AN INNOVATIVE APPROACH ON FUZZY CORRELATION COEFFICIENT WITH INTERVAL DATA

Hui-Li Hsu[1] and Berlin Wu[2]

[1]General Education Center
China University of Technology
No.56, Sec.3, Singlong Road, Wunshan District, Taipei, Taiwan
meghsu@cute.edu.tw

[2]Department of Mathematical Sciences
National Chengchi University, Taiwan

ABSTRACT. *We usually use the correlation coefficient to observe the relation of two variables. But if the data are exhibited by a fuzzy type, such as interval data, instead of the real number data, the traditional methods for evaluating correlation coefficients do not work. In this paper, we propose an innovative approach to measure the correlation for variables when the data are in the form of intervals. Such the correlation coefficient interval will be called the fuzzy correlation coefficient. Finally two empirical studies are given to compare the traditional correlation coefficient and the fuzzy correlation coefficient. The results show that the fuzzy correlation coefficient can provides an efficient and creative measurement in the perception measurement.*
Keywords: Fuzzy correlation coefficient, Interval data, Fuzzy statistics

1. **Introduction.** Correlation measurement and causality analysis between two variables are very important topics in the science research work. For any two random variables, we use *Person's correlation coefficient* estimator to compute the degree of correlation/dependence from sampled data. However, in a practical case, since uncertain or incomplete factors might interfere with the data collection so that the observed single-valued (real number) samples cannot fully describe the true situations of the samples from the population. To record the characteristics of such events, sometimes we use the interval data to represent the samples. For example, when we consider the daily temperature, we would like to know the temperature range for a day, and we record it in the interval form [*low, high*] by the low and high temperatures of the day.

Now the problem comes out: how do we calculate the correlation coefficient of interval data? It seems not easy to extend the evaluation method for *Person's correlation coefficient* straightly via the interval arithmetic. Hence, we need to consider applying the concept of the fuzzy operations to investigate the correlation of two variables. Nevertheless, it still is a challenging work to define the correlation for interval data via the concept of fuzzy set theory [19] as well as to evaluate the interrelation of fuzzy sets.

In the literatures, Bustince and Burillo [3], Chiang and Lin [4], Gerstenkorn and Manko [6], Hong and Hwang [10] and Yu [18] have discussed the correlation coefficient with fuzzy numbers. Liu and Kao [11] applied the extension principle of the fuzzy theory to calculate the correlation coefficient for a sample set of *n*-independent pairs of fuzzy observations. Nguyen and Wu [13] pointed out that intervals have the fuzzy characteristic due to their uncertainty. Hasuike and Ishii [7] considered the future return as a fuzzy number and proposed portfolio selection problems including them. Shinkai [15] proposed