Stony Brook University
The Graduate School

Doctoral Defense Announcement

Abstract

Power Analysis of Finite Mixtures of Poisson Distributions

By

Christine M. Brady

One of the most common types of studies that occur is the comparison of responses between a control group and treatment group. Typically, it is assumed that each group is homogeneous. However, when dealing with count data the phenomenon of overdispersion often occurs. This phenomenon may be due to heterogeneity that exists within the group. In such cases, a mixture of distributions is often used to account for such heterogeneity.

We developed a likelihood ratio test for comparing two groups assuming a two-component Poisson mixture exists within each group. We conducted a power study for the family of alternatives with a one parameter difference from the null hypothesis. For each model considered, we compared the power for the Likelihood Ratio Test to the Welch-Satterwheitte t-test, Wilcoxon Test, and Adjusted Wilcoxon Test.

The power study was done using user-friendly software that we developed which simulated our data. The software obtains the maximum likelihood estimates of the parameters under the null hypothesis, such that the control and treatment groups each follow a two-component Poisson mixture with equal mixing proportions and component means. As well, it can compute the MLE's for the two groups differing in either mixing proportions or exactly one component mean. In addition to simulating data, our program has the capability to input actual data and run similar studies.

We compared the power of the Likelihood Ratio Test using the asymptotic 95^{th} percentile critical value of the chi-squared distribution with one degree of freedom and the 95^{th} percentile asymptotic critical value of the standard normal distribution for the other three tests for sample sizes of 100 and 250 per group. As well, we investigated the empirical null distribution for the LRT. We conducted a similar power study for sample sizes of 100 per group using the 95^{th} percentile empirical value. Generally speaking, the LRT was found to be significantly more powerful than the other tests considered.

We applied our testing procedure for comparing two groups of two-component Poisson mixtures for two sets of count data that were provided. One data set that we studied consisted of the number of fibromas which existed on patients suffering from the disease tuberous sclerosis. The other data set that we applied our procedure to consisted of the number of deviant verbalizations from a study on schizophrenia.

Date: March 28, 2008
Time: 2:00 p.m.
Place: Math Tower, Room 1-122

Program: Applied Mathematics & Statistics
Dissertation Advisor: Dr. Nancy Mendell