GALAXY INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (GIIRJ)
ISSN (E): 2347-6915
Vol. 12, Issue 01, January (2024)

USING SIMULATION TO ESTIMATE A FUZZY REGRESSION MODEL

Fatima Othman Eatiah Al-Abadi
Master's Student at College of Administration and Economics, Basrah University, Iraq
Email: fothman996@gmail.com

Prof. Dr. Sahera Hussein Zain Al-Thalabi
Administration and Economics College, Basrah University, Iraq
Email: sahera.zain@uobasrah.edu.iq

ABSTRACT

The researcher faces a lot of problems when testing the accuracy of the model to estimate the parameters of the fuzzy regression model, and to remedy this problem, the prediction error was reduced by generating variables that follow a normal distribution using the most famous and common method, which is the (Box-Muller) method, which depends on the method of generating random variables that follow the standard uniform distribution U(0,1), and then these variables are converted into independent random variables that follow the standard normal distribution to estimate the parameters of the model and with the aim of reducing the prediction error between the expected and actual concentrations This indicates model accuracy and model blur that represents uncertainty in model predictions. The lower these values, the better the model performs in terms of accuracy and reliability.

Keywords: simulation concept, fuzzy regression,, Box-Muller method.

INTRODUCTION

The researcher faces several problems when estimating statistical regression functions, for example, when the model is undefined, the relationship between the parameters of the model is ambiguous, the data is hierarchically, our problem does not meet the probability regression assumptions (i.e., the regression coefficients must be constant), or when the available data is small, so fuzzy regression is used instead of statistical regression. In real life, we often face processes and problems that are difficult to solve or cannot be solved mathematically, because they are complex to change because it is not possible to control the variables involved and affecting the problem in question, and there are statistical theories that are complex to understand and not easy to analyze logically using mathematical proof. [3],[6]

From here comes the role of the simulation method, which describes these processes in a similar way to the real image with certain models, understanding the model achieves us a degree of awareness of the original process or real reality through simulating the model and that the degree of similarity between any simulation experience and the real reality depends on the extent to which the simulation model is identical or similar to the real system, and simulation experiments are characterized by shortening the time when implementing operations on the electronic calculator and the flexibility of choosing the sizes of different samples, as well as assuming levels of variation for random errors.

Some of what has been studied in previous years on the subject of the study will be addressed, according to the researcher's knowledge of it:

- A study (Elias and Sabbagh, 2006) dealt with fuzzy linear regression in the event that the data is fuzzy and unfuzzy and analyzed by several methods, namely fuzzy and modified linear regression Tanaka Fuzzy least squares With the employment of linear programming in the analysis, the study was applied in the medical field with real data on osteoporosis, the results proved that the use of the modified Tanka model is better than the fuzzy linear regression model to avoid the appearance of the estimated parameters are not fuzzy, while the fuzzy least squares method is better than Fuzzy linear regression method through the results of the model's degree of affiliation. [7]
- The study (Karakasidis and other,2012) dealt with the analysis of foggy regression and ambiguous similarity to study the relationship between different hardness measures and tensile strength of materials, and from the hardness measures that were studied with the tensile strength are (Vickers) and (Brinell) and (Rockwell C) and (Shore Scleroscope hardness scales) then the resulting foggy regression models showed a linear relationship, and the results also showed that there is a similarity in the foggy regression that reflects the basic similarity of the physical quantities of materials whose hardness was measured, and when comparing the results with the results of the analysis Conventional linear regression The comparison showed that the ambiguities and blurs in the model are better reflected in the case of fuzzy regression models. [8]
- The study (Charfeddine and other, 2014) aimed to apply the proposed fuzzy linear regression to estimate air transport demand and build trapezoidal foggy sets for the estimated variable and take into account all data samples, as the demand for the air transport market (frequency of flights, number of seats, flight price), that the effects of foggy factors to predict this demand in the classical way turned out to be very dangerous, and the study showed how regressions based on fuzzy logic that combines statistics and expert trends can be used to improve the estimation of transport demand. Aerial by relying on the Tanaka method. [4]
- A study (Abbas and Al-Metwally, 2021) used the formulation of the fuzzy relationship between the Iraqi GDP and the dollar exchange rate based on a fuzzy linear regression model with fuzzy inputs and outputs, and then using the fuzzy least squares method in estimating the impact of the dollar exchange rate (explanatory variable) on GDP (response variable), as the observations of both variables were represented by trigonometric fuzzy numbers, and the results reached the possibility of applying the fuzzy least squares method in estimating the fuzzy linear regression model with inputs and outputs Blur in a way that describes and represents the relationship well. [1]

