
Séance d'affichage II

JUSTIN ANGEVAARE, University of Guelph

Utilité de la variance des captures par unité d'effort

Les modèles de population jouent un rôle central dans la gestion des pêches. Les données sur la pêche commerciale, particulièrement les captures par unité d'effort (CPUE), sont utilisées pour prédire les paramètres et l'abondance des populations. Une revue de la littérature récente révèle que la variabilité des CPUE est généralement ignorée. La pratique courante consiste à regrouper les données de capture et d'effort sur une base annuelle. Nous présentons une étude de simulation qui examine les modèles traditionnels et ceux qui ont été modifiés pour y intégrer la variabilité des CPUE. Les populations de poissons et les événements de pêche sont simulés sur une base quotidienne pour une période de 60 ans. Les données de capture et d'effort sont regroupées en fonction de plusieurs échelles temporelles. Les estimés des paramètres de population sont comparés entre les différents modèles et échelles de regroupement temporelles.

SHAHEENA BASHIR, University Health Network, Toronto

Tests simultanés de l'ampleur de l'effet dans une méta-analyse multivariée

La méta-analyse multivariée combine les estimateurs de plusieurs paramètres reliés dans plusieurs études. Ce travail examine la possibilité d'obtenir une statistique F qui peut tester les différences parmi plus de deux groupes dans un contexte de méta-analyse. Cette question s'est posée dans une méta-analyse des structures d'expression génétique émergeant de données comportant quatre groupes de traitements et six comparaisons possibles par paires dans deux études. Nous évaluons la performance de notre test statistique en le comparant, au moyen de simulations, à des approches standard en matière de méta-analyse à effets aléatoires. Les simulations démontrent que la forme quadratique de la statistique t par paire du modèle à effets aléatoires suit une distribution de Fisher.

VICTORIA BORG DEBONO, McMaster University

Qualité des rapports sur les essais contrôlés randomisés (ECR) dans la gestion de la douleur postopératoire

Les rapports de qualité sur les ECR sont importants puisqu'ils permettent aux chercheurs d'évaluer la rigueur méthodologique des ECR. Cette étude vise à évaluer la qualité des rapports (QR) sur les ECR dans la gestion de la douleur postopératoire à l'aide du critère CONSORT, et à déterminer les facteurs associés à une QR élevée. Globalement, il a été constaté que la QR sur les essais en anesthésie était faible à modérée, et ces résultats concordent avec les observations d'autres disciplines cliniques. Les renseignements ainsi obtenus peuvent être utilisés par des revues pour marquer l'urgence d'améliorer la présentation des résultats dans les rapports en vue de rendre les conclusions des ECR accessibles et comparables.

ZHENGFEI CHEN, University of Toronto

MCMC hamiltonien sur des variétés de Riemann pour des processus de Cox log-gaussiens

Les processus de Cox log-gaussiens (PCLG) ponctuels forment une riche classe de modèles pour les motifs ponctuels en grappes. Ils sont utiles et populaires pour les données de localisation distribuées de façon inhomogène. La méthode de Monte Carlo hamiltonienne sur des variétés de Riemann (MMCHVR) produit un échantillonnage hautement efficace même dans des cas de très haute dimensionnalité, où différents ajustements d'échelle sont nécessaires lors des phases transitoires et stationnaires de la chaîne de Markov. Nous examinons de belles propriétés des PCLG avec une covariance de Matérn utilisant le MCMC hamiltonien (MCMCH), de même que l'application de la MMCHVR à des PCLG avec une matrice de covariance de Matérn sur un jeu de données de haute dimensionnalité.

JING DING, Acadia University

Évaluations des populations au moyen des données sur l'âge des prises et l'effort de pêche

Les données sur l'âge des prises et l'effort de pêche sont des quantités facilement accessibles dans la gestion de la faune. Elles fournissent des renseignements essentiels pour la modélisation des populations en vue de révéler la situation des espèces exploitées. Nous introduisons trois méthodes applicables. La reconstruction de Downing (1980) est une technique utilisée couramment qui consiste en des ajouts antéchronologiques aux cohortes afin de procéder à la reconstruction des populations dans le temps. Une autre méthode applicable élaborée par Paloheimo (1980) consiste en un modèle de régression linéaire incorporant l'effort de pêche. La méthode de l'effort de pêche stochastique (Dupont, 1983) constitue la méthode la plus avancée, puisqu'elle requiert des hypothèses plus faibles sur le processus de naissance-mort-capture. Des modèles de risques sont définis et les moindres carrés ordinaires peuvent être utilisés pour en estimer les paramètres.

