Roteiro de aulas - Estimação

Disciplina: *Probabilidade e Estatística*

Professor: André Oliveira

21 de outubro de 2025

Estimação

Um parâmetro é uma constante que descreve uma característica de uma população, em geral é desconhecido. Quando é de interesse **estimar** um determinado **parâmetro** (média, variância, proporção etc) a partir de uma amostra há duas **metodologias**.

- Pode-se fornecer uma **estimativa pontual**. No caso de uma média populacional, a estimativa pontual é a média amostral. Consiste em atribuir um único valor como estimativa do parâmetro. Por exemplo, para estimar a média populacional (μ) utiliza-se a média amostral ($\widehat{\mu}$) como estimativa pontual.
- Pode-se fornecer uma **estimativa intervalar**. No caso da média populacional é estimada pela média amostral e adiciona-se uma **margem de erro** à estimativa. Todo levantamento amostral está sujeito ao erro. Em vez de apresentar apenas um valor, fornece-se um intervalo de confiança, que indica uma faixa de valores dentro da qual o parâmetro populacional deve estar, com determinado nível de confiança. No caso da média populacional, esse intervalo é construído a partir da média amostral somada e subtraída de uma **margem de erro**.

Uma **estimativa** é o valor de um parâmetro populacional desconhecido obtido a partir de uma amostra. Estimativa é uma **constante** inerente à amostra e representa o valor assumido por uma estatística em uma amostra específica. Um **estimador** (ou estatística) é a expressão **algébrica** que permite obter uma estimativa ou a variável aleatória que é usada no processo de estimação é função dos elementos amostrais.

- ***O estimador da média populacional*** é $\widehat{\mu}_x = \dfrac{\sum\limits_{i=1}^{n} X_i}{n}$
- ***O estimador da variância populacional*** é $\widehat{\sigma}_x^2 = \dfrac{\sum\limits_{i=1}^{n} (X_i - \widehat{\mu}_x)^2}{n-1}$

Estimação por ponto

O procedimento para estimar a média populacional varia conforme o tipo de amostragem utilizado. Quando a amostragem é aleatória simples. Tanto para populações finitas como para populações infinitas o parâmetro (μ_x) pode ser estimado pela média amostral, ou seja, pelo estimador.

$$\widehat{\mu}_x = \frac{\sum_{i=1}^{n} X_i}{n}$$

O procedimento para estimar a variância populacional se a população for finita ou infinita

$$\widehat{\sigma}^2_x = \frac{\sum_{i=1}^{n} (X_i - \widehat{\mu}_x)^2}{n-1}$$

Exemplo

Um empresário deseja estudar (investigar) qual é o comprimento **médio** e a **variabilidade** dos cabos de rede produzidos por sua empresa. O padrão especificado é de 8 metros. Para isso, ele coletou uma amostra de 10 cabos e mediu o comprimento de cada um como descrito abaixo.

8.11, 7.82, 7.76, 8.21, 7.91, 8.26, 7.74, 7.99, 8.05, 8.13

A estimativa **pontual** para a média populacional é:

$$\widehat{\mu}_x = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{8,11 + 7,82 + 7,76 + 8,21 + 7,91 + 8,26 + 7,74 + 7,99 + 8,05 + 8,13}{10}$$

$\widehat{\mu}_x = \frac{79{,}98}{10} = 7{,}998$ metros $\leftarrow$ Estimativa pontual da média

A estimativa **pontual** para a variância populacional é:

$\widehat{\sigma}^2_x = \frac{\sum_{i=1}^{n} (X_i - \widehat{\mu}_x)^2}{n-1} = \frac{(8{,}11 - 7{,}998)^2 + (7{,}82 - 7{,}998)^2 + \cdots + (8{,}13 - 7{,}998)^2}{10-1} =$ $0{,}034\ metros^2$ $\leftarrow$ Estimativa pontual da variância

Assim a estimativa pontual da média e da variância populacional são $\widehat{\mu}_x = 7{,}998$ metros e $\widehat{\sigma}^2_x = 0{,}034\ metros^2$

