Found 6 Records

CONTROL ID: 3296407
PRESENTER: Collins Ojwang Odhiambo
PRESENTER (INSTITUTION ONLY): Strathmore University
TITLE: Assessing Efficient Risk Ratios: An Application to Surgical Stage Prediction in Cervical Cancer
ABSTRACT BODY:
Abstract Body: Background: Cervical cancer remains the second most commonly diagnosed cancer and the third leading cause of cancer death in developing countries. Improving clinicians' knowledge and understanding of surgico-pathological staging is critical in the fight against the disease. However, a systematic evaluation of different ordinal regression models based on diverse predicted outcomes has not been given its due share in literature.
Objective: To systematically assess the flexibility of odds ratios for three popular ordinal regression models i.e. the Cumulative Proportional Odds model (CPOM), The Continuation Ratio model (CRM) and Adjacent Category Logistic Model (ACLM) when applying cervical cancer data in surgical stage prediction.
Method: We systematically, compared the performance of CRM, CPOM and the ACLM as the predictive mechanisms, and evaluate the most appropriate model in the cervical cancer setting. The study considered women who visited the Oncology department at the Moi Teaching and Referral Hospital's Chandaria Cancer and Chronic Diseases Center and were diagnosed and surgically treated for cervical cancer from January 2014 to December 2018.
Results and conclusion: We presented the comparison between 3 different regression models for ordinal data with respect to goodness-of-fit under cervical cancer setting. We found that the CRM model without proportional odds yielded better results. All the 5 independent features selected for classifying the patients into surgical stages had significant adjusted odds ratio with the following predictor variables the FIGO clinical stage, the presence or absence of cancer in the vaginal and parametrial regions, the presence or absence of symptomatic vaginal discharge and lower abdominal pain.
AUTHORS/INSTITUTIONS: C.O. Odhiambo, SIMS, Strathmore University, Nairobi, KENYA|

CONTROL ID: 3358416

PRESENTER: Tom Duchemin

PRESENTER (INSTITUTION ONLY): Conservatoire National des Arts et Métiers, Malakoff Humanis

TITLE: A statistical algorithm for sick leave outbreak detection at the workplace

ABSTRACT BODY:

Abstract Body: The workplace is a major vector for the spread of infectious diseases as influenza or gastroenteritis. It is also a place of development of emerging diseases related to working conditions and environment such as burnout or musculoskeletal disorders. Sick leave data are easily accessible to companies and an abnormal amount of sick leave spells can be the consequence of a high incidence of this type of phenomenon. Thus, the identification of sick leave outbreaks could be a useful contribution to help decision makers detect indirectly any potential changes that might affects employees' physical or mental health.

In this context, we propose an adaptation of the Farrington algorithm [1],[2] for monitoring sick-leave data at the company level. The Farrington algorithm, used successfully in routine infectious disease surveillance systems, is based on a Quasi-Poisson regression model to predict potential outbreak after adjusting for linear trend and seasonality and underweighting past alerts. Our adaptation extends this model to consider also the within- and between-company characteristics (i.e. socio-demographic structure, multiplicity of sites, employees' insurance type, management). We developed a generalized linear model with covariates and random effect that leads to a specific computation of the threshold limit for the outbreak detection. We evaluated the model through extensive simulations and propose an application based on sick leave data from 1785 companies with more than 50 employees, followed for at least 6 years between 2010 and 2017.

Preliminary simulations provide promising results since it suggests that the performance of our model is close to the performance of the first Farrington model: we obtain a mean Probability of Detection of 58% and a False Positive Rate of 1.24% for outbreaks simulated as in [2]. This algorithm could then be used to help companies to detect drifts in sick leave but could also be used to address any surveillance issue in a multi-site setting.

[1] Farrington CP, Andrews NJ, Beale AJ, Catchpole MA, A statistical algorithm for the early detection of outbreaks of infectious disease. Journal of the Royal Statistical Society Series A. 1996; 159:547-563

[2] Noufaily A, Enki DG, Farrington CP, Garthwaite P, Andrews, NJ, Charlett A, An improved algorithm for outbreak detection in multiple surveillance systems. Statistics in Medicine, 2012 7:1206-1222.

