Dans les questions faisant intervenir des instructions en langage Python, on prendra soin d'importer les bibliothèques nécessaires lors de leur première utilisation.

Exercice 1 (d'après Ecricome)

Soit n un entier naturel non nul.
Une urne contient n boules indiscernables au toucher et numérotées de 1 à n. On tire une boule au hasard dans l'urne. Si cette boule tirée porte le numéro k, on place alors dans une seconde urne toutes les boules suivantes : une boule numérotée 1, deux boules numérotées 2, et plus généralement pour tout $j \in [\![1,k]\!]$, j boules numérotées j, jusqu'à k boules numérotées k. Les boules de cette deuxième urne sont aussi indiscernables au toucher. On effectue alors un tirage au hasard d'une boule dans cette seconde urne.
Et on note X la variable aléatoire égale au numéro de la première boule tirée et on note Y la variable aléatoire égale au numéro de la deuxième boule tirée.

1. Reconnaître la loi de X et donner son espérance et sa variance.

2. Déterminer $Y(\Omega)$.

3. Soit $k \in [\![1,n]\!]$.

a. On suppose que l'événement $[X=k]$ est réalisé.
Déterminer, en fonction de k, le nombre total de boules présentes dans la seconde urne.

b. Pour tout entier j de $[\![1,n]\!]$, exprimer $P_{[X=k]}(Y=j)$ en fonction de k et j.
On distinguera les cas $j \leq k$ et $j \geq k+1$.

4. **a.** Déterminer deux réels a et b tels que, pour tout entier naturel k non nul,

$$\frac{1}{k(k+1)} = \frac{a}{k} + \frac{b}{k+1}.$$

b. En déduire que, pour tout élément j de $Y(\Omega)$,

$$P(Y=j) = \frac{2(n+1-j)}{n(n+1)}.$$

5. Justifier que Y admet une espérance et montrer que $E(Y) = \dfrac{n+2}{3}$.

6. Les variables X et Y sont-elles indépendantes ?

7. **a.** Montrer que $E(XY) = \dfrac{(n+1)(4n+5)}{18}$.

b. En déduire que $\mathrm{Cov}(X,Y) = \dfrac{n^2-1}{18}$.

8. **a.** Écrire une fonction en langage Python, nommée `seconde_urne`, prenant en entrée un entier naturel `k` non nul, et renvoyant une liste contenant 1 élément valant 1, 2 éléments valant 2, ... , j éléments valant j, ... , jusqu'à k éléments valant k.
Par exemple, l'appel de `seconde_urne(4)` renverra `[1,2,2,3,3,3,4,4,4,4]`.

b. Recopier et compléter la fonction en langage Python suivante pour qu'elle prenne en entrée un entier naturel n non nul, et qu'elle renvoie une réalisation du couple de variables aléatoires (X,Y).

```
import numpy.random as rd

def simul_XY(n):
    X = _____
    urne2 = seconde_urne(_____)
    nb = len(urne2)
    i = rd.randint(0, nb)
    Y = _____
    return X,Y
```

c. On considère la fonction en langage Python suivante, prenant en entrée un entier naturel `n` non nul.

```
def fonction(n):
    liste = [0]*n
    for i in range(10000):
        j = simul_XY(n)[1]
        liste[j-1] = liste[j-1] + 1/10000
    return liste
```

Quelles valeurs les éléments de la liste renvoyée permettent-ils d'estimer ?

9. Dans toute cette question, on suppose $n = 20$. On simule 50 réalisations du couple de variables aléatoires (X, Y) à l'aide de la fonction `simul_XY` définie à la question **8b**. On représente alors les valeurs obtenues sous forme d'un nuage de points, où les valeurs des réalisations de X sont représentées en abscisse et les valeurs des réalisations de Y en ordonnées. On trace également, sur la même figure, la droite de régression linéaire associée à ce nuage de points.

a. Déterminer par un calcul une valeur approchée des coordonnées du point moyen du nuage de points. Quel théorème de probabilités permet de justifier cette approximation ?

b. Parmi les figures représentées ci-dessous, en justifiant soigneusement votre réponse, indiquer celle qui correspond au nuage de points et à la droite de régression linéaire étudiés.

Figure 1

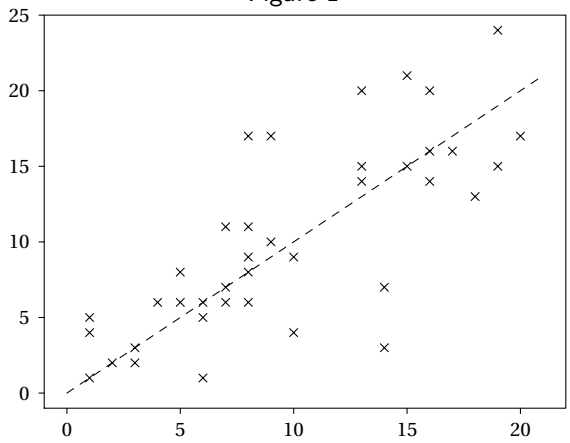

Figure 2

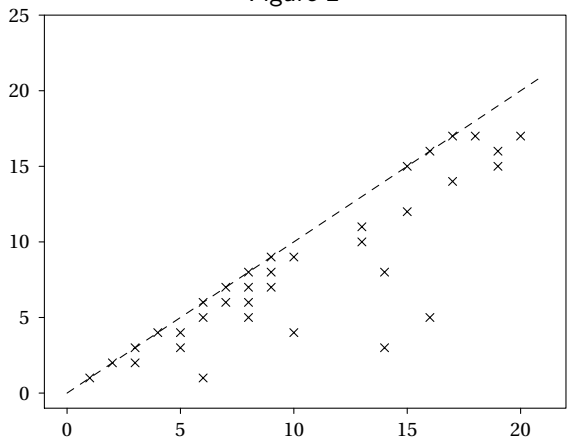

Figure 3

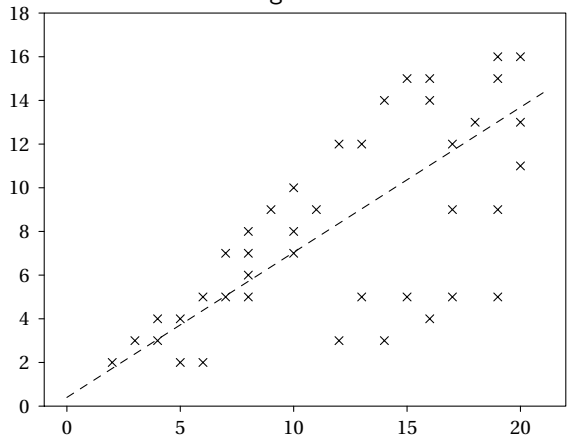
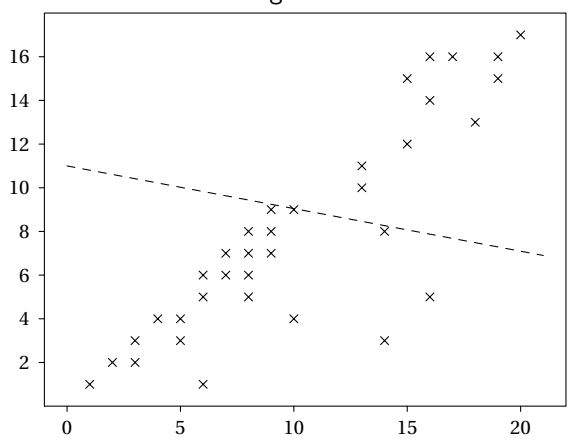
Figure 4