# POLI - ENGENHARIA DE PETRÓLEO
# ESTATÍSTICA
# INTERVALOS DE CONFIANÇA

Aulas 6 e 7

Prof. Regina Meyer Branski

# O que veremos...



- Intervalos de confiança para:
  - Média (amostras grandes)
  - Média (amostras pequenas)
  - Proporções populacionais
  - Variância e Desvio Padrão

# Construção de uma Estimativa de Intervalo



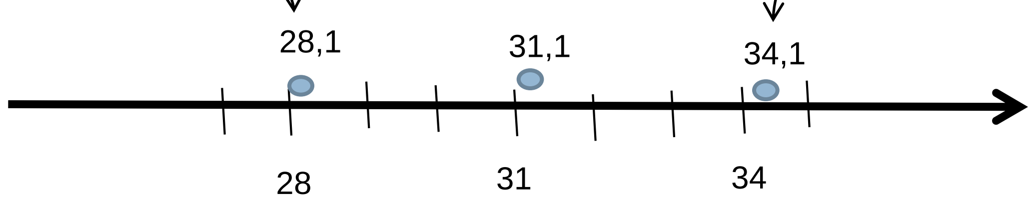
A partir da média da amostra da linha de carros de passeio, fabricante pode estimar o índice médio de economia de combustível como sendo de 31,1 km por galão **para toda** linha de carro de passeio.

Encontre a média da amostra aleatória
Média Amostra
$\bar{x} = 31,1$

Encontre a margem de erro
E=3,0

Encontre os pontos finais do intervalo
Esquerda: 31,1 – 3,0 = 28,1
Direita: 31,1 + 3,0 = 34,1

Encontre a estimativa de intervalo
$28,1 < \mu < 34,1$




Quero estar 90% confiante de que minha estimativa do índice médio de economia de combustível para toda a linha de carros de passeio está entre 28,1 e 34,1.

# Objetivos

- Encontrar uma estimativa pontual e uma margem de erro
- Construir e interpretar intervalos de confiança para a média populacional
- Determinar o tamanho mínimo da amostra necessária na estimativa de $\mu$

# INTERVALOS DE CONFIANÇA PARA A MÉDIA (AMOSTRAS GRANDES)

Prof. Regina Meyer Branski

# Estimativa pontual para população $\mu$

- Valor único estimado para um parâmetro populacional
- A estimativa pontual menos tendenciosa de uma média populacional $\mu$ é a média amostral $\overline{x}$

| Parâmetro de estimativa populacional | Com amostra estatística |
|---|---|
| Média: $\mu$ | $\overline{x}$ |

# Exemplo: estimativa pontual para população

Pesquisadores de mercado usam o número de frases por anúncio como medida de legibilidade de anúncios de revistas. A seguir, representamos uma amostra aleatória do número de frases encontrado em 50 anúncios. Encontre a estimativa pontual da média populacional $\mu$. (*Fonte: Journal of Advertising Research.*)

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | 20 | 18 | 16 | 9 | 9 | 11 | 13 | 22 | 16 | 5 | 18 | 6 | 6 | 5 |
| 12 | 25 | 17 | 23 | 7 | 10 | 9 | 10 | 10 | 5 | 11 | 18 | 18 | 9 | 9 |
| 17 | 13 | 11 | 7 | 14 | 6 | 11 | 12 | 11 | 6 | 12 | 14 | 11 | 9 | 18 |
| 12 | 12 | 17 | 11 | 20 | | | | | | | | | | |

# Estimativa pontual para população $\mu$

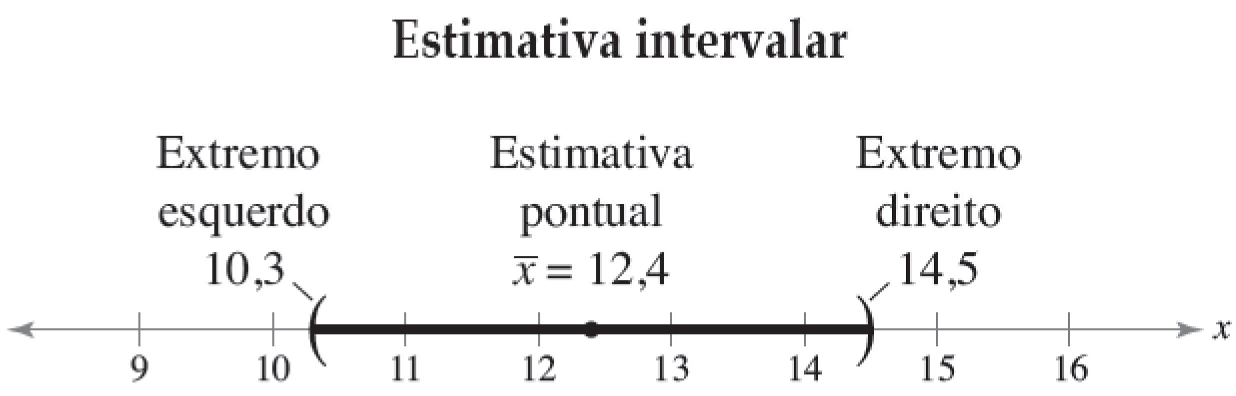
A média amostral dos dados é

$$\boxed{\bar{x} = \frac{\Sigma x}{n} = \frac{620}{50} = 12.4}$$

Então, a estimativa pontual para a média do comprimento de todos os anúncios de revista é 12,4 frases.

# Estimativa intervalar

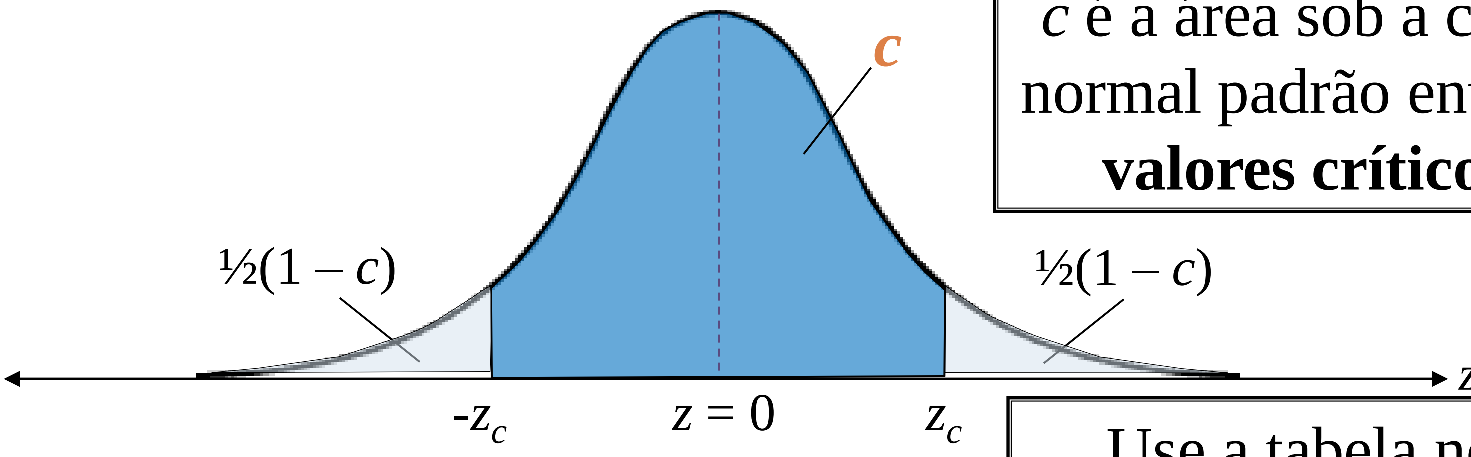
- Intervalo, ou amplitude de valores, usado para estimar um parâmetro populacional




Qual é o nível de confiança que queremos ter para a estimativa intervalar conter a média populacional $\mu$?

