Estatística II

Licenciatura em Gestão do Desporto
2.º Ano/2.º Semestre
2023/2024

Aulas Teórico-Práticas N.º 2 e 3 (Semana 2)

Docente: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt

https://doity.com.br/estatistica-aplicada-a-nutricao

https://basiccode.com.br/produto/informatica-basica/

Conteúdos Programáticos

Aulas Teórico-Práticas (Semanas 1 a 5)	Aulas Teórico-Práticas (Semanas 5 a 7)	Aulas Teórico-Práticas (Semanas 7 a 9)	Aulas Teórico-Práticas (Semanas 10 a 13)
• **Capítulo 1**: Revisões e Distribuições de Amostragem	• **Capítulo 2**: Estimação	• **Capítulo 3**: Testes de Hipóteses	• **Capítulo 4**: Modelo de Regressão Linear Múltipla

Material didático: Exercícios do Livro Murteira et al (2015), Formulário e Tabelas Estatísticas

Bibliografia: B. Murteira, C. Silva Ribeiro, J. Andrade e Silva, C. Pimenta e F. Pimenta; *Introdução à Estatística*, 2ª ed., Escolar Editora, 2015.

1

Distribuição F-Snedcor

Variáveis Aleatórias Contínuas

Distribuição F-Snedcor

- A distribuição F-Snedecor é uma distribuição de probabilidade contínua.

- A distribuição F-Snedecor é dada pelo quociente entre duas variáveis aleatórias com distribuição do qui-quadrado, cada uma dividida pelos respectivos graus de liberdade.

  I.e., se $X \sim \chi^2_{(m)}$ e $Y \sim \chi^2_{(n)}$, duas variáveis aleatórias independentes, então:

$$\boxed{F \equiv \frac{X/m}{Y/n} \cap F_{(m,n)}}$$

- Se X tem uma distribuição F-Snedecor com m e n graus de liberdade, escreve-se:

$$\boxed{X \sim F_{(m,\, n)}}$$

- A distribuição F-Snedecor tem dois parâmetros: m e n – *i.e.*, o nº de graus de liberdade do numerador e do denominador (m, n $\in$ N).

Distribuição F-Snedcor

- É uma distribuição positiva e não simétrica.
- E[X] = n/(n-2), quando n > 2

$$VAR[X] = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \quad , \text{ quando n > 4}$$

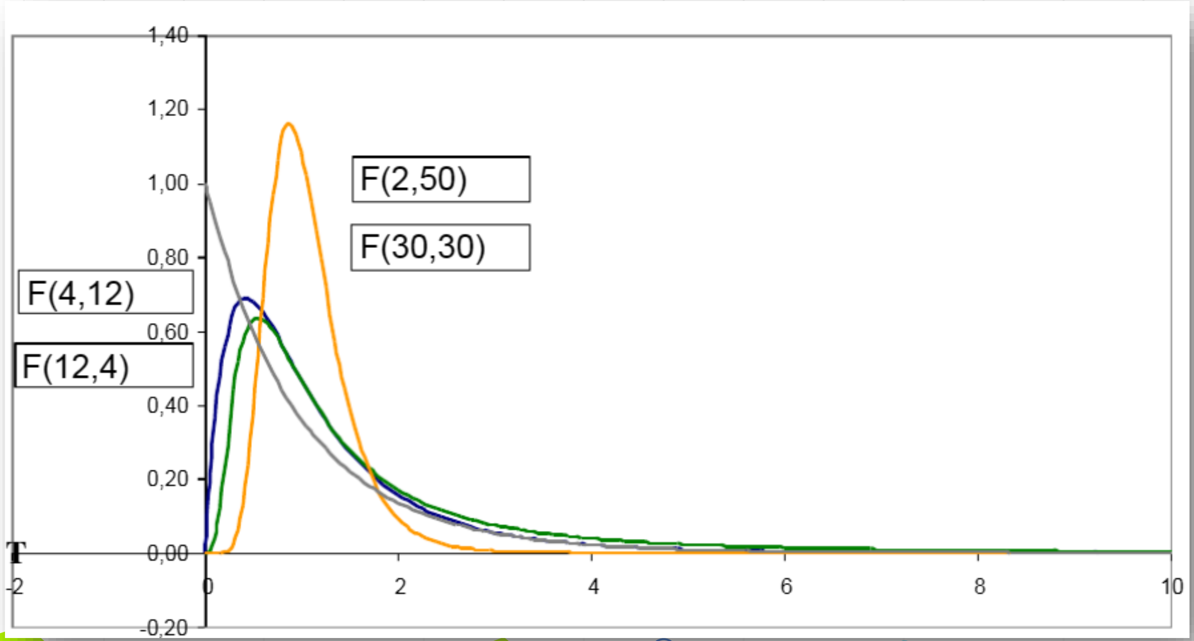
- Se $X \sim F_{(m,\,n)}$, então:

$$\frac{1}{X} \cap F_{(n,m)}$$

- Se $X \sim t_{(n)}$, então $X^2 \sim F_{(1,\,n)}$
- A distribuição F-Snedecor está tabelada para algumas probabilidade e alguns m e n

(m, n = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20, 30, 60, 120, ∞)

Distribuição F-Snedcor



Distribuição Qui-quadrado, Distribuição t-Student e Distribuição F-Snedecor - Distribuições - Studocu

Distribuição F-Snedcor

Formulário

- ***F*-SNEDCOR**

$$F = \frac{U/m}{V/n} \sim F(m,n) \text{ com } U \sim \chi^2(m),\ V \sim \chi^2(n) \text{ (independentes)}$$

$$E(X) = \frac{n}{n-2} \quad (n > 2);\ \operatorname{Var}(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \quad (n > 4)$$

Propriedades:
- $X \sim F(m,n) \Rightarrow \frac{1}{X} \sim F(n,m)$
- $T \sim t_{(n)} \Rightarrow T^2 \sim F(1,n)$

Distribuição do F-Snedcor: Exercícios

Variáveis Aleatórias Contínuas

2

Suponha que $X \sim F_{(10;\,5)}$.

a) Calcule o valor de <u>a</u> tal que $P[X > a] = 0{,}025$;

b) Calcule o valor de <u>b</u> tal que $P[X < b] = 0{,}05$;

c) Calcule os valores de <u>c</u> e <u>d</u> tal que $P[c < X < d] = 0{,}9$.

Exercício a)

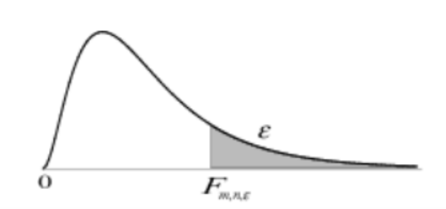
Suponha que X ~ F $_{(10;\,5)}$.

a) Calcule o valor de a tal que P[X > a] = 0,025;

b) Calcule o valor de b tal que P[X < b] = 0,05;

c) Calcule os valores de c e d tal que P[c < X < d] = 0,9.

$$F_{m,n,\varepsilon} : P(X > F_{m,n,\varepsilon}) = \varepsilon$$



m - graus de libe[illegible] do numerador

n	ε	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
	.050	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.90	245.95	248.02	249.05	250.10	251.14	252.20	253.25	254.32
	.025	647.79	799.48	864.15	899.60	921.83	937.11	948.20	956.64	963.28	968.63	976.72	984.87	993.08	997.27	1001.40	1005.60	1009.79	1014.04	1018.26
	.010	4052.18	4999.34	5403.53	5624.26	5763.96	5858.95	5928.33	5980.95	6022.40	6055.93	6106.68	6156.97	6208.66	6234.27	6260.35	6286.43	6312.97	6339.51	6365.59
2	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
	.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
	.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
	.010	98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.48	99.48	99.49	99.50
3	.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
	.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
	.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
	.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.11
	.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
	.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
	.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02

libe[illegible]de do denominador

P(X > a) = 0,025 => a = 6,62

Exercício b)

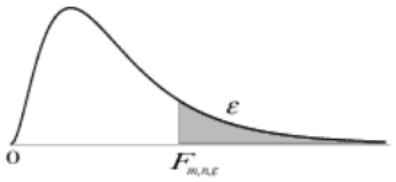
Suponha que X ~ F $_{(10;\,5)}$.

a) Calcule o valor de a tal que P[X > a] = 0,025;

b) Calcule o valor de b tal que P[X < b] = 0,05;

c) Calcule os valores de c e d tal que P[c < X < d] = 0,9.

