

Statistique - L1 SV - 2009/2010
Feuille 1 : paramètres d'une population

Exercice 1/

Un test quantitatif de la charge virale permet de mesurer la quantité de virus contenu dans un millilitre de sang. Ce test est utilisé pour déterminer une infection du VHC (Virus Hépatite C). Le test par PCR (réaction en chaîne de la polymérase) détecte la présence de l'ARN du VHC dans le sang, révélant ainsi une infection active. Ce type de test quantitatif par PCR est très sensible, il peut mesurer une quantité aussi petite que 100 UI/mL (UI = Unité Internationale).

La série de données suivante représente les charges virales de dix patients exprimées en 10^4 UI/mL :

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
43	63	57	65	95	62	82	41	90	52

On donne $x_1 + \dots + x_{10} = 650$ et $x_1^2 + \dots + x_{10}^2 = 45370$.

- 1) Quelles sont les variables mesurées ?
- 2) Paramètres de position: calculez la moyenne, une médiane, les quartiles, les extrêmes et éventuellement le(s) mode(s).
- 3) Paramètres de dispersion : calculez la variance, l'écart-type, l'amplitude, le coefficient de variation.
- 4) Tracez l'histogramme.
- 5) Construire le boxplot.
- 6) Que se passe-t-il si on remplace x_1 par $x_1 = 120$?

On décide de considérer comme "faible" une charge virale inférieure à 500000 UI/mL; de "moyenne" une charge virale entre 500000 et 800000; et comme "élevée" une charge supérieure à 800000.

- 7) Construire une nouvelle variable en recodant la variable "charge virale".
- 8) Représenter (en pourcentage) cette nouvelle variable.

8) Qu'a-t-on perdu comme information ?

Exercice 2/

Une étude du taux d'hémoglobine dans le sang sur un échantillon de 100 personnes âgées de plus de soixante ans a donné les résultats du tableau suivant (en gr/L) :

[100; 110[6
[110; 120[13
[120; 130[16
[130; 140[22
[140; 150[18
[150; 160[10
[160; 170[6
[170; 180[4
[180; 190[3
[190; 200[2

- 1) Quelle est la population statistique ? Quelle est la variable statistique ?
- 2) Tracer l'histogramme associé à ces données .
- 3) Tracer la courbe des fréquences cumulées.
- 4) Déterminer le mode et la moyenne de cette distribution.
- 5) Déterminer les trois premiers quartiles.

Exercice 3/

Les histogrammes suivants représentent les moyennes (fictives) des étudiants dans différentes filières universitaires.

- 1) Quelles sont les variables mesurées et les populations statistiques ?
- 2) Comparer ces populations.

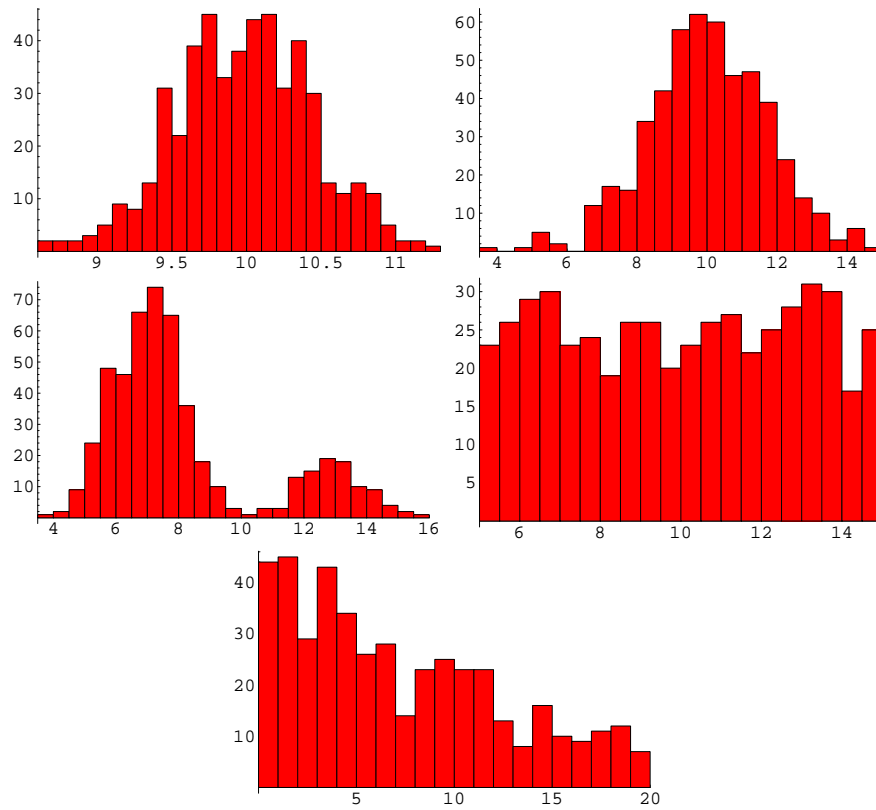


Figure 1: Moyennes dans cinq filières différentes

Exercice 4/

Les boxplots suivants représentent les poids (en kg) mesurés dans cinq cliniques différentes pour 100 nouveaux nés dans chaque clinique.

- 1) Quelles sont les variables étudiées et les populations statistiques ?
- 2) Comparer ces populations.

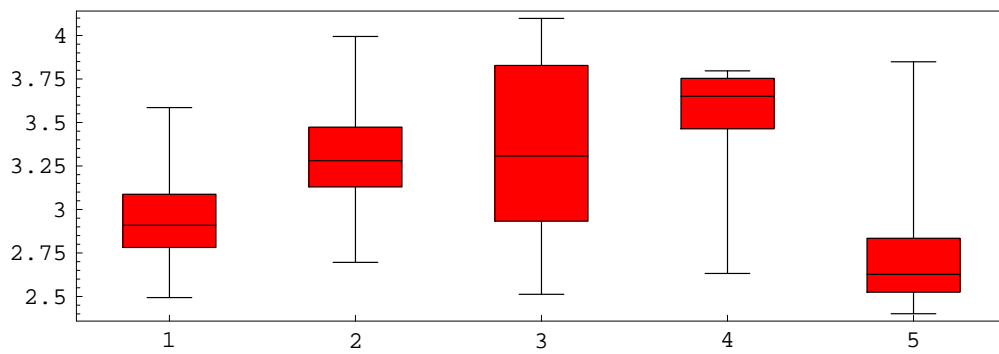


Figure 2: Poids de naissance dans cinq cliniques