Biostatistics

Bioinformatics

Biomathematics

Bio^3 Seminar Series

On the Asymptotic Distribution of Likelihood Ratio Test when Parameters Lie on the Boundary

Leonid Kopylev, PhD
U.S. EPA/ORD/NCEA

Bimal Sinha, PhD
University of Maryland at Baltimore County

Abstract:
This talk discusses statistical inference dealing with the asymptotic theory of likelihood ratio tests when some parameters may lie on boundary of the parameter space. We first discuss nature of problems when parameters are on the boundary. Then, following seminal paper by Self and Liang (1987), we derive a closed form solution for the case when one parameter of interest and one nuisance parameter lie on the boundary.

The asymptotic distribution is not always a mixture of several chi-square distributions. For the cases when one parameter of interest and two nuisance parameters or two parameters of interest and one nuisance parameter are on the boundary, we provide an explicit solution which can be easily computed by simulation. These results can be used in many applications, e.g. one-sided confidence intervals in environmental risk assessment and testing for random effects in genetics. We discuss claims of some authors in the applied literature and the difficulty of dealing with parameters on the boundary in applications.

Friday, September 24, 2010 10:00-11:00 am

Warwick Evans Conference Room, Building D

Refreshments will be provided at 9:45am

Sponsored by the Department of Biostatistics, Bioinformatics, and Biomathematics
This seminar series meets the 2nd and 4th Friday of every month.

Biostatistics

Bioinformatics

Biomathematics

Bio^3 Seminar Series

Estimating effects by combining instrumental variables with case-control designs

Constantine Frangakis, PhD
Professor
Department of Biostatistics
Johns Hopkins University

Abstract:
The instrumental variables framework is commonly used for the estimation of causal effects from cohort samples. However, the combination of instrumental variables with more efficient designs such as case-control sampling requires new methodological consideration. For example, as the use of Mendelian randomization studies is increasing and the cost of genotyping and gene expression data can be high, the analysis of data gathered from more cost-effective sampling designs is of prime interest. We show that the standard instrumental variables analysis does not appropriately estimate the causal effects of interest when the instrumental variables design is combined with the case-control design. We also propose a method that can estimate the causal effects in such combined designs. We illustrate the method with a study in oncology.

The work is joint with Russell T. Shinohara, Konstantinos Tsilidis and Elizabeth A. Platz.

Friday, October 8, 2010 10:00-11:00 am

Warwick Evans Conference Room, Building D

Refreshments will be provided at 9:45am

Sponsored by the Department of Biostatistics, Bioinformatics, and Biomathematics
This seminar series meets the 2nd and 4th Friday of every month.

Lombardi Comprehensive Cancer Center, Georgetown University Medical Center

Biostatistics

Bioinformatics

Biomathematics

Bio^3 Seminar Series

Challenges in next generation sequencing data analysis

Yunlong Liu, PhD
Assistant Professor of Medical and Molecular Genetics
Indiana University School of Medicine

Abstract:
Next generation sequencing technology promises to revolutionize life science in fields as diverse as human genomics and transcriptome, species discovery, personalized medicine, and any other field in which advancement depends on a deeper understanding of genomic variation and transcriptome complexity. In this talk, I will review the variety of biological applications enabled by this technology, and present the statistical challenges while interpreting the data. I will also introduce one statistical approach we developed to identify microRNA promoters using ChIP-seq-derived RNA polymerase II binding data.

Friday, October 22, 2010 10:00-11:00 am
Warwick Evans Conference Room, Building D
Refreshments will be provided at 9:45am

Sponsored by the Department of Biostatistics, Bioinformatics, and Biomathematics
This seminar series meets the 2nd and 4th Friday of every month.

Biostatistics

Bioinformatics

Biomathematics

Bio^3 Seminar Series

Censored Quantile Regression Redux

Roger Koenker, PhD
Professor of Statistics, William B. McKinley Professor of Economics
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Abstract:

Quantile regression for censored survival (duration) data offers a more flexible alternative to the Cox proportional hazard model for some applications. We describe three estimation methods for such applications that have been recently incorporated into the R package quantreg: the Powell (1986) estimator for fixed censoring, and two methods for random censoring, one introduced by Portnoy (2003), and the other by Peng and Huang (2008). The Portnoy and Peng-Huang estimators can be viewed, respectively, as generalizations to regression of the Kaplan-Meier and Nelson- Aalen estimators of univariate quantiles for censored observations. Some asymptotic and simulation comparisons are made to highlight advantages and disadvantages of the three methods.

