

# Travaux Pratiques de Biométrie 3

*Benoît Simon-Bouhet*

2019-03-02

## Table des matières

<b>1</b>	<b>Préambule</b>	<b>1</b>
<b>2</b>	<b>Introduction</b>	<b>2</b>
<b>3</b>	<b>Séance 1 : statistiques descriptives et tests d'hypothèses</b>	<b>3</b>
3.1	Packages et données . . . . .	3
3.2	Statistiques descriptives . . . . .	4
3.3	Comparaison de moyennes et conditions d'application . . . . .	4
3.4	Comparaison de la moyenne d'une population à une valeur théorique . . . . .	4
3.5	Comparaison de la moyenne de 2 populations : données appariées . . . . .	4
3.6	Comparaison de la moyenne de 2 populations : données indépendantes . . . . .	4
3.7	Tests bilatéraux et unilatéraux . . . . .	4
<b>4</b>	<b>Séance 2 : analyse de variance</b>	<b>4</b>
<b>5</b>	<b>Séance 3 : corrélations et régressions</b>	<b>4</b>
<b>6</b>	<b>Séance 4 : applications et corrections</b>	<b>4</b>

## 1 Préambule

Ce livre contient l'ensemble du matériel (contenus, exemples, exercices...) nécessaire à la réalisation des travaux pratiques et TEA de biométrie 3. Ces travaux pratiques ont un seul objectif principal : vous permettre de mettre en œuvre, dans RStudio les méthodes statistiques découvertes en cours magistral et en TD de biométrie 2 (au semestre précédent) et en biométrie 3 depuis début janvier.

Je considère qu'à ce stade, vous devez être à l'aise dans RStudio pour effectuer les tâches suivantes :

1. Importer des jeux de données dans RStudio
2. Manipuler des tableaux de données avec `tidyr` pour les mettre dans un format permettant les analyses statistiques et les représentations graphiques
3. Faire des graphiques exploratoires avec `ggplot2` pour visualiser des données
4. Filtrer des lignes, sélectionner des colonnes, trier, créer de nouvelles variables et calculer des résumés des données avec les fonction `filter()`, `select()`, `arrange()`, `mutate()`, `summarise()` et `group_by()` du package `dplyr`

5. Utiliser le pipe %>% afin d'enchaîner plusieurs commandes
6. Créer des scripts parlants contenant des commandes et des commentaires utiles
7. Spécifier/modifier votre répertoire de travail
8. Installer des packages additionnels.

Si vous pensez avoir besoin de rappels sur ces notions, je vous encourage vivement à consulter [le livre en ligne](#) dédié aux travaux pratiques de Biométrie 2 pour vous rafraîchir la mémoire.

L'organisation des TP et TEA de biométrie 3 sera la suivante :

- Séance 1 : 1h30 de TP suivie d'une séance de 1h30 de TEA. Rappels concernant les statistiques descriptives et les visualisations graphiques utiles pour démêler la complexité de certains jeux de données. Comparaisons (paramétriques et non paramétriques) de la moyenne de 2 populations.
- Séance 2 : 1h30 de TP suivie d'une séance de 1h30 de TEA. Comparaisons (paramétriques et non paramétriques) la moyenne de plus de 2 populations : analyse de variance, hypothèses et conditions d'application.
- Séance 3 : 1h30 de TP suivie d'une séance de 1h30 de TEA. Étude de la liaison entre 2 variables. Corrélation (paramétrique et non paramétrique) et régression linéaire. Tests d'hypothèses, estimation et conditions d'application.
- Séance 4 : 1h30 de TP. Exercices d'application et corrections en guise de préparation pour l'examen.

