
Translations of Abstracts

BIOMETRIC METHODOLOGY**DISCUSSION PAPER**

O. Saarela, D. A. Stephens, E. E. M. Moodie, and M. B. Klein **279**
On Bayesian Estimation of Marginal Structural Models

L'objectif de la pondération de la probabilité inverse de traitement (IPT) dans l'estimation des effets marginaux du traitement est de construire une pseudo-population avec déséquilibres dans les covariables mesurées, écartant ainsi les effets de confusion et de censure informative pour la réalisation de l'inférence. Dans cet article, nous formalisons la notion d'une telle pseudo-population comme processus générateur de données à caractéristiques particulières et nous montrons que ceci conduit à une interprétation bayésienne naturelle de l'estimation pondérée par IPT. En utilisant cette interprétation, nous pouvons proposer la première procédure totalement bayésienne d'estimation des paramètres de modèles structurels marginaux à l'aide d'une pondération IPT. Notre approche suggère que les poids devraient être obtenus à partir de l'affectation de traitement prédictive a posteriori et des probabilités de censure, résolvant ainsi la question de savoir si et comment l'incertitude dans l'estimation des poids doit être incorporée dans l'inférence bayésienne des effets marginaux de traitement. L'approche proposée est comparée aux méthodes existantes sur des données simulées et appliquée également à l'analyse d'une cohorte canadienne dans un cadre de co-infection.

Y.-J. Cheng and M.-C. Wang **302**
Causal Estimation Using Semiparametric Transformation Models under Prevalent Sampling

Cet article développe des méthodes d'inférence pour l'estimation causale dans des modèles de transformation semi-paramétriques pour données de survie prévalentes. Grâce à l'estimation des modèles de transformation et de la distribution des covariables, nous proposons des procédures analytiques pour estimer la fonction de survie causale. Avec des données observationnelles, la réponse potentielle non observée (durée de survie) peut être associée à l'attribution du traitement et, donc, il peut exister un déséquilibre systématique entre les données observées dans chaque bras du traitement. En outre, en raison d'un recrutement prévalent, les sujets ne sont observés que s'ils n'ont pas expérimenté l'événement étudié avant le début de la collecte des données, provoquant ainsi un biais de sélection. Nous proposons une approche unifiée qui corrige simultanément le biais de sélection dû au recrutement prévalent et le déséquilibre systématique dû aux données observationnelles. Nous montrons dans une étude de simulation qu'une analyse standard sans un ajustement correct de ces aspects conduirait à des résultats biaisés pour l'inférence causale. Les propriétés asymptotiques des

procédures d'estimation sont établies par des techniques de processus empiriques et évaluées par des études de simulation. Les méthodes proposées sont appliquées aux données SEER-Medicare sur les femmes diagnostiquées avec un cancer du sein.

L. Li, T. Hanson, and J. Zhang

313

Spatial Extended Hazard Model with Application to Prostate Cancer Survival

Cet article développe une approche semi-paramétrique bayésienne pour le modèle à risque étendu, avec une généralisation à des données groupées spatialement de grande dimension. La corrélation spatiale au niveau d'un comté se traite marginalement à travers le modèle de transformation normale de Li et Lin (2006), en utilisant une structure de corrélation envisagée selon un a priori intrinsèque conditionnellement autorégressif. Des algorithmes efficaces de Monte Carlo par chaînes de Markov sont développés, spécialement appliqués à l'ajustement de jeux de données de survie très larges, aréolaires et hautement censurés. Des tests par variable pour les modèles à risques proportionnels, à durée de vie accélérée et à risques accélérés sont réalisés efficacement avec ou sans corrélation spatiale à travers des facteurs bayésiens. Les modèles spatiaux résultants, réduits et interprétables, s'ajustent significativement mieux que le modèle de Cox additif standard à fragilités spatiales.

T. M. Erhardt, C. Czado, and U. Schepsmeier

323

R-Vine Models for Spatial Time Series with an Application to Daily Mean Temperature

Nous présentons une extension des modèles de copule de R-Vine pour tenir compte des dépendances spatiales et un modèle de prédictions pour des zones non observées. Le modèle spatial R-Vine proposé combine la flexibilité des copules de Vine avec l'idée géographique classique de modéliser les dépendances à partir des distances entre les zones. En particulier, le modèle est capable de capturer les dépendances spatiales non gaussiennes. Pour développer et illustrer notre approche, nous considérons des données journalières des moyennes de températures observées à 54 stations d'enregistrement en Allemagne. Nous identifions les relations entre les paramètres de la copule de Vine et les distances des stations et les exploitons afin de réduire le nombre important de paramètres nécessaires pour paramétrer la R-Vine de dimension 54 adaptée aux données. La nouvelle distance basée sur la paramétrisation du modèle conduit à une réduction importante du nombre de paramètres et rend l'estimation des paramètres et la prédiction pour des zones non observées possibles. Les capacités de prédiction sont validées en utilisant des techniques de score adaptées et montrent une meilleure performance du modèle spatial de copule de R-vine par rapport au modèle spatial gaussien.

