ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 06, Issue, 01, pp. 6444-6446, January, 2016

International Journal of
DEVELOPMENT RESEARCH

Full Length Review Article

A SURVEY ON SAMPLING THEORY

***Kalpanapriya, D. and Sujatha, V.**

Department of Mathematics, School of Advanced Sciences, VIT University,Vellore-632 014, India

ARTICLE INFO

Article History:

Received 17th October, 2015
Received in revised form
26th November, 2015
Accepted 11th December, 2015
Published online 31st January, 2016

Key Words:

Fuzzy sets, Testing hypotheses,
Analysis of variance.

ABSTRACT

A large number of developments have been made so far in Fuzzy sets, and Statistical techniques which played a vital role in sampling theory and Analysis of Variance (ANOVA). In spite of growing literature, the development of an application of statistical techniques in fuzzy analysis, a review in need and it is felt for a more systematic insight into the potentialities of cross fertilization between Statistics and Fuzzy Logic. In this article we review papers which is looking for several aspects of study design, data handling, statistical analysis on fuzzy set theory and the results.

INTRODUCTION

The statistical hypotheses testing under fuzzy environments has been studied by many researchers using the fuzzy set theory concepts introduced by Zadeh (1965). The application by using fuzzy set theory to statistics has been studied (Manto *et al*. (1994) and Buckley (2005) and Viertl (2006)). The literature survey on fuzzy statistical data analysis in imprecise environments is summarised. These would be complement of the statistical guidelines recommended by several authors. Statistics is an application-oriented mathematical branch which deals with methods for collection, classification and analysis of data for drawing valid conclusions and making appropriate decisions. It is classified into two group's namely descriptive statistics and inferential statistics. Both the groups are used to transform data into knowledge that leads to better decision making. Descriptive statistics is used to understand the given data. Inferential statistics helps us to measure behaviour in samples to learn more about the behaviour in populations that are often too large or inaccessible. A process used to gain the inside information from the data in order to extract knowledge is known as statistical data analysis.

Corresponding author: Kalpanapriya, D.,
Department of Mathematics, School of Advanced Sciences, VIT University,Vellore-632 014, India.

Sampling theory and testing of hypothesis

Sampling theory is a study of relationships existing between a population and samples drawn from the population. It is concerned with estimating the parameters of the population from those of the sample and also with determining the precision of the estimate. It includes estimation, hypothesis testing, and analysis of relationships or forecasting.

Statistical hypothesis is usually considered as the principal instrument in statistical data research. Its main function is to suggest new experiments and observations. Statistical hypothesis testing or significance testing is a method for testing a claim or hypothesis about a parameter in a population, using data measured in a sample. Statistical hypothesis should be capable of being tested and it should state relationship between variables, if it happens to be a relational hypothesis. Statistical hypothesis testing plays an important role to find the significance of parameters for populations based on their samples, but the difficulty arises when we happen to examine the significance of the difference amongst more than two samples / populations at the same time involved in hypotheses testing, ANOVA (analysis of variance) is applied. In conventional hypotheses testing Paul New bold *et al*. (2007) and Devore (2008), considering samples are crisp and the significance test leads to the binary decision. In real life situations, the sample data cannot be recorded precisely always. So, impressive data sample may be got for testing hypotheses.

Arnold (1998) discussed the fuzzy hypotheses testing with crisp data. The Neyman–Pearson type testing hypotheses were proposed by Son *et al.* (1992). Saade (1990) introduced fuzzy hypothesis testing with hybrid data. Casals and Gil (1989 & 1994) considered the Bayesian sequential tests for fuzzy parametric hypotheses from fuzzy information and discussed the operativeness of Neyman-Pearson tests with fuzzy information. Casals *et al.*.(1986) proposed the fuzzy decision problem of testing statistical hypotheses with fuzzy information. The fuzzy tests for hypotheses testing with vague data were proposed by Grzegorzewski (2000) and Watanabe and Imaizumi (1993). Niskanen (2001) discussed the applications of soft statistical hypotheses. Montenegro *et al.* (2001) discussed two sample hypothesis tests of means of a fuzzy random variable.

The statistical hypotheses testing for fuzzy data by proposing the notions of degrees of optimism and pessimism was studied by Wu (2005). Charles J. Geyer *et al*. (2005) discussed fuzzy and randomized confidence intervals and p-values. Akbari and Rezaei (2009) investigated a bootstrap method for inference about the variance based on fuzzy data. Jamkhaneh and Nadi Ghara (2010 & 2011) proposed testing statistical hypotheses for comparison means with vague data and statistical hypotheses for compare fuzzy test statistic. Viertl (2006 and 2011) investigated some methods to construct confidence intervals and statistical tests for fuzzy data. Wu (2009) proposed some approaches to construct fuzzy confidence intervals for the unknown fuzzy parameter. Arefi and Taheri (2011) developed an approach to test fuzzy hypotheses upon fuzzy test statistic for vague data. A new approach to the problem of testing statistical hypotheses for fuzzy data using the relationship between confidence intervals and testing hypotheses is introduced by Chachi *et al*. (2012). Miyandoab *et al*. (2012) proposed the fuzzy hypothesis testing with fuzzy data by using fuzzy p-value. A new test procedure of one- sample statistical hypotheses tests for means in normal population with interval data is proposed by Kalpanapriya (2012).

