
Translations of Abstracts

BIOMETRIC METHODOLOGY**W. Sun****1**

A Statistical Framework for eQTL Mapping Using RNA-seq Data

Le séquençage ARN pourrait dans un avenir proche remplacer les puces à ADN pour évaluer le niveau d'expression des gènes. Ce séquençage permet d'estimer le niveau d'expression d'un gène à partir du nombre total de lectures de séquence localisées à l'emplacement de ce gène dans le génome, une mesure qu'on appelle le nombre total de lectures (NTL). Les méthodes traditionnelles de cartographie eQTL, telles que la régression linéaire, peuvent s'appliquer aux NTL une fois ces mesures convenablement normalisées. Dans cet article nous montrons qu'en cartographie eQTL la modélisation directe des NTL par des distributions discrètes a une puissance statistique supérieure à la procédure en deux étapes : modélisation des données puis régression linéaire. De plus, le séquençage ARN donne une information sur l'expression spécifique de chaque allèle (ESA), ce que ne donnent pas les puces à ADN. En combinant l'information venant des NTL et des ESA nous pouvons distinguer par le calcul les *cis*- et les *trans*-eQTL et augmenter encore la puissance de la cartographie e-QTL. La puissance accrue de nos nouvelles méthodes est confirmée par simulation et par l'analyse de données réelles. Nous discutons aussi les problèmes de planification des expériences de séquençage ARN. Nous montrons en particulier qu'en combinant des mesures NTL et ESA on peut minimiser les coûts et garder la puissance statistique de la cartographie *cis*-eQTL en réduisant le nombre d'échantillons et en augmentant le nombre de lectures de séquence par échantillon. Outre les données de séquençage ARN, notre méthode peut être appliquée à l'étude de la base génétique d'autres données de séquençage, telles que le séquençage ADN après immunoprécipitation de la chromatine. Dans cet article nous mettons l'accent sur la cartographie e-QTL monogénique par l'étude des associations. Mais notre méthode fournit un cadre statistique pour de futures améliorations des méthodes de cartographie eQTL par le séquençage ARN (e. g. l'établissement de cartes de liaison d'eQTL), et l'étude conjointe de plusieurs marqueurs génétiques et/ou de plusieurs gènes.

W. Sun and L. Li**12**

Multiple Loci Mapping via Model-free Variable Selection

En dépit de l'abondance des méthodes récemment développées pour la sélection de variables, la cartographie pangénomique de loci multiples reste un défi actuel. La majorité des méthodes de sélection de variables existantes impose un modèle, le plus souvent une relation linéaire entre variables, avant la sélection. Cependant, la vraie nature de l'association existant entre caractère phénotypique et marqueurs génétiques est rarement connue *a priori*, et la présence d'interactions épistatiques rend cette association plus complexe qu'une relation linéaire. La sélection de variables sous modèle libre offre une alternative intéressante dans ce contexte, mais le fait que le nombre de marqueurs p soit souvent bien plus grand que le nombre d'unités expérimentales n rend inapplicables tous les modèles libres qui requièrent $n > p$. Dans cet article, nous examinons quelques méthodes de sélection de variables sous modèle libre pour des régressions petit- n -grand- p dans ce contexte de cartographie pangénomique de loci multiples. Nous proposons et recommandons la solution appelée mGAP (consistant en une pénalisation multivariée adaptative par groupe), qui ne requiert aucune spécification préalable de modèle et qui est ainsi adaptée à la description d'associations complexes entre marqueurs et caractère, et permet de manipuler les variables une par une, ce qui convient pour une situation $n < p$. L'efficacité de cette nouvelle méthode est montrée grâce à la fois à des simulations et une analyse détaillée d'un jeu de données réel portant sur 6100 caractères d'expression de gènes.

T. H. McCormick, A. E. Raftery, D. Madigan, and R. S. Burd **23**
Dynamic Logistic Regression and Dynamic Model Averaging for Binary Classification

Nous proposons une procédure de classification binaire dans les cas où il y a une incertitude sur le modèle et quand les paramètres change au cours du cours du temps. Nous prenons en compte l'incertitude du modèle par Dynamic Model Averaging (DMA), une extension dynamique du Bayesian Model Averaging (BMA) dans laquelle les probabilités a posteriori peuvent aussi changer au cours du temps. On considère un modèle espace-états pour les paramètres de chaque modèle et on permet au modèle générant les données de changer au cours du temps selon une chaîne de Markov. En calibrant un facteur d'oubli, différents niveaux de modification du mécanisme générateur des données sont pris en compte. Nous proposons un modèle qui ajuste le niveau d'oubli de manière synchronisée en utilisant la distribution prédictive a posteriori, ce qui permet de prendre en compte différents niveaux de changement à différentes dates. Nous appliquons notre méthodologie à des données concernant des enfants dont l'appendicite est traitée par chirurgie classique ou selon une procédure laparoscopique. Les facteurs associés avec le type de traitement reçu par les enfants change de manière substantielle au cours des sept années de collecte des données, une caractéristique non prise en compte par les modèles de régression standards. Du fait que notre procédure peut être implémentée de manière complètement synchrone, la future collecte de données pour des études similaires ne nécessiterait que le stockage temporaire d'informations personnelles des patients, limitant ainsi les risques de rupture de confidentialité.