# Nível de confiança c

- A probabilidade de que o intervalo estimado contenha o parâmetro populacional

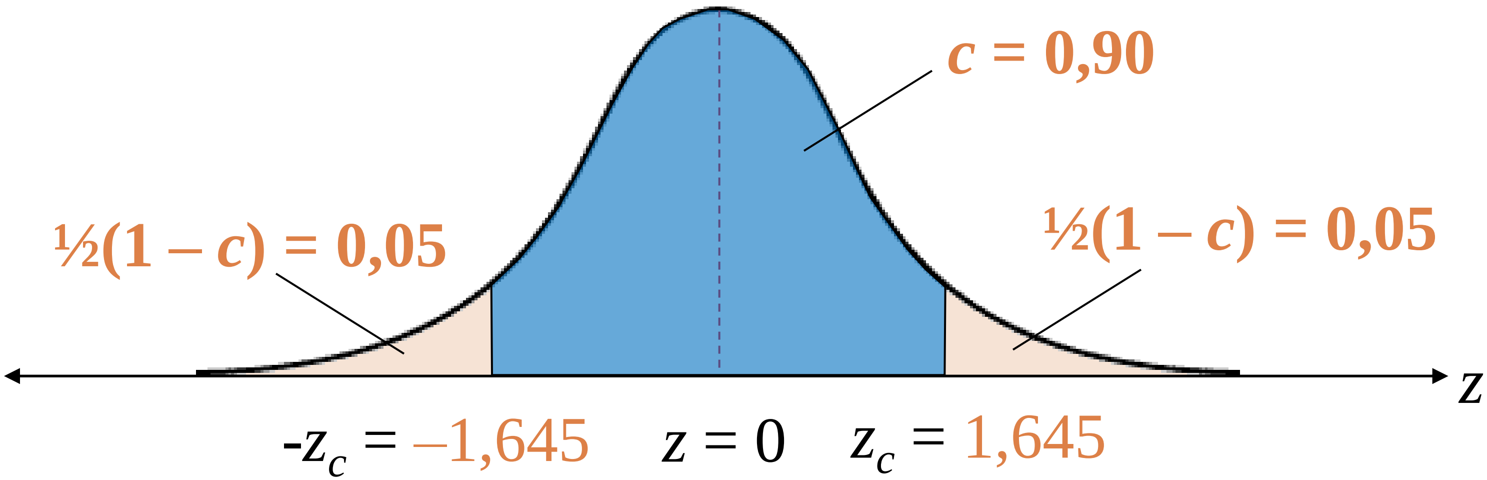



$c$ é a área sob a curva normal padrão entre os **valores críticos**.

Use a tabela normal padrão para encontrar os $z$ escores correspondentes.

A área restante nas caudas é $1 – c$ .

- Se o nível de confiança é 90%, isso significa que temos 90% de confiança que o intervalo contém a média populacional $\mu$




Os escores $z$ correspondentes são $\pm 1{,}645$.

# Erro de amostragem

- A diferença entre a estimativa pontual e o valor real do parâmetro populacional
- Para $\mu$:
  - Erro de amostragem é a diferença $\overline{X} - \mu$
  - μ geralmente é desconhecido
  - $\overline{X}$ varia de amostra para amostra

# Margem de erro

- Maior distância possível entre o ponto de estimativa e o valor do parâmetro que está estimando para um dado nível de confiança, c
- Denotado por *E*

$$\boxed{E = z_c \sigma_{\overline{x}} = z_c \frac{\sigma}{\sqrt{n}}} \longrightarrow \boxed{\text{Quando } n \geq 30\text{, o desvio padrão da amostra, } s\text{, pode ser usado para } \sigma.}$$

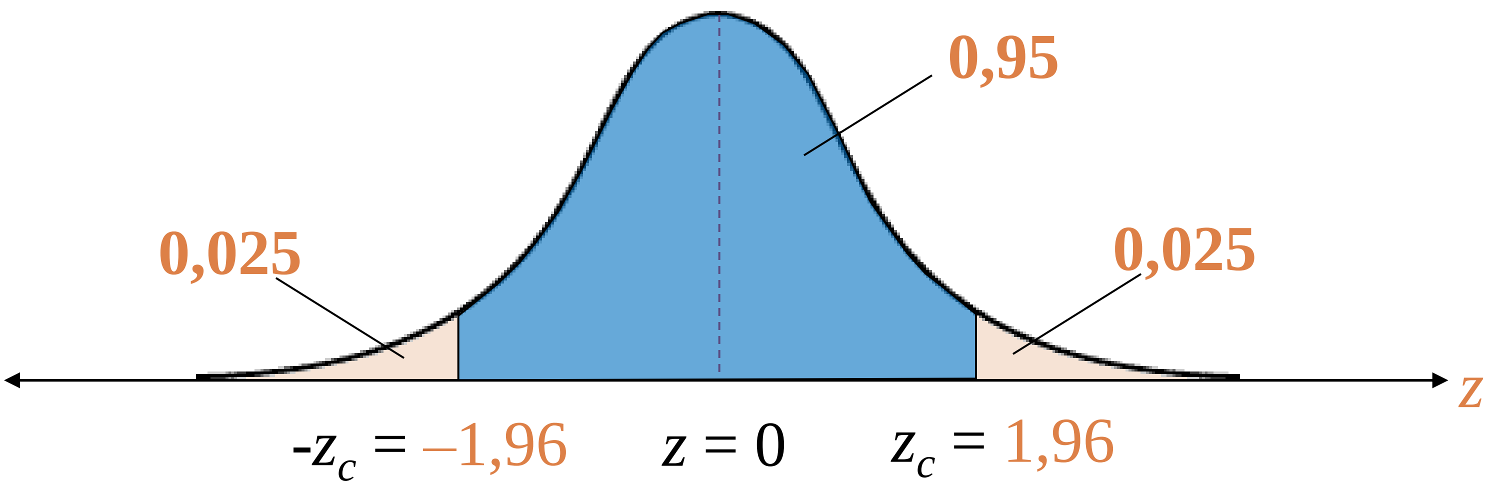
- Às vezes chamado de erro máximo ou tolerância de erro

# Exemplo: encontrando a margem de erro

Use os dados das propagandas das revistas e um nível de confiança de 95% para encontrar a margem de erro do número de frases em todos os anúncios de revistas. Assuma que o desvio padrão da amostra seja aproximadamente 5,0.

# Solução: encontrando a margem de erro

- Primeiro, encontre os valores críticos




Confiante de que 95% da área sob a curva normal padrão cai dentro de 1,96 desvio padrão da média. (Você pode aproximar a distribuição das médias amostrais com uma curva normal pelo Teorema do Limite Central, já que $n = 50 \geq 30$.)

$$E = z_c \frac{\sigma}{\sqrt{n}} \approx z_c \frac{s}{\sqrt{n}}$$

$$\approx 1.96 \cdot \frac{5.0}{\sqrt{50}}$$

$$\approx 1.4$$

Você não conhece σ, mas já que $n \geq 30$, você pode usar $s$ no lugar de σ.

Você tem 95% de confiança que a margem de erro para a média populacional é de aproximadamente 1,4 frases.

# Intervalos de confiança para a média populacional

Um **intervalo de confiança c** para a média populacional $\mu$ é:

$$\bar{x} - E < \mu < \bar{x} + E.$$

A probabilidade de que o intervalo de confiança contenha $\mu$ é **c**.

$$E = z_c \frac{\sigma}{\sqrt{n}}$$

# Intervalos de confiança para a média populacional

Encontrando um intervalo de confiança para a média populacional (n > ou igual a 30 ou ***σ*** é conhecido com uma população normalmente distribuída).

| *Em palavras* | *Em símbolos* |
| --- | --- |
| 1. Encontre a estatística amostral *n* e $\overline{X}$ | $\overline{x} = \frac{\Sigma x}{n}$ |
| 2. Especifique σ, se for conhecido. Caso contrário, encontre o desvio padrão amostral *s* e use-o como uma estimativa para *σ*. | $s = \sqrt{\frac{\Sigma(x-\overline{x})^2}{n-1}}$ |

| Em palavras | Em símbolos |
| --- | --- |
| 3. Encontre o valor crítico $z_c$ que corresponda ao nível de confiança dado. | Use a tabela normal padrão |
| 4. Encontre a margem de erro *E*. | $\boxed{E = z_c \frac{\sigma}{\sqrt{n}}}$ |
| 5. Encontre os extremos esquerdo e direito e forme o intervalo de confiança. | Extremo esquerdo: $\overline{x} - E$<br>Extremo direito: $\overline{x} + E$<br>Intervalo: $\overline{x} - E < \mu < \overline{x} + E$ |

# Exemplo: construindo um intervalo de confiança

Construa um intervalo de confiança de 95% para a média do número de frases em todos os anúncios de revista.

Lembre-se: $\bar{x} = 12.4$ e $E = 1,4$

# Exemplo: construindo um intervalo de confiança

Construa um intervalo de confiança de 95% para a média do número de frases em todos os anúncios de revista.

