# Probabilidade e Estatística (EAD)

*Tradução e adaptação:*

Priscilla Priscilla Gnewuch, Márcia Helena Barbian e Maitê Mückler

Slides baseados no material desenvolvido por Mine Çetinkaya-Rundel of OpenIntro.



UFRGS
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Instituto de MATEMÁTICA E ESTATÍSTICA UFRGS
SEAD UFRGS
EDUCAÇÃO A DISTÂNCIA

# Capítulo 5: Inferência para dados numéricos

Slides desenvolvidos por Mine Çetinkaya-Rundel of OpenIntro.

# 5.1. Teste de uma amostra para a média com a distribuição $t$

# Sexta-Feira 13

Entre 1990 e 1992, pesquisadores do Reino Unido coletaram dados sobre fluxo de trânsito, acidentes e internações hospitalares numa sexta-feira 13 e na sexta-feira imaeidatamente anterior, dia 6. Abaixo está um resumo deste conjunto de dados. Podemos supor que o fluxo de tráfego em determinado dia nos locais 1 e 2 seja independente.

# Sexta-Feira 13

| | tipo | data | dia 6 | dia 13 | qtde | localização |
|---|---|---|---|---|---|---|
| 1 | tráfego | 1990, Julho | 139246 | 138548 | 698 | loc 1 |
| 2 | tráfego | 1990, Julho | 134012 | 132908 | 1104 | loc 2 |
| 3 | tráfego | 1991, Setembro | 137055 | 136018 | 1037 | loc 1 |
| 4 | tráfego | 1991, Setembro | 133732 | 131843 | 1889 | loc 2 |
| 5 | tráfego | 1991, Dezembro | 123552 | 121641 | 1911 | loc 1 |
| 6 | tráfego | 1991, Dezembro | 121139 | 118723 | 2416 | loc 2 |
| 7 | tráfego | 1992, Março | 128293 | 125532 | 2761 | loc 1 |
| 8 | tráfego | 1992, Março | 124631 | 120249 | 4382 | loc 2 |
| 9 | tráfego | 1992, Novembro | 124609 | 122770 | 1839 | loc 1 |
| 10 | tráfego | 1992, Novembro | 117584 | 117263 | 321 | loc 2 |

Scanlon, T.J., Luben, R.N., Scanlon, F.L., Singleton, N. (1993), “Is Friday the 13th Bad For Your Health?,"BMJ, 307, 1584-1586.

## Sexta-feira 13

- Queremos investigar se o comportamento das pessoas é diferente na sexta-feira 13 em comparação com a sexta-feira do dia 6.

## Sexta-feira 13

- Queremos investigar se o comportamento das pessoas é diferente na sexta-feira 13 em comparação com a sexta-feira do dia 6.
- Uma abordagem é comparar o fluxo de tráfego nesses dois dias.

# Sexta-feira 13

- Queremos investigar se o comportamento das pessoas é diferente na sexta-feira 13 em comparação com a sexta-feira do dia 6.
- Uma abordagem é comparar o fluxo de tráfego nesses dois dias.
- $H_0$ : O fluxo médio de tráfego nas duas sextas-feiras é igual.
  $H_A$ : O fluxo médio de tráfego nas duas sextas-feiras é diferente.

# Sexta-feira 13

Cada caso no conjunto de dados representa o fluxo de tráfego registrado no mesmo local no mesmo mês do mesmo ano: uma contagem na sexta-feira 6 e a outra sexta-feira 13. Essas duas contagens são independentes?

## Sexta-feira 13

Cada caso no conjunto de dados representa o fluxo de tráfego registrado no mesmo local no mesmo mês do mesmo ano: uma contagem na sexta-feira 6 e a outra sexta-feira 13. Essas duas contagens são independentes?

*Não*

## Hipóteses

Quais são as hipóteses para testar uma diferença entre o fluxo médio de tráfego entre sexta-feira 6 e 13?

(a) $H_0 : \mu_6 = \mu_{13}$
    $H_A : \mu_6 \neq \mu_{13}$

(b) $H_0 : p_6 = p_{13}$
    $H_A : p_6 \neq p_{13}$

(c) $H_0 : \mu_{qtde} = 0$
    $H_A : \mu_{qtde} \neq 0$

(d) $H_0 : \bar{x}_{qtde} = 0$
    $H_A : \bar{x}_{qtde} = 0$

# Hipóteses

Quais são as hipóteses para testar uma diferença entre o fluxo médio de tráfego entre sexta-feira 6 e 13?

(a) $H_0 : \mu_6 = \mu_{13}$

$H_A : \mu_6 \neq \mu_{13}$

(b) $H_0 : p_6 = p_{13}$

$H_A : p_6 \neq p_{13}$

(c) $H_0 : \mu_{qtde} = 0$

$H_A : \mu_{qtde} \neq 0$

(d) $H_0 : \bar{x}_{qtde} = 0$

$H_A : \bar{x}_{qtde} = 0$

# Condições

- *Independência:* assume que os casos (linhas) são independentes.

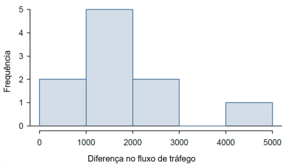
## Condições

- *Independência:* assume que os casos (linhas) são independentes.
- *Tamanho de amostra/assimetria:*

# Condições

- *Independência:* assume que os casos (linhas) são independentes.
- *Tamanho de amostra/assimetria:*

- A distribuição da amostra não parece ser extremamente assimétrica, mas é muito difícil avaliar com um tamanho de amostra tão pequeno. Podemos pensar se é possível que a distribuição da população seja assimétrica ou não, porém, provavelmente não deve ser assimétrica, já que deve ser igualmente provável que ocorram dias com tráfego abaixo da média e dias com o tráfego acima da média.

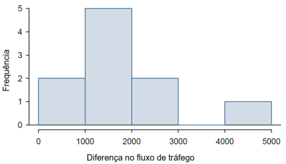



- Nós não conhecemos $\sigma$ e $n$ é muito pequeno para que $s$ seja uma estimativa confiável para $\sigma$.

# Condições

- *Independência:* assume que os casos (linhas) são independentes.
- *Tamanho de amostra/assimetria:*

- A distribuição da amostra não parece ser extremamente assimétrica, mas é muito difícil avaliar com um tamanho de amostra tão pequeno. Podemos pensar se é possível que a distribuição da população seja assimétrica ou não, porém, provavelmente não deve ser assimétrica, já que deve ser igualmente provável que ocorram dias com tráfego abaixo da média e dias com o tráfego acima da média.




- Nós não conhecemos $\sigma$ e $n$ é muito pequeno para que $s$ seja uma estimativa confiável para $\sigma$.

Então, o que fazemos quando o tamanho da amostra é pequeno?

## Revisão: para que serve um tamanho de amostra grande?

Enquanto as observações forem independentes e a distribuição da população não for extremamente assimétrica, uma grande amostra irá garantir que ...

- a distribuição amostral da média é quase normal
- a estimativa do erro padrão, como $\frac{s}{\sqrt{n}}$, é confiável

# A condição de normalidade

- O TCL, que afirma que as distribuições da amostra serão quase normais, vale para *qualquer* tamanho da amostra, desde que a distribuição da população seja quase normal.

# A condição de normalidade

- O TCL, que afirma que as distribuições da amostra serão quase normais, vale para *qualquer* tamanho da amostra, desde que a distribuição da população seja quase normal.
- Embora esse seja um caso especial bastante útil, é inerentemente difícil verificar a normalidade em conjuntos de dados pequenos.

# A condição de normalidade

- O TCL, que afirma que as distribuições da amostra serão quase normais, vale para *qualquer* tamanho da amostra, desde que a distribuição da população seja quase normal.
- Embora esse seja um caso especial bastante útil, é inerentemente difícil verificar a normalidade em conjuntos de dados pequenos.
- Devemos ter cautela ao verificar a condição de normalidade para amostras pequenas. É importante não só examinar os dados, mas também pensar sobre a origem dos dados.
  - Por exemplo, se pergunte: eu esperaria que essa distribuição fosse simétrica e estou confiante de que os outliers são raros?

## A distribuição $t$

- Quando o desvio padrão da população é desconhecido (o que quase sempre ocorre), a incerteza da estimativa do erro padrão é resolvida usando uma nova distribuição: a *distribuição $t$*.

# A distribuição $t$

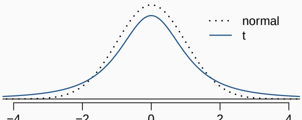
- Quando o desvio padrão da população é desconhecido (o que quase sempre ocorre), a incerteza da estimativa do erro padrão é resolvida usando uma nova distribuição: a *distribuição $t$*.
- Essa distribuição também tem formato de sino, mas suas caudas são *mais pesadas* do que as do modelo normal.

# A distribuição $t$

- Quando o desvio padrão da população é desconhecido (o que quase sempre ocorre), a incerteza da estimativa do erro padrão é resolvida usando uma nova distribuição: a *distribuição $t$*.
- Essa distribuição também tem formato de sino, mas suas caudas são *mais pesadas* do que as do modelo normal.
- Portanto, é mais provável que as observações caiam além de dois SDs da média do que sob a distribuição normal.

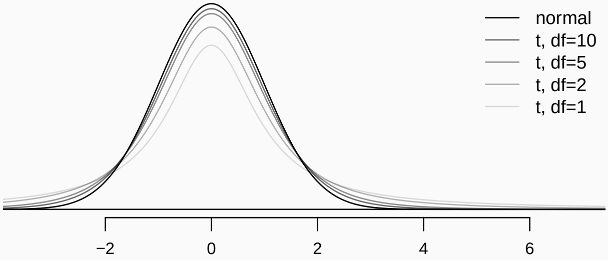
# A distribuição $t$

- Quando o desvio padrão da população é desconhecido (o que quase sempre ocorre), a incerteza da estimativa do erro padrão é resolvida usando uma nova distribuição: a *distribuição $t$*.
- Essa distribuição também tem formato de sino, mas suas caudas são *mais pesadas* do que as do modelo normal.
- Portanto, é mais provável que as observações caiam além de dois SDs da média do que sob a distribuição normal.
- Essas caudas mais pesadas são úteis para resolver nosso problema com uma estimativa menos confiável do erro padrão (já que $n$ é pequeno).



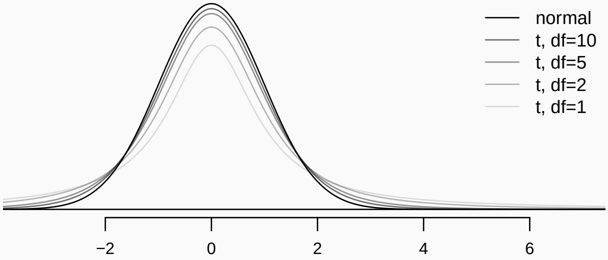
## A distribuição $t$ (cont.)

- Sempre centralizado em zero, como a distribuição normal padrão ($z$).
- Tem um único parâmetro *graus de liberdade*: ($df$).



# A distribuição $t$ (cont.)

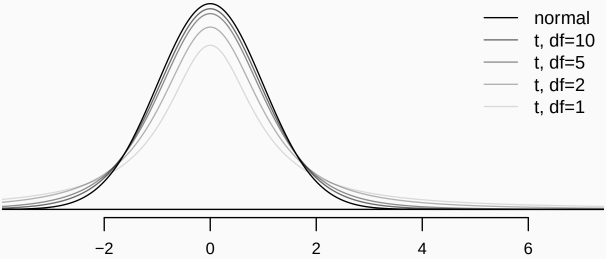
- Sempre centralizado em zero, como a distribuição normal padrão ($z$).
- Tem um único parâmetro *graus de liberdade*: ($df$).




O que acontece com a forma da distribuição $t$ quando $df$ aumenta?

# A distribuição $t$ (cont.)

- Sempre centralizado em zero, como a distribuição normal padrão ($z$).
- Tem um único parâmetro *graus de liberdade*: ($df$).




O que acontece com a forma da distribuição $t$ quando $df$ aumenta?

*Aproxima-se da normal.*

# De volta à sexta-feira 13

| | tipo | data | dia 6 | dia 13 | qtde | localização |
|---|---|---|---|---|---|---|
| 1 | trafégo | 1990, Julho | 139246 | 138548 | 698 | loc 1 |
| 2 | trafégo | 1990, Julho | 134012 | 132908 | 1104 | loc 2 |
| 3 | trafégo | 1991, Setembro | 137055 | 136018 | 1037 | loc 1 |
| 4 | trafégo | 1991, Setembro | 133732 | 131843 | 1889 | loc 2 |
| 5 | trafégo | 1991, Dezembro | 123552 | 121641 | 1911 | loc 1 |
| 6 | trafégo | 1991, Dezembro | 121139 | 118723 | 2416 | loc 2 |
| 7 | trafégo | 1992, Março | 128293 | 125532 | 2761 | loc 1 |
| 8 | trafégo | 1992, Março | 124631 | 120249 | 4382 | loc 2 |
| 9 | trafégo | 1992, Novembro | 124609 | 122770 | 1839 | loc 1 |
| 10 | trafégo | 1992, Novembro | 117584 | 117263 | 321 | loc 2 |

$\downarrow$

$\bar{x}_{qtde} = 1836$

$s_{qtde} = 1176$

$n = 10$

# Encontrando a estatística de teste

## Teste estatístico: inferência para média em uma amostra pequena

A estatística de teste para inferência sobre a média em uma amostra pequena ($n < 50$) é a estatística $T$ com $df = n - 1$.

$$T_{df} = \frac{\text{esmativa pontual – valor da hipótese nula}}{SE}$$

# Encontrando a estatística de teste

## Teste estatístico: inferência para média em uma amostra pequena

A estatística de teste para inferência sobre a média em uma amostra pequena ($n < 50$) é a estatística $T$ com $df = n - 1$.

$$T_{df} = \frac{\text{esmativa pontual} - \text{valor da hipótese nula}}{SE}$$

*Neste contexto...*

$$estimativa\ pontual \quad = \quad \bar{x}_{qtde} = 1836$$

# Encontrando a estatística de teste

## Teste estatístico: inferência para média em uma amostra pequena

A estatística de teste para inferência sobre a média em uma amostra pequena ($n < 50$) é a estatística $T$ com $df = n - 1$.

$$T_{df} = \frac{\text{esmativa pontual} - \text{valor da hipótese nula}}{SE}$$

*Neste contexto...*

$$\begin{aligned} \textit{estimativa pontual} &= \bar{x}_{qtde} = 1836 \\ SE &= \frac{s_{qtde}}{\sqrt{n}} = \frac{1176}{\sqrt{10}} = 372 \end{aligned}$$

# Encontrando a estatística de teste

## Teste estatístico: inferência para média em uma amostra pequena

A estatística de teste para inferência sobre a média em uma amostra pequena ($n < 50$) é a estatística $T$ com $df = n - 1$.

$$T_{df} = \frac{\text{esmativa pontual} - \text{valor da hipótese nula}}{SE}$$

*Neste contexto...*

$$\begin{aligned}
\textit{estimativa pontual} &= \bar{x}_{qtde} = 1836 \\
SE &= \frac{s_{qtde}}{\sqrt{n}} = \frac{1176}{\sqrt{10}} = 372 \\
T &= \frac{1836 - 0}{372} = 4.94
\end{aligned}$$

# Encontrando a estatística de teste

## Teste estatístico: inferência para média em uma amostra pequena

A estatística de teste para inferência sobre a média em uma amostra pequena ($n < 50$) é a estatística $T$ com $df = n - 1$.

$$T_{df} = \frac{\text{esmativa pontual} - \text{valor da hipótese nula}}{SE}$$

*Neste contexto...*

$$\begin{aligned}
estimativa\ pontual &= \bar{x}_{qtde} = 1836 \\
SE &= \frac{s_{qtde}}{\sqrt{n}} = \frac{1176}{\sqrt{10}} = 372 \\
T &= \frac{1836 - 0}{372} = 4.94 \\
df &= 10 - 1 = 9
\end{aligned}$$

*Note: O valor da hipótese nula é 0 porque definimos $\mu_{qtde} = 0$.*

# Encontrando o valor p

- O valor p é, mais uma vez, calculado como sendo a área de cauda sob a distribuição $t$.

