uniformément dans $[\![1, 8]\!]$.
On a :
$$\mathbb{P}(Z_n \leq 5) = \frac{5}{8}$$
d'où, le nombre de paquets de 3 bits utilisés par la méthode A suit une loi géométrique $\mathcal{G}(5/8)$, car les tirages sont indépendants.

D'où, $\mathbb{E}[\text{Nombre de bits utilisés}] = \frac{8}{5} \times 3 = 4{,}8$.

B) Par calcul de la représentation binaire, on sait que :

- $1/5 = 0{,}0011\,\overline{0011}$,
- $2/5 = 0{,}011\,\overline{0011}$,
- $3/5 = 0{,}1\,\overline{0011}$,
- $4/5 = 0{,}11\,\overline{0011}$.

On lit 2 bits pour se placer dans la situation suivante : 

.

Dans chacun des intervalles induits par les 2 premiers bits, on peut être dans deux intervalles de 1/5. Ainsi, à chaque nouveau bit, on se trouve dans une situation 

: soit on est sûr d'être dans un intervalle, soit le doute plane encore. Ceci suit une loi de Bernoulli $\mathcal{B}(1/2)$.

D'où, après les 2 premiers bits, l'espérance du nombre de bits utilisés est 2 car loi géométrique $\mathcal{G}(1/2)$. On conclut :
$$\mathbb{E}[\text{Nombre de bits utilisés}] = 2 + 2 = 4$$

C) Prendre 5 bits de loi $\mathcal{B}(1/2)$ est équivalent à tirer uniformément dans $[\![0, 2^5 - 1]\!]$ car les tirages des bits sont indépendants.
Soit $Y \sim \mathcal{U}([\![0, 2^5-1]\!])$.
$$\begin{aligned}\mathbb{P}(\text{les 5 bits sont identiques}) &= \mathbb{P}(Y = 0 \overset{\text{disjoint}}{\text{ ou }} Y = 2^5 - 1) \\ &= \mathbb{P}(Y=0) + \mathbb{P}(Y = 2^5-1) \\ &= 2/2^5 = 2^{-4}.\end{aligned}$$

Ainsi, l'espérance du nombre de paquets de 5 utilisés est $\frac{1}{1-2^{-4}} = \frac{16}{15}$ car elle suit une loi géométrique $\mathcal{G}(1-2^{-4})$.

De plus, on a

2 bits identiques
$$\mathbb{P}(\text{on tire 1 bit supplémentaire}) = \frac{2 \cdot \binom{5}{2}}{2^5 - 2} = \frac{2}{3}$$
tous les nombre identiques exclu

d'où $\mathbb{E}[\text{nombre de bits utilisés}] = \frac{16}{15}\cdot 5 + \underbrace{\frac{2}{3}}_{\substack{\text{loi de Bernoulli} \\ \text{bits supplémentaires}}} = \frac{18}{3} = 6$

__Mieux__ : on tire 3 bits, pour obtenir $x \in [\![1,8]\!]$. Si $x \in \{1,2,3,4,5\}$, on renvoie x.
Sinon, on tire un bit supplémentaire : on a donc 6 possibilités $\{(6,0),(6,1),(7,0),(7,1),(8,0),(8,1)\}$.
On renvoie la pos° dans l'ordre donné ci-avant, sauf si on a $(6,0)$, où dans ce cas, on recommence. $\mathbb{E}[\text{Nombre de bits utilisés}] = \frac{1}{1-\mathbb{P}(\text{avoir }(6,0))}\times 4 - \mathbb{P}(\text{n'utiliser que 3 bits à la fin})$

$= \frac{16}{15}\times 4 - \frac{10}{15} = 3{,}6$

Exercice 2. Bits aléatoires II

Q1. Notons A, B les éventuels 2 bits générés par l'algorithme (3 cas possibles : $A = B = \Lambda$; $A \in \{0,1\}$ et $B = \Lambda$; et $A, B \in \{0,1\}$). On note $q := 1-p$.

$\mathbb{P}(A = \Lambda = B) = \mathbb{P}(X_1 \cdots X_4 \in \{0000, 1111\}) = (p^4+q^4)/16 \}$ ignoré

$\mathbb{P}(A=0, B=0) = \frac{1}{16}pq(p^2+q^2+pq) = \frac{pq}{16}\left(\underbrace{\cancel{(p+q)}^2}_{1} - pq\right)$ $\quad \mathbb{P}(A=0, B=1) = \frac{1}{16}pq(p^2+q^2+pq) = \frac{pq(1-pq)}{16}$

$\mathbb{P}(A=0, B=\Lambda) = p^2q^2/16$ $\quad \mathbb{P}(A=1, B=0) = \frac{pq}{16}(1-pq)$

$\mathbb{P}(A=1, B=\Lambda) = p^2q^2/16.$ $\quad \mathbb{P}(A=1, B=1) = pq(1-pq)/16$

même probabilité donc non biaisé. $\quad$ même probabilité donc non biaisé

D'où $(y_k)_{k\geq 1}$ est une suite de bits non biaisés. Chaque bit individuel est non-biaisé.

Notons N le nombre de bits produits après la lecture de 4 bits biaisés.

On calcule :

$\mathbb{E}[\text{Nombre de bits produits}] = \frac{1}{4}\cdot\left(\frac{2}{16} + 2\cdot\frac{12}{16}\right) = \frac{26}{4\cdot 16} = \underset{\substack{\| \\ 0{,}40625}}{\frac{13}{32}} \geq \frac{8}{32} = \frac{1}{4} \geq p\cdot(1-p).$

Q2. On lit les bits biaisés 8 par 8 et on produit $(y_k)_{k\geq 1}$ selon la règle suivante :

$0000\ 0000 \longrightarrow \Lambda$
$1111\ 1111 \longrightarrow \Lambda$
$0000\ 1111 \longrightarrow 0000$
$1111\ 0000 \longrightarrow 1111$
$\underbrace{\quad}_{A}\ \underbrace{\quad}_{B} \longrightarrow x_A\, x_B$ où x_A est le résultat de l'algorithme en Q1 sur le mot de 4 bits A (de même avec x_B).

