On some assumptions of Null Hypothesis Statistical Testing (NHST)

Alexandre Galvão Patriota

Departamento de Estatística
Instituto de Matemática e Estatística
Universidade de São Paulo
The main goals of this presentation are
The main goals of this presentation are

- to discuss the **classical statistical model and statistical hypotheses**, 


Alexandre G. Patriota (patriota@ime.usp.br)
The main goals of this presentation are

- to discuss the **classical statistical model** and **statistical hypotheses**,  

- to present some **limitations of the classical p-value** with numerical examples,
Main goals

The main goals of this presentation are

- to discuss the **classical statistical model** and **statistical hypotheses**, 

- to present some **limitations of the classical p-value** with numerical examples,

- to introduce **an alternative measure of evidence**, called **s-value**, that overcomes some limitations of the p-value.
The classical statistical model is:

$$(\Omega, \mathcal{F}, \mathcal{P})$$,

where:

- $\Omega$ is the space of possible experiment outcomes,
- $\mathcal{F}$ is a $\sigma$-field of $\Omega$,
- $\mathcal{P}$ is a family of non-random probability measures that possibly explain the experiment outcomes.

**Remark:** a random vector $Z$ is a measurable function from $(\Omega, \mathcal{F})$ to $(\mathcal{Z}, \mathcal{B})$. 
The classical statistical model

The classical statistical model is:

\[(\Omega, \mathcal{F}, \mathcal{P}),\]

where:

- \(\Omega\) is the space of possible experiment outcomes,
- \(\mathcal{F}\) is a \(\sigma\)-field of \(\Omega\),
- \(\mathcal{P}\) is a family of non-random probability measures that possibly explain the experiment outcomes.

**Remark:** a random vector \(Z\) is a measurable function from \((\Omega, \mathcal{F})\) to \((Z, \mathcal{B})\)

The quantity of interest is \(g(P)\). For instance:

\[g(P) = EP(Z),\]
\[g(P) = P(Z_1 \in B | Z_2 \in A),\]
and so on.
A particular model

Take $Z = (X, \gamma)$, where $X$ is the **observable** random vector and $\gamma$ is the **unobservable** one.
Take $Z = (X, \gamma)$, where $X$ is the observable random vector and $\gamma$ is the unobservable one.

**Conditional, marginal and joint** distributions can be used to make inferences about $\gamma$. 
A particular model

Take $Z = (X, \gamma)$, where $X$ is the observable random vector and $\gamma$ is the unobservable one.

Conditional, marginal and joint distributions can be used to make inferences about $\gamma$.

Take $\mathcal{P} = \{P_0\}$ and build your joint probability $P_0$ from:

- $\gamma \sim f_0(\cdot)$ (with no unknown constants),
- $X|\gamma \sim f_1(\cdot|\gamma)$

Now, you are ready to be a hard core Bayesian!
Can we reduce the family $\mathcal{P}$ to a subfamily $\mathcal{P}_0$, where $\mathcal{P}_0 \subset \mathcal{P}$?
Can we reduce the family $\mathcal{P}$ to a subfamily $\mathcal{P}_0$, where $\mathcal{P}_0 \subset \mathcal{P}$?

The positive claim can be written by means of a null hypothesis:

$$H_0: \text{“at least one measure in } \mathcal{P}_0 \text{ could generate the observed data”}$$

(or simply $H_0: \text{“} P \in \mathcal{P}_0 \text{”}$)
Hypothesis testing

**Classical null hypotheses**

Can we reduce the family \( \mathcal{P} \) to a subfamily \( \mathcal{P}_0 \), where \( \mathcal{P}_0 \subset \mathcal{P} \)?

