

# Introduction et organisation du cours

Andreas Zanzoni ([orcid.org/0000-0002-4818-6161](https://orcid.org/0000-0002-4818-6161))

Jacques van Helden ([orcid.org/0000-0002-8799-8584](https://orcid.org/0000-0002-8799-8584))

Lou BERGOGNE ([orcid.org/0009-0004-9409-0154](https://orcid.org/0009-0004-9409-0154))



Andreas Zanzoni

<https://orcid.org/0000-0002-4818-6161>



Jacques van Helden

<https://orcid.org/0000-0002-8799-8584>



Lou BERGOGNE

[orcid.org/0009-0004-9409-0154](https://orcid.org/0009-0004-9409-0154)

# Table des matières

---

1. Objectifs pédagogiques
2. Logistique (matériel + logiciels)
3. Organisation du cours
4. Diagnostic de vos connaissances préalables
5. Contenus couverts par le cours
6. Modalités de contrôle des connaissances
7. Accès aux supports: Ametice + github

## Objectifs pédagogiques

---

- Appréhender des données correspondant aux pratiques actuelles de la biologie (données tabulaires, résultats de technologies “omiques”)
- Formuler des questions scientifiques concernant ces données
- Connaître les principes du raisonnement statistique (modélisation statistique)
- Choisir les approches d’analyse statistique appropriées pour répondre à ces questions (visualisation, découverte réduction des données, tests d’hypothèses, ...)
- Comprendre les fondements théoriques sous-jacents à ces analyses (échantillonnage, estimation, lois de probabilités, formulation des hypothèses, ...)
- Implémenter des scripts pour mettre en oeuvre les analyses (langage R)
- Interpréter les résultats en terme de signification statistique et biologique

# Logistique

---

## Matériel

Un ordinateur portable sera mis à disposition des étudiants qui n'en ont pas.

## Logiciel

Avant la première séance, nous vous demandons d'installer sur votre ordinateur les logiciels **R** et **RStudio**.

Site d'installation: <https://posit.co/download/rstudio-desktop/>

- Commencez par installer R (langage spécialisé pour l'analyse statistique)
  - Suivez le lien **DOWNLOAD AND INSTALL R** (il vous redirige vers [cran.rstudio.com](http://cran.rstudio.com))
  - Installez la version 4.5.1 de votre système opérateur (Mac, Windows, Linux)
- Installez ensuite RStudio (environnement spécialisé pour l'utilisation du langage R)

En cas de problème, contactez-nous au plus tard le jeudi 6 novembre pour que nous puissions vous dépanner.

## Organisation du module d'enseignement

---

- Séances de cours-TP
- Alternance de théorie et de pratique

# Sondage : diagnostic de vos connaissances préalables

## 5. Avez-vous déjà suivi des cours de probabilités et statistiques dans votre parcours universitaire ?

Oui



43 répondants

41 votes

Non



2 votes

## 8. Avez-vous utilisé des outils informatiques pour les travaux pratiques de probabilités et statistiques?

Oui



42 répondants

32 votes

Non



10 votes

## 9. Si oui, cochez les outils utilisés

R / RStudio



42 répondants

32 votes

Python



6 votes

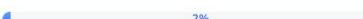
Excel



21 votes

1 vote

Jamovi



0 votes

0 votes

SPSS



0 votes

0 votes

Autres (precisez dans les commentaires)



4 votes

4 votes

## R / RStudio - Commentaires 9

tableau de données brutes pour analyse via commande R création de graphique

R studio + R markdown + Quarto

À peine employé

Traitement de données + production de graphiques

Traitement donnée, production de graphique

Analyse à Composantes Principales

En statistiques/probabilités pour la génomique des populations

Dans l'autre question j'ai envoyé la réponse incorrecte. J'ai utilisé R de manière basique pour calculer matrices ou boucles

