Uma medida de evidência clássica para testar hipóteses gerais chamada valor-s

Alexandre Galvão Patriota

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

Canal no YouTube:
"A Ciência da Estatística"

Outline

O modelo estatístico clássico

O modelo estatístico clássico

O modelo estatístico clássico é definido pela trinca

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

em que:

- Ω é o espaço de **resultados** do experimento,
- $\mathcal{F}$ é uma σ-álgebra de Ω,
- $\mathcal{P}$ é uma **família** de medidas de probabilidade.

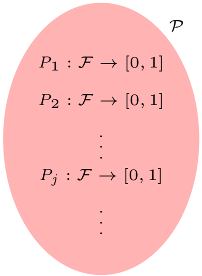
O modelo estatístico clássico

O modelo estatístico clássico é definido pela trinca

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

em que:

- Ω é o espaço de **resultados** do experimento,
- $\mathcal{F}$ é uma σ-álgebra de Ω,
- $\mathcal{P}$ é uma **família** de medidas de probabilidade.



Quantidades de interesse:

$$\begin{aligned} g(P) &= E_P(Z) \\ g(P) &= P(Z_1 \in B | Z_2 \in A) \end{aligned}$$

O modelo estatístico clássico

O modelo estatístico clássico é definido pela trinca

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

em que:

- Ω é o espaço de **resultados** do experimento,
- $\mathcal{F}$ é uma σ-álgebra de Ω,
- $\mathcal{P}$ é uma **família** de medidas de probabilidade.

$\mathcal{P}$

$$P_1 : \mathcal{F} \to [0, 1]$$
$$P_2 : \mathcal{F} \to [0, 1]$$
$$\vdots$$
$$P_j : \mathcal{F} \to [0, 1]$$
$$\vdots$$

Quantidades de interesse:

$$\begin{aligned} g(P) &= E_P(Z) \\ g(P) &= P(Z_1 \in B | Z_2 \in A) \end{aligned}$$

Observação: um vetor aleatório $Z : \Omega \to \mathbb{R}^s$ é uma função mensurável

Um modelo particular

Um modelo particular

Tome $Z = (X, \gamma)$, em que X é a amostra aleatória (**observável**) e γ é o vetor de variáveis aleatórias (**não-observável**).

Um modelo particular

Tome $Z = (X, \gamma)$, em que X é a amostra aleatória (**observável**) e γ é o vetor de variáveis aleatórias (**não-observável**).

As distribuições **Conditionais**, **marginais** e **conjuntas** podem ser usadas para descrever probabilisticamente os valores de γ.

Um modelo particular

Tome $Z = (X, \gamma)$, em que X é a amostra aleatória (**observável**) e γ é o vetor de variáveis aleatórias (**não-observável**).

As distribuições **Conditionais**, **marginais** e **conjuntas** podem ser usadas para descrever probabilisticamente os valores de γ.

Defina $\mathcal{P} = \{P_0\}$ e construa a sua conjunta P_0 induzida por Z definindo:

- $\gamma \sim f_0(\cdot)$,
- $X|\gamma \sim f_1(\cdot|\gamma)$

Você está pronto para ser um **Bayesiano do núcleo Duro**!

O modelo estatístico nos cursos de graduação

O modelo estatístico nos cursos de graduação

Em geral, o modelo estatístico é enunciado da seguinte forma:

O modelo estatístico nos cursos de graduação

Em geral, o modelo estatístico é enunciado da seguinte forma:

"Sejam $X_1, \ldots, X_n$ v.a.'s com função (densidade) de probabilidade $f_\theta(x_1, \ldots, x_n)$, $\theta \in \Theta$"

O modelo estatístico nos cursos de graduação

Em geral, o modelo estatístico é enunciado da seguinte forma:

"Sejam $X_1, \dots, X_n$ v.a.'s com função (densidade) de probabilidade $f_\theta(x_1, \dots, x_n)$, $\theta \in \Theta$"

Esse enunciado pode ser escrito em termos da trinca:

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

em que $Z = (X_1, \dots, X_n)$ é o vetor mensurável e $\mathcal{P} = \{P_\theta : \ \theta \in \Theta\}$

Testes de hipóteses

Hipótese nula

Objetivo: verificar se podemos

1. reduzir $\mathcal{P}$ a uma subfamília $\mathcal{P}_0 \subset \mathcal{P}$,
2. reduzir o espaço paramétrico Θ a um subconjunto $\Theta_0 \subset \Theta$

Hipótese nula

Objetivo: verificar se podemos

1. reduzir $\mathcal{P}$ a uma subfamília $\mathcal{P}_0 \subset \mathcal{P}$,
2. reduzir o espaço paramétrico Θ a um subconjunto $\Theta_0 \subset \Theta$

Hipótese nula:

H_0: *"pelo menos uma medida em $\mathcal{P}_0$ descreve probabilisticamente os dados experimentais"* $\Leftrightarrow$ $\boxed{H_0 : \text{“}P \in \mathcal{P}_0\text{”}}$

Hipótese nula

Objetivo: verificar se podemos

1. reduzir $\mathcal{P}$ a uma subfamília $\mathcal{P}_0 \subset \mathcal{P}$,
2. reduzir o espaço paramétrico Θ a um subconjunto $\Theta_0 \subset \Theta$

Hipótese nula:

H_0: "pelo menos uma medida em $\mathcal{P}_0$ descreve probabilisticamente os dados experimentais" $\Leftrightarrow$ $\boxed{H_0 : \text{“}P \in \mathcal{P}_0\text{”}}$

Em termos do vetor de parâmetros

- $\boxed{H_0 : \theta \in \Theta_0}$, em que $\Theta_0 \subset \Theta$ e $\mathcal{P}_0 = \{P_\theta : \ \theta \in \Theta_0\}$.

