On using Different Distance Measures for Fuzzy Numbers in Fuzzy Linear Regression Models

Duygu İçen [1] Marco E.G.V. Cattaneo[2]

[1]Hacettepe University,
Department of Statistics, 06800, Ankara, Turkey

[2]Ludwig-Maximilians-University,
Department of Statistics, 80539, Munich, Germany

31^{th} March, 2014

Outline

- In many cases in real life, most of data are approximately known.

- In many cases in real life, most of data are approximately known.
- Fuzzy set theory introduced by Zadeh (1965) has been applied to many areas which need to manage uncertain and vague data.

- In many cases in real life, most of data are approximately known.
- Fuzzy set theory introduced by Zadeh (1965) has been applied to many areas which need to manage uncertain and vague data.
- Such areas include approximate reasoning, decision making, time series, control and regression analysis where the difference of two fuzzy numbers plays an important role in the decision process.

- In many cases in real life, most of data are approximately known.
- Fuzzy set theory introduced by Zadeh (1965) has been applied to many areas which need to manage uncertain and vague data.
- Such areas include approximate reasoning, decision making, time series, control and regression analysis where the difference of two fuzzy numbers plays an important role in the decision process.

- A fuzzy number is a quantity whose value is imprecise and it depicts the physical world more realistically than single-valued numbers (Gao et al., 2009).

- A fuzzy number is a quantity whose value is imprecise and it depicts the physical world more realistically than single-valued numbers (Gao et al., 2009).
- Many research articles have been published in order to define a distance between fuzzy numbers. Several distance measures for fuzzy numbers are well established in the literature.

- A fuzzy number is a quantity whose value is imprecise and it depicts the physical world more realistically than single-valued numbers (Gao et al., 2009).
- Many research articles have been published in order to define a distance between fuzzy numbers. Several distance measures for fuzzy numbers are well established in the literature.
- In this study, the distance measures for fuzzy numbers by

- A fuzzy number is a quantity whose value is imprecise and it depicts the physical world more realistically than single-valued numbers (Gao et al., 2009).
- Many research articles have been published in order to define a distance between fuzzy numbers. Several distance measures for fuzzy numbers are well established in the literature.
- In this study, the distance measures for fuzzy numbers by
  - Kaufmann and Gupta (1991)

- A fuzzy number is a quantity whose value is imprecise and it depicts the physical world more realistically than single-valued numbers (Gao et al., 2009).
- Many research articles have been published in order to define a distance between fuzzy numbers. Several distance measures for fuzzy numbers are well established in the literature.
- In this study, the distance measures for fuzzy numbers by
    - Kaufmann and Gupta (1991)
    - Heilpern (1997)

- A fuzzy number is a quantity whose value is imprecise and it depicts the physical world more realistically than single-valued numbers (Gao et al., 2009).
- Many research articles have been published in order to define a distance between fuzzy numbers. Several distance measures for fuzzy numbers are well established in the literature.
- In this study, the distance measures for fuzzy numbers by
    - Kaufmann and Gupta (1991)
    - Heilpern (1997)
    - Chen and Hsieh (1998)

- A fuzzy number is a quantity whose value is imprecise and it depicts the physical world more realistically than single-valued numbers (Gao et al., 2009).
- Many research articles have been published in order to define a distance between fuzzy numbers. Several distance measures for fuzzy numbers are well established in the literature.
- In this study, the distance measures for fuzzy numbers by
    - Kaufmann and Gupta (1991)
    - Heilpern (1997)
    - Chen and Hsieh (1998)

- In the Monte Carlo method, several random crisp or fuzzy vectors are generated as regression coefficient vector. Then using these random vectors, the dependent variable is calculated.

- In the Monte Carlo method, several random crisp or fuzzy vectors are generated as regression coefficient vector. Then using these random vectors, the dependent variable is calculated.
- Two error measures are obtained by the difference of observed and estimated values of dependent variable to decide the best random vector for parameter estimation.

- In the Monte Carlo method, several random crisp or fuzzy vectors are generated as regression coefficient vector. Then using these random vectors, the dependent variable is calculated.
- Two error measures are obtained by the difference of observed and estimated values of dependent variable to decide the best random vector for parameter estimation.
    - One of these error measures depends on the error measure defined by Kim and Bishu (1998).

- In the Monte Carlo method, several random crisp or fuzzy vectors are generated as regression coefficient vector. Then using these random vectors, the dependent variable is calculated.
- Two error measures are obtained by the difference of observed and estimated values of dependent variable to decide the best random vector for parameter estimation.
    - One of these error measures depends on the error measure defined by Kim and Bishu (1998).
      In this error measure, distance of two fuzzy numbers has to be calculated.

- In the Monte Carlo method, several random crisp or fuzzy vectors are generated as regression coefficient vector. Then using these random vectors, the dependent variable is calculated.
- Two error measures are obtained by the difference of observed and estimated values of dependent variable to decide the best random vector for parameter estimation.
  - One of these error measures depends on the error measure defined by Kim and Bishu (1998).

    In this error measure, distance of two fuzzy numbers has to be calculated.

    Therefore, distance measure between two fuzzy numbers plays an important role in fuzzy regression with Monte Carlo method.

- In the Monte Carlo method, several random crisp or fuzzy vectors are generated as regression coefficient vector. Then using these random vectors, the dependent variable is calculated.
- Two error measures are obtained by the difference of observed and estimated values of dependent variable to decide the best random vector for parameter estimation.
  - One of these error measures depends on the error measure defined by Kim and Bishu (1998).

    In this error measure, distance of two fuzzy numbers has to be calculated.

    Therefore, distance measure between two fuzzy numbers plays an important role in fuzzy regression with Monte Carlo method.

Aim of the study

Aim of the study

- Highlight the utility of distance measures

Aim of the study

- Highlight the utility of distance measures
- Calculate different distance measures in fuzzy linear regression with Monte Carlo method.

