Journal of Economics
and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq

Semi Parametric Logistic Regression Model with the Outputs Representing Trapezoidal Intuitionistic Fuzzy Number

Ayad H. shemail
Department of Statistics, College of Administration & Economics, University of Diyala, Iraq.
mailto:ayadstatistic@uodiyala.edu.iq

Mohammed J. Mohammed
Department of Statistics, College of Administration & Economics, University of Baghdad, Iraq
mailto:m.jasim@coadec.uobaghdad.edu.iq

Received:29/5/2022 **Accepted:28/6/2022** **Published: September / 2022**

Abstract:

In this paper, the fuzzy logic and the trapezoidal fuzzy intuitionistic number were presented, as well as some properties of the trapezoidal fuzzy intuitionistic number and semi- parametric logistic regression model when using the trapezoidal fuzzy intuitionistic number. The output variable represents the dependent variable sometimes cannot be determined in only two cases (response, non-response)or (success, failure) and more than two responses, especially in medical studies; therefore so, use a semi parametric logistic regression model with the output variable (dependent variable) representing a trapezoidal fuzzy intuitionistic number.

the model was estimated on simulation data when sample sizes 25,50 and 100, as the parametric part was estimated by two methods of estimation, are fuzzy ordinary least squares estimators FOLSE method and suggested fuzzy weighted least squares estimators SFWLSE , while the non-parametric part is estimated by Nadaraya Watson estimation and Nearest Neighbor estimator. The results were the fuzzy ordinary least squares estimators method was better than the suggested fuzzy weighted least squares estimators while, in the non-parametric portion, the Nadaraya Watson estimators had better than Nearest Neighbor estimators to estimate the model.

Keywords: **semiparametric logistic regression model, trapezoidal intuitionistic fuzzy number, SFWLSE, FOLSE.**

1. Introduction

Fuzzy logic can be defined as modeling events or phenomena in an inaccurate style, which is in a style that recognizes the existence of ambiguity; one of the objectives of fuzzy logic is to find a method to solve very complex questions and mysterious, which cannot analyze by using pure mathematical methods. Fuzzy logic is to make guesses and estimates dependent on a mathematical-statistical model that shows the effect of variables on each other (Klir and Yuan, 1996).

The fuzzy logic was created by the researcher lotfi zadeh in the year 1965 in the university of California when display the fuzzy set is a set of the element which have a degree of belong. It's called the membership function (Zadeh, 1965).

Tanaka et al (1982) studied the fuzzy sets on a linear regression model whose objective was to model ambiguous or inaccurate events using fuzzy parameters, and this model was developed.

The semiparametric logistic regression model is one of the most important regression models, which is composed of two parts. Firstly, the partition is parametric, and secondly is the non-parametric partition represented by any non-parametric smoothing estimator, while the semiparametric logistic regression model has become more applicable in the last few years on the fuzzy sets.

When the output variable (dependent variable) is a binary response variable, logistic regression is used to study the effect between the output variable and the explanatory variables (independent variable). As the results are sometimes imprecise, the output variable (dependent variable) sometimes can't define in only two terms, (response, non-response) or (success, failure) and as well more than two responses. Particularly in medical side, the patient's response to treatment or the rate of injury to the person is in the form of periods, for example, from (5% to 25%) or (20% to 40%) (35% to 55) or (50% to 70) or (65 to 85) and to assess the severity of a patient's condition or suffering, the terms (low, very low, medium, high, and very high) are used, These terms are usually considered to be a fuzzy sets which is expressed by the intuitionistic fuzzy number.

This paper discusses, that the output variable (dependent variable) represents the trapezoidal Intuitionistic fuzzy numbers. As a result, the semiparametric logistic regression model will be used to cope with fuzzy or vagueness situations.

Several researchers have investigated fuzzy regression models; Diamond (1988) studied the development of several fuzzy regression models for constructing simple least squares for fuzzy observation data. Carroll and Wand (1991) studied the semi parametric logistic regression model with the measurement error when using the prediction. Cheng and Lee (1999) verified the fuzzy nonparametric regression by modifying the model through kernel smoothing method and nearest neighbor smoothing method. Wu et al (2010) suggested a novel technique based on fuzzy logic data and study of the logistic regression model to detect financial crises on the banking industry in Taiwan. Namdari et al (2015) studied the fuzzy logistic regression model for crisp inputs and fuzzy output data. Hesamian et al (2017) studied the logistic regression effect between inputs and outputs when you are fuzzy through intuitive fuzzy groups instead of the usual fuzzy groups. Ahmadini (2021) suggested a new fuzzy approach for logistic regression model while dealing with fuzzy parameters with degrees difference of uncertainty and fuzzy simultaneously. Shemail and Mohammed (2022) studied the semiparametric

logistic regression model on the fuzzy intuitionistic data representing Coronavirus when using the triangular intuitionistic fuzzy number.

2. Trapezoidal Fuzzy Intuitionistic Number

Before starting to the trapezoidal fuzzy intuitionistic number, we must explain the intuitionistic fuzzy set, Let X universal set and $\mu_{\acute{A}}(x)$ membership function to the fuzzy set $\acute{A}$.

The fuzzy intuitionistic set A can be expressed mathematically in X through a set represent A= {the element x, the membership function $\mu_A(x)$ of element x, the non-membership function $\nu_A(x)$ of element x} according to the formula of the equation (Szmidt and Kacprzyk, 2002):

$$A = \{\langle x, \mu_A(x), \nu_A(x)\rangle / x \in X\} \qquad (1)$$

When $\mu_A(x)$ is the membership function of variable x to the fuzzy intuitionistic set, $\nu_A(x)$ is the non- membership function of x (De et al, 2001).

