International Journal of Approximate Reasoning 152 (2023) 297–309

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Logical coherence in Bayesian simultaneous three-way hypothesis tests

Luís G. Esteves [a], Rafael Izbicki [b], Julio M. Stern [a], Rafael B. Stern [a,*]

[a] *Universidade de São Paulo, Instituto de Matemática e Estatística, São Paulo, São Paulo, Brazil*
[b] *Universidade Federal de São Carlos, Departamento de Estatística, São Carlos, São Paulo, Brazil*

ARTICLE INFO

Article history:
Received 4 April 2022
Received in revised form 24 October 2022
Accepted 28 October 2022
Available online 8 November 2022

Keywords:
Three-way decisions
Agnostic hypothesis tests
Bayesian hypothesis tests

ABSTRACT

Hypothesis testing is a key component in Statistics. In practice, it is common for a practitioner to test several hypotheses, the so-called simultaneous hypothesis testing problem. Unfortunately, simultaneous hypothesis tests can be logically incoherent: for instance, if hypothesis H_1 implies hypothesis H_2, a procedure that rejects H_2 should also reject H_1, a property not always met by multiple test procedures. Indeed, previous results show that standard two-way hypothesis tests cannot be logically coherent. Three-way tests allow more nuanced data-based decisions. This paper studies whether Bayesian simultaneous three-way hypothesis tests can be logically coherent. Two types of results are obtained. First, under the standard error-wise constant loss, only for a limited set of models a Bayes simultaneous test can be logically coherent. Second, if more general loss functions are used, then it is possible to obtain Bayes simultaneous tests that are always logically coherent. An explicit example of such a loss function is provided. These results provide guidelines on how to build a logically coherent posterior probability three-way hypothesis test or more general Bayesian three-way hypothesis tests.

1. Introduction

In a three-way decision problem [1–6], one must classify objects into three categories. While a two-way decision necessarily leads to an affirmation or a negation, a three-way decision also allows non-commitment or pause to gather more evidence. Such a flexible approach has led to advances in areas such as clustering [7], classification [8–13], cognitive analytics [14], conflict analysis [15–17], crowdsourcing [18], game theory [19–21], granular computing [22], investment decisions [23], medicine [24], multi-agent decisions [25–27], pattern discovery [28], and recommender systems [29,30].

Among the frameworks for three-way decisions [1,22], the first approach is based on Pawlak's rough set model [31], a special case of the TAO model [4]. In this framework, one wishes to determine a concept related to $X \subseteq \mathscr{X}$ based on $\sim$, an equivalence relation over $\mathscr{X}$. Letting $[x] = \{y \in \mathscr{X} : y \sim x\}$, the upper and lower approximations of X are defined, respectively, as $\overline{\text{appr}}(X) = \{x \in \mathscr{X} : [x] \cap X \neq \emptyset\}$ and $\underline{\text{appr}}(X) = \{x \in \mathscr{X} : [x] \cap X \neq \emptyset\}$. These definitions divide $\mathscr{X}$ into three regions: $POS(X) := \overline{\text{appr}}(X)$ are the elements that certainly belong to the concept, $BND(X) := \overline{\text{appr}}(X) - \underline{\text{appr}}(X)$ are the elements that might belong to the concept, and $NEG(X) := \mathscr{X} - \overline{\text{appr}}(X)$ are the elements that do not belong to the concept.

* Corresponding author.
E-mail address: rbstern@gmail.com (R.B. Stern).

https://doi.org/10.1016/j.ijar.2022.10.019
0888-613X/

One generalization of Pawlak's rough set framework is given by decision-theoretic rough set theory [32]. In the latter, $POS(X) = \{x \in \mathcal{X} : \mathbb{P}(X|[x] > \alpha\}$, $BND(X) = \{x \in \mathcal{X} : \beta \leq \mathbb{P}(X|[x] \leq \alpha\}$, and $NEG(X) = \{x \in \mathcal{X} : \mathbb{P}(X|[x] < \beta\}$. The use of probability in these definitions makes them susceptible to application in Statistics.

A particular application occurs in statistical hypothesis testing [33–40]. In this context, one gathers data, $x \in \mathcal{X}$ and wishes to decide whether x corroborates a given statistical hypothesis, H. More formally, data is used to learn about an unobserved quantity, $\theta \in \Theta$ and a statistical hypothesis is an assertion of the type $\theta \in H$. The goal of hypothesis testing is to decide whether one believes that θ belongs to H in the light of data x. This problem can use the rough set framework by taking each dataset x as an object and H as a concept. While standard hypothesis tests allow only the rejection or non-rejection of H, three-way (agnostic) tests also allow H to remain undecided. That is an agnostic test is such that, for each data, $x \in \mathcal{X}$, $x \in$ POS(H), if x confirms H, $x \in$ NEG(H), if x negates H, and $x \in$ BND(H), otherwise.[1]

In the statistical literature, such tests are usually represented by a function, $\varphi_H : \mathcal{X} \to \{0, \frac{1}{2}, 1\}$. In this context, $\varphi_H(x) = 0$, $\varphi_H(x) = 1$, and $\varphi_H(x) = \frac{1}{2}$ mean that one decides, respectively, to accept, reject and remain undecided about H after observing data x. This definition can be identified with the standard three-decision regions:

$$POS(H) := \{x \in \mathcal{X} : \varphi_H(x) = 0\}, \tag{1}$$

$$NEG(H) := \{x \in \mathcal{X} : \varphi_H(x) = 1\}, \text{ and} \tag{2}$$

$$BND(H) := \left\{x \in \mathcal{X} : \varphi_H(x) = \frac{1}{2}\right\}. \tag{3}$$

This framework can be extended to the problem of testing several hypotheses at the same time, the so-called simultaneous (or multiple) hypothesis testing problem [43]. Unfortunately, simultaneous hypothesis tests can be logically incoherent. For instance, an incoherent test might be such that $x \in$ NEG$(\theta \geq 0)$ and $x \in$ POS$(\theta = 0)$, even though $\theta = 0$ implies $\theta \geq 0$ [44]. Thus, a large body of work is devoted to understanding whether binary hypothesis tests can be logically coherent while retaining statistical optimality [44–47].

Logical coherence can also be of importance in more general multi-concept rough set models. For instance, one might wish to classify images as cats and as animals. Since every cat is an animal, one might wish the logical requirement that POS(cat) $\subseteq$ POS(animal). Otherwise, if $x \in$ POS(cat) but $x \notin$ POS(animal), one has the challenging task of explaining how it is certain that x represents a cat, but not an animal.

Example 1.1, below, illustrates how logical incoherence can be an important issue in a common three-way decision hypothesis test.

Example 1.1 *(Multiple means comparisons).* Consider the Analysis of Variance (ANOVA) data from Example 3.12 in Izbicki and Esteves [44]. It consists of independent samples from three Gaussian distributions with different means and variances. Let μ_i be the mean of the i-th group, $i = 1, 2, 3$. The goal is to test the following hypotheses: $H_1 : \mu_1 > \mu_2$, $H_2 : \mu_2 > \mu_3$, and $H_3 : \mu_1 > \mu_2 > \mu_3$. Assume that means of each group are independent and that $\mu_i \sim N(0, 10^2)$. Also, for simplicity, assume that σ_i is the known standard deviation of the i-th group, where $\sigma_1 = 1.09$, $\sigma_2 = 0.5$, and $\sigma_3 = 0.79$. The following sample means are observed: $\bar{X}_1 = 0.15$, $\bar{X}_2 = -0.13$, and $\bar{X}_3 = -0.38$.

Under the model above, one can obtain $\mathbb{P}(H_1|x) = 70.1\%$, $\mathbb{P}(H_2|x) = 72.5\%$, and $\mathbb{P}(H_3|x) = 48.3\%$.[2] Consider a posterior probability three-way test [48] that rejects every hypothesis with posterior probability lesser than 50% and accepts every hypothesis with probability higher than 70%. Such a test accepts both H_1 and H_2, but rejects H_3. This test is logically incoherent since $H_3 : \mu_1 > \mu_2 > \mu_3$ is rejected, but believing in $H_1 : \mu_1 > \mu_2$ and in $H_2 : \mu_2 > \mu_3$ entails the logical conclusion that $\mu_1 > \mu_2 > \mu_3$. On the other hand, the 80%-level GFBST three-way test [39] is not incoherent, since it remains undecided about the three hypotheses. This conclusion can be observed in Fig. 1, since the oval Highest Posterior Density (HPD) set intercepts all hypotheses and their complements. △

In Example 1.1, the usual posterior probability three-way hypothesis test is logically incoherent for performing multiple comparisons. Specifically, the test accepts $H_1 : \mu_1 > \mu_2$ and $H_2 : \mu_2 > \mu_3$, but rejects the logical deduction that $H_3 : \mu_1 >$

[1] This goal is similar to the one in Bayesian confirmation theory [41], in which one wishes to determine whether $x \in \mathcal{X}$ confirms or disconfirms an hypothesis, $H \subseteq \Theta$. Absolute confirmation theory states that, for some constant $k \in [0, 1]$, x confirms H if $\mathbb{P}(H|x) > k$, disconfirms H if $\mathbb{P}(H|x) < k$, and is neutral if $\mathbb{P}(H|x) = k$. Similarly, incremental according to incremental confirmation theory, x confirms H if $\mathbb{P}(H|x) > \mathbb{P}(H)$, disconfirms H if $\mathbb{P}(H|x) < \mathbb{P}(H)$, and is neutral if $\mathbb{P}(H|x) = \mathbb{P}(H)$. One might argue that these approaches are too restrictive by stating that x is neutral only if $\mathbb{P}(H|x) = k$ or $\mathbb{P}(H|x) = \mathbb{P}(H)$. From a practical perspective, x might be neutral if it provides solely weak evidence for or against H. The approach in [32] deals with this concern and generalizes absolute confirmation theory by stating that x confirms H, if $\mathbb{P}(H|x) > \alpha$, disconfirms H, if $\mathbb{P}(H|x) < \beta$, and is neutral if $\beta \leq \mathbb{P}(H|x) \leq \alpha$.