Hipótese alternativa

Neyman e Pearson definiram uma suposta "hipótese alternativa" para encontrar o teste mais poderoso.

Hipótese alternativa

Neyman e Pearson definiram uma suposta "hipótese alternativa" para encontrar o teste mais poderoso.

Fisher criticou: uma hipótese alternativa H_1 só faria sentido no caso em que temos **CERTEZA** que $\mathcal{P}$ inclui a medida que gera os dados.

Hipótese alternativa

Neyman e Pearson definiram uma suposta “hipótese alternativa” para encontrar o teste mais poderoso.

Fisher criticou: uma hipótese alternativa H_1 só faria sentido no caso em que temos **CERTEZA** que $\mathcal{P}$ inclui a medida que gera os dados.

- Para Neyman e Pearson: podemos decidir entre H_0 e H_1.

Hipótese alternativa

Neyman e Pearson definiram uma suposta “hipótese alternativa” para encontrar o teste mais poderoso.

Fisher criticou: uma hipótese alternativa H_1 só faria sentido no caso em que temos **CERTEZA** que $\mathcal{P}$ inclui a medida que gera os dados.

- Para Neyman e Pearson: podemos decidir entre H_0 e H_1.
- Para Fisher: podemos apenas rejeitar ou não rejeitar H_0.

Hipótese alternativa

Neyman e Pearson definiram uma suposta "hipótese alternativa" para encontrar o teste mais poderoso.

Fisher criticou: uma hipótese alternativa H_1 só faria sentido no caso em que temos **CERTEZA** que $\mathcal{P}$ inclui a medida que gera os dados.

- Para Neyman e Pearson: podemos decidir entre H_0 e H_1.
- Para Fisher: podemos apenas rejeitar ou não rejeitar H_0.
- Hoje em dia as abordagens de Fisher e Neyman-Pearson são misturadas (define-se H_1, mas conclui-se pela rejeição ou não de H_0).

Definição do valor-p

Definição do valor-p

Seja T_{H_0} uma estatística tal que quanto maior for a discrepância entre H_0 e x maior é o seu valor observado[1]

[1] i.e., T_{H_0} pode ser $-2\log RV$.

Definição do valor-p

Seja T_{H_0} uma estatística tal que quanto maior for a discrepância entre H_0 e x maior é o seu valor observado[1]

O valor-p (para $\Theta_0 \neq \varnothing$):

$$p_x(\Theta_0) = \sup_{\theta \in \Theta_0} P_\theta\Big(T_{H_0}(X) \geq T_{H_0}(x)\Big)$$

[1] i.e., T_{H_0} pode ser $-2 \log RV$.

Definição do valor-p

Seja T_{H_0} uma estatística tal que quanto maior for a discrepância entre H_0 e x maior é o seu valor observado[1]

O valor-p (para $\Theta_0 \neq \varnothing$):

$$p_x(\Theta_0) = \sup_{\theta \in \Theta_0} P_\theta\Big(T_{H_0}(X) \geq T_{H_0}(x)\Big)$$

$p_x(\Theta_0) \approx 0$ indica que:

[1] i.e., T_{H_0} pode ser $-2\log RV$.

Definição do valor-p

Seja T_{H_0} uma estatística tal que quanto maior for a discrepância entre H_0 e x maior é o seu valor observado[1]

O valor-p (para $\Theta_0 \neq \varnothing$):

$$p_x(\Theta_0) = \sup_{\theta \in \Theta_0} P_\theta\big(T_{H_0}(X) \geq T_{H_0}(x)\big)$$

$p_x(\Theta_0) \approx 0$ indica que:

- mesmo sob o melhor caso em H_0, os eventos tão ou mais extremos quanto o observado são raros.

[1] i.e., T_{H_0} pode ser $-2\log RV$.

Definição do valor-p

Seja T_{H_0} uma estatística tal que quanto maior for a discrepância entre H_0 e x maior é o seu valor observado[1]

O valor-p (para $\Theta_0 \neq \varnothing$):

$$\boxed{p_x(\Theta_0) = \sup_{\theta \in \Theta_0} P_\theta\big(T_{H_0}(X) \geq T_{H_0}(x)\big)}$$

$p_x(\Theta_0) \approx 0$ indica que:

- mesmo sob o melhor caso em H_0, os eventos tão ou mais extremos quanto o observado são raros.
- como o evento ocorreu: "A hipótese nula deve ser falsa ou um evento raro ocorreu" (Modus Tollens Fisheriano).