Let $(\acute{l}_a > l_a > 0; \acute{r}_b > r_b > 0)$ the trapezoidal fuzzy intuitionistic number TrFIN is$\bar{A} = (a, b; l_a, r_b; \acute{l}_a, \acute{r}_b)$; TrIFN be in the following form (Garai et al, 2018):

$$\mu_A(x_i) = \begin{cases} \frac{x_i - a + l_a}{l_a} & if \quad a - l_a \le x_i < a \\ 1 & if \quad a \le x_i \le b \\ \frac{b + r_b - x_i}{r_b} & if \quad b < x_i \le b + r_b \\ 0 & otherwise \end{cases}$$

$$\nu_A(x_i) = \begin{cases} 1 - \frac{x_i - a + \acute{l}_a}{\acute{l}_a} & if \quad a - \acute{l}_a \le x_i < a \\ 0 & if \quad a \le x_i \le b \\ 1 - \frac{b + \acute{r}_b - x_i}{\acute{r}_b} & if \quad b < x_i \le b + \acute{r}_b \\ 1 & otherwise \end{cases} \qquad (2)$$

If there are two numbers trapezoidal intuitionistic fuzzy number $\bar{A} = (a_2, a_3; l_a, r_a; \acute{l}_a, \acute{r}_a)$, so that $\acute{l}_a < l_a < a_2 < a_3 < r_a < \acute{r}_a$ and the trapezoidal intuitionistic fuzzy number$\bar{B} = (b_2, b_3; l_b, r_b; \acute{l}_b, \acute{r}_b)$, so that $\acute{l}_b < l_b < b_2 < b_3 < r_b < \acute{r}_b$, the some arithmetic operations are as follows (Chakraborty et al, 2015):

1. $\bar{A} \oplus \bar{B} = (a_2 + b_2, a_3 + b_3;\ l_a + l_b, r_a + r_b;\ \acute{l}_a + \acute{l}_b, \acute{r}_a + \acute{r}_b)$
2. $\bar{A} \ominus \bar{B} = (a_2 - b_3, a_3 - b_2; l_a + r_b, r_a + l_a; \acute{l}_a + \acute{r}_b, \acute{r}_a + \acute{l}_a)$
3. $\bar{A} \otimes \bar{B} \approx \left(a_2.b_2, a_3.b_3; a_2 l_b + b_2 l_a, a_3 r_b + b_3 r_a; a_2 \acute{l}_b + \acute{b_2} l_a, a_3 r_b + \acute{b_3} \acute{r}_a\right)$
4. $\bar{A} \oslash \bar{B} \approx \left(\frac{a_2}{b_3}, \frac{a_3}{b_2}; \frac{a_2 r_b + b_3 l_a}{b_3^2}, \frac{a_3 l_b + b_2 r_a}{b_2^2}; \frac{a_2 \acute{r}_b + b_3 \acute{l}_a}{b_3^2}, \frac{a_3 \acute{l}_b + b_2 \acute{r}_a}{b_2^2}\right)$

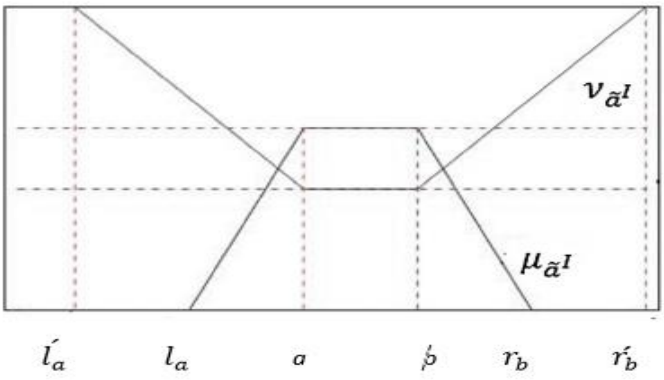

Figure 1: membership function and non-membership function to TrIFN.

3. Semiparametric Logistic Regression Model

The semiparametric logistic regression model with exact inputs and trapezoidal intuitionistic fuzzy number outputs is defined as (Hesamian, and Akbari, 2017).

$$\mathrm{V}_i = \ln\left(\frac{\overline{p_\iota}}{1 \ominus \overline{p_\iota}}\right) = \oplus_{j=1}^{k}\left(\underline{\beta} \otimes x_{ij}\right) \oplus g(t_i) \oplus \epsilon_i \qquad (3)$$

Where $i = 1, 2, \dots, n$ The number of observations, $j = 1, 2, \dots\dots k$ the number of input variables, $\overline{p_\iota}$ he fuzzy logistic regression function represents the trapezoidal intuitionistic fuzzy $\overline{p_\iota} = (\overline{p_{2\iota}}, \overline{p_{3\iota}}; l_{\overline{p_\iota}}, r_{\overline{p_\iota}}; l'_{\overline{p_{\iota_\iota}}}, r'_{\overline{p_\iota}})$, $\underline{\beta}$ the vector of unidentified parameters represent the trapezoidal fuzzy intuitionistic parameters $\beta_j = (b_{j2}, b_{j3}; l_{b_j}, r_{b_j}; l'_{b_j}, r'_{b_j})$, $g(t_i)$ Smooth function when estimate represent the trapezoidal intuitionistic fuzzy $\hat{g}(t_i) = ((t_{i2}), (t_{i3}); l_{t_i}, r_{t_i}; l'_{t_\iota}, r'_{t_\iota})$, ϵ_i error term.

