Énoncé

On désigne par n un entier naturel supérieur ou égale à 2. On dispose de n urnes, numérotées de 1 à n, contenant chacune n boules toutes identiques. On répète n épreuves, chacune consistant à choisir une urne au hasard et à en extraire une boule au hasard. On suppose que les choix des urnes sont indépendants les uns des autres.
Pour tout $i \in [\![1;n]\!]$, on note X_i la variable aléatoire prenant la valeur 1 si l'urne numérotée i contient toujours n boules à l'issue des n épreuves ; et qui prend la valeur 0 sinon.
Pour tout i et tout k de $[\![1;n]\!]$, on note $U_{i,k}$ l'évènement "l'urne numérotée i est choisie à la k-ième épreuve".

1. 1.a. Soit $i \in [\![1;n]\!]$. Écrire l'évènement $[X_i = 1]$ à l'aide de certains des évènements $U_{i,k}$, puis montrer que :
$$\mathbb{P}\big([X_i = 1]\big) = \left(1 - \frac{1}{n}\right)^n$$
   1.b. Soient i et j deux éléments distincts de $[\![1;n]\!]$. Démontrer : $\mathbb{P}\big([X_i = 1] \cap [X_j = 1]\big) = \left(1 - \frac{2}{n}\right)^n$.
   1.c. Comparer $\left(1 - \frac{2}{n}\right)$ et $\left(1 - \frac{1}{n}\right)^2$ et en déduire que, si i et j sont deux entiers distincts de $[\![1;n]\!]$, alors les variables aléatoires X_i et X_k ne sont pas indépendantes.
2. On pose $Y_n = \displaystyle\sum_{i=1}^{n} X_k$.
   2.a. Déterminer l'espérance de Y_n.
   2.b. En déduire $\displaystyle\lim_{n \to +\infty} \frac{\mathbb{E}(Y_n)}{n}$ et donner un équivalent de $\mathbb{E}(Y_n)$ lorsque n est au voisinage de $+\infty$.
3. Pour tout $i \in [\![1;n]\!]$, on note N_i la variable aléatoire égale au nombre de boules manquantes dans l'urne numérotée i à l'issue des n épreuves.
   3.a. Donner la loi de N_i ainsi que son espérance.
   3.b. Que vaut le produit $N_i X_i$?
   3.c. Les variables aléatoires N_i et X_i sont-elles indépendantes ?
4. Recopier et compléter les lignes manquantes du programme ci-dessous de sorte que l'exécution de **simul(n)** simule l'expérience décrite au début de cet exercice et renvoie deux listes, la première contenant les réalisations des variables aléatoires $X_1, \ldots, X_n$; et la seconde les réalisations des variables aléatoires $N_1, \ldots, N_n$.

```
import numpy.random as rd

def simul(n):
    X=......
    N=......
    for k in range(n):
        i =......
        X[i-1]=......
        N[i-1]=......
    return X,N
```

Corrigé

On désigne par n un entier naturel supérieur ou égale à 2. On dispose de n urnes, numérotées de 1 à n, contenant chacune n boules toutes identiques. On répète n épreuves, chacune consistant à choisir une urne au hasard et à en extraire une boule au hasard. On suppose que les choix des urnes sont indépendants les uns des autres.
Pour tout $i \in [\![1;n]\!]$, on note X_i la variable aléatoire prenant la valeur 1 si l'urne numérotée i contient toujours n boules à l'issue des n épreuves ; et qui prend la valeur 0 sinon.
Pour tout i et tout k de $[\![1;n]\!]$, on note $U_{i,k}$ l'évènement "l'urne numérotée i est choisie à la k-ième épreuve".

1. 1.a. Soit $i \in [\![1;n]\!]$. Écrire l'évènement $[X_i = 1]$ à l'aide de certains des évènements $U_{i,k}$, puis montrer que :

$$\mathbb{P}\big([X_i = 1]\big) = \left(1 - \frac{1}{n}\right)^n$$

$[X_i = 1]$ est réalisé si, et seulement si, l'urne i contient toujours n boules à l'issue des n épreuves
si, et seulement si, aucune boule n'a été piochée dans l'urne i au cours des n épreuves
si, et seulement si, des épreuves 1 à n, toutes les boules ont été piochées dans d'autres urnes que l'urne i

Important !
Cette étape permet de reformuler à volonté ce que traduit l'évènement $[X_i = 1]$ afin de pouvoir l'écrire ensuite comme demandé.