Akbari (2012) developed statistical hypotheses testing in the fuzzy environment. Blanco-fernandez *et al*. (2013) in random fuzzy sets, a mathematical tool to develop statistical fuzzy data analysis aims to review a significant part of the recent literature concerning the statistical data analysis with fuzzy data and is developed around the concept of random fuzzy numbers/sets.

Alizadeh *et al*. (2013) investigated the fuzzy hypothesis testing with vague data using likelihood ratio test. A method for constructing fuzzy test statistics with application was introduced by Assel S. Mohammad (2013). Tests of statistical hypotheses with crisp data using small samples are extended to with membership function of the fuzzy sets was developed by Pandian (2014). Tests of fuzzy hypotheses based on linear correlation using small sample data with membership grades are presented which are tests of characteristics or attributes was proposed by Pandian (2014).

Analysis of variance

ANOVA is a particular form of statistical hypothesis testing deeply used in the analysis of experimental data. The ANOVA technique was originally used in the analysis of agricultural research data. ANOVA is essentially a procedure for testing the difference among different groups of data for homogeneity. In a set of data, there may be variation between samples and also within sample items. ANOVA is a method of analyzing the variance to which a response is subject into its various components corresponding to various sources of variation since it consists in splitting the variance for analytical purposes.

ANOVA is a statistical hypothesis testing method used when the population samples are two or more. Kempton (1984) discussed the inadequacies of this approach, summarized the alternatives available at that time and noted that all of these approaches could be classified as multiplicative models. In traditional statistical testing by Paul *et al.* (2007) and Devore (2008), the observations of sample are crisp and a statistical test leads to the binary decision. Due to insufficient knowledge and information, the data sometimes cannot be recorded or collected precisely in the real life situation.

Armstrong (2000) aimed primarily at eye care practitioners who are undertaking advanced clinical research, and who wish to apply analysis of variance (ANOVA) to their data. Smith (2005) developed a general formulation that encompasses all of these methods is described, and then individual methods are considered in detail. Bharathi (2010) found the smallest set of genes that can ensure highly accurate classification of cancer from micro array data by using supervised machine learning algorithms. Jean Ashby (2011) present a research study which compared student success in a Developmental Math course offered in three different learning environments (online, blended, and face-to-face). One factor ANOVA test of population with crisp data is extended to with membership function of the fuzzy sets over the samples of the populations by Pandian (2013).

Konishi *et al*. (2006) proposed the method of ANOVA for the fuzzy interval data by using the concept of fuzzy sets. Hypothesis testing of one factor ANOVA model for fuzzy data was proposed by Wu (2007) by using the h-level set and the notions of pessimistic degree and optimistic degree by solving optimization problems. Tingley (2012) studied a Bayesian ANOVA scheme for calculating climate anomalies, with applications to the instrumental temperature record. Gil González-Rodríguez *et al*. (2012) developed an approach to the Fuzzy ANOVA test for analyzing fuzzy data. A new statistical technique for testing the fuzzy hypotheses of one factor ANOVA model using their samples having fuzzy data is propose without using h-level concept by Kalpanapriya (2012). Alireza Jiryaei *et al*. (2013), have proposed one-way ANOVA technique in fuzzy environment in which the least squares method is employed . Nourbakhsh *et al*. (2013) have discussed one-way ANOVA for a sample having fuzzy observations.

REFERENCES

Akbari, M.G. 2012. Statistical Hypotheses Testing in the Fuzzy Environment, *Journal of Uncertain Systems*, Vol.6, No.3, pp.186-199.

Akbari, M.G. and Rezaei, A. 2009. Bootstrap statistical inference for the variance based on fuzzy data, *Austrian Journal of Statistics*, Vol.38, No.2, pp. 121-130.

Arefi, M. and Taheri, S.M. 2011. Testing fuzzy hypotheses using fuzzy data based on fuzzy test statistic, *Journal of Uncertain Systems*, Vol. 5, No.1, pp. 45-61.

Arnold, B.F. 1995. Statistical tests optimally meeting certain fuzzy requirements on the power function and on the sample size, *Fuzzy Sets and Systems*,Vol. 75, pp.365–372.

Arnold, B.F. 1998. Testing fuzzy hypotheses with crisp data, *Fuzzy Sets and Systems*, Vol. 94, pp.323–333.

Assel S. Mohammad. 2013. A Method for Constructing Fuzzy Test Statistics with Application, *Mathematical Theory and Modeling* ,Vol.3, No.13,pp. 24-29.