C. Croux, I. Gijbels, and I. Prosdocimi **31**
Robust Estimation of Mean and Dispersion Functions in Extended Generalized Additive Models

Les Modèles Linéaires Généralisés sont une méthode largement utilisée pour obtenir des

estimateurs paramétriques de la fonction de moyenne. Ils ont été ensuite étendus par les Modèles Additifs Généralisés pour plus de flexibilité en prenant en compte les relations entre la fonction de moyenne et les covariables. Cependant la structure fixée de la variance s'avère trop restrictive dans de nombreux cas. Le contexte de la quasi-vraisemblance étendue (EQL) permet l'estimation aussi bien de la moyenne que celle de la dispersion/variance comme fonctions des covariables. Cependant, comme pour d'autres méthodes de maximum de vraisemblance, les estimateurs EQL sont sensibles aux valeurs éloignées ; on a besoin de méthodes pour obtenir des estimateurs robustes et de la fonction de moyenne et de celle de dispersion. Dans cet article, nous obtenons des estimateurs fonctionnels à la fois robustes et lisses pour la moyenne et pour la dispersion. Les performances des méthodes proposées sont illustrées par une étude de simulation et quelques exemples sur données réelles.

Y. Malinovsky, P. S. Albert, and E. F. Schisterman

45

Pooling Designs for Outcomes under a Gaussian Random Effects Model

A cause du coût croissant des mesures en laboratoire, il est devenu commun dans les études épidémiologiques de regrouper les échantillons biologiques. C'est particulièrement vrai dans les études longitudinales où le coût de réalisation de mesures multiples à travers le temps est devenu prohibitif. Dans ce papier, nous examinons le problème de l'estimation des paramètres d'un modèle à effets aléatoires gaussien, où la mesure répétée est sujette à regroupements. Nous considérons différents plans de regroupement pour l'estimation efficace par maximum de vraisemblance des composantes de la variance, avec une attention particulière à l'estimation du coefficient de corrélation intra-classe (ICC). Nous évaluons l'efficacité de différentes stratégies de regroupement en utilisant des résultats analytiques et d'études de simulation. Nous examinons la robustesse de ces plans par rapport à l'écart à la normalité des distributions et considérons des plans non-équilibrés. La méthodologie de l'essai est illustrée par une étude longitudinale sur des femmes préménopausées en s'attachant à établir la reproductibilité de la mesure du F2-isoprostane, un biomarqueur du stress oxydatif durant le cycle menstruel.

J. Braun, L. Held, and B. Ledergerber

53

Predictive Cross-validation for the Choice of Linear Mixed-Effects Models with Application to Data from the Swiss HIV Cohort Study

Le choix d'un modèle dans le cadre des modèles linéaires mixtes pour des données longitudinales reste un challenge. Outre la sélection des covariables, le choix des effets aléatoires et de la structure de corrélation des résidus devraient être possible. L'utilisation des critères classiques de choix de modèles tel que AIC ou BIC n'est pas évidente, et beaucoup de versions existent. Dans ce papier, une approche prédictive de validation croisée est proposée, elle est fondée sur le logarithme et le score de probabilité de rang continu. En contraste avec la validation croisée totale, le modèle doit être estimé une seule fois, ce qui permet des calculs rapides, même pour des grands jeux de données. Les relations avec le récent critère AIC conditionnel sont discutées. La méthode est appliquée pour identifier le meilleur modèle pour prédire l'évolution des comptages de CD4+ à partir des données de l'étude de cohorte Suisse sur le VIH (SHCS).

B. Claggett, S. W. Lagakos, and R. Wang

62

Augmented Cross-Sectional Studies with Abbreviated Follow-up for Estimating HIV Incidence

L'estimation transversale de l'incidence du VIH basée sur un test sensible et moins sensible offre de grands avantages par rapport à l'étude de cohorte traditionnelle. Cependant, son utilisation a été limitée en raison du taux de faux négatifs du test moins sensible, reflétant le phénomène que certains sujets peuvent rester négative de manière permanente par le test le moins sensible. Wang et Lagakos (2010) ont proposé un schéma d'étude transversale augmentée qui fournit un moyen d'estimer la taille de la population des infectés qui restent négatifs de façon permanente et par la suite incorporer cette information dans l'estimateur de l'incidence transversale. Dans une étude transversale augmentée, les sujets dont le test est négatif sur le critère moins sensible dans l'enquête transversale sont suivis jusqu'à la transition vers l'état, où ils auraient des résultats positifs sur le test moins sensible. Toutefois, une incertitude considérable demeure quant à la longueur appropriée de suivi et de la taille de la population infectée qui reste non réactive en permanence au test moins sensible. Dans ce papier, nous évaluons l'impact de la variation du temps de suivi sur les estimateurs des incidences dans le cadre d'une étude transversale augmentée, nous évaluons la robustesse des estimateurs transversaux aux hypothèses sur l'existence et la taille de la sous-population qui restera négative en permanence, et nous proposons un nouvel estimateur basé sur un temps de suivi abrégé (AF). Par rapport à l'estimateur original d'une étude transversale augmentée, l'estimateur AF permet des temps de suivi plus courts et ne nécessite pas l'estimation de la période moyenne, définie comme le temps moyen entre l'identification de l'infection à VIH avec les tests sensible et moins sensible. Il est montré de bons résultats dans un large éventail de situations. Nous discutons lorsque l'on peut s'attendre à de bons résultats de l'estimateur AF et nous proposons des considérations pour le schéma d'une étude transversale augmentée avec un temps de suivi abrégé.