Other approaches to confirmation theory, such as [42], suggest that a high value of $\mathbb{P}(H|x)$ is a necessary but not a sufficient condition for stating that x confirms H. That is, there exists α such that, if $\mathbb{P}(H|x) \leq \alpha$, then x does not confirm H. However, $\mathbb{P}(H|x) > \alpha$ does not necessarily imply that x confirms H. Such an approach is motivated by the fact that, using absolute confirmation theory, a value of x might confirm mutually contradictory hypothesis, $H_1, \ldots, H_n$. This issue is showcased in Example 1.1 and further developed in section 4.

[2] These probabilities were computed via Monte Carlo integration; the code can be found on https://github.com/rizbicki/three_way_hypothesis_tests/blob/main/example_anova.R.

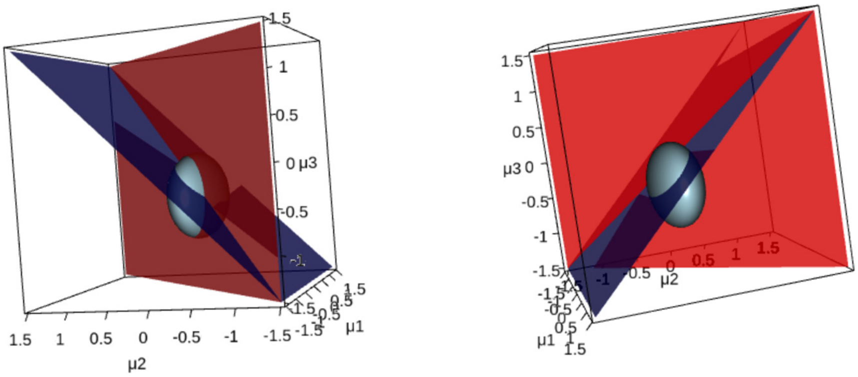

Fig. 1. Two perspectives for the HPD region (gray ellipse) and the boundaries of hypotheses H_1 (red) and H_2 (blue) for the data of Example 1.1. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

$\mu_2 > \mu_3$. Such a set of conclusions is hard to explain to most researchers. Hence, one might wish to use a logically coherent test. However, the existing literature provides no guidelines on how to build a logically coherent posterior probability three-way hypothesis test or more general Bayesian three-way hypothesis tests.

The goal of this paper is to study whether decision-theoretic [49] simultaneous three-way hypothesis tests can be logically coherent. Section 2 reviews key concepts in decision-theoretic three-way hypothesis testing (Section 2.1) and logical coherence (Section 2.2). Under this setting, Section 3 explores the relation between the EC loss, posterior probability three-way tests, and logical coherence. This section shows that it is impossible to fully reconcile the traditional posterior probability three-way tests with logical coherence. This argument can be extended to procedures based on more general probability thresholds. Section 4 explores decision-theoretic three-way tests under more general loss functions than the EC loss. This section defines the GFBST loss and shows that, under this loss, the decision-theoretic three-way test is always logically coherent.

2. Background review

This section reviews the main concepts regarding probability-based hypothesis tests and logical coherence that are used in this paper. In particular, the following subsection reviews key concepts that are required for defining simultaneous posterior probability three-way hypothesis tests.

2.1. Review of hypothesis tests and loss functions

In order to determine optimal decision regions, one can use Bayesian decision theory. This flexible theory can be used to justify probabilistic and decision-theoretic rough sets [50–52] and also other types of Bayesian procedures [49,53–55], including hypothesis tests. In principle, these other procedures can be useful for investigating other types of three-way decisions, an approach discussed in Section 4.

In the context of Bayesian decision theory, we define a loss function,

$$L_H : \Theta \times \left\{0, \frac{1}{2}, 1\right\} \longrightarrow \mathbb{R}.$$

L_H measures how bad each decision in $\left\{0, \frac{1}{2}, 1\right\}$ is for each parameter value, $\theta \in \Theta$. Next, we define a prior distribution over Θ, which denotes one's prior uncertainty about θ. We denote this distribution by $f(\theta)$. The prior distribution, together with the data distribution, $f(x|\theta)$, yields the posterior distribution, $f(\theta|x)$. This quantity encodes one's uncertainty about the θ after observing x. Finally, the Bayes decision consists of picking $d \in \{0, \frac{1}{2}, 1\}$ that minimizes the expected posterior loss, $\mathbf{E}[L_H(\theta, d)|\mathbf{x}] = \int L_H(\theta, d) f(\theta|x) d\theta$. We call such decision the *optimal Bayes decision*, or simply the *Bayes decision*. In three-way tests, the Bayes decision is the decision-theoretic three-way test.

A common choice of L_H is the error-type constant (EC) loss function:

Definition 2.1 *(Error-type constant loss function).* Let H be a hypothesis. The error-type constant (EC) loss function, L_H, is given by Table 1, where $0 < \lambda^H_{BP} < \lambda^H_{NP}$, $0 < \lambda^H_{BN} < \lambda^H_{PN}$, and $(\lambda^H_{PN} - \lambda^H_{BN})\lambda^H_{NP} > \lambda^H_{BP}\lambda^H_{PN}$. These restrictions are made so that the loss for each type of error corresponds to its intuitive meaning. For instance, when H is true, accepting H is better than not deciding at all, which in turn is better than rejecting H. Also, the last inequality warrants that not deciding is an admissible decision.

The loss in Table 1 is a special case of Yao [48]. Since loss functions are invariant to translations [49], there is no loss in generality in setting the loss to 0 when the hypothesis is correct and is accepted. Similar to other references in

Table 1
Error-type constant loss function.

	$\theta \in H$	$\theta \notin H$
$x \in POS(H)$	0	λ^H_{PN}
$x \in BND(H)$	λ^H_{BP}	λ^H_{BN}
$x \in NEG(H)$	λ^H_{NP}	0

hypothesis tests [54], we also set the loss of accepting the hypothesis when it is correct as equal to the loss of rejecting the hypothesis when it is incorrect. There is no loss of generality in the optimal decision rules that can be obtained, as shown in Example 2.2 below.

The decision-theoretic three-way test with respect to the error-type constant loss function is shown in the following example.

Example 2.2 *(Posterior probability three-way tests).* Under the EC loss (Definition 2.1), the optimal three-way decision regions for hypothesis tests can be obtained from the results in Yao [48] by setting $\lambda^H_{PP} = \lambda^H_{NN} = 0$ and using the restriction $(\lambda^H_{PN} - \lambda^H_{BN})\lambda^H_{NP} > \lambda^H_{BP}\lambda^H_{PN}$:

$$POS(H) = \{x \in \mathcal{X} : \mathbb{P}(\theta \in H|x) > \beta^H\},$$
$$NEG(H) = \{x \in \mathcal{X} : \mathbb{P}(\theta \in H|x) < \alpha^H\}, \text{ and}$$
$$BND(H) = \{x \in \mathcal{X} : \alpha^H \leq \mathbb{P}(\theta \in H|x) \leq \beta^H\},$$

where

$$\beta^H = \frac{\lambda^H_{PN} - \lambda^H_{BN}}{(\lambda^H_{PN} - \lambda^H_{BN}) + \lambda^H_{BP}} < 1, \text{ and } \alpha^H = \frac{\lambda^H_{BN}}{(\lambda^H_{BN} - \lambda^H_{BP}) + \lambda^H_{NP}} > 0 \quad \triangle \tag{4}$$

Alternative approaches for determining decision thresholds involve, for instance, Bayesian rough sets [56], Bayesian confirmation measures [57], the optimization approach [58], granular shadowed sets [59], information-theory [60], game-theory [61], or linguistic intuitionistic fuzzy information [62].

In simultaneous hypothesis testing, one wishes to test a collection of hypotheses, $\sigma(\Theta)$, at the same time [63,64]. Definition 2.3 describes Bayesian optimality in this context.

Definition 2.3 *(Bayesian optimality for simultaneous hypothesis tests).* For each hypothesis $H \in \sigma(\Theta)$, let L_H be a loss function. A simultaneous hypothesis test, φ, is Bayes-optimal if, for each hypothesis, H, φ_H is a Bayes test for testing H against L_H. In words, a simultaneous hypothesis test is Bayes-optimal if each of its hypotheses is being optimally tested.

