Available online at www.sciencedirect.com
SciVerse ScienceDirect
Fuzzy Sets and Systems 243 (2014) 70–83
FUZZY sets and systems
www.elsevier.com/locate/fss

Inclusion and exclusion hypothesis tests for the fuzzy mean

A.B. Ramos-Guajardo *, A. Colubi, G. González-Rodríguez

Department of Statistics, Operational Research and Mathematic Didactics, University of Oviedo, Spain

Received 16 November 2012; received in revised form 24 June 2013; accepted 25 June 2013

Available online 1 July 2013

Abstract

A hypothesis test for the inclusion of the expected value of a fuzzy-valued random element in a given fuzzy set is provided. The exclusion, or the empty intersection of the expected value of a random fuzzy set and a fuzzy set, is also tested, that allows us to define one-sided tests for this expected value without the need of considering any restrictive ordering. Both tests are developed by using a measure of the intersection between fuzzy sets. Asymptotic and bootstrap techniques are established. Some simulations are included to show the performance of the bootstrap approaches. Finally, the methodology proposed is applied on a real-life situation related to the field of sensorial analysis.

Keywords: Random fuzzy set; Hypothesis testing; Expected value; Bootstrap techniques; Sensorial analysis

1. Introduction

Many real-life experiments involve two different kinds of uncertainty: randomness or stochastic uncertainty, which is due to sampling, and fuzziness or non-stochastic uncertainty, which is due to partial ignorance, imprecision, aggregation processes, etc. Random fuzzy sets (for short RFSs [17]) have been introduced to model these situations associating fuzzy values with experimental outcomes.

This work is focused on those situations in which a random process generating fuzzy set-valued outcomes is considered, and classical statistics are applied on them. Specifically, the approach proposed is soundly applicable when the fuzzy sets generated by the random process either represent precise objects or they are imprecise descriptions or perceptions provided by an observer [4,11,17,19]. This includes, for instance, the use of fuzzy sets for describing the set of languages more or less well spoken by individuals in a population or a region, and also situations where the interest is centered on the random behavior of the observer who delivers fuzzy set-valued observations. This is the case of the description of the opinions or perceptions of an expert about the quality of a product, the quality of the trees, etc. The real-life situation analyzed in this work can be placed in this context.

Concerning imprecise data, hypothesis testing procedures and regression approaches have been recently introduced [2,5,6,9,10,18,19]. Most of the hypotheses tests have been developed in practice for centralization or dispersion

* Corresponding author.
E-mail addresses: ramosana@uniovi.es (A.B. Ramos-Guajardo), colubi@uniovi.es (A. Colubi), gil@uniovi.es (G. González-Rodríguez).

0165-0114/$ – see front matter
http://dx.doi.org/10.1016/j.fss.2013.06.015

measures, among which are pointed out hypotheses tests for the mean [6,11,12,15] and for the variance [13,18,19] of one or several RFSs.

The population mean of a random fuzzy set is also a fuzzy set. Some null hypotheses about this fuzzy mean could be established either based on the equality between this value and a hypothetical given fuzzy set [6,12] or, due to the imprecision gathered in fuzzy data, based on the inclusion of the fuzzy mean in a fuzzy set. The first aim of this paper is to test the null hypothesis that the population mean of an RFS is fully included in a given fuzzy set.

A hypothesis test for checking if the population mean of an RFS and a given fuzzy set have empty intersection will also be provided. This procedure will lead to define one-sided tests for the expected value of a random fuzzy set in a strict sense, that is, to analyze if the expected value of an RFS is greater or lower than a given fuzzy set having both sets empty intersection. In this line a one-side test for the expected value of RFSs has been developed in [14]. However, the last test was centered on simple RFSs (i.e., those taking on a finite number of different values) in a particular class and it considered a parameterized ranking function while the test proposed here is useful in a more general setting and it is not necessary to consider a priory any restrictive ordering.

