# Testes de Hipóteses para Média de Populações Normais- Variância conhecida e desconhecida

Universidade Estadual de Santa Cruz

Ivan Bezerra Allaman

# Cronograma

1. Considerando variância conhecida
    1. Introdução
    2. Regra baseada no critério do valor crítico
        - `Exemplo: Hipótese bilateral`
        - `Exemplo: Hipótese unilateral à direita`
        - `Exemplo: Hipótese unilateral à esquerda`
    3. Regra baseada no *p-valor*
        - `Exemplo: Hipótese bilateral`
        - `Exemplo: Hipótese unilateral à direita`
        - `Exemplo: Hipótese unilateral à esquerda`
    4. Aplicações

1. Considerando variância desconhecida
    1. Introdução
    2. Aplicações

# Considerando variância conhecida

# Introdução

- Nestes casos utiliza-se a seguinte estatística de teste:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

- Uma vez calculada a estatística de teste, chegou-se a hora de tomarmos uma decisão.
- Há duas regras para tomadas de decisão:
  - a regra de rejeição baseada no critério do valor crítico;
  - a regra de rejeição baseada no critério do *p-valor*.
    - Você pode optar por uma delas.

# Regra baseada no critério do valor crítico

- Obviamente que a tomada de decisão irá depender do tipo de hipótese estabelecida (bilateral, unilateral à direita ou a esquerda).
- Logo, teremos as seguintes regras de acordo com as hipóteses estabelecidas:

## Hipótese bilateral

Rejeitar $H_0$ se $Z \leq -Z_{\alpha/2}$ ou se $Z \geq Z_{\alpha/2}$

# Aplicação

1. Um metalúrgico decide testar a pureza de um certo metal, que supõe ser constituído exclusivamente de manganês. Adota para isso o critério da verificação do ponto de fusão. Experiências anteriores mostraram que esse ponto de fusão se distribuía normalmente com média de 1260° e desvio padrão de 2°. O metalúrgico realizou 4 experiências, obtendo 1267°, 1269°, 1261° e 1263°. Poderá ele aceitar que o metal é puro ao nível de 5% de significância?

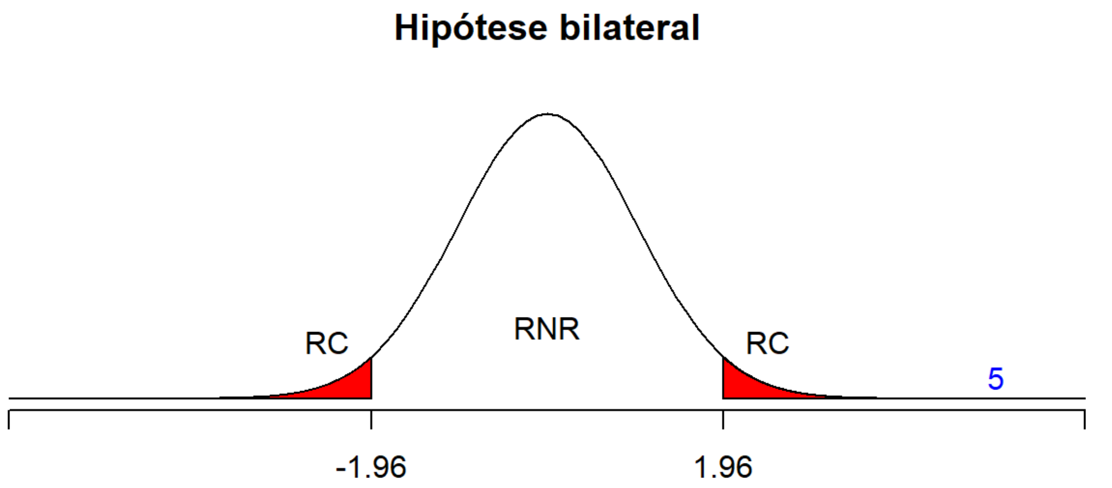
Vamos estabelecer primeiramente as hipóteses. Se o metal for puro, então não devemos refutar a hipótese do ponto de fusão médio ser de 1260°. Caso contrário, será diferente de 1260°. Logo, o tipo de hipótese em questão é bilateral. Logo,

$$H_0 : \mu = 1260$$
$$H_a : \mu \neq 1260$$

Vamos então aos cálculos das estatísticas.

$$\bar{X} = \frac{1267 + \cdots + 1263}{4} = 1265$$
$$\sigma = 2$$
$$n = 4$$
$$\alpha = 5\%$$

$$\sigma_{\bar{X}} = \frac{2}{\sqrt{4}} = 1$$
$$Z_{calc.} = \frac{1265 - 1260}{1} = 5$$
$$Z_{5\%/2} = abs(qnorm(0,025)) = 1,96$$




Resp: Como a estatística Z(5) está dentro da região crítica ($Z_{calc.} > Z_{\alpha/2} = 1,96$), então rejeita-se $H_0$ com $95\%$ de confiança, ou seja, podemos inferir com $95\%$ de confiança que a média da população é diferente de $1260$, ou seja, o metal não é puro.