Lembre-se: $\bar{x} = 12.4$ e $E$ = 1,4

Extremo esquerdo:

$$\bar{x} - E$$
$$= 12.4 - 1.4$$
$$= 11.0$$

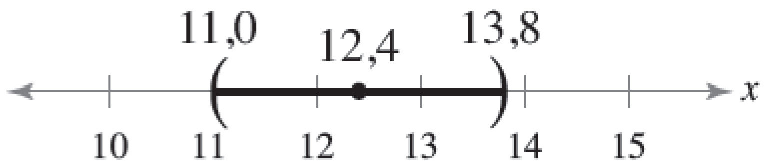
Extremo direito:

$$\bar{x} + E$$
$$= 12.4 + 1.4$$
$$= 13.8$$

**11,0 < μ < 13,8**

# Construindo um intervalo de confiança

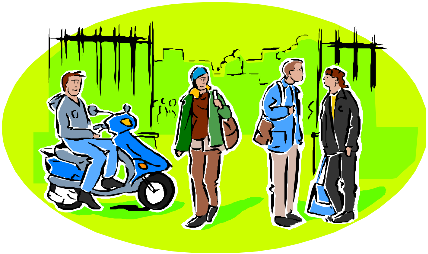
**11,0 < μ < 13,8**




Com 95% de confiança, você pode dizer que a média populacional do número de frases está entre 11,0 e 13,8.

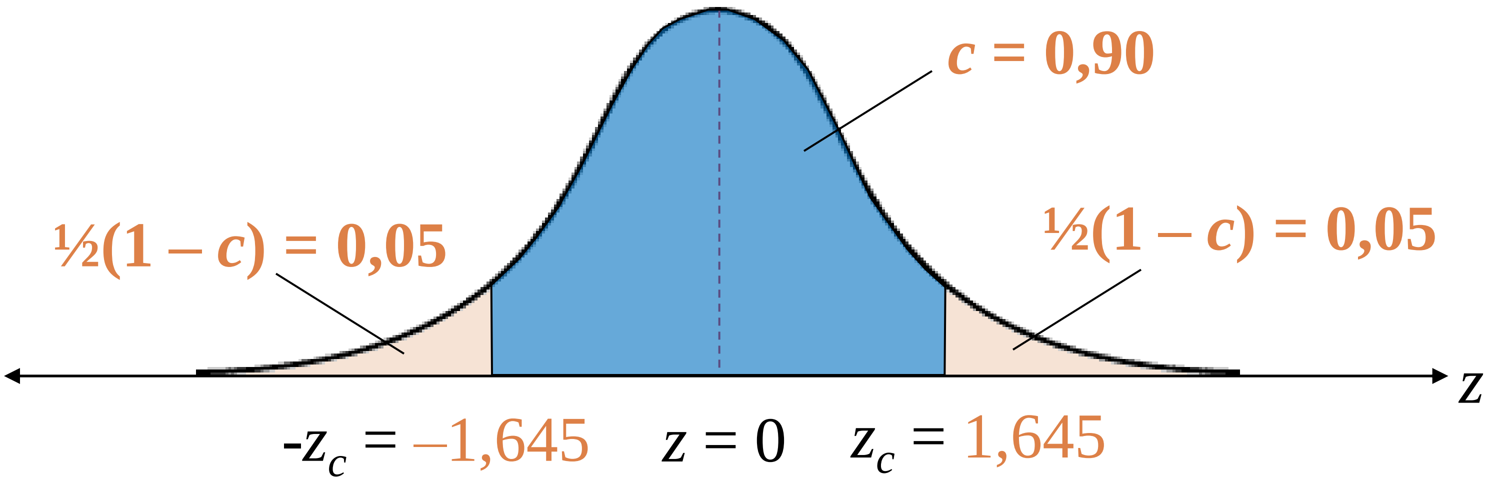
# Construindo um intervalo de confiança $\sigma$ conhecido

O diretor de admissão de uma faculdade deseja estimar a idade média de todos os estudantes matriculados. Em uma amostra aleatória de 20 estudantes, a idade média encontrada é de 22,9 anos. Baseado em estudos anteriores, o desvio padrão conhecido é 1,5 ano e a população é normalmente distribuída. Construa um intervalo de confiança de 90% para a média de idade da população.



# Construindo um intervalo de confiança σ conhecido

- Primeiro encontre os valores críticos




$$z_c = 1,645$$

# Construindo um intervalo de confiança σ conhecido

- Margem de erro:

$$E = z_c \frac{\sigma}{\sqrt{n}} = 1.645 \cdot \frac{1.5}{\sqrt{20}} \approx 0.6$$

- Intervalo de confiança:

Extremo esquerdo

$$\bar{x} - E$$
$$= 22.9 - 0.6$$
$$= 22.3$$

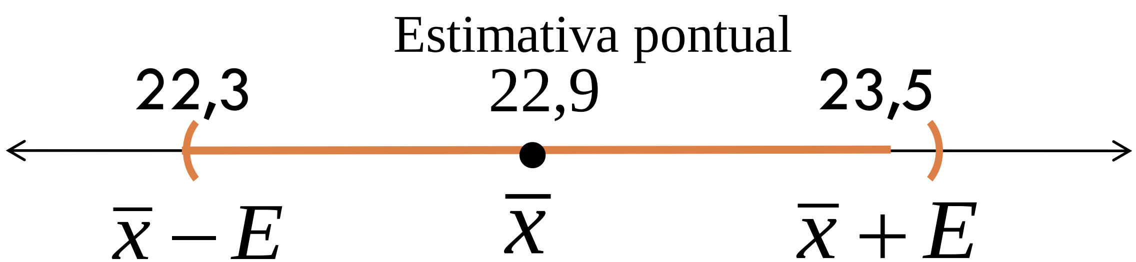
Extremo direito

$$\bar{x} + E$$
$$= 22.9 + 0.6$$
$$= 23.5$$

$$22{,}3 < \mu < 23{,}5$$

# Construindo um intervalo de confiança σ conhecido

**22,3 < μ < 23,5**



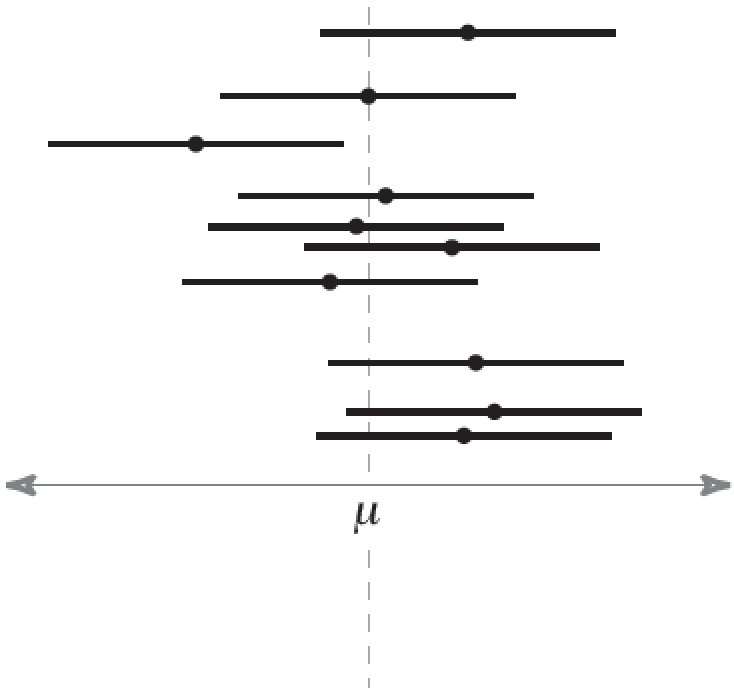

Com 90% de confiança, você pode dizer que a idade média de todos os estudantes está entre 22,3 e 23,5 anos.

# Interpretando os resultados

- $\mu$ é um número conhecido. Ou é um intervalo de confiança ou não.
- **Incorreto:** "Existe uma probabilidade de 90% que a média real esteja no intervalo (22,3, 23,5)."
- **Correto:** "Se um número grande de amostras é coletado e um intervalo de confiança é criado para cada uma, aproximadamente 90% desses intervalos conterão $\mu$.

# Interpretando os resultados

Os segmentos horizontais representam 90% de intervalos de confiança para diferentes amostras do mesmo tamanho. A longo prazo, 9 de cada 10 intervalos destes conterão μ.