To solve both tests the notion of intersection kernel introduced in [22] will be used. Asymptotic techniques will be applied due to the lack of realistic parametric distributions for RFSs. Some methods based on bootstrap techniques will also be developed in order to get a better performance for moderate sample sizes taking inspiration on the bootstrap tests for the fuzzy mean of an RFS (see, for instance, [6,9,15]).

Finally, the methodology will be applied in order to determine if the quality of a specific kind of blue cheese (called Gamonedo cheese) remains into the limits required by the Protected Designation of Origin (PDO) category by checking whether the mean value of the subjective opinions given by the expert tasters is either contained or not into these limits. Representing the feelings of the experts about the features of the cheese by means of precise numbers on a numerical scale looks arbitrary for such kind of data, since they underly no precise objective feature values. We try to account for this state of facts by considering trapezoidal fuzzy sets as an approximate rendering of such subjective perceptions.

The rest of the manuscript is organized as follows. In Section 2, some preliminaries about RFSs are gathered. The asymptotic procedures for testing the inclusion and the empty intersection of the expected value of a random fuzzy set with respect to a given fuzzy set are presented in Sections 3 and 4. Some special cases are analyzed in Section 5. A bootstrap procedure for solving the proposed tests in practice is stated in Section 6. Some simulations are carried out to show the empirical performance of the tests in Section 7 and their application is illustrated on a case study in Section 8. In Section 9 some remarks and open problems are provided.

2. Preliminary concepts

Let $\mathcal{F}_c(\mathbb{R}^p)$ denote the class of fuzzy sets $U : \mathbb{R}^p \to [0,1]$ such that $U_\alpha \in \mathcal{K}_c(\mathbb{R}^p)$ for all $\alpha \in (0,1]$, where $\mathcal{K}_c(\mathbb{R}^p)$ is the family of all non-empty compact convex subsets of $\mathbb{R}^p$. The α-levels of U are defined as $U_\alpha = \{x \in \mathbb{R}^p \mid U(x) \geqslant \alpha\}$ if $\alpha \in (0,1]$, and U_0 is the closure of the support of U.

The studies in next sections will be focused on the space of fuzzy sets

$$\mathcal{F}_c^2(\mathbb{R}^p) = \{U \in \mathcal{F}_c(\mathbb{R}^p):\ s_U \in \mathcal{L}^2(\mathbb{S}^{p-1} \times [0,1]), \lambda_p \times \lambda)\},$$

where s_U is the *support function* of U (see, for instance, [16]) defined such that $s_U(u,\alpha) = \sup_{v \in U_\alpha} \langle u, v\rangle$ for $(u,\alpha) \in \mathbb{S}^{p-1} \times [0,1]$, $\mathbb{S}^{p-1}$ is the unit sphere in $\mathbb{R}^p$, λ_p is the uniform surface measure on $\mathbb{S}^{p-1}$ and λ denotes the Lebesgue measure on $(0,1]$.

The usual arithmetic between fuzzy sets in the space $\mathcal{F}_c^2(\mathbb{R}^p)$ is based on Zadeh's extension principle which agrees levelwise with the Minkowski addition and the product by a real number for compact convex sets.

The metric used in the formalizations is based on the concepts of *generalized mid and spread*, defined in Trutschnig et al. [23] by using the support function of $U \in \mathcal{F}_c^2(\mathbb{R}^p)$ as $\operatorname{mid}_U(u,\alpha) = \operatorname{mid}_{U_\alpha}(u) = (s_U(u,\alpha) - s_U(-u,\alpha))/2$ and $\operatorname{spr}_U(u,\alpha) = \operatorname{spr}_{U_\alpha}(u) = (s_U(u,\alpha) + s_U(-u,\alpha))/2$, for $u \in \mathbb{S}^{p-1}$ and $\alpha \in (0,1]$.

Some properties of these functions can be found in [23]. If $\|\cdot\|_2$ denotes the usual functional L_2-norm with respect to the measure λ_p, then the *distance between two fuzzy sets* $U, V \in \mathcal{F}_c^2(\mathbb{R}^p)$ (see [23]) is defined by