1. O salário médio por hora para empregados em indústrias de produção de bens é atualmente de US$ 24,57 (site do Bureau of Labor Statistics, 12 de abril de 2012). Suponha que consideremos uma amostra de empregados da indústria de manufatura para ver se o salário médio por hora difere da média reportada, de US$ 24,57, para as indústrias produtoras de bens.

   a. Declare as hipóteses nula e alternativa que devemos utilizar para testar se o salário horário médio da população na indústria manufatureira difere do salário médio por hora da população nas indústrias produtoras de bens.

   b. Suponha que uma amostra de 30 empregados da indústria de manufatura tenha uma média amostral de US$ 23,89 por hora. Assuma um desvio padrão populacional de US$ 2,40 por hora. Com $\alpha = 0,05$ como nível de significância, qual é a sua conclusão?

# Hipótese unilateral à direita

Rejeitar $H_0$ se $Z \geq Z_\alpha$

# Aplicação

2. A tensão de ruptura de cabos fabricados por uma empresa apresenta distribuição normal, com média de 1800 kg e desvio padrão de 100 kg. Mediante uma nova técnica de produção, proclamou-se que a tensão de ruptura teria aumentado. Para testar essa declaração, ensaio-se uma amostra de 50 cabos, obtendo-se como tensão média de ruptura 1850 kg. Pode-se aceitar a proclamação ao nível de 5%?

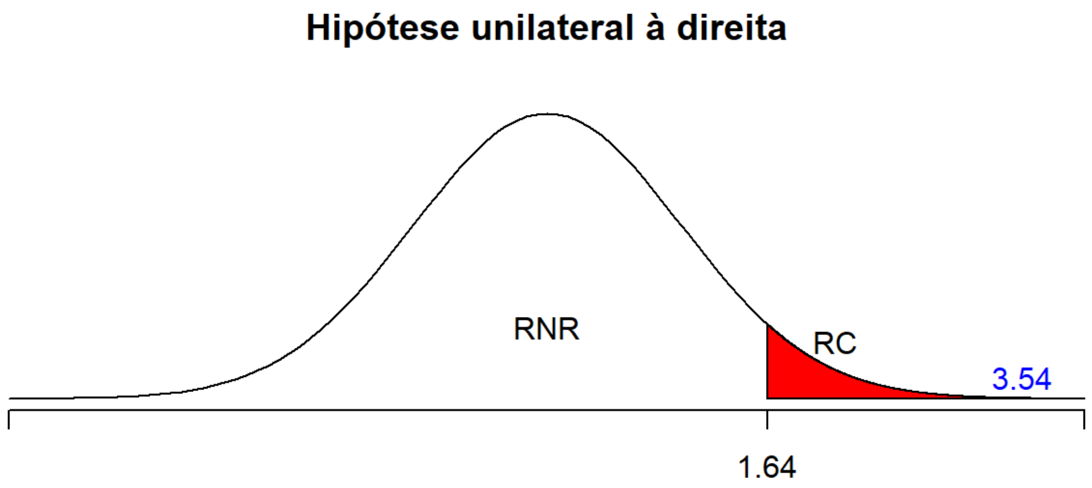
   Elaborando as hipóteses tem-se:

$$H_0 : \mu \leq 1800$$
$$H_a : \mu > 1800$$

Vamos então aos cálculos das estatísticas.

$$\bar{X} = 1850$$
$$\sigma = 100$$
$$n = 50$$
$$\alpha = 5\%$$

$$\sigma_{\bar{X}} = \frac{100}{\sqrt{50}} = 14,14$$
$$Z_{calc.} = \frac{1850 - 1800}{14,14} = 3,54$$
$$Z_{5\%} = abs(qnorm(0,05)) = 1,64$$




Resp: Como a estatística Z (3,54) está fora da região de não-rejeição ($Z > Z_{alpha} = 1,64$), então rejeita-se $H_0$ com $95\%$ de confiança, ou seja, podemos inferir com $95\%$ de confiança que a tensão de ruptura média da população é maior que $1800$.