K. Zhang, M. Traskin, and D. S. Small

75

A Powerful and Robust Test Statistic for Randomization Inference in Group-Randomized Trials with Matched Pairs of Groups

Dans les essais avec groupes randomisés, l'inférence de randomisation, basée sur les statistiques de rangs, fournit une inférence exacte robuste vis-à-vis des distributions non normales. Cependant dans les plans en paires appareillées, les statistiques basées sur les rangs actuellement disponibles font perdre significativement de la puissance en comparaison des statistiques de test du modèle linéaire mixte normal (LMM) quand le LMM est vrai. Dans la présente recherche nous investiguons et développons une statistique de test optimal parmi toutes les statistiques formées de sommes pondérées de statistiques signées de Mann-Whitney Wilcoxon sous certaines conditions. Ce test est presque aussi puissant que celui du LMM quand ce modèle est vrai, mais il est bien plus puissant pour les distributions à queues lourdes. Une étude de simulation est effectuée pour examiner la puissance.

R. T. Shinohara, C. E. Frangakis, and C. G. Lyketsos

85

A Broad Symmetry Criterion for Nonparametric Validity of Parametrically Based Tests in Randomized Trials

Les phases pilotes d'un essai clinique randomisé suggèrent souvent qu'un modèle paramétrique puisse donner une description précise des trajectoires longitudinales de l'essai. Cependant les modèles paramétriques sont fréquemment non utilisés de peur qu'ils n'invalident les tests dans lesquels les hypothèses nulles sont celles d'égalité entre les groupes expérimentaux. Des travaux préalables ont montré que lorsque pour certains types de données, certains modèles paramétriques sont utilisés, la validité du test de l'hypothèse nulle est préservée même si le modèle paramétrique est incorrect. Nous proposons une caractérisation élargie et plus aisée des modèles paramétriques pouvant être utilisés pour (a) préserver la validité non-paramétrique du test de l'hypothèse nulle, c'est-à-dire lorsque les modèles sont incorrects, et (b) augmenter la puissance par rapport à des bornes non- ou semi-paramétrique lorsque les modèles sont presque corrects. Nous illustrons par des résultats sur une étude clinique de la dépression chez des patients atteints de la maladie d'Alzheimer.

L. F. Burgette and J. P. Reiter

92

Modeling Adverse Birth Outcomes via Confirmatory Factor Quantile Regression

Nous décrivons un modèle bayésien de régression par quantiles qui utilise une structure factorielle confirmatoire au sein de la matrice design. Ce modèle est approprié lorsque les covariables sont des indicateurs des facteurs latents, et ce sont ces facteurs latents que l'analyse vise à inclure comme prédicteurs dans la régression par quantiles. Nous appliquons le modèle à une étude sur les poids de naissance dans laquelle on s'intéresse aux variables latentes représentant la santé psychosociale et l'utilisation réelle du tabac sur les faibles quantiles de la distribution des poids. Les modèles peuvent être appliqués en utilisant un package R, nommé factorOR.

S. Ji, L. Peng, Y. Cheng, and H. Lai

101

Quantile Regression for Doubly Censored Data

La double censure apparaît fréquemment dans les études cliniques reposant sur des registres, lorsque sont présentes à la fois une censure à gauche et une censure à droite. Dans ce travail, nous proposons une nouvelle stratégie d'analyse pour les données doublement censurées en utilisant un modèle de régression quantile. Nous développons des procédures d'estimation et d'inférence à faible coût computationnel, en utilisant de façon appropriée la structure de martingale sous-jacente. Les propriétés asymptotiques des estimateurs –y compris la consistance uniforme et la convergence faible- sont démontrées. De plus, afin de traiter les questions d'identifiabilité liées à la double censure, nous proposons une inférence conditionnelle. Nous montrons en outre que la méthode développée peut-être facilement adaptée à la censure à gauche. Les études de simulation démontrent les bonnes propriétés de ces procédures inférentielles pour un échantillon de taille finie. L'utilité pratique de la méthode est illustrée par une analyse de l'apparition de *Pseudomonas aeruginosa* -infection respiratoire la plus communément étudiée- chez des enfants atteints de mucoviscidose, grâce à l'utilisation du registre US Cystic Fibrosis.

M. P. Perme, J. Stare, and J. Estève

113

On Estimation in Relative Survival

L'estimation de la survie relative est devenu la première et la plus basique étape quant on

rapporte des statistiques de survie dans le cancer. Des estimateurs standard sont utilisés en routine dans les registres du cancer. Cependant il a récemment été noté que ces estimateurs ne fournissent pas une information sur la mortalité par cancer qui soit indépendante de la mortalité dans la population nationale générale. Ainsi ils ne sont pas appropriés pour une comparaison entre pays. De plus, l'interprétation usuelle de la mortalité relative est vague et induit en erreur. L'article présent tente de remédier à ces problèmes. Les quantités liées à la population dans les estimateurs traditionnels sont soigneusement décrites et leur interprétation discutée. Nous proposons ensuite un nouvel estimateur de la probabilité de survie nette qui permet la comparaison souhaitable entre pays. Le nouvel estimateur n'exige pas de modélisation et est assorti d'un estimateur simple de la variance. Les méthodes sont décrites sur des données tant réelles que simulées.

W. Yan, Y. Hu, and Z. Geng

121

Identifiability of Causal Effects for Binary Variables with Baseline Data Missing Due to Death

Nous discutons le problème que pose l'identifiabilité et l'estimation des effets causaux d'un traitement dans des sous-groupes définis par une covariable qui est parfois manquante à cause du décès, ce qui est différent du problème des événements censurés par le décès. Frangakis et al. (2007) ont proposé une approche pour estimer les effets causaux sous une hypothèse de monotonie forte (SM). Dans cet article, nous nous concentrons sur l'identifiabilité de la distribution jointe de la covariable, du traitement et des événements potentiels. Nous montrons des conditions suffisantes pour le problème de l'identifiabilité et nous relâchons l'hypothèse SM vers des hypothèses de monotonie (M) et de non-interaction (NI). Nous dérivons les algorithmes EM pour trouver les estimateurs du maximum de vraisemblance des paramètres de la distribution jointe sous différentes hypothèses. Ensuite, nous enlevons les hypothèses M et NI, et nous montrons que des signes des effets causaux d'un traitement dans les sous-groupes sont identifiables, ce qui implique que leurs bornes ne recouvrent pas la valeur zéro. Nous réalisons des simulations et des analyses de sensibilité pour évaluer nos approches. Enfin, nous appliquons les méthodes aux données NSCOT qui ont été également analysées par Frangakis et al. (2007) et Xie et Murphy (2007).

