On some assumptions of Null Hypothesis Statistical Testing (NHST)

Alexandre Galvão Patriota

Departamento de Estatística
Instituto de Matemática e Estatística
Universidade de São Paulo

Outline

Main goals

The main goals of this presentation are

Main goals

The main goals of this presentation are

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

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,

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

The classical statistical model is:

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

where:

- Ω is the space of possible experiment outcomes,
- $\mathcal{F}$ is a σ-field of Ω,
- $\mathcal{P}$ 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 $(\mathcal{Z}, \mathcal{B})$

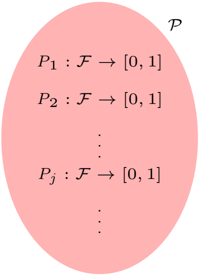
The classical statistical model

The classical statistical model is:

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

where:

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



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

$$g(P) = E_P(Z),$$

$$g(P) = P(Z_1 \in B | Z_2 \in A),$$

etc.

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

A particular model

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

A particular model

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

Conditional, **marginal** and **joint** distributions can be used to make inferences about γ.

A particular model

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

Conditional, **marginal** and **joint** distributions can be used to make inferences about γ.

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**!

Classical null hypotheses

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

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: "at least one measure in $\mathcal{P}_0$ could generate the observed data"
(or simply $\boxed{H_0 : \text{“}P \in \mathcal{P}_0\text{”}}$*)*

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: "at least one measure in $\mathcal{P}_0$ could generate the observed data"
(or simply $\boxed{H_0 : \text{"}P \in \mathcal{P}_0\text{"}}$*)*

Under a parametric model, there exists a finite dimensional set Θ such that:

- $\boxed{\boldsymbol{\mathcal{P} \equiv \{P_\theta : \ \theta \in \Theta\}}}$, where $\Theta \subseteq \mathbb{R}^p$, $p < \infty$,
- $\boxed{\boldsymbol{H_0 : \theta \in \Theta_0}}$, where $\Theta_0 \subset \Theta$ and $\mathcal{P}_0 \equiv \{P_\theta : \ \theta \in \Theta_0\}$.

Alternative hypotheses

According to Fisher, the negation of H_0 **cannot be expressed** in terms of probability measures.

[1] since they would be mutually exclusive and exhaustive

Alternative hypotheses

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 $\mathcal{P}$: $\boxed{H_1 : P \in (\mathcal{P} - \mathcal{P}_0)}$.

[1] since they would be mutually exclusive and exhaustive

Alternative hypotheses

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 $\mathcal{P}$: $\boxed{H_1 : P \in (\mathcal{P} - \mathcal{P}_0)}$.

In the last context, we can **choose**[1] between H_0 and H_1 — Neyman and Pearson approach.

[1] since they would be mutually exclusive and exhaustive

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:

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

Bayesian hypotheses

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

The general null and alternative hypotheses are:

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

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:

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

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.

P-value definition

The p-value for testing the classical null hypothesis H_0 is defined as follows

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

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.

P-value definition

The p-value for testing the classical null hypothesis H_0 is defined as follows

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

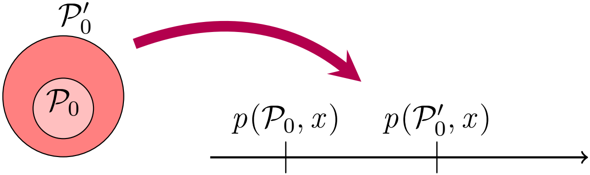
where T_{H_0} is a statistic such that the more discrepant is H_0 from x, the larger is its observed value.[2]

$p(\mathcal{P}_0, x) \approx 0$ indicates that the best case in H_0 provides a small probability to more "extreme events" than the observed one.

[2] i.e., T_{H_0} could be $-2\log$ of the likelihood-ratio statistic.

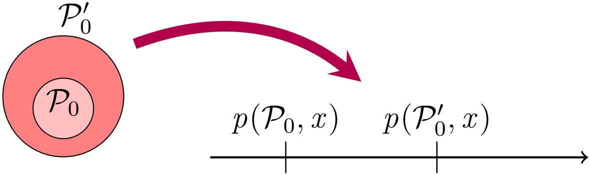
P-value limitations

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-value limitations

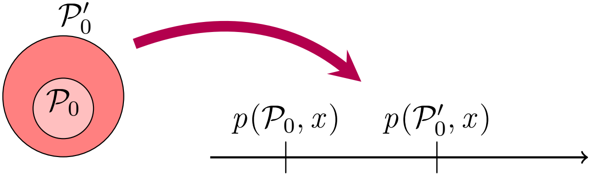
Consider two null hypotheses H_0 : "$P \in \mathcal{P}_0$" and H_0' : "$P \in \mathcal{P}_0'$" such that $H_0 \implies H_0'$. Then, we would expect that:



But it is not always the case!