$$F_{m,n,\varepsilon} : P(X > F_{m,n,\varepsilon}) = \varepsilon$$



	ε	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
		m – graus de liberdade do numerador																		
10	.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
	.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
	.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
	.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
	[illegible]	[illegible]	[illegible]	[illegible]	3.36	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.61	2.57	2.53	2.49	2.45	2.40
	[illegible]	[illegible]	[illegible]	[illegible]	4.28	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.17	3.12	3.06	3.00	2.94	2.88
	[illegible]	[illegible]	[illegible]	[illegible]	5.67	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.02	3.94	3.86	3.78	3.69	3.60
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.48	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.04	2.01	1.99	1.96	1.93	1.90
	[illegible]	[illegible]	[illegible]	[illegible]	3.26	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.51	2.47	2.43	2.38	2.34	2.30
	.025	6.55	5.10	4.47	4.12	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.02	2.96	2.91	2.85	2.79	2.72
	.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36

Quantil da distribuição F-Snedecor de probabilidade 0,05 com 10 e 5 graus de liberdade

Quantil da distribuição F-Snedecor de probabilidade 0,95 com 5 e 10 graus de liberdade. No caso da tabela é o valor d tal que P(F > d) = 0,05

P(X < b) = 0,05 ⇔ 1-P(X > b) = 0,05 ⇔ P(X > b) = 0,95

=> b = F0,05; 10,5 = 1/ F0,95; 5,10 = 1/3,33 = 0,30

$$F_{n_1-1,n_2-1;\frac{\alpha}{2}} = \frac{1}{F_{n_2-1,n_1-1;1-\frac{\alpha}{2}}}$$

Exercício c)

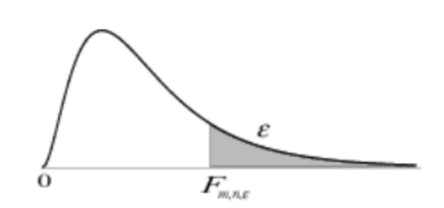
Suponha que X ~ F $_{(10;\,5)}$.

a) Calcule o valor de a tal que P[X > a] = 0,025;

b) Calcule o valor de b tal que P[X < b] = 0,05;

c) Calcule os valores de c e d tal que P[c < X < d] = 0,9.

$$F_{m,n,\varepsilon} : P(X > F_{m,n,\varepsilon}) = \varepsilon$$



liberdade do denominador	ε	m - graus de libe… do numerador 1	2	3	4	5	6	7	8	9	10	12	15							∞
1	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.2							63.33
	.050	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.90	245.9							254.32
	.025	647.79	799.48	864.15	899.60	921.83	937.11	948.20	956.64	963.28	968.63	976.72	984.8							1018.26
	.010	4052.18	4999.34	5403.53	5624.26	5763.96	5858.95	5928.33	5980.95	6022.40	6055.93	6106.68	6156.9							6365.59
2	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
	.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.4				
	.025	38.51	39.00	39.17	39.25	39.30														
	.010	98.50	99.00	99.16	99.25	99.30														
3	.100	5.54	5.46	5.39	5.34	5.31														
	.050	10.13	9.55	9.28	9.12	9.01														
	.025	17.44	16.04	15.44	15.10	14.88														
	.010	34.12	30.82	29.46	28.71	28.24														
4	.100	4.54	4.32	4.19	4.11	4.05														
	.050	7.71	6.94	6.59	6.39	6.26														
	.025	12.22	10.65	9.98	9.60	9.36														
	.010	21.20	18.00	16.69	15.98	15.52														
5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.11
	.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
	.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
	.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02

Quantil da distribuição F-Snedecor de probabilidade 0,05 com 10 e 5 graus de liberdade

$$F_{n_1-1,n_2-1;\frac{\alpha}{2}} = \frac{1}{F_{n_2-1,n_1-1;1-\frac{\alpha}{2}}}$$

P(c < X < d) = 0,9

P(X > d) = 0,05 => d = 4,74

P(X > c) = 0,95 => c = F0,05; 10,5 = 1/ F0,95; 5,10 = 1/3,33 = 0,30

(ver slide a seguir)

Exercício c)