Fuzzy linear regression model FLR [9]

It is one of the statistical methods used to represent the relationship between two or more phenomena when these phenomena are characterized by inaccuracy or reliability of their data, and fuzzy regression is also used to estimate the functional relationship between the response variable and the explanatory variable in a fuzzy ocean with a linear function.

Fuzzy linear regression aims to model an inaccurate or ambiguous phenomenon using fuzzy model parameters, and if the hypotheses of the ordinary least squares method are realized, the fuzzy regression is more effective and more flexible for different problems as an alternative to classical regression.

Fuzzy aggregate theory is the appropriate and effective means of formulating statistical models where it is the preeminent tool for addressing inaccuracies or ambiguities when observations are blurry. Ambiguity (uncertainty) in model companion regression models is caused by blur rather than randomness or both (randomness and blur). The formula for the general linear model of fuzzy linear regression is as follows:[10]

$$Y_i = \tilde{\beta}_0 + \tilde{\beta}_1\tilde{x}_1 + \cdots + \tilde{\beta}_n\tilde{x}_n + \varepsilon_i \quad \ldots \quad (1)$$

Whereas:

$\tilde{y}_i$ The fuzzy response variable represents

$\tilde{\beta}_0, \ldots, \tilde{\beta}_n$ Represents fuzzy regression parameters.

$\tilde{x}_1, \ldots, \tilde{x}_n$ Illustrative variables represent fuzzy.

ε_i represents a random error that is distributed normally.

The fuzzy regression according to the methods of estimation is of two types:[2]

1. The possible regression depending on the existence of the object and uses mathematical programming in estimating
2. Possible regression The fuzzy least squares (FLSLR) method is used in estimating

The impetus for the development of fuzzy regression analysis results from the realization that sometimes observations cannot be known or quantified exactly and we can only give a rough description of them, or periods to include in the fuzzy regression.

Simulation models [11]

The models vary according to the phenomena they represent, so it is not possible to display all models at the same time, but these models have been completed from actually published research that is in line with the practical application of this research, so some of these models that are linear were chosen, and these models are:

$$y_i = \beta_0 + x_{i1}\,\beta_1 + x_{i2}\,\beta_2 + x_{i3}\,\beta_3 + x_{i4}\,\beta_4 + e_i \quad \ldots\ldots\ldots (2)$$

where consists of a parametric regression function $X_i'\beta$ and Random error limit ε_i

The multi-target linear fuzzy regression model will be:

$$Y_i^* = A_i X_{ij} \quad \ldots\ldots\ldots (3)$$

Whereas

$$Y_i = (y_i, e_i)$$
$$A_i = (a, \alpha)$$

Whereas

a: Central parameter to A_i , α Propagation parameter to A_i

The following default values have been adopted:

$$\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4) = (1.5, 2, -0.5, 0.3, 1)$$

Four sample sizes were selected:

$$(n = 20, 50, 150)$$

The values of the illustrative variables were generated X_{ij} According to the standard normal distribution

$$X_{ij} \sim N(0,1), \quad i = 1, \ldots\ldots n, \quad j = 1,2,3,4 \quad \ldots.(4)$$

For the threshold limit or segment of the belonging function

$$(\alpha_\circ) = (0.3, 0.4, 0.5)$$

Random error vector generation [5]