# Tamanho da amostra

- Dado um nível de confiança $c$ e uma margem de erro $E$, o tamanho amostral mínimo $n$ necessário para estimar a média populacional $\mu$ é

$$n = \left( \frac{z_c \sigma}{E} \right)^2$$

- Se $\sigma$ é desconhecido, você pode estimar seu valor usando $s$ caso tenha uma amostra preliminar de pelo menos 30 membros.

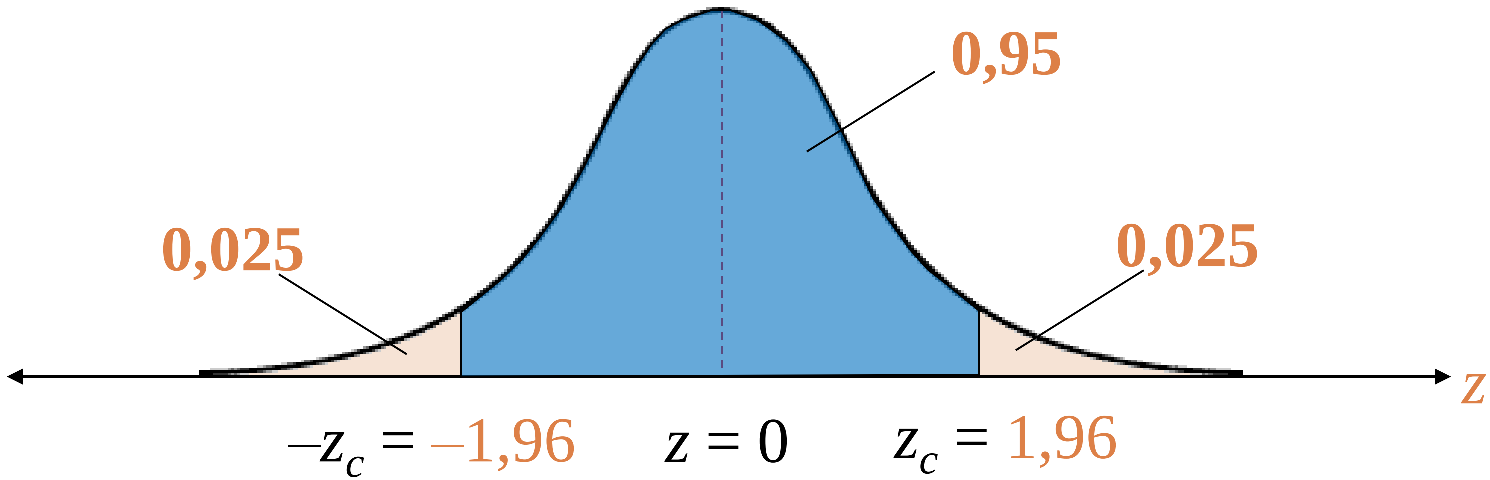
# Exemplo: tamanho da amostra

Você quer estimar o número médio de frases em anúncios de revista. Quantos anúncios de revista devem ser incluídos na amostra se você quer estar 95% confiante de que a média amostral esteja dentro de uma frase da média populacional? Assuma que o desvio padrão é aproximadamente 5,0.

Quero E=1

# Solução: tamanho da amostra

- Primeiro encontre os valores críticos




$$z_c = 1{,}96$$

# Tamanho da amostra

$z_c = 1{,}96 \qquad \sigma \approx s = 5{,}0 \qquad E = 1$

$$n = \left(\frac{z_c \sigma}{E}\right)^2 \approx \left(\frac{1.96 \cdot 5.0}{1}\right)^2 = 96.04$$

Quando necessário, **arredonde para cima** para obter um número inteiro.

Você deve incluir **pelo menos 97** anúncios de revistas em sua amostra.

# Tamanho da amostra



- Quantos anúncios de revista devem ser incluídos na amostra se você quiser estar 95% confiante de que a média amostral está dentro de duas frases da média populacional?

# Objetivos

- Encontrar uma estimativa pontual e uma margem de erro
- Construir e interpretar intervalos de confiança para a média populacional
- Determinar o tamanho mínimo da amostra necessária na estimativa de $\mu$

# INTERVALO DE CONFIANÇA PARA A MÉDIA (AMOSTRAS PEQUENAS)

Prof. Regina Meyer Branski

# Objetivos

- Interpretar a distribuição *t*
- Usar uma tabela de distribuição *t*
- Construir intervalos de confiança quando $n < 30$, a população é normalmente distribuída e $\sigma$ é desconhecido

# Distribuição *t*

- Quando o desvio padrão da população é desconhecido, o tamanho da amostra é menor que 30, e a variável *x* é normalmente distribuída; ela segue uma **distribuição *t***

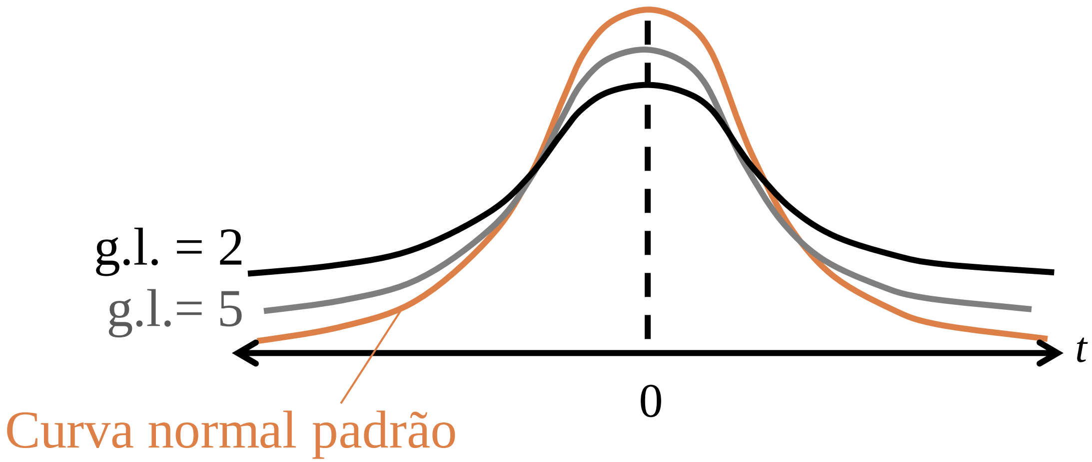
$$\boxed{t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}}$$

- Valores críticos de *t* são denotados por $t_c$

# Propriedades da distribuição *t*

1. A distribuição *t* tem formato de sino e é simétrica em relação à média.
2. A distribuição *t* é uma família de curvas, cada uma determinada por um parâmetro chamado de graus de liberdade.
3. Os **graus de liberdade** são o número de escolhas livres deixadas depois que uma amostra estatística, como $\overline{X}$, é calculada. Quando usamos a distribuição *t* para estimar a média da população, os graus de liberdade são iguais ao tamanho da amostra menos um.
4. g.l. $= n - 1$ Graus de liberdade

5. A área total sob a curva *t* é 1 ou 100%.
6. A média, a mediana e a moda da distribuição *t* são iguais a zero.
7. Conforme os graus de liberdade aumentam, a distribuição *t* aproxima-se da distribuição normal. Depois de 30 g.l., a distribuição *t* está muito próxima da distribuição normal padrão z.




As caudas na distribuição *t* são "mais grossas" que aquelas da distribuição normal padrão.

