REVSTAT – Statistical Journal
Volume 22, Number 4, October 2024, 411–432
https://doi.org/10.57805/revstat.v22i4.542

Bootstrap Resampling Method for Estimation of Fuzzy Regression Parameters and a Sample Application

Authors: Derviş Topuz ✉
– Niğde Zübeyde Hanim School of Health, Niğde Ömer Halisdemir University, Niğde, Türkiye
topuz@ohu.edu.tr

Volkan Özkaya
– Faculty of Health Sciences, Department of Nutrition and Dietetics, İstanbul Medipol University, İstanbul, Türkiye
volkan.ozkaya@medipol.edu.tr

Betül Çiçek
– Faculty of Health Sciences, Department of Nutrition and Dietetics, Erciyes University, Kayseri, Türkiye
bcicek@erciyes.edu.tr

Received: January 2022 Revised: October 2022 Accepted: October 2022

Abstract:

- In fuzzy regression modeling, the fuzzy least squares technique is based on minimizing the squares of the total difference between observed and predicted outcomes. When the sample size is small, bootstrap resampling method is appropriate and useful for improving model estimation. The bootstrap resampling technique relies on resampling observations and resampling errors for bootstrap regression. The aim of this study is to investigate the use of Bootstrap in fuzzy regression modeling to estimate mean prediction with smaller errors at a particular α-segment, and apply it on a clinical data set. The behavior and properties of the least-squares estimators are affected when deviations or fuzziness arise in the sample and/or by slight changes in the data set. Bootstrap technique, on the other hand, provide robust estimators of the parameters which avoid such adverse effects.

Keywords:

- *bootstrapping; fuzzines; fuzzy linear regression; mean prediction error; fuzzy level; fuzzy confidence interval.*

AMS Subject Classification:

- 49A05, 78B26.

✉ Corresponding author

1. INTRODUCTION

When examining any problem, more than one independent variable may be related to the dependent variable. In order to explain the relationship between such variables, linear regression analysis is the most frequently used technique in statistical models. On the basis of the technique; while evaluating an observed event, it is essential to investigate which events are affected. These events may be one or more, as well as indirectly or directly affected. The technique expresses to what extent the observation values and the affected events are related with the help of a function. Generally, what percentage of the total change in an observed event can be explained by the affected event is evaluated according to the coefficient of determination. However, due to the occurrence of diagnostic procedures related to events in health services (disease, birth, death, etc.) at several levels, complex clinical uncertainties may arise in the relationships due to the insufficiency and uncertainty of information due to the nature of the data and various measurement methods used. In addition, as the sources of uncertainties in clinical relationships, reasons such as the subjective nature of medical history information, objectivity in the examination method, the fact that patient information may contain falsehood, measurement errors in the results of laboratory and other diagnostic tests, due to various restrictive factors, reasons such as the formation of sample sizes in the form of small data sets is shown [21]. In these situations involving uncertainties of medical applications, general facts are that the decisions made by experts can often cause contradictions since valid and reliable sampling and analysis techniques are not preferred. In this case, the calculations made have become questionable.

Various estimation techniques have been developed in order to solve these problems. These techniques are the bootstrap resampling method developed by Efron [21] and fuzzy least squares regression (FLSR) analysis technique by Tanaka *et al.* [59], [24]. The techniques can also be used as a correction method in cases where the assumptions about the error values of the regression model are not realized [55]. These can ensure that there is no difference between the actual observation value and the estimated values or that the difference is minimal. Fuzzy regression analysis is a valid and reliable technique for investigating and predicting data sets by measuring a concept that contains some degree of ambiguity or uncertainty [56]. In models where the data are insufficient or imperfect, caused by the imprecision or vagueness, it has been proven to be useful to use fuzzy models [9], [64]. The importance of using the bootstrap resampling technique and fuzzy regression technique in analysis for estimating model parameters has been increasingly recognized in recent years. Bootstrap resampling technique and fuzzy least squares technique are alternative techniques used in many areas such as time series and simulation in estimating linear and non-linear regression parameters.

This study aims to give illustration and application of bootstrap resampling technique in fuzzy least squares regression analysis in examining the clinical relationships between variables. Because classical least-squares technique is influenced by the outliers, therefore the presents of outliers may distort the estimates. Accordingly, bootstrap fuzzy regression methods have been created to modify least-squares methods so that the outliers have much less influence on the final estimates. We proposed the bootstrap fuzzy least squares regression technique as the estimation approach. The proposed model fitting approach is highly robust to the presence of outliers and properly determines outlier points and neutralizes the negative influence of outliers in the estimation procedure. In addition, we can provide more general approaches that can consider different estimation scenarios. Some hierarchical algorithms concerning bootstrap technique in OLS and fuzzy least squares regression analysis are demonstrated. The basics of the bootstrap resampling technique and their applications to the clinic numerical example that can be described by fuzzy least squares regression model were discussed and compared the results with ordinary least squares regression technique results. It was also aimed to estimate the bias, standard error and confidence interval of the regression coefficients calculated by the techniques and to compare the performance of bootstrap ordinary least squares technique (BOLS) with the related estimates using some comparison criteria. The expectation for the future research topic on fuzzy regression is that many other new proposals and applications will appear in this context. The extension of the proposed procedure to the case of fuzzy input-fuzzy output observations, potential subjects for future researches. It turns out that traditional/common methods in the literature, as well as several other robust methods of fuzzy regression, can be formulated as special instances of bootstrap fuzzy regression.

2. BOOTSTRAP RESAMPLING TECHNIQUE

One of the most important purposes of statistical analysis is that the sample taken from the population must represent the population. The bootstrap resampling technique was developed by Efron [20] as a general technique for assessing the statistical accuracy of an estimator. The main purpose of the technique is to calculate the predictive θ value by choosing random samples with width n volumes, independent of a certain unknown distribution $f(x;\theta)$ and accept it as the predictor of the parameter θ. The bootstrap resampling technique is theoretically used to estimate values associated with the sampling distribution of estimators and test statistics.

The ordinary sampling techniques use some assumptions related to the form of the estimator distribution. These are the cases where standard assumptions are invalid, e.g. n volume is small, data contains uncertainty, data shows non-standard twist. In these situations, the use of these standard techniques may not give reliable and valid results. When these assumptions are doubtful or when the calculation of standard errors is necessary when parametric inference is impossible, the bootstrap resampling technique makes calculations without the need for these distributive assumptions because the sample population is considered [21]. The results calculated by the estimators can be used as an experimental distribution for statistics [11], [20]. The technique has been rarely used, although it is used to generate the estimation of the standard error of a statistic, confidence intervals and distributions by repeated use of the observed data [21], [23], [24], [53].

With the application of the in 4.1. bootstrap algorithm, the bias between population parameters and estimators will be reduced without increasing the sample size, and by obtaining the sampling distributions of the estimators, it will be provided to calculate the standard error of the estimators more accurately [10], [15], [16].

3. FUZZY LEAST SQUARES REGRESSION (FLSR) ANALYSIS

Fuzzy regression analysis is a fuzzy (or possibility) type of ordinary regression analysis. Fuzzy regression analysis studies the relationships between a response variable and a set of explanatory variables in complex systems involving imprecise data. The approach is one of the most widely used statistical techniques for evaluating the functional relationship between dependent and independent variables in uncertainty situations. In fuzzy regression analysis, the relationship between dependent variables and independent variables is not as precise as in ordinary regression analysis [64]. In these uncertain cases, fuzzy techniques can explain the effects of independent variables in a more realistic way. A commonly used technique of the parameter estimation of the fuzzy regression model is the least-squares method. The fuzzy least squares (FLS) technique, which is an extension of the least squares technique to fuzzy set theory, was used by to estimate fuzzy parameters [9], [17], [64]. These methods are very important because sometimes even a single observation can change the value of the parameter estimates, and omitting this observation from the data may lead to totally different estimates [12], [19], [31].