S. R. Seaman, I. R. White, A. J. Copas, and L. Li

129

Combining Multiple Imputation and Inverse-Probability Weighting

Deux approches communément utilisées concernant les données manquantes sont l'imputation multiple (MI) et la pondération par probabilité inverse (IPW). La méthode IPW est aussi utilisée pour ajuster à des fractions d'échantillonnage inégales. La méthode MI est généralement plus efficace que IPW mais elle est plus complexe. Alors que IPW requiert seulement un modèle pour la probabilité qu'un individu soit à donnée complète (résultat unidimensionnel), MI nécessite un modèle pour la distribution jointe des données manquantes (résultat multivarié) sachant les données observées. Des inadéquations pour chaque modèle peuvent conduire à des biais importants si de grandes quantités de données sont manquantes. Une troisième approche combine MI et IPW pour obtenir un estimateur doublement robuste (DR). Une quatrième approche (IPW/MI) combine MI et IPW, mais à l'inverse des méthodes DR, elle impute seulement des

valeurs manquantes isolées et utilise des pondérations pour tenir compte de plus larges blocs de données manquantes non imputées, telles que celles survenant par exemple dans une étude de cohorte sujette à attrition de l'échantillon (perdus de vue), et/ou à fractions d'échantillon inégales. Dans cet article, nous examinons les performances, en terme de biais et d'efficacité, de IPW/MI par rapport à MI et à IPW isolés, et nous étudions si l'estimateur de variance des règles de Rubin est valide pour IPW/MI. Nous prouvons que cet estimateur est valide pour IPW/MI dans la régression linéaire avec un résultat imputé ; nous présentons des simulations en faveur de l'utilisation de cet estimateur de variance dans des situations plus générales, et nous démontrons que IPW/MI est une méthode avantageuse sur ses alternatives. La méthode IPW/MI est appliquée à des données de l'Etude Nationale sur le Développement de l'Enfant (NCDS).

T. Holland-Letz, H. Dette, and D. Renard

138

Efficient Algorithms for Optimal Designs with Correlated Observations in Pharmacokinetics and Dose-Finding Studies

Les modèles à effets aléatoires sont largement utilisés dans les études de pharmacocinétique de population et de recherche de dose. Cependant, lorsque plus d'une observation est collectée par patient, la présence d'observations corrélées (due aux effets aléatoires en commun et éventuellement à de la corrélation résiduelle en série) rend habituellement difficile la détermination explicite de plans optimaux. Dans cet article nous introduisons une classe d'algorithmes multiplicatifs pour prendre en compte les données corrélées et par la même permettre le calcul numérique de plans expérimentaux optimaux dans ce type de situations. En particulier, nous démontrons son application dans un cas concret d'une étude croisée de recherche de dose, ainsi que dans un exemple typique de pharmacocinétique de population. En outre, nous dérivons une borne inférieure pour l'efficacité de n'importe quel plan donné dans ce contexte, ce qui nous permet d'un côté de contrôler le progrès de l'algorithme, et d'un autre côté d'investiguer l'efficacité d'un plan donné sans connaître le plan optimal. Finalement, nous étendons la méthodologie pour qu'elle puisse être utilisée pour la détermination de plans optimaux s'il existe certaines conditions requises sur le nombre minimal de traitements pour plusieurs (dans certains cas toutes les) conditions expérimentales.

J. Bertrand, E. Comets, M. Chenel, and F. Mentré

146

Some Alternatives to Asymptotic Tests for the Analysis of Pharmacogenetic Data Using Nonlinear Mixed Effects Models

Les modèles non linéaires à effets mixtes permettent d'explorer les différences entre les individus en termes de profils de concentrations et de réponse au traitement alors que la pharmacogénétique se concentre sur la part de cette variabilité qui peut être expliquée par les polymorphismes génétiques. Dans ce contexte, les deux tests principalement utilisés pour déterminer l'existence d'un effet génétique sur un paramètre pharmacocinétique sont i) le test de Wald qui évalue si les coefficients d'effet associés au polymorphisme à l'étude sont différents de 0 et ii) le test du rapport de vraisemblance qui compare les modèle avec et sans effet gène. Parce que ces tests ont une erreur de type I significativement augmentée sur les petits échantillons et/ou lorsque les groupes génotypiques sont déséquilibrés, nous proposons deux alternatives que nous évaluons par une étude de simulation. Dans un premier temps, nous évaluons l'apport du test de

permutation en utilisant les statistiques de Wald et du rapport de vraisemblance. Puis, nous proposons pour le test de Wald l'utilisation d'une distribution de Fisher comme distribution de référence en considérant quatre calculs différents pour le degré de liberté au dénominateur. L'impact du choix de la méthode d'estimation est aussi exploré à travers l'utilisation d'un algorithme de linéarisation du modèle au premier ordre et d'un algorithme basé sur une approximation stochastique de l'algorithme d'expectation-maximization. Les différentes méthodes de tests ont été appliquées à l'analyse de la pharmacogénétique de l'indinavir chez des patients VIH inclus dans l'essai COPHAR2. Les résultats de l'étude de simulation montrent que le test de permutation peut être employé dans ce contexte mais au prix de temps de calculs importants. Une des quatre approches basées sur l'utilisation d'une distribution de Fisher permet de corriger l'inflation de l'erreur de type I du test de Wald mais des travaux supplémentaires sont nécessaires pour confirmer cette performance.