Random errors are generated and distributed naturally with an average of (0) and variance σ^2, as the generation of variables that follow a normal distribution is done using the most famous and common method, which is the (Box-Muller) method, which depends on the method of generating random variables that follow the standard regular distribution U(0,1), and then these variables are converted into independent random variables that follow the standard normal distribution according to the following formula:

$$Z_1 = (-2LnU_1)^{\frac{1}{2}}Cos\,(2\pi U_2)$$

$$Z_2 = (-2LnU_1)^{\frac{1}{2}}Sin(2\pi U_2)$$

$$Z = \frac{(Z_1 + Z_2)}{2}$$

$$\varepsilon_i = Z\sigma^2 \quad ,i = 1,2,\dots,n$$

Simulation Experiment Results

Simulation experiments were carried out using three sample sizes n = 20,50,150 and replicates and contrast levels $\sigma^2 = 1.5$ and different cutting levels (0.3,0.4,0.5) for each simulation experiment.

Table (1) Estimation of Multi-target Fuzzy Regression Parameters Using LS Method when $\sigma^2 = 1.5$ Different threshold limits and sample sizes:

Threshold value	Sample Size	Parameters	X1	X2	X3	X4
α = 0.3	20	a	1.0909	2.2596	2.8805	1.1727
		α	2.0088	2.5661	1.5560	1.8515
	50	a	2.0123	1.2063	2.2740	2.1332
		α	2.2993	1.5889	2.2934	1.4556
	150	a	1.2416	1.3781	2.6670	1.3218
		α	2.6937	2.1614	2.4405	2.3734
Threshold value	Sample Size	Parameters	X1	X2	X3	X4
α = 0.4	20	a	1.9956	1.2884	1.0270	1.3430
		α	1.5114	2.1868	2.7703	2.2634
	50	a	1.0366	1.2824	1.6189	2.2766
		α	2.0418	1.7658	2.8984	1.6640
	150	a	2.5310	2.6776	1.3967	1.2510
		α	1.6958	2.4223	2.6685	1.0980
Threshold value	Sample Size	Parameters	X1	X2	X3	X4
α = 0.5	20	a	2.6767	2.0692	2.5709	2.8717
		α	1.9389	1.1130	1.2148	1.8669
	50	a	2.2417	2.2304	2.8607	2.4483
		α	1.6926	1.8794	1.3479	1.9388
	150	a	2.5318	2.6878	2.5442	2.3827
		α	1.7246	1.7870	1.9218	1.1743

Table (1) shows the estimation of the parameters of the multi-target fuzzy linear regression model according to the different sizes of the samples and according to the cutting limit of the belonging function and using the method of maximum possibility (MLE), the method of least squares (LS), the method of moments (Moment) and the method of Bayesian, where we note the fluctuation of the parameters with the rise and fall of their values with increasing variance, except in the case of α = 0.5, the estimators portfolio on the value ranges of the parameters.

Table (2) represents the values of the mean squares of error standard for the multi-target fuzzy regression model when the value of $\sigma^2 = 1.5$according to the sample sizes and the ellipsis limit of the affiliation function

alpha	Sample Size	LS	MLE	MOM	Bayesian	Best Method
0.3	20	5.4173	3.7236	4.7835	3.5287	Bayesian
	50	4.4885	3.1761	2.7046	2.6354	Bayesian
	150	2.7494	2.0346	2.1012	2.4617	MLE
0.4	20	3.8112	1.6135	2.8327	2.6118	MLE
	50	2.3453	1.4859	2.5245	2.4090	MLE
	150	2.3361	1.2920	1.6760	2.2421	MLE
0.5	20	2.8912	1.5060	3.8345	1.3046	Bayesian
	50	1.8860	1.3130	1.6646	1.2155	Bayesian
	150	1.2831	1.2908	1.3981	1.2138	Bayesian

For the purpose of giving a clear picture of the results of the simulation experiments, the results were analyzed to estimate the multi-objective fuzzy regression model based on the mean test of error squares and the results of Table (2) showed that the estimator is an estimator (Bayesian), when the sample size is n = 20 and the cutting parameter alpha = 0.3 and the level of variance $\sigma^2 = 1.5$and when the sample size is increased n = 50, the estimator (Bayesian) was also the best and the sample size was increased to n = 150, so the estimator (MLE) became the best among the rest of the estimators The results also showed that there is no fluctuation in the values of (MSE) with the increase in the sample size, it was noted that the values of (MSE) decrease with increasing sample size, and the results also showed that the best estimator is an estimator (MLE), when the sample size is n=20, the cut-off parameter alpha=0.4, and the level of variance $\sigma^2 = 1.5$, and when the sample size is increased n=50, the estimator (MLE) is also the best.

