

# Probabilités et Statistiques

## Projet noté

MADANI Abdenour  
TRIOLET Hugo

Licence 3  
2021 - 2022

### Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Objectifs . . . . .	2
1.2	Définitions . . . . .	2
1.3	Résumé de notre approche . . . . .	2
<b>2</b>	<b>Régression linéaire</b>	<b>2</b>
2.1	Régression Linéaire simple . . . . .	2
2.1.1	Modèle vectoriel . . . . .	3
2.1.2	Résultats obtenus . . . . .	3
2.2	Régression linéaire et descente de gradient . . . . .	4
<b>3</b>	<b>Étude et manipulation de lois de probabilités</b>	<b>4</b>
3.1	Loi Binomiale . . . . .	4
3.2	Loi Normale univariée . . . . .	4
3.3	Simulation de données à partir d'une loi . . . . .	4
3.3.1	Cas de la loi normale . . . . .	4
3.4	Estimation de densité . . . . .	4
3.4.1	Cas de la loi normale . . . . .	4
3.4.2	Cas de la loi exponentielle . . . . .	4
<b>4</b>	<b>Intervalles de confiance</b>	<b>4</b>
4.1	Fonctions . . . . .	5
4.2	Problème 1 . . . . .	5
4.3	Problème 2 . . . . .	5
4.4	Problème 3 . . . . .	5
<b>5</b>	<b>Exemples d'utilisation du code</b>	<b>5</b>
5.1	Comment utiliser le code . . . . .	5

# 1 Introduction

## 1.1 Objectifs

Les objectifs de ce TPs sont :

- implémenter nous-mêmes plusieurs algorithmes de régression linéaire et les comparer à des fonctions issues de bibliothèques scientifiques
- manipuler différentes lois vues en cours via leur implémentation issues de bibliothèques scientifiques
- déterminer des intervalles de confiance et effectuer des applications sur quelques exemples

On utilisera pour ceci **Python** et les bibliothèques de fonctions : Numpy, Scipy, Matplotlib, et Statsmodels, entre autres.

## 1.2 Définitions

Hugo : tu peux virer ça si t'as aucune définition à mettre (tu peux la réutiliser plus bas et virer cette partie aussi)

### DÉFINITION

#### Mot défini

Définition ici

## 1.3 Résumé de notre approche

Nous avons 3 fichiers, 1 pour chaque TP.

Vis-à-vis du code, nous l'avons documenté à l'aide de la docstring de Python, ainsi que des commentaires normaux : les fonctions se comprennent donc naturellement grâce à ceux-ci.

# 2 Régression linéaire

## 2.1 Régression Linéaire simple

La fonction calculant la régression linéaire simple est "regression\_lineaire".

Étant donné deux listes  $x$  et  $y$  de même taille, elle calcule la régression linéaire

$$y = \beta_1 \cdot x + \beta_0$$

### 2.1.1 Modèle vectoriel

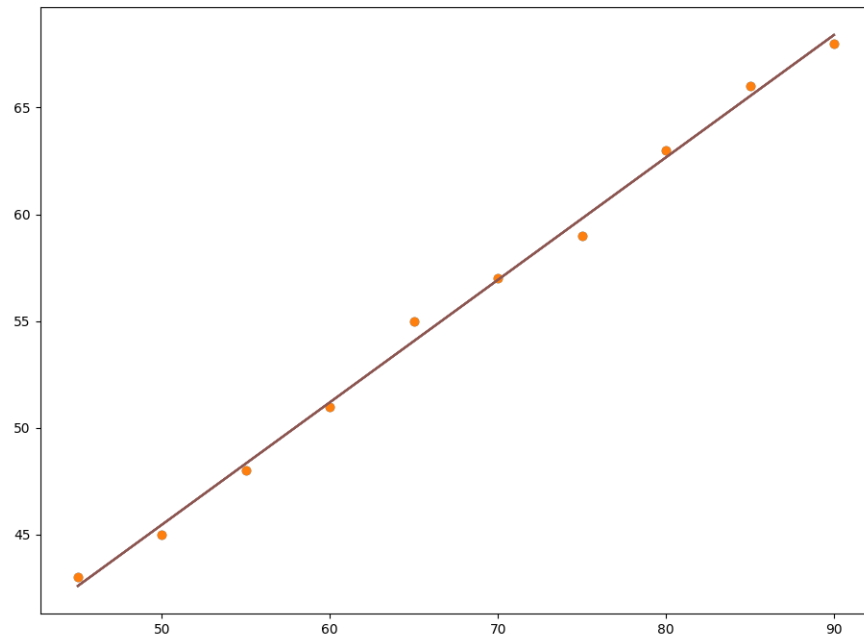
On applique simplement la formule donnée dans le TP.

La fonction calculant la régression linéaire simple est "regression\_lineaire\_vec".

Étant donné deux listes  $x$  et  $y$  de même taille, elle calcule la régression linéaire en utilisant la méthode vectorielle :

$$y = \beta_1 \cdot x + \beta_0$$

### 2.1.2 Résultats obtenus



Représentation graphique obtenue avec Matplotlib

En orange sont affichés les points de  $(x_i, y_i)$ , et on voit plusieurs droites superposées de couleurs différentes, quasiment indiscernables : ce sont nos deux régressions linéaires ainsi que celle de Numpy (polyfit).

Les résultats sont donc concordants : visuellement, toutes les régressions linéaires donnent le même résultat sur ce jeu de donnée.

Les coefficients sont de mêmes très proches voire égaux.

## 2.2 Régression linéaire et descente de gradient

Hugo : todo

## 3 Étude et manipulation de lois de probabilités

### 3.1 Loi Binomiale

texte

Si tu veux mettre une image Hugo

### 3.2 Loi Normale univariée

texte

### 3.3 Simulation de données à partir d'une loi

texte

#### 3.3.1 Cas de la loi normale

texte

### 3.4 Estimation de densité

texte

#### 3.4.1 Cas de la loi normale

texte

#### 3.4.2 Cas de la loi exponentielle

texte

## 4 Intervalles de confiance

Le but de cette partie (correspondant au fichier *tp3.py*) est de déterminer les intervalles de confiance de différentes lois de probabilité. Dans les 2 premiers problèmes, nous ne connaissons pas la variance de l'échantillon, ainsi la fonction calculant l'intervalle de confiance est de type empirique et le fractile d'ordre  $1 - \frac{\alpha}{2}$  de la loi de Student.

## 4.1 Fonctions

## 4.2 Problème 1

*On possède deux échantillons de taille 16 à notre disposition : un échantillon de masses de 16 pots de confitures mesurée en kilogramme (kg), et l'autre de masses d'avocats provenant du Mexique mesurée en grammes (g).*

## 4.3 Problème 2

*texte*

## 4.4 Problème 3

*texte*

# 5 Exemples d'utilisation du code

## 5.1 Comment utiliser le code

*Concernant le code, il est séparé en trois fichier, chacun correspondant par son indice à la partie du TP correspondante (tp3.pyaveclapartie3, TP2.pyaveclapartie2, ...). Si le code est exécuté sur la tête" du code, c'est à dire de la première ligne au premier) d'exécuter séparément chaque sous-partie du code, délimitée. Autrement, le code s'exécute normalement et entier, si Pyzou' est pas le logiciel où celui-ci est exécuté.*