A. Arribas-Gil, R. De la Cruz, E. Lebarbier, and C. Meza

333

Classification of Longitudinal Data Through a Semiparametric Mixed-Effects Model Based on Lasso-Type Estimators

Nous proposons une méthode de classification pour données longitudinales. Le classificateur de Bayes est classiquement utilisé pour déterminer une règle de classification où la densité sous-jacente dans chaque classe doit être bien modélisée et estimée. Ce travail est motivé par un jeu de données réelles de niveaux hormonaux mesurés en début de grossesse qui peuvent être utilisés pour prédire des issues de

grossesse « normales » versus « anormales ». Le modèle proposé, qui est un modèle linéaire mixte semi-paramétrique (SLMM), est un cas particulier de la classe des modèles mixtes non-linéaires semi-paramétriques (SNMM) dans lequel une dimension finie (effets fixes et composantes de la variance) et une dimension infinie (une fonction inconnue) de paramètres doivent être estimés. Dans les SNMM, une estimation par maximum de vraisemblance est réalisée itérativement alternant des procédures paramétriques et non paramétriques. Cependant, si l'hypothèse que les effets aléatoires et la fonction inconnue interagissent de manière linéaire, des méthodes d'estimations plus efficaces peuvent être utilisées. Notre contribution porte sur la proposition d'une procédure d'estimation unifiée basée sur un algorithme de type EM pénalisé. Les étapes de l'Espérance et de la Maximisation sont explicites. A la dernière étape, la fonction inconnue est estimée de manière non paramétrique en utilisant une procédure de type lasso. Une étude de simulation et une application sont effectuées.

J. Goldsmith, V. Zipunnikov, and J. Schrack

344

Generalized Multilevel Function-on-Scalar Regression and Principal Component Analysis

Cet article concerne les modèles de régression pour des réponses fonctionnelles généralisées multi-niveaux : les fonctions sont *généralisées* en ce sens qu'elles suivent une distribution de la famille exponentielle et *multi-niveaux* en ce sens qu'elles sont rassemblées à l'intérieur de groupes ou de sujets. Une telle structure de données est de plus en plus répandue dans les domaines scientifiques et l'on en donne un exemple dans lequel des courbes binaires indiquant l'activité ou l'inactivité physique sont observées chez près de six cent sujets pendant cinq jours. Nous utilisons un modèle linéaire généralisé pour incorporer des covariables scalaires dans la structure de la moyenne et décomposer les déviations spécifiques des sujets ou spécifiques du couple sujet-jour au moyen d'une analyse fonctionnelle multi-niveaux en composantes principales. Ainsi, les effets fonctionnels fixes sont estimés en tenant compte des corrélations intra-fonctions et intra-sujets, et les directions principales de variabilité intra-sujets et inter-sujets sont identifiées. Les fonctions coefficients des effets fixes et les fonctions de base des composantes principales sont estimées au moyen de splines pénalisés ; les paramètres du modèle sont estimés dans un cadre bayésien avec **Stan**, un langage de programmation qui met en œuvre un échantillonneur Hamiltonien de Monte-Carlo. Les simulations adaptées à l'application ont de bonnes propriétés pour l'estimation et l'inférence avec des temps de calculs raisonnables pour des tailles de données modérées, à la fois pour des scénarios transversaux et multi-niveaux ; le code de **Stan** est disponible librement. Dans l'application, nous identifions les effets de l'âge et de l'indice de masse corporelle sur la modification dans le temps de la probabilité d'être actif sur une période de vingt-quatre heures ; de plus, l'analyse en composantes principales identifie les profils d'activité qui différencient les sujets et les jours pour un sujet.

Y. Li, B. Nan, and J. Zhu

354

Multivariate Sparse Group Lasso for the Multivariate Multiple Linear Regression with an Arbitrary Group Structure

Nous proposons une méthode d'estimation et de sélection de groupes de variables par lasso groupé épars multivarié pour des données comportant des prédicteurs de grande

dimension et des réponses de grande dimension également. La méthode est développée dans le cadre de la régression linéaire multiple multivariée pénalisée avec une structure de groupe arbitraire sur la matrice des coefficients de régression. Ceci est adapté à beaucoup d'études biologiques qui cherchent à détecter des associations entre des caractéristiques multiples et des prédicteurs multiples, avec chaque caractéristique ou prédicteur associé à un groupe biologique fonctionnel comme les gènes, des chemins ou des régions du cerveau. L'approche permet d'éliminer de manière efficace des groupes non importants ainsi que des coefficients individuels non importants dans des groupes importants, en particulier dans des problèmes avec p grand et n petit. Elle est flexible dans le traitement de différentes structures complexes de groupes comme les structures hiérarchiques imbriquées, multiniveaux ou superposées. L'approche est évaluée sur des simulations intensives avec comparaison aux méthodes conventionnelles du lasso et du lasso groupé et est appliquée à une étude d'association sur l'expression de locus de caractères quantitatifs.

J. de Uña-Álvarez and L. Meira-Machado

364

Nonparametric Estimation of Transition Probabilities in the Non-Markov Illness-Death Model: A Comparative Study

Les modèles multi-états sont souvent utilisés pour modéliser les données complexes d'histoires d'événements. Dans ces modèles, l'estimation des probabilités de transition est d'un intérêt particulier car ces probabilités permettent des prédictions à long terme des processus. Ces quantités ont été estimées traditionnellement par les estimateurs d'Aalen-Johansen qui sont consistants si le processus est Markovien. Dans le cas non-Markovien, plusieurs estimateurs ont été proposés dans la littérature récente et leur supériorité par rapport à l'estimateur d'Aalen-Johansen a été prouvée dans des conditions où l'hypothèse de Markov est fortement violée. Toutefois, les estimateurs existants présentent l'inconvénient de nécessiter que le support de la distribution de la censure contienne le support de la distribution de la durée de vie, ce qui n'est pas souvent le cas. Dans cet article, nous proposons deux nouvelles méthodes d'estimation des probabilités de transition dans le modèle sain-malade-mort progressif. Certains résultats asymptotiques sont proposés. Les estimateurs proposés sont consistants indépendamment de l'hypothèse de Markov et de la condition sur le support de la distribution de la censure. Nous explorons les propriétés en échantillon fini par des simulations. La principale conclusion de ce travail de recherche est que les estimateurs proposés sont beaucoup plus efficaces dans la plupart des cas que les autres estimateurs existants pour le cas non-Markovien. Une application à un essai clinique est proposée. Les extensions aux modèles progressifs au-delà du modèle à trois états sain-malade-mort sont discutées.

