Connection of flow graphs and contingency tables

Péter Takács, Zoltán Ernő Csajbók[a]

[a]Faculty of Health, University of Debrecen
takacs.peter@foh.unideb.hu, csajbok.zoltan@foh.unideb.hu

Abstract

In statistics, analysis of categorical data commonly uses contingency tables and related methods. Contingency table investigations in which occur more than three variables are relatively rare. The reason for this is that the calculations are complicated and interpretation of the results will become more complex.

The original rough set theory based on - among others - the concept indiscernibilty - Z. Pawlak, 1982. Main aim of the research was to approximate a fixed set with another sets. The theory introduces the concepts of lower and upper approximation sets and border domain. Pawlak introduced the rough set flow graphs - a graphical framework for reasoning from data - at the millennium. These methodological steps extend the theory of rough sets. Pawlak showed that the information flow in flow graphs is governed by classical Bayes' rule but the interpretation of the new theory distance itself from basis of probability theory. The research and practical applications have demonstrated that the flow graphs is a useful mathematical tool of data analysis.

This paper presents research in decision rules, flow graphs and statistical contingency tables. It can be shown that such concepts as certainty, coverage factors can be contacted with contingency tables. These results make it possible to approach the classical statistical concepts and methods from a different perspective and enable analysis of data on new foundations.

Keywords: flow graphs, rough set theory, statistics, contingency tables

MSC: MSC68R10, 05C21