

Similarity based one-sided tests for the expected value and interval data

Przemysław Grzegorzewski^{1,2} Ana Belén Ramos-Guajardo³

¹Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

²Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland

³Department of Statistics, Operational Research and Mathematics Teaching, University of Oviedo, Spain

Abstract

A novel way for addressing one-sided hypothesis tests for the mean of unable to observe real random variables from which we have information provided in terms of interval data is proposed. For this purpose, a measure of similarity between intervals is considered and the classical one-sided hypothesis test to check whether the similarity between the expectation of a random interval and a subset of a specific unbounded interval is high enough. Asymptotic and bootstrap techniques are applied in order to study the limit distribution of the proposed test statistic. Some simulations have been carried out in order to show the empirical behavior of the test.

Keywords: One-sided hypothesis test, expected value, interval data, hypothesis testing, bootstrap approach

1. Introduction

Most activity in human cognition is related to stating and verifying hypotheses. Testing hypotheses is also one of the primary purposes of statistical inference. In classical statistics, although we admit some uncertainty caused by randomness, all elements of the statistical model (i.e. data, hypotheses, test requirements etc.) should be given precisely. In particular, data used in hypothesis testing are usually considered as real numbers which seems to be quite natural in many applications. Actually, such results as weight, height, strength, temperature, and so on are real.

However, in many situations we cannot observe the experimental results precisely. It might be caused by many reasons, like measurement problems, inaccurate devices and subjective assessments that are more perceptions than actual measurements. In such cases an expectation to obtain the results in a form of precise real numbers is either unrealistic or requires unjustified subjectivity. Thus, here we deal with two kinds of uncertainty: besides randomness we also sustain some epistemic uncertainty connected with a data perception.

In situations like mentioned above, quite often a reasonable and well-grounded solution is to consider interval data. However, to handle such data we need new mathematical models, different than

used in the classical statistics, both to describe observations and to perform statistical reasoning. In this paper we suggest a novel approach to hypothesis testing based on the interval data. Our approach utilizes the concept of similarity well-established in reasoning admitting different sources of uncertainty. The problem will be addressed from a theoretical point of view although its application to different fields is to be tackled in the future.

The paper is organized as follows. In Section 2 we introduce a notation and basic concepts used for modeling interval data. We recall there also an idea of similarity based statistical test for random intervals proposed in [12]. In Section 3 we discuss how to generalize a classical one-sided hypothesis testing problem about the population mean for interval data. Our method utilizes a similarity measure which is a particular case of the well-known Tversky coefficient (see [15]). Next, in Section 4, we derive a one-sided statistical test for the mean and interval data. The properties of the suggested test are examined in Section 5 through the simulation studies including the bootstrap approach. Finally, some conclusions and remarks are given.

2. Preliminaries

2.1. Random intervals

A starting point of the majority of problems in classical statistics is a random sample V_1, \dots, V_n , where V_1, \dots, V_n are usually independent and identically distributed random variables from a distribution under study. If (Ω, \mathcal{A}, P) denote a probability space then each random variable $V_i : \Omega \rightarrow \mathbb{R}$ is a function which assigns to a random event a real number, hence the outcome of the experiment is traditionally a set of real values. Unfortunately, in many situations we cannot measure the results of the experiment precisely and the only thing we could do is to specify an interval for each outcome which contains the true measurement.

Denote by $\mathcal{K}_c(\mathbb{R})$ the family of all non-empty closed and bounded intervals in the real line \mathbb{R} . The formalizations required for the theoretical developments of this research are based on the parametrization (mid, spr) of the intervals, i.e. $A = [\text{mid } A \pm \text{spr } A]$ for $A \in \mathcal{K}_c(\mathbb{R})$, where $\text{mid } A \in \mathbb{R}$ is the midpoint (or center) of the interval A , while $\text{spr } A \geq 0$

is the spread (or radius) of A . This characterization has been shown to be a valuable tool for different statistical purposes (see, e.g., [2, 5]).

The arithmetic between intervals taking into account the (mid, spr) parametrization is settled as follows:

$$A_1 + \lambda A_2 = [(\text{mid } A_1 + \lambda \text{mid } A_2) \pm (\text{spr } A_1 + |\lambda| \text{spr } A_2)], \quad (1)$$

where $A_1, A_2 \in \mathcal{K}_c(\mathbb{R})$ and $\lambda \in \mathbb{R}$.

