Teste de Hipóteses para Tomada de Decisão

Magalhães e Lima - Seções 8.2 e 8.4
Montgomery et al – Seções 4.3 e 4.4

Objetivos de Aprendizagem

Os alunos devem ser capazes de:

- Construir hipóteses em termos do problema e em termos estatísticos;
- Construir a estatística do teste;
- Concluir o teste de hipóteses via região crítica ou via valor-p.

Acompanhe, previamente, o PLANO DE AULA no BLACKBOARD!

Tomada de decisão – Aula 19:

Comprar ou não comprar um lote de resistores de um determinado fabricante?

Sua empresa está desenvolvendo um projeto de larga escala e o seu departamento ficou responsável por indicar um fabricante de resistores para o projeto.

Serão compradas centenas de milhões de unidades para o produto final e ficaria inviável medir todos os componentes, garantindo assim um mínimo de qualidade.

Para a tomada de decisão, um fabricante lhe enviou uma pequena amostra do produto. Assuma que a tolerância do resistor indica o seu desvio padrão $\sigma = 50\ k\Omega$ (5% de $1000\ k\Omega$).

Teste de Hipóteses

O procedimento básico de um teste de hipóteses consiste em **supor verdadeira uma das hipóteses** em questão e verificar se a amostra observada leva à rejeição ou não desta hipótese, ou seja, verificar se os dados coletados trazem evidências a favor ou não de uma hipótese formulada.

Teste de Hipóteses

A teoria do teste de hipóteses fornece métodos para a tomada de decisão a respeito de hipóteses formuladas, informando também a probabilidade de erro que acompanha a decisão.

O erro de decisão não pode ser evitado, mas sua probabilidade de ocorrência pode ser controlada ou mensurada, obtendo-se uma medida de validade das conclusões obtidas.

Chamamos de α, **nível de significância**, a probabilidade de cometer o erro complementar ao coeficiente de confiança γ, isto é,

$$\alpha = P(\text{rejeitar } H_0 \mid H_0 \text{ é verdadeira})$$

Buscar uma regra de decisão que aponte quais resultados amostrais te levam a rejeição de H_0, ou seja, te levam a concluir pelo descrito na hipótese alternativa H_1.

A probabilidade α de cometer um erro de primeira espécie é um valor arbitrário e recebe o nome de **nível de significância do teste**.

O resultado da amostra é cada vez mais significante para rejeitar H_0 quanto menor for esse nível α. Usualmente, esses valores são fixados em 1%, 5% ou 10%.

Teste de hipóteses para uma média populacional

Caso 1

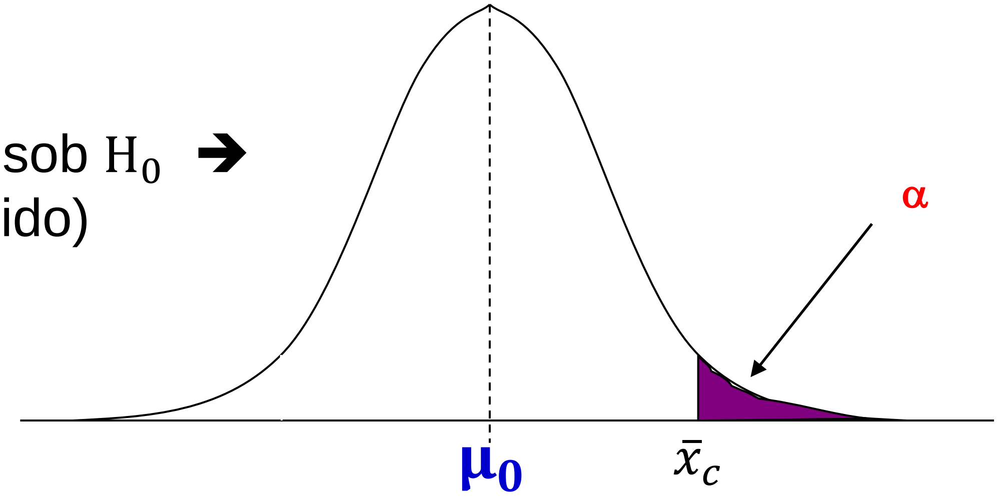
(variância populacional conhecida)

