

## I EXERCICE-1

Ce tableau donne la répartition des contribuables selon la tranche de revenu imposable en francs, dans le département de la Drôme, en 1988.

$a_i$	$b_i$	$n_i$
0	40	4500
40	50	8700
50	60	8800
60	70	8500
70	80	6700
80	90	7500
90	100	700
100	125	13100
125	150	7700
150	200	7100
200	250	2600
250	500	2700
500	500 et plus	700

1. Représenter cette série par un histogramme.
2. Préciser la classe modale, en justifiant votre résultat, puis calculer le mode.
3. Donner à l'aide de la calculatrice les valeurs de la moyenne, la variance et de l'écart-type, en rappelant les formules du cours.
4. Estimer la proportion d'observations situées dans l'intervalle  $[\bar{x} - 2\sigma(x); \bar{x} + 2\sigma(x)]$ .
5. Calculer la médiane.
6. Représenter la boîte à moustache.
7. Calculer l'indice de gini après avoir représenté la courbe de Lorenz.

## II EXERCICE-2

L'entreprise Kado a demandé à Monsieur Noël, le statisticien de l'entreprise, de tester le lien entre son chiffre d'affaires  $y$  (en millions d'euros) et ses dépenses de publicité  $x$  (en millions d'euros). Sur la base des résultats des 7 dernières années, les données se présentent comme suit ( $x_i$  et  $y_i$ , pour  $i$  entier variant de 1 à 7, représentent les modalités respectives des deux caractères) :

$x_i$	1	2	5	7	9	12	16
$y_i$	100	200	600	1200	1500	2000	2500

1. **Le nuage**
  - a. Calculer les moyennes  $\bar{x}$  et  $\bar{y}$ , à  $10^{-3}$  près.
  - b. Tracer le nuage de points représentant cette série double et placer le point moyen G sur le graphique.
2. **Premier ajustement**
  - a. Indiquer le type d'ajustement que permet de conjecturer l'observation graphique.
  - b. Déterminer le coefficient de corrélation linéaire entre  $x$  et  $y$  à  $10^{-4}$  près (donner directement le résultat fourni par la calculatrice).
  - c. Rappeler la formule donnant le coefficient de corrélation linéaire et commenter le résultat.
  - d. Donner l'équation de la droite de régression de  $y$  en  $x$ , obtenue par la méthode des moindres carrés, sous la forme  $y = ax + b$ . Rappeler les formules exprimant les coefficients  $a$  et  $b$  et donner directement le résultat fourni par la calculatrice.
  - e. Formuler l'équation d'analyse de la variance et calculer la somme des carrés résiduels.
  - f. Calculer le coefficient de détermination et donner son interprétation.

3. **Deuxième ajustement**

Monsieur Noël indique que l'ajustement donné par :

$$y = 94,762x^{1,21924}$$

est de meilleure qualité que le précédent.

- Montrer que l'ajustement précédent peut être obtenu à l'aide d'un ajustement linéaire entre  $t = \ln x$  et  $z = \ln y$ .
- Donner sur votre copie le tableau de valeurs de la série double  $(t_i ; z_i)$ .
- Donner sans justification une équation de la droite de régression de  $z$  en  $t$ , sous la forme  $z = at + b$ ,  $a$  et  $b$  étant donnés à  $10^{-5}$  près.
- Retrouver l'ajustement proposé par Monsieur Noël.
- Etes vous d'accord avec Monsieur Noël quand il affirme que le deuxième ajustement est de meilleure qualité que le premier ? (justifier votre réponse).

### III EXERCICE-3

On considère les données suivantes, concernant l'âge et le sexe des employés d'une entreprise

Age		
	F – Femmes	H – Hommes
[16; 25[	147	283
[25; 35[	61	275
[35; 45[	33	62
[45; 55[	20	55
[55; 60[	22	41

- Calculer les fréquences (relatives) marginales de cette série double.
- Calculer la fréquence des femmes de moins de 25 ans
- Calculer la fréquence des [35; 45[ parmi les hommes
- Calculer l'âge moyen des employés.