Y. Yuan and V. E. Johnson

156

Goodness-of-Fit Diagnostics for Bayesian Hierarchical Models

Cet article propose une méthodologie pour mesurer l'ajustement dans les modèles hiérarchiques bayésiens. La méthodologie est basée sur la comparaison des valeurs de mesures de divergence pivotale, calculées en utilisant des valeurs de paramètres tirées de la distribution a posteriori, versus les distributions de référence connues. Parce que les diagnostics qui en résultent peuvent être calculés à partir de sorties standards d'algorithme de Monte Carlo par chaînes de Markov, leurs coûts de calculs sont minimaux. Plusieurs études de simulation sont montrées, chacune d'elles suggère que les diagnostics basés sur les mesures de divergence pivotale ont une puissance statistique plus grande que des vérifications de diagnostic prédictif postérieur comparables. La méthodologie proposée est illustrée sur une application clinique ; une application à des données discrètes est décrite en matériel supplémentaire.

K. St. Clair and D. O'Connell

165

A Bayesian Model for Estimating Population Means Using a Link-Tracing Sampling Design

Des plans d'échantillonnage de type « link-tracing » peuvent être utilisés pour étudier des populations humaines contenant des groupes cachés qui tendent à être liées les unes aux autres par un caractère social commun. Ces liens peuvent être utilisés pour augmenter l'intensité d'échantillonnage du domaine caché en reliant ces caractères sur les individus sélectionnés dans une première vague d'échantillonnage à des membres additionnels du domaine caché. Chow and Thompson (2003, *Survey Methodology*, 29, 197-205) ont développé un modèle bayésien pour estimer la taille ou la proportion d'individus de la population cachée pour certains plans d'échantillonnage de type « link-tracing ». Nous proposons une extension à leur modèle qui permet de modéliser une réponse quantitative. Nous évaluons les propriétés de notre modèle sur une population construite et sur une population réelle d'individus à risque, ces deux populations contenant chacune deux domaines d'individus cachés et non cachés. Nos résultats montrent que notre modèle produit de bonnes estimations ponctuelles et d'intervalles pour la moyenne de la population et des deux domaines quand les hypothèses sur la population sont satisfaites.

R. E. Gangnon

174

Local Multiplicity Adjustment for the Spatial Scan Statistic Using the Gumbel Distribution

La statistique de balayage spatial est un outil important et largement utilisé dans la détection de clusters. Elle est basée sur l'évaluation simultanée de la significativité statistique de la statistique du test du rapport des vraisemblances sur un large éventail de clusters potentiels. Dans la plupart des problèmes de détection de clusters, il y a une variation dans le degré de multiplicité locale à travers la région d'étude. Par exemple, en utilisant un rayon géographique maximum fixe pour les clusters, les zones urbaines ont typiquement beaucoup de clusters potentiels qui se superposent, alors que les zones rurales en ont peu. La statistique de balayage spatial ne tient pas compte de la variation de multiplicité locale. Nous décrivons un ajustement de la multiplicité locale précédemment proposé basé sur une correction de Bonferroni emboîtée et proposons un nouvel ajustement basé sur une approximation par distribution de Gumbel de la distribution d'une statistique de balayage local. Nous comparons la performance des trois statistiques en termes de puissance et d'un nouveau critère de détection de clusters non biaisé. Ces méthodes sont ensuite appliquées aux célèbres données de leucémie à New-York et à des données d'incidence du cancer du sein dans le Wisconsin.

M. C. Jones and A. Pewsey

183

Inverse Batschelet Distributions for Circular Data

Nous donnons des familles de distributions à quatre paramètres sur un cercle, unimodales et montrant de plus larges étendues à la fois d'asymétrie et d'aplatissement que celles déjà disponibles. Notre approche consiste à transformer l'échelle d'une fonction génératrice, telle que celle de Von Mises, en utilisant des extensions variées et non triviales d'une approche utilisée pour la première fois dans le livre de Batschelet (1981).

La clé en est l'emploi d'inverses de transformations de type Batschelet de certaines façons ; ceci présente des avantages considérables par rapport aux transformations directes de Batschelet. La transformation de l'asymétrie est spécialement attractive car elle n'a pas d'effet sur la constante de normalisation. Aussi bien qu'une variété de propriétés théoriques, ces distributions donnent l'orthogonalité entre les éléments de paires de paramètres de (position, asymétrie) et (concentration, aplatissement) quand on explore l'inférence de la vraisemblance. De plus, le paramètre de position peut parfois être construit approximativement orthogonal à tous les autres paramètres. Les profils de vraisemblance apparaissent en pratique. Deux applications, l'une concernant le déplacement d'une larve de *Drosophila*, l'autre analysant un grand échantillon d'enfants victimes du syndrome de la mort subite du nourrisson, sont étudiées comme illustration.

