Intervalo de Confiança – Aula 02

Statistics for Business and Economics 7 edição, by Paul Newbold , William Carlson , Betty Thorne (cap. Estimation: Single Population)

Cap 11 Bussab e Morettin

Statistics for Economics, Accounting and Business Studies, capítulo 4, Barrow

Marislei Nishijima

Lembre-se

Parâmetro populacional (alguma característica da população que queremos conhecer)

Estimador amostral (Medida obtida da amostra para inferir algum parâmetro populacional)

População X, assume valores xi conforme uma distribuição de probabilidade (é uma variável aleatória)

Estatística amostral de X (media amostral, por exemplo, é um estimador amostral), também é uma variável aleatória porque varia conforme a amostra.

Intervalo de Confiança

- Intervalo de Confiança para a media Populacional, μ
  - Quando a variância populacional σ^2 é conhecida
  - Quando a variância populacional σ^2 é desconhecida
- Intervalo de Confiança para a proporcão Populacional, $\hat{p}$ (grandes amostras)

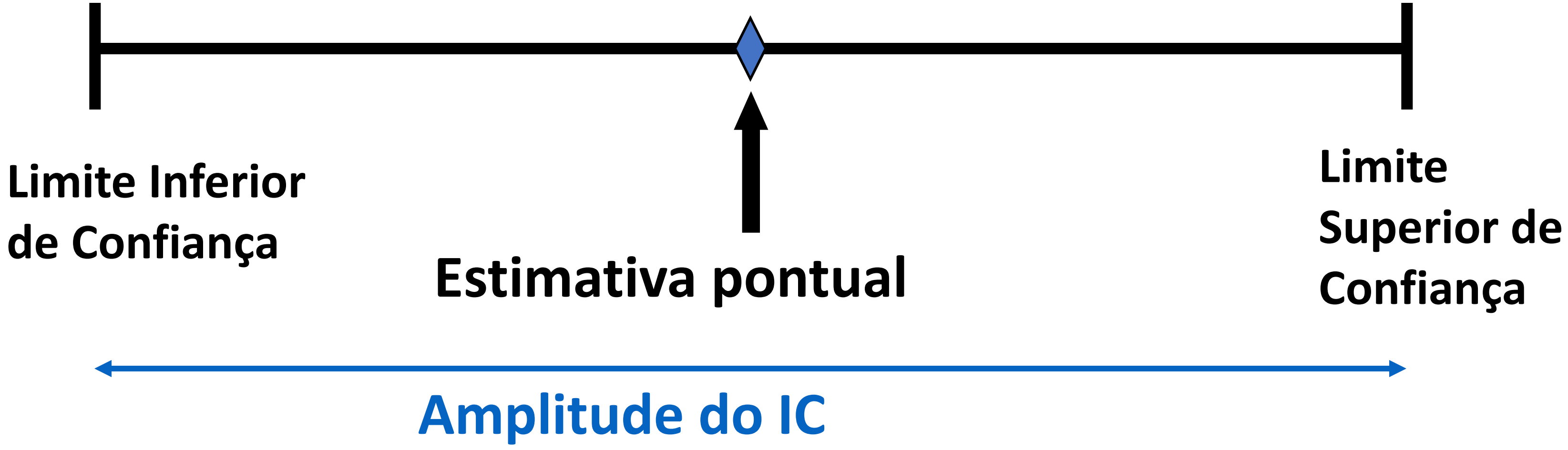
Definição

- Um estimador de um parâmetro populacional é
  - Uma variável aleatória que depende da informação da amostra . . .
  - Seu valor provê uma aproximação deste parâmetro desconhecido.

- Um valor específico desta variável aleatória é chamado de estimativa

Estimativas pontuais e por intervalo

- Uma estimative pontual é um número,
- Um intervalo de confiança provê informação adicional sobre variabilidade



Estimativas Pontuais

Podemos estimar um parâmetro populacional…		com uma estatística amostral (uma estimativa de ponto)
Média	μ	$\bar{x}$
Proporção	P	$\hat{p}$

Imparcialidade (ausência de viés)

- Um estimador pontual $\hat{\theta}$ é um estimador não viesado do *parâmetro* θ se o valor esperado, ou média, da distribuição amostral de $\hat{\theta}$ é θ,

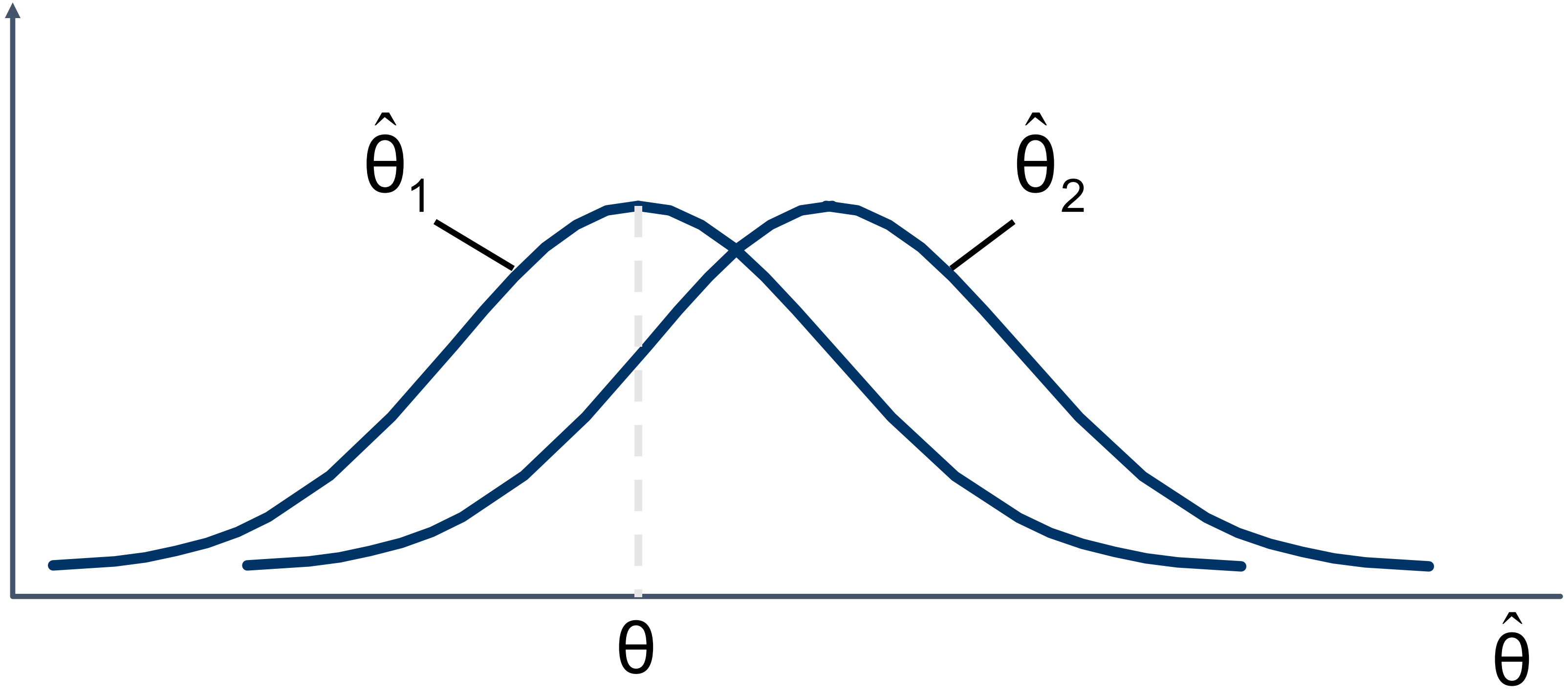
$$\boxed{E(\hat{\theta}) = \theta}$$

- Exemplos:
  - A média da amostra $\bar{x}$ é um estimador não-viesado de μ
  - A variância da amostra s^2 é um estimador não-viesado de σ^2
  - A proporção da amostra $\hat{p}$ é um estimador não-viesado de P

Imparcialidade (ausência de viés)

(cont.)

