

XMAC02

Métodos Matemáticos para Análise de Dados

Motivação

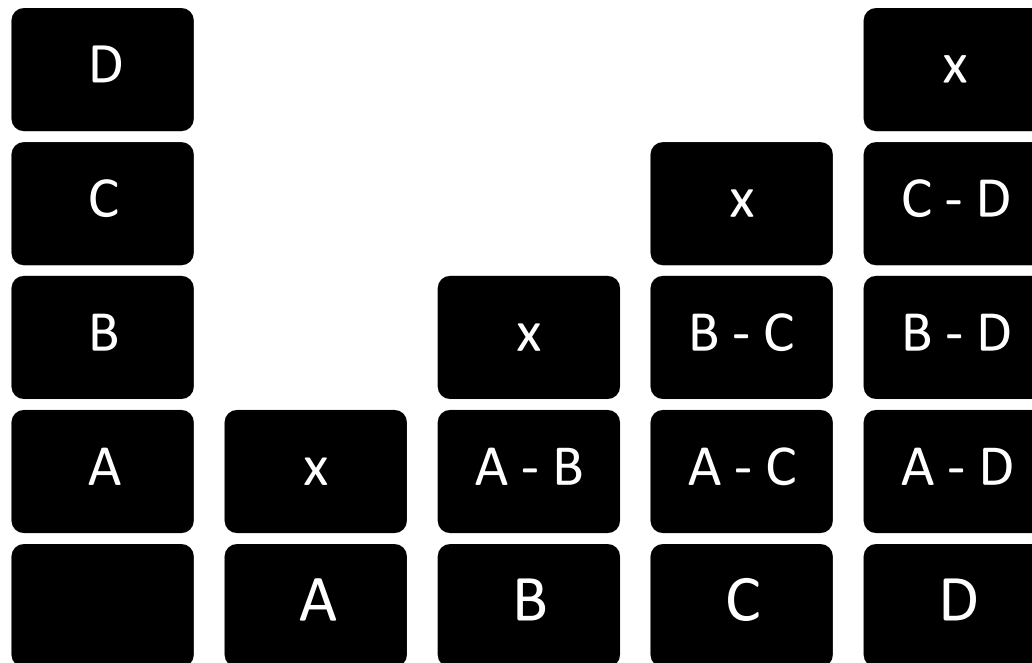
2

- Suponha que eu queira comparar a média de 3 populações: A, B e C
 - ▣ Posso usar o teste t, mas seria necessário realizar 3 testes
 - A-B
 - B-C
 - C-A

Motivação

3

□ Para 4 populações teríamos:



▣ 6 testes seriam necessários

Motivação

4

- ❑ Ao realizar 6 testes t com confiança de 95%, teríamos o seguinte:

$$0.95 \times 0.95 \times 0.95 \times 0.95 \times 0.95 \times 0.95 = 0.735$$

- ❑ Assim, ao invés de termos 1 erro a cada 20 testes, teríamos 1 erro a cada 4 testes!

ANOVA

5

- Permite que testemos a média de múltiplas populações num único teste

$$H_0: \mu_A = \mu_B = \mu_C = \mu_D \dots = \mu_k$$

H_a : Ao menos uma das médias
é diferente das outras

ANOVA

6

- ❑ Vamos examinar 3 máquinas de envase de perfume

Máquina 1	Máquina 2	Máquina 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

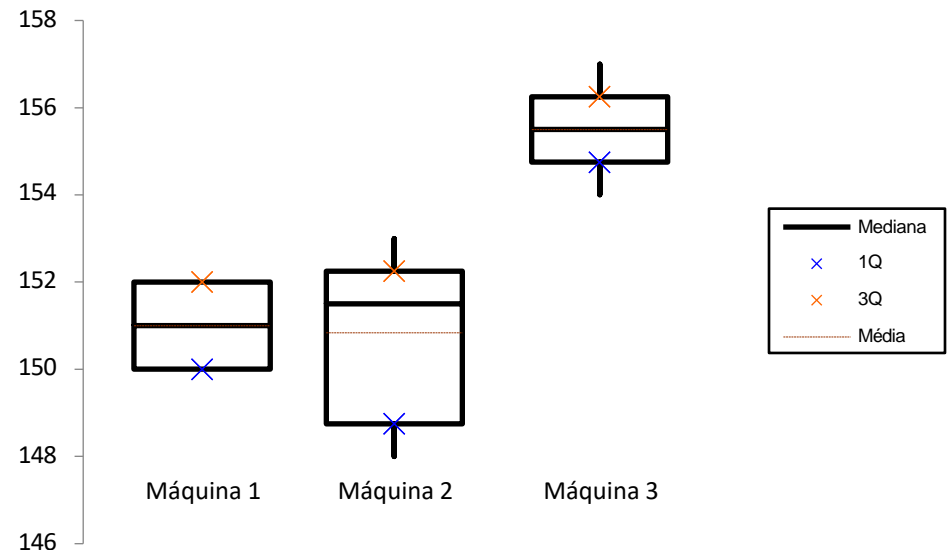
- ❑ $H_0: \mu_1 = \mu_2 = \mu_3$?
- ❑ À primeira vista, parece que μ_3 pode ser diferente