JAYDE EUSTACE, Memorial University of Newfoundland

Modèle à espace d'états de Gumbel pour des valeurs extrêmes

La théorie des valeurs extrêmes offre plusieurs applications dans des domaines tels que la météorologie, la finance et l'ingénierie. Au cours des dernières années, l'intérêt a repris concernant l'analyse de séries de valeurs extrêmes. Dans cet exposé, certaines propriétés des mélanges de Gumbel et des variables aléatoires α -stables sont utilisées pour construire un modèle à espace d'états avec des lois marginales de type Gumbel. La densité des variables aléatoires α -stables ne peut pas être exprimée sous forme analytique et leur espérance n'existe pas. Cet aspect pose des défis en matière d'inférence. Les méthodes proposées sont comparées aux modèles actuels par des simulations avec des données réelles.

REMYA M. JOHN, McMaster University

Analyse statistique de données d'ARN-Seq : comparaison de trois librairies R

Le séquençage d'ARN de nouvelle génération (ARN-Seq) est un nouvel outil permettant d'analyser l'expression différentielle de gènes. Plusieurs librairies R/BioConductor ont été élaborées au cours des dernières années pour gérer ce type de données. Dans ce projet, nous examinons la performance de trois librairies largement utilisées (DESeq, edgeR et DESeq2) sur deux jeux de données avec répétitions biologiques, nous montrons que l'utilisation d'un modèle de Poisson dans DESeq est inappropriée à la présence de réplicats biologiques et qu'il y a un manque de cohérence entre les deux autres méthodes, même si leur performance générale est similaire.

XIAOQING NIU, University of Alberta

Test d'homogénéité pour des modèles gaussiens à dispersion aléatoire

Les modèles gaussiens à dispersion aléatoire sont un type particulier de modèle de mélange ayant plusieurs applications. Nous proposons d'évaluer leur homogénéité à l'aide d'un test EM révisé. Deux approches sont mises de l'avant pour l'initialisation des moyennes communes : utiliser les moyennes échantillonales ou maximiser la log-vraisemblance pénalisée. Nous montrons que la statistique du test EM suit la distribution simple $\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$. Des études de simulation montrent l'exactitude du seuil du test EM et la puissance supérieure de la seconde approche face à la première. Les données de deux exemples concrets sont analysées.

MATTHEW RUEFFER, University of Guelph

Modélisation spatiale et temporelle de la variation des pêches

Entre 1979 et 2010, des données relatives aux captures du grand corégone (*Coregonus clupeaformis*) ont été recueillies pour la portion canadienne du lac Huron. Ces données, qui proviennent de pêcheurs autochtones et du ministère des ressources naturelles de l'Ontario, comprennent l'emplacement géographique, les captures, l'effort relatif et le maillage. Cette affiche

examine les captures par unité d'effort (CPUE = quotient des captures sur l'effort) à travers les régions et les années au moyen de méthodes exploratoires et de modèles temporels et spatiaux.

KATHLEEN RYAN, University of Guelph

*Distribution en profondeur des larves de grand corégone (*Coregonus clupeaformis*) dans une baie des Grands Lacs.*

Cette étude examine la relation entre la densité de larves du grand corégone (*Coregonus clupeaformis*) et la profondeur de l'eau à Stokes Bay dans le lac Huron. L'hypothèse a été émise selon laquelle les eaux littorales (1-3 m) des Grands Lacs offriraient un habitat favorable aux larves du grand corégone, ce qui a mené les chercheurs à concentrer leur travail à cet endroit. La distribution des larves à Stokes Bay a été examinée au moyen de filets de 500 microns remorqués en surface entre avril et mai 2011. Un nouveau plan d'échantillonnage aléatoire stratifié a assigné des sites d'échantillonnage à des profondeurs variant de 1 à 11 mètres (n=71). Les analyses statistiques suggèrent que la profondeur n'est pas significative pour prédire la densité de larves.

OLGA SHESTAKOVSKA, Population Health Research Institute, McMaster University and Hamilton Health Sciences

De grands ratios de risque ne signifient pas toujours une grande exactitude prédictive : étude de simulation

En analyse de survie, le ratio de risque est une mesure du risque relatif fréquemment utilisée pour évaluer les effets de la démographie, de caractéristiques cliniques et de biomarqueurs sur les taux d'événements. Nous explorons la relation entre le ratio de risque et les mesures de l'exactitude prédictive comme la statistique c de Harrell et les courbes ROC dépendantes du temps dans une étude de simulation. La motivation est venue de l'analyse d'un score de risque calculé pour prévoir les accidents vasculaires cérébraux chez les patients atteints de fibrillation auriculaire. Nous observons une forte association entre le score et les taux d'événements (HR=2,45; 95% CI=1,66-3,75) alors que la statistique c de Harrell donne 0,567.