The positive claim can be written by means of a null hypothesis:

\[
H_0: \text{“at least one measure in } \mathcal{P}_0 \text{ could generate the observed data”}
\]

(or simply \( H_0 : “P \in \mathcal{P}_0” \))

Under a parametric model, there exists a finite dimensional set \( \Theta \) such that:

- \( \mathcal{P} \equiv \{ P_\theta : \theta \in \Theta \} \), where \( \Theta \subseteq \mathbb{R}^p \), \( p < \infty \),
- \( H_0 : \theta \in \Theta_0 \), where \( \Theta_0 \subset \Theta \) and \( \mathcal{P}_0 \equiv \{ P_\theta : \theta \in \Theta_0 \} \).
According to Fisher, the negation of $H_0$ cannot be expressed in terms of probability measures.

\[ H_1 : P \in (P - P_0) \]

\(^1\)since they would be mutually exclusive and exhaustive.
According to Fisher, the negation of $H_0$ cannot be expressed in terms of probability measures.

The alternative hypothesis $H_1$ makes sense if we are certain about the family $\mathcal{P}$: $H_1 : P \in (\mathcal{P} - \mathcal{P}_0)$. 

---

1 since they would be mutually exclusive and exhaustive.
According to Fisher, the negation of $H_0$ cannot be expressed in terms of probability measures.

The alternative hypothesis $H_1$ makes sense if we are certain about the family $\mathcal{P}$: $H_1 : P \in (\mathcal{P} - \mathcal{P}_0)$.

In the last context, we can choose\footnote{since they would be mutually exclusive and exhaustive} between $H_0$ and $H_1$ — Neyman and Pearson approach.
Bayesian hypotheses

Recall the hard core Bayesian approach, where $\mathcal{P} = \{P_0\}$ and $Z = (X, \gamma)$.
Bayesian hypotheses

Recall the hard core Bayesian approach, where \( \mathcal{P} = \{ P_0 \} \) and \( Z = (X, \gamma) \).

The general null and alternative hypotheses are:

- \( H_0 : \"\gamma \in \Gamma_0\" \)
- \( H_1 : \"\gamma \in (\Gamma - \Gamma_0)\" \)

A classical statistician may also test Bayesian hypotheses. Rather than p-values, they would use estimated conditional probabilities.
Bayesian hypotheses

Recall the hard core Bayesian approach, where $\mathcal{P} = \{P_0\}$ and $Z = (X, \gamma)$.

The general null and alternative hypotheses are:

$$H_0 : \"\gamma \in \Gamma_0\" \quad \text{and} \quad H_1 : \"\gamma \in (\Gamma - \Gamma_0)\".$$  

The focus is not on the family of probability measures $\mathcal{P}$, since $P_0$ is given.
Bayesian hypotheses

Recall the hard core Bayesian approach, where $\mathcal{P} = \{P_0\}$ and $Z = (X, \gamma)$.

The general null and alternative hypotheses are:

\[
H_0 : \, \gamma \in \Gamma_0 \quad \text{and} \quad H_1 : \, \gamma \in (\Gamma - \Gamma_0)
\]

The focus is not on the family of probability measures $\mathcal{P}$, since $P_0$ is given.

A classical statistician may also test Bayesian hypotheses. Rather than p-values, they would use estimated conditional probabilities.
The p-value for testing the classical null hypothesis $H_0$ is defined as follows

\[ p(P_0, x) = \sup_{P \in P_0} P(T_{H_0}(X) > T_{H_0}(x)) \]

where $T_{H_0}$ is a statistic such that the more discrepant is $H_0$ from $x$, the larger is its observed value.\(^2\)

\(^2\)i.e., $T_{H_0}$ could be $-2\log$ of the likelihood-ratio statistic.
The p-value for testing the classical null hypothesis $H_0$ is defined as follows

$$p(\mathcal{P}_0, x) = \sup_{P \in \mathcal{P}_0} P\left(T_{H_0}(X) > T_{H_0}(x)\right)$$

where $T_{H_0}$ is a statistic such that the more discrepant is $H_0$ from $x$, the larger is its observed value.\(^2\)

\[^2\text{i.e., } T_{H_0} \text{ could be } -2 \log \text{ of the likelihood-ratio statistic.}\]
Consider two null hypotheses $H_0 : \{ P \in \mathcal{P}_0 \}$ and $H'_0 : \{ P \in \mathcal{P}'_0 \}$ such that $H_0 \implies H'_0$. Then, we would expect that:

$$p(\mathcal{P}_0, x) \leq p(\mathcal{P}'_0, x)$$

But it is not always the case! The previous p-value is not monotone over the set of null hypotheses/Sets.
Consider two null hypotheses $H_0 : \ "P \in \mathcal{P}_0"$ and $H_0' : \ "P \in \mathcal{P}_0'"$ such that $H_0 \implies H_0'$. Then, we would expect that:

$$p(\mathcal{P}_0, x) \leq p(\mathcal{P}_0', x).$$

But it is not always the case!
Consider two null hypotheses $H_0 : \{P \in \mathcal{P}_0\}$ and $H'_0 : \{P \in \mathcal{P}'_0\}$ such that $H_0 \implies H'_0$. Then, we would expect that:

$$p(\mathcal{P}_0, x) \leq p(\mathcal{P}'_0, x)$$

But it is not always the case!

The previous p-value is not monotone over the set of null hypotheses/Sets.
Example: Bivariate Normal distribution

Let $X = (X_1, \ldots, X_n)$ be a sample from a bivariate normal distribution with mean vector $\mu = (\mu_1, \mu_2)^\top$ and identity variance matrix.
Example: Bivariate Normal distribution

Let \( X = (X_1, \ldots, X_n) \) be a sample from a bivariate normal distribution with mean vector \( \mu = (\mu_1, \mu_2)^\top \) and identity variance matrix.

Notice that \((-2 \log \text{ of})\) the likelihood-ratio statistic

- under \( H_0 : \mu = 0 \) is

\[
T_{H_0}(X) = n \bar{X}^\top \bar{X} \sim \chi^2_2,
\]
Example: Bivariate Normal distribution

Let \( X = (X_1, \ldots, X_n) \) be a sample from a bivariate normal distribution with mean vector \( \mu = (\mu_1, \mu_2)^\top \) and identity variance matrix.

Notice that \((-2 \log \text{ of})\) the likelihood-ratio statistic

- under \( H_0 : \mu = 0 \) is

\[
T_{H_0}(X) = n \bar{X}^\top \bar{X} \sim \chi^2_2,
\]

- under \( H'_0 : \mu_1 = \mu_2 \) is

\[
T_{H'_0}(X) = \frac{n}{2} (\bar{X}_1 - \bar{X}_2)^2 \sim \chi^2_1,
\]

where \( \bar{X} = (\bar{X}_1, \bar{X}_2)^\top \) is the maximum likelihood estimator for \( \mu \).
P-values do not respect monotonicity

<table>
<thead>
<tr>
<th>Observed sample $(\bar{x}_1, \bar{x}_2)$</th>
<th>$\bar{x}_1 - \bar{x}_2$</th>
<th>$H_0: \mu = 0$ p-value</th>
<th>$H'_0: \mu_1 = \mu_2$ p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.05,-0.05)</td>
<td>0.1</td>
<td>0.9753</td>
<td>0.8231</td>
</tr>
<tr>
<td>(0.09,-0.11)</td>
<td>0.2</td>
<td>0.9039</td>
<td>0.6547</td>
</tr>
<tr>
<td>(0.14,-0.16)</td>
<td>0.3</td>
<td>0.7977</td>
<td>0.5023</td>
</tr>
<tr>
<td>(0.19,-0.21)</td>
<td>0.4</td>
<td>0.6697</td>
<td>0.3711</td>
</tr>
<tr>
<td>(0.23,-0.27)</td>
<td>0.5</td>
<td>0.5331</td>
<td>0.2636</td>
</tr>
<tr>
<td>(0.28,-0.32)</td>
<td>0.6</td>
<td>0.4049</td>
<td>0.1797</td>
</tr>
<tr>
<td>(0.33,-0.37)</td>
<td>0.7</td>
<td>0.2926</td>
<td>0.1175</td>
</tr>
<tr>
<td>(0.37,-0.43)</td>
<td>0.8</td>
<td>0.2001</td>
<td>0.0736</td>
</tr>
<tr>
<td>(0.42,-0.48)</td>
<td>0.9</td>
<td>0.1308</td>
<td>0.0442</td>
</tr>
<tr>
<td>(0.47,-0.53)</td>
<td>1.0</td>
<td>0.0813</td>
<td>0.0253</td>
</tr>
</tbody>
</table>
Level curves (contour curves)

