E-ISSN: 2717-3453 | P-ISSN: 2783-1442

Journal of Fuzzy Extension and Applications

www.journal-fea.com

J. Fuzzy. Ext. Appl. Vol. 3, No. 4 (2022) 279–289.

Paper Type: Research Paper

Fuzzy Simple Linear Regression Using Gaussian Membership Functions Minimization Problem

Besma Belhadj*

LaREQuaD, FSEGT, University of ElManar, Tunisia; besmabelhaj14@gmail.com.

Citation:

Belhadj, B. (2022). Fuzzy simple linear regression using Gaussian membership functions minimization problem. *Journal of fuzzy extension and applications, 3*(4), 279-289.

Received: 02/06/2022 Reviewed: 30/06/2021 Revised: 19/08/2022 Accepted: 22/08/2022

Abstract

Under the additional assumption that the errors are normally distributed, the Ordinary Least Squares (OLS) method is the maximum likelihood estimator. In this paper, we propose, for a simple regression, an estimation method alternative to the OLS method based on a so-called Gaussian membership function, one that checks the validity of the verbal explanation suggested by the observer. The fuzzy estimation approach demonstrated here is based on a suitable framework for a natural behavior observed in nature. An application based on a group of MENA countries in 2015 is presented to estimate the employment poverty relationship.

Keywords: Mathematical modeling, Fuzzy regression, Gaussian fuzzy responses, Gaussian membership function.

1 | Introduction

The simple linear regression concerns sample points with one independent variable and one dependent variable and finds a linear function (a non-vertical straight line) that, as accurately as possible, predicts the dependent variable values as a function of the independent variable. One of the most used regression methods is the Ordinary Least Squares (OLS) method [4].

OLS is a type of linear least squares method for estimating the unknown parameters in a linear regression model. The OLS estimator is consistent when the regressors are exogenous and, by the Gauss Markov theorem, optimal in the class of linear unbiased estimators when the errors are homoscedastic and serially uncorrelated. Under these conditions, the method of OLS provides minimum-variance mean-unbiased estimation when the errors have finite variances. Under the additional assumption that the errors are normally distributed, OLS is the maximum likelihood estimator. The errors after modeling, however, should be normal to draw a valid conclusion by hypothesis testing.

Corresponding Author: besmabelhaj14@gmail.com

https://doi.org/10.22105/jfea.2022.345298.1222

Our data can be normal or not; we can check data distributions to understand their behaviors. Our data might not be normal for a reason. Usually, in such cases, you may want to transform it or use other analysis methods (e.g., generalized linear models or nonparametric methods).

Belhadj [2] proves using the theory of fuzzy subsets that any dependent or independent variable follows a particular real behavior. In particular, according to this author, poverty can have a trapezoidal behavior. However, the analysis and estimation of regression must change depending on the behavior of the regression variables. For example, Belhadj and Kaabi [3] proposed a method for estimating a simple regression where the behavior of the dependent variable is trapezoidal.

This paper aims to develop further and refine another case study on the same strand of research of [3]. First, we assume our data is normal and set the so-called Gaussian membership functions. They are fuzzy versions of the classic Gaussian distribution, those who verify the validity of the verbal explanation suggested by the observer. Secondly, we construct fuzzy mathematical modelings of a simple linear regression model using these Gaussian membership functions. Fuzzy modeling approaches demonstrated here are based on mimicking a natural behavior observed in nature. Fuzzy modeling approaches demonstrated here change according to the distribution of our data. It keeps reality, unlike the OLS and any other estimation method, in its entirety.

This paper is structured as follows: Section 2 presents some unique properties of Gaussian distribution and briefly presents the OLS method using the normal distribution. Section 3 shows the difference between the Gaussian distribution and membership function. Section 4 puts forward an alternative estimation method called fuzzy regression using Gaussian membership functions. In this section, estimators of a linear fuzzy regression model are constructed, and the consistency of these estimators is established. Section 5 illustrates the use of this proposed method to estimate the employment poverty relationship in MENA in 2015. Section 6 concludes.

2 | Gaussian Distribution, OLS Method

A Gaussian distribution, said normally distributed, is a continuous probability distribution for a real-valued random variable. Its probability density function [4] is

$$(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-1}{2}\left(\frac{x-m}{\sigma}\right)^2}, x\in R.$$

m is the mean and also its median and mode, while the parameter σ is its standard deviation.

The normal distribution is important in analytic studies. It is the only distribution with zero cumulants beyond the mean and variance. It is also the continuous distribution with the maximum entropy for a specified mean and variance. And is the only distribution where the mean and variance calculated from a set of independent draws are independent of each other. The normal distribution is a subclass of the elliptical distributions. It is symmetric about its mean and is non-zero over the entire real line. It is one of the few stable distributions with probability density functions that can be expressed analytically. Many results and methods, such as propagation of uncertainty and least squares parameter fitting, can be derived analytically in explicit form when the relevant variables are normally distributed.