P-value limitations

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:



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 $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ 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 $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ and identity variance matrix.

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

- under $H_0 : \boldsymbol{\mu} = \mathbf{0}$ 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 $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ and identity variance matrix.

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

- under $H_0 : \boldsymbol{\mu} = \mathbf{0}$ is

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

- under $H_0' : \mu_1 = \mu_2$ is

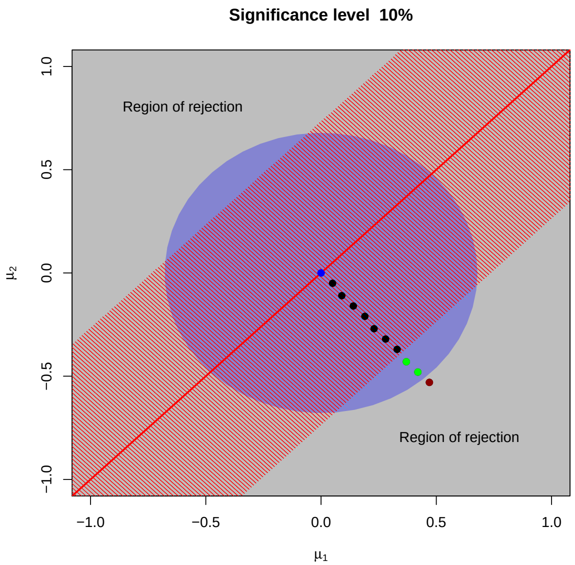
$$T_{H_0'}(X) = \frac{n}{2}(\bar{X}_1 - \bar{X}_2)^2 \sim \chi^2_1,$$

where $\bar{X} = (\bar{X}_1, \bar{X}_2)^\top$ is the maximum likelihood estimator for $\boldsymbol{\mu}$.

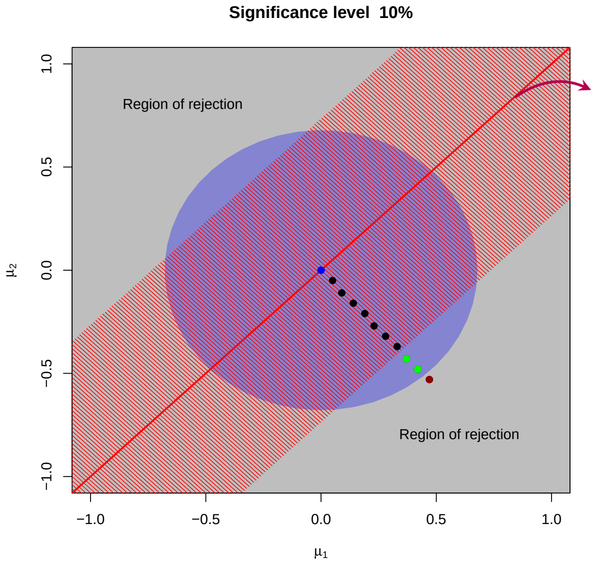
P-values do not respect monotonicity

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

Level curves (contour curves)

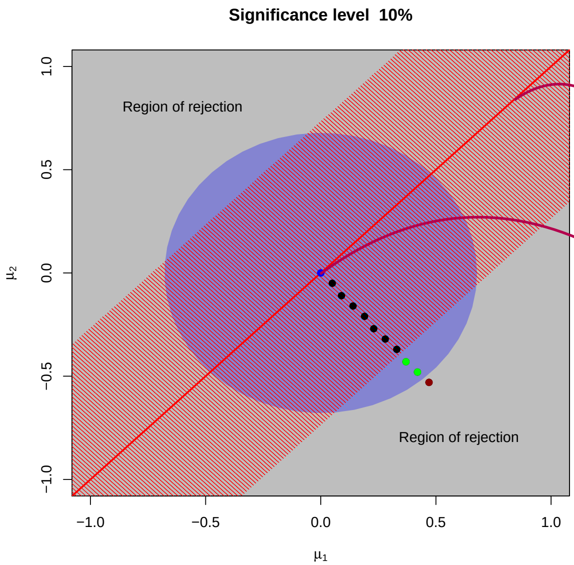


Level curves (contour curves)



