
Translations of Abstracts

Biometric Methodology**P. Saha and P. J. Heagerty****999***Time-Dependent Predictive Accuracy in the Presence of Competing Risks*

La prise en compte des risques compétitifs se pose très fréquemment dans les études où l'on mesure le délai jusqu'à un événement.

Dans cet article nous proposons des mesures dépendantes du temps des capacités prédictives d'un marqueur en présence de données censurées avec des risques compétitifs.

Des équivalents dépendants du temps de la sensibilité ou du taux de vrais positifs (TP) amènent naturellement à considérer soit les cas cumulés (ou prévalents) qui apparaissent au cours d'une période fixée, soit les cas incidents qui sont observés parmi les sujets n'ayant pas présenté l'événement à un temps donné. Une spécificité dépendant du temps (spécificité dynamique, $1 - \text{Faux Positifs (FP)}$), peut être basée sur la distribution du marqueur parmi les sujets n'ayant pas présenté l'événement. Nous proposons une extension de ces définitions afin d'introduire la cause de l'échec en cas de risques compétitifs. L'estimation proposée pour estimer le ratio cause-spécifiques TP/ taux dynamique de FP est basée sur une estimation par la méthode des plus proches voisins de la fonction bivariée de la distribution du marqueur et du temps jusqu'à événement. Par ailleurs, le ratio taux incident de TP /taux dynamique de FP peut être estimée à l'aide d'un modèle de Cox qui peut être à risques non proportionnels pour les risques cause-spécifiques avec une repondération de la distribution du marqueur. Les méthodes proposées sont une extension des mesures de la qualité prédictive en présence d'événements dépendant du temps proposées par Heagerty, Lumley, et Pepe (2000), et Heagerty et Zheng (2005).

J. H. Shih and P. S. Albert**1012***Modeling Familial Association of Ages at Onset of Disease in the Presence of Competing Risk*

Dans les études génétiques familiales, les âges de début des maladies sont couramment recueillis. On est en effet souvent intéressé par l'association au sein des familles des âges de début d'une maladie. Cependant, lorsqu'un risque compétitif est présent et est lié à la maladie, la mesure usuelle de l'association, qui suppose l'événement compétitif comme un événement indépendant de censure, est biaisée. Nous proposons un modèle bivarié qui considère deux types d'association : une entre les temps de survenue au premier événement pour des paires d'individus, et l'autre entre les types des événements survenus sachant le temps au premier événement. Nous considérons des mesures flexibles pour ces deux types d'association et nous estimons les paramètres correspondants en adoptant la stratégie en deux étapes de Shih et Louis (1995), et de Nan et al. (2006). La méthode proposée est illustrée avec des données d'apparentés venant de l'étude « Washington Ashkenazi Study ».

Y. Zheng, P. J. Heagerty, L. Hsu, and P. A. Newcomb

1024

On Combining Family-Based and Population-Based Case–Control Data in Association Studies

Pour mesurer l'effet de facteurs génétiques ou environnementaux, ou leur interaction, la combinaison de données provenant de sources différentes est susceptible d'améliorer la précision des estimations. Cette combinaison est cependant compliquée lorsque les données ont été recueillies selon des schémas différents, par exemple dans le cas d'une étude familiale et d'une enquête cas-témoin où les sujets sont indépendants. Nous décrivons dans cet article des méthodes, fondées sur la vraisemblance, qui permettent l'estimation conjointe des effets de covariables sur le risque de maladie sous des schémas d'étude incluant des cas, des parents de cas et des sujets non apparentés. Nos méthodes prennent en compte les corrélations familiales et différents modes de recrutement des cas. Des simulations établissent les bonnes performances de ces méthodes d'estimation et d'inférence sous des conditions réalistes et permettent de comparer l'efficacité de différents schémas d'étude. Nous appliquons les méthodes à des données d'un registre familial de cancers colorectaux.

E. Park, and Y.-c. I. Chang

1034

Sequential Analysis of Longitudinal Data in a Prospective Nested Case-Control Study

Le schéma cas-témoin emboîté est un type relativement nouveau d'étude observationnelle où une approche cas-témoin est utilisée dans le cadre d'une cohorte. Dans ce schéma, on observe longitudinalement les cas et les témoins en retenant tous les cas lorsqu'ils surviennent mais les témoins à certains temps. Les témoins peuvent être obtenus à des temps choisis aléatoirement ou fixés à l'avance selon ce qui est le plus commode. Ce schéma à observations longitudinales est efficace en terme de coût et de durée, spécialement lorsque la maladie est rare et l'évaluation de l'exposition difficile. Dans notre schéma, nous proposons des méthodes d'échantillonnage séquentiel et étudions à la fois les tests et estimations séquentielles de manière à arrêter l'étude dès que les règles d'arrêt sont satisfaites. Afin de rendre cet échantillonnage séquentiel plus efficace en terme de nombre de sujets et de réplifications, nous proposons d'appliquer simultanément ces méthodes d'échantillonnages aux sujets et aux réplifications jusqu'à ce que le critère d'information soit satisfait. Cet échantillonnage séquentiel simultané sur les sujets et les réplifications est plus flexible pour que les utilisateurs définissent leur plan d'échantillonnage et est différent des approches classiques utilisées pour les études longitudinales. Nous définissons un nouvel ensemble σ adapté à notre schéma d'échantillonnage, qui contient des mélanges d'observations indépendantes et corrélées et démontrons l'optimalité asymptotique d'une estimation séquentielle reposant sur la théorie des martingales. Nous démontrons aussi que la structure à incréments indépendants est retenue si bien que la méthode séquentielle groupée est applicable. Enfin, nous présentons des résultats en utilisant l'estimation séquentielle et la procédure de tests séquentiels groupés à la fois sur des données simulées et des données de diarrhées infantiles.