The following example shows that simultaneous tests based on posterior probabilities are obtained from the EC loss similarly to Example 2.2:

Example 2.4 *(Simultaneous test based for error-type constant (EC) losses).* Let L be a loss function such that, for each hypothesis, H, L_H is the loss function presented in Table 1. Then, the Bayes simultaneous test for L is such that, for each H, it satisfies Equation (4).

Definition 2.5 *(Simultaneous test based on trivial error-type constant (TEC) losses).* If for each $H \in \sigma(\Theta)$, L_H is such that the constants in Table 1 do not depend on H, then L is said to be a trivial error-type constant loss (TEC). In this case, the Bayes simultaneous test given by Equation (4) is such that $\alpha^H = \alpha$ and $\beta^H = \beta$ do not depend on H, $\alpha, \beta \in (0, 1)$.

Despite satisfying Bayesian optimality, simultaneous posterior probability three-way hypothesis tests might not be logically coherent. The following subsection defines logical coherence and provides examples of tests meeting each requirement.

2.2. Review of logical coherence

In the context of simultaneous tests, one is often interested in an overall interpretation of all the tests. One condition that is required for the interpretability of the tests is their logical coherence. For instance, if $x \in POS(\theta > 1)$[3] and also

[3] For simplicity, we write $\theta > 1$ as a shorthand for $\{\theta \in \Theta : \theta > 1\}$.

$x \in NEG(\theta > 0)$, then, after observing x, one would believe both that "$\theta > 1$" is true and that "$\theta > 0$" is false, a logical contradiction. Such contradictory conclusions are hard to interpret and should be avoided.

Based on this challenge and previous proposals for logical requirements [44–47,65–67], the concept of logical coherence in simultaneous hypothesis testing is proposed [39]:

Definition 2.6 *(Logical coherence).* A simultaneous hypothesis test is logically coherent if it satisfies the following:

1. (Propriety) $POS(\Theta) = \mathcal{X}$.
2. (Monotonicity) If $H_1 \subseteq H_2$, then $x \in POS(H_1)$ implies that $x \in POS(H_2)$ and $x \in NEG(H_2)$ implies that $x \in NEG(H_1)$.[4]
3. (Intersection consonance) If $x \in POS(H_1)$ and $x \in POS(H_2)$, then it also holds that $x \in POS(H_1 \cap H_2)$.
4. (Invertibility) $x \in POS(H)$ if, and only if, $x \in NEG(H^c)$.

These requirements can be interpreted as follows. Propriety states that the hypothesis $H : \theta \in \Theta$ must be accepted for every possible data $x \in \mathcal{X}$. That is natural since one typically designs the parameter space Θ in a way such that it contains all possible parameter values. Thus, $\theta \in \Theta$ holds by design. Monotonicity states that if $H_1 \subseteq H_2$ are nested hypotheses, then the conclusion obtained for H_2 needs to be always at least as favorable as the conclusion obtained for H_1. In particular, if one remains undecided about H_1 after observing x, then one either remains undecided about H_2 or accepts H_2. Intersection consonance implies that if two hypotheses are accepted, then so should be their intersection – this requirement is not met by the posterior probability three-way hypothesis test of Example 1.1. Finally, invertibility states that, no matter whether H or its complement H^c is being tested, the conclusions obtained should be the same.

Logically coherent tests can be characterized in terms of region estimators [39], $R(x)$, which are subsets of the parameter space. These sets are typically $R(x)$ build such that they contain likely values for θ (in light of the observed data x) - this is the case of the ellipsoids centered close to the observed means in Fig. 1, for instance. Region estimators are formalized below.

Definition 2.7 *(Region estimator).* A region estimator is a function $R : \mathcal{X} \to \mathcal{P}(\Theta)$, where $\mathcal{P}(\Theta)$ is the collection of all subsets of Θ.

A particular type of region estimator is the highest posterior density (HPD) set. The HPD contains the parameter values with posterior density above a given threshold. If Θ is finite, then the posterior density of each element of Θ is often taken as its posterior probability. Thus, in this case, the HPD contains the most probable values for θ.

Definition 2.8 *(Highest posterior density set).* A region estimator, $R(x)$, is a highest posterior density set with respect to a posterior density, $f(\theta|x)$, if there exists k such that

$$R(x) = \{\theta \in \Theta : f(\theta|x) \geq k\}. \tag{5}$$

The cutoff k in Definition 2.8 is chosen by the practitioner. Two common ways of choosing it are (i) via loss functions and (ii) by setting k such that the posterior probability of $R(x)$ achieves a prespecified value (e.g. 95%) [54,68].

Below, Example 2.9 illustrates how to obtain the HPD in an election poll.

Example 2.9. Assume that n individuals are sampled without replacement from a population in which θ individuals vote for candidate A and $N - \theta$ individuals vote for candidate B, where N is larger than n and θ is unknown. Our goal is to infer θ using the total number of sampled individuals who vote for candidate A, X. Since individuals are sampled without replacement, $X|\theta \sim \text{Hypergeometric}(\theta, N - \theta, n)$.

If assume that, a priori, $\theta \sim \text{Binomial}(N, 0.5)$, then it can be shown that the posterior distribution is such that $\theta - X|X \sim \text{Binomial}(N - n, 0.5)$. Indeed, since $\theta \sim \text{Binomial}(n, 0.5)$, one can imagine that each individual in the population votes for candidate A independently with probability 0.5. After observing that X out of n vote for candidate A, this part of the population is fixed and the remainder part is such that each individual votes independently with probability 0.5 in candidate A. Hence, $\theta - X|X \sim \text{Binomial}(N - n, 0.5)$.

Since a Binomial distribution with probability 0.5 is concave and symmetric around its average, the HPD is an interval that is symmetric around the average of the Binomial. That is, for every choice of $k > 0$ and observed data $1 \leq x \leq n$, the HPD region for θ is given by

$$R(x) = \left\{0 \leq i \leq N : \left|\frac{(N-n)}{2} + x - i\right| \leq k\right\}, \tag{6}$$

[4] In Esteves et al. [39], Monotonicity was defined differently. Lemma A.2 shows the equivalence between the previous definition and the more elegant one presented in this paper, which was suggested by an anonymous referee.

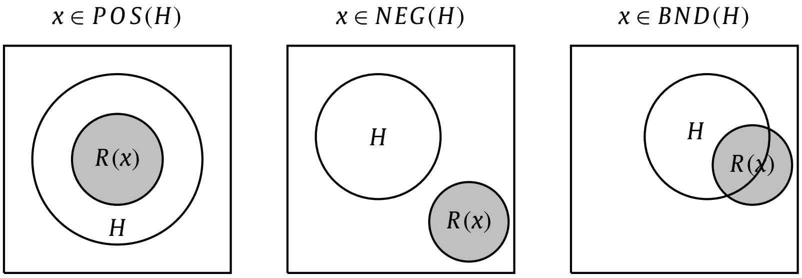

Fig. 2. φ is a region-based test for testing H.

that is, $R(x)$ contains the most probable values for θ, which in this case correspond to the values close to $x + \frac{N-n}{2}$. △

Using region estimators, one can construct a simultaneous test, as illustrated in Fig. 2. A test based on a region estimator, $R(x)$, accepts H if $R(x) \subseteq H$, that is, all likely values for θ reside in H. Similarly, it rejects H if $H \cap R(x) = \emptyset$, that is no likely value of θ resides in H. Otherwise, the test remains agnostic about H.

Definition 2.10 *(Region-based three-way test).* φ is a region-based test if there exists a region estimator, R, such that $x \in POS(H)$ if $R(x) \subseteq H$, $x \in NEG(H)$ if $R(x) \cap H = \emptyset$, and $x \in BND(H)$, otherwise, that is,

$$\varphi_H(x) = \begin{cases} 0 & \text{, if } R(x) \subseteq H \\ 1 & \text{, if } R(x) \cap H = \emptyset \\ \frac{1}{2} & \text{, otherwise} \end{cases} \tag{7}$$

The (non-invariant) Generalized Full Bayesian Significance Test (GFBST; Stern et al. [40]) is a particular type of test based on a region estimator. It uses a HPD as region estimator.

Definition 2.11. The GFBST is a region-based test in which $R(x)$ is a HPD.

Example 2.12 illustrates the GFBST for a particular problem.

Example 2.12. Consider the hypergeometric model in Example 2.9. The GFBST accepts a hypothesis H when H contains $R(x)$, that is, all points close to $\frac{N-n}{2} + x$ (where how close depends on the choice of k). Also, the GFBST rejects H when H is disjoint from $R(x)$, that is, H contains none of the points close to $\frac{N-n}{2} + x$. Finally, the GFBST remains agnostic whenever H contains some but not all of the points close to $\frac{N-n}{2} + x$. △

Although region-based tests are three-way tests, they generally differ from posterior-probability three-way tests. Hence, they provide a different class of three-way tests that necessarily provides logical coherence. Several applications of the GFBST, such as testing Hardy–Weinberg equilibrium, bioequivalence and linear regression have been developed [69,70].

Under special circumstances all logically coherent simultaneous tests are based on region estimators [39]. In particular, this relation is valid when Θ is a finite set.