ANOVA

7

Examinando as máquinas 1, 2 e 3

Máquina 1	Máquina 2	Máquina 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$



ANOVA

8

- Agora, vamos examinar as máquinas 4, 5 e 6

Machine 4	Machine 5	Machine 6
130	163	166
155	152	154
160	143	155
158	141	151
152	149	152
145	157	155
$\bar{x}_4 = 151.00$	$\bar{x}_5 = 150.83$	$\bar{x}_6 = 155.50$

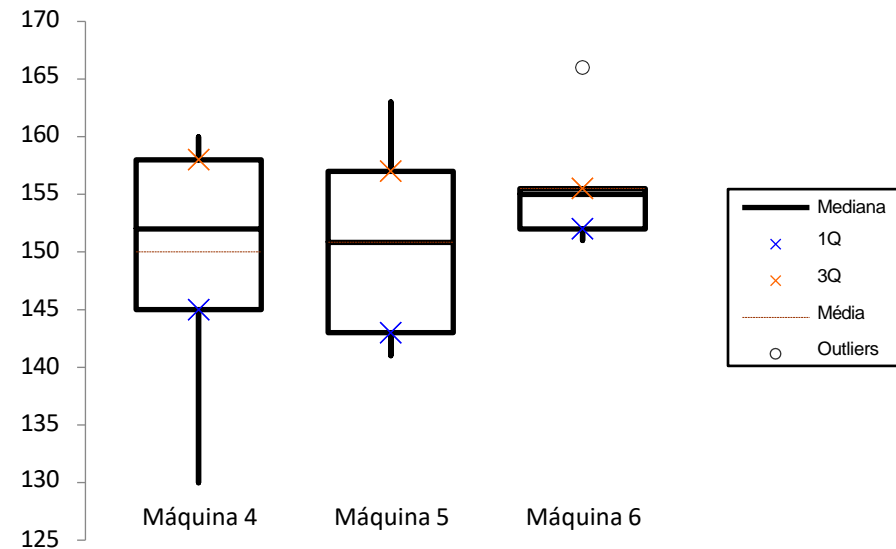
- $H_0: \mu_4 = \mu_5 = \mu_6$?
- À primeira vista, parece que μ_6 pode ser diferente

ANOVA

9

□ Agora, vamos examinar as máquinas 4, 5 e 6

Machine 4	Machine 5	Machine 6
130	163	166
155	152	154
160	143	155
158	141	151
152	149	152
145	157	155
$\bar{x}_4 = 151.00$	$\bar{x}_5 = 150.83$	$\bar{x}_6 = 155.50$