Unilateral ou unicaudal à direita

$H_0: \ \mu = \mu_0$ (ou $\mu \leq \mu_0$)

$H_A: \ \mu > \mu_0$

Distribuição de $\bar{X}$ sob H_0 ➔
(com σ conhecido)



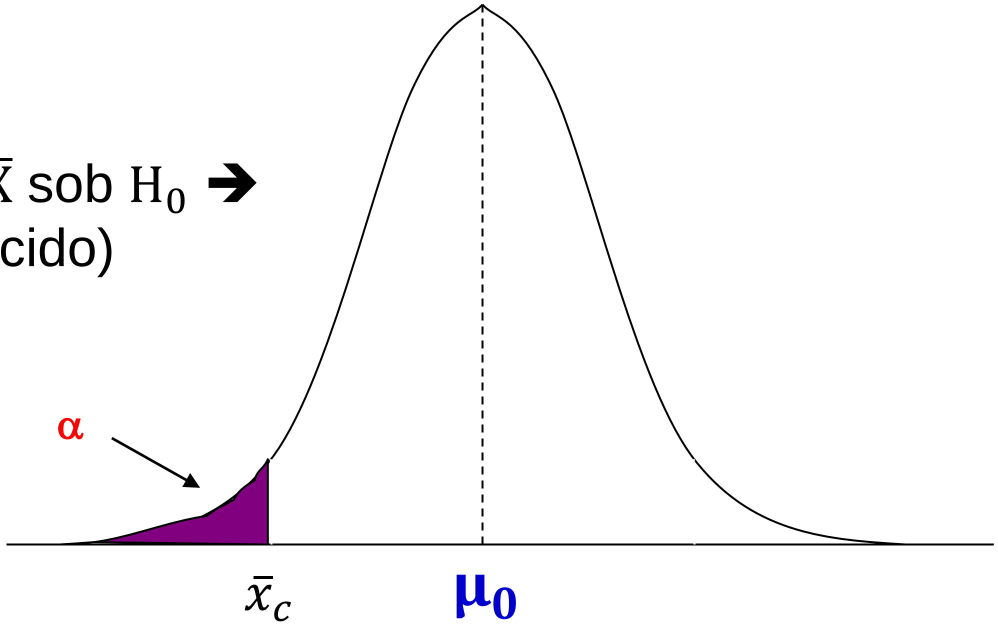
Região crítica: Rejeita-se H_0 para qualquer média amostral $\bar{x}_{obs} > \bar{x}_c$

Unilateral ou unicaudal à esquerda

$H_0: \ \mu = \mu_0$ (ou $\mu \geq \mu_0$)

$H_A: \ \mu < \mu_0$

Distribuição de $\bar{X}$ sob H_0 ➔
(com σ conhecido)

