
Translations of Abstracts

DISCUSSION PAPER**Tyler J. VanderWeele****561**

Discussion on “Surrogate Measures and Consistent Surrogates”

Lorsqu’il est difficile ou coûteux de mesurer un critère principal, il devient de toute première importance de pouvoir recourir à un critère de substitution permettant de prédire l’effet d’un traitement sur le critère d’intérêt à partir de l’effet du traitement sur le critère de substitution. Malheureusement, l’emploi de ces critères de substitution peut parfois donner lieu à des situations paradoxales, où, alors que l’effet du traitement sur le critère de substitution est positif, et que les critères de substitution et d’intérêt sont fortement corrélés positivement, l’effet du traitement sur le critère d’intérêt s’avère, quant à lui, négatif : il s’agit du fameux « paradoxe des critères de substitution ». Nous proposons de nouveaux résultats concernant la cohérence des critères de substitution, prolongeant ainsi la littérature existant déjà sur les conditions suffisantes à prévenir ce paradoxe. Tout particulièrement, nous montrons qu’une condition nécessaire à l’avènement du paradoxe des critères de substitution est qu’il y ait 1) soit un effet direct du traitement sur le critère d’intérêt, qui ne passe pas par le critère de substitution et qui soit de sens opposé à la partie de l’effet qui survient via le critère de substitution, 2) soit une corrélation positive entre le critère de substitution et le critère d’intérêt qui soit seulement liée à la présence d’un facteur de confusion non mesuré (le critère de substitution n’ayant, par lui-même, aucun effet sur le critère d’intérêt), 3) soit un manque de transitivité, c’est-à-dire que les individus pour lesquels le traitement affecte positivement le critère de substitution ne sont pas les mêmes que ceux pour lesquels le critère de substitution affecte positivement le critère d’intérêt. Les conditions de la cohérence d’un critère de substitution et les résultats de cet article sont importants car ils permettent aux investigateurs de prédire le sens de l’effet d’un traitement sur un critère principal simplement à partir du sens de l’effet de ce traitement sur le critère de substitution. Nous relierons ensuite nos résultats sur la cohérence des critères de substitution aux quatre approches décrites par Joffe et Greene (2009, *Biometrics* **65**, 530-538), afin d’établir si le critère standard employé par lesdites approches, pour évaluer si un critère de substitution est « bon », suffit à éviter le fameux paradoxe.

BIOMETRIC METHODOLOGY**Sunkyung Kim, Wei Pan, and Xiaotong Shen****582***Network-Based Penalized Regression with Application to Genomic Data*

Les approches de régression pénalisée sont intéressantes dans le traitement des données de grande dimension comme celles provenant d’études génomiques à haut rendement. De nouvelles méthodes ont été introduites pour utiliser la structure du réseau des prédicteurs,

par exemple réseaux de gènes, afin d'améliorer l'estimation des paramètres et la sélection de variables. Toutes les méthodes existantes fondées sur le réseau pénalisé sont basées sur l'hypothèse que les paramètres, par exemple les coefficients de régression, de nœuds voisins dans un réseau sont proches en valeur, ce qui, cependant peut ne pas être le cas. Nous proposons ici une nouvelle méthode de régression pénalisée fondée sur une hypothèse plus faible que les paramètres de nœuds voisins dans un réseau sont susceptibles d'être nuls (ou différents de zéro) en même temps, quelles que soient leurs valeurs spécifiques. Nous proposons une nouvelle fonction de pénalité non convexe pour incorporer cet-a priori, et un algorithme fondé sur la programmation convexe. Nous utilisons les données simulées et un ensemble de données d'expression génique pour montrer les avantages de la méthode proposée par rapport à quelques méthodes existantes. Nos méthodes proposées peuvent être appliquées plus généralement aux problèmes de sélection de groupes de variables.