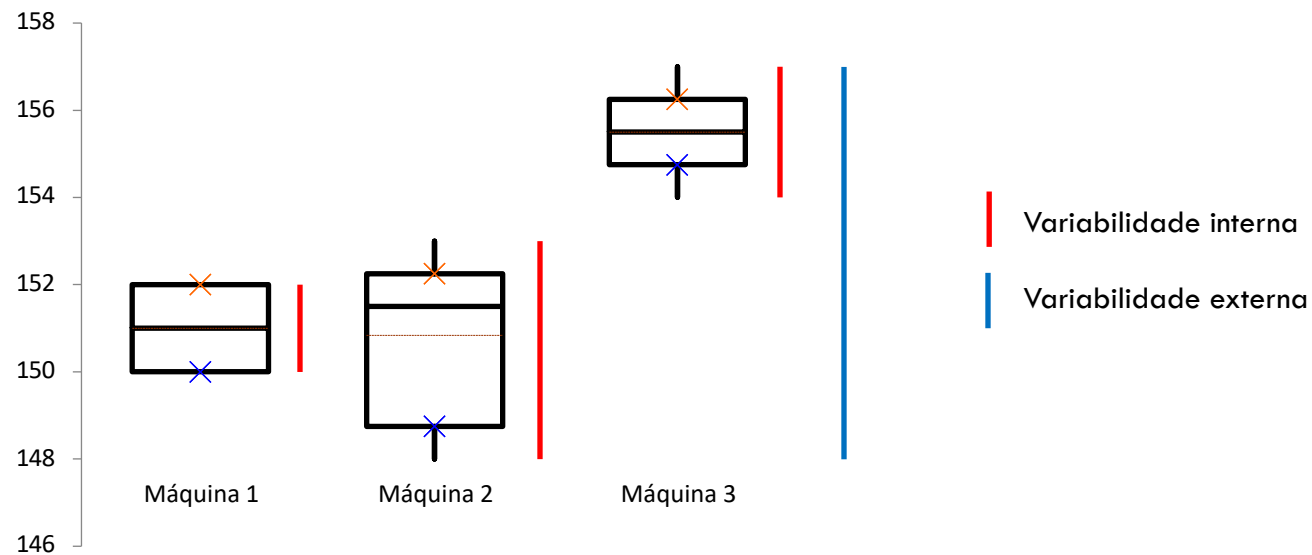


ANOVA

Variabilidade interna vs externa

10

❑ Máquinas 1, 2 e 3

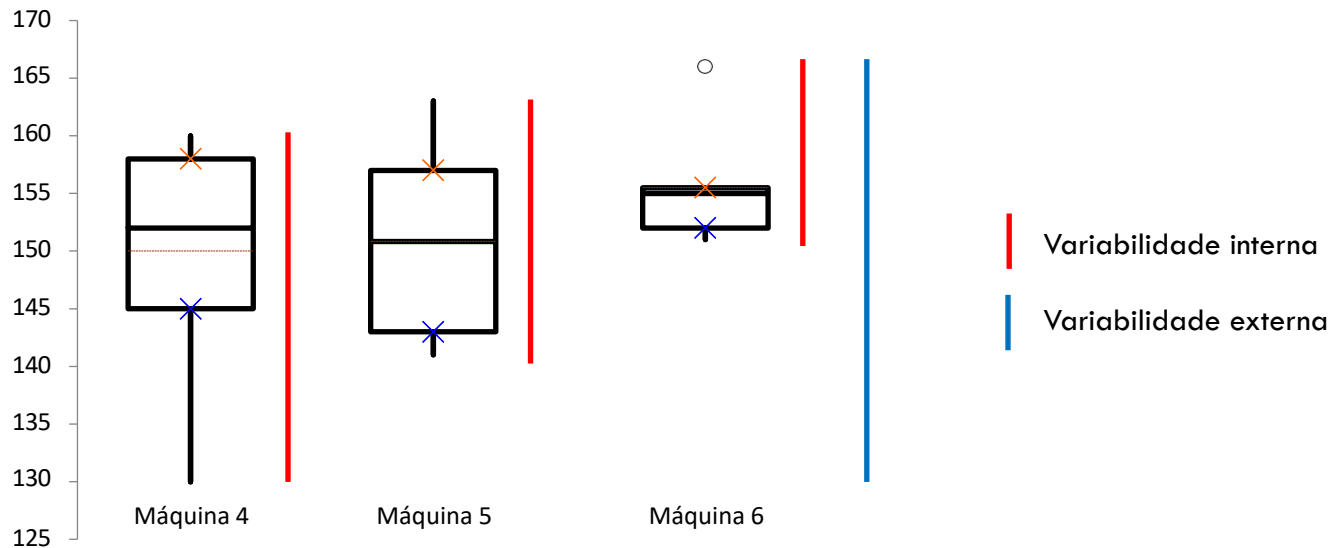


ANOVA

Variabilidade interna vs externa

11

❑ Máquinas 4, 5 e 6



ANOVA

12

- ❑ ANOVA = ANalysis Of Variance
- ❑ Variância

$$s^2 = \frac{\sum (x_i - \bar{X})^2}{n-1}$$

- ❑ Numerador: soma de quadrados (SQ)
- ❑ Denominador: degree of freedom

ANOVA

13

□ Variância de 2 pop.

