Neutrosophic Sets and Systems, Vol. 67, 2024
University of New Mexico

Neutrosophic Statistics for Enhanced Time Series Analysis of Unemployment Trends in Ecuador

Stephanie M. Delgado Estrada [1*], Katia del Rocio Ruiz Molina [2], Fernando Enrique Ponce Orellana [3], Jorge Luis Chabusa Vargas [4]

[1] Universidad de Guayaquil, Guayas, Ecuador, Email: stephanie.delgadoe@ug.edu.ec*
[2] Universidad de Guayaquil, Guayas, Ecuador, Email: katia.ruizmo@ug.edu.ec
[3] Universidad de Guayaquil, Guayas, Ecuador, Email: fernando.ponceor@ug.edu.ec
[4] Universidad de Guayaquil, Guayas, Ecuador, Email: jorge.chabusav@ug.edu.ec

Abstract. This study harnesses advanced time series models ARIMA, ETS, and SARIMA, coupled with neutrosophic statistics, to forecast unemployment trends through interval-based predictions. Transforming these predictions into neutrosophic forms enables the quantification of indeterminacy, providing a nuanced interpretation of potential economic scenarios. The integration of neutrosophic statistics enhances the interpretative power and accuracy of these models, offering a deeper insight into the inherent uncertainties of economic forecasting. The approach reveals not only the variabilities and potential outcomes within the unemployment rates but also strengthens the decision-making processes by presenting data that encompass both precision and indeterminacy. This paper underscores the importance of advanced statistical methods in economic predictions, suggesting further exploration into other economic metrics and advocating for a broader application of neutrosophic statistics to enhance the reliability of economic forecasting across diverse contexts.

Keywords: Neutrosophic Statistics, Time Series Forecasting, Economic Forecasting, Uncertainty Quantification.

1 Introduction

Unemployment stands as a crucial economic issue, receiving extensive attention in scholarly discussions. It is defined as a condition where individuals actively seeking employment for more than three months remain unable to secure a job. The unemployment rate quantifies this by expressing the proportion of unemployed individuals relative to the total labor force. This economic issue not only undermines living standards but also contributes to rising crime rates and insecurity, particularly noted in places like Ecuador [1].

Time series data, characterized by the dependency of successive observations, supports the analysis of such phenomena by acknowledging the sequence in which data appears. Time series modeling is widely applied in various fields like sales, meteorology, and inventory management, proving critical in scenarios involving uncertainty about the future. These models are particularly effective in forecasting [2].

Predicting time series data, especially when expressed in intervals rather than specific numbers, allows for more accurate yet uncertain outcomes [3]. This method falls under Neutrosophic Statistics, which incorporates interval data into statistical predictions. Enhancing the precision of these forecasts involves integrating multiple models, a growing and significant research area, especially applied to predicting unemployment rates for the future [4].

Unemployment remains a critical economic issue that garners extensive attention in scholarly discussions, largely due to its profound impact on societal welfare and economic stability. Defined as the condition in which individuals who are actively seeking employment for more than three months remain jobless, the unemployment rate quantifies this phenomenon, reflecting the proportion of the unemployed within the total labor force. The ramifications of high unemployment are severe, undermining living standards and contributing to increased crime rates and insecurity, with notable effects observed in regions like Ecuador [5]. Utilizing time series data, which acknowledges the sequential dependency of observations, enhances the analysis of such phenomena. Time series modeling, widely applied across various domains such as sales, meteorology, and inventory management, proves indispensable in scenarios laden with future uncertainties.

This paper explores the application of predictive models enhanced by neutrosophic statistics to project unemployment rates, offering a novel approach in the realm of economic forecasting by integrating interval-based predictions with a neutrosophic framework to manage and interpret uncertainty more effectively.

2 Preliminaries

2.1 Time series and Neutrosophic Statistics

Time series analysis involves the examination of data points arranged in chronological order. This methodology is crucial for understanding trends, cycles, and seasonal variations inherent in various datasets across time. Time series can be represented as [6]:

$$Y_t = f(t) + \varepsilon_t \quad (1)$$

Where:
Y_t is the is the value of the series at time t,
$f(t)$ represents the deterministic components like trends or seasonal effects,
ε_t is the random error component.
To predict future values of a time series and address the uncertainty of these predictions, interval forecasting is used. Interval predictions provide a range (interval) within which future observations are expected to fall, rather than pinpointing a single value. This can be formulated in the equation editor as [7]:

$$\hat{Y}_{(t+h|t)} = \hat{f}(t+h) \pm z * \hat{\sigma} \quad (2)$$

Where:
$\hat{Y}_{(t+h|t)} =$ is the predicted value at time $t+h$, based on the information up to time t,
f̂(t+h) is the predicted deterministic component,
z is the z-score from the normal distribution corresponding to the desired confidence level,
$\hat{\sigma}$ is the estimated standard deviation of the forecast errors.