[1] i.e., T_{H_0} pode ser $-2 \log RV$.

Valor-p não é medida monotônica

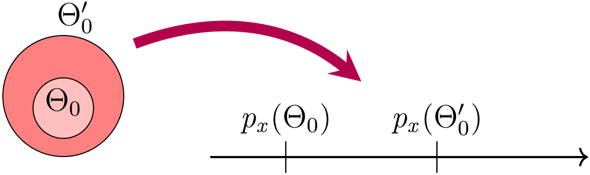
Limitações do valor-p

Considere duas hipóteses nulas H_0 : "$\theta \in \Theta_0$" e H_0' : "$\theta \in \Theta_0'$" tais que $H_0 \implies H_0'$, Ou seja, $\Theta_0 \subset \Theta_0'$.

Limitações do valor-p

Considere duas hipóteses nulas H_0 : "$\theta \in \Theta_0$" e H_0' : "$\theta \in \Theta_0'$" tais que $H_0 \implies H_0'$, Ou seja, $\Theta_0 \subset \Theta_0'$.

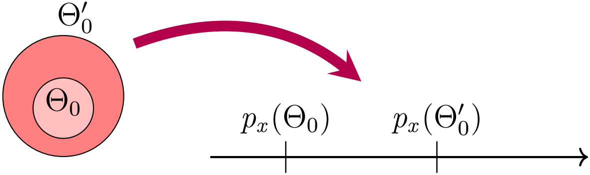
Esperamos que



Limitações do valor-p

Considere duas hipóteses nulas H_0 : "$\theta \in \Theta_0$" e H_0' : "$\theta \in \Theta_0'$" tais que $H_0 \implies H_0'$, Ou seja, $\Theta_0 \subset \Theta_0'$.

Esperamos que



Problema: isso nem sempre ocorre: **valor-p não é uma medida monotônica**.

Exemplo numérico

Normal bivariada

Seja $X = (X_1, \dots, X_n)$ uma amostra aleatória de uma normal bivariada com média $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ e matriz de variância-covariância identidade.

Normal bivariada

Seja $X = (X_1, \ldots, X_n)$ uma amostra aleatória de uma normal bivariada com média $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ e matriz de variância-covariância identidade.

Considere H_0 : "$\boldsymbol{\mu} = \mathbf{0}$" e H_0' : "$\mu_1 = \mu_2$"

Normal bivariada

Seja $X = (X_1, \ldots, X_n)$ uma amostra aleatória de uma normal bivariada com média $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ e matriz de variância-covariância identidade.

Considere H_0 : "$\boldsymbol{\mu} = \mathbf{0}$" e H_0' : "$\mu_1 = \mu_2$"

Observe que $H_0 \implies H_0'$,

pois se $\boldsymbol{\mu} = \mathbf{0}$, então $\mu_1 = \mu_2$.

Normal bivariada

A distribuição de $(-2\log)$ estatística da razão de verossimilhanças:

- sob $H_0 : \boldsymbol{\mu} = \mathbf{0}$ é

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

em que $\bar{X} = (\bar{X}_1, \bar{X}_2)^\top$ é o EMV de $\boldsymbol{\mu}$.

Normal bivariada

A distribuição de $(-2\log)$ estatística da razão de verossimilhanças:

- sob $H_0 : \boldsymbol{\mu} = \mathbf{0}$ é

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

em que $\bar{X} = (\bar{X}_1, \bar{X}_2)^\top$ é o EMV de $\boldsymbol{\mu}$.