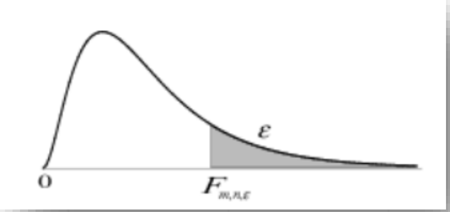
Suponha que X ~ F $_{(10;\,5)}$.

a) Calcule o valor de a tal que P[X > a] = 0,025;

b) Calcule o valor de b tal que P[X < b] = 0,05;

c) Calcule os valores de c e d tal que P[c < X < d] = 0,9.

$$F_{m,n,\varepsilon} : P(X > F_{m,n,\varepsilon}) = \varepsilon$$



		m – graus de liberdade do numerador																		
	ε	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
10	.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
	.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
	.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
	.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
	.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
	.025	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
	.010	9.65	7.21	6.22	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.60
12	.100	3.18	2.81	2.61																1.90
	.050	4.75	3.89	3.49																2.30
	.025	6.55	5.10	4.47																2.72
	.010	9.33	6.93	5.95																3.36

P(c < X < d) = 0,9

P(X > c) = 0,95 => c = F0,05; 10,5 = 1/ F0,95; 5,10 = 1/3,33 = 0,30

$$F_{n_1-1,n_2-1;\frac{\alpha}{2}} = \frac{1}{F_{n_2-1,n_1-1;1-\frac{\alpha}{2}}}$$

3 Inferência Estatística

Conceitos: Universo, Amostra Aleatória, Distribuição de uma Amostra Aleatória e Estatísticas

Probabilidade vs. Inferência Estatística

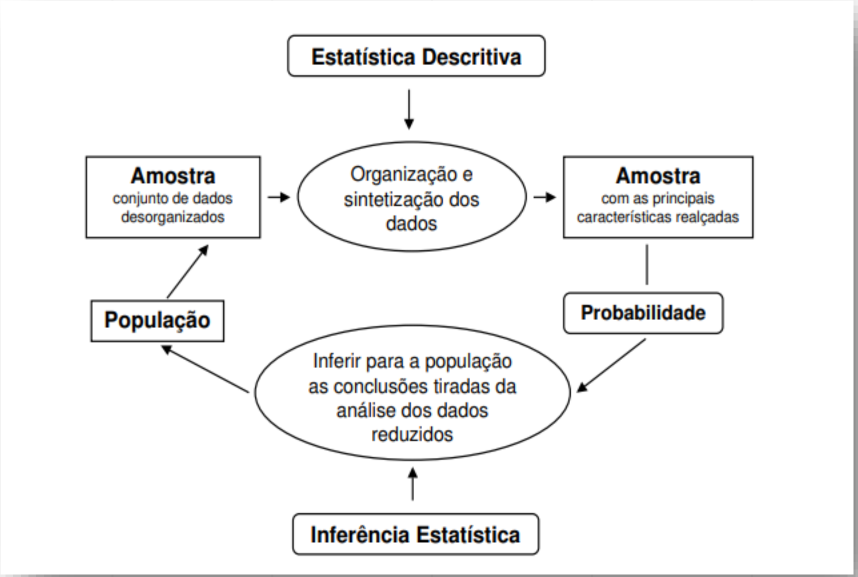
Pode dizer-se que a Probabilidade e a Inferência têm objectivos diferentes: enquanto na Probabilidade se parte de um dado esquema ou modelo para calcular a probabilidade de certos resultados ou acontecimentos se realizarem; na Inferência parte-se de dados ou observações e procura saber-se ou inferir-se algo sobre o modelo.