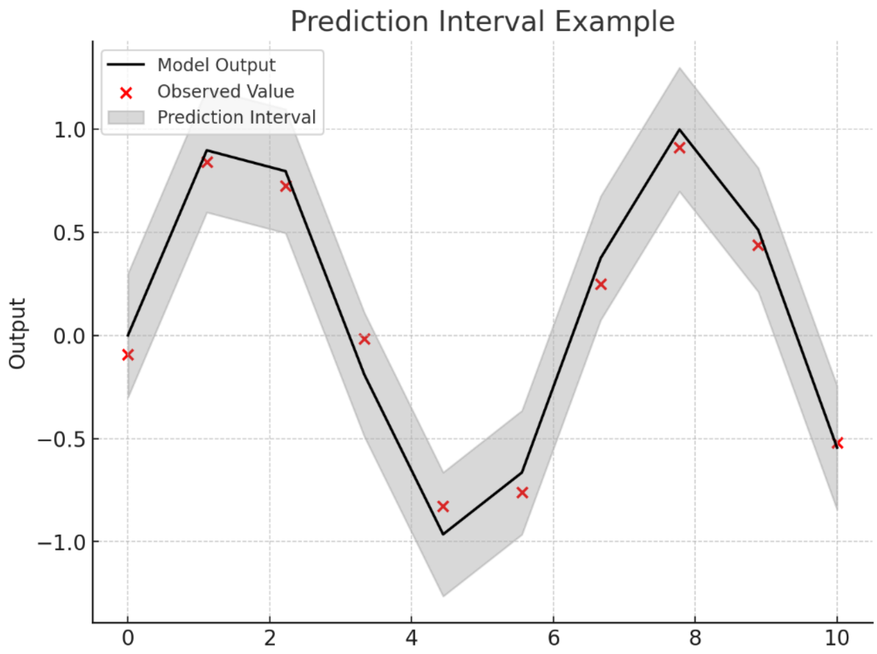

Figure 1. Components of prediction intervals around model outputs [8].

This approach to forecasting accommodates the inherent uncertainty in future predictions and provides a more realistic representation of the expected outcomes, making it invaluable in fields such as economics, finance, and environmental science.
To enhance this model using neutrosophic statistics, which allows for handling data ambiguity and indeterminacy more effectively, the interval can be transformed into a neutrosophic number. This transformation involves expanding the classical interval to include an indeterminacy component, reflecting the uncertainty and imprecision inherent in real-world data. The neutrosophic treatment of the interval is as follows [9]:

$\hat{f}(t+h) - z * \hat{\sigma} \times (\hat{f}(t+h) + z) \cdot I_N$ (3)

Here, I_N represents the indeterminacy factor associated with the prediction, where $I_N \in [I_l, I_u]$, This notation introduces the bounds of indeterminacy.

I_l (lower indeterminacy) and I_u (upper indeterminacy), which define the range of possible deviations due to uncertain elements affecting the forecast.

This refined representation not only aligns with the principles of neutrosophic statistics but also provides a more nuanced and realistic portrayal of the uncertainties inherent in time series forecasting. It makes the prediction interval more robust and informative, particularly useful in scenarios where data ambiguity and fuzziness are prevalent. This approach is instrumental for researchers and practitioners in fields where data quality and precision are variable and often not strictly deterministic.

2 Material and Methods

The analysis presented in this article is based on data sourced from the National Survey of Employment, Unemployment, and Underemployment (ENEMDU), provided by the National Institute of Statistics and Censuses (INEC). This study integrates advanced time series models to enhance the accuracy and understanding of labor trends. Specifically, it combines the established models such as ARIMA[10], the Exponential Smoothing Model (ETS)[11] , and SARIMA (Seasonal ARIMA)[12]. The fusion of these methods is achieved through the use of neutrosophic means, an approach that adeptly handles the uncertainty and indetermination inherent in predictions. This integration not only enhances the robustness of the predictive models but also offers a more profound framework for interpreting the complex dynamics of the labor market[13,14].