MONICA SIRSKI, University of Manitoba

À propos de l'analyse statistique de données fonctionnelles découlant de plans d'expérience

Nous examinons plusieurs méthodes pour vérifier si deux groupes de courbes sont significativement différents d'un point de vue statistique, avec la motivation d'appliquer les méthodes d'analyse des données découlant des plans d'expérience. Nous proposons un ensemble de tests basés sur des différences par paire entre les courbes individuelles. Notre objectif est de comparer la puissance et la robustesse d'une variété de tests, y compris des tests de permutation. Nous illustrons l'application de ces tests dans le contexte d'un plan d'expérience factoriel 2^4 avec une étude de cas utilisant des données fournies par la NASA.

GARY SNEDDON, Mount Saint Vincent University

Modèles pour données longitudinales et succès électoral d'anciens ministres

Les modèles pour données longitudinales sont d'intérêt dans une variété de disciplines où des données transversales sont observées à répétition sur les mêmes individus au fil du temps. Nous examinons ces modèles et leurs méthodes d'estimation en étudiant des données sur le succès électoral d'anciens ministres fédéraux aux élections subséquentes. Nous examinons aussi des approches statistiques plus courantes dans la littérature des sciences sociales.

MARY THIESS, Pêches et Océans Canada

Analyse de séries chronologiques : évaluation de "l'effet saumon rose"

Le saumon sockeye du fleuve Fraser a récemment fait l'objet d'une grande couverture médiatique en raison d'une montaison d'adultes très variable au cours des dernières années dans un contexte de déclin généralisé. L'une des nombreuses hypothèses pouvant expliquer cette variabilité suggère la compétition potentielle entre le saumon sockeye et le saumon rose. L'isolation de la relation entre leur série chronologique respective nécessite l'évaluation (et la quantification) de l'autocorrélation et de

la tendance non monotone dans le temps. Cette affiche expose les défis soulevés par l'évaluation de "l'effet saumon rose" et propose certaines approches pour l'analyse de ces données.

MONICA H.T. WONG, University of Guelph

Analyse de regroupement pour évaluer le maintien d'une perte de poids : application des analyses de regroupements à la nutriginomique

En nutriginomique, une analyse de regroupements des profils d'expression génique provenant de biopuces peut aider à identifier des sous-populations dont la réponse aux interventions diététiques est différente. Alors que les plans de nutrition personnalisés gagnent en popularité, l'utilisation d'analyses de regroupements foisonne dans la recherche sur l'obésité. Cette étude vise à identifier les individus dont la réponse diffère à partir d'une analyse de regroupements de profils d'expression génique obtenus durant une diète faible en calories, ainsi qu'à identifier les voies biologiques affectées chez ces individus durant la diète.

HAIYAN YANG, Memorial University of Newfoundland

Modélisation conjointe de la liaison et de l'association génétique

Un modèle conjoint est présenté pour la liaison et l'association génétique à l'intérieur des familles ayant un ancêtre commun éloigné. Ce modèle conjoint est basé sur la vraisemblance, ce qui permet l'inclusion de covariables quantitatives. Pour les traits binaires, un test de score est proposé. L'étude de simulation montre que le test conjoint est plus puissant qu'un test sur la liaison ou l'association seulement où les deux sources de variations sont présentes. De plus, la méthode proposée permet des contre-hypothèses spécifiques, par exemple la significativité de la liaison sachant qu'il n'y a pas d'association.

LI HUA YUE, University of Western Ontario

Prévisions logistiques de l'état du cancer de la prostate à l'aide de données d'imagerie

De nouvelles techniques d'imagerie médicale, telles que la tomographie par émission de positons et la résonance magnétique, sont devenues couramment disponibles pour le diagnostic et la gestion du cancer de la prostate. Il est essentiel d'identifier précisément la nature du cancer pour l'élaboration de traitements. Les données sur l'état pathologique du cancer de la prostate disponibles pour cette étude nous permettent de construire un modèle prédictif basé sur les caractéristiques des données d'imagerie. Nous pouvons ainsi prédire le statut du cancer de la prostate, et nous nous attendons à ce que ce modèle soit utilisé dans le futur pour le diagnostic du cancer de la prostate basé sur l'imagerie médicale. Nous présentons des résultats préliminaires de prévision du statut du cancer basée sur des propriétés d'images obtenues par tomographie par émission de positons, T2WMR et DCEMR. L'exactitude des prévisions est évaluée par l'aire sous la courbe ROC, déterminée par une validation croisée séparant les données en dix morceaux.

WEI YUHONG, University of Guelph

Moyennage de modèles de mélange

Les critères couramment utilisés pour choisir le nombre de composantes d'un modèle de mélange fini comprennent l'AIC, le BIC, la pondération approximative des informations probantes et l'intégration de la vraisemblance complète. Nous avançons que ces approches ne sont pas raisonnables lorsque la différence entre la valeur optimale du critère et les autres valeurs est petite. Par conséquent, il est préférable d'envisager une approche de moyennage des modèles. Nous considérons le moyennage des M meilleurs modèles de mélange et présentons des applications pour la classification et les analyses de regroupements.