J. Goldstein, M. Haran, I. Simeonov, J. Fricks, and F. Chiaromonte

376

An Attraction-Repulsion Point Process Model for Respiratory Syncytial Virus Infections

Comment la progression d'un virus est-elle influencée par les propriétés intrinsèques de cellules individuelles ? Nous nous intéressons à cette question en étudiant la susceptibilité de cellules infectées par deux souches du virus respiratoire syncytial humain (VRS-A et VRS-B) dans une étude expérimentale *in vitro*. L'observation de motifs spatiaux de cellules infectées nous fournit un éclairage sur la façon dont les conditions locales

influent la susceptibilité au virus. Nous observons un comportement complexe d'attraction et de répulsion, une tendance des cellules infectées à s'agréger ou au contraire à rester séparées. Nous développons un nouveau processus ponctuel spatial afin de décrire ce comportement. L'inférence sur les processus ponctuels spatiaux est difficile parce que les fonctions de vraisemblance de ces modèles contiennent des constantes de normalisation impossibles à évaluer. Nous adaptons un algorithme MCMC appelé double Metropolis-Hastings afin de surmonter ces difficultés calculatoires. Nos méthodes sont efficaces sur le plan calculatoire même pour des motifs ponctuels de grande taille consistant en plus de 10 000 points. Nous illustrons l'application de notre modèle et l'approche inférentielle développée avec des données simulées et ajustons notre modèle à plusieurs études expérimentales sur le VRS. Grâce à la facilité d'interprétation des paramètres de notre modèle, nous sommes en mesure de tirer des conclusions scientifiques intéressantes des modèles ajustés obtenus.

D. Follmann and C.-Y. Huang

386

Incorporating Founder Virus Information in Vaccine Field Trials

Les essais cliniques de vaccins avec surveillance active de l'infection utilisent souvent le délai de survenue de l'infection comme critère principal. Une méthode d'analyse courante pour de tels essais est de comparer ces délais d'infection entre des groupes vaccin et placebo au moyen du modèle de Cox. Grâce aux nouvelles technologies, nous pouvons en plus enregistrer le nombre de virions qui causent l'infection plutôt que seulement l'indicateur que l'infection s'est produite. Dans cet article nous développons une approche unifiée pour les essais de vaccins qui associe le délai de survenue de l'infection et le nombre de virus infectants ou fondateurs. Nous supposons que le risque instantané d'exposition potentiellement infectieuse des individus dans les groupes placebo et vaccin suit le même modèle à intensité proportionnelle. Après exposition, le nombre de virus fondateurs X^* est supposé issu d'une distribution sur $0, 1, \dots$, qui peut être différente pour les deux groupes. Les expositions qui conduisent à $X^*=0$ ne sont pas observables. Nous notons μ et $\mu\Delta$ les moyennes sous placebo et vaccin, si bien que $1-\Delta$ mesure la réduction en proportion du nombre moyen de virions infectants du fait du vaccin. Nous développons différentes méthodes semi-paramétriques pour estimer Δ . Nous autorisons la distribution de X^* à être de Poisson ou bien non spécifiée, et discutons comment incorporer des covariables qui influent sur le temps d'exposition et/ou sur X^* . Curieusement Δ , qui est un rapport de deux moyennes non tronquées, peut être estimé d'une façon fiable au moyen de données tronquées ($X^*>0$), même si les distributions placebo et vaccin sont complètement non-spécifiées. Des simulations d'essais cliniques de vaccins montrent que la méthode permet d'estimer Δ de façon fiable dans des conditions réalistes. Nous appliquons nos méthodes à un essai de vaccin HIV conduit chez des utilisateurs de drogues injectées.

X. Cheng, W. Lu, and M. Liu

397

Identification of Homogeneous and Heterogeneous Variables in Pooled Cohort Studies

Les analyses par mise en commun (« pooled ») intègrent des données de multiples études et obtiennent ainsi un échantillon élargi en vue d'améliorer la puissance statistique. Lorsqu'on est en présence d'hétérogénéité entre études pour les effets des variables sur le résultat étudié, la stratégie de simple mise en commun échoue à obtenir une

représentation correcte et complète des effets des variables hétérogènes. Par conséquent, il est important d'étudier l'homogénéité et l'hétérogénéité pour la structure des variables dans les études par mise en commun. Dans cet article, nous nous intéressons à la mise en commun d'études de cohortes avec des critères d'évaluation de type « délai jusqu'à l'événement » et nous proposons une approche de vraisemblance partielle pour modèle de Cox pénalisé avec des pénalités adaptatives pondérées sur les effets homogènes et hétérogènes des variables. Nous montrons que notre méthode peut caractériser les variables selon des effets hétérogènes, homogènes ou nuls et estimer les effets non nuls. Les résultats sont facilement extensibles aux applications en plus haute dimension lorsque le nombre de paramètres est plus grand que la taille d'échantillon. La procédure proposée de sélection et d'estimation peut être implémentée en utilisant l'algorithme de tir itératif. Nous effectuons des études numériques extensives pour évaluer la performance de la méthode que nous proposons et nous l'appliquons en utilisant une étude par mise en commun sur des patientes atteintes de cancer de l'ovaire et concernant l'expression de gènes.

O. Komori, S. Eguchi, and J. B. Copas

404

Generalized t-Statistic for Two-Group Classification

Dans le modèle discriminant classique de deux distributions normales avec des matrices de variances égales, la fonction discriminante linéaire est optimale à la fois en termes de rapport de vraisemblances et en termes de maximisation de la différence standardisée (statistique t) entre des moyennes de deux distributions. Dans une étude cas-témoins typique, supposer la normalité peut être raisonnable pour l'échantillon des témoins mais l'hétérogénéité et l'incertitude dans le diagnostic peuvent suggérer qu'un modèle plus flexible est nécessaire pour les cas. Nous généralisons l'approche par la statistique t en déterminant la fonction linéaire qui maximise une différence standardisée mais avec des données de l'un des groupes (les cas) filtrées par une fonction U qui peut être non-linéaire. Nous étudions les conditions de consistance de la méthode et trouvons la fonction U qui est optimale au sens de l'efficacité asymptotique. L'optimalité peut également s'étendre à d'autres mesures d'efficacité discriminante telle que l'aire sous la courbe ROC (Receiver Operating Characteristic). La fonction optimale U dépend d'une fonction de densité de probabilité scalaire qui peut être estimée de manière non-paramétrique en utilisant un algorithme numérique standard. Une version de lasso pour la sélection des variables est implémentée en ajoutant une régularisation L1 à la statistique t généralisée. Deux jeux de données de puces à ADN dans l'étude de l'asthme et de nombreux cancers sont utilisés comme exemples d'illustration.

C. Baayen, P. Hougaard, and C. B. Phipps

417

Testing Effect of a Drug Using Multiple Nested Models for the Dose-Response

Au cours du développement d'un médicament, la détermination de la dose optimale se fait lors d'un essai clinique de phase II où les doses évaluées sont incluses avec un placebo. Deux méthodes statistiques sont classiquement utilisées pour analyser ces essais de recherche de dose : la comparaison séparée de chaque dose au placebo (en utilisant une procédure de comparaisons multiples) ou une stratégie basée sur un modèle (en ajustant un modèle dose-effet sur toutes les données). La première approche fonctionne mieux lorsque les patients sont concentrés sur quelques doses mais ne peut pas conclure

pour les doses non testées. Les méthodes basées sur un modèle dose-effet permettent l'interpolation entre les doses mais la validité dépend de l'exactitude du modèle supposé. Bretz et al. (2005; *Biometrics* 61, 738-748) ont suggéré une approche combinée qui sélectionne un ou plusieurs modèles appropriés à partir d'un ensemble de modèles candidats en utilisant une procédure de comparaisons multiples. La méthode nécessite une estimation a priori de tous les paramètres non-linéaires des modèles candidats de sorte qu'il y a toujours un certain degré possible de mauvaise spécification du modèle. De plus, seule une ou plusieurs restrictions au modèle général peuvent être testées. Nous proposons une procédure alternative de tests multiples qui évalue un ensemble de modèles dose-effet plausibles les uns contre les autres pour aboutir à la sélection d'un modèle final. Notre méthode ne nécessite pas d'estimation a priori et contrôle le taux d'erreur de type I de choisir un modèle trop complexe.

T. Chekouo, F. C. Stingo, J. D. Doeck, and K.-A. Do

428

miRNA–Target Gene Regulatory Networks: A Bayesian Integrative Approach to Biomarker Selection with Application to Kidney Cancer

La disponibilité de données génomiques à grande échelle provenant de multiplateformes a permis l'étude de relations biologiques complexes dans beaucoup de cancers. L'identification de biomarqueurs fiables dans le domaine du cancer nécessite la caractérisation de multiples interactions dans des réseaux génétiques complexes. Les micro-ARN sont des ARN non codants de petite taille qui régulent l'expression des gènes. Cependant, la relation directe entre un micro-ARN et ses gènes cibles est difficile à mesurer. Nous proposons une méthode bayésienne innovante afin d'identifier les micro-ARN et leurs gènes cibles qui sont associés à la survie en incorporant le réseau de régulation des micro-ARN dans l'expression de la distribution a priori. Nous supposons que les biomarqueurs impliqués dans les réseaux de régulation sont susceptibles d'être associés à la survie. Nous utilisons des distributions a priori non locales et une méthode de recherche stochastique pour sélectionner les biomarqueurs associés à la survie. Nous utilisons la voie d'information KEGG afin d'incorporer les effets génétiques corrélés dans les réseaux de régulation. Nous évaluons la performance de notre méthode au moyen de simulations et l'appliquons à des données expérimentales sur le carcinome rénal obtenues à partir de l'atlas génomique du cancer (« Cancer Genome Atlas »). Notre méthode innovante permet de valider des biomarqueurs précédemment identifiés et d'identifier des biomarqueurs spécifiques de la progression du carcinome rénal qui n'avaient pas encore été identifiés. À partir des données sur le carcinome rénal, nous confirmons que les biomarqueurs impliqués dans les réseaux de régulation sont plus susceptibles d'être associés à la survie en montrant des connexions dans un réseau de régulation pour cinq de ces six gènes identifiés.

W. Shen, J. Ning, and Y. Yuan

439

A Direct Method to Evaluate the Time-Dependent Predictive Accuracy for Biomarkers

Les courbes ROC qui dépendent du temps et leur aire sous la courbe (AUC) sont des mesures importantes pour évaluer la qualité de la prédiction faite à partir de biomarqueurs pour des critères liés au temps (par exemple le délai jusqu'à la progression de la maladie ou au décès). Dans cet article, nous proposons une méthode directe pour estimer la fonction AUC(t) en utilisant des modèles polynomiaux fractionnés flexibles,

sans l'étape intermédiaire de la modélisation de la courbe ROC dépendante du temps. Nous développons une procédure basée sur une pseudo-vraisemblance partielle pour l'estimation des paramètres et fournissons une procédure de test pour comparer la performance prédictive entre biomarqueurs. Nous établissons les propriétés asymptotiques de l'estimateur proposé et des statistiques de test. Un avantage majeur de la méthode proposée est sa facilité à réaliser une inférence et comparer la précision de la prédiction entre biomarqueurs, ce qui la rend particulièrement attrayante pour les études qui nécessitent la comparaison et la sélection d'un grand nombre de biomarqueurs candidats. Nous évaluons la performance de la méthode proposée dans le cas d'échantillons finis au moyen d'études de simulation et illustrons notre méthode par une application à des données du groupe coopérateur « AIDS Clinical Trials Group 175 ».

BIOMETRIC PRACTICE

Y. K. Cheung, B. Chakraborty, and K. W. Davidson

450

Sequential Multiple Assignment Randomized Trial (SMART) with Adaptive Randomization for Quality Improvement in Depression Treatment Program

Les études d'implémentation représentent un outil important dans la mise en place des nouveaux traitements pour passer des études cliniques d'évaluation de l'efficacité de ces traitements à des programmes thérapeutiques ayant le double objectif de mieux connaître l'efficacité réelle de ces traitements et d'améliorer la qualité des soins des patients inclus dans ces programmes. Dans cet article, nous nous intéressons aux programmes avec un schéma dynamique de traitement (DTR) chez des patients ayant une dépression survenant après un syndrome coronarien aigu. Nous introduisons un schéma adaptatif innovant de randomisation pour une attribution séquentielle multiple dans un essai randomisé avec un schéma DTR. Notre approche adapte les probabilités de randomisation pour favoriser les séquences de traitement ayant des Q-fonctions utilisées dans le Q-apprentissage comparativement supérieures. L'approche proposée résout trois problèmes principaux des études d'implémentation : elle permet l'incorporation de données historiques ou d'opinions d'experts, elle inclut le tirage au sort à des fins d'apprentissage et elle vise à améliorer le soin des patients à travers l'adaptation tout au long de la durée du programme. Nous montrons comment appliquer notre méthode afin de concevoir un programme de traitement de la dépression en utilisant des données d'une précédente étude. Au moyen de simulations, nous illustrons que l'apport de données historiques est important pour la performance du programme telle qu'elle est mesurée par les réponses attendues des patients participant au programme mais montrons également que le schéma de randomisation adaptative est capable de compenser la mauvaise spécification de données historiques par l'amélioration des résultats des patients dans un horizon de temps raisonnable. Les résultats de simulations confirment également que le schéma proposé permet un apprentissage efficace des traitements en réduisant la grande difficulté que représente la grande dimension du problème.

A. S. C. Conlon, J. M. G. Taylor, and D. J. Sargent

460

Improving Efficiency in Clinical Trials Using Auxiliary Information: Application of a Multi-State Cure Model

Dans les études cliniques, un marqueur intermédiaire mesuré après randomisation peut

souvent apporter des informations précoces à propos de l'effet du traitement sur la valeur finale de la variable d'intérêt. Dans cet article, nous explorons l'utilisation des délais de rechute comme une variable auxiliaire pour estimer l'effet du traitement sur la survie globale dans les études cliniques de Phase III sur le cancer du côlon, avec randomisation des traitements. Un modèle multi-états tenant compte d'une fraction de population « guérie » est utilisé pour modéliser conjointement délai de rechute et délai de survie. Dans ce cadre, nous explorons différentes voies pour lesquelles l'information sur le délai de rechute et les hypothèses sur les modèles statistiques utilisés puissent conduire à une amélioration de l'efficacité. Les estimateurs de la survie globale et de la survie sans récurrence peuvent être obtenus de ce modèle avec un gain d'efficacité comparativement aux estimateurs de Kaplan-Meier. De façon alternative, des gains en efficacité peuvent être obtenus en utilisant ce même modèle au sein d'une démarche plus simple dans une procédure d'imputation multiple capable d'imputer des délais de survie aux patients à suivi censuré. En utilisant le « modèle joint », la rechute est utilisée comme une variable auxiliaire prédictive du délai de survie. Nous démontrons que l'utilisation éventuelle des méthodes proposées conduirait à raccourcir la durée d'une étude clinique ainsi que son nombre de patients.

N. Jeffries and N. L. Geller

469

Longitudinal Clinical Trials with Adaptive Choice of Follow-Up Time

Dans les études longitudinales comparant deux traitements avec un temps de suivi maximum, il peut être intéressant d'évaluer l'effet du traitement pour des suivis intermédiaires. Une motivation à cela est d'identifier le moment où la différence entre traitements est la plus grande lorsque l'effet du traitement n'est pas monotone au cours du temps. Une autre motivation est de permettre de prendre une décision plus rapidement quant à l'efficacité ou l'inefficacité du nouveau traitement. Ici, nous développons un test avec une hypothèse nulle composite - il n'y a pas de différence entre les traitements quel que soit le temps de suivi - versus l'hypothèse alternative - il existe une différence à au moins l'un des temps de suivi. Les méthodes sont pertinentes quand peu de mesures sont réalisées au cours du temps. C'est le cas dans les essais thérapeutiques précoces avec peu de suivi ou dans les études ancillaires. On suppose que la statistique de test Z_{t_k} est utilisée pour tester l'hypothèse qu'il n'y a pas d'effet traitement à un temps de suivi fixé t_k . Dans ce contexte, une approche classique est de réaliser une étude pilote sur N_1 sujets, d'évaluer l'effet traitement à différents temps fixés t_1, \dots, t_K et de choisir t^* comme étant la valeur de t_k pour laquelle Z_{t_k} est maximisée. Une fois t^* choisi, un second essai peut être planifié. Dans le cadre d'une procédure de tests séquentiels groupés, nous avons considéré plusieurs alternatives à l'approche classique. Il s'agit d'alternatives adaptatives qui traitent l'étude pilote et le second essai comme une entité combinée homogène. Nous avons évalué l'erreur de type I et la puissance pour ces alternatives. Les schémas adaptatifs considérés ici ont en général amélioré la puissance par rapport à l'approche classique des essais séparés.

B. A. Goldstein, T. Assimes, W. C. Winkelmayr, and T. Hastie

478

Detecting Clinically Meaningful Biomarkers with Repeated Measurements: An Illustration with Electronic Health Records

Les données de mesures répétées sont une ressource remarquable pour comprendre la

relation entre les changements d'un marqueur biologique et le risque d'un événement clinique.

Alors que les données longitudinales fournissent l'opportunité d'observer des changements de ce risque au cours du temps, ces analyses peuvent être très complexes lorsque les mesures de marqueurs sont rares ou irrégulières, rendant inapplicables les méthodes statistiques usuelles.

Dans cet article, nous utilisons des données du dossier médical informatisé (DMI) pour proposer une procédure créant un échantillon d'analyse et y associant une méthode d'analyse dans le but de détecter des marqueurs significatifs de l'apparition d'un infarctus du myocarde (MI). Utilisant les DMI d'un grand groupe de patients dialysés, nous avons extrait les dossiers de 64 318 individus parmi lesquels 4 769 MI ont été identifiés durant la période d'étude. Nous décrivons le schéma d'étude cas-témoins nichée que nous avons développé pour échantillonner un groupe témoin approprié et l'approche analytique utilisant des régressions par splines.

En utilisant un modèle mixte avec des fonctions splines de puissance tronquée, nous proposons une série de tests d'adéquation afin de déterminer quels sont les prédicteurs utiles en clinique parmi 11 marqueurs de laboratoire régulièrement collectés. Nous testons la pertinence clinique de chacun des marqueurs en utilisant un jeu de données indépendant. Nos résultats suggèrent que les données du DMI peuvent être ainsi facilement utilisées pour identifier les marqueurs pertinents de la survenue d'événements cliniques aigus.

Aucun programme particulier, pas plus que des outils d'analyse spécifiques ne sont nécessaires, même dans le cas de données irrégulièrement recueillies dans le DMI.

B. Jiang, M. R. Elliott, M. D. Sammel, and N. Wang

487

Joint Modeling of Cross-Sectional Health Outcomes and Longitudinal Predictors via Mixtures of Means and Variances

Les méthodes de modélisation conjointes sont devenues des outils populaires pour relier des caractéristiques importantes extraites de données longitudinales à un événement primaire. Alors que la plupart des stratégies de modélisation ont porté sur l'association entre les trajectoires moyennes longitudinales et le risque d'un événement, nous considérons des modèles conjoints qui intègrent les informations des tendances à long terme et de la variabilité

à court terme dans un sous-modèle longitudinal. Nous considérons également à la fois un effet aléatoire et une approche par classe latente dans le modèle primaire pour prédire un résultat binaire intéressant. Nous développons des études de simulation afin de comparer ces deux stratégies de modélisation ; en particulier, nous étudions en détail les effets d'une mauvaise spécification du premier modèle. Parmi les autres résultats, nous constatons que lorsque nous analysons les données à partir d'un effet aléatoire partagé en utilisant un modèle de structure latente tandis que l'information à partir des données longitudinales est faible, l'approche par classe latente est plus sensible à une telle mauvaise spécification du modèle. En paramétrant de cette façon, le modèle de structure latente a une performance supérieure dans la prédiction au sein de l'échantillon qui ne peut être reproduite pour prédire de nouveaux échantillons. C'est une caractéristique unique de l'approche de structure latente qui est nouvelle, autant que nous le sachions par rapport à la littérature existante. Enfin, nous utilisons les modèles proposés pour étudier comment les trajectoires de l'hormone folliculo-stimulante (FSH) sont liées au risque de

développer des bouffées de chaleur sévères pour les femmes participant à l'étude de vieillissement ovarien de Penn.