## 2 Introduction

Sur votre disque dur, créez un dossier nommé "Biometrie3" (sans accent, sans espace). Au début de chaque nouvelle séance de TP, vous devrez ensuite effectuer les opérations suivantes :

1. Créez, dans votre dossier "Biometrie3", un sous-dossier nommé "TP\_1", "TP\_2", etc.
2. Téléchargez les fichiers utiles disponibles sur l'ENT et placez-les dans le dossier du TP correspondant.
3. Lancez RStudio.
4. Dans l'onglet "Files" de RStudio, naviguez jusqu'au sous-dossier "TP\_X" que vous venez de créer et indiquez à RStudio qu'il s'agit de votre répertoire de travail. Si vous ne savez plus comment faire, consultez [la section 2.2.2](#) du livre en ligne de Biométrie 2. Si votre répertoire de travail a été correctement spécifié, vous devriez constater qu'une commande ressemblant à ceci est apparue dans la console de RStudio :

```
setwd("C:/...../Biometrie3/TP_X")
```

5. Dans la console, tapez :

```
list.files()
```

La liste des fichiers contenus dans votre répertoire de travail (donc le nom des fichiers que vous avez téléchargé sur l'ENT) devrait apparaître dans la console. Si ce n'est pas le cas, recommencez depuis le début. Vous pouvez également vérifier à tout moment si le répertoire de travail utilisé par RStudio est bien celui que vous pensez en tapant :

```
getwd()
```

6. Créez un nouveau script dans votre répertoire de travail et sauvegardez-le. Si vous ne savez plus comment faire, consultez [la section 2.2.3](#) du livre en ligne de Biométrie 2.
7. Dans l'onglet "History" de RStudio, cliquez sur la commande commençant par `setwd()` puis cliquez sur le bouton "To source" (une flèche verte dirigée vers la gauche). Cela a pour effet de copier dans votre script la commande permettant de spécifier le répertoire de travail correct. Ainsi, lors de votre prochaine session de travail, vous n'aurez pas besoin de spécifier manuellement quel est votre répertoire de travail comme nous l'avons fait à l'étape 4 ci-dessus : il vous suffira d'ouvrir votre script et d'envoyer cette commande dans la console en pressant les touches `ctrl` + `Entrée`.
8. N'oubliez pas de sauvegarder votre script très régulièrement et d'y ajouter autant de commentaires que nécessaire avec le symbole `#`.

Si vous suivez rigoureusement ces étapes, vous devriez être dans la situation idéale pour commencer à travailler efficacement dans RStudio. Avec un minimum d'habitude, mettre tout ça en place ne devrait pas vous demander plus de 2 ou 3 minutes. À partir de maintenant, toutes vos analyses et commentaires doivent figurer dans vos scripts.

## 3 Séance 1 : statistiques descriptives et tests d'hypothèses

### 3.1 Packages et données

Pour chacune des 4 séances de travaux pratiques (et TEA) qui viennent, vous aurez besoin d'utiliser des packages spécifiques et d'importer des données depuis des fichiers externes fournis.

Les packages dont vous aurez besoin pour cette séance sont les suivants :

```
library(tidyverse)
library(skimr)
```

Si ces commandes (que vous devez taper dans vos scripts avant de les exécuter dans la console de RStudio) renvoient des messages d'erreur, c'est que les packages que vous essayez de charger en mémoire ne sont pas installés sur votre ordinateur. Il vous faudra alors installer les packages manquants avec la fonction :

```
install.packages("nom_du_package")
```

Comme d'habitude, si tout ça est un peu flou pour vous, relisez [la section 2.3](#) du livre de biométrie 2 disponible en ligne.

### **3.2 Statistiques descriptives**

### **3.3 Comparaison de moyennes et conditions d'application**

### **3.4 Comparaison de la moyenne d'une population à une valeur théorique**

#### **3.4.1 Le test paramétrique**

#### **3.4.2 Le test non paramétrique**

### **3.5 Comparaison de la moyenne de 2 populations : données appariées**

#### **3.5.1 Le test paramétrique**

#### **3.5.2 Le test non paramétrique**

### **3.6 Comparaison de la moyenne de 2 populations : données indépendantes**

#### **3.6.1 Le test paramétrique**

#### **3.6.2 Le test non paramétrique**

### **3.7 Tests bilatéraux et unilatéraux**

#### **3.7.1 Le test paramétrique**

#### **3.7.2 Le test non paramétrique**

## **4 Séance 2 : analyse de variance**

## **5 Séance 3 : corrélations et régressions**

## **6 Séance 4 : applications et corrections**