Proceedings of KFIS 2002 Spring Conference, 2002. 5. 25

퍼지수 데이터를 가지는 퍼지선형회귀직선의 여러 가지 성질들

SOME PROPERTIES OF FUZZY LINEAR REGRESSION WITH FUZZY NUMBER DATA

손미정, 이부영, 박동근, 권영철
동아대학교 수학과
Mi-Jung Son, Bu-Young Lee, Dong-Gun Park and Young-chel Kwun
Major in Mathematics Dong-a University
E-mail : deltasemi@hanmail.net, bylee@daunet.donga.ac.kr
dgpark@daunet.donga.ac.kr, yckwun@daunet.donga.ac.kr

Abstract. In this paper, we consider the fuzzy linear regression line for fuzzy number data.

1. Introduction

The theory of fuzzy mathematics has gained more and more recognition from many researches over the wide range of scientific field. Recently, M. K. Kang, J. R. Choi and E. S. Bae([4]) are studied some properties of fuzzy sample correlation coefficient with fuzzy data.

In this paper, we discuss the fuzzy linear regression line and some properties of estimated fuzzy regression line.

2. Preliminaries

A fuzzy number A in the real line R is a fuzzy set characterized by a membership function μ_A as

$$\mu_A : R \rightarrow [0,1] .$$

A fuzzy number A is expressed as

$$A = \int_{x \in R} \mu_A(x)/x ,$$

with understanding that $\mu_A(x) \in [0,1]$ represents the grade of membership of x in A and $\int$ denotes the union of $\mu_A(x)/x$'s.

A fuzzy number A in R is said to be convex if for any real numbers $x, y, z \in R$ with $x \leq y \leq z$,

$$\mu_A(y) \geq \mu_A(x) \wedge \mu_A(z)$$

with $\wedge$ standing for min.

A fuzzy number A is called normal if the following holds.

$$\max_x \mu_A(x) = 1 .$$

A fuzzy number which is normal and convex is referred to as a normal convex fuzzy number.

An α-level set of a fuzzy number A is a nonfuzzy set denoted by $[A]^\alpha$ and is defined by

$$[A]^\alpha = \{ x \mid \mu_A(x) \geq \alpha,\ 0 < \alpha \leq 1 \}.$$

An α-level set of a fuzzy number A is convex fuzzy set then which is a closed bounded interval which we denoted by

$$[A]^\alpha = [A_l^\alpha, A_r^\alpha] .$$

Two fuzzy numbers A and B are called equal $A = B$, if $\mu_A(x) = \mu_B(x)$ for all $x \in R$. It follows that

$$A = B \Leftrightarrow [A]^\alpha = [B]^\alpha, \text{ for all } \alpha \in (0,1]$$

A fuzzy number A may be decomposed into its level sets through the resolution identity

$$A = \int_0^1 \alpha [A]^\alpha ,$$

where, $\alpha [A]^\alpha$ is the product of a scalar

α with the set $[A]^{\alpha}$ and $\int$ is the union of $[A]^{\alpha}$'s with α ranging from 0 to 1.

Let, A and B be fuzzy numbers in R and let $\odot$ be a binary operation defined in R. Then the operation $\odot$ can be extended to the fuzzy numbers A and B by definding the relation (the extension principle).

$$A \odot B = \int_{x,y \in R} (\mu_A(x) \wedge \mu_B(y)) / (x \cdot y)$$

Theorem 2.1. ([4]) If A and B are convex fuzzy numbers in the real line R, then $A \oplus B$, $A \ominus B$ and $A \otimes B$ are also convex fuzzy numbers.

Theorem 2.2. ([4]) If A is convex fuzzy number and B is a positive(or negative) convex fuzzy number, then $A \ominus B$ is a convex fuzzy number.

