



The University of Chicago

Department of Statistics

Master's Seminar

HUIMIN GU

Department of Statistics
The University of Chicago

**Exploring Insurance Abusive Claim Identification
with Model-Based Clustering Using the EM Algorithm**

**FRIDAY, October 26, 2007 at 9:00 AM
110 Eckhart Hall, 5734 S. University Avenue**

ABSTRACT

Based on the expectation-maximization (EM) algorithm and the auto bodily injury liability claims in the paper “Modeling Hidden Exposures in Claim Severity Via the EM Algorithm”, we re-create the results in this paper and use the model-based clustering method (implemented in software R as “Mclust”) to confirm the results we obtain. According to Bayesian Information Criterion (BIC), a 2-component mixture model is constructed to fit the claim size data. Using the information on the abusive claims in the original paper, the accuracy of the classification is evaluated through simulation. The results show that the 2-component mixture model is reasonable and may be effective to distinguish the abusive claims. Moreover, a threshold is added to prevent huge losses of abusive claims.