

International Conference on Advances in Interdisciplinary Statistics and Combinatorics,
October 12–14, 2007, University of North Carolina at Greensboro, Greensboro, North
Carolina, USA

**A CORRECTION OF JACCARD SIMILARITY INDEX FOR
CHANCE AGREEMENT IN CLUSTER ANALYSIS**

AHMED N. ALBATINEH

One distinction of Jaccard similarity index is that its expectation under fixed marginal totals can not be found easily because it is not linear in the matching counts, hence it can not be corrected for chance agreement. In this talk an approximation of Jaccard's expectation will be presented, and hence a correction for chance agreement can be computed. Simulation results before and after correction for a null and structured data will be presented using data generated from a mixture of bivariate normal distributions. Simulations show that the proposed approximation and hence correction for chance agreement improves the performance of the index.

NOVA SOUTHEASTERN UNIVERSITY, FORT LAUDERDALE, FLORIDA
E-mail address: `albatine@nsu.nova.edu`