Yaqing Si and Peng Liu

594

An Optimal Test with Maximum Average Power While Controlling FDR with Application to RNA-Seq Data

La récente technologie RNA-seq est une méthode attractive pour étudier l'expression des gènes. Un des objectifs les plus importants dans l'analyse des données RNA-seq est de détecter des gènes différentiellement exprimés entre traitements. Bien que plusieurs méthodes statistiques aient été publiées, il n'y a pas de justification théorique de l'optimalité de ces méthodes ou comment chercher un test optimal. En outre, la plupart des tests proposés sont conçus pour tester si les niveaux d'expression moyens sont exactement les mêmes ou non entre les traitements, tandis que parfois, les biologistes sont intéressés par la détection de gènes dont le changement d'expression est plus grand qu'un certain seuil. Une autre question avec ces méthodes est que le contrôle du taux de faux positif (FDR) n'est pas bien étudié. Dans ce manuscrit, nous proposons un test qui aborde toutes ces questions. Sous les hypothèses du modèle, nous dérivons un test optimal qui atteint le maximum de puissance moyenne parmi ceux qui contrôlent le FDR à un même niveau. Nous fournissons aussi une version approximée, le test *approximated most average powerful* (AMAP), pour une implémentation pratique. La méthode proposée permet de tester des hypothèses plus générales que celles que la majorité des études précédentes ont considérées et conduit à une façon naturelle de contrôler le FDR. A travers des études de simulation, nous montrons que notre test a une puissance supérieure aux autres méthodes, incluant les méthodes très utilisées edgeR, DESeq et baySeq, et un meilleur contrôle du FDR que deux autres procédures de contrôles du FDR utilisées en pratique de façon courante. En démonstration, nous appliquons aussi la méthode proposée sur un jeu réel de données RNA-seq obtenu à partir du maïs.

A. Farcomeni, and L. Finos

606

FDR Control with Pseudo-Gatekeeping Based on a Possibly Data Driven Order of the Hypotheses

Nous proposons une procédure de tests multiples qui contrôle le taux de fausses découvertes. Elle repose sur un ordonnancement des hypothèses, éventuellement guidé par les données. Les hypothèses sont testées successivement à un niveau de risque non modifié q , jusqu'à ce qu'un nombre suffisant d'hypothèses ne soient pas rejetées. Quand

l'ordre est guidé par les données, les effets de plus grande taille sont testés les premiers, afin que les hypothèses biologiquement les plus intéressantes aient une probabilité de sélection plus élevée que les autres. La procédure est valide lorsque les statistiques de test sont indépendantes. Nous présentons une modification de la procédure qui en garantit la validité en présence de dépendances arbitraires. Une évaluation par simulation de notre méthode et de deux alternatives montre l'intérêt de notre approche, en particulier sur de petits échantillons. Nous présentons en conclusion une application à l'identification des signatures moléculaires de l'épendymome. Ces méthodes ont été implémentées sous R dans un paquet de programmes accessible gratuitement sur le CRAN. Des suppléments à cet article sont accessibles en ligne.

Simina M. Boca, Héctor Céorrada Bravo, Brian Caffo, Jeffrey T. Leek, and Giovanni Parmigiani 614

A Decision-Theory Approach to Interpretable Set Analysis for High-Dimensional Data

Un problème clé dans l'analyse de significativité en grande dimension est de trouver des groupes prédéfinis qui montrent un enrichissement pour le signal statistique d'intérêt ; l'exemple classique est l'enrichissement de groupes de gènes pour des gènes différentiellement exprimés. Ici, nous proposons une approche issue de la théorie de la décision pour l'analyse de groupes de gènes qui se focalise sur l'estimation de la fraction de variables non-nulles dans un groupe. Nous introduisons l'idée d'« atomes », des groupes disjoints basés sur les annotations de groupes prédéfinis. Notre approche se concentre sur la recherche de l'union d'atomes qui minimise une moyenne pondérée du nombre de fausses découvertes et de découvertes manquées. Nous introduisons un nouveau taux de fausses découvertes pour des groupes, appelé le taux de fausse découverte atomique (afdr), et prouvons que l'estimateur optimal dans notre approche issue de la théorie de la décision est le seuil de l'afdr. Ces résultats donnent un cadre cohérent et interprétable pour l'analyse de groupes qui aborde les questions clés des annotations chevauchantes et la difficulté d'interprétation des p -valeurs dans les tests compétitifs (en analyse de groupes, tests où l'hypothèse nulle exprime que les variables dans un groupe donné S sont au plus aussi fréquemment non-nulles que les variables dans le complément de S) et les tests autonomes (en analyse de groupes, tests où l'hypothèse nulle exprime que toutes les variables du groupe correspondent à la distribution nulle). Nous illustrons notre méthode et la comparons à une méthode existante courante à l'aide d'exemples simulés, et avec des analyses de groupes de gènes et des données de régions d'intérêt (ROI) du cerveau.