# Encontrando o valor p

- O valor p é, mais uma vez, calculado como sendo a área de cauda sob a distribuição $t$.
- Usando R:

```
> 2 * pt(4.94, df = 9, lower.tail = FALSE)

[1] 0.0008022394
```

# Encontrando o valor p

- O valor p é, mais uma vez, calculado como sendo a área de cauda sob a distribuição $t$.
- Usando R:

```
> 2 * pt(4.94, df = 9, lower.tail = FALSE)

[1] 0.0008022394
```

- Usando a web app: https://gallery.shinyapps.io/dist_calc/

# Encontrando o valor p

- O valor p é, mais uma vez, calculado como sendo a área de cauda sob a distribuição $t$.
- Usando R:

```
> 2 * pt(4.94, df = 9, lower.tail = FALSE)

[1] 0.0008022394
```

- Usando a web app: https://gallery.shinyapps.io/dist_calc/
- Se o app não estiver disponível, podemos usar uma tabela $t$.

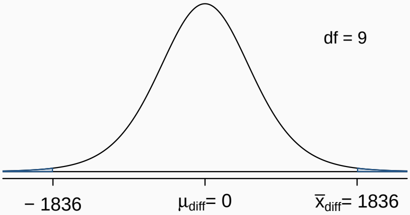
# Encontrando o valor p

Localize a estatística $T$ calculada na linha $df$ apropriada, obtenha o valor p do cabeçalho da coluna correspondente (uma ou duas linhas, dependendo da hipótese alternativa).

| one tail | | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
|---|---|---|---|---|---|---|
| two tails | | 0.200 | 0.100 | 0.050 | 0.020 | 0.010 |
| $df$ | 1 | 3.08 | 6.31 | 12.71 | 31.82 | 63.66 |
| | 2 | 1.89 | 2.92 | 4.30 | 6.96 | 9.92 |
| | 3 | 1.64 | 2.35 | 3.18 | 4.54 | 5.84 |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | |
| | 17 | 1.33 | 1.74 | 2.11 | 2.57 | 2.90 |
| | 18 | 1.33 | 1.73 | 2.10 | 2.55 | 2.88 |
| | 19 | 1.33 | 1.73 | 2.09 | 2.54 | 2.86 |
| | 20 | 1.33 | 1.72 | 2.09 | 2.53 | 2.85 |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | |
| | 400 | 1.28 | 1.65 | 1.97 | 2.34 | 2.59 |
| | 500 | 1.28 | 1.65 | 1.96 | 2.33 | 2.59 |
| | $\infty$ | 1.28 | 1.64 | 1.96 | 2.33 | 2.58 |

# Encontrando o valor p (cont.)

| one tail | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
|---|---|---|---|---|---|
| two tails | 0.200 | 0.100 | 0.050 | 0.020 | 0.010 |
| df 6 | 1.44 | 1.94 | 2.45 | 3.14 | 3.71 |
| 7 | 1.41 | 1.89 | 2.36 | 3.00 | 3.50 |
| 8 | 1.40 | 1.86 | 2.31 | 2.90 | 3.36 |
| 9 | 1.38 | 1.83 | 2.26 | 2.82 | 3.25 |
| 10 | 1.37 | 1.81 | 2.23 | 2.76 | 3.17 |




$T = 4.94$

# Conclusão do teste

Qual é a conclusão do teste de hipóteses?

# Conclusão do teste

Qual é a conclusão do teste de hipóteses?

*Os dados fornecem evidências convincentes de uma diferença entre o fluxo de tráfego na sexta-feira 6 e 13.*

# Qual é a diferença?

- Concluímos que há uma diferença no fluxo de tráfego entre sexta-feira 6 e 13.

## Qual é a diferença?

- Concluímos que há uma diferença no fluxo de tráfego entre sexta-feira 6 e 13.
- Mas seria mais interessante descobrir exatamente que diferença é essa.

# Qual é a diferença?

- Concluímos que há uma diferença no fluxo de tráfego entre sexta-feira 6 e 13.
- Mas seria mais interessante descobrir exatamente que diferença é essa.
- Podemos usar um intervalo de confiança para estimar essa diferença.

# Intervalo de confiança para uma média em uma amostra pequena

- Os intervalos de confiança são sempre da forma

$$\text{estimativa pontual} \pm ME$$

# Intervalo de confiança para uma média em uma amostra pequena

- Os intervalos de confiança são sempre da forma

$$\text{estimativa pontual} \pm ME$$

- ME é sempre calculado como o produto entre o valor crítico e o SE.

## Intervalo de confiança para uma média em uma amostra pequena

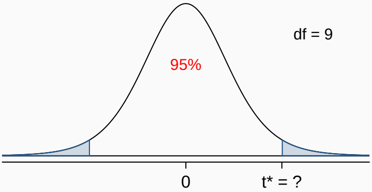
- Os intervalos de confiança são sempre da forma

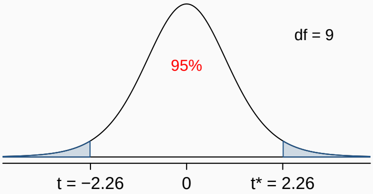
$$\text{estimativa pontual} \pm ME$$

- ME é sempre calculado como o produto entre o valor crítico e o SE.
- Uma vez que uma amostra pequena significa uma distribuição de $t$ (e não uma distribuição de $z$), o valor crítico é $t^\star$ (ao contrário de $z^\star$).

$$\text{estimativa pontual} \pm t^\star \times SE$$

# Encontrando o valor crítico $t$ ($t^\star$)




$n = 10$, $df = 10 - 1 = 9$, $t^\star$ está na interseção da linha $df = 9$ e duas probabilidades da cauda 0.05.

# Encontrando o crítico $t$ ($t^\star$)

| Uma cauda | | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
|---|---|---|---|---|---|---|
| duas caudas | | 0.200 | 0.100 | 0.050 | 0.020 | 0.010 |
| df | 6 | 1.44 | 1.94 | 2.45 | 3.14 | 3.71 |
| | 7 | 1.41 | 1.89 | 2.36 | 3.00 | 3.50 |
| | 8 | 1.40 | 1.86 | 2.31 | 2.90 | 3.36 |
| | 9 | 1.38 | 1.83 | 2.26 | 2.82 | 3.25 |
| | 10 | 1.37 | 1.81 | 2.23 | 2.76 | 3.17 |

# Encontrando o crítico $t$ ($t^{\star}$)

| Uma cauda | | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
|---|---|---|---|---|---|---|
| duas caudas | | 0.200 | 0.100 | 0.050 | 0.020 | 0.010 |
| df | 6 | 1.44 | 1.94 | 2.45 | 3.14 | 3.71 |
| | 7 | 1.41 | 1.89 | 2.36 | 3.00 | 3.50 |
| | 8 | 1.40 | 1.86 | 2.31 | 2.90 | 3.36 |
| | 9 | 1.38 | 1.83 | 2.26 | 2.82 | 3.25 |
| | 10 | 1.37 | 1.81 | 2.23 | 2.76 | 3.17 |

# Encontrando o crítico $t$ ($t^\star$)

| Uma cauda | | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
|---|---|---|---|---|---|---|
| duas caudas | | 0.200 | 0.100 | *0.050* | 0.020 | 0.010 |
| df | 6 | 1.44 | 1.94 | 2.45 | 3.14 | 3.71 |
| | 7 | 1.41 | 1.89 | 2.36 | 3.00 | 3.50 |
| | 8 | 1.40 | 1.86 | 2.31 | 2.90 | 3.36 |
| | 9 | 1.38 | 1.83 | 2.26 | 2.82 | 3.25 |
| | 10 | 1.37 | 1.81 | 2.23 | 2.76 | 3.17 |

# Encontrando o crítico $t$ ($t^\star$)

| Uma cauda | | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
|---|---|---|---|---|---|---|
| duas caudas | | 0.200 | 0.100 | 0.050 | 0.020 | 0.010 |
| df | 6 | 1.44 | 1.94 | 2.45 | 3.14 | 3.71 |
| | 7 | 1.41 | 1.89 | 2.36 | 3.00 | 3.50 |
| | 8 | 1.40 | 1.86 | 2.31 | 2.90 | 3.36 |
| | 9 | 1.38 | 1.83 | 2.26 | 2.82 | 3.25 |
| | 10 | 1.37 | 1.81 | 2.23 | 2.76 | 3.17 |

# Construindo um IC para uma média pequena de amostra

Qual dos seguintes é o cálculo correto de um intervalo de confiança de 95% para a diferença entre o fluxo de tráfego entre sexta-feira 6 e 13?

$$\bar{x}_{qtde} = 1836 \qquad s_{qtde} = 1176 \qquad n = 10 \qquad SE = 372$$

(a) $1836 \pm 1.96 \times 372$

(b) $1836 \pm 2.26 \times 372$

(c) $1836 \pm -2.26 \times 372$

(d) $1836 \pm 2.26 \times 1176$

# Construindo um IC para uma média pequena de amostra

Qual dos seguintes é o cálculo correto de um intervalo de confiança de 95% para a diferença entre o fluxo de tráfego entre sexta-feira 6 e 13?

$$\bar{x}_{qtde} = 1836 \qquad s_{qtde} = 1176 \qquad n = 10 \qquad SE = 372$$

(a) $1836 \pm 1.96 \times 372$

(b) $1836 \pm 2.26 \times 372$ $\rightarrow$ *(995, 2677)*

(c) $1836 \pm -2.26 \times 372$

(d) $1836 \pm 2.26 \times 1176$

# Interpretando o IC

Qual das seguintes opções é a *melhor* interpretação para o intervalo de confiança que acabamos de calcular?

$$\mu_{qtde:6-13} = (995, 2677)$$

Estamos 95% confiantes de que ...

(a) a diferença entre o número médio de carros na estrada na sexta-feira 6 e 13 é entre 995 e 2.677.

(b) na sexta-feira 6 há 995 a 2.677 carros a menos na estrada do que na sexta-feira 13, em média.

(c) na sexta-feira 6 há, em média, 995 carros a menos e na sexta-feira 13 há, em média, 2.677 carros a mais na estrada.

(d) na sexta-feira 13 há de 995 a 2.677 carros a menos na estrada do que na sexta-feira 6, em média.

# Interpretando o IC

Qual das seguintes opções é a *melhor* interpretação para o intervalo de confiança que acabamos de calcular?

$$\mu_{qtde:6-13} = (995, 2677)$$

Estamos 95% confiantes de que ...

(a) a diferença entre o número médio de carros na estrada na sexta-feira 6 e 13 é entre 995 e 2.677.

(b) na sexta-feira 6 há 995 a 2.677 carros a menos na estrada do que na sexta-feira 13, em média.

(c) na sexta-feira 6 há, em média, 995 carros a menos e na sexta-feira 13 há, em média, 2.677 carros a mais na estrada.

(d) *na sexta-feira 13 há de 995 a 2.677 carros a menos na estrada do que na sexta-feira 6, em média.*

# Síntese

A conclusão do teste de hipótese concorda com o que encontramos para o intervalo de confiança?

Você acha que as descobertas deste estudo sugerem que as pessoas acreditam que sexta-feira 13 é um dia de má sorte?

## Síntese

A conclusão do teste de hipótese concorda com o que encontramos para o intervalo de confiança?

*Sim, o teste de hipótese encontrou uma diferença significativa e o IC não contém o valor da hipótese nula que é igual zero.*

Você acha que as descobertas deste estudo sugerem que as pessoas acreditam que sexta-feira 13 é um dia de má sorte?

# Síntese

A conclusão do teste de hipótese concorda com o que encontramos para o intervalo de confiança?

*Sim, o teste de hipótese encontrou uma diferença significativa e o IC não contém o valor da hipótese nula que é igual zero.*

Você acha que as descobertas deste estudo sugerem que as pessoas acreditam que sexta-feira 13 é um dia de má sorte?

*Não, este é um estudo observacional. Acabamos de observar uma diferença significativa entre o número de carros na estrada nesses dois dias. Nós não testamos as crenças das pessoas.*

# Recapitulando: Inferência usando a distribuição $t$

- Se $\sigma$ for desconhecido, use a distribuição $t$ com $SE = \frac{s}{\sqrt{n}}$.