J.-J. Daudin, L. Pierre, and C. Vacher

1043

Model for Heterogeneous Random Networks Using Continuous Latent Variables and an Application to a Tree-Fungus Network

Les modèles de mélanges sont particulièrement bien appropriés pour modéliser les réseaux hétérogènes parce qu'ils intègrent naturellement la plupart des structures d'hétérogénéité, comme les nœuds centraux, les structures hiérarchique ou les sous-groupes homogènes. Actuellement les méthodes statistiques permettant d'estimer les paramètres de ces modèles ne sont pas entièrement satisfaisantes, à cause des difficultés de calcul ou de l'absence de propriétés théoriques démontrées. De plus, les modèles de mélange imposent à chaque sommet d'appartenir à un seul groupe et on ne peut donc pas modéliser des noeuds ayant une situation intermédiaire entre plusieurs groupes.

Le modèle proposé dans cet article suppose que chaque noeud est lui-même un mélange de noeuds extrêmes abstraits. Les propriétés de connectivité de chaque noeud sont une moyenne pondérée de celle des noeuds extrêmes. Les paramètres de ce modèle sont, d'une part la matrice de connectivité reliant les noeuds extrêmes, et d'autre part les poids des noeuds extrêmes pour chaque sommet. Ils sont estimés par la méthode du maximum de vraisemblance.

Ce modèle est utilisé pour analyser un réseau écologique hôte/parasite.

W.-H. Hwang and T.-J. Shen

1052

Small-Sample Estimation of Species Richness Applied to Forest Communities

Il existe de nombreuses méthodes bien établies pour estimer le nombre d'espèces d'une communauté forestière. Cependant, la plupart de ces méthodes produisent un biais négatif considérable en pratique, car les études de terrain ont très souvent un effort d'échantillonnage très faible. Nous proposons ici une nouvelle méthode basée sur l'échantillonnage avec remplacement pour estimer la richesse spécifique, via la procédure du jackknife généralisé. L'estimateur que nous proposons produit un petit biais et un intervalle d'estimation raisonnablement précis même pour de petits échantillons. Nous comparons les performances de ce nouvel estimateur avec celles de plusieurs estimateurs classiques par simulation sur deux jeux de données de recensements complets à Panama et en Malaisie.

A. Alonso, A. Laenen, G. Molenberghs, H. Geys, and T. Vangeneugden

1061

A Unified Approach to Multi-item Reliability

La fiabilité des échelles multi-items a été très étudiée dans la littérature psychométrique, où une myriade de mesures telles que l'alpha de Cronbach ou la formule de Spearman-Brown ont été proposées. La plupart de ces mesures, cependant, se basent sur des modèles très restrictifs qui s'appliquent seulement à des instruments unidimensionnels. Dans cet article nous introduisons deux mesures pour quantifier la fiabilité d'échelles multi-items en nous basant sur un modèle plus général. Nous montrons qu'ils enveloppent deux aspects différents du problème de fiabilité et satisfont un ensemble minimal de propriétés intuitives. L'intérêt et l'apport complémentaire des mesures sont étudiés, et on replace les approches antérieures dans un cadre théorique élargi. Enfin, nous appliquons ces mesures dans l'étude de la fiabilité de l'échelle « Positive and Negative Syndrome Scale », une échelle de cotation dans l'évaluation de la sévérité de la schizophrénie.

H. D. Bondell, A. Krishna, and S. K. Ghosh

1069

Joint Variable Selection for Fixed and Random Effects in Linear Mixed-Effects Models

Il est d'un grand intérêt pratique d'identifier simultanément les prédicteurs importants qui correspondent à la fois aux effets fixes et aléatoires dans un modèle linéaire mixte. Les approches classiques procèdent séparément sur chacune des composantes fixes et aléatoires des effets. Cependant la modification de la structure d'un des deux ensembles d'effets peut amener à des choix différents des variables de l'autre ensemble d'effets. Nous proposons une sélection simultanée des facteurs fixes et aléatoires dans un modèle linéaire mixte à l'aide d'une version modifiée de la décomposition de Cholesky. Notre méthode est basée sur une vraisemblance conjointe pénalisée avec une pénalité adaptative pour la sélection et l'estimation à la fois des effets fixes et aléatoires. Elle fournit une sélection de modèles en permettant aux effets fixes ou aux variances des effets aléatoires d'être nuls. Un algorithme EM contraint est alors utilisé pour obtenir les estimations finales. Il est de plus montré que les estimateurs pénalisés proposés satisfont à la propriété d'Oracle en ce que, asymptotiquement, leurs estimations sont aussi bonnes que si le vrai modèle était connu. Nous démontrons les qualités de notre méthode à partir d'une étude de simulation et d'un exemple réel.