If the experiment outcomes are not reals but intervals, we need another mathematical model than a random variable to describe this random phenomenon. Let (Ω, \mathcal{A}, P) denote, as before, a probability space. A *random interval* is a Borel measurable mapping $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ w.r.t. the well-known Hausdorff metric on $\mathcal{K}_c(\mathbb{R})$ (see [10]). In addition, if $\text{mid } X, \text{spr } X \in L_1(\Omega, \mathcal{A}, P)$ the expected value of a random interval X is defined in terms of the Aumann expectation (see [1]) and it satisfies the equality $E([\text{mid } X \pm \text{spr } X]) = [E(\text{mid } X) \pm E(\text{spr } X)]$.

Thus, a random interval X might be seen as a perception of an unknown usual random variable V , called the *original* of X (see [9]). It is clear that also fuzzy sets [16] can be used for describing such a perception. However, in this work we will center our attention in those situations in which either the perception of the real random variable is an interval or the observation of the random experiment is any possible value of an interval. Of course, an interval could be perceived also as a special fuzzy set (i.e. a rectangular fuzzy number) where imprecision appears not through the degree of membership but the area of possible values. The extension of the problem to the framework of fuzzy sets with partial degrees of membership will be addressed in future studies.

Since we have no direct access to the originals and the only information about V is delivered by X , the statistical reasoning on the parameters of the underlying distribution would be also based on the analysis of random intervals. In particular, further on we will show how to construct statistical tests for one-sided hypotheses on the parameters of the distribution utilizing interval data modeled by the random intervals.

2.2. Similarity measures

In this section we will show different similarity measures that can be considered when dealing with sets.

A. B. Ramos-Guajardo [12] proposed a statistical test for the interval data to verify a hypothesis for the similarity between the expected value of a random interval and a fixed interval formulated as follows:

$$\begin{cases} H_0 : S_J(E(X), A) \geq d \\ H_1 : S_J(E(X), A) < d, \end{cases} \quad (2)$$

where $d \in [0, 1]$ is a fixed a priori chosen value and S_J denotes the Jaccard similarity measure defined

by

$$S_J(A, B) = \frac{\lambda(A \cap B)}{\lambda(A \cup B)}, \quad (3)$$

for any two intervals $A, B \in \mathcal{K}_c(\mathbb{R})$ and where $\lambda(A)$ stands for the Lebesgue measure of A .

It is clear that the more similar $E(X)$ to A is, the higher the value of d . Thus, it seems that the Jaccard measure (3) appears quite useful in formulating hypotheses for interval data which generalizes the classical two-sided hypothesis testing problem about the mean with crisp real data, i.e. $H_0 : m_V = m_0$ vs. $H_1 : m_V \neq m_0$, where m_V denotes the mean of the sample distribution and m_0 is a fixed real value. Unfortunately, (3) cannot be useful for the generalization of the one-sided hypothesis testing problem because of its symmetry.

One can find in the literature many coefficients used as similarity measures (see, e.g., [4] and references given there). Most of them are symmetric, as (3). However, Tversky in his seminal paper [15] noticed that although similarity has been viewed traditionally as a typical symmetric relation, there exist many examples of asymmetric similarities (some examples can be also found in [15]). Hence for such situations asymmetric coefficients are needed to quantify the amount of similarity of the objects under study. In this very paper he also suggested a general idea of the similarity measure given by

$$S_{\alpha, \beta}(A, B) = \frac{f(A \cap B)}{f(A \cap B) + \alpha f(A \setminus B) + \beta f(B \setminus A)}, \quad (4)$$

where $\alpha, \beta \geq 0$ are some fixed constants and $f(A)$ is a measure of a set A . One may easily notice that for any two intervals $A, B \in \mathcal{K}_c(\mathbb{R})$ by taking $\alpha = \beta = 1$ and assuming that $f \equiv \lambda$ we get the Jaccard measure (3). If $\alpha \neq \beta$ then (4) is no longer symmetric.

3. Interval data and one-sided hypotheses

Let us focus our attention on the generalization of the classical one-sided test for the mean, i.e. either $H'_0 : m_V \leq m_1$ vs. $H'_1 : m_V > m_1$, or $H''_0 : m_V \geq m_2$ vs. $H''_1 : m_V < m_2$, where m_V denotes, as before, the mean of the sample distribution while m_1, m_2 are fixed real values. It is obvious that so defined H'_0 is equivalent to $H'_0 : m_V \in (-\infty, m_1]$, while H''_0 is equivalent to $H''_0 : m_V \in [m_2, +\infty)$. Thus, in both cases the null hypothesis claims that the true mean belongs to the infinite interval. Or, in other words, one may perceive H'_0 as a statement that the true mean m_V is equal to some real value m less than or equal to m_1 . Analogously, H''_0 may be considered as a statement that the true mean m_V is equal to some real value m greater than or equal to m_2 .