Denis Heng-Yan Leung, Dylan S. Small, Jing Qin, and Min Zhu 624

Shrinkage Empirical Likelihood Estimator in Longitudinal Analysis with Time-Dependent Covariates—Application to Modeling the Health of Filipino Children

La méthode d'équations d'estimation généralisées (GEE) est un outil populaire pour l'analyse de données longitudinales (en panel). Les covariables recueillies sont souvent dépendantes du temps par nature, comme par exemple, l'âge, l'état de récidence et le revenu mensuel. Lors de l'utilisation de GEE pour analyser des données longitudinales avec des covariables dépendantes du temps, des hypothèses cruciales sur les covariables sont nécessaires pour produire une inférence valide. Lorsque ces hypothèses ne sont pas ou ne peuvent pas être vérifiées, Pepe et Anderson (1994) préconisent comme approche robuste le recours à des matrices de travail indépendantes pour les corrélations dans les

modèles GEE. Cependant, utiliser la GEE avec cette l'hypothèse d'indépendance peut amener à une perte significative d'efficacité (Fitzmaurice, 1995). Dans cet article, nous proposons une méthode qui extrait des informations supplémentaires des équations d'estimation exclues par l'hypothèse d'indépendance. Le procédé repose toujours sur les équations d'estimation sous l'hypothèse d'indépendance mais la contribution des équations d'estimation restantes est pondérée en fonction de la vraisemblance de chaque équation à être une équation d'estimation consistante et de l'information qu'elle apporte. Nous appliquons la méthode à une étude longitudinale sur la santé d'un groupe d'enfants philippins.

Anestis Touloumis, Alan Agresti, and Maria Kateri

633

GEE for Multinomial Responses Using a Local Odds Ratios Parameterization

Dans cet article, nous proposons une approche à équations d'estimation généralisées (GEE) pour des réponses corrélées multinomiales ordinales ou catégorielles avec une paramétrisation en rapports de côtes locaux. Notre motivation repose sur l'observation que : (i) modéliser la dépendance entre des réponses multinomiales corrélées via les rapports de côtes locaux a du sens à la fois pour les échelles réponses ordinales et celles catégorielles, et (ii) les méthodes GEE ordinaires ne peuvent pas assurer l'existence conjointe des estimations des paramètres de régression marginale et de celles de la structure de dépendance. Pour éviter (ii), nous traitons le vecteur d'association η dit 'de travail' comme un vecteur de paramètres de nuisance qui définit la structure des rapports de côtes locaux aux tableaux de contingence marginalisés après agrégation sur tous les sujets de la variable réponse sans ajustement sur les variables explicatives et à chaque couple de temps. Pour estimer η et simultanément approcher correctement les éventuelles structures de dépendance sous-jacente, nous utilisons la famille des modèles d'association proposée par Goodman. Dans des simulations, pour un modèle probit cumulé marginal, les estimateurs des paramètres avec la méthode GEE proposée se révèlent être moins biaisés et plus efficaces que ceux obtenus avec le modèle 'de travail' indépendant, en particulier dans les études qui ont des variables explicatives dépendantes du temps et une forte corrélation.

Gabriela V. Cohen Freue, Hernan Ortiz-Molina, and Ruben H. Zamar

641

A Natural Robustification of the Ordinary Instrumental Variables Estimator

Les estimateurs de variables instrumentales sont conçus pour fournir des estimations consistantes des paramètres des modèles de régression linéaire lorsque certaines covariables sont corrélées avec le terme d'erreur. Nous proposons un nouvel estimateur des variables instrumentales *robuste* (RIV) qui est une extension naturelle de l'estimateur des variables instrumentales ordinaire (OIV). Spécifiquement, nous construisons RIV en utilisant un *S*-estimateur multivarié robuste de position et dispersion pour rendre la solution des équations d'estimation qui définissent OIV robuste. RIV est non coûteux en temps de calcul et facilement disponible pour des applications grâce à la bibliothèque R intitulée *riv*. Il présente des propriétés de robustesse et asymptotiques intéressantes, dont une grande résistance aux valeurs extrêmes, une fonction d'influence bornée, la consistance sous des hypothèses de distribution faibles, la normalité asymptotique sous des conditions de régularité peu contraignantes et l'équivariance. En outre, nous dotons RIV d'un algorithme itératif qui permet l'estimation de modèles avec covariables

endogènes continues et covariables exogènes en indicatrices. Nous étudions la performance de RIV quand les données contiennent des valeurs aberrantes par une vaste étude de simulation de Monte Carlo et en l'appliquant à la base de données d'accès limité de la cohorte de l'étude Framingham Cœur pour estimer l'effet de la pression artérielle systolique à long terme sur la taille auriculaire gauche.