$H_0' : \mu_1 = \mu_2$

Level curves (contour curves)

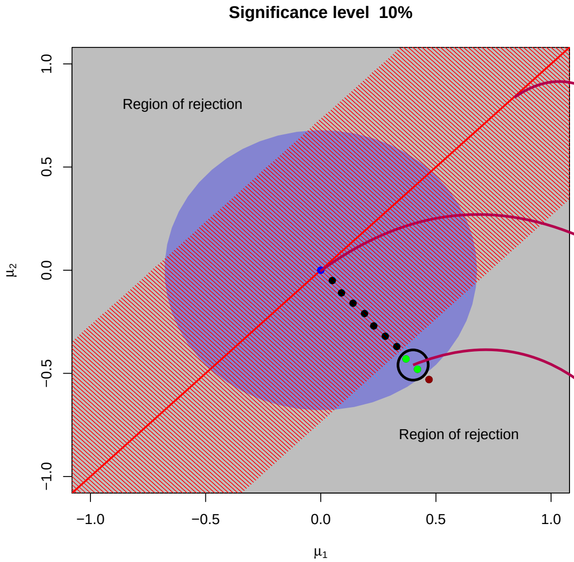


$$H_0' : \mu_1 = \mu_2$$

$$\Uparrow$$

$$H_0 : \boldsymbol{\mu} = \mathbf{0}$$

Level curves (contour curves)



$H_0' : \mu_1 = \mu_2$

$\Uparrow$

$H_0 : \boldsymbol{\mu} = \mathbf{0}$

and

Rejection of H_0' but no rejection of H_0

$\Downarrow$

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).

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 $s : 2^{\Theta} \times \mathcal{X} \to [0, 1]$ such that

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

where Λ_α is a confidence set for θ with confidence level $1 - \alpha$ with some “nice” properties.

Interpretation

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

Interpretation

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

Large values of s indicate that **there exists at least one** element in Θ_0 close to the center of Λ_α (e.g., close to the ML estimate).

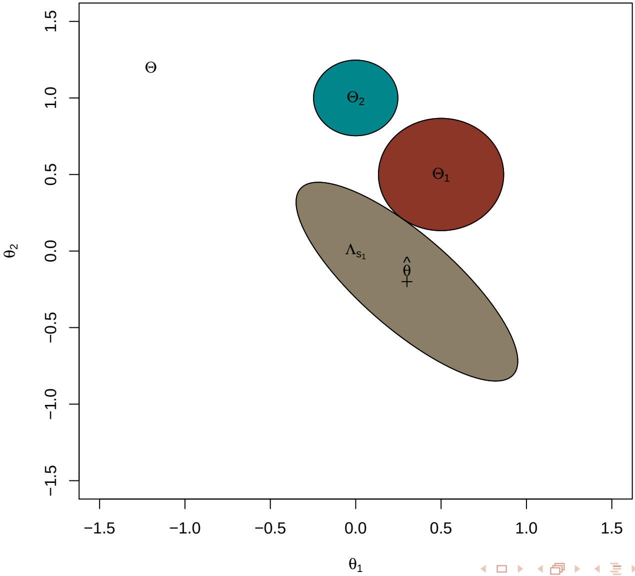
Interpretation

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

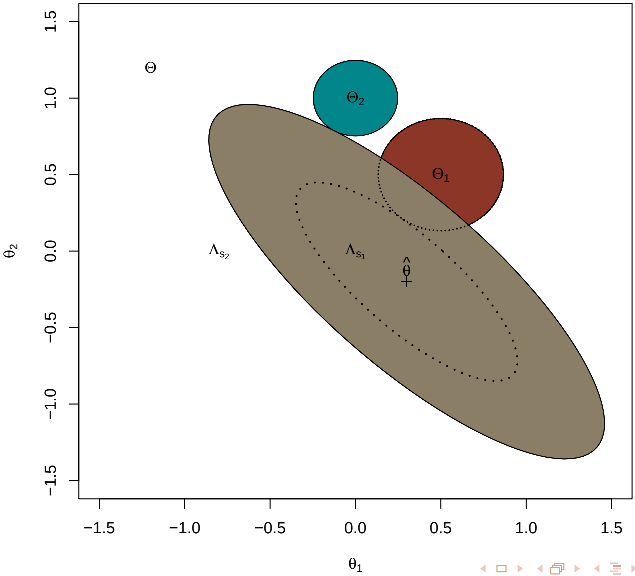
Large values of s indicate that **there exists at least one** element in Θ_0 close to the center of Λ_α (e.g., close to the ML estimate).