Let the sample data $\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_n$ be convex fuzzy numbers with membership function $\mu_{\tilde{X}_1}, \mu_{\tilde{X}_2}, \cdots, \mu_{\tilde{X}_n}$ respectively, then the fuzzy sample mean denotes $\overline{\tilde{X}}$, defined by

$$\overline{\tilde{X}} = \frac{1}{n} \sum_{i=1}^{n} \tilde{X}_i = \frac{1}{n} (\tilde{X}_1 \oplus \tilde{X}_2 \oplus \cdots \oplus \tilde{X}_n)$$

Let $\tilde{X}_i$ and $\tilde{Y}_i$ for $i \in (0,1]$ be convex fuzzy numbers. Then we define that

$$\frac{1}{n-1} \sum_{i=1}^{n} \{ (\tilde{X}_i \ominus \overline{\tilde{X}}) \otimes (\tilde{Y}_i \ominus \overline{\tilde{Y}}) \} = \tilde{S}_{XY}$$

$$\frac{1}{n-1} \sum_{i=1}^{n} (\tilde{X}_i \ominus \overline{\tilde{X}})^2 = \tilde{S}_{XX},$$

$$\frac{1}{n-1} \sum_{i=1}^{n} (\tilde{Y}_i \ominus \overline{\tilde{Y}})^2 = \tilde{S}_{YY}$$

3. Fuzzy linear regression line

In this section, we study the fuzzy regression line.

First, we consider $\min \tilde{Q}$ where

$$\tilde{Q} = \sum_{i=1}^{n} \tilde{\varepsilon}_i^2 = \sum_{i=1}^{n} (\tilde{Y}_i \ominus \tilde{\beta}_0 \ominus \tilde{\beta}_1 \tilde{X}_i)^2 .$$

Put $[\tilde{Q}]^{\alpha} = [\tilde{Q}_l^{\alpha}, \tilde{Q}_r^{\alpha}]$, $[\tilde{\beta}_0]^{\alpha} = [\tilde{\beta}_{0l}^{\alpha}, \tilde{\beta}_{0r}^{\alpha}]$, $[\tilde{\beta}_1]^{\alpha} = [\tilde{\beta}_{1l}^{\alpha}, \tilde{\beta}_{1r}^{\alpha}]$, $[\tilde{X}_i]^{\alpha} = [\tilde{X}_{il}^{\alpha}, \tilde{X}_{ir}^{\alpha}]$ and $[\tilde{Y}_i]^{\alpha} = [\tilde{Y}_{il}^{\alpha}, \tilde{Y}_{ir}^{\alpha}]$ $(i = 1, 2, \cdots, n)$.

We take

$\forall \tilde{Q}_p^{\alpha} \in [\tilde{Q}_l^{\alpha}, \tilde{Q}_r^{\alpha}]$,

$\forall \tilde{\beta}_{0p}^{\alpha} \in [\tilde{\beta}_{0l}^{\alpha}, \tilde{\beta}_{0r}^{\alpha}]$,

$\forall \tilde{\beta}_{1p}^{\alpha} \in [\tilde{\beta}_{1l}^{\alpha}, \tilde{\beta}_{1r}^{\alpha}]$,

$\forall \tilde{X}_{ip}^{\alpha} \in [\tilde{X}_{il}^{\alpha}, \tilde{X}_{ir}^{\alpha}]$,

$\forall \tilde{Y}_{ip}^{\alpha} \in [\tilde{Y}_{il}^{\alpha}, \tilde{Y}_{ir}^{\alpha}]$, $i = 1, 2, \cdots, n$.

Then

$$\tilde{Q}_p^{\alpha} = \sum_{i=1}^{n} (\tilde{Y}_{ip}^{\alpha} - \tilde{\beta}_{0p}^{\alpha} - \tilde{\beta}_{1p}^{\alpha} \tilde{X}_{ip}^{\alpha})^2 .$$