Brian J. Reich and Luke B. Smith

651

Bayesian Quantile Regression for Censored Data

Dans cet article, nous proposons un modèle de régression quantile semi-paramétrique pour données de survie censurées. La régression quantile permet aux covariables d'affecter différemment la survie à différentes étapes de la période de suivi, fournissant ainsi une étude approfondie de la distribution de survie. Nous adoptons une approche semi-paramétrique, représentant le processus quantile comme une combinaison linéaire de fonctions de base. Les fonctions de base sont choisies de sorte que l'a priori pour le processus quantile est centré sur un simple modèle position-échelle, mais suffisamment souple pour s'adapter à un large éventail de processus quantile. Nous montrons dans une étude de simulation que cette approche est une alternative concurrente des autres méthodes existantes. La méthode est illustrée en utilisant des données d'une étude de traitement de la toxicomanie, où nous constatons qu'un modèle Bayésien donne souvent des mesures d'incertitude plus petites que ses concurrents, et donc identifie des effets plus significatifs.

Bradley W. McEvoy, Rajesh R. Nandy, and Ram C. Tiwari

661

Bayesian Approach for Clinical Trial Safety Data Using an Ising Prior

Dans le domaine de la sécurité du médicament, le développement de méthodes statistiques pour les problèmes d'ajustement de la multiplicité a exploité les relations potentielles entre effets adverses et caractères médicaux sous-jacents. En raison de la grosse granularité des caractéristiques biologiques utilisée pour regrouper les effets adverses, granularité qui sert de base à l'ajustement, il est possible qu'un effet adverse puisse en fait être décrit par plusieurs caractéristiques biologiques simultanément. Cependant les méthodes existantes sont limitées du fait qu'elles ne sont pas suffisamment flexibles structurellement pour prendre en compte le caractère multidimensionnel de l'effet adverse. Dans le but de préserver les dépendances complexes des données issues d'essais cliniques, une approche bayésienne modélisant les différentiels de risque de l'effet adverse entre le traitement testé et le traitement de base ou placebo. La procédure proposée fait appel à une loi *a priori* de Ising pour unifier les effets adverses médicalement liés. Cette méthode et une méthode bayésienne existante sont comparées sur un jeu de données d'essais cliniques, les résultats des deux méthodes étant confrontés. De plus, les résultats d'une petite étude sur données simulées sont aussi présentés.

Peter F. Thall, Hoang Q. Nguyen, Thomas M. Braun, and Muzaffar H. Qazilbash

673

Using Joint Utilities of the Times to Response and Toxicity to Adaptively Optimize Schedule-Dose Regimes

On propose un plan expérimental bayésien pour optimiser les chronologies et les doses

dans les essais de phase I-II d'un agent expérimental, en se basant sur les dates d'apparition de la réponse et de la toxicité dans le cas où les chronologies ne sont pas emboîtées et sont qualitativement différentes. La prise de décision séquentielle et adaptative est basée sur l'utilité conjointe des dates des deux événements. Pour construire une fonction d'utilité on divise en rectangles le quadrant positif du plan des dates possibles des événements et on attribue une valeur d'utilité à chacun de ces rectangles, puis on ajuste une fonction paramétrique régulière sur ces valeurs. On suppose que les dates de chaque événements sont distribuées selon une loi Gamma dont les paramètres de forme et d'échelle sont l'un et l'autre dépendant des dates et doses de la chronologie. On utilise une copule pour obtenir une distribution bivariable. Pour garantir le caractère éthique de l'essai on impose des conditions d'acceptabilité, en dates et en doses, en termes de sécurité et d'efficacité, sur la chronologie. Dans l'étape 1 de la planification les patients sont randomisés de façon équilibrée entre les chronologies et, pour chacune d'elles, on choisit une dose en utilisant un algorithme hybride qui maximise l'utilité moyenne à posteriori ou on randomise selon les doses acceptables. Dans l'étape 2 la randomisation équilibrée est remplacée par l'algorithme hybride. On utilise une version modifiée de cet algorithme pour les chronologies emboîtées. On adapte la démarche au cas des interruptions d'essais pour cause de décès par des extensions du modèle et de la fonction d'utilité. La méthode est illustrée par une application à un essai de transplantation de cellules souches autologues dans le traitement de myélome multiple, incluant une étude par simulation.