All these intuitions given above indicate the way of the generalization required for the interval data. Namely, having interval data instead of $H'_0 : m_V \in (-\infty, m_1]$ we will consider the problem whether the

interval expected value $\mathbb{E}(X)$ is similar to any interval $B \subset L = (-\infty, m_1]$. Or, in other words, our goal may be perceived as a testing problem whether the similarity between $\mathbb{E}(X)$ and $\bigcup_{B \subset L} B$ is high enough.

The discussed problem is typical to situation when A is a subject of the comparison and B is the referent. In such case one naturally focuses rather on the subject of the comparison and hence the features of the subject are weighted more heavily than the features of the referent. Tversky notices that in this case consequently “similarity is reduced more by the distinctive features of the subject than by the distinctive features of the referent” (see [15]). This implies that the adequate version of the similarity measure (4) should have $\alpha > \beta$.

Having in mind all those remarks let us consider as a similarity measure the Tversky coefficient (4) with $\alpha = 1$ and $\beta = 0$. This way we obtain

$$S^*(A, B) = S_{1,0}(A, B) = \frac{\lambda(A \cap B)}{\lambda(A)}. \quad (5)$$

It is worth noticing that such similarity measure was considered by Bush and Mosteller [3]. In addition, the similarity measure in (5) can be also viewed as a kind of inclusion degree of A in B (see [8, 13]).

Some hypothesis tests for checking the inclusion degree of the expectation of an interval in a previously pre-fixed interval have been developed in [11]. Specifically, the two sided-test for the expected value of a random interval is a particular case of the test in [11] when the inclusion degree equals 1. Thus, the corresponding one-side tests can be easily addressed taking into account an analogous procedure as we will show below.

Now, going back to the one-sided test for the mean based on the interval data and taking into account that our problem can be seen as a testing problem whether the similarity between $\mathbb{E}(X)$ and $\bigcup_{B \subset L} B$ is high enough, we can utilize the similarity measure in (5) to formulate our testing problem as follows.

Let (Ω, \mathcal{A}, P) be a probability space, $X : \Omega \rightarrow K_c(\mathbb{R})$ denote a random interval so that $\text{spr}\mathbb{E}(X) > 0$ and $L = (-\infty, m_1]$ be an interval bounded by an a priori fixed real number m_1 . Then for any given $d \in [0, 1]$ the aim it to test

$$\begin{cases} H'_0 : \max_{B \subset L} S^*(\mathbb{E}(X), B) \geq d; \\ H'_1 : \max_{B \subset L} S^*(\mathbb{E}(X), B) < d. \end{cases} \quad (6)$$

However, one may easily notice that to compute $\max_{B \subset L} S^*(\mathbb{E}(X), B)$ it is enough to examine the relation between $\mathbb{E}(X) = [\text{mid}\mathbb{E}(X) \pm \text{spr}\mathbb{E}(X)]$ and m_1 . Actually, if $\text{mid}\mathbb{E}(X) + \text{spr}\mathbb{E}(X) \leq m_1$ then for any interval $B \subset L$ such that $\mathbb{E}(X) \subseteq B$ we get $S^*(\mathbb{E}(X), B) = 1$ which means that $\mathbb{E}(X)$ is surely similar to some $B \subset L$. On the other hand, if $m_1 < \text{mid}\mathbb{E}(X) - \text{spr}\mathbb{E}(X)$ then there is no interval $B \subset L$ such that $\mathbb{E}(X) \cap B \neq \emptyset$ so we get $\max_{B \subset L} S^*(\mathbb{E}(X), B) = 0$ and

one may conclude that $\mathbb{E}(X)$ surely is not similar to any $B \subset L$. Finally, when $\text{mid}\mathbb{E}(X) - \text{spr}\mathbb{E}(X) < m_1 \leq \text{mid}\mathbb{E}(X) + \text{spr}\mathbb{E}(X)$ we obtain $0 < \max_{B \subset L} S^*(\mathbb{E}(X), B) < 1$ which indicates a partial similarity.