We can find the logit function to the dependent variable (fuzzy output) to represent the trapezoidal fuzzy intuitionistic number $\overline{p_\iota} = (\overline{p_{2\iota}}, \overline{p_{3\iota}}; l_{\overline{p_\iota}}, r_{\overline{p_\iota}}; l'_{\overline{p_{\iota_\iota}}}, r'_{\overline{p_{\iota_\iota}}})$ and depending on the points (2, 4) from some arithmetic operations of fuzzy intuitionistic set as follows:

$$V_i = ln\frac{\overline{p_\iota}}{1-\overline{p_\iota}} = \left(v_{2i}, v_{3i};\ l_{v_i}, r_{v_i}; l'_{v_\iota}, r'_{v_\iota}\right) \qquad (4)$$

Thus, the output (dependent variable) representing the trapezoidal intuitionistic the fuzzy number to the semiparametric logistic regression model is:

$$V_i = ln\frac{\overline{p_\iota}}{1-\overline{p_\iota}} =$$

$$\left(ln\left(\frac{\overline{p_{2\iota}}}{1-\overline{p_{2\iota}}}\right), ln\left(\frac{\overline{p_{3\iota}}}{1-\overline{p_{3\iota}}}\right); ln\frac{\overline{p_{2\iota}}+l_{\overline{p_\iota}}}{(1-\overline{p_{2\iota}})^2}, ln\frac{\overline{p_{3\iota}}+r_{\overline{p_\iota}}}{(1-\overline{p_{3\iota}})^2}; ln\frac{\overline{p_{2\iota}}+l'_{\overline{p_\iota}}}{(1-\overline{p_{2\iota}})^2}, ln\frac{\overline{p_{3\iota}}+r'_{\overline{p_\iota}}}{(1-\overline{p_{3\iota}})^2}\right) \qquad (5)$$

This model is estimated by the Hesamian, and Akbari method from two steps, The parametric portion is estimated in the first step in the model then estimate the non-parametric portion in the second step as follows (Wang et al, 2007) (Hesamian, and Akbari, 2017).:

$$\hat{g}(t) = \sum_{i=1}^{n} w_{hi}(t_j) \otimes \left(\mathrm{V}_i - \oplus_{j=1}^{k}\left(\underline{\beta} \otimes x_{ij}\right)\right) \qquad (6)$$

This leads to an estimate of the semi parametric logistic regression model when using the trapezoidal intuitionistic fuzzy number as follows:

$$\widehat{V}_i = \left(\sum_{i=1}^{n} w_{hi}(t_j) \otimes V_i\right) \oplus \left(\oplus_{j=1}^{k} \underline{\beta} \otimes x_{ij}^{**}\right) \qquad (7)$$

Where $w_{hi}(t_j)$ the smoothing weights are estimated based on the Nadaraya-Watson estimator and the Nearest-Neighbor , and the total of a sequence of weights equals one, $\sum_{i=1}^{n} w_{hi}(t_j) = 1$, and $x_{ij}^{**} = x_{ij} - \sum_{i=1}^{n} w_{hi}(t_j)\, x_{ij}$.

The estimated output, which represents the trapezoidal intuitionistic fuzzy number in semiparametric logistic regression model, is as follows:

$$\hat{v}_{2i} = \left(\sum_{i=1}^{n} w_{hi}(t_j) \otimes v_{2i}\right) \oplus \left(\oplus_{j=1}^{k} b_{2j} \otimes x_{ij}^{**}\right)$$

$$\hat{v}_{3i} = \left(\sum_{i=1}^{n} w_{hi}(t_j) \otimes v_{3i}\right) \oplus \left(\oplus_{j=1}^{k} b_{3j} \otimes x_{ij}^{**}\right)$$

$$l_{\widehat{v_i}} = \left(\sum_{i=1}^{n} w_{hi}(t_j) \otimes l_{v_i}\right) \oplus \left(\oplus_{j=1}^{k} l_{b_j} \otimes s_{ji} x_{ij}^{**}\right) \ominus \left(\oplus_{j=1}^{k} r_{b_j} \otimes (1 - s_{ji}) x_{ij}^{**}\right)$$

$$r_{\widehat{v_i}} = \left(\sum_{i=1}^{n} w_{hi}(t_j) \otimes r_{v_i}\right) \oplus \left(\oplus_{j=1}^{k} r_{b_j} \otimes s_{ji} x_{ij}^{**})\right) \ominus \left(\oplus_{j=1}^{k} l_{b_j} \otimes (1 - s_{ji}) x_{ij}^{**}\right)$$

$$l'_{\widehat{v_i}} = \left(\sum_{i=1}^{n} w_{hi}(t_j) \otimes l'_{v_i}\right) \oplus \left(\oplus_{j=1}^{k} l'_{b_j} \otimes s_{ji} x_{ij}^{**}\right) \ominus \left(\oplus_{j=1}^{k} r'_{b_j} \otimes (1 - s_{ji}) x_{ij}^{**}\right)$$

$$r'_{\widehat{v_i}} = \left(\sum_{i=1}^{n} w_{hi}(t_j) \otimes r'_{v_i}\right) \oplus \left(\oplus_{j=1}^{k} r'_{b_j} \otimes s_{ji} x_{ij}^{**}\right) \ominus \left(\oplus_{j=1}^{k} l'_{b_j} \otimes (1 - s_{ji}) x_{ij}^{**}\right)$$