Aim of the study

- Highlight the utility of distance measures
- Calculate different distance measures in fuzzy linear regression with Monte Carlo method.
- Estimate the parameters of fuzzy linear regression with Monte Carlo method according to the different distance measures

Aim of the study

- Highlight the utility of distance measures
- Calculate different distance measures in fuzzy linear regression with Monte Carlo method.
- Estimate the parameters of fuzzy linear regression with Monte Carlo method according to the different distance measures

Outline

Definition 2.1. $\mu_A(x)$ is the membership function of an element x belonging to a fuzzy set $\tilde{A}$, where $0 \leq \mu_A(x) \leq 1$.

Definition 2.1. $\mu_A(x)$ is the membership function of an element x belonging to a fuzzy set $\tilde{A}$, where $0 \leq \mu_A(x) \leq 1$.

Definition 2.2. A general fuzzy number $\tilde{A}$ is a normal convex fuzzy set of $\Re$ with a piecewise continuous membership function. The left and right sides of fuzzy numbers are $L(x) = \frac{a_2 - x}{a_2 - a_1}$ and $R(x) = \frac{x - a_3}{a_4 - a_3}$ respectively.

Definition 2.1. $\mu_A(x)$ is the membership function of an element x belonging to a fuzzy set $\tilde{A}$, where $0 \leq \mu_A(x) \leq 1$.

Definition 2.2. A general fuzzy number $\tilde{A}$ is a normal convex fuzzy set of $\Re$ with a piecewise continuous membership function. The left and right sides of fuzzy numbers are $L(x) = \frac{a_2 - x}{a_2 - a_1}$ and $R(x) = \frac{x - a_3}{a_4 - a_3}$ respectively.

Definition 2.3. The α-cut of a fuzzy number $\tilde{A}$ is a non-fuzzy set defined as $\tilde{A}(\alpha) = \{x \in \Re, \mu_A(\alpha) \geq \alpha\}$.

Definition 2.1. $\mu_A(x)$ is the membership function of an element x belonging to a fuzzy set $\tilde{A}$, where $0 \leq \mu_A(x) \leq 1$.

Definition 2.2. A general fuzzy number $\tilde{A}$ is a normal convex fuzzy set of $\Re$ with a piecewise continuous membership function. The left and right sides of fuzzy numbers are $L(x) = \frac{a_2 - x}{a_2 - a_1}$ and $R(x) = \frac{x - a_3}{a_4 - a_3}$ respectively.

Definition 2.3. The α-cut of a fuzzy number $\tilde{A}$ is a non-fuzzy set defined as $\tilde{A}(\alpha) = \{x \in \Re, \mu_A(\alpha) \geq \alpha\}$.
$\{\tilde{A}(\alpha) = [A^L(\alpha), A^U(\alpha)]\}$

Definition 2.1. $\mu_A(x)$ is the membership function of an element x belonging to a fuzzy set $\tilde{A}$, where $0 \leq \mu_A(x) \leq 1$.

Definition 2.2. A general fuzzy number $\tilde{A}$ is a normal convex fuzzy set of $\Re$ with a piecewise continuous membership function. The left and right sides of fuzzy numbers are $L(x) = \frac{a_2 - x}{a_2 - a_1}$ and $R(x) = \frac{x - a_3}{a_4 - a_3}$ respectively.

Definition 2.3. The α-cut of a fuzzy number $\tilde{A}$ is a non-fuzzy set defined as $\tilde{A}(\alpha) = \{x \in \Re, \mu_A(\alpha) \geq \alpha\}$.
$\{\tilde{A}(\alpha) = [A^L(\alpha), A^U(\alpha)]\}$

Definition 2.4. $v_k = (v_{0k}, ..., v_{mk})$ is called random crisp vector.

Definition 2.4. $v_k = (v_{0k}, ..., v_{mk})$ is called random crisp vector.
v_{ik} are all real numbers in intervals I_i , $i = 0, 1, ..., m$.

Definition 2.4. $v_k = (v_{0k}, ..., v_{mk})$ is called random crisp vector.
v_{ik} are all real numbers in intervals I_i , $i = 0, 1, ..., m$.
Firstly, random crisp vectors
$v_k = (x_{ok}, ..., x_{mk})$ with all $x_{ik} \in [0, 1]$ are generated.

Definition 2.4. $v_k = (v_{0k}, ..., v_{mk})$ is called random crisp vector.
v_{ik} are all real numbers in intervals I_i , $i = 0, 1, ..., m$.
Firstly, random crisp vectors
$v_k = (x_{ok}, ..., x_{mk})$ with all $x_{ik} \in [0, 1]$ are generated.
Then all x_{ik} are put in the interval
$I_i = [c_i, d_i]$ by $v_{ik} = c_i + (d_i - c_i)x_{ik}, i = 0, 1, ..., m$.

Definition 2.4. $v_k = (v_{0k}, ..., v_{mk})$ is called random crisp vector.
v_{ik} are all real numbers in intervals I_i , $i = 0, 1, ..., m$.
Firstly, random crisp vectors
$v_k = (x_{ok}, ..., x_{mk})$ with all $x_{ik} \in [0, 1]$ are generated.
Then all x_{ik} are put in the interval
$I_i = [c_i, d_i]$ by $v_{ik} = c_i + (d_i - c_i)x_{ik}, i = 0, 1, ..., m$.

Definition 2.5. $\widetilde{V}_k = (\widetilde{V}_{0k}, ..., \widetilde{V}_{mk})$ is called random fuzzy vector

Definition 2.4. $v_k = (v_{0k}, ..., v_{mk})$ is called random crisp vector.
v_{ik} are all real numbers in intervals I_i , $i = 0, 1, ..., m$.
Firstly, random crisp vectors
$v_k = (x_{ok}, ..., x_{mk})$ with all $x_{ik} \in [0, 1]$ are generated.
Then all x_{ik} are put in the interval
$I_i = [c_i, d_i]$ by $v_{ik} = c_i + (d_i - c_i)x_{ik}, i = 0, 1, ..., m$.

Definition 2.5. $\widetilde{V}_k = (\widetilde{V}_{0k}, ..., \widetilde{V}_{mk})$ is called random fuzzy vector
$\widetilde{V}_{ik}$ are all triangular fuzzy numbers.