Small values of s indicate that **ALL** elements of Θ_0 are far away from the center of Λ_α.

Graphical illustration: $s_1 = s(\Theta_1, x)$



Graphical illustration: $s_2 = s(\Theta_2, x)$



Properties: the s-value is a possibility measure

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

Properties: the s-value is a possibility measure

1. $s(\varnothing, x) = 0$ 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. $s(\varnothing, x) = 0$ 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. $s(\varnothing, x) = 0$ 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. $s(\varnothing, x) = 0$ 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 $\widehat{\theta} \in \overline{\Theta_1}$ (closure of Θ_1), then $s(\Theta_1, x) = 1$,
    - if $\widehat{\theta} \in \overline{\Theta_1^c}$, then $s(\Theta_1^c, x) = 1$.

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

Decisions about H_0

Let Φ be a function such that:

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

Then,

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

Decisions about H_0

Let Φ be a function such that:

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

Then,

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

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

Decisions about H_0

Let Φ be a function such that:

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

Then,

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

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

$\Phi(\Theta_0) = \langle 1, 1 \rangle \implies$ **total ignorance** about H_0.

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

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 *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:

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

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

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

Example: Bivariate Normal distribution

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

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

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

The confidence set Λ_α is given by

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

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

Numerical illustration

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

Numerical illustration

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

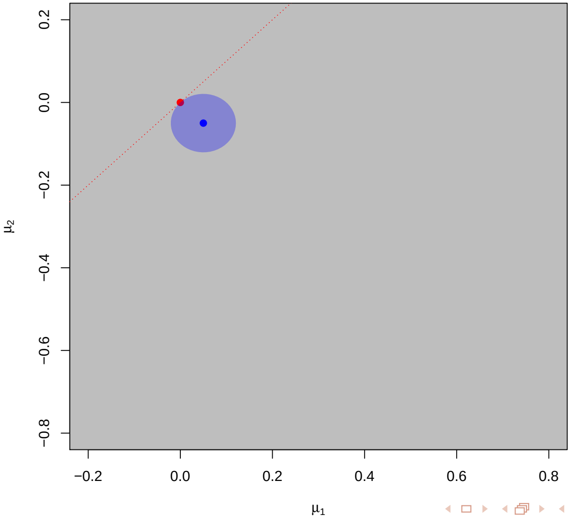
Numerical illustration

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

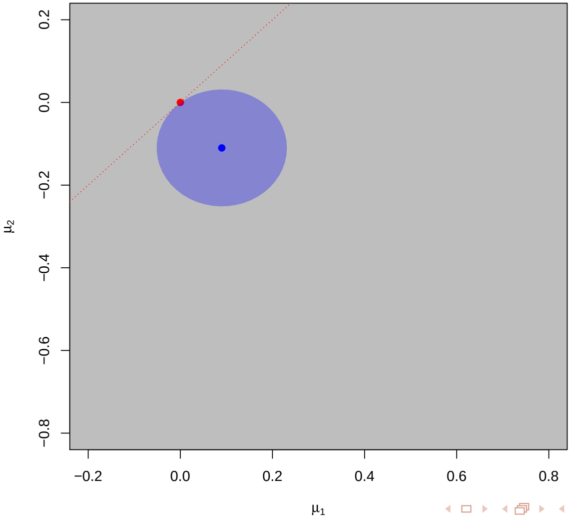
Numerical illustration

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

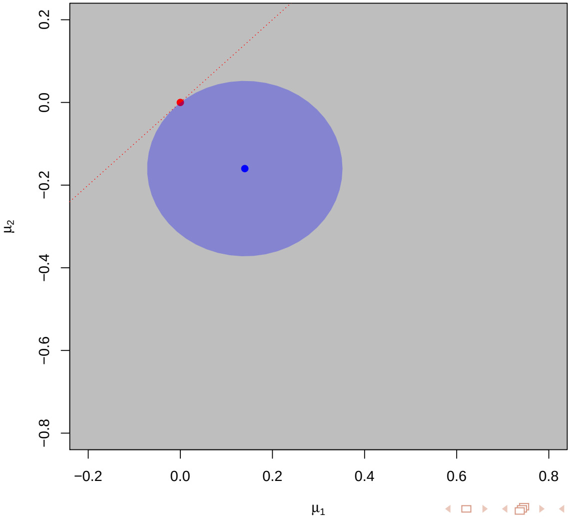
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_1) = 0.9753$