- $\hat{\theta}_1$ é um estimador não-viesado, $\hat{\theta}_2$ é viesado:



Viés

- Seja $\hat{\theta}$ um estimador de θ
- O viés em $\hat{\theta}$ é definido como a diferença entre sua média e θ

$$\boxed{\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta}$$

- O viés de um estimador não-viesado é zero

Estimator Mais Eficiênte

- Suponha que existam vários estimadores de θ
- O estimador mais eficiênte ou o estimador de mínima variância e não-viesado de θ é o estimador não-viesado com a menor variância

- Seja $\hat{\theta}_1$ e $\hat{\theta}_2$ dois estimadores não-viesados de θ, calculados com o mesmo número de observações de amostra. Então,
  - $\hat{\theta}_1$ é mais eficiênte do que $\hat{\theta}_2$ se $\boxed{\mathrm{Var}(\hat{\theta}_1) < \mathrm{Var}(\hat{\theta}_2)}$
  - A eficiência relativa de $\hat{\theta}_1$ com respeito a $\hat{\theta}_2$ é a taxa de suas variâncias:

$$\boxed{\text{Relative Efficiency} = \frac{\mathrm{Var}(\hat{\theta}_2)}{\mathrm{Var}(\hat{\theta}_1)}}$$

Intervalo de Confiança

- Quanta incerteza está associada a uma estimativa pontual de um parâmetro populacional?
- Uma estimative em intervalo provê mais informação sobre uma característica da população do que uma estimativa pontual
- Estas estimativas em intervalo são os intervalos de confiança

Estimativa do intervalo de confiança

- Um intervalo fornece uma faixa de valores:
  - Considere a variação nas estatísticas amostrais de cada amostra
  - Com base nas observações de uma amostra
  - Fornece informação sobre proximidade sobre um parâmetro populacional desconhecido
  - É estabelecido em termos de nível de confiança
    - Nunca terá 100% de confiança

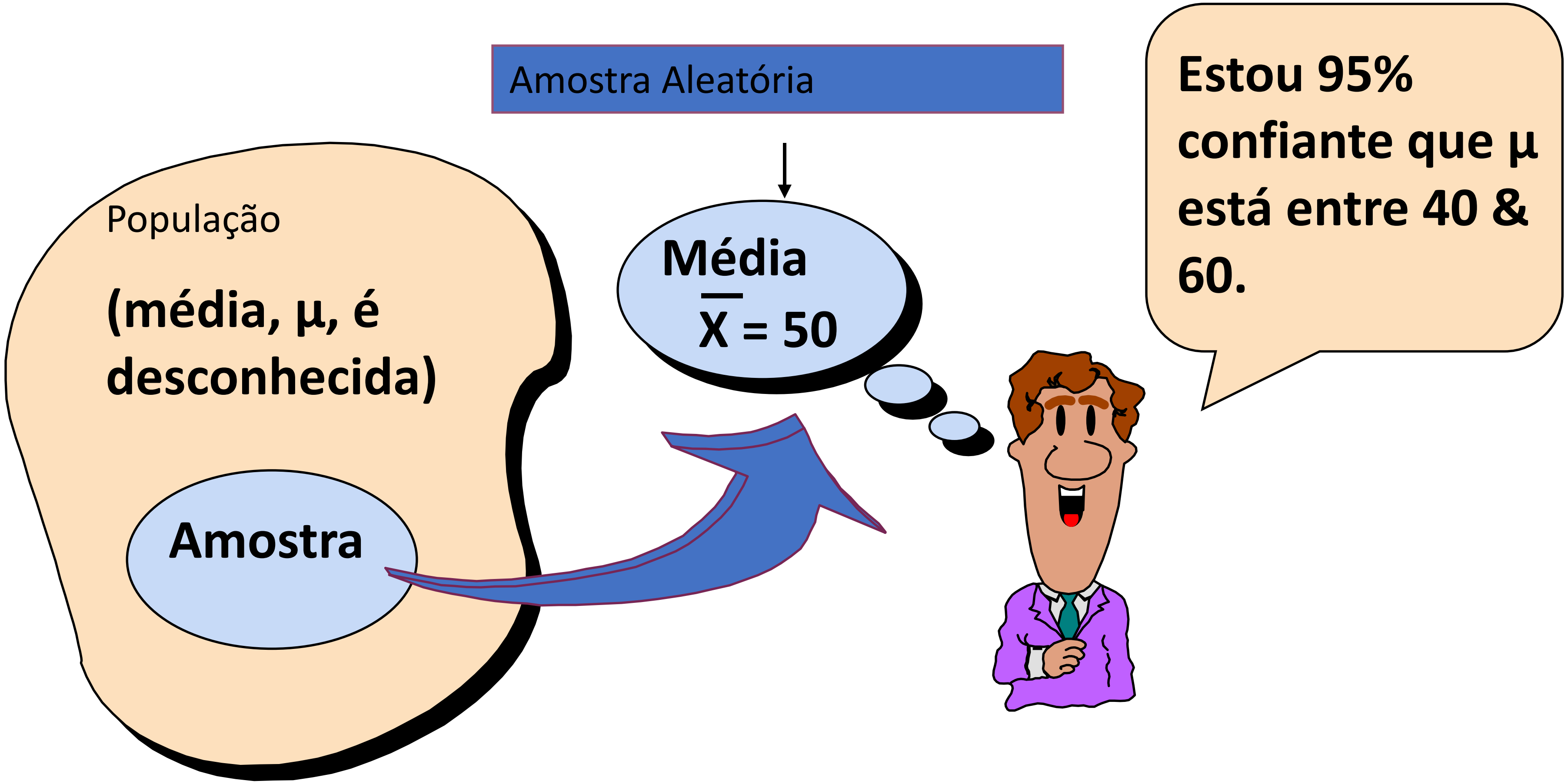
Intervalo de confiança e nível de confiança

- Se $P(a < \theta < b) = 1 - \alpha$ então, o interval de confiança de θ entre a e b é $100(1 - \alpha)\%$.

- A quantidade $(1 - \alpha)$ é o nível de confiança do intervalo (α varia de 0 a 1)

  - Em amostras repetidas da população, o valor verdadeiro do parâmetro θ deve estar contido em $100(1 - \alpha)\%$ dos intervalos calculados dessa maneira.
  - O intervalo de confiança calculado dessa maneira é escrito como $a < \theta < b$ com $100(1 - \alpha)\%$ de confiança

Processo de Estimação



Nível de Confiança, $(1-\alpha)$

(cont.)