Theorem 2.13. *If Θ is finite and φ is a logically coherent simultaneous test, then φ is based on a region estimator.*

The next section studies under what circumstances a Bayes test against an EC loss can be logically coherent. That is, the section studies whether a region-based three-way test can be a posterior-probability three-way test. Whenever such circumstances exist, probability-based tests are logically-coherent.

3. Can posterior-probability based three-way tests be logically coherent?

A logically coherent three-way test, φ, that is Bayes against an EC loss admits further characterization. In such a case, not only is φ a region-based test, but also based on a HPD. That is, every logically coherent posterior-probability three-way test is a GFBST, as presented in Theorem 3.1.[5]

[5] Lemma A.5, in the Appendix, is used to prove Theorem 3.1. Recall that if a test is logically coherent and Θ is a finite set, then the test is based on a region estimator. Lemma A.5 shows that, if a Bayes test is based on a region estimator, then there exists a loss such that the region estimator is Bayes. That is, a Bayes logically coherent test is necessarily based on a region estimator which is also Bayes.

Theorem 3.1. *Let Θ be a finite set. If there exists a probability, $\mathbb{P}$, and a TEC loss, L, such that a logically coherent simultaneous test, φ, is Bayes against L according to $\mathbb{P}$, then φ is a GFBST.*

Theorem 3.1 shows that, if a posterior-probability three-way test is logically coherent, then the test is a GFBST. That is, if the former conditions are desirable, then attention can be focused on the class of GFBST tests. For instance, if the thresholds in posterior probability three-way tests are chosen so that the test is logically coherent, then the test assumes the form of a GFBST. This result provides a way of obtaining logically coherent three-way tests.

However, do there exist actual cases in which a posterior-probability three-way test is logically coherent? Section 2 shows that a logically coherent test must be based on a region estimator. Despite this strong restriction, Theorem 3.2 shows that every logically coherent test is a probability-based three-way test for some probability measure.

Theorem 3.2. *Let Θ and $\mathcal{X}$ be finite sets. If φ is a logically coherent simultaneous test, then there exists a probability, $\mathbb{P}$, and a TEC loss function, L, such that φ is Bayes against L.*

Theorem 3.2 shows that, for each logically coherent test, there exists a choice of $\mathbb{P}$ and a TEC loss L such that the test is Bayes with respect to L.[6] That is, every logically coherent test is a posterior-probability three-way test for a specific choice of boundaries (α, β) and probability function, $\mathbb{P}$. Example 3.3 illustrates such a choice when the parameter space has 4 elements.

Example 3.3. Let $\Theta = \{1, 2, 3, 4\}$, $\mathcal{X} \in \{0, 1\}$, $R(0) = \{1, 2\}$, $R(1) = \{3, 4\}$, and φ be a test based on R. Let L be a TEC loss so that $\beta = \frac{7}{10}$ and $\alpha = \frac{3}{10}$. Also, let $\mathbb{P}(1|x) = \mathbb{P}(2|x) = \frac{4^{1-x}}{10}$ and $\mathbb{P}(3|x) = \mathbb{P}(4|x) = \frac{4^x}{10}$. Let φ^* be the Bayes test according to L. Note that the two least probable outcomes sum up a probability of $\frac{2}{10}$. Hence, every hypothesis that contains none of the most probable outcomes is rejected by φ^*. Next, if a hypothesis contains both of the most probable outcomes, than its probability is at least $\frac{8}{10}$, so it is accepted by φ^*. Finally, if a hypothesis contains only one of the most probable outcomes, than its probability is between $\frac{4}{10}$ and $\frac{6}{10}$, so φ^* remains agnostic about H. From the previous conclusions, obtain that $\varphi \equiv \varphi^*$, that is, φ is a logically coherent test that is Bayes against L according to $\mathbb{P}$. Finally, note that when using $\mathbb{P}$, R is a HPD, that is, φ is a GFBST, as also known from Theorem 3.1. △

Example 3.3 shows that, for a given three-way region-based test, a specific choice of TEC loss and $\mathbb{P}$ are required so that the test is Bayes. However, in most settings $\mathbb{P}$ is given and one wishes to choose L so that the Bayes test is logically coherent. Theorem 3.4 shows that there is no choice of an EC loss such that the Bayes test is logically coherent for every $\mathbb{P}$.

Theorem 3.4. *Let $|\Theta| \geq 3$. For each $\mathbb{P}$ and L, let $\varphi_{\mathbb{P},L}$ be a Bayes simultaneous test against L according to $\mathbb{P}$. If L is an EC loss, then there exists $\mathbb{P}$ such that $\varphi_{\mathbb{P},L}$ is not logically coherent.*

Theorem 3.4 shows that, if L is an EC loss, then there exists a probability, $\mathbb{P}$, such that the resulting Bayes test is not logically coherent. In particular, for every probability-based three-way test, there exists a probability such that the test is not logically coherent. Hence, a procedure that yields Bayes tests that are logically coherent against every probability $\mathbb{P}$ cannot be a probability-based three-way test and must be based on more general loss functions. The next section explores these losses and their resulting three-way tests.

4. A logically coherent Bayesian procedure

This section develops a loss function such that, for every probability, $\mathbb{P}$, the resulting Bayes three-way test is logically coherent. This loss is presented in Definition 4.1:

Definition 4.1 *(GFBST loss).* Let μ be a measure over Θ such that $\mathbb{P}(\theta|x)$ is absolutely continuous with respect to μ for every $x \in \mathcal{X}$ and $f(\theta|x) := \frac{dP(\theta|x)}{d\mu}$. For every $x \in \mathcal{X}$, the tangent set to hypothesis H according to μ, T_x^H, is defined as $T_x^H := \{\theta \in \Theta : f(\theta|x) > \sup_{\theta' \in H} f(\theta'|x)\}$. The GFBST loss according to μ for testing H is given by Table 2.

The GFBST loss, which generalizes the two-way counterpart in Madruga et al. [55], admits an intuitive interpretation [71]. Observe that $T_x^H \subseteq H^c$ is the collection of values in Θ that are more likely than every point in H. Hence, T_x^H and $T_x^{H^c}$ can be interpreted as the set of points that are strong contenders for, respectively, H and H^c. The GFBST loss is lowest, 0, when either H is rejected and θ is a strong contender for H or H is accepted and θ is a strong contender for H^c. Also the GFBST loss is largest, $b + c$, when either H is rejected and θ is a strong contender for H^c or H is accepted and θ is a strong

[6] Under mild assumptions, Theorems 3.1 and 3.2 also hold when Θ is a countable set.

Table 2
The GFBST loss. T_x^H are the points in the parameter space that are strong contenders against H after observing x. If θ is among the strong contenders against H, then one receives a large penalty for accepting H, an intermediate penalty for remaining undecided, and a small penalty for rejecting H. Similarly, if θ is among the strong contenders against H^c, then one receives a large penalty for rejecting H, an intermediate penalty for remaining undecided, and a small penalty for accepting H.

decision	state of the nature		
	$\theta \in T_x^H$	$\theta \notin T_x^H \cup T_x^{H^c}$	$\theta \in T_x^{H^c}$
0	$b+c$	b	0
$\frac{1}{2}$	$v+c$	v	$v+c$
1	0	b	$b+c$

contender for H. Finally, the GFBST loss assumes intermediate values, when either θ is not a strong contender for H or H^c or when the agnostic decision is chosen.

In the following, Theorem 4.2 shows that the optimal test against the GFBST loss is necessarily a GFBST three-way test and, therefore, logically coherent.

Theorem 4.2. *For every probability, $\mathbb{P}$, if φ is a Bayes simultaneous three-way test against the GFBST loss, then φ is a GFBST.*

Proof. The posterior expected losses for each decision are given by:

$$\mathbf{E}[L_A(0,(\theta,x)|x] = b\mathbb{P}(\theta \notin T_x^A \cup T_x^{A^c}|x) + (b+c)\mathbb{P}(\theta \in T_x^A|x), \tag{8}$$

$$\mathbf{E}\left[L_A\left(\frac{1}{2},(\theta,x\right)\bigg|x\right] = v + c\mathbb{P}(\theta \in T_x^A \cup T_x^{A^c}|x), \tag{9}$$

$$\mathbf{E}[L_A(1,(\theta,x)|x] = b\mathbb{P}(\theta \notin T_x^A \cup T_x^{A^c}|x) \;+\; (b+c)\mathbb{P}(\theta \in T_x^{A^c}|x)\;. \tag{10}$$

Next, it follows from definition that $T_x^A \subseteq A^c$ and $T_x^{A^c} \subseteq A$. Hence, $T_x^A \cap T_x^{A^c} = \emptyset$:

$$\mathbf{E}[L_A(0,(\theta,x)|x] - \mathbf{E}\left[L_A\left(\frac{1}{2},(\theta,x\right)\bigg|x\right] = (b+c)\mathbb{P}(\theta \notin T_x^{A^c}|x) - (v+c) \tag{11}$$

$$\mathbf{E}[L_A(0,(\theta,x)|x] - \mathbf{E}[L_A(1,(\theta,x)|x] = (b+c)\left(\mathbb{P}(\theta \in T_x^A|x) - \mathbb{P}(\theta \in T_x^{A^c}|x)\right) \tag{12}$$

$$\mathbf{E}[L_A(1,(\theta,x)|x] - \mathbf{E}\left[L_A\left(\frac{1}{2},(\theta,x\right)\bigg|x\right] = (b+c)\mathbb{P}(\theta \notin T_x^A|x) - (v+c) \tag{13}$$

