CORRECTION DM 1 - PROBABILITÉS DISCRÈTES

Exercice 1 - EDHEC ECS 2011

1. (a) T_1 est le rang du premier succès de « tirer deux boules avec le même numéro ».
Pour un unique tirage, notons A_k l'événement : « le numéro de la première boule est k ». Et notons B_k l'événement : « le numéro de la seconde boule est k ». Si on note le succès S, on a :

$$\begin{aligned} P(S) &= P\left(\bigcup_{k=1}^{n}(A_k \cap B_k)\right) \\ &= \sum_{k=1}^{n} P(A_k)P_{A_k}(B_k) \\ &\quad \text{(formule des probabilités totales avec } (A_k)_k \text{ comme s.c.e.)} \\ &= \sum_{k=1}^{n} \frac{1}{n} \times \frac{1}{2n-1} \boxed{= \frac{1}{2n-1}.} \end{aligned}$$

Donc T_1 suit une loi géométrique de paramètre $\frac{1}{2n-1}$. Formellement :

$$\boxed{T_1 \hookrightarrow \mathcal{G}\left(\frac{1}{2n-1}\right).}$$

On a donc :

$$\boxed{\forall k \in \mathbb{N}^\star, P(T_1 = k) = \left(1 - \frac{1}{2n-1}\right)^{k-1} \times \frac{1}{2n-1}.}$$

(b) On a donc :

$$\boxed{E(T_1) = \frac{1}{\frac{1}{2n-1}} = 2n-1.}$$

2. (a) `rd.randint(2*n)` simule le tirage d'une variable aléatoire suivant la loi $\mathcal{U}([\![0, 2n-1]\!])$. Il y a donc $2n$ valeurs différentes correspondant au nombre de boules dans l'urne. Ainsi les boules sont représentées par un numéro de 0 à $2n-1$. Le but de la boucle est de simuler le second tirage qui ne suit pas une remise. Il faut donc que `b` soit différent de `a`. Et donc on boucle tant que c'est égal (ce qui explique l'initialisation de `b` à la valeur de `a`).

```python
def tirage(n):
  a = rd.randint(2*n)
  b = a
  while b == a:
    b = rd.randint(2*n)
  return a,b
```

(b) Le code utilise l'instruction `%` qui calcule le reste de la division euclidienne. Ainsi `n1` et `n2` sont des entiers entre 0 et $n-1$. À un décalage prêt, cela représente le numéro sur la boule. On tire donc tant que les numéros sont différents.

```python
def T1(n):
  n1 = 0
  n2 = 1
  t = 0
  while n1 != n2:
    a,b = tirage(n)
    n1 = a % n
```

```
        n2 = b % n
        t = t + 1
    return t
```

(c) Le sujet attendant visiblement la condition `a != b`. Dans ce cas, `a` représente le numéro sur la première boule et `b` le numéro sur la seconde boule et on regarde quand on obtient la première paire.

Les valeurs possibles pour `a` et `b` sont bonnes. En revanche, leurs probabilités sont mauvaises. Lors du tirage de la première boule, le numéro suit bien une loi uniforme. Mais puisqu'il n'y a pas de remise avant la seconde boule, le numéro de la seconde boule ne suit pas une telle loi. En effet, le numéro de la première boule est légèrement moins probable. C'est ça l'erreur de l'énoncé.

3. (a) $\boxed{X_i \text{ est le nombre de tirages supplémentaires nécessaires pour constituer la } i^{\text{ème}} \text{ paire.}}$

(b) Soit $i \in [\![1, n]\!]$.

- Si $i = 1$, alors $X_i = X_1 = T_1$ qui $\boxed{\text{suit une loi géométrique de paramètre } \frac{1}{2n-1}}$ et qui $\boxed{\text{a pour espérance } 2n-1.}$
- Si $i \neq 1$, alors $X_i = T_i - T_{i-1}$. On peut raisonner comme dans la première question : X_i est le rang du premier succès (à partir de la $i^{\text{ème}}$ paire constituée) avec une probabilité de succès $\frac{1}{2(n-(i-1))-1}$ (raisonnement identique à la question 1 mais seul $n-(i-1)$ numéros sont encore dans l'urne). Donc :

$$\boxed{X_i \hookrightarrow \mathcal{G}\left(\frac{1}{2n-2i+1}\right) \quad \text{et} \quad E(X_i) = 2n-2i+1.}$$

On remarque que la formule donnée dans le second cas s'applique également au premier cas.

