

Propositions de sujets de stage, parcours Biostatistique, 2010-2011

Sujet 1

Titre :

Réduction de dimension en analyse de survie

Contact/encadrement :

Marie-Luce Taupin, Laboratoire Statistique & Génome, marie-luce.taupin@genopole.cnrs.fr

Sujet 2

Titre :

Modélisation de séquences d'ARN par réseaux Bayésiens et application à la distribution de motifs dans des séquences aléatoires contraintes.

Contact/encadrement :

Grégory Nuel, Laboratoire MAP5, gregory.nuel@parisdescartes.fr

Brève description :

En biologie moléculaire, l'acide ribonucléique est un polymère linéaire constitué d'un enchaînement de nucléotides (au nombre de quatre : adénine, guanine, cytosine et uracile) qui peut donc s'assimiler à une séquence écrite dans un alphabet à quatre lettres (A, C, G, U). Du fait des propriétés de complémentarité de ces nucléotides (A-U et C-G), les séquences d'ARN ont une tendance naturelle à se replier sur elle-même pour former des structures dites secondaires (tiges, boucles, etc.). L'objectif de ce stage est la modélisation de tels séquences fortement contraintes à l'aide de réseau bayésien avec pour objectif final la recherche statistique de motifs d'intérêt dans les séquences d'ARN en tenant compte de leurs structures secondaires.

Sujets 3 et 4

Titre 1 :

Débruitage de séquences DCE-CT pour l'aide au suivi des traitements anti-angiogéniques dans le cancer. Approche par tests d'équivalence pour la phase de classification automatique des tissus.

Titre 2 :

Débruitage de séquences DCE-CT pour l'aide au suivi des traitements anti-angiogéniques dans le cancer. Approche par ondelettes, réduction de la multiplicité des tests et classification.

Contact/encadrement :

Yves Rozenholc, Laboratoire MAP5, yves.rozenholc@parisdescartes.fr

Brève description :

Le sujet général est le débruitage de séquences DCE-CT pour l'aide au suivi des traitements anti-angiogéniques dans le cancer. Suivre "in vivo" la réponse tissulaire à l'injection d'un

agent de contraste (Dynamic Contrast Enhanced) via un scanner tomographique (CT) permet d'accéder aux paramètres de perfusion tissulaire et de permettre un suivi des effets d'un traitement angiogénique. Malheureusement, pour des raisons techniques mais aussi éthiques (usage de radiations en CT), ce type d'examen procure des images de mauvaise qualité difficilement utilisables pour une évaluation directe.

Récemment [1] a introduit une nouvelle approche basée sur des tests multiples pour débruiter ce type de séquence. Ce travail propose suite d'utiliser les résultats de cette phase de débruitage comme entrée d'une classification automatique des types de tissus. Cette seconde phase offre aux médecins la construction automatique de régions d'intérêt (ROI) homogènes eut égard aux rehaussements tissulaires.

Deux stages peuvent être proposés pour poursuivre ces travaux :

Stage 1 : S'appuyant sur la notion de test d'équivalence [2], ce stage veut mettre en place une technique qui offrira aux médecins une classification des tissus en ROI homogènes en s'émancipant de la phase de débruitage.

Stage 2 : S'appuyant sur une approche de type décomposition en ondelette pour laquelle il est connu que la représentation des signaux est "sparse" en nombre de coefficients, l'idée est de comparer directement les développement de type ondelette pour réduire le nombre de tests multiples nécessaires à la comparaison des signaux et ainsi offrir un débruitage et une classification plus efficace. Cette approche s'inspire des travaux [3] et [4].

[1] Y. Rozenholc, M. Reiß, D. Balvay, and C.-A. Cuenod. Growing time-homogeneous neighborhoods for denoising and clustering dynamic contrast enhanced-ct sequences. preprint, <http://www.math-info.univ-paris5.fr/~rozen/RRBC-preprint/last-version.pdf>, submitted.

[2] Stefan Wellek. Testing Statistical Hypotheses of Equivalence. Chapman & Hall, 2002.

[3] Yoav Benjamini and Ruth Heller. False discovery rates for spatial signals. J. Amer. Statist. Assoc., 102(480):1272–1281, 2007.

[4] R. Heller., D. Stanley, D. Yekutieli, N. Rubin, and Y. Benjamini. Cluster-based analysis of fmri data. Neuroimage, 33(2):599–608, 2006.

