
Translations of Abstracts

BIOMETRIC METHODOLOGY**J. Horrocks, D. C. Hamilton, and H. Whitehead****681***A Likelihood Approach to Estimating Animal Density from Binary Acoustic Transects*

Nous proposons une méthode basée sur la maximisation d'une approximation de la vraisemblance, pour estimer l'abondance et la densité d'une espèce animale à partir de transects acoustiques linéaires, lorsque la probabilité de détection et le rayon de détection sont inconnus. Une étude par transect est planifiée de telle manière que les points de données successifs soient statistiquement liés, et cette dépendance est utilisée pour estimer conjointement la densité, le rayon de détection, et la probabilité de détection. On suppose les données distribuées selon un processus de Poisson spatialement homogène, et on utilise une approximation Markovienne de second ordre de la vraisemblance. Une étude par simulations nous permet de montrer que cette méthode a un biais faible sous les hypothèses utilisées pour obtenir la vraisemblance, mais les résultats sont meilleurs lorsque la probabilité de détection est proche de 1. Nous évaluons aussi les effets de violation de ces hypothèses, et nous trouvons que la méthode est sensible à une structuration spatiale de la densité, et à l'existence de groupes. Nous illustrons la méthode par une étude réelle de détection acoustique de grands cachalots et de baleines à bosse.

Z. Wen, K. Pollock, J. Nichols, and P. Waser**691***Augmenting Superpopulation Capture-Recapture Models with Population Assignment Data*

Les écologues qui appliquent les modèles de capture-recapture aux populations animales ont quelquefois des informations complémentaires sur l'origine des individus des populations (information sur la génétique, sur les isotopes stables, etc.). Des tests qui affectent le génotype d'un individu à une population d'origine la plus probable sont de plus en plus utilisés. Ici, nous montrons comment améliorer un modèle de capture-recapture d'une super-population avec une telle information. Nous considérons un seul modèle de super-population sans structure d'âge, et nous séparons la probabilité d'entrée en deux composantes dues aux naissances *in situ* et à l'immigration. Nous montrons qu'il est possible d'estimer ces deux probabilités séparément. Nous considérons d'abord le cas d'une information parfaite sur les populations d'origine, où nous pouvons distinguer les individus nés *in situ* des immigrants avec certitude. Après, nous considérons le cas plus réaliste d'une information imparfaite, où nous utilisons une information génétique ou autre pour affecter des probabilités à chaque origine d'individu qu'elle soit *in situ* ou extérieure à la population. Nous utilisons une approche par ré-échantillonnage pour imputer la vraie population d'origine à partir d'une information d'affectation imparfaite. L'intégration des données sur l'origine d'une population avec des valeurs de capture-recapture nous permet de déterminer les contributions de l'immigration et de la

reproduction *in situ* sur la croissance de la population, une question importante pour les écologues. Nous illustrons nos nouveaux modèles avec des données de capture-recapture et d'affectation génétique d'une population de rats kangourous bannière de Virginie *Dipodomys spectabilis* en Arizona.

R. Li and L. Peng

701

Quantile Regression for Left-Truncated Semicompeting Risks Data

Les risques semi-compétitifs sont souvent rencontrés dans les études biomédicales où des événements absorbants censurent des événements non absorbants, mais pas inversement. En pratique, la troncature à gauche sur un événement absorbant peut survenir et fortement compliquer l'analyse de régression sur l'événement non-absorbant. Dans ce travail, nous proposons une méthode de régression par quantile sur des données tronquées à gauche et avec des risques semi-compétitifs : elle fournit des interprétations faciles et permet la prise en compte de covariables avec différents effets. Nous développons des procédures d'estimation et d'inférence qui peuvent être facilement implémentées dans les logiciels statistiques existants. Les propriétés asymptotiques des estimateurs retrouvés sont établies : ils convergent lentement et sont consistants. La performance de la méthode proposée sur un échantillon fini est évaluée via des simulations. Une application à des données issues de registre illustre nos propos.

Y. Guan, J. Yan, and R. Sinha

711

Variance Estimation for Statistics Computed from Single Recurrent Event Processes

Cet article traite de l'estimation de la variance pour des statistiques calculées à partir de processus à événement récurrent. De telles statistiques sont importantes pour le diagnostic pour chaque processus à événement récurrent. La méthode proposée suppose seulement une forme semi-paramétrique pour la structure d'ordre un du processus et non pour la structure d'ordre deux (c.à.d. la dépendance). Il est montré que le nouvel estimateur de variance est consistant pour le paramètre cible sous des conditions très légères. L'estimateur peut être utilisé dans de nombreuses applications pour l'analyse par régression semi-paramétrique d'événements récurrents telles que la détection de valeurs aberrantes, le diagnostic des résidus ainsi que des régressions robustes. Une étude par simulation et une application sur deux jeux de données réelles ont été utilisées pour démontrer l'utilisation de la méthode proposée.

J.-J. Hsieh, A. A. Ding, and W. Wang

719

Regression Analysis for Recurrent Events Data under Dependent Censoring

Des données de récurrence sont communément vues dans les études longitudinales. Une censure dépendante se produit, à cause du décès ou de l'exclusion de l'étude pour des raisons liées au processus de la maladie. Dans ce papier nous supposons des modèles de régression marginaux flexibles sur le processus de récurrence et les temps de censure dépendants sans spécifier la structure de dépendance. Le modèle proposé généralise l'approche de Ghosh et Lin (2003). La technique de censure artificielle procure une façon de maintenir l'homogénéité des variables d'erreurs d'hypothèses sous censure dépendante. Ici, nous proposons d'appliquer cette technique à deux statistiques de type Gehan. L'une considère seulement une information sur l'ordre sur les paires tandis que l'autre approche

utilise une information additionnelle des temps de censure observés sur les données disponibles de récurrence. Une procédure d'examen de modèle est aussi proposée pour examiner l'adéquation du modèle ajusté. Les estimateurs proposés ont des propriétés asymptotiques agréables. Leurs performances à distance finie sont examinées via des simulations. Pour finir les méthodes proposées sont utilisées pour analyser des données de cohorte sur le SIDA en lien avec des usages intraveineux (ALIVE).

Y. Guan

730

Second-Order Analysis of Semiparametric Recurrent Event Processes

Un jeu de données d'évènements récurrents typique consiste en un nombre souvent grand de processus d'évènements récurrents, chacun contenant des temps d'évènements multiples observés chez un individu pendant une période de suivi. De telles données sont devenues disponibles de façon croissante dans les études médicales et épidémiologiques. Dans ce papier, nous introduisons de nouvelles procédures pour mener une analyse de second ordre pour une classe flexible de processus d'évènements récurrents semi-paramétriques. Une telle analyse peut fournir une information utile concernant la structure de dépendance entre chaque processus d'évènements récurrents. En particulier, nous utiliserons les procédures proposées pour tester si les processus d'évènements récurrents individuels sont tous des processus de Poisson et pour suggérer des modèles alternatifs judicieux pour ceux qui n'en sont pas. Nous appliquons ces procédures à un jeu de données bien connu d'évènements récurrents sur la granulomatose chronique et sur un jeu de données épidémiologique sur des cas d'infection à méningocoque à Merseyside (Royaume Uni) pour illustrer leur valeur pratique.

M. Zhang and D. E. Schaubel

740

Estimating Differences in Restricted Mean Lifetime Using Observational Data Subject to Dependent Censoring

Dans les études épidémiologiques de survie, la durée de vie moyenne a souvent un intérêt direct. Nous proposons des méthodes pour estimer des différences dans la durée de vie moyenne restreinte spécifique au groupe de traitement pour des études où le traitement n'est pas randomisé et les survies sont sujettes à la fois à des censures dépendantes et indépendantes. Les méthodes proposées peuvent être vues comme une méthode hybride de deux approches générales qui tiennent compte des facteurs confondants. En particulier, des modèles à risque proportionnel par groupe de traitement sont utilisés pour prendre en compte les covariables initiales, tandis que la probabilité inverse de censure pondérée est utilisée pour accommoder les covariables dépendantes du temps qui sont corrélées à la censure. L'effet causal moyen est alors obtenu en moyennant sur les différences des valeurs estimées fondées sur les modèles à risque proportionnel. Des propriétés sur de grands échantillons des estimateurs proposés sont dérivées et des études de simulations sont conduites pour estimer leur applicabilité sur des échantillons finis. Nous appliquons les méthodes proposées sur les données de mortalité de la liste d'attente du Registre Scientifique des Bénéficiaires de Transplantation du foie.