Let the model $y_i = a + bx_i + \varepsilon_i$. The variable that is supposed to be normally distributed is just the prediction error ε_i. Prediction error should follow a normal distribution with a mean of 0. The confidence interval and variable significance calculation are based on this assumption [4]. For example, the effect of unemployment on poverty based on a 5% significance level requires following a normal

Belhadj |J. Fuzzy. Ext. Appl. 3(4) (2022) 279-289

distribution with a mean of 0. If the error distribution significantly deviates from the mean 0 normal distribution, the effect may not actually be significant enough to explain poverty.

The least squares estimates in the case $y_i = a + bx_i + \varepsilon_i$ are given by simple formulas [4]:

$$\hat{b} = \frac{n\sum_i x_i y_i - \sum_i x_i \sum_i y_i}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2},$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x}.$$

The OLS estimators $\hat{a}$ and $\hat{b}$ are Best Linear Unbiased Estimator (BLUE) and Consistent.

3 | Gaussian Membership Function

Fuzzy sets (see, e.g. [11], [21] for application of this approach) are extensions of the classical sets whose elements have degrees of membership [23]. In fuzzy set theory, classical bivalent sets are usually called crisp sets. The fuzzy set theory allows describing situations in which the data are imprecise. Fuzzy sets handle such situations by attributing degrees to which elements belong to a set, called membership function, having intervals [0,1].

The membership function fully defines the fuzzy set. It measures the degree of similarity of an element to a fuzzy set. Membership functions can: either be chosen by the user arbitrarily, based on the user's experience. Or be designed using machine learning methods (e.g., artificial neural networks, genetic algorithms). Membership functions have different shapes, mainly Triangular, Trapezoidal, and Gaussian forms [14].

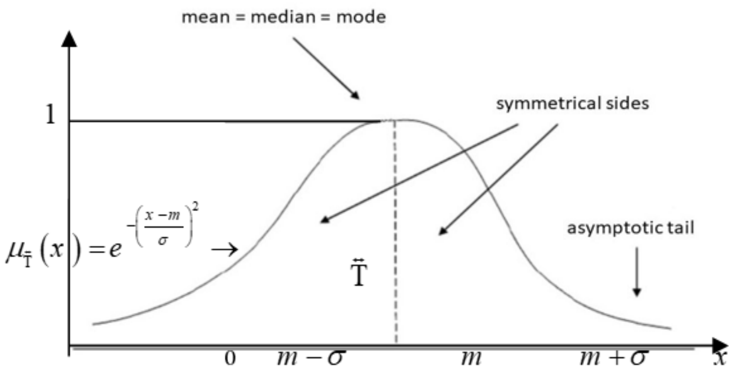

Fig. 1. Gaussian membership function.

In fuzzy logic, the Gaussian membership function is a generalization of the Gaussian distribution for classical sets. It represents the degree of truth often confused with probability. However, it is conceptually distinct because fuzzy truth represents membership in vaguely defined sets, not the likelihood of some event or condition. The general form of the Gaussian membership function is $\mu_{\tilde{T}}(x) = e^{-\left(\frac{x-m}{\sigma}\right)^2}$ (*Fig. 1*). For an element x of X, $\mu_{\tilde{T}}(x)$ quantifies the grade of membership of the element x to the fuzzy set $\tilde{T}$ (*Fig. 1*). The value 0 means that x is not a member of the fuzzy set; the value 1 indicates that x is fully a member of the fuzzy set. The values between 0 and 1 characterize fuzzy members, which belong to the fuzzy set only partially. The Gaussian membership function is employed in several domains (e.g., [16]-[18]).

4 | Fuzzy Linear Regression Models

A fuzzy linear regression model was first introduced by [21]. He formulated a linear regression model with fuzzy response data, crisp predictor data, and fuzzy parameters as a mathematical programming problem. Their approach was later improved to give birth to many other methods: linear-programming-based methods said possibilistic approach (e.g. [1], [6], [9], [10], [19], [20]), fuzzy least-squares methods (e.g. [5], [7], [8], [22]) and fuzzy minimizing method said dissemblance method [3].

Fuzzy simple linear regression is different from simple linear regression in the sense that it is a no statistical method [12]. However, in some cases, we may need to consider that the relationship expressed as $y_i = a + bx_i + \varepsilon_i$ may be fuzzy. Indeed, three cases are possible:

In the case where the predictor variable is fuzzy, but the parameters are crisp:

$$\tilde{y}_i = a + b\tilde{x}_i + \tilde{\varepsilon}_i. \quad (1)$$

The case of a crisp predictor and fuzzy parameters:

$$\tilde{y}_i = \tilde{a} + \tilde{b}x_i + \tilde{\varepsilon}_i. \quad (2)$$

And finally, the case of a fuzzy predictor and fuzzy parameters:

$$\tilde{y}_i = \tilde{a} + \tilde{b}\tilde{x}_i + \tilde{\varepsilon}_i, \quad (3)$$

where $ith, i = 1, ..., n,$ are fuzzy responses.

