DS 03 - Statistiques, informatiques et relation de négligeabilité

▷ **Exercice 1 :**

1. Énoncer la définition de la covariance.
2. Énoncer et démontrer le théorème de Koenig Huygens pour la covariance.

Voir le cours !

▷ **Exercice 2 :** On considère la série statistique

x_i	-1	0	1	2	3
y_i	0.5	1	2.5	4	4

1. Tracer le nuage de points.
2. Calculer la covariance de la série statistique double.
3. Recopier et compléter les lignes de codes permettant de calculer le coefficient de corrélation linéaire (Uniquement la commande permettant d'appeler la racine carré est autorisé dans la bibliothèque numpy.)

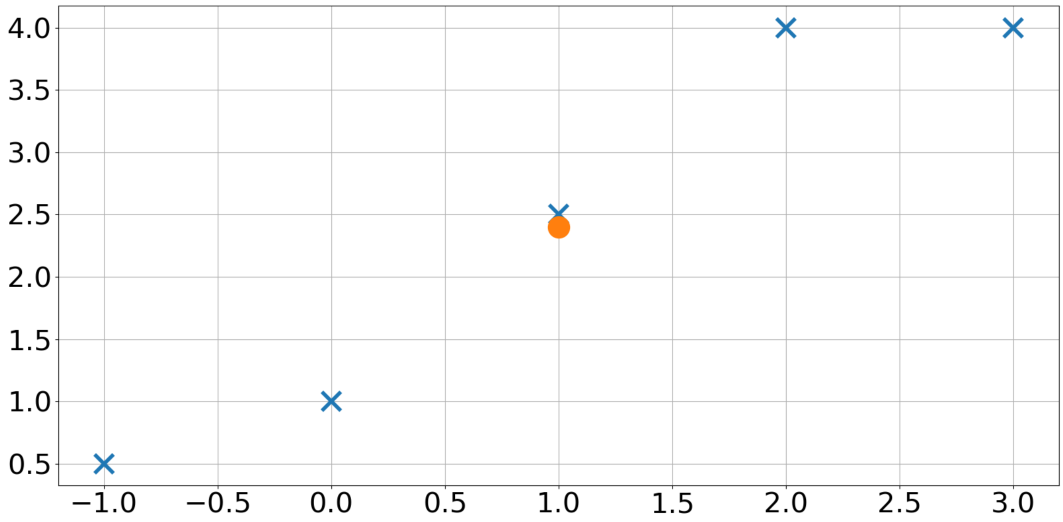
```
import numpy as np
def EcartType(X) :
        Moy = ······
        S = 0
        for k in ······ :
                S = S + ······
        return ······

def Covariance(X,Y) :
        S = 0
        for k in ······ :
                S = S + ······
        return ············

X = ······
Y = ······
······
```

4. On trouve à l'aide de python $s_x^2 = 2$. En déduire l'expression de la droite de corrélation linéaire de $(y_i)_i$ en $(x_i)_i$ sous la forme $y = ax + b$.
5. Tracer sur votre nuage de points la droite de corrélation linéaire.

1. Nous avons besoin de calculer le point moyen avant de tracer le nuage de points. On a $\overline{x} = \dfrac{-1+0+1+2+3}{5} = \dfrac{5}{5} = 1$ et $\overline{y} = \dfrac{0.5+1+2.5+4+4}{5} = \dfrac{12}{5} = 2.4$. On en déduit le graphique suivant :



2. Par la formule de Koenig Huygens, on a :

$$\begin{aligned} s_{x,y} &= \frac{1}{5}(-1\times 0.5 + 0\times 1 + 1\times 2.5 + 2\times 4 + 3\times 4) - 1\times 2.4 \\ &= \frac{1}{5}(-0.5 + 0 + 2.5 + 8 + 12) - 2.4 = \frac{22}{5} - 2.4 \\ &= 4.4 - 2.4 = 2 \end{aligned}$$

3.