J. Stare, M. P. Perme, and R. Henderson

750

A Measure of Explained Variation for Event History Data

De nombreuses adaptations du coefficient de détermination du modèle linéaire, R^2 , ont été proposées dans le cadre des modèles de survie. Cependant, aucun consensus ou recommandation n'existe sur l'utilisation de ces mesures. Une mesure pouvant être utilisée dans le modèle multi-états est proposée, elle peut être interprétée en temps discrets ou continus, et en présence de censure aléatoire. Cette extension de l'index de concordance et de la proportion de variation de temps de survie, est fondée sur la comparaison entre les rangs des intensités de transition observés à chaque temps d'événements et les rangs estimés par le modèle considéré pour les intensités. Cette mesure de *rangs expliqués* permet d'inclure des facteurs pronostiques dépendants du temps dans le modèle des intensités, et permet la comparaison de modèle non emboîtés pour les fonctions de risques. La versatilité de cet index est notamment illustrée pour des événements récurrents, situation où à ce jour aucune mesure de concordance n'avait été développée.

C.-C. Wen and C.-T. Lin

760

Analysis of Current Status Data with Missing Covariates

L'inférence statistique fondée sur les données censurées à droite pour les modèles à risques proportionnels (PH) avec des covariables manquantes a reçue une attention considérable, contrairement aux données censurées par intervalles ou les données d'état ponctuel manquant qui n'ont encore pas été étudiées. Notre étude est en partie motivée par l'analyse de données de fracture de la Base de Données Originale d'Enquêtes et d'Interview de Santé Nationale de Taiwan en 2005, où les événements de fractures étaient censurés par intervalles et la covariable ostéoporose n'a pas été rapportée pour tous les résidents. Sous l'hypothèse que les données ont été analysées par un modèle PH, un estimateur du maximum de vraisemblance semi-paramétrique (SPMLE) implémenté par un algorithme hybride est proposé pour analyser les données d'état ponctuel avec des covariables manquantes. Une comparaison de la performance de notre méthode avec une analyse de cohorte entière, une analyse de données des cas complets, et une analyse avec un critère de substitution est faite via une simulation avec des échantillons de taille modérée. Les données de fractures sont alors analysées.

X. Zhao and J. Sun

770

Nonparametric Comparison for Panel Count Data with Unequal Observation Processes

Cet article traite des comparaisons non-paramétriques entre plusieurs groupes de traitement sur des données de comptage issues d'un panel d'observation tel que ceux rencontrés dans les études de suivi médical ou les expériences de fiabilité pour des événements récurrents. Dans notre problématique, la plupart des procédures existantes nécessitent, entre autres pré-requis, des processus d'observation identiques d'un groupe à l'autre. On propose une nouvelle classe de procédures de test non-paramétriques permettant d'analyser des processus d'observation différents. Les nouveaux tests statistiques sont construits à partir des différences pondérées et intégrées, entre les fonctions moyennes estimées des processus d'événements récurrents sous-jacents. Les distributions asymptotiques des statistiques de tests sont données et leurs propriétés en échantillonnage fini sont examinées à travers des simulations de Monte-Carlo qui indiquent que l'approche proposée est bien adaptée aux situations pratiques. Un exemple illustratif est fourni.

Les méthodes pour ajuster des modèles de Markov non homogènes à des données observées de panel, en utilisant une solution numérique directe aux équations forward de Kolmogorov, sont développées. Les modèles de Markov non homogènes apparaissent plus souvent lorsque les intensités de transitions dépendent du temps calendaire, mais également avec des covariables déterministes dépendantes du temps tel que l'âge. Nous proposons des intensités de transitions basées sur des B-splines comme une alternative lissée aux intensités constantes par morceaux et aussi comme une généralisation des modèles de transformation du temps. Un développement du système d'équations différentielles permet d'obtenir des dérivées premières de la vraisemblance qui peuvent être utilisées dans l'algorithme des scores de Fisher pour des estimations par maximum de vraisemblance. La méthode est évaluée via une petite étude de simulation et décrite sur des données liées au développement de la vasculopathie d'allogreffe cardiaque chez des patients transplantés.

Nous nous intéressons au problème de l'estimation d'un modèle de hasards proportionnels avec covariables manquantes. En nous basant sur des ensembles de données validés et non-validés, nous proposons une méthode conditionnelle jointe qui est une extension du travail de Wang et al. (2002). La méthode proposée est nonparamétrique puisqu'elle ne nécessite ni modèle additionnel du mécanisme conduisant aux covariables manquantes, ni spécification de la distribution conditionnelle des covariables manquantes étant données les variables observées. Sous l'hypothèse que les covariables observées et que la variable subrogée sont catégorielles, nous obtenons une propriété pour grands échantillons. Les études de simulation montrent que dans diverses situations, la méthode conditionnelle jointe est plus efficace que la méthode d'estimation conditionnelle et que la méthode pondérée. Nous utilisons également un ensemble de données réelles provenant d'une étude de satisfaction de TV par câble pour illustrer ces approches.

En analyse statistique, l'affectation multiple est une approche utile en pratique pour traiter des données incomplètement observées. La règle de Rubin a rendu aisé l'estimation des paramètres et l'inférence quand celles-ci reposent sur les données complétées. Cependant, imaginer une procédure d'affectation qui puisse être utilisée de façon flexible et opérationnelle constitue un véritable défi. La procédure que nous proposons dans cet article s'inscrit dans un cadre qui utilise les modèles d'odds ratio semiparamétriques et conditionnels. Ce cadre est plus flexible et plus robuste que celui basé sur le modèle normal. Comparé à l'approche qui repose sur les modèles totalement conditionnels, notre

cadre est compatible. Les algorithmes d'affectation que nous utilisons font appel à des méthodes MCMC faciles à implémenter. Des études par simulation montrent que l'approche que nous préconisons améliore les approches communément utilisées. Nous appliquons notre procédure d'imputation des valeurs manquantes à des données de fractures.

C. Wang and M. J. Daniels

810

A Note on MAR, Identifying Restrictions, Model Comparison, and Sensitivity Analysis in Pattern Mixture Models with and without Covariates for Incomplete Data

Les modèles de mélanges constituent une approche populaire pour traiter les données longitudinales incomplètes. De tels modèles ne sont pas identifiables par construction. Pour résoudre cette difficulté, une approche possible consiste à rechercher des contraintes qui les rendent identifiables (Little, 1995; Little et Wang, 1996; Thijs et al., 2002; Kenward et al., 2003; Daniels et Hogan, 2008); cette façon de procéder constitue un point de départ naturel pour effectuer une analyse de sensibilité en présence de données non aléatoirement manquantes (Thijs et al., 2002; Daniels et Hogan, 2008). Cependant, dans un cadre multivarié et gaussien, de telles contraintes d'identifiabilité peuvent ne pas exister en présence de données aléatoirement manquantes. De plus, les stratégies qui permettent d'identifier ces contraintes peuvent être problématiques à mettre en oeuvre dans les modèles avec covariables (par exemple avec des covariables affectées de coefficients constants au cours du temps). Quand les données manquantes le sont aléatoirement, nous recherchons des conditions nécessaires sur les contraintes d'identifiabilité en faisant une hypothèse de multinormalité. Nous explorons également des stratégies pour effectuer une analyse de sensibilité ou pour réaliser une analyse totalement bayésienne, et ce, en présence d'a priori informatifs. De plus, nous proposons des stratégies d'analyse de sensibilité quand les hypothèses faites sur la variable réponse sont peu contraignantes. Nous utilisons le critère DIC pour comparer les modèles et réaliser une étude par simulation qui vise à évaluer les performances des différentes approches. Puis, nous appliquons ces méthodes à des données longitudinales provenant d'un essai clinique. Les problèmes induits par la présence de covariables affectées de coefficients constants au cours du temps sont examinés, et une contrainte d'identifiabilité reposant sur les résidus est proposée en solution.

D. Rizopoulos

819

Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data

Dans les études longitudinales, l'objectif est souvent d'évaluer comment un marqueur mesuré de façon répétée au cours du temps est associé au délai jusqu'à un événement d'intérêt. Ce type de question de recherche a permis l'émergence d'un domaine de la recherche en biostatistique autour de la modélisation conjointe de données longitudinales et de données de survie. Dans ce papier, nous considérons ce type de modélisation et nous focalisons particulièrement sur l'évaluation du pouvoir prédictif du marqueur longitudinal par rapport au temps d'événement. En particulier, nous commençons par présenter comment les probabilités de survie peuvent être estimées sur de futurs patients à partir de leurs mesures répétées disponibles et les estimations d'un modèle conjoint. Nous dérivons ensuite des mesures de précision dans le contexte du modèle conjoint et évaluons la capacité du marqueur à discriminer des sujets qui ont l'événement dans une

période de temps médicalement pertinente de ceux qui ne l'ont pas. Nous illustrons nos propositions sur un jeu de données réel de patients infectés par le VIH dans lequel l'objectif est de prédire le temps jusqu'au décès en utilisant les mesures répétées du nombre de cellules CD4.