The neutrosophic mean, denoted as X_n, is calculated by considering the neutrosophic inclusion $I_N$that belongs to the interval $[I_l, I_u]$. This mean consists of two main elements: X_l, which is the mean of the lower part of the neutrosophic samples, and X_u, which is the mean of the upper part. The respective definitions are:

$$X_l = \frac{\sum_{i=1}^{n_l} x_{il1}}{n_l} \quad (4)$$

$$X_u = \frac{\sum_{i=1}^{n_u} x_{iu}}{n_u} \quad (5)$$

where n_l and $n_u$represent the number of elements in the lower and upper parts of the neutrosophic samples, respectively. Therefore, the neutrosophic mean X_n, is expressed as the sum of X_l and X_u, adjusted by the interval of indetermination I_n:

$$X_N = X_l + X_u I_N;\ I_N \in [I_l, I_u] \quad (6)$$

$I_{l,}$=0, and I_u

$$I_u = \frac{X_u - X_l}{X_u} \quad (7)$$

3 Results

we applied the three methods—ARIMA, ETS, and SARIMA—to the dataset, yielding interval-based predictions. These intervals were then transformed into neutrosophic forms, which effectively capture the measures of indeterminacy associated with each prediction. This transformation allows for a nuanced interpretation of the data, emphasizing the inherent uncertainties within the labor market forecasts. For detailed interval and indeterminacy values, refer to Table 1. This table showcases how each model's predictions are expressed in neutrosophic terms, providing a clear depiction of the range and reliability of the forecasts.

	First Quarter 2024	Second Quarter 2024	Third quarter 2024
ARIMA	[2.97, 4.25]	[2.72, 4.49]	[2.52, 4.69]
ETS	[3.05, 3.72]	[2.91, 3.58]	[2.78, 3.45]
SARIMA	[2.87, 3.30]	[3.16, 3.61]	[2.54, 3.01]
Media	[2.96, 3.76]	[2.93,3.89]	[2.61,3.72
Neutrosophic forms	2.96+3.76I; $I \in$ [0,0.21]	2.93+3.89I; $I \in$ [0,0.247]	2.61+3.72I; $I \in$ [0,0.298]

Table 1: Prediction Intervals and Neutrosophic Forms for Time Series Models.

Table 1 presents the prediction intervals and corresponding neutrosophic forms for three different time series models: ARIMA, ETS, and SARIMA, across three future quarters of 2024. These intervals reflect the range of potential outcomes, illustrating the inherent uncertainty in our predictions. The neutrosophic forms further quantify this uncertainty, highlighting the degree of indeterminacy associated with each interval. This representation allows for a more nuanced understanding of the forecasted data, facilitating more informed decision-making in uncertain conditions.

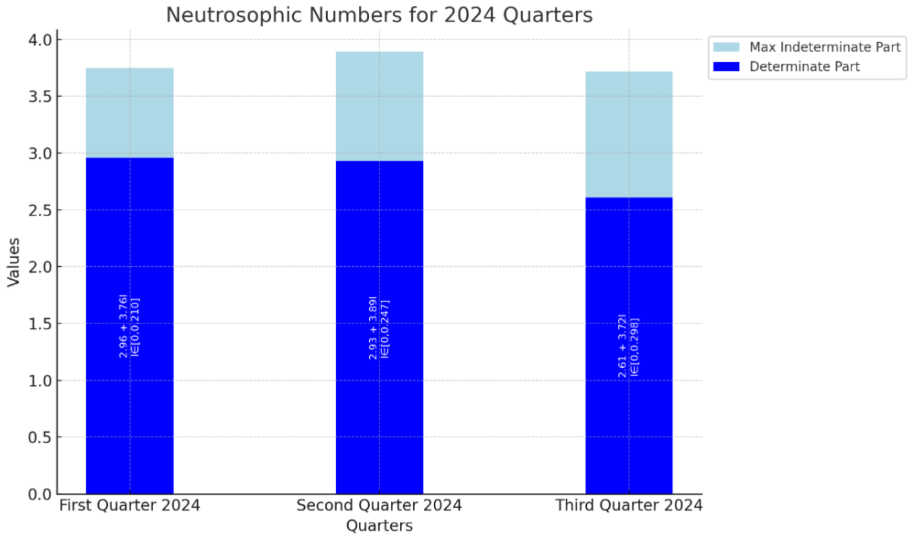

Figure 2. Neutrosophic Number Representation

Figure 2 illustrates the neutrosophic representation of prediction intervals for the first three quarters of 2024, clearly depicting an increasing trend in uncertainty as the year progresses. This rise in indeterminacy, captured through the enlarging segments of the neutrosophic intervals, is crucial for understanding the dynamics at play within the environmental data being analyzed.

The increase in the indeterminate component of the neutrosophic numbers suggests a growing complexity or variability in the underlying data as the year advances. This could be attributed to several factors, including seasonal variations, changes in data collection methodologies, or external economic or environmental impacts that become more pronounced over time.