Friday, November 12, 2010 10:00-11:00 am

Warwick Evans Conference Room, Building D

Refreshments will be provided at 9:45am

Sponsored by the Department of Biostatistics, Bioinformatics, and Biomathematics
This seminar series meets the 2nd and 4th Friday of every month.

Biostatistics

Bioinformatics

Biomathematics

Bio^3 Seminar Series

Epidemiologic methods are useless. They can only give you answers

Miguel Hernán, PhD

Associate Professor of Epidemiology

School of Public Health, Harvard University

Abstract:

The first duty of any epidemiologist is to ask a relevant question. Learning and applying sophisticated epidemiologic methods is of little help if the methods are used to answer irrelevant questions. This talk will discuss the formulation of research questions in the presence of time-varying treatments and treatments with multiple versions, including pharmacological treatments and lifestyle exposures. Several examples will show that discrepancies between observational studies and randomized trials are often not due to confounding, but to the different questions asked.

Friday, December 10, 2010 10:00-11:00 am

Warwick Evans Conference Room, Building D

Refreshments will be provided at 9:45am

Sponsored by the Department of Biostatistics, Bioinformatics, and Biomathematics

This seminar series meets the 2nd and 4th Friday of every month.

Lombardi Comprehensive Cancer Center, Georgetown University Medical Center

Biostatistics

Bioinformatics

Biomathematics

Bio^3 Seminar Series

Modeling Protein Mass Spectra TOF Data Using an Ion Flight Mixture

Mei-Ling Ting Lee, PhD

Professor, Department of Epidemiology and Biostatistics
Director, Biostatistics and Risk Assessment Center
University of Maryland, College Park

Abstract:

MALDI-TOF mass spectrometry represents a very fast high-throughput technology for proteomic research. A MALDI-TOF spectrum of a tryptic digest can be acquired in less than one minute and once the sample is spotted on the plate, it can be reanalyzed several times if needed. However, peptide fingerprinting using MALDI-TOF has many limitations. It cannot distinguish between two different sequences of the same mass unless the mass analyzer has sufficient resolution to resolve the two peptides and it cannot identify several proteins in a mixture. We aim to develop better statistical methods to improve MALDI data interpretation with the goal of overcoming some of the limitations of time-of-flight (TOF) analyzers. A mathematical mixture of first hitting time distributions is proposed as a unifying statistical model for the analysis of TOF mass spectrometry data. The model recognizes the time of flight of an ion as a first hitting time and, additionally, models the ion stream as a Poisson process. The model guides how a target protein mass spectrum may be deconvolved into signatures of known ions from protein and peptide data bases. The inference methods and ideas are illustrated using a data set.

Friday, January 28, 2011 10:00-11:00 am

Warwick Evans Conference Room, Building D

Refreshments will be provided at 9:45am

Sponsored by the Department of Biostatistics, Bioinformatics, and Biomathematics
This seminar series meets the 2nd and 4th Friday of every month.

Biostatistics

Bioinformatics

Biomathematics

Bio^3 Seminar Series

Designs for phase I trials of combinations of agents

Mark Conaway, PhD
Professor, Biostatistics and Epidemiology
University of Virginia

Abstract:
Most methods for the design of Phase I trials in oncology are intended for studies involving a single cytotoxic agent. The goal of these studies is to estimate the maximally tolerated dose, the highest dose that can be administered with an acceptable level of toxicity. A key assumption of these methods is the monotonicity of the dose-toxicity curve; administration of greater doses of the agent can be expected to produce dose limiting toxicities (DLT's) in increasing proportions of patients. In this case, the dose-toxicity curve is said to follow a 'simple order' because the ordering of the probabilities of a DLT for any pair of doses is known.

It is becoming increasingly common for combinations of agents to be tested in phase I trials. In these studies, the probabilities of a DLT associated with the dose combinations often follow a 'partial order' in that there are pairs of dose combinations for which the ordering of the probabilities is not known. This talk will review existing methods and outline new methods that combine features of the continual reassessment method and order restricted inference to develop designs for partially ordered phase I trials. Even though the emphasis is on phase I trials of multiple agents in combination, the methods we develop can shed light on other issues in phase I trial design, including the study of ordered groups and phase I trials that combine measures of toxicity and efficacy.

Friday, February 11, 2011 10:00-11:00 am

Warwick Evans Conference Room, Building D

Refreshments will be provided at 9:45am

Sponsored by the Department of Biostatistics, Bioinformatics, and Biomathematics
This seminar series meets the 2nd and 4th Friday of every month.