In this section, we retain a curve representing the modification of the confidence interval from 0 to 1. It may have a deflection in its slope, resulting in a flat region, as shown in *Fig. 2*. This curve is an L-R Gaussian membership function with a flat.

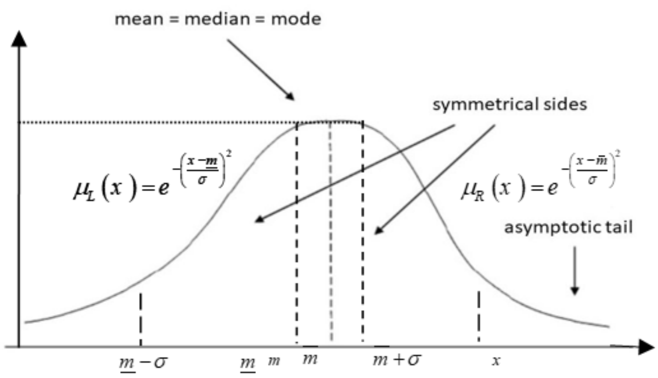

Fig. 2. L-R Gaussian fuzzy number with a flat.

We estimate *Model (2)* parameters when unknown parameters are L-R Gaussian fuzzy numbers with a flat. The *Model (2)* was also treated by [3] using Trapezoidal fuzzy numbers.

We choose four significant numbers to represent the L-R Gaussian fuzzy number nonunimodally. Let $\tilde{S} = (\underline{m}-\sigma, \underline{m}, \bar{m}, \bar{m}+\sigma)$ an L-R Gaussian fuzzy number with a Flat where $\underline{m}-\sigma$ and $\bar{m}+\sigma$ are the left and right "end" points of the corresponding bell, and $\underline{m}$ and $\bar{m}$ are the left and right "middle" points (*Fig. 2*).

The membership of the fuzzy response $\tilde{y}_i$ of the *Model (2)* is as follows:

Belhadj |J. Fuzzy. Ext. Appl. 3(4) (2022) 279-289

$$\mu_{\tilde{y}_i}(u_i)=\begin{cases} e^{-\left(\frac{u_i-\underline{m}}{\sigma}\right)^2}, & -\infty \le u_i < \underline{m},\\ 1, & \underline{m} \le u_i < \bar{m},\\ e^{-\left(\frac{u_i-\bar{m}}{\sigma}\right)^2}, & \bar{m} \le u_i < +\infty. \end{cases} \tag{4}$$

We suppose that the parameters of the *Model (2)*, $\tilde{a}$ and $\hat{b}$ are L-R Gaussian fuzzy numbers with a flat, where $\tilde{a}=\left(\underline{a}-\sigma_a,\underline{a},\bar{a},\bar{a}+\sigma_a\right)$ and $\tilde{b}=\left(\underline{b}-\sigma_b,\underline{b},\bar{b},\bar{b}+\sigma_b\right)$. Then according to the [13]:

$$a(+)bx=\left(\left(\underline{a}-\sigma_a\right)+\left(\underline{b}-\sigma_b\right)x,\underline{a}+\underline{b}x,\bar{a}+\bar{b}x,\left(\bar{a}+\sigma_a\right)+\left(\bar{b}+\sigma_b\right)x\right),$$

from which we have

$$\forall u \in R:$$

$$\begin{aligned} \mu_{\tilde{a}(+)\tilde{b}x_i}(u_i) &= e^{-\left(\frac{u_i-\left(\left(\underline{a}-\sigma_a\right)+\left(\underline{b}-\sigma_b\right)x_i\right)}{\sigma_a+\sigma_b}\right)^2} && -\infty < u_i < \left(\underline{a}-\sigma_a\right)+\left(\underline{b}-\sigma_b\right)x_i\\ &=1 && \underline{a}+\underline{b}x_i \le u_i < \bar{a}+\bar{b}x_i\\ &= e^{-\left(\frac{u_i-\left(\left(\bar{a}-\sigma_a\right)+\left(\bar{b}-\sigma_b\right)x_i\right)}{\sigma_a+\sigma_b}\right)^2} && \left(\bar{a}-\sigma_a\right)+\left(\bar{b}-\sigma_b\right)x_i \le u_i < +\infty. \end{aligned} \tag{5}$$

