

# Tests d'hypothèses

23 septembre 2020

Ce laboratoire doit être remis le **30 septembre à 17h sur Moodle**. Dans votre réponse pour chaque question, veuillez inclure le code R utilisé (s'il y a lieu) et les résultats obtenus.

## 1. Concentration d'ozone dans trois jardins

Pour cet exercice, nous utiliserons le tableau de données `gardens.csv`, qui provient du manuel *Statistics: An Introduction Using R* de Michael Crawley. Ces données représentent les concentrations d'ozone (en parties par 100 millions ou pphm) mesurées dans trois jardins (A, B et C) lors de différentes journées.

```
gardens <- read.csv("gardens.csv")
head(gardens)
```

```
##   Ozone Garden
## 1     3      A
## 2     4      A
## 3     4      A
## 4     3      A
## 5     2      A
## 6     3      A
```

- Visualisez les mesures d'ozone en fonction du jardin avec un graphique de type `geom_jitter` dans *ggplot2*. Quel est l'avantage de ce type de graphique ici, comparé à `geom_point` ou `geom_boxplot`?
- Quelle est la moyenne et l'écart-type de la concentration d'ozone dans chaque jardin? Est-ce que la moyenne est une bonne indicatrice de la valeur "typique" dans chaque jardin?
- À partir de ces données, testez l'hypothèse nulle selon laquelle les jardins A et B reçoivent la même concentration d'ozone en moyenne. Quel est votre estimé de la différence entre les moyennes et son intervalle de confiance à 99%? Est-ce que ce test donne une bonne idée de la différence entre les deux jardins? Expliquez votre réponse.
- Répétez l'exercice précédent pour l'hypothèse nulle selon laquelle les jardins A et C reçoivent la même concentration d'ozone en moyenne. Commentez sur la différence entre ce résultat et le résultat précédent.

## 2. Effet des feuilles mortes sur l'apport en azote de semis

Le tableau de données `nconc.csv` présente les résultats d'une expérience (fictive) visant à estimer l'effet des feuilles mortes sur l'apport en azote de semis.

```
nconc <- read.csv("nconc.csv")
nconc
```

```
##   plot  litter no_litter
## 1     1 1.859543 1.8073724
## 2     2 1.461364 0.7367826
## 3     3 1.488136 1.6332546
## 4     4 1.325093 1.1615338
```

```
## 5    5 1.600666 0.9864743
## 6    6 2.038625 1.9011473
## 7    7 1.788214 1.3162220
## 8    8 1.994081 1.7849742
```

Huit placettes (**plot**) ont été délimitées dans une forêt et divisées en deux moitiés. Dans chaque placette, une moitié (choisie au hasard) a reçu le traitement qui consistait à enlever systématiquement les feuilles mortes au sol. Les deux dernières colonnes du tableau montrent la concentration d'azote (en % de la biomasse) pour les semis de la moitié non-traitée (**litter**) et de la moitié traitée (**no\_litter**).

- a) Quel est l'avantage d'avoir effectué l'expérience de cette façon plutôt que d'avoir enlevé complètement les feuilles mortes de quatre placettes sur huit? Quel type de test  $t$  est approprié ici pour déterminer l'effet du traitement sur la concentration moyenne d'azote des semis?
- b) Effectuez le test  $t$  choisi et déterminez si le traitement a un effet significatif à un seuil  $\alpha = 0.05$ . Quel est l'effet moyen estimé (n'oubliez pas d'interpréter le signe de la différence) et son intervalle de confiance?