Definition 2.4. $v_k = (v_{0k}, ..., v_{mk})$ is called random crisp vector.
v_{ik} are all real numbers in intervals I_i , $i = 0, 1, ..., m$.
Firstly, random crisp vectors
$v_k = (x_{ok}, ..., x_{mk})$ with all $x_{ik} \in [0, 1]$ are generated.
Then all x_{ik} are put in the interval
$I_i = [c_i, d_i]$ by $v_{ik} = c_i + (d_i - c_i)x_{ik}, i = 0, 1, ..., m$.

Definition 2.5. $\widetilde{V}_k = (\widetilde{V}_{0k}, ..., \widetilde{V}_{mk})$ is called random fuzzy vector
$\widetilde{V}_{ik}$ are all triangular fuzzy numbers.
First crisp vectors $v_k = (v_{1k}, \ldots, v_{(3m+3,k)})$ with all the x_{ik} in $[0, 1]$, $k = 1, ..., N$ are generated.

Definition 2.4. $v_k = (v_{0k}, ..., v_{mk})$ is called random crisp vector.
v_{ik} are all real numbers in intervals I_i , $i = 0, 1, ..., m$.
Firstly, random crisp vectors
$v_k = (x_{ok}, ..., x_{mk})$ with all $x_{ik} \in [0, 1]$ are generated.
Then all x_{ik} are put in the interval
$I_i = [c_i, d_i]$ by $v_{ik} = c_i + (d_i - c_i)x_{ik}, i = 0, 1, ..., m$.

Definition 2.5. $\widetilde{V}_k = (\widetilde{V}_{0k}, ..., \widetilde{V}_{mk})$ is called random fuzzy vector
$\widetilde{V}_{ik}$ are all triangular fuzzy numbers.
First crisp vectors $v_k = (v_{1k}, \ldots, v_{(3m+3,k)})$ with all the x_{ik} in $[0, 1]$, $k = 1, ..., N$ are generated.
Then the first three numbers in v_k are chosen and ordered from smallest to largest.

Definition 2.4. $v_k = (v_{0k}, ..., v_{mk})$ is called random crisp vector.
v_{ik} are all real numbers in intervals I_i , $i = 0, 1, ..., m$.
Firstly, random crisp vectors
$v_k = (x_{ok}, ..., x_{mk})$ with all $x_{ik} \in [0, 1]$ are generated.
Then all x_{ik} are put in the interval
$I_i = [c_i, d_i]$ by $v_{ik} = c_i + (d_i - c_i)x_{ik}, i = 0, 1, ..., m$.

Definition 2.5. $\widetilde{V}_k = (\widetilde{V}_{0k}, ..., \widetilde{V}_{mk})$ is called random fuzzy vector
$\widetilde{V}_{ik}$ are all triangular fuzzy numbers.
First crisp vectors $v_k = (v_{1k}, \ldots, v_{(3m+3,k)})$ with all the x_{ik} in $[0, 1]$, $k = 1, ..., N$ are generated.
Then the first three numbers in v_k are chosen and ordered from smallest to largest.
Let us assume that $x_{3k} < x_{1k} < x_{2k}$,

Definition 2.4. $v_k = (v_{0k}, ..., v_{mk})$ is called random crisp vector.
v_{ik} are all real numbers in intervals I_i , $i = 0, 1, ..., m$.
Firstly, random crisp vectors
$v_k = (x_{ok}, ..., x_{mk})$ with all $x_{ik} \in [0, 1]$ are generated.
Then all x_{ik} are put in the interval
$I_i = [c_i, d_i]$ by $v_{ik} = c_i + (d_i - c_i)x_{ik}, i = 0, 1, ..., m$.

Definition 2.5. $\widetilde{V}_k = (\widetilde{V}_{0k}, ..., \widetilde{V}_{mk})$ is called random fuzzy vector
$\widetilde{V}_{ik}$ are all triangular fuzzy numbers.
First crisp vectors $v_k = (v_{1k}, \ldots, v_{(3m+3,k)})$ with all the x_{ik} in $[0, 1]$, $k = 1, ..., N$ are generated.
Then the first three numbers in v_k are chosen and ordered from smallest to largest.
Let us assume that $x_{3k} < x_{1k} < x_{2k}$,
then the first triangular fuzzy numbers is $\widetilde{V}_{0k} = (x_{3k}/x_{1k}/x_{2k})$.

Outline

Choi and Buckley (2008) classified fuzzy regression models in three categories:

- Input and output data are both crisp (First Category)

Choi and Buckley (2008) classified fuzzy regression models in three categories:

- Input and output data are both crisp (First Category)
- Input data is crisp and output data is fuzzy (Second Category)

Choi and Buckley (2008) classified fuzzy regression models in three categories:

- Input and output data are both crisp (First Category)
- Input data is crisp and output data is fuzzy (Second Category)
- Input and output data are both fuzzy (Third Category)

Fuzzy linear regression model (Second Category)

$$\widetilde{Y}_l = \widetilde{A}_0 + \widetilde{A}_1 x_{1l} + \widetilde{A}_2 x_{2l} + \ldots + \widetilde{A}_m x_{ml} \quad l = 1, 2, .., n \tag{1}$$

Fuzzy linear regression model (Third Category)

$$\widetilde{Y}_l = a_0 + a_1 \widetilde{X}_{1l} + a_2 \widetilde{X}_{2l} + \ldots + a_m \widetilde{X}_{ml} \quad l = 1, 2, .., n \tag{2}$$

Predicted values

Predicted values

Fuzzy linear regression model (Second Category)

$$\widetilde{Y}^*_{lk} = \widetilde{V}_{0k} + \widetilde{V}_{1k} x_{1l} + \widetilde{V}_{2k} x_{2l} + ... + \widetilde{V}_{mk} x_{ml} \quad l = 1, 2, .., n \tag{3}$$

Predicted values

Fuzzy linear regression model (Second Category)

$$\widetilde{Y}_{lk}^{*} = \widetilde{V}_{0k} + \widetilde{V}_{1k}x_{1l} + \widetilde{V}_{2k}x_{2l} + ... + \widetilde{V}_{mk}x_{ml} \quad l = 1, 2, .., n \tag{3}$$

Fuzzy linear regression model (Third Category)

$$\widetilde{Y}_{lk}^{*} = v_{0k} + v_{1k}\,\widetilde{X}_{1l} + v_{2k}\widetilde{X}_{2l} + ... + v_{mk}\widetilde{X}_{ml}; \quad l = 1, 2, .., n \tag{4}$$

- Since the dependent variable has a membership function, the estimated fuzzy output, which is also represented by a membership function, should be close to the membership function of the given data.

