Editorial

New Kind of Statistical Methods

Cagdas Hakan Aladag

Department of Statistics, Hacettepe University, Ankara, Turkey

Florence Nightingale said that the Statistics is the most important science in the whole world: For upon it depends the practical application of every other science and every art: The one science essential to all political and social administration, all education, all organization based on experience, for it only gives results of our experience. Basic forms of statistics have been used since the beginning of civilization. It is obvious that we need statistical methods since Statistics appears in almost all areas of science, technology, research and wherever data is obtained for the purpose of finding information. Statistics is the study of data and how it can be collected, analyzed and presented in order to answer questions pertaining to the world around us. Statistics was also described as the science of making conclusions in the presence of uncertainty. To deal with the uncertainty, statistical conventional methods based on probability theory have been utilized since 17th century. The improvements in computer technology over recent decades make it possible to use intelligent approaches for analyzing the data and dealing with the uncertainty (Aladag *et al*., 2014a). Using such techniques in Statistics has provided some important advantages (Yolcu *et al*., 2013). At the present time, a method based on the philosophy of intelligent approaches is a new kind of statistical methods.

In a scientific point of view, the question always stays the same but the answer keeps changing in order to go further. Statistics through the years has sought an answer for the question, how the data can be analyzed in the presence of uncertainty. The first answer was to use deterministic approaches before the early 17th century. Then, these approaches have become inadequate and the probability theory has become the fundamental technique in the development of Statistics (Stigler, 1986). The methods based on probability theory have been utilized since 17th century (Senesen, 2007). These methods were really a good answer. The conventional methods such as factor analysis, regression analysis and analysis of variance are still so important and are successfully used in many applications. However,

according to today's conditions, the answer is beginning to change. This change refers to an improvement in statistical methods to analyze real-world data better. In recent decades, intelligent techniques such as heuristic algorithms, Artificial Neural Networks (ANN) and fuzzy logic based systems have been proposed to analyze the data due to the improvements in computer technology (Aladag, 2011; Celikyilmaz and Turksen, 2009; Zhang *et al*., 1998). Although the conventional methods based on probability theory have important advantages, they include some major drawbacks. To get over these problems, the intelligent methods have been utilized in the literature. Hence, using intelligent techniques instead of the conventional ones provides some important advantages such as eliminating restrictions arising from some assumptions, an increase in prediction accuracy and applicability to real-world problems (Aladag and Egrioglu, 2012). At the present time, to be unaware of usage of the intelligent methods in data analysis is not acceptable.

It should be noted that making a direct comparison between the conventional methods and the intelligent techniques is nonsense. It would be a crucial mistake to think these approaches separately. They are pieces of a whole since they are just different answers for the same question. Instead of comparing them against each other, both can be used together to reach better results. In other words, a better answer can be obtained by taking advantage of both approaches. Thus, the problems included in these approaches can be dealt with and better methods can be constructed. A key point is that a simple hybrid method, which combines some of these approaches, should not be considered only as a better method for data analysis. Instead of this, philosophies of these approaches should be combined. To do this, the philosophies of these approaches have to be known well. For example, a standard ANN model for time series forecasting do not have a Moving Average (MA) term. This method works like an Autoregressive (AR) model because of the characteristics of the method. A

hybrid method can be generated by combining MA model and ANN method in order to analyze time series which have both AR and MA structures. It is expected that a hybrid approach will give more accurate forecasts than those obtained from MA models or standard ANN models (Egrioglu *et al*., 2013). However, it would be wiser to generate a new ANN model which includes MA term. To do this, a new *ANN architecture that includes MA term* should be constructed and may be a new training algorithm should be improved to train this new architecture. Such a forecasting approach will produce very accurate results for real-world time series including both AR and MA structures. Another example can be given for fuzzy logic and probability theory based methods. Probability theory concerns about random variables which take crisp values. In a fuzzy system, variables take fuzzy values. Instead of using a hybrid method combines a fuzzy logic based method and a method based on probability theory, a new *probabilistic fuzzy approach* can be generated. To do this, may be the definition of *random fuzzy variable* can be described firstly.