Moreover, let us notice that $\lambda(\mathbb{E}(X)) = 2\text{spr}\mathbb{E}(X)$ and the Lebesgue measure of $A \cap B$ can be expressed (cf. [14]) as follows:

$$\begin{aligned} \lambda(\mathbb{E}(X) \cap B) &= \\ &= \max \left\{ 0, \min \{ 2\text{spr}\mathbb{E}(X), 2\text{spr}B, \right. \\ &\quad \left. \text{spr}\mathbb{E}(X) + \text{spr}B - |\text{mid}\mathbb{E}(X) - \text{mid}B| \} \right\} \end{aligned} \quad (7)$$

Hence, to maximize $S^*(\mathbb{E}(X), B)$ we have to maximize $\lambda(\mathbb{E}(X) \cap B)$ for any $B \subset L$. By the considerations given above it can be shown that

$$\begin{aligned} \max_{B \subset L} \lambda(\mathbb{E}(X), B) &= \\ &= \max \left\{ 0, \min \{ 2\text{spr}\mathbb{E}(X), \right. \\ &\quad \left. m_1 - \text{mid}\mathbb{E}(X) + \text{spr}\mathbb{E}(X) \} \right\}. \end{aligned} \quad (8)$$

Hence, by (9) our hypothesis testing problem can be equivalently expressed as

$$H'_0 : \text{mid}\mathbb{E}(X) + (2d - 1)\text{spr}\mathbb{E}(X) \leq m_1 \quad (9)$$

vs.

$$H'_1 : \text{mid}\mathbb{E}(X) + (2d - 1)\text{spr}\mathbb{E}(X) > m_1. \quad (10)$$

It is clear that to generalize the hypothesis $H'_0 : m_V \in [m_2, +\infty)$ for the interval data we will check whether the interval expected value $\mathbb{E}(X)$ is similar to any interval $B \subset U = [m_2, +\infty)$. Therefore, following steps analogous to the left-sided null hypothesis testing problem, our right-sided testing problem can be described as follows: let (Ω, \mathcal{A}, P) be a probability space, $X : \Omega \rightarrow K_c(\mathbb{R})$ denote a random interval so that $\text{spr}\mathbb{E}(X) > 0$ and $U = [m_2, +\infty)$ be an interval bounded by an a priori fixed real number m_2 . Then for any given $d \in [0, 1]$ the aim it to test

$$H''_0 : \max_{B \subset U} S^*(\mathbb{E}(X), B) \leq d \quad (11)$$

vs.

$$H''_1 : \max_{B \subset U} S^*(\mathbb{E}(X), B) > d, \quad (12)$$

which can be equivalently expressed as

$$H''_0 : \text{mid}\mathbb{E}(X) - (2d - 1)\text{spr}\mathbb{E}(X) \geq m_2 \quad (13)$$

vs.

$$H''_1 : \text{mid}\mathbb{E}(X) - (2d - 1)\text{spr}\mathbb{E}(X) < m_2. \quad (14)$$

4. One-sided testing based on interval data

4.1. Theoretical results

The testing procedure to be developed in this section concerns the test with hypotheses (9) and (10).

It should be remarked that test with hypotheses (13) and (14) can be addressed analogously.

Let $\{X_i\}_{i=1}^n$ be a set of random intervals independent and identically distributed as X . Then, to solve the test composed by the hypotheses given in (9) and (10) the following test statistic is defined:

$$T_n = \sqrt{n} \left(\text{mid} \overline{X_n} + (2d-1) \text{spr} \overline{X_n} - m_1 \right), \quad (15)$$

where $\text{mid} \overline{X_n}$ and $\text{spr} \overline{X_n}$ are the corresponding classical sample means $\overline{\text{mid} X_n}$ and $\overline{\text{spr} X_n}$, respectively.

In order to avoid trivial situations it is necessary to guarantee that the variances of $\text{mid} X$ and $\text{spr} X$ are finite, that the variance of $\text{spr} X$ is positive and that $\sigma_{\text{mid} X, \text{spr} X}^2 \neq \sigma_{\text{mid} X}^2 \sigma_{\text{spr} X}^2$, where $\sigma_{\text{mid} X, \text{spr} X}^2$ stands for the covariance of $\text{mid} X$ and $\text{spr} X$.

Let $Z = (z_1, z_2)^T \equiv \mathcal{N}_2(\vec{0}, \Sigma)$ be a bivariate normal distribution being Σ the covariance matrix for the random vector $(\text{mid} X, \text{spr} X)$. The asymptotic distribution of T_n is provided as follows.