From a decision-making perspective, this trend underscores the need for adaptive strategies that can accommodate an expanding range of outcomes. It also highlights the importance of continuous monitoring and updating of predictive models to better align with the evolving data landscape, ensuring that decision-making remains robust in the face of increasing uncertainty.

Such observations not only validate the utility of incorporating neutrosophic statistics in the analysis of time series data but also emphasize the critical role of these techniques in enhancing our comprehension of uncertainty in predictive modeling.

Conclusion

This study has demonstrated the effective integration of ARIMA, ETS, and SARIMA models, enhanced by neutrosophic statistics, to forecast unemployment trends through interval-based predictions. By transforming these intervals into neutrosophic forms, we were able to capture and quantify the underlying indeterminacy, providing a more nuanced view of the predictive landscape. This approach not only addresses the uncertainties inherent in economic forecasting but also enhances the interpretative power of the results, offering a more comprehensive understanding of potential future scenarios. The utilization of neutrosophic statistics has proven instrumental in refining the precision of these predictions, allowing for a detailed depiction of uncertainty and variability that traditional models might overlook.
Given the promising results observed in this study, future research could explore the extension of neutrosophic statistics to other areas of economic forecasting, such as inflation rates, GDP growth, or market volatility. Additionally, further refinement of the neutrosophic models to include more dynamic elements of indeterminacy could provide even greater accuracy and reliability in predictions. It would also be beneficial to conduct comparative

studies across different economies to validate the effectiveness of neutrosophic statistics in diverse economic contexts. Such investigations would not only enhance our understanding of neutrosophic methods but also potentially lead to the development of standardized approaches for handling uncertainty in economic time series analysis.

References

[1]. Tulcanaza-Prieto, A. B., Báez Salazar, P. A., & Aguilar-Rodríguez, I. E. (2023). Determinants of Youth Unemployment in Ecuador in 2019. Economies, 11(2), 59.

[2]. Cerqueira, V., Torgo, L., & Mozetič, I. (2020). Evaluating time series forecasting models: An empirical study on performance estimation methods. Machine Learning, 109(11), 1997-2028.

[3]. Cruzaty, L. E. V., Tomalá, M. R., & Gallo, C. M. C. (2020). A neutrosophic statistic method to predict tax time series in Ecuador. Neutrosophic Sets and Systems, 34, 33-39.

[4]. Zhongda, T., Shujiang, L., Yanhong, W., & Yi, S. (2017). A prediction method based on wavelet transform and multiple models fusion for chaotic time series. Chaos, Solitons & Fractals, 98, 158-172.

[5]. Munguía-Cortés, J. A., & Torres-Valdovinos, M. M. (2018). Social representations about employment and unemployment of university students from Quito, Ecuador. Revista Reflexiones, 97(2), 7-22.

[6]. Kočenda, E., & Černý, A. (2015). Elements of time series econometrics: An applied approach. Charles University in Prague, Karolinum Press.

[7]. Chatfield, C. (2001). Prediction intervals for time-series forecasting. Principles of forecasting: A handbook for researchers and practitioners, 475-494.

[8]. Shrestha, D. L., & Solomatine, D. P. (2006). Machine learning approaches for estimation of prediction interval for the model output. Neural networks, 19(2), 225-235.

[9]. Smarandache, F., & Aslam, M. (Eds.). (2023). Cognitive Intelligence with Neutrosophic Statistics in Bioinformatics. Elsevier.

[10]. Khan, S., & Alghulaiakh, H. (2020). ARIMA model for accurate time series stocks forecasting. International Journal of Advanced Computer Science and Applications, 11(7).

[11]. Seong, B. (2020). Smoothing and forecasting mixed-frequency time series with vector exponential smoothing models. Economic Modelling, 91, 463-468.

[12]. Dimri, T., Ahmad, S., & Sharif, M. (2020). Time series analysis of climate variables using seasonal ARIMA approach. Journal of Earth System Science, 129, 1-16.

[13]. Saqlain, M., Garg, H., Kumam, P., & Kumam, W. (2023). Uncertainty and decision-making with multi-polar interval-valued neutrosophic hypersoft set: A distance, similarity measure and machine learning approach. Alexandria Engineering Journal, 84, 323-332.

[14]. Abdullah, W. (2024). A Study of Neutrosophic Sets and Deep Learning Models for Breast Cancer Classification. Multicriteria Algorithms with Applications, 3, 50-59.

Received: Feb 12, 2024. Accepted: April 28, 2024