

Testes Estatísticos e Nível de Significância

João Carlos Xavier Júnior

jcxavier@imd.ufn.br

Introdução

❑ Testes paramétricos:

- ❖ Exigem que as amostras tenham uma distribuição normal;
- ❖ Também são chamados de testes t;
- ❖ Não se pode garantir que as amostras possuam uma distribuição normal de valores;
- ❖ Indicado para amostras grandes (≥ 30).

Introdução

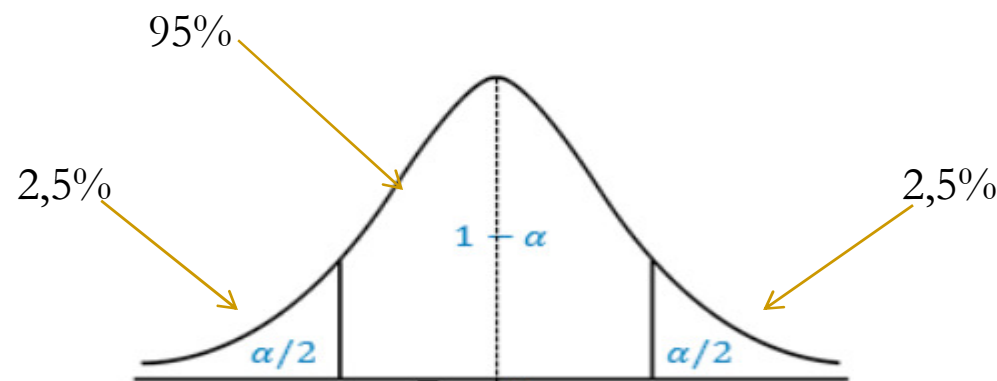
□ Testes não-paramétricos:

- ❖ Não exigem requisitos tão fortes (distribuição normal);
- ❖ São indicados para amostras pequenas (< 30);
- ❖ São mais indicados para a área da Ciências dos Dados;
- ❖ São não tão potentes quanto os paramétricos;

Nível de Significância

❑ Nível de Significância (NS):

- ❖ Tipo de estimativa por intervalo de um parâmetro populacional desconhecido.
- ❖ Podem ser mostrados em vários níveis de confiança como 90%, 95% e 99%.
- ❖ Porém, 95% é o mais comum.



Teste de Hipóteses

❑ Hipótese nula (H_0):

- ❖ Hipótese tida como verdadeira até que provas estatísticas indiquem o contrário.
- ❖ Em geral, consiste em afirmar que os parâmetros ou características matemáticas de duas ou mais populações são idênticos.
- ❖ Exemplo: “a média das alturas da cidade A é igual à da cidade B” ou não há diferença significativa.

❑ Hipótese alternativa (H_A):

- ❖ Deve ser contrária, oposta, antagônica à hipótese nula.
- ❖ Também designada por H_1 .

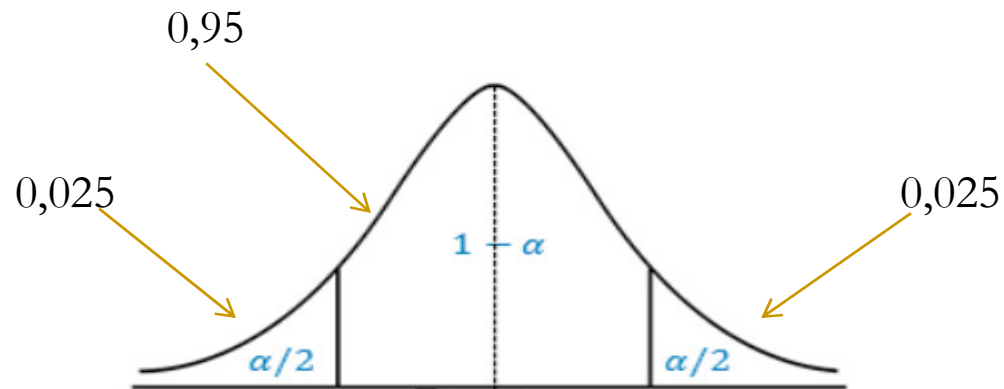
Decisão do Teste

□ p-Value:

- ❖ Valor de significância do resultado.
- ❖ Utilizado para aceitar ou rejeitar uma hipótese nula.

$$\text{p-value} = 0,025 + 0,025 \Rightarrow 0,05$$

$$\text{p-value} < 0,05 \Rightarrow \text{rejeitar } H_0$$



Decisão do Teste

□ p-Value:

p-value	Descrição
$> 0,05$	Não significativa
0,01 até 0,05	Significante
0,001 até 0,005	Muito significativa
$< 0,001$	Extremamente significativa

Testes não Paramétricos

- Segue os seguintes passos:
 - ❖ Formular as hipóteses (nula $\Rightarrow \mathbf{H}_0$ ou alternativa);
 - ❖ Estabelecer a estatística (estimador) a ser utilizada;
 - ❖ Fixar o nível de significância do teste;
 - ❖ Calcular a estatística teste (a estimativa).
 - ❖ Tomar a decisão:
 - Se o valor da estatística observada na amostra estiver na região crítica rejeitar \mathbf{H}_0 , caso contrário aceitar \mathbf{H}_0 .

