

# Cinquièmes rencontres de statistique Avignon/Marseille:

## Programme détaillé

*9h30-10h00: accueil des participants*

- (1) 10h00-11h00: Catherine Matias (Université Pierre et Marie Curie, CNRS, Paris)

Titre: Statistical clustering of temporal networks through a dynamic stochastic block model (travail en collaboration avec V. Miele)

Résumé: Statistical node clustering in discrete time dynamic networks is an emerging field that raises many challenges. Here, we explore statistical properties and frequentist inference in a model that combines a stochastic block model (SBM) for its static part with independent Markov chains for the evolution of the nodes groups through time. We model binary data as well as weighted dynamic random graphs (with discrete or continuous edges values). Our approach, motivated by the importance of controlling for label switching issues across the different time steps, focuses on detecting groups characterized by a stable within group connectivity behavior. We study identifiability of the model parameters, propose an inference procedure based on a variational expectation maximization algorithm as well as a model selection criterion to select for the number of groups. We carefully discuss our initialization strategy which plays an important role in the method and compare our procedure with existing ones on synthetic datasets. We also illustrate our approach on two dynamic contact networks, one of encounters among high school students and the second on birds interactions. An implementation of the method is available as an R package called `dynsbm`.

*11h00-11h20: pause*

- (2) 11h20-12h20: Stephanie Allasonnière (CMAP, Ecole Polytechnique)

Titre: Mixed-effect model for the spatiotemporal analysis of longitudinal manifold-valued data

Résumé: In this work, we propose a generic hierarchical spatiotemporal model for longitudinal manifold-valued data, which consist in repeated measurements over time for a group of individuals. This model allows us to estimate a group-average trajectory of progression, considered as a geodesic of a given Riemannian manifold. Individual trajectories of progression are obtained as random variations, which consist in parallel shifting and time reparametrization, of the average trajectory. These spatiotemporal transformations allow us to characterize changes in the direction and in the pace at which trajectories are followed. We propose to estimate the parameters of the model using a stochastic version of the expectation-maximization (EM) algorithm, the Monte Carlo Markov Chain Stochastic Approximation EM (MCMC SAEM) algorithm. This generic spatiotemporal model is used to analyze the temporal progression of a family of biomarkers. This progression model estimates a normative scenario of the progressive impairments of several cognitive functions, considered here as biomarkers, during the course of Alzheimer's disease. The estimated average trajectory provides a normative scenario of disease progression. Random effects provide unique insights into the variations in the ordering and timing of the succession of cognitive impairments across different individuals.

*12h20-14h00: repas offert aux participants inscrits à la rencontre*

- (3) 14h00-15h00: Romain Azais (INRIA Nancy)

Titre: Deux stratégies non-paramétriques pour estimer le taux de saut d'un processus markovien déterministe par morceaux

Résumé: Les processus markoviens déterministes par morceaux ont été introduits dans les années 80 par Davis comme une généralisation des processus de renouvellement marqués. Entre deux sauts, le comportement de ces processus est régi par une équation différentielle dépendant éventuellement de la marque souvent appelée régime dans ce contexte. Le taux de saut d'un tel processus est une fonction de la trajectoire déterministe suivie. Je présenterai deux stratégies non-paramétriques pour tenter d'estimer cette caractéristique du modèle: la première s'appuie sur le modèle à intensité multiplicative d'Aalen traditionnellement utilisé en analyse de survie; la seconde, basée sur des méthodes à noyau récursives, permet de construire une famille d'estimateurs consistants parmi lesquels on choisit celui de variance minimale.

(4) 15h00-15h40: Thomas Opitz (INRA Avignon, BioSp)

Titre: Inférence statistique des valeurs extrêmes spatiales: modèle asymptotique classique ou modèle tendant vers l'indépendance asymptotique?

Résumé: La théorie des valeurs extrêmes spatiales propose des outils statistiques pour caractériser, modéliser et prédire la fréquence et l'ampleur de phénomènes extrêmes tels que les tempêtes de vent, les pluies torrentielles ou les vagues de chaleur ou de froid. Dans la première partie de l'exposé, je présenterai cette théorie, les lois asymptotiques qui en découlent et les modèles couramment utilisés dans le cadre spatial. En particulier, nous distinguons entre les processus limites max-stables, obtenus pour des observations de maxima, et les processus limites de Pareto généralisés, obtenus pour des dépassements de seuil. Ces modèles sont bien adaptés à capter la dépendance asymptotique, caractérisée par la co-occurrence de très fortes valeurs sur les sites proches en cas d'un événement extrême. Sur un exemple pratique, j'illustrerai l'ajustement d'un modèle de Pareto généralisé à des données de précipitation journalière observées sur 44 sites autour de Zurich. Or, dans beaucoup de cas, l'exploration empirique de la dépendance spatiale extrême des variables climatiques suggère que l'hypothèse de dépendance asymptotique est pour le moins contestable. Cette observation se manifeste dans une étendue spatiale des valeurs les plus fortes qui devient de plus en plus confinée pour les phénomènes les plus rares. C'est sur ce fond que j'évoquerai certains modèles actuellement étudiés, capables de mieux capter et caractériser l'indépendance asymptotique que les modèles asymptotiques classiques.

*15h40-16h00: pause*

(5) 16h00-16h40: Frédéric Richard (Institut de Mathématiques de Marseille)

Titre: Indices d'anisotropie pour l'analyse de textures Browniennes d'images.

Résumé: L'anisotropie est l'une des propriétés essentielles des textures d'images, qui sert notamment pour leur classification. Dans cet exposé, j'aborderai la question de la caractérisation de cette propriété en construisant des indices d'anisotropie. Je me placerai dans un cadre aléatoire où l'image est vue comme la réalisation d'un champ aléatoire défini par extension du modèle Brownien fractionnaire. Je présenterai ensuite une méthode pour inférer des caractéristiques directionnelles du champ sous-jacent. Cette méthode repose sur des sommes quadratiques d'incrémentes, appelées variations quadratiques, qui ont la particularité d'être calculées à différentes échelles et dans différentes directions. En m'appuyant sur une étude asymptotique de ces variations, je dégagerai les relations qui existent entre elles et les propriétés directionnelles du champ. A partir de ces relations, je construirai des indices d'anisotropie et en proposerai des estimateurs. Pour finir, je présenterai une application de ces indices à la détection de tumeurs dans les mammographies.

(6) 16h40-17h30: Quentin Paris (Laboratory of Stochastic Analysis and its Application, Moscow)

Titre: Comportement en temps fini de la probabilité de nouveauté.

Résumé: Etant donnée une suite de  $n$  observations  $X_1, \dots, X_n$  indépendantes et identiquement distribuées selon une loi discrète, la probabilité de nouveauté au temps  $(n+1)$  est définie comme la probabilité que la  $(n+1)$ -ième observation  $X_{n+1}$  soit distincte de toutes les observations précédentes. Cette quantité simple, parfois appelée masse manquante, est d'intérêt pratique, par exemple, en écologie, Génomique, Théorie de l'information et Linguistique. Du point de vue Théorique, un certain nombre de caractéristiques de la masse manquante sont aujourd'hui bien comprises (comportement asymptotique, concentration). Suite à un rappel des résultats connus, cet exposé portera sur certains résultats récents concernant le comportement à  $n$  fixé de la probabilité de nouveauté. En fin d'exposé, nous discuterons d'une extension de nos résultats au cas d'observations i.i.d. à valeurs dans un espace métrique général.

*17h20: fin des rencontres*