Significance level 10%
Level curves (contour curves)

Significance level 10%

Region of rejection

$H_0' : \mu_1 = \mu_2$
Level curves (contour curves)

Significance level 10%

Region of rejection

\[ H'_0 : \mu_1 = \mu_2 \]

\[ H_0 : \mu = 0 \]
Level curves (contour curves)

Significance level 10%

Region of rejection

H₀' : \( \mu_1 = \mu_2 \)

\( \uparrow \)

H₀ : \( \mu = 0 \)

and

Rejection of \( H₀' \) but no rejection of \( H₀ \)

non-coherent conclusion

Rejection of \( H₀' \) but no rejection of \( H₀ \)

non-coherent conclusion
An alternative measure of evidence (parametric case)

In what follows, we present an alternative measure called **s-value** to overcome the previous issue (Patriota, 2013, FSS, 233).

$$s(\Theta_0, x) = \begin{cases} \sup \{\alpha \in (0, 1) : \Lambda_\alpha(x) \cap \Theta_0 \neq \emptyset\}, & \text{if } \Theta_0 \neq \emptyset, \\ 0, & \text{if } \Theta_0 = \emptyset. \end{cases}$$

where $\Lambda_\alpha$ is a confidence set for $\theta$ with confidence level $1 - \alpha$ with some "nice" properties.
An alternative measure of evidence (parametric case)

In what follows, we present an alternative measure called s-value to overcome the previous issue (Patriota, 2013, FSS, 233).

The s-value is a function $s : 2^{\Theta} \times X \rightarrow [0, 1]$ such that

$$s(\Theta_0, x) = \begin{cases} 
\sup\{\alpha \in (0, 1) : \Lambda_\alpha(x) \cap \Theta_0 \neq \emptyset\}, & \text{if } \Theta_0 \neq \emptyset, \\
0, & \text{if } \Theta_0 = \emptyset.
\end{cases}$$

where $\Lambda_\alpha$ is a confidence set for $\theta$ with confidence level $1 - \alpha$ with some “nice” properties.
Interpretation: \( s = s(\Theta_0, x) \) is the largest significance level \( \alpha \) (or \( 1 - s \) is the smallest confidence level \( 1 - \alpha \)) for which the confidence set and the set \( \Theta_0 \) have at least one element in common.
Interpretation: \( s = s(\Theta_0, x) \) is the largest significance level \( \alpha \) (or \( 1 - s \) is the smallest confidence level \( 1 - \alpha \)) for which the confidence set and the set \( \Theta_0 \) have at least one element in common.

Large values of \( s \) indicate that **there exists at least one** element in \( \Theta_0 \) close to the center of \( \Lambda_\alpha \) (e.g., close to the ML estimate).
**Interpretation**

Interpretation: \( s = s(\Theta_0, x) \) is the largest significance level \( \alpha \) (or \( 1 - s \) is the smallest confidence level \( 1 - \alpha \)) for which the confidence set and the set \( \Theta_0 \) have at least one element in common.

Large values of \( s \) indicate that there exists at least one element in \( \Theta_0 \) close to the center of \( \Lambda_\alpha \) (e.g., close to the ML estimate).

Small values of \( s \) indicate that **ALL** elements of \( \Theta_0 \) are far away from the center of \( \Lambda_\alpha \).
Graphical illustration: \( s_1 = s(\Theta_1, x) \)
An alternative measure of evidence and some of its properties

**Graphical illustration:** $s_2 = s(\Theta_2, x)$
Properties: the s-value is a possibility measure

1. $s(\emptyset, x) = 0$ and $s(\Theta, x) = 1$, 
Properties: the s-value is a possibility measure

1. \( s(\emptyset, x) = 0 \) and \( s(\Theta, x) = 1 \),

2. If \( \Theta_1 \subseteq \Theta_2 \), then \( s(\Theta_1, x) \leq s(\Theta_2, x) \),
Properties: the s-value is a possibility measure

