

## Exercices de TD- Feuille 3 (Statistiques descriptives)

### Exercice 1

1. Si  $x$  désigne un caractère statistique, et avec les notations du cours, que représente  $\frac{\sum_{k=1}^i n_k}{n}$  ?
2. toujours avec les mêmes conventions, montrer que  $\sum_{i=1}^p f_i(x_i - 2) = (\sum_{i=1}^p f_i x_i) - 2$ .  
Comment se résultat peut-il s'interpréter ?

### Exercice 2

Voici les loyers mensuels (en euros, 0 pour une occupation gratuite) et les surfaces (en  $m^2$ ) de quelques appartements:

|         |     |     |     |    |     |     |     |    |     |
|---------|-----|-----|-----|----|-----|-----|-----|----|-----|
| Loyer   | 400 | 340 | 500 | 0  | 450 | 600 | 410 | 0  | 520 |
| Surface | 20  | 18  | 28  | 47 | 24  | 40  | 22  | 27 | 30  |

1. Préciser la population et les caractères statistiques. Ces caractères sont-ils qualitatifs, quantitatifs, discrets, continus ?
2. Calculer la moyenne et la médiane des loyers et des surfaces.
3. Dans cette situation, quel avantage de la médiane sur la moyenne peut-on observer ?
4. Que penser d'un appartement dont la surface est la surface moyenne et le prix le prix moyen ?

### Exercice 3

On a relevé le nombre de dents cariées chez 100 enfants âgés de 7 ans dans une région de France.

Les résultats obtenus sont les suivants :

|                     |    |    |    |    |   |   |   |   |
|---------------------|----|----|----|----|---|---|---|---|
| Nb de dents cariées | 0  | 1  | 2  | 3  | 4 | 5 | 6 | 7 |
| Nb d'enfants        | 30 | 34 | 14 | 10 | 4 | 5 | 1 | 2 |

1. Faire l'étude statistique de cette série, c'est-à-dire :
  - (a) Représenter graphiquement les données.
  - (b) Tracer le diagramme des fréquences cumulées.
  - (c) Déterminer les caractéristiques statistiques attachées à cette série : mode, médiane, moyenne empirique, variance empirique et quartiles.

2. On ne désire pas une étude aussi précise et on se contente d'une étude faite en regroupant les données par classes :

|                     |       |       |       |
|---------------------|-------|-------|-------|
| Nb de dents cariées | [0,2[ | [2,4[ | [4,8[ |
| Nb d'enfants        | ?     | ?     | ?     |

- (a) Tracer l'histogramme de cette nouvelle série.  
 (b) Déterminer la moyenne empirique et la variance empirique. Retrouve-t-on les mêmes valeurs ? Pourquoi ?

## Exercice 4

Dans une copropriété on a relevé pour chaque logement sa surface ( notée  $x$  ) .

|               |         |          |          |          |            |        |
|---------------|---------|----------|----------|----------|------------|--------|
| $x_i$         | [20,30[ | [30, 40[ | [40, 60[ | [60, 75[ | plus de 75 | Somme  |
| centres $c_i$ | 25      | 35       | 50       | 67,5     | 85         |        |
| $n_i$         | 16      | 28       | 56       | 44       | 56         | 200    |
| $c_i n_i$     | 400     | 980      | 2800     | 2970     | 4760       | 11910  |
| $c_i^2 n_i$   | 10000   | 34300    | 140000   | 200475   | 404600     | 789375 |
| $c_i n_i^2$   | 6400    | 27440    | 156800   | 130680   | 266560     | 587880 |

1. Tracer le diagramme représentant les fréquences cumulées.
2. Déterminer la médiane et le premier quartile .
3. Calculer l'écart type.
4. Ces logements ont tous été construits avec un cellier de  $7m^2$  attenant, qui n'a pas été pris en compte dans la surface. Si on ajoute ces  $7m^2$  à la surface de chaque appartement, en quoi cela modifie-t-il : la moyenne, la médiane, la variance et l'écart-type ?

## Exercice 5

On a relevé la consommation hebdomadaire de sucre (en grammes) de 300 familles:

|                        |            |            |            |            |            |            |
|------------------------|------------|------------|------------|------------|------------|------------|
| Consommation ( $x_i$ ) | [420, 480[ | [480, 540[ | [540, 570[ | [570, 630[ | [630, 690[ | Somme      |
| Nombre de familles     | 8          | 83         | 71         | 114        | 24         | 300        |
| centres $c_i$          | 450        | 510        | 555        | 600        | 660        |            |
| $c_i n_i$              | 3 600      | 42 330     | 39 405     | 68 400     | 15 840     | 169 575    |
| $c_i^2 n_i$            | 1 620 000  | 21 588 300 | 21 869 775 | 41 040 000 | 10 454 400 | 96 572 475 |
| $c_i n_i^2$            | 28 800     | 3 513 390  | 2 797 755  | 7 797 600  | 380 160    | 14 517 705 |

1. Calculer la moyenne.
2. Tracer le diagramme représentant les fréquences cumulées.
3. Déterminer la médiane et le troisième quartile (faire le calcul exact par la méthode de l'interpolation lineaire).
4. Calculer l'écart type.
5. On considère maintenant le caractère "consommation hebdomadaire de sucre **en kilogramme**". Calculer la moyenne et la variance de ce caractère.

## Exercice 6

Le tableau suivant concerne des exploitations agricoles françaises:

| Taille (en ha) | % d'exploitations | % de superficie |
|----------------|-------------------|-----------------|
| [1,20[         | 51.4              | 14.3            |
| [20,50[        | 31.4              | 34.8            |
| [50 et +[      | 17.2              | 50.9            |

1. Tracer la courbe de concentration de Lorentz.
2. Calculer l'indice de Gini.
3. Calculer la médiane.

## Exercice 7

On observe les moyennes trimestrielles en mathématiques des 36 élèves d'une classe de seconde comprenant 8 redoublants. La note moyenne des redoublants est de 9,8 avec un écart-type de 3, la note moyenne des autres élèves est de 11 avec un écart type de 2,4. Calculer les variance intra et inter-population et conclure.

## Exercice 8

Soient  $x_1, x_2, \dots, x_n$   $n$  nombres réels quelconques, et  $\bar{x}$  la moyenne de ces nombres.

1. Montrer que

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2.$$

2. En déduire que la variance d'une série statistique est toujours un nombre positif.
3. Dans quel cas la variance d'une série statistique est-elle nulle ?