J. Zhang and F. Liang

1078

Robust Clustering Using Exponential Power Mixtures

Le regroupement est une méthode largement utilisée pour extraire l'information utile dans les données d'expressions géniques pour lesquelles il est suspecté que des structures de corrélation inconnues de gènes persistent après normalisation. Ces structures de corrélation posent de grands problèmes dans les méthodes de regroupement traditionnelles comme les modèles de mélange gaussiens (GM), les k-moyennes (KM) et le partitionnement autour de centroïdes (PAM) qui ne sont pas robustes à des dépendances générales dans les données. Dans ce travail, nous utilisons les modèles de mélange de puissances exponentielles pour augmenter la robustesse du regroupement aux dépendances générales et à la non-normalité des données. Un algorithme d'espérance - maximisation conditionnelle (ECM) est développé pour calculer les estimateurs du maximum de vraisemblance des paramètres inconnus de ces mélanges. Le critère d'information bayésien (BIC) est ensuite utilisé pour déterminer le nombre de composantes de mélange. Nous montrons que les estimateurs du maximum de vraisemblance sont consistants sous l'hypothèse de dépendance faible. Nos résultats numériques indiquent que la procédure proposée donne de meilleurs résultats que les méthodes GM, KM et PAM en cas de dépendance forte ou de composantes non-Gaussiennes dans les données.

M. Lee, H. Shen, J. Z. Huang, and J. S. Marron

1087

Biclustering via Sparse Singular Value Decomposition

Nous proposons d'utiliser la décomposition sur valeurs singulières creuses (SSVD) comme instrument d'analyse exploratoire en biclassification ou pour interpréter des associations lignes-colonnes dans des matrices de données de dimension élevée. Cette approche SSVD recherche une approximation de la matrice des données sous forme d'une matrice de faible rang, et structurée en damier. La structure en damier recherchée est obtenue en contraignant les vecteurs singuliers à droite comme à gauche à être creux, c'est-à-dire à contenir de nombreux zéros. En interprétant les vecteurs singuliers comme des vecteurs de coefficients de régression, des pénalisations régularisantes sont imposées

pour obtenir des vecteurs singuliers creux. Un algorithme itératif efficace est proposé pour le calcul des vecteurs singuliers, en parallèle à une discussion de la sélection des paramètres de pénalisation. Nous utilisons un jeu de données provenant de l'utilisation de microprocesseurs dans des cancers du poumon et un jeu de données de nutrition pour illustrer l'utilisation de la méthode SSVD en biclassification. Nous comparons également l'approche SSVD à d'autres méthodes existantes en biclassification à l'aide de données simulées.

S. Huang, T. Tong, and H. Zhao

1096

Bias-Corrected Diagonal Discriminant Rules for High-Dimensional Classification

Les règles de discrimination diagonales ont été utilisées avec succès pour les problèmes de classement de dimension élevée, mais elles souffrent du sérieux inconvénient d'aboutir à des scores de discrimination biaisés. Dans cet article, nous proposons des règles de discrimination diagonales améliorées, avec des scores de discrimination à correction de biais pour les classements de dimension élevée. Nous montrons que les scores de discrimination proposés sont meilleurs que les scores standards sous une fonction de perte quadratique. Nous donnons également des résultats analytiques sur les raisons de l'amélioration potentielle de la précision de prédiction donnée par les règles à correction de biais. Enfin, nous démontrons l'amélioration apportée par les règles proposées par rapport aux règles classiques à l'aide de simulations extensives et d'études sur données réelles.

R. Iaci, T. N. Sriram, and X. Yin

1107

Multivariate Association and Dimension Reduction: A Generalization of Canonical Correlation Analysis

Dans cet article, nous proposons un nouvel indice généralisé pour détecter les relations entre deux ensembles de vecteurs aléatoires, en recherchant les projections vectorielles qui minimisent une distance L_2 entre chaque vecteur projeté et une fonction inconnue de l'autre. Les fonctions inconnues sont estimées à l'aide du lisseur de Nadaraya-Watson. Nous discutons également des extensions à des ensembles multiples et à des groupes d'ensembles multiples, et nous proposons une procédure bootstrap pour détecter le nombre de relations significatives. Toutes les méthodes proposées sont évaluées au travers de simulations extensives et d'analyses de données réelles. En particulier, nous avons appliqué notre méthodologie multi-ensemble à un ensemble de données environnementales du Comté de Los Angeles pour étudier les relations entre mortalité, climat et polluants. Nous avons détecté l'existence de relations linéaires et non-linéaires entre les vecteurs de dimension réduite, qui sont alors utilisés pour construire des modèles de régression temporelle non-linéaire pour le vecteur de mortalité à dimension réduite. Ces résultats illustrent aussi l'utilisation potentielle de notre méthode à beaucoup d'autres applications. Une présentation de l'évaluation de nos méthodes ainsi que leurs propriétés théoriques est proposée dans un appendice Web.

D. Liu and X.-H. Zhou

1119

A Model for Adjusting for Nonignorable Verification Bias in Estimation of the ROC Curve and Its Area with Likelihood-Based Approach

Pour estimer une courbe ROC quand l'état de santé des sujets est incomplètement vérifié et que les données manquantes ne sont pas ignorables, on peut écrire la vraisemblance en utilisant un modèle de sélection qui fait intervenir les probabilités d'observation complète. Nous proposons ici un estimateur de la courbe ROC et de l'aire sous la courbe ROC en présence d'un processus de vérification non ignorable. Pour que le modèle de sélection non ignorable soit identifiable, nous modélisons paramétriquement l'état de santé. A partir des probabilités estimées de vérification et de maladie, nous obtenons quatre estimations empiriques de la courbe ROC et de l'aire associée, fondées sur des méthodes d'imputation ou de pondération. Un échantillon de relativement grande taille est nécessaire pour pouvoir estimer le modèle de sélection. Des simulations ont montré que les quatre estimateurs de l'aire sous la courbe présentaient des performances satisfaisantes et que l'approche par imputation était généralement la plus efficace. Nous illustrons notre démarche avec des données concernant la maladie d'Alzheimer.

