Self-consistent confidence sets and tests of composite hypotheses applicable to restricted parameters

August 19, 2014

David R. Bickel

Ottawa Institute of Systems Biology

Department of Biochemistry, Microbiology, and Immunology

Department of Mathematics and Statistics

University of Ottawa; 451 Smyth Road; Ottawa, Ontario, K1H 8M5

Abstract

Frequentist methods, without the coherence guarantees of fully Bayesian methods, are known to yield self-contradictory inferences in certain settings. The framework introduced in this paper provides a simple adjustment to p-values and confidence sets to ensure the mutual consistency of all inferences without sacrificing frequentist validity. Based on a definition of the compatibility of a composite hypothesis with the observed data given any parameter restriction and on the requirement of self-consistency, the adjustment leads to the possibility and necessity measures of possibility theory rather than to the posterior probability distributions of Bayesian and fiducial inference.

Keywords: bounded parameter; confidence distribution; confidence curve; deductive closure; deductive cogency; empty confidence set; foundations of statistics; possibility theory; p-value function; ranking function; ranking theory; restricted parameter space; significance function; surprise measure

1 Introduction

A common criticism of frequentist statistical methods is that they lead to contradictory conclusions in situations in settings where Bayesian methods cannot. Following Kaplan (1996), a method of hypothesis testing or set estimation will be called *deductively cogent* if it cannot make mutually contradictory rejections of hypotheses. Minimal requirements for a deductively cogent method of hypothesis testing are the following:

1. It is *restriction-respecting* in the sense that it cannot reject every hypothesis that is consistent with the restriction imposed and in that it rejects all hypotheses that are inconsistent with the restriction.

2. It is *coherent* in the sense that a hypothesis can only be rejected if every hypothesis implying it is also rejected (Gabriel, 1969).

Standard confidence procedures often fail to meet the first requirement in the presence of parameter restrictions, which are often encountered in physics. For example, if the parameter restriction is a bound on the parameter of interest, then inference should proceed conditional on that bound. However, confidence intervals can be partially or entirely outside the bound (Mandelkern, 2002; Fraser, 2011); cf. Zhang and Woodroofe (2003); Marchand and Strawderman (2004); Wang (2007); Marchand and Strawderman (2013). Taking the intersection of the parameter restriction set and the confidence set leads in the former case to truncating the confidence set at the bound, and in the latter case to an empty confidence set. Since parameter values outside a confidence set are considered rejected, an empty confidence set is equivalent to rejecting the entire set of possible parameter values, contradicting the condition that the parameter value lies in that set.

Empty confidence sets also occur for an epidemiological model, a branching process, and Brownian motion (Ball et al., 2002). While an empty confidence is often interpreted as an indication of model inadequacy, procedures leading to them also lead to very small confidence sets, misleadingly indicating accurate knowledge of the parameter value (Ball et al., 2002). As a result, such confidence sets do not give the estimates of uncertainty that are needed in practice (Mandelkern, 2002; Wang, 2006).

For an example of violating coherence, one-sided p-values are interpreted as attained confidence levels of composite hypotheses, including those concerning the value of an unbounded parameter. Since such attained confidence levels can be smaller for a region than for a region it contains (Efron and Tibshirani, 1998; Polansky, 2007, pp. 224-227), they do not correspond to coherent hypothesis tests. The fact that frequentist approaches can violate coherence has led many to develop methods complying with the strong likelihood principle, whether using prior distributions (e.g., Schervish (1996); Lavine and Schervish (1999)) or not (e.g., Royall (1997); Bickel (2012); Zhang and Zhang (2013)).

To render existing frequentist methods deductively cogent, this paper instead presents an alternative framework of hypothesis testing and confidence sets. The framework is based on the concept of the compatibility between a hypothesis and the observed data rather than on any likelihood principle.

That data-compatibility measure is specified by the first principles of Section 2. Section 3 derives properties of the data compatibility of a hypothesis, including the fact that the data compatibility of a point null hypothesis is the p-value divided by the highest p-value corresponding to the point null hypotheses in the parameter space or in the parameter restriction, if any. As a result, the corresponding set estimate is a conservative confidence set. Two simple examples are provided in Section 4, one of which features a bounded parameter problem. Section 5 introduces the concept of the acceptability of a hypothesis in order to indicate when to accept the hypothesis, when to reject it, and when to take neither of those actions. The restriction-respecting and coherence aspects of that procedure are proven in Section 6. Finally, Section 7 remarks on the place of the proposed framework in possibility theory and ranking theory.

2 First principles of data-hypothesis compatibility

2.1 Preliminary notation

The unknown values θ and γ of the parameter of interest and of the nuisance parameter are members of the sets Θ and Γ, respectively. The observed tuple x is a member of some set $\mathcal{X}$ of possible observations.

A function $p(\bullet;\bullet) : \Theta \times \mathcal{X} \to]0,1[$ is a p-value *function* if

$$P_{\theta_0,\gamma}\left(p\left(\theta_0; X\right) < \alpha\right) = \alpha \tag{1}$$

for all $\theta_0 \in \Theta$ and $0 \le \alpha \le 1$. Each $p(\theta_0; x)$ is the p-value for testing the hypothesis that $\theta = \theta_0$ given the observation that $X = x$. While usual p-value functions are isomorphic to confidence distributions (Bickel and Padilla, 2014; cf. Schweder and Hjort, 2002; Xie and Singh, 2013), the concept of the observed confidence level (Polansky, 2007), a belief-type probability according to a confidence distribution, plays no role in the current paper, in which probability is always of the frequency type (see Hacking, 2001).

For the purpose of representing hypotheses, $\mathfrak{H}$ will denote a σ-field of subsets of Θ. For any $\mathcal{H}_0 \in \mathfrak{H}$, the hypothesis that $\theta \in \mathcal{H}_0$ is *simple* if $|\mathcal{H}_0| = 1$ and *composite* if $|\mathcal{H}_0| > 1$.

2.2 Whether a composite hypothesis is compatible with data

The phrase "The hypothesis that $\theta \in \mathcal{H}_0$ is compatible" herein abbreviates "The hypothesis that θ is a member of $\mathcal{H}_0$ is compatible" rather than "The hypothesis that θ, which is a member of $\mathcal{H}_0$, is compatible." A hypothesis about a parameter value, not a parameter value itself, may be compatible, rejected, etc.

What it means for a hypothesis to be "compatible" with data is defined in analogy with confidence intervals. Recall that, for any restriction of θ to a set $\mathcal{R} \in \mathfrak{H} \setminus \{\emptyset\}$, the set

$$\mathrm{CS}\left(\alpha; x | \mathcal{R}\right) = \left\{\theta_0 \in \mathcal{R} : p\left(\theta_0; x\right) \geq \alpha\right\} \tag{2}$$

is called an *exact* $(1-\alpha)\,(100\%)$-*confidence set* for any $\theta_0 \in \mathcal{R}$ since

$$P_{\theta_0,\gamma}\left(\theta_0 \in \mathrm{CS}\left(\alpha; X | \mathcal{R}\right)\right) = 1 - \alpha$$

for all $\alpha \in]0, 1]$ and $\gamma \in \Gamma$ results from equation (1).

What it means for a hypothesis to be sufficiently compatible with data is defined in terms of the still-undefined concept of compatibility.

Definition 1. For any $\mathcal{H}_0, \mathcal{R} \in \mathfrak{H} \setminus \{\emptyset\}$, $x \in \mathcal{X}$, and $\alpha \in]0, 1]$, the hypothesis that $\theta \in \mathcal{H}_0$ is α-*compatible* with the observation that $X = x$, conditional on the restriction that $\theta \in \mathcal{R}$, if there is a $\theta_0 \in \mathcal{H}_0$ such that $c\left(\theta_0; x | \mathcal{R}\right) \geq \alpha$, where $c\left(\theta_0; x | \mathcal{R}\right)$ is the *compatibility* of the hypothesis that $\theta = \theta_0$ with the observation that $X = x$ conditional on the restriction that $\theta \in \mathcal{R}$. The α-*compatibility set* given $X = x$ and $\theta \in \mathcal{R}$ is

$$\mathcal{H}\left(\alpha; x | \mathcal{R}\right) = \left\{\theta_0 \in \Theta : c\left(\theta_0; x | \mathcal{R}\right) \geq \alpha\right\} \tag{3}$$

for all $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$, $x \in \mathcal{X}$, and $\alpha \in \left]0, 1\right]$.

The definition formally explicates the imprecise idea of whether a hypothesis is compatible with the data given any restrictions on the basis of the compatibility function to be defined in Section 2.3. As will be seen in Section 3.2, $c(\theta_0; x|\Theta) = p(\theta_0; x)$ often holds when there are no restrictions on θ.

2.3 Degrees of data-hypothesis compatibility

The next definition applies the α-compatible concept to composite hypotheses as well as simple hypotheses. Just as a p-value can be defined in terms of whether the null hypothesis is rejected at a fixed significance level α, the degree of compatibility with data is defined in terms of whether the null hypothesis is α-compatible with the data at a fixed value of α.