## Recapitulando: Inferência usando a distribuição $t$

- Se $\sigma$ for desconhecido, use a distribuição $t$ com $SE = \frac{s}{\sqrt{n}}$.
- Condições:
  - independência das observações (usar amostragem aleatória, e se a amostragem for sem reposição, $n <$ 10% da população)
  - nenhum desvio extremo

# Recapitulando: Inferência usando a distribuição $t$

- Teste de hipóteses:

$$T_{df} = \frac{\text{estimativa pontual – valor da hipótese nula}}{SE}, \text{ onde } df = n-1$$

## Recapitulando: Inferência usando a distribuição $t$

- Teste de hipóteses:

$$T_{df} = \frac{\text{estimativa pontual} - \text{valor da hipótese nula}}{SE}, \text{ onde } df = n{-}1$$

- Intervalo de confiança:

$$\text{estimativa pontual} \pm t^{\star}_{df} \times SE$$

# Recapitulando: Inferência usando a distribuição $t$

- Teste de hipóteses:

$$T_{df} = \frac{\text{estimativa pontual} - \text{valor da hipótese nula}}{SE}, \text{ onde } df = n - 1$$

- Intervalo de confiança:

$$\text{estimativa pontual} \pm t^{\star}_{df} \times SE$$

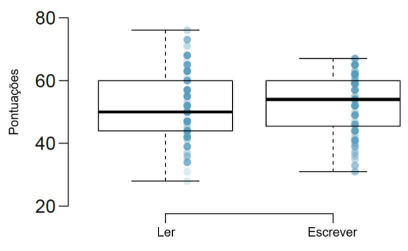
---

*Note: O exemplo que usamos foi para médias pareadas (diferença entre grupos dependentes). Tomamos a diferença entre as observações e usamos apenas essas diferenças (uma amostra) em nossa análise, portanto, a mecânica é a mesma de quando estamos trabalhando com apenas uma amostra.*

## 5.2. Dados pareados

# Prática

Foram amostradas aleatoriamente 200 observações em uma pesquisa na High School and Beyond. Os alunos fizeram um teste de leitura e escrita. À primeira vista, parece haver uma diferença entre a pontuação média do teste de leitura e escrita?



# Prática

As notas de leitura e escrita de cada aluno são independentes umas das outras?

| | id | leitura | escrita |
|---|---|---|---|
| 1 | 70 | 57 | 52 |
| 2 | 86 | 44 | 33 |
| 3 | 141 | 63 | 44 |
| 4 | 172 | 47 | 52 |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| 200 | 137 | 63 | 65 |

(a) Sim

(b) Não

## Prática

As notas de leitura e escrita de cada aluno são independentes umas das outras?

| | id | leitura | escrita |
|---|---|---|---|
| 1 | 70 | 57 | 52 |
| 2 | 86 | 44 | 33 |
| 3 | 141 | 63 | 44 |
| 4 | 172 | 47 | 52 |
| ⋮ | ⋮ | ⋮ | ⋮ |
| 200 | 137 | 63 | 65 |

(a) Sim

(b) *Não*

# Analisando dados pareados

- Quando dois conjuntos de observações têm essa correspondência especial (não independente), diz-se que são *pareados*.

# Analisando dados pareados

- Quando dois conjuntos de observações têm essa correspondência especial (não independente), diz-se que são *pareados*.
- Para analisar dados pareados, é frequentemente útil observar a diferença nos resultados de cada par de observações.

diferenca = leitura - escrita

# Analisando dados pareados

- Quando dois conjuntos de observações têm essa correspondência especial (não independente), diz-se que são *pareados*.
- Para analisar dados pareados, é frequentemente útil observar a diferença nos resultados de cada par de observações.

  diferenca = leitura - escrita
- É importante subtrairmos sempre usando uma ordem consistente.

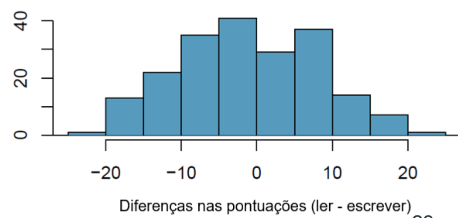
# Analisando dados pareados

- Quando dois conjuntos de observações têm essa correspondência especial (não independente), diz-se que são *pareados*.
- Para analisar dados pareados, é frequentemente útil observar a diferença nos resultados de cada par de observações.

  diferenca = leitura - escrita

- É importante subtrairmos sempre usando uma ordem consistente.

| | id | leitura | escrita | diferença |
|---|---|---|---|---|
| 1 | 70 | 57 | 52 | 5 |
| 2 | 86 | 44 | 33 | 11 |
| 3 | 141 | 63 | 44 | 19 |
| 4 | 172 | 47 | 52 | -5 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 200 | 137 | 63 | 65 | -2 |



# Estimativa de parâmetro e de ponto

- *Parâmetro de interesse:* Diferença média entre as pontuações de leitura e escrita de *todos* estudantes do ensino médio.

$$\mu_{diferenca}$$

# Estimativa de parâmetro e de ponto

- *Parâmetro de interesse:* Diferença média entre as pontuações de leitura e escrita de *todos* estudantes do ensino médio.

$$\mu_{diferenca}$$

- *Estimativa pontual:* Diferença média entre as pontuações de leitura e escrita de estudantes do ensino médio *amostrado*.

$$\bar{x}_{diferenca}$$

# Definindo as hipóteses

Se, de fato, não houve diferença entre as pontuações nos exames de leitura e escrita, como você esperaria que a diferença média fosse?

Se, de fato, não houve diferença entre as pontuações nos exames de leitura e escrita, como você esperaria que a diferença média fosse?

*0*

# Definindo as hipóteses

Se, de fato, não houve diferença entre as pontuações nos exames de leitura e escrita, como você esperaria que a diferença média fosse?

*0*

Quais são as hipóteses para testar se existe uma diferença entre as pontuações médias de leitura e escrita?

## Definindo as hipóteses

Se, de fato, não houve diferença entre as pontuações nos exames de leitura e escrita, como você esperaria que a diferença média fosse?

*0*

Quais são as hipóteses para testar se existe uma diferença entre as pontuações médias de leitura e escrita?

$H_0$: Não há diferença entre a pontuação média de leitura e escrita.

$$\mu_{diferenca} = 0$$

$H_A$: Existe uma diferença entre a pontuação média de leitura e escrita.

$$\mu_{diferenca} \neq 0$$

## Nada de novo aqui

- A análise não é diferente do que fizemos antes.
- Temos dados de *uma* amostra: diferenças.
- Estamos testando para ver se a diferença média é diferente de 0.

# Verificação de suposições & condições

## Qual dos seguintes é verdadeiro?

(a) Como os alunos são amostrados aleatoriamente e são menores que 10% de todos os alunos do ensino médio, podemos supor que a diferença entre os escores de leitura e escrita de um aluno na amostra é independente do outro.

(b) A distribuição das diferenças é bimodal, portanto não podemos continuar com o teste de hipóteses.

(c) Para que as diferenças sejam aleatórias, devemos ter amostras com reposição.

(d) Como os alunos são amostrados aleatoriamente e são menores que 10% de todos os alunos, podemos supor que a distribuição amostral da diferença média será quase normal.

# Verificação de suposições & condições

## Qual dos seguintes é verdadeiro?

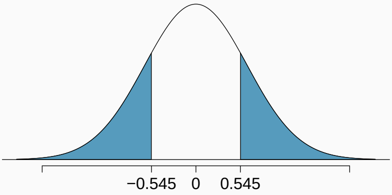
(a) *Como os alunos são amostrados aleatoriamente e são menores que 10% de todos os alunos do ensino médio, podemos supor que a diferença entre os escores de leitura e escrita de um aluno na amostra é independente do outro.*

(b) A distribuição das diferenças é bimodal, portanto não podemos continuar com o teste de hipóteses.

(c) Para que as diferenças sejam aleatórias, devemos ter amostras com reposição.

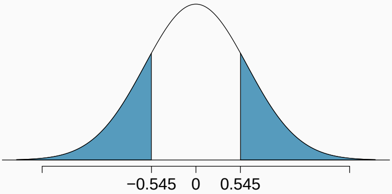
(d) Como os alunos são amostrados aleatoriamente e são menores que 10% de todos os alunos, podemos supor que a distribuição amostral da diferença média será quase normal.

A diferença média observada entre as duas pontuações é de -0,545 pontos e o desvio padrão da diferença é de 8,887 pontos. Esses dados fornecem evidências convincentes de uma diferença entre as pontuações médias dos dois exames? Usar $\alpha = 0.05$.



## Calculando a estatística de teste e o valor p

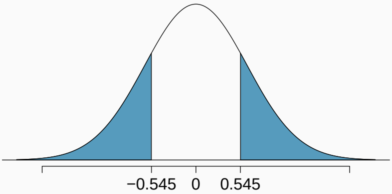
A diferença média observada entre as duas pontuações é de -0,545 pontos e o desvio padrão da diferença é de 8,887 pontos. Esses dados fornecem evidências convincentes de uma diferença entre as pontuações médias dos dois exames? Usar $\alpha = 0.05$.




$$
\begin{aligned}
T &= \frac{-0.545 - 0}{\frac{8.887}{\sqrt{200}}} \\
&= \frac{-0.545}{0.628} = -0.87 \\
df &= 200 - 1 = 199
\end{aligned}
$$

# Calculando a estatística de teste e o valor p

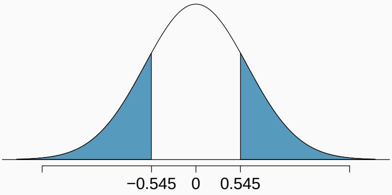
A diferença média observada entre as duas pontuações é de -0,545 pontos e o desvio padrão da diferença é de 8,887 pontos. Esses dados fornecem evidências convincentes de uma diferença entre as pontuações médias dos dois exames? Usar $\alpha = 0.05$.




$$
\begin{aligned}
T &= \frac{-0.545 - 0}{\frac{8.887}{\sqrt{200}}} \\
&= \frac{-0.545}{0.628} = -0.87 \\
df &= 200 - 1 = 199 \\
valor - p &= 0.1927 \times 2 = 0.3854
\end{aligned}
$$

## Calculando a estatística de teste e o valor p

A diferença média observada entre as duas pontuações é de -0,545 pontos e o desvio padrão da diferença é de 8,887 pontos. Esses dados fornecem evidências convincentes de uma diferença entre as pontuações médias dos dois exames? Usar $\alpha = 0.05$.




$$\begin{aligned} T &= \frac{-0.545 - 0}{\frac{8.887}{\sqrt{200}}} \\ &= \frac{-0.545}{0.628} = -0.87 \\ df &= 200 - 1 = 199 \\ valor - p &= 0.1927 \times 2 = 0.3854 \end{aligned}$$

Como o valor de p $>$ 0,05, não é possível rejeitar, os dados <u>não</u> fornecem evidências convincentes de uma diferença entre as pontuações médias de leitura e escrita.

# Interpretação do valor p

## Qual das seguintes é a interpretação correta do valor p?

(a) Probabilidade de que as pontuações médias nos exames de leitura e escrita sejam iguais.

(b) Probabilidade de que as pontuações médias nos exames de leitura e escrita sejam diferentes.

(c) Probabilidade de obter uma amostra aleatória de 200 alunos, onde a diferença média entre as pontuações de leitura e escrita é de pelo menos 0.545 (em qualquer direção), se de fato a diferença média real entre as pontuações é 0.

(d) Probabilidade de rejeitar incorretamente a hipótese nula se, de fato, a hipótese nula for verdadeira.

# Interpretação do valor p

Qual das seguintes é a interpretação correta do valor p?

(a) Probabilidade de que as pontuações médias nos exames de leitura e escrita sejam iguais.

(b) Probabilidade de que as pontuações médias nos exames de leitura e escrita sejam diferentes.

(c) *Probabilidade de obter uma amostra aleatória de 200 alunos, onde a diferença média entre as pontuações de leitura e escrita é de pelo menos 0.545 (em qualquer direção), se de fato a diferença média real entre as pontuações é 0.*

(d) Probabilidade de rejeitar incorretamente a hipótese nula se, de fato, a hipótese nula for verdadeira.

# TH $\leftrightarrow$ IC

Suponha que fôssemos construir um intervalo de confiança de 95% para a diferença média entre as pontuações de leitura e escrita. Você esperaria que esse intervalo incluísse 0?

(a) sim
(b) não
(c) não posso dizer a partir da informação dada

# TH ↔ IC

Suponha que fôssemos construir um intervalo de confiança de 95% para a diferença média entre as pontuações de leitura e escrita. Você esperaria que esse intervalo incluísse 0?

(a) *sim*
(b) não
(c) não posso dizer a partir da informação dada

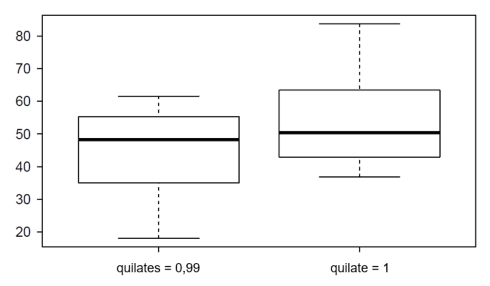
$$\begin{aligned}
-0.545 \pm 1.97 \frac{8.887}{\sqrt{200}} &= -0.545 \pm 1.97 \times 0.628 \\
&= -0.545 \pm 1.24 \\
&= (-1.785, 0.695)
\end{aligned}$$

# 5.3. Diferença de duas médias

# Diamantes

- Pesos de diamantes são medidos em quilates.
- 1 quilate = 100 pontos, 0,99 quilates = 99 pontos, etc.
- A diferença entre o tamanho de um diamante de 0,99 quilates e um diamante de 1 quilate é indetectável a olho nu, mas o preço de um diamante de 1 quilate tende a ser maior do que o preço de um diamante de 0,99 quilates?
- Vamos testar para ver se existe uma diferença entre os preços médios de diamantes de 0,99 e 1 quilates.
- Para podermos comparar unidades equivalentes, dividimos os preços de 0,99 quilates em 99 e de 1 quilate em 100, e comparamos os preços médios em pontos.

# Dados




| | *0.99 quilate* | *1 quilate* |
|---|---|---|
| | pt99 | pt100 |
| $\bar{x}$ | 44.50 | 53.43 |
| $s$ | 13.32 | 12.22 |
| $n$ | 23 | 30 |

Estes dados são uma amostra aleatória do conjunto de dados `diamonds` no pacote do R `ggplot2`.