Lemma 1 *If $\{X_i\}_{i=1}^n$ are n random intervals independent and equally distributed as X defined on the probability space (Ω, \mathcal{A}, P) , and satisfying the conditions indicated above, then:*

- a) *Whenever $d = 1$ and $\text{mid} \mathbb{E}(X) + \text{spr} \mathbb{E}(X) = \sup \mathbb{E}(X) = m_1$, it is fulfilled that*

$$T_n \xrightarrow{\mathcal{L}} z_1 + z_2. \quad (16)$$

- b) *Whenever $d \in (0, 1)$ and $\text{mid} \mathbb{E}(X) + (2d-1) \text{spr} \mathbb{E}(X) = m_1$, it is fulfilled that*

$$T_n \xrightarrow{\mathcal{L}} z_1 + (2d-1)z_2. \quad (17)$$

- c) *Whenever $d = 0$ and $\text{mid} \mathbb{E}(X) - \text{spr} \mathbb{E}(X) = \inf \mathbb{E}(X) = m_1$, it is fulfilled that*

$$T_n \xrightarrow{\mathcal{L}} z_1 - z_2. \quad (18)$$

Proof. The statistic T_n , firstly defined in (15), can be equivalently written as

$$T_n = \sqrt{n} \left(\begin{aligned} &\text{mid} \overline{X_n} + (2d-1) \text{spr} \overline{X_n} \\ &- \text{mid} \mathbb{E}(X) - (2d-1) \text{spr} \overline{X_n} \\ &+ (\text{mid} \mathbb{E}(X) + (2d-1) \text{spr} \overline{X_n} - m_1) \end{aligned} \right).$$

- a) If $d = 1$ and $\sup \mathbb{E}(X) = m_1$, then $\text{mid} \mathbb{E}(X) + \text{spr} \mathbb{E}(X) - m_1 = 0$ and the other term of T_n converges in law to $z_1 + z_2$ as $n \rightarrow \infty$ by the CLT for real variables.
- b) If $d \in (0, 1)$ and $\text{mid} \mathbb{E}(X) + (2d-1) \text{spr} \mathbb{E}(X) = \sup \mathbb{E}(X) = m_1$, then $\text{mid} \mathbb{E}(X) + (2d-1) \text{spr} \mathbb{E}(X) + \sup \mathbb{E}(X) - m_1 = 0$ and the other term of T_n converges in law to $z_1 + (2d-1)z_2$ as $n \rightarrow \infty$ by the CLT.
- c) If $d = 1$ and $\inf \mathbb{E}(X) = m_1$, then $\text{mid} \mathbb{E}(X) - \text{spr} \mathbb{E}(X) - m_1 = 0$ and the other term of T_n converges in law to $z_1 - z_2$ as $n \rightarrow \infty$ by the CLT. \square

It is straightforward that other situations under H_0 different than the ones specified in Lemma 1 implies that $\text{mid} \mathbb{E}(X) + (2d-1) \text{spr} \mathbb{E}(X) < m_1$ and the divergence of the statistic T_n to $-\infty$.

The consistency of the test is established in Theorem 1.

Theorem 1 *Let $\alpha \in [0, 1]$ and $k_{1-\alpha}$ be the $(1-\alpha)$ -quantile of the asymptotic distribution of T'_n . If H_0 in (9) is true, then it is satisfied that*

$$\limsup_{n \rightarrow \infty} P(T'_n > k_{1-\alpha}) \leq \alpha \quad (19)$$

and the equality is achieved whenever conditions in a), b) and c) in Lemma 1 are fulfilled. In addition, if H_0 is not fulfilled then

$$\lim_{n \rightarrow \infty} P(T'_n > k_{1-\alpha}) = 1.$$

Therefore, the test which rejects H_0 in (9) at the significance level α whenever $T'_n > k_{1-\alpha}$ is asymptotically correct and consistent.

Proof. Denote $\text{mid} \mathbb{E}(X) + (2d-1) \text{spr} \mathbb{E}(X) - m_1$ by $A(d)$. Given $k \in \mathbb{R}$ we will prove that

$$\begin{aligned} P(T_n > k) &= P(T_n > k \mid A(d) \geq 0) \leq \\ &P(T_n > k \mid A(d) = 0). \end{aligned} \quad (20)$$

We define $T_n^X = \sqrt{n} \hat{A}(d)$, where $\hat{A}(d) = \text{mid} \overline{X_n} + (2d-1) \text{spr} \overline{X_n} - m_1$. Suppose that X satisfies that $A(d) < 0$. Given the random interval $Y = X + B$ where

$$B = \begin{bmatrix} -m_1 + \text{mid} \mathbb{E}(X) + (2d-1) \text{spr} \mathbb{E}(X), \\ m_1 - \text{mid} \mathbb{E}(X) - (2d-1) \text{spr} \mathbb{E}(X) \end{bmatrix},$$

we have that $\text{mid} \overline{X_n} + (2d-1) \text{spr} \overline{X_n} - m_1 = 0$.