Also, recall from definition that either $T_x^A = \emptyset$ or $T_x^{A^c} = \emptyset$. Hence, since $0 < v < b$ and $c > 0$, if φ is Bayes, then $\varphi_H(x) = 0$ if and only if $\mathbb{P}(\theta \notin T_x^{A^c}|x) < \frac{v+c}{b+c}$ and $\varphi_H(x) = 1$ if and only if $\mathbb{P}(\theta \notin T_x^A|x) < \frac{v+c}{b+c}$. It follows from Esteves et al. [39] that φ is the GFBST. □

Theorem 4.2 shows that, if the GFBST loss is used, then the Bayes test is a GFBST. Therefore, for every probability measure, the Bayes test against the GFBST loss is logically coherent. Hence, using loss functions that are more general than the EC loss, it is possible to always reconcile Bayesian optimality with logical coherence. In particular, if one wishes to obtain logical coherence, then performing decision-theoretic three-way tests based on Table 2 might be preferable to those based on Table 1. The three-way tests obtained from Table 2 assure logically coherent and generally differ from probability-based tests.

Example 4.3 illustrates the differences between posterior-probability and GFBST three-way tests through a simple case of observations from a normal population with unknown mean and known variance.

Example 4.3. Let $\mathbf{X} = (X_1, \ldots, X_n)$ be independent instances of a normal population with unknown mean, $\mu \in \mathbb{R}$, and known variance, $\sigma_0^2 = \frac{100}{3}$. Furthermore, consider that, a priori, one believes that μ is close to 0 and, furthermore, $\mu \sim N(0,1)$. Using Bayes theorem, one obtains that $\mu|\mathbf{X} \sim N\left(\frac{n\sigma_0^{-2}\bar{X}}{1+n\sigma_0^{-2}}, \frac{1}{1+n\sigma_0^{-2}}\right)$, where $\bar{X}$ is the sample mean. For the sake of illustration, if $n = 100$ and $\bar{X} = 4$, then one obtains that $\mu|\mathbf{X} \sim N(3, 0.25)$.

When applying posterior-probability three-way tests it is possible to obtain logical incoherence. For instance, consider such a three-way test with thresholds 0.05 and 0.95. Note that $\mathbb{P}(\mu \in [2.15, \infty[|\mathbf{x}) > 0.95$ and $\mathbb{P}(\mu \in]-\infty, 3.85]|\mathbf{x}) > 0.95$. Hence, both $\mu \in [2.15, \infty[$ and $\mu \in]-\infty, 3.85]$ are supported by $\mathbf{x}$. By combining both propositions, one obtains the logical deduction that $\mu \in [2.15, 3.85]$. Hence, one might expect such a proposition to be accepted. However, using the posterior distribution, one obtains that $\mathbb{P}(\mu \in [2.15, 3.85]|\mathbf{x}) \approx 0.91$. That is, one remains undecided whether $\mu \in [2.15, 3.85]$, a logical contradiction.

Such a contradiction cannot occur when using the GFBST three-way test. Given the posterior distribution, the 95% HPD set is approximately $[2, 4]$. Since $\mu \in [2.15, \infty[$ and $\mu \in]-\infty, 3.85]$ do not contain the HPD and are not disjoint from the HPD, one remains undecided about both propositions. Hence, no logical contradiction ensues. In this case, the GFBST would accept that μ lies in every set that contains $[2, 4]$, reject that μ lies in every set that is disjoint from $[2, 4]$, and remain undecided about the remaining cases. For instance, one would accept that $\mu \in [1, 10]$, reject that $\mu \in [-3, -1]$, and remain undecided about $\mu \in [1, 3]$.

5. Final remarks

Simultaneous three-way decisions may require more constraints than are typically used in individual decision problems. In particular, when performing simultaneous hypothesis tests, one might expect logical coherence between conclusions. This paper presents results on whether it is possible to obtain logically coherent three-way tests that are also Bayes optimal.

Two types of results are obtained. If an error-type constant loss is used, then only for a limited set of models can a Bayes simultaneous test be logically coherent. Specifically, a posterior probability three-way test can only be logically coherent if it is a GFBST test. This result motivated the investigation of other types of loss functions which might provide a better reconciliation between decision-theoretic three-way tests and logical coherence. We propose the GFBST loss and show that every Bayes test against this loss is a GFBST. Since every GFBST is logically coherent, the GFBST loss yields Bayes tests that are always logically coherent.

The above results show that the GFBST three-way test yields conclusions which are more interpretable than posterior probability three-way tests. The results also show that simultaneous three-way decisions can yield a layer of complexity that is not present in individual decision problems. Further investigation would determine statistical properties of logically consistent three-way hypothesis tests in more general settings, such as high-dimensional parameter spaces, functional spaces, or nonparametric models. Logically coherent three-way decisions can also be explored in other types of problems, such as in hierarchical classification. Given the hierarchical nature of the classes, one might wish for the classifier to respect the logical relations between the labels.

CRediT authorship contribution statement

The authors shared equal responsibilities in the development.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data and code is publicly available at github. Link is available in paper.

Acknowledgements

The author would like to thank the referees and editor who provided useful and detailed comments on an earlier version of the manuscript. The authors have received direct and indirect support from the following institutions: University of São Paulo (JMS and LGE); Federal University of São Carlos (RI and RBS); Brazilian Jurimetrics Association (JMS and RBS); Brazilian National Counsel of Technological and Scientific Development, grants PQ 303290/2021-8 (JMS), PQ 309607/2020-5 and 422705/2021-7 (RI); State of São Paulo Research Foundation, through CEPID Center for Mathematical Sciences Applied to Industry (grant 2013/07375-0, JMS), CEPID Shell-RCGI, Research Center for Greenhouse Gas Innovation (grant 2014/50279-4, JMS); CEPID Research, Innovation and Dissemination Center for Neuromathematics (NeuroMat, grant 2013/07699-0, RBS), and grant 2019/11321-9 (RI); Coordination of Superior Level Staff Improvement, financial code 001 (RI and RBS).

Appendix A. Proofs

Definition A.1. A simultaneous hypothesis test, φ, satisfies Monotonicity* if, for every $H_1 \subseteq H_2$, $x \in POS(H_1)$ implies that $x \in POS(H_2)$ and $x \in BND(H_1)$ implies that $x \in BND(H_2) \cup POS(H_2)$.

Lemma A.2. *Monotonicity* and Monotonicity (Definition 2.6.2) are equivalent.*

Proof. First, note that both definitions of monotonicity impose that $x \in POS(H_1)$ implies that $x \in POS(H_2)$.

Next, we show that, if Monotonicity* holds, then Monotonicity also holds:

$$\begin{cases} x \in POS(H_1) \rightarrow x \in POS(H_2) & \text{(Monotonicity)} \\ x \in BND(H_1) \rightarrow x \in (BND(H_2) \cup POS(H_2)) & \text{(Monotonicity)} \end{cases}$$

$$\begin{cases} x \notin POS(H_2) \rightarrow x \notin POS(H_1) & \text{(Counter-positive)} \\ x \notin (BND(H_2) \cup POS(H_2)) \rightarrow x \notin BND(H_1) & \text{(Counter-positive)} \end{cases}$$

$$\begin{cases} x \in NEG(H_2) \rightarrow x \notin POS(H_1) & (x \in NEG(H_2) \rightarrow x \notin POS(H_2)) \\ x \in NEG(H_2) \rightarrow x \in POS(H_1) \cup NEG(H_1) & \end{cases}$$

$$\begin{cases} x \in NEG(H_2) \rightarrow x \in NEG(H_1) \end{cases}$$

Finally, we show that, if Monotonicity holds, then Monotonicity* also holds. It follows from Monotonicity that $x \in NEG(H_2) \rightarrow x \in NEG(H_1)$. Hence, $x \notin NEG(H_1) \rightarrow x \notin NEG(H_2)$. Equivalently,

$$x \in (BND(H_1) \cup POS(H_1)) \rightarrow x \in (BND(H_2) \cup POS(H_2)).$$

Since $x \in BND(H_1) \rightarrow x \in (BND(H_1) \cup POS(H_1))$, the previous expression yields that $x \in BND(H_1) \rightarrow x \in (BND(H_2) \cup POS(H_2))$. □

Proof of Theorem 2.13. Let $\mathscr{F} = \{H \in \sigma(\Theta) : \forall H^* \in \sigma(\Theta), H \cap H^* \in \{\emptyset, H\}\}$. Since Θ is finite and $\sigma(\Theta)$ is a σ-field, $\mathscr{F}$ partitions Θ. Define the equivalence relation $\sim$ such that $\theta_1 \sim \theta_2$ if there exists $F \in \mathscr{F}$ such that $\theta_1 \in F$ and $\theta_2 \in F$. Define Θ^* as the quotient space $\Theta \backslash \sim$. Also, let $\sigma(\Theta^*)$ and φ^* be the quotient σ-field of $\sigma(\Theta)$ and the quotient test of φ over $\sim$. It follows from construction that $\sigma(\Theta^*)$ includes the singleton. Hence, Esteves et al. [39] obtains that φ^* is based on a region estimator, R^*. Conclude that φ is based on a region estimator, R. □