Na Cui, Yuguo Chen, and Dylan S. Small

683

Modeling Parasite Infection Dynamics when there Is Heterogeneity and Imperfect Detectability

Dans les maladies parasitaires, la connaissance des taux d'infection et de guérison représente un outil de planification précieux en Santé Publique. La modélisation de ces taux rencontre deux obstacles (1) le statut des individus n'est observé qu'à des temps discrets alors que les processus d'infection et de guérison se déroulent en temps continu ; (2) la détection de l'infection est entachée d'erreur. Nous traitons ces deux difficultés dans un modèle hiérarchique bayésien avec une distribution de Weibull sur les effets aléatoires. Ce modèle prend en compte l'hétérogénéité des taux d'infection et de guérison dans la population ainsi qu'une détection imparfaite de l'infection. Le modèle est estimé à l'aide d'un algorithme de Monte Carlo par chaîne de Markov avec augmentation des données. Nous présentons les résultats d'une évaluation par simulations et une application à une étude sur l'infection d'enfants par le parasite *Giardia Lamblia* au Kenya.

Christoph Bernau, Thomas Augustin, and Anne-Laure Boulesteix

693

Correcting the Optimal Resampling-Based Error Rate by Estimating the Error Rate of Wrapper Algorithms

Des tâches de classification binaire en grande dimension, par exemple la classification d'échantillons de puce à ADN en tissus normal ou cancéreux, impliquent habituellement un paramètre de réglage. En rapportant la performance de la seule valeur du meilleur paramètre de réglage, des erreurs de prédiction trop optimistes sont obtenues. Pour corriger ce biais de réglage, nous développons une nouvelle méthode qui est basée sur une décomposition du taux d'erreur inconditionnelle impliquant la procédure de réglage,

c'est-à-dire nous évaluons le taux d'erreur d'algorithmes « *wrapper* » comme présenté dans le contexte de validation croisée interne (ICV) par Varma et Simon (2006). Notre estimateur basé sur du sous-échantillonnage peut être écrit comme une moyenne pondérée des erreurs obtenues utilisant différentes valeurs de paramètre de réglage, et peut ainsi être interprété comme une version lissée d'ICV, qui est l'approche standard pour éviter ce biais de réglage. Par contraste avec l'ICV, notre méthode garantit des bornes intuitives pour l'erreur corrigée. De plus, nous suggérons d'utiliser des méthodes de correction de biais aussi pour aborder le biais de sélection de méthodes conceptuellement semblables qui résulte du choix optimal de la méthode de classification elle-même quand plusieurs méthodes sont évaluées successivement. Nous démontrons la performance de notre méthode sur des données de puces à ADN et des données simulées et le comparons à l'ICV. Cette étude suggère que notre approche conduit à des estimateurs compétitifs à un coût computationnel plus bas.

D. L. Borchers, W. Zucchini, M. P. Heide-Jørgensen, A. Cãnadás, and R. Langrock 703

Using Hidden Markov Models to Deal with Availability Bias on Line Transect Surveys

Nous développons des estimateurs pour des échantillonnages par transects en ligne d'animaux qui sont indisponibles à la détection de façon stochastique alors qu'ils sont présents dans la zone d'échantillonnage (animaux marins en plongée par exemple). Le processus de détection est formulé comme un modèle de Markov caché, avec un modèle d'observation binaire qui dépend des distances vers l'avant dans l'axe du transect et perpendiculairement à cet axe. Cette approche fournit une méthode paramétrique pour tenir compte de biais de détectabilité quand des estimations des paramètres du processus d'indisponibilité existent, même en l'absence de séries d'événements de disponibilité eux-mêmes.

Nous appliquons les estimateurs à un suivi de baleines aérien et par bateau, et nous en examinons les propriétés par simulation. Nous montrons qu'ils sont plus généraux et flexibles que les estimateurs existants basés sur des modèles paramétriques du processus de disponibilité. Nous montrons aussi que les méthodes utilisant des facteurs de corrections pour la disponibilité peuvent être sérieusement biaisés quand les suivis ne sont pas presque instantanés, comme peuvent l'être des estimateurs assumant une indépendance temporelle quand ce n'est pas le cas.