X. Wang, K. Wang, and J. Lim

194

Isotonized CDF Estimation from Judgment Poststratification Data with Empty Strata

Certaines applications qui requièrent une efficacité pour le coût, les tailles d'échantillons sont typiquement petites et le problème des strates vides donc apparaît souvent dans le jugement post-stratification (JPS Judgment Post-Stratification), qui est une variante importante de l'échantillonnage équilibré dans un ensemble ordonné. Dans cet article,

nous étudions l'estimation de la fonction de répartition (CDF Cumulative Distribution Function) à partir d'échantillons JPS avec strate vide. Dans la littérature, les estimateurs standards et restreints des CDF (Stokes et Sager 1988, Frey et Ozturk 2011) ne sont pas performants quand on les utilise en simplement ignorant la strate vide. Dans cet article, nous montrons que l'estimateur original isotonsé (Ozturk 2007) peut traiter les strates vides automatiquement par deux méthodes MinMax et MaxMin. Cependant, leur utilisation à l'aveugle peut conduire à des résultats indésirables dans l'une ou l'autre queue de distribution de la CDF. Motivés par ces résultats, nous proposons des estimations isotonsées modifiées pour améliorer l'efficacité de l'estimation. Par des simulations et des études expérimentales, nous montrons que nos estimateurs marchent bien dans différentes régions de la CDF et améliorent aussi la performance globale de l'estimation de la fonction toute entière.

BIOMETRIC PRACTICE

L. Trippa, G. L. Rosner, and P. Müller

203

Bayesian Enrichment Strategies for Randomized Discontinuation Trials

Nous proposons, pour optimiser le choix des paramètres régissant les plans d'expérience où la randomisation porte sur l'arrêt ou non du traitement étudié (*RDD* : *Random Discontinuation Designs*), de recourir à l'approche bayésienne de la théorie de la décision, et nous nous intéressons à l'application de ces plans *RDD* aux essais de phase II qui, en cancérologie, évaluent l'activité d'agents cytostatiques. Ces essais se déroulent en deux étapes. On traite d'abord – première étape – tous les patients en ouvert avec l'agent à l'étude, afin d'identifier une sous-population éventuellement sensible à cet agent. Puis – deuxième étape –, on randomise, on traite et on suit les patients du sous-groupe identifié, chez lesquels on compare les critères d'évaluation, la randomisation ayant porté sur la continuation du traitement à l'étude ou, au contraire, son arrêt avec remplacement par un traitement contrôle. Trois paramètres caractérisent le *RDD* : le nombre de patients dans l'essai, la durée du stade préliminaire et la durée de suivi après randomisation. Après avoir défini un modèle probabiliste de la croissance tumorale et spécifié une fonction d'utilité appropriée, nous développons un algorithme calculatoire permettant de choisir, de façon optimale, ces trois paramètres.

D. Ghosh, J. M. G. Taylor, and D. J. Sargent

226

Meta-analysis for Surrogacy: Accelerated Failure Time Models and Semicompeting Risks Modeling

Un intérêt récent dans la littérature médicale et statistique concerne l'évaluation et la validation de critères substitutifs comme critères intermédiaires pour les critères de jugements cliniques dans les études médicales. Plus récemment, des chercheurs se sont concentrés sur l'utilisation de méthodes meta-analytiques pour quantifier la valeur de cette substitution. Dans cet article, nous étendons des procédures existantes pour l'analyse fondée sur un modèle de survie accéléré dans ce contexte. Cette approche, relative au modèle à hasards proportionnels, a l'avantage de pouvoir prendre en compte l'analyse de risques semi-compétitifs, où nous modélisons la région où le critère de jugement substitutif survient avant le vrai critère de jugement. Plusieurs méthodes d'estimation et de procédures déductives sont présentées. De plus, des méthodes inter- et intra-essai pour

évaluer la valeur du critère substitutif sont développées ; une procédure innovante à composantes principales est développée pour quantifier la substitution au niveau de l'essai. Les méthodes sont illustrées par une application sur des données de plusieurs études dans le cancer colorectal. Les procédures proposées tiennent compte de l'évaluation du critère de substitution sans nécessairement créer des critères de jugements composites.

S. G. Baker, D. J. Sargent, M. Buyse, and T. Burzykowski **248**
Predicting Treatment Effect from Surrogate Endpoints and Historical Trials: An Extrapolation Involving Probabilities of a Binary Outcome or Survival to a Specific Time

En utilisant de multiples essais historiques avec des points terminaux réels et subrogés, nous considérons des modèles variés pour prédire l'effet d'un traitement sur un point terminal réel d'un essai cible dans lequel seul un point terminal subrogé est observé. Le résultat prédit est calculé en utilisant (1) un modèle de prédiction (mélange, linéaire, ou stratifié) estimé à partir des essais historiques et du point terminal subrogé, et (2) une erreur d'extrapolation aléatoire estimée en écartant successivement un essai parmi les essais historiques. La méthode s'applique aussi bien pour des issues binaires qu'à la survie à un temps particulier calculé à partir des données de survie censurées. Nous calculons un intervalle de confiance à 95% pour le résultat prédit et nous le validons par simulation. Pour couvrir l'incertitude additionnelle sur l'effet estimé du traitement, provenant de l'utilisation d'une prévision à la place du véritable résultat, nous calculons les coefficients multiplicateurs de l'erreur standard (moyenne du rapport des erreurs standard du résultat prédit et du résultat réel). Un package est disponible pour téléchargement.