1. \( s(\emptyset, x) = 0 \) and \( s(\Theta, x) = 1 \),

2. If \( \Theta_1 \subseteq \Theta_2 \), then \( s(\Theta_1, x) \leq s(\Theta_2, x) \),

3. For any \( \Theta_1, \Theta_2 \subseteq \Theta \), \( s(\Theta_1 \cup \Theta_2, x) = \max\{ s(\Theta_1, x), s(\Theta_2, x) \} \),
Properties: the s-value is a possibility measure

1. $s(\emptyset, x) = 0$ and $s(\Theta, x) = 1$,

2. If $\Theta_1 \subseteq \Theta_2$, then $s(\Theta_1, x) \leq s(\Theta_2, x)$,

3. For any $\Theta_1, \Theta_2 \subseteq \Theta$, $s(\Theta_1 \cup \Theta_2, x) = \max\{s(\Theta_1, x), s(\Theta_2, x)\}$,

4. If $\Theta_1 \subseteq \Theta$, then $s(\Theta_1, x) = \sup_{\theta \in \Theta_1} s(\{\theta\}, x)$,
Properties: the s-value is a possibility measure

1. $s(\emptyset, x) = 0$ and $s(\Theta, x) = 1$,

2. If $\Theta_1 \subseteq \Theta_2$, then $s(\Theta_1, x) \leq s(\Theta_2, x)$,

3. For any $\Theta_1, \Theta_2 \subseteq \Theta$, $s(\Theta_1 \cup \Theta_2, x) = \max\{s(\Theta_1, x), s(\Theta_2, x)\}$,

4. If $\Theta_1 \subseteq \Theta$, then $s(\Theta_1, x) = \sup_{\theta \in \Theta_1} s(\{\theta\}, x)$,

5. $s(\Theta_1, x) = 1$ or $s(\Theta_1^c, x) = 1$:
   - if $\hat{\theta} \in \overline{\Theta_1}$ (closure of $\Theta_1$), then $s(\Theta_1, x) = 1$,
   - if $\hat{\theta} \in \overline{\Theta_1^c}$, then $s(\Theta_1^c, x) = 1$.

where $\hat{\theta}$ is an element of the center of $\Lambda_\alpha$, i.e., $\hat{\theta} \in \bigcap_\alpha \Lambda_\alpha(x)$. 
Decisions about $H_0$

Let $\Phi$ be a function such that:

$$\Phi(\Theta_0) = \langle s(\Theta_0), s(\Theta_0^c) \rangle.$$

Then,

$$\Phi(\Theta_0) = \langle a, 1 \rangle \quad \implies \quad \text{rejection of } H_0 \text{ if } a \text{ is “small” enough},$$

$$\Phi(\Theta_0) = \langle 1, b \rangle \quad \implies \quad \text{acceptance of } H_0 \text{ if } b \text{ is “small” enough},$$

$$\Phi(\Theta_0) = \langle 1, 1 \rangle \quad \implies \quad \text{total ignorance about } H_0.$$
Decisions about $H_0$

Let $\Phi$ be a function such that:

$$\Phi(\Theta_0) = \langle s(\Theta_0), s(\Theta_0^c) \rangle.$$ 

Then,

- $\Phi(\Theta_0) = \langle a, 1 \rangle \iff \text{rejection of } H_0 \text{ if } a \text{ is “small” enough}$,

- $\Phi(\Theta_0) = \langle 1, b \rangle \iff \text{acceptance of } H_0 \text{ if } b \text{ is “small” enough}$.
Decisions about $H_0$

Let $\Phi$ be a function such that:

$$\Phi(\Theta_0) = \langle s(\Theta_0), s(\Theta_0^c) \rangle.$$ 

Then,

- $\Phi(\Theta_0) = \langle a, 1 \rangle \implies \text{rejection of } H_0 \text{ if } a \text{ is "small" enough,}$

- $\Phi(\Theta_0) = \langle 1, b \rangle \implies \text{acceptance of } H_0 \text{ if } b \text{ is "small" enough.}$

- $\Phi(\Theta_0) = \langle 1, 1 \rangle \implies \text{total ignorance about } H_0.$
An alternative measure of evidence and some of its properties

How to find the thresholds for $a$ and $b$ to decide about $H_0$?