Therefore for $\alpha \in (0,1]$, we can find out the partial differentiation for $\tilde{\beta}_{0p}^{\alpha}$ and $\tilde{\beta}_{1p}^{\alpha}$ respectively, as follows

$$\frac{\partial \tilde{Q}_p^{\alpha}}{\partial \tilde{\beta}_{0p}^{\alpha}} = -2 \sum_{i=1}^{n} (\tilde{Y}_{ip}^{\alpha} - \tilde{\beta}_{0p}^{\alpha} - \tilde{\beta}_{1p}^{\alpha} \tilde{X}_{ip}^{\alpha})$$

$$\frac{\partial \tilde{Q}_p^{\alpha}}{\partial \tilde{\beta}_{1p}^{\alpha}} = -2 \sum_{i=1}^{n} (\tilde{Y}_{ip}^{\alpha} - \tilde{\beta}_{0p}^{\alpha} - \tilde{\beta}_{1p}^{\alpha} \tilde{X}_{ip}^{\alpha}) \tilde{X}_{ip}^{\alpha} .$$

Let the right side of above equation equal to 0, then

$$\sum_{i=1}^{n} \tilde{Y}_{ip}^{\alpha} - n \tilde{\beta}_{0p}^{\alpha} - \tilde{\beta}_{1p}^{\alpha} \sum_{i=1}^{n} \tilde{X}_{ip}^{\alpha} = 0$$

$$- \sum_{i=1}^{n} \tilde{X}_{ip}^{\alpha} \tilde{Y}_{ip}^{\alpha} + \tilde{\beta}_{0p}^{\alpha} \sum_{i=1}^{n} \tilde{X}_{ip}^{\alpha} + \tilde{\beta}_{0p}^{\alpha} \sum_{i=1}^{n} (\tilde{X}_{ip}^{\alpha})^2 = 0 .$$

Hence,

$$n \tilde{\beta}_{0p}^{\alpha} + \tilde{\beta}_{1p}^{\alpha} \sum_{i=1}^{n} \tilde{X}_{ip}^{\alpha} = \sum_{i=1}^{n} \tilde{Y}_{ip}^{\alpha} \qquad (3.1),$$

$$\sum_{i=1}^{n} \tilde{X}_{ip}^{\alpha}\ \tilde{Y}_{ip}^{\alpha} = \tilde{\beta}_{0p}^{\alpha} \sum_{i=1}^{n} \tilde{X}_{ip}^{\alpha}$$

$$+ \tilde{\beta}_{0p}^{\alpha} \sum_{i=1}^{n} (\tilde{X}_{ip}^{\alpha})^2 \qquad (3.2).$$

$\tilde{\beta}_{0p}^{\alpha}$ and $\tilde{\beta}_{1p}^{\alpha}$ replaced by point estimator $\tilde{b}_{0p}^{\alpha}$ and $\tilde{b}_{1p}^{\alpha}$, respectively then for $\alpha \in (0,1]$,

$$n\ \tilde{b}_{0p}^{\alpha} + \tilde{b}_{1p} \sum_{i=1}^{n} \tilde{X}_{ip}^{\alpha}$$

$$= \sum_{i=1}^{n} \tilde{Y}_{ip}^{\alpha} \qquad (3.3),$$

$$\tilde{b}_{0p}^{\alpha} \sum_{i=1}^{n} \tilde{X}_{ip}^{\alpha} + \tilde{b}_{1p}^{\alpha} \sum_{i=1}^{\alpha} (\tilde{X}_{ip}^{2})^{\alpha}$$

$$= \sum_{i=1}^{n} \tilde{X}_{ip}^{\alpha}\ \tilde{Y}_{ip}^{\alpha} \qquad (3.4).$$

Theorem 3.1. $\tilde{b}_0 = \overline{\tilde{Y}} \ominus (\tilde{b} \otimes \tilde{X})$,

$$\tilde{b}_1 = \frac{\tilde{S}_{XY}}{\tilde{S}_{XX}}, \quad \hat{\tilde{Y}}_i = \tilde{b}_0 \oplus (\tilde{b}_1 \otimes \tilde{X}_i),$$

where,

$$\tilde{S}_{XY} = \sum_{i=1}^{n} (\tilde{X}_i \otimes \tilde{Y}_i) \ominus n (\overline{\tilde{X}} \otimes \overline{\tilde{Y}}),$$

$$\tilde{S}_{XX} = \sum_{i=1}^{n} \tilde{X}_i^2 \ominus n\ \overline{\tilde{X}}^2 .$$