# Exemplo: valores críticos de *t*

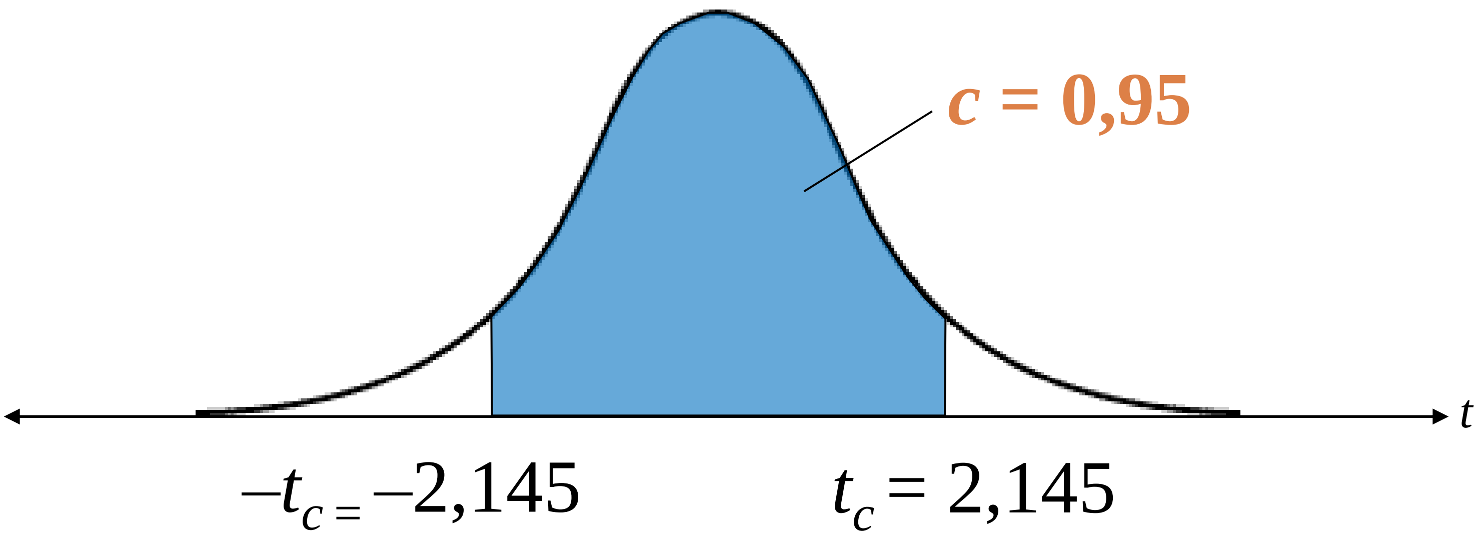
Encontre o valor crítico de $t_c$ para uma confiança de 95% quando o tamanho da amostra é 15.

$$\text{g.l.} = n - 1 = 15 - 1 = 14$$

| | Nível de confiança, *c* | 0,50 | 0,80 | 0,90 | 0,95 | 0,98 |
|---|---|---|---|---|---|---|
| | Uma cauda, α | 0,25 | 0,10 | 0,05 | 0,025 | 0,01 |
| g.l. | Duas caudas, α | 0,50 | 0,20 | 0,10 | 0,05 | 0,02 |
| 1 | | 1,000 | 3,078 | 6,314 | 12,706 | 31,821 |
| 2 | | 0,816 | 1,886 | 2,920 | 4,303 | 6,965 |
| 3 | | 0,765 | 1,638 | 2,353 | 3,182 | 4,541 |
| 12 | | 0,695 | 1,356 | 1,782 | 2,179 | 2,681 |
| 13 | | 0,694 | 1,350 | 1,771 | 2,160 | 2,650 |
| 14 | | 0,692 | 1,345 | 1,761 | 2,145 | 2,624 |
| 15 | | 0,691 | 1,341 | 1,753 | 2,131 | 2,602 |
| 16 | | 0,690 | 1,337 | 1,746 | 2,120 | 2,583 |
| 28 | | 0,683 | 1,313 | 1,701 | 2,048 | 2,467 |
| 29 | | 0,683 | 1,311 | 1,699 | 2,045 | 2,462 |
| ∞ | | 0,674 | 1,282 | 1,645 | 1,960 | 2,326 |

# Solução: valores críticos de *t*

95% da área sob a curva da distribuição *t* com 14 graus de liberdade está entre $t = \pm 2{,}145$.



# Intervalos de confiança para a média populacional

Um intervalo de confiança c para a média populacional $\mu$

- $$\boxed{\bar{x} - E < \mu < \bar{x} + E \quad \text{where } E = t_c \frac{s}{\sqrt{n}}}$$

- A probabilidade de que o intervalo de confiança contenha $\mu$ é c

# Intervalos de confiança e a distribuição *t*

| *Em palavras* | *Em símbolos* |
|---|---|
| 1. Identifique as amostras estatísticas $n$, $\overline{x}$ e $s$. | $\overline{x} = \frac{\Sigma x}{n}$ $\quad s = \sqrt{\frac{\Sigma(x-\overline{x})^2}{n-1}}$ |
| 2. Identifique os graus de liberdade, o nível de confiança c e o valor crítico $t_c$. | $\text{g.l.} = n - 1$ |
| 3. Encontre a margem de erro $E$. | $E = t_c \frac{s}{\sqrt{n}}$ |

| *Em palavras* | *Em símbolos* |
|---|---|
| 4. Encontre os extremos esquerdo e direito e forme um intervalo de confiança. | Extremo esquerdo: $\overline{x} - E$<br>Extremo direito: $\overline{x} + E$<br>Intervalo: $\overline{x} - E < \mu < \overline{x} + E$ |

# Exemplo: construindo um intervalo de confiança

Você seleciona aleatoriamente 16 cafeterias e mede a temperatura do café vendido em cada uma delas. A média de temperatura da amostra é 162,0°F com desvio padrão da amostra de 10,0°F. Encontre um intervalo de confiança de 95% para a temperatura média. Assuma que as temperaturas são normalmente distribuídas.

# Exemplo: construindo um intervalo de confiança

Você seleciona aleatoriamente 16 cafeterias e mede a temperatura do café vendido em cada uma delas. A média de temperatura da amostra é 162,0°F com desvio padrão da amostra de 10,0°F. Encontre um intervalo de confiança de 95% para a temperatura média. Assuma que as temperaturas são normalmente distribuídas.

Use a distribuição $t$ ($n$ < 30, σ é desconhecido, temperaturas são normalmente distribuídas.)

# Solução: construindo um intervalo de confiança

$t$. Usando $n = 16$, $\bar{x} = 162{,}0$, $s = 10{,}0$, $c = 0{,}95$ e g.l. = 15, você pode usar a Tabela 5 para encontrar que $t_c = 2.131$. A margem de erro no intervalo de confiança de 95% é:

| | Nível de confiança, $c$ | 0,50 | 0,80 | 0,90 | 0,95 | 0,98 |
|---|---|---|---|---|---|---|
| | Uma cauda, $\alpha$ | 0,25 | 0,10 | 0,05 | 0,025 | 0,01 |
| g.l. | Duas caudas, $\alpha$ | 0,50 | 0,20 | 0,10 | 0,05 | 0,02 |
| 1 | | 1,000 | 3,078 | 6,314 | 12,706 | 31,821 |
| 2 | | 0,816 | 1,886 | 2,920 | 4,303 | 6,965 |
| 3 | | 0,765 | 1,638 | 2,353 | 3,182 | 4,541 |
| 12 | | 0,695 | 1,356 | 1,782 | 2,179 | 2,681 |
| 13 | | 0,694 | 1,350 | 1,771 | 2,160 | 2,650 |
| 14 | | 0,692 | 1,345 | 1,761 | 2,145 | 2,624 |
| 15 | | 0,691 | 1,341 | 1,753 | 2,131 | 2,602 |
| 16 | | 0,690 | 1,337 | 1,746 | 2,120 | 2,583 |
| 28 | | 0,683 | 1,313 | 1,701 | 2,048 | 2,467 |
| 29 | | 0,683 | 1,311 | 1,699 | 2,045 | 2,462 |
| $\infty$ | | 0,674 | 1,282 | 1,645 | 1,960 | 2,326 |

$t_c = 2.131$

- Margem de erro:

$$E = t_c \frac{s}{\sqrt{n}} = 2.131 \cdot \frac{10}{\sqrt{16}} \approx 5.3$$

- Intervalo de confiança:

Extremo esquerdo:

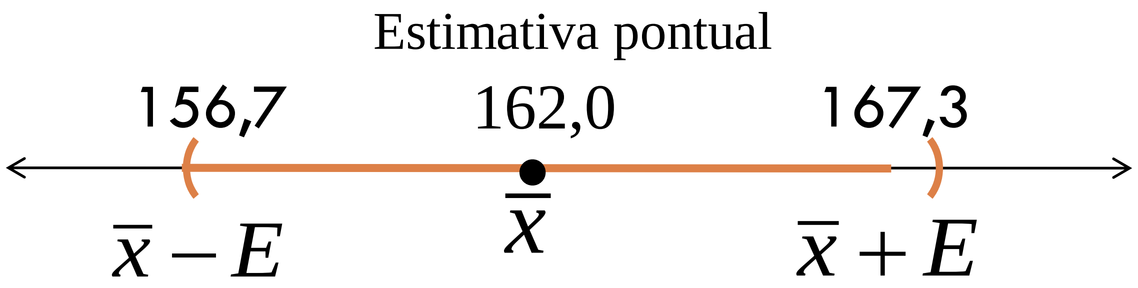
$$\bar{x} - E$$
$$= 162 - 5.3$$
$$= 156.7$$

Extremo direito:

$$\bar{x} + E$$
$$= 162 + 5.3$$
$$= 167.3$$

**156,7 < μ < 167,3**

- **156,7 < μ < 167,3**




Com 95% de confiança, você pode dizer que a temperatura média do café vendido está entre 156.7°F e 167.3°F.