The approach is based on blurring the coefficients. Blurring can be done in two ways. It is possible by 1) blurring the model coefficients estimated by the ordinary least squares technique at a specified "h level", or 2) estimating the coefficients as fuzzy numbers [29], [45], [48]. However, in 1988, Diamond [18] concluded that "Tanaka *et al.* [61] used linear programming techniques to develop a model superficially resembling linear regression, but it is unclear what the relation is to a least squares concept, or that any measure of best fit by residuals is present". Most of the researches on fuzzy regression analysis focuses on the possibilistic regression [59], [61] and on the fuzzy Least-Squares (LS) regression [8], [18]. Recently, robust approaches to fuzzy regression have been considered as alternative approaches to fuzzy regression analysis [13], [15]. The bootstrap resampling technique using fuzzy data is developed in different approaches [46], [58] have considered the problem of hypothesis testing about the mean of a fuzzy random variable. Akbari and Rezaei [1] present a bootstrap fuzzy test for variance. Ferraro *et al.* (2010) [27] "International Journal of Approximate Reasoning 51 (2010) 759–770" proposed to use of a bootstrap procedure to evaluate the accuracy of the estimators in FLS regression. This idea is also investigated and proposed by many authors like "Akbari *et al.* (2012) [2], [24]. In this regard, Peters (1994) [51] considered outliers in Tanaka's possibilistic approach [59] with crisp input-output data which was later extended by Chen (2001) [10] to the model with fuzzy output-crisp input data. Hung and Yang (2006) [31] proposed an omission approach for Tanaka's approach [59] which had the ability to consider the effect of each observation while omitted on the value of the objective function of the model. Nasrabadi *et al.* (2007) [49] proposed an LP-based approach to outliers detection in fuzzy regression analysis. Varga (2007) [63] presented robust estimation approaches to fuzzy and non-fuzzy regression models. Nasrabadi and Hashemi (2008) [50] suggested a robust nonlinear fuzzy regression model using multilayered feedforward neural networks. Kula and Apaydin (2008) [36] proposed a robust fuzzy regression analysis based on the ranking of fuzzy sets. D'Urso and Massari (2013) [19] proposed weighted least-squares and least-median squares estimation for fuzzy linear regression analysis. Yang, Yin and Chen (2013) [66] present a robustified fuzzy varying coefficient model for fuzzy input-fuzzy output variables. Shakouri and Nadimi (2013) [57] investigated a method for outlier detection in fuzzy linear regression problems. Ferraro and Giordani (2013) [28] dealt with robustness in the field of regression analysis for imprecise information managed in terms of fuzzy sets. Leski and Kotas (2015) [41], by introducing an objective function based on Huber's M-estimators and Yager's OWA operators, proposed a robust fuzzy-regression model. Chachi (2019) [12] introduced a weighted objective function to overcome the disadvantages of the LS fuzzy regression approaches in the presence of outliers. Arefi (2020) [5] investigated a quantile fuzzy regression based on fuzzy outputs and fuzzy parameters. Akbari and Hesamian (2019) [3] investigated a partial-robust-ridge-based regression model with fuzzy predictors-responses. Bootstrap fuzzy resampling technique tests for the mean and variance with Dp, q-distance [52]. They proposed bootstrap fuzzy linear regression model (BFLRM), a linear regression model with fuzzy dependent, crisp explanatory and fuzzy coefficients [59], [60]. Most of these developed fuzzy regression models are evaluated with fuzzy outputs and fuzzy parameters but non-fuzzy (net) inputs. Fuzzy least squares regression (FLSR) analysis technique, which is generally based on linear programming (LP), is proposed in order to minimize the fuzziness of the analyzed data and the total spread of the output (see, for example [12], [17] [29]). Hesamian and Akbari (2020) [32] proposed a robust varying coefficient approach to fuzzy multiple regression model. Hesamian and Akbari (2021) [33] adopted a two-stage robust procedure to propose and estimate the components of a robust multiple regression model with fuzzy intercepts and crisp coefficients. Khammar *et al.* (2020) [37], Khammar *et al.* (2021) [38], Khammar *et al.* (2021) [39] presented general approaches to fit fuzzy regression models crisp/fuzzy input and fuzzy output. Asadolahi *et al.* (2021) [6] proposed a robust support vector regression with exact predictors and fuzzy responses. Taheri and Chachi (2021) [62] investigated a robust variable-spread fuzzy regression model. Chachi and Chaji (2021) [14] considered quantile fuzzy regression using OWA operators. In the context of multi-attribute decision-making problems, Chachi *et al.* (2021) [13] developed a multi-objective two-stage optimization and decision technique for fuzzy regression modeling problems in order to handle both of the weak performances analysis of fuzzy regression models and their sensitivity to outliers. As mentioned above, during the last years, considerable attention was given to robust estimation problems in fuzzy environments, and several methodologies were developed in the literature [15].

According to the FLSR approach, it is assumed that the deviations between the observed values and the predicted values are caused by the uncertainty of the system structure or the blurring of the regression coefficients, not from measurement and observation errors, contrary to the OLSR analysis method [9]. That is, it assumes that the coefficients of the regression analysis model are related to its blur. For this purpose, the formula below is employed to estimate parameters of FLSR:

$$f = X \times \widetilde{\beta} \rightarrow \widetilde{Y}_i, \widetilde{Y}_i = f\left(\widetilde{\beta}, X\right) \tag{3.1}$$

It is given by the function 1. Here, $\widetilde{Y}_i$, denotes the fuzzy dependent variable in the symmetric triangular property structure estimated and is shown as $\widetilde{Y}_i = (\widetilde{y}_c,\ \widetilde{e}_s)$, $\widetilde{y}_c$ denotes the mean value (center), and $\widetilde{e}_s$ denotes the spread value.

In the case of fuzzy observations, consider a fuzzy linear regression for crisp explanatory and fuzzy response observations as follows:

$$\widetilde{Y}_i = f\left(\widetilde{\beta}, X\right) = \widetilde{\beta}_0 + \widetilde{\beta}_1 X_{i1} + ... + \widetilde{\beta}_{p-1} X_{i(p-1)} = \widetilde{\beta}_0 + \sum_{i=1}^{n} \widetilde{\beta}_i X_i \tag{3.2.a}$$

(3.2.b) $\widetilde{Y}_i = \{c_0,\ s_0\} + \{c_1,\ s_1\}X_{i1} + \{c_2,\ s_2\}X_{i2} + ... + \{c_{p-1},\ s_{p-1}\}X_{i(p-1)}$

in which $\widetilde{\beta}_j = \left[\widetilde{\beta}_0\ and\ \ \widetilde{\beta}_1, \widetilde{\beta}_2, \widetilde{\beta}_3, ...\widetilde{\beta}_j......, \widetilde{\beta}_{p-1}\right]^t$ are the coefficient values of the independent variables in the function and it is a set of dependent and independent variables formed in the form of $\left\{Y_i,\ X_{i1}, X_{i2},\ X_{i3}, ..., X_{(p-1)n}\ \right\} = \{Y_i, X_{ij}\ \}$, and each dependent variable observation is expressed as $x\epsilon X\ (i = 1, ..., n,\quad j = 1, 2, ..., p-1)$. That is, they are crisp values of the explanatory variables. It is defined by $(i = 1, 2, 3, ..., n)$. In the fuzzy least squares regression model, the data of the dependent $\widetilde{Y}_i$ variable can be real numbers or fuzzy numbers. It is generally assumed that the data for the dependent Y_i variable are symmetrical fuzzy numbers of interval type [35].

$\widetilde{\beta}_j = \left[\widetilde{\beta}_0\ ve\ \ \widetilde{\beta}_1, \widetilde{\beta}_2, \widetilde{\beta}_3, ...\widetilde{\beta}_j......, \widetilde{\beta}_{p-1}\right]^t$ are fuzzy regression coefficients vectors with a symmetric triangular fuzzy number structure and they are fuzzy numbers in the form of $\widetilde{\beta}_j = (c_j,\ s_j)\widetilde{\beta}_j, (j : 0, 1, 2, 3, ..., p-1).c_j$, is the $\mu_{\widetilde{\beta}_i}(c_j) = 1$ value representing the midpoint of the coefficients, that is, the center value, and has the form $c_j = [c_1, c_2 c_3, ..., c_n]^t$. s_j, shows the spread of the coefficients belonging to the fuzzy regression analysis model and is $s_j = [s_1, s_2, s_3, ..., s_n\ ]^t$ shaped [64].

Each coefficient value $\widetilde{\beta}_i = \{c_j,\ s_j\} = \left\{\widetilde{\beta}_i : c_j - s_j \leq \widetilde{\beta}_i \leq c_j + s_j\right\}$ has a symmetric triangular property structure and is $\widetilde{\beta}_i(j : 0, 1, 2, 3, ..., p-1)$ [59].