P. Gustafson, L. C. McCandless, A. R. Levy, and S. Richardson **1129**

Simplified Bayesian Sensitivity Analysis for Mismeasured and Unobserved Confounders

On étudie les situations où l'intérêt se porte sur les associations conditionnelles entre variables d'exposition et effet observé, en présence de possibles facteurs de confusion. Des problèmes surgissent quand ces facteurs de confusion ne sont pas précisément déterminés, voire pas mesurés du tout. On peut utiliser une forme d'analyse de sensibilité pour évaluer les conséquences sur les conclusions de cette limitation des données disponibles. Une approche bayésienne de l'analyse de sensibilité est conceptuellement immédiate : une loi à priori est construite pour englober les relations plausibles entre variables observées et non observées et une inférence à posteriori sur la relation entre exposition et maladie en résulte. En pratique, cependant, il peut ne pas être si immédiat de construire de façon simple une loi à priori réaliste. De plus, il peut être difficile de développer un algorithme MCMC qui fonctionne efficacement avec un modèle aussi largement sous-identifié. Dans cet article, on développe une loi à priori simple qui prend en compte des variables de confusion mesurées et non mesurées pour laquelle l'utilisateur n'a besoin de définir qu'un petit nombre d'hyperparamètres. Enfin, une approche calculatoire particulière a été développée pour l'inférence à posteriori car l'application standard de la MCMC s'est révélée inefficace pour ce problème.

E. J. Tchetgen Tchetgen and J. Robins **1138**

The Semiparametric Case-Only Estimator

Nous proposons un estimateur semi-paramétrique de type cas exclusive pour des interactions gène-environnement ou gène-gène, sous l'hypothèse d'indépendance conditionnelle des deux facteurs étant donné un vecteur de variables de confusion potentiel. Notre estimateur permet des inférences valides sur la fonction d'interaction si une (mais pas nécessairement les deux) fonctions « baseline » des confusions est correctement modélisée. De plus, lorsque les deux modèles sont corrects, notre estimateur a la plus petite variance asymptotique possible pour estimer le paramètre d'interaction dans un modèle semi-paramétrique supposant correct au moins un des deux (mais pas nécessairement les deux) modèles « baseline ».

L. Lakhal-Chaieb, R. J. Cook, and X. Lin **1145**

Inverse Probability of Censoring Weighted Estimates of Kendall's τ for Gap Time Analyses

Dans les études d'histoire d'événements, on s'intéresse souvent aux intervalles de temps entre les événements et à l'association entre ces intervalles. Les analyses d'intervalles de temps sont complexes même si la longueur du suivi est déterminée indépendamment des événements car l'association entre les intervalles de temps implique une censure dépendante sur le deuxième intervalle de temps et les suivants. Ce papier discute l'estimation non paramétrique de l'association entre les intervalles successifs de temps basée sur le tau de Kendall en présence de ce type de censure dépendante. Un estimateur non paramétrique qui utilise des poids égaux aux inverses des probabilités de censure est présenté. Un estimateur de la distribution conditionnelle de l'intervalle de temps est présenté sous l'hypothèse d'une copule de Clayton. Des simulations montrent la bonne performance des estimateurs présentés par rapport à des estimateurs existants. Une généralisation à une copula de Clayton par morceaux est présentée. Plusieurs études de simulation et illustrations avec des jeux de données réels sont aussi présentées.

J. Wolfson and P. Gilbert

1153

Statistical Identifiability and the Surrogate Endpoint Problem, with Application to Vaccine Trials

Etant donné un traitement randomisé Z , une issue de traitement Y , et un marqueur biologique S mesuré à un temps fixé une fois que Z ait été administré, nous pouvons être intéressés par le problème du critère de substitution en évaluant si S peut être utilisé pour prédire de manière fiable l'effet de Z sur Y . Plusieurs propositions récentes pour l'évaluation statistique de la valeur du critère de substitution étaient basées sur une stratification principale. Dans ce papier nous considérons deux estimations de stratification principale : les risques joints et les risques marginaux. Les risques joints mesurent les associations causales des effets du traitement sur S et Y , fournissant un aperçu de la valeur de substitution du marqueur biologique, mais ne sont pas statistiquement identifiables à partir de données d'un essai vaccinal. Alors que les risques marginaux ne mesurent pas d'associations causales des effets du traitement, ils fournissent néanmoins une indication pour la recherche future et nous décrivons un schéma de recueil de données et les hypothèses pour lesquelles les risques marginaux sont statistiquement identifiables. Nous montrons comment différents ensembles d'hypothèses affectent l'identifiabilité des estimations, en particulier nous nous écartons des travaux précédents en considérant les conséquences de relâcher l'hypothèse de l'absence d'un effet individuel du traitement sur Y avant que S ne soit mesuré. Basé sur des relations algébriques entre les risques joints et marginaux, nous proposons une approche d'analyse de sensibilité pour mesurer la valeur de substitution et nous montrons que dans beaucoup de cas, la valeur de substitution d'un marqueur biologique peut être difficile à établir même lorsque la taille d'échantillon est grande.