# Estimativa de parâmetro e ponto

- *Parâmetro de interesse:* Diferença média entre os preços em pontos de *todos* os diamantes de 0,99 quilates e de 1 quilate.

$$\mu_{pt99} - \mu_{pt100}$$

# Estimativa de parâmetro e ponto

- *Parâmetro de interesse:* Diferença média entre os preços em pontos de *todos* os diamantes de 0,99 quilates e de 1 quilate.

$$\mu_{pt99} - \mu_{pt100}$$

- *Estimativa pontual:* Diferença média entre os preços em pontos da *amostra* de diamantes de 0,99 quilates e 1 quilate.

$$\bar{x}_{pt99} - \bar{x}_{pt100}$$

# Hipótese

Qual das alternativas a seguir é o conjunto correto de hipóteses para testar se o preço médio dos diamantes de 1 quilate ($_{pt100}$) é maior que o preço médio dos diamantes de 0,99 quilates ($_{pt99}$)?

(a) $H_0 : \mu_{pt99} = \mu_{pt100}$

$H_A : \mu_{pt99} \neq \mu_{pt100}$

(b) $H_0 : \mu_{pt99} = \mu_{pt100}$

$H_A : \mu_{pt99} > \mu_{pt100}$

(c) $H_0 : \mu_{pt99} = \mu_{pt100}$

$H_A : \mu_{pt99} < \mu_{pt100}$

(d) $H_0 : \bar{x}_{pt99} = \bar{x}_{pt100}$

$H_A : \bar{x}_{pt99} < \bar{x}_{pt100}$

## Hipótese

Qual das alternativas a seguir é o conjunto correto de hipóteses para testar se o preço médio dos diamantes de 1 quilate ($_{pt100}$) é maior que o preço médio dos diamantes de 0,99 quilates ($_{pt99}$)?

(a) $H_0 : \mu_{pt99} = \mu_{pt100}$

$H_A : \mu_{pt99} \neq \mu_{pt100}$

(b) $H_0 : \mu_{pt99} = \mu_{pt100}$

$H_A : \mu_{pt99} > \mu_{pt100}$

(c) $H_0 : \mu_{pt99} = \mu_{pt100}$

$H_A : \mu_{pt99} < \mu_{pt100}$

(d) $H_0 : \bar{x}_{pt99} = \bar{x}_{pt100}$

$H_A : \bar{x}_{pt99} < \bar{x}_{pt100}$

# Condições

Qual dos seguintes itens <u>não</u> precisa ser satisfeito para conduzir este teste de hipótese usando métodos teóricos?

(a) O preço do ponto de um diamante de 0,99 quilates na amostra deve ser independente do outro, e o preço em pontos de um diamante de 1 quilate deve ser independente do outro.

(b) Os preços pontuais de 0,99 quilates e 1 quilate de diamantes na amostra devem ser independentes.

(c) Distribuições de preços pontuais de 0,99 e 1 quilate de diamantes não devem ser extremamente assimétrico.

(d) Ambos os tamanhos de amostra devem ser pelo menos 30.

# Condições

Qual dos seguintes itens não precisa ser satisfeito para conduzir este teste de hipótese usando métodos teóricos?

(a) O preço do ponto de um diamante de 0,99 quilates na amostra deve ser independente do outro, e o preço em pontos de um diamante de 1 quilate deve ser independente do outro.

(b) Os preços pontuais de 0,99 quilates e 1 quilate de diamantes na amostra devem ser independentes.

(c) Distribuições de preços pontuais de 0,99 e 1 quilate de diamantes não devem ser extremamente assimétrico.

(d) *Ambos os tamanhos de amostra devem ser pelo menos 30.*

# Estatística de teste

## Teste estatístico para inferência sobre a diferença de duas médias amostrais pequenas

A estatística de teste para inferência sobre a diferença de duas médias onde $\sigma_1$ e $\sigma_2$ são desconhecidos é a estatística $T$.

$$T_{df} = \frac{\text{estimativa pontual} - \text{valor da hipótese nula}}{SE}$$

onde

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \qquad \text{e} \qquad df = min(n_1 - 1, n_2 - 1)$$

---

*Note: O cálculo do $df$ é realmente muito mais complicado. Para simplificar, usaremos a fórmula acima para estimar o verdadeiro $df$ ao conduzir a análise à mão.*

# Estatística de teste (cont.)

| | *0.99 quilate* | *1 quilate* |
|---|---|---|
| | pt99 | pt100 |
| $\bar{x}$ | 44.50 | 53.43 |
| $s$ | 13.32 | 12.22 |
| $n$ | 23 | 30 |

*Neste contexto...*

# Estatística de teste (cont.)

| | *0.99 quilate* | *1 quilate* |
|---|---|---|
| | pt99 | pt100 |
| $\bar{x}$ | 44.50 | 53.43 |
| $s$ | 13.32 | 12.22 |
| $n$ | 23 | 30 |

*Neste contexto...*

$$T = \frac{\text{estimativa pontual} - \text{valor da hipótese nula}}{SE}$$

# Estatística de teste (cont.)

| | *0.99 quilate* | *1 quilate* |
|---|---|---|
| | pt99 | pt100 |
| $\bar{x}$ | 44.50 | 53.43 |
| $s$ | 13.32 | 12.22 |
| $n$ | 23 | 30 |

*Neste contexto...*

$$\begin{aligned} T &= \frac{\text{estimativa pontual} - \text{valor da hipótese nula}}{SE} \\ &= \frac{(44.50 - 53.43) - 0}{\sqrt{\frac{13.32^2}{23} + \frac{12.22^2}{30}}} \end{aligned}$$

# Estatística de teste (cont.)

| | *0.99 quilate* | *1 quilate* |
|---|---|---|
| | pt99 | pt100 |
| $\bar{x}$ | 44.50 | 53.43 |
| $s$ | 13.32 | 12.22 |
| $n$ | 23 | 30 |

*Neste contexto...*

$$
\begin{aligned}
T &= \frac{\text{estimativa pontual} - \text{valor da hipótese nula}}{SE} \\
&= \frac{(44.50 - 53.43) - 0}{\sqrt{\frac{13.32^2}{23} + \frac{12.22^2}{30}}} \\
&= \frac{-8.93}{3.56}
\end{aligned}
$$

# Estatística de teste (cont.)

| | *0.99 quilate* | *1 quilate* |
|---|---|---|
| | pt99 | pt100 |
| $\bar{x}$ | 44.50 | 53.43 |
| $s$ | 13.32 | 12.22 |
| $n$ | 23 | 30 |

*Neste contexto...*

$$
\begin{aligned}
T &= \frac{\text{estimativa pontual} - \text{valor da hipótese nula}}{SE} \\
&= \frac{(44.50 - 53.43) - 0}{\sqrt{\frac{13.32^2}{23} + \frac{12.22^2}{30}}} \\
&= \frac{-8.93}{3.56} \\
&= -2.508
\end{aligned}
$$

## Estatística de teste (cont.)

Qual dos seguintes é o $df$ correto para este teste de hipótese?

(a) 22

(b) 23

(c) 30

(d) 29

(e) 52

## Estatística de teste (cont.)

Qual dos seguintes é o $df$ correto para este teste de hipótese?

(a) *22*

(b) 23

(c) 30

(d) 29

(e) 52

$$\rightarrow df = min(n_{pt99} - 1, n_{pt100} - 1)$$
$$= min(23 - 1, 30 - 1)$$
$$= min(22, 29) = 22$$

# valor-p

Qual dos seguintes é o valor p correto para este teste de hipótese?

$$T = -2.508 \qquad df = 22$$

(a) entre 0.005 e 0.01

(b) entre 0.01 e 0.025

(c) entre 0.02 e 0.05

(d) entre 0.01 e 0.02

| uma cauda | | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
|---|---|---|---|---|---|---|
| duas caudas | | 0.200 | 0.100 | 0.050 | 0.020 | 0.010 |
| df | 21 | 1.32 | 1.72 | 2.08 | 2.52 | 2.83 |
| | 22 | 1.32 | 1.72 | 2.07 | 2.51 | 2.82 |
| | 23 | 1.32 | 1.71 | 2.07 | 2.50 | 2.81 |
| | 24 | 1.32 | 1.71 | 2.06 | 2.49 | 2.80 |
| | 25 | 1.32 | 1.71 | 2.06 | 2.49 | 2.79 |

# valor-p

Qual dos seguintes é o valor p correto para este teste de hipótese?

$$T = -2.508 \qquad df = 22$$

(a) entre 0.005 e 0.01

(b) *entre 0.01 e 0.025*

(c) entre 0.02 e 0.05

(d) entre 0.01 e 0.02

| | uma cauda | 0.100 | 0.050 | *0.025* | *0.010* | 0.005 |
|---|---|---|---|---|---|---|
| | duas caudas | 0.200 | 0.100 | 0.050 | 0.020 | 0.010 |
| df | 21 | 1.32 | 1.72 | 2.08 | 2.52 | 2.83 |
| | 22 | 1.32 | 1.72 | *2.07* | *2.51* | 2.82 |
| | 23 | 1.32 | 1.71 | 2.07 | 2.50 | 2.81 |
| | 24 | 1.32 | 1.71 | 2.06 | 2.49 | 2.80 |
| | 25 | 1.32 | 1.71 | 2.06 | 2.49 | 2.79 |

# Síntese

Qual é a conclusão do teste de hipóteses? Como essa conclusão mudaria seu comportamento se você comprasse diamantes?

# Síntese

Qual é a conclusão do teste de hipóteses? Como essa conclusão mudaria seu comportamento se você comprasse diamantes?

- *O valor p é pequeno, portanto, rejeite $H_0$. Os dados fornecem evidências convincentes para sugerir que o preço em pontos de 0,99 quilates é menor do que o preço em pontos de 1 quilate de diamantes.*
- *Talvez compre um diamante de 0,99 quilates? É quase um quilate, mas é significativamente mais barato.*

# Nível de confiança equivalente

Qual é o nível de confiança equivalente para um teste de hipótese unilateral para $\alpha = 0.05$?

(a) 90%

(b) 92.5%

(c) 95%

(d) 97.5%

# Nível de confiança equivalente

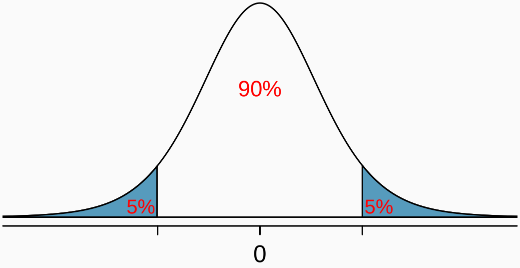
Qual é o nível de confiança equivalente para um teste de hipótese unilateral para $\alpha = 0.05$?

(a) *90%*

(b) 92.5%

(c) 95%

(d) 97.5%



## Valor crítico

Qual é o $t^\star$ apropriado para um intervalo de confiança para a diferença média entre os preços pontuais de 0,99 e 1 quilate de diamantes?

(a) 1.32
(b) 1.72
(c) 2.07
(d) 2.82

| | | | | | | |
|---|---|---|---|---|---|---|
| uma cauda | | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
| duas caudas | | 0.200 | 0.100 | 0.050 | 0.020 | 0.010 |
| df | 21 | 1.32 | 1.72 | 2.08 | 2.52 | 2.83 |
| | 22 | 1.32 | 1.72 | 2.07 | 2.51 | 2.82 |
| | 23 | 1.32 | 1.71 | 2.07 | 2.50 | 2.81 |
| | 24 | 1.32 | 1.71 | 2.06 | 2.49 | 2.80 |
| | 25 | 1.32 | 1.71 | 2.06 | 2.49 | 2.79 |

# Valor crítico

Qual é o $t^\star$ apropriado para um intervalo de confiança para a diferença média entre os preços pontuais de 0,99 e 1 quilate de diamantes?

(a) 1.32
(b) *1.72*
(c) 2.07
(d) 2.82

| uma cauda | | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
|---|---|---|---|---|---|---|
| duas caudas | | 0.200 | *0.100* | 0.050 | 0.020 | 0.010 |
| df | 21 | 1.32 | 1.72 | 2.08 | 2.52 | 2.83 |
| | 22 | 1.32 | *1.72* | 2.07 | 2.51 | 2.82 |
| | 23 | 1.32 | 1.71 | 2.07 | 2.50 | 2.81 |
| | 24 | 1.32 | 1.71 | 2.06 | 2.49 | 2.80 |
| | 25 | 1.32 | 1.71 | 2.06 | 2.49 | 2.79 |

# Intervalo de confiança

Calcule o intervalo e interprete-o no contexto.

Calcule o intervalo e interprete-o no contexto.

$$\textit{estimativa pontual} \pm ME$$

# Intervalo de confiança

Calcule o intervalo e interprete-o no contexto.

$$\textit{estimativa pontual} \pm ME$$

$$(\bar{x}_{pt99} - \bar{x}_{pt1}) \pm t^{\star}_{df} \times SE \quad = \quad (44.50 - 53.43) \pm 1.72 \times 3.56$$

# Intervalo de confiança

Calcule o intervalo e interprete-o no contexto.

$$\textit{estimativa pontual} \pm ME$$

$$\begin{aligned}(\bar{x}_{pt99} - \bar{x}_{pt1}) \pm t^\star_{df} \times SE &= (44.50 - 53.43) \pm 1.72 \times 3.56 \\ &= -8.93 \pm 6.12\end{aligned}$$

# Intervalo de confiança

Calcule o intervalo e interprete-o no contexto.

$$\textit{estimativa pontual} \pm ME$$

$$\begin{aligned}
(\bar{x}_{pt99} - \bar{x}_{pt1}) \pm t^{\star}_{df} \times SE &= (44.50 - 53.43) \pm 1.72 \times 3.56 \\
&= -8.93 \pm 6.12 \\
&= (-15.05, -2.81)
\end{aligned}$$

# Intervalo de confiança

## Calcule o intervalo e interprete-o no contexto.

$$\textit{estimativa pontual} \pm ME$$

$$\begin{aligned}
(\bar{x}_{pt99} - \bar{x}_{pt1}) \pm t^{\star}_{df} \times SE &= (44.50 - 53.43) \pm 1.72 \times 3.56 \\
&= -8.93 \pm 6.12 \\
&= (-15.05, -2.81)
\end{aligned}$$

*Temos 90% de confiança de que o preço médio de um diamante de 0,99 quilates é \$ 15,05 a \$ 2,81 abaixo do preço médio de um diamante de 1 quilate.*

## Recapitulação: Inferência usando diferença de duas médias amostrais pequenas

- Se $\sigma_1$ ou $\sigma_2$ for desconhecido, a diferença entre a amostra segue uma distribuição $t$ com $SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_1}}$.

