

Estadística Aplicada II

▶ Estimación e Intervalos de Confiança

Aula de hoje

▶ Tópicos

- ▶ Revisão
- ▶ Estimação
- ▶ Intervalos de Confiança

▶ Referências

- ▶ Barrow, M. Estatística para economia, contabilidade e administração. São Paulo: Ática, 2007, Cap. 4
- ▶ Morettin, P. e W. Bussab. Estatística básica. 5. ed. São Paulo: Saraiva, 2005. Cap. I I

Objetivo

- ▶ **Objetivo:** Inferir os parâmetros populacionais a partir da informação amostral
- ▶ **Por exemplo, estimar a média μ :**
- ▶ Da renda domiciliar em uma localidade
- ▶ Dos gastos domiciliares com educação no estado de São Paulo
- ▶ Da altura dos indivíduos no país



Revisão

Média amostral

- ▶ Teorema: A **média amostral**, extraída de uma população normalmente distribuída $X \sim \mathbf{N}(\mu ; \sigma^2)$, é uma variável aleatória com **distribuição normal**, com **média** μ e **variância** σ^2 / n

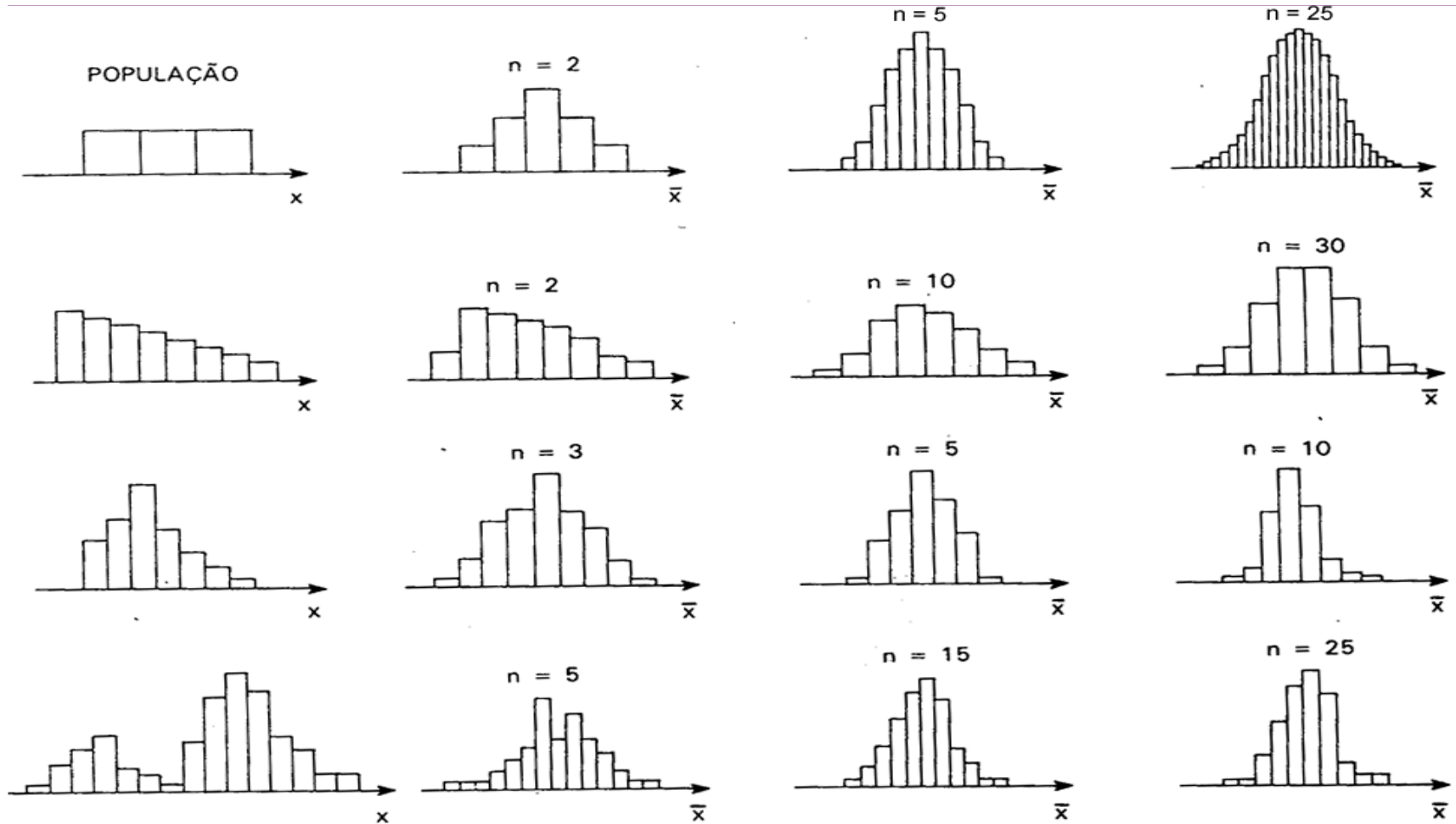
Média amostral

- ▶ Suponha que a altura média tenha uma distribuição Normal, com média 167 cm e desvio-padrão 9cm
- ▶ Pergunta: Qual é a probabilidade de que uma amostra de 9 indivíduos obtida ao acaso tenha uma altura média superior a 170cm?

Média amostral

- ▶ **Teorema do Limite Central:** A média amostral, extraída de uma população com **média** μ e **variância** σ^2 , tem uma distribuição que se aproxima de uma distribuição normal com **média** μ e **variância** σ^2 / n , à medida que o tamanho da amostra tende ao infinito.

Exemplo Bussab & Morettin



Fonte: BUSSAB & MORETTIN. *Estatística Básica*. São Paulo, Atual, 4ª edição, 2002, pp. 273.

Média amostral