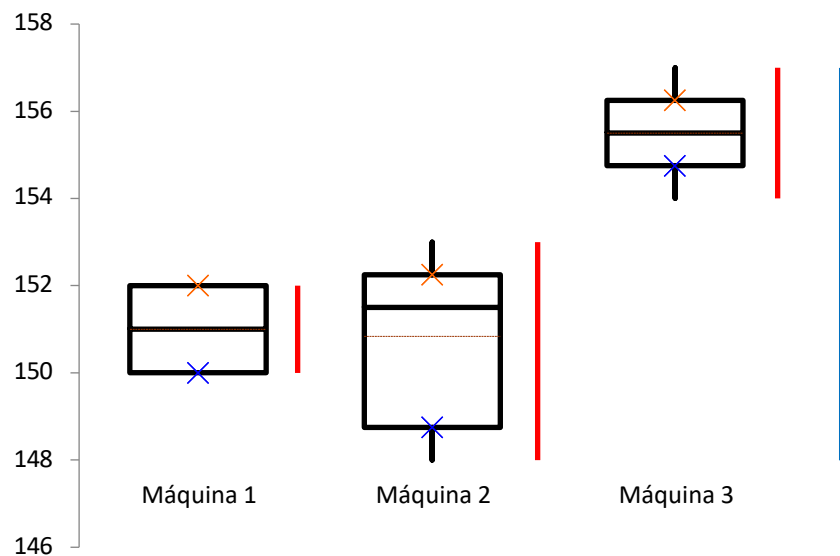
$$F = \frac{s_1^2}{s_2^2}$$

$$F = \frac{\frac{\sum (x - \bar{x}_1)^2}{n_1 - 1}}{\frac{\sum (x - \bar{x}_2)^2}{n_2 - 1}}$$

$$F = \frac{\frac{SQ_1}{df_1}}{\frac{SQ_2}{df_2}} \quad F = \frac{MSQ_1}{MSQ_2}$$

□ Variância de 3+ pop.

$$F = \frac{MSQ_{\text{externa}}}{MSQ_{\text{interna}}} \quad F = \frac{\frac{SQ_{\text{externa}}}{df_{\text{externa}}}}{\frac{SQ_{\text{interna}}}{df_{\text{interna}}}}$$



ANOVA

Exemplo

14

Máquina 1	Máquina 2	Máquina 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

$$H_0: \mu_1 = \mu_2 = \mu_3$$

H_a : Médias não são iguais

Nível de confiança: 95%

$$F = \frac{\frac{SQ_{\text{externa}}}{df_{\text{externa}}}}{\frac{SQ_{\text{interna}}}{df_{\text{interna}}}}$$

ANOVA

Exemplo

15

Máquina 1	Máquina 2	Máquina 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

$$SQ_{\text{interna}} = 4.00 + 18.83 + 5.50 = 28.33$$

Máq 1	$x_1 - \bar{x}_1$	$Sqr(x_1 - \bar{x}_1)$	Máq 2	$x_2 - \bar{x}_2$	$Sqr(x_2 - \bar{x}_2)$	Máq 3	$x_3 - \bar{x}_3$	$Sqr(x_3 - \bar{x}_3)$	
150.00	-1.00	1.00	153.00	2.17	4.69	156.00	0.50	0.25	
151.00	0.00	0.00	152.00	1.17	1.36	154.00	-1.50	2.25	
152.00	1.00	1.00	148.00	-2.83	8.03	155.00	-0.50	0.25	
152.00	1.00	1.00	151.00	0.17	0.03	156.00	0.50	0.25	
151.00	0.00	0.00	149.00	-1.83	3.36	157.00	1.50	2.25	
150.00	-1.00	1.00	152.00	1.17	1.36	155.00	-0.50	0.25	
151.00			150.83			155.50			152.44
		4.00			18.83			5.50	

ANOVA

Exemplo

16

Máquina 1	Máquina 2	Máquina 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

