

## **Discretization and Mutual Information Estimation**

Moon Yul HUH

Sungkyunkwan University, Seoul, Korea.

### **Abstract**

I am going to talk about the properties of discretization in estimating mutual information (MI) and suggest a discretization method LU-Chi. The method applies discretization only to a subset of the whole data set, and the Chi-square homogeneity test of a contingency table is used to determine stopping further splitting an interval. The efficiency of the method will be demonstrated using the Monte-Carlo experiments and the real data sets. Also shown will be the efficiency of applying the Laplace adjustment to the cell frequencies of the discretized data sets when the sample size is not large (less than 300).