utilisé r studio en L2 SV à saint charles

## Python - Commentaires 2

Pas mal employé

En cours cette années en programmation

## Excel - Commentaires 5

utilisation basique, juste pour visualiser les données

En autonomie

Mais plus souvent google sheet

Graphiques, traitement de données

Graphique

## Autres (precisez dans les commentaires) - Commentaires 3

rien

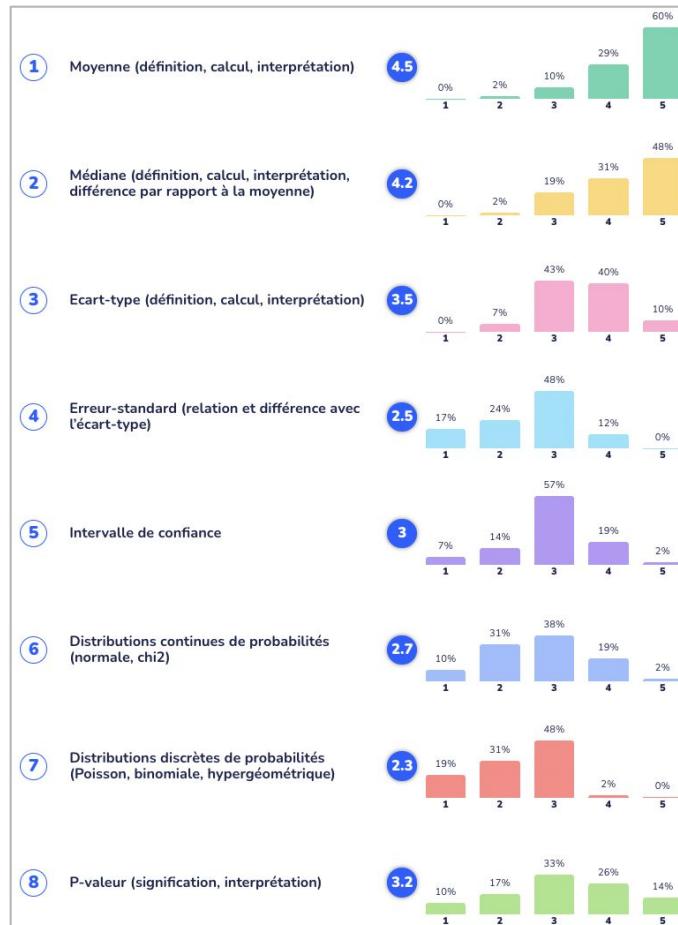
On a pas utilisé des outils informatiques

aucun

# Sondage : diagnostic de vos connaissances préalables

Evaluez votre maîtrise des concepts suivants sur une échelle de 1 à 5, avec les niveaux suivants :

1. Concept inconnu
2. Concept appris, mais mal compris ou retenu
3. Concept appris et mis en pratique, mais un petit rappel serait utile
4. Bonne compréhension du concept et capacité à le mettre en oeuvre lors d'une analyse de données
5. Maîtrise suffisante pour mettre en oeuvre et pour expliquer le concept et sa mise en oeuvre à un autre étudiant



1. Organisation du cours
2. Apprivoiser ses données: statistiques descriptives et exploratoires
3. Des populations aux échantillons et retour : échantillonnage et estimation
4. Tests de comparaison de moyennes
5. Tests d'ajustement
6. Tests d'indépendance
7. Tests d'indépendance / statistiques de classement
8. Révisions et préparation à l'examen

## Notes

- La structure des chapitres est susceptible d'évoluer
- Le temps consacré peut varier entre chapitres, il n'y a donc pas forcément une correspondance de 1 à 1 entre chapitres et séances

## Modalités de contrôle des connaissances

---

- Examen terminal sur table
- Documents et calculatrice ne seront pas nécessaires
- Les exercices couvriront l'ensemble des chapitres abordés pendant le cours

## Ametice

- <https://ametice.univ-amu.fr/course/view.php?id=140161>
- Interactions entre enseignants et étudiants
  - Annonces et forum
  - Sondages
- Supports pédagogiques (diaporamas, publications scientifiques)

## Dépôt github

- [https://github.com/ivanheld/proba-stat-bio\\_SBGAU14L](https://github.com/ivanheld/proba-stat-bio_SBGAU14L)
- Jeux de données à analyser
- Exemples de code