## Recapitulação: Inferência usando diferença de duas médias amostrais pequenas

- Se $\sigma_1$ ou $\sigma_2$ for desconhecido, a diferença entre a amostra segue uma distribuição $t$ com $SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_1}}$.
- Condições:
  - independência dentro dos grupos (frequentemente verificada por uma amostra aleatória, e se amostragem sem reposição, $n < 10$ % da população) e entre grupos.
  - nenhum desvio extremo em nenhum dos grupos.

# Recapitulação: Inferência usando diferença de duas médias amostrais pequenas

- Se $\sigma_1$ ou $\sigma_2$ for desconhecido, a diferença entre a amostra segue uma distribuição $t$ com $SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_1}}$.
- Condições:
  - independência dentro dos grupos (frequentemente verificada por uma amostra aleatória, e se amostragem sem reposição, $n <$ 10 % da população) e entre grupos.
  - nenhum desvio extremo em nenhum dos grupos.
- Teste de hipóteses:

$$T_{df} = \frac{\text{estimativa pontual} - \text{valor da hipótese nula}}{SE}$$

onde df = min(n$_1 - 1, n_2 - 1)$

## Recapitulação: Inferência usando diferença de duas médias amostrais pequenas

- Se $\sigma_1$ ou $\sigma_2$ for desconhecido, a diferença entre a amostra segue uma distribuição $t$ com $SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_1}}$.
- Condições:
  - independência dentro dos grupos (frequentemente verificada por uma amostra aleatória, e se amostragem sem reposição, $n <$ 10 % da população) e entre grupos.
  - nenhum desvio extremo em nenhum dos grupos.
- Teste de hipóteses:

$$T_{df} = \frac{\text{estimativa pontual} - \text{valor da hipótese nula}}{SE}$$

onde df = min(n$_1 - 1, n_2 - 1)$

- Intervalo de confiança:

$$\text{estimativa pontual} \pm t^{\star}_{df} \times SE$$

## 5.4. Calculando o poder para um teste de 2 amostras

| | | **Decisão** | |
|---|---|---|---|
| | | não rejeitar $H_0$ | rejeitar $H_0$ |
| **Verdade** | $H_0$ verdade | | |
| | $H_A$ verdade | | |

| | | **Decisão** | |
|---|---|---|---|
| | | não rejeitar $H_0$ | rejeitar $H_0$ |
| **Verdade** | $H_0$ verdade | | *Erro tipo 1, $\alpha$* |
| | $H_A$ verdade | | |

- O erro tipo 1 acontece quando se rejeita $H_0$ quando ela é verdadeira, e a probabilidade de fazer isso é $\alpha$ (nível de significância).

| | | **Decisão** | |
|---|---|---|---|
| | | não rejeitar $H_0$ | rejeitar $H_0$ |
| **Verdade** | $H_0$ verdade | | *Erro tipo 1, $\alpha$* |
| | $H_A$ verdade | *Erro tipo 2, $\beta$* | |

- O erro tipo 1 acontece quando se rejeita $H_0$ quando ela é verdadeira, e a probabilidade de fazer isso é $\alpha$ (nível de significância).
- O erro tipo 2 acontece quand se falha em rejeitar $H_0$ ela não é verdadeira, e a probabilidade de fazer isso é $\beta$ (um pouco mais complicado de calcular).

| | | **Decisão** | |
|---|---|---|---|
| | | não rejeitar $H_0$ | rejeitar $H_0$ |
| **Verdade** | $H_0$ verdade | $1-\alpha$ | *Erro tipo 1,* $\alpha$ |
| | $H_A$ verdade | *Erro tipo 2,* $\beta$ | |

- O erro tipo 1 acontece quando se rejeita $H_0$ quando ela é verdadeira, e a probabilidade de fazer isso é $\alpha$ (nível de significância).
- O erro tipo 2 acontece quand se falha em rejeitar $H_0$ ela não é verdadeira, e a probabilidade de fazer isso é $\beta$ (um pouco mais complicado de calcular).
- *Poder* de um teste é a probabilidade de rejeitar corretamente $H_0$, e a probabilidade de fazer isso é $1-\beta$.

| | | **Decisão** | |
|---|---|---|---|
| | | não rejeitar $H_0$ | rejeitar $H_0$ |
| **Verdade** | $H_0$ verdade | $1 - \alpha$ | *Erro tipo 1,* $\alpha$ |
| | $H_A$ verdade | *Erro tipo 2,* $\beta$ | *poder,* $1 - \beta$ |

- O erro tipo 1 acontece quando se rejeita $H_0$ quando ela é verdadeira, e a probabilidade de fazer isso é $\alpha$ (nível de significância).
- O erro tipo 2 acontece quand se falha em rejeitar $H_0$ ela não é verdadeira, e a probabilidade de fazer isso é $\beta$ (um pouco mais complicado de calcular).
- *Poder* de um teste é a probabilidade de rejeitar corretamente $H_0$, e a probabilidade de fazer isso é $1 - \beta$.
- Nos testes de hipóteses, queremos manter $\alpha$ e $\beta$ baixo, mas existem compensações inerentes.

## Taxa de erro tipo 2

Se a hipótese alternativa é realmente verdadeira, qual é a chance de fazermos um Erro Tipo 2, ou seja, deixamos de rejeitar a hipótese nula mesmo quando deveríamos rejeitá-la?

- A resposta não é obvia.
- Se a média real da população estiver muito próxima do valor da hipótese nula, será difícil detectar uma diferença (e rejeitar $H_0$).
- Se a média real da população for muito diferente do valor da hipótese nula, será mais fácil detectar uma diferença.
- Claramente, $\beta$ depende do *tamanho do efeito* ($\delta$)

# Exemplo - Pressão Arterial (PA), hipóteses

Suponha que uma empresa farmacêutica tenha desenvolvido uma nova droga para baixar a pressão sanguínea e esteja preparando um ensaio clínico para testar a eficácia da droga. Eles recrutam pessoas que estão tomando uma medicação padrão de pressão sanguínea, e metade dos indivíduos recebe a nova droga (tratamento) e a outra metade continua a tomar a medicação atual com pílulas genéricas para garantir a cegueira (controle). Quais são as hipóteses para um teste de hipóteses bilateral neste contexto?

## Exemplo - Pressão Arterial (PA), hipóteses

Suponha que uma empresa farmacêutica tenha desenvolvido uma nova droga para baixar a pressão sanguínea e esteja preparando um ensaio clínico para testar a eficácia da droga. Eles recrutam pessoas que estão tomando uma medicação padrão de pressão sanguínea, e metade dos indivíduos recebe a nova droga (tratamento) e a outra metade continua a tomar a medicação atual com pílulas genéricas para garantir a cegueira (controle). Quais são as hipóteses para um teste de hipóteses bilateral neste contexto?

$$H_0 : \mu_{tratamento} - \mu_{controle} = 0$$
$$H_A : \mu_{tratamento} - \mu_{controle} \neq 0$$

# Exemplo - BP, erro padrão

Suponha que os pesquisadores gostariam de executar o ensaio clínico em pacientes com pressão arterial sistólica entre 140 e 180 mmHg. Suponha que estudos publicados anteriormente sugiram que o desvio padrão das pressões sanguíneas dos pacientes é de cerca de 12 mmHg e a distribuição das pressões sanguíneas dos pacientes será aproximadamente simétrica. Se tivéssemos 100 pacientes por grupo, qual seria o erro padrão aproximado para diferença nas médias amostrais dos grupos de tratamento e controle?

# Exemplo - BP, erro padrão

Suponha que os pesquisadores gostariam de executar o ensaio clínico em pacientes com pressão arterial sistólica entre 140 e 180 mmHg. Suponha que estudos publicados anteriormente sugiram que o desvio padrão das pressões sanguíneas dos pacientes é de cerca de 12 mmHg e a distribuição das pressões sanguíneas dos pacientes será aproximadamente simétrica. Se tivéssemos 100 pacientes por grupo, qual seria o erro padrão aproximado para diferença nas médias amostrais dos grupos de tratamento e controle?

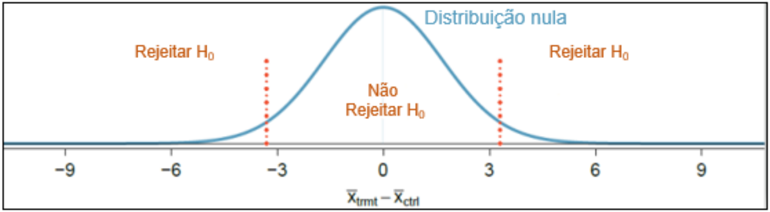
$$SE = \sqrt{\frac{12^2}{100} + \frac{12^2}{100}} = 1.70$$

## Exemplo - BP, tamanho de efeito mínimo necessário para rejeitar $H_0$

Para que valores da diferença entre as médias observadas da pressão arterial nos grupos tratamento e controle (tamanho do efeito) rejeitaríamos a hipótese nula no nível de significância de 5%?

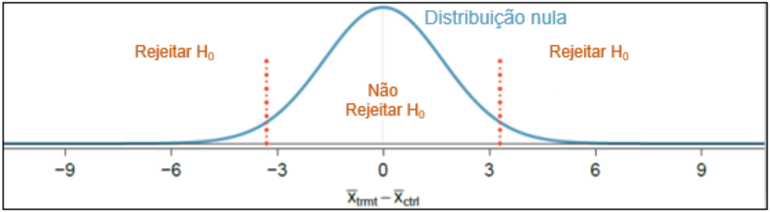
## Exemplo - BP, tamanho de efeito mínimo necessário para rejeitar $H_0$

Para que valores da diferença entre as médias observadas da pressão arterial nos grupos tratamento e controle (tamanho do efeito) rejeitaríamos a hipótese nula no nível de significância de 5%?



# Exemplo - BP, tamanho de efeito mínimo necessário para rejeitar $H_0$

Para que valores da diferença entre as médias observadas da pressão arterial nos grupos tratamento e controle (tamanho do efeito) rejeitaríamos a hipótese nula no nível de significância de 5%?




A diferença deve ser pelo menos

$$1.96 * 1.70 = 3.332$$

ou no máximo

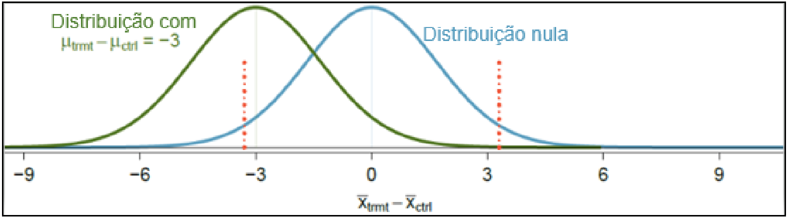
$$-1.96 * 1.70 = 3.332.$$

# Exemplo - BP, poder

Suponha que os pesquisadores da empresa se preocupem em encontrar qualquer efeito na pressão arterial que seja de 3 mmHg ou maior em relação à medicação padrão. Qual é o poder do teste que pode detectar esse efeito?

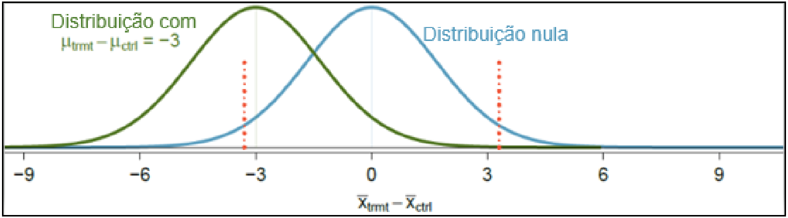
# Exemplo - BP, poder

Suponha que os pesquisadores da empresa se preocupem em encontrar qualquer efeito na pressão arterial que seja de 3 mmHg ou maior em relação à medicação padrão. Qual é o poder do teste que pode detectar esse efeito?



# Exemplo - BP, poder

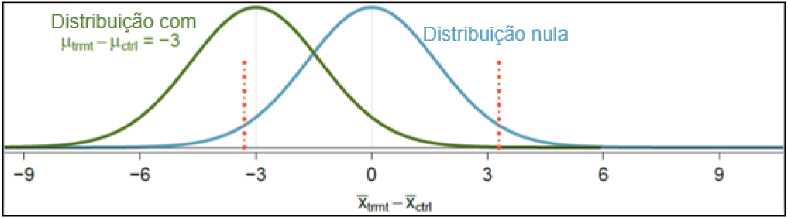
Suponha que os pesquisadores da empresa se preocupem em encontrar qualquer efeito na pressão arterial que seja de 3 mmHg ou maior em relação à medicação padrão. Qual é o poder do teste que pode detectar esse efeito?




$$Z = \frac{-3.332 - (-3)}{1.70} = -0.20$$

# Exemplo - BP, poder

Suponha que os pesquisadores da empresa se preocupem em encontrar qualquer efeito na pressão arterial que seja de 3 mmHg ou maior em relação à medicação padrão. Qual é o poder do teste que pode detectar esse efeito?




$$Z = \frac{-3.332 - (-3)}{1.70} = -0.20$$

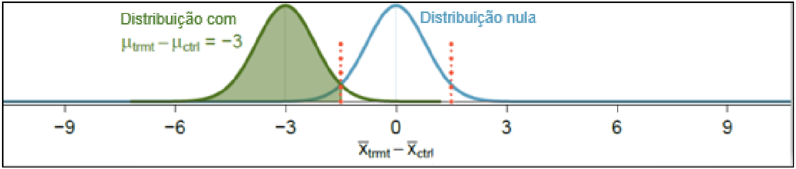
$$P(Z < -0.20) = 0.4207$$

## Exemplo - BP, tamanho de amostra necessário para 80% de poder

Qual tamanho de amostra levará a uma poder de 80% para este teste?