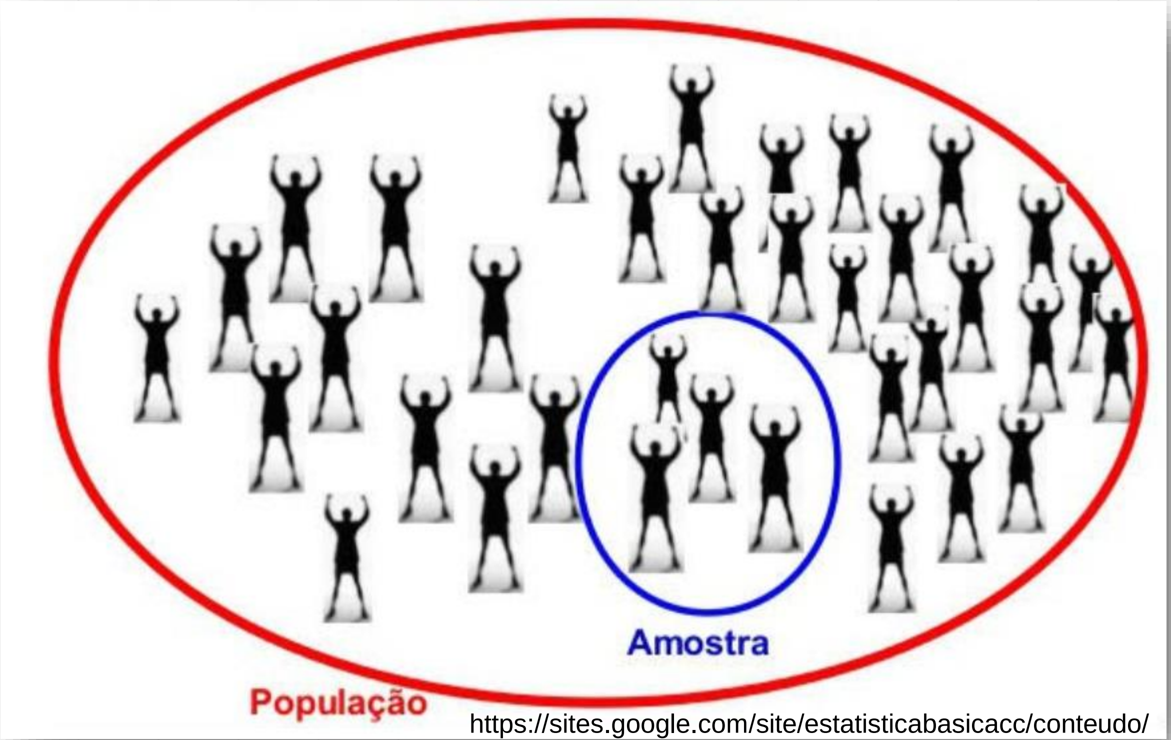
A Inferência é a "passagem do particular ao geral."

A Inferência Estatística tem como objectivo definir procedimentos que aplicados a uma amostra extraída da população, nos permitam estimar parâmetros desconhecidos dessa população ou algo sobre o modelo da população.

De facto uma amostra particular é apenas uma das muitas amostras (em nº infinito se a população for infinita) que se podem obter por um processo de amostragem.

Estatística Descritiva vs. Inferência Estatística





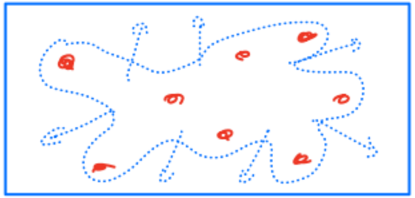
https://sites.google.com/site/estatisticabasicacc/conteudo/

- **POPULAÇÃO:**
  - é uma coleção completa de todos os elementos a serem estudados
- **AMOSTRA:**
  - é um subconjunto da população
- **CENSO:**
  - é uma coleção da dados relativos a todos os elementos de uma população:

https://slideplayer.com.br/slide/2627398

Amostra Aleatória

Perante "falta de informação" sobre a v.a. (população), vamos recolher uma amostra, estudar intensamente esta amostra e tentar extrapolar p/a população.



Ω

• – elementos da população que vão ser estudados

$\Downarrow$

amostra

$(x_1, x_2, \ldots, x_n)$

População vs Amostra Aleatória

população → recolher um subconjunto representativo da população →

① $(X_1, X_2, \ldots, X_n)$

(amostra aleatória)

(X)

↓

v.a. interesse

② estudar a amostra, através da estatística descritiva

→ ③ inferir ("extrapolar") os resultados da amostra p/ a população.