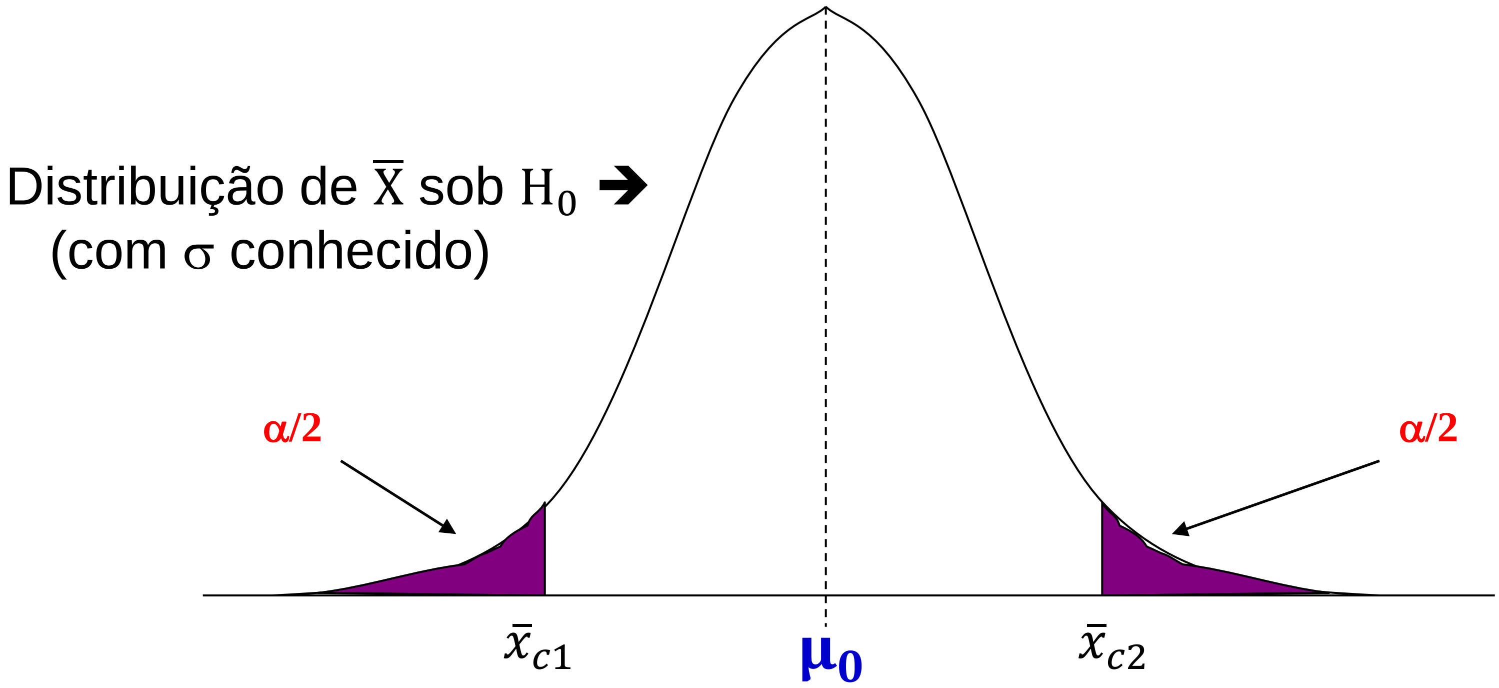


Região crítica: Rejeita-se H_0 para qualquer média amostral $\bar{x}_{obs} < \bar{x}_c$

