Available online at www.sciencedirect.com
ScienceDirect
Fuzzy Sets and Systems 157 (2006) 2608 – 2613
FUZZY sets and systems
www.elsevier.com/locate/fss

Bootstrap techniques and fuzzy random variables: Synergy in hypothesis testing with fuzzy data

Gil González-Rodríguez, Manuel Montenegro, Ana Colubi, María Ángeles Gil*

Departamento de Estadística, I.O. y D.M., Facultad de Ciencias, Universidad de Oviedo, E-33071 Oviedo, Spain

Available online 5 June 2006

Abstract

In previous studies we have stated that the well-known bootstrap techniques are a valuable tool in testing statistical hypotheses about the means of fuzzy random variables, when these variables are supposed to take on a finite number of different values and these values being fuzzy subsets of the one-dimensional Euclidean space. In this paper we show that the one-sample method of testing about the mean of a fuzzy random variable can be extended to general ones (more precisely, to those whose range is not necessarily finite and whose values are fuzzy subsets of finite-dimensional Euclidean space). This extension is immediately developed by combining some tools in the literature, namely, bootstrap techniques on Banach spaces, a metric between fuzzy sets based on the support function, and an embedding of the space of fuzzy random variables into a Banach space which is based on the support function.

Keywords: Bootstrap; Embedding; Fuzzy random variables; Generalized metric between fuzzy sets; Support function of a fuzzy set

1. Introduction

Fuzzy random variables in Puri and Ralescu's sense were introduced ([14], see also Gil et al. [6] in this issue for a recent survey) as fuzzy set-valued random elements extending the notion of compact (and usually convex) random sets. The (fuzzy) expected value of a fuzzy random variable is defined level-wise in terms of the Aumann integral of random sets [2].

To test the statistical hypothesis that the fuzzy mean of a fuzzy random variable is given by a specific fuzzy set, some procedures have been established.

In particular, Körner [9], and Montenegro et al. [13] have presented asymptotic one-sample procedures. Montenegro et al. [11,12] have developed also asymptotic methods for the two-sample case, and Gil et al. [7] have discussed the asymptotic ANOVA test.

Körner's asymptotic developments [9] concern general fuzzy random variables (taking on any—either finite or infinite—number of values in the space of compact convex fuzzy sets of a finite-dimensional Euclidean space) and they

* Corresponding author. Fax: +34 98 510 3356.
E-mail address: magil@uniovi.es (M.Á. Gil).

0165-0114/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.fss.2003.11.021

are based on two key ideas, namely,

- the embedding of the space of fuzzy sets of a finite-dimensional Euclidean space into a certain Banach space of functions through the support function of a fuzzy set;
- the central limit theorems for Banach space-valued random elements by Araujo and Giné [1].

Montenegro et al.'s asymptotic one-sample studies [13] concern fuzzy random variables taking on a finite number of different values (in the space of compact convex fuzzy sets of the one-dimensional Euclidean space) and they are based on two key ideas, namely,

- an operational generalized metric between fuzzy subsets which allow to express the hypothesis about the fuzzy mean as a quadratic form of certain real-valued random variables;
- well-known results from large sample theory (more precisely, those taking advantages of the convenient properties of the consistent and asymptotically normal estimators).

Although the method by Körner is more widely applicable than Montenegro et al.'s one, this last one shows several valuable advantages, namely,

- in the one-dimensional case the metric considered by Montenegro et al. is more general, which allows us to choose suitable metrics for each statistical problem;
- the testing procedure is easy to apply, whereas Körner's method involves some population parameters which are usually unknown in practice;
- the assumption of the cardinality of the range of the fuzzy random variable to be finite is not too restrictive, since fuzzy random variables in most of real-life examples take on a finite number of different values.

Asymptotic tests above guarantee that the larger the sample size the closer the probability of error I type to the nominal significance level is, but this nominal level is only achieved for very large sample sizes. In previous studies we have adapted bootstrap techniques to approximate asymptotic techniques, since for bootstrap procedures a faster convergence speed is expected (and hence the nominal significance level would be mostly achieved for moderate and even small samples). Thus, in Montenegro et al. [13] we have stated that bootstrap techniques are very valuable in dealing with testing about the mean of a fuzzy random variable taking on a finite number of different values, so that the accuracy of bootstrap approaches is greater than that of asymptotic ones for most of cases we have examined.

In this paper we are going to prove that, as it has been asserted in the previous papers, the extension of the bootstrap study by Montenegro et al. to general fuzzy random variables can be immediately developed on the basis of the bootstrap developments for Banach space-valued random elements by Giné and Zinn [8]. In fact, we will consider an extra extension by incorporating the generalized metric by Körner and Näther [10].

The structure of the paper is as follows: in Section 2 we present some preliminaries on the fuzzy set-valued random elements we will deal with, and the generalized metric by Körner and Näther; in Section 3, after recalling the supporting results (an embedding theorem as well as a central limit theorem for Banach space-valued random elements), the announced extension is presented; in Section 4 we illustrate the application of the generalized test by means of an example; finally, a few remarks are included.

2. Preliminaries on fuzzy values

Consider the p-dimensional Euclidean space $\mathbb{R}^p$ and let $\|\cdot\|$ denote the usual associated norm. If $\mathcal{K}_c(\mathbb{R}^p)$ denote the class of the nonempty compact convex subsets of $\mathbb{R}^p$, then we will denote

$$\mathcal{F}_c(\mathbb{R}^p) = \{U : \mathbb{R}^p \to [0,1] \mid U_\alpha \in \mathcal{K}_c(\mathbb{R}^p) \text{ for all } \alpha \in [0,1]\}$$

(where U_α denotes the α-level of fuzzy set U for all $\alpha \in (0,1]$, and U_0 being the closure of the support of U).

$\mathcal{F}_c(\mathbb{R}^p)$ can be endowed with a semilinear structure by means of the sum and the product by a scalar based on Zadeh's extension principle [15] (see the overview by Gil et al. in this issue).

Given a probability space $(\Omega, \mathcal{A}, P)$, a mapping $\mathcal{X} : \Omega \to \mathcal{F}_c(\mathbb{R}^p)$ is said to be a *fuzzy random variable* (also called *random fuzzy set*, and referred to as FRV for short) if for all $\alpha \in [0,1]$ the set-valued mappings $\mathcal{X}_\alpha : \Omega \to \mathcal{K}_c(\mathbb{R}^p)$, defined so that $\mathcal{X}_\alpha(\omega) = (\mathcal{X}(\omega))_\alpha$ for all $\omega \in \Omega$, are random sets (that is, Borel-measurable mappings with the Borel σ-field generated by the topology associated with the well-known Hausdorff metric d_H on $\mathcal{K}(\mathbb{R}^p)$). Alternatively, an