Where $s_{ji} = \begin{cases} 1 & x_{ij}^{**} \geq 1 \\ 0 & x_{ij}^{**} < 0 \end{cases} \qquad (8)$

We can write the estimation model in a matrix as follows:

$$\underline{\widehat{V}} = \left(W\underline{V_2} + X^*\underline{\widehat{\beta}}_2, W\underline{V_3} + X^*\underline{\widehat{\beta}}_3; W\underline{L_V} + X_s^*\underline{L} - X_{1-s}^*\underline{R}, W\underline{R_V} + X_s^*\underline{R} - X_{1-s}^*\underline{L}; W\underline{L'_V} + X_s^*\underline{L'} - X_{1-s}^*\underline{R'}, W\underline{R'_V} + X_s^*\underline{R'} - X_{1-s}^*\underline{L'}\right) \qquad (9)$$

Where

$$W = \begin{bmatrix} w_{11} & w_{12} & . & . & . & w_{1n} \\ w_{21} & w_{22} & . & . & . & w_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ . & . & . & . & . & . \\ w_{n1} & w_{n2} & . & . & . & w_{nn} \end{bmatrix} \quad X = \begin{bmatrix} x_{11} & x_{12} & . & . & . & x_{1p} \\ x_{21} & x_{22} & . & . & . & x_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ . & . & . & . & . & . \\ x_{n1} & x_{n2} & . & . & . & x_{np} \end{bmatrix}$$

$$X^{**} = \begin{bmatrix} x_{11}^{**} & x_{12}^{**} & . & . & . & x_{1k}^{**} \\ x_{21}^{**} & x_{22}^{**} & . & . & . & x_{2k}^{**} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ . & . & . & . & . & . \\ x_{n1}^{**} & x_{n2}^{**} & . & . & . & x_{nk}^{**} \end{bmatrix} \quad X_s^{**} = \begin{bmatrix} s_{11}x_{11}^{**} & s_{12}x_{12}^{**} & . & . & . & s_{1k}x_{1k}^{**} \\ s_{21}x_{21}^{**} & s_{22}x_{22}^{**} & . & . & . & s_{2k}x_{2k}^{**} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ . & . & . & . & . & . \\ s_{n1}x_{n1}^{**} & s_{n2}x_{n2}^{**} & . & . & . & s_{nk}x_{nk}^{**} \end{bmatrix}$$

$$X_{(1-s)}^{**} = \begin{bmatrix} (1 - s_{11})x_{11}^{**} & (1 - s_{12})x_{12}^{**} & . & . & . & (1 - s_{1k})x_{1k}^{**} \\ (1 - s_{21})x_{21}^{*} & s_{22}x_{22}^{*} & . & . & . & s_{2k}x_{2k}^{*} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ . & . & . & . & . & . \\ (1 - s_{n1})x_{n1}^{**} & (1 - s_{n2})x_{n2}^{**} & . & . & . & (1 - s_{nk})x_{nk}^{**} \end{bmatrix}$$

$$\underline{\widehat{\beta_2}} = \begin{bmatrix} b_{21} \\ b_{22} \\ \cdot \\ \cdot \\ \cdot \\ b_{2j} \end{bmatrix}, \underline{\widehat{\beta_3}} = \begin{bmatrix} b_{31} \\ b_{32} \\ \cdot \\ \cdot \\ \cdot \\ b_{3j} \end{bmatrix}, \ \underline{L} = \begin{bmatrix} l_{b_1} \\ l_{b_2} \\ \cdot \\ \cdot \\ \cdot \\ l_{b_n} \end{bmatrix}, \underline{R} = \begin{bmatrix} r_{P_1} \\ r_{P_2} \\ \cdot \\ \cdot \\ \cdot \\ r_{P_n} \end{bmatrix}, \underline{\acute{L}} = \begin{bmatrix} l'_{P_1} \\ l'_{P_2} \\ \cdot \\ \cdot \\ \cdot \\ l'_{P_n} \end{bmatrix}, \underline{\acute{R}} = \begin{bmatrix} r'_{P_1} \\ r'_{P_2} \\ \cdot \\ \cdot \\ \cdot \\ r'_{P_n} \end{bmatrix}$$

$$V_2 = \begin{bmatrix} v_{21} \\ v_{22} \\ \cdot \\ \cdot \\ \cdot \\ v_{2n} \end{bmatrix} V_3 = \begin{bmatrix} v_{31} \\ v_{32} \\ \cdot \\ \cdot \\ \cdot \\ v_{3n} \end{bmatrix}, L_V = \begin{bmatrix} l_{v_1} \\ l_{v_2} \\ \cdot \\ \cdot \\ \cdot \\ l_{v_n} \end{bmatrix}, R_V = \begin{bmatrix} r_{v_1} \\ r_{v_2} \\ \cdot \\ \cdot \\ \cdot \\ r_{v_n} \end{bmatrix}, \acute{L_V} = \begin{bmatrix} l'_{v_1} \\ l'_{v_2} \\ \cdot \\ \cdot \\ \cdot \\ l'_{v_n} \end{bmatrix}, \acute{R_V} = \begin{bmatrix} r'_{v_1} \\ r'_{v_2} \\ \cdot \\ \cdot \\ \cdot \\ r'_{v_n} \end{bmatrix}$$

4. Fuzzy Ordinary Least Squares Estimators (FOLSE)

Spekman researcher was the first to discuss the partial semi-parametric regression in 1988 through the following model (Speckman, 1988):

$$\underline{V} = \mathrm{X}\underline{\beta} + \mathrm{Z}\underline{\gamma} + \underline{\mathrm{U}} \qquad (10)$$

Where $\underline{V}$ vector dependent variable, X independent variable matrix (n*p) and $\underline{\beta}$ vertical vector of parameters (p*1), and where Z elements of kernel smoothing function from rank (n*p), $\underline{\gamma}$ added parameters vector, $\underline{\mathrm{U}}$ vector of the error term.