Como a estimativa **pontual** da média populacional foi obtida a partir de uma amostra ($n = 10$) o erro padrão ($SE_{\widehat{\mu}}$) da associado a estimativa da média é:

$SE_{\widehat{\mu}} = \frac{\widehat{\sigma}_x}{\sqrt{n}} = \frac{0,185}{\sqrt{10}} = 0{,}059$ metros $\leftarrow$ Erro padrão associado à média

Exemplo

Uma pesquisa foi conduzida com o objetivo de investigar a proporção de empresas que utilizam a linguagem *Python* para visualizar e processar dados. Para isso, uma amostra de 40 empresas foi selecionada, e os resultados estão descritos a seguir.

Python, Python, Python, Outra linguagem, Outra linguagem, Outra linguagem, Python, Outra linguagem, Outra linguagem, Python, Python, Python, Outra linguagem, Outra linguagem, Python, Python, Python, Python, Python, Python, Python, Python, Python, Python, Python, Python, Python, Python, Outra linguagem, Outra linguagem, Python, Python, Python, Outra linguagem, Python, Outra linguagem, Python, Python, Python, Outra linguagem

A estimativa **pontual** da proporção de empresas que usam *Python* é total observado de *Python* entre as 40 observações:

$\widehat{\rho} = \dfrac{28}{40} = 0{,}7 = 70{,}00\,\% \leftarrow$ Estimativa pontual da proporção de empresas que usam *Python*

Como a estimativa pontual da proporção populacional de empresas que usam *Python* foi obtida a partir de uma amostra ($n = 40$) tem um erro padrão ($SE_{\widehat{\rho}}$) associado a esta estimativa da proporção de:

$SE_{\widehat{p}} = \sqrt{\dfrac{\widehat{p}.(1-\widehat{p})}{n}} = \sqrt{\dfrac{0,700 \times (1-0,700)}{40}} = 0,0725 = 7,25\% \leftarrow$ Erro padrão associado à estimativa

Estimação intervalar - Média

A característica de interesse da população é chamada de parâmetro e a característica da amostra correspondente é a estatística da amostra ou parâmetro de estimativa. Como a estatística é um resumo das informações sobre um parâmetro obtido a partir da amostra, o valor de uma estatística depende da amostra particular que foi extraída da população. Os seus valores mudam aleatoriamente a partir de uma amostra aleatória para a seguinte, por conseguinte, uma estatística é uma quantidade aleatória (variável). A distribuição de probabilidade desta variável aleatória é chamada distribuição amostral. A distribuição amostral de uma (amostra) estatística é importante porque nos permite tirar conclusões (inferências) sobre o parâmetro de população correspondente com base em uma amostra aleatória.

Para obter uma estimativa intervalar para a média (μ) de uma população utilizamos informações tais como a média amostral, o *desvio-padrão* populacional ou amostral, e algumas informações referentes às distribuições Normal e *t de Student*. Considerando ***populações infinitas***, $\dfrac{n}{N} < 0,05$, existem duas metodologias estatísticas para a construção do intervalo de confiança para média.

Quando a amostra observada segue distribuição normal e se conhece o valor do *desvio-padrão* populacional (σ), o intervalo de confiança (IC) com confiança de $\gamma = 100.(1-\alpha)$ é dado por:

$$IC(\mu;\gamma):\left[\widehat{\mu}-z_{\frac{\alpha}{2}}.\left(\frac{\sigma}{\sqrt{n}}\right);\;\;\widehat{\mu}+z_{\frac{\alpha}{2}}.\left(\frac{\sigma}{\sqrt{n}}\right)\right],\sigma^2\ conhecida$$

Se os dados observados têm distribuição normal, mas a variância populacional não é conhecida, devemos utilizar a variância amostral para estimar a variância populacional (ver variância e *desvio-padrão*) e o intervalo de com confiança de $\gamma = 100.(1-\alpha)$ é dado por:

$$IC(\mu;\gamma):\left[\widehat{\mu}-t_{\left(\frac{\alpha}{2},\,\nu\right)}.\left(\frac{\widehat{\sigma}}{\sqrt{n}}\right);\;\;\widehat{\mu}+t_{\left(\frac{\alpha}{2},\,\nu\right)}.\left(\frac{\widehat{\sigma}}{\sqrt{n}}\right)\right],\sigma^2\ desconhecida$$

Em que ν é o grau de liberdade e α é o nível de significância adotado.

