

CCSU
DEPARTMENT OF MATHEMATICAL SCIENCES

COLLOQUIUM

Friday, April 11

2:00 – 3:00 PM

Maria Sanford, Room 101

INFERENCES FOR A SINGLE PROPORTION USING CLUSTERED AND NON-CLUSTERED BINARY RESPONSE DATA

KRISHNA SAHA

CENTRAL CONNECTICUT STATE UNIVERSITY

(Joint work with Daniel Miller)

Abstract: In many applied fields, the confidence interval approach is often preferred for making inferences about a population proportion. In elementary statistics courses, we obtain a 95% confidence interval for a single proportion P based on non-clustered binary data as

$$\hat{p} - Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq P \leq \hat{p} + Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}},$$

with $Z_{\alpha/2} = 1.96$. However, this approach leads to two obvious defects, namely *overshoot* and *degeneracy*. In addition, this interval can result in serious under-coverage, that is, the pre-assigned confidence level of 95% may not be achieved, especially for small and large values of the true proportion P when the sample size n is small. Extensive work has been done to propose many alternative intervals, and some of them are already implemented in standard software packages (SAS, SPSS, STATA, StatXact). In this talk, I will discuss some of the alternative methods, and show how the computations can be done using R. In addition, I will discuss the extensions of these methods to cluster studies which arise in many biological and biomedical applications.

For further information:

gotchevi@ccsu.edu 860-832-2839

<http://www.math.ccsu.edu/gotchev/colloquium/>