4.1 | Error's Gaussian Membership Function

Starting from *Model (2)*, we can write:

$$\mu_{\tilde{\varepsilon}_i}(\varepsilon_i)=\mu_{\tilde{y}_i(-)\,\tilde{a}+\tilde{b}x_i}(\varepsilon_i). \tag{6}$$

for the term to the right of *Eq. (6)*, it's about the subtraction of Gaussian fuzzy numbers, which is written as follows:

$$\mu_{\tilde{y}_i(-)\,\tilde{a}+\tilde{b}x_i}(\varepsilon_i)=\bigvee_{\varepsilon_i}\left(\mu_{\tilde{a}+\tilde{b}x_i}(a+bx_i)\wedge\mu_{\tilde{y}_i}(y_i)\right), \tag{7}$$

which corresponds to

$$\tilde{y}_i(-)\tilde{a}+\tilde{b}x_i=\left(\underline{m}-\sigma-\left(\bar{a}+\sigma_a\right)-\left(\bar{b}+\sigma_b\right)x_i,\underline{m}-\bar{a}-\bar{b}x_i,\bar{m}-\underline{a}-\underline{b}x_i,\bar{m}+\sigma-\left(\underline{a}-\sigma_a\right)-\left(\underline{b}-\sigma_b\right)x_i\right).$$

Now using *Eq. (4)*, let $\forall\alpha\in[0,1]$, $\alpha=\left(e^{-\left(\frac{u_i-\underline{m}}{\sigma}\right)^2}\right)^{(\alpha)}$, $\alpha=\underline{m}^{(\alpha)}$, $\alpha=\bar{m}^{(\alpha)}$, $\alpha=\left(e^{-\left(\frac{u_i-\bar{m}}{\sigma}\right)^2}\right)^{(\alpha)}$, from which

$$\begin{aligned} \breve{y}_{i\alpha} &= \left(y_{1i}^{(\alpha)},y_{2i}^{(\alpha)},y_{3i}^{(\alpha)},y_{4i}^{(\alpha)}\right)\\ &= \left[\alpha-e^{-\left(\frac{u_i-\underline{m}}{\sigma}\right)^2},\underline{m},\bar{m},\alpha-e^{-\left(\frac{u_i-\bar{m}}{\sigma}\right)^2}\right]. \end{aligned} \tag{8}$$

Now using *Eq. (5)*, we obtain $\alpha=\left(e^{-\left(\frac{u_i-\left((\underline{a}-\sigma_a)+(\underline{b}-\sigma_b)x_i\right)}{\sigma_a+\sigma_b}\right)^2}\right)^{(\alpha)}$, $\alpha=\left(\underline{a}+\underline{b}x_i\right)^{(\alpha)}$, $\alpha=\left(\bar{a}+\bar{b}x_i\right)^{(\alpha)}$,

$\alpha=\left(e^{-\left(\frac{u_i-\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)}{\sigma_a+\sigma_b}\right)^2}\right)^{(\alpha)}$, from which

$$\begin{aligned}\left(\tilde{a}+\tilde{b}x_i\right)_\alpha &=\left(v_{1i}^{(\alpha)},v_{2i}^{(\alpha)},v_{3i}^{(\alpha)},v_{4i}^{(\alpha)}\right)\\ &=\left[-\alpha+e^{-\left(\frac{u_i-\left((\underline{a}-\sigma_a)+(\underline{b}-\sigma_b)x_i\right)}{\sigma_a+\sigma_b}\right)^2},\underline{a}+\underline{b}x_i,\bar{a}+\bar{b}x_i,-\alpha+e^{-\left(\frac{u_i-\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)}{\sigma_a+\sigma_b}\right)^2}\right].\end{aligned} \tag{9}$$

We must then have

$$\breve{y}_{i\alpha}(-)\left(\breve{a}+\breve{b}x_i\right)_\alpha=\left(y_{1i}^{(\alpha)},y_{2i}^{(\alpha)},y_{3i}^{(\alpha)},y_{4i}^{(\alpha)}\right)(-)\left(v_{1i}^{(\alpha)},v_{2i}^{(\alpha)},v_{3i}^{(\alpha)},v_{4i}^{(\alpha)}\right)=\left[y_{1i}^{(\alpha)}-v_{4i}^{(\alpha)},y_{2i}^{(\alpha)}-v_{3i}^{(\alpha)},y_{3i}^{(\alpha)}-v_{2i}^{(\alpha)},y_{4i}^{(\alpha)}-v_{1i}^{(\alpha)}\right].$$

Thus if we define $y_{1i}^{(\alpha)}-v_{4i}^{(\alpha)}=\alpha-e^{-\left(\frac{u_i-\underline{m}}{\sigma}\right)^2}+\alpha-e^{-\left(\frac{u_i-\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)}{\sigma_a+\sigma_b}\right)^2}$,