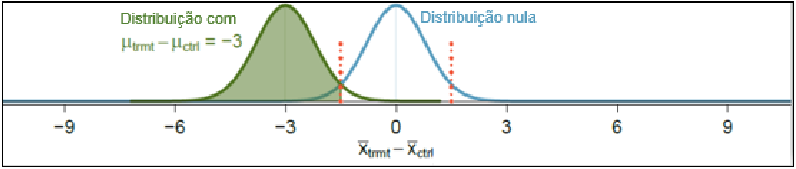
# Exemplo - BP, tamanho de amostra necessário para 80% de poder

Qual tamanho de amostra levará a uma poder de 80% para este teste?



# Exemplo - BP, tamanho de amostra necessário para 80% de poder

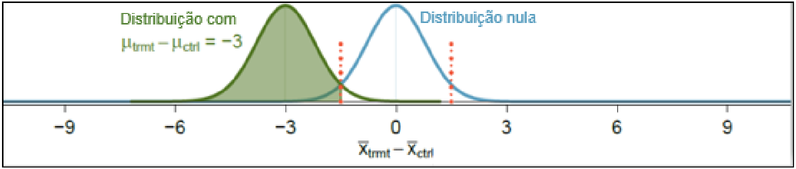
Qual tamanho de amostra levará a uma poder de 80% para este teste?




$$SE = \frac{3}{2.8} = 1.07142$$

# Exemplo - BP, tamanho de amostra necessário para 80% de poder

Qual tamanho de amostra levará a uma poder de 80% para este teste?

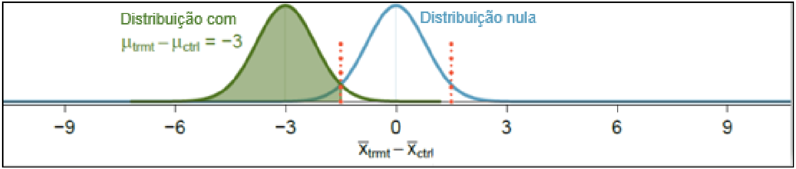



$$SE = \frac{3}{2.8} = 1.07142$$

$$1.07142 = \sqrt{\frac{12^2}{n} + \frac{12^2}{n}}$$

# Exemplo - BP, tamanho de amostra necessário para 80% de poder

Qual tamanho de amostra levará a uma poder de 80% para este teste?



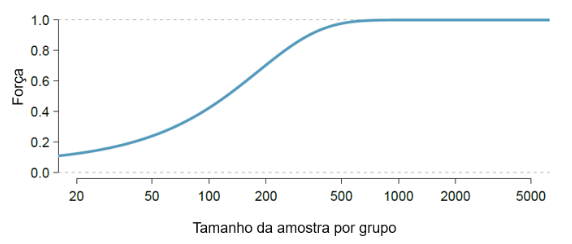

$$SE = \frac{3}{2.8} = 1.07142$$

$$1.07142 = \sqrt{\frac{12^2}{n} + \frac{12^2}{n}}$$

$$n = 250.88 \rightarrow n \geq 251$$

# Recapitulando

- Calcule o tamanho de amostra necessário para um nível desejado de poder.
- Calcule a poder para uma faixa de tamanhos de amostra e, em seguida, escolha o tamanho da amostra que produz a poder desejado (geralmente 80% ou 90%).



# Alcançar o poder desejado

Existem várias maneiras de aumentar o poder (e, portanto, diminuir a taxa de erro do tipo 2):

# Alcançar o poder desejado

Existem várias maneiras de aumentar o poder (e, portanto, diminuir a taxa de erro do tipo 2):

1. Aumentar o tamanho da amostra

# Alcançar o poder desejado

Existem várias maneiras de aumentar o poder (e, portanto, diminuir a taxa de erro do tipo 2):

1. Aumentar o tamanho da amostra
2. Diminuir o desvio padrão da amostra, que essencialmente tem o mesmo efeito que aumentar o tamanho da amostra (diminuirá o erro padrão). Com um $s$ menor, temos uma chance maior de distinguir o valor nulo da estimativa pontual observada. Isso é difícil de assegurar, mas um processo de medição cauteloso e a limitação da população para que ela seja mais homogênea podem ajudar.

## Alcançar o poder desejado

3. Aumente $\alpha$, o que aumentará a probabilidade de rejeitar $H_0$ (mas observe que isso tem o efeito colateral de aumentar a taxa de erro de tipo 1).

## Alcançar o poder desejado

3. Aumente $\alpha$, o que aumentará a probabilidade de rejeitar $H_0$ (mas observe que isso tem o efeito colateral de aumentar a taxa de erro de tipo 1).

4. Considere um tamanho de efeito maior. Se a verdadeira média da população estiver na hipótese alternativa, mas próxima do valor nulo, será mais difícil detectar uma diferença.

# 5.5. Comparando médias com ANOVA

# Prática


- O rio Wolf no Tennessee passa por um local abandonado antes utilizado pela indústria de pesticidas para despejar resíduos, incluindo clordano (pesticida), aldrina e dieldrina (ambos inseticidas).

# Prática


- O rio Wolf no Tennessee passa por um local abandonado antes utilizado pela indústria de pesticidas para despejar resíduos, incluindo clordano (pesticida), aldrina e dieldrina (ambos inseticidas).
- Estes compostos orgânicos altamente tóxicos podem causar vários tipos de câncer e problemas congênitos.

## Prática

- Os métodos padrão para testar se essas substâncias estão presentes em um rio é coletar amostras a seis décimos de profundidade.

# Prática

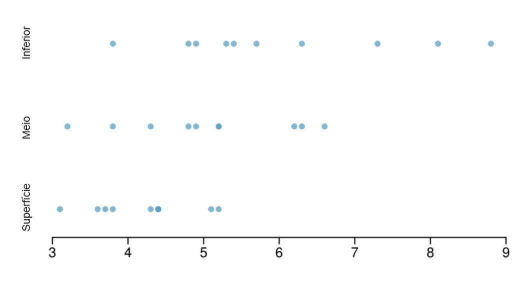
- Os métodos padrão para testar se essas substâncias estão presentes em um rio é coletar amostras a seis décimos de profundidade.
- Mas, uma vez que esses compostos são mais densos que a água e suas moléculas tendem a aderir a partículas de sedimento, eles são mais propensos a serem encontrados em concentrações mais altas perto do fundo do que perto da metade da profundidade.

## Dados

Concentração de aldrina (nanogramas por litro) em três níveis de profundidade.

| | aldrin | depth |
|---|---|---|
| 1 | 3.80 | inferior |
| 2 | 4.80 | inferior |
| ... | | |
| 10 | 8.80 | inferior |
| 11 | 3.20 | meio |
| 12 | 3.80 | meio |
| ... | | |
| 20 | 6.60 | meio |
| 21 | 3.10 | superfície |
| 22 | 3.60 | superfície |
| ... | | |
| 30 | 5.20 | superfície |

Concentração de aldrina (nanogramas por litro) em três níveis de profundidade.




| | n | média | sd |
|---|---|---|---|
| inferior | 10 | 6.04 | 1.58 |
| meio | 10 | 5.05 | 1.10 |
| superficie | 10 | 4.20 | 0.66 |
| no geral | 30 | 5.1 0 | 1.37 |

## Questão de pesquisa

Existe diferença entre as concentrações médias de aldrina entre os três níveis?

# Questão de pesquisa

Existe diferença entre as concentrações médias de aldrina entre os três níveis?

- Para comparar médias de dois grupos usamos uma estatística Z ou T.

# Questão de pesquisa

Existe diferença entre as concentrações médias de aldrina entre os três níveis?

- Para comparar médias de dois grupos usamos uma estatística Z ou T.
- Para comparar médias de 3+ grupos, usamos um novo teste chamado *ANOVA* e uma nova estatística chamada *F*.

# ANOVA

A ANOVA é usada para avaliar se a média da variável é diferente para diferentes níveis de uma variável categórica.

# ANOVA

A ANOVA é usada para avaliar se a média da variável é diferente para diferentes níveis de uma variável categórica.

$H_0$ : O resultado médio é o mesmo em todas as categorias,

$$\mu_1 = \mu_2 = \cdots = \mu_k,$$

onde $\mu_i$ representa a média do resultado para observações na categoria $i$.

$H_A$ : Pelo menos uma média é diferente das outras.

# Condições

1. As observações devem ser independentes dentro e entre os grupos
    - Se os dados são provenientes de amostra aleatória simples de menos de 10% da população, esta condição é satisfeita.
    - Verifique com cuidado se os dados podem ser independentes (por exemplo, sem pareamento).
    - Sempre importante, mas às vezes difícil de verificar.

# Condições

1. As observações devem ser independentes dentro e entre os grupos
    - Se os dados são provenientes de amostra aleatória simples de menos de 10% da população, esta condição é satisfeita.
    - Verifique com cuidado se os dados podem ser independentes (por exemplo, sem pareamento).
    - Sempre importante, mas às vezes difícil de verificar.
2. As observações dentro de cada grupo devem ser quase normais.
    - Especialmente importante quando os tamanhos das amostras são pequenos.

Como podemos verificar a normalidade?

## Condições

1. As observações devem ser independentes dentro e entre os grupos
   - Se os dados são provenientes de amostra aleatória simples de menos de 10% da população, esta condição é satisfeita.
   - Verifique com cuidado se os dados podem ser independentes (por exemplo, sem pareamento).
   - Sempre importante, mas às vezes difícil de verificar.
2. As observações dentro de cada grupo devem ser quase normais.
   - Especialmente importante quando os tamanhos das amostras são pequenos.

Como podemos verificar a normalidade?

3. A variabilidade entre os grupos deve ser aproximadamente igual.
   - Especialmente importante quando os tamanhos das amostras diferem entre os grupos.

Como podemos verificar essa condição?

# $z/t$ teste vs. ANOVA - Objetivo

*$z/t$ teste*

Comparar mpedias de *dois* grupos para ver se elas estão tão distantes que a diferença observada não pode ser atribuída à variabilidade amostral.

$$H_0 : \mu_1 = \mu_2$$

*ANOVA*

Comparar as médias de *dois ou mais* grupos para ver se elas estão tão distantes que as diferenças observadas não podem ser todas atribuídas à variabilidade amostral.

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

# $z/t$ teste vs. ANOVA - Método

*$z/t$ teste*

Calcule uma estatística de teste (uma proporção).

$$z/t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE(\bar{x}_1 - \bar{x}_2)}$$

*ANOVA*

Calcule uma estatística de teste (uma proporção).

$$F = \frac{\text{variabilidade entre grupos}}{\text{variabilidade dentro de grupos}}$$

# $z/t$ teste vs. ANOVA - Método

*$z/t$ teste*

Calcule uma estatística de teste (uma proporção).

$$z/t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SE(\bar{x}_1 - \bar{x}_2)}$$

*ANOVA*

Calcule uma estatística de teste (uma proporção).

$$F = \frac{\text{variabilidade entre grupos}}{\text{variabilidade dentro de grupos}}$$

- Valores grandes das estatísticas de teste levam a valores de p pequenos.
- Se o valor p for pequeno o suficiente, $H_0$ é rejeitada, concluímos que as médias populacionais não são iguais.

# $z/t$ teste vs. ANOVA

- Com apenas dois grupos, o teste t e a ANOVA são equivalentes, mas apenas se usarmos uma variância padrão agrupada no denominador da estatística de teste.

# $z/t$ teste vs. ANOVA

- Com apenas dois grupos, o teste t e a ANOVA são equivalentes, mas apenas se usarmos uma variância padrão agrupada no denominador da estatística de teste.
- Com mais de dois grupos, a ANOVA compara as médias amostrais com uma *grande média geral*.

## Hipóteses

Quais são as hipóteses corretas para testar a diferença entre as concentrações médias de aldrina entre os três níveis?

(a) $H_0 : \mu_B = \mu_M = \mu_S$

$H_A : \mu_B \neq \mu_M \neq \mu_S$

(b) $H_0 : \mu_B \neq \mu_M \neq \mu_S$

$H_A : \mu_B = \mu_M = \mu_S$

(c) $H_0 : \mu_B = \mu_M = \mu_S$

$H_A$ : Pelo menos uma média é diferente.

(d) $H_0 : \mu_B = \mu_M = \mu_S = 0$

$H_A$ : Pelo menos uma média é diferente.

(e) $H_0 : \mu_B = \mu_M = \mu_S$

$H_A : \mu_B > \mu_M > \mu_S$

## Hipóteses

Quais são as hipóteses corretas para testar a diferença entre as concentrações médias de aldrina entre os três níveis?

(a) $H_0 : \mu_B = \mu_M = \mu_S$

$H_A : \mu_B \neq \mu_M \neq \mu_S$

(b) $H_0 : \mu_B \neq \mu_M \neq \mu_S$

$H_A : \mu_B = \mu_M = \mu_S$

(c) $H_0 : \mu_B = \mu_M = \mu_S$

$H_A$ : *Pelo menos uma média é diferente.*

(d) $H_0 : \mu_B = \mu_M = \mu_S = 0$

$H_A$ : Pelo menos uma média é diferente.

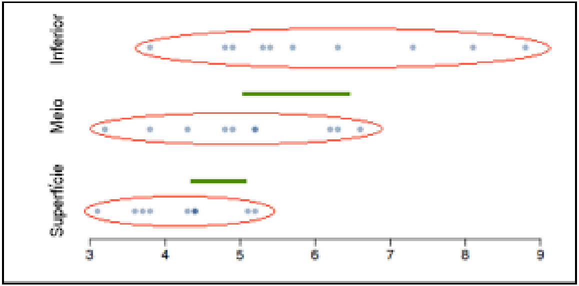
(e) $H_0 : \mu_B = \mu_M = \mu_S$

$H_A : \mu_B > \mu_M > \mu_S$

# Estatística de teste

Parece haver muita variabilidade dentro dos grupos? E que tal entre grupos?

$$F = \frac{\text{variabilidade entre grupos}}{\text{variabilidade dentro dos grupos}}$$



# Estatística de teste (cont.)

$$F = \frac{\text{variabilidade entre grupos}}{\text{variabilidade dentro dos grupos}} = \frac{MSG}{MSE}$$

- *MSG* é o quadrado médio entre os grupos

$$df_G = k - 1$$

  onde $k$ é o número de grupos

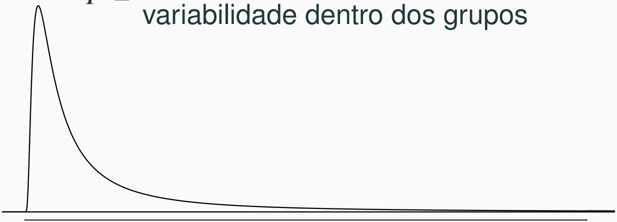
- *MSE* é erro quadrático médio - variabilidade dos resíduos

$$df_E = n - k$$

  onde $n$ é o número de observações.