Definition 2. The functions $C(\bullet; \bullet|\bullet) : \mathfrak{H} \times \mathcal{X} \times \mathfrak{H} \to [0, 1]$ and $C(\bullet; \bullet) = C(\bullet; \bullet|\Theta) : \mathfrak{H} \times \mathcal{X} \to [0, 1]$ are *compatibility set functions*, and $C(\mathcal{H}_0; x|\mathcal{R})$ is the *compatibility* of the hypothesis that $\theta \in \mathcal{H}_0$ with the observation that $X = x$ conditional on the restriction that $\theta \in \mathcal{R}$ if these conditions hold for all $x \in \mathcal{X}$, $\mathcal{H}_0 \in \mathfrak{H}$, and $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$:

- *Axiom of minimal compatibility.* If $\mathcal{H}_0 \cap \mathcal{R} = \emptyset$, then $C(\mathcal{H}_0; x|\mathcal{R}) = 0$.
- *Axiom of maximal compatibility.* $C(\Theta; x|\Theta) = 1$.
- *Axiom of conditional compatibility.* If $\mathcal{H}_0 \cap \mathcal{R} \neq \emptyset$, then $C(\mathcal{H}_0; x|\mathcal{R}) = K(C(\mathcal{H}_0 \cap \mathcal{R}; x), C(\mathcal{R}; x))$ for some function $K(\bullet, \bullet) : [0, 1]^2 \to [0, 1]$.
- *Axiom of conditional probability.* For any $\theta_0 \in \Theta$ and $\alpha \in [p(\theta_0; x), 1]$,

$$C(\{\theta_0\}; x|\mathcal{R}(\alpha; x)) = Q_{\theta_0,\gamma}(p(\theta_0; X) < p(\theta_0; x), \epsilon|\theta_0 \in \mathcal{R}(\alpha; X)),$$

where $\mathcal{R}(\alpha; x_0) = \{\theta_1 \in \Theta : p(\theta_1; x) \leq \alpha\}$ for all $x_0 \in \mathcal{X}$ and $Q_{\theta_0,\gamma}$ is a joint distribution of X and an auxiliary event ϵ such that $P_{\theta_0,\gamma}$ is a version of the conditional distribution $Q_{\theta_0,\gamma}(\bullet|\epsilon)$ and such that there is a constant $\kappa > 0$ satisfying $\kappa = Q_{\theta_0,\gamma}(\epsilon)$ for all $\theta_0 \in \Theta$ and $\gamma \in \Gamma$.

- *Axiom of compatible hypotheses.* With "$\mathcal{H}_0 \overset{\alpha|\mathcal{R}}{\sim} x$" denoting the hypothesis that "$\theta \in \mathcal{H}_0$ is α-compatible with the observation that $X = x$, conditional on the restriction that $\theta \in \mathcal{R}$,"

$$C(\mathcal{H}_0; x|\mathcal{R}) = \sup\left\{\alpha \in]0,1] : \mathcal{H}_0 \overset{\alpha|\mathcal{R}}{\sim} x\right\}. \tag{4}$$

The functions $c(\bullet;\bullet|\bullet) : \Theta \times \mathcal{X} \times \mathfrak{H} \to [0,1]$ and $c(\bullet;\bullet) = c(\bullet;\bullet|\Theta) : \Theta \times \mathcal{X} \to [0,1]$ are *compatibility point functions* if $c(\theta_0; x|\mathcal{R}) = C(\{\theta_0\}; x|\mathcal{R})$ for all $\theta_0 \in \Theta$, $x \in \mathcal{X}$, and $\mathcal{R} \in \mathfrak{H} \setminus \{\emptyset\}$.

The compatibility $C(\mathcal{H}_0; x|\mathcal{R})$ is the degree to which the hypothesis that $\theta \in \mathcal{H}_0$ is compatible with x under the restriction that $\theta \in \mathcal{R}$. This definition ties the loose end in Definition 1 by connecting the compatibility functions to the p-value function.

The absence of a restriction is represented by $\mathcal{R} = \Theta$. Since the degenerate restriction that $\theta \in \Theta$ is necessarily true, $C(\mathcal{H}_0; x|\Theta)$ and $c(\theta_0; x|\Theta)$ are marginal compatibilities. They are abbreviated by $C(\mathcal{H}_0; x)$ and $c(\theta_0; x)$ and called a C*-value* and c*-value*, respectively. Each is the marginal degree to which its hypothesis is compatible with x.

3 Properties of data-hypothesis compatibility

3.1 Relations between concepts

The following lemma connects the concepts of a compatible hypothesis and a compatibility set.

Lemma 1. *For any $\mathcal{H}_0 \in \mathfrak{H}$, $\mathcal{R} \in \mathfrak{H} \setminus \{\emptyset\}$, $x \in \mathcal{X}$, and $\alpha \in]0,1]$, the hypothesis that $\theta \in \mathcal{H}_0$ is α-compatible with the observation that $X = x$, conditional on the restriction that $\theta \in \mathcal{R}$, if and only if $\mathcal{H}_0 \cap \mathcal{H}(\alpha; x|\mathcal{R}) \neq \emptyset$, where $\mathcal{H}(\alpha; x|\mathcal{R})$ is the α-compatibility set given $X = x$ and $\theta \in \mathcal{R}$.*

Proof. By definition, the hypothesis is α-compatible if and only if $\emptyset \neq \{\theta_0 \in \mathcal{H}_0 : c(\theta_0; x|\mathcal{R}) \geq \alpha\} = \mathcal{H}_0 \cap \{\theta_0 \in \Theta : c(\theta_0; x|\mathcal{R}) \geq \alpha\}$. □

The compatibility of a hypothesis is now seem to be proportional to the p-value.

Lemma 2. *For any $\theta_0 \in \Theta$ and $x \in \mathcal{X}$, the marginal compatibility of the hypothesis that $\theta = \theta_0$ with the observation that $X = x$ is*

$$c(\theta_0; x) = \kappa p(\theta_0; x) \tag{5}$$

for some $\kappa \in]0,1]$. The corresponding conditional compatibility given $p(\theta; x) \leq \alpha$ is

$$c(\theta_0; x|\mathcal{R}(\alpha; x)) = \kappa p(\theta_0; x) / \alpha \tag{6}$$

for any $\alpha \in]0,1]$.

Proof. From the axiom of conditional probability, $\kappa = Q_{\theta_0,\gamma}(\epsilon) > 0$ for all $\theta_0 \in \Theta$ and $\gamma \in \Gamma$. The axiom of conditional probability and the usual definition of conditional probability give

$$\begin{aligned} C\left(\{\theta_0\}; x|\mathcal{R}(\alpha; x)\right) &= Q_{\theta_0,\gamma}(\epsilon)\, Q_{\theta_0,\gamma}\left(p(\theta_0; X) < p(\theta_0; x)\,|\theta_0 \in \mathcal{R}(\alpha; X), \epsilon\right) \\ &= \kappa \frac{P_{\theta_0,\gamma}\left(p(\theta_0; X) < p(\theta_0; x), \theta_0 \in \mathcal{R}(\alpha; X), \epsilon\right)}{P_{\theta_0,\gamma}\left(\theta_0 \in \mathcal{R}(\alpha; X)\right)} \qquad (7) \\ &= \kappa \frac{P_{\theta_0,\gamma}\left(p(\theta_0; X) < p(\theta_0; x), p(\theta_0; X) \leq \alpha\right)}{P_{\theta_0,\gamma}\left(p(\theta_0; X) \leq \alpha\right)} \\ &= \kappa \frac{P_{\theta_0,\gamma}\left(p(\theta_0; X) < p(\theta_0; x)\right)}{P_{\theta_0,\gamma}\left(p(\theta_0; X) \leq \alpha\right)} = \kappa \frac{p(\theta_0; x)}{\alpha} \qquad (8) \end{aligned}$$

for all $\alpha \in [p(\theta_0; x), 1]$, with the last step a consequence of equation (1). Since $\Theta = \mathcal{R}(1)$, it follows that $C(\{\theta_0\}; x) = C(\{\theta_0\}; x|\mathcal{R}(1)) = \kappa p(\theta_0; x)$. □

The next result indicates a sense in which conditional compatibility is to joint and marginal compatibility what conditional probability is to joint and marginal probability.

Lemma 3. *Given some $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$, the compatibility $C(\mathcal{H}_0; x|\mathcal{R})$ of the hypothesis that $\theta \in \mathcal{H}_0$ with the observation that $X = x$ conditional on the restriction that $\theta \in \mathcal{R}$ satisfies*

$$C(\mathcal{H}_0; x|\mathcal{R}) = \frac{C(\mathcal{H}_0 \cap \mathcal{R}; x)}{C(\mathcal{R}; x)} \qquad (9)$$

for all $x \in \mathcal{X}$ and $\mathcal{H}_0 \in \mathfrak{H}$, where $C(\bullet; \bullet|\bullet)$ is a compatibility set function on $\mathfrak{H} \times \mathcal{X} \times \mathfrak{H}$ and $C(\bullet; \bullet) = C(\bullet; \bullet|\Theta)$.

Proof. Equation 9 is immediately seen to be true by the axiom of minimal compatibility in the case that $\mathcal{H}_0 \cap \mathcal{R} = \emptyset$. The remainder of the proof addresses the other case, $\mathcal{H}_0 \cap \mathcal{R} \neq \emptyset$. According to Lemma 2 and $\{\theta_0\} \cap \mathcal{R}(\alpha; x) = \{\theta_0\}$, the axiom of conditional compatibility requires that

$$\begin{aligned} C\left(\{\theta_0\};x|\mathcal{R}\left(\alpha;x\right)\right) &= K\left(C\left(\{\theta_0\}\cap\mathcal{R}\left(\alpha;x\right);x\right),C\left(\mathcal{R}\left(\alpha;x\right);x\right)\right) \\ &= K\left(\kappa p\left(\theta_0;x\right),C\left(\mathcal{R}\left(\alpha;x\right);x\right)\right). \end{aligned}$$

Thus, equation (6) implies that $C\left(\mathcal{R}\left(\alpha;x\right);x\right)=\alpha$ and that

$$K\left(q_1,q_2\right)=\frac{q_1}{q_2}$$

for all $q_1, q_2 \in [0,1]$ satisfying $q_1 \leq q_2$. Equation (9) then follows from the axiom of conditional compatibility. □

3.2 Deriving data-hypothesis compatibility

The compatibility is easily derived from the p-value function using the simple equations of the next two results.

Theorem 1. *The compatibility of the hypothesis that $\theta \in \mathcal{H}_0$ with the observation that $X = x$ conditional on the restriction that $\theta \in \mathcal{R}$ is*

$$C\left(\mathcal{H}_0;x|\mathcal{R}\right)=\begin{cases}0 & \text{if } \mathcal{H}_0\cap\mathcal{R}=\emptyset \\ \frac{\sup_{\theta_0\in\mathcal{H}_0\cap\mathcal{R}}p(\theta_0;x)}{\sup_{\theta_1\in\mathcal{R}}p(\theta_1;x)} & \text{if } \mathcal{H}_0\cap\mathcal{R}\neq\emptyset\end{cases}$$

for all $x \in \mathcal{X}$, $\mathcal{H}_0 \in \mathfrak{H}$, and $\mathcal{R} \in \mathfrak{H}\backslash\{\emptyset\}$.