- Since the dependent variable has a membership function, the estimated fuzzy output, which is also represented by a membership function, should be close to the membership function of the given data.
- The sum of the differences is calculated as

- Since the dependent variable has a membership function, the estimated fuzzy output, which is also represented by a membership function, should be close to the membership function of the given data.
- The sum of the differences is calculated as

$$D = \int |\mu_{\tilde{Y}}(x) - \mu_{\tilde{Y}^*_{lk}}(x)| dx$$

- Since the dependent variable has a membership function, the estimated fuzzy output, which is also represented by a membership function, should be close to the membership function of the given data.
- The sum of the differences is calculated as

$$D = \int |\mu_{\tilde{Y}}(x) - \mu_{\tilde{Y}^*_{lk}}(x)| dx$$

$$E = \frac{\int_{S_{\widetilde{Y}} \cup S_{\widetilde{Y}^*_{lk}}} |\mu_{\widetilde{Y}(x)} - \mu_{\widetilde{Y}^*_{lk}(x)}| dx}{\int_{S_{\widetilde{Y}}} \mu_{\widetilde{Y}}(x) dx}$$

Error Measure (Abdalla & Buckley (2007))

$$E_1 = \frac{\sum_{l=1}^{n} \left[\int_{-\infty}^{\infty} |\widetilde{Y}_l(x) - \widetilde{Y}^{*}_{lk}(x)| dx \right]}{\left[\int_{-\infty}^{\infty} \widetilde{Y}_l(x) dx \right]} \tag{5}$$

Error Measure (Abdalla & Buckley (2007))

$$E_1 = \frac{\sum_{l=1}^{n}\left[\int_{-\infty}^{\infty}|\widetilde{Y}_l(x) - \widetilde{Y}^*_{lk}(x)|dx\right]}{\left[\int_{-\infty}^{\infty}\widetilde{Y}_l(x)dx\right]} \tag{5}$$

- $\widetilde{Y}_l = (y_{l1}/y_{l2}/y_{y3})$ and $\widetilde{Y}^*_{lk} = (y_{lk1}/y_{lk2}/y_{lk3})$

Error Measure (Abdalla & Buckley (2007))

$$E_1 = \frac{\sum_{l=1}^{n} \left[\int_{-\infty}^{\infty} |\widetilde{Y}_l(x) - \widetilde{Y}^*_{lk}(x)| dx \right]}{\left[\int_{-\infty}^{\infty} \widetilde{Y}_l(x) dx \right]} \tag{5}$$

- $\widetilde{Y}_l = (y_{l1}/y_{l2}/y_{y3})$ and $\widetilde{Y}^*_{lk} = (y_{lk1}/y_{lk2}/y_{lk3})$

$$\widetilde{V}_k \in \{\widetilde{V}_1, ..., \widetilde{V}_N\} \text{ and } v_k \in \{v_1, ..., v_N\}$$

Outline

The methods of measuring the distance between fuzzy numbers have become important due to the significant applications in diverse fields like data mining, pattern recognition, multivariate data analysis and so on.

The methods of measuring the distance between fuzzy numbers have become important due to the significant applications in diverse fields like data mining, pattern recognition, multivariate data analysis and so on.

- Kaufmann (1991)

The methods of measuring the distance between fuzzy numbers have become important due to the significant applications in diverse fields like data mining, pattern recognition, multivariate data analysis and so on.

- Kaufmann (1991)
- Heilpern (1997)

The methods of measuring the distance between fuzzy numbers have become important due to the significant applications in diverse fields like data mining, pattern recognition, multivariate data analysis and so on.

- Kaufmann (1991)
- Heilpern (1997)
  - Heilpern-1 (1997)

The methods of measuring the distance between fuzzy numbers have become important due to the significant applications in diverse fields like data mining, pattern recognition, multivariate data analysis and so on.

- Kaufmann (1991)
- Heilpern (1997)
    - Heilpern-1 (1997)
    - Heilpern-2 (1997)

The methods of measuring the distance between fuzzy numbers have become important due to the significant applications in diverse fields like data mining, pattern recognition, multivariate data analysis and so on.

- Kaufmann (1991)
- Heilpern (1997)
  - Heilpern-1 (1997)
  - Heilpern-2 (1997)
  - Heilpern-3 (1997)

The methods of measuring the distance between fuzzy numbers have become important due to the significant applications in diverse fields like data mining, pattern recognition, multivariate data analysis and so on.

- Kaufmann (1991)
- Heilpern (1997)
  - Heilpern-1 (1997)
  - Heilpern-2 (1997)
  - Heilpern-3 (1997)
- Chen & Hsieh (1998)

The methods of measuring the distance between fuzzy numbers have become important due to the significant applications in diverse fields like data mining, pattern recognition, multivariate data analysis and so on.

- Kaufmann (1991)
- Heilpern (1997)
  - Heilpern-1 (1997)
  - Heilpern-2 (1997)
  - Heilpern-3 (1997)
- Chen & Hsieh (1998)

Kaufmann (1991)

$$d(\widetilde{A}, \widetilde{B}) = \int_0^1 \left(|A^L(\alpha) - B^L(\alpha)| + |A^U(\alpha) - B^U(\alpha)|\right) d\alpha$$

Kaufmann (1991)

$$d(\widetilde{A}, \widetilde{B}) = \int_0^1 \left(|A^L(\alpha) - B^L(\alpha)| + |A^U(\alpha) - B^U(\alpha)|\right) d\alpha$$

- $\left[A^L(\alpha), A^U(\alpha)\right]$ and $\left[B^L(\alpha), B^U(\alpha)\right]$ are the closed intervals of α-cuts

Heilpern-1 (1997)

Heilpern-1 (1997)

$$\widetilde{A} = (a_1, a_2, a_3, a_4)$$

Heilpern-1 (1997)

$$\widetilde{A} = (a_1, a_2, a_3, a_4)$$

- $E_*(\widetilde{A}) = a_2 - (a_2 - a_1) \int_0^\infty L(x)dx$
- $E^*(\widetilde{A}) = a_3 + (a_4 - a_3) \int_0^\infty R(x)dx$

