

Exercice d'auto-évaluation - Session 5

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 variations de taux de matières grasses dépendent-elles de celles des teneurs en trigonelline ?

1. Dans cette problématique, quelle est la variable réponse ? Quelle est la variable explicative ? Vous donnerez aussi la nature (quantitative ou catégorielle) de ces variables.
2. Proposez un modèle statistique adapté à cette question. Vous donnerez le nom courant de ce modèle.
3. Exprimez la question sous forme de test d'hypothèses dont vous donnerez les hypothèses H_0 et H_1 .
4. Quelle statistique de test permet de répondre à la question ? Calculez cette statistique de test.
5. Quelle est la distribution de cette statistique de test sous l'hypothèse H_0 ?
6. Les variations de taux de matières grasses dépendent-elles de celles des teneurs en trigonelline de manière significative au seuil $\alpha = 0.05$?
7. Selon vous, peut-on prédire avec précision le taux de matières grasses à partir de la teneur en trigonelline ? Argumentez votre réponse.
8. Si l'on sait qu'un café a un taux de trigonelline valant 0.8, donnez un intervalle de confiance de niveau 0.95 pour son taux de matières grasses.