B. Chen and X.-H. Zhou

830

Doubly Robust Estimates for Binary Longitudinal Data Analysis with Missing Response and Missing Covariates

Les études longitudinales comportent souvent des données incomplètes sur les réponses et les covariables. Les méthodes basées sur la vraisemblance telles que l'algorithme EM donnent des estimateurs consistants pour les paramètres du modèle quand les données sont manquantes au hasard à partir du moment où le modèle de réponse et le modèle des covariables manquantes sont correctement spécifiées ; mais nous n'avons pas besoin de spécifier le mécanisme de données manquantes. Une méthode alternative est l'équation d'estimation pondérée qui donne des estimateurs consistants si les modèles de données manquantes et de la réponse sont correctement spécifiées ; mais nous n'avons pas besoin de spécifier la distribution des covariables qui ont des données manquantes. Dans ce papier, nous développons une méthode d'estimation doublement robuste pour données longitudinales avec réponse manquante et covariable manquante quand les données sont manquantes au hasard. Cette méthode est attractive étant donné qu'elle fournit des estimateurs consistants si l'un des deux modèles de données manquantes ou des covariables manquantes est correctement spécifié. Des études de simulation démontrent que cette méthode fonctionne bien dans une variété de situations.

Y. Huang and B. Leroux

843

Informative Cluster Sizes for Subcluster-Level Covariates and Weighted Generalized Estimating Equations

La méthode « CWGEE » des équations d'estimation généralisées pondérées par classe proposée par Williamson et al. (2003) est efficace pour prendre en compte un biais lié à la taille des classes pour des covariables de classe à niveaux. Nous montrons que la CWGEE ne donne cependant pas toujours de bons résultats dans le cas où les covariables peuvent prendre des valeurs différentes au sein d'une classe si les nombres d'observations pour chaque niveau de la covariable sont informatifs. Par contre, la probabilité inverse de pondération du traitement tient compte de la propension du traitement à informer mais pas de l'information

concernant la taille des classes. Motive par l'évaluation des effets d'une exposition binaire avec ces deux types d'information, nous proposons plusieurs estimateurs d'équations d'estimation généralisées pondérés, avec des pondérations liées aussi bien à la taille de la classe qu'à la distribution de l'exposition binaire au sein de la classe. Le choix des poids de la population étudiée et de la nature de l'exposition. Par des simulations, nous montrons la performance accrue des estimateurs proposés en comparaison à des estimateurs existants comme GEE, CWGEE et GEE à probabilité inverse de traitement. Notre méthode est appliquée à un exemple concernant les effets de covariables sur le risque de caries dentaires chez les jeunes enfants.

S. Wu and H.-G. Müller

852

Nous proposons un modèle de réponse adaptative pour une régression linéaire fonctionnelle qui est adaptée aux réponses longitudinales partiellement mesurées. Notre méthode vise à prédire les trajectoires de réponse et à modéliser la relation de régression en conditionnant directement les observations partielles et irrégulières de la réponse sur le prédicteur, qui peut être un scalaire, un vecteur ou une fonctionnelle. Ceci évite de devoir modéliser les trajectoires de réponse, une tâche qui peut être difficile pour les données longitudinales partielles and qui était précédemment nécessaire pour l'implémentation d'une régression fonctionnelle pour des données longitudinales. L'approche proposée se montre supérieure en terme d'erreur de prédiction comparativement aux anciennes approches de régressions fonctionnelles. Elle comprend une variété de classes de régression qui sont pertinentes pour la modélisation fonctionnelle de données longitudinales dans les sciences de la vie. L'amélioration de la prédiction des trajectoires de réponse avec l'approche de réponse adaptative proposée est illustrée par une étude longitudinale de croissance de la masse de Kiwi and par l'analyse de la relation dynamique entre la charge virale et le nombre de cellules CD4 observés dans les essais cliniques du SIDA.

H. Chen and Y. Wang

861

A Penalized Spline Approach to Functional Mixed Effects Model Analysis

Dans ce travail, nous proposons des méthodes basées sur des splines pénalisés pour des modèles fonctionnels à effets mixtes avec coefficients variables. Nous décomposons des résultats longitudinaux par une somme de plusieurs termes : une fonction moyenne de la population, des covariables avec des coefficients variant avec le temps, des effets aléatoires fonctionnels spécifiques aux sujets et des processus de mesures résiduelles de l'erreur. Utilisant des splines pénalisés, nous proposons une estimation non paramétrique de la moyenne de population, des coefficients variables, des courbes aléatoires spécifiques aux sujets et de la fonction covariance associée représentant la variation entre sujets et de la fonction variance des mesures résiduelles de l'erreur représentant la variation intra sujets. Les méthodes proposées offrent une estimation flexible des courbes à la fois au niveau de la population et au niveau des sujets. En outre, décomposer la variabilité des résultats en sources entre et intra sujets est utile pour identifier la composante dominante de la variance et, par conséquent, pour modéliser de manière optimale une fonction de covariance. Nous utilisons une méthode basée sur la vraisemblance pour sélectionner des paramètres de lissage multiples. En outre, nous étudions les estimateurs asymptotiques de la ligne de base de la P-spline avec des données longitudinales. Nous effectuons des études de simulation pour étudier les performances des méthodes proposées. Le bénéfice de la décomposition de covariance entre et intra sujets est illustré par une analyse de données de croissance de Berkeley où nous identifions clairement des structures distinctes des fonctions de covariance entre et intra sujets de tailles d'enfants. Nous appliquons aussi les méthodes proposées pour estimer l'effet de traitements anti-hypertenseur à des données d'une étude cardiaque de Framingham.

T. Nummi, J. Pan, T. Siren, and K. Liu

871

Testing for Cubic Smoothing Splines under Dependent Data

Dans la plupart des recherches sur les lissages de splines, l'accent est mis sur l'estimation, alors que l'inférence statistique et en particulier les tests d'hypothèses, sont moins sources d'attention. Avec la définition appropriée de matrices « Design » pour les effets fixes et aléatoires et la structure des matrices de variance covariance des erreurs aléatoires, le lissage par spline cubique admet une formulation type modèle mixte, qui place cette approche non paramétrique dans un cadre paramétrique.

Les courbes non linéaires peuvent ainsi être prises en compte dans des effets et des coefficients aléatoires. Le paramètre de lissage est le rapport entre le coefficient aléatoire et les variances d'erreur, les tests de régression linéaire se réduisent à des tests de nullité de variances des coefficients aléatoires. Nous proposons un F-test exact adapté à la situation que nous évaluons à partir de données réelles de souches de pin puis par des simulations. Sous certaines conditions, les méthodes proposées peuvent également être appliquées dans le cas de données non indépendantes.

H. Zhou, G. Qin, and M. P. Longnecker

876

A Partial Linear Model in the Outcome-Dependent Sampling Setting to Evaluate the Effect of Prenatal PCB Exposure on Cognitive Function in Children

Le principe d'échantillonnage basé sur le critère d'intérêt (*outcome-dependent sampling, ODS*) a largement été mis en œuvre dans les études biomédicales pour des raisons de coût-efficacité. Cependant, en présence d'un critère continu, la représentation de la variable d'exposition s'est toujours limitée au cadre des modèles linéaires, compte tenu de la gageure à la fois théorique et calculatoire. Or les modèles linéaires partiels (*partial linear models, PLM*) constituent un outil inférentiel puissant pour modéliser, de façon non paramétrique, la relation entre un critère et une variable d'exposition. Nous considérons donc, dans cet article, une étude de cas où un modèle linéaire partiel est appliqué aux données issues d'un plan d'expérience de type ODS. Nous proposons une méthode semi-paramétrique de maximisation de la vraisemblance permettant de tirer des inférences à partir d'un PLM. Nous en développons les propriétés asymptotiques et conduisons des études de simulation pour montrer que l'estimateur ODS proposé peut produire une estimation plus efficace que celle issue d'un traditionnel plan d'échantillonnage aléatoire de même taille. La nouvelle méthode que nous proposons nous a permis d'explorer une question épidémiologique encore ouverte, à savoir si l'exposition *in utero* à certains niveaux de PCB est associée à un déficit intellectuel chez les enfants. Notre modèle fournit de nouveaux éléments d'appréciation de la relation entre l'exposition à des faibles niveaux de PCB et la fonction cognitive des enfants. Les résultats apportent un éclairage nouveau à un corpus de découvertes épidémiologiques jusque-là incohérentes.

B. J. Reich, H. D. Bondell, and L. Li

886

Sufficient Dimension Reduction via Bayesian Mixture Modeling

La réduction de dimension est au cœur de l'analyse de données avec nombreux prédicteurs. Une réduction de dimension suffisante vise à identifier la plus petite (en nombre) combinaison linéaire possible des prédicteurs, appelés prédicteurs suffisants, retenant toute l'information des prédicteurs concernant la distribution de la réponse. Dans cet article nous proposons une solution bayésienne pour une réduction de dimension

suffisante. Nous modélisons directement la densité de réponse avec les prédicteurs suffisants en utilisant un modèle de mélange fini. Cette approche est efficace sur le plan des calculs et offre un cadre unifié pour tenir compte de prédicteurs catégoriels, de prédicteurs manquants, et de sélection de variable bayésienne. Nous illustrons cette méthode en utilisant à la fois une étude de simulation et l'analyse d'un ensemble de données portant sur le VIH.