Slides Professora Claúdia Nunes

Amostragem

① Amostragem (Regras p/ assegurar a "bondade" dos resultados)

Amostra é representativa da população se tiver as seguintes características:

$(X_1, X_2, \ldots, X_n)$: conjunto de variáveis aleatórias independentes entre si e têm todas a mesma distribuição, que é a distribuição da v.a. (população) X

a.a. = amostra aleatória
i.i.d. = independentes e identicamente distribuídas

Slides Professora Claúdia Nunes

Distribuição de uma a.a.: V.a.´s Discretas

Qual é a distribuição da a.a. $(X_1, \ldots, X_n)$, sendo as v.a.´s X_i´s $(i = 1, \ldots, n)$ iid a uma v.a. discreta X?

Distribuição de uma a.a.: V.a.´s Discretas

1ª Consequência: Qual a distribuição de uma a.a.?

i) se X for uma v.a. discreta

$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$ [função de prob. conjunta da a.a.]

$\underset{\text{independentes}}{=} P(X_1 = x_1)\, P(X_2 = x_2) \ldots P(X_n = x_n)$

$\underset{\text{ident. distribuídas}}{=} P(X = x_1)\, P(X = x_2) \ldots P(X = x_n)$

$= \prod_{i=1}^{n} P(X = x_i)$

Slides Professora Claúdia Nunes

Distribuição de uma a.a.: V.a.´s Poisson

Determine a seguinte função de probabilidade conjunta da a.a. $(X1, X2, X3)$: $P(X1=4, X2=3, X3=5)$, sendo Xi´s $(i = 1, 2, 3)$ iid a $X \cap P(\lambda)$.

Distribuição de uma a.a.: V.a.´s Poisson

exemplo: $X \sim Poisson(\lambda) \qquad \lambda = ?$

$$P(X_1 = 4, X_2 = 3, X_3 = 5) = \prod_{i=1}^{3} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$$

$$= \frac{e^{-\lambda} \lambda^4}{4!} \frac{e^{-\lambda} \lambda^3}{3!} \frac{e^{-\lambda} \lambda^5}{5!} = \frac{e^{-3\lambda} \lambda^{12}}{4!\,3!\,5!}$$

Slides Professora Claúdia Nunes

Distribuição de uma a.a.: V.a.´s Contínuas

Qual é a distribuição da a.a. (X1, …, Xn), sendo Xi´s (i = 1, …, n) iid a uma v.a. Contínua X?

Distribuição de uma a.a.: V.a.´s Contínuas

ii) se X for uma v.a. contínua

$$f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n)$$ [função de densidade de probabilidade conjunta]

$$\underset{\text{independentes}}{=} f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_n}(x_n)$$

$$= f_X(x_1) f_X(x_2) \ldots f_X(x_n)$$

identicamente distribuídas

$$= \prod_{i=1}^{n} f_X(x_i)$$

Slides Professora Claúdia Nunes

Distribuição de uma a.a.: V.a.´s Exponenciais

Determine a seguinte função densidade de probabilidade conjunta da a.a. $(X1, X2, X3, X4)$: $f(1.3, 0.5, 2.5, 0.78)$, sendo Xi´s $(i = 1, 2, 3, 4)$ iid a $X \cap Exp(\lambda)$.

Distribuição de uma a.a.: V.a.´s Exponenciais

exemplo: $X \sim Exp(\mu)$

$x_1 = 1.3$; $x_2 = 0.5$; $x_3 = 2.5$; $x_4 = 0.78$

$$f_{X_1, X_2, X_3, X_4}(1.3, 0.5, 2.5, 0.78) = \mu e^{-\mu \times 1.3} \times \mu e^{-\mu \times 0.5}$$

$$\times \mu e^{-\mu \times 2.5} \times \mu e^{-\mu \times 0.78}$$

$$= \mu^4 e^{-\mu(1.3 + 0.5 + 2.5 + 0.78)} //$$

Slides Professora Claúdia Nunes

Inferência Estatística

③ Inferência estatística

- X ✓ (v.c. de interesse)
- com distribuição conhecida $\mathcal{D}$
- Todos ou alguns parâmetros da distribuição são desconhecidos

parâmetros desconhecidos (θ)

Objectivo: "adivinhar" o valor de θ, com base na amostra

i.e

estimar θ, com base na amostra

⇓

Estimação { Pontual (cap. 6) ; Intervalar (cap. 7) }

Slides Professora Claúdia Nunes

Estatística

i) **Estatística**: uma estatística é uma função (qualquer) que depende da amostra e não depende de nenhum parâmetro desconhecido.