The $\widetilde{\beta}_i = \{c_j,\ s_j\}$ value of the fuzzy coefficients was estimated by the minimum blur method proposed by Tanaka. The method is given in the following equation. In least squares regression analysis proposed by Tanaka and Watada (1988) [60], the linear programming (LP) formulation considers triangular membership functions (not necessarily symmetric). The spreads of the calculated fuzzy coefficients are calculated with the help of equation. The LP formulation is as follows (3):

$$\text{(3.3)} \qquad \min\ \ Z(x)\ \ s^t|X_i| = \min_{c,\,s}\left[s_0 + \sum_{j=0}^{n} s_j\ |X_{ij}|\ \right]$$

$$\min_{c,\,s} J = c_1, c_2, .., c_n\ ,\ \ c_j \geq 0,\ \forall i,\ \ i = 1, 2, .., m \text{ and}$$

$$\min_{c,\,s} J = s_1, s_2, ..., s_n,\ \ s_j \geq 0\ \ \forall i,\ \ j = 0, 1, 2, ..., n$$

$$\sum_{j=0}^{n} c_j X_{ij} + (1-h)\left[\sum_{j=0}^{n} s_j\ |X_{ij}|\ \right] \geq \widetilde{y}_c + (1-h)\widetilde{e}_s \quad \forall i,\ \ i = 1, 2, ..., n$$

$$\sum_{j=0}^{n} c_j X_{ij} - (1-h)\left[\sum_{j=0}^{n} s_j\ |X_{ij}|\ \right] \leq \widetilde{y}_c - (1-h)\widetilde{e}_s \quad \forall i,\ \ i = 1, 2, ..., n$$

Here $Z(x)$: function shows the total blur in the model. m: is the number of observations regarding the dependent variable. j : the number of independent variable x_{ij} : is the i-th observation value of the j-th independent variable. For each predicted $\widetilde{Y}_i$ observation value, the constraint number must be $2xn$ [43]. In order to minimize the total spread, the level h, $\widetilde{Y}_i$, the predictor of each observation value Y_i, is assumed to have a turbidity tolerance

$\mu_{\widetilde{Y}_i}(Y_i) \geq h \;\; i = 1, 2, .., m$ [30]. In Equation 3, the objective function is weighted with the absolute values of the measurements of the distributions of the independent variables. The application of bootstrap resampling technique in fuzzy least squares regression analysis is given below.

4. BOOTSTRAP FUZZY REGRESSION ANALYSIS

In this section, we introduce bootstrap resampling technique procedure. In general, regression technique for bootstrap is divided into two approaches: the first is based on the resampling observations approach and the second is based on the resampling errors. Bootstrap technique based on resampling errors is known as more suitable for the case of deterministic, whereas bootstrap resampling technique based on the drawing i.i.d. sample from the observations pairs is more appropriate for the case of random. However, bootstrap resampling technique pairs can also be used for deterministic [1]. The bootstrap is a "model-dependent" technique in terms of its implementation and performance although the bootstrap requires no theoretical formula for the quantity to be estimated and is less model-dependent than the traditional approach. In this paper, we use bootstrap technique based on the resampling errors. The bootstrap fuzzy regression analysis procedure is as follows:

Method: To describe the resampling methods we start with an n sized sample $w_i = (Y_i, X_{ji})'$ and assume that w_i's are drawn independently and identically from a distribution of F, where $Y_i = (y_1, y_2, .., y_n)'$ contains the responses, $X_{ji} = (x_{j1}, x_{j2}, ..., x_{jn})'$ is a matrix of dimension nxk, where $j = 1, 2, ...k, i = 1, 2, 3, ..., n$.

4.1. Bootstrapping Regression Algorithm

Here, two approaches for bootstrapping regression methods were given. The choice of either methods depends upon the regressors are fixed or random. If the regressors are fixed, the bootstrap uses resampling of the error term. If the regressors are random, the bootstrap uses resampling of observation sets w_i [55].

4.2. Bootstrap Based On The Resampling Errors

If the regressors are fixed, as in desing experiment, then the bootstrap resampling must preserve that structure. The bootstrap procedure based on the resampling errors as follows [55]:

$1^{(e)}$. Fit the full-sampling least-squares regression equation to estimate the regression coefficients of the model (6.a).

$2^{(e)}$. Calculate the e_i values $\left(e_i = Y_i - \hat{Y}_i\right)$

$3^{(e)}$. Draw an n sized bootstrap random sample with replacement $e_1^{(b)}, e_2^{(b)}, e_3^{(b)}, ..., e_n^{(b)}$ from the e_i values calculated in step $2^{(e)}$ giving $1/n$ probability each e_i values and Calculate the centered residual of $\bar{\tilde{e}}_i^{(b)}$ [42] [53] [65]:

$4^{(e)}$. Compute the bootstrap $Y_i^{(b)}$ values by adding resampled residuals onto the ordinary least squares regression fit, holding the regression desing fixed [16] [55].

$$Y_i^{(b)} = X\hat{\beta} + \bar{\tilde{e}}_i^{(b)} \tag{4.1}$$

$5^{(e)}$. Obtain least squares estimates from the 1-th bootstrap sample:

$$\widetilde{\beta}^{(b1)} = (X'X)^{-1}X'Y^{(b)} \quad \text{(we need } Y^*\text{)} \tag{4.2}$$

$$\widetilde{\beta}^{(b1)} = \hat{\beta}(X'X)^{-1}X'e^{(b)} \quad \text{(we do not need } Y^*\text{)} \tag{4.3}$$

$6^{(e)}$. Repeat steps $3^{(e)}, 4^{(e)}$ and $5^{(e)}$ for $r = 1, 2, ..., B$, and proceed as in resampling with random regressors $7^{(e)}$ and $8^{(e)}$.

An illustrative example that presents how the regression parameters are estimated from the bootstrap based on the resampling observations was given in Table 1.

By resampling residuals and randomly reattaching them to fitted values, the procedure implicitly assumes that the errors are identically distributed. Bootstrapping draws an analogy between the fitted value $\hat{Y}_i$ in the sample and Y in the population, and between the residual e_i in the sample and the error ε_i in the population [21]. In bootstrap resampling technique principle, the sample represents the population as the bootstrap samples. According to the weak law of large numbers, the empirical distribution function converges in probability to the true distribution function [42]. Note that define the bootstrap observation $Y_i^{(b)}$, by treating $\hat{\beta}$ as the "true" parameter and $e_i^{(b)}$ as the "population" of errors [54].

$7^{(e)}$. Obtain the probability distribution $\left(F\left(\widetilde{\beta}^{(b)}\right)\right)$ of bootstrap estimates $\widetilde{\beta}^{(b1)}$, $\widetilde{\beta}^{(b2)}, ..., \widetilde{\beta}^{(bB)}$ and use the $\left(F\left(\widetilde{\beta}^{(b)}\right)\right)$ to estimate regression coefficients, variances and confidence intervals as follows. The bootstrap estimate of regression coefficient is the mean of the distribution $\left(F\left(\widetilde{\beta}^{(b)}\right)\right)$ [25] [55].

$$\beta^{(b)} = \sum_{b=1}^{B} \beta^{(br)}/B = \beta^{(br)} \tag{4.4}$$

$8^{(e)}$. Thus, the bootstrap regression equation is

$$\min_{a_c, a_s} J1 = \begin{bmatrix} c0 \ + X_{i1} * c1 + X_{i2} * c2 + X_{i3} * c3 + X_{i4} * c4 - s0 - X_{i1} * s1 - X_{i2} * s2 - X_{i3} * s3 - X_{i4} * s4 <= Y_1 \\ c0 \ + X_{i1} * c1 + X_{i2} * c2 + X_{i3} * c3 + X_{i4} * c4 - s0 + X_{i1} * s1 + X_{i2} * s2 + X_{i3} * s3 + X_{i4} * s4 >= Y_1 \end{bmatrix}$$

$$\min_{a_c, a_s} Jn = \begin{bmatrix} c0 \ + X_{i150} * c1 + X_{i150} * c2 + X_{i150} * c3 + X_{i150} * c4 - s0 - X_{i150} * s1 - X_{i150} * s2 - X_{i150} * s3 - X_{i150} * s4 <= Y_1 \\ c0 \ + X_{i150} * c1 + X_{i150} * c2 + X_{i150} * c3 + X_{i150} * c4 - s0 + X_{i150} * s1 + X_{i150} * s2 + X_{i150} * s3 + X_{i150} * s4 >= Y_1 \end{bmatrix}$$

$$\widetilde{Y}_i = f\left(\widetilde{\beta}, x\right) = \widetilde{\beta}_0^{(b)} + \widetilde{\beta}_1^{(b)} x_{j1}^{(b)} + \widetilde{\beta}_2^{(b)} x_{j2}^{(b)} + ... + \widetilde{\beta}_s^{(b)} x_{jn}^{(b)} \tag{4.5}$$

where $\widetilde{\beta}_0^{(b)}$ is unbiased estimator of β [40].

