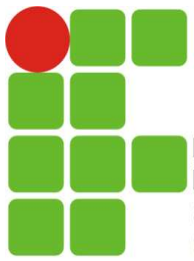


INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
SUL-RIO-GRANDENSE
Campus Passo Fundo

ANÁLISE DE VARIÂNCIA - COMPARAÇÃO ENTRE VÁRIOS GRUPOS

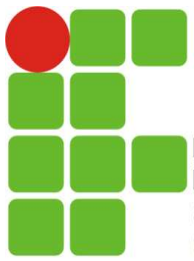
Introdução à Pesquisa Experimental

Professora Sabrina Elicker Hagemann



ANÁLISE DE VARIÂNCIA

- A ideia, na análise de variância, é comparar a **variância devida aos tratamentos** com a **variação devida ao acaso**, ou resíduo.
- Os fatores propostos podem ser variáveis quantitativas ou qualitativas, enquanto que a variável dependente é uma quantitativa e é observada dentro das classes dos fatores,
- O objetivo da análise de variância é analisar as diferenças entre as médias aritméticas dos grupos, a partir de uma análise na variação dos dados, entre os grupos. Ou seja, toma-se a variação total e subdivide-se a em **variação entre os grupos** e a **variação dentro do grupo**, a qual considera –se como um **erro experimental**, mas se a variação ocorrer **entre os grupos** ela é atribuída ao efeito do tratamento recebido.



ANÁLISE DE VARIÂNCIA

- **Teste F:** utilizado para comparar a variância entre grupos e a variância dentro do grupo.

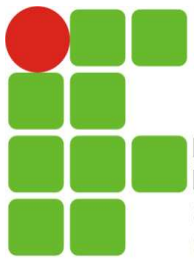


De um modo geral, os testes estatísticos trabalham com duas hipóteses:

- H_0 – hipótese nula
- H_1 – hipótese alternativa

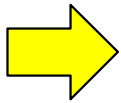
O ponto de partida de um teste estatístico consiste em considerar **a H_0 como sendo a hipótese verdadeira**. Dependendo do resultado do teste, pode-se cometer dois tipos diferentes de erro:

- **Erro tipo I – rejeitar H_0 sendo H_0 verdadeira**
- **Erro tipo II – Não rejeitar H_0 sendo H_0 falsa**



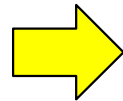
ANÁLISE DE VARIÂNCIA

α



Nível de significância de um teste estatístico e corresponde à probabilidade de cometer o erro do tipo I, ou seja, rejeitar a hipótese nula sendo ela verdadeira.

Valor - p

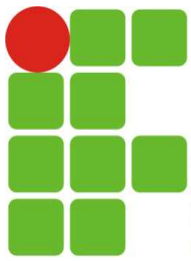


o menor valor do teste para rejeitar H_0



- Se $p\text{-valor} < \alpha$, então H_0 é rejeitada
- Se $p\text{-valor} > \alpha$, então H_0 não é rejeitada

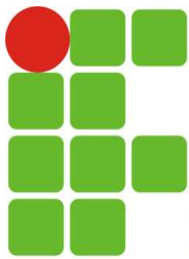
Se F calculado for maior que F tabelado, (ou $\text{valor-p} < \alpha$), descarta-se H_0 , ou seja, existe diferenças significativas entre os grupos provocada pelo fator controlável em estudo.



COMPARAÇÃO ENTRE GRUPOS

- **Experimentos que envolvem:** Variável de resposta e um fator controlável a vários níveis
- **Objetivo:** Identificar se os valores da variável de resposta medidos nos diversos níveis diferem entre si.
- **Disposição dos dados:**

Fator A		A1	A2	...	Ak	
		y ₁₁	y ₁₂	...	y _{1k}	
		y ₂₁	y ₂₂	...	y _{2k}	
		:	:	y _{ij}	:	
		:	:	...	:	
		y _{n1,1}	y _{n2,2}	...	y _{nk,k}	
Totais	T _{.j}	T _{.1}	T _{.2}	...	T _{.k}	T _{..} =
No.Obs.	n _j	n ₁	n ₂	...	n _k	N =
Médias	$\bar{Y}_{.j}$	$\bar{Y}_{.1}$	$\bar{Y}_{.2}$...	$\bar{Y}_{.k}$	$\bar{\bar{Y}}_{..}$ =



EXEMPLO

Um profissional deseja estudar se a temperatura de cura influencia na resistência à compressão de corpos de prova de concreto. Para isso, avaliou 3 temperaturas de curas diferentes.

	Temperatura (°C)		
	15	20	25
Resistências (MPa)	12	16	18
	13	18	20
	11	17	19

→ Fator controlável

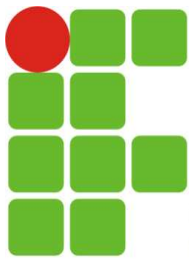
→ Níveis do fator controlável

Repetições

Modelo Estatístico:

$$Y_{ij} = \mu + \tau_j + \varepsilon_{ij}$$

onde: μ é a média geral;
 τ_j é o efeito do grupo j ;
 ε_{ij} é um erro aleatório.