Y. Wang, H. Chen, R. Li, N. Duan, and R. Lewis-Fernández

896

Prediction-Based Structured Variable Selection through the Receiver Operating Characteristic Curves

Dans beaucoup de situations cliniques, un problème communément rencontré est l'évaluation de la qualité d'un test de criblage pour la détection précoce d'une maladie. Dans ces applications, ce qui est intéressant est la qualité de prédiction du test. Une sélection de variables peut être utile pour la construction d'un test médical. Un exemple en est une étude de recherche conduite pour construire un nouveau test de criblage en sélectionnant les variables d'un cribleur existant avec une structure hiérarchique sur les variables : il y a plusieurs questions racines suivies de leurs questions tiges. Les questions tiges sont posées seulement après que le sujet ait répondu à la question racine. Il est donc déraisonnable de sélectionner un modèle qui contiendrait seulement les variables tiges et non leurs variables racines. Dans ce travail, nous proposons des méthodes permettant une sélection de variables avec des variables structurées quand l'objectif principal de l'analyse est la qualité prédictive d'un test de diagnostic. Nous prenons une combinaison linéaire des variables individuelles pour construire un test combiné. Nous maximisons alors une mesure résumée de la performance prédictive du test, l'aire en dessous (AUC) d'une courbe "receiver operating characteristic" (ROC), assujettie à une fonction de pénalité pour contrôler le surajustement. Comme la maximisation de l'AUC empirique de la ROC d'un test combiné est un problème non convexe compliqué (Pepe et al. 2006), nous explorons la connexion entre l'AUC empirique et une machine à support vectoriel (SVM). Nous présentons le problème de la maximisation de la performance prédictive d'un test combiné comme un problème SVM pénalisé et nous appliquons une reparamétrisation pour imposer la structure hiérarchique des variables. Nous décrivons aussi une procédure de sélection de variables progressive pénalisée d'une régression logistique pour des variables structurées et nous la comparons aux approches fondées sur la ROC. Nous utilisons des études de simulation fondées sur des données réelles pour examiner la performance des méthodes proposées. Enfin, nous appliquons les méthodes développées pour construire un criblage structuré à utiliser dans les cliniques de premiers soins pour éventuellement envoyer les patients potentiellement psychotiques dans des lieux plus spécialisés dans leurs diagnostics et traitements.

D. Liu, and X.-H. Zhou

906

Semiparametric Estimation of the Covariate-Specific ROC Curve in Presence of Ignorable Verification Bias

Les courbes ROC covariables-spécifiques sont souvent utilisées pour évaluer la qualité d'un test de diagnostic médical ou d'un biomarqueur, quand la qualité du test est associée à certaines covariables. Dans beaucoup de tests de criblage à grande échelle, l'étalon or est assujetti à des données manquantes dues à des coûts élevés ou aux

nuisances pour le patient. Dans cet article, nous proposons une estimation semi-paramétrique des courbes ROC particulières avec covariables avec un étalon d'or partiel conditionnellement aux données manquantes. Un modèle avec des paramètres de position et d'échelle est construit pour le test résultant de la modélisation de l'effet des covariables, mais les distributions résiduelles sont laissées indéterminées. Alors la ligne de base et les fonctions de lien de la courbe ROC ont toutes deux des formes flexibles. Avec l'étalon or et l'hypothèse de données manquantes au hasard (MAR: Missing At Random), nous considérons des équations pondérées permettant l'estimation des paramètres d'échelle et de position, et des équations d'estimation à noyau pondérées pour les distributions résiduelles. Nous proposons et comparons trois estimateurs de la courbe ROC, à savoir basée sur l'imputation, probabilité inverse pondérée et estimateurs doublement robustes. Nous en déduisons la normalité asymptotique de la courbe ROC estimée ainsi que la forme analytique de l'estimateur de l'écart type résiduel. La méthode proposée est motivée par et appliquée à des données de recherche sur la maladie d'Alzheimer.

J. Ghosh, A. H. Herring, and A. M. Siega-Riz

917

Bayesian Variable Selection for Latent Class Models

Dans cet article nous développons un modèle à classes latentes avec des probabilités de classes dépendant de covariables sujet-spécifiques. L'un de nos objectifs majeurs est l'identification des prédicteurs importants des classes latentes. Nous envisageons une méthodologie qui permet l'estimation des classes latentes tout en prenant en compte l'incertitude de la sélection de variables. Nous proposons une approche bayésienne de la sélection de variables, et nous implémentons un échantillonneur de Gibbs à recherche stochastique pour le calcul a posteriori afin de modéliser les estimateurs des quantités moyennes intéressantes telles que les probabilités marginales d'inclusion des prédicteurs. Nos méthodes sont illustrées par des études de simulation et une application à des données sur la prise de poids durant la grossesse, où il est intéressant d'identifier les prédicteurs importants des classes latentes de prise de poids.

Y. Guan

926

Bias-Corrected Variance Estimation and Hypothesis Testing for Spatial Point and Marked Point Processes Using Subsampling

Pour le traitement des processus ponctuels spatiaux et de points marqués, nous présentons, d'une part de nouvelles méthodes de régression basées sur l'extrapolation, afin de corriger les biais souvent importants dans l'estimation de la variance par sous-échantillonnage, et d'autre part un nouveau test d'hypothèse. Pour l'estimation de la variance, les estimateurs que nous proposons sont des combinaisons linéaires de l'estimateur habituel de sous-échantillonnage, basées sur des tailles de sous-blocs situées dans un intervalle continu. Nous montrons qu'ils peuvent être meilleurs en erreur quadratique moyenne que l'estimateur de variance habituel de sous-échantillonnage. En particulier, si n est la dimension de la fenêtre d'observation, le taux optimal de convergence d'ordre n^{-2} peut être atteint si la dépendance entre les points se situe dans un intervalle fini. Pour le test d'hypothèse, nous appliquons directement l'extrapolation de régression proposée à la statistique de test basée sur les différentes tailles de sous-blocs et évitons ainsi le besoin d'une correction de biais pour chaque élément de la matrice de covariance utilisée pour construire la statistique de test. Nous évaluons par simulation les

mérites numériques des méthodes proposées et nous les appliquons à des données de forêt tropicale.

Y. R. Yue and J. M. Loh

937

Bayesian Semiparametric Intensity Estimation for Inhomogeneous Spatial Point Processes

Dans cet article, on propose une méthode semi-paramétrique entièrement Bayésienne pour estimer l'intensité d'un processus ponctuel spatial non-homogène. L'idée de base est d'abord de transposer l'estimation de l'intensité dans un cadre de régression de Poisson en utilisant des points d'accumulation de données sur un treillis régulier, et ensuite de procéder à une modélisation semi-paramétrique de la log-intensité en utilisant une version adaptative des champs de Markov aléatoires gaussiens pour lisser les quantités correspondantes. L'inférence est réalisée par un algorithme de simulation MCMC efficace. Par rapport aux méthodes existantes pour estimer l'intensité, e.g., modélisation paramétrique et lissage par noyau, l'estimateur proposé ne fait pas que fournir l'inférence à travers le lien entre la fonction d'intensité et les covariables possibles, mais utilise également l'information des données pour déterminer, localement et en mode adaptatif, le niveau de lissage. L'efficacité de notre méthode est démontrée à travers des études de simulation et une application sur des données de pluviométrie sur une forêt.

W. F. Christensen

947

Filtered Kriging for Spatial Data with Heterogeneous Measurement Error Variances

Il est communément supposé, lors de la prédiction de la composante sans erreur de mesure d'un processus ponctuel spatial observé, que le processus a une variance commune des erreurs de mesure. Cependant, il est fréquemment observé dans les données spatiales que chaque point de mesure est affecté d'une variance de l'erreur de mesure propre à chaque site, rendant ainsi le processus observé non stationnaire. Nous proposons une approche simple pour estimer le semi-variogramme du processus sous-jacent sans erreur de mesure impossible à observer, utilisant un ajustement du biais dans la formule classique du semi-variogramme. Nous développons sur cette base un nouveau prédicteur de krigeage filtrant les erreurs de mesure. Pour les cas où la variance d'erreur de mesure de chaque site est une fonction du processus étudié, nous recommandons de combiner cette approche avec une transformation de stabilisation de la variance. Les propriétés du prédicteur de krigeage à variance hétérogène filtrant l'erreur de mesure (HFK) et l'amélioration de cette approche par rapport au krigeage avec filtrage standard de l'erreur de mesure sont étudiées par simulation. La méthode est illustrée par des résultats de modèles climatiques concernant le détroit d'Hudson, dans le Nord Canada. Dans cette application, les sites à faible ou forte erreur de mesure se voient attribuer un poids inversement proportionnel dans le processus de prédiction, permettant de produire une image raisonnablement lissée du phénomène étudié.