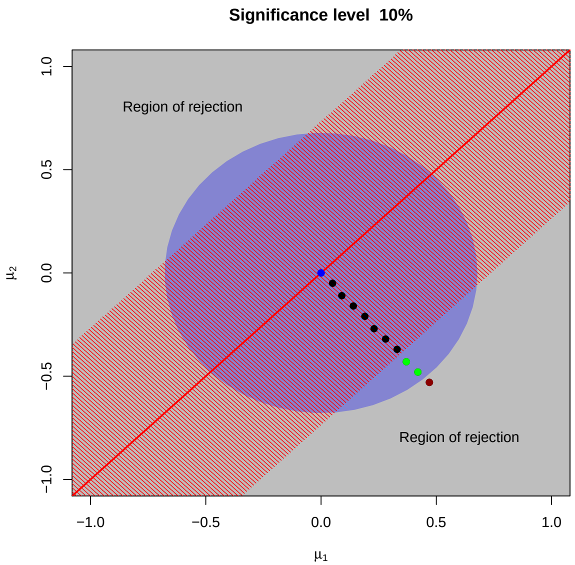
- sob $H_0' : \mu_1 = \mu_2$ é

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

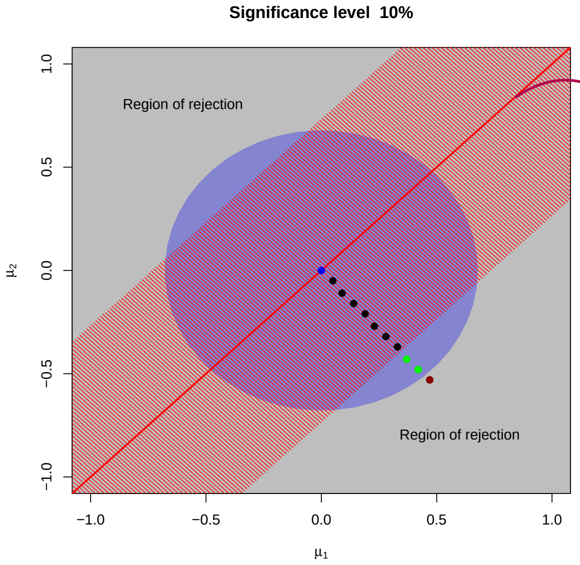
Valores observados $n = 10$

Amostra observada $(\bar{x}_1, \bar{x}_2)$	$\bar{x}_1 - \bar{x}_2$	$H_0 : \boldsymbol{\mu} = \mathbf{0}$ valor-p	$H_0' : \mu_1 = \mu_2$ valor-p
(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

Curvas de nível

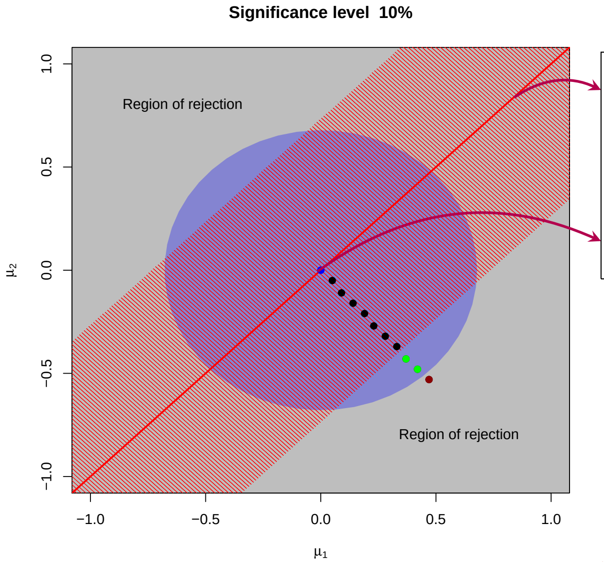


Curvas de nível



$H_0' : \mu_1 = \mu_2$

Curvas de nível

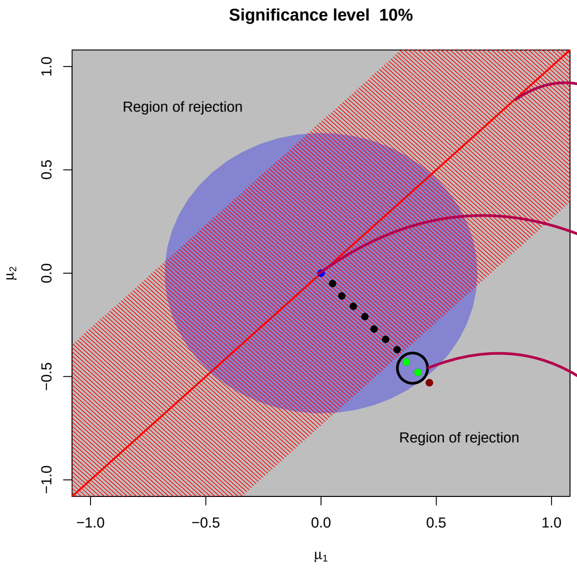


$$H_0' : \mu_1 = \mu_2$$

$$\Uparrow$$

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

Curvas de nível



$H_0' : \mu_1 = \mu_2$

$\Uparrow$

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

e

Rejeição de H_0' mas não-rejeição de H_0

$\Downarrow$

Conclusão incoerente

Uma medida de evidência alternativa

Uma medida de evidência alternativa

Seja Λ_α um conjunto de confiança para θ com coeficiente de confiança de $1 - \alpha$ (nível de significância α) com algumas propriedades “boas”.

Uma medida de evidência alternativa

Seja Λ_α um conjunto de confiança para θ com coeficiente de confiança de $1 - \alpha$ (nível de significância α) com algumas propriedades "boas".

O **valor-s** é uma função $s_x : 2^\Theta \to [0, 1]$ tal que

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

para cada amostra observada x.

Uma medida de evidência alternativa

Seja Λ_α um conjunto de confiança para θ com coeficiente de confiança de $1 - \alpha$ (nível de significância α) com algumas propriedades "boas".

O **valor-s** é uma função $s_x : 2^\Theta \to [0, 1]$ tal que

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

para cada amostra observada x.

$s = s_x(\Theta_0)$ é o maior nível de significância (ou $1 - s$ é o menor coeficiente de confiança) para o qual o conjunto de confiança e o conjunto $\overline{\Theta_0}$ tenha ao menos um ponto em comum.

