
Stage M2 – Analyse de mesures répétées en grande dimension

Description

L'équipe METRICS (Univ. Lille, CHU Lille) est une équipe de recherche pluri-disciplinaire ressemblant des cliniciens et des méthodologistes dans différents domaines, dont la biostatistique. Notre équipe a en particulier développé une expertise sur l'analyse statistique des expériences à haut débit, par exemple en transcriptomique ou protéomique (données omiques). Ces données présentent beaucoup plus de variables que d'individus. Dans ce contexte, les régressions pénalisées se sont avérées très utiles pour sélectionner des biomarqueurs à un temps donné. Dans le cadre de mesures répétées de données omiques sur 3-4 temps, la sélection des biomarqueurs reste encore un sujet de recherche à explorer.

L'objectif de ce stage est d'appliquer, sur des données réelles, des méthodes conceptuellement différentes pour l'analyse de données répétées en grande dimension :

- rJCLM - modèle de classe latente régularisé (Sun et al., 2019)
- analyse de corrélation canonique généralisée multivoie (A. Gloaguen et al., 2022)

Ce stage pourra être suivi d'une thèse de doctorat.

Missions

- Analyser des données réelles à l'aide du code R des auteurs associés aux publications précédentes
- Réaliser une revue bibliographique pour suggérer d'autres analyses afin de sélectionner des marqueurs dans une étude longitudinale d'expériences à haut débit

Compétences nécessaires

- Programmation R avancée
- Apprentissage statistique dans le contexte de la grande dimension
- Analyse de survie

Informations pratiques

Durée : 6 mois

Localisation : ULR2694 METRICS – Pôle recherche de la faculté de médecine – 1 place de Verdun, 59000 Lille

Contact

Envoyer CV, lettre de motivation et notes de Licence 3 et Master 1 à Génia Babykina evgeniya.babykina@univ-lille.fr et Guillemette Marot guillemette.marot@univ-lille.fr

Bibliographie

Gloaguen A, Philippe C, Frouin V, Gennari G, Dehaene-Lambertz G, Le Brusquet L, Tenenhaus A. Multiway generalized canonical correlation analysis. *Biostatistics*. 2022 Jan 13;23(1):240-256. doi: 10.1093/biostatistics/kxaa010. PMID: 32451525.

Sun J, Herazo-Maya JD, Molyneaux PL, Maher TM, Kaminski N, Zhao H. Regularized Latent Class Model for Joint Analysis of High-Dimensional Longitudinal Biomarkers and a Time-to-Event Outcome. *Biometrics*. 2019 Mar;75(1):69-77. doi: 10.1111/biom.12964. Epub 2018 Dec 5. PMID: 30178494.