X. Jiang, D. Gold, and E. D. Kolaczyk

958

Network-based Auto-probit Modeling for Protein Function Prediction

La prédiction de la fonction de protéines à partir de données pangénomiques, comme par exemple des réseaux d'associations entre protéines, est maintenant un problème

canonique en biologie computationnelle. Nous inspirant des méthodes de classification binaire, nous développons dans cet article un modèle spatial auto-Probit basé sur des réseaux. Plus particulièrement, nous développons un cadre bayésien hiérarchique pour le modèle Probit permettant de modéliser des processus binaires indexés sur des réseaux, à l'aide d'un processus Gaussien latent, conditionnel et autorégressif (CAR). Ce dernier permet d'incorporer facilement la topologie d'associations entre les protéines considérées – que les poids des associations soient binaires ou continus – dans la modélisation de la similarité entre fonctions. Nous utilisons ce cadre pour prédire la fonction de protéines, les fonctions étant définies par la base de données *Gene Ontology* (GO), qui fournit pour les fonctionnalités biologiques un vocabulaire à la fois répandu et rigoureux. De plus, nous montrons qu'une extension naturelle de notre cadre bayésien peut être utilisée pour modéliser et corriger le haut taux d'associations entre protéines prédites à tort (faux positifs) sur un échantillon d'apprentissage par des outils basés sur GO, désavantage important caractéristique de l'utilisation de ce genre de bases de données. Les performances de notre méthode sont évaluées et comparées à celles des méthodes de la littérature sur des réseaux pondérés d'associations entre protéines caractéristiques de levures, réseaux extraits d'une base de données intégrative récemment développée appelée STRING. Les résultats obtenus montrent que notre méthode est compétitive par rapport aux autres méthodes. De plus, la méthode étendue, qui incorpore à l'information contenue dans un échantillon d'apprentissage le fait que certaines associations entre protéines sont mal connues, montre un potentiel d'amélioration de la précision des prédictions remarquable.

S. Lee, F. A. Wright, and F. Zou

967

Control of Population Stratification by Correlation-Selected Principal Components

Dans le cadre d'études d'associations portant sur le génome entier, il est reconnu que la stratification des populations échantillonnées induit un gonflement des statistiques, ce qui a pour effet d'augmenter le nombre d'erreurs de Type I. Des méthodes basées sur la construction de composantes principales sont utilisées sur les génotypes afin de contrôler cet effet de stratification et permettent de plus d'apporter de l'information sur la structure des données. Nous explorons ici le lien entre le gonflement des statistiques et les composantes principales des génotypes, ce qui permet par la même occasion de créer un lien entre le contrôle de la stratification basé sur les composantes principales et les autres approches du contrôle génomique. Les résultats que nous obtenons permettent de justifier de façon inhérente l'utilisation de composantes principales, mais aussi remettent en question la pratique répandue consistant à sélectionner les principales composantes uniquement sur la base de la significativité des valeurs propres associées. Nous proposons une nouvelle méthode, appelée EigenCorr, qui permet la sélection des composantes principales selon la valeur propre qui leur est associée et leur corrélation au phénotype considéré. Notre approche tend à sélectionner un nombre plus restreint de composantes principales pour le contrôle de la stratification que ce que permet une approche qui ne testerait que les valeurs propres, ce qui permet une amélioration substantielle de la puissance et du temps de calcul. Nous montrons l'utilité de notre méthode sur des données simulées ainsi que des données réelles.

T. Cai, G. Tonini, and X. Lin

975

Kernel Machine Approach to Testing the Significance of Multiple Genetic Markers for Risk Prediction

Il existe une évidence grandissante que la recherche en génomique et protéomique détient de grandes potentialités pour changer irrévocablement la pratique de la médecine. La capacité à identifier des marqueurs génomiques et biologiques pour l'évaluation du risque peut avoir un impact important en santé publique pour la prévention des maladies, leur détection et le choix du traitement. Cependant, le nombre potentiellement large de marqueurs et la complexité des relations entre les marqueurs et le critère d'intérêt entraînent une grande difficulté pour développer des modèles de prédiction fiables. L'approche habituelle pour identifier les marqueurs importants est souvent d'évaluer l'effet marginal des marqueurs individuellement sur le phénotype d'intérêt. Quand plusieurs marqueurs sont reliés au phénotype simultanément par une structure complexe, une telle analyse marginale peut ne pas être effective. Pour surmonter ces difficultés, nous utilisons le cadre de la régression de Cox avec machine à noyaux et utilisons un test du score efficient pour évaluer l'effet global d'un jeu de marqueurs, tels que les gènes à l'intérieur d'une voie ou d'un réseau, sur la survie. Le test proposé a l'avantage de capturer les effets non linéaires potentiels sans spécifier explicitement une forme fonctionnelle non linéaire potentielle. Pour approximer la distribution de la statistique du score sous l'hypothèse nulle, nous proposons une procédure simple de rééchantillonnage qui peut être implémentée facilement en pratique. Des études numériques suggèrent que le test fonctionne bien en fonction à la fois de la taille empirique et de la puissance, même si le nombre de variables dans le set de gènes n'est pas petit comparé à la taille de l'échantillon.

X. Wang, X. Guo, M. He, and H. Zhang

987

Statistical Inference in Mixed Models and Analysis of Twin and Family Data

Les enquêtes familiales de jumeaux constituent l'un des fondements des études sur l'héritabilité génétique des maladies. Les développements théoriques sur le modèle linéaire général et la disponibilité de logiciels ont suscité un intérêt considérable pour l'analyse des enquêtes familiales de jumeaux à l'aide de modèles linéaires mixtes plutôt que de modèles à équation structurelle (SEM). Malgré une histoire déjà longue, l'analyse des enquêtes familiales de jumeaux pose des questions restées sans réponse. Nous en abordons deux. L'une est celle des conditions nécessaires et suffisantes d'identifiabilité des paramètres. L'autre est celle de la distribution asymptotique du test du rapport des vraisemblances dans une situation où les conditions standard de régularité ne sont pas remplies. Nous considérons une série d'exemples spécifiques sur lesquels nous montrons comment formuler les modèles mixtes pour qu'ils reflètent correctement les données. La clef de notre approche réside dans l'utilisation de la décomposition de Cholesky. Enfin nous appliquons notre méthode à deux jeux de données et nous obtenons une mesure d'héritabilité plus précise que dans les publications antérieures.

W. Pan and D. Zeng

996

Estimating Mean Cost Using Auxiliary Covariates

Nous étudions l'estimation de coûts médicaux moyens dans des situations où le processus de censure est dépendant, et en présence d'une grande quantité d'information auxiliaire.

Sous l'hypothèse de données manquantes aléatoires (« missing at random ») nous proposons des modèles semi-paramétriques pour obtenir des scores résumés de plus faible dimension. Un estimateur du coût moyen total peut être obtenu, par méthode non paramétrique, conditionnellement aux scores résumés. Nous montrons que, lorsque les deux modèles pour le processus coût-survie ou le modèle pour la distribution de censure est correct, alors l'estimateur est consistant et asymptotiquement normal. La performance de la méthode proposée, pour de petits échantillons, est évaluée par simulation. Enfin, notre approche est appliquée à l'analyse d'un ensemble de données réelles en économie de la santé.

S. Hiriote and V. M. Chinchilli

1007

Matrix-based Concordance Correlation Coefficient for Repeated Measures

Le coefficient de corrélation de concordance (CCC) de Lin est fréquemment utilisé dans les études cliniques pour mesurer l'accord entre deux méthodes ou deux évaluateurs. On peut cependant avoir besoin de mesurer une concordance dans des situations plus complexes, en particulier quand les réponses sont observées à plusieurs reprises. Nous proposons un nouveau CCC, adapté aux mesures répétées, le coefficient de corrélation de concordance matriciel (CCCM), qui utilise une norme matricielle possédant les propriétés requises pour mesurer la concordance entre deux vecteurs aléatoires de dimension p . On montre que pour $p=1$, le CCCM coïncide avec le CCC de Lin. Nous proposons un estimateur du CCCM fondé sur des statistiques-U. Nous établissons la normalité de la distribution asymptotique de l'estimateur du CCCM et nous fournissons ses moments. Une étude de simulation montre que l'estimateur du CCCM a un comportement satisfaisant, en termes de biais, de précision et de pourcentage de recouvrement, en particulier quand la taille de l'échantillon est supérieure à 40. Enfin, nous illustrons cette méthode nouvelle avec des données d'un réseau de recherche clinique sur l'asthme et celles de l'enquête de l'Université Penn State sur la santé des jeunes femmes.