$SQ_{\text{externa}} = ?$

Máq 1	$x_1 - \bar{x}_1$	$Sqr(x_1 - \bar{x}_1)$	Máq 2	$x_2 - \bar{x}_2$	$Sqr(x_2 - \bar{x}_2)$	Máq 3	$x_3 - \bar{x}_3$	$Sqr(x_3 - \bar{x}_3)$	
150.00	-1.00	1.00	153.00	2.17	4.69	156.00	0.50	0.25	
151.00	0.00	0.00	152.00	1.17	1.36	154.00	-1.50	2.25	
152.00	1.00	1.00	148.00	-2.83	8.03	155.00	-0.50	0.25	
152.00	1.00	1.00	151.00	0.17	0.03	156.00	0.50	0.25	
151.00	0.00	0.00	149.00	-1.83	3.36	157.00	1.50	2.25	
150.00	-1.00	1.00	152.00	1.17	1.36	155.00	-0.50	0.25	
151.00			150.83			155.50			152.44
		4.00			18.83			5.50	
	-1.44	2.07		-1.61	2.58		3.06	9.36	

ANOVA

Exemplo

17

Máquina 1	Máquina 2	Máquina 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

$$SQ_{\text{externa}} = (2.07 + 2.58 + 9.36) \times 6 = 84.06$$

Máq 1	$x_1 - \bar{x}_1$	$Sqr(x_1 - \bar{x}_1)$	Máq 2	$x_2 - \bar{x}_2$	$Sqr(x_2 - \bar{x}_2)$	Máq 3	$x_3 - \bar{x}_3$	$Sqr(x_3 - \bar{x}_3)$	
150.00	-1.00	1.00	153.00	2.17	4.69	156.00	0.50	0.25	
151.00	0.00	0.00	152.00	1.17	1.36	154.00	-1.50	2.25	
152.00	1.00	1.00	148.00	-2.83	8.03	155.00	-0.50	0.25	
152.00	1.00	1.00	151.00	0.17	0.03	156.00	0.50	0.25	
151.00	0.00	0.00	149.00	-1.83	3.36	157.00	1.50	2.25	
150.00	-1.00	1.00	152.00	1.17	1.36	155.00	-0.50	0.25	
151.00			150.83			155.50			152.44
		4.00			18.83			5.50	
	-1.44	2.07		-1.61	2.58		3.06	9.36	

ANOVA

Exemplo

18

Máquina 1	Máquina 2	Máquina 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

$$SQ_{\text{interna}} = 4.00 + 18.83 + 5.50 = 28.33$$

$$SQ_{\text{externa}} = (2.07 + 2.58 + 9.36) \times 6 = 84.06$$

ANOVA

Exemplo

19

Máquina 1	Máquina 2	Máquina 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	$\bar{x}_2 = 150.83$	$\bar{x}_3 = 155.50$

□ $df_{\text{total}} = 18 - 1 = 17$

□ $df_{\text{externa}} = 3 - 1 = 2$

□ $df_{\text{interna}} = df_{\text{total}} - df_{\text{externa}} = 17 - 2 = 15$

ANOVA

Exemplo

20

- $SQ_{\text{interna}} = 28.33$

- $SQ_{\text{externa}} = 84.06$

- $df_{\text{externa}} = 2$

- $df_{\text{interna}} = 15$

$$F = \frac{\frac{SQ_{\text{externa}}}{df_{\text{externa}}}}{\frac{SQ_{\text{interna}}}{df_{\text{interna}}}}$$

$$F = \frac{MSQ_{\text{externa}}}{MSQ_{\text{interna}}}$$

- $MSQ_{\text{externa}} = 84.06 / 2 = 42.03$

- $MSQ_{\text{interna}} = 28.33 / 15 = 1.89$

- $F = 42.03 / 1.89 = 22.24$

ANOVA

Exemplo

21

F - Distribution ($\alpha = 0.05$ in the Right Tail)

df ₂	df ₁	Numerator Degrees of Freedom								
		1	2	3	4	5	6	7	8	9
1		161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2		18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3		10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4		7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5		6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6		5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7		5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8		5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9		5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10		4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11		4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12		4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13		4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14		4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15		4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16		4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17		4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18		4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19		4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20		4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21		4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22		4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23		4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24		4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25		4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26		4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27		4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28		4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29		4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30		4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40		4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60		4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120		3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞		3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799

□ $df_{\text{externa}} = 2$ (num)

□ $df_{\text{interna}} = 15$ (den)

□ $\alpha = 0.05$ uma cauda

□ $F_{0.05, 2, 15} = 3.68$

ANOVA

Exemplo

22

- $F = 22.24$
- $F_{0.05, 2, 15} = 3.68$

$$H_0: \mu_1 = \mu_2 = \mu_3$$

H_a : Médias não são iguais

- Rejeitamos a hipótese nula
- Médias não são iguais