A. E. Raftery and L. Bao

1162

Estimating and Projecting Trends in HIV/AIDS Generalized Epidemics Using Incremental Mixture Importance Sampling

Le programme joint des Nations Unies sur le HIV/AIDS (UNAIDS) a décidé d'utiliser une agglomération (melding) Bayésienne comme base pour ses projections en probabilité

de la prévalence du HIV dans les pays d'épidémie généralisée. Elle combine un modèle épidémiologique mécaniste, des données de prévalence et des opinions d'experts. Initialement, la distribution a posteriori était approchée par l'algorithme d'échantillonnage-importance-rééchantillonnage, qui est simple à mettre en pratique, facile à interpréter, transparent pour les utilisateurs et qui donnait des résultats acceptables pour la plupart des pays. Cependant, pour certains pays, ce n'est pas efficace d'un point de vue calculatoire car la distribution a posteriori tend à être concentrée sur des arrêtes non linéaires et peuvent aussi être multimodales. Nous proposons plutôt un *Incremental Mixture Importance sampling* (IMIS), qui, de façon itérative, construit une meilleure fonction d'échantillonnage d'importance. Ceci conserve la simplicité et la transparence de l'échantillonnage-importance-rééchantillonnage mais est plus efficace d'un point de vue calculatoire. Il aboutit aussi à un estimateur plus simple de la vraisemblance intégrée qui est la base de la comparaison de modèles bayésiens et le moyennage des modèles.

Dans des expériences de simulation et avec des données réelles il surpasse à la fois l'algorithme échantillonnage-importance-rééchantillonnage et les trois algorithmes publiquement disponibles MCMC (Markov Chain Monte Carlo).

A. C. Tamhane, C. R. Mehta, and L. Liu

1174

Testing a Primary and a Secondary Endpoint in a Group Sequential Design

Nous considérons un essai clinique avec un point d'arrêt primaire et un point d'arrêt secondaire où le secondaire est testé seulement si le primaire est significatif. L'essai utilise une procédure séquentielle par groupes avec deux étapes. Le taux d'erreur sur le regroupement (familywise error rate : FWER) de conclure à tort que l'un ou l'autre point d'arrêt est significatif, doit être contrôlé au niveau nominal α . Le taux d'erreur de Type I pour le point d'arrêt primaire est contrôlé en choisissant n'importe quelle limite d'arrêt de niveau α , par exemple, les limites standard O'Brien-Flemming ou Pocock. Étant donnée une quelconque limite de niveau α pour le point d'arrêt primaire, nous étudions le problème de déterminer la limite pour le second point d'arrêt de façon à contrôler le FWER. Nous étudions cette FWER de façon analytique et numérique et trouvons qu'il est maximisé quand le coefficient de corrélation ρ entre les deux points d'arrêt est égal à 1. Considérant les quatre combinaisons possibles avec les limites O'Brien-Flemming et Pocock pour les points d'arrêt primaire et secondaire, les constantes critiques nécessaires pour le contrôle de la FWER sont calculées pour différentes valeurs de ρ . Une limite ad-hoc est proposée pour le point d'arrêt secondaire dans un intérêt pratique qui peut être l'enjeu dans une application. Des études numériques indiquent que la frontière d'O'Brien-Flemming pour le point d'arrêt primaire et celle de Pocock pour le secondaire, donnent généralement les meilleures performances pour les puissances aussi bien primaire que secondaire. La frontière de Pocock peut être remplacée par une frontière ad-hoc pour le point d'arrêt secondaire avec une très faible perte de puissance secondaire si l'enjeu est pratique. Un exemple d'essai clinique est donné pour illustrer les méthodes.

P. H. Westfall, J. F. Troendle, and G. Pennello

1185

Multiple McNemar Tests

Des méthodes pour réaliser des tests multiples de proportions appariées sont décrites. Une méthode largement applicable utilisant le test exact de Mac Nemar et les distributions exactes de tous les tests statistiques est développée. La méthode contrôle le

taux d'erreur associé à une famille de tests dans un sens fort sous des hypothèses minimales. On procure un algorithme exact (qui n'est pas basé sur des simulations) pour réaliser la méthode. Une alternative bootstrap est développée pour prendre en compte la structure de corrélation. Les caractéristiques opérationnelles de ces méthodes et d'autres sont évaluées via une étude de simulations. On donne des applications aux comparaisons multiples de modèles prédictifs pour la classification de maladies et la surveillance de phase quatre d'évènements adverses .

Biometric Practice

R. Henderson, P. Ansell, and D. Alshibani

1192

Regret-Regression for Optimal Dynamic Treatment Regimes

Nous considérons la détermination du régime optimal de traitement dynamique dans un cas pratique. La construction du modèle, sa validation et comparaison n'ont reçu que peu d'intérêt dans la littérature jusqu'à maintenant. Motivés par une application sur le dosage optimal d'anticoagulants, nous proposons un modèle et une stratégie d'estimation associée intégrant les fonctions de regret de Murphy (2003) dans un modèle de régression des réponses observées. L'estimation est rapide et des outils de diagnostic sont proposés, permettant la comparaison de plusieurs modèles. Cette méthode est testée sur des données simulées puis appliquée au dosage d'anticoagulants.

