

Exercice d'auto-évaluation - Session 2

Un groupe industriel commercialisant du café souhaite comparer les cafés provenant de différents sites de production à partir de leur profil de composition physico-chimique : teneurs en acide chlorogénique (CGA), caféine (Cafeine), matières grasses (Fat), trigonelline (Trigonelline) et matières sèches (DM).

Les données sont fournies dans le fichier `coffee.txt`. Un script R important ces données est également à disposition. Dans un premier temps, on se concentre sur la question suivante : les taux de matières sèches sont-ils différents selon le site de production ?

1. Proposez un modèle statistique adapté à cette question. Vous donnerez le nom courant de ce modèle.
2. Exprimez la question sous forme de test d'hypothèses dont vous donnerez les hypothèses H_0 et H_1 .
3. Estimez les paramètres du modèle de la question 1.
4. Que vaut la somme des carrés des résidus de ce modèle ?
5. Que vaut la somme des carrés des résidus du modèle sous l'hypothèse nulle ?
6. Quelle statistique de test permet de répondre à la question ? Calculez cette statistique de test.
7. Quelle est la distribution de cette statistique de test sous l'hypothèse H_0 ?
8. Les teneurs moyennes en matières sèches sont-elles significativement différentes au seuil $\alpha = 0.05$?