Baloui Jamkhaneh, E. and Nadi Ghara, A. 2010. Testing statistical hypotheses for compare means with vague data, *International Mathematical Forum*, Vol.5, No.13, pp. 615 – 620.

Baloui Jamkhaneh, E. and Nadi Ghara, A. 2011. Testing statistical hypotheses for compare fuzzy test statistic, *Journal of Uncertain Systems*, Vol.5 , pp.45-61.

Blanco-fernandez, A., Casals, M.R. Colubi, A. Corral, N. Garcia-barzana, M. Gil, M. A. Gonzalez-rodriguez, G. Lopez, M. T. Lubiano, M. A. Montenegro, M. Ramos-guajardo, A. B. Delarosadesaa, S. and Sinova, B. 2013. Random fuzzy sets: a mathematical tool to develop statistical fuzzy data analysis, *Iranian journal of fuzzy systems*, Vol. 10, No. 2, pp. 1-28.

Buckley, J.J. 2005. Fuzzy Statistics, *Springer-Verlag*, New York.

Casals, M.R. and Gil, M.A. 1989. A note on the operativeness of Neyman–Pearson tests with fuzzy information, *Fuzzy Sets and Systems,* Vol.30, Issue 2, pp. 215–220.

Casals, M.R. and Gil, P. 1994. Bayesian sequential test for fuzzy parametric hypotheses from fuzzy information, *Information Science,* Vol.80, pp. 283–298.

Casals, M.R., Gil, M.A. and Gil, P. 1986. The Fuzzy decision problem: an approach to the problem of testing statistical hypotheses with fuzzy information, *European Journal of Operational Research* ,Vol.27 ,pp. 371–382.

Ch. Son, J., Song, I. and Kim, H.Y. 1992. A fuzzy decision problem based on the generalized Neymen– Pearson criterion, *Fuzzy Sets and Systems*, Vol.47 , pp.65–75.

Chachi, J., Taheri, S.M. and Viertl, R. 2012. Testing statistical hypotheses based on fuzzy confidence intervals, Forschungsbericht SM-2012-2, *Technische Universitat Wien,* Austria.

Charles J, Geyer, and Glen D. Meeden (2005), Fuzzy and Randomized Confidence Intervals and P-Values, *Statistical Science*, Vol. 20, No. 4, pp. 358–366.

Devore, J.L. 2008. Probability and Statistics for Engineers, Seventh edition, Cengage Learning, New Delhi.

Grzegorzewski, P. 2000. Testing statistical hypotheses with vague data, *Fuzzy Sets and Systems,* Vol.112, pp. 501–510.

Klir, G. and Bo Yuan, 2008. Fuzzy sets and fuzzy logic, *Theory and Applications*, Prentice- Hall, New Jersey.

Konishi, M., Okuda, T. and Asai, K. 2006. Analysis of variance based on fuzzy interval data using moment correction method, *International Journal of Innovative Computing*, Information and Control, Vol. 2, No.1, pp. 83-99.

Kruse, R. 1984. Statistical estimation with linguistic data, *Information Sciences*, Vol. 33, pp.197–207.

Manto, K.G., Woodbury, M.A. and Tolley, H.D. 1994. Statistical Applications Using Fuzzy Sets, John Wiley & Sons, New York.

Miyandoab, M., Arzideh, K. and Davood Farbod, 2012. Fuzzy hypothesis testing with fuzzy data by using fuzzy p-value, *Acta Universitatis Apulensis,*No.32,pp.293-307.

Moheb Alizadeh, H., Arshadi Khamseh, S. and Fatemi Ghomi, 2013. Fuzzy hypotheses testing with vague data using likelihood ratio test, *Soft computing*, Vol.17, No.9, pp.1629-1641.

Montenegro, M., Casals, M.R. Lubiano, M.A. and Gil, M.A. 2001. Two-sample hypothesis tests of means of a fuzzy random variable, *Information Sciences*, Vol. 133, pp. 89- 100.

Niskanen, V.A. 2001. Prospects for soft statistical computing: describing data and inferrings from data with words in the human sciences, *Information Sciences*, Vol.132, pp. 83– 131.

Paul Newbold, William L.Carlson and Betty Thorne 2007, Statistics for Business and Economics, *Pearson Education*, Eighth edition.

Saade, J.J. and Schwarzlander, H. 1990. Fuzzy hypothesis testing with hybrid data, *Fuzzy Sets and Systems*, Vol.35, pp.197–212.

Son, J. Ch., Song, I. and Kim, H.Y. 1992. A fuzzy decision problem based on the generalized Neymen-Pearson criterion, *Fuzzy Sets and Systems*, Vol. 47, pp.65–75.

Torabi, H. and Mirhosseini, S.M. 2009. The Most Powerful Tests for Fuzzy Hypotheses Testing with Vague Data, *Applied Mathematical Sciences*, Vol. 3,No. 33, pp.1619 – 1633.

Viertl, R. 2006. Univariate statistical analysis with fuzzy data, *Computational Statistics and Data Analysis*, Vol.51, pp. 33-147.

****