Journal of the Korean
Data & Information Science Society
2001, Vol. 12, No.1, pp. 99 ~ 101

A Note on Fuzzy Linear Regression Analysis of Fuzzy Valued Variables

DUG HUN HONG [1]

Abstract

In this note, we show that a linear regression model, using entropy and degree of nearness of fuzzy numbers, suggested by Wang and Li[FSS 36, 125-136] seems to be unreasonable by an example.

Key Words and Phrases:Estimation of parameter; Fuzzy linear regression.

1. Introduction

In general, there are two possible cases in systems in which human intelligence participates : (1) The relations of the variables are subject to fuzziness, and (2) the variables themselves are fuzzy. Wang and Li[2] constructed two different fuzzy linear regression models for the second case, based on possibility theory[3]. But, a linear regression model($\tilde{E}$-D estimation method) using entropy and degree of nearness of fuzzy numbers does not make sense. We will show this by an example.

To begin with, we introduce the $\tilde{E}$-D estimation method. For convenience, we use the same notations as they used.

Let $X_0, X_1, \cdots, X_n$ be variables which take values in $F(R)$ (we call them fuzzy variables) and suppose there exists a linear relation among them:

$$X_0 = \beta_1 X_1 + \cdots + \beta_{n-1} X_{n-1} + \beta_n X_n,$$

where $\beta_i,\ i = 1, \cdots, n$, are unknown real coefficients, and

$$\begin{aligned} X_n(x) &= \begin{cases} 1, & x = 1, \\ 0 & x \neq 1 \end{cases} \\ &\equiv I_1. \end{aligned}$$

[1]Associate Professor, School of Mechanical and Automotive Engineering, Catholic University of Taegu Kyungbuk, 712 - 702, Korea

Generally, let $X_0, X_1, \cdots, X_n$ be fuzzy variables and take values in $\mathcal{L}_{\bar{\mu}}$. We observe these fuzzy variables m times, and in the j-th observation ($j = 1, \cdots, m$), we obtain the sample values x_{ijk} ($k = 1, \cdots, s(i,j)$) of $X_i (i = 0, 1, \cdots, n)$. we call the data $\{x_{ijk} | i = 0, 1, \cdots, n;\ j = 1, \cdots, m;\ k = 1, \cdots, s(i,j)\}$ initial data. Then we determine a $\bar{\mu}$-type fuzzy number X^*_{ij}, the j-th fuzzy observed value of $X_i, i = 0, \cdots, n, j = 1, \cdots, m$, using the method of maximin μ/E estimation. Thus we obtain fuzzy observed values $(X^*_{0j}, X^*_{1j}, \cdots, X^*_{nj})$ of $(X_0, X_1, \cdots, X_n)$, $j = 1, \cdots m$.

Let $\tilde{E} = \sum_{j=1}^{m} |E(X^*_{0j}) - E(\hat{X}^*_{0j})|$, where $E(X^*_{0j})$, $E(\hat{X}^*_{0j})$ denote the entropy of the j-th fuzzy observed value, and the j-th fuzzy estimate value of X_0, respectively. If $\beta_1, \cdots, \beta_n$ have a change, it can lead to a change of $\tilde{E}$ and affect the degree of nearness between the fuzzy observed values and the fuzzy estimate values. Owing to these two aspects, we advance the following principle:

Choose $\beta^*_1, \cdots, \beta^*_n$, such that

$$\sum_{j=1}^{m} |E(X^*_{0j}) - E(\hat{X}^*_{0j})| = \min_{\beta_1, \cdots, \beta_n} \sum_{j=1}^{m} |E(X^*_{0j}) - E(\bar{X}^*_{0j})|$$

subject to

$$\bigwedge_{j=1}^{m} D(X^*_{0j}, \hat{X}^*_{0j}) \geq h, \quad j = 1, \cdots, m,$$

where D is the degree of nearness of fuzzy numbers, h is some given standard of nearness ($h \in (0,1]$), $\hat{X}^*_{0j} = \sum_{i=1}^{n} \beta^*_i X^*_{ij}, j = 1, \cdots, m$.

$\beta^*_1, \cdots, \beta^*_n$ are called $\tilde{E}$-D estimation of $\beta_1, \cdots, \beta_n$. This method is called $\tilde{E}$-D estimation method.

Example. We define a fuzzy number H as follow:

$$H(x) = \begin{cases} 1 - \frac{1}{2}|x| & \text{if } x \in [-1,1], \\ \frac{1}{2} & \text{if } |x| \in [1,5], \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathcal{L}_{\bar{\mu}} = \mathcal{L}_H$ and consider the following model

$$X_0 = \beta_1 X_1 + \beta_2 X_2$$

with standard of nearness $h = 0.5$. Now we consider the data, $(X^*_{0j}, X^*_{ij}) = (H(10 - j, 1), H(j, 1)), j = 1, 2, \cdots, 10$. Then we can easily see that $\beta^*_1 = -1, \beta^*_2 = 10$ is a reasonable possible coefficients to fit the data. But $\beta^*_1 = 1, \beta^*_2 = 0$ is another possible answer, since $E(X^*_{0j}) = E(\bar{X}^*_{0j}) = 5.5$ for all $j = 1, 2, \cdots, 10$ and $D(X^*_{0j}, \bar{X}^*_{0j}) = 0.5$ for $j = 1, 2, \cdots, 10$. As we can see from this fact, the slope of one possible answer is 1 and that of another possible answer is -1. So this model seems to be unreasonable.

Acknowledgments

This work was supported by grant No. 2000-2-10400-002-3 from the Basic Research Program of the Korea Science & Engineering Foundation.

References

[1] Tanaka, H. ,Uejima, S. ,and Asia, K. (1982). Linear regression analysis with fuzzy model, IEEE Trans. Man. Cybernet, Vol. 12(6), pp. 903-907.

[2] Wang, Z. ,and Li, S. (1990). Fuzzy linear regression analysis of fuzzy valued variables, *Fuzzy Sets and Systems*, Vol. 36, pp. 125-136.

[3] Zadeh, L. A. (1978). Fuzzy sets as basis for theory of possibility, *Fuzzy Sets and Systems*, Vol. 1, pp. 3-28.