$y_{4i}^{(\alpha)}-v_{1i}^{(\alpha)}=\alpha-e^{-\left(\frac{u_i-\bar{m}}{\sigma}\right)^2}+\alpha-e^{-\left(\frac{u_i-\left((\underline{a}-\sigma_a)+(\underline{b}-\sigma_b)x_i\right)}{\sigma_a+\sigma_b}\right)^2}$, from which we have

$$\forall u\in R:$$

$$\begin{aligned}\mu_{\tilde{\varepsilon}}(u_i)&=e^{-\left(\frac{u_i-\left(\underline{m}-\sigma-(\underline{a}-\sigma_a)-(\underline{b}-\sigma_b)x_i\right)}{\sigma_a+\sigma_b+\sigma}\right)^2} && -\infty<u_i<\underline{m}-\sigma-\left(\underline{a}-\sigma_a\right)-\left(\underline{b}-\sigma_b\right)x_i\\ &=1 && \underline{m}-\underline{a}-\underline{b}x_i\le u_i<\bar{m}-\bar{a}-\bar{b}x_i\\ &=e^{-\left(\frac{u_i-\left(\bar{m}+\sigma-(\bar{a}+\sigma_a)-(\bar{b}+\sigma_b)x_i\right)}{\sigma_a+\sigma_b+\sigma}\right)^2} && \bar{m}+\sigma-\left(\bar{a}+\sigma_a\right)-\left(\bar{b}+\sigma_b\right)x_i\le u_i<+\infty\end{aligned} \tag{10}$$

The estimate of a and b of *Model (2)* consists of solving the following minimization problem:

$$\min\mu_{\tilde{\varepsilon}_i}\left(\tilde{a},\tilde{b}\right). \tag{11}$$

Using *Eq. (7)*, *Eq. (11)* becomes

$$\min\mu_{\tilde{\varepsilon}_i}\left(\tilde{a},\tilde{b}\right)=\min\bigvee_{\varepsilon_i}\left(\mu_{\tilde{a}+\tilde{b}x_i}(u_i)\wedge\mu_{\tilde{y}_i}(u_i)\right). \tag{12}$$

We show that to solve *Eq. (12)* is to solve the following system

$$\partial/\partial\underline{a},\partial\underline{b}\frac{\sum_{i=1}^{n}\left(u_i-\left(\underline{m}-\sigma-\left(\underline{a}+\sigma_a\right)-\left(\underline{b}+\sigma_b\right)x_i\right)\right)^2}{\left(\sigma_a+\sigma_b+\sigma\right)^2}=0\ . \tag{13}$$

$$\partial/\partial\bar{a},\partial\bar{b}\frac{\sum_{i=1}^{n}\left(u_i-\left(\bar{m}-\sigma-\left(\bar{a}+\sigma_a\right)-\left(\bar{b}+\sigma_b\right)x_i\right)\right)^2}{\left(\sigma_a+\sigma_b+\sigma\right)^2}=0. \tag{14}$$

Belhadj |J. Fuzzy. Ext. Appl. 3(4) (2022) 279-289

We prove that the error's membership function $\mu_{\tilde{\varepsilon}_i}(\tilde{a},\tilde{b})$ is also the (L)-R hatched area of *Fig. 3*, given by the difference between the membership function of the observed value, $\mu_{\tilde{y}_i}(u_i)$, and the membership function of the estimated value of y_i, $\mu_{\tilde{a}+\tilde{b}x_i}(u_i)$. Therefore, solving *Eq. (12)* is equivalent to minimizing the hatched areas of *Fig. 3*, hence the following method.

4.2 | Fuzzy Numbers Gaps Square Method

Starting from *Eqs. (8)* and *(9)*, we define $\forall i = 1,\ldots,n$ two Gaussian fuzzy numbers A and B, as follows: $A=[y_{1i},y_{2i},y_{3i},y_{4i}]$ and $B=[v_{1i},v_{2i},v_{3i},v_{4i}]$. Or, from *Eq. (4)*, $\forall\alpha\in[0,1]$ $\alpha=e^{-\left(\frac{u_i-\underline{m}}{\sigma}\right)^2} \Rightarrow u_i=\sigma\sqrt{-\log\alpha}+\underline{m}$, $\alpha=e^{-\left(\frac{u_i-\bar{m}}{\sigma}\right)^2} \Rightarrow u_i=\sigma\sqrt{-\log\alpha}+\bar{m}$, $A_\alpha=\left[\sigma\sqrt{-\log\alpha}+\underline{m},\sigma\sqrt{-\log\alpha}+\bar{m}\right]$. And from *Eq. (5)*, $B_\alpha=\left[\left((\underline{a}-\sigma_a)+(\underline{b}-\sigma_b)x_i\right)+(\sigma_a+\sigma_b)\sqrt{-\log\alpha},\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)+(\sigma_a+\sigma_b)\sqrt{-\log\alpha}\right]$.