Definition A.3 *(Proper loss function).* A loss functions, L, is proper if, for every $A \in \sigma(\Theta)$,

$$L_A(0,\theta) < L_A\left(\frac{1}{2},\theta\right) < L_A(1,\theta), \text{if } \theta \in A \tag{A.1}$$

$$L_A(0,\theta) > L_A\left(\frac{1}{2},\theta\right) > L_A(1,\theta), \text{if } \theta \notin A \tag{A.2}$$

$$L_A\left(\frac{1}{2},\theta\right) < \frac{L_A(0,\theta) + L_A(1,\theta)}{2}, \forall\, \theta \in \Theta \tag{A.3}$$

Lemma A.4. *If L is a proper loss, then* $\min\left(\boldsymbol{E}\left[L_{\{\theta'\}}\left(\frac{1}{2},\theta\right)\middle|x\right], (\boldsymbol{E}\left[L_{\{\theta'\}}(0,\theta)\middle|x\right]\right) \leq \boldsymbol{E}\left[L_{\{\theta'\}}(1,\theta)\middle|x\right]$ *implies that* $\boldsymbol{E}\left[L_{\{\theta'\}}\left(\frac{1}{2},\theta\right)\middle|x\right] \leq \boldsymbol{E}\left[L_{\{\theta'\}}(1,\theta)\middle|x\right]$.

Proof. It is sufficient to prove that, if $\mathbf{E}\left[L_{\{\theta'\}}(0,\theta)\middle|x\right] \leq \mathbf{E}\left[L_{\{\theta'\}}(1,\theta)\middle|x\right]$, then $\mathbf{E}\left[L_{\{\theta'\}}\left(\frac{1}{2},\theta\right)\middle|x\right] \leq \mathbf{E}\left[L_{\{\theta'\}}(1,\theta)\middle|x\right]$. Let $\mathbf{E}\left[L_{\{\theta'\}}(0,\theta)\middle|x\right] \leq \mathbf{E}\left[L_{\{\theta'\}}(1,\theta)\middle|x\right]$. Since L is proper,

$$\mathbf{E}\left[L_{\{\theta'\}}\left(\frac{1}{2},\theta\right)\middle|x\right] \leq \frac{\mathbf{E}\left[L_{\{\theta'\}}(0,\theta)\middle|x\right]}{2} + \frac{\mathbf{E}\left[L_{\{\theta'\}}(1,\theta)\middle|x\right]}{2} \tag{A.4}$$

$$\leq \mathbf{E}\left[L_{\{\theta'\}}(1,\theta)\middle|x\right]. \quad \square \tag{A.5}$$

Lemma A.5. *Let Θ be finite, $\sigma(\Theta)$ include the unitary sets, and φ be generated by the region estimator, R. If there exists a probability, $\mathbb{P}$, and a proper loss, L, such that φ is Bayes against L according to $\mathbb{P}$, then R is a Bayes region estimator against $\bar{L}$ according to $\mathbb{P}$, where*

$$\bar{L}(A,\theta) = \sum_{\theta' \in A}\left[L_{\{\theta'\}}\left(\frac{1}{2},\theta\right) - L_{\{\theta'\}}(1,\theta)\right]. \tag{A.6}$$

Table A.3
Loss function used in the proof of Theorem 3.2.

Decision	state of the nature	
	$\theta \in A$	$\theta \notin A$
0 (accept A)	0	k
$\frac{1}{2}$ (remain agnostic about A)	1	1
1 (reject A)	k	0

Proof. The Bayes region estimator against $\bar{L}$, R^*, satisfies:

$$R^*(x) := \left\{\theta' \in \Theta : \mathbf{E}\left[L_{\{\theta'\}}\left(\frac{1}{2},\theta\right)\middle|x\right] \leq \mathbf{E}\left[L_{\{\theta'\}}(1,\theta)\middle|x\right]\right\}. \tag{A.7}$$

Hence, it is sufficient to prove that $R \equiv R^*$. Since φ is Bayes against L, $\varphi_{\{\theta'\}}(x) < 1$ if and only if $\min\left(\mathbf{E}\left[L_{\{\theta'\}}(\frac{1}{2},\theta)\middle|x\right], (\mathbf{E}\left[L_{\{\theta'\}}\left(0,\theta\right)\middle|x\right]\right) < \mathbf{E}\left[L_{\{\theta'\}}(1,\theta)\middle|x\right]$. Using Lemma A.4, conclude that $\varphi_{\{\theta'\}}(x) < 1$ if and only if $\mathbf{E}\left[L_{\{\theta'\}}(\frac{1}{2},\theta)\middle|x\right] \leq \mathbf{E}\left[L_{\{\theta'\}}(1,\theta)\middle|x\right]$. Since φ is generated by R, it follows that $R(x) = \{\theta' : \varphi_{\{\theta'\}}(x) < 1\}$, that is,

$$R(x) = \left\{\theta' : \varphi_{\{\theta'\}}(x) < 1\right\} \tag{A.8}$$

$$= \left\{\theta' : \mathbf{E}\left[L_{\{\theta'\}}\left(\frac{1}{2},\theta\right)\middle|x\right] \leq \mathbf{E}\left[L_{\{\theta'\}}(1,\theta)\middle|x\right]\right\} \equiv R^*(x) \quad \square \tag{A.9}$$

Proof of Theorem 3.1. Since φ is logically coherent, it follows from Theorem 2.13 that φ is based on a region estimator, R. It follows from Lemma A.5 that R is a Bayes region estimator against $\bar{L}$. Since L is a TEC loss, which is proper, $\bar{L}(A,\theta) = \lambda_{BN}|A| - ((\lambda_{BN} - \lambda_{BP}) + \lambda_{NP})\mathbb{I}_A(\theta)$. That is,

$$R(x) = \left\{\theta \in \Theta : \mathbb{P}(\theta|x) \geq \frac{\lambda_{BN}}{(\lambda_{BN} - \lambda_{BP}) + \lambda_{NP}}\right\}. \tag{A.10}$$

Conclude that $R(x)$ is a HPD. $\square$

Lemma A.6 *(Union consonance). Let φ be logically coherent. If H_1 and H_2 are such that $\varphi_{H_1}(x) = 1$ and $\varphi_{H_2}(x) = 1$, then $\varphi_{H_1 \cup H_2}(x) = 1$.*

Proof. It follows from invertibility that $\varphi_{H_1^c}(x) = 0$ and $\varphi_{H_2^c}(x) = 0$. Hence, from intersection consonance, $\varphi_{H_1^c \cap H_2^c}(x) = 0$. Finally, conclude from invertibility that $\varphi_{H_1 \cup H_2}(x) = 1$. $\square$

Lemma A.7. *Let Θ be a finite set. If φ is a logically coherent simultaneous test, then:*

(a) For every $x \in \mathcal{X}$, there exists $\theta_0 \in \Theta$ such that $\varphi_{\{\theta_0\}}(x) < 1$.
(b) For every $x \in \mathcal{X}$, if $\varphi_{\{\theta_0\}}(x) = 0$, then $\varphi_{\{\theta\}}(x) = 1$, $\forall\theta \neq \theta_0$.

Proof. **(a)** Assume that there exists $x \in \mathcal{X}$ such that $\varphi_{\{\theta\}}(x) = 1$, for every $\theta \in \Theta$. It follows from Lemma A.6 that $\varphi_\Theta(x) = 1$, which contradicts the propriety of φ. **(b)** Let θ_0 be such that $\varphi_{\{\theta_0\}}(x) = 0$. It follows from invertibility that $\varphi_{\{\theta_0\}^c}(x) = 1$. Conclude from monotonicity that, for every $\theta \neq \theta_0$, $\varphi_{\{\theta\}}(x) = 1$. $\square$

Proof of Theorem 3.2. Since φ is logically coherent, it follows from Esteves et al. [39] that there exists $R(x)$ such that, $\varphi_H(x) = 1 \Leftrightarrow H \cap R(x) = \emptyset$, $\varphi_H(x) = 0 \Leftrightarrow R(x) \subseteq H$ and $\varphi_H(x) = \frac{1}{2}$, otherwise. Using Lemma A.7, conclude that $R(x) \neq \emptyset$. In the following, we determine a loss, L, and a joint probability, $\mathbb{P}(\theta, x)$, such that φ is Bayes.

Let $|\Theta| = k$. Also, let L be the TEC given by Table A.3. It follows from Yao [48] that φ is Bayes with respect to L when:

$$\varphi_H(x) = \begin{cases} 1 & \text{, if } \mathbb{P}(\theta \in H|x) < \frac{1}{k} \\ 0 & \text{, if } \mathbb{P}(\theta \in H|x) > \frac{k-1}{k} \\ \frac{1}{2} & \text{, otherwise.} \end{cases} \tag{A.11}$$

Next, we determine $\mathbb{P}(\theta, x)$ such that these conditions hold.