BIOMETRIC PRACTICE

Bibhas Chakraborty, Eric B. Laber, and Yingqi Zhao 714

Inference for Optimal Dynamic Treatment Regimes Using an Adaptive m-Out-of-n Bootstrap Scheme

Un régime dynamique de traitement consiste en un ensemble de règles de décision qui indiquent comment attribuer des traitements individualisés à des patients, connaissant l'histoire des traitements disponibles et des covariables. Une méthode usuelle pour estimer un régime dynamique optimal de traitement à partir des données est le Q-apprentissage qui implique des opérations non lisses sur les données. A cause de ce caractère non lisse, les approches asymptotiques standard pour l'inférence, telles que le bootstrap ou les arguments de séries de Taylor ne fonctionnent pas en l'absence de

corrections. Ici nous considérons le bootstrap *m de n* pour construire des intervalles de confiance pour les paramètres qui indexent le régime dynamique optimal. Nous proposons un choix adaptatif de *m* et montrons qu'il produit des ensembles de confiance asymptotiquement corrects sous alternatives fixées. De plus la méthode proposée a l'avantage d'être, conceptuellement et en termes de calcul, bien plus simple que les méthodes concurrentes qui possèdent la même propriété théorique. Nous produisons une étude de simulation extensive pour comparer la méthode proposée avec les procédures d'inférence actuellement disponibles. Les résultats suggèrent que la méthode proposée produit les taux de couvertures nominaux, tout en étant moins conservatrice que les méthodes concurrentes. Les méthodes proposées sont programmées en **R** dans le package **qlearn**, et sont disponibles sur le Comprehensive R-Archive Network (<http://cran.r-project.org/>). L'analyse de l'étude STAR*D sur les alternatives de traitements séquencés pour atténuer la dépression est utilisée comme exemple illustratif.

Changyu Shen, Jaesik Jeong, Xiaochun Li, Peng-Sheng Chen, and Alfred Buxton 724

Treatment Benefit and Treatment Harm Rate to Characterize Heterogeneity in Treatment Effect

Il est bien connu que l'estimation de l'effet du traitement sur la base d'une simple moyenne des patients a ses propres limites car elle ignore l'hétérogénéité des réponses au traitement dans la population cible. Cependant, il existe des métriques alternatives dans la littérature qui sont conçues pour capturer une telle hétérogénéité. Nous proposons le concept du taux de bénéfice du traitement (TBR) et le taux de dommages du traitement (THR) qui caractérisent à la fois l'effet global du traitement et de l'ampleur de l'hétérogénéité. Nous discutons d'une méthode pour estimer le TBR et le THR qui intègre facilement un schéma d'analyse de sensibilité, et illustrons l'idée à travers l'analyse d'une étude randomisée pour évaluer l'efficacité d'un défibrillateur implantable (ICD) dans la réduction de la mortalité. Une étude de simulation est présentée pour évaluer la performance de la méthode proposée.

Brent A. Johnson, Heather Ribaldo, Roy M. Gulick, and Joseph J. Eron, Jr. 732

Modeling Clinical Endpoints as a Function of Time of Switch to Second-line ART with Incomplete Data on Switching Times

La modélisation d'un paramètre clinique comme critère de changement de thérapie antivirale (ART) pour but de répondre à une question simple mais fondamentale : le changement d'ART est-il ou non bénéfique ? Nous proposons de répondre à une question scientifique similaire dans cet article, si ce n'est que nous nous intéressons à la modélisation du temps au bout duquel il faut changer d'ART plutôt qu'à la comparaison de deux ou plusieurs ART. La réponse à cette question scientifique est inconnue et a été difficile à exprimer cliniquement. Naturellement, le changement d'ART est décidé par le patient et son médecin, et la date du changement dépend ainsi des caractéristiques de chaque patient. Il y a une abondante littérature sur la prise en compte de facteurs de confusion potentiels et ces techniques sont fondamentales pour le succès de notre méthode. Un défi encore plus important est de construire une stratégie de modélisation systématique permettant de prendre en compte les délais manquants du fait que certains patients n'ont pas changé de traitement pour un ART de seconde ligne pendant la durée