Note that, in the right of *Fig. 2*, A and B intersect at the following two points (*Fig. 3*) $\sigma\sqrt{-\log\alpha}+\bar{m}=\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)+(\sigma_a+\sigma_b)\sqrt{-\log\alpha}$, gives

$$\alpha_1=e^{-\left(\frac{\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)-\bar{m}}{\left(\sigma-(\sigma_a+\sigma_b)\right)}\right)^2}. \tag{15}$$

$\bar{m}=\bar{a}+\bar{b}x_i$ gives

$$\alpha_2=1. \tag{16}$$

We then have

$$\begin{aligned} y_{4i}^{(\alpha)}-v_{4i}^{(\alpha)} &= \sigma\sqrt{-\log\alpha}+\bar{m}-\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)-(\sigma_a+\sigma_b)\sqrt{-\log\alpha} \\ &= \bar{m}-\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)+\left(\sigma-(\sigma_a+\sigma_b)\right)\sqrt{-\log\alpha}. \end{aligned}$$

Always on the right, if we proceed to integrate $R\delta(A_\alpha,B_\alpha)$ from $\alpha=0$ to $\alpha=1$, we obtain a distance $R\delta(A,B)$ by the summation of distances: $R\delta(A,B)=\int_{\alpha=0}^{1}R\delta(A_\alpha,B_\alpha)d\alpha.$

It is a distance between two Gaussian fuzzy numbers A and B.

$$\begin{aligned} R\delta(A,B) &= \int_{\alpha=0}^{\alpha_1}\left[\bar{m}-\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)+\left(\sigma-(\sigma_a+\sigma_b)\right)\sqrt{-\log\alpha}\right]d\alpha \\ &+\int_{\alpha_1}^{1}\left[\bar{m}-\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)+\left(\sigma-(\sigma_a+\sigma_b)\right)\sqrt{-\log\alpha}\right]d\alpha \end{aligned} \tag{17}$$

The calculation of *Eq. (17)* shows

$$R\delta(A,B)=\left[\bar{m}-\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)\right]\left([\alpha]_{\alpha=0}^{\alpha_1}+[\alpha]_{\alpha_1}^{1}\right)+\left(\sigma-(\sigma_a+\sigma_b)\right)\int_0^1\left[\sqrt{-\log\alpha}\right]d\alpha.$$

And,

$$R\delta(A,B)=\left[\bar{m}-\left((\bar{a}-\sigma_a)+(\bar{b}-\sigma_b)x_i\right)\right]\left([\alpha]_{\alpha=0}^{\alpha_1}+[\alpha]_{\alpha_1}^{1}\right)+\frac{\sqrt{\pi}}{2}\left(\sigma-(\sigma_a+\sigma_b)\right). \tag{18}$$

For Gaussian fuzzy numbers A and B, the distance $R\delta(A,B)$ is the sum of areas hatched in *Fig. 3*.

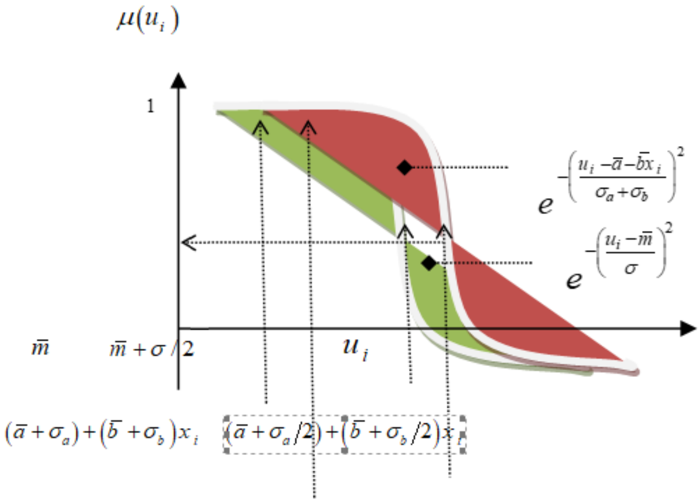

Fig. 3. R Gaussian fuzzy number with a flat.

The estimate of a and b of *Model (2)* consists of solving the following minimization problem:

$$\min\sum_i R\delta^2(A,B)=\min\sum_i R\delta^2(\tilde{a},\tilde{b}). \tag{19}$$

Problem (19) is equivalent to the *Problem (14)*. Its resolution consists of solving the *System (20)*.

$$\begin{cases}\dfrac{\partial R\delta^2(\tilde{a},\tilde{b})}{\partial \bar{a}}=0,\\[2ex] \dfrac{\partial R\delta^2(\tilde{a},\tilde{b})}{\partial \bar{b}}=0.\end{cases} \tag{20}$$

The solution of *Eq. (20)* is

$$\begin{cases}\bar{a}+\bar{b}\bar{X}=\dfrac{\sqrt{\pi}}{2}(\sigma-\sigma_a-\sigma_b)+\bar{m}+n\sigma_a+\bar{X}\sigma_b,\\ \bar{a}\bar{X}+\bar{b}\overline{X^2}=\dfrac{\sqrt{\pi}}{2}(\sigma-\sigma_a-\sigma_b)\bar{X}-\bar{m}\bar{X}+\bar{X}\sigma_a+\overline{X^2}\sigma_b.\end{cases} \tag{21}$$