L. A. Renfro, B. P. Carlin, and D. J. Sargent **258**
Bayesian Adaptive Trial Design for a Newly Validated Surrogate Endpoint

L'évaluation des critères de substitution pour une utilisation primaire dans les essais cliniques futurs est un domaine de recherche de plus en plus important en raison d'une demande pour des essais plus efficaces conjuguée avec l'acceptation récente par les autorités réglementaires de considérer comme 'valide' certains critères de substitution. Cependant, peu d'attention a été accordée à la manière dont une étude qui utilise un critère de substitution nouvellement validé comme critère primaire d'évaluation pourrait être conçu. Nous proposons un nouveau design bayésien d'essai adaptatif qui permet au nouveau critère de substitution de jouer un rôle dominant dans l'évaluation de l'effet d'une intervention tout en restant prudent et réaliste quant à son utilisation. En intégrant plusieurs informations historiques d'essais cliniques sur la relation validée entre les critères de substitution et clinique, puis ensuite en évaluant les données accumulées contre cette relation à mesure que le nouvel essai progresse, on se prémunit de manière adaptative contre une appréciation erronée de l'effet du traitement basée sur un critère de substitution réellement invalide. Lorsque les résultats communs dans le nouvel essai semblent plausibles au regard des informations historiques issues d'essais similaires, nous utilisons le critère de substitution comme critère principal, et ceci de façon adaptative par exemple en arrêtant l'essai précocement pour supériorité ou infériorité du traitement expérimental ou bien pour futilité. Sinon, nous supprimons le critère de substitution et procédons à l'analyse adaptative avec le critère primaire d'origine. Nous

utilisons la simulation pour tester les caractéristiques opératoires de ce nouveau design par rapport à l'approche standard d'O'Brien-Fleming ainsi que la capacité de notre design à discriminer les critères de substitution fiables et non fiables dans d'hypothétiques futurs essais. Par ailleurs, les bénéfices possibles sont étudiés à l'aide de données de patients provenant de 18 essais de thérapie adjuvante dans le cancer du côlon où la survie sans maladie est considérée comme un nouveau critère de substitution validé pour la survie globale.

P. Zhang, J. Liu, J. Dong, J. L. Holovati, B. Letcher, and L. E. McGann 268
A Bayesian Adjustment for Multiplicative Measurement Errors for a Calibration Problem with Application to a Stem Cell Study

Nous développons une approche bayésienne afin d'étudier la calibration lorsque une covariable est sujette à des erreurs de mesure multiplicatives. La motivation de notre travail est une étude dont le but est de déterminer la dose recommandée minimum de cellules souches à transplanter après une transfusion sanguine. Lorsque l'on détermine une dose sans danger de cellules souches à partir d'échantillons pré-congelés, le taux de récupération post cryoconservation rentre dans le modèle comme un terme multiplicatif d'erreur de mesure, comme indiqué dans le modèle (2). Nous étudions l'impact, en termes de biais asymptotique sur les coefficients de régression, de l'absence de prise en compte de l'erreur de mesure. Suivant la structure générale des données disponibles dans la pratique, nous proposons une méthode bayésienne en deux étapes pour estimer le modèle en utilisant R2WinBUGS (Sturtz, Ligges, et Gelman, 2005). Nous illustrons cette méthode sur l'étude citée précédemment. Les résultats de ce travail permettent aux laboratoires qui extraient des cellules souches à partir du sang, d'une part, d'établir la dose minimum de cellule souches à transplanter et d'autre part, de développer une approche systématique afin d'étudier si l'analyse post-décongélation est justifiée.

M. M. Joffe, W. P. Yang, and H. Feldman 275
G-Estimation and Artificial Censoring: Problems, Challenges, and Applications

En principe, l'estimation-G est une approche intéressante pour prendre en compte des facteurs de confusion dans l'évaluation des traitements. Elle a été rarement utilisée pour l'estimation des effets du traitement dans le cadre des modèles de survie. Une raison à cela est due à la censure artificielle, une technique d'analyse qui considère quelques sujets qui en fait sont observés en échec comme s'ils étaient censurés. La technique de censure artificielle mène à un manque de lissage dans la fonction d'estimation qui peut poser des problèmes dans l'estimation de la variance et dans l'optimisation. Cela peut également mener à l'impossibilité de trouver des solutions dans l'estimation habituelle des fonctions, ce qui soulève alors des questions à propos des critères appropriés pour l'optimisation. Pour améliorer la performance des procédures d'optimisation, nous considérons des approches qui réduisent le nombre de censures artificielles. Nous proposons de substituer des fonctions indicatrices par des fonctions de lissage, et d'utiliser des fonctions d'estimation ajustées à la mesure de l'information des données ; nous évaluons la performance de ces approches en utilisant des simulations. Nous considérons également des critères d'optimisation appropriés en présence d'une perte d'information due à la censure artificielle. Nous motivons et illustrons nos approches en utilisant des données observationnelles sur l'effet de l'érythropoïétine sur la mortalité des

sujets sous hémodialyse.

C. S. McMahan, J. M. Tebbs, and C. R. Bilder
Informative Dorfman Screening

287

Depuis le début des années 1940, le test par groupe a été utilisé pour réduire les coûts dans diverses applications incluant le dépistage de maladies infectieuses, la découverte de traitements et la génétique. Dans ces applications, l'objectif est souvent de classer les individus comme positifs ou négatifs à partir des résultats de tests par groupes initiaux et du processus ultérieur de décodage des regroupements positifs. Beaucoup d'algorithmes de décodage ont été proposés mais la plupart ne réussit pas à identifier et à exploiter plus avant la nature hétérogène des individus dépistés. Dans cet article, nous utilisons les probabilités de risque des individus pour formuler de nouveaux algorithmes informatifs de décodage qui appliquent le retesting de Dorfman en population hétérogène. Nous introduisons le concept de valeur-seuil pour classer les individus comme à « haut » ou « bas risque », si bien que des algorithmes séparés spécifiques à chaque groupe de risque peuvent être appliqués, tout en identifiant simultanément les tailles des regroupements qui minimisent l'espérance du nombre de tests. En comparant avec d'autres algorithmes concurrents qui traitent la population comme homogène, nous montrons que des gains significatifs dans l'efficacité du test peuvent être réalisés sans perte virtuelle de précision du dépistage. Un avantage supplémentaire important est que nos nouvelles procédures sont faciles à implémenter. Nous appliquons nos méthodes à des données de chlamydia et gonorrhéa collectées récemment dans le Nebraska au sein du projet de prévention de l'infertilité.