Proof. In the case that $\mathcal{H}_0\cap\mathcal{R}=\emptyset$, the axiom of minimal compatibility gives $C\left(\mathcal{H}_0;x|\mathcal{R}\right)=$

0. In the $\mathcal{H}_0 \cap \mathcal{R} \neq \emptyset$ case, Definition 1, equation (4), and Lemma 3 yield

$$\begin{aligned} C\left(\mathcal{H}_0; x|\mathcal{R}\right) &= \sup\left\{\alpha \in \left]0,1\right] : \theta_0 \in \mathcal{H}_0, c\left(\theta_0; x|\mathcal{R}\right) \geq \alpha\right\} && (10) \\ &= \sup\left\{\alpha \in \left]0,1\right] : \theta_0 \in \mathcal{H}_0 \cap \mathcal{R}, c\left(\theta_0; x\right) \geq \alpha C\left(\mathcal{R}; x\right)\right\} && (11) \\ &= \sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} c\left(\theta_0; x\right) / C\left(\mathcal{R}; x\right) && (12) \end{aligned}$$

Since $C\left(\mathcal{R}; x\right) = C\left(\mathcal{R}; x|\Theta\right)$, equation (12) entails that $C\left(\mathcal{R}; x\right) = \sup_{\theta_0 \in \mathcal{R}} c\left(\theta_0; x\right) / C\left(\Theta; x\right)$. By the axiom of maximal compatibility, $C\left(\mathcal{R}; x\right) = \sup_{\theta_0 \in \mathcal{R}} c\left(\theta_0; x\right)$. Thus, with Lemma 2, equation (12) reduces to $C\left(\mathcal{H}_0; x|\mathcal{R}\right) = \sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} p\left(\theta_0; x\right) / \sup_{\theta_1 \in \mathcal{R}} p\left(\theta_1; x\right)$. □

Corollary 1. *For any $\theta_0 \in \Theta$, $x \in \mathcal{X}$, and $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$, the compatibility of the hypothesis that $\theta = \theta_0$ with the observation that $X = x$ conditional on the restriction that $\theta \in \mathcal{R}$ is*

$$c\left(\theta_0; x|\mathcal{R}\right) = \begin{cases} 0 & \text{if } \theta_0 \notin \mathcal{R} \\ \frac{p(\theta_0; x)}{\sup_{\theta_1 \in \mathcal{R}} p(\theta_1; x)} & \text{if } \theta_0 \in \mathcal{R}. \end{cases}$$

Proof. By Definition 2, $c\left(\theta_0; x|\mathcal{R}\right) = C\left(\{\theta_0\}; x|\mathcal{R}\right)$ for all $\theta_0 \in \Theta$. The desired result follows from Theorem 1. □

In the usual setting of testing the simple hypothesis that $\theta = \theta_0$, the parameter is relatively unrestricted, and the compatibility is the p-value. That is formally stated as this direct result of Theorem 1 and Corollary 1.

Corollary 2. *For any $x \in \mathcal{X}$, $\theta_0 \in \mathcal{R}$, and $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$ such that $\sup_{\theta_1 \in \mathcal{R}} p\left(\theta_1; x\right) = 1$, the compatibility of the hypothesis that $\theta = \theta_0$ with the observation that $X = x$, conditional on the restriction that $\theta \in \mathcal{R}$, is $c\left(\theta_0; x|\mathcal{R}\right) = p\left(\theta_0; x\right)$. Under the same conditions, the compatibility of the hypothesis that $\theta \in \mathcal{H}_0$ with the observation that $X = x$ conditional on* *the restriction that* $\theta \in \mathcal{R}$ *is*

$$C(\mathcal{H}_0; x|\mathcal{R}) = \sup_{\theta_0 \in \mathcal{H}_0} p(\theta_0; x) \tag{13}$$

for all $x \in \mathcal{X}$ *and* $\mathcal{H}_0 \in \mathfrak{H}$ *such that* $\mathcal{H}_0 \subseteq \mathcal{R}$.

Corollary 2 justifies the practice of maximizing a p-value over all the parameter values of a composite null hypothesis (e.g., Wendell and Schmee, 1996; Silvapulle and Sen, 2011, p. 33; Patriota, 2013).

The next corollary highlights ways conditional compatibility is similar to and different from conditional probability.

Corollary 3. *Given some* $x \in \mathcal{X}$, $\mathcal{H}_0 \in \mathfrak{H}$, *and* $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$, *the compatibility* $C(\mathcal{H}_0; x|\mathcal{R})$ *of the hypothesis that* $\theta \in \mathcal{H}_0$ *with the observation that* $X = x$ *conditional on the restriction that* $\theta \in \mathcal{R}$ *satisfies* $C(\mathcal{H}_0; x|\mathcal{R}) = 1$ *if and only if* $\mathcal{H}_0 \cap \mathcal{R} \neq \emptyset$ *and*

$$\sup_{\theta_1 \in \mathcal{R}} p(\theta_1; x) = \sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} p(\theta_0; x). \tag{14}$$

Proof. In the $\mathcal{H}_0 \cap \mathcal{R} = \emptyset$ case, Theorem 1 gives $C(\mathcal{H}_0; x|\mathcal{R}) = 0 \neq 1$. On the other hand, in the case that $\mathcal{H}_0 \cap \mathcal{R} \neq \emptyset$, Theorem 1 implies that equation (14) holds if and only if $C(\mathcal{H}_0; x|\mathcal{R}) = 1$.

□

3.3 Conservative error rate control and coverage

This theorem demonstrates that compatibility controls the Type I error rate and that α-compatibility sets are $(1-\alpha)(100\%)$-confidence sets that are valid in that their coverage rates are conservative if not exact.

Theorem 2. *For every* $x \in \mathcal{X}$, $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$, *and* $\theta_0 \in \mathcal{R}$, *let* $p(\theta_0; x)$ *denote the* p*-value testing* $\theta = \theta_0$ *as the null hypothesis, and let* $c(\theta_0; x|\mathcal{R})$ *denote the compatibility of the hypothesis that* $\theta = \theta_0$ *with the observation that* $X = x$ *conditional on the restriction that* $\theta \in \mathcal{R}$, *let* $\mathrm{CS}(\alpha; x|\mathcal{R})$ *denote the exact confidence set given by equation* (2), *and let* $\mathcal{H}(\alpha; x|\mathcal{R})$ *denote the* α*-compatibility set given* $X = x$ *and* $\theta \in \mathcal{R}$ *for every* $\alpha \in]0, 1]$. *For any* $\gamma \in \Gamma$, *it follows that* $c(\theta_0; x|\mathcal{R}) \geq p(\theta_0; x)$, $\mathrm{CS}(\alpha; x|\mathcal{R}) \subseteq \mathcal{H}(\alpha; x|\mathcal{R})$, *and*

$$\mathcal{H}(\alpha; x|\mathcal{R}) = \mathrm{CS}\left(\alpha \sup_{\theta_1 \in \mathcal{R}} p(\theta_1; x); x|\mathcal{R}\right) \tag{15}$$

$$P_{\theta_0,\gamma}\left(c(\theta_0; X|\mathcal{R}) < \alpha\right) \leq \alpha \tag{16}$$

$$P_{\theta_0,\gamma}\left(\theta_0 \in \mathcal{H}(\alpha; X|\mathcal{R})\right) \geq 1 - \alpha, \tag{17}$$

with the formulas holding with exact equality if $\sup_{\theta_1 \in \mathcal{R}} p(\theta_1; x) = 1$.

Proof. Since $\theta_0 \in \mathcal{R}$, Corollary 1 entails that $c(\theta_0; x|\mathcal{R}) \geq p(\theta_0; x)$ for all $x \in \mathcal{X}$, from which $P_{\theta_0,\gamma}\left(c(\theta_0; X|\mathcal{R}) \geq p(\theta_0; X)\right) = 1$ follows, providing

$$P_{\theta_0,\gamma}\left(c(\theta_0; X|\mathcal{R}) < \alpha\right) \leq P_{\theta_0,\gamma}\left(p(\theta_0; X) < \alpha\right), \tag{18}$$

and equation (1) yields formula (16). Applying inequality (18) to equation (3),

$$\mathrm{CS}(\alpha; x|\mathcal{R}) = \{\theta_0 \in \Theta : p(\theta_0; x) \geq \alpha\} \subseteq \mathcal{H}(\alpha; x|\mathcal{R}) \tag{19}$$

for every $x \in \mathcal{X}$. Hence, by equation (1),

$$P_{\theta_0,\gamma}\left(p\left(\theta_0; X\right) \geq \alpha\right) = 1 - \alpha \leq P_{\theta_0,\gamma}\left(\theta_0 \in \mathcal{H}\left(\alpha; X|\mathcal{R}\right)\right),$$

proving formula (17). Corollary 1 and equation (19) imply that $\mathrm{CS}\left(\alpha \sup_{\theta_1 \in \mathcal{R}} p\left(\theta_1; x\right); x|\mathcal{R}\right) = \left\{\theta_0 \in \Theta : c\left(\theta_0; x|\mathcal{R}\right) \geq \alpha\right\}$ and thus that equation (15) holds. Finally, if $\sup_{\theta_1 \in \mathcal{R}} p\left(\theta_1; x\right) = 1$ for all $x \in \mathcal{X}$, then Lemma 1 requires that $c\left(\theta_0; x\right) = p\left(\theta_0; x\right)$ for all $\theta_0 \in \mathcal{R}$ and that $\mathrm{CS}\left(\alpha; x|\mathcal{R}\right) = \mathcal{H}\left(\alpha; x|\mathcal{R}\right)$. Consequently, $P_{\theta_0,\gamma}\left(c\left(\theta_0; X|\mathcal{R}\right) < \alpha\right) = \alpha$ and $P_{\theta_0,\gamma}\left(\theta_0 \in \mathcal{H}\left(\alpha; X|\mathcal{R}\right)\right) = 1 - \alpha$ follow from equation (1). □

4 Examples

The first example compares a simple null hypothesis to a simple alternative hypothesis (cf. Berger, 2003; Wang, 2004) to demonstrate the use of the proposed framework as simply as possible.