Exemplo: $X \sim Poisson(\lambda)$ $\quad \lambda = ?$

a.a. $(X_1, X_2, \ldots, X_{10})$

$T_1 = \bar{X} = \frac{1}{10}\sum_{i=1}^{n} X_i$: É UMA ESTATÍSTICA

$T_2 = 3$: É UMA ESTATÍSTICA

$T_3 = \frac{1}{10}\sum_{i=1}^{10}(X_i - \lambda)^2$ NÃO É ESTATÍSTICA!!

Slides Professora Claúdia Nunes

Estatística

DEF. 4: ESTATÍSTICA:

Uma estatística é uma função das variáveis observáveis, e é por si própria uma variável observável que não contém qualquer parâmetro desconhecido.

EXEMPLO 5: Seja a a.a. $[X_1, X_2, \ldots, X_n]$. A média amostral $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ é uma estatística.

EXEMPLO 6: Seja a v.a $X \sim N(\mu, \sigma^2)$ com μ e σ^2 desconhecidos, então $x - \mu$ não é estatística porque não é observável, é função do parâmetro desconhecido μ.

EXEMPLO 6: Seja uma v.a. com distribuição $N(\mu;\sigma^2)$. Quais são Estatísticas?

a) $X^2-\mu$

b) $\frac{X}{\sigma^2}$

c) X^2-3

d) $X-4$

e) $X - \log X^3$

Quais destas funções são estatísticas?

apostila.PDF (ufpr.br)

Média e Variância Amostrais

② Duas características da amostra frequentemente calculadas são:

<u>média amostral</u>: $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$

<u>variância amostral</u>: $S^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2 \quad (*)$$

(*) porque: $\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 =$

$$= \frac{1}{n} \sum_{i=1}^{n} (x_i^2 + \bar{x}^2 - 2x_i\bar{x}) = \frac{1}{n} \Big[\sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2\bar{x} \underbrace{\sum_{i=1}^{n} x_i}_{n\bar{x}} \Big]$$

$$= \frac{1}{n} \Big[\sum_{i=1}^{n} x_i^2 + n\bar{x}^2 - 2n\bar{x}^2 \Big]$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2$$

variância amostral corrigida: $S^2_{n-1} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

4

Média Amostral

Distribuições de Amostragem

Média Amostral

Seja $X_1, X_2, \ldots, X_n$, uma a. a. duma dada população com média μ e variância σ^2. A **média amostral** é definida por:

$$\overline{X} = \sum_{i=1}^{n} \frac{X_i}{n}.$$

Propriedades:

- $\mu_{\overline{X}} = E(\overline{X}) = \mu;$
- $\sigma^2_{\overline{X}} = Var(\overline{X}) = \frac{1}{n}\sigma^2.$

ProbabilidadesEstatistica2019.pdf

Média Amostral: Variância Conhecida

Se a distribuição da população é Normal com desvio padrão σ conhecido, então $\overline{X} \sim N\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$, ou seja,

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0; 1).$$

Se a distribuição da população não for Normal, mas a amostra é de grande dimensão então, pelo corolário do T. L. C., vem

$$Z \overset{\circ}{\sim} N(0; 1).$$

ProbabilidadesEstatistica2019.pdf

Média Amostral: Variância Desconhecida

Se a população é Normal mas o desvio padrão σ é desconhecido, e não rejeitando a hipótese de independência das distribuições por amostragem da média e da variância da amostra, então tem-se que:

$$T = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}.$$

Se a distribuição da população não for Normal, mas a amostra for de grande dimensão então, por extensão do corolário do T. L. C.,

$$\overline{X} \overset{\circ}{\sim} N\left(\mu; \frac{S}{\sqrt{n}}\right), \text{ ou seja, } Z = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \overset{\circ}{\sim} N(0; 1).$$

Observação: Para valores elevados de n, a distribuição *t*-Student toma valores muito próximos aos da $N(0; 1)$. Existem alguns programas estatísticos (por exemplo, SPSS) que, nestas condições, não utilizam a distribuição Normal mas a *t*-Student.