Testes não Paramétricos

Nível de mensuração	TESTES ESTATÍSTICOS NÃO-PARAMÉTRICOS					Medidas de correlação não-paramétricas
	Caso de uma amostra	Caso de duas Amostras		Caso de k amostras		
		Amostras relacionadas	Amostras independentes	Amostras relacionadas	Amostras independentes	
Nominal	Binomial e χ^2	McNemar	Fisher e χ^2	Q de Cochram	χ^2	De contingência
Ordinal	Kolmogorov-Smimov Iterações	Sinais <u>Wilcoxon</u>	Mediana U de Mann-Withney Kolmogorov-Smimov Iterações de Wald-Wolfowitz Moses	<u>Friedman</u>	Extensão da mediana Kruskal-Wallis	Por postos de Spearman Por postos de Kendall Parcial de postos de Kendall Concordância de Kendall
Intervalar		Walsh Aleatoriedade	Aleatoriedade			

Teste de Friedman

- Teste de Friedman:
 - ❖ Indicado para mais de duas amostras.
 - ❖ Pretende testar se as diferentes amostras provêm de uma mesma população ou de populações diferentes.
- O teste de Friedman também utiliza as duas hipóteses:
 - ❖ H_0 : as distribuições das k amostras são idênticas;
 - ❖ H_A : as distribuições das k amostras diferem.

Teste de Friedman

Treatment 1	Treatment 2	Treatment 3	Treatment 4	Treatment 5
0.3191	0.3467	0.2936	0.2882	
0.3002	0.323	0.2851	0.2739	
0.3094	0.3377	0.2844	0.2715	
0.3092	0.3158	0.2841	0.2938	
0.3104	0.3206	0.2677	0.3244	
0.3119	0.2976	0.2799	0.2829	
0.1441	0.0938	0.1736	0.0947	
0.2103	0.1914	0.1802	0.1236	
0.1962	0.1727	0.1558	0.1631	
0.0859	0.0153	0.0765	0.0014	
0.0857	0.0188	0.0773	0.0015	
0.0857	0.0196	0.0772	0.0016	
0.1599	0.2863	0.3126	0.1901	
0.1468	0.3524	0.3273	0.1278	
0.1704	0.3125	0.2713	0.1533	

Significance Level:

☐ .01

☒ .05

☐ .10

<https://www.socscistatistics.com/tests/friedman/Default.aspx>

Teste de Friedman

Treatment 1	Treatment 2	Treatment 3	Treatment 4
0.3191	0.3467	0.2936	0.2882
0.3002	0.323	0.2851	0.2739
0.3094	0.3377	0.2844	0.2715
0.3092	0.3158	0.2841	0.2938
0.3104	0.3206	0.2677	0.3244
0.3119	0.2976	0.2799	0.2829
0.1441	0.0938	0.1736	0.0947
0.2103	0.1914	0.1802	0.1236
0.1962	0.1727	0.1558	0.1631
0.0859	0.0153	0.0765	0.0014
0.0857	0.0188	0.0773	0.0015
0.0857	0.0196	0.0772	0.0016
0.1599	0.2863	0.3126	0.1901
0.1468	0.3524	0.3273	0.1278
0.1704	0.3125	0.2713	0.1533

Ranks T1	Ranks T2	Ranks T3	Ranks T4
3	4	2	1
3	4	2	1
3	4	2	1
3	4	1	2
2	3	1	4
4	3	1	2
3	1	4	2
4	3	2	1
4	3	1	2
4	2	3	1
4	2	3	1
4	2	3	1
1	3	4	2
2	4	3	1
2	4	3	1
Sum: 46	Sum: 46	Sum: 35	Sum: 23

Teste de Friedman

□ Estatística de Friedman:

$$\chi^2 = \frac{12}{nk(k+1)} \sum_{j=1}^k R_j^2 - 3n(k+1), \text{ onde}$$

n = número de linhas,

k = número de colunas,

R_j = soma dos postos da coluna j

The χ^2_r statistic is 14.44 (3, $N = 15$).

The p -value is .00236.