In order to determine $\mathbb{P}(\theta, x)$ it is sufficient to choose $\mathbb{P}(x)$ and $\mathbb{P}(\theta|x)$. For each $A \subset \mathcal{X}$, let $\mathbb{P}(x \in A) = \frac{|A|}{|\mathcal{X}|}$, that is, the uniform distribution over $\mathcal{X}$. Also, for $H \subset \Theta$,

$$\mathbb{P}(\theta \in H|x) = \frac{1}{2k} \cdot \frac{|H|}{|\mathcal{X}|} + \frac{2k-1}{2k} \cdot \frac{|H \cap R(x)|}{|R(x)|}. \tag{A.12}$$

It remains to show that φ is Bayes with respect to L and $\mathbb{P}$. We study three cases: **(i)** If $\varphi_H(x) = 1$, then $H \cap R(x) = \emptyset$. Using Equation (A.12), conclude that $\mathbb{P}(\theta \in H|x) \le \frac{1}{2k} \cdot 1 + \frac{2k-1}{2k} \cdot 0 < \frac{1}{k}$, **(ii)** If $\varphi_H(x) = 0$, then $R(x) \subseteq H$. Using Equation (A.12), conclude that $\mathbb{P}(\theta \in H|x) \ge \frac{1}{2k} \cdot 0 + \frac{2k-1}{2k} \cdot 1 > \frac{k-1}{k}$, **(iii)** If $\varphi_H(x) = \frac{1}{2}$, then $R(x) \cap H^c \neq \emptyset$ and $R(x) \cap H \neq \emptyset$, that is, $1 \le |H \cap R(x)| < |R(x)| \le k$. Using Equation (A.12), conclude that $\mathbb{P}(\theta \in H|x) \ge \frac{1}{2k} \cdot \frac{1}{k} + \frac{2k-1}{2k} \cdot \frac{1}{k} = \frac{1}{k}$. Also, $\mathbb{P}(\theta \in H|x) \le \frac{1}{2k} \cdot \frac{k-1}{k} + \frac{2k-1}{2k} \cdot \frac{k-1}{k} = \frac{k-1}{k}$. That is, $\frac{1}{k} \le \mathbb{P}(\theta \in H|x) \le \frac{k-1}{k}$. It follows from Equation (A.11) that φ is Bayes with respect to L using $\mathbb{P}$. □

Lemma A.8. *Let L be an EC loss Definition 2.1 and, for each $\mathbb{P}$, let $\varphi_{\mathbb{P},L}$ be a Bayes simultaneous test for $\mathbb{P}$ against L. If, for every $\mathbb{P}$, $\varphi_{\mathbb{P},L}$ is logically coherent, then $\varphi_{\mathbb{P},L}$ is a simultaneous test such as in Example 2.2 and:*

1. *for every $A, B \in \sigma(\Theta)$ such that $\emptyset \neq A \subseteq B \neq \Omega$, $\alpha^A \ge \alpha^B$.*
2. *for every $A, B \in \sigma(\Theta)$ such that $A - B \neq \emptyset$, $B - A \neq \emptyset$, and $A \cup B \neq \Omega$: $\alpha^A + \alpha^B \le \alpha^{A \cup B}$.*

Proof. Let $x \in \mathcal{X}$ be arbitrary.

If $\alpha^A < \alpha^B$, then for $\mathbb{P}$ such that $\mathbb{P}(\theta \in A|x) = \mathbb{P}(\theta \in B|x) = 0.5(\alpha^A + \alpha^B)$, $\varphi_{\mathbb{P},L}(A) < 1$ and $\varphi_{\mathbb{P},L}(B) = 1$, that is, $\varphi_{\mathbb{P},L}$ does not satisfy monotonicity. Conclude that, if $\varphi_{\mathbb{P},L}$ is logically coherent for every $\mathbb{P}$, then $\alpha^A \ge \alpha^B$ for every $\emptyset \neq A \subseteq B \neq \Omega$.

If $\alpha^A + \alpha^B > \alpha^{A \cup B}$, then let $\delta := (\alpha^A + \alpha^B) - \alpha^{A \cup B} > 0$. By taking $\mathbb{P}$ such that

$$\mathbb{P}(\theta \in A|x) = \max(0, \alpha^A - 0.4\delta), \tag{A.13}$$

$$\mathbb{P}(\theta \in B|x) = \max(0, \alpha^B - 0.4\delta), \tag{A.14}$$

$$\mathbb{P}(\theta \in A \cup B|x) = \min(1, \alpha^A + \alpha^B - 0.8\delta), \tag{A.15}$$

obtain $\varphi_{\mathbb{P},L}(A) = 1$, $\varphi_{\mathbb{P},L}(B) = 1$, and $\varphi_{\mathbb{P},L}(A \cup B) < 1$, that is, it follows from Lemma A.6 that $\varphi_{\mathbb{P},L}$ is not logically coherent. Conclude that, if $\varphi_{\mathbb{P},L}$ is logically coherent for every $\mathbb{P}$, then $\alpha^A + \alpha^B \le \alpha^{A \cup B}$. □

Proof of Theorem 3.4. Assume that, for every $\mathbb{P}$, $\varphi_{L,\mathbb{P}}$ is logically coherent. Let $\theta_1, \theta_2 \in \Theta$ and $A = \{\theta_1\}$, $B = \{\theta_2\}$. Since $|\Theta| \ge 3$, $A - B \neq \emptyset$, $B - A \neq \emptyset$ and $A \cup B \neq \Omega$. Hence, it follows from Lemma A.8 that

$$\alpha^A \ge \alpha^{A \cup B}, \tag{A.16}$$

$$\alpha^B \ge \alpha^{A \cup B}, \tag{A.17}$$

$$\alpha^{A \cup B} \ge \alpha^A + \alpha^B. \tag{A.18}$$

That is, $\alpha^A = \alpha^B = \alpha^{A \cup B} = 0$, a contradiction with Example 2.2. Conclude that there exists $\mathbb{P}$ such that $\varphi_{L,\mathbb{P}}$ is not logically coherent. □

References

[1] Yiyu Yao, An outline of a theory of three-way decisions, in: International Conference on Rough Sets and Current Trends in Computing, Springer, 2012, pp. 1–17.
[2] Dun Liu, Decui Liang, An overview of function based three-way decisions, in: International Conference on Rough Sets and Knowledge Technology, Springer, 2014, pp. 812–823.
[3] Yiyu Yao, Rough sets and three-way decisions, in: International Conference on Rough Sets and Knowledge Technology, Springer, 2015, pp. 62–73.
[4] Yiyu Yao, Tri-level thinking: models of three-way decision, Int. J. Mach. Learn. Cybern. 11 (5) (2020) 947–959.
[5] Yiyu Yao, Set-theoretic models of three-way decision, Granul. Comput. 6 (1) (2021) 133–148.
[6] Yiyu Yao, The geometry of three-way decision, Appl. Intell. 51 (9) (2021) 6298–6325.
[7] Hong Yu, A framework of three-way cluster analysis, in: International Joint Conference on Rough Sets, Springer, 2017, pp. 300–312.
[8] Bing Zhou, Multi-class decision-theoretic rough sets, Int. J. Approx. Reason. 55 (1) (2014) 211–224.
[9] Daryl H. Hepting, Hadeel Hatim Bin Amer, Yiyu Yao, Facial similarity analysis: a three-way decision perspective, in: Uncertainty Management with Fuzzy and Rough Sets, Springer, 2019, pp. 289–313.
[10] Yan Zhang, PengFei Liu, JingTao Yao, Three-way email spam filtering with game-theoretic rough sets, in: 2019 International Conference on Computing, Networking and Communications (ICNC), IEEE, 2019, pp. 552–556.
[11] Huaxiong Li, Libo Zhang, Bing Huang, Xianzhong Zhou, Cost-sensitive dual-bidirectional linear discriminant analysis, Inf. Sci. 510 (2020) 283–303.
[12] Yu Fang, Cong Gao, Yiyu Yao, Granularity-driven sequential three-way decisions: a cost-sensitive approach to classification, Inf. Sci. 507 (2020) 644–664.
[13] Mengjun Hu, Three-way Bayesian confirmation in classifications, Cogn. Comput. (2021) 1–20.
[14] JingTao Yao, Yiyu Yao, Davide Ciucci, Kaizhu Huang, Granular computing and three-way decisions for cognitive analytics, Cogn. Comput. (2022) 1–4.
[15] Guangming Lang, Duoqian Miao, Hamido Fujita, Three-way group conflict analysis based on Pythagorean fuzzy set theory, IEEE Trans. Fuzzy Syst. 28 (3) (2019) 447–461.
[16] Yiyu Yao, Three-way conflict analysis: reformulations and extensions of the Pawlak model, Knowl.-Based Syst. 180 (2019) 26–37.
[17] Guangming Lang, Yiyu Yao, New measures of alliance and conflict for three-way conflict analysis, Int. J. Approx. Reason. 132 (2021) 49–69.

