

Proposition de stage en biostatistique

Duré du stage : 2/3 mois

Maitre de stage : Nicola Kuczewski Centre de Recherche en Neurosciences de Lyon ; MCU Université Lyon1.

nicola.kuczewski@univ-lyon1.fr

Détermination de l'impact produit par l'élimination des « outliers » sur la probabilité d'erreurs de type I et de type II

L'élimination de l'échantillon expérimental des données supposées être aberrantes, les outliers, est une pratique qui est souvent utilisée pour réduire la probabilité d'erreurs statistiques lors de la comparaison entre échantillons. Il existe différentes méthodes de détection des outliers. L'objectif de ce stage est de comparer l'impact de plusieurs de ces méthodes sur la probabilité d'erreurs de type I et de type II, en fonction de la puissance statistique utilisée. Pour cela une approche empirique sera utilisée. Elle consiste à simuler des échantillonnages aléatoires des populations prédéfinies, i.e. pour lesquelles les distributions et la taille de l'effet sont connues. Cela permettra d'évaluer d'éventuelles déviations de l'occurrence des erreurs de type I et II, par rapport aux valeurs attendues, suite à l'élimination des outliers. Une connaissance des langages de programmation Python et /ou R est un prérequis important mais pas essentiel pour le bon déroulement du stage.