To sum up, statistical methods are evolving and improving in accordance with the needs of the day. Thanks to the intelligent methods, more effective statistical approaches have been recently developed in the literature (Aladag *et al*., 2014b). For future studies, the following issues have to be considered carefully:

- The intelligent approaches have also some crucial problems in spite of their success in data analysis. Some of these problems could be dealt with by utilizing the probability theory and the conventional approaches
- The intelligent techniques can generally produce better results than those obtained from the conventional methods. However, it does not mean that the conventional methods are not useful or unnecessary. Theory of the conventional approaches should be well understood since new methods such as the intelligent ones can be generated only if shortcomings and disadvantages of the conventional methods are known well. In addition, the conventional approaches can give desired results in some cases. Furthermore, the theory of the conventional approaches should be used as a guide to develop new methods
- All intelligent techniques are not based on the probability theory. However, this does not mean that the intelligent approaches are not based on any theory. These approaches work on the basis of advanced strategies. These strategies should be coded in a proper computer language. This requires the art of computer programming
- The conventional methods based on probability theory and the intelligent methods can be combined in order to develop new efficient data analysis techniques. When such a technique is being constructed, the philosophies of these methods should be taken into account as mentioned before. Therefore, the philosophies of both approaches have to be known well

At the present time, the intelligent techniques have to be utilized in Statistics. Ignoring the importance of the intelligent methods in data analysis would be a great mistake. Of course the intelligent methods have some problems as with the conventional methods based on the probability theory (Aladag *et al*., 2012). As G.E.P. Box said," All models are wrong, but some are useful." The question, how the data can be analyzed better, will stay the same while the answer will be changing. We, researchers and practitioners, will keep seeking a better answer for this question. In other words, the methods such as deterministic, methods based on probability theory, intelligent or hybrid approaches for data analysis will continue to evolve and improve. Also, new kinds of statistical methods will be proposed in the future.

References

Aladag, C.H. and E. Egrioglu, 2012. Advanced Time Series Forecasting Methods. In: Advances in Time Series Forecasting, Aladag, C.H. and E. Egrioglu (Eds.), Bentham Science Publishers, Oak Park, ISBN-10: 1608053733.

Aladag, C.H., 2011. A Hybrid Intelligent Technique Combines Neural Networks and Tabu Search Methods for Forecasting, Computer Search Algorithms. 1st Edn., Nova Publisher, ISBN: 978-1-61209-043-6.

Aladag, C.H., E Egrioglu, U. Yolcu and V.R. Uslu, 2014a. A high order seasonal fuzzy time series model and application to international tourism demand of Turkey. J. Intell. Fuzzy Syst., 26: 295-302. DOI: 10.3233/IFS-120738

Aladag, C.H., E. Egrioglu and U. Yolcu, 2014b. Robust multilayer neural network based on median neuron model. Neural Comput. Applic., 24: 945-956. DOI: 10.1007/s00521-012-1315-5

Aladag, C.H., E. Egrioglu, U. Yolcu and A.Z. Dalar, 2012. A new time invariant fuzzy time series forecasting method based on particle swarm optimization. Applied Soft Comput., 12: 3291-3299. DOI: 10.1016/j.asoc.2012.05.002

Celikyilmaz, A. and I.B. Turksen, 2009. Modeling Uncertainty with Fuzzy Logic: With Recent Theory and Applications. 1st Edn., Springer, Berlin, ISBN-10: 3540899235, pp: 400.

Egrioglu, E., C.H. Aladag and U. Yolcu, 2013. Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks. Expert Syst. Applic., 40: 854-857. DOI: 10.1016/j.eswa.2012.05.040

Senesen, U., 2007. Statistics: Understanding Behind the Numbers. 1st Edn., Lıteratur Press, ISBN: 975-04-0283-9.

Stigler, S.M., 1986. The History of Statistics: The Measurement of Uncertainty before 1900. 1st Edn., Belknap Press/Harvard University Press, ISBN: 0-674-40341-X.

Yolcu, U., C.H. Aladag, E. Egrioglu and V.R. Uslu, 2013. Time-series forecasting with a novel fuzzy time-series approach: An example for Istanbul stock market. J. Stat. Comput. Simulat., 83: 597-610. DOI: 10.1080/00949655.2011.630000

Zhang, G., B.E. Patuwo and Y.M. Hu, 1998. Forecasting with artificial neural networks:: The state of the art. Int. J. Forecast., 14: 35-62. DOI: 10.1016/S0169-2070(97)00044-7