Biostatistics

Bioinformatics

Biomathematics

Bio^3 Seminar Series

A New Class of Dantzig Selectors for Censored Linear Regression Models with Applications in Genomic Studies

Yi Li, PhD

Associate Professor of Biostatistics

Department of Biostatistics, Dana-Farber Cancer Institute

Harvard School of Public Health

Abstract:

The Dantzig variable selector has recently emerged as a powerful tool for fitting regularized regression models. A key advantage is that it does not pertain to a particular likelihood or objective function, as opposed to the existing penalized likelihood methods, and hence has the potential for wide applicability. To our knowledge, limited work has been done for the Dantzig selector when the outcome is subject to censoring. This paper proposes a new class of Dantzig variable selectors for linear regression models for right-censored outcomes. We first establish the finite sample error bound for the estimator and show the proposed selector is nearly optimal in the `2 sense. To improve model selection performance, we further propose an adaptive Dantzig variable selector and discuss its large sample properties, namely, consistency in model selection and asymptotic normality of the estimator. The practical utility of the proposed adaptive Dantzig selectors is verified via extensive simulations. We apply the proposed methods to a myeloma clinical trial and identify important predictive genes for patients' survival.

Friday, February 25, 2011 10:00-11:00 am

Warwick Evans Conference Room, Building D

Refreshments will be provided at 9:45am

Sponsored by the Department of Biostatistics, Bioinformatics, and Biomathematics

This seminar series meets the 2nd and 4th Friday of every month.

Lombardi Comprehensive Cancer Center, Georgetown University Medical Center

Biostatistics

Bioinformatics

Biomathematics

Bio^3 Seminar Series

Challenges in Program Evaluation

David Judkins

Senior Scientist, WESTAT

Fellow, American Statistical Association

Program Chair for the 2011 Joint Statistical Meetings

Abstract:

This talk will focus on the role of the statistician in program evaluation. Mr. Judkins has worked on evaluations of WIC, an anti-drug media campaign, a pro-exercise media campaign, a college access program for students from high poverty middle schools, a "waiver" from standard regulations allowing child welfare workers in the state of Ohio to use local initiatives to try to improve outcomes for abused and neglected children, and a curriculum reform program in preschools serving mostly children of high-school dropouts and recent immigrants. He will be urging statisticians to see their roles rather as being rather broad. It can be difficult to influence research questions, data collection design, and data processing, but he will argue that statisticians should struggle to be involved in all these areas to some extent. He will also talk about the role of randomization in program evaluation and the utility of some of the currently buzzworthy analytic procedures like propensity score matching, multiple imputation, and multi-level modeling. Finally, we will have some comments on choosing your camp in the statistical tribe: Bayesian, model-based frequentist, or design-based frequentist.

Friday, March 25, 2011 10:00-11:00 am

Warwick Evans Conference Room, Building D

Refreshments will be provided at 9:45am

Sponsored by the Department of Biostatistics, Bioinformatics, and Biomathematics

This seminar series meets the 2nd and 4th Friday of every month.

Lombardi Comprehensive Cancer Center, Georgetown University Medical Center

Biostatistics

Bioinformatics

Biomathematics

Bio^3 Seminar Series

A sub-set based approach to meta-analysis to improve power and interpretation of combined analysis of heterogeneous traits

Nilanjan Chatterjee, PhD
Chief of the Biostatistics Branch
Division of Cancer Epidemiology and Genetics (DECG)
National Cancer Institute (NCI)

Abstract:

Genome-wide association studies can identify regions in the genome that should lead to insights into both distinct and shared etiologies among heterogeneous, but possibly related, traits. Standard approaches for meta-analysis or pooled-analysis, however, may not be optimal for combined studies of distinct traits in which individual variants are associated with only a subset of the traits or demonstrate different directions for distinct traits. We propose an approach that 'agnostically' explores subsets of traits to identify the strongest association signal and then evaluates the significance of the detected association using efficient adjustment for multiple testing that accounts for correlation among the different tests. A two-sided version of the test combines association signals in opposite directions. Simulation studies illustrate that the method can achieve a major increase in power for detecting susceptibility loci over traditional methods. The method also performs well in detecting the correct subset of traits that are likely to be associated with a specific variant. An illustrative application of the method on a total of >20000 cases and >20000 controls for seven different cancers identified secondary effects for known cancer susceptibility loci in additional cancer types.

Friday, April 8, 2011 10:00-11:00 am

Warwick Evans Conference Room, Building D

Refreshments will be provided at 9:45am

Sponsored by the Department of Biostatistics, Bioinformatics, and Biomathematics

This seminar series meets the 2nd and 4th Friday of every month.

Lombardi Comprehensive Cancer Center, Georgetown University Medical Center