Ilustração gráfica: $s_1 = s_x(\Theta_1)$

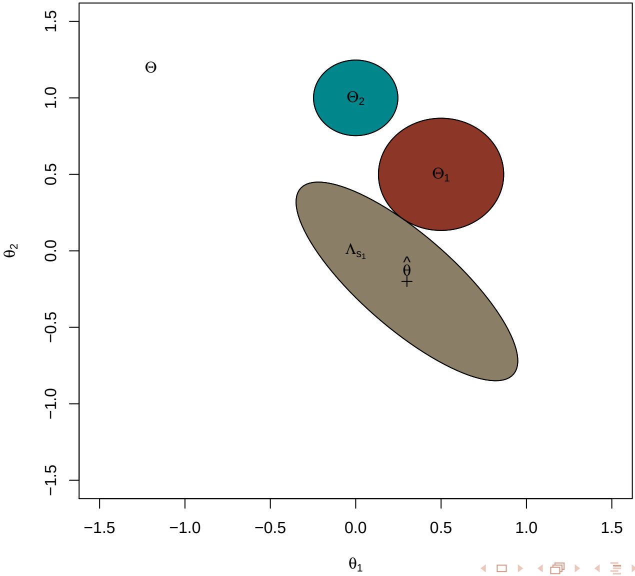
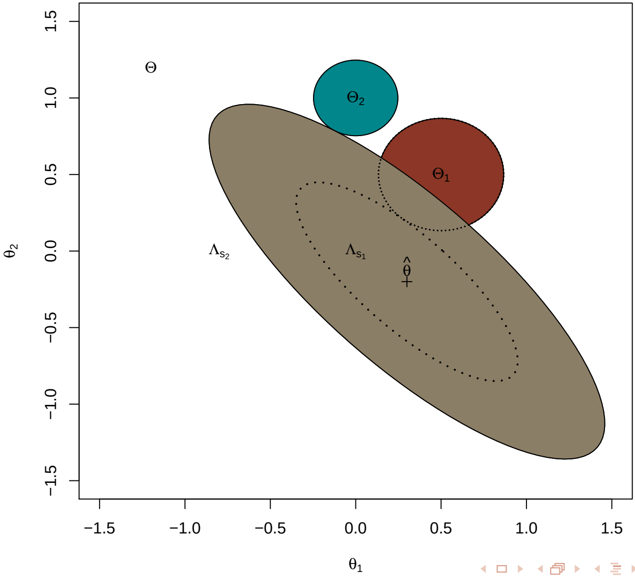


Ilustração gráfica: $s_2 = s_x(\Theta_2)$



Interpretação

Valores de s próximos de 1 indicam que **existe pelo menos um** elemento em Θ_0 perto do centro de Λ_α (e.g., perto da EMV).

Interpretação

Valores de s próximos de 1 indicam que **existe pelo menos um** elemento em Θ_0 perto do centro de Λ_α (e.g., perto da EMV).

Valores de s próximos de 0 indicam que **TODOS** os elementos de Θ_0 estão longe do centro de Λ_α.

Propriedades: o valor-s é uma medida de possibilidade

1. $s_x(\varnothing) = 0$ e $s_x(\Theta) = 1$,

Propriedades: o valor-s é uma medida de possibilidade

1. $s_x(\varnothing) = 0$ e $s_x(\Theta) = 1$,
2. Se $\Theta_1 \subseteq \Theta_2$, então $s_x(\Theta_1) \leq s_x(\Theta_2)$,

Propriedades: o valor-s é uma medida de possibilidade

1. $s_x(\varnothing) = 0$ e $s_x(\Theta) = 1$,
2. Se $\Theta_1 \subseteq \Theta_2$, então $s_x(\Theta_1) \leq s_x(\Theta_2)$,
3. Para quaisquer $\Theta_1, \Theta_2 \subseteq \Theta$,
   $s_x(\Theta_1 \cup \Theta_2) = \max\{s_x(\Theta_1), s_x(\Theta_2)\}$,

Propriedades: o valor-s é uma medida de possibilidade

1. $s_x(\varnothing) = 0$ e $s_x(\Theta) = 1$,
2. Se $\Theta_1 \subseteq \Theta_2$, então $s_x(\Theta_1) \leq s_x(\Theta_2)$,
3. Para quaisquer $\Theta_1, \Theta_2 \subseteq \Theta$, $s_x(\Theta_1 \cup \Theta_2) = \max\{s_x(\Theta_1), s_x(\Theta_2)\}$,
4. Para qualquer $\Theta_1 \subseteq \Theta$, $s_x(\Theta_1) = \sup_{\theta \in \Theta_1} s_x(\{\theta\})$,

Propriedades: o valor-s é uma medida de possibilidade

1. $s_x(\varnothing) = 0$ e $s_x(\Theta) = 1$,
2. Se $\Theta_1 \subseteq \Theta_2$, então $s_x(\Theta_1) \leq s_x(\Theta_2)$,
3. Para quaisquer $\Theta_1, \Theta_2 \subseteq \Theta$,
   $s_x(\Theta_1 \cup \Theta_2) = \max\{s_x(\Theta_1), s_x(\Theta_2)\}$,
4. Para qualquer $\Theta_1 \subseteq \Theta$, $s_x(\Theta_1) = \sup_{\theta \in \Theta_1} s_x(\{\theta\})$,
5. $s_x(\Theta_1) = 1$ ou $s_x(\Theta_1^c) = 1$.

