

# ANOVA à 1 facteur

## 1 Données et objectifs statistiques

En Juin 1986, une étude comparant la valeur nutritionnelle des différents types de hotdogs des principales marques américaines a été publiée dans *Consumer Reports*<sup>1</sup>. Dans l'étude, les hotdogs de 54<sup>2</sup> marques ont été analysés. On a mesuré sur chacun d'eux le nombre de calories par hotdog, la quantité de sodium en milligrammes par hotdog.

Une des questions de l'étude est de savoir si ces caractéristiques nutritionnelles sont liées au type de viande utilisée dans le hotdog. Et si c'est le cas, quel est le type de viande donnant les hotdogs aux meilleures caractéristiques nutritionnelles. Les différents types de viandes observées sont : le boeuf (codé 1), mélange (principalement du porc et du boeuf mais aussi de la viande de volaille jusqu'à 15%) (codé 2), de la volaille (codé 3) ou du soja (codé 4).

## 2 Importation et mise en forme des données

1. Spécifiez dans R votre répertoire de travail (`Session - Set working directory - choose directory`).

2. Chargez les données dans la variable `tab` à l'aide de la commande

```
tab <- read.table("hotdogs.txt",header=T)
```

3. Vérifiez que l'importation s'est effectuée correctement en utilisant les fonctions `head()`.

4. À l'aide de l'instruction `str(tab)` :

(a) vérifiez le nombre d'observations et le nombre de variables,

(b) précisez la nature de l'ensemble des variables et dans le cas de variables qualitatives, précisez-en aussi le nombre de modalités.

5. Vérifiez au moyen de l'instruction `table(tab$Type)` le nombre d'observations par type. Taper la commande `tab=tab[-which(tab$Type==4),]`, exécutez-la. Que change-t-elle? Pourquoi l'a-t-on effectuée?

## 3 Statistiques descriptives

6– Résumez l'information contenue dans les variables `Calories` et `Sodium`.

7– Croisez les variables `Calories` et `Sodium` avec la variable `tab$Type` afin de produire un résumé descriptif par type de viande.

8– Commentez les résultats obtenus au regard des questions posées dans la section 1.

---

1. Consumer Reports, June 1986, pp. 366-367.

2. Pour des raisons pédagogiques, un individu statistique artificiel a été ajouté au jeu de données.

## 4 Etude du nombre de Calories - Anova 1 facteur

9 – Ecrire le modèle associé à l'expérience.

10 – Exécutez l'instruction :

```
mod1=lm(Calories~Type, data=tab)
par(mfrow=c(2,2))
plot(mod1)
```

Les hypothèses du modèle sont-elles vérifiées?

11 – Exécutez l'instruction :

```
anova(mod1)
```

Dans le tableau ainsi obtenu :

- A quoi correspondent les différentes cases du tableau?
- Quelles sont les hypothèses (H0 et H1) du test effectué dans ce tableau?
- Quel est le résultat de ce test et comment l'interprétez-vous?

12 – Exécutez l'instruction :

```
summary(mod1)
```

- Quelles sont les valeurs des coefficients estimés? Comment les interprétez-vous?
- Quelle est la contrainte choisie dans ce modèle?
- Pour chacune des modalités du facteur Type, quel est le nombre de calories prédit par le modèle?
- Quelle est la valeur du coefficient de détermination du modèle? Comment l'interprétez-vous?
- Quel est l'écart-type des résidus estimés par le modèle?
- A quel test correspond la dernière ligne fournie par cette commande?

13 – Exécutez l'instruction :

```
library(lsmeans)
lsmeans(mod1, pairwise~Type, data=tab, adjust="bonferroni")
```

- A quoi correspond l'argument `adjust="bonferroni"`? A quoi sert-il?
- Quel est le type de hotdog le moins calorique?
- Est-il significativement moins calorique que les autres?

## 5 Etude de la quantité de sodium

Recommencez la même étude que pour les Calories (questions 9 à 15). Quel est le type de hotdog le moins salé?