ProbabilidadesEstatistica2019.pdf

Distribuições de Amostragem

Formulário

AMOSTRAGEM. DISTRIBUIÇÕES POR AMOSTRAGEM

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \; ; \qquad S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n} = \frac{\sum_{i=1}^{n} X_i^2}{n} - \overline{X}^2 \; ; \qquad (n-1)S'^2 = n\,S^2$$

$$E(\overline{X}) = \mu \; ; \qquad \mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n} \; ; \qquad E(S^2) = \frac{n-1}{n}\sigma^2 \; ; \qquad E(S'^2) = \sigma^2$$

Média Amostral: Exercícios

Distribuições de Amostragem

5

Um fabricante de automóveis defende que o novo modelo que vai ser lançado no próximo mês gasta em média 9,7 litros aos 100 km, em circuito urbano, com desvio padrão de 1 litro. Admita que o consumo segue uma distribuição Normal.

a) Qual a probabilidade de numa amostra aleatória de 20 automóveis o gasto médio ser superior a 10 litros.

b) Qual a deverá ser a dimensão da amostra para obter, com pelo menos 90% probabilidade, um gasto médio inferior a 10 litros.

ProbabilidadesEstatistica2019.pdf

Exercício: Média Amostral com Variância Conhecida

Seja X a v.a. que representa o número de litros consumidos pelo automóvel aos 100 km, em circuito urbano, com $X \sim N(\mu = 9{,}7;\ \sigma = 1)$.

a) $n = 20$.

$$\left.\begin{array}{l} X \text{ dist. Normal} \\ \sigma \text{ conhecido} \end{array}\right| \Rightarrow Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0;1).$$

$$P(\overline{X} > 10) = 1 - P(\overline{X} \le 10) = 1 - P\left(Z \le \frac{10 - 9{,}7}{\frac{1}{\sqrt{20}}}\right) = 1 - P(Z \le 1{,}34) = 1 - \Phi(1{,}34)$$

$$= 1 - 0{,}9099 = 0{,}0901.$$

b) $n = ?$

$$P(\overline{X} < 10) \ge 0{,}9 \Leftrightarrow P\left(Z < \frac{10 - 9{,}7}{\frac{1}{\sqrt{n}}}\right) \ge 0{,}9 \Leftrightarrow \Phi(0{,}3\sqrt{n}) \ge 0{,}9$$

como $\Phi(1{,}282) = 0{,}9$

$$\Leftrightarrow 0{,}3\sqrt{n} \ge 1{,}282 \Leftrightarrow \sqrt{n} \ge 4{,}2733 \Rightarrow n \ge 4{,}2733^2 = 18{,}3 \Rightarrow n \ge 19.$$

ProbabilidadesEstatistica2019.pdf

Um fabricante de automóveis defende que o novo modelo que vai ser lançado no próximo mês gasta em média 9,7 litros aos 100 km, em circuito urbano, e o desvio padrão é desconhecido.
Através de um esquema de amostragem estimou-se tal desvio padrão como sendo $s = 1$ litro. Admitindo que o consumo segue uma distribuição Normal, qual a probabilidade de, numa amostra aleatória de 20 automóveis, o consumo médio amostral ser superior a 10 litros? E inferior a 8,9 litros?

ProbabilidadesEstatistica2019.pdf

Exercício: Média Amostral com Variância Desconhecida

Seja X a v.a. que representa o número de litros consumidos pelo automóvel aos 100 km, em circuito urbano, com $X \sim N(\mu = 9{,}7;\ \sigma = ?\,)$.

$n = 20.$

$$\left.\begin{array}{l} X \text{ dist. Normal} \\ \sigma \text{ desconhecido} \end{array}\right| \Rightarrow T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1=19}.$$

$$P(\bar{X} > 10) = 1 - P(\bar{X} \leq 10) = 1 - P\left(T \leq \frac{10 - 9{,}7}{\frac{1}{\sqrt{20}}}\right) = 1 - P(T \leq 1{,}342) \approx 1 - 0{,}9 = 0{,}1.$$

$$P(\bar{X} < 8{,}9) = P\left(T < \frac{8{,}9 - 9{,}7}{\frac{1}{\sqrt{20}}}\right) = P(T < -3{,}578) = 1 - P(T < 3{,}578) \approx 1 - 0{,}999 = 0{,}001.$$

ProbabilidadesEstatistica2019.pdf

Obrigada!

Questões?