Example 1. Comparison of two simple hypotheses, $X \sim \mathrm{N}(0,1)$ and $X \sim \mathrm{N}(1,1)$, on the basis of a single observation x. In this example, $\mathcal{R} = \{0,1\} \subseteq \Theta \subseteq \mathbb{R}$, $P_{\theta_0} = \mathrm{N}(0,1)$ for $\theta_0 \in \{0,1\}$, and the two null hypotheses may be restated as $\theta = 0$ and $\theta = 1$. Figure 1 displays the following "significance values" of the hypothesis that $\theta = 1$:

1. The two-sided p-value $p(1; x) = 2\left(\Phi(x-1) \vee \left(1 - \Phi(x-1)\right)\right)$, where $\vee$ is the maximum and Φ is the standard normal distribution function. This does not depend on the hypothesis that $\theta = 0$.

2. The corresponding compatibility of the hypothesis that $\theta = 1$ with x conditional on

$\theta \in \{0,1\}$. According to Corollary 1, that compatibility is

$$c(1;x|\{0,1\}) = \begin{cases} \frac{p(1;x)}{p(0;x)} & \text{if } p(1;x) < p(0;x) \\ 1 & \text{if } p(1;x) \geq p(0;x) \end{cases},$$

where $p(0;x) = 2(\Phi(x) \vee (1 - \Phi(x)))$ is the p-value of the hypothesis that $\theta = 0$.

3. The posterior probability that $\theta = 1$ on the basis of 50% prior probability of each of the null hypotheses conditional on $\theta \in \{0,1\}$.

From Figure 1, it can be seen that, given any significance level $\alpha \in]0,1[$, the p-value would erroneously lead to the rejection of the better-supported null hypothesis for sufficiently large $x > 1$ but that the other two quantities take the other null hypothesis into account. Further, for all $x > 1/2$, there is not any $\alpha \in]0,1[$ such that the compatibility conditional on $\theta \in \{0,1\}$ is less than α, with the result that it impossible to reject the better-supported null hypothesis, regardless of how high the significance level is. The posterior probability does not share that feature: being strictly less than 1, it is less than sufficiently high values of α. In agreement with $c(1;x|\{0,1\})$, Chuaqui (1991, p. 97) recommended the ratio of p-values for comparing two hypotheses on the basis of the same observation. ▲

The next example is an idealized version of restricted parameter problem encountered, for example, in physics (§1).

Example 2. Bounded parameter. Fraser (2011) considered a $\mathrm{N}(\theta,1)$ observable variable $X \sim P_\theta = \mathrm{N}(\theta,1)$ with observed value x and the parameter restriction $\theta \geq 0$, and the left-tailed version of the two-tailed p-value

$$p(\theta_0;x) = 2(\Phi(x-\theta_0) \vee (1 - \Phi(x-\theta_0)))$$

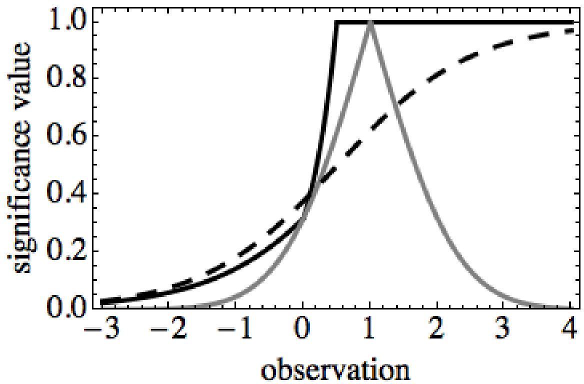

Figure 1: The p-value $p\left(1;x\right)$ in solid gray, the data compatibility $c\left(1;x|\left\{0,1\right\}\right)$ in solid black, and the posterior probability that $\theta=1$ in dashed black as functions of x, the value of the normal observation.

for every $\theta_0 \geq 0$. Thus, if $x \geq 0$, then $p\left(\theta_0;x\right)=1$ holds for a value of $\theta_0 \geq 0$, namely, $\theta_0 = x$. In that case, Corollary 2 applies, and $c\left(\theta_0;x|\left[0,\infty\right[\right)=p\left(\theta_0;x\right)$ for all $\theta_0 \geq 0$. On the other hand, if $x<0$, then Corollary 1 instead gives $c\left(\theta_0;x|\left[0,\infty\right[\right)=p\left(\theta_0;x\right)/\sup_{\theta_1\geq 0}p\left(\theta_1;x\right)$ for all $\theta_0 \geq 0$. This relationship between the compatibility and the p-value is seen in Figure 2 for the observation $x=-1$. The exact $\left(1-\alpha\right)\left(100\%\right)$-confidence interval

$$\mathrm{CI}\left(\alpha;x|\left[0,\infty\right[\right)=\left\{\theta_0\geq 0:p\left(\theta_0;x\right)\geq\alpha\right\}=\left[0\vee\left(x+\Phi^{-1}\left(\alpha/2\right)\right),x+\Phi^{-1}\left(1-\alpha/2\right)\right],$$

with Φ^{-1} denoting the quantile function. By contrast, equation (3) and Theorem 1 give the α-compatibility interval

$$\begin{aligned}\mathcal{H}\left(\alpha;x|\left[0,\infty\right[\right) &= \left\{\theta_0\in\Theta:p\left(\theta_0;x\right)\geq\alpha\sup_{\theta_1\in\mathcal{R}}p\left(\theta_1;x\right)\right\}\\ &= \left[0\vee\left(x+\Phi^{-1}\left(\frac{\alpha p^+\left(x\right)}{2}\right)\right),0\vee\left(x+\Phi^{-1}\left(1-\frac{\alpha p^+\left(x\right)}{2}\right)\right)\right],\end{aligned}$$

where $p^+\left(x\right)=\sup_{\theta_1\geq 0}p\left(\theta_1;x\right)$. As required by Theorem 2, $\mathcal{H}\left(\alpha;x|\left[0,\infty\right[\right)=\mathrm{CI}\left(\alpha p^+\left(x\right);x|\left[0,\infty\right[\right)$.

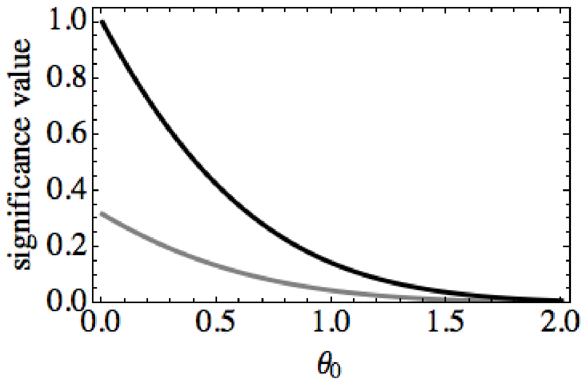

Figure 2: The p-value $p\,(\theta_0; -1)$ in gray and the data compatibility $c\,(\theta_0; -1|\,[0,\infty[)$ in black as functions of θ_0, the parameter value.

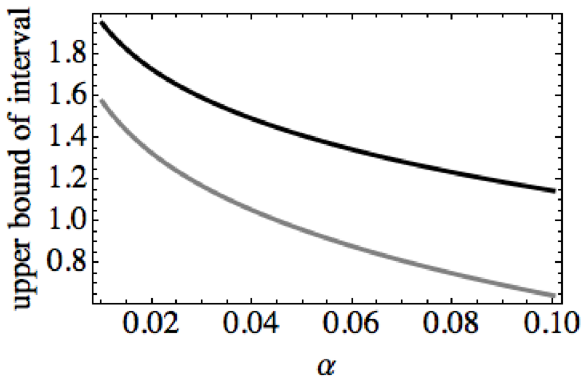

Figure 3: The upper bounds of the α-confidence interval $\mathrm{CI}\,(\alpha; -1|\,[0,\infty[)$ in gray and of the α-compatibility interval $\mathcal{H}\,(\alpha; -1|\,[0,\infty[)$ in black as functions of α, the threshold applied to the curves of Figure 2.

For the observation $x = -1$, the confidence intervals are compared to their compatibility counterparts in Figure 3. ▲

5 Hypothesis acceptance, rejection, or neither

5.1 Warrant for accepting a hypothesis

While the compatibility of a hypothesis with data does not warrant accepting the hypothesis, a lack of compatibility justifies rejecting it and accepting its negation. That idea leads to the following measure of the degree of warrant for accepting a hypothesis.

Definition 3. A function $W(\bullet;\bullet|\bullet) : \mathfrak{H} \times \mathcal{X} \times \mathfrak{H} \rightarrow [0,1]$, called the *warrant set function*, is defined as follows. For all $x \in \mathcal{X}$, $\mathcal{H}_0 \in \mathfrak{H}$, and $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$,

$$W(\mathcal{H}_0; x|\mathcal{R}) = 1 - C(\mathcal{R} \backslash \mathcal{H}_0; x|\mathcal{R})$$

is the *warrant* of the hypothesis that $\theta \in \mathcal{H}_0$ given the observation that $X = x$ conditional on the restriction that $\theta \in \mathcal{R}$, where $C(\mathcal{R} \backslash \mathcal{H}_0; x|\mathcal{R})$ is the compatibility of the hypothesis that $\theta \in \mathcal{R}$ but $\theta \notin \mathcal{H}_0$ with the observation that $X = x$ conditional on the restriction that $\theta \in \mathcal{R}$.

