

## Offre de stage de M2

# Modèles graphiques pour les extrêmes

**Laboratoires d'accueil** : IMAG et Inria Lemon, Montpellier

**Encadrants** : Marine Demangeot (Inria - Université Montpellier Paul Valéry), Nicolas Meyer (Inria - Université de Montpellier), Gwladys Toulemonde (Inria - Université de Montpellier)

**Durée** : 4 à 6 mois - début souhaité en mars/avril 2024

**Compétences recherchées** : Statistique/Probabilités

**Gratification** : selon les conditions actuelles de rémunération légale des stagiaires (environ 600 euros/mois)

### **Comment postuler ?**

Pour postuler ou si vous avez des questions, veuillez contacter Marine Demangeot ([marine.demangeot@umontpellier.fr](mailto:marine.demangeot@umontpellier.fr)) et/ou Nicolas Meyer ([nicolas.meyer@umontpellier.fr](mailto:nicolas.meyer@umontpellier.fr)) et/ou Gwladys Toulemonde ([gwladys.toulemonde@umontpellier.fr](mailto:gwladys.toulemonde@umontpellier.fr)) . Merci de nous envoyer un CV ainsi que les noms et adresses électroniques d'un ou deux contacts académiques.

### **Contexte scientifique**

La théorie des valeurs extrêmes propose un cadre asymptotique pour étudier les grandes valeurs d'un jeu de données, c'est-à-dire celles qui dépassent un certain seuil fixé. Dans un contexte multivarié, un enjeu crucial consiste à analyser la structure de dépendance extrême d'un vecteur aléatoire : si une composante de ce vecteur est grande, que peut-on dire des autres coordonnées ? Une démarche standard pour mettre en évidence la dépendance entre différentes coordonnées consiste à 1) regrouper les coordonnées qui ont un comportement extrême similaire en des petits groupes, appelés clusters, 2) étudier la corrélation des variables à l'intérieur de chaque groupe, 3) étudier les relations inter-groupes.

Récemment, des méthodes d'inférence de la dépendance extrême basées sur des modèles graphiques ont été proposées. De tels modèles constituent un outil efficace pour représenter une potentielle dépendance entre des coordonnées : chaque nœud du graphe correspond à une variable et une arête relie deux nœuds si ces deux variables sont corrélées, cette arête étant orientée en cas de causalité. Dans les articles récents sur le sujet, les modèles graphiques sont essentiellement développés pour les lois asymptotiques des dépassements de seuils (modèles de Pareto, Hüsler-Reiss, etc.).

### **Objectif de ce stage**

L'objectif de ce stage est d'étudier les différents modèles graphiques pour les extrêmes proposés dans la littérature, et de comprendre le lien entre les propriétés graphiques et probabilistes des modèles. Dans un deuxième temps, il s'agira d'étendre la notion de modèle graphique extrême à des modèles non-asymptotique, notamment le modèle gaussien avec marginales à queues lourdes. Étudier les propriétés graphiques conservées à la limite sera une étape-clé pour comprendre si les modèles asymptotiques peuvent être transposés au cadre non-asymptotique.

Une part importante du stage sera également consacrée à l'implémentation des modèles (sur R ou Python), et à leur application sur des données simulées ou réelles (par exemple des données de pluie).

Si le stage se déroule bien, une poursuite en thèse pourra être envisagée.

### **Références bibliographiques**

Engelke and Hitz, Graphical models for extremes (2020). *Journal of the Royal Statistical Society: Series B*, 82(4) 871 – 932

Meyer and Wintenberger, Discussion on "Graphical models for extremes" (2020). *Journal of the Royal Statistical Society: Series B*

Meyer and Wintenberger, Multivariate sparse clustering for extremes (2023). *Journal of the American Statistical Association*, in press

Engelke, Ivanovs, Strokorb, Graphical models for infinite measures with applications to extremes and Lévy processes, *arxiv: 2211.15769*

Röttger, F. Engelke, E. Zwiernik, P. Total positivity in multivariate extremes. *Annals of Statistics*, 51(3) 962 - 1004