4.3. The bootstrap bias, variance, confidence and percentile interval. The bootstrap bias equals,

$$\text{bias}_b = \widetilde{\beta}^{(b)} - \widetilde{\beta} \tag{4.6}$$

Further discussion are described in Efron and Tibshirani (1998) [23]. The bootstrap variance from the distribution $\left(F\left(\widetilde{\beta}^{(b)}\right)\right)$ are calculated by [53] [55].

$$\text{var}\left(\widetilde{\beta}^{(b)}\right) = \sum_{i=1}^{B}\left[\left(\widetilde{\beta}^{(br)} - \widetilde{\beta}^{(b)}\right)\left(\widetilde{\beta}^{(br)} - \widetilde{\beta}^{(b)}\right)'\right] \Big/ (B-1), \quad r = 1, 2, ..., B \tag{4.7}$$

The bootstrap confidence interval by normal approach is obtained by

$$\left(\widetilde{\beta}^{(b)} - t_{n-p,\ \frac{\alpha}{2}} * S_e\left(\widetilde{\beta}^{(b)}\right) < \beta < \widetilde{\beta}^{(b)} + t_{n-p,\ \frac{\alpha}{2}} * S_e(\widetilde{\beta}^{(b)})\right) = 1 - \alpha \tag{4.8}$$

where $t_{n-p,\ \frac{\alpha}{2}}$ is the critical value of t with probability $\alpha/2$ the right for $n-p$ degrees of freedom, and $S_e\left(\widetilde{\beta}^{(b)}\right)$ is the standard error of the $\widetilde{\beta}^{(b)}$. If the sample size is n $\geq$ 30, then Z distribution values are used instead of t in estimation of confidence intervals [22].

A non-parametric confidence interval named percentile Interval can be constructed from the quantiles of the bootstrap sampling distribution of $\widetilde{\beta}^{(b)}$. The $(\alpha/2)\%$ and $(1-\alpha/2)\%$ percentile interval is

$$\widetilde{\beta}^{(br)}_{(lower)} < \beta < \widetilde{\beta}^{(br)}_{(upper)} \tag{4.9}$$

where $\widetilde{\beta}^{(br)}$ is the ordered bootstrap estimates of regression coefficient from Equation 9 or 10, lower=$(\alpha/2)B$ and upper=$(1-\alpha/2)B$.

5. ILLUSTRATIVE EXAMPLE

Numerical examples are used to illustrate the fuzzy regression model that are summarized in previous sections. This example focuses on illustration and application of bootstrap technique in fuzzy regression analysis. Rapidly changing scientific and technological developments in recent years have negatively affected the health status of individuals by changing their nutritional habits. One of the main indicators of a healthy life is to have a stabile body composition. In recent years, along with the increasing prevalence of overweight and obesity worldwide, it has become even more crucial how to have a stabile body composition. In addition to overweight and obese individuals, it has become important to maintain the stability of body composition in the elderly, athletes and individuals with certain diseases. For such cases, anthropo-plyometric measurements can be used to evaluate the development-growth and nutritional status of individuals on body composition. In addition, the effects of dietary patterns of different diseases can be monitored and body composition can be determined.

In this study, in order to estimate total fat $(DEXATF)$ calculated according to DEXA method (Y) values with minimum error, Triceps values of independent variables such as Body Mass Index $(BMI)(kg/m^2)(X_1)$, age $(YEAR)$ (X_2) waist circumference fat percentage

$(WCFP)$ (X_3) were used as material in the model. These values were used in the classical bootstrap regression analysis method $(BOLSR)$ and fuzzy linear bootstrap regression analysis methods $(BFLSR)$, and the results were compared by calculating the estimated values of the coefficients and statistical values. The sample size was determined as 50 participants in order to determine whether more reliable results can be calculated in a shorter time with small data sets in cases where the constraints of the classical bootstrap regression analysis method cannot be met and the uncertainties in the datasets are not minimized.

The data used in the current study was obtained by the permissions of Drug Researches Local Ethics Committee of Erciyes University Faculty of Medicine (Date: 02.12.2008, Number: 2008/613) and Human Researches Ethics Committee of Kocaeli University (Date: 10.03.2009, Number: 2009/48) and the support of Scientific Research Project Coordination Unit of Erciyes University (Project code: TSY-09-772). The study was conducted in accordance with the principles of the Declaration of Helsinki. The study sample was consisted of randomly selected 137 voluntary participants, aged between 18-65 years and admitted to Kocaeli University, Faculty of Medicine, Department of Nuclear Medicine from May to July 2009. Of the participants, 67 (50%) were females and 67 (50%) were males, respectively. Women with pregnancy/suspected pregnancy and in the menstruation period, participants with metabolic and endocrine diseases and with any systemic diseases (liver, kidney, heart) and participants prescribed with hormonal drugs and anti-oedematous drugs were excluded.

The data pairs $w_i = (Y_i, X_{ji})'$ of Table 1 population, $(i = 1, ..., 50)$ are used to demonstrate the proposed procedure in case where the crisp input X and crisp output Yi .

Table 1: $n = 50$ volume original data set.

No	**DEXATF (Y)**	**BMI** (X_1)	**YEAR** (X_2)	**WCFP** (X_3)
1	31.20	28.30	44.00	34.82
2	26.50	21.30	26.00	44861
3	34.80	28.40	54.00	39.12
.	.	.	.	.
48	53.40	40.30	54.00	58.16
49	37.00	36.60	37.00	35.07
50	24.50	44859	24.00	19.07

DEXATF: total fat (DEXATF) calculated according to DEXA method; BMI: Body Mass Index (kg/m^2); YEAR; WCFP: Waist Circumference Fat Percentage.

The bootstrap algorithm based on error terms has been applied to the data in Table 1 as follows:

$1^{(e)}$. First, the ordinary least squares regression (OLSR) model was fitted to data given in and the results of the ordinary least squares regression was summarized in Table 2.

All of the regressions in Table 2 are significant (p<0.01) and the determination of coefficients $R^2 = 0.933$, respectively. The regression of total fat calculated according to DEXA method on the Body Mass Index (kg/m^2), YEAR and waist circumference Fat

Percentage is significant as result of variance analysis (P<0.01**). According to the t-tests for significance of regression coefficients, all of the regression coefficients are significant (P<0.01). Therefore, BOLSR can be substituted as an alternative modelling approach. The illustration of the bootstrap ($B = 1000$ bootstrap samples, each of size $n = 30$) regression procedure, from the data given in Table 1, calculation the bootstrap estimates of the regression parameters for each sample are shown in Table 3.

Table 2: The summary statistics of regression coefficients for OLS regression.

Variables	$\hat{\beta}$	S.E.($\hat{\beta}$)	t	Sig	95%Confidence Interval
Constant	−6.973	3.031	−2.300	.026	(−13.074)–(−0.871)
BMI (X_1)	0.984	0.149	6.612	.000	(0.685)–(1.284)
YEAR (X_2)	−0.111	0.061	−1.809	.07	(−2.234)–(0.013)
WCFP (X_3)	0.399	0.125	3.177	.003	(0.146)–(0.651)
R^2=0.933, N = 50, SSE = 3.720, F = 103.853**					

DEXATF: total fat (DEXATF) calculated according to DEXA method; BMI: Body Mass Index BMI: Body Mass Index (kg/m^2); YEAR: WCFP: Waist Circumference Fat Percentage; SSE: sum of squares of error.

$2^{(e)}$**.** The values in Table 3 are obtained by calculating the values of e_i with $e_i = Y_i - \hat{Y}_i$.

Table 3: Bootstrap residual instances created by assigning the probability 1/n to each e_i value.

No	Y_i	$\hat{Y}_i$	e_i	**1/50**	**r**	$e_1^{(b)}$	$e_2^{(b)}$	$e_3^{(b)}$	$e_4^{(b)}$	.	$e_{48}^{(b)}$	$e_{49}^{(b)}$	$e_{50}^{(b)}$	$\bar{e}_i^{(b)}$
1	31.20	29.88	1.32	0.03	**1**	0.02	−0.11	−0.08	0.03	.	−0.05	−0.02	−0.11	−0.07
2	26.50	21.91	4.59	0.09	**2**	−0.06	−0.05	−0.10	−0.05	.	−0.05	−0.08	−0.05	−0.04
3	34.80	30.59	4.21	0.08	**3**	0.11	0.07	−0.11	0.10	.	−0.12	−0.05	−0.05	−0.02
4	31.30	28.92	2.38	0.05	**4**	−0.04	−0.08	−0.05	−0.05	.	−0.01	0.02	−0.10	−0.03
5	20.40	26.02	−5.62	−0.11	**5**	−0.02	0.01	0.10	−0.02	.	0.07	0.07	−0.05	−0.06
6	30.30	29.17	1.13	0.02	**6**	0.06	0.09	−0.10	−0.05	.	0.05	−0.02	0.02	−0.04
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
48	53.40	49.89	3.51	0.07	**48**	−0.06	−0.02	−0.04	0.09	.	−0.02	−0.04	0.01	−0.02
49	37.00	38.93	−1.93	−0.04	**49**	0.07	−0.01	−0.01	0.05	.	0.02	0.07	0.11	0.00
50	24.50	22.67	1.83	0.04	**50**	−0.05	−0.05	−0.01	0.03	.	−0.11	−0.01	−0.08	0.00
					.	.	.	.	.	.	.	.	.	.
					997	−0.10	0.18	0.09	0.05	.	−0.13	−0.11	0.38	−0.01
					998	0.05	−0.10	−0.22	0.09	.	−0.11	−0.08	−0.02	−0.04
					999	0.18	0.07	−0.07	0.09	.	−0.02	−0.03	−0.03	−0.05
					1000	−0.10	0.18	0.09	0.05	.	−0.13	−0.11	0.38	0.00
$\bar{\hat{e}}_i^{(b)} = \frac{\sum_{b=1}^{1000} e_i^{(b)}}{1000}$						−0.001	0.000	−0.004	0.001		−0.002	0.000	−0.001	

$3^{(e)}$**.** Draw an n sized bootstrap random sample with replacement $e_1^{(b)}, e_2^{(b)}, e_3^{(b)}, ..., e_n^{(b)}$ from the e_i values calculated in step $2^{(e)}$ giving 1/n probability each e_i values.

