

## ED1 : Statistique descriptive des distributions univariées

### Exercice 1 :

Le tableau suivant donne les exportations d'énergie primaire de deux pays A et B.

Type d'énergie	Pays A	Pays B
Combustibles solides	20	32
Produits pétroliers	52	114
Gaz	33	86
Electricité	45	18
Total des exportations	150	250

Unité : millions de tonnes-équivalents pétrole

1. Pourquoi avoir exprimé toutes les exportations dans la même unité ?
2. Comparer graphiquement le volume et la structure des exportations d'énergie des pays A et B en utilisant :
  - un diagramme en bâton,
  - un graphique circulaire.

### Exercice 2 :

On dispose de la statistique suivante relative à 344 exploitations laitières d'une certaine région, classées d'après leur production de lait :

Production (en milliers de litres)	Nombre d'exploitations
$[0 ; 1[$	26
$[1 ; 2[$	35
$[2 ; 3[$	29
$[3 ; 5[$	53
$[5 ; 10[$	105
$[10 ; 20[$	71
$[20 ; 40[$	25

1. Construire l'histogramme relatif à cette distribution. Déterminer la classe modale.
2. Construire la courbe des fréquences cumulées.
3. Déterminer la médiane et la moyenne, puis l'écart interquartile, la variance et l'écart-type.

### **EXERCICE 3**

Soit  $X$  une variable statistique de moyenne  $\bar{x}$  de variance  $s_x^2$  et d'écart type  $s_x$

Soit  $Y = aX + b$  avec  $a > 0$

Calculer  $\bar{y}$ ,  $s_y^2$  et  $s_y$

### **EXERCICE 4**

Calculer la moyenne et la variance des  $n$  premiers nombres entiers 1, 2, ..n

Indication : 
$$\sum_1^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

### **EXERCICE 5**

Soient  $x_1, x_2, x_n$  les valeurs observées pour  $n$  individus, de la variable  $X$

1) Démontrer que la moyenne arithmétique  $\bar{x}$  minimise la quantité

$$Q(a) = \sum_{i=1}^n (x_i - a)^2$$

2) Démontrer que la médiane minimise la quantité

$$P(a) = \sum_{i=1}^n |x_i - a|$$

### **EXERCICE 6**

Durant 4 ans un consommateur a consacré la même quantité d'argent à l'achat de chocolat, aux prix successifs de 17 euros le kilo, la première année, 22 euros la deuxième année, 27 euros la troisième année, 30 euros la quatrième.

- 1) Quel est le prix moyen du kilo de chocolat acheté par ce consommateur sur les quatre ans ?
- 2) Durant ces mêmes quatre années, un autre consommateur a acheté une quantité identique, année par année, de ces mêmes chocolats dans le même magasin. Quel est pour lui le prix moyen payé ?

## **EXERCICE 7**

Ci dessous sont présentées les caractéristiques de la distribution des vitesses maximales de 40 voitures.

- 1) Rappeler la signification des différents paramètres affichés
- 2) Tracer un box-plot de cette distribution

### **Statistiques résumées pour VITESSE**

Effectif = 40  
Moyenne = 202,6  
Médiane = 200,5  
Mode = 250,0  
Moyenne géométrique = 199,842  
Variance = 1171,17  
Ecart-type = 34,2223  
Erreur-type = 5,41103  
Minimum = 150,0  
Maximum = 300,0  
Etendue = 150,0  
1er quartile = 171,0  
3ème quartile = 225,0  
Etendue inter-quart = 54,0  
Asymétrie = 0,466255  
Asymétrie std. = 1,20387  
Aplatissement = 0,0773924  
Aplatissement std. = 0,0999132  
Coef. de variation = 16,8916%

Ce rapport donne des statistiques résumées pour VITESSE. Il inclut des mesures de tendance centrale, des mesures de variabilité et des mesures de forme. Les coefficients d'aplatissement standardisé et d'asymétrie standardisée sont particulièrement utiles car ils peuvent être utilisés pour déterminer si l'échantillon suit une loi normale. Des valeurs de ces statistiques en dehors de la plage -2 à +2 indiquent un écart significatif à la normalité, ce qui rend non valides les tests statistiques concernant l'écart-type. Dans ce cas, la valeur de l'asymétrie standardisée est dans la plage attendue pour des données suivant une loi normale. La valeur de l'aplatissement standardisé est dans la plage attendue pour des données suivant une loi normale.