This is still an open problem. We could try to find those thresholds via loss functions or via frequentist criteria by employing the following asymptotic property:

Property: If the statistical model is regular and the confidence region is built from a statistics $T_{\theta}(X)$ that converges in distribution to $\chi^2_k$, then:

$$s_a = 1 - F_k(F_{H_0}^{-1}(1 - p_a)),$$

where $p_a = 1 - F_{H_0}(t)$ is the asymptotic p-value to test $H_0$. 

Alexandre G. Patriota (patriota@ime.usp.br)
An alternative measure of evidence and some of its properties

How to find the thresholds for $a$ and $b$ to decide about $H_0$?

This is still an open problem.
How to find the thresholds for \( a \) and \( b \) to decide about \( H_0 \)?

This is still an open problem.

We could try to find those thresholds \textit{via} loss functions.
How to find the thresholds for $a$ and $b$ to decide about $H_0$?

This is still an open problem.

We could try to find those thresholds via loss functions.

or via frequentist criteria by employing the following asymptotic property:

$$s_a = 1 - F_k(F^{-1}_{H_0}(1 - p_a)),$$

where $p_a = 1 - F_{H_0}(t)$ is the asymptotic p-value to test $H_0$. 
How to find the thresholds for $a$ and $b$ to decide about $H_0$?

This is still an open problem.

We could try to find those thresholds \textit{via} loss functions.

or \textit{via} frequentist criteria by employing the following asymptotic property:

\textbf{Property:} If the statistical model is regular and the confidence region is built from a statistics $T_{\theta}(X)$ that converges in distribution to $\chi^2_k$, then:

$$s_a = 1 - F_k(F_{H_0}^{-1}(1 - p_a)),$$

where $p_a = 1 - F_{H_0}(t)$ is the asymptotic p-value to test $H_0$. 
Example: Bivariate Normal distribution

Let $X = (X_1, \ldots, X_n)$ be a sample from a bivariate normal distribution with mean vector $\mu = (\mu_1, \mu_2)^\top$ and identity variance matrix.
Example: Bivariate Normal distribution

Let $X = (X_1, \ldots, X_n)$ be a sample from a bivariate normal distribution with mean vector $\mu = (\mu_1, \mu_2)^\top$ and identity variance matrix.

Notice that ($-2 \log$ of) the likelihood-ratio statistic is

$$T_\mu(x) = n(\bar{X} - \mu)^\top(\bar{X} - \mu) \sim \chi^2_2,$$
Example: Bivariate Normal distribution

Let $X = (X_1, \ldots, X_n)$ be a sample from a bivariate normal distribution with mean vector $\mu = (\mu_1, \mu_2)^\top$ and identity variance matrix.

Notice that $(-2 \log \text{ of})$ the likelihood-ratio statistic is

$$T_\mu(x) = n(\bar{X} - \mu)^\top(\bar{X} - \mu) \sim \chi^2_2,$$

The confidence set $\Lambda_\alpha$ is given by

$$\Lambda_\alpha(x) = \{\mu \in \mathbb{R}^2 : T_\mu(x) \leq F_2^{-1}(1 - \alpha)\},$$

where $F_2$ is the cumulative chi-squared distribution with two degrees of freedom.
Numerical illustration