M. Moreno-Betancur, G. Rey, and A. Latouche

498

Direct Likelihood Inference and Sensitivity Analysis for Competing Risks Regression with Missing Causes of Failure

En analyse de survie, les risques concurrents considèrent, en plus du délai de survenue d'un évènement, sa cause. Fréquemment, il est difficile de connaître les causes d'évènement pour tous les sujets. C'est pourquoi plusieurs auteurs ont proposé des stratégies pour ajuster des modèles semi-paramétriques pour des données de survie avec risques concurrents lorsque certaines causes d'évènement manquent, sous l'hypothèse qu'elles sont manquantes au hasard (MAH). Comme plusieurs auteurs l'ont noté, même si les modèles semi-paramétriques sont pratiques, les modèles (complètement) paramétriques pour les risques cause-spécifique (RCS) et les fonctions d'incidence cumulée (FIC) peuvent être pertinents pour la prédiction et pour mieux comprendre la dynamique temporelle des risques concurrents. Nous proposons d'employer une méthode de maximisation de la vraisemblance pour ajuster ces modèles sous l'hypothèse MAH. Cette hypothèse n'est cependant pas vérifiable à partir des données observées. Nous proposons donc une méthode pour mener des analyses de sensibilité afin d'évaluer la robustesse des inférences aux violations de cette hypothèse. Cette méthode repose sur les modèles de mélange de profils et a été évaluée dans une étude de simulation. Cette approche est applicable à plusieurs modèles de régression (paramétriques ou semi-paramétriques, pour le RCS ou la FIC). Nous illustrons les méthodes proposées avec l'analyse d'un essai clinique dans le cancer du sein et suggérons des méthodes graphiques ad hoc pour évaluer la qualité de l'ajustement sous l'hypothèse MAH.

L. B. Smith, B. J. Reich, A. H. Herring, P. H. Langlois, and M. Fuentes

508

Multilevel Quantile Function Modeling with Application to Birth Outcomes

Les nourrissons nés prématurément ou avec un poids de naissance faible par rapport à l'âge gestationnel ont des taux de morbidité et de mortalité élevés. En utilisant les archives des actes de naissance du Texas entre 2002 et 2004 et les estimations de l'Agence de Protection de l'Environnement (EPA) sur la pollution de l'air, nous étudions le lien entre les fonctions quantiles du poids de naissance et de l'âge gestationnel avec l'exposition à l'ozone ainsi que de multiples facteurs prédictifs dont l'âge des parents, l'ethnie, et le niveau d'éducation. Nous proposons une approche bayésienne semi-paramétrique basée sur les quantiles qui modélise la totalité de la fonction quantile plutôt que seulement quelques niveaux de quantiles. Notre modèle de la fonction quantile multi-niveaux établit des liens entre le poids de naissance et les facteurs prédictifs pris séparément pour chaque semaine d'âge gestationnel ainsi qu'entre l'âge gestationnel et les facteurs prédictifs pris séparément selon les régions de santé publique du Texas. Nous permettons à ces relations des variations non linéaires selon l'âge gestationnel, le domaine spatial, et le niveau de quantile et nous les réunissons dans un modèle hiérarchique via une expansion de base sur les coefficients de régression afin de préserver leur caractère interprétable. Le très faible poids de naissance étant une préoccupation majeure, nous utilisons la théorie de la valeur extrême pour compléter notre modèle au niveau de la queue de la distribution. Les âges gestationnels étant enregistrés en semaines

complètes de gestation (valeurs entières), nous présentons une méthodologie pour modéliser les fonctions quantiles de données discrètes. Dans une étude de simulation, nous montrons que regrouper l'information selon l'âge gestationnel et le niveau de quantile réduit considérablement l'erreur quadratique moyenne (MSE) des effets prédictifs. Nous constatons que l'ozone est négativement associé aux plus faibles valeurs de l'âge gestationnel dans le sud du Texas et à la distribution du poids de naissance pour les valeurs élevées de l'âge gestationnel. Nos méthodes sont disponibles dans la bibliothèque (« package ») « Bsquare » du logiciel R.

R. Langrock, T. Kneib, A. Sohn, and S. L. DeRuiter

520

Nonparametric Inference in Hidden Markov Models Using P-Splines

Les modèles de Markov cachés (HMM) sont des modèles de séries temporelles flexibles dans lesquels la distribution des observations dépend d'états autocorrélés non observés. Les distributions de ces états cachés sont sélectionnées en général dans une classe de distributions paramétriques. Le choix de cette classe peut être difficile et un mauvais choix peut avoir de graves conséquences par exemple sur les estimations des états cachés et plus généralement sur la complexité et l'interprétation du modèle résultant. Nous illustrons ces questions pratiques avec des données réelles sur la vitesse de plongée des baleines à bec. Nous démontrons que les approches paramétriques peuvent conduire à des processus trop complexes, empêchant ainsi une bonne interprétation biologique. En revanche, les estimations non paramétriques sur ces données donnent des modèles HMM beaucoup plus parcimonieux concernant le nombre d'états cachés et plus faciles à interpréter, tout en conservant un bon ajustement aux données. Notre méthode d'estimation non paramétrique est basée sur l'idée de représenter les densités de distribution des états cachés comme des combinaisons linéaires d'un grand nombre de fonctions de base B-splines standardisées, ce qui impose l'utilisation d'un terme de pénalité afin de maintenir un bon compromis entre ajustement et lissage.