K. Lu, L. Jiang, and A. A. Tsiatis

1202

Multiple Imputation Approaches for the Analysis of Dichotomized Responses in Longitudinal Studies with Missing Data

Souvent une variable binaire est générée en dichotomisant une variable continue sous-jacente mesurée à un temps spécifique par rapport à une valeur de seuil préspecifiée. Au cas où les mesures continues sous-jacentes proviennent d'une étude longitudinale on peut utiliser un modèle à mesures répétées pour imputer les valeurs manquantes du statut du répondeur comme le résultat d'un abandon du sujet et appliquer un modèle de régression logistique au statut observé ou imputé du répondeur. Des techniques d'imputations multiples bayésiennes standards (Rubin, 1987, *Multiple Imputation for Nonresponse Survey*) qui tirent les paramètres du modèle d'imputation à partir d'une distribution a posteriori et construisent la variance des estimations des paramètres pour le modèle d'analyse comme une combinaison de variances imputées intra et inter sont trouvées conservatives. L'approche imputation multiple fréquentiste qui fixe les paramètres pour le modèle d'imputation à l'estimation du maximum de vraisemblance et construit la variance de l'estimation des paramètres pour le modèle d'analyse en utilisant les résultats de Robins et Wang (2000, *Biometrika*, 87, 113-124) est montrée être plus efficace. Nous proposons d'avoir recours aux degrés de liberté (Kenward et Roger, 1997, *Biometrics*, 53, 983-997) pour prendre en compte l'incertitude associée à la structure de covariance des paramètres estimés pour le modèle à mesures répétées.

D. Rummel, T. Augustin, and H. Küchenhoff

1209

Correction for Covariate Measurement Error in Nonparametric Longitudinal Regression

Nous proposons une méthode pour corriger l'erreur de mesure de variables explicatives en régression non-paramétrique avec une application aux données longitudinales binaires

issues d'une étude sur le sommeil chez l'homme. Les données ont été recueillies pour explorer l'association entre des niveaux hormonaux et la probabilité d'être endormi. L'effet des hormones est modélisé de manière souple en tenant compte de l'erreur de mesure de la concentration de l'hormone dans le sang et le caractère longitudinal des données. Nous présentons une approche bayésienne utilisant les techniques d'inférence de Monte Carlo par chaînes de Markov (MCMC) et introduisons une mise à jour par bloc pour améliorer l'échantillonnage et les performances numériques dans le cas binaire. Notre modèle est partiellement inspiré des «relevance vector machine» avec des fonctions de base radiales, où généralement un très petit nombre de fonctions de base sont sélectionnées de façon automatique pour ajuster les données. Dans la méthode proposée, nous implémentons un tel ajustement de la complexité régi par les données en adoptant l'idée de calibration bayésienne. Outre la théorie générale et le schéma détaillé d'échantillonnage, nous fournissons une étude de simulation pour les cas Gaussien et binaire en comparant notre méthode à l'approche naïve qui consiste à ignorer l'erreur de mesure. Les résultats montrent une nette amélioration avec la méthode de correction proposée, en particulier dans le cas Gaussien avec des variances des erreurs de mesure moyennes à grandes, et ce même lorsque que le modèle pour la variable explicative est mal spécifié.

W. Sun, M. M. Joffe, J. Chen, and S. M. Brunelli

1220

Design and Analysis of Multiple Events Case-Control Studies

Dans les études cas-témoins qui s'intéressent à différents événements, définissant ainsi plusieurs groupes de cas, les analyses usuelles ne permettent pas d'utiliser toute l'information disponible. Les études cas-témoins à événements multiples (*Multiple Events Case-Control*, MECC) introduisent un nouveau mode d'échantillonnage des témoins à partir d'une cohorte, innovation particulièrement utile pour étudier différents types d'événements dans une même cohorte : les sujets ayant présenté au moins un type d'événements, ainsi qu'une partie des sujets restant, y sont sélectionnés. Cependant, comme nous le montrons ici, si les données d'une étude MECC sont analysées comme si elles étaient issues d'une simple étude cas-témoins, les résultats obtenus sont biaisés. Nous avons donc développé trois approches basées sur les équations d'estimation généralisées, approches dont nous comparons ici, par simulation, les efficacités respectives des unes par rapport aux autres et de chacune des trois par rapport aux analyses conventionnelles correspondantes : le gain s'avère substantiel dans de nombreuses situations. Une application est réalisée sur une étude cas-témoins stratifiée, ayant pour but d'évaluer l'effet du phosphate de sodium par voie orale sur des lésions rénales chroniques, étude où sont considérées différentes définitions des cas.

S. Teerenstra, B. Lu, J. S. Preisser, T. van Achterberg, and G. F. Borm

1230

Sample Size Considerations for GEE Analyses of Three-Level Cluster Randomized Trials

Dans le domaine de la santé, on réalise des essais randomisés en grappes pour comparer l'impact de différentes interventions en termes d'amélioration de la qualité des soins. Ces essais, parfois, impliquent non pas deux, mais trois niveaux : chaque type d'intervention mis en œuvre au sein des centres de soins (les «grappes») vise à modifier le comportement des professionnels de santé (les «sujets») travaillant dans chacun de ces centres, mais c'est au niveau des patients que les effets (les «évaluations») sont finalement mesurés. En travaillant dans le cadre de l'approche GEE (Generalized

Estimating Equations : Equations Estimantes Généralisées), nous proposons une formule de calcul d'effectif qui prend en compte deux niveaux hiérarchiques, en l'occurrence les sujets au sein des grappes et les évaluations au sein des sujets. Par rapport à un plan d'expérience qui serait composé d'évaluations totalement indépendantes, cette formule révèle que la taille de l'échantillon se voit multipliée par un terme exprimable comme le produit de deux facteurs d'inflation de la variance, le premier quantifiant l'impact de la corrélation intra-sujet des évaluations sur la variance des moyennes « sujets », et le second quantifiant l'impact de la corrélation inter-sujets sur la variance des moyennes « grappes ». Les puissances prédites par la formule du calcul d'effectif coïncident avec les puissances obtenues par simulations (pour des nombres de grappes supérieurs à dix), lorsque les données sont analysées en combinant des corrections de biais – dans l'équation estimante – pour les paramètres de corrélation avec soit un estimateur de la covariance issu d'un modèle, soit l'estimateur sandwich avec correction pour échantillon de taille finie.

