

Durée de l'épreuve : 3h00

Tous documents autorisés.

(Les trois exercices sont indépendants. Un soin tout particulier sera apporté à la rédaction des réponses)

Exercice 1**Partie I.**

1. Importer sous R le fichier de données mis à disposition dans le dossier « Master SIDE/ » en utilisant la fonction adéquate.
2. Afficher, une par une, les variables *taille*, *poids* et *sexe*. Créer 3 vecteurs taille, poids et sexe correspondant aux données de la table.
3. Quel est le nombre d'individus dans l'échantillon ?
4. Afficher le tableau des indicateurs de tendance centrale pour les deux variables quantitatives. Essayer également cette commande en donnant le nom de la table et non le nom d'une variable. Qu'obtient-on ?
5. Décomposer l'échantillon en un échantillon d'hommes et un échantillon de femmes.
6. Quel est le nombre d'hommes et le nombre de femmes dans l'échantillon ?

Partie II.

La masse corporelle d'un individu est mesurée à partir du poids et de la taille. L'indice le plus utilisé est celui proposé par Adolphe Quêtelet (1796-1874), astronome et mathématicien belge.

7. Construire la variable « indice de Quêtelet » à partir des données

$$\text{imc} = \frac{\text{poids}}{\text{taille}^2}$$

La taille est exprimée en mètre et le poids en kilogrammes. Cet indice s'appelle aujourd'hui « indice de masse corporelle » (imc) ou « body mass index » (bmi). Il permet de mesurer la corpulence de l'homme adulte. L'Organisation Mondiale de la Santé a défini les critères suivants : maigreur (inférieur à 18,5), normal (de 18,5 à 25), risque de surpoids (de 25 à 30), obésité (supérieur à 30). Mais l'indice de Quêtelet n'a qu'une valeur indicative. Pour déterminer l'existence d'une obésité réelle, il faut faire d'autres mesures destinées à établir exactement la proportion de masse grasse, car c'est l'excès de masse grasse qui représente un facteur de risque.

8. Calculer les paramètres statistiques élémentaires de cette nouvelle variable sur l'ensemble des individus et en fonction de chaque sexe.
9. Construire l'histogramme de cette nouvelle variable sur l'ensemble des individus.
10. Comparer graphiquement les indices de masse corporelle chez les hommes et chez les femmes à l'aide d'histogrammes et de fonctions de répartitions empiriques (utiliser la fonction « ecdf » de la librairie « stepfun »).
11. En utilisant la fonction « boxplot », représenter l'indice de Quêtelet en fonction du sexe.
12. Construire sur un seul graphique les nuages de points de l'indice de Quêtelet en fonction de la taille, en fonction du poids, pour chaque sexe. Commenter.
13. Représenter, sur un seul graphique, le nuage de points de l'indice de Quêtelet en spécifiant hommes et femmes. Placer les bornes des critères de l'O.M.S. Commenter.

Exercice 2 X suit une loi binomiale de paramètres $(50, 1/3)$. Calculer la probabilité des événements suivants

1. $\{X = 1\}$;
2. $\{X \leq 5\}$;
3. $\{X \geq 15\}$;
4. $\{X \notin \{15, 3, 4, 10\}\}$.

Exercice 3 Soit X une variable aléatoire qui suit une loi normale standard $\mathcal{N}(0, 1)$.

1. Calculer la probabilité des événements suivants : $\{X \leq 1\}$; $\{X \geq 26\}$; $\{0.5 < X \leq 1\}$.
2. Calculer le quantile d'ordre $a = 0.75$, c'est-à-dire la valeur de x telle que $P(X \leq x) = a$.
3. Représenter graphiquement la densité et la fonction de répartition de la loi de X .
4. Simuler un échantillon $(x_1; \dots; x_{100})$ de taille $n = 100$ suivant la loi de X .
5. Représenter les valeurs de l'échantillon simulé.
6. Créer une liste qui contient la moyenne empirique $\frac{1}{100} \sum_{i=1}^{100} x_i$, le minimum de (x_1, \dots, x_{100}) et le maximum de (x_1, \dots, x_{100}) .