P. R. Rosenbaum

1017

A New u -Statistic with Superior Design Sensitivity in Matched Observational Studies

Dans une étude d'observation ou non randomisée d'effets d'un traitement, une analyse de sensibilité indique la magnitude du biais lié à des covariables non mesurées dont la présence modifierait les conclusions d'une analyse naïve qui supposerait que les covariables observées suffisent à retirer tout biais. La puissance d'une analyse de sensibilité est la probabilité de rejeter à tort une hypothèse selon laquelle les effets du traitement comparés à ceux obtenus dans le cadre randomisé seraient d'une magnitude spécifique. En particulier, dans le cas d'une absence de différence, c'est équivalent à la puissance d'un test de randomisation dans une expérience randomisée. On propose ici une nouvelle famille d' U -statistiques incluant celle du test de rang signé de Wilcoxon mais également d'ordre avec potentiellement une meilleure puissance statistique dans le cas d'une analyse de sensibilité dans une étude d'observation. La statistique de Wilcoxon a une puissance élevée pour détecter de petits effets dans des expériences randomisées de grande taille, c'est-à-dire une bonne efficacité de Pitman, mais les petits effets sont invariablement plus sensibles aux biais inobservés. Les membres de cette famille de U -statistique portant sur des effets moyens a grands peuvent avoir une meilleure puissance dans une analyse de sensibilité. Par exemple, dans le cas de 250 paires issues d'une loi

normale de moyenne $\frac{1}{2}$ et de variance 1, la puissance d'une analyse de sensibilité basée sur la statistique de Wilcoxon est de 0.08 alors que la puissance d'autres statistiques membres de cette famille est de 0.66. Ce sujet est étudié en effectuant des analyses de sensibilité dans trois études d'observation en utilisant une mesure asymptotique appelée sensibilité au design et en simulant des échantillons de taille infiniment grande. Les trois exemples sont tirés de l'épidémiologie, de la médecine clinique et de la toxicologie génétique.

J. M. Albert and S. Nelson

1028

Generalized Causal Mediation Analysis

Le but de l'analyse de médiation est l'établissement des effets directs et indirects d'un traitement, d'une exposition ou d'une survenue. Plus généralement, on peut s'intéresser au contexte d'un modèle causal ainsi que décrit par un graphe orienté acyclique (DAG) où la médiation par un chemin spécifique allant de l'exposition à l'issue peut impliquer un nombre arbitraire de liens (ou étapes). Les méthodes pour estimer les effets de médiation existent pour des issues continues et un médiateur continu liés par un modèle linéaire, alors que pour une issue catégorielle ou un médiateur catégoriel, les méthodes sont limitées à une médiation en deux étapes. Nous présentons une méthode applicable à une médiation en plusieurs étapes et des variables de types mixtes, en utilisant les modèles linéaires généralisés. Nous définissons les effets de médiation à l'aide d'un cadre d'issues potentielles, et nous présentons une formule générale donnant l'effet d'une exposition à partir de n'importe quel chemin médiateur spécifié. Certains effets médiateurs sont non identifiables et leur estimation nécessite une hypothèse concernant la corrélation entre contrefaits. Nous développons une analyse de sensibilité pour établir l'impact de cette hypothèse. Nous obtenons des intervalles de confiance pour les effets médiateurs par une méthode de bootstrap. Nous appliquons notre méthode à une étude de cohorte portant sur les caries dentaires d'adolescents à petit poids de naissance. Une étude de simulation démontre le faible biais des estimateurs des effets de médiation et un taux de couverture des intervalles de confiance proches du taux nominal. Nous avons également obtenu une faible sensibilité à la corrélation des contrefaits dans la plupart des scénarios.

G. Gutjahr, W. Brannath, and P. Bauer

1039

An Approach to the Conditional Error Rate Principle with Nuisance Parameters

Le principe de l'erreur conditionnelle est difficile à appliquer en présence de paramètres de bruitage, car la fonction d'erreur conditionnelle du test préplanifié dépend des paramètres de bruitage. Pour utiliser le principe d'erreur conditionnelle avec paramètres de bruitage, nous proposons de rechercher parmi les tests qui garantissent le contrôle de l'erreur globale, le test qui maximise une combinaison pondérée des taux d'erreur conditionnelle parmi les valeurs possible des paramètres de bruitage. Nous montrons que le problème d'optimisation qui caractérise un tel test peut être résolu de façon efficiente par les algorithmes existants.

B. P. Hobbs, B. P. Carlin, S. J. Mandrekar, and D. J. Sargent

1047

Hierarchical Commensurate and Power Prior Models for Adaptive Incorporation of Historical Information in Clinical Trials

Les dispositifs bayesiens pour les essais cliniques offrent la possibilité de réduire substantiellement la taille de l'échantillon, d'accroître la puissance statistique, et de réduire le coût et le risque éthique. Cependant quand l'information a priori et l'information courante se contredisent, les méthodes bayésiennes peuvent conduire à une erreur de type 1 plus grande que prévu, et à la possibilité d'un essai plus coûteux et plus long. Cela motive une investigation de la faisabilité de méthodes bayésiennes hiérarchiques afin d'incorporer des données historiques qui soient adaptativement robustes pour une information a priori qui se révélerait incohérente avec les données expérimentales en cours d'accumulation. Dans cet article, nous présentons plusieurs modèles qui permettent à la commensurabilité des informations historique et courante de déterminer quelle quantité d'information historique est utilisée. Un outil de base consiste à élaborer l'approche traditionnelle de la puissance a priori basée sur une mesure de commensurabilité pour des données Gaussiennes. Nous comparons les performances fréquentistes de plusieurs méthodes par des simulations, et terminons avec un exemple d'un essai dans le cancer du colon qui illustre une extension de modèles linéaires de notre approche par emprunts adaptatifs. Les méthodes que nous proposons produisent des estimations plus précises des paramètres du modèle, conférant en particulier la signification statistique à la réduction observée de la taille de la tumeur avec le régime expérimental comparé au régime contrôle.

BIOMETRIC PRACTICE

R. M. Pfeiffer and M. H. Gail

1057

Two Criteria for Evaluating Risk Prediction Models

Nous proposons et nous étudions deux critères pour établir l'utilité de modèles prédisant le risque d'incidence de maladie aux fins de dépistage et de prévention, ou l'utilité de modèles de pronostic pour la stratégie suivant le diagnostic d'une maladie. Le premier critère, la proportion de cas suivis PCF(q), est la proportion d'individus inclus dans la proportion q d'individus au plus haut risque de la population qui vont développer la maladie. Le second critère est la proportion nécessaire au suivi, PNF(p), c'est-à-dire la proportion de la population générale à plus haut risque, que l'on a besoin de suivre pour qu'une proportion p de ceux qui vont être atteints puissent être suivis. PCF(q) établit l'efficacité d'un programme qui suivrait 100q % de la population à plus haut risque. PNF(p) établit la faisabilité de couvrir 100p % de cas en indiquant quelle partie de la population à plus haut risque est à suivre. Nous montrons la relation de ces deux indicateurs à la courbe de Lorenz et à son inverse, et nous présentons des éléments théoriques de distribution pour les estimateurs de PCF et de PNF. Nous développons de nouvelles méthodes, basées sur les fonctions d'influence, pour l'inférence dans un modèle à risque simple, et également pour comparer les PCF et les PNF de deux modèles de risque, lorsqu'ils sont tous deux évalués avec les mêmes données de validation.

J. Carpenter, G. Rücker, and G. Schwarzer

1066

Assessing the Sensitivity of Meta-analysis to Selection Bias: A Multiple Imputation Approach

Effectuer, à l'aide d'une méta-analyse, la synthèse tant qualitative que quantitative de

résultats, constitue un élément central du développement de la médecine dite « factuelle » (*evidence-based medicine*). Malheureusement, la méta-analyse est souvent sujette au soupçon que les études disponibles ne constituent, par le phénomène du biais de publication, qu'un sous-ensemble biaisé de résultats, problème auquel vient s'ajouter celui de l'observation systématique d'effets différents dans les études de plus petite taille. Des solutions statistiques, bien sûr, existent, la méthode *trim-and-fill* et le modèle de sélection de Copas faisant partie de celles qui ont été les plus discutées. Ces deux méthodes ont cependant des inconvénients : la méthode *trim-and-fill* repose sur le postulat que le *funnel plot* est nécessairement symétrique ; quant au modèle de sélection de Copas, outre qu'il se heurte parfois à des problèmes d'estimation, sa complexité le rend moins accessible aux comités de lecture. Dans cet article, nous modélisons la sélection des études par un modèle logistique, puis nous montrons comment les effets traitement peuvent être aisément estimés en recourant à l'imputation multiple : il s'agit en l'occurrence d'imputer des études sous l'hypothèse qu'elles sont MAR (*Missing At Random*) et d'effectuer une repondération pour calculer les estimateurs dans un modèle où la sélection des études n'est plus aléatoire. Simple d'un point de vue calculatoire, notre proposition permet aux utilisateurs de redresser la sélection en contrôlant l'ampleur de l'asymétrie résiduelle du *funnel plot*, lequel fournit de surcroît une visualisation des résultats. Nous illustrons notre approche par la méta-analyse d'un petit nombre d'études portant sur l'hypertrophie bénigne de la prostate.