(c) On a :

$$\begin{aligned} T_n &= (T_n - T_{n-1}) + (T_{n-1} - T_{n-2}) + \cdots + (T_2 - T_1) + T_1 \\ &= \sum_{k=2}^{n} (T_k - T_{k-1}) + T_1. \end{aligned}$$

Par linéarité de l'espérance, $\boxed{T_n \text{ admet une espérance}}$ et :

$$\boxed{E(T_n) = \sum_{k=2}^{n} (2n-2i+1) + 2n-1 = (2n+1)(n-1) - 2\frac{(n-1)(n+2)}{2} + 2n-1 = n^2.}$$

4. (a) On constituent 0 paires au bout de n tirages si et seulement si il faut strictement plus de n tirages pour obtenir la première paire. Formellement :

$$\boxed{[S_n = 0] = [T_1 > n].}$$

Donc :

$$\begin{aligned} P(S_n = 0) &= P(T_1 > n) = P(T_1 \geqslant n+1) \\ &= \sum_{k=n+1}^{+\infty} P(T_1 = k) \\ &= \sum_{k=n+1}^{+\infty} \left(1 - \frac{1}{2n-1}\right)^{k-1} \times \frac{1}{2n-1} \\ &= \frac{1}{2n-1} \times \left(1 - \frac{1}{2n-1}\right)^{n} \times \frac{1}{1 - \left(1 - \frac{1}{2n-1}\right)} \\ &= \left(1 - \frac{1}{2n-1}\right)^{n} \boxed{= \left(\frac{2n-2}{2n-1}\right)^{n}.} \end{aligned}$$

(b) On a :
$$P(S_n=0)=\left(\frac{2n-2}{2n-1}\right)^n=\exp\left(n\ln\frac{2n-2}{2n-1}\right).$$
Comme $\frac{2n-2}{2n-1}\xrightarrow[n\to+\infty]{}1$, on a : $\ln\frac{2n-2}{2n-1}\underset{n\to+\infty}{\sim}\frac{2n-2}{2n-1}-1\underset{n\to+\infty}{\sim}-\frac{1}{2n-1}\underset{n\to+\infty}{\sim}-\frac{1}{2n}$.
Ainsi :
$$\boxed{P(S_n=0)=\exp\left(n\ln\frac{2n-2}{2n-1}\right)\xrightarrow[n\to+\infty]{}\mathrm{e}^{-\frac{1}{2}}.}$$

(c) Notons A_k l'événement : « une paire a été constituée au $k^{\text{ème}}$ tirage ». On a alors :
$$[S_n=n]=A_1\cap A_2\cap\cdots\cap A_n.$$
D'après la formule des probabilités composées, on a :
$$P(S_n=n)=\underbrace{P(A_1)}_{=\frac{1}{2n-1}}\times\underbrace{P_{A_1}(A_2)}_{=\frac{1}{2(n-1)-1}}\times\underbrace{P_{A_1\cap A_2}(A_3)}_{=\frac{1}{2(n-2)-1}}\times\cdots\times\underbrace{P_{A_1\cap\cdots\cap A_{n-1}}(A_n)}_{=\frac{1}{2(n-(n-1))-1}}=\prod_{k=0}^{n-1}\frac{1}{2(n-k)-1}.$$
Commençons par le changement d'indice : $i=n-k$. On obtient :
$$P(S_n=n)=\prod_{i=1}^{n}\frac{1}{2i-1}.$$
Le produit ne contient que des termes impaires. Rajoutons les termes paires et compensons-les au numérateur :
$$\boxed{P(S_n=n)=\prod_{i=1}^{n}\frac{2i}{(2i-1)(2i)}=\left(\prod_{i=1}^{n}2i\right)\left(\prod_{k=1}^{2n}\frac{1}{k}\right)=\frac{2^n n!}{(2n)!}.}$$

5. Dans cette fonction, le nombre d'itérations de la boucle est fixé (à n). On simule donc n tirages successifs. La variable m est celle passée à `tirage` et diminue à chaque fois que `n1` et `n2` sont égaux. Elle représente donc le nombre de numéros distincts restant dans l'urne. La variable `z` part de 0 et augmente quand `n1` et `n2` sont égaux. Il s'agit donc du nombre de paires constituées. C'est aussi la valeur renvoyée. Donc la fonction renvoie une simulation du nombre de paires constitues sur n tirages. $\boxed{\text{Il s'agit donc d'une simulation de } S_n.}$