Heilpern-1 (1997)

$$\widetilde{A} = (a_1, a_2, a_3, a_4)$$

- $E_*(\widetilde{A}) = a_2 - (a_2 - a_1) \int_0^\infty L(x)dx$
- $E^*(\widetilde{A}) = a_3 + (a_4 - a_3) \int_0^\infty R(x)dx$

$$EV(\widetilde{A}) = \frac{1}{2}\left[E_*(\widetilde{A}) - E^*(\widetilde{A})\right]$$

Heilpern-1 (1997)

$$\widetilde{A} = (a_1, a_2, a_3, a_4)$$

- $E_*(\widetilde{A}) = a_2 - (a_2 - a_1) \int_0^\infty L(x)dx$
- $E^*(\widetilde{A}) = a_3 + (a_4 - a_3) \int_0^\infty R(x)dx$

$$EV(\widetilde{A}) = \frac{1}{2}\left[E_*(\widetilde{A}) - E^*(\widetilde{A})\right]$$

$$\sigma(\widetilde{A}, \widetilde{B}) = |EV(\widetilde{A}) - EV(\widetilde{B})| \tag{6}$$

Heilpern-2 (1997)

Heilpern-2 (1997)

$$d_p(\widetilde{A}, \widetilde{B}) = \int_0^1 d_p(\widetilde{A}(\alpha), \widetilde{B}(\alpha)d\alpha) \tag{7}$$

Heilpern-2 (1997)

$$d_p(\widetilde{A}, \widetilde{B}) = \int_0^1 d_p(\widetilde{A}(\alpha), \widetilde{B}(\alpha) d\alpha) \tag{7}$$

$\widetilde{A}(\alpha) = [A^L(\alpha), A^U(\alpha)]$ and $\widetilde{B}(\alpha) = [B^L(\alpha), B^U(\alpha)]$

Heilpern-2 (1997)

$$d_p(\widetilde{A}, \widetilde{B}) = \int_0^1 d_p(\widetilde{A}(\alpha), \widetilde{B}(\alpha)d\alpha) \tag{7}$$

$\widetilde{A}(\alpha) = [A^L(\alpha), A^U(\alpha)]$ and $\widetilde{B}(\alpha) = [B^L(\alpha), B^U(\alpha)]$

$$d_p\left(\widetilde{A}(\alpha), \widetilde{B}(\alpha)\right) =$$

$$\begin{cases} (0.5)(|A^L(\alpha) - B^L(\alpha)|^p + |A^U(\alpha) - B^U(\alpha)|^p)^{1/p}, & 1 \leq p \leq \infty; \\ max|A^L(\alpha) - B^L(\alpha)|, |A^U(\alpha) - B^U(\alpha)|, & p = \infty. \end{cases} \tag{8}$$

Heilpern-3 (1997)

- $\widetilde{A} = (a_1, a_2, a_3, a_4)$
- $\widetilde{B} = (b_1, b_2, b_3, b_4)$

Heilpern-3 (1997)

- $\widetilde{A} = (a_1, a_2, a_3, a_4)$
- $\widetilde{B} = (b_1, b_2, b_3, b_4)$

$$\delta_p(\widetilde{A}, \widetilde{B}) = \begin{cases} 0.25 \left(\sum_{i=1}^{4} |a_i - b_i|^p \right)^{1/p}, & 1 \leq p < \infty; \\ max(|a_i - b_i|), & p = \infty. \end{cases} \tag{9}$$

Chen & Hsieh (1998)

- $$P(A) = \frac{\int_0^w \alpha \left(\frac{L^{-1}(\alpha) + R^{-1}(\alpha)}{2} \right) d\alpha}{\int_0^w \alpha d\alpha}$$

Chen & Hsieh (1998)

- $P(A) = \dfrac{\int_0^w \alpha \left(\frac{L^{-1}(\alpha)+R^{-1}(\alpha)}{2} \right) d\alpha}{\int_0^w \alpha d\alpha}$

$$\widetilde{A} = (a_1, a_2, a_3, a_4)$$

Chen & Hsieh (1998)

- $P(A) = \dfrac{\int_0^w \alpha \left(\frac{L^{-1}(\alpha)+R^{-1}(\alpha)}{2} \right) d\alpha}{\int_0^w \alpha d\alpha}$

$$\widetilde{A} = (a_1, a_2, a_3, a_4)$$

$$P(A) = \frac{a_1+2a_2+2a_3+a_4}{6}$$

Chen & Hsieh (1998)

- $$P(A) = \frac{\int_0^w \alpha \left(\frac{L^{-1}(\alpha) + R^{-1}(\alpha)}{2} \right) d\alpha}{\int_0^w \alpha d\alpha}$$

$$\widetilde{A} = (a_1, a_2, a_3, a_4)$$

$$P(A) = \frac{a_1 + 2a_2 + 2a_3 + a_4}{6}$$

$$P(A) = \frac{a_1 + 4a_2 + a_4}{6}$$

Chen & Hsieh (1998)

- $P(A) = \dfrac{\int_0^w \alpha \left(\frac{L^{-1}(\alpha) + R^{-1}(\alpha)}{2} \right) d\alpha}{\int_0^w \alpha d\alpha}$

$$\widetilde{A} = (a_1, a_2, a_3, a_4)$$

$$P(A) = \frac{a_1 + 2a_2 + 2a_3 + a_4}{6}$$

$$P(A) = \frac{a_1 + 4a_2 + a_4}{6}$$

Chen & Hsieh (1998)

- $P(A) = \dfrac{\int_0^w \alpha\left(\frac{L^{-1}(\alpha)+R^{-1}(\alpha)}{2}\right)d\alpha}{\int_0^w \alpha d\alpha}$

$$\widetilde{A} = (a_1, a_2, a_3, a_4)$$

$$P(A) = \frac{a_1+2a_2+2a_3+a_4}{6}$$

$$P(A) = \frac{a_1+4a_2+a_4}{6}$$

$$|P(A) - P(B)| \tag{10}$$

Outline

- In this section, there are two different applications.