G. A. Whitmore, G. Zhang, and M.-L. Ting Lee

297

Constructing Normalcy and Discrepancy Indexes for Birth Weight and Gestational Age Using a Threshold Regression Mixture Model

Le poids de naissance et l'âge gestationnel sont des mesures importantes de la santé intrinsèque du nouveau-né, servant à la fois de critère de jugement et de variables explicatives dans les études de santé. Ces mesures sont fortement corrélées mais occasionnellement incohérentes.

Nous préfigurons que des indices synthétiques reposant sur le poids de naissance et l'âge gestationnel et fournissant des indications précises de l'état de santé du nouveau-né rendraient un grand service aux chercheurs oeuvrant dans le domaine de la santé ainsi qu'aux autres scientifiques.

Nous proposons deux indices appelés « mesure de normalité de naissance » (Birth Normalcy Index ou BNI) et « mesure d'incohérence de naissance » (Birth Discrepancy Index ou BDI)

Ces deux indices sont des fonctions simples du poids de naissance et de l'âge gestationnel ; par construction leurs transformations logarithmiques sont orthogonales. Le BNI décrit grossièrement si le poids et l'âge gestationnel pris conjointement sont anormaux. Le BDI décrit dans quelle mesure le poids de naissance et l'âge gestationnel sont incohérents. Nous présentons un modèle de mélange à trois composantes pour le BNI, les composantes représentant la prématurité, les naissances à risques et les naissances sans complications. La loi du BNI, déduite d'un modèle stochastique du développement fœtal proposé par Whitmore and Su (2007) prend la forme d'un mélange de lois gaussiennes inverses. Nous

modélisons le BDI par une loi de Student décentrée. Le BNI et le BDI sont aussi adaptés pour comparer les données de naissance (poids et âge gestationnel) dans différentes populations de références.

Pour ces comparaisons, nous proposons de simples z-scores et t-scores. La loi de probabilité des BNI et BDI peut-être estimée pour les naissances d'une population de référence grâce à un modèle de régression seuillée.

F. Han and W. Pan

307

A Composite Likelihood Approach to Latent Multivariate Gaussian Modeling of SNP Data with Application to Genetic Association Testing

Plusieurs tests statistiques ont été proposés pour des données cas-témoins en vue de détecter des associations de maladies avec de multiples polymorphismes mononucléotidiques (PMN) en déséquilibre de liaison. La principale raison de la multiplication des tests est que chaque test vise à détecter un ou deux aspects parmi de nombreuses possibilités de différence distributionnelle entre les cas et les témoins, faute d'un modèle général et cependant suffisamment simple pour traiter des données génotypiques discrètes. Nous proposons ici un modèle de variable latente pour représenter les données PMN : les données observées sont supposées obtenues en discrétisant une distribution multivariante gaussienne latente. De ce fait la distribution latente est complètement caractérisée par son vecteur de moyennes et sa matrice de variance-covariance, contrairement à des formes beaucoup plus complexes de distribution générale pour des données PMN discrètes multivariées. Nous proposons une méthode de vraisemblance composite pour estimer les paramètres. Une application directe de ce modèle de variable latente est le test d'une association avec des PMN multiples dans un gène (ou une région du génome) candidat. Contrairement à plusieurs tests existants, qui ne visent qu'à détecter un ou deux aspects parmi de nombreuses possibilités de différence distributionnelle pour des données PMN discrètes, nous pouvons mettre l'accent sur le test des paramètres, moyennes et covariances, de la distribution gaussienne latente pour les cas et pour les témoins. Nos résultats de simulation démontrent les gains de puissance potentiels de la méthode proposée par rapport à quelques méthodes existantes.

Z. J. Daye, J. Chen, and H. Li

316

High-Dimensional Heteroscedastic Regression with an Application to eQTL Data Analysis

Nous considérons le problème de la régression de grande dimension dans des conditions où la variance d'erreur n'est pas constante. En dépit d'être un phénomène courant dans les applications biologiques, l'hétéroscédasticité a été jusqu'ici largement ignorée dans les analyses génomiques à grande dimension. Nous proposons une nouvelle méthodologie qui permet d'avoir des variances d'erreur non constantes pour des estimations et des sélections de modèle en grande dimension.

Notre méthode prend en compte l'hétéroscédasticité par une modélisation simultanée de la moyenne et des composantes de variance à l'aide d'une nouvelle approche régularisée. Des simulations de Monte Carlo extensives indiquent que notre méthode peut donner lieu à une meilleure estimation et à une meilleure sélection de variables que les méthodes existantes lorsque l'hétéroscédasticité est due à la présence de prédicteurs expliquant les variances d'erreurs et les données atypiques. De plus nous démontrons la présence

d'hétéroscédasticité et nous appliquons notre méthode à une étude (eQTLs) sur 112 séparateurs de levure. La nouvelle procédure peut automatiquement tenir compte de l'hétéroscédaticité en identifiant les eQTLs associés à des variations d'expression des gènes et conduire à de plus petites erreurs de prédiction.

Ces résultats démontrent l'importance de considérer l'hétéroscédasticité dans des analyses de données eQTL.