

Modélisation statistique : TP 3

Avant de commencer,

- récupérer les données "diet" sur ma page web
- créer un répertoire dans vos documents pour ce TP, y mettre tous les fichiers associés.

Exercice 1

1. Charger les données. Vérifier le type des variables. Faire les changements nécessaires.
2. Créer une variable `weight.loss` qui indique la perte de poids.
3. Visualiser les relations entre la perte de poids et le régime. Puis entre la perte de poids, le régime et le genre
4. Via `dplyr`, calculer la moyenne des pertes de poids pour les différents régimes.

Exercice 2

Nous voulons répondre à la question : la perte de poids est-elle différente suivant le type de régime ?

1. Proposer un premier modèle linéaire pour répondre à la question.
2. Dans l'écriture classique $Y = X\beta + \varepsilon$. Quelle matrice X R a-t-il utilisée, (donner ses premières lignes) ?
3. A partir du `summary` du modèle linéaire, donner la moyenne de la perte de poids pour les différents régimes.
4. Quels diagnostics du modèle linéaire faut-il faire dans ce cas ? Faites les et analysez les.
5. La perte de poids est-elle différente suivant le type de régime ? Quel test utilisez vous ?
6. Y a-t-il une différence entre les régimes 1 et 2 ? Quel test utilisez vous ?
7. Créer une nouvelle variable `Diet2` qui vaut 1 si `Diet` vaut 1 ou 2 et 0 sinon. Refaire une analyse de la variance.

Exercice 3

1. Refaire l'étude pour le genre.
2. En ajoutant une interaction régime/genre, répondre à la question : les régimes ont-ils la même influence sur la perte de poids suivant le genre ?