Which gives

$$\bar{b}=\frac{-2\bar{m}\bar{X}+(1-n)\bar{X}\sigma_a}{\sigma_X^2}+\sigma_b. \tag{22}$$

And

$$\bar{a}=\frac{\sqrt{\pi}}{2}(\sigma-\sigma_b)+\frac{\bar{m}\left(\sigma_X^2+2\bar{X}^2\right)+\sigma_a\left[\left(n-\sqrt{\pi}/2\right)\sigma_X^2-(1-n)\bar{X}^2\right]}{\sigma_X^2}. \tag{23}$$

This paper's fuzzy estimators of the linear regression proposed are solutions to reality modeling. They may be of some use because of the complexity of phenomena, such as the social economy and the influence of various uncertain factors existing in the boundary environment around the phenomena. As justification for the proposed fuzzy modeling, we estimate the employment poverty relationship among a group of MENA countries in the following section in 2015. This estimation justifies the relationship unclear between poverty and unemployment.

Belhadj |J. Fuzzy. Ext. Appl. 3(4) (2022) 279-289

5 | Employment Poverty Relationship Using Fuzzy Numbers Gaps Square Method

We start from the observation that employment and the access that the poor have to decent earning opportunities will be crucial determinants of poverty reduction. This observation brings us back to estimating the relationship between the poverty headcount ratio at $1.90 and the unemployment rate in 2015 for a group of developing countries from the MENA region.

The estimation of this relation by the OLS method shows the following results

$$\begin{aligned} &P_i = 17{,}37 - 0{,}81U_i \\ &\quad (2{,}89)(-1{,}94) \\ &(R^2 = 0{,}201;\ N = 23;\ \text{t values in parentheses}). \end{aligned} \tag{24}$$

Looking at the results of *Eq. (24)*, we find that the application of the OLS method show, as [23], a negative relationship which can be explained by: the unemployment rate is generally higher in developing middle-income countries. In comparison, the poverty rate is higher in developing low-income countries [23]. But we show below that by changing the method, the result changes.

We retain the same sample from *Eq. (24)* and apply the fuzzy numbers gaps square method. We calculate, however, the solutions *Eq. (22)* and *Eq. (23)*. We, therefore, obtain the following results.

Table 1. Employment poverty relationship using the fuzzy numbers gaps square method.

	a	b	R^2
The left-Gaussian-fuzzy-number-estimates	23,45 (3,06)	-0,58 (-2,03)	0,879
The right-Gaussian-fuzzy-number-estimates	14,06 (3,97)	0,43 (3,77)	0,901

$R^2 = SSR / SST$ where SSR is the sum of squares due to regression, SST is the total sum of squares

As shown in *Table 1*, the value of the constant term depicts that when the unemployment rate is assumed to be zero, the incidence of poverty will still be visible, thus insinuating that the unemployment rate is not the only factor responsible for the incidence of poverty.

We also find that the parameter measuring unemployment is not always negative as in *Eq. (24)*. The relationship between the incidence of poverty and unemployment is positive in the right-Gaussian-fuzzy-number *Fig. 2*, but in the left-Gaussian-fuzzy-number, the relationship is negative. These results justify the relationship unclear between poverty and unemployment. However, being unemployed usually results in falling in one's living standard due to the absence of income, and one can be employed and still be poor.

The relationship between the incidence of poverty and unemployment is, then, subject to the linked labor force and income statistics path tracing. For example, the problem of interpretation becomes complicated among workers who experience an employment problem but whose family income does not fall below the poverty level.

However, the realistic depiction of the phenomenon is critical. For many natural effects, including fog, steam, smoke, and fire, a recent survey by [25] gives an overview of the best-known methods for importance sampling. Our approach illustrates an alternative importance sampling strategy that can yield more robust estimators. It is based on variable distribution to calculate the relationship between them validly.

Both positive and negative relationship between unemployment and poverty then returns to the distribution of each variable. A possible situation shows that MENA's 'working poor' still live in poverty.

The study results show that the employment rate has a statistically significant impact on reducing non-extreme poverty for the MENA sample. The results confirm that, in the right part of the distribution, a one-unit increase in the employment rate results in a fall in the poverty headcount ratio by about 43% for the MENA sample. Therefore, in the right part of the distribution, employment is a vector to lift out of poverty because job quality is sufficient, including adequate earnings, job security, and safe working environments. In the left part of the distribution, the working poverty rate reveals the proportion of employed people living in poverty despite being employed, implying that their employment-related incomes are insufficient to lift out of poverty and ensure decent living conditions.

The relationship between employment and poverty depends significantly on the extent to which decent work is ensured in the labor market. This means that having a job is not enough to keep out of poverty, pointing to job quality issues, particularly the inadequacy of earnings.