M. L. Neely, H. D. Bondell, and J.-Y. Tzeng

529

A Penalized Likelihood Approach for Investigating Gene–Drug Interactions in Pharmacogenetic Studies

La pharmacogénétique étudie la relation entre la variabilité génétique héritable et la variabilité dans la réponse des individus aux traitements médicamenteux. Souvent, les interactions gènes-médicaments jouent un rôle primordial dans cette réponse et l'identification de ces effets peut aider à l'élaboration de régimes de traitement individualisés. Les haplotypes peuvent détenir une information clé pour comprendre l'association entre la variabilité génétique et la réponse aux médicaments. Cependant, l'approche standard pour l'analyse d'association basée sur les haplotypes ne concerne pas directement les questions de recherche dictées par la médecine individualisée. Une analyse post-hoc complémentaire est nécessaire et cette analyse post-hoc est généralement sous puissante après ajustement pour les comparaisons multiples et peut conduire à des conclusions apparemment contradictoires. Dans ce travail, nous proposons une approche de vraisemblance pénalisée qui est capable de surmonter les inconvénients de l'approche standard et d'obtenir le résultat personnalisé souhaité. Nous démontrons l'utilité de notre méthode en l'appliquant à l'essai randomisé écossais du cancer de l'ovaire. Nous avons également mené des études de simulation et montré que

la méthode pénalisée proposée a une puissance comparable ou supérieure à celle de l'approche standard et maintient de faibles taux d'erreur de type I à la fois pour les réponses aux médicaments binaires et quantitatives. Les plus grands gains de performance sont observés lorsque la fréquence des haplotypes est faible, la différence de tailles d'effet est de petite taille ou la vraie relation entre les médicaments est plus complexe.

Y. Liu, Y. Chen, and H. Chu

538

A Unification of Models for Meta-Analysis of Diagnostic Accuracy Studies without a Gold Standard

Plusieurs méthodes statistiques pour la méta-analyse des études de précision de diagnostic ont été discutées en présence d'une référence. Cependant, dans la pratique, le test de référence sélectionné peut être imparfait en raison d'une erreur de mesure, de sa non-existence, de sa nature invasive ou d'un coût élevé de sa réalisation. Il a été suggéré que le traitement d'un test de référence imparfait comme référence peut conduire à un biais important dans l'estimation de la précision du test de diagnostic. Récemment, deux modèles ont été proposés pour tenir compte d'un test de référence imparfait, à savoir, un modèle linéaire mixte généralisé multivarié (MGLMM) et un modèle ROC (Receiver Operating Characteristic) de synthèse hiérarchique (HSROC). Les deux modèles sont très flexibles dans la prise en compte de l'hétérogénéité de précision des tests entre les études ainsi que dans celle de la dépendance entre les tests. Dans cet article, nous montrons que ces deux modèles, malgré différentes formulations, sont étroitement liés et sont équivalents en l'absence de variables au niveau de l'étude. En outre, nous fournissons les relations exactes entre les paramètres de ces deux modèles et les hypothèses sous lesquelles les deux modèles peuvent être réduits à des sous-modèles équivalents. D'autre part, nous montrons que certains sous-modèles des MGLMM n'ont pas de sous-modèles HSROC équivalents correspondants et vice versa. Avec trois exemples réels, nous illustrons les cas où ajuster des modèles MGLMM et HSROC conduit à des sous-modèles équivalents et donc à une inférence identique et les cas où les inférences des deux modèles sont légèrement différentes. Nos résultats généralisent les relations importantes entre le modèle linéaire mixte bivarié et le modèle HSROC lorsque le test de référence est une norme de référence.

N. Demetrashvili and E. R. Van den Heuvel

548

Confidence Intervals for Intra-class Correlation Coefficients in a Nonlinear Dose-Response Meta-Analysis in a nonlinear dose-response meta-analysis

Ce travail a été motivé par une méta-analyse « case study » sur les médicaments antipsychotiques. La courbe de Michaelis-Menten est utilisée pour modéliser la relation non-linéaire entre la dose et la réponse du récepteur D2 dans de multiples études. Un coefficient de corrélation intra-classe (CCI) est utilisé pour quantifier l'hétérogénéité entre les études. Pour interpréter la taille de cette hétérogénéité, une estimation juste de ce CCI et de son intervalle de confiance est nécessaire. Le but est d'appliquer une approche Bêta générique récemment proposée pour la construction d'intervalles de confiance du CCI pour des modèles à effets mixtes linéaires jusqu'à des modèles à effets mixtes non linéaires utilisant quatre méthodes d'estimation. Ces méthodes d'estimation sont basées

sur le maximum de vraisemblance, les équations d'estimation du second ordre généralisées et deux procédures pas à pas en deux étapes. L'approche Bêta est comparée à une approximation Normale sur grands échantillons (méthode « Delta ») et à une procédure bootstrap. Les estimations des intervalles de confiance par méthode « Delta » et par bootstrap non paramétrique sur les percentiles avec diverses stratégies de ré-échantillonnages ont échoué dans le cadre de nos essais. L'approche Bêta démontre, elle, une bonne adéquation avec les deux méthodes d'estimation pas à pas et, par conséquent, c'est cette approche que nous recommandons pour le calcul des intervalles de confiance des CCI dans le cas de modèles à effets mixtes non linéaires pour les études de petite taille.

READER REACTION

B. Wu and W. Guan

556

Reader Reaction on the Generalized Kruskal–Wallis Test for Genetic Association Studies Incorporating Group Uncertainty

Acar et Sun (2013, *Biometrics*, **69**, 427-435) ont présenté un test de Kruskal-Wallis généralisé pour des études d'association qui incorpore l'incertitude génotypique et ont montré ses performances robustes et compétitives par rapport aux méthodes existante. Nous présentons une autre façon intéressante de dériver ce même test en utilisant un modèle linéaire de rangs.