[18] Andrea Campagner, Davide Ciucci, Carl-Magnus Svensson, Marc Thilo Figge, Federico Cabitza, Ground truthing from multi-rater labeling with three-way decision and possibility theory, Inf. Sci. 545 (2021) 771–790.
[19] Joseph P. Herbert, JingTao Yao, Game-theoretic rough sets, Fundam. Inform. 108 (3–4) (2011) 267–286.
[20] Nouman Azam, JingTao Yao, Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets, Int. J. Approx. Reason. 55 (1) (2014) 142–155.
[21] Zia Bashir, Saima Mahnaz, Muhammad Ghulam Abbas Malik, Conflict resolution using game theory and rough sets, Int. J. Intell. Syst. 36 (1) (2021) 237–259.
[22] Yiyu Yao, Three-way decision and granular computing, Int. J. Approx. Reason. 103 (2018) 107–123.
[23] Haibo Jiang, Bao Qing Hu, A novel three-way group investment decision model under intuitionistic fuzzy multi-attribute group decision-making environment, Inf. Sci. 569 (2021) 557–581.
[24] Jinxing Zhu, Xueling Ma, Jianming Zhan, Yiyu Yao, A three-way multi-attribute decision making method based on regret theory and its application to medical data in fuzzy environments, Appl. Soft Comput. (2022) 108975.
[25] Xiaoping Yang, JingTao Yao, Modelling multi-agent three-way decisions with decision-theoretic rough sets, Fundam. Inform. 115 (2–3) (2012) 157–171.
[26] Jiubing Liu, Huaxiong Li, Bing Huang, Yu Liu, Dun Liu, Convex combination-based consensus analysis for intuitionistic fuzzy three-way group decision, Inf. Sci. 574 (2021) 542–566.
[27] Bingzhen Sun, Weimin Ma, Binjiang Li, Xiaonan Li, Three-way decisions approach to multiple attribute group decision making with linguistic information-based decision-theoretic rough fuzzy set, Int. J. Approx. Reason. 93 (2018) 424–442.
[28] Fan Min, Zhi-Heng Zhang, Wen-Jie Zhai, Rong-Ping Shen, Frequent pattern discovery with tri-partition alphabets, Inf. Sci. 507 (2020) 715–732.
[29] Heng-Ru Zhang, Fan Min, Bing Shi, Regression-based three-way recommendation, Inf. Sci. 378 (2017) 444–461.
[30] Xiaoqing Ye, Dun Liu, A cost-sensitive temporal-spatial three-way recommendation with multi-granularity decision, Inf. Sci. (2022).
[31] Zdzisław Pawlak, Rough Sets: Theoretical Aspects of Reasoning About Data, vol. 9, Springer Science & Business Media, 1991.
[32] Y.Y. Yao, S.K. Michael Wong, A decision theoretic framework for approximating concepts, Int. J. Man-Mach. Stud. 37 (6) (1992) 793–809.
[33] Abraham Wald, Sequential tests of statistical hypotheses, Ann. Math. Stat. 16 (2) (1945) 117–186.
[34] Henry F. Kaiser, Directional statistical decisions, Psychol. Rev. 67 (3) (1960) 160.
[35] John W. Tukey, Conclusions vs decisions, Technometrics 2 (4) (1960) 423–433.
[36] Richard J. Harris, Reforming significance testing via three-valued logic, in: What If There Were No Significance Tests?: Classic Edition, 2016, p. 131.
[37] Nathan Berg, No-decision classification: an alternative to testing for statistical significance, J. Socio-Econ. 33 (5) (2004) 631–650.
[38] Rob Goudey, Do statistical inferences allowing three alternative decisions give better feedback for environmentally precautionary decision-making?, J. Environ. Manag. 85 (2) (2007) 338–344.
[39] L.G. Esteves, R. Izbicki, J.M. Stern, R.B. Stern, The logical consistency of simultaneous agnostic hypothesis tests, Entropy 18 (7) (2016) 256.
[40] Julio Michael Stern, Rafael Izbicki, Luis Gustavo Esteves, Rafael Bassi Stern, Logically-consistent hypothesis testing and the hexagon of oppositions, Log. J. IGPL 25 (5) (2017) 741–757.
[41] Branden Fitelson, Studies in Bayesian Confirmation Theory, The University of Wisconsin-Madison, 2001.
[42] Hannes Leitgeb, Reducing belief simpliciter to degrees of belief, Ann. Pure Appl. Log. 164 (12) (2013) 1338–1389.
[43] Yoav Benjamini, Yosef Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc., Ser. B, Stat. Methodol. 57 (1) (1995) 289–300.
[44] R. Izbicki, L.G. Esteves, Logical consistency in simultaneous statistical test procedures, Log. J. IGPL 23 (5) (2015) 732–758.
[45] K. Ruben Gabriel, Simultaneous test procedures–some theory of multiple comparisons, Ann. Math. Stat. (1969) 224–250.
[46] Gerhard Hommel, Frank Bretz, Aesthetics and power considerations in multiple testing–a contradiction?, Biom. J., J. Math. Methods Biosci. 50 (5) (2008) 657–666.
[47] Joseph P. Romano, Azeem Shaikh, Michael Wolf, Consonance and the closure method in multiple testing, Int. J. Biostat. 7 (1) (2011).
[48] Yiyu Yao, Decision-theoretic rough set models, in: International Conference on Rough Sets and Knowledge Technology, Springer, 2007, pp. 1–12.
[49] M.H. DeGroot, Optimal Statistical Decisions, vol. 82, John Wiley & Sons, 2005.
[50] Yiyu Yao, Three-way decisions with probabilistic rough sets, Inf. Sci. 180 (3) (2010) 341–353.
[51] Yiyu Yao, The superiority of three-way decisions in probabilistic rough set models, Inf. Sci. 181 (6) (2011) 1080–1096.
[52] Jiubing Liu, Bing Huang, Huaxiong Li, Xiangzhi Bu, Xianzhong Zhou, Optimization-based three-way decisions with interval-valued intuitionistic fuzzy information, IEEE Trans. Cybern. (2022).
[53] Giovanni Parmigiani, Lurdes Inoue, Decision Theory: Principles and Approaches, John Wiley & Sons, 2009.
[54] M.J. Schervish, Theory of Statistics, Springer Science & Business Media, 2012.
[55] M. Regina Madruga, Luis G. Esteves, Sergio Wechsler, On the bayesianity of Pereira-Stern tests, Test 10 (2) (2001) 291–299.
[56] Dominik Slezak, Wojciech Ziarko, The investigation of the Bayesian rough set model, Int. J. Approx. Reason. 40 (1–2) (2005) 81–91.
[57] Salvatore Greco, Benedetto Matarazzo, Roman Słowiński, Parameterized rough set model using rough membership and Bayesian confirmation measures, Int. J. Approx. Reason. 49 (2) (2008) 285–300.
[58] Jiubing Liu, Huaxiong Li, Xianzhong Zhou, Bing Huang, Tianxing Wang, An optimization-based formulation for three-way decisions, Inf. Sci. 495 (2019) 185–214.
[59] Yiyu Yao, Jilin Yang, Granular rough sets and granular shadowed sets: three-way approximations in Pawlak approximation spaces, Int. J. Approx. Reason. 142 (2022) 231–247.
[60] Xiaofei Deng, Yiyu Yao, An information-theoretic interpretation of thresholds in probabilistic rough sets, in: International Conference on Rough Sets and Knowledge Technology, Springer, 2012, pp. 369–378.
[61] Joseph P. Herbert, JingTao Yao, Game-theoretic rough sets, Fundam. Inform. 108 (3–4) (2011) 267–286.
[62] Jiubing Liu, Jiaxin Mai, Huaxiong Li, Bing Huang, Yongjun Liu, On three perspectives for deriving three-way decision with linguistic intuitionistic fuzzy information, Inf. Sci. 588 (2022) 350–380.
[63] Juliet Popper Shaffer, Multiple hypothesis testing, Annu. Rev. Psychol. 46 (1) (1995) 561–584.
[64] Erich Leo Lehmann, Joseph P. Romano, George Casella, Testing Statistical Hypotheses, vol. 3, Springer, 2005.
[65] Mark J. Schervish, P values: what they are and what they are not, Am. Stat. 50 (3) (1996) 203–206.
[66] Michael Lavine, Mark J. Schervish, Bayes factors: what they are and what they are not, Am. Stat. 53 (2) (1999) 119–122.
[67] Spencer Hansen, Ken Rice, Coherent tests for interval null hypotheses, Am. Stat. (2022) 1–21.
[68] Giovanni Parmigiani, Lurdes Inoue, Decision Theory: Principles and Approaches, John Wiley & Sons, 2009.
[69] Victor Coscrato, Luís Gustavo Esteves, Rafael Izbicki, Rafael Bassi Stern, Interpretable hypothesis tests, arXiv preprint, arXiv:1904.06605, 2019.
[70] Luis Gustavo Esteves, Rafael Izbicki, Julio Michael Stern, Rafael Bassi Stern, Pragmatic hypotheses in the evolution of science, Entropy 21 (9) (2019) 883.
[71] Julio Michael Stern, Significance tests, belief calculi, and burden of proof in legal and scientific discourse, Front. Artif. Intell. Appl. 101 (2003).