Clearly, $T_n^Y > T_n^X$ for all $\omega \in \Omega$, which implies that $P(T_n^Y > k) \geq P(T_n^X > k)$ and the inequality (20) is fulfilled.

Lemma 1 implies that $P(T_n > k_{1-\alpha} \mid A(d)) \xrightarrow{n \rightarrow \infty} \alpha$, so taking into account the inequality (20), the desired result (19) is concluded.

Finally, if H_0 is not fulfilled then it is immediate to deduce that T_n diverges to ∞ as n tends to ∞ . Thus, $\lim_{n \rightarrow \infty} P(T_n > k_{1-\alpha}) = 1$ and the test is also consistent. \square

As it is not immediate to handle the asymptotic distribution of T_n in practice, bootstrap techniques are applied since they have been shown to be a very useful tool to get a better approximation to the sampling distribution [6, 7]. Specifically, we propose the use of a residual bootstrap approach.

First of all, let X be a random interval such that $\text{spr} \mathbb{E}(X) > 0$ and let $\{X_i\}_{i=1}^n$ be a simple random sample from X . If $\{X_i^*\}_{i=1}^n$ is a bootstrap sample drawn from $\{X_i\}_{i=1}^n$, the corresponding bootstrap statistic is defined as

$$T_n^{boot} = \sqrt{n} \left(\begin{aligned} &\text{mid} \overline{X_n^*} + (2d-1) \text{spr} \overline{X_n^*} \\ &- \text{mid} \overline{X_n} - (2d-1) \text{spr} \overline{X_n} \end{aligned} \right), \quad (21)$$

Lemma 2 shows the asymptotic distribution of T_n^{boot} .

Lemma 2 Suppose that X is an RI satisfying that the variances of $\text{mid } X$ and $\text{spr } X$ are finite, the variance of $\text{spr } X$ is positive and $\sigma_{\text{mid } X, \text{spr } X}^2 \neq \sigma_{\text{mid } X}^2 \sigma_{\text{spr } X}^2$.

a) Whenever $d = 1$ and $\text{mid } \mathbb{E}(X) + \text{spr } \mathbb{E}(X) = \text{sup } \mathbb{E}(X) = m_1$, it is fulfilled that

$$T_n^{boot} \xrightarrow{\mathcal{L}} z_1 + z_2 \text{ a.s.} - [P]. \quad (22)$$

b) Whenever $d \in (0, 1)$ and $\text{mid } \mathbb{E}(X) + (2d - 1)\text{spr } \mathbb{E}(X) = m_1$, it is fulfilled that

$$T_n^{boot} \xrightarrow{\mathcal{L}} z_1 + (2d - 1)z_2 \text{ a.s.} - [P]. \quad (23)$$

c) Whenever $d = 0$ and $\text{mid } \mathbb{E}(X) - \text{spr } \mathbb{E}(X) = \text{inf } \mathbb{E}(X) = m_1$, it is fulfilled that

$$T_n^{boot} \xrightarrow{\mathcal{L}} z_1 - z_2 \text{ a.s.} - [P]. \quad (24)$$

Besides, theorem below shows the correctness of the bootstrap test.

Theorem 2 Let $\alpha \in [0, 1]$ and let $k_{1-\alpha}^*$ be the $(1 - \alpha)$ -quantile of the distribution of T_n^{boot} . If H_0 in (9) is fulfilled and T_n is defined as in 15, then

$$\limsup_{n \rightarrow \infty} P(T_n > k_{1-\alpha}^*) \leq \alpha \text{ a.s.} - [P],$$

and the equality is achieved whenever conditions in a), b) and c) in Lemma 1 are fulfilled. As a consequence, the test rejecting H_0 in (9) when $T_n > k_{1-\alpha}^*$ is asymptotically correct.

If H_0 is fulfilled but conditions in a), b) and c) are not fulfilled, T_n^{boot} converges to certain limit distributions different from the ones of Lemma 1 and the bootstrap procedure can be proved to be consistent, that is, that the power tends to 1 under H_1 in (10).