The warrant for a hypothesis corresponding to a set estimate $\mathcal{H}(\alpha; x|\mathcal{R})$ is important as a lower bound on the coverage rate of the set estimator $\mathcal{H}(\alpha; X|\mathcal{R})$, as formally stated in the next theorem.

Theorem 3. *Let $x \in \mathcal{X}$, $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$, and $\theta_0 \in \mathcal{R}$, and let W denote a warrant function corresponding to $\mathcal{H}(\alpha; x|\mathcal{R})$, the α-compatibility set given $X = x$ for every $\alpha \in]0,1]$. For any $\alpha \in]0,1]$ and $\gamma \in \Gamma$,*

$$P_{\theta_0,\gamma}(\theta_0 \in \mathcal{H}(\alpha; X|\mathcal{R})) \geq W(\mathcal{H}(\alpha; x|\mathcal{R}); x|\mathcal{R}), \tag{20}$$

which holds with exact equality if $\sup_{\theta_1 \in \mathcal{R}} p(\theta_1; x) = 1$*, where* $p(\theta_0; x)$ *is the* p*-value testing* $\theta = \theta_0$ *as the null hypothesis for all* $\theta_0 \in \mathcal{R}$*.*

Proof. According to the definitions of warrant and the α-compatibility set,

$$\begin{aligned} W(\mathcal{H}(\alpha; x|\mathcal{R}); x|\mathcal{R}) &= 1 - C(\mathcal{R} \backslash \mathcal{H}(\alpha; x|\mathcal{R}); x|\mathcal{R}) \\ &= 1 - C(\{\theta_0 \in \Theta : c(\theta_0; x|\mathcal{R}) < \alpha\}; x|\mathcal{R}), \end{aligned}$$

Thus, since that C is the relevant compatibility set function,

$$W(\mathcal{H}(\alpha; x|\mathcal{R}); x|\mathcal{R}) = 1 - \sup\{c(\theta_0; x|\mathcal{R}) < \alpha : \theta_0 \in \Theta\} = 1 - \alpha, \quad (21)$$

Formula (20) then results from Theorem 2. The same theorem says $\sup_{\theta_1 \in \mathcal{R}} p(\theta_1; x) = 1$ implies that $P_{\theta_0,\gamma}(\theta_0 \in \mathcal{H}(\alpha; X|\mathcal{R})) = 1 - \alpha$, leading to $P_{\theta_0,\gamma}(\theta_0 \in \mathcal{H}(\alpha; X|\mathcal{R})) = W(\mathcal{H}(\alpha; x|\mathcal{R}); x|\mathcal{R})$ via equation (21). □

Equation (21) interprets the nominal confidence level $1 - \alpha$ as the degree of warrant for the hypothesis that the observed confidence set $\mathcal{H}(\alpha; x|\mathcal{R})$ contains the true value of the parameter.

5.2 Acceptability of a hypothesis

The information in the data-compatibility and warrant of a hypothesis will be combined into a single measure of acceptability in this section. Hypotheses of sufficiently high acceptability are accepted, those with sufficiently negative acceptability are rejected, and the remaining hypotheses are neither accepted nor rejected.

For any $x \in \mathcal{X}$, $\mathcal{H}_0 \in \mathfrak{H}$, and $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$, let $C(\mathcal{H}_0; x|\mathcal{R})$ denote the compatibility of the hypothesis that $\theta \in \mathcal{H}_0$ with the observation that $X = x$ conditional on the restriction that $\theta \in \mathcal{R}$, where $\mathfrak{H}$ is some σ-field of subsets of Θ. Let $\mathcal{H}_0'$ denote the complement of $\mathcal{H}_0$, that is, $\mathcal{H}_0' = \Theta \backslash \mathcal{H}_0$.

Definition 4. The *acceptability* of the hypothesis that $\theta \in \mathcal{H}_0$ given the observation that $X = x$ and the restriction that $\theta \in \mathcal{R}$ is the extended real number $A(\mathcal{H}_0; x|\mathcal{R}) \in \{-\infty, \infty\} \cup \mathbb{R}$ such that, for all $\alpha \in]0, 1]$,

$$(\theta_1 \in \mathcal{H}(\alpha; x|\mathcal{R}) \implies \theta_1 \in \mathcal{H}_0) \iff A(\mathcal{H}_0; x|\mathcal{R}) > \log {}^1\!/\!_\alpha \tag{22}$$

$$(\theta_2 \in \mathcal{H}(\alpha; x|\mathcal{R}) \implies \theta_2 \in \mathcal{H}_0') \iff A(\mathcal{H}_0; x|\mathcal{R}) < -\log {}^1\!/\!_\alpha \tag{23}$$

$$\exists \theta_1, \theta_2 \in \mathcal{H}(\alpha; x|\mathcal{R}); \, \theta_1 \in \mathcal{H}_0; \, \theta_2 \in \mathcal{H}_0' \iff |A(\mathcal{H}_0; x|\mathcal{R})| \leq \log {}^1\!/\!_\alpha, \tag{24}$$

where $\mathcal{H}(\alpha; x|\mathcal{R})$ is the α-compatibility set given $X = x$ and $\theta \in \mathcal{R}$. Here, the base of log might be 2 for best interpretability but can be any number greater than 1. At level α, the hypothesis that $\theta \in \mathcal{H}_0$, given the observation that $X = x$ and the restriction that $\theta \in \mathcal{R}$, is *accepted* if and only if $A(\mathcal{H}_0; x|\mathcal{R}) > \log {}^1\!/\!_\alpha$ and is *rejected* if and only if $A(\mathcal{H}_0; x|\mathcal{R}) < -\log {}^1\!/\!_\alpha$. In the absence of a restriction ($\mathcal{R} = \Theta$), the acceptability $A(\mathcal{H}_0; x|\Theta)$ is abbreviated as $A(\mathcal{H}_0; x)$.

In that way, the acceptability of a general hypothesis over its alternative hypothesis is defined in terms of which values of the parameter of interest are compatible with the observed data and with the given restrictions according to Section 2.2. Formula (22) says a hypothesis is accepted at level α if it is consistent with *all* of the α-compatible parameter values. Likewise, formula (23) says a hypothesis is rejected at level α if it is not consistent with *any* of the α-compatible parameter values. Finally, formula (24) means there is insufficient evidence to accept or reject the hypothesis at level α if it is consistent with some but not all of the α-compatible parameter values.

The possibility of the last case means there is no arbitrary requirement that every hypothesis be either rejected or accepted. At the same time, the rejection of a null hypothesis for lack of compatibility with other information necessarily implies acceptance of an alternative hypothesis, as this lemma makes clear.

Lemma 4. *These propositions are equivalent for any* $\mathcal{H}_0 \in \mathfrak{H}$, $\mathcal{R} \in \mathfrak{H}\backslash\{\emptyset\}$, $x \in \mathcal{X}$, *and* $\alpha \in]0,1]$*:*

1. $A\left(\mathcal{H}_0; x|\mathcal{R}\right) < -\log {}^1/_\alpha$.

2. *The hypothesis that* $\theta \in \mathcal{H}_0$*, given the observation that* $X = x$ *and the restriction that* $\theta \in \mathcal{R}$*, is rejected at level* α*.*

3. *The same hypothesis is not* α*-compatible with the observation that* $X = x$*, conditional on the restriction that* $\theta \in \mathcal{R}$*.*

4. $A\left(\mathcal{H}_0'; x|\mathcal{R}\right) > \log {}^1/_\alpha$.

5. *The hypothesis that* $\theta \in \mathcal{H}_0'$*, given the observation that* $X = x$ *and the restriction that* $\theta \in \mathcal{R}$*, is accepted at level* α*.*

Proof. Propositions 1 and 2 are equivalent by Definition 4: the hypothesis that $\theta \in \mathcal{H}_0$ is rejected if and only if $A\left(\mathcal{H}_0; x|\mathcal{R}\right) < -\log {}^1/_\alpha$. Similarly, Propositions 4 and 5 are equivalent. According to formula (23), Proposition 1 is equivalent to

$$\theta_2 \in \mathcal{H}\left(\alpha; x|\mathcal{R}\right) \implies \theta_2 \in \mathcal{H}_0', \tag{25}$$

which, by formula (22), holds if and only if $A\left(\mathcal{H}_0'; x|\mathcal{R}\right) > \log {}^1/_\alpha$, the definition of accepting the hypothesis that $\theta \in \mathcal{H}_0'$. That establishes the equivalence of Propositions 1 and 4. Lemma 1 entails that Proposition 3 is equivalent to $\mathcal{H}_0 \cap \mathcal{H}\left(\alpha; x|\mathcal{R}\right) = \emptyset$, and that equivalence makes the same assertion as formula (25). Therefore, Propositions 2 and 3 are equivalent. □

Thus, whereas the fact that a hypothesis is data-compatible is merely necessary for its acceptance, the fact that its denial is incompatible is sufficient. Calculating the acceptability is facilitated by the next theorem.

Theorem 4. *For any $x \in \mathcal{X}$, $\mathcal{H}_0 \in \mathfrak{H}$, and $\mathcal{R} \in \mathfrak{H}\backslash\left\{\emptyset\right\}$, the acceptability of the hypothesis that $\theta \in \mathcal{H}_0$, given the observation that $X = x$ and the restriction that $\theta \in \mathcal{R}$, is*

$$A\left(\mathcal{H}_0; x|\mathcal{R}\right) = \log \frac{C\left(\mathcal{H}_0; x|\mathcal{R}\right)}{C\left(\mathcal{H}_0'; x|\mathcal{R}\right)}; \tag{26}$$

$A\left(\mathcal{H}_0; x|\mathcal{R}\right) = -\infty$ *if* $\mathcal{H}_0 \cap \mathcal{R} = \emptyset$, $A\left(\mathcal{H}_0; x|\mathcal{R}\right) = \infty$ *if* $\mathcal{H}_0' \cap \mathcal{R} = \emptyset$, *or*

$$A\left(\mathcal{H}_0; x|\mathcal{R}\right) = \log \frac{\sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} p\left(\theta_0; x\right)}{\sup_{\theta_0 \in \mathcal{H}_0' \cap \mathcal{R}} p\left(\theta_0; x\right)} \tag{27}$$

if $\mathcal{H}_0 \cap \mathcal{R} \neq \emptyset$ *and* $\mathcal{H}_0' \cap \mathcal{R} \neq \emptyset$.