AUTHORS/INSTITUTIONS: T. Duchemin, M.N. Hocine, MESuRS lab, Conservatoire National des Arts et Métiers, Paris, FRANCE|T. Duchemin, Malakoff Humanis, Paris, FRANCE|A. Noufailly, Statistics and Epidemiology, Warwick Medical School, University of Warwick, Coventry, UNITED KINGDOM|

CONTROL ID: 3358700

PRESENTER: Thorsten Dickhaus

PRESENTER (INSTITUTION ONLY): University of Bremen

TITLE: Combining High-Dimensional Classification and Multiple Hypotheses Testing For the Analysis of Big Data in Genetics

ABSTRACT BODY:

Abstract Body: We present the so-called COMBI method for evaluating genome-wide association studies which we have developed in prior work. In contrast to traditional locus-by-locus analyses, COMBI is a multivariate procedure which takes dependencies between different genomic loci into account. This is done by combining methods from machine learning and multiple testing. In a first stage of data analysis, a support vector machine (which is an inherently multivariate classification method) is trained. In a second stage, only the genomic positions with the largest contributions to the resulting classification rule are explicitly tested for association with the phenotype of interest, yielding a drastic dimension reduction. The thresholding of the association p-values for the selected positions is performed by means of a resampling procedure. Some remarks on software implementations of COMBI are made, and real data analyses are presented. The presentation is based on [1] and [2].

References:

[1] Thorsten Dickhaus (2018). Combining high-dimensional classification and multiple hypotheses testing for the analysis of big data in genetics. In: Asis Kumar Chattopadhyay, Gaurangadeb Chattopadhyay (Eds.): Statistics and its Applications. Platinum Jubilee Conference, Kolkata, India, December 2016. Springer Proceedings in Mathematics & Statistics, Vol. 244, 47-50.

[2] Bettina Mieth, Marius Kloft, Juan Antonio Rodriguez, Sören Sonnenburg, Robin Vobruba, Carlos Morcillo-Suarez, Xavier Farre, Urko M. Marigorta, Ernst Fehr, Thorsten Dickhaus, Gilles Blanchard, Daniel Schunk, Arcadi Navarro, Klaus-Robert Müller (2016). Combining Multiple Hypothesis Testing with Machine Learning Increases the Statistical Power of Genome-wide Association Studies. Scientific Reports, Vol. 6, Article 36671.

AUTHORS/INSTITUTIONS: T. Dickhaus, Institute for Statistics, University of Bremen, Bremen, GERMANY|

CONTROL ID: 3367608
PRESENTER: Magnus Münch
PRESENTER (INSTITUTION ONLY): Leiden University, Amsterdam University Medical Centers
TITLE: Drug sensitivity prediction with normal inverse Gaussian shrinkage informed by external data
ABSTRACT BODY:
Abstract Body: Precision cancer medicine relies heavily on accurate drug sensitivity prediction models. These models are often (partly) developed through simultaneous drug screening of tumour derived cell lines. The well-characterised molecular profiles of these cell lines provide ample opportunity for predictive modelling. These types of prediction problems entail modelling multivariate drug responses on high dimensional molecular feature sets in typically > 1000 cell lines. Typically the number of drugs are in the hundreds and the number of molecular features in the thousands. The dimensions of the problem require specialised models and estimation methods. In addition, external information on both the drugs and the features is often available, such as the target molecular pathways and developmental stages of the drugs.

We propose to model the drug responses through a Bayesian linear regression with shrinkage enforced through a normal inverse Gaussian prior. We let the prior depend on the external information on drugs and features, and estimate the model and its hyperparameters in an empirical-variational Bayes framework. Estimation of the hyperparameters leads to a data-dependent procedure that automatically weighs the informativeness of the external drug and feature information. The resulting Bayesian shrinkage model addresses the issues due to the high dimensionality of the problem.

To asses the performance of our model, we present a simulation study. The simulation study highlights the accurate hyperparameter estimation of the empirical Bayes procedure and shows an increase in predictive performance compared to models that do not use external information on the drugs and features. In addition, we demonstrate the usefulness of this model in the publicly available Genomics of Drug Sensitivity in Cancer data.
AUTHORS/INSTITUTIONS: M. Münch, Mathematical Institute, Leiden University, Leiden, NETHERLANDS|M. Münch , M. van de Wiel, Epidemiology and Biostatistics, Amsterdam University Medical Centers, Amsterdam, NETHERLANDS|M. van de Wiel, S. Richardson, G.G. Leday, MRC Biostatistics Unit, Cambridge University, Cambridge, UNITED KINGDOM|