- ▶ Suponha que a renda média de cada brasileiro seja de R\$1113, com variância de 44.100 reais. Suponha escolhessemos uma amostra de 100 pessoas ao acaso. Qual é a probabilidade de que a renda média na amostra seja maior do que R\$1150?



Estimação

Estimação Pontual e em Intervalos

- ▶ Há dois modos de apresentação da estimativa de um parâmetro:
- ▶ **Estimativa Pontual:** um único valor correspondente a estimativa do parâmetro em questão
- ▶ **Estimativa em Intervalos:** A estimativa do parâmetro pode assumir qualquer valor dentro de um intervalo de valores possíveis.

Regras e Critérios para a Estimação

- ▶ Algumas definições importantes:
- ▶ **Estimador:** conjunto de regras que descrevem o procedimento para a obtenção da estimativa
- ▶ **Estimativa:** valor obtido a partir de um estimador para inferência sobre o valor do parâmetro populacional
- ▶ Exemplos de estimadores para a média da população
 - ▶ Média de uma amostra
 - ▶ Menor observação na amostra
- ▶ Pergunta: Qual estimador você recomendaria?

Regras e Critérios para a Estimação

- ▶ Sendo que o valor do parâmetro populacional é desconhecido, é impossível checar se uma estimativa é correta ou não.
- ▶ Critérios para julgamento de estimadores:
- ▶ **1º Critério: Viés**
 - ▶ Um estimador é não-enviesado quando fornece a resposta correta em média
 - ▶ Em outras palavras, um estimador é não-enviesado quando o seu valor esperado é equivalente ao parâmetro populacional de interesse

Regras e Critérios para a Estimação

- ▶ **Definição de viés:** diferença do valor esperado para o estimador e o valor do parâmetro
- ▶ **Exemplo:** Viés do estimador média amostral

$$\begin{aligned}\text{Viés} &= E(\text{média amostral}) - \mu \\ &= 0\end{aligned}$$