6 | Conclusion

In this paper, we propose new non-statistical methods - fuzzy regressions - as alternative methods of ordinary regression analysis. These methods deal with fuzziness; they estimate the parameters of a fuzzy simple regression model for the case of fuzzy parameters using Gaussian membership functions nonunimodally.

These methods are desirable since they are based on an integral abstraction of reality while keeping the same process as the classical linear regression. Moreover, the calculation technique for these methods keeps a Gaussian distribution in a real context. We can, in future work, extend our methodology for the case where predictor and parameters are fuzzy.

Funding

The author declares that no funds, grants, or other support were received during the preparation of this manuscript.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

The author contributed to the study's conception and design. The author read and approved the final manuscript.

References

[1] Arabpour, A. R., & Tata, M. (2008). Estimating the parameters of a fuzzy linear regression model. *Iranian journal of fuzzy systems, 5*(2), 1-19.

[2] Belhadj, B. (2011). A new fuzzy unidimensional poverty index from an information theory perspective. *Empirical economics, 40*(3), 687-704.

[3] Belhadj, B., & Kaabi, F. (2021). The relationship between employment and poverty using fuzzy regression. In *Analysis of socio-economic conditions* (pp. 264-275). Routledge.

[4] Bourbonnais, R. (2021). *Econométrie*. Dunod. (In French).

[5] Diamond, P. (1988). Fuzzy least squares. *Information sciences, 46*(3), 141-157.

[6] Diamond, P., & Tanaka, H. (1998). Fuzzy regression analysis. In *Fuzzy sets in decision analysis, operations research and statistics* (pp. 349-387). Springer, Boston, MA.

[7] D'Urso, P., & Gastaldi, T. (2000). A least-squares approach to fuzzy linear regression analysis. *Computational statistics and data analysis, 34*(4), 427-440.

[8] D'Urso, P., & Chachi, J. (2022). OWA fuzzy regression. *International journal of approximate reasoning*, *142*, 430-450.

[9] Hong, D. H., & Yi, H. C. (2003). A note on fuzzy regression model with fuzzy input and output data for manpower forecasting. *Fuzzy sets and systems, 138*(2), 301-305.

[10] Hong, D. H., Hwang, C., & Ahn, C. (2004). Ridge estimation for regression models with crisp inputs and Gaussian fuzzy output. *Fuzzy sets and systems, 142*(2), 307-319.

[11] Imeni, M. (2020). Fuzzy logic in accounting and auditing. *Journal of fuzzy extension and applications, 1*(1), 69-75.

[12] Kim, K. J., & Chen, H. R. (1997). A comparison of fuzzy and nonparametric linear regression. *Computers and operations research, 24*(6), 505-519.

[13] Kaufmann, A., & Gupta, M. M. (1991). *Introduction to fuzzy arithmetic: theory and applications*. Van Nostrand Reinhold.

[14] Leandry, L., Sosoma, I., & Koloseni, D. (In Press). Basic fuzzy arithmetic operations using α–cut for the Gaussian membership function. *Journal of fuzzy extension and applications*. https://doi.org/10.22105/jfea.2022.339888.1218

[15] Majid, N. (2001). Les travailleurs pauvres dans les pays en développement. *Revue internationale du travail*, *140*(3), 323-343.

[16] Pandit, M., Chaudhary, V., Dubey, H. M., & Panigrahi, B. K. (2015). Multi-period wind integrated optimal dispatch using series PSO-DE with time-varying Gaussian membership function based fuzzy selection. *International journal of electrical power & energy systems*, *73*, 259-272.

[17] Raab, M., Seibert, D., & Keller, A. (2008). Unbiased global illumination with participating media. In *Monte Carlo and Quasi-Monte Carlo methods 2006* (pp. 591-605). Springer, Berlin, Heidelberg.

[18] Radhakrishna, V., Aljawarneh, S. A., Veereswara Kumar, P., & Janaki, V. (2019). ASTRA-A Novel interest measure for unearthing latent temporal associations and trends through extending basic Gaussian membership function. *Multimedia tools and applications, 78*(4), 4217-4265.

[19] Tanaka, H. (1987). Fuzzy data analysis by possibilistic linear models. *Fuzzy sets and systems, 24*(3), 363-375.

[20] Tanaka, H., & Ishibuchi, H. (1991). Identification of possibilistic linear systems by quadratic membership functions of fuzzy parameters. *Fuzzy sets and systems, 41*(2), 145-160.

[21] Tanaka, H., Uejima, S., & Asai, K. (1982). Linear regression analysis with fuzzy model. *IEEE transactions systems man and cybernetics, 12*(6), 903–907.

[22] Yang, M. S., & Liu, H. H. (2003). Fuzzy least-squares algorithms for interactive fuzzy linear regression models. *Fuzzy sets and systems, 135*(2), 305-316.

[23] Zadeh, L. (1965). Probability theory and fuzzy logic are complementary rather than competitive. *Technometrics, 37*(3), 271–276.