

**TP – ANOVA à un facteur****Exercice 1.**

On considère la base de données **coagulation** du package **faraway**. Cette base contient deux variables : **coag**, qui donne le temps coagulation en secondes, et **diet**, qui donne le type de régime - A, B, C ou D. Les mesures ont été effectués sur 24 animaux.

1. Mettre en place une analyse de variance pour étudier l'effet du régime sur les temps de coagulation à l'aide de la fonction **aov**. Vérifier la pertinence du modèle.
2. Vérifier les hypothèses du modèle à l'aide des représentations graphiques des résidus. Vérifier également l'hypothèse de normalité des résidus à l'aide d'un test de Shapiro-Wilk et l'hypothèse d'homogénéité de la variance à l'aide d'un test de Bartlett (fonction **bartlett.test**).
3. Faire les tests de comparaison 2 à 2 avec une correction de la multiplicité adaptée. On utilisera la fonction **TukeyHSD** ou **lsmeans**, issue du paquet homonyme. Représenter graphiquement les intervalles de confiance des temps de coagulation moyens en fonction des régimes.
4. Proposer un regroupement des types de régime en deux catégories. On pourra utiliser la fonction **cld** du paquet **lsmeans**. Tester la validité de ce regroupement.

**Exercice 2.**

Récupérer les données **recovery** du paquet **multcomp**. Une entreprise a testé des nouveaux types de couverture chirurgicales aidant à maintenir la température corporelle. Elle a mesuré le temps de récupération après une opération (variable **minutes**) en fonction du type de couverture utilisée (variable **blanket**). Mettre en place un modèle d'analyse de variance afin de tester les différences entre les couvertures étudiées et 'b0', le standard utilisé. Quelle couverture conseilleriez-vous ?

**Exercice 3.**

On considère la base de données **anorexia** du package **MASS**.

Il s'agit de données d'un essai comparant différents traitements de l'anorexie chez les adolescentes. Dans cet essai, des jeunes filles ont été incluses et réparties dans 3 groupes de traitement. On a mesuré le poids de chaque jeune fille avant et après l'essai. Le but de l'étude est de savoir s'il y a une différence entre les trois groupes, et si un des traitements est plus efficace que les autres.

1. On s'intéresse à la différence de poids avant et après traitement (sans tenir compte des groupes de traitement).
  - (a) Quel test permet de comparer ces deux variables ? Donner son nom, l'hypothèse nulle et alternative.

- (b) Mettre en place ce test. Tester s'il y a une augmentation en moyenne du poids. Y-a-t-il un effet traitement (quelque soit le groupe de traitement) ?
- 2. On va maintenant comparer les différences entre les groupes de traitement. Dans un premier temps, on regarde successivement s'il y avait une différence entre les groupes avant l'étude puis s'il y a une différence après l'étude.
  - (a) Calculer les statistiques descriptives du poids avant et après l'étude dans chaque groupe. Proposer une représentation graphique.
  - (b) Quel test permet de tester s'il y a une différence de poids entre les 3 groupes avant traitement ?
  - (c) Mettre en place ce test, ainsi que celui testant une différence de poids entre les 3 groupes après traitement.
  - (d) Conclure.
- 3. On étudie maintenant si les groupes ont influencé la prise de poids.
  - (a) Créer la variable prise de poids qui vaut la différence entre le poids après traitement et le poids avant traitement.
  - (b) Calculer les statistiques descriptives de cette variable pour chaque groupe. Proposer une représentation graphique. Commenter.
  - (c) On cherche à savoir si le poids a augmenté dans le groupe contrôle.
    - i. Quel test permet de savoir s'il y a une augmentation significative du poids dans ce groupe ? Donner le nom du test, les hypothèses nulle et alternative.
    - ii. Réaliser ce test. Commenter le résultat.
  - (d) Faire de même pour le groupe *traitement du comportement*, puis le groupe *traitement familial*.
  - (e) On veut maintenant comparer l'efficacité des 3 traitements simultanément. Quel test permet de réaliser cette comparaison ? Rappeler l'hypothèse nulle, alternative et les conditions d'application du test.
  - (f) Le mettre en place sous R. Commenter.
  - (g) On réalise maintenant des tests de comparaison 2 à 2 de ces 3 groupes. Quelle est la précaution à prendre lorsqu'on fait des comparaisons multiples ?
    - i. Les deux traitements (*traitement du comportement* et *traitement familial*) sont-ils plus efficaces que le groupe *contrôle* ?
    - ii. Quel est le traitement le plus efficace ?