<table>
<thead>
<tr>
<th>Observed sample ((\bar{x}_1, \bar{x}_2))</th>
<th>(\bar{x}_1 - \bar{x}_2)</th>
<th>(H_0: \mu = 0) p/s-value</th>
<th>(H'_0: \mu_1 = \mu_2) p-value</th>
<th>s-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.05,-0.05)</td>
<td>0.1</td>
<td>0.9753</td>
<td>0.8231</td>
<td>0.9753</td>
</tr>
<tr>
<td>(0.09,-0.11)</td>
<td>0.2</td>
<td>0.9039</td>
<td>0.6547</td>
<td>0.9048</td>
</tr>
<tr>
<td>(0.14,-0.16)</td>
<td>0.3</td>
<td>0.7977</td>
<td>0.5023</td>
<td>0.7985</td>
</tr>
<tr>
<td>(0.19,-0.21)</td>
<td>0.4</td>
<td>0.6697</td>
<td>0.3711</td>
<td>0.6703</td>
</tr>
<tr>
<td>(0.23,-0.27)</td>
<td>0.5</td>
<td>0.5331</td>
<td>0.2636</td>
<td>0.5353</td>
</tr>
<tr>
<td>(0.28,-0.32)</td>
<td>0.6</td>
<td>0.4049</td>
<td>0.1797</td>
<td>0.4066</td>
</tr>
<tr>
<td>(0.33,-0.37)</td>
<td>0.7</td>
<td>0.2926</td>
<td>0.1175</td>
<td>0.2938</td>
</tr>
<tr>
<td>(0.37,-0.43)</td>
<td>0.8</td>
<td>0.2001</td>
<td>0.0736</td>
<td>0.2019</td>
</tr>
<tr>
<td>(0.42,-0.48)</td>
<td>0.9</td>
<td>0.1308</td>
<td>0.0442</td>
<td>0.1320</td>
</tr>
<tr>
<td>(0.47,-0.53)</td>
<td>1.0</td>
<td>0.0813</td>
<td>0.0253</td>
<td>0.0821</td>
</tr>
</tbody>
</table>
### Numerical illustration

<table>
<thead>
<tr>
<th>Observed sample</th>
<th>$\bar{x}_1 - \bar{x}_2$</th>
<th>$H_0 : \mu = 0$ p/s-value</th>
<th>$H'_0 : \mu_1 = \mu_2$ p-value</th>
<th>s-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.05, -0.05)</td>
<td>0.1</td>
<td>0.9753</td>
<td>0.8231</td>
<td>0.9753</td>
</tr>
<tr>
<td>(0.09, -0.11)</td>
<td>0.2</td>
<td>0.9039</td>
<td>0.6547</td>
<td>0.9048</td>
</tr>
<tr>
<td>(0.14, -0.16)</td>
<td>0.3</td>
<td>0.7977</td>
<td>0.5023</td>
<td>0.7985</td>
</tr>
<tr>
<td>(0.19, -0.21)</td>
<td>0.4</td>
<td>0.6697</td>
<td>0.3711</td>
<td>0.6703</td>
</tr>
<tr>
<td>(0.23, -0.27)</td>
<td>0.5</td>
<td>0.5331</td>
<td>0.2636</td>
<td>0.5353</td>
</tr>
<tr>
<td>(0.28, -0.32)</td>
<td>0.6</td>
<td>0.4049</td>
<td>0.1797</td>
<td>0.4066</td>
</tr>
<tr>
<td>(0.33, -0.37)</td>
<td>0.7</td>
<td>0.2926</td>
<td>0.1175</td>
<td>0.2938</td>
</tr>
<tr>
<td>(0.37, -0.43)</td>
<td>0.8</td>
<td>0.2001</td>
<td>0.0736</td>
<td>0.2019</td>
</tr>
<tr>
<td>(0.42, -0.48)</td>
<td>0.9</td>
<td>0.1308</td>
<td>0.0442</td>
<td>0.1320</td>
</tr>
<tr>
<td>(0.47, -0.53)</td>
<td>1.0</td>
<td>0.0813</td>
<td>0.0253</td>
<td>0.0821</td>
</tr>
</tbody>
</table>
## Numerical illustration