de l'étude, même après échec du premier ART. Dans cet article, nous adaptons et appliquons une méthode statistique qui avait été initialement proposée pour modéliser les données de perfusion dans des essais ou la durée de perfusion peut être associée à une censure informative, en expliquant comment cette stratégie peut être appliquée ici. Notre application de cette méthode au traitement des pathologies HIV/SIDA est nouvelle et intéressante. A partir des données de l'essai A5095 de l'AIDS Clinical Trials Group (ACTG), nous avons modélisé les paramètres immunologiques selon une fonction polynomiale du temps de passage à un ART de seconde ligne pour les 1832 patients qui ont d'ores et déjà été en échec du traitement de première ligne. Dans notre analyse, nous avons montré que les patients qui changent précocement de traitement ont un niveau plus élevé de suppression de l'ARN HIV-1 après l'échec virologique que ceux qui changent plus tardivement. Cependant, nous avons aussi trouvé que les patients qui changent très tardivement de traitement, et qui pourraient être censurés à cause de la fin de l'étude, ont un bon niveau de suppression du RNA HIV-1 en moyenne. Nous pensons que nos conclusions scientifiques contribuent de façon significative à la littérature sur le HIV et souhaitons que la stratégie basique de modélisation proposée ici soit utile à d'autres équipes confrontées à des analyses similaires avec des données partiellement manquantes concernant la durée du traitement.

M. S. Shotwell, K. J. Drake, V. Y. Sidorov, and J. P. Wikswo
Mechanistic Analysis of Challenge–Response Experiments

741

Nous présentons une application de modèles mécanistiques et de régression longitudinale non-linéaire aux expériences du type contrainte-réponse ou stimulation-réponse où elles sont sous-utilisées. Dans ce genre d'expérience on impose des contraintes expérimentales et on observe la réponse qui en résulte. La combinaison de la modélisation mécanistique et de la régression longitudinale non-linéaire met en évidence de nouveaux aspects et révèle des perspectives inattendues de planification optimale. Plus précisément l'abord mécanistique de notre approche permet l'optimisation des paramètres (i.e. intensité et durée) de la contrainte ou de la stimulation. Cet article pose quelques bases de cette approche. Nous étudions des séries d'expériences où un cœur isolé de lapin est soumis à des périodes d'anoxie. Le cœur répond à l'installation de la contrainte et récupère quand elle cesse. La réponse moyenne est modélisée par un système d'équations différentielles censé décrire la réponse cardiaque à l'anoxie. Le système cardiaque se comporte de façon plus variable au repos et sous la contrainte. Ainsi les observations faites avec cette expérience montrent une hétéroscédasticité complexe et des changements brutaux de la réponse moyenne. Nous présentons la preuve que la stratégie de l'inférence statistique asymptotique peut ne pas rendre un compte adéquat de l'incertitude statistique. Deux autres méthodes sont critiquées qualitativement (dans le présent contexte) et quantitativement en utilisant une méthode de Monte Carlo innovante. Nous concluons avec une discussion sur les possibilités intéressantes de la planification des expériences stimulation-réponse ou contrainte-réponse.

Elina Numminen, Lu Cheng, Mats Gyllenberg, and Jukka Corander
*Estimating the Transmission Dynamics of *Streptococcus pneumoniae* from Strain Prevalence Data*

748

Le *Streptococcus pneumoniae* est une bactérie commensale typique provoquant des

maladies graves. Sa prévalence est élevée chez les jeunes enfants fréquentant les services de soins de jour en raison d'un plus faible niveau d'immunité acquise et un taux élevé de contacts infectieux entre personnes. Comprendre la dynamique des populations de différentes souches de *S. pneumoniae* est nécessaire, par exemple, pour faire des prédictions de changements dans la répartition des souches d'une communauté de bactéries dans le cadre de politiques d'intervention.

Dans cet article, nous analysons les données sur les souches de *S. pneumoniae* de patients transportés dans des unités de soins de jour dans la région métropolitaine d'Oslo, en Norvège. Nous introduisons une variante des méthodes de Calcul Bayésien Approximé (ABC), qui est approprié pour l'estimation des paramètres qui régissent la dynamique de transmission dans un contexte où de petites populations locales d'hôtes sont soumises à des épidémies de différentes souches pathogènes dus à des infections acquises indépendamment dans la communauté.

Nous montrons une forte concurrence entre les souches, par exemple l'acquisition d'autres souches au sein d'un hôte déjà colonisé est estimée à un taux relatif de 0,09 (IC de crédibilité à 95% [0.06, 0.14]). Nous prédisons également la fréquence et la taille des épidémies au sein de l'unité de soins de jour, ainsi que d'autres caractéristiques épidémiologiques pertinentes. L'hypothèse de neutralité écologique entre les souches semble être compatible avec les données. La validation du modèle et la cohérence de nos résultats avec des recherches antérieures confirment la validité de nos conclusions.