# Exemplo: construindo um intervalo de confiança

Você seleciona aleatoriamente 20 instituições que realizam financimaneto pra compra da casa própria e determina o atual índice de juros do financiamento em cada. A média da amostra´dos juros é de 6,22% com desvio padrão de 0,42%. Encontre o intervalo de confiança de 99% par a média populacional do índice de juros do financimanento. Assuma que os índices de juros são aproximadamente normalmente distribuidos.

# Normal ou distribuição *t*?

$n \geq 30?$

**Sim** →

Use a distribuição normal com

$$E = z_c \frac{\sigma}{\sqrt{n}}$$

Se $\sigma$ for desconhecido, use *s*.

**Não** ↓

A população é distribuída normalmente, ou aproximadamente normal?

**Não** →

Não pode usar a distribuição normal ou a distribuição *t*.

**Sim** ↓

$\sigma$ é conhecido?

**Sim** →

Use a distribuição normal com

$$E = z_c \frac{\sigma}{\sqrt{n}}$$

**Não** ↓

Use a distribuição *t* com

$$E = t_c \frac{s}{\sqrt{n}}$$

.

# Exemplo: normal ou distribuição *t*?

Você seleciona aleatoriamente 25 casas construídas recentemente. A média amostral do custo da construção é $181.000 e o desvio padrão da população é de $28.000. Assumindo que os custos com a construção são normalmente distribuídos, você deve usar a distribuição normal, a distribuição *t* ou nenhuma delas para construir um intervalo de confiança de 95% para a média populacional dos custos de construção? Explique seu raciocínio.

# Exemplo: normal ou distribuição *t*?

Você seleciona aleatoriamente 25 casas construídas recentemente. A média amostral do custo da construção é \$181.000 e o desvio padrão da população é de \$28.000. Assumindo que os custos com a construção são normalmente distribuídos, você deve usar a distribuição normal, a distribuição *t* ou nenhuma delas para construir um intervalo de confiança de 95% para a média populacional dos custos de construção? Explique seu raciocínio.

**Use a distribuição normal** (a população é normalmente distribuída e o desvio padrão da população é conhecido).

# Objetivos



- Interpretamos a distribuição *t* e usamos uma tabela de distribuição *t*

- Construímos intervalos de confiança quando $n < 30$, a população é normalmente distribuída e σ é desconhecido

# INTERVALOS DE CONFIANÇA PARA PROPORÇÕES POPULACIONAIS

Prof. Regina Meyer Branski

# Objetivos

- Encontrar uma estimativa pontual para a proporção populacional
- Construir um intervalo de confiança para uma proporção populacional
- Determinar o tamanho mínimo da amostra quando estimamos uma proporção populacional

# Estimativa pontual para população *p*

- Proporção populacional
  - A probabilidade de **sucesso** em uma única tentativa de um experimento binomial denotado por p
- Estimar uma proporção populacional p usando intervalo de confiança
  - começa com uma estimativa pontual
  - Proporção de sucessos em uma amostra

**x – número de sucessos em um exemplo**

**n – número de exemplos**

**Leia como "p chapéu"**

$$\hat{p} = \frac{x}{n},$$

| Parâmetro populacional estimado | Com amostra estatística |
|---|---|
| Proporção: $p$ | $\hat{p}$ |

Estimativa pontual para $q$, a proporção das falhas

- Denotado por $\hat{q} = 1 - \hat{p}$
- Leia como "$q$ chapéu"

# Exemplo: estimativa pontual para *p*

Em uma pesquisa com 1.219 adultos norte-americanos, 354 disseram que seu esporte favorito era o futebol americano. Encontre uma estimativa pontual para a proporção populacional de adultos norte-americanos que dizem que seu esporte favorito é o futebol. (*Adaptado de The Harris Poll*.)

# Exemplo: estimativa pontual para *p*

Em uma pesquisa com 1.219 adultos norte-americanos, 354 disseram que seu esporte favorito era o futebol americano. Encontre uma estimativa pontual para a proporção populacional de adultos norte-americanos que dizem que seu esporte favorito é o futebol. (*Adaptado de The Harris Poll*.)

$$n = 1219 \quad \text{e} \quad x = 354$$

$$\hat{p} = \frac{x}{n} = \frac{354}{1219} \approx 0.290402 \approx 29.0\%$$

# Intervalos de confiança para *p*

Um intervalo de confiança c para a proporção populacional *p*

$$\hat{p} - E < p < \hat{p} + E \quad \text{where } E = z_c\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Distribuição Binomial pode ser aproximada a uma normal se

$$n\hat{p} \geq 5, \quad n\hat{q} \geq 5$$

# Construindo intervalos de confiança para *p*

| *Em palavras* | *Em símbolos* |
| --- | --- |
| 1. Identifique as estatísticas amostrais *n* e *x*. | |
| 2. Encontre a estimativa pontual $\hat{p}$. | $\hat{p} = \dfrac{x}{n}$ |
| 3. Verifique se a distribuição amostral de $\hat{p}$ pode ser aproximada por distribuição normal. | $n\hat{p} \geq 5, \quad n\hat{q} \geq 5$ |
| 4. Encontre o valor crítico $z_c$ que corresponda ao dado nível de confiança c. | Use a tabela normal padrão |

| *Em palavras* | *Em símbolos* |
| --- | --- |
| 5. Encontre a margem de erro *E*. | $E = z_c\sqrt{\frac{\hat{p}\hat{q}}{n}}$ |
| 6. Encontre os extremos esquerdo e direito e forme o intervalo de confiança. | Extremo esquerdo: $\hat{p} - E$<br>Extremo direito: $\hat{p} + E$<br>Intervalo: $\hat{p} - E < p < \hat{p} + E$ |

# Exemplo: intervalo de confiança para *p*

Em uma pesquisa com 1.219 adultos norte-americanos, 354 disseram que seu esporte favorito para assistir era o futebol americano. Construa um intervalo de confiança de 95% para a proporção de adultos nos Estados Unidos que dizem que seu esporte favorito é o futebol americano.

Lembre-se: $\hat{p} \approx 0.290402$

# Exemplo: intervalo de confiança para *p*

Em uma pesquisa com 1.219 adultos norte-americanos, 354 disseram que seu esporte favorito para assistir era o futebol americano. Construa um intervalo de confiança de 95% para a proporção de adultos nos Estados Unidos que dizem que seu esporte favorito é o futebol americano.

Lembre-se: $\hat{p} \approx 0.290402$

$$\hat{q} = 1 - \hat{p} = 1 - 0.290402 = 0.709598$$

# Solução: intervalo de confiança para *p*

- Verifique se a distribuição amostral de $\hat{p}$ pode ser aproximada pela distribuição normal

$$n\hat{p} \approx 1219 \cdot 0.290402 \approx 354 > 5$$

$$n\hat{q} \approx 1219 \cdot 0.709598 \approx 865 > 5$$

- Margem de erro:

$$E = z_c \sqrt{\frac{\hat{p}\hat{q}}{n}} \approx 1.96 \sqrt{\frac{(0.290402) \cdot (0.709598)}{1219}} \approx 0.025$$

# Solução: intervalo de confiança para *p*

- Intervalo de confiança para p:

Extremo esquerdo:

$$\hat{p} - E$$
$$= 0.29 - 0.025$$
$$= 0.265$$

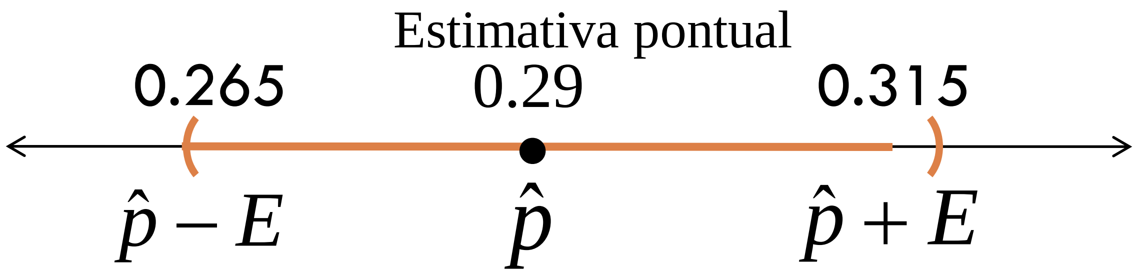
Extremo direito:

$$\hat{p} + E$$
$$= 0.29 + 0.025$$
$$= 0.315$$

$$\mathbf{0{,}265 < p < 0{,}315}$$

# Solução: intervalo de confiança para *p*

- $0{,}265 < p < 0{,}315$




Com 95% de confiança, você pode dizer que a proporção de adultos que dizem que o futebol americano é seu esporte favorito está entre 26,5% e 31,5%.