$4^{(e)}$**.** Calculated the bootstrap Y_i^* values by adding resampled residuals onto the ordinary least squares regression fit, holding the regression design fixed (Table 4).

Table 4: Bootstrap $Y^{(b)}$ values calculated with resampled residuals.

No	$Y^{(b)}$	$\hat{\beta}_0$	**B**MI	$\hat{\beta}_1$	YEAR	$\hat{\beta}_2$	**W**CFP	$\hat{\beta}_3$	$\bar{\hat{e}}_i^{(b)}$
1	29.88	−6.973	28.30	0.984	44.00	−0.111	34.82	0.399	−0.001
2	21.92	−6.973	21.30	0.984	26.00	−0.111	27.10	0.399	0.000
3	0.58	−6.973	28.40	0.984	54.00	−0.111	39.12	0.399	−0.004
.	.	.	.	.	.	.	.	.	.
48	9.89	−6.973	40.30	0.984	54.00	−0.111	58.16	0.399	−0.002
49	38.93	−6.973	36.60	0.984	37.00	−0.111	35.07	0.399	0.000
50	22.67	−6.973	25.10	0.984	24.00	−0.111	19.07	0.399	−0.001

$5^{(e)}$**.** Obtain least squares estimates from the 1^{th} bootstrap sample:

$$Y^{(b)} = -6.973 + 0.984 * BMI - 0.111 * YEAR + 0.399 * WCFP + \bar{\hat{e}}_i^{(b)} \tag{5.1}$$

$6^{(e)}$**.** Repeat steps $3^{(e)}$, $4^{(e)}$ and $5^{(e)}$ for r =1,2,...,B, and proceed as in resampling with random regressors $7^{(e)}$ and $8^{(e)}$ (Table 5).

Table 5: Some bootstrap descriptive statistics based on the resampling of the ($n = 50$) error term of the data in Table 2.

Variables	Observed	$\hat{\beta}^*_{ort}$	$S_e(\hat{\beta}^*)$	$SS(\hat{\beta}^*)$	**Confidence intervals**	
					95% Confidence Interval	
Constant	−6.973	−69.590	31.809	0.0133	−12.977	−13.997
BMI(X_1)	0.984	0.9862	0.1246	−0.0020	0.802	12.145
YEAR(X_2)	−0.111	−0.1129	0.0641	−0.0020	−0.224	−0.0050
WCFP(X_3)	0.399	0.3993	0.0974	0.0005	0.250	0.6100

DEXATF: total fat (DEXATF) calculated according to DEXA method; BMI: Body Mass Index BMI: Body Mass Index (kg/m^2); YEAR; WCFP: Waist Circumference Fat Percentage.

$7^{(e)}$**.** By using the new observation points that have been formed, the parameters are estimated with the FLS regression analysis method:

MIN = 50*s0+1488.2*s1+1897*s2+1913*s3;

$$\min_{a_c,\,a_s} J1 = \begin{bmatrix} c0 \ + 28.30*c1+44*c2+34.82*c3*0.5*s0+28.30*0.5*s1+44*0.5*s2+34.82*0.5*s3 \geq = 29.88; \\ \\ c0 \ + 28.30*c1+44*c2+34.82*c3-0.5*s0-28.30*0.5*s1-44*0.5*s2-34.82*0.5*s3 >= 29.88; \end{bmatrix}$$

(5.2)

$$\min_{a_c,\,a_s} J1 = \begin{bmatrix} c0 \ + 25.10*c1+24*c2+19.07*c3*0.5*s0+25.10*0.5*s1+24*0.5*s2+19.07*0.5*s3 \geq \ 22.67; \\ \\ c0 \ + 25.10*c1+24*c2+19.07*c3-0.5*s0-25.10*0.5*s1-24*0.5*s2-19.07*0.5*s3 \leq \ 22.67; \end{bmatrix}$$

$FREE(c0);\ FREE(c1);\ FREE(c2);\ FREE(c3);\ END$

(5.3) $\widetilde{Y}_i = (24.461; 12.40) + (0.116; 0.0) * BMI + (-0.197; 0.0) * YEAR + (0.369; 0.0) * WCFP$

The data pairs $w_i = (\widetilde{Y}_i, X_{ji})'$ of Table 6, (i = 1,...,50) are used to demonstrate the proposed procedure in case where the crisp input X_{ji}, risp output $\widetilde{Y}_i$ and fuzzy regression coefficients.

Table 6: Some Bootstrap Fuzzy Descriptor Statistics Based on the Resampling of the Error Term Belonging to the Data in Table 1 ($n = 50$).

Variables	Observed	c_j	s_j	**Confidence intervals**	
				95%Confidence Interval	
Constant	−6.973	24.461	12.408	−6.988	−6.964
BMI(X_1)	0.984	0.116	0.00	0.983	0.984
YEAR(X_2)	−0.111	−0.197	0.00	−0.111	0.111
WCFP(X_3)	0.399	0.369	0.00	0.399	0.400
$R^2 = 1.0$, N = 50, SSE = 0,0075, F = 25142495,53**					

DEXATF: total fat (DEXATF) calculated according to DEXA method; BMI: Body Mass Index BMI: Body Mass Index (kg/m^2); YEAR; WCFP: Waist Circumference Fat Percentage.

Estimates of the bootstrap regression coefficients in the form of were calculated. Also, this model explains response variable using fewer variables although there is no procedure available in FLR which can be used as variable selection method.

6. DISCUSSION AND CONCLUSIONS

In this study, using the samples obtained by bootstrap resampling technique in ordinary least squares and fuzzy least squares regression analysis techniques, it has been tried to reveal which of them is more effective to estimate parameters.

Parameter estimates as well as their standard errors and confidence intervals statistics from bootstrapping ordinary least squares regression and bootstrapping fuzzy regression coefficients are presented in Table 7 for prediction of DEXATF (Y) (gr).

Table 7: BOLS and BFLS regression ($n = 50$, $B = 1000$) parameter estimations and the regression coefficients statistics for estimation of some DEXATF (gr).

	Variables	**Observed**	**Average**	**S.E.**	**Bias**	**Confidence intervals**	
						95% Confidence Interval	
BOLSR	Constant	−6.973	−6.959	31.809	0.0133	−12.977	−13.997
	BMI(X_1)	0.984	0.9862	0.1246	−0.0020	0.802	12.145
	YEAR(X_2)	−0.111	−0.1129	0.0641	−0.0020	−0.224	−0.005
	WCFP(X_3)	0.399	0.3993	0.0974	0.0005	0.250	0.610
BFLSR	Constant	−6.973	24.461		12.408	−6.988	−6.964
	BMI(X_1)	0.984	0.116	0.00	0.00	0.983	0.984
	YEAR(X_2)	−0.111	−0.197	0.00	0.00	−0.111	0.111
	WCFP(X_3)	0.399	0.369	0.00	0.00	0.399	0.400

DEXATF: total fat (DEXATF) calculated according to DEXA method; BMI: Body Mass Index (kg/m^2); YEAR; WCFP: Waist Circumference Fat Percentage

$B = 10000$ bootstrap samples are generated randomly to reflect the exact behavior of the bootstrap procedure and the distributions of bootstrap regression parameter estimations ($\widetilde{\beta}^{(b)}$) are graphed in Figure 2(a), 2(b), 2(c) (Figures 1). The histograms of the bootstrap estimates conform quite well to the limiting normal distribution for all regression coefficients. Hence, the confidence intervals should be based on that distribution, where B is sufficiently large ($B = 1000$). And, bootstrap fuzzy regressions are generated by putting each one of the observation sets in place in the model and the regression coefficients are estimated as $\widetilde{\beta}^{(b)}$. To reflect the exact behavior of the bootstrap sample procedure the distributions of fuzzy regression parameter estimations $\widetilde{\beta}^{(b)}$ are graphed in Figures 2(d) (Figures 1). The histograms of the bootstrap fuzzy estimates are no similar to the normal distribution for bootstrap OLS regression coefficients.