Chang Xuan Mao, Nan Yang, and Jinhua Zhong

758

On Population Size Estimators in the Poisson Mixture Model

L'estimation de tailles de populations par des expérimentations de type capture-recapture a d'énormes applications. Le modèle de mélange poissonnien a été adopté pour des applications à liste unique dans lesquelles chaque individu apparaît une ou plusieurs fois. Nous comparons plusieurs estimateurs non-paramétriques, incluant l'estimateur de Chao, l'estimateur de Zelterman, deux estimateurs de jackknife et l'estimateur du bootstrap. Le paramètre cible de l'estimateur de Chao est une borne inférieure de la taille de population. Pour les quatre autres estimateurs, la cible n'est pas une borne inférieure, et ils peuvent donc conduire à de plus faibles limites de confiance de la taille de population avec de faibles probabilités de couverture. Une étude par simulation est présentée, et deux exemples sont étudiés.

Simon J. Bonner and Jason Holmberg

766

Mark-Recapture with Multiple, Non-Invasive Marks

Des marques non-invasives, comme des patrons de coloration, des cicatrices ou des marqueurs génétiques, sont fréquemment utilisées pour identifier des individus dans des suivis de capture-recapture. Si les animaux peuvent être identifiés à partir de plusieurs types de telles marques non-invasives, alors certains individus peuvent figurer deux fois dans les données. Analyser les histoires observées sans tenir compte de ces erreurs peut conduire à des inférences erronées sur la dynamique des populations. Parmi les approches précédentes de ce problème, on peut traiter les données obtenues à partir d'un seul type de marque, et combiner les estimateurs obtenus en supposant qu'ils sont indépendants. Motivés par l'analyse de données du catalogue en ligne ECOCEAN du Requin-baleine (*Rhincodon typus*), nous décrivons une méthode Bayésienne pour analyser des données de

marques non-invasives multiples basée sur le modèle multinomial latent de Link et al. (2010). Nous décrivons en outre une simplification de l'algorithme de Monte-Carlo par chaînes de Markov de Link et al. (2010) qui permet des calculs plus efficaces. Nous présentons les résultats de l'analyse des données ECOCEAN de Requin-baleine, et, à l'aide de simulations, comparons notre méthode avec des approches existantes.

B. L. Robertson, J. A. Brown, T. McDonald, and P. Jaksons

776

BAS: Balanced Acceptance Sampling of Natural Resources

Pour planifier une enquête efficace ou un programme de surveillance d'une ressource naturelle, il est important de tenir compte de la répartition spatiale de la ressource. En général, les plans d'échantillonnage qui sont spatialement équilibrés sont plus efficaces que ceux qui ne le sont pas. Un plan spatialement équilibré sélectionne un échantillon qui est distribué de façon uniforme sur l'étendue de la ressource. Dans cet article, nous présentons un nouveau plan équilibré dans l'espace qui peut être utilisé pour sélectionner un échantillon de populations discrètes et continues dans un espace multidimensionnel. Le plan, que nous appelons échantillonnage d'acceptation équilibrée (BAS), utilise la séquence de Halton afin d'assurer la diversité spatiale des emplacements sélectionnés. Des probabilités d'inclusion spécifiques sont réalisées par échantillonnage d'acceptation. La conception *BAS* est conceptuellement plus simple que les conceptions concurrentes spatialement équilibrées, d'exécution plus rapide, et atteignent un meilleur équilibre spatial si on les mesure par un certain nombre de quantités. L'algorithme a été programmé dans un package libre R, disponible en téléchargement.

D. B. Woodard, T. M. T. Love, S. W. Thurston, D. Ruppert, S. Sathyanarayana, and S. H. Swan

785

Latent Factor Regression Models for Grouped Outcomes

Nous considérons les modèles de régression pour de multiples résultats corrélés, lorsque ces résultats sont emboîtés en domaines. Nous montrons que les modèles à effets aléatoires pour cette situation d'emboîtement donne un ajustement dans un cadre qui nous conduit à voir les options de modélisation comme un spectre entre les modèles parcimonieux à effets aléatoires et les modèles continus à facteurs latents. Nous introduisons un ensemble de modèles identifiables sur ce spectre, qui étend les modèles existants à effets aléatoires pour des résultats multiples emboîtés en domaines. Nous caractérisons le compromis entre parcimonie et flexibilité dans cet ensemble de modèles, en les appliquant à la fois à des données simulées et à des données associant les traits de dimorphisme sexuel chez les nourrissons mâles à des variables explicatives. Un appendice accessible en ligne fournit des éléments supplémentaires.