EXEMPLO

Hipóteses:

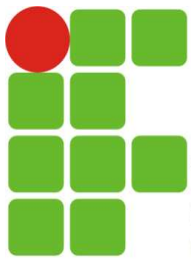
H_0 : não há diferenças significativas entre os grupos;

H_1 : há diferenças significativas entre os grupos provocada pelo fator controlável investigado

	Temperatura (°C)			
	15	20	25	
Resistências (MPa)	12	16	18	
	13	18	20	
	11	17	19	
$T_j =$	36	51	57	$T = 144$
$N =$	3	3	3	$N = 9$
$\hat{Y}_j =$	12	17	19	$\hat{Y} = 16$

Totais e média de cada grupo

Totais e média de todos os grupos



EXEMPLO

	Temperatura (°C)			
	15	20	25	
Resistências (MPa)	12	16	18	
	13	18	20	
	11	17	19	
T _j =	36	51	57	T _{..} =144
N=	3	3	3	N=9
Ŷ _j =	12	17	19	Ŷ _{..} =16

Modelo Estatístico:

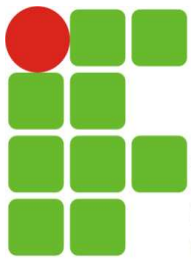
$$Y_{ij} = \mu + \tau_j + \varepsilon_{ij}$$

$$18 = 16 + 1 + 1$$

Média geral

$$\text{Média do grupo} - \text{média geral} = 17 - 16 = 1$$

$$\text{Valor individual} - \text{média do grupo} = 18 - 17 = 1$$



EXEMPLO

	Temperatura (°C)			
	15	20	25	
Resistências (MPa)	12	16	18	
	13	18	20	
	11	17	19	
T _j =	36	51	57	T=144
N=	3	3	3	N=9
Ŷ _j =	12	17	19	Ŷ=16

Modelo Estatístico:

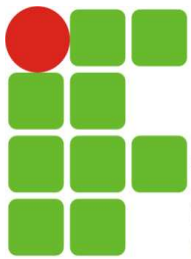
$$Y_{ij} = \mu + \tau_j + \varepsilon_{ij}$$

$$20 = 16 + 3 + 1$$

Média geral

$$\text{Média do grupo} - \text{média geral} = 19 - 16 = 3$$

$$\text{Valor individual} - \text{média do grupo} = 20 - 19 = 1$$



EXEMPLO

Decomposição dos resíduos:

$$(Y_{ij} - \bar{Y}_{..}) = (\bar{Y}_{.j} - \bar{Y}_{..}) + (Y_{ij} - \bar{Y}_{.j})$$

Diagram illustrating the decomposition of residuals. The equation is shown with arrows pointing from terms to their respective labels:

- Y_{ij} (Valor da repetição)
- $\bar{Y}_{..}$ (Média geral)
- $\bar{Y}_{.j}$ (Média do grupo)
- $\bar{Y}_{..}$ (Média geral)
- Y_{ij} (Valor da repetição)
- $\bar{Y}_{.j}$ (Média do grupo)

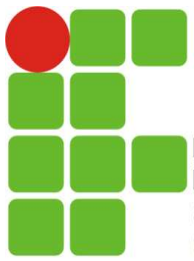
Elevando ao
quadrado e somando:

$$\sum (Y_{ij} - \bar{Y}_{..})^2 = \sum n (\bar{Y}_{.j} - \bar{Y}_{..})^2 + \sum (Y_{ij} - \bar{Y}_{.j})^2$$

$$SQT = SQG + SQR$$

Graus de Liberdade:

$$(N - 1) = (K - 1) + (N - K)$$



EXEMPLO

Médias quadradas: $MQG = SQG / (K - 1)$

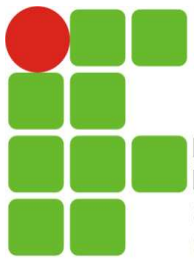
$$MQR = SQR / (N - K)$$

Se não há diferenças significativas entre os grupos: $E [MQG] = E [MQR]$

O teste F compara as
duas variâncias:

$$F_{calc} = \frac{\text{Variância entre grupos}}{\text{Variância dentro do grupo}} = \frac{MQG}{MQR}$$

Comparar **F calculado** com **F tabelado**, se o valor calculado for maior que o valor tabelado (ou valor-p <0,05), descarta-se H_0 , ou seja, existe diferenças significativas entre os grupos provocada pelo fator controlável em estudo.



EXEMPLO

Fórmulas:

$$TC = T_{..}^2 / N \longrightarrow \text{Termo de correção}$$

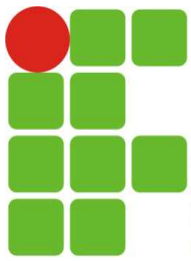
$$SQT = \sum(Y_{ij}^2) - TC$$

$$SQG = \sum(T_{.j}^2 / n_j) - TC$$

$$SQR = SQT - SQG$$

Tabela ANOVA:

Fonte	SQ	GDL	MQ	Teste F
Entre Grupos	SQG	K - 1	MQG	F = MQG / MQR
Dentro Grupos	SQR	N - K	MQR	
Total	SQT	N - 1		



EXEMPLO

O limite de decisão é estabelecido usando os valores tabelados da distribuição F , ou seja:

$$F_{\alpha, k-1, N-k}$$

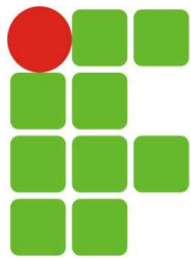
Onde:

α : nível de significância
(usualmente 0,05)

$k-1$: graus de liberdade do numerador (MQG)

$N-k$: graus de liberdade do denominador (MQR)

		Nível de significância - alfa = 0,05								
		1	2	3	4	5	6	7	8	9
n1 \ n2	n2									
1	1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5
2	2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38
3	3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81
4	4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00
5	5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77
6	6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10



BIBLIOGRAFIA

MONTGOMERY, Douglas. C. Design and Analysis of Experiments. 5. ed. Nova Iorque: John Wiley & Sons, 2001.

RIBEIRO, J.L.D.; CATEN, C.S.ten. Projeto de experimentos. Porto Alegre: Universidade Federal do Rio Grande do Sul, Escola de Engenharia, Programa de Pós-Graduação em Engenharia de Produção, 2011.