



Análise de Variância (ANOVA)

Bárbara Maria de Andrade Costa
Murillo Fernando Rodrigues

São Paulo
4 a 6 de outubro, 2016

Por que fazer ANOVA?

Vamos comparar múltiplas médias

volte dois passos

Tipos de variáveis

A natureza das nossas medidas

Tipos de variáveis

- **Variável DEPENDENTE (resposta):** mede o efeito sofrido (y)
- **Variável INDEPENDENTE (preditora):** causa o efeito que procuramos investigar (x)

Tipos de variáveis

Qual a variável dependente (resposta) e independente (preditora) nas duas questões abaixo?

- A temperatura afeta a o tamanho da região nasal em roedores?
- A gradiente altitudinal afeta o tamanho da região nasal em roedores?

Tipos de variáveis

Qual a variável dependente (resposta) e independente (preditora) nas duas questões abaixo?

- A temperatura afeta a o tamanho da região nasal em roedores?
dependente *independente*

- A gradiente altitudinal afeta o tamanho da região nasal em roedores?
dependente *independente*

A natureza das nossas medidas

- **Variáveis categóricas (qualitativas):** distinções de qualidade
- **Variáveis contínuas (quantitativas):** diferenças de quantidade

Cor (preto, branco, azul) // Comprimento de onda da luz emitida

Altitude (cotas) // Altitude em gradientes

Por que fazer ANOVA?

Porque queremos investigar o efeito de uma **variável categórica (preditora)** sobre a variação de uma **variável contínua (resposta)**

Por que fazer ANOVA?

Porque queremos investigar o efeito de uma **variável categórica (preditora)** sobre a variação de uma **variável contínua (resposta)**

Por que não o teste t?

Por que fazer ANOVA?

Porque queremos investigar o efeito de uma **variável categórica (preditora)** sobre a variação de uma **variável contínua (resposta)**

Hipótese nula(H_0): $\mu_1 = \mu_2 = \dots = \mu_n$

Por que fazer ANOVA?

Porque queremos investigar o efeito de uma **variável categórica (preditora)** sobre a variação de uma **variável contínua (resposta)**

Hipótese nula(H_0): $\mu_1 = \mu_2 = \dots = \mu_n$

Hipótese alternativa (H_a): $\mu_1 \neq \mu_2 \neq \dots \neq \mu_n$

Tabela de ANOVA

Variação	Soma dos Quadrados	Graus de Liberdade	Quadrados das medias	F
Entre Grupos (SQE)				
Dentro dos grupos (SDQ)				
Total (SQT)				

Será que as subespécies de andorinha tem comprimentos de crânio diferentes?

subespécies		
a	b	c
3	5	5
2	3	6
1	4	7

valores de comprimento de crânio para cada grupo

Será que as subespécies de andorinha tem comprimentos de crânio diferentes?

subespécies		
a	b	c
3	5	5
2	3	6
1	4	7

valores de comprimento de crânio para cada grupo

Média de cada subespécie:

a = 2 b = 4 c = 6

Será que essas diferenças nas médias de cada subespécie é aleatória ?

subespécies		
a	b	c
3	5	5
2	3	6
1	4	7

valores de comprimento de
crânio para cada grupo

Média de cada subespécie:

a = 2 b = 4 c = 6

A grande média ou média das médias

a	b	c
3	5	5
2	3	6
1	4	7

$$\text{Grande média } (\bar{X}) = \frac{3 + 2 + 1 + 5 + 3 + 4 + 5 + 6 + 7}{9} = 4$$

Soma dos Quadrados Total

a	b	c
3	5	5
2	3	6
1	4	7

Grande média (\bar{X}) = 4

$$\text{SQT} = (3 - 4)^2 + (2 - 4)^2 + (1 - 4)^2 + (5 - 4)^2 + (3 - 4)^2 + (4 - 4)^2 + (5 - 4)^2 + (6 - 4)^2 + (7 - 4)^2$$

$$\text{SQT} = 30$$

Soma dos Quadrados Total

a	b	c
3	5	5
2	3	6
1	4	7

Grande média (\bar{X}) = 4

$$\text{SQT} = (3 - 4)^2 + (2 - 4)^2 + (1 - 4)^2 + (5 - 4)^2 + (3 - 4)^2 + (4 - 4)^2 + (5 - 4)^2 + (6 - 4)^2 + (7 - 4)^2$$

$$\text{SQT} = 30$$

Graus de liberdade = números de elementos - 1

Graus de liberdade = 8

Tabela de ANOVA

Variação	Soma dos Quadrados	Graus de Liberdade	Quadrados das medias	F
Entre Grupos (SQE)				
Dentro dos grupos (SDQ)				
Total (SQT)	30	8		

Quanto da variação total se deve à variação dentro e entre grupos?

a	b	c
3	5	5
2	3	6
1	4	7

$$\text{SQTotal} = \text{SQ Entre} + \text{SQ Dentro}$$

Soma dos Quadrados Dentro dos Grupos

a	b	c
3	5	5
2	3	6
1	4	7

Média de cada grupo = $X_1 = 2$ $X_2 = 4$ $X_3 = 6$

$$\text{SQD} = (3 - 2)^2 + (2 - 2)^2 + (1 - 2)^2 + (5 - 4)^2 + (3 - 4)^2 + (4 - 4)^2 + (5 - 6)^2 + (6 - 6)^2 + (7 - 6)^2$$

$$\text{SQD} = 6$$

Graus de liberdade = 6

De 30, 6 se deve a variação dentro dos grupo

$$\text{SQTotal} = \text{SQ Entre} + \text{SQ Dentro}$$

$$\text{SQT} = 30$$

$$\text{SQD} = 6$$

Tabela de ANOVA

Variação	Soma dos Quadrados	Graus de Liberdade	Quadrados das medias	F
Entre Grupos (SQE)				
Dentro dos grupos (SDQ)	6	6		
Total (SQT)	30	8		