As it is done in the classical framework, the Monte Carlo method is employed to approximate the distribution of T_n^{boot} .

4.2. Algorithm for the bootstrap test

Step 1. For $\{X_i\}_{i=1}^n$ compute the value of the statistic T_n defined in (15).

Step 2. Obtain a bootstrap sample $\{X_i^*\}_{i=1}^n$ from $\{X_i\}_{i=1}^n$ and compute the value of the bootstrap statistic T_n^{boot} defined in (21).

Step 3. Repeat Step 2 a large number B of times to get a set E of B values of the bootstrap statistics.

Step 4. Compute the bootstrap p -value as the proportion of values in E which are greater than or equal to T_n .

5. Simulation studies

In order to check the behavior of the theoretical approach proposed in Section 4, some models are analyzed by means of simulation studies. Suppose that we have a real random variable that cannot be perceived but from which we know that its true mean m_X belongs to the interval $(-\infty, m_1]$. Thus, two models are simulated so that a sample of mid-points is generated having the same mean as m_X as well as the corresponding spreads that can be independent or dependent of the mid-points (first case refers to Model 1 and the second one refers to Model 2). Finally, we obtain a sample of intervals that can be considered as disturbances of the true but not observable real data. These two models are defined as follows:

- **Model 1:** $\text{mid } X \equiv \mathcal{N}(2, 5)$ and $\text{spr } X \equiv U(1, 5)$.
- **Model 2:** $\text{mid } X \equiv \mathcal{N}(2, 5)$ and $\text{spr } X \equiv \text{mid } X + U(0, 2)$.

The aim is to solve the following test for the two models above:

$$\begin{cases} H_0 : \text{mid } \mathbb{E}(X) + (2d - 1)\text{spr } \mathbb{E}(X) \leq m_1; \\ H_1 : \text{mid } \mathbb{E}(X) + (2d - 1)\text{spr } \mathbb{E}(X) > m_1. \end{cases}$$

Three different situations will be analyzed for each model depending on the similarity measure considered, mainly, $d = 1$, $d \in (0, 1)$ and $d = 0$. For this three situations, three different values of m_1 will be taken into account: For case $d = 1$, $m_1 = 5$; for case $d \in (0, 1)$, $m_1 = 1$; and for case $d = 0$, $m_1 = -1$.

The bootstrap approach in Section 4 has been implemented, providing for 10000 simulations with 1000 bootstrap replications at the usual significance levels ρ (.01, .05 and .1) for different sample sizes. Results are gathered in Tables 1, 2, 3, 4, 5 and 6.

		Model 1, $d = 1$		
$n \setminus 100 \cdot \rho$		1	5	10
10		3.02	7.52	13.02
30		1.78	6.08	11.54
50		1.18	5.28	10.56
100		1.11	5.21	10.34
200		1.04	5.06	10.08

Table 1: Empirical size of the bootstrap test, Model 1, $d = 1$

Tables 1, 2, 3, 4, 5 and 6 display that the empirical sample sizes are in all the cases quite close to the nominal significance levels for sample sizes greater than or equal to $n = 50$. In addition, there are no considerable differences between the independent (Model 1) and dependent (Model 2) situations. However, it would be interesting to address a deeper sensitive analysis as future work in order to check if there are remarkable influences, for instance, on

		Model 1, $d = 1/3$		
$n \setminus 100 \cdot \rho$		1	5	10
10		3.08	7.38	11.86
30		1.42	6.18	11.04
50		1.24	6.02	10.26
100		1.12	4.88	10.18
200		1	4.90	10.06

Table 2: Empirical size of the bootstrap test, Model 1, $d = 1/3$

		Model 1, $d = 0$		
$n \setminus 100 \cdot \rho$		1	5	10
10		3.26	6.90	13.08
30		1.66	5.94	10.84
50		1.42	5.58	10.18
100		1.28	5.16	10.14
200		1.08	5.06	10.08

Table 3: Empirical size of the bootstrap test, Model 1, $d = 0$

		Model 2, $d = 1$		
$n \setminus 100 \cdot \rho$		1	5	10
10		2.84	7.46	12.28
30		1.80	5.90	11.16
50		1.28	5.40	10.36
100		1.12	5.22	10.20
200		1.12	4.92	10.08