Example 3.1. We obtained five numbers observation data like a following for two variables $\tilde{X}$ and $\tilde{Y}$. Find out the fuzzy regression line by Theorem 3.1.

i	1	2	3	4	5
$\tilde{X}$	$\tilde{2}$	$\tilde{3}$	$\tilde{4}$	$\tilde{5}$	$\tilde{6}$
$\tilde{Y}$	$\tilde{4}$	$\tilde{7}$	$\tilde{6}$	$\tilde{8}$	$\tilde{10}$

The membership functions of above table are follows.

$$\mu_{X_1}(x) = \begin{cases} 100x - 199, & 1.99 \le x \le 2 \\ -100x + 201, & 2 \le x \le 2.01 \end{cases},$$

$$\mu_{X_2}(x) = \begin{cases} 100x - 299, & 2.99 \le x \le 3 \\ -100x + 301, & 3 \le x \le 3.01 \end{cases},$$

$$\mu_{X_3}(x) = \begin{cases} 100x - 399, & 3.99 \le x \le 4 \\ -100x + 401, & 4 \le x \le 4.01 \end{cases},$$

$$\mu_{X_4}(x) = \begin{cases} 100x - 499, & 4.99 \le x \le 5 \\ -100x + 501, & 5 \le x \le 5.01 \end{cases},$$

$$\mu_{X_5}(x) = \begin{cases} 100x - 599, & 5.99 \le x \le 6 \\ -100x + 601, & 6 \le x \le 6.01 \end{cases},$$

$$\mu_{Y_1}(x) = \begin{cases} 100x - 399, & 3.99 \le x \le 4 \\ -100x + 401, & 4 \le x \le 4.01 \end{cases},$$

$$\mu_{Y_2}(x) = \begin{cases} 100x - 699, & 6.99 \le x \le 7 \\ -100x + 701, & 7 \le x \le 7.01 \end{cases},$$

$$\mu_{Y_3}(x) = \begin{cases} 100x - 599, & 5.99 \le x \le 6 \\ -100x + 601, & 6 \le x \le 6.01 \end{cases},$$

$$\mu_{Y_4}(x) = \begin{cases} 100x - 799, & 7.99 \le x \le 8 \\ -100x + 801, & 8 \le x \le 8.01 \end{cases},$$

$$\mu_{Y_5}(x) = \begin{cases} 100x - 999, & 9.99 \le x \le 10 \\ -100x + 1001, & 10 \le x \le 10.01 \end{cases}.$$

Theorem 3.2. (1) $\sum_{i=1}^{n} \tilde{e}_i = \tilde{0}$.

(2) $\sum_{i=1}^{n} \tilde{X}_i\ \tilde{e}_i = \tilde{0}$.

(3) $\sum_{i=1}^{n} \hat{\tilde{Y}}_i\ \tilde{e}_i = \tilde{0}$.

(4) $(\overline{\tilde{X}}, \overline{\tilde{Y}})$ lies the estimated fuzzy regression line.

where $\tilde{0}$ is zero fuzzy number.

4. References

[1] B. George and B. Maria, "Fuzzy set, Fuzzy Logic, Application," World Scientific, 1995.

[2] Phil Diamond and Peter Kloeden, "Metric spaces of fuzzy sets," World Scientific, 1994.

[3] A. Kaufmann and M. M.Gupta, "Introduction to fuzzy arithmetic," Van Nostrand Rainhold, 1991.

[4] M. K. Kang, J. R. Choi and E. S. Bae, "Some properties of fuzzy sample correlation coefficient with fuzzy data," Far East J. Theo. Stat., Vol. 4(1), pp.165-177, 2000.

[5] R. Kruse and K. D. Meyer, "Statistic with vague data," D. Reidel Publishing Company, 1987.

[6] M. Mizumoto and K.Tanaka, "Some properties of fuzzy numbers," North Holland Publishing Company, pp.153-164, 1979.