Depending the model above the fuzzy ordinary least squares estimators in the case of the output represents the trapezoidal intuitionistic fuzzy number $(v_{2i}, v_{3i};\ l_{v_i}, r_{v_i};\ l'_{v_i}, r'_{v_i})$ as follow:

$$\begin{aligned} \underline{\widehat{\beta}_2} &= \left[\acute{\mathrm{X}}(I - P_W)\mathrm{X}\right]^{-1}\acute{\mathrm{X}}(I - P_W)\underline{V}_2 \\ \underline{\widehat{\beta}_3} &= \left[\acute{\mathrm{X}}(I - P_W)\mathrm{X}\right]^{-1}\acute{\mathrm{X}}(I - P_W)\underline{V}_3 \\ \underline{\hat{L}} &= \left[\acute{\mathrm{X}}(I - P_W)\mathrm{X}\right]^{-1}\acute{\mathrm{X}}(I - P_W)L_{\underline{V}} \\ \underline{\hat{R}} &= \left[\acute{\mathrm{X}}(I - P_W)\mathrm{X}\right]^{-1}\acute{\mathrm{X}}(I - P_W)R_{\underline{V}} \\ \underline{\hat{\acute{L}}} &= \left[\acute{\mathrm{X}}(I - P_W)\mathrm{X}\right]^{-1}\acute{\mathrm{X}}(I - P_W)\acute{L}_{\underline{V}} \\ \underline{\hat{\acute{R}}} &= \left[\acute{\mathrm{X}}(I - P_W)\mathrm{X}\right]^{-1}\acute{\mathrm{X}}(I - P_W)\acute{R}_{\underline{V}} \end{aligned} \qquad (11)$$

Where $P_W = \mathrm{Z}(\acute{\mathrm{Z}}\mathrm{Z})^{-1}\acute{\mathrm{Z}}$

5. Suggested Fuzzy Weighted Least Squares Estimators (SFWLSE)

We will suggest an estimator called suggested fuzzy weighted least squares estimators based the partial semi-parametric regression model that was introduced by the scientist Spekmain and the robust weights presented by (Şanli and paydin, 2004).

where the robust weights are entered on the fuzzy partial semi-parametric regression model according to the following formula:

$$p^{-1}V = P^{-1}\mathrm{X}\underline{\beta} + P^{-1}\mathrm{Z}\underline{\gamma} + P^{-1}\underline{\mathrm{U}} \qquad (12)$$

Where p^{-1} diagonal matrix dimensions n*n, its elements represent the roots of the robust weight matrix, where $W = \acute{P}^{-1}P^{-1}$ A diagonal matrix of dimensions n * n represents the robust weights.

$$\underline{\acute{\mathrm{U}}}\mathrm{W}\underline{\mathrm{U}} = \left(V - \mathrm{X}\underline{\beta} - \mathrm{Z}\underline{\gamma}\right)' \mathrm{W} \left(V - \mathrm{X}\underline{\beta} - \mathrm{Z}\underline{\gamma}\right) \qquad (13)$$

By derivation equation No. 13 to parameters:

$$\frac{\partial \acute{\underline{U}}W\underline{U}}{\partial \underline{\beta}} = -\acute{X}W\left(V - X\underline{\widehat{\beta}} - Z\underline{\widehat{\gamma}}\right) = 0$$

$$\frac{\partial \acute{\underline{U}}W\underline{U}}{\partial \underline{\gamma}} = -\acute{Z}W\left(V - X\underline{\widehat{\beta}} - Z\underline{\widehat{\gamma}}\right) = 0$$

The estimations will be as follows:

$$\underline{\widehat{\beta}} = \left[\acute{X}W(I - P_W W)X\right]^{-1}\acute{X}W(I - P_W W)V \qquad (14)$$

Where $P_W = Z(\acute{Z}WZ)^{-1}\acute{Z}$

The diagonal robust weights matrix W is calculated according to the following steps: (Şanli and paydin, 2004).

1- The estimated regression parameters are calculated based on ordinary least squares in the case of the dependent variable representing the trapezoidal intuitionistic fuzzy number.

2- The residuals $\underline{U}$ are found using the estimated logt function

$$\widehat{V}_i = \text{logit}\,(\bar{p}_i) = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_k X_k$$

Then $\underline{U} = \underline{V} - \underline{\widehat{V}}$

3- The distance is found based on the values of the absolute residuals and the median to the residuals, so the distance is according to the following formula:

$$d_t = \left\|\underline{U} - mead(U)\right\|$$

Where $\|.\|$ represents the Euclidean distance

4- The membership function is calculated according to the distance as follows:

$$M_d = \begin{cases} 1 & d_t \leq \ median(d_t) \\ \dfrac{\max(d_t) - |u|}{\max(d_t) - median(d_t)} & median(d_t) < \ d_t < \max(\ d_t) \\ 0 & o.w \end{cases}$$

5- Depending on the membership function above, determine the robust weights matrix W, which represents a diagonal matrix. The main diameter elements represent the values of the membership function.

6- If $\underline{\beta}^{k+1} - \underline{\beta}^{k} < \varepsilon$ stop, then go to step 2 when ε is a very small values.