D. Lee and G. Shaddick

1238

Spatial Modeling of Air Pollution in Studies of Its Short-Term Health Effects

Dans les études estimant les effets de court terme de la pollution de l'air sur la santé, les mesures quotidiennes de pollution sont souvent relevées en plusieurs endroits de la région étudiée. Par contre, les données de santé sont la plupart du temps fournies sous la forme de comptages quotidiens pour la région d'étude entière et ne peut par conséquent correspondre qu'à une mesure unique de pollution. L'approche usuelle est de considérer la moyenne des relevés effectués dans la région dans un modèle de santé log-linéaire. Cependant, comme la surface polluée varie, cette statistique simple ne peut refléter précisément la concentration moyenne de pollution sur la région, ce qui peut conduire à un biais en termes d'effets sur la santé. Dans ce papier, nous proposons l'approche alternative suivante : une modélisation conjointement les concentrations en pollution et la donnée de santé dans un modèle spatio-temporel bayésien. Nous comparons cette approche à l'approche usuelle dans une simulation en regardant l'impact de la variation spatiale et en analysant l'erreur de mesure et de placement dans les données de pollution. Nous présentons ensuite une étude épidémiologique réalisée sur le grand Londres dans laquelle le lien entre la mortalité liée à des maladies respiratoires et la présence de quatre polluants est analysée.

T. A. Marques, S. T. Buckland, D. L. Borchers, D. Tosh, and R. A. McDonald

1247

Point Transect Sampling Along Linear Features

L'échantillonnage par les distances est une méthode largement utilisée pour estimer l'abondance de la faune. Une condition de base de l'échantillonnage par les distances est que les échantillonneurs (lignes ou points) sont placés selon un plan randomisé qui assure que les échantillonneurs sont positionnés indépendamment des animaux. Souvent les échantillonneurs sont placés le long de tracés linéaires comme les routes, de sorte qu'un biais est attendu si les animaux ne sont pas uniformément distribués selon les distances de ce tracé linéaire. Nous présentons une approche d'analyse des données de distance issues d'une étude où les échantillonneurs sont des points placés le long d'un tracé linéaire. Basé sur les résultats d'une étude de simulation et d'une étude des lièvres irlandais dans le Nord de l'Irlande conduite à partir de points le long des routes, nous concluons qu'un

biais important peut exister si la position des échantillonneurs n'est pas randomisée et que les méthodes d'analyse ne réussissent pas à prendre en compte cette non uniformité.

S. J. Bonner, B. J. T. Morgan, and R. King

1256

Continuous Covariates in Mark-Recapture-Recovery Analysis: A Comparison of Methods

Les covariables individuelles temps-dépendantes posent des problèmes dans les expérimentations avec marquage d'animaux parce que la covariable ne peut être observée que lorsque l'animal est capturé. Nous examinons trois méthodes qui intègrent des covariables individuelles temps-dépendantes de probabilités de survie dans l'analyse de données d'expérimentations de capture-recapture: l'imputation déterministe, une démarche d'imputation bayésienne basée sur la modélisation de la distribution conjointe de la covariable et de l'historique des captures et une approche conditionnelle en considérant uniquement les événements pour lesquels la valeur de la covariable associée est complètement observée (méthode trinomiale). Après avoir décrit les trois méthodes nous comparons les résultats obtenus avec elles sur l'analyse de l'effet de la masse corporelle sur la survie de moutons Soay (*Ovis aries*) sur l'île de Hirta en Ecosse. Des simulations basées sur ces trois méthodes sont réalisées afin de procéder à des comparaisons supplémentaires. Nous concluons que les méthodes trinomiale et bayésienne fournissent toutes les deux les meilleurs résultats dans différentes situations. Si les probabilités de capture et de seconde prise sont élevées alors le modèle conditionnel fournit des estimations sans biais et précises qui ne dépendent d'aucunes suppositions concernant la distribution de la covariable. Par contre la méthode d'imputation bayésienne est notablement meilleure quand les probabilités de captures et de récupérations sont basses à la condition que le modèle posé sur la covariable soit une bonne approximation du véritable mécanisme générateur des données.

E. R. Brown

1266

Bayesian Estimation of the Time-Varying Sensitivity of a Diagnostic Test with Application to Mother-to-Child Transmission of HIV

Nous présentons un modèle bayésien pour estimer la sensibilité d'un test diagnostique répété au cours du temps, dans sa variation temporelle, lorsque de plus l'état pathologique évolue et que la référence est observée seulement partiellement. Le modèle repose sur des hypothèses paramétriques pour la distribution du temps (inobservé, latent) de début de la maladie, et la sensibilité dépendant du temps. De plus, nous illustrons l'inclusion de données historiques pour la construction de distributions a priori. Nous appliquons ces nouvelles méthodes à des données recueillies dans une étude sur la transmission mère-enfant du virus HIV, et nous incluons une covariable à la sensibilité pour évaluer si deux tests différents possèdent des profils de sensibilité différents.