## $F$ distribuição e valor p

$$F = \frac{\text{variabilidade entre grupos}}{\text{variabilidade dentro dos grupos}}$$



- Para podermos rejeitar $H_0$, precisamos de um pequeno valor-p, que requer uma estatística F grande.
- Para obter uma estatística F grande, a variabilidade entre as médias da amostra precisa ser maior que a variabilidade dentro das médias amostrais.

# Graus de liberdade associados à ANOVA

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (Grupo) | profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| (Erro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | Total | 29 | 54.29 | | | |

## Graus de liberdade associados à ANOVA

- grupos: $df_G = k - 1$, onde $k$ é o número de grupos
- total: $df_T = n - 1$, onde $n$ é o tamanho total da amostra
- erro: $df_E = df_T - df_G$

# Graus de liberdade associados à ANOVA

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (Grupo) | profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| (Erro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | Total | 29 | 54.29 | | | |

## Graus de liberdade associados à ANOVA

- grupos: $df_G = k - 1$, onde $k$ é o número de grupos
- total: $df_T = n - 1$, onde $n$ é o tamanho total da amostra
- erro: $df_E = df_T - df_G$

- $df_G = k - 1 = 3 - 1 = 2$

# Graus de liberdade associados à ANOVA

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (*G*rupo) | profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| (*E*rro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | *T*otal | 29 | 54.29 | | | |

## Graus de liberdade associados à ANOVA

- grupos: $df_G = k - 1$, onde $k$ é o número de grupos
- total: $df_T = n - 1$, onde $n$ é o tamanho total da amostra
- erro: $df_E = df_T - df_G$

- $df_G = k - 1 = 3 - 1 = 2$
- $df_T = n - 1 = 30 - 1 = 29$

# Graus de liberdade associados à ANOVA

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (*G*rupo) | profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| (*E*rro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | *T*otal | 29 | 54.29 | | | |

## Graus de liberdade associados à ANOVA

- grupos: $df_G = k - 1$, onde $k$ é o número de grupos
- total: $df_T = n - 1$, onde $n$ é o tamanho total da amostra
- erro: $df_E = df_T - df_G$

- $df_G = k - 1 = 3 - 1 = 2$
- $df_T = n - 1 = 30 - 1 = 29$
- $df_E = 29 - 2 = 27$

# Soma dos quadrados entre os grupos, SSG

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (Grupo) | profundidade | 2 | *16.96* | 8.48 | 6.13 | 0.0063 |
| (Erro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | Total | 29 | 54.29 | | | |

## Soma dos quadrados entre os grupos, SSG

Mede a variabilidade entre grupos

$$SSG = \sum_{i=1}^{k} n_i(\bar{x}_i - \bar{x})^2$$

onde $n_i$ é o tamanho de cada grupo, $\bar{x}_i$ é a média para cada grupo, $\bar{x}$ é a média geral (grande).

# Soma dos quadrados entre os grupos, SSG

| | n | média |
|---|---|---|
| inferior | 10 | 6.04 |
| meio | 10 | 5.05 |
| superfície | 10 | 4.2 |
| no geral | 30 | 5.1 |

# Soma dos quadrados entre os grupos, SSG

| | n | média |
|---|---|---|
| inferior | 10 | 6.04 |
| meio | 10 | 5.05 |
| superfície | 10 | 4.2 |
| no geral | 30 | 5.1 |

$$SSG = \left(10 \times (6.04 - 5.1)^2\right)$$

# Soma dos quadrados entre os grupos, SSG

| | n | média |
|---|---|---|
| inferior | 10 | 6.04 |
| meio | 10 | 5.05 |
| superfície | 10 | 4.2 |
| no geral | 30 | 5.1 |

$$
\begin{aligned}
SSG &= \left(10 \times (6.04 - 5.1)^2\right) \\
&+ \left(10 \times (5.05 - 5.1)^2\right)
\end{aligned}
$$

# Soma dos quadrados entre os grupos, SSG

| | n | média |
|---|---|---|
| inferior | 10 | 6.04 |
| meio | 10 | 5.05 |
| superfície | 10 | 4.2 |
| no geral | 30 | 5.1 |

$$\begin{aligned} SSG &= \left(10 \times (6.04 - 5.1)^2\right) \\ &+ \left(10 \times (5.05 - 5.1)^2\right) \\ &+ \left(10 \times (4.2 - 5.1)^2\right) \end{aligned}$$

# Soma dos quadrados entre os grupos, SSG

| | n | média |
|---|---|---|
| inferior | 10 | 6.04 |
| meio | 10 | 5.05 |
| superfície | 10 | 4.2 |
| no geral | 30 | 5.1 |

$$
\begin{aligned}
SSG &= \left(10 \times (6.04 - 5.1)^2\right) \\
&+ \left(10 \times (5.05 - 5.1)^2\right) \\
&+ \left(10 \times (4.2 - 5.1)^2\right) \\
&= 16.96
\end{aligned}
$$

# Soma do total de quadrados SST

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (Grupos) | Profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| (Erro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | Total | 29 | 54.29 | | | |

## Soma do total de quadrados SST

Mede a variabilidade entre grupos

$$SST = \sum_{i=1}^{n} (x_i - \bar{x})$$

onde $x_i$ representam cada observação no conjunto de dados.

# Soma do total de quadrados SST

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (Grupos) | Profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| (Erro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | Total | 29 | 54.29 | | | |

## Soma do total de quadrados SST

Mede a variabilidade entre grupos

$$SST = \sum_{i=1}^{n}(x_i - \bar{x})$$

onde $x_i$ representam cada observação no conjunto de dados.

$$SST = (3.8 - 5.1)^2 + (4.8 - 5.1)^2 + (4.9 - 5.1)^2 + \cdots + (5.2 - 5.1)^2$$

# Soma do total de quadrados SST

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (*G*rupos) | Profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| (*E*rro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | *T*otal | 29 | *54.29* | | | |

## Soma do total de quadrados SST

Mede a variabilidade entre grupos

$$SST = \sum_{i=1}^{n}(x_i - \bar{x})$$

onde $x_i$ representam cada observação no conjunto de dados.

$$\begin{aligned} SST &= (3.8 - 5.1)^2 + (4.8 - 5.1)^2 + (4.9 - 5.1)^2 + \cdots + (5.2 - 5.1)^2 \\ &= (-1.3)^2 + (-0.3)^2 + (-0.2)^2 + \cdots + (0.1)^2 \end{aligned}$$

# Soma do total de quadrados SST

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (Grupos) | Profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| (Erro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | Total | 29 | 54.29 | | | |

## Soma do total de quadrados SST

Mede a variabilidade entre grupos

$$SST = \sum_{i=1}^{n} (x_i - \bar{x})$$

onde $x_i$ representam cada observação no conjunto de dados.

$$\begin{aligned} SST &= (3.8 - 5.1)^2 + (4.8 - 5.1)^2 + (4.9 - 5.1)^2 + \cdots + (5.2 - 5.1)^2 \\ &= (-1.3)^2 + (-0.3)^2 + (-0.2)^2 + \cdots + (0.1)^2 \\ &= 1.69 + 0.09 + 0.04 + \cdots + 0.01 \end{aligned}$$

# Soma do total de quadrados SST

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (*G*rupos) | Profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| (*E*rro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | *T*otal | 29 | *54.29* | | | |

## Soma do total de quadrados SST

Mede a variabilidade entre grupos

$$SST = \sum_{i=1}^{n}(x_i - \bar{x})$$

onde $x_i$ representam cada observação no conjunto de dados.

$$\begin{aligned} SST &= (3.8 - 5.1)^2 + (4.8 - 5.1)^2 + (4.9 - 5.1)^2 + \cdots + (5.2 - 5.1)^2 \\ &= (-1.3)^2 + (-0.3)^2 + (-0.2)^2 + \cdots + (0.1)^2 \\ &= 1.69 + 0.09 + 0.04 + \cdots + 0.01 \\ &= 54.29 \end{aligned}$$

# Soma dos quadrados erro, SSE

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (*G*rupo) | Profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| (*E*rro) | Resíduos | 27 | *37.33* | 1.38 | | |
| | *T*otal | 29 | 54.29 | | | |

## Soma dos quadrados erro, SSE

Mede a variabilidade dentro de grupos:

$$SSE = SST - SSG$$

# Soma dos quadrados erro, SSE

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (*G*rupo) | Profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| (*E*rro) | Resíduos | 27 | *37.33* | 1.38 | | |
| | *T*otal | 29 | 54.29 | | | |

## Soma dos quadrados erro, SSE

Mede a variabilidade dentro de grupos:

$$SSE = SST - SSG$$

$$SSE = 54.29 - 16.96 = 37.33$$

# Erro quadrático médio

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (*G*rupo) | profundidade | 2 | 16.96 | *8.48* | 6.13 | 0.0063 |
| (*E*rro) | Resíduos | 27 | 37.33 | *1.38* | | |
| | *T*otal | 29 | 54.29 | | | |

## Erro quadrático médio

O erro quadrático médio é calculado como a soma dos quadrados divididos pelos graus de liberdade.

# Erro quadrático médio

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (*G*rupo) | profundidade | 2 | 16.96 | *8.48* | 6.13 | 0.0063 |
| (*E*rro) | Resíduos | 27 | 37.33 | *1.38* | | |
| | *T*otal | 29 | 54.29 | | | |

## Erro quadrático médio

O erro quadrático médio é calculado como a soma dos quadrados divididos pelos graus de liberdade.

$$MSG \quad = \quad 16.96/2 = 8.48$$

# Erro quadrático médio

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (*G*rupo) | profundidade | 2 | 16.96 | *8.48* | 6.13 | 0.0063 |
| (*E*rro) | Resíduos | 27 | 37.33 | *1.38* | | |
| | *T*otal | 29 | 54.29 | | | |

## Erro quadrático médio

O erro quadrático médio é calculado como a soma dos quadrados divididos pelos graus de liberdade.

$$
\begin{aligned}
MSG &= 16.96/2 = 8.48 \\
MSE &= 37.33/27 = 1.38
\end{aligned}
$$

# Estatística de teste, valor F

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (*G*rupo) | Profundidade | 2 | 16.96 | 8.48 | *6.14* | 0.0063 |
| (*E*rro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | *T*otal | 29 | 54.29 | | | |

## Estatística de teste, valor F

Como discutimos anteriormente, a estatística F é a razão entre o grupo e a variabilidade dentro do grupo.

$$F = \frac{MSG}{MSE}$$

# Estatística de teste, valor F

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (*G*rupo) | Profundidade | 2 | 16.96 | 8.48 | *6.14* | 0.0063 |
| (*E*rro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | *T*otal | 29 | 54.29 | | | |

## Estatística de teste, valor F

Como discutimos anteriormente, a estatística F é a razão entre o grupo e a variabilidade dentro do grupo.

$$F = \frac{MSG}{MSE}$$

$$F = \frac{8.48}{1.38} = 6.14$$

# valor-p

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (Grupo) | profundidade | 2 | 16.96 | 8.48 | 6.14 | 0.0063 |
| (Erro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | Total | 29 | 54.29 | | | |

## valor-p

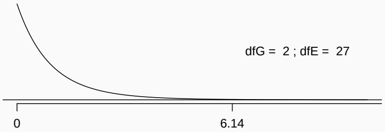
O valor-p é a probabilidade de pelo menos uma relação tão grande entre a variabilidade "entre grupo"e "dentro do grupo", se de fato os meios de todos os grupos são iguais. É calculado como a área sob a curva F, com graus de liberdade $df_G$ e $df_E$, acima da estatística F observada.

# valor-p

| | | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|---|
| (Grupo) | profundidade | 2 | 16.96 | 8.48 | 6.14 | 0.0063 |
| (Erro) | Resíduos | 27 | 37.33 | 1.38 | | |
| | Total | 29 | 54.29 | | | |

## valor-p

O valor-p é a probabilidade de pelo menos uma relação tão grande entre a variabilidade "entre grupo"e "dentro do grupo", se de fato os meios de todos os grupos são iguais. É calculado como a área sob a curva F, com graus de liberdade $df_G$ e $df_E$, acima da estatística F observada.



# Conclusão - no contexto

Qual é a conclusão do teste de hipóteses?

Os dados fornecem evidências de que a concentração média de aldrina

(a) é diferente para todos os grupos.

(b) na superfície é menor que os outros níveis.

(c) é diferente para pelo menos um grupo.

(d) é o mesmo para todos os grupos.

# Conclusão - no contexto

## Qual é a conclusão do teste de hipóteses?

Os dados fornecem evidências de que a concentração média de aldrina

(a) é diferente para todos os grupos.

(b) na superfície é menor que os outros níveis.

(c) *é diferente para pelo menos um grupo.*

(d) é o mesmo para todos os grupos.

## Conclusão

- Se valor-p for pequeno (menor que $\alpha$), rejeite $H_0$. Os dados fornecem evidências de que pelo menos uma média é diferente (mas não podemos dizer qual).

# Conclusão

- Se valor-p for pequeno (menor que $\alpha$), rejeite $H_0$. Os dados fornecem evidências de que pelo menos uma média é diferente (mas não podemos dizer qual).
- Se o valor p for grande, não rejeite $H_0$. Os dados não fornecem evidências de que pelo menos um par de médias são diferentes umas das outras, as diferenças observadas nas médias amostrais são atribuíveis à variabilidade amostral (ou acaso).

# (1) Independência

Esta condição parece estar satisfeita?