The suggested fuzzy weighted least squares estimators in the case the log function of the output variable (response variable) represents the trapezoidal fuzzy intuitionistic number $V_i = \left(v_{2i}, v_{3i};\ l_{v_i}, r_{v_i}; \acute{l}_{v_i}, \acute{r}_{v_i}\right)$ be according to the following formulas:

$$\begin{aligned} \underline{\widehat{\beta}}_2 \ &= \left[\acute{X}W_2(I - P_{W2}W_2)X\right]^{-1}\acute{X}W_2(I - P_{W2}W_2)V_2 \\ \underline{\widehat{\beta}}_3 \ &= \left[\acute{X}W_3(I - P_{W3}W_3)X\right]^{-1}\acute{X}W_3(I - P_{W3}W_3)V_2 \\ \underline{\widehat{L}} &= \left[\acute{X}L_W\left(I - L_{P_W}L_W\right)X\right]^{-1}\acute{X}L_W\left(I - L_{P_W}L_W\right)L_V \\ \underline{\widehat{R}} &= \left[\acute{X}R_W\left(I - R_{P_W}R_W\right)X\right]^{-1}\acute{X}R_W\left(I - R_{P_W}R_W\right)R_V \\ \underline{\widehat{\acute{L}}} &= \left[\acute{X}\acute{L}_W\left(I - \acute{L}_{P_W}\acute{L}_W\right)X\right]^{-1}\acute{X}\acute{L}_W\left(I - \acute{L}_{P_W}\acute{L}_W\right)\acute{L}_V \\ \underline{\widehat{\acute{R}}} &= \left[\acute{X}\acute{R}_W\left(I - \acute{R}_{P_W}\acute{R}_W\right)X\right]^{-1}\acute{X}\acute{R}_W\left(I - \acute{R}_{P_W}\acute{R}_W\right)\acute{R}_V \end{aligned} \qquad (15)$$

Where

$P_{W2} = Z(\acute{Z}W_2Z)^{-1}\acute{Z}$

$P_{W3} = Z(\acute{Z}W_3Z)^{-1}\acute{Z}$

$L_{P_W} = Z(\acute{Z}L_WZ)^{-1}\acute{Z}$

$R_{P_W} = Z(\acute{Z}R_WZ)^{-1}\acute{Z}$

$\acute{L}_{P_W} = Z(\acute{Z}\acute{L}_WZ)^{-1}\acute{Z}$

$\acute{R}_{P_W} = Z(\acute{Z}\acute{R}_WZ)^{-1}\acute{Z}$

6. Simulation

In this section; Simulations were obtained out using four sample sizes (n=25, n=50, n=100) and iterations 1000 through the program R. the simulation experience is described through the following steps

1- The independent variables were generated according to the normal distribution, with 11 variables ($X_{i1}, X_{i2}, \dots X_{i11}$) .

2- The non-parametric variable t_i is generated depending on the following equation. (Wu, 2016):

$$t_i = \left(\frac{i-0.5}{n}\right) \quad , i = 1, 2, \dots, n \qquad (16)$$

Where the nonparametric component it is most used with non-linear models as follows:

$$g(t_i) = 9\left(4.26e^{-3.25t_i} - 4e^{-6.5t_i} + 3e^{-9.75t_i}\right) \qquad (17)$$

3- The initial values for the parameters are:

β_0 = 0.8	β_1=0 .5	β_2=0 .02	β_3 =- 0.3	β_4=0 .02	β_5=0 .03	β_6=0 .04	β_7 =- 0.1	β_8=0 . 3	β_9 =- 0.5	β_{10}= 0.1	β_{11}= 0.02

4- The dependent variable is generated through the logistic function depending on the parametric portion and the non- parametric portion as follows:

$$p\left(y_i = 1/\underline{X} = \underline{x}, t\right) = \frac{e^{0.8+0.5X_{i1}+\cdots+0.02X_{i11}+g(t_i)}}{1+e^{0.8+0.5X_{i1}+\cdots+0.02X_{i11}+g(t_i)}} + ei \qquad (18)$$

When ei the error term is generated by the Bernoulli distribution.

7. Results and Discussion

The values of the dependent variable lie between zero and one.This dependent variable will be converted into a trapezoidal intuitionistic fuzzy dependent variable according to the division in the table no. 1 as follows:

Table 1: Intuitionistic trapezoidal fuzzy dependent variable

Cases	Trapezoidal Intuitionistic dependent variable
Very low	(0.00,0.05,0.16,0.173,0.2,0.25)
Low	(0.2,0.25,0.31,0.313,0.4,0.45)
Medium	(0.4,0.45,0.49,0.497,0.55,0.60)
High	(0.55,0.60,0.667,0.67,0.75,0.80)
Very high	(0.75,0.79,0.80,0.824,0.95,1.00)

Table No. 1 is that of the researcher's diligence by looking at previous studies such as the researcher's studies Hesamian & Akbari. The semi parametric logistic regression model when using trapezoidal intuitionistic fuzzy number is estimated by depending on fuzzy ordinary least squares estimators and suggested fuzzy weighted least squares estimators in the parametric part, while the Nadaraya Watson estimator is used in the case non-parametric part, when using Gaussian kernel function and band width represent (refined plug, Silverman rule of thumb) (Härdle et al, 2004).

