2012 Seminars

11/14/2012

Overview of Agreement Statistics for Continuous, Binary, and Ordinal Data

Lawrence Lin, PhD

Biography: Dr. Lawrence I. Lin has recently retired after 33 years of distinguished tenure at Baxter International Inc. He is a Principal Consultant at JBS Consulting Services. He is an Adjunct Professor in the Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago. Dr. Lin is a Fellow of the American Statistical Association, and an elected member of the International Statistical Institute. He has served on as a referee of many international journals.

Abstract: This will be a general overview presentation with practical examples and without much statistical formulas. We will introduce the concepts of un-scaled and scaled agreement statistics based on the basic case between two raters with paired samples for continuous, binary, and ordinal data. We will then progress into more complex cases when we have multiple raters and each rater has multiple readings per sample. Here, we can assess intra-rater and inter-rater agreement, compare inter-rater deviation to intra-rater deviation, and compare precision of a rater against another. We will explore the meaning of the two-stage criteria presented in the FDA guidance UCM070244: Statistical Approaches to Establishing Bioequivalence. The content is largely based on the materials presented in the newly published book by Springer, entitled “Statistical Tools for Assessing Agreement”.

04/18/2012

Some problems in the analysis of high-dimensional models

Anthony Gamst, Ph.D

Biography: Dr. Gamst is a Professor in the Division of Biostatistics and Bioinformatics, Department of Family Medicine and Public Health at UCSD. He is the author of over 90 papers in various areas of biostatistical applications and methodology, with predilection in imaging analysis and Alzheimer's disease.

Abstract: Models with large numbers of nuisance parameters are common in modern statistics, having applications in laboratory medicine, genomics, clinical trials, medical imaging, epidemiology, and many other areas. Classical techniques, including Bayes and Maximum Likelihood, tend to produce sub-optimal or even inconsistent estimates of the parameters of interest in these models, when naively applied, while approximately unbiased estimating equations work rather generally. We study several such models, identify the sources of bias and spurious correlation which lead to inconsistency or sub-optimality, and compute the minimal smoothness required for the existence of root-n consistent (and efficient) parameter estimates. We also examine simultaneous estimation of nuisance parameters and parameters of interest. The results of the study are related to every-day practice, particularly to the fitting of regression models with many predictors, and some heuristics are given.

04/06/2012

Integrated Statistical Methodology for the Analysis of High-Dimensional Data with Applications to Translational Cancer Research

Kim-Anh Do, Ph.D

Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston

Abstract: Early detection is critical in disease control and prevention. The long term translational research goal is that if drugs can be targeted to specific tissues in the body, then dosage can be altered to achieve the desired effect while minimizing side effects such as toxicity. Motivated by specific problems involving high throughput data in the form of phage peptides, we have developed nonparametric and semiparametric mixture models for real-time analysis in the context of correlated phage experiments. Our main focus is to address the multiplicity issue automatically by incorporating a false discovery rate or utility function. We will highlight direct applications of both frequentist and Bayesian methods to cancer research challenges that address our long term translational goal. Specifically, the developed statistical methodology can assist in isolating ligand peptides and identify their corresponding tissue-specific receptors in rodent models and in patients, including discovery and validation of a ligand-receptor tumor targeting system in human metastatic prostate cancer.

03/23/2012

Residual Life; A Useful Summary Measure for Survival Data?

Jong-Hyeon Jeong, PhD

Department of Biostatistics, University of Pittsburgh

Abstract: The hazard function is a popular summary measure of time-to-event or survival data from medical studies. However, translation of the study results based on the hazard function might not be straightforward for the stakeholders like patients and physicians. Therefore, consideration of the remaining life years to events of interest might be more useful. In time-to-event data, the issue of competing risks is often encountered, whenever the events of interest are precluded from being observed, due to some competing events. In this talk, statistical methods that recently have been developed to infer quantile residual life under competing risks will be presented. Some issues to be overcome for further generalization of the proposed methods will be also discussed. The proposed methods will be illustrated with a real dataset from a phase III clinical study on breast cancer with a long-term follow-up of more than 30 years.

03/21/2012

Bayesian Survival Trees for Clustered Observations, Applied to Tooth Prognosis

Richard Levine, PhD

Professor and Chair, Department of Mathematics and Statistics, San Diego State University

Abstract: Tooth loss from periodontal disease or dental caries (decay) afflicts most adults over the course of their lives. Survival tree methods for correlated observations have shown potential for developing objective tooth prognosis systems, however the current technology suffers either from prohibitive computational expense or unrealistic simplifying assumptions to overcome computational demands. In this talk Bayesian tree methods are developed for correlated survival data, relying on a computationally feasible, yet flexible, frailty model with piecewise constant hazard function. Bayesian stochastic search methods, using a Laplace approximated marginal likelihood, are detailed for tree construction and posterior ensemble averaged variable importance ranking and amalgamation procedures are developed to identify indicators of tooth prognostic groups from a forest of trees. The proposed methods are used to assign each tooth from the VA Dental Longitudinal Study to one of five prognosis categories and evaluate the effects of clinical factors and genetic polymorphisms in predicting tooth loss. The prognostic rules established may be used in clinical practice to optimize tooth retention and devise periodontal treatment plans.

03/07/2012

Estimating Abundances of Retroviral Insertion Sites from DNA Fragment Length Data.

Chuck Berry, PhD

Chuck Berry is Professor Emeritus and Interim Division Chief of the Division of Biostatistics and Bioinformatics, Department of Family Medicine and Public Health at UCSD. He authored over 180 papers on the methodology and applications of Biostatistics in Medical Sciences, and he is actively involved in several research projects, with a particular emphasis on statistical genetics.

Abstract: The relative abundance of retroviral insertions in a host genome is important in understanding the persistence and pathogenesis of both natural retroviral infections and retroviral gene therapy vectors. When host genomic DNA is randomly broken via sonication and then amplified, amplicons of varying lengths are produced. A likelihood function is proposed for these lengths along with a hybrid Expectation-Maximization algorithm. Patient data illustrate the method and simulations show that relative abundance can be estimated with little bias, but that variation in highly abundant sites can be large. Reference: Charles C. Berry, Nicolas A. Gillet, Anat Melamed, Niall Gormley, Charles R.M Bangham, and Frederic Bushman Estimating Abundances of Retroviral Insertion Sites from DNA Fragment Length Data. Bioinformatics. first published online January 11, 2012 https://www.ncbi.nlm.nih.gov/pubmed/22238265