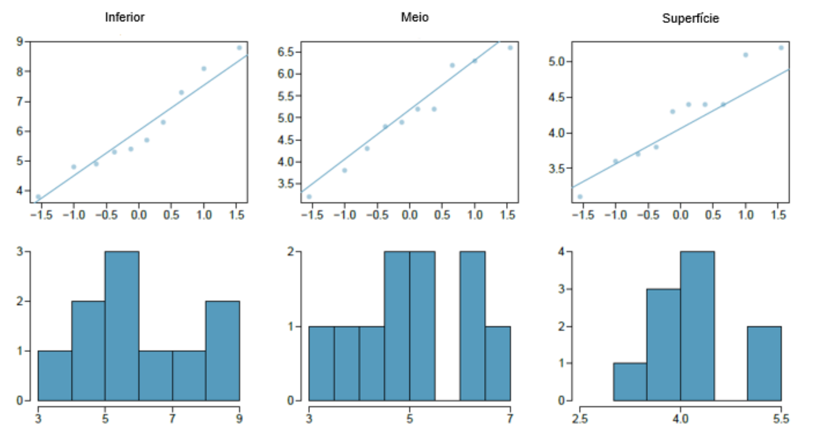
## (1) Independência

Esta condição parece estar satisfeita?

*Neste estudo, não temos razão para acreditar que a concentração de Aldrin não seja independente uma da outra...*

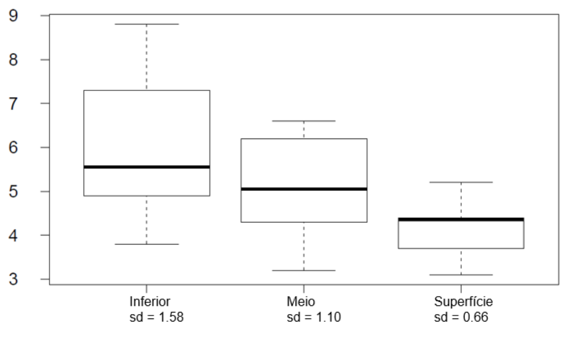
# (2) Aproximadamente normal

Esta condição parece estar satisfeita?



# (3) Variância constante

Esta condição parece estar satisfeita?



## O que significa diferir?

- Anteriormente, concluímos que pelo menos um par de médias diferem. A questão natural que se segue é "quais?"

# O que significa diferir?

- Anteriormente, concluímos que pelo menos um par de médias diferem. A questão natural que se segue é "quais?"
- Podemos fazer dois testes $t$ de amostra para diferenças em cada par de grupos possível.

## O que significa diferir?

- Anteriormente, concluímos que pelo menos um par de médias diferem. A questão natural que se segue é "quais?"
- Podemos fazer dois testes $t$ de amostra para diferenças em cada par de grupos possível.

Você consegue ver alguma armadilha com essa abordagem?

# O que significa diferir?

- Anteriormente, concluímos que pelo menos um par de médias diferem. A questão natural que se segue é "quais?"
- Podemos fazer dois testes $t$ de amostra para diferenças em cada par de grupos possível.

Você consegue ver alguma armadilha com essa abordagem?

- Quando executamos muitos testes, a taxa de erro do tipo 1 aumenta.
- Esse problema é resolvido usando um nível de significância modificado.

# Comparações Múltiplas

- O cenário de testar muitos pares de grupos é chamado *comparações múltiplas*.

## Comparações Múltiplas

- O cenário de testar muitos pares de grupos é chamado *comparações múltiplas*.
- A *correção de Bonferroni* sugere que um nível de significância mais *rigoroso* é mais apropriado para estes testes:

$$\alpha^{\star} = \alpha/K$$

onde $K$ é o número de comparações consideradas.

## Comparações Múltiplas

- O cenário de testar muitos pares de grupos é chamado *comparações múltiplas*.
- A *correção de Bonferroni* sugere que um nível de significância mais *rigoroso* é mais apropriado para estes testes:

$$\alpha^\star = \alpha/K$$

  onde $K$ é o número de comparações consideradas.
- Se existem grupos $k$, então todos os pares possíveis são comparados e $K = \frac{k(k-1)}{2}$.

## Determinando o modificado $\alpha$

No aldrin, a profundidade do conjunto de dados tem 3 níveis: inferior, média e superfície. Se $\alpha = 0.05$, qual deve ser o nível de significância modificado para dois testes $t$ de amostra para determinar quais pares de grupos possuem médias significativamente diferentes?

(a) $\alpha^* = 0.05$

(b) $\alpha^* = 0.05/2 = 0.025$

(c) $\alpha^* = 0.05/3 = 0.0167$

(d) $\alpha^* = 0.05/6 = 0.0083$

# Determinando o modificado $\alpha$

No aldrin, a profundidade do conjunto de dados tem 3 níveis: inferior, média e superfície. Se $\alpha = 0.05$, qual deve ser o nível de significância modificado para dois testes $t$ de amostra para determinar quais pares de grupos possuem médias significativamente diferentes?

(a) $\alpha^* = 0.05$

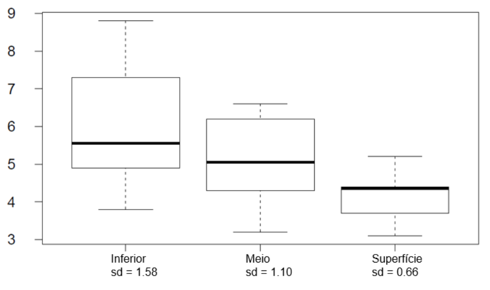
(b) $\alpha^* = 0.05/2 = 0.025$

(c) $\alpha^* = 0.05/3 = 0.0167$

(d) $\alpha^* = 0.05/6 = 0.0083$

# O que significa diferir?

Com base nos gráficos abaixo, o que significa que você esperaria ser significativamente diferente?




(a) inferior & superfície

(b) inferior & meio

(c) meio & superfície

(d) inferior & meio; meio & superfície

(e) inferior & meio; inferior & superfície; meio & superfície

## O que significa diferir? (cont.)

Se a suposição de variabilidade igual entre os grupos for satisfeita para a ANOVA, podemos usar os dados de todos os grupos para estimar a variabilidade:

- Estime qualquer desvio padrão dentro do grupo com $\sqrt{MSE}$, que é $s_{agrupado}$
- Use os graus de liberdade do erro, $n - k$, para a distribuição $t$.

Diferença entre duas médias: após ANOVA

$$SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \approx \sqrt{\frac{MSE}{n_1} + \frac{MSE}{n_2}}$$

# Prática

Existe uma diferença entre a concentração média de aldrina na parte inferior e a meia profundidade?

| | n | média | sd |
|---|---|---|---|
| inferior | 10 | *6.04* | 1.58 |
| meio | 10 | *5.05* | 1.10 |
| superfície | 10 | 4.2 | 0.66 |
| no geral | 30 | 5.1 | 1.37 |

| | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|
| profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| Resíduos | *27* | 37.33 | *1.38* | | |
| Total | 29 | 54.29 | | | |

$$T_{df_E} = \frac{(\bar{x}_{inferior} - \bar{x}_{meio})}{\sqrt{\frac{MSE}{n_{inferior}} + \frac{MSE}{n_{meio}}}}$$

# Prática

Existe uma diferença entre a concentração média de aldrina na parte inferior e a meia profundidade?

| | n | média | sd |
|---|---|---|---|
| inferior | 10 | *6.04* | 1.58 |
| meio | 10 | *5.05* | 1.10 |
| superfície | 10 | 4.2 | 0.66 |
| no geral | 30 | 5.1 | 1.37 |

| | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|
| profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| Resíduos | *27* | 37.33 | *1.38* | | |
| Total | 29 | 54.29 | | | |

$$T_{df_E} = \frac{(\bar{x}_{inferior} - \bar{x}_{meio})}{\sqrt{\frac{MSE}{n_{inferior}} + \frac{MSE}{n_{meio}}}}$$

$$T_{27} = \frac{(6.04 - 5.05)}{\sqrt{\frac{1.38}{10} + \frac{1.38}{10}}} = \frac{0.99}{0.53} = 1.87$$

# Prática

Existe uma diferença entre a concentração média de aldrina na parte inferior e a meia profundidade?

| | n | média | sd |
|---|---|---|---|
| inferior | 10 | *6.04* | 1.58 |
| meio | 10 | *5.05* | 1.10 |
| superfície | 10 | 4.2 | 0.66 |
| no geral | 30 | 5.1 | 1.37 |

| | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|
| profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| Resíduos | *27* | 37.33 | *1.38* | | |
| Total | 29 | 54.29 | | | |

$$T_{df_E} = \frac{(\bar{x}_{inferior} - \bar{x}_{meio})}{\sqrt{\frac{MSE}{n_{inferior}} + \frac{MSE}{n_{meio}}}}$$

$$T_{27} = \frac{(6.04 - 5.05)}{\sqrt{\frac{1.38}{10} + \frac{1.38}{10}}} = \frac{0.99}{0.53} = 1.87$$

$$0.05 < valor - p < 0.10$$ (frente e verso)

# Prática

Existe uma diferença entre a concentração média de aldrina na parte inferior e a meia profundidade?

| | n | média | sd |
|---|---|---|---|
| inferior | 10 | *6.04* | 1.58 |
| meio | 10 | *5.05* | 1.10 |
| superfície | 10 | 4.2 | 0.66 |
| no geral | 30 | 5.1 | 1.37 |

| | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|
| profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| Resíduos | *27* | 37.33 | *1.38* | | |
| Total | 29 | 54.29 | | | |

$$T_{df_E} = \frac{(\bar{x}_{inferior} - \bar{x}_{meio})}{\sqrt{\frac{MSE}{n_{inferior}} + \frac{MSE}{n_{meio}}}}$$

$$T_{27} = \frac{(6.04 - 5.05)}{\sqrt{\frac{1.38}{10} + \frac{1.38}{10}}} = \frac{0.99}{0.53} = 1.87$$

$$0.05 < valor - p < 0.10$$ (frente e verso)

$$\alpha^{\star} = 0.05/3 = 0.0167$$

# Prática

Existe uma diferença entre a concentração média de aldrina na parte inferior e a meia profundidade?

| | n | média | sd |
|---|---|---|---|
| inferior | 10 | *6.04* | 1.58 |
| meio | 10 | *5.05* | 1.10 |
| superfície | 10 | 4.2 | 0.66 |
| no geral | 30 | 5.1 | 1.37 |

| | Df | Soma Sq | Média Sq | F valor | Pr(>F) |
|---|---|---|---|---|---|
| profundidade | 2 | 16.96 | 8.48 | 6.13 | 0.0063 |
| Resíduos | *27* | 37.33 | *1.38* | | |
| Total | 29 | 54.29 | | | |

$$T_{df_E} = \frac{(\bar{x}_{inferior} - \bar{x}_{meio})}{\sqrt{\frac{MSE}{n_{inferior}} + \frac{MSE}{n_{meio}}}}$$

$$T_{27} = \frac{(6.04 - 5.05)}{\sqrt{\frac{1.38}{10} + \frac{1.38}{10}}} = \frac{0.99}{0.53} = 1.87$$

$$0.05 < valor - p < 0.10$$ (frente e verso)

$$\alpha^{\star} = 0.05/3 = 0.0167$$

Não rejeite $H_0$, os dados não fornecem evidências convincentes de uma diferença entre as concentrações médias de aldrina na profundidade inferior e média.

# Prática

## Comparações entre pares

Existe uma diferença entre a concentração média de aldrina no fundo e na superfície?

# Prática

## Comparações entre pares

Existe uma diferença entre a concentração média de aldrina no fundo e na superfície?

$$T_{df_E} = \frac{(\bar{x}_{inferior} - \bar{x}_{superf cie})}{\sqrt{\frac{MSE}{n_{inferior}} + \frac{MSE}{n_{superficie}}}}$$

## Comparações entre pares

Existe uma diferença entre a concentração média de aldrina no fundo e na superfície?

$$
\begin{aligned}
T_{df_E} &= \frac{(\bar{x}_{inferior} - \bar{x}_{superf\,cie})}{\sqrt{\frac{MSE}{n_{inferior}} + \frac{MSE}{n_{superficie}}}} \\
T_{27} &= \frac{(6.04 - 4.02)}{\sqrt{\frac{1.38}{10} + \frac{1.38}{10}}} = \frac{2.02}{0.53} = 3.81
\end{aligned}
$$

# Prática

## Comparações entre pares

Existe uma diferença entre a concentração média de aldrina no fundo e na superfície?

$$
\begin{aligned}
T_{df_E} &= \frac{(\bar{x}_{inferior} - \bar{x}_{superf\,cie})}{\sqrt{\frac{MSE}{n_{inferior}} + \frac{MSE}{n_{superficie}}}} \\
T_{27} &= \frac{(6.04 - 4.02)}{\sqrt{\frac{1.38}{10} + \frac{1.38}{10}}} = \frac{2.02}{0.53} = 3.81 \\
valor - p &< 0.01 \qquad \textit{(frente e verso)}
\end{aligned}
$$

# Prática

## Comparações entre pares

Existe uma diferença entre a concentração média de aldrina no fundo e na superfície?

$$
\begin{aligned}
T_{df_E} &= \frac{(\bar{x}_{inferior} - \bar{x}_{superf cie})}{\sqrt{\frac{MSE}{n_{inferior}} + \frac{MSE}{n_{superficie}}}} \\
T_{27} &= \frac{(6.04 - 4.02)}{\sqrt{\frac{1.38}{10} + \frac{1.38}{10}}} = \frac{2.02}{0.53} = 3.81 \\
valor - p &< 0.01 \qquad \textit{(frente e verso)} \\
\alpha^{\star} &= 0.05/3 = 0.0167
\end{aligned}
$$

# Prática

## Comparações entre pares

Existe uma diferença entre a concentração média de aldrina no fundo e na superfície?

$$
\begin{aligned}
T_{df_E} &= \frac{(\bar{x}_{inferior} - \bar{x}_{superfcie})}{\sqrt{\frac{MSE}{n_{inferior}} + \frac{MSE}{n_{superficie}}}} \\
T_{27} &= \frac{(6.04 - 4.02)}{\sqrt{\frac{1.38}{10} + \frac{1.38}{10}}} = \frac{2.02}{0.53} = 3.81 \\
valor - p &< 0.01 \qquad \textit{(frente e verso)} \\
\alpha^{\star} &= 0.05/3 = 0.0167
\end{aligned}
$$

*Rejeite $H_0$, os dados fornecem evidências convincentes de uma diferença entre as concentrações médias de aldrina no fundo e na superfície.*