Ilustração numérica

Normal bivariada

Seja $X = (X_1, \ldots, X_n)$ uma amostra aleatória de uma normal bivariada com média $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ e matriz de variâncias-covariâncias identidade.

Normal bivariada

Seja $X = (X_1, \ldots, X_n)$ uma amostra aleatória de uma normal bivariada com média $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ e matriz de variâncias-covariâncias identidade.

Para cada, $\boldsymbol{\mu} \in \mathbb{R}^2$ fixado, temos que

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

Normal bivariada

Seja $X = (X_1, \ldots, X_n)$ uma amostra aleatória de uma normal bivariada com média $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ e matriz de variâncias-covariâncias identidade.

Para cada, $\boldsymbol{\mu} \in \mathbb{R}^2$ fixado, temos que

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

O conjunto de confiança Λ_α pode ser definido como

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

em que F_2 é a função distribuição acumulada da quiquadrado com 2 graus de liberdade.

Cálculo do valor-s

Para o exemplo, cosneguimos calcular o valor-s da seguinte forma:

Cálculo do valor-s

Para o exemplo, cosneguimos calcular o valor-s da seguinte forma:

Se $\Theta_0 \subset \mathbb{R}^2$ é não-vazio, então

$$s_x(\Theta_0) = 1 - F_2\big(\inf_{\boldsymbol{\mu} \in \Theta_0} T_{\boldsymbol{\mu}}(x)\big)$$

em que F_2 é a função distribuição acumulada da quiquadrado com 2 graus de liberdade.

Ilustração numérica

Amostra observada $(\bar{x}_1, \bar{x}_2)$	$\bar{x}_1 - \bar{x}_2$	$H_0 : \boldsymbol{\mu} = \mathbf{0}$ valor-p=valor-s	$H_0' : \mu_1 = \mu_2$ valor-p	valor-s
(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

Ilustração numérica

Amostra observada		$H_0 : \boldsymbol{\mu} = \mathbf{0}$	$H_0' : \mu_1 = \mu_2$	
$(\bar{x}_1, \bar{x}_2)$	$\bar{x}_1 - \bar{x}_2$	valor-p=valor-s	valor-p	valor-s
(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

Ilustração numérica

Amostra observada		$H_0 : \boldsymbol{\mu} = \mathbf{0}$	$H_0' : \mu_1 = \mu_2$	
$(\bar{x}_1, \bar{x}_2)$	$\bar{x}_1 - \bar{x}_2$	valor-p=valor-s	valor-p	valor-s
(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

Ilustração numérica

Amostra observada		$H_0 : \boldsymbol{\mu} = \mathbf{0}$	$H_0' : \mu_1 = \mu_2$	
$(\bar{x}_1, \bar{x}_2)$	$\bar{x}_1 - \bar{x}_2$	valor-p=valor-s	valor-p	valor-s
(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