Teste bilateral ou bicaudal

$$H_0 : \mu = \mu_0$$

$$H_A : \mu \neq \mu_0$$



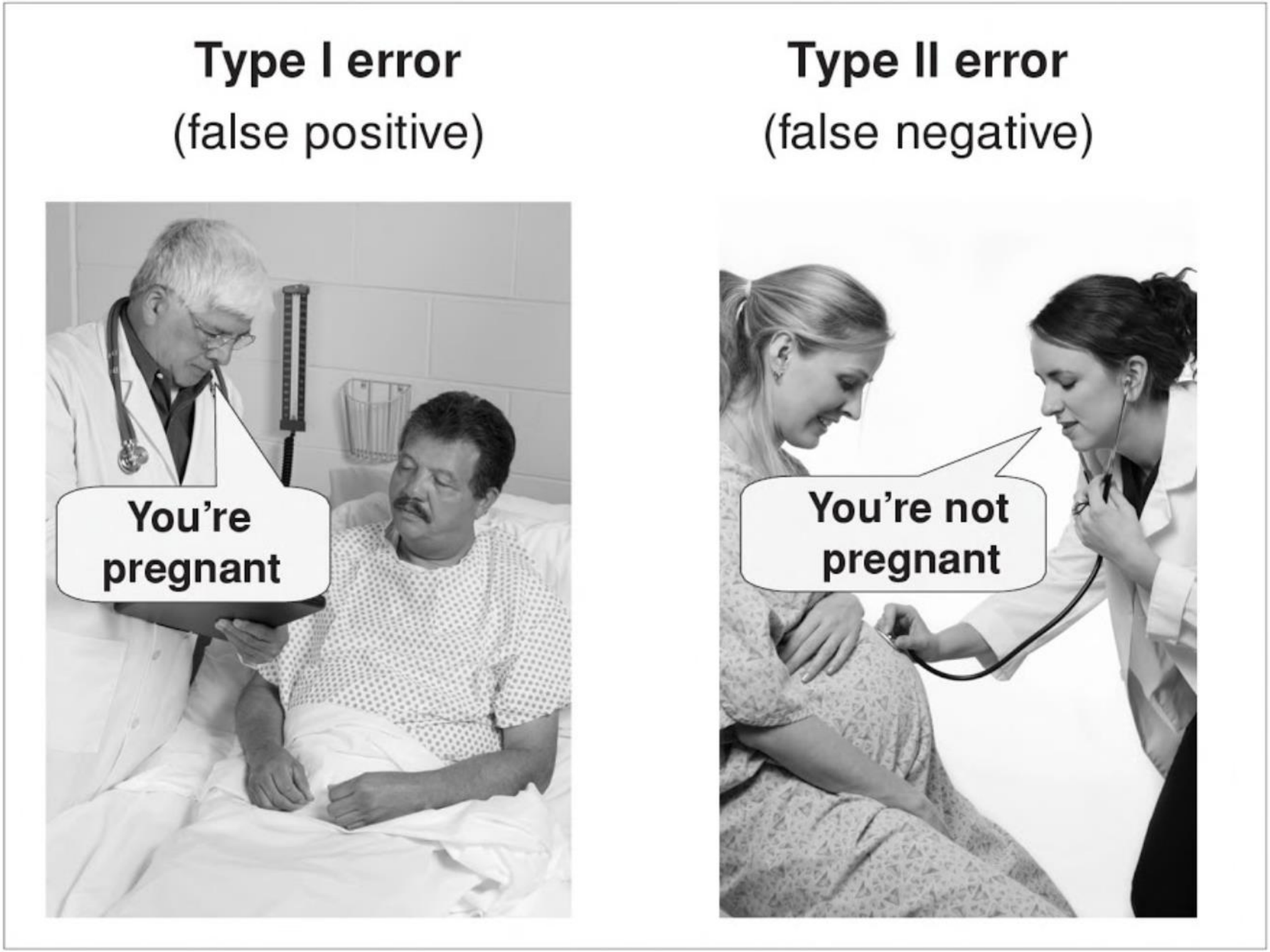
Região crítica: Rejeita-se H_0 para qualquer média amostral $\bar{x}_{obs} < \bar{x}_{c1}$ ou $\bar{x}_{obs} > \bar{x}_{c2}$

Tipos de Erros

Erro Tipo I: rejeitar H0 quando H0 é verdadeira
Erro Tipo II: não rejeitar H0 quando H0 é falsa

Erros tipo I e tipo II

	$\mathrm{H_0}$ é Verdadeiro	$\mathrm{H_0}$ é Falso
Rejeito $\mathrm{H_0}$	• Erro tipo 1 • Falso Positivo • Probabilidade α	Decisão Correta
Não Rejeito $\mathrm{H_0}$	Decisão Correta	• Erro tipo 2 • Falso Negativo • Probabilidade β



Fonte: The essential guide to effect sizes. Statistical Power, Meta-analysis and Interpretation of Results. Paul D. Ellis. pg. 50

Teste de Hipóteses para Tomada de Decisão

VALOR-P

Valor-p do Teste

Caso Unilateral a Esquerda:

$$H_0: \mu = \mu_0$$
$$H_1: \mu < \mu_0$$

É razoável rejeitar H_0 se a média amostral observada ($\bar{x}_{obs}$) for muito menor que μ_0.

Define-se o valor-p em um teste unilateral a esquerda como sendo

$$\text{valor} - \text{p} = P(\bar{X} \leq \bar{x}_{obs} | \mu = \mu_0) =$$
$$= \text{stats.norm.cdf}(\bar{x}_{obs}, \text{loc} = \mu_0, \text{scale} = \frac{\sigma}{\sqrt{n}})$$

Insper

Valor-p do Teste

Caso Unilateral a Direita:

$$H_0: \mu = \mu_0$$
$$H_1: \mu > \mu_0$$

É razoável rejeitar H_0 se a média amostral observada ($\bar{x}_{obs}$) for muito maior que μ_0.