CONTROL ID: 3367692

PRESENTER: Laura Bertien Zwep

PRESENTER (INSTITUTION ONLY): Leiden University, Leiden University

TITLE: A Novel Goodness-of-Fit Measure Enables Identification of Collateral Effects of Antibiotic Resistance in Large-Scale Population Surveillance Data

ABSTRACT BODY:

Abstract Body: The global increase in antimicrobial resistance (AMR) represents a major public health crisis. The current project concerns analysis of collateral effects of AMR, which occur when resistance against one antibiotic leads to either increased resistance, or increased sensitivity against a second antibiotic. Knowledge of such collateral effects is of clinical relevance as these may enable the design of antibiotic dosing schedules that limit AMR. Sensitivity or resistance of pathogens against antibiotics is quantified using the minimum inhibitory concentration (MIC). The MIC reflects the lowest concentration that leads to inhibition of growth, and is typically measured at discrete and exponentially increasing antibiotic concentrations, with an underlying continuous distribution. MIC data for most antibiotics in hospital-isolated pathogens is collected in population surveillance studies. Statistical methodologies to effectively identify both the occurrence, effect size and effect directionality of collateral effects of AMR in pathogens that are associated with clinical infections are however lacking. We describe a novel goodness-of-fit measure to address this challenge.

We consider the situation where the effect of changes in MIC for two antibiotics is tested as a goodness-of-fit measure. All available MIC observations for an antibiotic (A) are used to estimate a marginal MIC distribution. A subset of these MIC observations, with a specified threshold value for the MIC of an antibiotic B indicating resistance, is used to estimate the conditional distribution of antibiotic A. The difference between the means is tested using a dependent sample T-test, where all observations in the conditional sample are contained in the marginal sample. Antibiotic B can also be tested conditionally on resistance to antibiotic A, which can yield very different results, setting this test apart from simple correlation tests. We apply this measure to a large clinical AMR surveillance database containing MIC data measured in the pathogen Escherichia coli. We identify several antibiotic combinations that show distinct collateral effects. We conclude that the developed methodology is of relevance for the systematic identification of collateral effects in other pathogens and population surveillance databases.

AUTHORS/INSTITUTIONS: L.B. Zwep, K.L. Duisters, Mathematical Institute, Leiden University, Leiden, NETHERLANDS|L.B. Zwep, Y. Haakman, J. van Hasselt, Leiden Academic Centre for Drug Research, Leiden University, Leiden, NETHERLANDS|

CONTROL ID: 3368164

PRESENTER: Gerhard Schulze

PRESENTER (INSTITUTION ONLY): Independent Researcher

TITLE: Clinical Outcome Prediction Based On Multi-Omics Data: The Adaptive IPF-Lasso

ABSTRACT BODY:

Abstract Body: Predicting clinical outcomes using multi-omics data, which consists for example of genomics, proteomics or metabolomics data, requires statistical methods that can deal with the situation of a huge number of covariates versus a relatively small number of observations. One option is the application of penalized regression like the Lasso, which results in sparse models because variable selection is part of the model building process.
In contrast to the standard Lasso model, the IPF-Lasso individually penalizes the various groups (modalities) of multi-omics data based on a set of manually entered candidate penalty-factors. However, this advantage has to be paid for with exponentially increasing computation time when there are more than two or three modalities involved.
In order to avoid this problem, we propose an extension of the IPF-Lasso called adaptive IPF-Lasso which follows a two-step approach:
In step 1 the penalty factors are generated based on the outcome of a penalized regression model, non-relevant modalities are removed from the input data, and the penalty factors are fed into the original IPF-Lasso in step 2.
Several variants of the adaptive IPF-Lasso model are evaluated and compared to the original IPF-Lasso and other competitors by analyzing simulated and real data sets. The results show that for a given data set, the new adaptive IPF-Lasso needs only a fraction of the computation time of the original IPF-Lasso model and nevertheless can compete in terms of prediction performance.
An R-function for implementing the adaptive IPF-Lasso is included into version 1.1 of the R-package ipflasso.

AUTHORS/INSTITUTIONS: G. Schulze, Independent Researcher, Buehl, GERMANY|A. Boulesteix, R. Hornung, Department of Medical Information Processing, Biometry and Epidemiology (IBE), Ludwig-Maximilians-University Munich, Munich, GERMANY|