V. G. Hennessey, G. L. Rosner, R. C. Bast Jr., and M.-Y. Chen

1275

A Bayesian Approach to Dose-Response Assessment and Synergy and Its Application to In Vitro Dose-Response Studies

Dans cet article, nous proposons une approche bayésienne pour l'estimation de la relation dose-effet et celle de la synergie issue de la combinaison de deux agents. Nous nous intéressons à une étude *in vitro* sur le cancer de l'ovaire conçue pour explorer les activités antiprolifératives de quatre agents, seuls ou combinés par paires, sur deux lignées de

cellules humaines de cancer de l'ovaire. Dans cette étude, des expériences dose-effet indépendantes ont été répétées trois fois. Chacune de ces expériences a été répliquée en plusieurs versions, à des niveaux de dose différents, dont un niveau témoin (sans traitement). Nous proposons un modèle bayésien hiérarchique de régression non linéaire qui prend en compte la variabilité entre les expériences, la variabilité à l'intérieur d'une même expérience entre les différentes versions et la variabilité dans les effets observés au niveau des contrôles. Nous utilisons un algorithme MCMC (Markov Chain Monte Carlo) pour calibrer le modèle aux données avant de d'estimer *a posteriori* les quantités d'intérêt (la concentration d'inhibition médiane, CI50) ; Par ailleurs, nous proposons une méthode, basée sur l'additivité de Loewe, qui permet d'évaluer la présence d'une synergie en prenant raisonnablement bien en compte l'incertitude. D'importantes simulations permettent de faire apparaître que l'approche que nous proposons est plus fiable pour repérer une synergie que les analyses standards actuelles telles que le Principe de l'Effet Médian ou méthode de l'Indice de Combinaison (Chou-Talalay, 1984), qui ignorent d'importantes sources de variabilité ou d'incertitude.

Q. Mo and F. Liang

1284

Bayesian Modeling of ChIP-chip Data Through a High-Order Ising Model

Des procédures d'immunoprécipitation de chromatine avec puces à ADN (appelées ChIP-chip) permettent d'étudier des processus biologiques variés, tels que les interactions ADN-protéines, la modification des histones et la méthylation de l'ADN. La particularité la plus importante de ces techniques est que les mesures d'intensité par les sondes sont corrélées dans l'espace du fait que les fragments d'ADN s'hybrident avec des sondes voisines dans les expériences. Nous proposons une méthode bayésienne hiérarchique, simple mais puissante, basée sur un modèle d'Ising avec des interactions d'ordre élevé. La méthode prend naturellement en compte la structure spatiale intrinsèque des données, et peut être utilisée pour analyser des données provenant de plusieurs plateformes avec des résolutions génomiques différentes. Les paramètres du modèle sont estimés à l'aide d'un échantillonneur de Gibbs. La méthode est illustrée par deux ensembles de données des plateformes Affymetrix et Agilent, d'accès public, et comparées à trois méthodes bayésiennes également applicables, soit BAC, HGMM et Tilemap HMM. Les résultats numériques indiquent que la méthode proposée fait aussi bien que les trois autres méthodes avec les données Affymetrix mais significativement mieux que les trois autres méthodes avec les données Agilent. En outre, nous trouvons que la méthode proposée a des caractéristiques opératoires supérieures en termes de sensibilités et de taux de fausse détection sous des scénarios variés.

L. Held, K. Rufibach, and F. Balabdaoui

1295

A Score Regression Approach to Assess Calibration of Continuous Probabilistic Predictions

La calibration définie comme la cohérence statistique entre les prévisions et les observations est primordiale pour les prévisions probabilistes. Dans le cas de distributions continues, cette calibration est typiquement évaluée en utilisant la transformée intégrale de l'histogramme. Dans ce papier, on propose des tests significatifs fondés sur des règles de score dans le but est celui d'évaluer la calibration de distributions prédictives continues.

Dans le cas d'une prévision parfaite de loi Gaussienne, on établit l'expression des deux

premiers moments de deux règles de score très communément utilisées : le score logarithmique et le score continu des probabilités ordonnées. Ces scores donnent lieu à la construction de deux tests inconditionnels pour les prévisions Gaussiennes. D'une manière plus générale, on propose une nouvelle méthode basée sur la régression du score, où les scores individuels sont régressés par rapport à certaines fonctions appropriées de la variance prédictive. Cette approche conditionnelle a l'avantage d'être utilisable même dans le cas de certaines prévisions non Gaussiennes basées sur le score de David-Sebastiani. Deux cas d'études montrent bien que la régression du score a typiquement une puissance plus grande pour détecter des prévisions non calibrées par rapport aux autres approches considérées, parmi les quelles figure la technique basée sur les courbes de probabilités conditionnelles de dépassement.

Reader Reaction

M. Mandel and Y. Ritov

1306

The Accelerated Failure Time Model Under Biased Sampling

Chen (2009, *Biometrics*) a étudié le modèle semi-paramétrique de défaillance à temps accéléré (AFT) pour des données biaisées en taille. Chen considère seulement le cas non-censuré et utilise des méthodes d'estimation à partir de fonction de risque initialement proposées pour des données censurées. Cependant pour des données non-censurées, une simple régression linéaire sur échelle logarithmique est plus naturelle et fournit de meilleurs estimateurs.