Regras e Critérios para a Estimação

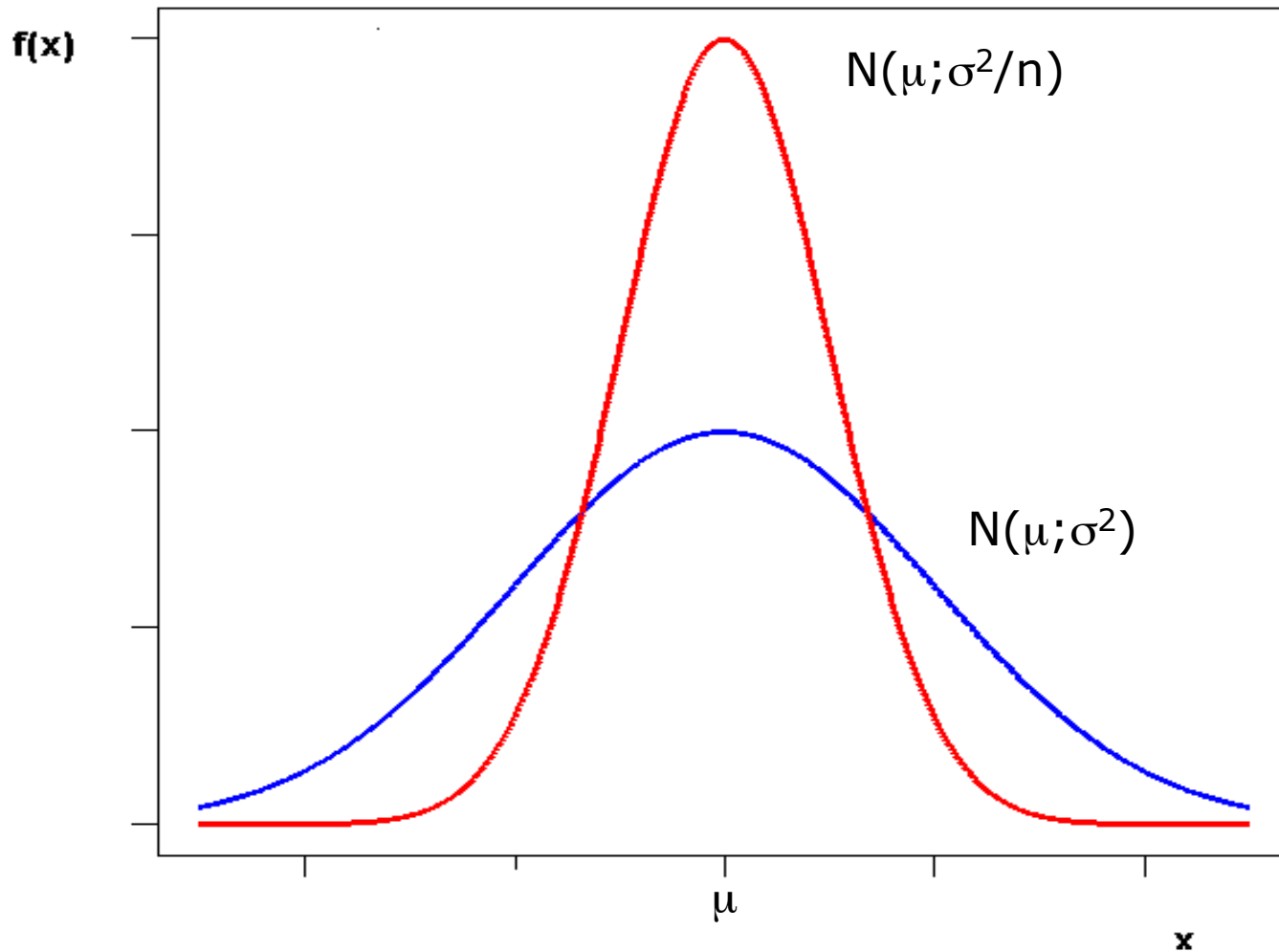
▶ 2º Critério: Precisão

- ▶ Diferentemente do viés, a precisão é um conceito relativo pois envolve a comparação entre estimadores
- ▶ Um estimador A é mais preciso do que um estimador B quando a variância de A for menor do que a variância de B. Em outras palavras, quando a dispersão das estimativas produzida por A é menor do que a dispersão produzida por B.
- ▶ Exemplo: Na aula passada, vimos que a distribuição da média amostral, sob determinados pressupostos, é dada por:

$$\mathbf{N}(\mu ; \sigma^2/n)$$

- ▶ Portanto, quanto maior o tamanho da amostra, maior a precisão desse estimador

Média amostral



Comparação entre viés e precisão

- ▶ Suponha que você possua um estimador A enviesado, porém mais preciso do que o estimador B não-enviesado.
- ▶ **Pergunta:** Qual você escolheria?

Estimação por intervalos

- ▶ **Suponha que estamos trabalhando com amostras grandes**
- ▶ **Nesse caso, vimos na aula passada, que a média amostral, extraída de uma população com média μ e variância σ^2 , tem uma distribuição que se aproxima de uma distribuição normal com média μ e variância σ^2/n**

Estimação por intervalos

- ▶ Suponha que não soubéssemos a renda média da população brasileira e escolhessemos uma amostra de 100 pessoas, com renda média amostral de R\$1150 e variância amostral de 44.100.
- ▶ **Pergunta: Qual seria uma estimativa pontual da renda média da população brasileira?**

Estimação por intervalos

- ▶ **Problema:** Essa estimativa não nos dá uma idéia da incerteza associada a essa estimativa, o que justificaria a obtenção de uma estimativa por intervalo.
- ▶ **Desafio:** **Obtenha uma estimativa por intervalo da renda média da população brasileira.**

Estimação por intervalos

- ▶ **Nota:** Se o valor da variância populacional σ^2 for desconhecido, s^2 pode ser utilizado na construção do intervalo de confiança **apenas em amostras grandes**. Nesse caso, s^2 é uma estimativa suficientemente precisa de σ^2 .

Estimação por intervalos

- ▶ Quanto maior a incerteza associada a uma estimativa pontual, mais largo é o intervalo de confiança
- ▶ Essa incerteza será maior, quanto:
 - ▶ **1. Menor for o tamanho da amostra**
 - ▶ **2. Maior for o desvio-padrão**

Estimação por intervalos

- ▶ Quanto maior o nível de confiança exigido, mais largo é o intervalo de confiança
- ▶ Por convenção, os níveis de confiança de 95% e 99% são os mais utilizados na literatura.

Estimação da diferença entre duas médias

- ▶ Problema: Temos agora duas amostras e queremos saber se há diferença entre suas respectivas populações
- ▶ Exemplo: Sessenta alunos da escola 1 obtiveram nota média de 6,2, com um desvio-padrão de 1,8; 35 alunos da escola 2 obtiveram nota média de 7,0, com um desvio-padrão de 1,2. Qual é a verdadeira diferença entre as notas médias obtidas por todos os alunos nas duas escolas?

Revisão

▶ **Vimos nas aulas passadas que:**

I. Se a população tiver distribuição normal, a média amostral também terá.

Nesse caso, os intervalos de confiança podem ser construídos com base em:

$$\mathbf{Média\ Amostrал\ \sim\ N(\mu ; \sigma^2/n)}$$

Revisão

▶ Vimos nas aulas passadas que:

2. Independentemente da distribuição populacional se a amostra for suficientemente grande, a média amostral terá distribuição normal.