Proof. For any $\mathcal{H}_0 \in \mathfrak{H}$, let

$$\widetilde{A}\left(\mathcal{H}_0\right) = \log \frac{C\left(\mathcal{H}_0; x|\mathcal{R}\right)}{C\left(\mathcal{H}_0'; x|\mathcal{R}\right)}, \tag{28}$$

and let $A\left(\mathcal{H}_0; x|\mathcal{R}\right)$ denote the acceptability of the hypothesis that $\theta \in \mathcal{H}_0$, given the observation that $X = x$ and the restriction that $\theta \in \mathcal{R}$. Assume, contrary to the claim, that $A\left(\mathcal{H}_0; x|\mathcal{R}\right) \neq \widetilde{A}\left(\mathcal{H}_0\right)$. In the case that relation (14) holds, $\widetilde{A}\left(\mathcal{H}_0\right) = \log {}^1/_{C\left(\mathcal{H}_0'; x|\mathcal{R}\right)}$ by

Corollary 3. From equation (4) and Lemma 4,

$$\begin{aligned}
\widetilde{A}(\mathcal{H}_0) &= \log\left(1/\sup\left\{\alpha \in \left]0,1\right] : \mathcal{H}_0' \overset{\alpha|\mathcal{R}}{\sim} x\right\}\right) \\
&= \log\left(1/\sup\left(\left]0,1\right] \setminus \left\{\alpha \in \left]0,1\right] : \neg\mathcal{H}_0' \overset{\alpha|\mathcal{R}}{\sim} x\right\}\right)\right) \\
&= \log\left(1/\sup\left\{\alpha \in \left]0,1\right] : A\left(\mathcal{H}_0'; x|\mathcal{R}\right) \geq -\log {}^1\!/\!_\alpha\right\}\right) \\
&= \inf\left\{\log {}^1\!/\!_\alpha \geq 0 : \log {}^1\!/\!_\alpha \geq -A\left(\mathcal{H}_0'; x|\mathcal{R}\right)\right\} \\
&= -A\left(\mathcal{H}_0'; x|\mathcal{R}\right) = A\left(\mathcal{H}_0; x|\mathcal{R}\right),
\end{aligned}$$

the last equality following from the equivalence of Propositions 1 and 4 of Lemma 4. In the case that relation (14) does not hold, $\sup_{\theta_1 \in \mathcal{R}} p(\theta_1; x) > \sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} p(\theta_0; x)$, yielding

$$\sup_{\theta_1 \in \mathcal{R}} p(\theta_1; x) = \sup_{\theta_0 \in \mathcal{H}_0' \cap \mathcal{R}} p(\theta_0; x).$$

Thus, Corollary 3 now gives $C\left(\mathcal{H}_0'; x|\mathcal{R}\right) = 1$ and $\widetilde{A}(\mathcal{H}_0) = \log C\left(\mathcal{H}_0; x|\mathcal{R}\right)$ by implication. From equation (4) and Lemma 4,

$$\begin{aligned}
\widetilde{A}(\mathcal{H}_0) &= \log\left(\sup\left\{\alpha \in \left]0,1\right] : \mathcal{H}_0 \overset{\alpha|\mathcal{R}}{\sim} x\right\}\right) \\
&= \log\left(\sup\left\{\alpha \in \left]0,1\right] : A\left(\mathcal{H}_0; x|\mathcal{R}\right) \geq -\log {}^1\!/\!_\alpha\right\}\right) \\
&= \sup\left\{\log\alpha \leq 0 : A\left(\mathcal{H}_0; x|\mathcal{R}\right) \geq \log\alpha\right\} \\
&= \sup\left\{\log\alpha \leq 0 : \log\alpha \leq A\left(\mathcal{H}_0; x|\mathcal{R}\right)\right\} \\
&= A\left(\mathcal{H}_0; x|\mathcal{R}\right).
\end{aligned}$$

Therefore, $\widetilde{A}(\mathcal{H}_0) = A\left(\mathcal{H}_0; x|\mathcal{R}\right)$ in both possible cases, contradicting the assumption and establishing equation (26). The rest of the claims follow from Theorem 1. □

Breaking that into the three major cases sheds light on the interpretation of acceptability.

Corollary 4. *For any* $x \in \mathcal{X}$*,* $\mathcal{H}_0 \in \mathfrak{H}$*, and* $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$ *such that* $\mathcal{H}_0 \cap \mathcal{R} \neq \emptyset$ *and* $\mathcal{H}_0' \cap \mathcal{R} \neq \emptyset$*, the acceptability of the hypothesis that* $\theta \in \mathcal{H}_0$*, given the observation that* $X = x$ *and the restriction that* $\theta \in \mathcal{R}$*, is*

$$A\left(\mathcal{H}_0; x|\mathcal{R}\right) = \begin{cases} -\log \sup_{\theta_0 \in \mathcal{H}_0' \cap \mathcal{R}} c\left(\theta_0; x|\mathcal{R}\right) & \text{if } \mathcal{H}_0 \cap \widehat{\mathcal{H}}\left(x|\mathcal{R}\right) \neq \emptyset, \mathcal{H}_0' \cap \widehat{\mathcal{H}}\left(x|\mathcal{R}\right) = \emptyset \\ \log \sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} c\left(\theta_0; x|\mathcal{R}\right) & \text{if } \mathcal{H}_0 \cap \widehat{\mathcal{H}}\left(x|\mathcal{R}\right) = \emptyset, \mathcal{H}_0' \cap \widehat{\mathcal{H}}\left(x|\mathcal{R}\right) \neq \emptyset \\ 0 & \text{if } \mathcal{H}_0 \cap \widehat{\mathcal{H}}\left(x|\mathcal{R}\right) \neq \emptyset, \mathcal{H}_0' \cap \widehat{\mathcal{H}}\left(x|\mathcal{R}\right) \neq \emptyset; \end{cases}$$

$$\widehat{\mathcal{H}}\left(x|\mathcal{R}\right) = \left\{\theta_1 \in \mathcal{R} : \forall \theta_0 \in \Theta, p\left(\theta_0; x\right) \leq p\left(\theta_1; x\right)\right\}. \tag{29}$$

Proof. Corollary 1 implies that $\widehat{\mathcal{H}}\left(x|\mathcal{R}\right) = \left\{\theta_0 \in \mathcal{R} : c\left(\theta_0; x|\mathcal{R}\right) = 1\right\}$. Thus, by equation (27),

$$\begin{aligned} A\left(\mathcal{H}_0; x|\mathcal{R}\right) &= \log \frac{\sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} c\left(\theta_0; x|\mathcal{R}\right)}{\sup_{\theta_0 \in \mathcal{H}_0' \cap \mathcal{R}} c\left(\theta_0; x|\mathcal{R}\right)} \\ &= \begin{cases} \log\left(1/\sup_{\theta_0 \in \mathcal{H}_0' \cap \mathcal{R}} c\left(\theta_0; x|\mathcal{R}\right)\right) & \text{if } \mathcal{H}_0 \cap \widehat{\mathcal{H}}\left(x|\mathcal{R}\right) \neq \emptyset \\ \log\left(\sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} c\left(\theta_0; x|\mathcal{R}\right)/1\right) & \text{if } \mathcal{H}_0' \cap \widehat{\mathcal{H}}\left(x|\mathcal{R}\right) \neq \emptyset. \end{cases} \end{aligned}$$

If both $\mathcal{H}_0 \cap \widehat{\mathcal{H}}\left(x|\mathcal{R}\right) \neq \emptyset$ and $\mathcal{H}_0' \cap \widehat{\mathcal{H}}\left(x|\mathcal{R}\right) \neq \emptyset$, then $-\log \sup_{\theta_0 \in \mathcal{H}_0' \cap \mathcal{R}} c\left(\theta_0; x|\mathcal{R}\right) = \log \sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} c\left(\theta_0; x|\mathcal{R}\right)$, which is only possible if $\sup_{\theta_0 \in \mathcal{H}_0' \cap \mathcal{R}} c\left(\theta_0; x|\mathcal{R}\right) = \sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} c\left(\theta_0; x|\mathcal{R}\right) = 1$.

□

In the typical case of testing a simple hypothesis (Corollary 2), its acceptability cannot be positive. That agrees with the idea that evidence might be against a simple hypothesis but can never support it. The necessary conditions are stated in this corollary of Theorem 4. Its brief proof is omitted.

Corollary 5. *For any* $x \in \mathcal{X}$, $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$, *and* $\theta_0 \in \Theta$ *such that* $\sup_{\theta_1 \in \mathcal{R}} p(\theta_1; x) = 1$, *given the observation that* $X = x$ *and the restriction that* $\theta \in \mathcal{R}$, *the acceptability of the hypothesis that* $\theta = \theta_0$ *is* $A(\mathcal{H}_0; x|\mathcal{R}) = \log p(\theta_0; x) \leq 0$.

5.3 Connections with possibility theory

In agreement with the classical idea of inference to the best explanation (Peirce, 1998, p. 234), the acceptability $A(\mathcal{H}_0; x|\mathcal{R})$ may be understood as the degree to which the data would evoke surprise were the hypothesis that $\theta \in \mathcal{H}_0$ known to be false. While that should not be confused with Shackle's degree of potential surprise in the revealed truth of a hypothesis (Shackle, 1961), the concepts share many properties at the mathematical level.

