ALAN KER, University of Guelph

Semiparametric Estimation of the Link Function in Binary-Choice Single-Index Models

In this manuscript we propose a new semiparametric estimator for binary-choice single-index models which uses parametric information in the form of a known link function and nonparametrically corrects it. Asymptotic properties are derived and the finite sample performance of the proposed estimator is compared to those of the parametric probit and semiparametric single-index model estimator of Klein and Spady.

FRANÇOIS LAROCHELLE, Research and Development Canada, Department of National Defence

A Complementary Perspective on the Use of ROC Curve to Assess Logistic Regression Models

The Receiver Operating Characteristics (ROC) curve is commonly used to assess the power of a logistic regression model for predicting a binary outcome. In this presentation, we consider questions such as ‘how do the estimated model parameters affect the Area under the ROC Curve (AUC)’ and ‘what is the maximum AUC that can empirically be observed’. Based on the answers, we discuss the benefits of assessing the goodness-of-fit of a fitted model by comparing the model’s AUC with the AUC for the saturated model. Examples from the analysis of non-response for the Canadian Forces ‘Your-Say’ Survey are discussed.

JINHUI MA, McMaster University

Comparison of Marginal and Cluster-Specific Models in Analysis of Cluster Randomized Trials with Missing Binary Outcomes

Marginal and cluster-specific models are commonly used to analyze binary outcomes in cluster randomized trials to estimate the population-averaged and the cluster-specific treatment effect, respectively. However, limited attention has been paid to their performance when there are missing outcomes and some multiple imputation strategies are used to handle the missing data. Under the assumption that the probability of having missing outcomes is covariate dependent, we compare the performance of the generalized estimating equations and random-effects logistic models using a simulation study. The methods are compared in terms of standardized bias, coverage, root mean squared error and average standard error.

BINGRUI SUN, Memorial University of Newfoundland

Correlations Versus Odds Ratios Based Inferences in Bivariate Binary Models

When bivariate binary responses are collected from a large number of independent individuals in a cross-sectional setup, it is standard to use a multinomial distribution with a cell probability modeled as a function of two marginal parameters and an association parameter. This approach makes the interpretation of marginal probabilities in terms of the model parameters difficult. In this talk, we model the marginal probabilities as functions of corresponding marginal parameters only and a correlation based conditional probability function is modeled to obtain the joint probabilities. Likelihood and quasi-likelihood estimation for inference about this model are discussed.

RONG CHU, McMaster University

Methodological Issues on Prognostic Imbalance in Randomized Controlled Trials (RCTs) Evaluating A Binary Outcome

Chance imbalance in baseline prognostic factors (PF) can lead to over or underestimation of treatment effects in RCTs. We simulated data from RCTs by varying risk of the outcome, effect of the treatment, power and prevalence of the PF, and sample size. Logistic regression models with and without adjustment for the PF were compared in terms of bias, standard error,
coverage of confidence interval and statistical power. Covariate adjustment improves estimation accuracy and efficiency. The probability of prognostic imbalance in small trials can be substantial. Two thousand patients may be needed to minimize the chance and impact of prognostic imbalance.

RAFIQUL CHOWDHURY, University of Western Ontario
A Prediction Model to Identify Elderly Who Get Hospitalized – An Application Using Longitudinal Data

One of the most expensive services used by older persons is hospitalization or nursing home use in developed countries. It is a great challenge of the policy-makers for the cost-effective allocation of health care resources for aging populations. It is important to identify that elderly people living in the community are at risk for a variety of adverse outcomes which might lead to hospitalization. In this paper we presented a predictive model based on a regressive model approach to predict future hospitalizations among the elderly. The application is based on repeated measure data from Health and Retirement Study (HRS) in USA.