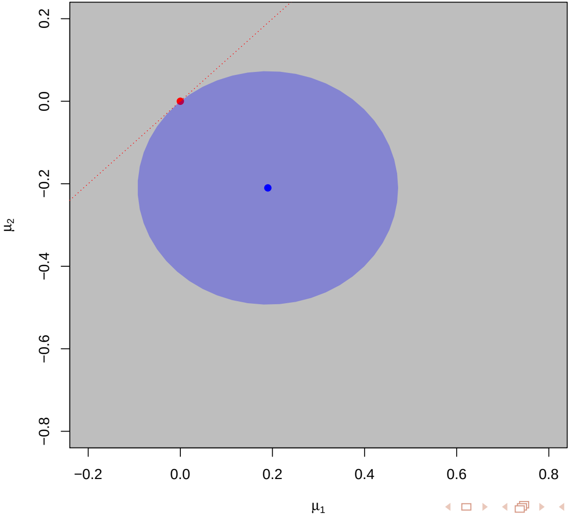
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_2) = 0.9048$



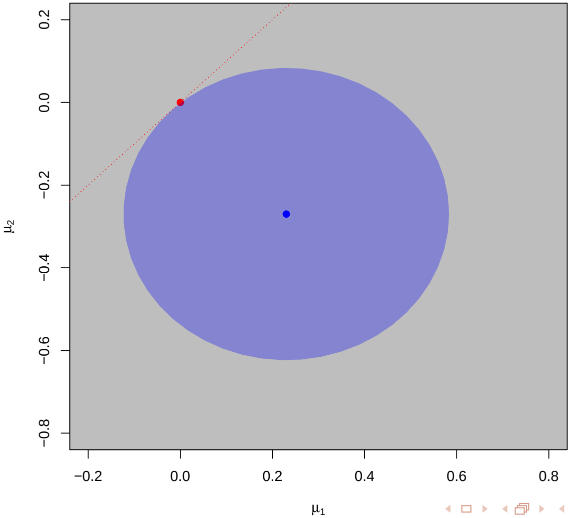
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_3) = 0.7985$



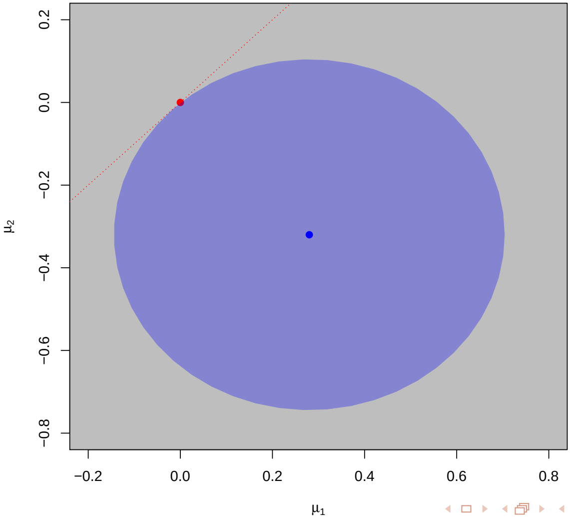
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_4) = 0.6703$



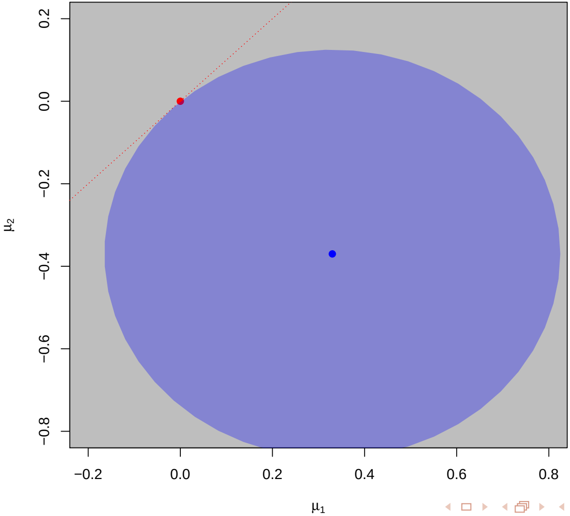
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_5) = 0.5353$