Exemplo

Um produtor deseja estudar (investigar) qual é o ***BRIX médio*** da laranja produzida em sua propriedade. Para isso, ele coletou uma amostra de 20 laranjas e mediu o *BRIX* com auxílio de um refratômetro como descrito abaixo. Construa o $IC(\mu;95\%)$ considerando um nível de significância $\alpha = 5\%$.

11.99, 12.35, 11.89, 12.28, 11.76, 12.3, 11.9, 12.12, 11.56, 12.04, 12.04, 12.13, 11.88, 12.05, 12.14, 11.98, 12.55, 12.23, 12.09, 12.1

- A estimativa **pontual** é $\widehat{\mu}_x = \dfrac{\sum_{i=1}^{n} X_i}{n} = \dfrac{241.38}{20}$

$\widehat{\mu}_x = 12.069\ BRIX \leftarrow$ Estimativa pontual da média populacional.

- A estimativa **pontual** da variância é $\widehat{\sigma}^2_x = \dfrac{\sum_{i=1}^{n}(X_i-\widehat{\mu}_x)^2}{n-1}$

$$\widehat{\sigma}^2_x = \frac{(11.99-12.069)^2+(12.35-12.069)^2+\cdots+(12.1-12.069)^2}{20-1}$$

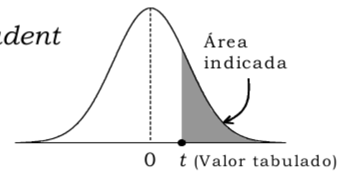
$\widehat{\sigma}^2_x = 0.048\ Brix^2$, $\widehat{\sigma}_x = 0.219\ Brix$,

$$IC(\mu;\gamma):\left[\widehat{\mu}-t_{\left(\frac{\alpha}{2},\,19\right)}.\left(\frac{\widehat{\sigma}}{\sqrt{n}}\right);\;\;\widehat{\mu}+t_{\left(\frac{\alpha}{2},\,19\right)}.\left(\frac{\widehat{\sigma}}{\sqrt{n}}\right)\right],\sigma^2\ desconhecida\ (foi\ estimada\ na\ amostra)$$

$$IC(\mu;95\%):\left[12.069-t_{\left(\frac{5,00\%}{2},\,19\right)}.\left(\frac{0.219}{\sqrt{20}}\right);\;\;12.069+t_{\left(\frac{5,00\%}{2},\,19\right)}.\left(\frac{0.219}{\sqrt{20}}\right)\right]$$

- $t_{\left(\frac{5,00\%}{2},\,19\right)} = 2,093 \rightarrow$ Consulta na tabela $t-Student$ **Figura 1**
- $-t_{\left(\frac{2,50\%}{2},\,19\right)} = -2,093 \rightarrow$ Consulta na tabela $t-Student$ **Figura 1**

Tabela 5 Distribuição ***t*** de *Student*



	Área na cauda superior								
gl	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
1	1,000	3,078	6,314	12,71	31,82	63,66	127,3	318,3	636,6
2	0,816	1,886	2,920	4,303	6,965	9,925	14,09	22,33	31,60
3	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,21	12,92
4	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,894	6,869
6	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	0,700	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	0,697	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	0,695	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	0,694	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	0,692	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	0,691	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	0,690	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	0,689	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	0,688	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	0,688	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	0,687	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	0,686	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	0,686	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792

Figura 1: Tabela t-student

$$IC(\mu; 95\%) : \left[12.069 - 2,093.\left(\frac{0.219}{\sqrt{20}}\right); \quad 12.069 + 2,093.\left(\frac{0.219}{\sqrt{20}}\right)\right]$$

$IC(\mu; 95\%) : [11.97; \quad 12.17] \leftarrow$ Estimativa intervalar para média populacional.

Estimação intervalar - Proporção

Quando há interesse de estimar a proporção de elementos de uma população que possui certo atributo (categorias). Como exemplo pode citar a proporção de técnicos que comprariam lançamento de um celular, proporção de pessoas com diabetes em um distrito. Uma proporção é estimada por $\widehat{\rho} = \dfrac{k}{n}$ onde k é o número de indivíduos com a característica de interesse e n é o tamanho da amostra selecionada. Estimação da proporção populacional de uma ***população infinita.***

$$IC(p, \gamma) : \left[\widehat{p} - Z_{\frac{\alpha}{2}}.\sqrt{\frac{p.(1-p)}{n}}; \widehat{p} + Z_{\frac{\alpha}{2}}.\sqrt{\frac{p.(1-p)}{n}}\right]$$

Exemplo

Uma pesquisa foi conduzida com o objetivo de investigar a proporção de empresas que utilizam o $Windows$ como sistema padrão. Para isso, uma amostra de 50 empresas foi selecionada, e os resultados estão descritos a seguir.

Outro Sistema, Windows, Windows, Outro Sistema, Windows, Windows, Windows, Outro Sistema, Windows, Windows, Windows, Windows, Win-

dows, Windows, Windows, Windows, Windows, Windows, Windows, Windows, Windows, Windows, Windows, Windows, Windows, Windows, Windows, Windows, Outro Sistema, Windows, Windows, Windows, Windows, Windows, Outro Sistema, Windows, Windows, Windows, Windows, Windows, Outro Sistema, Windows, Outro Sistema, Windows, Windows, Windows, Windows, Outro Sistema, Windows, Outro Sistema

Considere $\alpha = 5\%$ e construa o $IC(\mu; 95\%)$

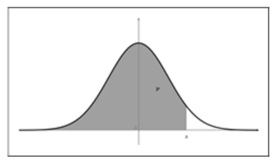
A estimativa pontual é $\widehat{p} = \dfrac{41}{50} = 0,82 = 82,00\ \%$

$$IC(p,\gamma):\left[\widehat{p} - Z_{\frac{\alpha}{2}}.\sqrt{\frac{p.(1-p)}{n}}; \widehat{p} + Z_{\frac{\alpha}{2}}.\sqrt{\frac{p.(1-p)}{n}}\right]$$

$$IC(p,\ 95\%):\left[0{,}82 - 1,96.\sqrt{\frac{0{,}82.(1-0{,}82)}{50}};\quad 0{,}82 + 1,96.\sqrt{\frac{0{,}82.(1-0{,}82)}{50}}\right]$$

$IC(p,\ 95\%) : [0,71;\quad 0,93] \leftarrow$ Estimativa intervalar para proporção.

Tabela da distribuição acumulada da normal padrão
$p = P(Z \leq z)$



Casa inteira e 1a.decimal	2a. casa decimal 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

Figura 2: Tabela Normal padrão

Lista de exercícios

1) Um empresário deseja investigar o atraso **médio** e a **variabilidade** dos tempos de seus colaboradores. O padrão tolerado é de 10 minutos. Para isso, ele coletou uma amostra de 15 tempos como descrito abaixo.

9,7, 10,32, 9,81, 9,82, 9,6, 9,95, 9,94, 9,87, 9,98, 10,09, 9,84, 9,74, 9,84, 10, 9,97

a) Obtenha a estimativa **pontual** da média populacional; Resposta: $\widehat{\mu}_y = 9{,}898$ min.
b) Obtenha a estimativa **pontual** da variância populacional; Resposta: $\widehat{\sigma}_y^2 = 0{,}030\ min.^2$
c) Obtenha erro padrão associado à estimativa **pontual** da média populacional. Resposta: $SE_{\widehat{\mu}} = \dfrac{\widehat{\sigma}_y}{\sqrt{n}} = \dfrac{0,172}{\sqrt{15}} = 0{,}045$ min.