Soma dos Quadrados Entre os Grupos

$$\text{SQT} = 30$$

$$\text{SQD} = 6$$

1	2	3
3	5	5
2	3	6
1	4	7

$$\text{Média de cada grupo} = X_1 = 2 \quad X_2 = 4 \quad X_3 = 6$$

$$\text{Grande média } (\bar{X}) = 4$$

Distância ao quadrado entre a média de cada grupo e a média das médias, para cada elemento.

$$\text{SQE} = (2 - 4)^2 + (2 - 4)^2 + (2 - 4)^2 + (4 - 4)^2 + (4 - 4)^2 + (4 - 4)^2 + (6 - 4)^2 + (6 - 4)^2 + (6 - 4)^2$$

$$\text{SQE} = 24$$

Soma dos Quadrados Entre os Grupos

$$\text{SQT} = 30$$

$$\text{SQD} = 6$$

1	2	3
3	5	5
2	3	6
1	4	7

$$\text{Média de cada grupo} = X_1 = 2 \quad X_2 = 4 \quad X_3 = 6$$

$$\text{Grande média } (\bar{X}) = 4$$

Distância ao quadrado entre a média de cada grupo e a média das médias, para cada elemento.

$$\text{SQE} = (2 - 4)^2 + (2 - 4)^2 + (2 - 4)^2 + (4 - 4)^2 + (4 - 4)^2 + (4 - 4)^2 + (6 - 4)^2 + (6 - 4)^2 + (6 - 4)^2$$

$$\text{SQE} = 24$$

$$\text{Graus de liberdade} = \text{números de subespécies} - 1$$

$$\text{Graus de liberdade} = 2$$

De 30, 24 se deve a variação entre grupos e 6 dentro dos grupos

$$\text{SQTotal} = \text{SQ Entre} + \text{SQ Dentro}$$

$$\text{SQT} = 30$$

8 g.l

$$\text{SQE} = 24$$

2 g.l

$$\text{SQD} = 6$$

6 g.l

Tabela de ANOVA

Variação	Soma dos Quadrados	Graus de Liberdade	Quadrados das medias	F
Entre Grupos (SQE)	24	2		
Dentro dos grupos (SDQ)	6	6		
Total (SQT)	30	8		

Será que as subespécies de andorinha tem comprimentos de crânio diferentes?

subespécies		
a	b	c
3	5	5
2	3	6
1	4	7

valores de comprimento de crânio para cada grupo

Será que essas diferenças nas médias de cada subespécie é aleatória ?

subespécies		
a	b	c
3	5	5
2	3	6
1	4	7

valores de comprimento de
crânio para cada grupo

Média de cada subespécie:

a = 2 b = 4 c = 6

Será que essas diferenças nas médias de cada subespécie é aleatória ?

subespécies		
a	b	c
3	5	5
2	3	6
1	4	7

valores de comprimento de
crânio para cada grupo

Média de cada subespécie:

a = 2 b = 4 c = 6

H0: $\bar{X}_a = \bar{X}_b = \bar{X}_c$ não faz diferença

H alternativa: há diferença entre comprimento do crânio para cada subespécies

Teste de hipótese com a estatística F

H0: $\bar{X}_a = \bar{X}_b = \bar{X}_c$ não faz diferença

H alternativa: as subespécies apresentam comprimento do crânio diferentes

$$\text{Estatística F} = \frac{\text{SQE} / \text{g.l}}{\text{SQD} / \text{g.l}}$$

Teste de hipótese com a estatística F

H0: $\bar{X}_a = \bar{X}_b = \bar{X}_c$ não faz diferença

H alternativa: as subespécies apresentam comprimento do crânio diferentes

$$\text{Estatística F} = \frac{\text{SQE} / m - 1}{\text{SQD} / m(n - 1)}$$

Teste de hipótese com a estatística F

H0: $\bar{X}_a = \bar{X}_b = \bar{X}_c$ não faz diferença

H alternativa: as subespécies apresentam comprimento do crânio diferentes

$$\text{Estatística F} = \frac{24 / 2}{6 / 6} = 12$$

Tabela de ANOVA

Variação	Soma dos Quadrados	Graus de Liberdade	Quadrados das medias	F
Entre Grupos (SQE)	24	2	12	12
Dentro dos grupos (SDQ)	6	6	1	
Total (SQT)	30	8		

TABLE E

F critical values

		Degrees of freedom in the numerator									
p		1	2	3	4	5	6	7	8	9	
Degrees of freedom in the denominator	1	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
		.050	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
		.025	647.79	799.50	864.16	899.58	921.85	937.11	948.22	956.66	963.28
		.010	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4	5981.1	6022.5
		.001	405284	500000	540379	562500	576405	585937	592873	598144	602284
	2	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
		.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
		.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
		.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
		.001	998.50	999.00	999.17	999.25	999.30	999.33	999.36	999.37	999.39
	3	.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
		.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
		.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
		.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
		.001	167.03	148.50	141.11	137.10	134.58	132.85	131.58	130.62	129.86
	4	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
		.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
		.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
		.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
		.001	74.14	61.25	56.18	53.44	51.71	50.53	49.66	49.00	48.47
	5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
		.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
		.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68
		.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
		.001	47.18	37.12	33.20	31.09	29.75	28.83	28.16	27.65	27.24
	6	.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
		.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
		.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52
		.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
		.001	35.51	27.00	23.70	21.92	20.80	20.03	19.46	19.03	18.69
	7	.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
		.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
		.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82
		.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
		.001	29.25	21.69	18.77	17.20	16.21	15.52	15.02	14.63	14.33

