

# Information identity and applications

Michelle Liou (劉長萱) and Philip E. Cheng\* (程爾觀)

Institute of Statistical Science, Academia Sinica

## Abstract

We proposed information identity to partition the mutual information of categorical variables into orthogonal and interpretable components. In this study, we will apply the information identity to a few real data examples. In these examples, the identities are used for two purposes: (i) measuring conditional association between three categorical variables, as compared with the conventional and the generalized Mantel-Haenszel statistics, and (ii) modeling multivariate categorical variables, as a concise alternative to the conventional hierarchical log-linear model and the logistic regression analysis. The data sets in the empirical examples are those reported in Lahkola *et al.* (2005), Azen & Walker (2011) and one of item responses in a mathematics achievement examination.

Keywords: logistic regression, log-linear model, Mantel-Haenszel test, mutual information identity