# Tamanho da amostra

- Dado um nível de confiança c e uma margem de erro *E*, o tamanho mínimo da amostra *n* necessário para estimar *p* é

$$n = \hat{p}\hat{q}\left(\frac{z_c}{E}\right)^2$$

- Essa fórmula assume que você tem uma estimativa para $\hat{q}$ e $\hat{p}$
- Senão, use $\hat{p} = 0.5$ e $\hat{q} = 0.5$.

# Exemplo: tamanho da amostra

Você está analisando uma campanha política e quer estimar, com 95% de confiança, a proporção dos eleitores registrados que irão votar no seu candidato. Sua estimativa deve ter uma margem de erro de 3% da população real. Encontre o número da amostragem mínimo necessário se:

Não há estimativas preliminares disponíveis.

# Exemplo: tamanho da amostra

Você está analisando uma campanha política e quer estimar, com 95% de confiança, a proporção dos eleitores registrados que irão votar no seu candidato. Sua estimativa deve ter uma margem de erro de 3% da população real. Encontre o número da amostragem mínimo necessário se:

1. Não há estimativas preliminares disponíveis.


Porque você não tem uma estimativa preliminar para $\hat{p}$ use $\hat{p} = 0.5$ e $\hat{q} = 0.5$.

# Solução: tamanho da amostra



- c = 0,95 $\quad z_c = 1{,}96 \quad E = 0{,}03$

$$n = \hat{p}\hat{q}\left(\frac{z_c}{E}\right)^2 = (0.5)(0.5)\left(\frac{1.96}{0.03}\right)^2 \approx 1067.11$$

**Arredonde para cima** para o próximo número inteiro.

Sem estimativas preliminares, o tamanho amostral mínimo seria de **pelo menos 1.068 votantes.**

# Exemplo: tamanho da amostra

Você está analisando uma campanha política e quer estimar, com 95% de confiança, a proporção dos eleitores registrados que irão votar no seu candidato. Sua estimativa deve ter uma margem de erro de 3% da população real. Encontre o número da amostragem mínimo necessário se:

Uma estimativa preliminar dá $\hat{p} = 0.31$

# Exemplo: tamanho da amostra

Você está analisando uma campanha política e quer estimar, com 95% de confiança, a proporção dos eleitores registrados que irão votar no seu candidato. Sua estimativa deve ter uma margem de erro de 3% da população real. Encontre o número da amostragem mínimo necessário se:

2. Uma estimativa preliminar dá $\hat{p} = 0.31$.


Use a estimativa preliminar $\hat{p} = 0.31$

$$\hat{q} = 1 - \hat{p} = 1 - 0.31 = 0.69$$

# Solução: tamanho da amostra

- c = 0,95 $\quad z_c = 1{,}96 \quad E = 0{,}03$

$$n = \hat{p}\hat{q}\left(\frac{z_c}{E}\right)^2 = (0.31)(0.69)\left(\frac{1.96}{0.03}\right)^2 \approx 913.02$$

**Arredonde para cima** para o próximo número inteiro. Com uma estimativa preliminar de $\hat{p} = 0.31$ , o tamanho amostral mínimo deveria ser de **pelo menos 914 votantes.**

Precisa de uma amostra maior se não houver estimativas preliminares disponíveis.

# Objetivos

- Encontrar uma estimativa pontual para a proporção populacional
- Construir um intervalo de confiança para uma proporção populacional
- Determinar o tamanho amostral mínimo, quando estimamos uma proporção populacional

# INTERVALOS DE CONFIANÇA PARA VARIÂNCIA E DESVIO PADRÃO

Prof. Regina Meyer Branski

# Objetivos

- Interpretar a distribuição qui-quadrado
- Usar tabela de distribuição qui-quadrado
- Usar a distribuição qui-quadrado para construir um intervalo de confiança para a variância e o desvio padrão

# Distribuição qui-quadrado

- A estimativa pontual para $\sigma^2$ é $s^2$
- A estimativa pontual para $\sigma$ é $s$
- $s^2$ é a estimativa menos tendenciosa para $\sigma^2$

| Parâmetro populacional estimado | Com estatística amostral |
| --- | --- |
| Variância: $\sigma^2$ | $s^2$ |
| Desvio padrão: $\sigma$ | $s$ |

# Distribuição qui-quadrado

- Pode usar uma distribuição qui-quadrado para construir um intervalo de confiança para a variância e desvio padrão
- Se a variável aleatória *x* tem distribuição normal, então a distribuição de:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

forma uma **distribuição qui-quadrado** para amostras de qualquer tamanho $n > 1$

# Propriedades da distribuição qui-quadrado

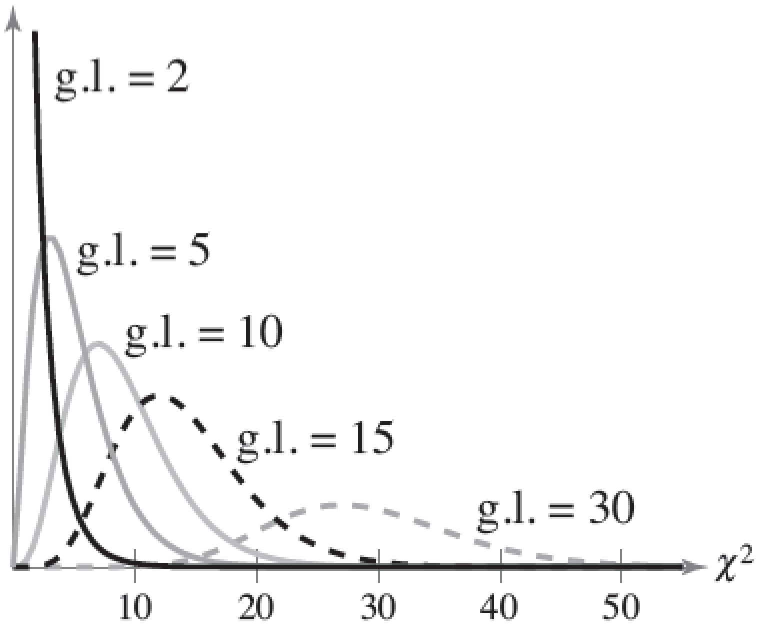
1. Todos valores qui-quadrado $\chi^2$ são maiores ou iguais a zero.

2. A distribuição qui-quadrado é uma família de curvas, cada uma determinada pelos graus de liberdade. Para formar um intervalo de confiança para $\sigma^2$, use a distribuição *Qui*² com graus de liberdade iguais a g.l. $= n - 1$

3. A área abaixo da curva da distribuição qui-quadrado é igual a um.

# Propriedades da distribuição qui-quadrado

4. As distribuições qui-quadrado são assimétricas positivas.

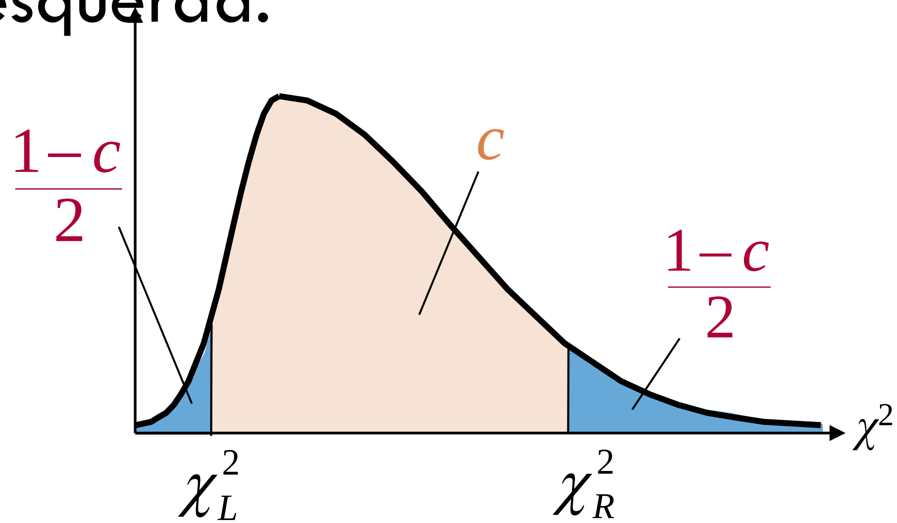



Distribuições qui-quadrado

# Valores críticos de $\chi^2$



- Há dois valores críticos para cada nível de confiança.
- O valor $\chi^2_R$ representa o valor crítico da cauda direita
- O valor $\chi^2_L$ representa o valor crítico da cauda esquerda.