<table>
<thead>
<tr>
<th>Observed sample $(\bar{x}_1, \bar{x}_2)$</th>
<th>$\bar{x}_1 - \bar{x}_2$</th>
<th>$H_0: \mu = 0$ p-value</th>
<th>$H'_0: \mu_1 = \mu_2$ p-value</th>
<th>s-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.05, -0.05)</td>
<td>0.1</td>
<td>0.9753</td>
<td>0.8231</td>
<td>0.9753</td>
</tr>
<tr>
<td>(0.09, -0.11)</td>
<td>0.2</td>
<td>0.9039</td>
<td>0.6547</td>
<td>0.9048</td>
</tr>
<tr>
<td>(0.14, -0.16)</td>
<td>0.3</td>
<td><strong>0.7977</strong></td>
<td>0.5023</td>
<td><strong>0.7985</strong></td>
</tr>
<tr>
<td>(0.19, -0.21)</td>
<td>0.4</td>
<td>0.6697</td>
<td>0.3711</td>
<td>0.6703</td>
</tr>
<tr>
<td>(0.23, -0.27)</td>
<td>0.5</td>
<td>0.5331</td>
<td>0.2636</td>
<td>0.5353</td>
</tr>
<tr>
<td>(0.28, -0.32)</td>
<td>0.6</td>
<td>0.4049</td>
<td>0.1797</td>
<td>0.4066</td>
</tr>
<tr>
<td>(0.33, -0.37)</td>
<td>0.7</td>
<td>0.2926</td>
<td>0.1175</td>
<td>0.2938</td>
</tr>
<tr>
<td>(0.37, -0.43)</td>
<td>0.8</td>
<td>0.2001</td>
<td><strong>0.0736</strong></td>
<td>0.2019</td>
</tr>
<tr>
<td>(0.42, -0.48)</td>
<td>0.9</td>
<td>0.1308</td>
<td><strong>0.0442</strong></td>
<td>0.1320</td>
</tr>
<tr>
<td>(0.47, -0.53)</td>
<td>1.0</td>
<td><strong>0.0813</strong></td>
<td><strong>0.0253</strong></td>
<td><strong>0.0821</strong></td>
</tr>
</tbody>
</table>
### Numerical illustration

<table>
<thead>
<tr>
<th>Observed sample</th>
<th>$H_0 : \mu = 0$</th>
<th>$H'_0 : \mu_1 = \mu_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\bar{x}_1, \bar{x}_2)$</td>
<td>$\bar{x}_1 - \bar{x}_2$</td>
<td>p/s-value</td>
</tr>
<tr>
<td>(0.05, -0.05)</td>
<td>0.1</td>
<td>0.9753</td>
</tr>
<tr>
<td>(0.09, -0.11)</td>
<td>0.2</td>
<td>0.9039</td>
</tr>
<tr>
<td>(0.14, -0.16)</td>
<td>0.3</td>
<td>0.7977</td>
</tr>
<tr>
<td>(0.19, -0.21)</td>
<td>0.4</td>
<td>0.6697</td>
</tr>
<tr>
<td>(0.23, -0.27)</td>
<td>0.5</td>
<td>0.5331</td>
</tr>
<tr>
<td>(0.28, -0.32)</td>
<td>0.6</td>
<td>0.4049</td>
</tr>
<tr>
<td>(0.33, -0.37)</td>
<td>0.7</td>
<td>0.2926</td>
</tr>
<tr>
<td>(0.37, -0.43)</td>
<td>0.8</td>
<td>0.2001</td>
</tr>
<tr>
<td>(0.42, -0.48)</td>
<td>0.9</td>
<td>0.1308</td>
</tr>
<tr>
<td>(0.47, -0.53)</td>
<td>1.0</td>
<td><strong>0.0813</strong></td>
</tr>
</tbody>
</table>
Graphical illustration: \( s(\{\mu_1 = \mu_2\}, x_1) = 0.9753 \)
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_2) = 0.9048$
Graphical illustration: \( s(\{\mu_1 = \mu_2\}, x_3) = 0.7985 \)
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_4) = 0.6703$
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_5) = 0.5353$
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_6) = 0.4066$
Graphical illustration: \( s(\{\mu_1 = \mu_2\}, x_7) = 0.2938 \)
Graphical illustration: $s\left(\{\mu_1 = \mu_2\}, x_8\right) = 0.2019$
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_9) = 0.1320$
Graphical illustration: $s\left(\{\mu_1 = \mu_2\}, x_{10}\right) = 0.0821$
Final remarks

The s-value:

- can be applied directly whenever the log-likelihood function is concave by the formula $s = 1 - F(F_{H_0}(1 - p))$

- is a possibilistic measure and can be studied by means of the Abstract belief Calculus ABC (Darwiche, Ginsberg, 1992).

- can be justified by desiderata (more basic axioms).

- avoids the p-value problem of non-monotonicity.

- is a classic alternative to the FBST (Pereira, Stern, 1998).
References:


