

Vendredi 26 juin 2015, à l'Université d'Avignon dans l'amphi 2E07

## Quatrièmes rencontres de statistique Avignon/Marseille: Programme détaillé

*9h-9h30: accueil des participants*

- (1) **9h30-10h30: Anne-Laure Fougères** (Institut Camille Jordan, Université Lyon 1)

Titre: Catastrophes naturelles et autres événements extrêmes : quelques modèles à l'oeuvre.

Résumé: Quelle est la probabilité qu'une tempête du type Xynthia ne se reproduise dans les 5 ans à venir? Comment évaluer la probabilité qu'une compagnie d'assurances, comptant plusieurs branches (généralement dépendantes) de contrats, n'ait à faire face à un sinistre provoquant sa ruine? Comment estimer la garantie que peut couvrir un constructeur? La théorie des valeurs extrêmes propose une famille de modèles permettant de répondre à des problèmes de ce type. Nous présenterons les hypothèses sous lesquelles de tels modèles existent, et montrerons comment des approximations peuvent être obtenues via ces modèles. Une application en lien avec des données climatiques illustrera notre propos.

*10h30-11h00: pause*

- (2) **11h00-12h00: Xavier Gendre** (Institut de Mathématiques de Toulouse, Université Toulouse 3)

Titre : Modèles déformables et moyenne de Fréchet pour le recalage de courbes

Résumé : Nous nous intéressons au problème d'estimation d'une forme moyenne à partir d'observations d'un ensemble de courbes similaires dans un cadre où la variabilité des données est due à des déformations géométriques et à un bruit additif. Ce problème nécessite de définir des distances non Euclidiennes en considérant l'action d'un groupe de Lie sur l'espace des courbes. Cette approche mène à la construction d'estimateurs basés sur la notion de moyenne de Fréchet qui est une généralisation de la moyenne classique dans des espaces non Euclidiens. A l'aide de ces outils, nous montrons le lien qui existe entre la théorie minimax pour la statistique non paramétrique et le problème d'estimation d'une forme moyenne à partir d'un ensemble de courbes.

*12h00-13h30: buffet offert à tous les participants inscrits à la rencontre*

- (3) **13h30-14h30: Arnaud Guyader** (Laboratoire de Statistique Théorique et Appliquée, Université Pierre et Marie Curie Paris 6)

Titre: Analyse non paramétrique de l'algorithme ABC

Résumé: Apparues à la fin des années 1990, les méthodes dites ABC (pour Approximate Bayesian Computation) entrent dans la catégorie des techniques bayésiennes. Elles ont typiquement pour but d'estimer la densité a posteriori du paramètre pour une observation donnée. Dans cet exposé, nous présentons une analyse mathématique

détaillée d'un algorithme ABC typique, en formalisant son lien avec des techniques non paramétriques d'estimation de la densité conditionnelle par plus proches voisins. Nous montrons en particulier qu'un estimateur de la densité correctement choisi et calibré permet d'approcher la loi a posteriori, tout en donnant une heuristique pour le choix du nombre de voisins à conserver.

- (4) **14h30-15h10: Thibaut Le Gouic** (Institut de Mathématiques de Marseille, Ecole Centrale de Marseille)

Titre: reconstruire une métrique à partir d'informations ordinales.

Résumé: On observe un échantillon de  $n$  variables aléatoires à valeurs dans un espace métrique  $(E, d)$ . On ne connaît pas la métrique  $d$  de l'espace, mais pour chaque quadruplé  $(w, x, y, z)$  de l'échantillon, on sait si  $d(w, x) < d(y, z)$ . Peut-on retrouver la distance  $d$ ? Cette question suscite un intérêt particulier dans la communauté de l'apprentissage automatique et nous présenterons des éléments de réponse.

*15h10-15h40: pause*

- (5) **15h40-16h40: Nicolas Verzelen** (Institut National de Recherche Agronomique, Montpellier)

Titre: détection de modèle de mélange gaussien parcimonieux et classification non supervisée

Résumé: L'objectif de cet exposé est de comparer la difficulté de deux problèmes statistiques: (i) classification supervisée (ii) classification non supervisée. Pour ce faire, on adoptera une approche "model-based" en considérant des modèles de mélange gaussien en grande dimension. Après une revue sélective de la littérature sur les modèles de mélange, on s'intéressera aux vitesses optimales de classification, ce afin de caractériser les régimes dans lesquels une classification consistante est réalisable.

- (6) **16h40-17h20: Delphine Blanke** (Laboratoire de Mathématiques d'Avignon, Université d'Avignon et des Pays du Vaucluse)

Titre: Détection de sauts pour des processus ARMAD

Résumé: Nous considérons des fonctions aléatoires réelles dépendantes et de même loi, définies sur  $[0, 1]$  qui admettent un nombre fini de sauts, à des instants fixes ou aléatoires. Ces fonctions, considérées sur des intervalles de temps successifs, forment alors des processus ARMA fonctionnels à valeurs dans  $D$ , où  $D$  est l'espace de Skorokhod associé aux fonctions càdlàg. Nous commençons par donner les propriétés des processus ARMAD puis nous étudions différents scénarii : instants fixes de sauts avec intensités ordonnées (cas "déterministe"), instants aléatoires de sauts avec intensités ordonnées (cas "pseudo-Poisson") et instants aléatoires de sauts avec intensités non ordonnées (cas "complètement aléatoire"). En se basant sur des observations discrétisées de ces processus, nous identifions les instants de saut pour chaque scénario puis nous estimons leur intensité.

*17h30: fin des rencontres*