A área entre os valores críticos esquerdo e direito é $c$.

# Exemplo: encontrando valores críticos para $\chi^2$

Encontre os valores críticos $\chi_R^2$ e $\chi_L^2$ para um intervalo de confiança de 90% quando o tamanho da amostra for 20.

# Exemplo: encontrando valores críticos para $\chi^2$

Encontre os valores críticos $\chi^2_R$ e $\chi^2_L$ para um intervalo de confiança de 90% quando o tamanho da amostra for 20.

$$\text{g.l.} = n - 1 = 20 - 1 = 19 \text{ g.l.}$$

- Cada área na tabela representa a região sob a curva qui-quadrado à *direita do valor crítico.*

- Área à direita de $\chi^2{}_R$ = $\dfrac{1-c}{2} = \dfrac{1-0.90}{2} = 0.05$

- Área à direita de $\chi^2{}_L$ = $\dfrac{1+c}{2} = \dfrac{1+0.90}{2} = 0.95$

# Solução: encontrando valores críticos para $\chi^2$

| Graus de liberdade | α | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0,995 | 0,99 | 0,975 | 0,95 | 0,90 | 0,10 | 0,05 |
| 1 | — | — | 0,001 | 0,004 | 0,016 | 2,706 | 3,841 |
| 2 | 0,010 | 0,020 | 0,051 | 0,103 | 0,211 | 4,605 | 5,991 |
| 3 | 0,072 | 0,115 | 0,216 | 0,352 | 0,584 | 6,251 | 7,815 |
| 15 | 4,601 | 5,229 | 6,262 | 7,261 | 8,547 | 22,307 | 24,996 |
| 16 | 5,142 | 5,812 | 6,908 | 7,962 | 9,312 | 23,542 | 26,296 |
| 17 | 5,697 | 6,408 | 7,564 | 8,672 | 10,085 | 24,769 | 27,587 |
| 18 | 6,265 | 7,015 | 8,231 | 9,390 | 10,865 | 25,989 | 28,869 |
| 19 | 6,844 | 7,633 | 8,907 | 10,117 | 11,651 | 27,204 | 30,144 |
| 20 | 7,434 | 8,260 | 9,591 | 10,851 | 12,443 | 28,412 | 31,410 |

$\chi^2_L$ $\chi^2_R$

Por meio da tabela, você pode ver que $\chi^2_R = 30{,}144$ e $\chi^2_L = 10{,}117$.

90% da área abaixo da curva está entre 10,117 e 30,144.

# Intervalos de confiança para σ² e σ

Intervalo de confiança para σ²:

$$\boxed{\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}}$$

Intervalo de confiança para σ:

$$\boxed{\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}}$$

A probabilidade de que o intervalo de confiança contenha σ² ou σ é c.

| ***Em palavras*** | ***Em símbolos*** |
| --- | --- |
| 1. Verifique se a população tem uma distribuição normal. | |
| 2. Identifique a amostra estatística *n* e os graus de liberdade. | $\text{g.l.} = n - 1$ |
| 3. Encontre a estimativa pontual $s^2$. | $s^2 = \dfrac{\sum (x - \bar{x})^2}{n-1}$ |
| 4. Encontre o valor crítico $\chi^2_R$ e $\chi^2_L$ que corresponda ao dado nível de confiança c. | Use a tabela |

| *Em palavras* | *Em símbolos* |
| --- | --- |
| 5. Encontre os extremo esquerdo e direito e forme o intervalo de confiança para a variância populacional. | $$\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$ |
| 6. Encontre o intervalo de confiança para o desvio padrão da população tomando a raiz quadrada de cada extremo. | $$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$ |

# Exemplo: construindo um intervalo de confiança

Você seleciona aleatoriamente 30 amostras de um antialérgico e as pesa. O desvio padrão da amostra é 1,20 miligrama. Supondo que os pesos são normalmente distribuídos, construa intervalos de confiança de 99% para a variância e o desvio padrão da população.

$$\text{g.l.} = n - 1 = 30 - 1 = 29 \text{ g.l.}$$

# Solução: construindo um intervalo de confiança

- Área à direita de $\chi^2_R = \dfrac{1-c}{2} = \dfrac{1-0.99}{2} = 0.005$
- Área à direita de $\chi^2_L = \dfrac{1+c}{2} = \dfrac{1+0.99}{2} = 0.995$
- Os valores críticos são

$$\chi^2_R = 52.336 \quad \text{e} \quad \chi^2_L = 13.121$$

# Solução: construindo um intervalo de confiança

Intervalo de confiança para $\sigma^2$:

Extremo esquerdo: $\dfrac{(n-1)s^2}{\chi_R^2} = \dfrac{(30-1)(1.20)^2}{52.336} \approx 0.80$

Extremo direito: $\dfrac{(n-1)s^2}{\chi_L^2} = \dfrac{(30-1)(1.20)^2}{13.121} \approx 3.18$

$$0,80 < \sigma^2 < 3,18$$

Com 99% de confiança você pode dizer que a variância da população está entre 0,80 e 3,18 miligramas.

# Solução: construindo um intervalo de confiança

Intervalo de confiança para σ:

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

$$\sqrt{\frac{(30-1)(1.20)^2}{52.336}} < \sigma < \sqrt{\frac{(30-1)(1.20)^2}{13.121}}$$

$$0{,}89 < \sigma < 1{,}78$$

Com 99% de confiança você pode dizer que o desvio padrão da população está entre 0,89 e 1,78 miligrama.

# Objetivos

- Interpretar a distribuição qui-quadrado
- Usar a tabela de distribuição qui-quadrado
- Usar a distribuição qui-quadrado para construir um intervalo de confiança para a variância e o desvio padrão

# Exercício



- Você é um agente de viagens que quer estimar, com 95% de confiança, a proporção de pessoas em férias que planejam viajar para fora do Brasil nos próximos 12 meses. Sua estimativa deve ter uma margem de erro de 3% da proporção real.
- A) não há estimativas prévias. Encontre o tamanho mínimo da amostra necessária
- B) encontre o tamanho minimo da amostra necessária usando um estudo anterior que mostrou que 26% dos pesquisados declararam que planejavam viajar para fora do Brasil nos próxmos 12 meses
- C) compare os resultados

# Exercício



**21.** (a) $n = \hat{p}\hat{q}\left(\frac{z_c}{E}\right)^2 = 0.5 \cdot 0.5\left(\frac{1.96}{0.03}\right)^2 \approx 1067.111 \to 1068$ vacationers

(b) $n = \hat{p}\hat{q}\left(\frac{z_c}{E}\right)^2 = 0.26 \cdot 0.74\left(\frac{1.96}{0.03}\right)^2 \approx 821.249 \to 822$ vacationers

(c) Having an estimate of the proportion reduces the minimum sample size needed.

# Exercício



- Um fabricante de máquinas para cortar grama está tentando determinar o desvio padrão da vida de um de seus modelos de máquinas. Para fazê-lo, ele seleciona aleatoriamente 12 máquinas que foram vendidas anos atrás e descobre que o desvio padrão da amostra é 3,25 anos. Use um nível de 95% de confiança para calcular os intervalos de confiança para a variância e o desvio padrão da população.

# Exercício



**11.** (a) $\left(\dfrac{(n-1)s^2}{\chi_R^2}, \dfrac{(n-1)s^2}{\chi_L^2}\right) = \left(\dfrac{11 \cdot (3.25)^2}{26.757}, \dfrac{11 \cdot (3.25)^2}{2.603}\right) \approx (4.34, 44.64)$

(b) $\left(\sqrt{4.342}, \sqrt{44.636}\right) \approx (2.08, 6.68)$