The fuzzy bootstrap regression standard errors of the BMI and $YEAR$ coefficients are substantially small than the estimated asymptotic OLS and bootstrap OLS standard errors, because of the inadequacy of the bootstrap in small samples. The confidence intervals based on the bootstrap fuzzy regression standard errors are very similar to the percentile intervals of the BMI and GS coefficients; however, the confidence intervals based on the OLS and bootstrap OLS standard errors are quite different from the percentile and confidence intervals based on the bootstrap standard errors. Comparing the bootstrap fuzzy coefficients averages $\widetilde{\widetilde{\beta}}_0^{(br)}$, $\widetilde{\widetilde{\beta}}_1^{(br)}$ and $\widetilde{\widetilde{\beta}}_2^{(br)}$ with the corresponding OLS and bootstrap OLS estimates $\hat{\beta}_0^{(br)}$, $\hat{\beta}_1^{(br)}$, $\hat{\beta}_2^{(br)}$ and $\widetilde{\beta}_0^{(br)}$, $\widetilde{\beta}_1^{(br)}$, $\widetilde{\beta}_2^{(br)}$ and β shows that there is a little bias in the bootstrap coefficients.

The shape of these graphs show that a histogram of the replicates with an overlaid smooth density estimate and the skewness of the distribution of regression parameter estimate from the OLS bootstrap and fuzzy bootstrap replicate. The shape of these graphs show that a histogram of the replicates with an overlaid smooth density estimate and the skewness of the distribution of regression parameter estimate from the OLS bootstrap and fuzzy bootstrap replicate.

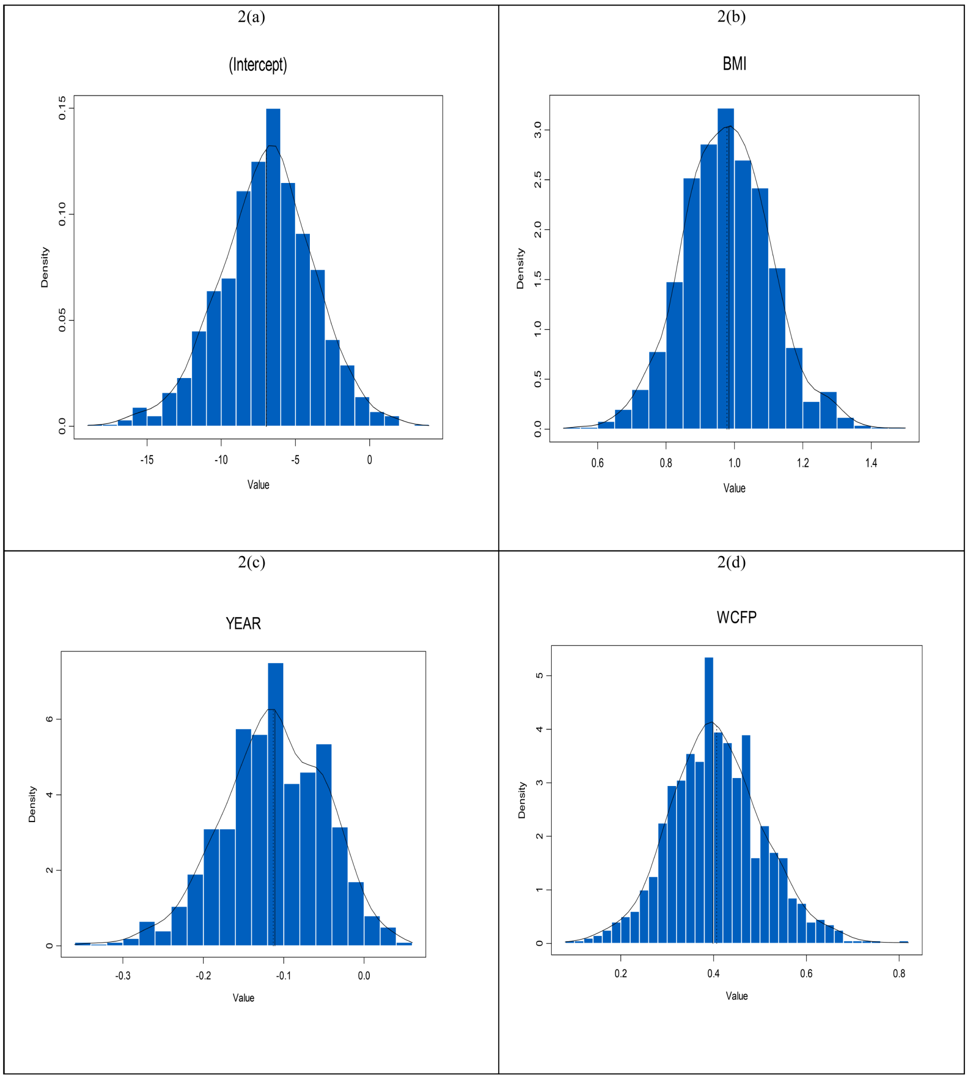

Figure 1: Histogram of bootstrap ($B = 1000$, (a), (b), (c), (d)) regression parameter estimates.

To examine the appearence of the distribution $\left(F\left(\beta^{(b)}\right)\right)$ of the replicates (B=10 000), the distribution plots of $\widetilde{\beta}^{(b)}$ from Equation 4.2 are given in Figures 1. The vertical lines of these plots give the mean of the B bootstrap parameter estimates ($\widetilde{\beta}^{(b)}$) and show the shape of distribution of bootstrap parameter estimates. Although, the larger bootstrap replicates (B) are used, the smoother distribution of $\widetilde{\beta}^{(b)}$ could usually be obtained in these plots (Fox, 1997) [25]. The number of bootstrap replications B depends on the application and size of sample. The bootstrap replications sufficient to be $B = 100$ for standard error estimates, for confidence interval estimates B≅1000, and for standard deviation estimate 50≤ B ≤ 100 were suggested by Leger *et al.* (1992) [40] and Efron (1990) [22]. In fact, it is known from the statistical theory of the bootstrap that a finite total of n^n possible bootstrap samples exist. If it was computed the parameter estimates for each of these n^n samples, it would obtain the true bootstrap estimates of parameters however, such extreme computation is wasteful and unnecessary [53].

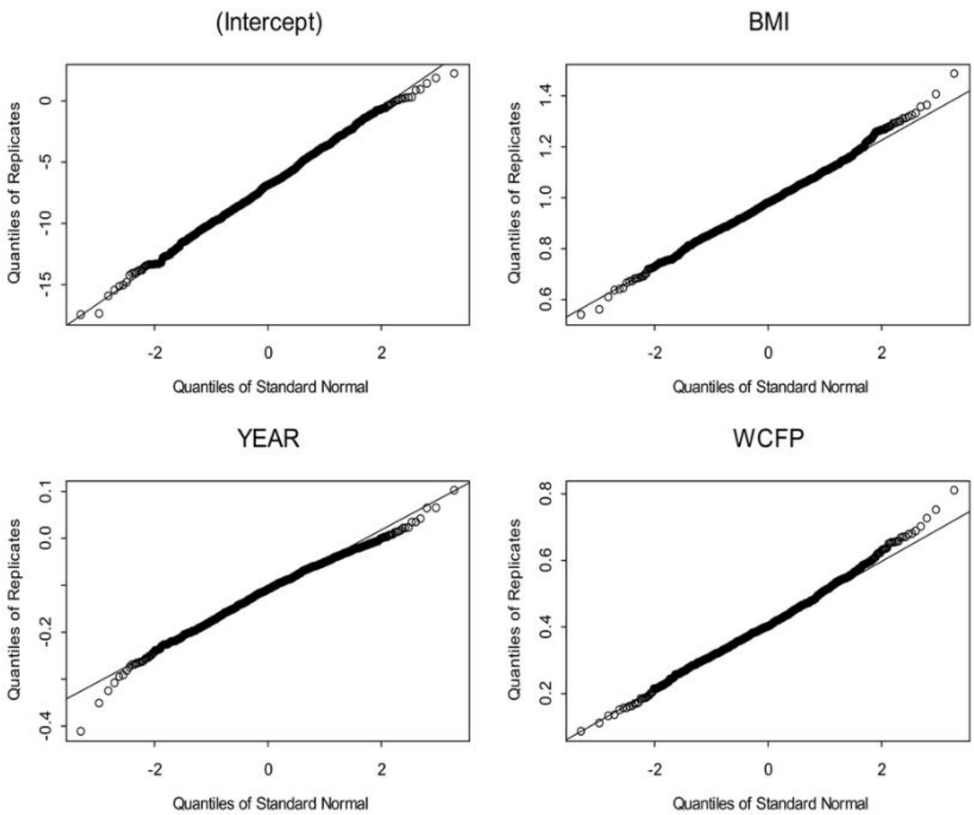

Figure 2: Normal Quantile – Quantile.

The bootstrap resampling technique is one of the most important concepts in statistics introduced. In classical techniques, the bootstrap resampling technique has become a very powerful tool used to estimate quantities associated with the sampling distribution of estimators and test statistics. In application of bootstrap resampling technique, there is often some uncertainty about the certain error structure, and a well-chosen resampling technique can give robust inferences to the certain error structure of the data. Indeed, it is harmful to pretend that mere calculation can replace thought about central issues such as the structure of a problem, the type of answer required, the sampling design and data quality. In these cases, for linear regression with normal random errors εj having constant variance, the least squares theory of regression estimation and inference provides clean, exact and optimal methods for analysis.

For generalizations to non-normal errors and non-constant variance, precise methods seldom exist, and we are faced with approximate techniques based on linear approximations to estimators and central limit theorems. Bootstrap resampling technique have the potential to provide more accurate and more valid analysis for modelling in complex problems. With ordinary least squares linear regression, in ideal conditions resampling essentially not only reproduces the exact theoretical analysis, but also offers the potential to deal with non-ideal circumstances such as non-constant variance. Despite its extent and usefulness, resampling technique should be carefully applied. Unless certain basic ideas are understood, it is all too easy to produce a wrong solution to the problem. Bootstrap resampling techniques are intended to help avoid tedious calculations based on questionable assumptions.

In conclusion, in this study it is aimed to describe: basic ideas, the standard error of bootstrap fuzzy regression technique, confidence intervals of the regression coefficients, application to bootstrap ordinary least squares technique and bootstrap fuzzy linear regression technique. The fuzzy regression technique is a new statistical technique that combines the classical regression technique with the theory of fuzzy logic. When functional relationship is not known in advance, fuzzy regression technique is introduced as an alternative technique which helps model crisp/crisp or crisp/fuzzy data. Also, the correct functional relationship between dependent variable and independent variable is not known. Bootstrap resampling technique is preferable in fuzzy least squares regression analysis and ordinary least squares regression techniques because of some theoretical properties like having any distributional assumptions on the residuals and hence, allows for inference even if the errors do not follow normal distribution.

The most important advantages of the bootstrap fuzzy least squares regression technique and bootstrap ordinary least squares technique are:

- they need smallers sample than ordinary least squares technique,
- they can be used when there are doubts about the distribution of the population,
- they can be used in cases of insufficient sample size and parametric assumptions are not realized,
- they can also be used in cases where the sample selection is not random,
- in cases of very large sample sizes, the methods can be applied by creating subgroups.

The bootstrap fuzzy least squares regression and bootstrap ordinary least squares techniques estimate the variation of a statistic from the variation of that statistic between subsamples, rather than from parametric assumptions and may yield similar results in many situations. However, it is a mistake to expect that bootstrap fuzzy least squares regression technique and bootstrap ordinary least squares regression technique always give valid and confident results. The confidence of results depend on the structure of the data and distribution function. Application of both regression techniques depend on development of statistical computer packages featured these analyses.

The estimations of the bootstrap standard error and confidence intervals of the regression coefficients are nearly equal to the standard error of regression coefficient estimates of the bootstrap fuzzy least squares regression technique. However, bootstrap fuzzy least squares regression technique gives regression coefficients, which generally have smaller standard errors and narrow confidence intervals than bootstrap regression technique. If the OLSR model did not satisfy the related model assumptions, the bootstrap regression technique and bootstrap fuzzy least squares regression techniques could be used for fitting the model and provide better estimates. Because the bootstrap resampling technique and bootstrap fuzzy least squares regression technique do not require above assumptions [7], [27]. Due to the computation of the standard errors and since confidence intervals are based on the distribution of bootstrap samples, not on assumptions about normal distributions. The assumption guessed behind bootstrap resampling technique is to treat the sample as if it were the entire population [4], [7].

In this research, we have presented model of fuzzy least squares regression for the literature. We have shown that the development of an adequate bootstrap resampling theory in the fuzzy context would be very profitable because in this context the asymptotic approximations are, in most cases, difficult to handle and hence, they are useless to make inferences. A real application to predict $DEXATF(Y)(g)$ in clinical data obtained was shown. $BOLS$ and $BFLS$ regression were obtained and also as can be seen from the statistical values calculated from a clinical numerical sample, the error of the $BFLS$ method regarding the estimates calculated according to the error criteria was detected to be lower than the errors calculated from the $BOLS$ method. Due to these results, we trust the results obtained with the $BFLSR$ method more than the results obtained with the $BOLSR$ method. It can be concluded that $BOLSR$ and $BFLSR$ methods have similar performance. Among these models, the $BFLSR$ method is proposed to be preferred. Although the bootstrap resampling technique is sometimes mentioned as a replacement for "standard statistics techniques", it is concluded that this thought is wrong, since the bootstrap resampling technique depends on the theoric elements of classic logic.

ACKNOWLEDGMENTS

The author is thankful to the referee and the editor for their constructive comments that led to a significant improvement of the paper.

ETHICS COMMITTEE APPROVAL

The study was conducted in accordance with the principles of the Declaration of Helsinki. The research permission was obtained from Erciyes University Ethics Committee (dated 12/02/2008 and number 2008/613) and Kocaeli University Ethics Committee (Date: 10.03.2009, Number: 2009/48). Informed consent was obtained from study participants.

DATA AVAILABILITY

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

REFERENCES

[1] Akbari, M.G. and Rezaei, A. (2009). Bootstrap statistical inference for the variance based on fuzzy data, *Austrian Journal of Statistics*, **38**(2), 121–130.

[2] Akbari, H. and Matthews, H.D. (2012). Global cooling updates: Reflective roofs and pavements, *Energy and Buildings*, **55**, 2–6.

[3] Akbari, M. and Hesamian, H. (2019). A partial-robust-ridge-based regression model with fuzzy predictors-responses, *Journal of Computational and Applied Mathematics*, **351**, 290–301.

[4] Ariani, D.; Yuki, N.N. and Yuniarti, D. (2017). Perbandingan Metode Bootstrap Dan Jackknife Resampling Dalam Menentukan Nilai Estimasi Dan Interval Konfidensi Parameter Regresi., *EKSPONENSIAL*, **8**(1), 43–49.

[5] Arefi, M. (2020). Quantile fuzzy regression based on fuzzy outputs and fuzzy parameters, *Soft Computing*, **24**(1), 311–320.

[6] Asadolahi, M.; Akbari, M.; Hesamian, G. and Arefi, M. (2021). A robust support vector regression with exact predictors and fuzzy responses, *International Journal of Approximate Reasoning*, **132**, 206–225.

[7] Beran, R. (2003). The Impact of the bootstrap on statistical algorithms and theory, *Statistical Science*, **18**(2), 175–184

[8] Celmins, A. (1987). Least squares model fitting to fuzzy vector data, *Fuzzy Sets and Systems*, **22**(3), 245–269.

[9] Chang, Y.H.O. and Ayyub, B.M. (2001). Fuzyy regression methotds-a comparative assessment, *Fuzzy Sets and Systems*, **119**(2), 187–203.

[10] Chen, Y.S. (2001). Outliers detection and confidence interval modification in fuzzy regression, *Fuzzy Sets and Systems*, **119**(2), 259–272.

[11] Casella, G. (2003). Introduction to the silver anniversary of the bootstrap, *Statistical Science*, **18**(2), 133–134.

[12] Chachi, J. (2019). A weighted least squares fuzzy regression for crisp input-fuzzy output data, *IEEE Transactions on Fuzzy Systems*, **27**(4), 739–748.

[13] Chachi, J.; Kazemifard, A. and Jalalvand, M.J. (2021). A multi-attribute assessment of fuzzy regression models, *Iranian Journal of Fuzzy Systems*, **18**(4), 131–148.

[14] Chachi, J. and Chaji, A. (2021). An OWA-based approach to quantile fuzzy regression, *Computers & Industrial Engineering*, **159**, 107498.

[15] Chukhrova, N. and Johannssen, A. (2019). Fuzzy regression analysis: Systematic review and bibliography, *Applied Soft Computing*, **84**, 105708.

[16] Diciccio, T. and Tibshirani, R. (1987). Bootstrap confidence intervals and bootstrap approximations, *Journal of the American Statistical Association*, **82**(397), 163–170.

[17] Diamond, P. (1987). *Least squares fitting of several fuzzy variables*, Preprints of Second IFSA World Congress, Tokyo, Japan, pp. 329–331.