R. Fu, D. K. Dey, and K. E. Holsinger

1073

A Beta-Mixture Model for Assessing Genetic Population Structure

Une proportion importante des données moléculaires récemment générées consiste en des marqueurs dominant. Ils contiennent une information substantielle sur la variation génétique mais l'effet de dominance rend impossible l'utilisation des outils standard pour estimer des mesures de différenciation comme la statistique type F_{ST} . Ici, nous proposons un nouveau modèle bayésien de mélange de lois bêta qui décrit plus précisément la structure génétique à partir de marqueurs dominant et estime plusieurs F_{ST} à partir de l'échantillon. Ce modèle a aussi d'importantes applications pour les marqueurs codominant et les polymorphismes mono-nucléotidiques. Le nombre de statistiques F_{ST} est préalablement supposé inconnu et suit une distribution aléatoire. Un algorithme à saut réversible est utilisé pour estimer ce nombre de statistiques et nous évaluons les performances de trois propositions de partage ainsi que la performance globale de la méthode sur des données simulées de marqueurs dominants. Notre modèle a pu de manière fiable identifier et estimer un spectre large de degré de différenciation génétique présent à de multiple loci. Les estimations de F_{ST} prennent également en compte l'incertitude sur l'ampleur du coefficient de consanguinité intra-populationnel. Enfin, nous illustrons notre méthode à l'aide de deux exemples, l'un utilisant des données de marqueurs dominants chez une orchidée rare et l'autre des données de marqueurs codominant dans des populations humaines.

B. E. Shepherd and C. Yu

1083

Accounting for Data Errors Discovered from an Audit in Multiple Linear Regression

Une équipe de coordination des données a effectué des audits sur sites et ont découvert des différences entre les données envoyées au centre de coordination et celles

enregistrées sur place. Nous présentons des méthodes statistiques pour incorporer les résultats de l'audit dans les analyses. Cela peut être envisagé comme un problème d'erreur de mesure, où la distribution des erreurs est un mélange avec une masse ponctuelle en zéro. Si le taux d'erreur est non nul, alors même si la moyenne des différences entre les valeurs enregistrée et correcte d'un prédicteur est zéro, les estimations de l'association entre deux variables continues seront biaisées. Nous considérons des scénarios où il y a 1) des erreurs pour le prédicteur, 2) des erreurs pour la réponse, et 3) des erreurs éventuellement corrélées pour le prédicteur et pour la réponse. Nous montrons comment incorporer le taux et la magnitude de l'erreur, estimés à partir d'un sous-ensemble aléatoire (les enregistrements de l'audit), pour calculer des estimations non biaisées de l'association et des intervalles de confiance corrects. Ensuite nous étendons ces résultats à la régression linéaire multiple où plusieurs covariables peuvent être incorrectes dans la base de données et le taux et la magnitude des erreurs peuvent dépendre du site. Nous étudions les propriétés d'échantillons finis de nos estimateurs à l'aide de simulations, discutons quelques considérations pratiques, et illustrons nos méthodes avec les données de 2815 patients infectés par le HIV en Amérique Latine, parmi lesquels 234 ont eu un audit de leurs données au moyen d'un plan d'audit séquentiel.

L. E. Dodd, E. L. Korn, B. Freidlin, R. Gray, and S. Bhattacharya
An Audit Strategy for Progression-Free Survival

1092

Dans les essais cliniques randomisés, l'utilisation de critères pouvant comporter une part de subjectivité entraîne le recours fréquent à des lectures centralisées en aveugle (*blinded independent central review* : BICR) et à la création de comités d'adjudication des événements : il s'agit, de cette façon, de réduire de possibles biais dans l'estimation d'un effet traitement mesuré sur la base d'évaluations locales (*local evaluation* : LE) réalisées dans les centres investigateurs. En oncologie, la survie sans progression constitue l'un de ces critères problématiques : elle requiert l'interprétation d'une image médicale pour déterminer si le cancer d'un patient a progressé. De ce fait, une lecture BICR a été préconisée, dans le but de réduire le risque de biais lié à la connaissance, par l'évaluateur local, du traitement randomisé. Il y a débat, cependant, sur la valeur de ce type d'évaluation centralisée pour un critère de temps de survenue comme la survie sans progression. Comme alternative à une lecture BICR complète, nous proposons une stratégie plus légère d'audit BICR partiel capable de valider la présence d'un effet traitement. Nous développons un estimateur du hazard ratio BICR qui, en utilisant les évaluations LE comme variable auxiliaire, gagne en efficacité par rapport à un estimateur du hazard ratio BICR qui ne prendrait en compte que les valeurs auditées (c'est-à-dire échantillonnées). Ainsi notre estimateur, en utilisant à la fois l'information LE sur tous les patients et l'information BICR sur les patients échantillonnés, s'avère-t-il constituer un estimateur asymptotiquement sans biais du hazard ratio BICR. Son efficacité, par rapport à l'estimateur des seules données BICR échantillonnées, augmente considérablement s'il y a une forte corrélation entre données LE et BICR. Une stratégie d'audit en deux étapes est également proposée et évaluée à l'aide de simulations. L'application rétrospective de notre méthode à un important essai en oncologie ayant fait l'objet d'une lecture BICR complète illustre son efficacité potentielle.

B. E. Shepherd, P. B. Gilbert, and C. T. Dupont

1100

Sensitivity Analyses Comparing Time-to-Event Outcomes Only Existing in a Subset Selected Postrandomization and Relaxing Monotonicity

Dans les études randomisées, il arrive que les chercheurs s'intéressent à l'effet traitement sur un critère de type temps jusqu'à événement, critère n'ayant de sens que sur une sous-population nécessairement identifiée après la randomisation. Par exemple, dans les essais de vaccin préventif du HIV, on cherche à déterminer si la randomisation d'un sujet dans le groupe vaccin influe sur le temps depuis le diagnostic d'infection jusqu'à l'initiation d'une thérapie antirétrovirale. Précédemment, certains travaux ont évalué ce type d'effet traitement au sein de celle des strates principales composée des individus qui auraient appartenu à la sous-population quel que soit le traitement alloué. Ces travaux supposaient un certain type de monotonie, à savoir que l'une des principales strates était vide (i.e. toute personne infectée dans le bras vaccin aurait également été infectée dans le bras placebo). Nous présentons ici une approche par analyse de sensibilité permettant de relâcher l'hypothèse de monotonie dans le cas d'un critère de type temps jusqu'à événement. Nous considérons également des scénarios où l'appartenance à la sous-population n'est pas connue, pour certains individus, en raison d'une censure non-informative (i.e. le statut « infecté : oui/non » k années après la randomisation est inconnu pour certains sujets en raison d'une entrée tardive dans l'étude). Nous illustrons notre méthode sur des données issues d'un essai de vaccin préventif du HIV.

L. Wang, and D. B. Dunson

1111

Semiparametric Bayes' Proportional Odds Models for Current Status Data with Underreporting

Les données concernant l'état observé des patients sont de type censurées par intervalle, pour lesquelles tous les sujets sont censurés soit à gauche, soit à droite. Notre intérêt pour ce thème provient d'une étude transversale au cours de laquelle nous avons étudié l'occurrence de fibromes utérins en fonction de l'âge à l'aide d'examen échographiques systématiques.

Nous proposons un modèle semi-paramétrique Bayésien à risques proportionnels, dans lequel la fonction de base des temps jusqu'à événement est estimée de façon non paramétrique à l'aide de splines monotones adaptatifs dans un modèle de régression logistique, les facteurs de risque potentiels étant introduits dans la partie paramétrique de la structure moyenne. L'approche proposée présente l'avantage de pouvoir être directement implémentée à l'aide d'un échantillonneur de Gibbs simple et efficace, alors que les autres modèles du temps jusqu'à événement semi-paramétriques Bayésiens posent des problèmes dans le cas de données concernant l'état courant. Le modèle est généralisé pour permettre de prendre en compte une sous-déclaration systématique dans un sous-ensemble des données. Ces méthodes

sont appliquées dans le cadre d'une étude épidémiologique sur le fibrome utérin.