Those relationships may be succinctly expressed in terms of possibility theory and ranking theory, the successors of the the theory of potential surprise (Remark 1). A function $\text{Poss} : \mathfrak{H} \to [0, 1]$ is a *possibility measure* on $\mathfrak{H}$ if $\text{Poss}(\emptyset) = 0$, $\text{Poss}(\Theta) = 1$, and $\text{Poss}\left(\bigcup_{j \in \mathcal{J}} \mathcal{H}_{0j}\right) = \sup_{j \in \mathcal{J}} \text{Poss}(\mathcal{H}_{0j})$ for any index set $\mathcal{J}$ such that $\bigcup_{j \in \mathcal{J}} \mathcal{H}_{0j} \in \mathfrak{H}$ and $\mathcal{H}_{0j} \in \mathfrak{H}$ for all $j \in \mathcal{J}$ (Wang, 2008, §4.6). Further, a function $\pi : \Theta \to [0, 1]$ such that $\text{Poss}(\mathcal{H}) = \sup_{\theta \in \Theta} \pi(\theta)$ is called a *possibility profile*, and a function $\text{Nec} : \mathfrak{H} \to [0, 1]$ is a *necessity measure* on $\mathfrak{H}$ if $\text{Nec}(\mathcal{H}) = 1 - \text{Poss}(\mathcal{H}')$ for all $\mathcal{H} \in \mathfrak{H}$ (Wang, 2008, §4.6). Thus, $C(\mathcal{H}_0; x|\mathcal{R}) = \sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} c(\theta_0; x|\mathcal{R})$ as a function of $\mathcal{H}_0$ is a possibility measure corresponding to the possibility profile $c(\bullet; x|\mathcal{R})$. Similarly, in view of Definition 3, $W(\mathcal{H}_0; x|\mathcal{R})$ as a function of $\mathcal{H}_0$ is a necessity measure. The conditionalization of equation (9) has often been considered as a rule for conditional possibility (e.g., Dubois and Prade, 1998; De Baets et al., 1999; Marchioni, 2006). As precursors to this transformation of the compatibility function of a parameter into a possibility measure, the p-value function of a parameter and the likelihood function had been transformed into possibility measures (Remark 2).

6 Deductive cogency of acceptability

As stated in Section 1, every deductively cogent statistical procedure is both restriction-respecting and coherent. Those properties will be proven of the acceptability method (Definition 4) in the next two subsections.

6.1 Acceptability is restriction-respecting

Recall that a restriction-respecting statistical method does not permit the rejection of all hypotheses that are consistent with the restriction but requires the rejection of all hypotheses that are inconsistent with the restriction (§1). Conditional acceptability is now seen to be restriction-respecting.

Theorem 5. *For any $\alpha \in]0,1]$, conditional on the restriction that $\theta \in \mathcal{R}$ for some $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$, the procedure in Definition 4 rejects the hypothesis that $\theta \in \mathcal{H}_0$ for every $\mathcal{H}_0 \in \mathfrak{H}$ such that $\mathcal{H}_0 \cap \mathcal{R} = \emptyset$ and does not reject every hypothesis that $\theta \in \mathcal{H}_1$ for all $\mathcal{H}_1 \in \mathfrak{H}$ such that $\mathcal{H}_1 \cap \mathcal{R} \neq \emptyset$.*

Proof. Theorem 4 says $A(\mathcal{H}_0; x|\mathcal{R}) = -\infty$ for every $\mathcal{H}_0 \in \mathfrak{H}$ such that $\mathcal{H}_0 \cap \mathcal{R} = \emptyset$. Thus, $A(\mathcal{H}_0; x|\mathcal{R}) < -\log {}^1/_\alpha$, which means $\theta \in \mathcal{H}_0$ is rejected, for all $\alpha \in]0,1]$. To prove the other claim, let $\widehat{\mathcal{H}}(x|\mathcal{R})$ be defined according to equation (29), and denote its complement by $\widehat{\mathcal{H}}'(x|\mathcal{R}) = \Theta \backslash \widehat{\mathcal{H}}(x|\mathcal{R})$. If $\widehat{\mathcal{H}}(x|\mathcal{R}) = \mathcal{R}$, then $\widehat{\mathcal{H}}'(x|\mathcal{R}) \cap \mathcal{R} = \emptyset$ and, according to Theorem

4, $A\left(\widehat{\mathcal{H}}\,(x|\mathcal{R})\,;x|\mathcal{R}\right)=\infty$. On the other hand, if $\widehat{\mathcal{H}}\,(x|\mathcal{R})\neq\mathcal{R}$, then $\widehat{\mathcal{H}}'\,(x|\mathcal{R})\cap\mathcal{R}\neq\emptyset$, and Theorem 4, with equation (29), yields

$$A\left(\widehat{\mathcal{H}}\,(x|\mathcal{R})\,;x|\mathcal{R}\right)=\log\frac{\sup_{\theta_0\in\mathcal{R}}p\,(\theta_0;x)}{\sup_{\theta_0\in\widehat{\mathcal{H}}'(x|\mathcal{R})\cap\mathcal{R}}p\,(\theta_0;x)}\geq 0.$$

Thus, since $A\left(\widehat{\mathcal{H}}\,(x|\mathcal{R})\,;x|\mathcal{R}\right)\geq 0$ in both cases, there is no $\alpha\in\,]0,1]$ such that $A\left(\widehat{\mathcal{H}}\,(x|\mathcal{R})\,;x|\mathcal{R}\right)<-\log{}^{1}/_{\alpha}$, which means $\theta\in\widehat{\mathcal{H}}\,(x|\mathcal{R})$ cannot be rejected. □

6.2 Acceptability is coherent

In the context of multiple comparisons, Gabriel (1969) called a statistical procedure "coherent" if, for every hypothesis that it rejects, it also rejects all of the hypotheses that imply the truth of the rejected hypothesis (§1). Thus, for every $\mathcal{H}_0\in\mathfrak{H}$, any *rejection-coherent* procedure rejects the hypothesis that $\theta\in\mathcal{H}_1$ for every $\mathcal{H}_1\in\mathfrak{H}$ such that $\mathcal{H}_1\subseteq\mathcal{H}_0$ if it rejects the hypothesis that $\theta\in\mathcal{H}_0$. Likewise, for every $\mathcal{H}_1\in\mathfrak{H}$, any *acceptance-coherent* procedure accepts the hypothesis that $\theta\in\mathcal{H}_0$ for every $\mathcal{H}_0\in\mathfrak{H}$ such that $\mathcal{H}_1\subseteq\mathcal{H}_0$ if it accepts the hypothesis that $\theta\in\mathcal{H}_1$.

The concepts are applied to compatibility and acceptability in the next two results.

Lemma 5. *Conditional on the restriction that $\theta\in\mathcal{R}$ for some $\mathcal{R}\in\mathfrak{H}\setminus\{\emptyset\}$, the compatibility of the hypothesis that $\theta\in\mathcal{H}_1$ with the observation that $X=x$ is at most the compatibility of any hypothesis that it implies with the same observation, that is,*

$$C\,(\mathcal{H}_1;x|\mathcal{R})\leq C\,(\mathcal{H}_0;x|\mathcal{R})$$

for every $\mathcal{H}_0,\mathcal{H}_1\in\mathfrak{H}$ such that $\mathcal{H}_1\subseteq\mathcal{H}_0$.

Proof. According to Theorem 1, either $C(\mathcal{H}_1; x|\mathcal{R}) = 0$, in which case $C(\mathcal{H}_1; x|\mathcal{R}) \leq C(\mathcal{H}_0; x|\mathcal{R})$, or $C(\mathcal{H}_1; x|\mathcal{R}) > 0$, in which case $\mathcal{H}_1 \cap \mathcal{R} \neq \emptyset$. Thus, since $\mathcal{H}_1 \subseteq \mathcal{H}_0$, it follows from $\mathcal{H}_1 \cap \mathcal{R} \neq \emptyset$ that $\mathcal{H}_0 \cap \mathcal{R} \neq \emptyset$ and, by Theorem 1, that

$$\frac{C(\mathcal{H}_0; x|\mathcal{R})}{C(\mathcal{H}_1; x|\mathcal{R})} = \frac{\sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} p(\theta_0; x)}{\sup_{\theta_0 \in \mathcal{H}_1 \cap \mathcal{R}} p(\theta_0; x)}.$$

That ratio satisfies $C(\mathcal{H}_0; x|\mathcal{R}) / C(\mathcal{H}_1; x|\mathcal{R}) \geq 1$ given that $\mathcal{H}_1 \cap \mathcal{R} \subseteq \mathcal{H}_0 \cap \mathcal{R}$. □

Theorem 6. *Conditional on the restriction that $\theta \in \mathcal{R}$ for some $\mathcal{R} \in \mathfrak{H} \backslash \{\emptyset\}$, the procedure in Definition 4 is both rejection-coherent and acceptance-coherent for any $\alpha \in]0,1]$, and the acceptability of the hypothesis that $\theta \in \mathcal{H}_1$ is at most the acceptability of any hypothesis that it implies, that is,*

$$A(\mathcal{H}_1; x|\mathcal{R}) \leq A(\mathcal{H}_0; x|\mathcal{R}) \quad (30)$$

for every $\mathcal{H}_0, \mathcal{H}_1 \in \mathfrak{H}$ such that $\mathcal{H}_1 \subseteq \mathcal{H}_0$.