- Suponha nível de confiança = 95%
- Também descrito por $(1 - \alpha) = 0.95$
- Uma interpretação de frequência relativa:
  - A partir de amostras repetidas, 95% de todos os intervalos de confiança que podem ser construídos conterão o parâmetro verdadeiro desconhecido
  - Um intervalo específico conterá ou não o parâmetro verdadeiro
  - Não há probabilidade envolvida em um intervalo específico

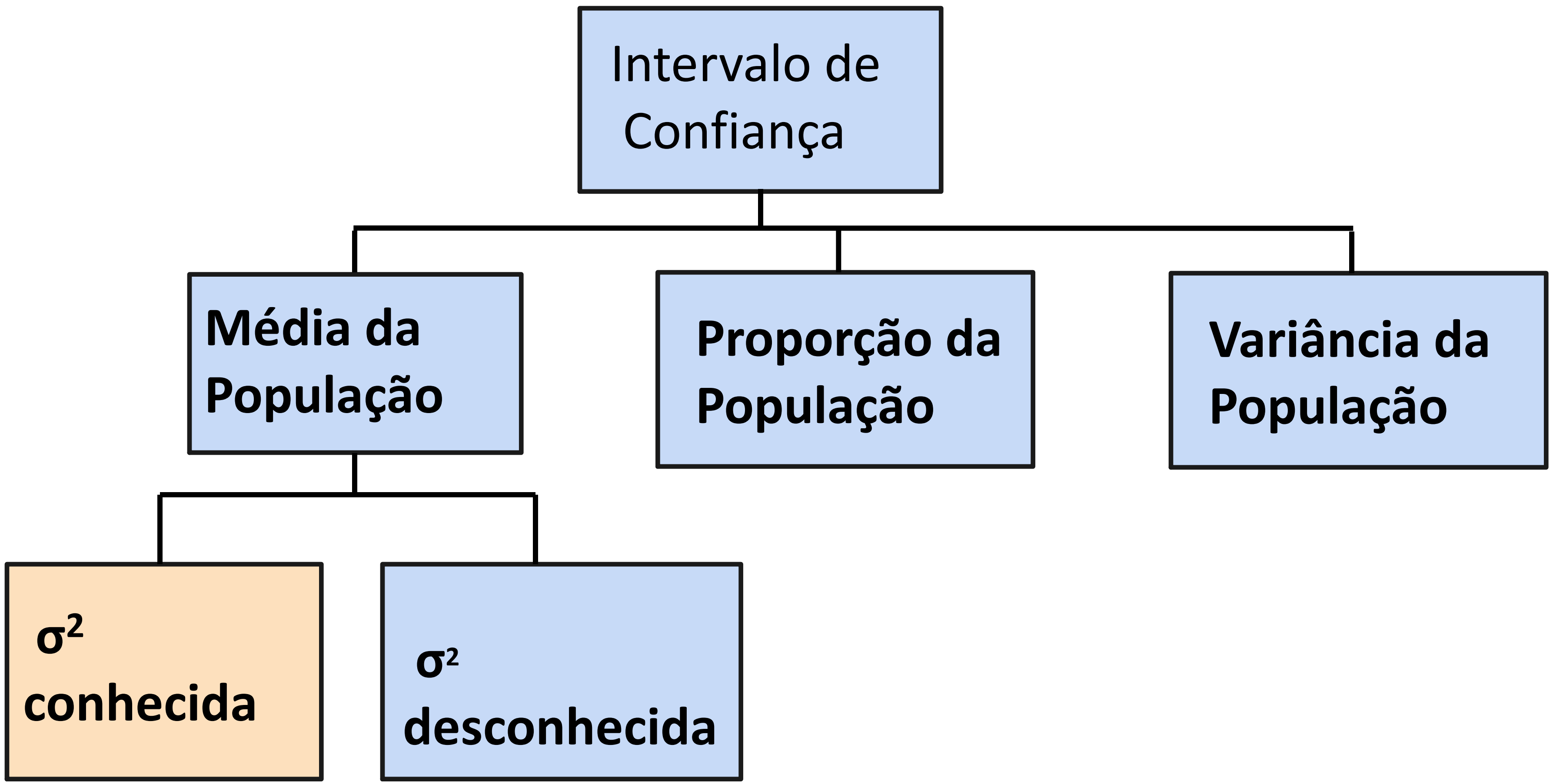
Fórmula Geral

- A fórmula geral para todos os intervalos de confiança é:

Valor pontual estimado $\pm$ (fator de confiabilidade)(Erro Padrão)

- O valor do fator de confiabilidade depende do nível de confiança desejado

Intervalo de Confiança



Intervalo de Confiança para μ (σ^2 Conhecida)

- Pressupostos
  - Variância da população σ^2 é conhecida
  - População é normalmente distribuída
  - Se população não é normal, use amostra grande

- Estimativa de Intervalo de Confiança :

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

(sendo $z_{\alpha/2}$ o valor da distribuicão normal para uma probabilidade de $\alpha/2$ em cada calda

Margem de erro

- O Intervalo de Confiança,

$$\boxed{\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}$$

- Também pode ser escrito $\boxed{\bar{x} \pm ME}$

  em que ME é a margem de erro

$$\boxed{ME = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}$$

- A amplitude do intervalo, w, corresponde a 2 vezes a margem de erro

Reduzindo a Margem de Erro

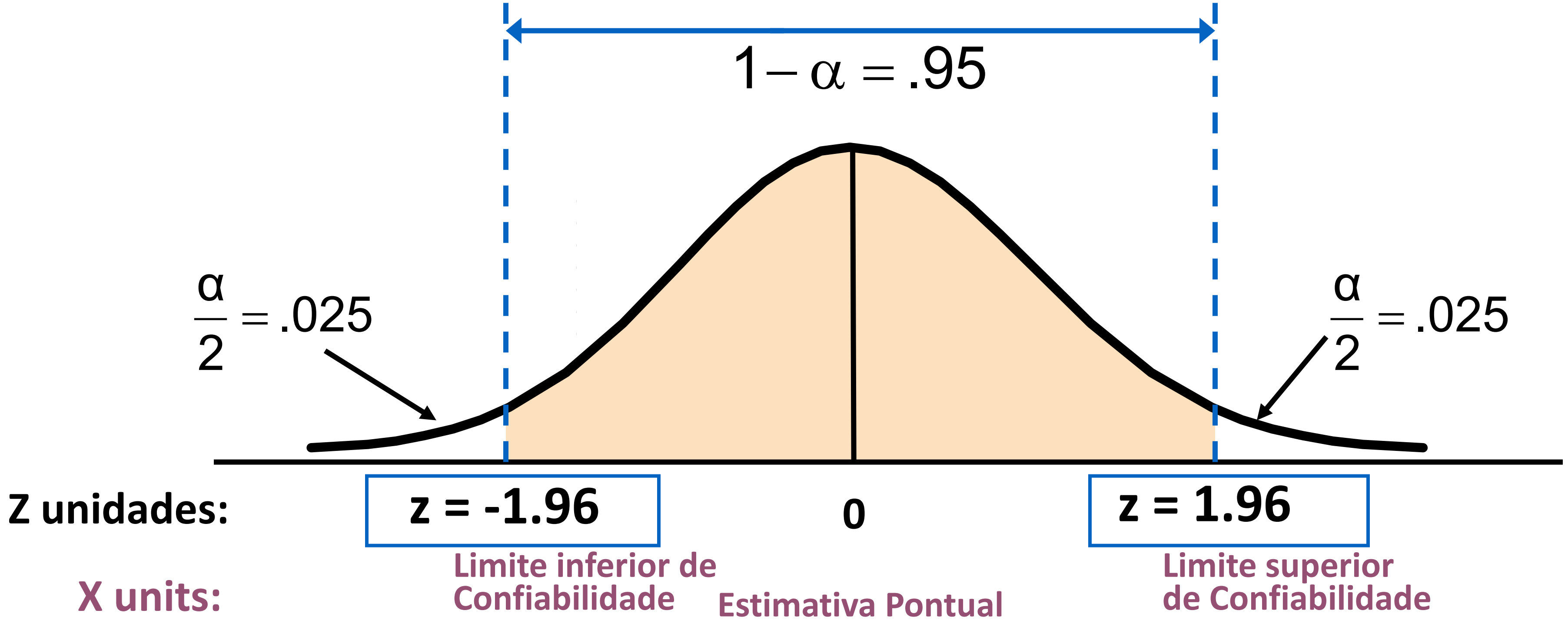
$$ME = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

A Margem de Erro pode ser reduzida se

- Se o desvio padrão da population puder ser reduzido ($\sigma\downarrow$)
- Se o tamanho da amostra aumenta ($n\uparrow$)
- Se o nível de confiança reduzir, $(1-\alpha)\downarrow$

Encontrando o Fator de confiabilidade, $z_{\alpha/2}$

- Conside um IC de 95% :



- Encontre $z_{.025} = \pm 1.96$ na tabela normal padrão

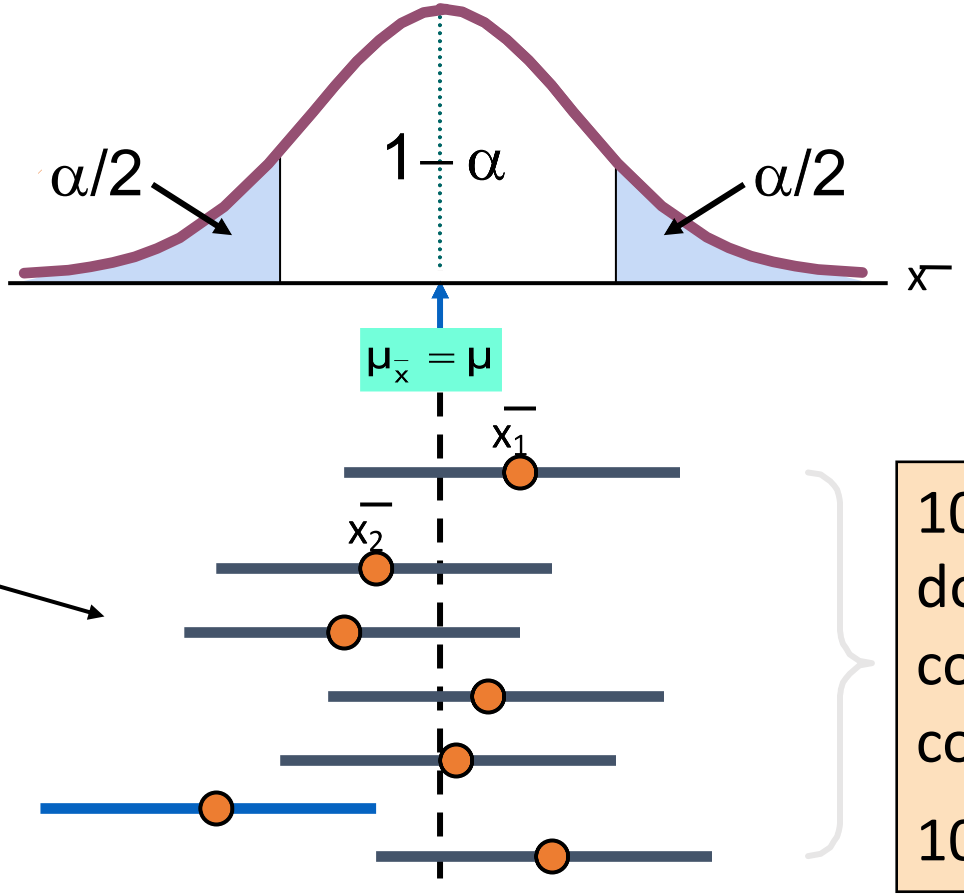
Níveis de confiança mais frequentes

- Os mais frequentes são 90%, 95%, e 99%

Nível de confiança	*Coeficiente de confiança,* $1-\alpha$	$Z_{\alpha/2}$ *valor*
80%	.80	1.28
90%	.90	1.645
95%	.95	1.96
98%	.98	2.33
99%	.99	2.58
99.8%	.998	3.08
99.9%	.999	3.27

Intervalos e níveis de *confiança*

Distribuição amostral da média



Intervalos de *confiança*

Os Intervalos se estendem de

$$LCL = \bar{x} - z\frac{\sigma}{\sqrt{n}}$$

para

$$UCL = \bar{x} + z\frac{\sigma}{\sqrt{n}}$$

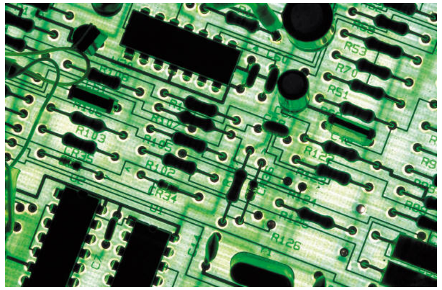
100(1-α)% dos intervalos construídos contém μ;

100(α)% não.

Exemplo

- Uma amostra de 11 circuitos de uma grande população com distribuição normal tem média de resistência de 2,20 ohms e desvio padrão da população é de 0,35 ohms.

- Determine um interval de confiança de 95% para a media de resistência da população.



Exemplo

- Uma amostra de 11 circuitos de uma grande população com distribuição normal tem média de resistência de 2.20 ohms e desvio padrão da população é de 0.35 ohms.

(cont.)

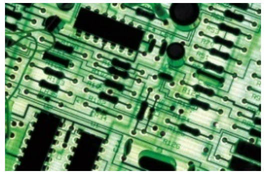
- Solução:

$$\bar{x} \pm z \frac{\sigma}{\sqrt{n}}$$

$$= 2.20 \pm 1.96\ (.35/\sqrt{11})$$

$$= 2.20 \pm .2068$$

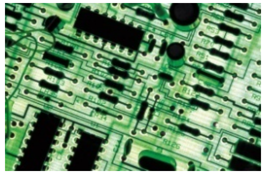
$$\boxed{1.9932 < \mu < 2.4068}$$



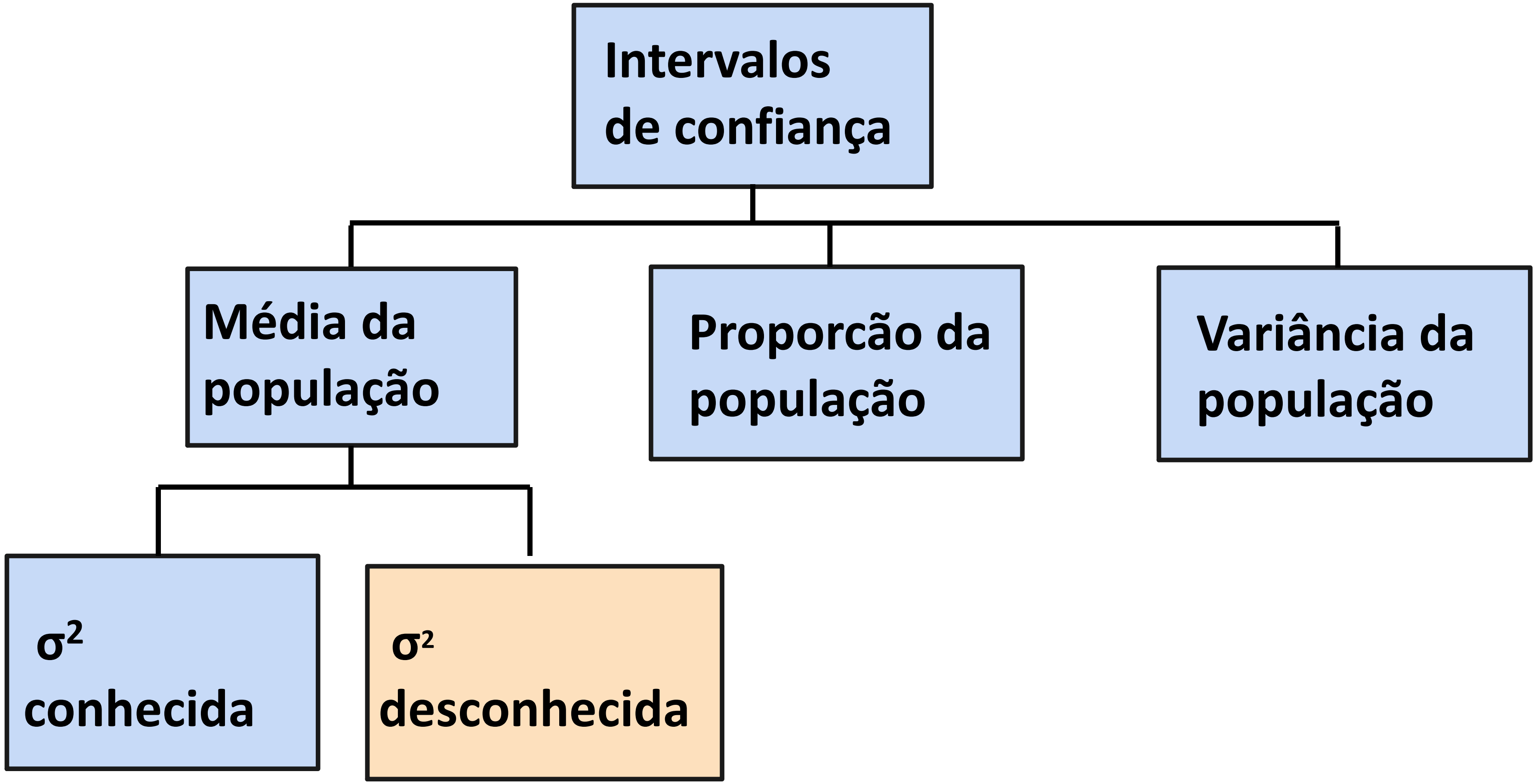
Interpretação

- Estamos 95% confiantes que a verdadeira média de resistência está entre 1.9932 e 2.4068 ohms

- Apesar da verdadeira media poder estar ou não no intervalo, 95% dos intervalos calculados dessa maneira conterão a media verdadeira.



Intervalos de confiança



Distribuição t de Student

- Considere uma amostra aleatória de n observações
  - Com media $\bar{x}$ e desvio padrão s
  - Proveniente de uma distribuição populacional normalmente distribuída com média μ

- A variável t, então,

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

segue uma distribuição t de Student com (n - 1) graus de liberdade

Intervalo de Confiança para μ (σ^2 Desconhecido)

- Se o desvio padrão da população σ for desconhecido, podemos substituir pelo desvio padrão da amostra, s
- Isto introduz maior incerteza, pois s é variável de amostra para amostra
- Então usamos a distribuição t ao invés da distribuição normal

Intervalo de Confiança para μ (σ^2 Desconhecido)

(cont.)

- Hipótese
  - O desvio padrão da população é desconhecido
  - A população é normalmente distribuída
  - Se a população não é normal, usar amostra grande
- Use a Distribuição t de Student
- Estimativa do IC:

$$\boxed{\bar{x} - t_{n-1,\alpha/2} \frac{S}{\sqrt{n}} < \mu < \bar{x} + t_{n-1,\alpha/2} \frac{S}{\sqrt{n}}}$$

sendo $t_{n-1,\alpha/2}$ o valor crítico da distribuição t com $n-1$ g.l. e uma área de $\alpha/2$ em cada calda:

$$\boxed{P(t_{n-1} > t_{n-1,\alpha/2}) = \alpha/2}$$

Margem de Erro

- O Intervalo de Confiança,

$$\bar{x} - t_{n\text{-}1,\alpha/2}\frac{S}{\sqrt{n}} < \mu < \bar{x} + t_{n\text{-}1,\alpha/2}\frac{S}{\sqrt{n}}$$

- Também pode ser escrito: $\bar{x} \pm ME$

em que ME é a margem de erro:

$$ME = t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}$$

Distribuição t de Student

- A t é uma família de distributições
- O t valor depende dos graus de liberdade (g.l.)
  - Número de observações que são livres para variar depois da media amostral ser calculada.

$$\boxed{d.f. = n - 1}$$

Distribuição t de Student

Note: t $\longrightarrow$ Z quando n cresce

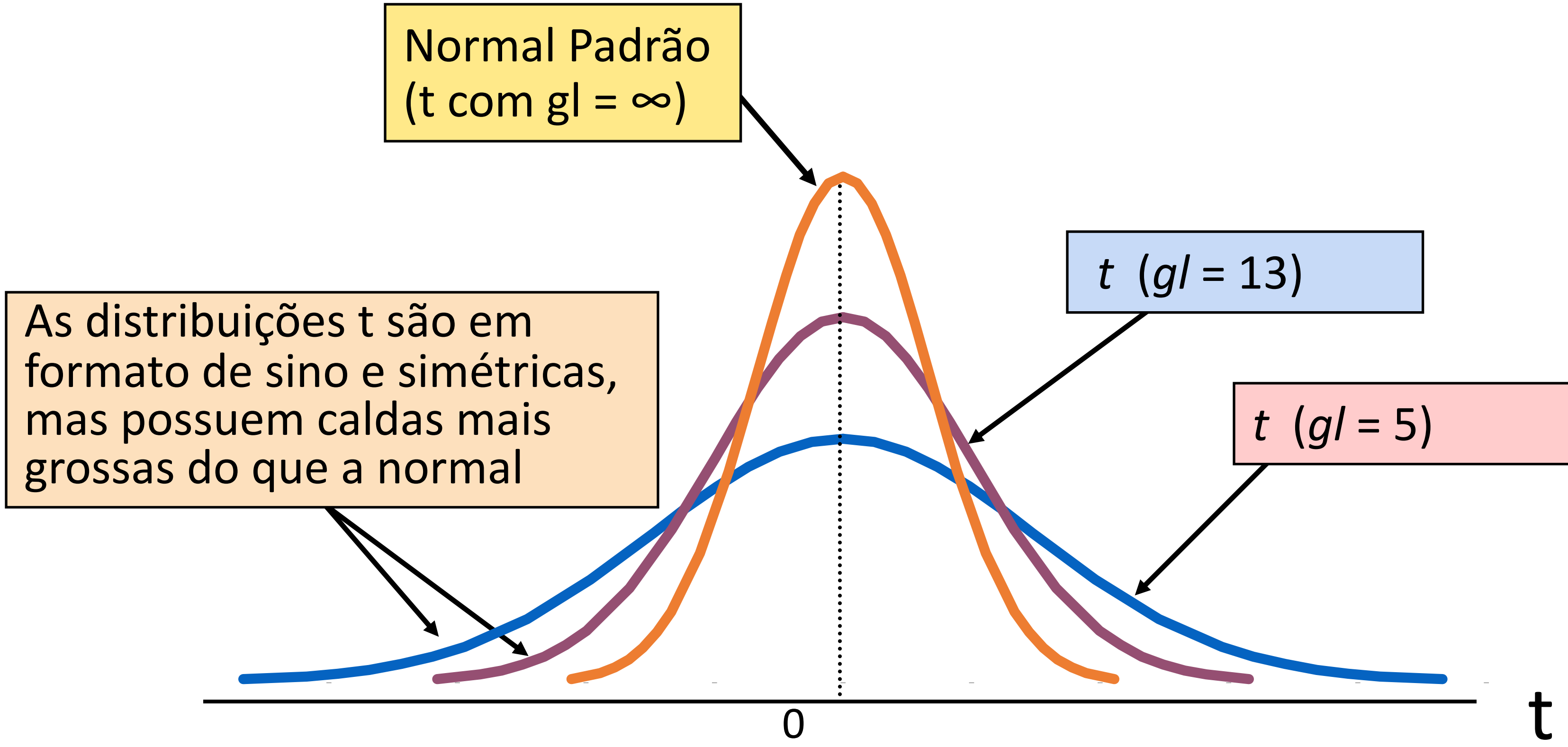
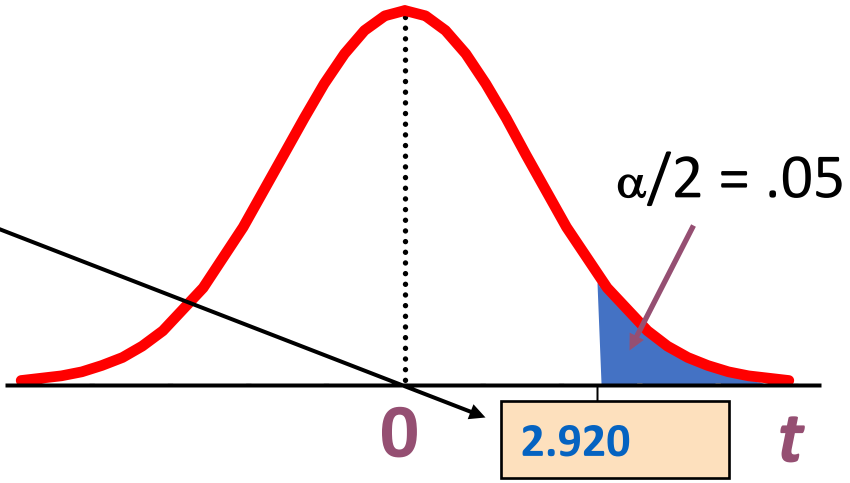


Tabela da Distribuição t de Student

	Área da calda superior		
gl	.10	**.05**	.025
1	3.078	6.314	12.706
2	1.886	**2.920**	4.303
3	1.638	2.353	3.182

Seja: n = 3
gl = *n* - 1 = 2
α = .10
$\alpha/2$ =.05

O corpo da tabela contém os t valores, não as probabilidades



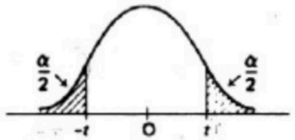
Valores da distribuição t

Comparando com Z valores

Confiança Level	t (10 g.l.)	t (20 g.l.)	t (30 g.l.)	Z
.80	1.372	1.325	1.310	1.282
.90	1.812	1.725	1.697	1.645
.95	2.228	2.086	2.042	1.960
.99	3.169	2.845	2.750	2.576

Note: $t \longrightarrow Z$ quando n cresce

TABELA DA DISTRIBUIÇÃO t de STUDENT



φ \ α	0,50	0,25	0,10	0,05	0,025	0,01	0,005
1	1,00000	2,4142	6.3138	12.706,	25,542	63,657	127,32
2	0,81650	1,6036	2,9200	4,3127	6,2053	9,9248	14,089
3	0,76489	1,4226	2,3534	3,1825	4,1765	5,8409	7,4533
4	0,74070	1,3444	2,1318	2,7764	3,4954	4,6041	5,5976
5	0,72669	1,3009	2,0150	2,5706	3,1634	4,0321	4,7733
6	0,71756	1,2733	1,9432	2,4469	2,9687	3,7074	4,3168
7	0,71114	1,2543	1,8946	2,3646	2,8412	3,4995	4,0293
8	0,70639	1,2403	1,8595	2,3060	2,7515	3,3554	3,8325
9	0,70272	1,2297	1,8331	2,2622	2,6850	3,2498	3,6897
10	0,69981	1,2213	1,8125	2,2281	2,6338	3,1693	3,5814
11	0,69745	1,2145	1,7959	2,2010	2,5931	3,1058	3,4966
12	0,69548	1,2089	1,7823	2,1788	2,5600	3,9545	3,4284
13	0,69384	1,2041	1,7709	2,1604	2,5326	3,0123	3,3725
14	0,692	1,2001	1,7613	2,1448	2,5096	2,9768	3,3257
15	0,69120	1,1967	1,7530	2,1315	2,4899	2,9467	3,2860
16	0,69013	1,1937	1,7459	2,1199	2,4729	2,9208	3,2520
17	0,68919	1,1910	1,7396	2,1098	2,4581	2,8982	3,2225
18	0,68837	1,1887	1,7341	2,1009	2,4450	2,8784	3,1966
19	0,68763	1,1866	1,7291	2,0930	2,4334	2,8609	3,1737
20	0,68696	1,1848	1,7247	2,0860	2,4231	2,8453	3,1534
21	0,68635	1,1831	1,7207	2,0796	2,4138	2,8314	3,1352
22	0,68580	1,1816	1,7171	2,0739	2,4055	2,8188	3,1188
23	0,68531	1,1802	1,7139	2,0687	2,3979	2,8073	3,1040
24	0,68485	1,1789	1,7109	2,0639	2,3910	2,7969	3,0905
25	0,68443	1,1777	1,7081	2,0595	2,3846	2,7874	3,0782
26	0,68405	1,1766	1,7056	2,0555	2,3788	2,7787	3,0669
27	0,68370	1,1757	1,7033	2,0518	2,3734	2,7707	3,0565
28	0,68335	1,1748	1,7011	2,0484	2,3685	2,7633	3,0469
29	0,68304	1,1739	1,6991	2,0452	2,3638	2,7564	3,0380
30	0,68276	1,1731	1,6973	2,0423	2,3596	2,7500	3,0298
40	0,68066	1,1673	1,6839	2,0211	2,3289	2,7045	2,9712
60	0,67862	1,1616	1,6707	2,0003	2,2991	2,6603	2,9146
120	0,67656	1,1559	1,6577	1,9799	2,2699	2,6174	2,8599
∞	0,67449	1,1503	1,6449	1,9600	2,2414	2,5758	2,8070

Extraída do livro "Curso de Estatística"-Jairo Simon da Fonseca & Gilberto de Andrade Martins-Editora Atlas

Exemplo

Uma amostra aleatória n = 25 tem $\bar{x}$ = 50 e s = 8. Construa um IC a 95% para μ

- g.l. = n – 1 = 24, so $t_{n-1,\alpha/2} = t_{24,.025} = 2.0639$

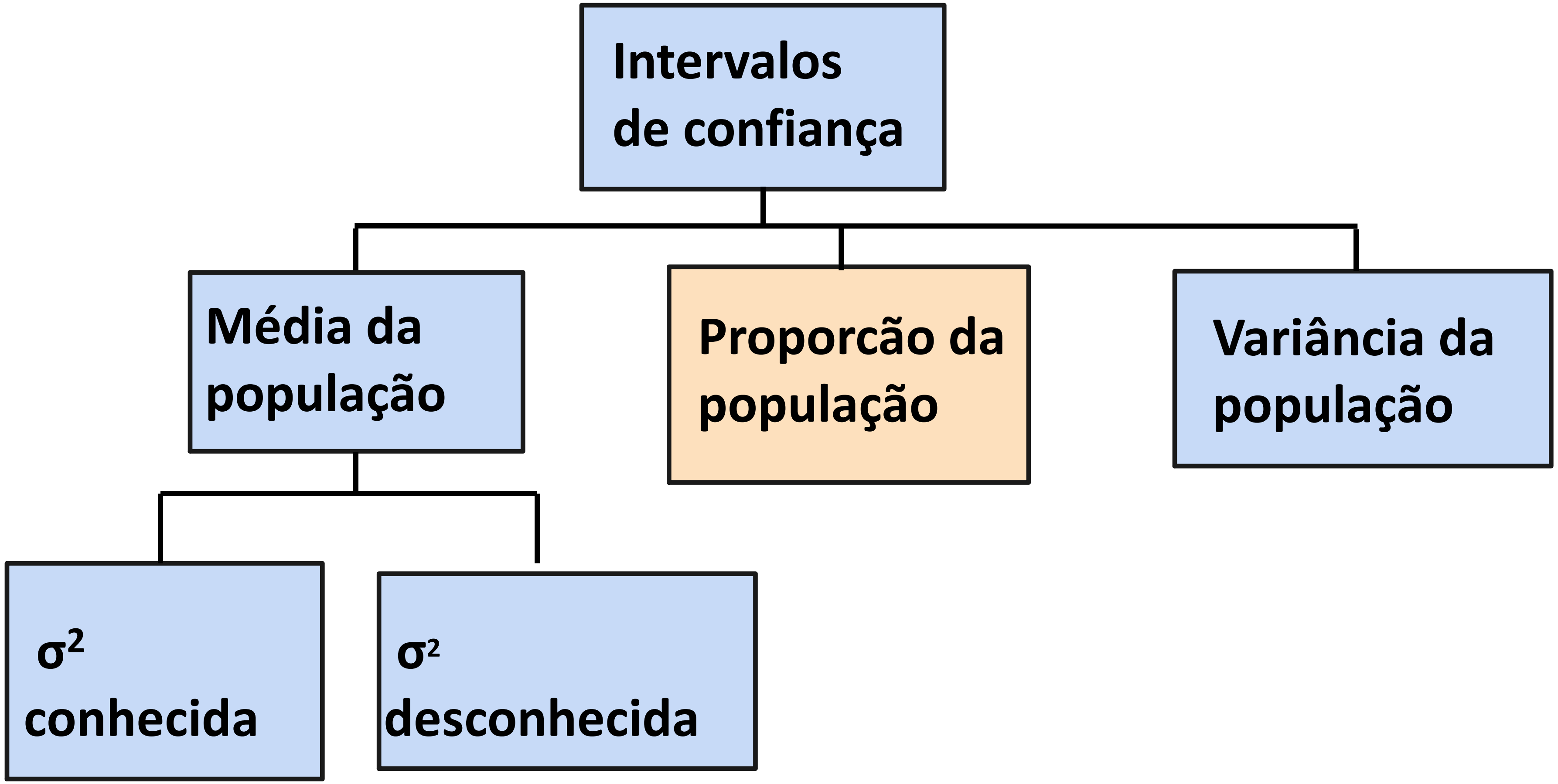
O Intervalo de Confiança

$$\bar{x} - t_{n\text{-}1,\alpha/2}\frac{S}{\sqrt{n}} < \mu < \bar{x} + t_{n\text{-}1,\alpha/2}\frac{S}{\sqrt{n}}$$

$$50 - (2.0639)\frac{8}{\sqrt{25}} < \mu < 50 + (2.0639)\frac{8}{\sqrt{25}}$$

$$\boxed{46.698 < \mu < 53.302}$$

Intervalos de confiança



Intervalo de Confiança para a Proporção Populacional

- Uma estimativa de IC para a proporção populacional (P) pode ser calculado adicionando uma tolerância para a incerteza da proporção amostral ($\hat{p}$)

Intervalo de Confiança para a Proporção Populacional, p

(cont.)

- Relembre que a distribuição da proporção da amostra é aproximadamente normal se o tamanho da amostra for grande, com desvio padrão

$$\sigma_P = \sqrt{\frac{P(1-P)}{n}}$$

- Estimaremos este IC com dados da amostra:

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Pontos finais de intervalo de confiança

- Os limites de confiança superior e inferior para a proporção da população são calculados com a fórmula

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < P < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- sendo
  - $z_{\alpha/2}$ é o valor normal padrão para o nível de confiança desejado
  - $\hat{p}$ é a proporção da amostra
  - n é o tamanho da amostra
  - $nP(1-P) > 5$

Exemplo

- Uma amostra aleatória de 100 pessoas mostra que 25 são canhotas.
- Construa um IC de 95% para a verdadeira proporção de canhotos

Exemplo

- Uma amostra aleatória de 100 pessoas mostra que 25 são canhotas. Construa um IC de 95% para a verdadeira proporção de canhotos

(cont.)

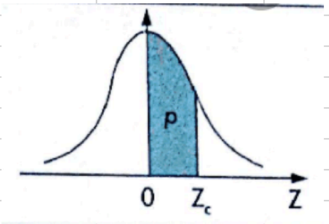
$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \;<\; P \;<\; \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\frac{25}{100} - 1.96\sqrt{\frac{.25(.75)}{100}} \;<\; P \;<\; \frac{25}{100} + 1.96\sqrt{\frac{.25(.75)}{100}}$$

$$\boxed{0.1651 \;<\; P \;<\; 0.3349}$$

z	0	1	2	3	4	5	6	7	8	9
0.0	0.000000	0.003989	0.007978	0.011966	0.015953	0.019939	0.023922	0.027903	0.031881	0.035856
0.1	0.039828	0.043795	0.047758	0.051717	0.055670	0.059618	0.063559	0.067495	0.071424	0.075345
0.2	0.079260	0.083166	0.087064	0.090954	0.094835	0.098706	0.102568	0.106420	0.110261	0.114092
0.3	0.117911	0.121720	0.125516	0.129300	0.133072	0.136831	0.140576	0.144309	0.148027	0.151732
0.4	0.155422	0.159097	0.162757	0.166402	0.170031	0.173645	0.177242	0.180822	0.184386	0.187933
0.5	0.191462	0.194974	0.198468	0.201944	0.205401	0.208840	0.212260	0.215661	0.219043	0.222405
0.6	0.225747	0.229069	0.232371	0.235653	0.238914	0.242154	0.245373	0.248571	0.251748	0.254903
0.7	0.258036	0.261148	0.264238	0.267305	0.270350	0.273373	0.276373	0.279350	0.282305	0.285236
0.8	0.288145	0.291030	0.293892	0.296731	0.299546	0.302337	0.305105	0.307850	0.310570	0.313267
0.9	0.315940	0.318589	0.321214	0.323814	0.326391	0.328944	0.331472	0.333977	0.336457	0.338913
1.0	0.341345	0.343752	0.346136	0.348495	0.350830	0.353141	0.355428	0.357690	0.359929	0.362143
1.1	0.364334	0.366500	0.368643	0.370762	0.372857	0.374928	0.376976	0.379000	0.381000	0.382977
1.2	0.384930	0.386861	0.388768	0.390651	0.392512	0.394350	0.396165	0.397958	0.399727	0.401475
1.3	0.403200	0.404902	0.406582	0.408241	0.409877	0.411492	0.413085	0.414657	0.416207	0.417736
1.4	0.419243	0.420730	0.422196	0.423641	0.425066	0.426471	0.427855	0.429219	0.430563	0.431888
1.5	0.433193	0.434478	0.435745	0.436992	0.438220	0.439429	0.440620	0.441792	0.442947	0.444083
1.6	0.445201	0.446301	0.447384	0.448449	0.449497	0.450529	0.451543	0.452540	0.453521	0.454486
1.7	0.455435	0.456367	0.457284	0.458185	0.459070	0.459941	0.460796	0.461636	0.462462	0.463273
1.8	0.464070	0.464852	0.465620	0.466375	0.467116	0.467843	0.468557	0.469258	0.469946	0.470621
1.9	0.471283	0.471933	0.472571	0.473197	0.473810	0.474412	0.475002	0.475581	0.476148	0.476705
2.0	0.477250	0.477784	0.478308	0.478822	0.479325	0.479818	0.480301	0.480774	0.481237	0.481691
2.1	0.482136	0.482571	0.482997	0.483414	0.483823	0.484222	0.484614	0.484997	0.485371	0.485738
2.2	0.486097	0.486447	0.486791	0.487126	0.487455	0.487776	0.488089	0.488396	0.488696	0.488989
2.3	0.489276	0.489556	0.489830	0.490097	0.490358	0.490613	0.490863	0.491106	0.491344	0.491576
2.4	0.491802	0.492024	0.492240	0.492451	0.492656	0.492857	0.493053	0.493244	0.493431	0.493613
2.5	0.493790	0.493963	0.494132	0.494297	0.494457	0.494614	0.494766	0.494915	0.495060	0.495201
2.6	0.495339	0.495473	0.495604	0.495731	0.495855	0.495975	0.496093	0.496207	0.496319	0.496427
2.7	0.496533	0.496636	0.496736	0.496833	0.496928	0.497020	0.497110	0.497197	0.497282	0.497365
2.8	0.497445	0.497523	0.497599	0.497673	0.497744	0.497814	0.497882	0.497948	0.498012	0.498074
2.9	0.498134	0.498193	0.498250	0.498305	0.498359	0.498411	0.498462	0.498511	0.498559	0.498605
3.0	0.498650	0.498694	0.498736	0.498777	0.498817	0.498856	0.498893	0.498930	0.498965	0.498999
3.1	0.499032	0.499065	0.499096	0.499126	0.499155	0.499184	0.499211	0.499238	0.499264	0.499289
3.2	0.499313	0.499336	0.499359	0.499381	0.499402	0.499423	0.499443	0.499462	0.499481	0.499499
3.3	0.499517	0.499534	0.499550	0.499566	0.499581	0.499596	0.499610	0.499624	0.499638	0.499651
3.4	0.499663	0.499675	0.499687	0.499698	0.499709	0.499720	0.499730	0.499740	0.499749	0.499758
3.5	0.499767	0.499776	0.499784	0.499792	0.499800	0.499807	0.499815	0.499822	0.499828	0.499835
3.6	0.499841	0.499847	0.499853	0.499858	0.499864	0.499869	0.499874	0.499879	0.499883	0.499888
3.7	0.499892	0.499896	0.499900	0.499904	0.499908	0.499912	0.499915	0.499918	0.499922	0.499925
3.8	0.499928	0.499931	0.499933	0.499936	0.499938	0.499941	0.499943	0.499946	0.499948	0.499950
3.9	0.499952	0.499954	0.499956	0.499958	0.499959	0.499961	0.499963	0.499964	0.499966	0.499967
4.0	0.499968	0.499970	0.499971	0.499972	0.499973	0.499974	0.499975	0.499976	0.499977	0.499978
4.1	0.499979	0.499980	0.499981	0.499982	0.499983	0.499983	0.499984	0.499985	0.499985	0.499986
4.2	0.499987	0.499987	0.499988	0.499988	0.499989	0.499989	0.499990	0.499990	0.499991	0.499991
4.3	0.499991	0.499992	0.499992	0.499993	0.499993	0.499993	0.499993	0.499994	0.499994	0.499994
4.4	0.499995	0.499995	0.499995	0.499995	0.499996	0.499996	0.499996	0.499996	0.499996	0.499996

Parte Inteira do Primeiro Decimal de Z



0.475

Interpretação

- Estamos 95% confiantes que a verdadeira percentagem de canhotos na população está entre

  16.51% e 33.49%.

- Embora o intervalo de 0,1651 a 0,3349 possa conter ou não a proporção verdadeira, 95% dos intervalos construídos a partir de amostras de tamanho 100 conterão a proporção verdadeira.