Define-se o valor-p em um teste unilateral a direita como sendo

$$\text{valor} - \text{p} = P(\bar{X} \geq \bar{x}_{obs} | \mu = \mu_0) =$$
$$= 1 - \text{stats.norm.cdf}(\bar{x}_{obs}, \text{loc} = \mu_0, \text{scale} = \frac{\sigma}{\sqrt{n}})$$

Insper

Valor-p do Teste

Caso Bilateral:

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

É razoável rejeitar H_0 se a média amostral observada ($\bar{x}_{obs}$) for muito maior ou muito menor que μ_0.

Como calcular o valor-p neste caso?

Valor-p do Teste

Caso Bilateral: se $\bar{x}_{obs} > \mu_0$

Verifique se a média amostral observada é superior ou inferior a μ_0.

Vamos supor que a **média amostral observada ($\bar{x}_{obs}$) tenha sido superior a μ_0**, então, calculamos

$$p' = P(\bar{X} \geq \bar{x}_{obs} | \mu = \mu_0) =$$

$$= 1 - \text{stats.norm.cdf}(\bar{x}_{obs}, \text{loc} = \mu_0, \text{scale} = \frac{\sigma}{\sqrt{n}})$$

Define-se o valor-p em um teste bilateral como sendo

$$\text{valor-p} = 2p'$$

Insper

Valor-p do Teste

Caso Bilateral: se $\bar{x}_{obs} < \mu_0$

Verifique se a média amostral observada é superior ou inferior a μ_0.

Vamos supor que a **média amostral observada ($\bar{x}_{obs}$) tenha sido inferior a μ_0**, então, calculamos

$$p' = P(\bar{X} \geq \bar{x}_{obs} | \mu = \mu_0) =$$

$$= \text{stats.norm.cdf}(\bar{x}_{obs}, \text{loc} = \mu_0, \text{scale} = \frac{\sigma}{\sqrt{n}})$$

Define-se o valor-p em um teste bilateral como sendo

$$\text{valor-p} = 2p'$$

Exercícios

Exercício 1

Uma empresa de serviços na área de soluções de pagamentos eletrônicos fornece máquinas a estabelecimentos comerciais para processarem o pagamento por cartão de crédito efetuado por seus clientes. Entretanto, os donos desses estabelecimentos comerciais reclamaram que o tempo de processamento das transações de cartão de crédito nessas máquinas estava, em média, muito alto.

Para evitar a perda de clientes, a empresa estabeleceu como meta reduzir o tempo médio de processamento das transações de cartão de crédito para menos do que 4 segundos até o fim de março de 2017 (assuma σ conhecido e igual a 1 segundo).

Após a realização de diversas modificações, a empresa gostaria de checar se a meta foi atingida, ao nível de significância de 1%.

Para isso, ela mediu o tempo de processamento de uma amostra aleatória de 49 transações, cuja média amostral foi igual a 3,6 s.

$RC = \{\bar{x}_{obs} < 3{,}74\}$. Como $\bar{x}_{obs} \in RC$, então rejeita H_0. Concluir para problema.

Insper

Exercício 2

Um estudo foi desenvolvido para avaliar o salário de empregadas domésticas na cidade de São Paulo. Foram sorteadas e entrevistadas 121 trabalhadoras. Admita que o desvio padrão dessa variável na cidade é de 0,88 salários mínimos.

a) Você conhece a distribuição do estimador $\overline{X}$? Se não, é possível fazer alguma suposição?
b) Desconfia-se que a média salarial seja superior a 2 salários mínimos. Baseado nessa informação, formule as hipóteses e interprete os erros do tipo I e do tipo II.
c) Para um nível de significância de 2%, construa a RC.
d) Se a amostra forneceu média de 2,38 salários mínimos, qual a conclusão deste teste?
e) Que suposições você fez para resolver os itens anteriores?