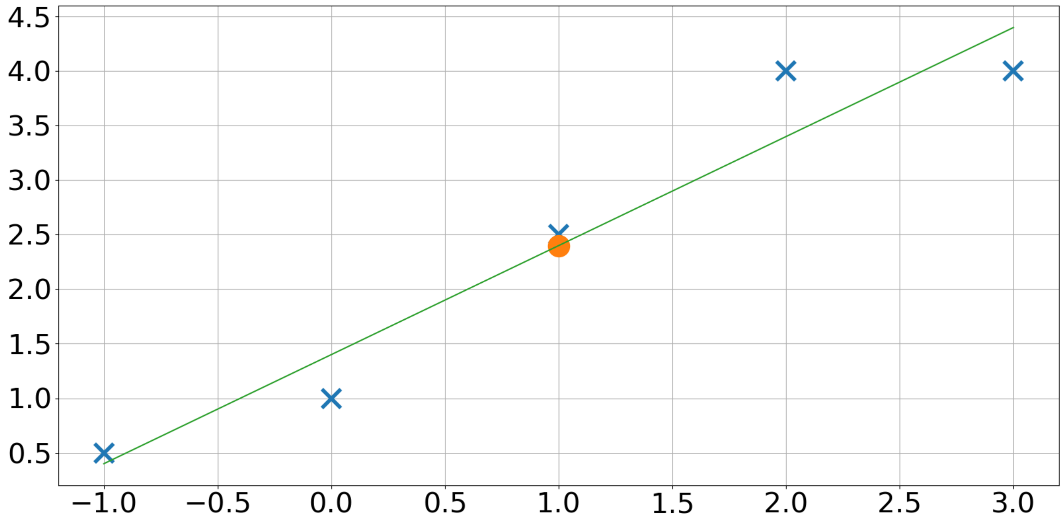
```
import numpy as np
def EcartType(X) :
    Moy = sum(X)/len(X)
    S = 0
    for k in range(len(X)) :
        S = S + (X[k]-Moy)**2
    return np.sqrt(S/len(X))

def Covariance(X,Y) :
    S = 0
    for k in range(len(X)) :
        S = S + X[k]*Y[k]
    return S/len(X) - (sum(X)/len(X))*(sum(Y)/len(Y))

X = [-1,0,1,2,3]
Y = [0.5,1,2.5,4,4]
Covariance(X,Y)/(EcartType(X)*EcartType(Y))
```

4. La droite de régression linéaire recherchée est

$$y = \frac{s_{x,y}}{s_x^2}(x - \overline{x}) + \overline{y} = \frac{2}{2}(x-1) + 2.4 = x - 1 + 2.4 = x + 1.4.$$



5.

▷ **Exercice 3 :**

1. Rappeler la définition et la caractérisation des relations de négligeabilité des suites.
2. Dans chaque cas dire si il existe une relation de négligeabilité entre les deux suites :
   (a) Pour tout $n \in \mathbb{N}\backslash\{0;1\}$, $u_n = \dfrac{1}{n^2}$ et $v_n = \dfrac{n}{n^4-1}$
   (b) Pour tout $n \in \mathbb{N}^*$, $u_n = \dfrac{n^2}{n+1}$ et $v_n = \dfrac{n^2+1}{n}$
   (c) Pour tout $n \in \mathbb{N}^*$, $u_n = e^n$ et $v_n = n^n$

1. Voir le cours
2. On observe que pour chaque cas pour tout $n \in \mathbb{N}\backslash\{0;1\}$, u_n et v_n sont des réels non nuls.
   (a) On a
   $$\frac{v_n}{u_n} = \frac{\frac{n}{n^4-1}}{\frac{1}{n^2}} = \frac{n^3}{n^4-1} = \frac{1}{n\left(1-\frac{1}{n^4}\right)}$$
   On a alors $\lim\limits_{n\to+\infty} \dfrac{v_n}{u_n} = 0$ donc on a $v_n \underset{n\to+\infty}{=} o(u_n)$.
   (b) On a
   $$\frac{u_n}{v_n} = \frac{\frac{n^2}{n+1}}{\frac{n^2+1}{n}} = \frac{n^3}{(n+1)(n^2+1)} = \frac{1}{\left(1+\frac{1}{n}\right)\left(1+\frac{1}{n^2}\right)}$$
   et
   $$\frac{v_n}{u_n} = \frac{\frac{n^2+1}{n}}{\frac{n^2}{n+1}} = \frac{(n+1)(n^2+1)}{n^3} = \left(1+\frac{1}{n}\right)\left(1+\frac{1}{n^2}\right)$$

On a alors $\lim_{n\to+\infty} \frac{u_n}{v_n} = \lim_{n\to+\infty} \frac{v_n}{u_n} = 1 \neq 0$ donc on a aucune relation de négligeabilité entre ces deux suites.

(c) On a

$$\frac{u_n}{v_n} = \frac{e^n}{n^n} = \frac{e^n}{e^{n\ln(n)}} = e^{n(1-\ln(n))}$$

On a alors $\lim_{n\to+\infty} \frac{u_n}{v_n} = \lim_{X\to-\infty} e^X = 0$ donc on a $u_n \underset{n\to+\infty}{=} o(v_n)$.