2. Em um estudo intitulado ***How undergraduate students use credit cards*** (como estudantes universitários utilizam cartões de crédito) foi relatado que estudantes universitários têm um saldo médio de cartão de crédito de US$ 3.173,00. Estes números representam uma grande marca e aumentaram 44% nos cinco anos anteriores. Suponha que um estudo atual esteja sendo realizado para determinar se é possível concluir que o saldo médio de cartões de crédito para estudantes universitários tem aumentado em comparação com o relatório original. Com base em estudos anteriores, utilize um desvio padrão populacional $\sigma =$ US$ 1.000,00.

   a. Declare as hipóteses nula e alternativa

   b. Considerando um amostra de 180 estudantes universitários com um saldo médio de cartão de crédito de US$ 3.325,00, qual é a sua conclusão considerando um nível de significância de 3%?

# Hipótese unilateral à esquerda

Rejeitar $H_0$ se $Z \leq -Z_\alpha$

# Aplicação

3. Um comprador de blocos de cimento acredita que a qualidade dos produtos da marca A esteja se deteriorando. Sabe-se, por experiência passada, que a força média de esmagamento desses blocos era de 400 libras, com desvio padrão de 20 libras. Uma amostra de 100 blocos da marca A forneceu uma força média de esmagamento de 390 libras (supor distribuição normal). Testar ao nível de 2,5% que a qualidade média dos blocos tenha diminuído.

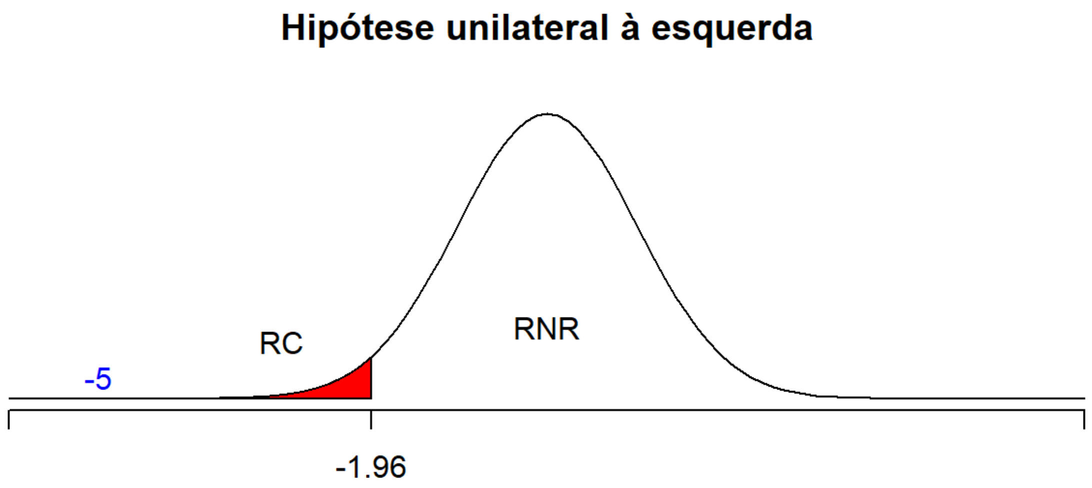
   Elaborando as hipóteses tem-se:

$$H_0 : \mu \geq 400$$
$$H_a : \mu < 400$$

Vamos então aos cálculos das estatísticas.

$$\bar{X} = 390$$
$$\sigma = 20$$
$$n = 100$$
$$\alpha = 2,5\%$$

$$\sigma_{\bar{X}} = \frac{20}{\sqrt{100}} = 2$$
$$Z_{calc.} = \frac{390 - 400}{2} = -5$$
$$Z_{2,5\%} = qnorm(0.025) = -1,96$$




Resp: Como a estatística Z (-5) está dentro da região crítica ($Z < Z_{alpha} = -1,96$), então rejeita-se $H_0$ com $97,5\%$ de confiança, ou seja, podemos inferir com $97,5\%$ de confiança que a qualidade média dos blocos é menor do que $400$.

3. O gasto anual com medicamentes prescritos foi de US$ 838 por pessoa no nordeste dos Estados Unidos (site do Hospital Care Cost Institute, 7 novembro de 2012). Uma pesquisa recente com 60 pessoas mostrou uma despesa anual de US$ 745 por pessoa com medicamentes prescritos. Use um desvio padrão populacional de US$ 300 para responder às seguintes perguntas.

   a. Formule adequadamente uma hipótese que sustente o argumento de que o gasto anual com medicamentos prescritos diminuiu.

   b. Considerando um nível de significância de 1% qual é a sua conclusão?