D'où :

$$[X_i = 1] = \bigcap_{k=1}^{n} \overline{U_{i,k}}$$

Par conséquent :

$$\begin{aligned}
\mathbb{P}\big([X_i = 1]\big) &= \mathbb{P}\left(\bigcap_{k=1}^{n} \overline{U_{i,k}}\right) && \text{les choix des urnes sont indépendants les uns des autres, donc les évènements } \overline{U_{i,1}}, \ldots \overline{U_{i,n}} \text{ sont indépendants}\\
&= \prod_{k=1}^{n} \mathbb{P}\left(\overline{U_{i,k}}\right) && \text{équiprobabilité du choix de l'urne}\\
&= \prod_{k=1}^{n} \frac{n-1}{n}\\
&= \left(\frac{n-1}{n}\right)^n
\end{aligned}$$

Conclusion : $\mathbb{P}\big([X_i = 1]\big) = \left(1 - \frac{1}{n}\right)^n$.

1.b. Soient i et j deux éléments distincts de $[\![1;n]\!]$. Démontrer : $\mathbb{P}\big([X_i = 1] \cap [X_j = 1]\big) = \left(1 - \frac{2}{n}\right)^n$.

$[X_i = 1] \cap [X_j = 1]$ est réalisé si, et seulement si, les urnes i et j contiennent toujours n boules à l'issue des n épreuves
si, et seulement si, aucune boule n'a été piochée dans les urnes i et j au cours des n épreuves
si, et seulement si, des épreuves 1 à n, toutes les boules ont été piochées dans d'autres urnes que les urnes i et j

D'où :

$$[X_i = 1] \cap [X_j = 1] = \bigcap_{k=1}^{n} \left(\overline{U_{i,k}} \cap \overline{U_{j,k}}\right)$$

Par conséquent :

$$\begin{aligned}
\mathbb{P}\big([X_i = 1] \cap [X_j = 1]\big) &= \mathbb{P}\left(\bigcap_{k=1}^{n} \left(\overline{U_{i,k}} \cap \overline{U_{j,k}}\right)\right) && \text{les choix des urnes sont indépendants les uns des autres}\\
&= \prod_{k=1}^{n} \mathbb{P}\left(\overline{U_{i,k}} \cap \overline{U_{j,k}}\right) && \text{équiprobabilité du choix de l'urne}\\
&= \prod_{k=1}^{n} \frac{n-2}{n}\\
&= \left(\frac{n-2}{n}\right)^n
\end{aligned}$$

Petite remarque
On pourrait détailler le calcul de $\mathbb{P}\left(\overline{U_{i,k}} \cap \overline{U_{j,k}}\right)$ en passant par l'évènement contraire et en utilisant le fait que $U_{i,k}$ et $U_{j,k}$ sont incompatibles...

Conclusion : si i et j sont deux éléments distincts de $[\![1;n]\!]$, alors $\mathbb{P}\big([X_i = 1] \cap [X_j = 1]\big) = \left(1 - \frac{2}{n}\right)^n$.

1.c. Comparer $\left(1-\frac{2}{n}\right)$ et $\left(1-\frac{1}{n}\right)^2$ et en déduire que, si i et j sont deux entiers distincts de $[\![1;n]\!]$, alors les variables aléatoires X_i et X_k ne sont pas indépendantes.

➠Réflexe ! Pour comparer deux réels, on étudie le signe de leur différence.