- In this section, there are two different applications.
- First application is for the second fuzzy regression model category and the other one is for the third fuzzy regression model category.

- In this section, there are two different applications.
- First application is for the second fuzzy regression model category and the other one is for the third fuzzy regression model category.
- We consider different distance measures for fuzzy numbers given in Section 4 in the error measure (E_1) for fuzzy linear regression models with Monte Carlo approach.

- In this section, there are two different applications.
- First application is for the second fuzzy regression model category and the other one is for the third fuzzy regression model category.
- We consider different distance measures for fuzzy numbers given in Section 4 in the error measure (E_1) for fuzzy linear regression models with Monte Carlo approach.

Table: Data for the application (Second category)

Fuzzy Output	x_{1l}	x_{2l}	x_{3l}
(2.27/5.83/9.39)	2.00	0.00	15.25
(0.33/0.85/1.37)	0.00	5.00	14.13
(5.43/13.93/22.43)	1.13	1.50	14.13
(1.56/4.00/6.44)	2.00	1.25	13.63
(0.64/1.65/2.66)	2.19	3.75	14.75
(0.62/1.58/2.54)	0.25	3.50	13.75
(3.19/8.18/13.17)	0.75	5.25	15.25
(0.72/1.85/2.98)	4.25	2.00	13.50

- Before the application we have to decide the intervals for $I_i, i = 0, 1, 2, 3$ to obtain the model coefficients as explained in Definition 2.5.

- Before the application we have to decide the intervals for $I_i, i = 0, 1, 2, 3$ to obtain the model coefficients as explained in Definition 2.5.
- We use same intervals in order to compare the results we have with the results from Abdalla and Buckley (2007) in the literature.

- Before the application we have to decide the intervals for $I_i, i = 0, 1, 2, 3$ to obtain the model coefficients as explained in Definition 2.5.
- We use same intervals in order to compare the results we have with the results from Abdalla and Buckley (2007) in the literature.
- Four separate intervals $(MCI, MCII, MCIII, MCIV)$ that they studied are given with Table 2.

- Before the application we have to decide the intervals for $I_i, i = 0, 1, 2, 3$ to obtain the model coefficients as explained in Definition 2.5.
- We use same intervals in order to compare the results we have with the results from Abdalla and Buckley (2007) in the literature.
- Four separate intervals $(MCI, MCII, MCIII, MCIV)$ that they studied are given with Table 2.

Table: Intervals for $I_i, i = 0, 1, 2, 3$ for second category

Interval	*MCI*	*MCII*	*MCIII*	*MCIV*
I_0	[-1,0]	[0,1]	[-18.174,-18.174]	[28.000,47.916]
I_1	[-1,0]	[-1,0]	[-1.083,-1.083]	[-2.542,-2.542]
I_2	[-1.5,-0.5]	[-1.5,-0.5]	[-1.150,-1.150]	[-2.323,-2.323]
I_3	[0,1]	[0,1]	[1.733,2.149]	[-1.354,-1.354]

Results for using different definitions of distance measures in fuzzy linear regression with MC method for minimizing E_1

Definitions	Parameters	Intervals											
		MCI			MCII			MCIII			MCIV		
Kaufmann *(1991)*	$\tilde{A}_0$	-0.8530	-0.5900	-0.2935	0.0607	0.3163	0.3414	-18.1740	-18.1740	-18.1740	31.0713	31.5636	32.1763
	$\tilde{A}_1$	-0.6934	-0.6033	-0.3096	-0.2712	-0.2684	-0.1293	-1.0830	-1.0830	-1.0830	-2.5420	-2.5420	-2.5420
	$\tilde{A}_2$	-1.4064	-1.3966	-1.3162	-0.8220	-0.7265	-0.7210	-1.1500	-1.1500	-1.1500	-2.3230	-2.3230	-2.3230
	$\tilde{A}_3$	0.5474	0.5727	0.5923	0.2591	0.2938	0.3359	1.7337	1.7519	1.8307	-1.3540	-1.3540	-1.3540
Heilpern-1 (1997)	$\tilde{A}_0$	-0.8472	-0.7690	-0.1782	0.0653	0.3254	0.3424	-18.1740	-18.1740	-18.1740	28.6932	30.4576	35.6408
	$\tilde{A}_1$	-0.8527	-0.3606	-0.0810	-0.8627	-0.4147	-0.0858	-1.0830	-1.0830	-1.0830	-2.5420	-2.5420	-2.5420
	$\tilde{A}_2$	-1.4198	-1.1616	-0.5778	-1.4075	-1.2370	-0.6181	-1.1500	-1.1500	-1.1500	-2.3230	-2.3230	-2.3230
	$\tilde{A}_3$	0.0251	0.6431	0.7575	0.1463	0.4066	0.7275	1.7339	1.7583	1.7678	-1.3540	-1.3540	-1.3540
Heilpern-2 (1997)	$\tilde{A}_0$	-0.8530	-0.5900	-0.2935	0.0607	0.3163	0.3414	-18.1740	-18.1740	-18.1740	31.0713	31.5636	32.1763
	$\tilde{A}_1$	-0.6934	-0.6033	-0.3096	-0.2712	-0.2684	-0.1293	-1.0830	-1.0830	-1.0830	-2.5420	-2.5420	-2.5420
	$\tilde{A}_2$	-1.4064	-1.3966	-1.3162	-0.8220	-0.7265	-0.7210	-1.1500	-1.1500	-1.1500	-2.3230	-2.3230	-2.3230
	$\tilde{A}_3$	0.5474	0.5727	0.5923	0.2591	0.2938	0.3359	1.7337	1.7519	1.8307	-1.3540	-1.3540	-1.3540
Heilpern-3 (1997)	$\tilde{A}_0$	-0.8530	-0.5900	-0.2935	0.0607	0.3163	0.3414	-18.1740	-18.1740	-18.1740	31.0713	31.5636	32.1763
	$\tilde{A}_1$	-0.6934	-0.6033	-0.3096	-0.2712	-0.2684	-0.1293	-1.0830	-1.0830	-1.0830	-2.5420	-2.5420	-2.5420
	$\tilde{A}_2$	-1.4064	-1.3966	-1.3162	-0.8220	-0.7265	-0.7210	-1.1500	-1.1500	-1.1500	-2.3230	-2.3230	-2.3230
	$\tilde{A}_3$	0.5474	0.5727	0.5923	0.2591	0.2938	0.3359	1.7334	1.7552	1.8369	-1.3540	-1.3540	-1.3540
Chen and Hsieh (1998)	$\tilde{A}_0$	-0.7617	-0.7454	-0.5821	0.0716	0.4464	0.5536	-18.1740	-18.1740	-18.1740	28.9831	31.8476	33.2103
	$\tilde{A}_1$	-0.6857	-0.4063	-0.3824	-0.9107	-0.4521	-0.0816	-1.0830	-1.0830	-1.0830	-2.5420	-2.5420	-2.5420
	$\tilde{A}_2$	-1.3294	-1.1576	-0.5469	-1.3458	-1.1448	-0.6135	-1.1500	-1.1500	-1.1500	-2.3230	-2.3230	-2.3230
	$\tilde{A}_3$	0.2521	0.4794	0.8036	0.2596	0.3323	0.9166	1.7443	1.7445	1.7981	-1.3540	-1.3540	-1.3540