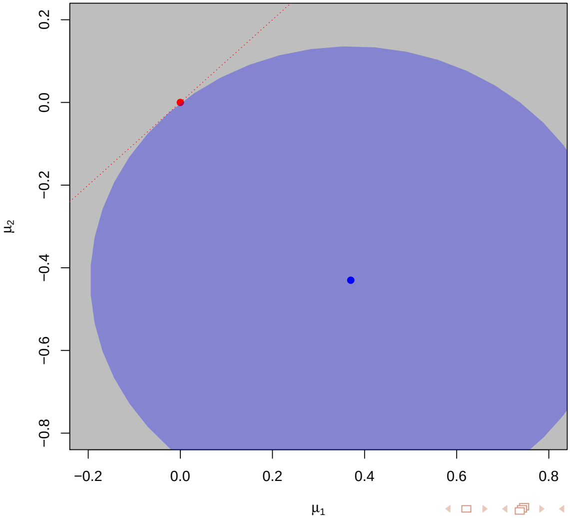
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_6) = 0.4066$



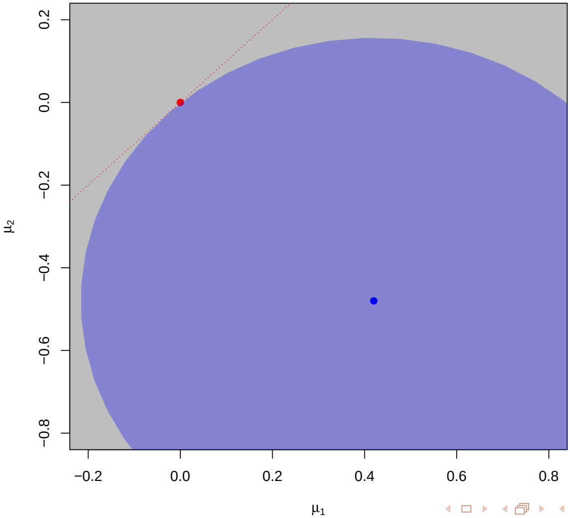
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_7) = 0.2938$



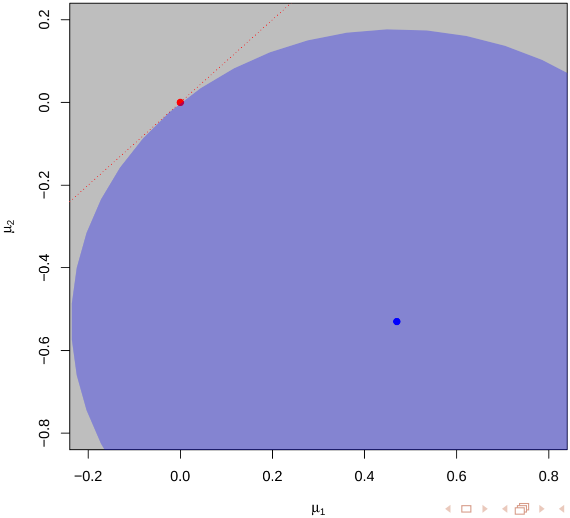
Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_8) = 0.2019$



Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_9) = 0.1320$



Graphical illustration: $s(\{\mu_1 = \mu_2\}, x_{10}) = 0.0821$



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:

Darwiche, A.Y., Ginsberg, M.L. (1992). A symbolic generalization of probability theory, AAAI-92, Tenth National Conference on Artificial Intelligence.

Patriota, AG (2017). On some assumptions of Null Hypothesis Statistical Testing, *Educational and Psychological Measurement*, 77, 507–524.

Patriota, AG (2013). A classical measure of evidence for general null hypotheses, Fuzzy Sets and Systems, 233, 74-88.

Pereira, C.A.B., Stern, J.M. (1999). Evidence and credibility: Full Bayesian significance test for precise hypotheses, Entropy, 1, 99–110.