2) Uma pesquisa foi conduzida com o objetivo de investigar a proporção de empresas que utilizam a linguagem $SoftwareR$ para visualizar e processar dados. Para isso, uma amostra de 80 empresas foi selecionada, e os resultados estão descritos a seguir.

Outra linguagem, Outra linguagem, Software R, Software R, Software R, Software R, Outra linguagem, Software R, Outra linguagem, Software R, Outra linguagem, Software R, Software R, Outra linguagem, Outra linguagem, Software R, Software R, Software R, Outra linguagem, Software R, Software R, Software R, Software R, Software R, Software R, Outra linguagem, Outra linguagem, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Outra linguagem, Software R, Software R, Outra linguagem, Software R, Software R, Software R, Software R, Outra linguagem, Software R, Software R, Software R, Outra linguagem, Outra linguagem, Software R, Outra linguagem, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Software R, Outra linguagem, Software R, Outra linguagem, Outra linguagem, Software R, Outra linguagem, Outra linguagem, Software R, Software R, Software R, Software R

a) Obtenha a estimativa **pontual** da proporção populacional de empresas que usam $SoftwareR$; Resposta: $\widehat{p} = 0{,}738$
b) Obtenha erro padrão associado à estimativa **pontual** da proprorção populacional de empresas que usam $SoftwareR$. Resposta: $SE_{\widehat{p}} = 0{,}0492 = 4{,}92\%$

3) Em certo lago com algo na ordem de 40000 peixes, uma amostra aleatória de 1000 peixes constatou-se que 290 tilápias. Obtenha a estimativa pontual e construa um intervalo de 95% de confiança para verdadeira proporção de tilápias na população piscosa do lago. Resposta: $\widehat{p} = 0,29$; $IC(p, 95\%) = [0,262; 0,318]$.

4) Um analista de infraestrutura de TI deseja estimar o tempo médio de resposta (em milissegundos) de um servidor web após uma atualização no sistema. Para isso, ele coleta uma amostra aleatória de 20 requisições, obtendo:

238,52, 241,58, 239,04, 239,08, 238, 239,73, 239,68, 239,37, 239,89, 240,43, 239,22, 238,71, 239,22, 240,01, 239,85, 239,3, 241,19, 240,34, 240,51, 239,71

Supõe-se que o tempo de resposta segue distribuição normal. Construa o intervalo de confiança de 95% para o tempo médio (μ) de resposta do servidor. Resposta: $\widehat{\mu} = 239,67$ milissegundos; $IC(\mu,\ 95\%) : [239,29;\ \ 240,05]$

Referências

[1] DEVORE, Jay L. *Probabilidade e Estatística para Engenharia e Ciências*. Tradução da 9ª edição norte-americana. São Paulo: Cengage Learning Brasil, 2018.

[2] FERREIRA, D. F. *Estatística Básica*. 2. ed. Lavras: Ed. UFLA, 2009.

[3] FERREIRA, Daniel Furtado. *Fundamentos de Probabilidade*. Lavras: Editora UFLA, 2020.

[4] MAGALHÃES, M. N. *Probabilidade e Variáveis Aleatórias*. 2ª ed. São Paulo: Edusp, 2006.

[5] MAGALHÃES, M. N.; LIMA, A. C. P. *Noções de Probabilidade e Estatística*. 6ª ed. São Paulo: USP, 2006.

[6] MEYER, Paul L. *Probabilidade: Aplicações à Estatística*. Rio de Janeiro: Grupo GEN, 2010.

[7] MORETTIN, Luiz Gonzaga. *Estatística Básica: Probabilidade e Inferência*. São Paulo: Pearson, 2010.

[8] MORETTIN, Pedro A. *Estatística Básica*. Rio de Janeiro: Grupo GEN, 2023.