The comparison of estimation methods is made in the semi parametric logistic regression model when using trapezoidal fuzzy intuitionistic number based on the mean square error $\mathrm{MSE}(\mathrm{V})$ and the criterion goodness of fit $\mathrm{S}(\widehat{\mathrm{V}},\mathrm{V})$ according to the following formulas (Ye, 2012)(Hesamian and Akbari, 2017):

$$\left(\widehat{V},V\right)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+d\left(\widehat{V},V\right)} \qquad (17)$$

$$MSE(V)=\frac{1}{n}\sum_{i=1}^{n}d\left(\widehat{\mathrm{V}},\mathrm{V}\right) \qquad (18)$$

Where $d\left(\widehat{\mathrm{V}},\mathrm{V}\right)$ represents the Euclidean distance between V and $\widehat{\mathrm{V}}$. The mean square error $\mathrm{MSE}(\mathrm{V})$ and the measure goodness of fit $\mathrm{S}\left(\widehat{\mathrm{V}},\mathrm{V}\right)$ to the semi parametric logistic regression model when the output represent trapezoidal intuitionistic fuzzy number is as follow:

Table 2: Mean Square Error & Goodness of Fit when use Nadaraya Watson

Sample size	Nadaraya Watson		*FOLSE*		*SFWLSE*	
	Kernel Function	Bind Width	$S(\widehat{V},V)$	$MSE(V)$	$S(\widehat{V},V)$	$MSE(V)$
25	Gaussian	Silverman's Rule $\widehat{h}_{rot}$=3.648	0.0827786	11.2459988	0.0585878	16.2262725
		Refined Plug $\widehat{h}_{pm1}$=2.007	0.0828066	11.2482382	0.0586077	16.2282990
50	Nadaraya Watson		*FOLSE*		*SFWLSE*	
	Kernel Function	Bind Width	$S(\widehat{V},V)$	$MSE(V)$	$S(\widehat{V},V)$	$MSE(V)$
	Gaussian	Silverman's Rule $\widehat{h}_{rot}$ = 3.250	0.2612136	3.1564893	0.1191453	7.7703558
		Refined Plug $\widehat{h}_{pm1}$= 1.63	0.2613108	3.1537742	0.1191459	7.7721266
100	Nadaraya Watson		*FOLSE*		*SFWLSE*	
	Kernel Function	Bind Width	$S(\widehat{V},V)$	$MSE(V)$	$S(\widehat{V},V)$	$MSE(V)$

	Gaussian	Silverman's Rule $\hat{h}_{rot}$=2.861	0.2162708	4.1549358	0.1285971	7.1561395
		Refined Plug $\hat{h}_{pm1}$=1.315	0.2162618	4.1554337	0.1285981	7.1566201

Table 3: Mean Square Error & Goodness of Fit when use Nearest Neighbor

Sample size	Nearest-Neighbor		*FOLSE*		*SFWLSE*	
25	Kernel function	Bind width	$S(\hat{V},V)$	*MSE(V)*	$S(\hat{V},V)$	*MSE(V)*
	Gaussian function	R	0.0627090	11.2425682	0.0385457	17.2229983
50	Nearest-Neighbor		*FOLSE*		*SFWLSE*	
	Kernel function	Bind width	$S(\hat{V},V)$	*MSE(V)*	$S(\hat{V},V)$	*MSE(V)*
	Gaussian function	R	0.1909242	3.5651107	0.1891221	8.7656727
100	Nearest-Neighbor		*FOLSE*		*SFWLSE*	
	Kernel function	Bind width	$S(\hat{V},V)$	*MSE(V)*	$S(\hat{V},V)$	*MSE(V)*
	Gaussian function	R	0.1062577	4.1524891	0.0985803	9.1533002

Through the above tables, we find the best-estimated model with a sample size of 50 when using the fuzzy ordinary least square estimators in the parametric portion while in the non-parametric portion use the Nadaraya Watson estimators. And when applying the semi-parametric logistic regression model in sample sizes 25,50 and 100 , we find the fuzzy ordinary least squares estimators were better than the suggested fuzzy weighted least squares estimators and the Nadaraya-Watson estimators are the best from Nearest-Neighbor estimators in the non-parametric portion. In addition, we show that there is no significant difference between the uses of bandwidth in all sample sizes.

8. Conclusions

- We notice the fuzzy ordinary least square estimators best the suggested fuzzy weighted least square estimator.
- We notice the Nadaraya Watson estimators in the non-parametric portion are better from Nearest-Neighbor estimator.
- The mean square error and the criterion of the goodness of fit have an inverse relationship.

Reference

1-Ahmadini, A. A. H. (2021). A novel technique for parameter estimation in intuitionistic fuzzy logistic regression model. Ain Shams Engineering Journal, 13(1), 101518.

2-Shemail, A. H., & Mohammed, M. J. (2022). Comparison between some estimation methods for an intuitionistic fuzzy semi-parametric logistic regression model with practical application about covid-19. International Journal of Nonlinear Analysis and Applications, 13(1), 3723-3732.

3- Klir, G. J., & Yuan, B. (1996). Fuzzy sets and fuzzy logic: theory and applications. Possibility Theory versus Probab. Theory, 32(2), 207-208.