The results also showed that the best estimator is the estimator (Bayesian), when the sample size is n = 20 and the cutting parameter alpha = 0.5 and the level of variance $\sigma^2 = 1.5$and when the sample size is increased n = 50, the estimator (Bayesian) is also the best, and the sample size is increased to n = 150, so the estimator (Bayesian) is the best among the estimators as the results showed that there is no fluctuation In the values of (MSE) with increasing sample size, it was noted that the values of (MSE) decrease with increasing sample size and the results showed that there is no fluctuation in the values of (MSE) with increasing sample size, it was noted that the values of (MSE) decrease with increasing sample size.

CONCLUSIONS

From it we conclude that the best estimator when $\sigma^2 = 1.5$and the cutting parameter alpha=0.3 is the Bayesian estimator) and when alpha = 0.4 is the estimator (ME) is the best and by increasing the level of the cutting limit or threshold to alpha=0.5, the Bayesian estimator is also the best.

From the above, and depending on the criterion of mean squares of error for all cases, we note that the best estimator for estimating the multi-objective fuzzy regression model is the Bayesian estimator, followed by the second place estimator (MLE) and the worst estimator was the moment estimator (MOM) and (LS).

REFERENCES

1. Abbas, M. S. & Al-Metwally, A. S., (2021) "Using the Fuzzy Linear Regression Model in Estimating the Impact of the Dollar Exchange Rate on the Gross Domestic Product in Iraq", Journal of Administration and Economics - Al-Mustansiriya University, Volume and Issue 129.
2. Bajaj, R. K., Garg, G., Hooda, D.S., 2009, " Estimating Regression Coefficients in a Restricted Fuzzy Linear Regression Model", https://www.researchgate.net/publication/260714699, 27/10/2023, 7:27P.M.
3. Bisserier. A,& Galichet.S,Boukezzoula. R, 2008, "Fuzzy Piecewise Linear Regression" , HongKong, China. pp. 2089-2094.
4. Charfeddine, S. , Mora-Camino, F. & Coligny , M. (2014) ," Fuzzy linear regression : application to the estimation ofair transport demand" , hal-enac.archives-ouvertes.fr ,Russia. pp 350-359. hal-01022443.
5. Dahmani,s., Maamoun, M., Zerar,G., Chabini, N.& Beguenane,R., 2022, " An Efficient FPGA-Based Gaussian Random Number Generator Using an Accurate Segmented Box-Muller Method ", IEEE Access, Citation information: DOI 10.1109/ACCESS.2023.3289432
6. Donoso. S, 2006, "Quadratic Programming Models for Fuzzy Regression", Nicol´asMar´ın and M. Amparo Vila Department of Computer Science and A. I. University of Granada - 18071 - Granada – Spain
7. Elias, H. M. & Sabbagh, Heba Ali Taha (2006), "Fuzzy Regression Analysis", Iraqi Journal of Statistical Sciences, No. 10, pp. (61-84.)
8. Karakasidis.T,Georgiou.D, Nieto.J, 2012," Fuzzy regression analysis: An application on tensile strength of materials and hardness scales", Journal of Intelligent & Fuzzy Systems 23 ,177–186
9. Mohammed,J.M., Abbas. M.S,2018, "Estimation Nonparametric Fuzzy Regression model using simulation", Economics & Administration of Journal The ,41,Volume115.
10. Mosleh. M, Otadi. M, Abbasbandy. S, 2011,". Fuzzy polynomial regression with fuzzy neural networks", Applied Mathematical Modelling, Volume 35, Issue 11, Pages 5400-5412
11. Nasrabadi, M.M, Nasrabadi, E.& Nasrabady, A.R., (2005) " Fuzzy linear regression analysis: a multi-objective programming approach", Applied Mathematics and Computation 163, P: 245–251.