Les bits ne sont pas biaisés car, avec Q1, on sait que le résultat de x_A n'est pas biaisé, et les règles ci-dessus ne "favorisent" pas plus la prod° de 0 que de 1.

$\mathbb{E}[\text{Nombre de bits produits}] = \frac{1}{8}\left(\frac{2}{2^8}\cdot 4\right) + \frac{13}{32} = \frac{105}{256} \simeq 0{,}41$ $\quad$ C'est mieux.

Exercice 3. Liste à sauts aléatoire.

Q1. Soit $n \in \mathbb{N}$. Soit $C \in \mathbb{R}_*^+$ quelconque.

$$\mathbb{P}(M_n \geq C \cdot \log n) = \mathbb{P}(\exists k \leq n,\ X_k \geq C \cdot \log n)$$

par indépendance

$$= 1 - \prod_{k=1}^{n} (1 - \mathbb{P}(X_k \geq C \log n))$$
$$= 1 - \prod_{k=1}^{n} (1 - \mathbb{P}(X_k > C \log n - 1))$$
$$= 1 - \prod_{k=1}^{n} \left(1 - \left(\tfrac{1}{2}\right)^{C \log n - 1}\right)$$
$$= 1 - \prod_{k=1}^{n} \left(1 - 2n^{-C}\right)$$
$$= 1 - (1 - 2n^{-C})^n$$
$$= 1 - 1 + 2n \cdot n^{-C} + o(n \cdot n^{-C})$$

On a :

$$\lim_{n \to \infty} \mathbb{P}(M_n > C \log n) = 0 \iff 2n^{1-C} \xrightarrow[n \to \infty]{} 0$$
$$\iff C > 1.$$

D'où, avec $C = \frac{\pi}{e} > 1$, on a bien

$$\mathbb{P}\left(M_n > \frac{\pi}{e} \log_2 n\right) \xrightarrow[n \to \infty]{} 0$$

Q2. La variable S_k représente le temps d'attente pour k succès.
Soit $B_n \sim \mathcal{B}(n, 1/2)$, où $n := \lfloor (1+\lambda) 2k \rfloor$.

car c'est un temps d'attente pour k succès

On a : $\mathbb{P}(S_k > (1+\lambda)2k) = \mathbb{P}(S_k > n) = \mathbb{P}(B_n < k) \leq \mathbb{P}(B_n \leq k)$.

Or, par Chernoff I, on a : $\mathbb{P}\Big(B_n \leq \underbrace{\tfrac{n}{2}}_{\mu} - \underbrace{(\tfrac{n}{2} - k)}_{\geq 2\lambda k > 0}\Big) \leq \exp\left(-2 \frac{(\frac{n}{2} - k)^2}{n}\right) \leq \exp\left(-2 \frac{(2k)^2}{(1+\lambda) \cdot 2k}\right)$

$$= \exp\left(-\frac{\lambda^2}{1+\lambda} k\right).$$

D'où le résultat demandé.

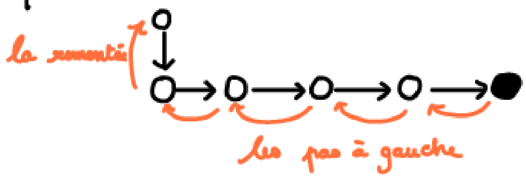
Q3. La "hauteur" de chaque élément x, c'est-à-dire $\max\{j \in \mathbb{N} \mid x \in L_j\}$, suit une loi géométrique $\mathcal{G}(1/2)$: succession de loi de Bernoulli $\mathcal{B}(1/2)$ jusqu'au premier succès (on supprime x du niveau).

Ainsi, M_n décrit le nombre de niveaux d'une liste à n éléments.
Par Q1,

$$\mathbb{P}(M_n > O(\log n)) \xrightarrow[n \to \infty]{} 0.$$

D'où, avec grande probabilité, la hauteur est en $O(\log n)$.

Q4. De manière équivalente, partons de la fin (du noeud x) et remontons jusqu'au noeud $-\infty$ le plus haut. Ceci est un enchaînements de pas à gauche puis une remontée.



Le nombre de pas à gauche et la remontée finale suit une loi géométrique $\mathcal{G}(1/2)$ car c'est le temps d'attente avant un succès (on a un noeud au dessus) et $\mathbb{P}(\text{un noeud est au dessus}) = \frac{1}{2}$.

D'où le nombre de pas total est une somme de lois géométriques $\mathcal{G}(1/2)$ qu'on note S_h où h est la hauteur. On utilise la question 2 :

$$\mathbb{P}(S_h > 2h) \leq \exp\left(-\frac{1}{2} h\right)$$

et par la question précédente, $h = O(\log n)$ avec grande probabilité.
D'où,

$$\mathbb{P}(S_h > 2h) \xrightarrow[n \to \infty]{} 0 \qquad \text{car } \exp\left(-\frac{1}{2}\log n\right) = \frac{\log_2 e}{\sqrt{n}} \longrightarrow 0$$

On en conclut qu'avec très grande probabilité le parcours se fait en $O(\log n)$.
En effet :

$$\text{Nombre d'opérations} = \text{Nombre de pas} + \underbrace{\text{Nombre de comparaisons}}_{O(\text{nombre de pas})}$$

Fin du DM.