



Dans une étude sur les baies de fruits rouges, 37 individus ont été pesés. Les résultats, exprimés en g sont les suivants :

1,05	1,38	1,36
1,96	1,07	0,90
2,00	1,07	1,81
1,26	1,37	1,28
2,11	1,23	0,86
2,53	1,28	1,01
1,49	0,61	1,18
0,97	0,90	0,66
0,93	1,02	0,87
0,73	1,84	1,83
4,38	1,05	0,86
1,12	1,13	0,96
1,12		

Questions

1. Construire le tableau de distribution de fréquence des données.
2. Représenter graphiquement ce tableau.
3. Calculer le mode de la distribution.
4. Calculer la médiane.
5. Calculer la moyenne.
6. Comparer les résultats obtenus sur les trois paramètres de position précédents et interpréter les différences.
7. Pour des données très fortement asymétriques la moyenne arithmétique est souvent remplacée par la moyenne géométrique, qui revient à calculer une moyenne arithmétique sur le logarithme des données. Construire le tableau de distribution de fréquence des données transformées \log^1 .

¹ Si la base du logarithme n'est pas indiquée, il s'agit toujours du logarithme à base 10.

8. Représenter graphiquement ce dernier tableau.
9. Comparativement à celui de la question 2, s'avère-t-il beaucoup moins asymétrique ?
10. Calculer le mode et la médiane sur les données transformées.
11. Calculer la moyenne géométrique des données.
12. S'avère-t-elle moins affectée par les valeurs extrêmes ? Justifier votre réponse
13. Calculer la variance des données non transformées et transformées.
14. La transformation logarithmique a-t-elle réduit la dispersion des données ?
15. Calculer l'écart type, l'étendue de la variation et le coefficient de variation des données transformées et non transformées.
16. L'interprétation des résultats reste-t-elle identique ? ou du moins compatible ?
17. Un des paramètres fortement affectés par la transformation logarithmique est le coefficient d'asymétrie. Pourquoi ?
18. Quelle signification peut-on donner aux paramètres calculés sur des résultats transformés en logarithmes ?

2eme partie

Supposons d'abord que la distribution est proche de la normalité.

1. Donner l'intervalle de confiance sur la moyenne (95% confiance)
2. Donner l'intervalle de confiance sur la moyenne (99% confiance)
3. Calculer la moyenne sur les dix dernières valeurs seulement, puis calculer les intervalles de confiance à 95% et 99%).