
Colles série 2 : Statistiques descriptives pour des variables quantitatives

Les cinq premiers exercices portent sur le fichier `mtcars` qui fait partie de la librairie de base de R. Il suffit d'exécuter `mtcars` pour voir défiler le fichier. Un début de script est disponible pour une première description des données et quelques exemples d'extraction d'échantillons, dans `intro-colles-serie2.R`

Sujet 1:

1. Extraire les échantillons de `mpg` pour les véhicules automatiques et pour les véhicules manuels et les nommer resp. `mpga` et `mpgm`. En donner les résumés numériques standards (y compris l'écart-type) et les interpréter.
2. Représenter sur un même graphique les boxplot de chaque échantillon (avec titre et label sous chaque boxplot) et les comparer.

Sujet 2:

1. Représenter la répartition de la variable `mpg` en indiquant un titre et des légendes sur les deux axes. Y ajouter une droite verticale en rouge qui passe par la moyenne observée.
2. Rajouter sur le précédent graphe le tracé de la densité d'une loi normale en vert dont on choisira au mieux les paramètres. Commentaire.

Sujet 3:

1. Extraire les échantillons de `qsec` pour les véhicules automatiques et pour les véhicules manuels et les nommer resp. `qseca` et `qsecm`. En donner les résumés numériques standards (y compris l'écart-type) et les interpréter.
2. Représenter sur un même graphique les boxplot de chaque échantillon (avec titre et label sous chaque boxplot) et les comparer.

Sujet 4:

1. Représenter la répartition de la variable `qsec` en indiquant un titre et des légendes sur les deux axes. Y ajouter une droite verticale en rouge qui passe par la moyenne observée.
2. Rajouter sur le précédent graphe le tracé de la densité d'une loi normale en vert dont on choisira au mieux les paramètres. Commentaire.

Sujet 5:

1. Extraire les échantillons de `cyl` pour les véhicules automatiques et pour les véhicules manuels et les nommer resp. `cyla` et `cylm`. En donner les résumés numériques et commenter
2. Donner le tableau des répartitions conditionnelles de `cyl` selon `am`. Les représenter côte à côte avec deux diagrammes en barres. Paraissent-elles semblables ? Qu'en conclure ?

Sujet 6:

1. Fabriquer un échantillon de n tirages indépendants d'une variable X avec `sample(c(1,2,2,3,3,3,4,4,4,4),n,replace=T)->x` en ayant préalablement affecté à n la valeur 10. Calculer les effectifs observés pour chaque modalité rencontrée et représenter les fréquences avec un diagramme en barres en imposant l'intervalle $[0,0.45]$ sur l'axe des ordonnées.
2. Ajouter une horizontale en rouge qui passe par 0.4 et une horizontale en vert qui passe par 0.1. Exécuter plusieurs fois le script complet et observer le graphique. Quelle est selon vous la loi de X (on pourra augmenter n à 100 puis 1000 pour voir)