# Regra baseada no *p-valor*

- O *p-valor* é a probabilidade de se obter uma estatística de teste igual ou mais extrema que aquela observada em uma amostra, sob a hipótese nula. Em termos coloquiais, ele avalia o quão bem os dados da amostra apoiam o argumento de que a hipótese nula é verdadeira, ou seja, o quão compatíveis os seus dados são com a hipótese nula.
- Logo, tem-se a seguinte regra de decisão:

## Hipótese bilateral

Rejeitar $H_0$ se o p-valor $\leq \alpha/2$

# Aplicação

4. Considerando a mesma aplicação 1, calcule o p-valor e faça as devidas conclusões.

Neste caso precisamos de um programa para tal cálculo. Utilizando o R temos:

$$p - valor = pnorm(Z_{calc.}) = pnorm(5, lower.tail = F) * 2 = 5,73 \cdot 10^{-7}$$

O nível de significância adotado foi de 5%. Como o p-valor foi menor do que o nível de significancia, então rejeitamos $H_0$, ou seja, a probabilidade dos dados apoiarem a hipótese nula foi menor do que aquele proposto pelo pesquisador como mínimo para apoiar tal hipótese.

# Hipótese unilateral à direita

Rejeitar $H_0$ se o p-valor $\leq \alpha$

# Aplicação

5. Considerando a mesma aplicação 2, calcule o p-valor e faça as devidas conclusões.

Utilizando o R temos:

$$p - valor = pnorm(Z_{calc.}) = pnorm(3.54, lower.tail = F) = 2 \cdot 10^{-4}$$

Como o p-valor é menor do que $\alpha$, rejeitamos $H_0$ com 95% de confiança.

# Hipótese unilateral à esquerda

Rejeitar $H_0$ se o p-valor $\leq \alpha$

# Aplicação

6. Considerando a mesma aplicação 3, calcule o p-valor e faça as devidas conclusões.

Utilizando o R temos:

$$p - valor = pnorm(Z_{calc.}) = pnorm(-5) = 2,86 \cdot 10^{-7}$$

Como o p-valor é menor do que $\alpha$, rejeitamos $H_0$ com 97,5% de confiança.

# Considerando variância desconhecida

# Introdução

- Neste caso só irá mudar a função densidade de probabilidade que será a *t-Student*, pois todo o raciocínio anterior é válido.
- Lembrem-se que além da média você deverá também estimar a variância da amostra.
- Logo, tem-se a seguinte estatística de teste:

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

- Lembrando que a distribuição t tem $v$ graus de liberdade que é $n-1$.

# Aplicação

7. Uma máquina de misturar fertilizantes é adaptada para fornecer 10g de nitrato para cada 100g de fertilizante. Dez porções de 100g são examinadas, com as seguintes porcentagens de nitrato:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 9 | 12 | 11 | 10 | 11 | 9 | 11 | 12 | 9 | 10 |

Há razões para crer que a porcentagem de nitrato não é 10g ao nível de 10%?

Elaborando as hipóteses tem-se:

$$H_0 : \mu = 10$$
$$H_a : \mu \neq 10$$

Vamos aos cálculos das estatísticas.

$$\bar{X} = 10,4$$
$$S = 1,174$$
$$n = 10$$
$$S_{\bar{X}} = \frac{1,174}{\sqrt{10}} = 0,371$$
$$t_{calc} = \frac{10,4 - 10}{0,371} = 1,078$$
$$v = n - 1 = 10 - 1 = 9$$
$$p - valor = pt(t_{calc}, v, lower.tail = F) * 2 = 0,309$$

Logo, como o p-valor (30,9%) é maior que 10%, não rejeitamos $H_0$, ou seja, não há razões para crer que a porcentagem de nitrato não é 10%.

7. O tempo que os homens casados dedicam aos cuidados de seus filhos é de 6,4 horas por semana (Time, 12 de março de 2012). Você pertence a um grupo profissional de estudos sobre práticas familiares que gostaria de fazer sua própria pesquisa para determinar se o tempo que os homens casados, que moram em sua região, gastam com o cuidado de seus filhos por semana difere da média relatada acima. Uma amostra de 40 casais será utilizada, com os dados coletados mostrando as horas por semana que os maridos gastam cuidando de seus filhos. Os dados amostrais estão contidos no seguinte link: https://lec.pro.br//download/R/dados/ChildCare.xlsx
   a. Quais são as hipóteses se o seu grupo quiser determinar se o número médio de horas que a população de homens casados gasta no cuidado de seus filhos é diferente da média relatada pela ***Time*** em sua região?
   b. Considerando um nível de significância de 4% qual é a sua conclusão?