

Nom, prénom :

/10

Durée : 15 minutes. Accès à python et cours sur internet.

IMPORTANT ! Pour l'exercice 1, vous devez nous envoyer la réponse sous la forme d'un script python (coeff_regression_lineaire.py) par mail (albenhenni@gmail.com).

- **Pour avoir tous les points, le script doit tourner sans erreurs et retourner les bonnes valeurs.**
- **Vous ne pouvez utiliser numpy que lorsque c'est précisé (1.3)**

Question 1 : (6 points)

Soit le lot de données (x, y) défini de la façon suivante :

```
import numpy as np

params=(5,3)
x = 10*np.random.random(100)
y = params[0] + params[1]*x + np.random.normal(size=len(x))
```

Le résultat de la résolution analytique pour les coefficients de la régression linéaire est la suivante :

$$\hat{\theta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{cov(x, y)}{var(x)}$$

$$\hat{\theta}_0 = \bar{y} - \hat{\theta}_1 \bar{x}$$

$\hat{\theta}_1$ et $\hat{\theta}_0$ sont les coefficients que l'on souhaite estimer, tel que

$$\hat{y} = \hat{\theta}_0 + \hat{\theta}_1 x$$

\bar{x} et \bar{y} sont les moyennes respectives des tableaux x et y.

La covariance et la variance sont définies par :

$$cov(x, y) = \frac{1}{N} * \sum (x_i - \bar{x})(y_i - \bar{y})$$

$$var(x) = \frac{1}{N} * \sum (x_i - \bar{x})^2$$

avec $N = len(x) - 1$

Nous allons vous demander de créer plusieurs fonctions pour calculer au final les paramètres $\hat{\theta}_1$ et $\hat{\theta}_0$ de façon analytique. Les questions de cet exercice 1 sont dépendantes.

Question 1.1 : Créer la fonction `moyenne(x)` retournant la valeur de la moyenne du tableau `x`. (1 point)

Question 1.2 : Créer la fonction `produit_scalaire(x,y)` permettant de retourner le résultat du produit scalaire du tableau `x` par le tableau `y`. (1 point)

Question 1.3 : Créer la fonction `variance(x)` permettant de retourner le résultat du calcul de la variance, définie plus haut. (1 point)

Indice : vous pouvez utiliser `np.square(...)` pour mettre un élément au carré.

Question 1.4 : Créer la fonction `covariance(x,y)` permettant de retourner le résultat du calcul de la covariance, définie plus haut. (1.5 point)

Question 1.5 : Créer la fonction `linear_parameters(x,y)` qui retourne les résultats de l'estimation des paramètres $\hat{\theta}_1$ et $\hat{\theta}_2$, d'après l'équation définie plus haut, et l'appeler sur le jeu de données généré par le code fourni plus haut (1.5 point)

Question 2 : Parmi les affirmations suivantes, cochez celles qui sont correctes (3 points)

- ☐ `git copy` permet de copier localement un dépôt distant
- ☒ `git add` ajoute un fichier modifié à la liste du prochain `commit`
- ☒ `git commit -m "message"` sauvegarde l'état actuel du code avec en commentaire le message "message"
- ☐ `git commit` pousse automatiquement les modifications sur le serveur distant du dépôt
- ☐ `git get` récupère les modifications poussées par les autres développeurs
- ☒ `git` permet de mieux gérer et partager les différents modifications au code d'un projet

Question 3 : Quelle proposition permet de générer un nombre aléatoire distribué de façon uniforme entre 5 et 10 ? (1 points)

- ```
from random import random
```
- ☐ `5*random()`
  - ☐ `10*random()`
  - ☒ `(10-5)*random()+5`
  - ☐ `10*random()+5`