S. K. Sinha, A. B. Troxel, S. R. Lipsitz, D. Sinha, G. M. Fitzmaurice, G. Molenberghs, and J. G. Ibrahim

1119

A Bivariate Pseudolikelihood for Incomplete Longitudinal Binary Data with Nonignorable Nonmonotone Missingness

Pour analyser des données longitudinales binaires avec des réponses manquantes non-

ignorable et non-monotone, une méthode basée sur la vraisemblance totale est compliquée algébriquement et nécessite souvent des calculs intensifs en particulier quand il y a beaucoup de temps de suivi. L'approche par pseudo-vraisemblance est une alternative avec un minimum d'hypothèses paramétriques proposée dans la littérature. Cette formulation ne nécessite que la définition des distributions marginales et du mécanisme de données manquantes et utilise une hypothèse de travail d'indépendance. Cependant, cet estimateur peut être inefficace pour estimer à la fois des effets dépendants du temps et des effets stationnaires dans le temps lorsque il y a une association modérée à forte entre les réponses répétées au sein des sujets. Dans cet article, nous proposons un estimateur alternatif, basé sur une pseudo-vraisemblance bivariée and nous démontrons par des simulations que la méthode proposée peut être bien plus efficace que la précédente méthode de pseudo-vraisemblance obtenue sous l'hypothèse d'indépendance. Nous illustrons la méthode en utilisant des données de nombre de CD4 issues de deux essais cliniques de patients infectés par le VIH.

D. Follmann and M. Nason

1127

An Augmented Probit Model for Missing Predictable Covariates in Quantal Bioassay with Small Sample Size

Les essais de dosage concernent la quantité ou le potentiel de certains composés tels que le

poison, les anticorps ou les drogues qui conduisent à un résultat binaire comme la mort ou l'infection chez les animaux. Pour les maladies infectieuses, on utilise couramment la régression probit pour l'inférence, et une mesure clé du potentiel est donnée par le ID_P , quantité qui résulte en un pourcentage P d'animaux infectés. Dans certaines expériences, un ensemble de validation peut être utilisé lorsque les mesures directes et indirectes de la dose sont disponibles sur un sous-ensemble d'animaux, la mesure indirecte étant disponible sur l'ensemble de ces animaux.

La variable de « mesure indirecte » peut être considérée comme une mauvaise image de la variable de mesure directe, conduisant à un problème avec erreurs sur variables. Nous développons un modèle pour l'ensemble de validation et utilisons un modèle sous contrainte de « seemingly unrelated regression » (SUR) pour obtenir la distribution de la variable de « mesure directe » conditionnellement à la variable de « mesure indirecte ».

Nous utilisons la distribution conditionnelle pour dériver une pseudo-vraisemblance basée sur une régression probit et nous utilisons le bootstrap paramétrique pour l'inférence statistique. Nous réévaluons une ancienne expérience portant sur vingt et un singes où des anticorps neutralisant le HIV ont été mesurés en utilisant un ancien dosage (indirect) chez tous les singes avec un nouveau dosage (direct) en test sur un ensemble de validation de onze singes qui avait suffisamment de plasma stockés. Grâce à nos méthodes, nous obtenons une estimation de l' ID_1 pour le nouveau dosage, qui est un objectif important pour les vaccins candidats contre le HIV. Par des simulations, nous comparons les estimations issues de la pseudo-vraisemblance avec la régression et avec une approche traditionnelle de vraisemblance.

M. Proschan, E. Brittain, and L. Kammerman

1135

Minimize the Use of Minimization with Unequal Allocation

La minimisation comme une alternative à la randomisation gagne en popularité pour des

essais cliniques de petite taille. En réponse aux questions des critiques sur l'analyse appropriée d'un tel essai, ses partisans ont argumenté qu'une approche par re-randomisation, apparentée à un test de permutation avec la randomisation conventionnelle, peut être utilisée. Cependant, ils ajoutent que cette approche par calcul intensif n'est pas nécessaire car ses résultats sont très similaires à ceux d'un test de Student ou d'un test de proportions, à moins que la taille d'échantillon ne soit très faible. Nous montrons que la minimisation appliquée avec une attribution inégale engendre des problèmes qui défont cette opinion conventionnelle.

C. X. Feng, J. Cao, and L. Bendell

1142

Exploring Spatial and Temporal Variations of Cadmium Concentrations in Pacific Oysters from British Columbia

Les huîtres de la côte pacifique nord-ouest de la Colombie Britannique (Canada) présentent de hautes teneurs en cadmium, parfois supérieures aux normes sanitaires internationales. Un objectif principal de cet article est l'étude des variations spatiales et temporelles de la concentration en cadmium d'huîtres échantillonnées le long de la côte de la Colombie Britannique. Cette information est nécessaire pour l'élaboration de recommandations sur les périodes et lieux d'ostréiculture à sélectionner dans le but de minimiser la teneur en cadmium des huîtres. Plusieurs méthodes statistiques modernes sont utilisées pour atteindre ce but, dont le lissage par splines monotones, l'analyse en composantes principales fonctionnelle, et la modélisation additive semi-paramétrique. Les taux de croissance des huîtres sont estimés comme les dérivées premières de courbes de croissances lissées monotones. Différents types d'accumulation du cadmium par les huîtres sont observés. Par exemple, les régions les plus intérieures tendent à avoir des concentrations en cadmium élevées que les régions les plus côtières, ce qui implique que plus de précautions devraient accompagner l'ostréiculture dans ces régions. La modélisation additive semi-paramétrique montre que la concentration en cadmium des huîtres décroît avec la longueur des huîtres, et que les huîtres échantillonnées à 7m de profondeur ont des concentrations en cadmium supérieures que celles prélevées à 1m de profondeur.

L. A. Hatfield, S. Gutreuter, M. A. Boogaard, and B. P. Carlin

1153

Multilevel Empirical Bayes Modeling for Improved Estimation of Toxicant Formulations to Suppress Parasitic Sea Lamprey in the Upper Great Lakes

Les estimations de réponses tout-ou-rien extrêmes, telles que la concentration qui peut tuer 99,9% des individus (LC99.9), restent un défi pour des dispositifs expérimentaux complexes ou en présence de nombreuses covariables. Une estimation exacte et précise de cette LC99.9 pour un mélange de toxiques est un élément critique du contrôle de la prolifération d'une espèce envahissante comme la lamproie dans les Grands Lacs de l'Amérique du Nord. La toxicité des espèces chimiques est modifiée par la variation dans le temps et dans l'espace de la chimie de l'eau qui doit être prise en compte par le modèle. Nous avons développé un modèle bayésien empirique multi-niveaux pour traiter les résultats expérimentaux de plusieurs laboratoires. Notre approche fournit une estimation de la LC99.9 plus exacte et plus précise que les méthodes auxquelles on l'a comparée. Cette étude démontre qu'en incorporant correctement la structure hiérarchique des données venant de plusieurs laboratoires on obtient de meilleures estimations de la LC99.9 pour le

traitement en eau-vive qui sont critiques pour le contrôle de la prolifération des larves dans la nature. De plus, la comparaison des résultats de prédictions avec des observations indépendantes in-situ révèle la présence d'un effet saisonnier latent, non apparent dans les données de laboratoire, ce qui suggère des modifications pour les études ultérieures et illustre l'importance de l'examen conjoint des résultats expérimentaux et d'observations.

M.-H. Chen, J. G. Ibrahim, P. Lam, A. Yu, and Y. Zhang

1163

Bayesian Design of Noninferiority Trials for Medical Devices Using Historical Data

Nous développons une nouvelle approche bayésienne de détermination de taille d'échantillon (SSD) pour la construction d'essais cliniques de non-infériorité. Nous étendons l'ajustement et les a priori d'échantillonnage de Wang et Gelfand (2002) à la détermination bayésienne de taille d'échantillon avec une attention particulière sur le contrôle de l'erreur de première espèce et de la puissance. Des données historiques sont incorporées par une approche de modélisation hiérarchique, ainsi que l'approche proposée par Ibrahim et Chen (2000) concernant la puissance. Nous examinons diverses propriétés de la méthodologie bayésienne proposée de détermination de taille d'échantillon et un algorithme de calcul basé sur la simulation est développé. La méthodologie proposée est appliquée au schéma d'un essai clinique de non infériorité d'une instrumentation médicale avec des données historiques issues d'essais antérieurs.

READER REACTION

L. Madsen and Y. Fang

1171

Joint Regression Analysis for Discrete Longitudinal Data

Nous présentons une approximation de la vraisemblance d'une copule gaussienne de Song, Li, et Yuan (2009, *Biometrics* **65**, 60-68) pour estimer les paramètres de régression de réponses corrélées, bivariées ou trivariées, et discrètes ou mixtes. Notre approximation permet l'estimation des paramètres à partir de vecteurs réponses de longueur bien plus grande que trois, et est asymptotiquement équivalente à la vraisemblance d'une copule gaussienne. Nous estimons des paramètres de régression à partir des données de De Backer, Keyser, De Vroey et Lesaffre (1996, *British Journal of Dermatology* **134**, 16-17) sur l'infection des ongles d'orteils. Ces données consistent en des vecteurs réponses binaires de longueur inférieure ou égale à sept sur un total de 294 sujets. Bien que la maximisation de la vraisemblance d'une copule gaussienne fournisse des estimateurs asymptotiquement plus efficaces que ceux obtenus par les équations d'estimation généralisées (GEE), notre simulation indique que pour des échantillons de taille finie, les estimateurs obtenus par GEE peuvent en réalité être jusqu'à 20% plus efficaces.