Nesse caso, os intervalos de confiança podem ser construídos com base em:

***Média Amostral* $\sim N(\mu ; \sigma^2/n)$ ou**

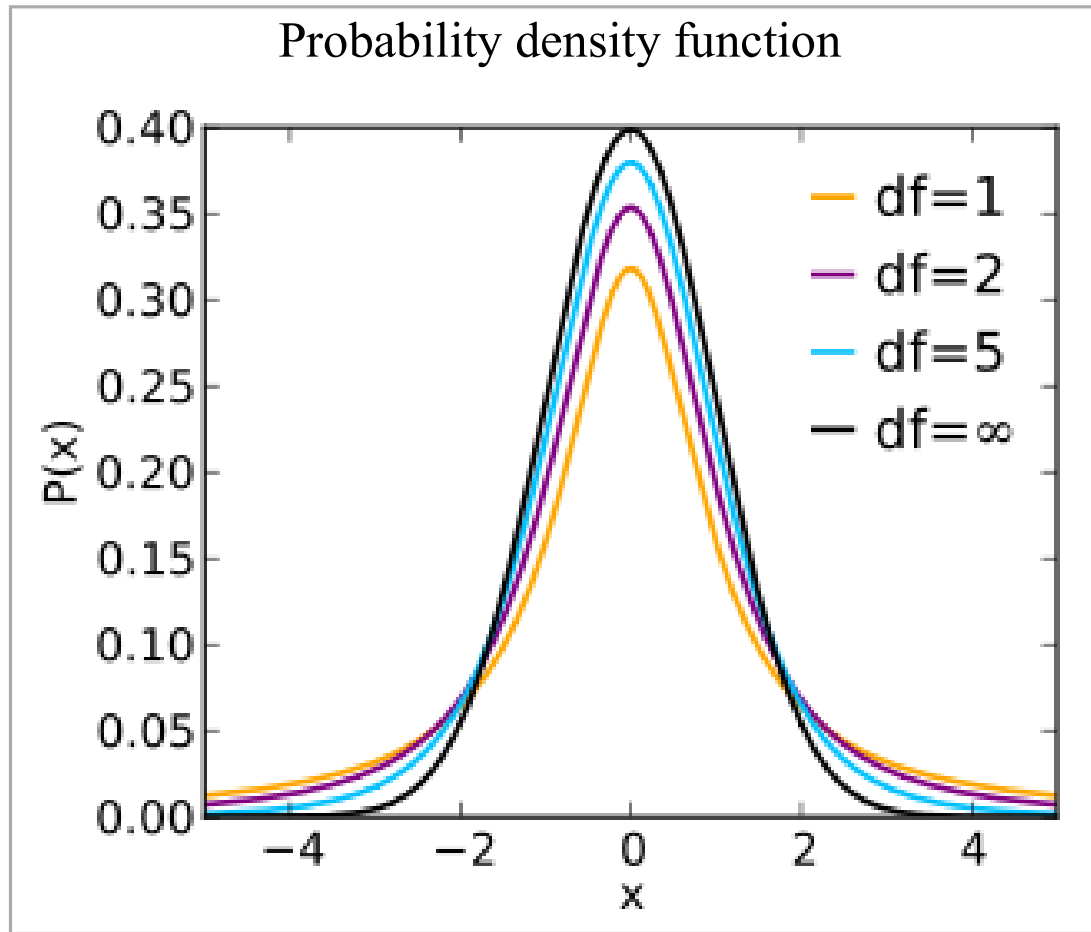
***Média Amostral* $\sim N(\mu ; s^2/n)$**

Estimação com amostras pequenas

- ▶ **A distribuição t deve ser utilizada quando:**
 - ▶ O tamanho da amostra é pequeno; e
 - ▶ A distribuição da população de origem é normal; e
 - ▶ A variância da amostra é utilizada para se estimar a variância da população

Estimação com amostras pequenas

- Devemos utilizar a distribuição t

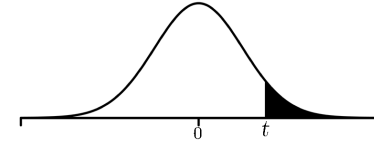


Estimação com amostras pequenas

- ▶ **Características da distribuição t:**
 - ▶ É caracterizada por apenas um parâmetro: o número de graus de liberdade, equivalentes ao tamanho da amostra menos 1 ($n-1$)
 - ▶ É simétrica
 - ▶ É centrada no valor 0
 - ▶ Tem forma de sino
 - ▶ Estende-se de $-$ infinito a $+$ infinito

Estimação com amostras pequenas

Critical Values for Student's t -Distribution.



df	Upper Tail Probability: $\Pr(T > t)$									
	0.2	0.1	0.05	0.04	0.03	0.025	0.02	0.01	0.005	0.0005
1	1.376	3.078	6.314	7.916	10.579	12.706	15.895	31.821	63.657	636.619
2	1.061	1.886	2.920	3.320	3.896	4.303	4.849	6.965	9.925	31.599
3	0.978	1.638	2.353	2.605	2.951	3.182	3.482	4.541	5.841	12.924
4	0.941	1.533	2.132	2.333	2.601	2.776	2.999	3.747	4.604	8.610
5	0.920	1.476	2.015	2.191	2.422	2.571	2.757	3.365	4.032	6.869
6	0.906	1.440	1.943	2.104	2.313	2.447	2.612	3.143	3.707	5.959
7	0.896	1.415	1.895	2.046	2.241	2.365	2.517	2.998	3.499	5.408
8	0.889	1.397	1.860	2.004	2.189	2.306	2.449	2.896	3.355	5.041
9	0.883	1.383	1.833	1.973	2.150	2.262	2.398	2.821	3.250	4.781
10	0.879	1.372	1.812	1.948	2.120	2.228	2.359	2.764	3.169	4.587
11	0.876	1.363	1.796	1.928	2.096	2.201	2.328	2.718	3.106	4.437
12	0.873	1.356	1.782	1.912	2.076	2.179	2.303	2.681	3.055	4.318
13	0.870	1.350	1.771	1.899	2.060	2.160	2.282	2.650	3.012	4.221
14	0.868	1.345	1.761	1.887	2.046	2.145	2.264	2.624	2.977	4.140
15	0.866	1.341	1.753	1.878	2.034	2.131	2.249	2.602	2.947	4.073
16	0.865	1.337	1.746	1.869	2.024	2.120	2.235	2.583	2.921	4.015
17	0.863	1.333	1.740	1.862	2.015	2.110	2.224	2.567	2.898	3.965
18	0.862	1.330	1.734	1.855	2.007	2.101	2.214	2.552	2.878	3.922
19	0.861	1.328	1.729	1.850	2.000	2.093	2.205	2.539	2.861	3.883
20	0.860	1.325	1.725	1.844	1.994	2.086	2.197	2.528	2.845	3.850

21	0.859	1.323	1.721	1.840	1.988	2.080	2.189	2.518	2.831	3.819
22	0.858	1.321	1.717	1.835	1.983	2.074	2.183	2.508	2.819	3.792
23	0.858	1.319	1.714	1.832	1.978	2.069	2.177	2.500	2.807	3.768
24	0.857	1.318	1.711	1.828	1.974	2.064	2.172	2.492	2.797	3.745
25	0.856	1.316	1.708	1.825	1.970	2.060	2.167	2.485	2.787	3.725
26	0.856	1.315	1.706	1.822	1.967	2.056	2.162	2.479	2.779	3.707
27	0.855	1.314	1.703	1.819	1.963	2.052	2.158	2.473	2.771	3.690
28	0.855	1.313	1.701	1.817	1.960	2.048	2.154	2.467	2.763	3.674
29	0.854	1.311	1.699	1.814	1.957	2.045	2.150	2.462	2.756	3.659
30	0.854	1.310	1.697	1.812	1.955	2.042	2.147	2.457	2.750	3.646
31	0.853	1.309	1.696	1.810	1.952	2.040	2.144	2.453	2.744	3.633
32	0.853	1.309	1.694	1.808	1.950	2.037	2.141	2.449	2.738	3.622
33	0.853	1.308	1.692	1.806	1.948	2.035	2.138	2.445	2.733	3.611
34	0.852	1.307	1.691	1.805	1.946	2.032	2.136	2.441	2.728	3.601
35	0.852	1.306	1.690	1.803	1.944	2.030	2.133	2.438	2.724	3.591
36	0.852	1.306	1.688	1.802	1.942	2.028	2.131	2.434	2.719	3.582
37	0.851	1.305	1.687	1.800	1.940	2.026	2.129	2.431	2.715	3.574
38	0.851	1.304	1.686	1.799	1.939	2.024	2.127	2.429	2.712	3.566
39	0.851	1.304	1.685	1.798	1.937	2.023	2.125	2.426	2.708	3.558
40	0.851	1.303	1.684	1.796	1.936	2.021	2.123	2.423	2.704	3.551
41	0.850	1.303	1.683	1.795	1.934	2.020	2.121	2.421	2.701	3.544
42	0.850	1.302	1.682	1.794	1.933	2.018	2.120	2.418	2.698	3.538
43	0.850	1.302	1.681	1.793	1.932	2.017	2.118	2.416	2.695	3.532
44	0.850	1.301	1.680	1.792	1.931	2.015	2.116	2.414	2.692	3.526
45	0.850	1.301	1.679	1.791	1.929	2.014	2.115	2.412	2.690	3.520
46	0.850	1.300	1.679	1.790	1.928	2.013	2.114	2.410	2.687	3.515
47	0.849	1.300	1.678	1.789	1.927	2.012	2.112	2.408	2.685	3.510
48	0.849	1.299	1.677	1.789	1.926	2.011	2.111	2.407	2.682	3.505
49	0.849	1.299	1.677	1.788	1.925	2.010	2.110	2.405	2.680	3.500
50	0.849	1.299	1.676	1.787	1.924	2.009	2.109	2.403	2.678	3.496
60	0.848	1.296	1.671	1.781	1.917	2.000	2.099	2.390	2.660	3.460
70	0.847	1.294	1.667	1.776	1.912	1.994	2.093	2.381	2.648	3.435
80	0.846	1.292	1.664	1.773	1.908	1.990	2.088	2.374	2.639	3.416
90	0.846	1.291	1.662	1.771	1.905	1.987	2.084	2.368	2.632	3.402
100	0.845	1.290	1.660	1.769	1.902	1.984	2.081	2.364	2.626	3.390
120	0.845	1.289	1.658	1.766	1.899	1.980	2.076	2.358	2.617	3.373
140	0.844	1.288	1.656	1.763	1.896	1.977	2.073	2.353	2.611	3.361
180	0.844	1.286	1.653	1.761	1.893	1.973	2.069	2.347	2.603	3.345
200	0.843	1.286	1.653	1.760	1.892	1.972	2.067	2.345	2.601	3.340
500	0.842	1.283	1.648	1.754	1.885	1.965	2.059	2.334	2.586	3.310
1000	0.842	1.282	1.646	1.752	1.883	1.962	2.056	2.330	2.581	3.300
∞	0.842	1.282	1.645	1.751	1.881	1.960	2.054	2.326	2.576	3.291

Distribuição Normal : Valores de $P(Z \leq z) = A(z)$

Segunda decimal de z

	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Parte inteira e primeira decimal de z

Estimação com amostras pequenas

- ▶ Suponha que não soubéssemos a altura média da população brasileira, distribuída normalmente, e escolhessemos uma amostra de 10 pessoas, com altura média de 168 cm e desvio-padrão de 9cm
- ▶ Nesse caso, como poderíamos obter o intervalo de confiança de 95% para a altura média?

Estimação com amostras pequenas

- ▶ Suponha que temos duas amostras de duas populações normalmente distribuídas, a primeira com idade média igual a 45, $n_1=15$, $s_1=18$ e a segunda com idade média igual a 55, $n_2=20$, $s_2=21$. Nesse caso, como poderíamos obter o intervalo de confiança para a diferença das idades médias das duas populações?