4-Zadeh, L. A. (1965). Fuzzy Sets*. University of California. Department of Electrical Engineering. Information control voll 8, 338—353.
5-Asai, H. T. S. U. K., Tanaka, S., & Uegima, K. (1982). Linear regression analysis with fuzzy model. IEEE Trans. Systems Man Cybern, 12, 903-907.
6-Diamond, P. (1988). Fuzzy least squares. Information Sciences, 46(3), 141-157.
7-Carroll, R. J., & Wand, M. P. (1991). Semiparametric estimation in logistic measurement error models. Journal of the Royal Statistical Society: Series B (Methodological), 53(3), 573-585.
8-Cheng, C. B., & Lee, E. S. (1999). Nonparametric fuzzy regression—k-NN and kernel smoothing techniques. Computers & Mathematics with Applications, 38(3-4), 239-251. cheng and lee.
9-Wu, S. J., Lin, S. L., Wu, D. B., & Ma, H. L. (2010, October). Predicting Risk Parameters Using Intelligent Fuzzy/Logistic Regression. In 2010 International Conference on Artificial Intelligence and Computational Intelligence (Vol. 3, pp. 397-401). IEEE.
10-Namdari, M., Yoon, J. H., Abadi, A., Taheri, S. M., & Choi, S. H. (2015). Fuzzy logistic regression with least absolute deviations estimators. Soft Computing, 19(4), 909-917.
11-Hesamian, G., & Akbari, M. G. (2017). Semi-parametric logistic regression model with exact inputs and intuitionistic fuzzy outputs. Applied Soft Computing, 58, 517-526.
12-Härdle, W. K., Müller, M., Sperlich, S., & Werwatz, A. (2004). Nonparametric and semiparametric models. Springer Science & Business Media.
13-Szmidt, E., & Kacprzyk, J. (2002). Using intuitionistic fuzzy sets in group decision making. Control and cybernetics, 31, 1055-1057.
14-De, S. K., Biswas, R., & Roy, A. R. (2001). An application of intuitionistic fuzzy sets in medical diagnosis. Fuzzy sets and Systems, 117(2), 209-213.
15-Garai, T., Chakraborty, D., & Roy, T. K. (2018). Possibility–necessity–credibility measures on generalized intuitionistic fuzzy number and their applications to multi-product manufacturing system. Granular Computing, 3(4), 285-299.
16-Chakraborty, D., Jana, D. K., & Roy, T. K. (2015). Arithmetic operations on generalized intuitionistic fuzzy number and its applications to transportation problem. Opsearch, 52(3), 431-471.
17-Wang, N., Zhang, W. X., & Mei, C. L. (2007). Fuzzy nonparametric regression based on local linear smoothing technique. Information sciences, 177(18), 3882-3900.
18-Speckman, P. (1988). Kernel smoothing in partial linear models. Journal of the Royal Statistical Society: Series B (Methodological), 50(3), 413-436.
19-Şanli, K., & Apaydin, A. (2004). The fuzzy robust regression analysis, the case of fuzzy data set has outlier. Gazi University Journal of Science, 17(3), 71-84.
20-Ye, J. (2012). Multicriteria group decision-making method using the distances-based similarity measures between intuitionistic trapezoidal fuzzy numbers. International Journal of General Systems, 41(7), 729-739.
21-Wu, J. (2016). Performance of the difference-based almost unbiased Liu estimator in partial linear model. Journal of Statistical Computation and Simulation, 86(14), 2874-2887.

أنموذج الانحدار اللوجيستي شبه المعلمي الضبابي مع المخرجات التي تمثل الرقم الضبابي الحدسي شبة المنحرف

م. أياد حبيب شمال
قسم الاحصاء - كلية الادارة والاقتصاد
جامعة ديالى - العراق
[1]mailto:ayadstatistic@uodiyala.edu.iq

ا. د. محمد جاسم محمد
قسم الاحصاء - كلية الادارة والاقتصاد
جامعة بغداد – العراق
[2]mailto:m.jasim@coadec.uobaghdad.edu.iq

Received:29/5/2022 Accepted:28/6/2022 Published: September / 2022

مستخلص البحث:

في هذا البحث تم دراسة المنطق الصبابي والرقم الضبابي شبة المنحرف الحدسي بالإضافة الى ذلك تم دراسة بعض خصاص الرقم الضبابي شبه المنحرف الحدسي وانموذج الانحدار اللوجستي شبة المعلمي الضبابي عند استعمال الرقم الضبابي الحدسي شبة المنحرف ، لا يمكن في بعض الأحيان تحديد متغير المخرجات (المتغير التابع)في حالتين فقط هما (الاستجابة وعدم الاستجابة) أو(النجاح والفشل) و أكثر من حالتين خاصة في الدراسات الطبية ؛ لذلك استخدم نموذج انحدار لوجستي شبه المعلمي الضبابي مع متغير الإخراج(المتغير التابع) يمثل الرقم الضبابي الحدسي شبة المنحرف.

انموذج الانحدار اللوجستي شبة المعلمي الضبابي يتم تقديره بالاعتماد على بيانات المحاكاة عند حجم العينة (25، 50، 100) الجزء المعلمي في الانموذج يتم تقديره من خلال طريقتين للتقدير هما مقدرات المربعات الصغرى الاعتيادية الضبابية ومقدرات المربعات الصغرى الموزونة الضبابية المقترحة، بينما الجزء اللامعلمي يتم تقديره من خلال مقدرات نداريه واتسون ومقدرات الجار الأقرب. وكانت النتائج تدل على مقدرات المربعات الصغرى الاعتيادية الضبابية أفضل من مقدرات المربعات الصغرى الموزونة الضبابية المقترحة بينما في الجزء غير المعلمي كانت مقدرات Nadaraya Watson أفضل من مقدرات Nearest-Neighbor لتقدير النموذج.

المصطلحات الرئيسة للبحث: انموذج الانحدار اللوجستي شبة المعلمي الضبابي، الرقم الضبابي شبة المنحرف الحدسي، مقدرات المربعات الصغرى الاعتيادية الضبابية، مقدرات المربعات الصغرى الموزونة الضبابية المقترحة.

*بحث مستل من اطروحة دكتوراه