Sujet 5

Titre :

Evaluation des modèles non linéaires mixtes appliqués à la dynamique virale dans le traitement VIH et/ou VHC

Contact/encadrement :

France Mentré, Inserm UMR 738, france.mentre@inserm.fr

Brève description :

Les modèles non linéaires mixtes sont utilisés pour modéliser l'évolution des charges virales dans la réponse au traitement VIH et VHC. Ces modèles sont décrits par des équations différentielles et reposent sur de nombreuses hypothèses statistiques concernant la variabilité inter et intra patients et l'erreur résiduelle. Différents outils d'évaluation interne des modèles

existent et les plus performants reposent sur des simulations intensives. Toutefois les tests d'adéquation de modèles sont peu étudiés dans ce contexte et permettent seulement de "rejeter" un modèle. Il s'agira d'étudier comment on peut inverser les hypothèses pour montrer qu'un modèle est "correct" en fonction de bornes d'équivalence à définir.

Sujet 6

Titre :

Modèles causaux pour l'étude des facteurs génétiques et environnementaux de l'asthme

Contact/encadrement :

Raphaëlle Varraso et Jean Maccario (INSERM U1018, Epidémiologie respiratoire et environnementale, raphaelle.varraso@inserm.fr et jean.maccario@inserm.fr), [Antoine Chambaz](mailto:Antoine.Chambaz@parisdescartes.fr) (laboratoire MAP5, antoine.chambaz@parisdescartes.fr)

Brève description :

L'asthme est un problème majeur de santé publique, qui concerne 5 à 10% de la population générale. La forte augmentation de la prévalence de l'asthme au cours des 30 dernières années, révèle l'importance de facteurs environnementaux et comportementaux ayant changé durant cette période. Parmi ces facteurs, on évoque particulièrement la diminution du niveau d'activité physique, l'augmentation de l'obésité et la modification des habitudes alimentaires (diminution de la consommation des fruits et légumes et augmentation de la consommation de plats préparés).

On sait que la composition corporelle est influencée par des choix alimentaires, le niveau d'activité physique et des facteurs génétiques. Cependant, les relations causales entre obésité, alimentation et activité physique ne sont pas clairement établies.

En ce qui concerne l'asthme, l'obésité a été montrée comme un facteur de risque de la maladie et de sa sévérité. Le rôle favorable d'une activité physique régulière sur les exacerbations d'asthme a été évoquée ; cependant ne sont clairement établis ni le sens ni la nature de ces relations. L'existence dans l'alimentation de facteurs délétères ou protecteurs pour l'asthme est suspectée, cependant les aliments/nutriments spécifiquement en cause restent inconnus. Les influences croisées de tous les facteurs mentionnés restent à étudier car on s'est peu intéressé aux relations entre activité physique, alimentation, composition corporelle et asthme.

L'utilisation de modèles causaux pourrait être particulièrement pertinente et utile pour aborder ce problème. Le stage pourrait porter sur l'établissement et la validation de tels modèles, et leur application aux données disponibles dans l'enquête EGEA (étude épidémiologique des facteurs génétiques et environnementaux de l'asthme, <http://ifr69.vjf.inserm.fr/~egeanet/>). Les données de l'enquête EGEA sont adaptées à ce type de travail car elles contiennent des informations détaillées sur la composition corporelle (indice de masse corporelle, mais aussi tour de taille, tour de hanche, masses grasse et maigre, silhouettes au cours de la vie), sur les différents types d'activité physique, sur l'alimentation ainsi que sur différents phénotypes d'asthme.

Sujet 7

Titre :

Modélisation des données censurées à droite et tronquées à gauche dans les essais cliniques et fonctions de survie qui se croisent

Contact/encadrement :

Mikhaïl Nikouline, IMB (Bordeaux), mikhail.nikouline@u-bordeaux2.fr

Sujet 8

Titre :

Modélisation mathématique de la conversion rénale citrulline-arginine

Contact/encadrement :

Ioannis Nicolis, Emmanuel Curis, Simone Bénazeth, Laboratoire de Biomathématiques – EA 4466 – Faculté de Pharmacie – Université Paris Descartes, simone.benazeth@parisdescartes.fr

Détails disponibles à l'URL

<http://www.math-info.univ-paris5.fr/~chambaz/M2Rbiostatistique/projetM2-modelmetab.pdf>