

# Contrôle continu statistiques numero : 1

Nom :  
Prénom:

## Présentation des données

Le fichier OSA.csv décrit 103 patients venus consulter pour des troubles du sommeil.  
Les variables disponibles sont les suivantes :

- age : l'âge des patients
- sexe : le sexe des patients 1 Homme 2 Femme
- taille : la taille en cm
- poids : le poids en kg
- hypercholesterolemie : la présence d'hypercholestérolémie (1)
- IAH : l'IAH par heure (nombres d'apnées et/ou hypopnées par heure)
- refluxGastro : reflux gastro-œsophagien (1)
- depression : l'existence d'antécédents de dépression (1)
- bmi : L'IMC en catégorie (Normal < 25, overweight ]25 ; 30], obese moderate ]30 ; 35], obese severe ]35 ; 40], obese very severe > 40)
- asthme : l'existence d'antécédents d'asthme (1)
- bcpo : l'existence d'antécédents de Bronchopneumopathie chronique obstructive (1)
- hypertensionArterielle : l'existence d'antécédents d'hypertension artérielle (1)
- hypercholesterolemie : l'existence d'antécédents d'hypercholesterolemie (1)
- glaucome : l'existence d'antécédents de glaucome (1)
- arteriopathie : l'existence d'antécédents d'arteriopathie (1)
- Depression : L'échelle de dépression de pichot (0-13) 0 : non déprimé ,13 : très déprimé
- Epworth : L'échelle de somnolence d'Epworth (0-24) 0 : non somnolent ,24 : très somnolent

Les variables quantitatives sont toutes potentiellement « variables d'intérêts » ; les variables catégorielles vont servir à définir les sous populations.

Une réponse ne sera considérée comme correcte que si vous répondez avec la/les commande(s) R adaptées

ainsi que la/les valeur(s) numérique(s) arrondie(s) à 2 chiffres significatifs

Attention, pour chaque test il est obligatoire de présenter les données (Vérifier la présence d'outliers), définir les hypothèses, les conditions d'application et de conclure. Précisez nous lorsque vous faites un Student si la statistique suit une loi Normale ou une loi de Student. Quand cela n'est pas précisé les tests statistiques sont à réaliser au seuil  $\alpha = 5 \%$ .

Dans le cas où une donnée semble fausse, il est fortement recommandé de retirer la ligne complète.

Le sujet est volontairement très long afin de vous laisser le droit à l'erreur et de choisir vos questions. La notation tient compte de cette longueur, **pas de panique**.

### Question 1 :

Calculer l'écart interquartile de la colonne IAH uniquement pour les lignes où la colonne asthme prend la valeur 0

### Question 2 :

Calculer le quantile 80 % de la colonne poids uniquement pour les lignes où la colonne depression prend la valeur 1

### Question 3 :

Calculer l'intervalle de confiance à 98 % de la moyenne de la colonne Epworth uniquement pour les lignes où la colonne hypertensionArterielle prend la valeur 0

### Question 4 :

Tester si la variable hypercholesterolemie est distribuée selon les probabilités suivantes : 0 : 0.648 ; 1 : 0.352

### Question 5 :

Tester si la colonne age est différente selon les modalités de la variable bmi

### Question 6 :

Tester si la colonne taille uniquement pour les individus ayant la modalité 1 dans la colonne sexe est différente de 180.4

### **Question 7 :**

Tester si la colonne IAH uniquement pour les individus ayant la modalité 0 dans la colonne asthme est distribuée selon une loi normale

### **Question 8 :**

Tester si les variables sexe et arteriopathie sont indépendantes.

### **Question 9 :**

Tester si les variables refluxGastro et glaucome sont indépendantes.

### **Question 10 :**

Tester si la variable depression est distribuée selon les probabilités suivantes : 0 : 0.643 ; 1 : 0.357

### **Question 11 :**

Tester si la colonne IAH uniquement pour les individus ayant la modalité 0 dans la colonne refluxGastro est différente de 26.95

### **Question 12 :**

Tester si la colonne age pour les individus ayant la modalité 0 dans la colonne glaucome est différente ( de la colonne age) pour les individus avec la modalité 1 dans la colonne glaucome

### **Question 13 :**

Tester si la colonne Depression uniquement pour les individus ayant la modalité 1 dans la colonne sexe est distribuée selon une loi normale

Question bonus :

### **Question bonus 1 :**

Présenter les effectifs par classe pour l'IMC catégorisé.

### **Question bonus 2 :**

Calculer le mode pour une variable qualitative de votre choix.