Exercício 2 - respostas

a) A distribuição do salário médio das empregadas domésticas é $\overline{X} \sim N\left(\mu, \frac{0{,}88^2}{121}\right)$, sendo μ o verdadeiro salário da classe na cidade de São Paulo. Sob a suposição de n suficientemente grande para utilizar o TLC.

b) $\begin{cases} H_0\text{: desconfiança não procede} \\ H_1\text{: desconfiança procede} \end{cases} \Leftrightarrow \begin{cases} H_0: \ \mu \leq 2 \\ H_1: \ \mu > 2 \end{cases}$.

Erro I: concluir que desconfiança procede, quando na verdade não procede.

Erro II: concluir que desconfiança não procede, quando na verdade procede.

c) $RC = \{\bar{x}_{obs} > 2{,}164\}$ para um nível de significância de 2%.

d) Como $\bar{x}_{obs} = 2{,}38 \in RC$, então rejeita H_0. Concluir para problema.

e) Descrita no item (a).

Insper

Exercício 3

A duração de pilhas falsificadas segue uma distribuição normal com média de 15 ut e variância 16 ut^2. Pilhas autênticas têm uma duração média maior e mesmo desvio-padrão. Um lote de pilhas apreendido será leiloado e para definir seu preço é preciso decidir se são produtos falsificados ou não. O leiloeiro adotou como regra de decisão que se a duração média de uma amostra aleatória de 16 pilhas for maior que 18 ut, o lote será considerado autêntico. Critique essa regra.

Considerar a $RC = \{\bar{x}_{obs} > 18\}$, obtem $\alpha = 0{,}13\%$, o qual é muito baixo. Explicar...

Insper

Exercício 4

Um estudo foi desenvolvido para avaliar a renda de pedreiros autônomos na cidade de São Paulo, o qual desconfia que a renda média seja superior a 3 salários mínimos. Admita que o desvio padrão dessa variável na cidade é de 0,7 salários mínimos. Foram sorteados e entrevistados 100 trabalhadores, cuja amostra forneceu média de 3,2 salários mínimos.

a) Baseado nessa informação, formule as hipóteses H_0 e H_1 e interprete o erro do tipo I e erro do tipo II, sendo (erro I = rejeitar H_0 | H_0 verdade) e (erro II = não rejeitar H_0 | H_0 falsa).

b) Conclua com base na RC com 1% de significância. $RC = \{\bar{x}_{obs} > 3{,}1631\}$

c) Se a verdadeira média é 3,3 s.m., qual a probabilidade de se tomar uma decisão errada sob H1? β=0,025=2,5%

d) Supondo que a hipótese nula seja verdadeira, qual deve ser a probabilidade de observar um valor tão ou mais desfavorável à hipótese nula quanto a particular média amostral estimada. Qual a conclusão deste teste com 1% de significância ? Valor-p=1-0,9979=0,0021=0,21%

e) Compare as conclusões obtidas no item b (com base na RC) e no item d (no valor-p). Há diferenças? Por quê?

Exercício 5

Em 2012, antes dos supermercados deixarem de distribuir sacolas plásticas gratuitamente para os consumidores, o consumo mensal de sacolas plásticas na cidade de São Paulo por adulto era, em média, de 25 sacolas plásticas, com desvio padrão de 6 unidades (valores populacionais).

No mesmo ano, os estabelecimentos passaram a cobrar pelo uso das mesmas e depois voltaram atrás. Alguns anos se passaram e uma ONG da área de sustentabilidade deseja avaliar se essas idas e vindas deixaram o consumidor um pouco mais consciente com o meio ambiente.

Para avaliar se consumo está mais consciente, verifique se o consumo de sacolas plásticas reduziu, em média, em mais do que 15% após todo esse período, ao nível de significância de 1%?

A ONG monitorou, por um mês, 100 adultos moradores da cidade e observou que o consumo médio amostral de sacolas plásticas entre eles foi de 20,01 unidades.

$RC = \{\bar{x}_{obs} < 19{,}852\}$. Como $\bar{x}_{obs} = 20{,}01 \notin RC$, então não rejeita H_0. Concluir para problema.