

# TD-TP2

## Représentations graphiques d'une variable statistique

---

### • Partie TD (sur papier)

#### Exercice 1

*Les questions 1 à 5 de cet exercice correspondent aux questions 1 à 5 de l'exercice 2 du TD-TP1... certaines d'entre vous les ont donc déjà faites.*

On s'intéresse à la variable statistique  $X$  du taux de chômage par régions en 2016, donnée dans le tableau suivant :

Bretagne	7,9
Normandie	9,4
Hauts-de-France	11,7
Grand Est	9,3
ile de France	8,1
Nouvelle-Aquitaine	8,8
Pays de la Loire	7,8
Centre-Val de Loire	8,8
Bourgogne-Franche-Comté	8,3
Auvergne-Rhône-Alpes	8,1
Occitanie	11
Provence-Alpes-Côte d'Azur	10,8
Corse	9,9

- Quelle est la population ?
  - Quel est l'effectif total ?
- Quelle est la moyenne de  $X$  sur l'ensemble des régions ?
- Cette moyenne correspond-elle au taux de chômage moyen en France ?
- Quelle(s) région(s) correspond(ent) à la médiane ? Aux quartiles ?
- Ecrire le calcul permettant de trouver la variance de  $X$ . (On exécutera ce calcul en Python dans la partie TP).

6. Réaliser la boîte à moustaches de Tukey de la variable  $X$ .

On définit le découpage en classes suivant :

$$[7.5, 8[, [8, 8.5[, [8.5, 9.5[, [9.5, 10.5[, [10.5, 12[.$$

7. Faire le tableau statistique des effectifs de la variable  $X$  en fonction de ces classes.
8. Tracer l'histogramme associé. A-t-on besoin de calculer les effectifs corrigés ?

### Exercice 2 - *Rab pour s'entraîner*

On s'intéresse à la série statistique des notes des étudiant.es d'un groupe-TD (le + sympa) lors du DS de Maths Discrètes :

11, 9, 14, 13, 11, 20, 9, 12, 16, 17, 7, 16, 15, 12, 12, 15, 15, 11, 14, 11.

1. (a) Quelle est la population ?  
(b) Quel est l'effectif total ?
2. (a) Faire le tableau statistique avec les effectifs pour toutes les valeurs prises par la série.  
(b) Tracer le diagramme en bâtons associé.  
(c) Calculer la moyenne, la variance et l'écart-type.
3. Dans cette question, on étudie notre série statistique (celle donnée ci-dessus) avec le découpage en classes suivant :  
$$[0, 8[, [8, 10[, [10, 12[, [12, 13[, [13, 15[, [15, 18[, [18, 21[.$$
  - (a) Faire le tableau statistique associé, avec les effectifs, les fréquences et les fréquences cumulées.
  - (b) Tracer l'histogramme associé.
4. (a) Quelle est la médiane, et quels sont les quartiles de cette série de notes ?  
(b) Tracer la boîte à moustaches correspondant à cette série statistique.

## • Partie TP (sur Python)

On importe à nouveau les librairies suivantes :

```
import numpy as np
import numpy.random as rd
```

Pour pouvoir faire des représentation graphiques, on importe de plus la librairie `pyplot`, déjà rencontrée en Méthodes Numériques :

```
import matplotlib.pyplot as plt
```

### Exercice 3 - *Les notes des étudiants*

On fait ce TP dans un fichier à la suite du TP1, car on aura besoin des fonctions **Moyenne** et **Variance** (à faire maintenant si elles ne sont pas finies).

1. Rentrer dans une liste **X** la série statistique des notes de l'exercice 2.
2. Vérifier vos calculs de la moyenne et de la variance (question 2c de l'exercice 2) grâce aux fonctions **Moyenne** et **Variance**.

Sur [cpge.frama.io/fiches-cpge/Python](http://cpge.frama.io/fiches-cpge/Python), vous trouverez des tutoriels de Python faits pour les classes prépa. On va utiliser un tutoriel qui introduit la fonction `plt.hist`, fonction qui permet de tracer des histogrammes en Python : [cpge.frama.io/fiches-cpge/Python/Graphiques/2-Histogramme/](http://cpge.frama.io/fiches-cpge/Python/Graphiques/2-Histogramme/)

3. A l'aide du tutoriel ci-dessus, tracer avec Python l'histogramme de la question 3b de l'exercice 2. Vous aurez normalement besoin de rajouter l'option **density=True** pour que votre histogramme aie la même forme que celui fait sur papier.
4. Tester la commande suivante :

```
plt.boxplot(X)
```

Comparer au résultat obtenu à la question 4b de l'exercice 2.

#### Exercice 4 - *Lancers de dés*

On réutilise la série **Serie100** de 100 lancers de dés, fabriquée dans le TD-TP1.

1. Tracer avec Python le diagramme en bâtons associé à votre série statistique **Serie100**.

*Indication* : voici le code utilisé pour tracer le diagramme en bâtons des absences des étudiant.es, apparaissant dans le Cours 1 :

```
X=[2,4,2,1,0,0,3,1,3,0,2,3,0,3,1,0,2,0,2,1,0,3,1,2,2,
3,1,3,2,2,0,2,0,3,2,1,2,2,3,2,0,2,0,2,0,2,1,2,0,0,3,
1,0,2,0,2,1,2,2,2,1,2,0,2,1,0,0,4,1,0,3,0,2,0,0,3,1,
3,0,0,1,2,1,2,6,2,2,1,1,2,3,1,0,1,1,1,3,1,3,2]
Vals=[0,1,2,3,4,5,6,7]
plt.hist(X,Vals,histtype='bar',align='left',rwidth=0.1)
```

#### Exercice 5 - *Pour les + rapides : pleein de lancers de dés*

1. Faire une fonction **Moy100()** qui réalise 100 tirages de dés et renvoie la moyenne de ces 100 tirages.
2. Faire une fonction **TabMoy100(N)** qui renvoie un tableau contenant  $N$  fois le résultat de l'expérience de la question 1.
3. Tester votre fonction **TabMoy100(N)** avec  $N = 50$ . Quelle semble être la moyenne théorique d'un tirage de dés ?
4. Stocker dans une variable **Tab10000** le résultat de votre fonction **TabMoy100** pour  $N = 10000$ . Tracer un histogramme de votre série statistique **Tab10000**.

La forme de cet histogramme vous rappelle-t-elle quelque chose ?

*Indication* : La partie "Personnalisation des intervalles" du tutoriel de `plt.hist` indiqué dans l'exercice 3 pourra vous servir.