Table: Data for the application (Third category)

Fuzzy output	X_{1l}	X_{2l}
(55.4/61.6/64.7)	(5.7/6.0/6.9)	(5.4/6.3/7.1)
(50.5/53.2/58.5)	(4.0/4.4/5.1)	(4.7/5.5/5.8)
(55.7/65.5/75.3)	(8.6/9.1/9.8)	(3.4/3.6/4.0)
(61.7/64.9/74.7)	(6.9/8.1/9.3)	(5.0/5.8/6.7)
(69.1/71.7/80.0)	(8.7/9.4/11.2)	(6.5/6.8/7.1)
(49.6/52.2/57.4)	(4.6/4.8/5.5)	(6.7/7.9/8.7)
(47.7/50.2/55.2)	(7.2/7.6/8.7)	(4.0/4.2/4.8)
(41.8/44.0/48.4)	(4.2/4.4/4.8)	(5.4/6.0/6.3)
(45.7/53.8/61.9)	(8.2/9.1/10.0)	(2.7/2.8/3.2)
(45.4/53.5/58.9)	(6.0/6.7/7.4)	(5.7/6.7/7.7)

- Before the application we have to decide the intervals for $I_i, i = 0, 1, 2$ to obtain the model coefficients as explained in Definition 2.4.

- Before the application we have to decide the intervals for $I_i, i = 0, 1, 2$ to obtain the model coefficients as explained in Definition 2.4.
- We use same intervals in order to compare the results we have with the results from Abdalla and Buckley (2008) in the literature.

- Before the application we have to decide the intervals for $I_i, i = 0, 1, 2$ to obtain the model coefficients as explained in Definition 2.4.
- We use same intervals in order to compare the results we have with the results from Abdalla and Buckley (2008) in the literature.
- Four separate intervals $(MCI, MCII, MCIII, MCIV)$ that they studied are given with Table 5.

- Before the application we have to decide the intervals for $I_i, i = 0, 1, 2$ to obtain the model coefficients as explained in Definition 2.4.
- We use same intervals in order to compare the results we have with the results from Abdalla and Buckley (2008) in the literature.
- Four separate intervals $(MCI, MCII, MCIII, MCIV)$ that they studied are given with Table 5.

Table: Intervals for $I_i, i = 0, 1, 2$ for third category

Interval	*MCI*	*MCII*	*MCIII*	*MCIV*
I_0	[0,5]	[0,37]	[16.528,16.528]	[33.808,36.601]
I_1	[0,6]	[0,6]	[3.558,3.982]	[1.294,3.756]
I_2	[0,4]	[0,6]	[2.575,2.575]	[0.423,0.473]

Table: Results for using different definitions of distance measures in fuzzy linear regression with MC method for minimizing E_1.

		Intervals			
Distance Measures	Parameters	*MCI*	*MCII*	*MCIII*	*MCIV*
Kaufmann (1991)	a_0	1.9138	1.8114	16.5280	33.8108
	a_1	4.7655	4.7820	3.5733	3.1333
	a_2	3.6687	3.6775	2.5750	0.4730
Heilpern-1 (1997)	a_0	2.4841	0.3650	16.5280	33.8106
	a_1	4.9058	4.8024	3.5580	2.7181
	a_2	3.4424	3.9099	2.5750	0.7430
Heilpern-2 (1997)	a_0	1.9138	1.8114	16.5280	33.8108
	a_1	4.7655	4.7820	3.5733	3.1333
	a_2	3.6687	3.6775	2.5750	0.4730
Heilpern-3 (1997)	a_0	4.4812	5.5354	16.5280	33.8111
	a_1	4.5835	4.5590	3.5580	3.0608
	a_2	3.4776	3.3425	2.5750	0.4730
Chen and Hsieh (1998)	a_0	2.1047	0.5538	16.5280	33.8086
	a_1	5.0605	5.0276	3.5580	3.0994
	a_2	3.3305	3.6148	2.5750	0.4730

Table: Error measures for application (second category)

E_1	MCI	$MCII$	$MCIII$	$MCIV$
Abdalla and Buckley (2008)	6.169	5.812	7.125	8.201
Kaufmann (1991)	32.63132	31.0182	24.1279	110.6466
Heilpern-1 (1997)	4.5126	6.8999	12.202	50.9251
Heilpern-2 (1997)	16.31566	15.5091	12.06395	55.3233
Heilpern-3 (1997)	16.3649	15.104	9.2622	40.2581
Chen and Hsieh (1998)	6.1242	4.8169	11.7306	58.7061

Table: Error measures for application (third category)

E_1	MCI	$MCII$	$MCIII$	$MCIV$
Abdalla and Buckley (2008)	10.017	9.389	12.7267	9.5933
Kaufmann (1991)	52.7943	83.9582	19.0558	24.3161
Heilpern-1 (1997)	26.2680	42.0170	9.4604	13.4241
Heilpern-2 (1997)	26.3971	41.9791	9.5279	12.1581
Heilpern-3 (1997)	19.8377	31.5128	7.2577	9.4778
Chen and Hsieh (1998)	26.3563	41.9412	9.4544	11.6395

Outline

1. Introduction
2. Preliminaries
3. Fuzzy Regression with Monte Carlo Method
4. Distance Measure for Fuzzy Numbers
5. Application
   - Application for Second Category
   - Application for Third Category
   - Solutions
6. Conclusion

Why we did this study!!!

- Monte Carlo methods in fuzzy regression is a very new and potential area that is easy to calculate model parameters without any long and complex mathematical equations, also no need for any regression assumptions.

Why we did this study!!!

- Monte Carlo methods in fuzzy regression is a very new and potential area that is easy to calculate model parameters without any long and complex mathematical equations, also no need for any regression assumptions.
- Distance measure between fuzzy numbers have gained importance due to the widespread applications in diverse fields like decision making, machine learning and market prediction.

Why we did this study!!!

- Monte Carlo methods in fuzzy regression is a very new and potential area that is easy to calculate model parameters without any long and complex mathematical equations, also no need for any regression assumptions.
- Distance measure between fuzzy numbers have gained importance due to the widespread applications in diverse fields like decision making, machine learning and market prediction.
- There are several different definitions of distance measure between two fuzzy numbers in the literature

Why we did this study!!!

- Monte Carlo methods in fuzzy regression is a very new and potential area that is easy to calculate model parameters without any long and complex mathematical equations, also no need for any regression assumptions.
- Distance measure between fuzzy numbers have gained importance due to the widespread applications in diverse fields like decision making, machine learning and market prediction.
- There are several different definitions of distance measure between two fuzzy numbers in the literature

Why we did this study!!!

- Monte Carlo methods in fuzzy regression is a very new and potential area that is easy to calculate model parameters without any long and complex mathematical equations, also no need for any regression assumptions.
- Distance measure between fuzzy numbers have gained importance due to the widespread applications in diverse fields like decision making, machine learning and market prediction.
- There are several different definitions of distance measure between two fuzzy numbers in the literature

Reason

- Only one definition of distance measure has been used in fuzzy regression with Monte Carlo method until now.
- Hence, we investigate using different definitions of distance measure between fuzzy numbers in estimating the parameters of fuzzy regression with Monte Carlo method.

Future Works !!!

Future Works !!!

- Making a simulation above the intervals according to the distance measures. For deciding which distance measure is the best for estimating the parameters.

Future Works !!!

- Making a simulation above the intervals according to the distance measures. For deciding which distance measure is the best for estimating the parameters.
- Investigating different definitions of distance measure between fuzzy numbers in different types of fuzzy regression models, such as nonparametric regression, exponential regression and

Future Works !!!

- Making a simulation above the intervals according to the distance measures. For deciding which distance measure is the best for estimating the parameters.
- Investigating different definitions of distance measure between fuzzy numbers in different types of fuzzy regression models, such as nonparametric regression, exponential regression and
- Considering different types of fuzzy numbers, such as trapezoidal, Gaussian in these regression models are potential future works.

Future Works !!!

- Making a simulation above the intervals according to the distance measures. For deciding which distance measure is the best for estimating the parameters.
- Investigating different definitions of distance measure between fuzzy numbers in different types of fuzzy regression models, such as nonparametric regression, exponential regression and
- Considering different types of fuzzy numbers, such as trapezoidal, Gaussian in these regression models are potential future works.

References

- Abdalla A, Buckley JJ (2007) Monte Carlo methods in fuzzy linear regression. Soft Comput, 11:991-996
- Abdalla A, Buckley JJ (2008) Monte Carlo methods in fuzzy linear regression II. Soft Comput, 12:463-468
- Bardossy, A (1990) Note on fuzzy regression. Fuzzy Sets Syst., 37:65-75
- Chen SH, Hsieh CH (2000) Representation, Ranking, Distance, and Similarity of L-R type fuzzy number and Application, Australia Journal of Intelligent Information Processing Systems, 6(4):217-229
- Choi HS, Buckley JJ (2007) Fuzzy regression using least absolute deviation estimators, Soft Comput 12:257-263
- Choi HS, Buckley JJ. (2008) Fuzzy regression using least absolute deviation estimators. Soft Comput, 12:257-263.
- Diamond P (1987) Least squares fitting of several fuzzy variables. In Proc of Second IFSA Congress. IFSA, Tokyo,. p. 20-25.
- Diamond, P, Korner, R (1997) Extended Fuzzy linear models and least squares estimate, Comput Math Appl 33:15-32
- Dubois, D, Prade, H (1978) Operations on fuzzy numbers, International Journal of Systems Science, vol.9, no.6, .613-626
- Gao, S, Zhang, Z, Cao, C (2009) Multiplication operation on fuzzy numbers, Journal of Software, 4,4, 331-338.
- Hajjari T (2010) Measuring Distance of Fuzzy Numbers by Trapezoidal Fuzzy Numbers, AIP Conference Proceedings 1309, 346
- Heilpern S (1997) Representation and application of fuzzy numbers, Fuzzy sets and Systems, 91(2):259-268

References

- Hsieh CH, Chen SH (1998) Graded mean representation distance of generalized fuzzy number, Proceeding of sixth Conference on Fuzzy Theory and its Applications, Chinese Fuzzy Systems Association
- Kaufmann A, Gupta MM (1991) Introduction to fuzzy arithmetic theory and applications, Van Nostrand Reinhold
- Kim B, Bishu RR (1998) Evaluation of fuzzy linear regression models by comparing membership functions, Fuzzy sets and systems, 100, 343-352
- Savic DA, Pedryzc W (1991) evaluation of fuzzy linear regression models. Fuzzy Sets Syst 39:51-63.
- Tanaka H (1987) Fuzzy Data analysis by possibilistic linear regression models. Fuzzy Sets Syst. 24:363-375.
- Tanaka H, Lee H (1999) Exponential possibility regression analysis by identification method of possibilistic coefficients. Fuzzy Sets Syst., 106:155-165
- Tanaka H, Uejima S, Asai K. (1982) Linear regression analysis with fuzzy model. IEEE Trans. Systems Man Cybernet, 12: 903-907
- Zadeh LA (1965) Fuzzy Sets. Information and control, 8, 338-353