- On a :

$$\left(1-\frac{2}{n}\right)-\left(1-\frac{1}{n}\right)^2 = \left(1-\frac{2}{n}\right)-\left(1-\frac{2}{n}+\frac{1}{n^2}\right) = \frac{-1}{n^2} < 0$$

Par conséquent :

$$\left(1-\frac{2}{n}\right) < \left(1-\frac{1}{n}\right)^2$$

- Soient $i, j \in [\![1;n]\!]$. Supposons que $i \neq j$.
  - ⋆ D'après la question précédente :

$$\mathbb{P}([X_i=1]\cap[X_j=1]) = \left(1-\frac{2}{n}\right)^n$$

  - ⋆ Or X_i et X_j ont même loi, donc d'après la question **1.a** :

$$\mathbb{P}\big([X_i=1]\big)\mathbb{P}\big([X_j=1]\big) = \left(1-\frac{1}{n}\right)^n\left(1-\frac{1}{n}\right)^n = \left(1-\frac{1}{n}\right)^{2n}$$

Or, d'après ce qui précède :

$$\left(1-\frac{2}{n}\right) < \left(1-\frac{1}{n}\right)^2$$

Donc, par stricte croissance de la fonction $.^n$ sur $\mathbb{R}^+$:

$$\left(1-\frac{2}{n}\right)^n < \left(1-\frac{1}{n}\right)^{2n}$$

D'où :

$$\mathbb{P}\big([X_i=1]\cap[X_j=1]\big) \neq \mathbb{P}\big([X_i=1]\big)\mathbb{P}\big([X_j=1]\big)$$

Conclusion : les variables aléatoires X_i et X_j ne sont pas indépendantes.

2. On pose $Y_n = \sum_{i=1}^{n} X_k$.

2.a. Déterminer l'espérance de Y_n.

Les variables aléatoires $X_1, \dots, X_n$ admettent toutes une espérance (elles suivent une loi de Bernoulli), donc Y_n admet une espérance et :

$$\mathbb{E}(Y_n) = \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\mathbb{E}(X_i) = n\left(1-\frac{1}{n}\right)^n$$

linéarité de l'espérance

pour tout $i \in [\![1;n]\!]$, X_i suit la loi de Bernoulli de paramètre $\mathbb{P}\big([X_i=1]\big)$ qui vaut $\left(1-\frac{1}{n}\right)^n$

Conclusion : $\mathbb{E}(Y_n) = n\left(1-\frac{1}{n}\right)^n$.

2.b. En déduire $\lim_{n\to+\infty}\frac{\mathbb{E}(Y_n)}{n}$ et donner un équivalent de $\mathbb{E}(Y_n)$ lorsque n est au voisinage de $+\infty$.

- Pour n suffisamment proche de $+\infty$, on a ainsi, d'après la question précédente :

$$\frac{\mathbb{E}(Y_n)}{n} = \left(1-\frac{1}{n}\right)^n = \exp\left(n\ln\left(1-\frac{1}{n}\right)\right)$$

licite car pour n suffisamment proche de $+\infty$, $1-\frac{1}{n} > 0$

Or :

✓ $\forall n \in \mathbb{N}^*, \ \frac{-1}{n} \neq 0$

✓ $\lim_{n\to+\infty}\frac{-1}{n} = 0$

D'où :

$$\ln\left(1-\frac{1}{n}\right) \underset{n\to+\infty}{\sim} \frac{-1}{n}$$

Et ainsi :

$$n\ln\left(1-\frac{1}{n}\right) \underset{n\to+\infty}{\sim} -1$$

Par conséquent $\lim\limits_{n\to+\infty} n\ln\left(1-\frac{1}{n}\right) = -1$ et donc, par composition de limites :

$$\lim_{n\to+\infty} \exp\left(n\ln\left(1-\frac{1}{n}\right)\right) = \mathrm{e}^{-1}$$

Conclusion : $\lim\limits_{n\to+\infty} \dfrac{\mathbb{E}(Y_n)}{n} = \mathrm{e}^{-1}$.

- On en déduit :

$$\lim_{n\to+\infty} \frac{\mathbb{E}(Y_n)}{n\mathrm{e}^{-1}} = 1$$

Conclusion : $\mathbb{E}(Y_n) \underset{n\to+\infty}{\sim} \dfrac{n}{\mathrm{e}}$.

3. Pour tout $i \in [\![1;n]\!]$, on note N_i la variable aléatoire égale au nombre de boules manquantes dans l'urne numérotée i à l'issue des n épreuves.