[18] Diamond, P. (1988). Fuzzy least squares, *Information Science*, **46**(3), 141–157.

[19] D'Urso, P. and Massari, R. (2013). Weighted least squares and least median squares estimation for the fuzzy linear regression analysis, *Metron*, **71**(3), 279–306.

[20] Efron, B. (1979). Bootstrap methods: Anoher look at the jackknife, *The Annals of Statistics*, **7**, 1–26.

[21] Efron, B. (1982). The jackknife, the bootstrap and other resampling plans, *Society For Industrial And Applied Mathematics*, **38**, 29–35.

[22] Efron, B. (1990). More efficient bootstrap computations, *Journal of the American Statistical Association*, **85**(409), 79–89.

[23] Efron, B. and Tibshirani, R.J. (1994). *An Introduction to the Bootstrap*, CRC Press, Boca Raton, Florida, pp. 60–82.

[24] Efron, B. (2003). Second thought on bootstrapping, *Statistical Science*, **18**(2), 135–140.

[25] Fox, J. (1997). *Applied Regression Analysis, Linear Models, and Related Methods*, Sage Publications, Thousand Oaks, CA.

[26] Friedl, H. and Stampfer, E. (2002). *Resampling Methods*, Encyclopedia of Environmetrics, 3 (A. El-Shaarawi, W. Piegorsch, Eds.), Wiley, Chichester, 1768–1770.

[27] Ferraro, M.; Coppi, R.; Rodriguez, G.G. and Colubi, A. (2010). A linear regression model for imprecise response, *International Journal of Approximate Reasoning*, **51**(7), 759–770.

[28] Ferraro, M. and Giordani, P. (2013). A proposal of robust regression for random fuzzy sets, *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*, Springer, Berlin, Heidelberg, 115–123.

[29] Hong, D.H.; Song, J.K. and Young, H. (2001). Fuzzy least-squares linear regression analysis using shape preserving operations, *Information Sciences*, **138**, 185–193.

[30] Hojati, M.; Bector, C.R. and Smimou, K.A. (2005). Simple method for computation of fuzzy linear regression, *European Journal of Operational Research*, **166**, 172–184.

[31] Hung, W.L. and Yang, M.S. (2019). n omission approach for detecting outliers in fuzzy regressions models, *Fuzzy Sets and Systems*, **157**, 3109–3122.

[32] Hesamian, G. and Akbari, M.G. (2021). A robust multiple regression model based on fuzzy random variables, *Journal of Computational and Applied Mathematics*, **388**, 1–13.

[33] Hesamian, G. and Akbari, M.G. (2020). A robust varying coefficient approach to fuzzy multiple regression model, *Journal of Computational and Applied Mathematics*, **375**, 113270.

[34] Jimenez-Gamero, M.D.; Pino-Mejias, R. and Rojas-Medar, M.A. (2005). A bootstrap test for the expectation of fuzzy random variables, *UNICAMP*.

[35] Kim, B. and Bishu, R.R. (1998). Evaluation of fuzzy linear regression models by comparing membership functions. *Fuzzy Sets and Systems*, **100**(1–3), 343–352.

[36] Kula, K. and Apaydin, A. (2008). Fuzzy robust regression analysis based on the ranking of fuzzy sets, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **22**(3), 245–269.

[37] Khammar, A.H.; Arefi, M. and Akbari, M.G. (2020). A robust least-squares fuzzy regression model based on kernel function, *Iranian Journal of Fuzzy Systems*, **17**(4), 105–119.

[38] Khammar, A.H.; Arefi, M. and Akbari, M.G. (2021). A general approach to fuzzy regression models based on different loss functions, *Soft Computing*, **25**(2), 835–849.

[39] Khammar, A.H.; Arefi, M. and Akbari, M.G. (2021). Quantile fuzzy varying coefficient regression based on kernel function, *Soft Computing*, **107**, 107313.

[40] Leger, C.; Politis, D.N. and Romano, J.P. (1992). Bootstrap technology and applications, *Technometrics*, **34**(4), 378–397.

[41] Leski, J.M. and Kotas, M. (2015). The statistical inferences of fuzzy regression based on bootstrap techniques, *Fuzzy Sets and Systems*, **279**, 112–129.

[42] Lee, W.J.; Jung, H.Y.; Yoon, J.H. and CHOI, S.H. (2015). A partial-robust-ridge-based regression model with fuzzy predictors-responses, *Soft Computing*, **19**, 883–890.

[43] Moskowitz, H. and Kim, K. (1993). On assesing the H value in fuzzy linear regression, *Fuzzy Sets and Systems*, **58**(3), 303–327.

[44] Mooney, C.Z. and Duval, R.D. (1993). Bootstrapping: A nonparametric approach to statistical inference, *SAGE Pub.*, London, 80.

[45] Ming, M.; Friedman, M. and Kandel, A. (1997). General fuzzy least squares, *Fuzzy Sets and Systems*, **88**, 107–118.

[46] Montenegro, M.; Colubi, A.; Rosa Casals, M. and Gil, M.A. (2004). Asymptotic and bootstrap techniques for testing the expected value of a fuzzy random variable, *Metrika*, **59**(1), 31–49.

[47] Mohammed, Z.R. (2008). Bootstrapping: A nonparametric approach to identify the effect of sparsity of data in the binary regression models, *Journal of Applied Sciences, Network for Scientific Information, Asian*, **17**(8), 2991–2997.

[48] Nasrabadi, M.M. and Nasrabadi, E.A. (2004). Mathematical-programming approach to fuzzy linear regression analysis, *Applied Mathematics and Computation*, **155**(3), 873–881.

[49] Nasrabadi, E.; Hashemi, S.M. and Ghatee, M. (2007). An LP-based approach to outliers detection in fuzzy regression analysis. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **15**, 441–456.

[50] Nasrabadi, E. and Hashemi, S.M. (2008). Robust fuzzy regression analysis using neural networks, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **16**, 597–598.

[51] Peters, G. (1994). Fuzzy linear regression with fuzzy intervals, *Fuzzy Sets and Systems*, **63**, 45–55.

[52] Sadeghpour Gildeh, B. and Rahimpour, S. (2013). Fuzzy bootstrap test for the mean and variance with Dp, q-distance, *International Journal of Machine Learning and Computing*, **3**, 21–23.

[53] Stine, R. (1990). Modern methods of data analysis, *SAGE Pub.*, Scotland, 325–373.

[54] Shao, J. (1996). Bootstrap model selection, *Journal of the American statistical Association*, **91**(434), 655–665.

[55] Sahinler, S. and Topuz, D. (2007). Bootstrap and Jackknife resampling algorithms for estimation of regression parameters, *Journal of Applied Quantitative Methods*, **2**, 188–199.

[56] Shakouri, H.; Nadimi, R. and Ghaderi, F. (2009). A hybrid TSK-FR model to study short-term variations of the electricity demand versus the temperature changes, *Expert Systems with Applications*, **36**, 1765–1772.

[57] Shakouri, H. and Nadimi, R. (2013). Outlier detection in fuzzy linear regression with crisp input-output by linguistic variable view, *Applied Soft Computing*, **13**, 734–742.

[58] Sungkono, J. (2015). Bootstrap resampling observasi pada estimasi parameter regresi menggunakan Software R, *MAGISTRA*, **27**, 92.

[59] Tanaka, H.; Uejima, S. and Asai, K. (1982). Linear regression analysis with fuzzy model, IEEE Transactions on Systems, *Man and Cyberneties*, **12**(6), 903–907.

[60] Tanaka, H. and Watada, J. (1988). Possibilistic linear systems and their application to the linear regression model, *Fuzzy Sets and Systems*, **27**(3), 275–289.

[61] Tanaka, H.; Hayashi, I. and Watada, J. (1989). Possibilistic linear regression analysis for fuzzy data *European Journal of Operational Research*, **40**, 389–396.

[62] Taheri, S.M. and Chachi, J. (2020). A robust variable-spread fuzzy regression model. In "Recent Developments and the New Direction in Soft-Computing Foundations and Applications", Springer, Cham. 309–320.

[63] Varga, S. (2007). Robust estimations in classical regression models versus robust estimations in fuzzy regression models, *Kybernetika*, **43**, 503–508.

[64] Wang, H.F. and Tsaur, R.C. (2000). Insight of a fuzzy regression model, *Fuzzy Sets and Systems*, **112**, 355–369.

[65] Wu, C.F.J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis, *Annals of Statistics*, **14**, 1261–1295.

[66] Yang, Z.; Yin, Y. and Chen, Y. (2013). Robust fuzzy varying coefficient regression analysis with crisp inputs and gaussian fuzzy output, *Journal of Computing Science and Engineering*, **7**, 263–271.