Table 4: Empirical size of the bootstrap test, Model 2, $d = 1$

		Model 2, $d = 1/3$		
$n \setminus 100 \cdot \rho$		1	5	10
10		2.88	8.18	13.06
30		1.70	6.20	11.04
50		1.32	5.58	10.62
100		1.25	5.10	10.16
200		1.14	4.96	9.84

Table 5: Empirical size of the bootstrap test, Model 2, $d = 1/3$

		Model 2, $d = 0$		
$n \setminus 100 \cdot \rho$		1	5	10
10		4.56	8.08	11.96
30		1.70	5.90	10.86
50		1.54	5.30	10.58
100		1.28	4.86	10.30
200		1	5.02	9.86

Table 6: Empirical size of the bootstrap test, Model 2, $d = 0$

the choice of the distributions involved in the simulations.

6. Conclusions

The classical one-sided hypothesis testing for the mean of real random variables which values cannot be observed as real values but intervals in which the true value is included in has been proposed in this contribution. Thus, the problem has been translated as a similarity test between the expectation of a random interval and a specific interval. The limit distribution of a suggested test statistic has been analyzed by means of both asymptotic and bootstrap approaches. Later, some simulations have been achieved to show that the bootstrap approach is suitable for moderate sample sizes.

In the future, a deep study of the influence of the variables chosen for the simulations should be carried out as well as a deep study regarding the power of the proposed test. In addition, it would be interesting to consider other similarity measures proposed in the literature and to extend the results provided in this work to the fuzzy framework. Some applications of the methodology suggested in this paper will be also tackled as well as the extension of the problem to the fuzzy sets framework.

Acknowledgments. We would like to thank the grant “Estadística robusta y flexible para datos intervalares, de conjunto y de conjunto difuso: localización, variabilidad y regresión lineal” (MTM2013-44212-P, Spanish Ministry of Education and Science) for the financial support.

References

- [1] R.J. Aumann, Integrals of set-valued functions, *Journal of Mathematical Analysis and Applications*, 12: 1–12, 1965.
- [2] A. Blanco-Fernández, N. Corral and G. González-Rodríguez, Estimation of a flexible simple linear model for interval data based on set arithmetic, *Computational Statistics and Data Analysis*, 55(9): 2568–2578, 2011.
- [3] R.R. Bush and F.A. Mosteller, A model for stimulus generalization and discrimination, *Psychological Review*, 58: 413–423, 1951.
- [4] I. Couso, L. Garrido and L. Sánchez, Similarity and dissimilarity measures between fuzzy sets. A formal relational study, *Information Sciences* 229: 122–141, 2013.
- [5] P. D’Urso and L. De Giovanni, Midpoint radius self-organizing maps for interval-valued data with telecommunications application, *Applied Soft Computing*, 11(5): 3877–3886, 2011.
- [6] M.B. Ferraro, R. Copi, G. González-Rodríguez and A. Colubi, A linear regression model for imprecise response, *International Journal of Approximating Reasoning*, 51(7): 759–770, 2010.
- [7] M.A. Gil, M. Montenegro, G. González-Rodríguez, A. Colubi and M.R. Casals, Bootstrap approach to the multi-sample test of

- means with imprecise data, *Computational Statistics and Data Analysis*, 51:148–162, 2006.
- [8] B. Kosko, Fuzzy entropy and conditioning, *Information Sciences*, 40: 165–174, 1986.
- [9] R. Kruse, K.D. Meyer, *Statistics with Vague Data*, Riedel, Dordrecht, 1987.
- [10] G. Matheron. *Random Sets and Integral Geometry*, Wiley, New York, 1975.
- [11] A.B. Ramos-Guajardo, A. Colubi, G. González-Rodríguez, Inclusion degree tests for the Aumann expectation of a random interval, *Information Sciences*, 288:412–422, 2014.
- [12] A.B. Ramos-Guajardo, Similarity Test for the expectation of a random interval and a fixed interval, In: P. Grzegorzewski et al. (eds.), *Strengthening Links between Data Analysis & Soft Computing*, Advances in Intelligent Systems and Computing 315, Springer, 2014, pp. 175-182.
- [13] E. Sánchez, Inverses of fuzzy relations, application to possibility distributions and medical diagnosis, *Fuzzy Sets and Systems*, 2:75–86, 1979.
- [14] J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*, Cambridge University Press, Cambridge, 2004.
- [15] A. Tversky, Features of similarity, *Psychological Review*, 84: 327–352, 1977.
- [16] L.A. Zadeh, Fuzzy sets, *Information Control*, 8: 338–353, 1965.