3.a. Donner la loi de N_i ainsi que son espérance.

- **Expérience.** L'expérience s'assimile à n répétitions indépendantes de la même épreuve de Bernoulli dont le succès "piocher une boule dans l'urne i" est de probabilité $\frac{1}{n}$ (par équiprobabilité du choix de l'urne).
- **Variable aléatoire.** La variable aléatoire N_i prend comme valeur le nombre de boules manquantes dans l'urne i ; autrement dit, elle prend comme valeur le nombre de succès de l'expérience ainsi décrite.

Conclusion : $N_i \hookrightarrow \mathscr{B}\left(n;\frac{1}{n}\right)$ et donc $\mathbb{E}(N_i) = 1$.

3.b. Que vaut le produit N_iX_i ?

Soit $\omega \in \Omega$. On a déjà :

$$(N_iX_i)(\omega) = N_i(\omega) \times X_i(\omega)$$

Distinguons deux cas :

- si $\omega \in [X_i = 0]$:
  Dans ce cas, $X_i(\omega) = 0$, et donc
  $$N_i(\omega)X_i(\omega) = 0$$
- si $\omega \notin [X_i = 0]$:
  Alors, puisque $X_i(\Omega) = \{0;1\}$, on a nécessairement $X_i(\omega) = 1$.
  Par conséquent, l'urne i contient toujours n boules à l'issue des n épreuves. D'où : $N_i(\omega) = 0$. Et ainsi :
  $$N_i(\omega)X_i(\omega) = 0$$

Dans tous les cas :

$$N_i(\omega)X_i(\omega) = 0$$

Conclusion : N_iX_i est constante égale à 0.

3.c. Les variables aléatoires N_i et X_i sont-elles indépendantes ?

D'après la question précédente, la variable aléatoire N_iX_i est constante égale à 0, donc :

$$\mathbb{E}(N_iX_i) = 0$$

Or :

$$\mathbb{E}(N_i) \times \mathbb{E}(X_i) = 1 \times \left(1-\frac{1}{n}\right)^n \neq 0$$

Par conséquent :

$$\mathbb{E}(N_iX_i) \neq \mathbb{E}(N_i)\mathbb{E}(X_i)$$

Conclusion : les variables aléatoires N_i et X_i ne sont pas indépendantes.

☞ **Rappel...**
Si X et Y sont deux variables aléatoires indépendantes et admettent une espérance, alors la variable aléatoire XY admet une espérance et $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. On utilise ici la contraposée...

4. Recopier et compléter les lignes manquantes du programme ci-dessous de sorte que l'exécution de **`simul(n)`** simule l'expérience décrite au début de cet exercice et renvoie deux listes, la première contenant les réalisations des variables aléatoires $X_1, \dots, X_n$; et la seconde les réalisations des variables aléatoires $N_1, \dots, N_n$.

```
import numpy.random as rd

def simul(n):
    X = ......
    N = ......
    for k in range(n):
        i = ......
        X[i − 1] = ......
        N[i − 1] = ......
    return X,N
```

Voici :

```
1  import numpy.random as rd
2  
3  def simul(n):
4      X=[1 for i in range(n)]
5      N=[0 for i in range(n)]
6      for k in range(n):
7          i=rd.randint(1,n+1)
8          X[i-1]=0
9          N[i-1]=N[i-1]+1
10     return X,N
```

Explications :

- On initialise `X` avec une liste contenant n fois la valeur 1 : le i-ème élément de la liste `X` prendra la valeur 0 si au moins une boule est extraite de l'urne i (donc fournira une réalisation de X_i).
- On initialise `N` avec une liste contenant n fois la valeur 0 : le i-ème élément de la liste `N` représentera le nombre de boules manquantes dans l'urne i (donc fournira une réalisation de N_i).
- En ligne 7, on choisit un entier au hasard entre 1 et n : on choisit donc une urne dans laquelle on pioche une boule...
- Ayant choisi l'urne i, on enlève une boule dans cette urne (X_i prend donc la valeur 0 et N_i est augmenté de 1).

✘ Attention !

Les urnes sont numérotées de 1 à n alors que l'indexation des listes `X` et `N` débute à 0...