



The University of Chicago
Department of Statistics

Ph.D. Seminar

IRINA OSTROVNAYA

Department of Statistics
The University of Chicago

**Estimating Error Rates for Independent
and Dependent Test Statistics**

Tuesday, July 25, 2006 at 3:00 PM
110 Eckhart Hall, 5734 S. University Avenue

ABSTRACT

Modern technology has allowed us to produce and process very large quantities of data making simultaneous testing of multiple hypotheses standard practice in many fields. One of the main problems in estimating simultaneous error rates for large sets of test statistics, including False Discovery Rate, is to evaluate the proportion of null hypotheses. This proportion is directly related to the minimum of the density of the p-value distribution. We have developed a new estimator for the minimum of the density that is based on constrained multinomial likelihood functions. The proposed method involves partitioning the support of the density into several intervals, calculating the number of observations in each interval, and estimating multinomial probabilities which are functions of the density. We evaluate theoretical properties of the estimator, and discuss optimal choices for the partition. A simulation study was performed and the results show that this method outperforms other available procedures designed for estimating the proportion of null hypotheses.

In many multiple testing settings, the test statistics are dependent. The proposed methodology can be extended to the case of the dependent observations by using weighted univariate likelihoods. Several criteria of optimality are considered, and optimal weights are obtained using the theory of estimating equations. The optimal weights depend on the discretized pairwise joint distributions of the observations. Approximations of these bivariate joint distributions can be obtained in many applications, and we discuss the estimation of the optimal weights for a two-sample problem motivated by testing for differential expression using microarray data.