Proof. The following statements hold for any $\alpha \in]0,1]$. According to Definition 4, the hypothesis that $\theta \in \mathcal{H}_0$, given the observation that $X = x$ and the restriction that $\theta \in \mathcal{R}$ is rejected at level α if and only if $A(\mathcal{H}_0; x|\mathcal{R}) < -\log 1/\alpha$. That requires that $A(\mathcal{H}_0; x|\mathcal{R}) < 0$, which only obtains when either $\mathcal{H}_0 \cap \mathcal{R} = \emptyset$, in which case $A(\mathcal{H}_0; x|\mathcal{R}) = -\infty$ by Theorem 4, or

$$A(\mathcal{H}_0; x|\mathcal{R}) = \log \sup_{\theta_0 \in \mathcal{H}_0 \cap \mathcal{R}} c(\theta_0; x|\mathcal{R}) = -\log 1/C(\mathcal{H}_0; x|\mathcal{R}) \quad (31)$$

by Corollary 4 and Theorem 1. If, on the other hand $A(\mathcal{H}_0; x|\mathcal{R}) > 0$, as required for acceptance ($A(\mathcal{H}_0; x|\mathcal{R}) > \log 1/\alpha$) then either $\mathcal{H}_0' \cap \mathcal{R} = \emptyset$, in which case $A(\mathcal{H}_0; x|\mathcal{R}) = \infty$

by Theorem 4, or

$$A\left(\mathcal{H}_0;x|\mathcal{R}\right)=-\log\sup_{\theta_0\in\mathcal{H}_0'\cap\mathcal{R}}c\left(\theta_0;x|\mathcal{R}\right)=\log{}^1\!/_{C\left(\mathcal{H}_0';x|\mathcal{R}\right)} \tag{32}$$

by Corollary 4 and Theorem 1. Whether equation (31) or equation (32) applies, equation (30) follows from Lemma 5. Both rejection coherence and acceptance coherence are immediate consequences of equation (30).

□

7 Remarks

The following remarks elaborate on the relationship between acceptability and possibility that was noted in Section 5.3.

Remark 1. If Poss is a possibility measure, then $-\log\mathrm{Poss}\left(\mathcal{H}_0\right)$ as a function of $\mathcal{H}_0$ is a *negative ranking function* (Spohn, 2012, §11.8). It follows that

$$\mathrm{Rank}\left(\mathcal{H}_0\right)=\log\frac{\mathrm{Poss}\left(\mathcal{H}_0\right)}{\mathrm{Poss}\left(\mathcal{H}_0'\right)}$$

as a function of $\mathcal{H}_0$ is a *two-sided ranking function* (Spohn, 2012, §5.2). Both $-\log C\left(\mathcal{H}_0;x|\mathcal{H}_1\right)$ and the potential surprise of $\mathcal{H}_0$ (Shackle, 1961) as functions of $\mathcal{H}_0$ are negative ranking functions. While $-\log C\left(\mathcal{H}_0;x|\mathcal{H}_1\right)$ does not measure the potential surprise of learning that $\theta\in\mathcal{H}_0$, it might be seen as the level of surprise of observing that $X=x$ were it known that $\theta\in\mathcal{H}_0$, in accordance with the comments on surprise in Section 5.3.

Since $C(\bullet; x|\mathcal{H}_1)$ is a possibility measure and since

$$A(\mathcal{H}_0; x|\mathcal{H}_1) = \log \frac{C(\mathcal{H}_0; x|\mathcal{H}_1)}{C(\mathcal{H}_0'; x|\mathcal{H}_1)}$$

by equation (26), $A(\bullet; x|\mathcal{H}_1)$ qualifies mathematically as a conditional two-sided ranking function. However, the interpretation encoded in Definition 4 differs from that of Spohn (2012), who developed ranking theory to model degrees of belief.

Remark 2. When $\sup_{\theta_1 \in \Theta} p(\theta_1; x) = 1$, possibility theory provides useful interpretations of p-values and confidence levels. First, Corollary 2 interprets the p-value as the level of compatibility of the null hypothesis with the data or how possible it is that the the null hypothesis is true in light of the data. Second, Theorem 3 interprets the confidence level as the degree of warrant for the hypothesis or how necessarily true it is given the data. These interpretations in terms of possibility and necessity measures are related to previous work. Under broad conditions, the confidence-based methods of Mauris et al. (2001, §2.2), Dubois et al. (2004), Masson and Denœux (2006), and Ghasemi Hamed et al. (2012) likewise lead to interpreting p-values as possibility values. Dubois et al. (1997) and Giang and Shenoy (2005) instead used the likelihood function in place of the p-value function $p(\bullet; x)$ for the special case in which Θ is countable and $\mathfrak{H} = 2^{\Theta}$. Patriota (2013) defines a large-sample possibility measure using both likelihood and confidence concepts.

Acknowledgments

This research was partially supported by the Canada Foundation for Innovation, by the Ministry of Research and Innovation of Ontario, and by the Faculty of Medicine of the

University of Ottawa.

References

Ball, F., Britton, T., O'Neill, P., 2002. Empty confidence sets for epidemics, branching processes and brownian motion. Biometrika 89 (1), 211–224.

Berger, J., 2003. Could Fisher, Jeffreys and Neyman have agreed on testing? Statistical Science 18 (1), 1–12.

Bickel, D. R., 2012. The strength of statistical evidence for composite hypotheses: Inference to the best explanation. Statistica Sinica 22, 1147–1198.

Bickel, D. R., Padilla, M., 2014. A prior-free framework of coherent inference and its derivation of simple shrinkage estimators. Journal of Statistical Planning and Inference 145, 204–221.

Chuaqui, R., 1991. Truth, Possibility and Probability: New Logical Foundations of Probability and Statistical Inference. North-Holland Mathematics Studies. Elsevier Science, Oxford.

De Baets, B., Tsiporkova, E., Mesiar, R., 1999. Conditioning in possibility theory with strict order norms. Fuzzy Sets and Systems 106, 221–229.

Dubois, D., Foulloy, L., Mauris, G., Prade, H., 2004. Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. Reliable Computing 10 (4), 273–297.

Dubois, D., Moral, S., Prade, H., Jan. 1997. A semantics for possibility theory based on likelihoods. Journal of Mathematical Analysis and Applications 205 (2), 359–380.

Dubois, D., Prade, H., 1998. Possibility theory: Qualitative and quantitative aspects. In: Smets, P. (Ed.), Quantified Representation of Uncertainty and Imprecision. Vol. 1 of Handbook of Defeasible Reasoning and Uncertainty Management Systems. Springer Netherlands, pp. 169–226.

Efron, B., Tibshirani, R., 1998. The problem of regions. Annals of Statistics 26, 1687–1718.

Fraser, D. A. S., 2011. Is Bayes posterior just quick and dirty confidence? Statistical Science 26, 299–316.

Gabriel, K. R., 1969. Simultaneous test procedures - some theory of multiple comparisons. Ann. Math. Statist. 40, 224–250.

Ghasemi Hamed, M., Serrurier, M., Durand, N., 2012. Representing uncertainty by possibility distributions encoding confidence bands, tolerance and prediction intervals. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 7520, 233–246.

Giang, P. H., Shenoy, P. P., 2005. Decision making on the sole basis of statistical likelihood. Artificial Intelligence 165, 137–163.

Hacking, I., 2001. An introduction to probability and inductive logic. Cambridge University Press, Cambridge.

Kaplan, M., 1996. Decision Theory As Philosophy. Cambridge University Press, Cambridge.

Lavine, M., Schervish, M. J., 1999. Bayes factors: What they are and what they are not. American Statistician 53, 119–122.

Mandelkern, M., 2002. Setting confidence intervals for bounded parameters. Statistical Science 17, 149–172.

Marchand, E., Strawderman, W. E., 2004. Estimation in restricted parameter spaces: A review. Lecture Notes-Monograph Series 45, 21–44.

Marchand, É., Strawderman, W., 2013. On bayesian credible sets, restricted parameter spaces and frequentist coverage. Electronic Journal of Statistics 7 (1), 1419–1431.

Marchioni, E., 2006. Possibilistic conditioning framed in fuzzy logics. International Journal of Approximate Reasoning 43 (2), 133–165.

Masson, M.-H., Denœux, T., 2006. Inferring a possibility distribution from empirical data. Fuzzy Sets and Systems 157 (3), 319–340.

Mauris, G., Lasserre, V., Foulloy, L., Apr. 2001. A fuzzy approach for the expression of uncertainty in measurement. Measurement 29 (3), 165–177.

Patriota, A. G., 2013. A classical measure of evidence for general null hypotheses. Fuzzy Sets and Systems 233, 74 – 88.

Peirce, C., 1998. The Essential Peirce: Selected Philosophical Writings (1893-1913). The Essential Peirce: Selected Philosophical Writings. Indiana University Press.

Polansky, A. M., 2007. Observed Confidence Levels: Theory and Application. Chapman and Hall, New York.

Royall, R., 1997. Statistical Evidence: A Likelihood Paradigm. CRC Press, New York.

Schervish, M. J., 1996. P values: What they are and what they are not. American Statistician 50, 203–206.

Schweder, T., Hjort, N. L., 2002. Confidence and likelihood. Scandinavian Journal of Statistics 29, 309–332.

Shackle, G., 1961. Decision, Order and Time in Human Affairs. Cambridge University Press, Cambridge.

Silvapulle, M., Sen, P., 2011. Constrained Statistical Inference: Order, Inequality, and Shape Constraints. Wiley Series in Probability and Statistics. John Wiley and Sons, New York.

Spohn, W., 2012. The Laws of Belief: Ranking Theory and Its Philosophical Applications. Oxford University Press.

Wang, H., 2004. Improved estimation of accuracy in simple hypothesis versus simple alternative testing. Journal of Multivariate Analysis 90 (2), 269 – 281.

Wang, H., 2006. Modified p-value of two-sided test for normal distribution with restricted parameter space. Communications in Statistics-Theory and Methods 35 (8), 1361–1374.

Wang, H., 2007. Modified p-values for one-sided testing in restricted parameter spaces. Statistics and Probability Letters 77, 625–631.

Wang, Q., 2008. Probability distribution and entropy as a measure of uncertainty. Journal of Physics A: Mathematical and Theoretical 41, 065004.

Wendell, J. P., Schmee, J., 1996. Exact inference for proportions from a stratified finite population. Journal of the American Statistical Association 91 (434), pp. 825–830.

Xie, M.-G., Singh, K., 2013. Confidence distribution, the frequentist distribution estimator of a parameter: A review. International Statistical Review 81 (1), 3–39.

Zhang, T., Woodroofe, M., 2003. Credible and confidence sets for restricted parameter spaces. Journal of Statistical Planning and Inference 115, 479–490.

Zhang, Z., Zhang, B., 2013. A likelihood paradigm for clinical trials. Journal of Statistical Theory and Practice 7, 157–177.