Ilustração gráfica: $s_{x_1}(\{\mu_1 = \mu_2\}) = 0.9753$

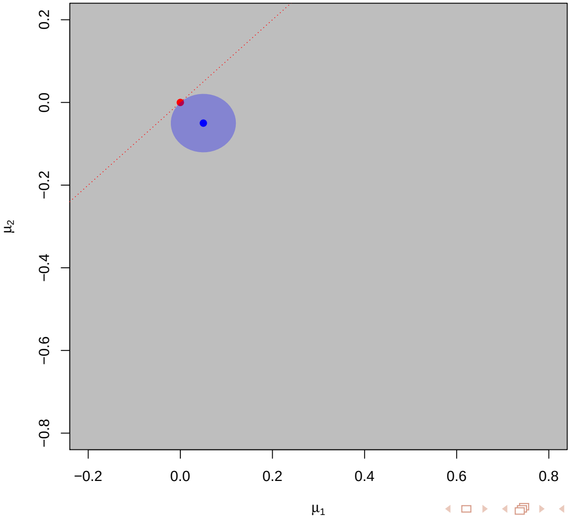


Ilustração gráfica: $s_{x_2}(\{\mu_1 = \mu_2\}) = 0.9048$

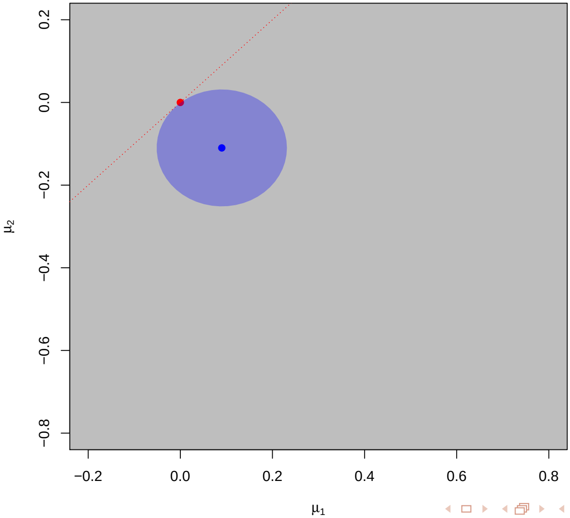


Ilustração gráfica: $s_{x_3}(\{\mu_1 = \mu_2\}) = 0.7985$

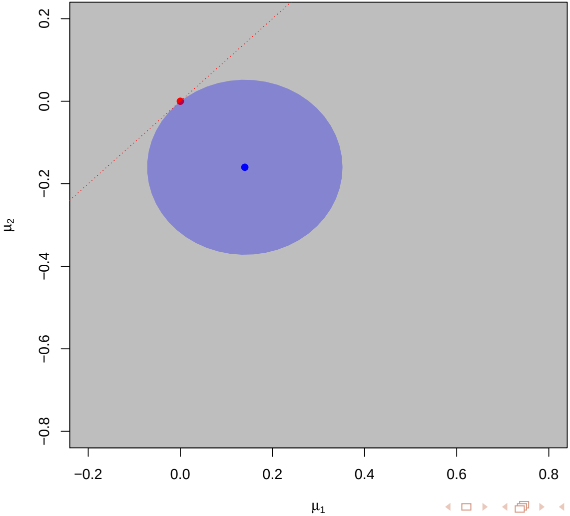


Ilustração gráfica: $s_{x_4}(\{\mu_1 = \mu_2\}) = 0.6703$

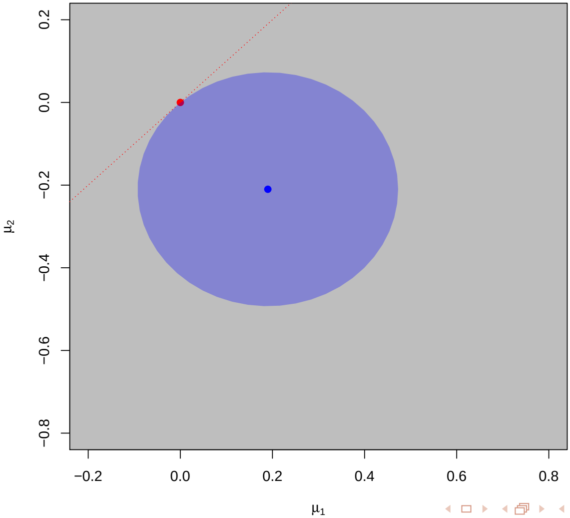


Ilustração gráfica: $s_{x_5}(\{\mu_1 = \mu_2\}) = 0.5353$

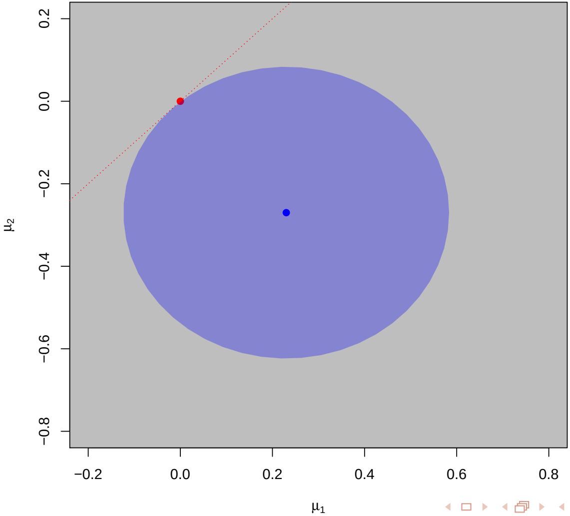


Ilustração gráfica: $s_{x_6}(\{\mu_1 = \mu_2\}) = 0.4066$

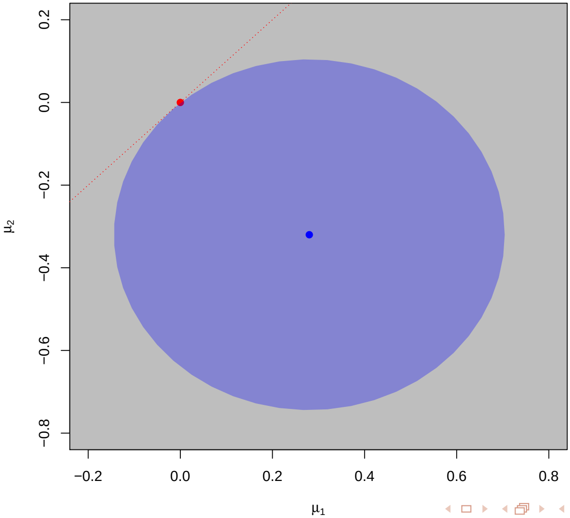


Ilustração gráfica: $s_{x_7}(\{\mu_1 = \mu_2\}) = 0.2938$

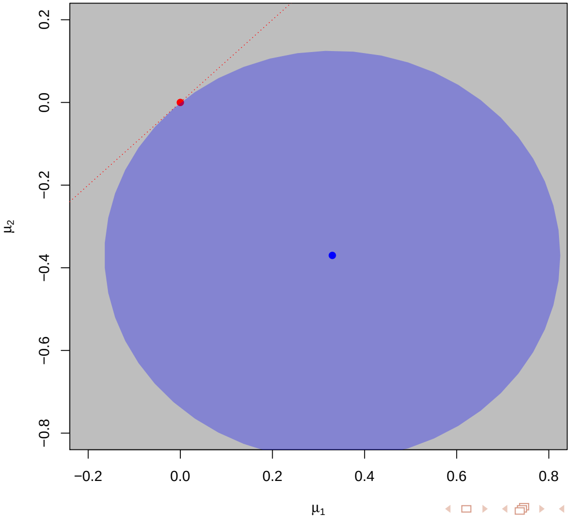


Ilustração gráfica: $s_{x_8}(\{\mu_1 = \mu_2\}) = 0.2019$

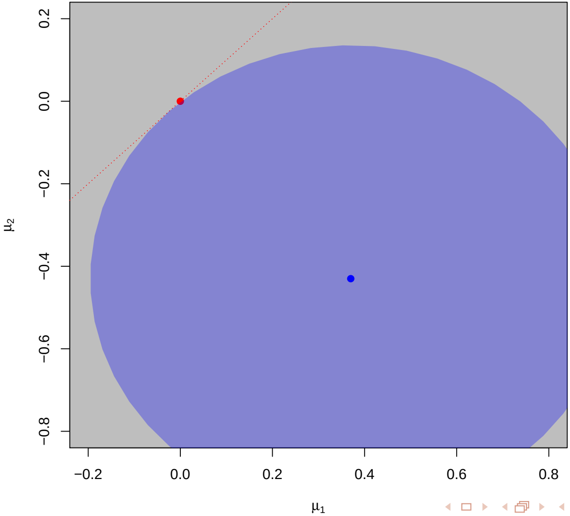


Ilustração gráfica: $s_{x_9}(\{\mu_1 = \mu_2\}) = 0.1320$

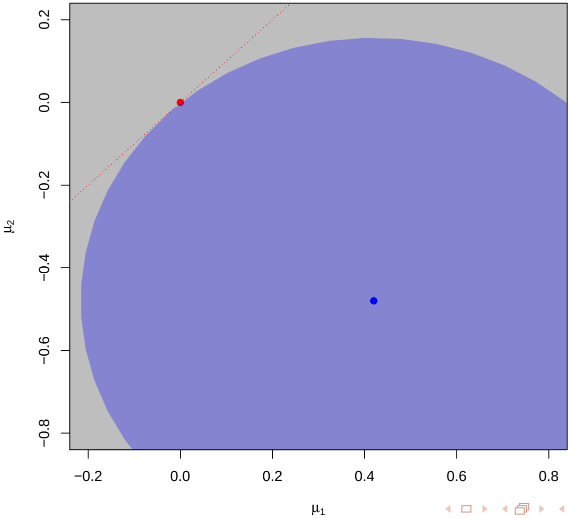
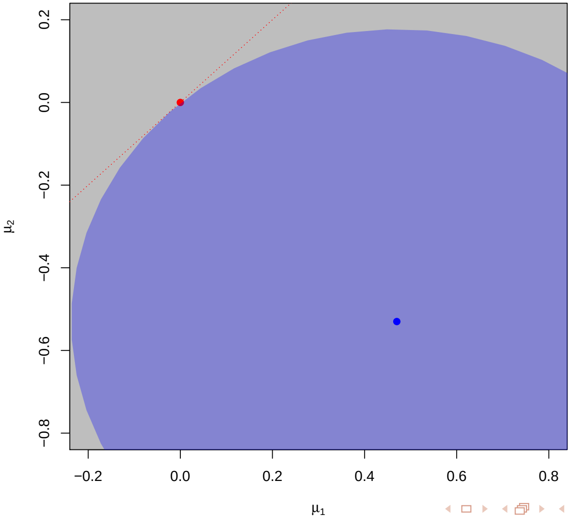


Ilustração gráfica: $s_{x_{10}}(\{\mu_1 = \mu_2\}) = 0.0821$



Palavras finais

Palavras finais

O valor-s:

- pode ser aplicado diretamente sempre que a função de log-verossimilhança for concava: $s_x(\Theta) = 1 - F(F_{H_0}(1 - p_x(\Theta_0)))$,
- é uma medida de possibilidade e pode ser justificado por *desiderata* (axiomas mais básicos),
- evita o problema de monotonicidade do valor-p,
- tem nível de significância α:

$$\sup_{\theta \in \Theta_0} P_\theta(s_X(\Theta_0) \leq \alpha) \leq \alpha.$$

Referências

Referências:

Neyman, J, Pearson, E. (1933). On the problem of the most efficient testsof statistical hypotheses, Philosophical Transactions of the Royal Society of London., A, 289–337.

Fisher, R. (1955). Statistical Methods and Scientific Induction, Journal of the Royal Statistical Society: Series B (Methodological), 17(1), 69–78

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

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

Patriota, AG (2018). Is NHST logically flawed? Commentary on: "NHST is still logically flawed", *Scientometrics*, 116, 2189–2191.

OBRIGADO