The result is significant at $p < .05$.

Teste de Friedman

O que isso significa?

The χ^2_r statistic is 14.44 (3, $N = 15$).

The p -value is .00236.

The result is significant at $p < .05$.

Teste de Friedman

O que isso significa?

The χ^2_r statistic is 14.44 (3, $N = 15$).

The p -value is .00236.

The result is significant at $p < .05$.

Você vai precisar de um teste Pos-Hoc

<https://www.rdocumentation.org/packages/PMCMR/versions/4.3/topics/posthoc.kruskal.nemenyi.test>

Teste de Friedman

□ Teste Pos-hoc:

Config	Base	K-NN	J48	NB	MLP
10 fold	Credit	0,3191	0,3467	0,2936	0,2882
70/30	Credit	0,3002	0,323	0,2851	0,2739
60/40	Credit	0,3094	0,3377	0,2844	0,2715
10 fold	Diabetes	0,3092	0,3158	0,2841	0,2938
70/30	Diabetes	0,3104	0,3206	0,2677	0,3244
60/40	Diabetes	0,3119	0,2976	0,2799	0,2829
10 fold	Ionosphere	0,1441	0,0938	0,1736	0,0947
70/30	Ionosphere	0,2103	0,1914	0,1802	0,1236
60/40	Ionosphere	0,1962	0,1727	0,1558	0,1631
10 fold	Nursery	0,0859	0,0153	0,0765	0,0014
70/30	Nursery	0,0857	0,0188	0,0773	0,0015
60/40	Nursery	0,0857	0,0196	0,0772	0,0016
10 fold	Sonar	0,1599	0,2863	0,3126	0,1901
70/30	Sonar	0,1468	0,3524	0,3273	0,1278
60/40	Sonar	0,1704	0,3125	0,2713	0,1533
Média		0,2097	0,2269	0,2231	0,1728
Desvio Padrão		0,0922	0,1301	0,0909	0,1141

Nemenyi test

	K-NN	J48	NB
J48	1,0000	-	-
NB	0,4043	0,4043	-
MLP	0,0063	0,0063	0,3252



Qual o Melhor Modelo?

□ Resultados

Config	Base	K-NN	J48	NB	MLP
10 fold	Credit	0,3191	0,3467	0,2936	0,2882
70/30	Credit	0,3002	0,323	0,2851	0,2739
60/40	Credit	0,3094	0,3377	0,2844	0,2715
10 fold	Diabetes	0,3092	0,3158	0,2841	0,2938
70/30	Diabetes	0,3104	0,3206	0,2677	0,3244
60/40	Diabetes	0,3119	0,2976	0,2799	0,2829
10 fold	Ionosphere	0,1441	0,0938	0,1736	0,0947
70/30	Ionosphere	0,2103	0,1914	0,1802	0,1236
60/40	Ionosphere	0,1962	0,1727	0,1558	0,1631
10 fold	Nursery	0,0859	0,0153	0,0765	0,0014
70/30	Nursery	0,0857	0,0188	0,0773	0,0015
60/40	Nursery	0,0857	0,0196	0,0772	0,0016
10 fold	Sonar	0,1599	0,2863	0,3126	0,1901
70/30	Sonar	0,1468	0,3524	0,3273	0,1278
60/40	Sonar	0,1704	0,3125	0,2713	0,1533
Média		0,2097	0,2269	0,2231	0,1728
Desvio Padrão		0,0922	0,1301	0,0909	0,1141

Nemenyi test

	K-NN	J48	NB
J48	1,0000	-	-
NB	0,4043	0,4043	-
MLP	0,0063	0,0063	0,3252



Qual o Melhor Modelo?

- ❑ Salvando o melhor Modelo:

The screenshot shows a software interface for model evaluation. On the left, there is a panel with a dropdown menu labeled "(Nom) class", "Start" and "Stop" buttons, and a "Result list (right-click for options)" section. A context menu is open over the result list, with "Save model" highlighted. A red arrow points to this option. The main area displays "Node 1" and "Time taken to build model: 0.77 seconds". Below this is a "Stratified cross-validation Summary" table. A red arrow points to the "0.2938" value in the table, which represents the best model's performance.

Stratified cross-validation Summary			
Instances	577		
Instances	191		
	0.4445		
	0.2938		
	0.4236		
	64.6434 %		
	88.8752 %		
	768		

Confusion Matrix			
FP Rate	Precision	Recall	
0,174	0,653	0,612	
0,388	0,799	0,826	
0,313	0,748	0,751	

Qual o Melhor Modelo?

- ❑ Salvando o melhor Modelo:



Qual o Melhor Modelo?

- ❑ Salvando o segundo melhor Modelo:

