



# Teste U

## Teste de Mann-whitney

Karla szczypkovski Silva  
Lilian Sayuri Sakamoto



# Testes Não-paramétricos





## ➤ VANTAGENS DOS MÉTODOS NÃO-PARAMÉTRICOS

1. Aplicado a uma grande variedade de situações ;
2. Não exige populações normalmente distribuídas;
3. Podem ser aplicados a dados categóricos;
4. Cálculos mais simples.

## ➤ DESVANTAGENS DOS MÉTODOS NÃO-PARAMÉTRICOS

1. Tendem a desperdiçar informações;
2. Evidências mais fortes para rejeitar a hipótese nula.

# Teste de Mann-Whitney

- O teste de Mann-Whitney é usado para a comparação de dois grupos independentes.
- Principalmente quando não se sabe a distribuição segue o padrão normal , ou os dados não são homogêneos.
- O teste U seria um alternativa equivalente ao teste T.



Frank Wilcoxon



Henry Berthold Mann



Donald Ransom Whitney.



# Testes de Mann-whitney ou Teste U

- › Tem por função comprovar se os dados de dois grupos independentes foram ou não extraídos de uma mesma população.
- › O conjunto de dados não necessita ter a mesma dimensão.
- › Para isso são comparadas as medianas entre as populações.

# Hipóteses do Teste

› As hipóteses do teste podem ser bilaterais quando:

$$H_0: \mu_X = \mu_Y;$$

$$H_1: \mu_X \neq \mu_Y.$$

# Hipóteses do Teste :

› Ou também podem ser unilaterais :

$$H_1 : \mu_X < \mu_Y \quad \text{ou} \quad H_1 : \mu_X > \mu_Y.$$

# Como montar o teste :

- › A eficácia de publicidade para 2 determinados produtos concorrentes ( Marca X e Marca Y). Uma pesquisa de mercado realizado em um centro comercial em que depois de degustar o produto cada participante deveria dar uma nota de 1 a 10.

Marca X	Marca Y
3	9
4	7
2	5
6	10
2	6
5	8





# Primeiro passo:



- › Formular Hipóteses :
- ›  $H_0$  – As medianas populacionais são iguais ,ou seja, não há diferença entre os duas marcas de café.
- ›  $H_1$ - As medianas populacionais não são iguais, ou seja, há diferença entre as duas marcas de café.

## Segundo Passo :

- › Se deve ordenar a amostra de forma ordinal atribuindo um posto (ranking) para cada valor, mas não esquecendo de atribuir de onde vem cada dado.

Valor	Posto	Marca
2	1	x
2	2	x
3	3	x
4	4	x
5	5	y
5	6	x
6	7	y
6	8	x
7	9	y
8	10	y
9	11	y
10	12	y

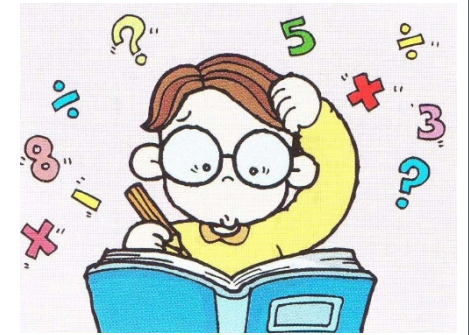




Valor	Posto
2	1
2	2
3	3
4	4
5	5
5	6
6	7
6	8
7	9
8	10
9	11
10	12

Como podemos observar na tabela temos valores repetidos para os postos.

O que fazer ?



Nesse caso devemos tirar uma média os postos e então atribuir o valor a esse novo posto

# Tabela final

Valor	Posto	Marca
2	1,5	x
2	1,5	x
3	3	x
4	4	x
5	5,5	y
5	5,5	x
6	7,5	y
6	7,5	x
7	9	y
8	10	y
9	11	y
10	12	y



# Calcular R1 e R2

- › R1 e R2 é nada mais que a somatória dos postos.
- › R1- Marca X.
- › R2- Marca Y.

Posto	Valor
1,5	2
1,5	2
3	3
4	4
5,5	5
7,5	6
23	

Posto	Valor
5,5	5
7,5	6
9	7
10	8
11	9
12	10
55	

# Calcular a estatística do teste

› Teste U :

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1,$$

$$U = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

› **N1** - é o número total de postos no grupo 1.

› **N2** - é o número total de postos no grupo 2.

## Calculando a estatística do teste:

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 = 6 \cdot 6 + \frac{6 \cdot 7}{2} - 23 = 34,$$

$$U = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 = 6 \cdot 6 + \frac{6 \cdot 7}{2} - 55 = 2$$



- › Escolher o menor valor de U
- › Para valores em que  $n < 20$  utilizamos a tabela de U do teste de Mann-whitney em que o para achar o U crítico.

# Tomada de Decisão



- › Tomar decisão no nível de significância que ira ser utilizado, isso fica a critério do pesquisador.
- › No nosso caso o valor calculado  $U = 2$  é menor aos valores da tabela.
- › Aqui utilizamos o valor de  $\alpha = 5\%$  para buscar a decisão se aceita ou rejeita o teste, para  $\alpha(5\%) = 5$  na U Tabela. Nesse caso **rejeita-se** a hipótese nula de igualdade entre as medianas populacionais .



## Critical Values of the Mann-Whitney U (Two-Tailed Testing)

$n_2$	$\alpha$	$n_1$												
		3	4	5	6	7	8	9	10	11	12	13	14	15
3	.05	--	0	0	1	1	2	2	3	3	4	4	5	5
	.01	--	0	0	0	0	0	0	0	0	1	1	1	2
4	.05	--	0	1	2	3	4	4	5	6	7	8	9	10
	.01	--	--	0	0	0	1	1	2	2	3	3	4	5
5	.05	0	1	2	3	5	6	7	8	9	11	12	13	14
	.01	--	--	0	1	1	2	3	4	5	6	7	7	8
6	.05	1	2	3	5	6	8	10	11	13	14	16	17	19
	.01	--	0	1	2	3	4	5	6	7	9	10	11	12
7	.05	1	3	5	6	8	10	12	14	16	18	20	22	24
	.01	--	0	1	3	4	6	7	9	10	12	13	15	16



```
> marcax<-c(3,4,2,6,2,5)
> marcay<-c(9,7,5,10,6,8)
> wilcox.test(marcax,marcay)
```

wilcoxon rank sum test with continuity correction

data: marcax and marcay

W = 2, p-value = 0.01259

alternative hypothesis: true location shift is not equal to 0



# EXEMPLO 1

A tabela 1 inclui os escores de Facilidade de Leitura de Flesch para páginas selecionadas aleatoriamente de dois livros: **Harry Potter** e a Pedra Filosofal, de J.K Rowling, e **Guerra e Paz**, de Lev Tolstoi. Use os dois conjuntos independentes de dados amostrais da tabela 1, com um nível de significância de 0,05, para testar a afirmativa de que os escores de leitura dos dois livros têm a mesma distribuição.

TABELA 1 – Escores de Leituras		
	Harry Potter	Guerra e Paz
1	85,3	69,4
2	84,3	64,2
3	79,5	71,4
4	82,5	71,6
5	80,2	68,4
6	84,6	51,9
7	79,2	72,2
8	70,9	74,4
9	78,6	52,8
10	86,2	58,4
11	74,0	65,4
12	83,7	73,6
13	71,4	

- $H_0$ : Os dois livros têm escore de Facilidade de leitura de Flesch com a mesma distribuição.
- $H_1$ : As duas populações têm distribuição de escore de Facilidade de leitura de Flesch que são diferentes de alguma maneira.



TABELA FINAL

Livro	Valor	Posição	Posto
Guerra e paz	51,9	1	1
Guerra e paz	52,8	2	2
Guerra e paz	58,4	3	3
Guerra e paz	64,2	4	4
Guerra e paz	65,4	5	5
Guerra e paz	68,5	6	6
Guerra e paz	69,4	7	7
Harry Potter	70,9	8	8
Harry Potter	71,4	9	9,5
Guerra e paz	71,4	10	9,5
Guerra e paz	71,6	11	11
Guerra e paz	72,2	12	12
Guerra e paz	73,6	13	13
Harry Potter	74,0	14	14
Guerra e paz	74,4	15	15
Harry Potter	78,6	16	16
Harry Potter	79,2	17	17
Harry Potter	79,5	18	18
Harry Potter	80,2	19	19
Harry Potter	82,5	20	20
Harry Potter	83,7	21	21
Harry Potter	84,3	22	22
Harry Potter	84,6	23	23
Harry Potter	85,3	24	24
Harry Potter	86,2	25	25

› Amostra maior que vinte

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\sigma_T = \sqrt{\frac{n_2 \mu_T}{6}}$$

$$Z = \frac{T - \mu_T}{\sigma_T}$$

## › Dados

### - Harry Potter

›  $n_1$ : 13

›  $T_1$ : 236,5

### - Guerra e paz

›  $n_2$ : 12

›  $T_2$ : 88,5

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\mu_T = 169$$

$$\sigma_T = \sqrt{\frac{n_2 \mu_T}{6}}$$

$$\sigma_T = 18,385$$

$$Z = \frac{T - \mu_T}{\sigma_T}$$

$$Z = 3,67$$





# Mann Whitney U-test Calculator

<http://scistatcalc.blogspot.co.uk/2013/10/mann-whitney-u-test-calculator.html>

Group 1 values: (enter comma separated numbers)

85.3,84.3,79.5,82.5,80.2,84.6,79.2,70.9,78.6,86.2,74.0,83.7,71.4

Group 2 values: (enter comma separated numbers)

69.4,64.2,71.4,71.6,68.4,51.9,72.2,74.4,52.8,58.4,65.4,73.6

Clear Group 1

Clear Group 2

Escolher arquivo Nenhum arquivo selecionado

Perform Mann Whitney U-test

Sum of ranks for Group 1 is 236.5, for 13 samples

Sum of ranks for Group 2 is 88.5, for 12 samples

Value of U statistic is 10.5

For 5% two-tailed level, Critical value of U (from tables) is 41

As calculated  $U < U_{critical}$ , we reject the Null Hypothesis - i.e. the result is significant

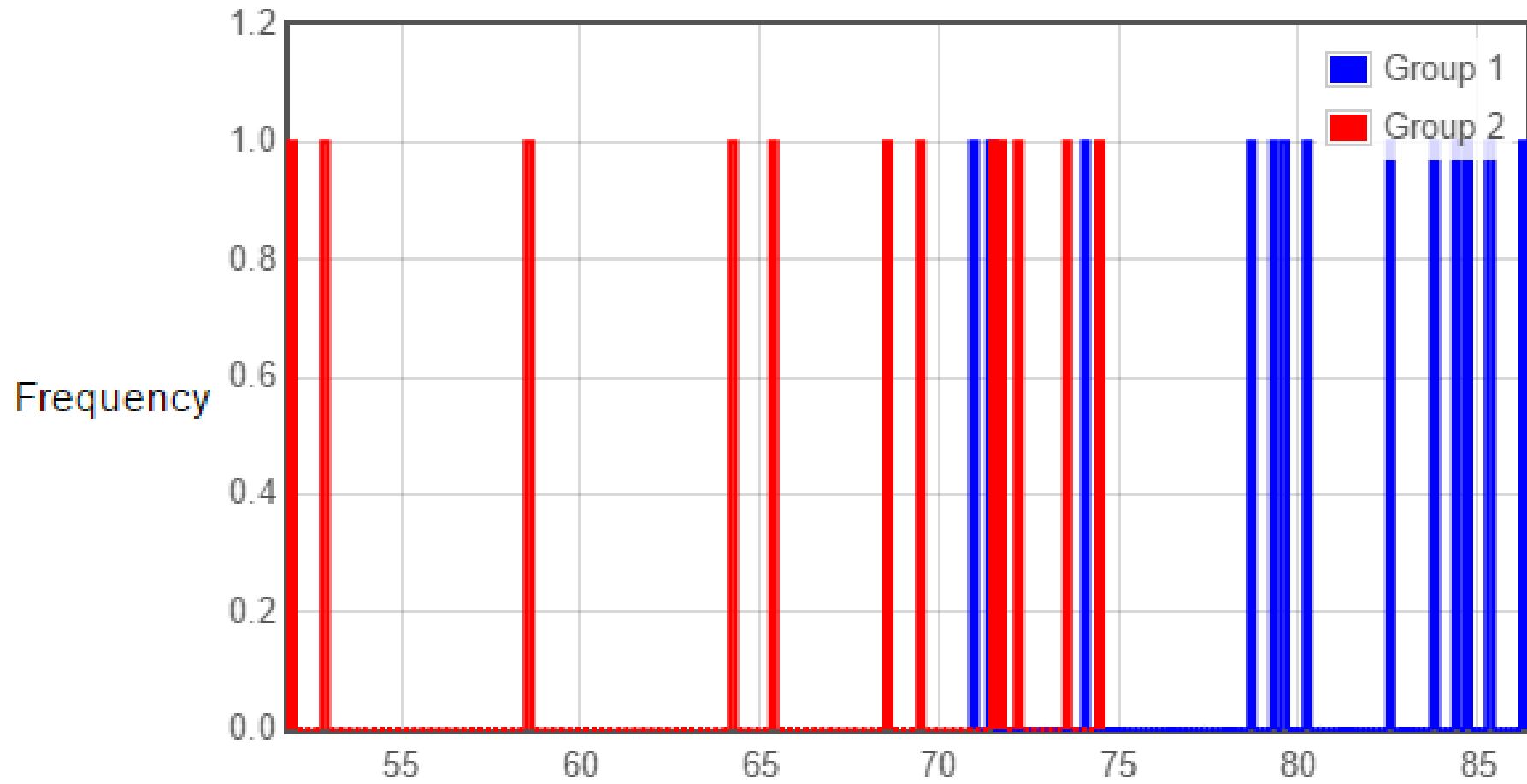
z score is 3.671516, will be quite inaccurate

z critical (5%, two-tailed) is 1.959964

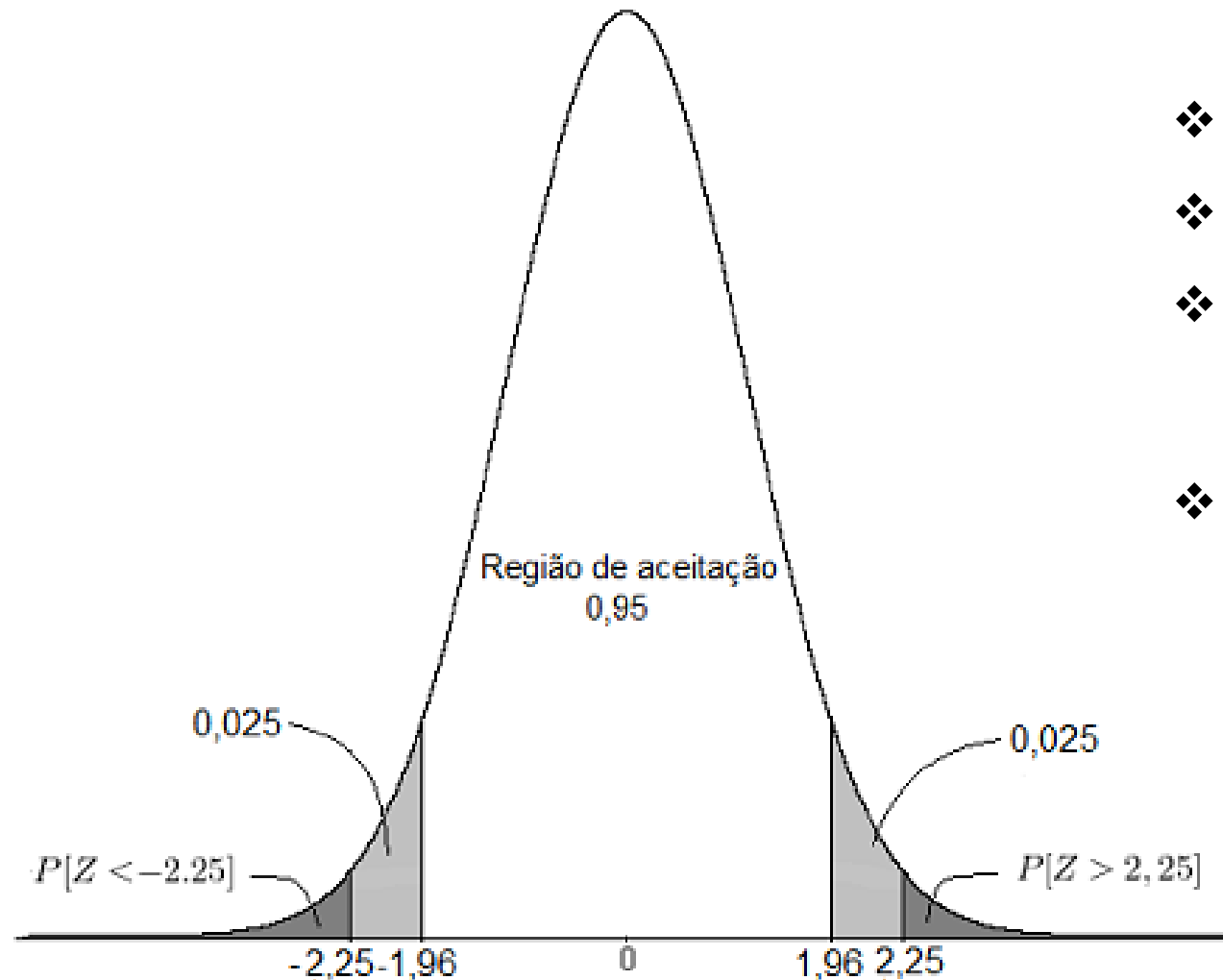
pvalue is 0.000241



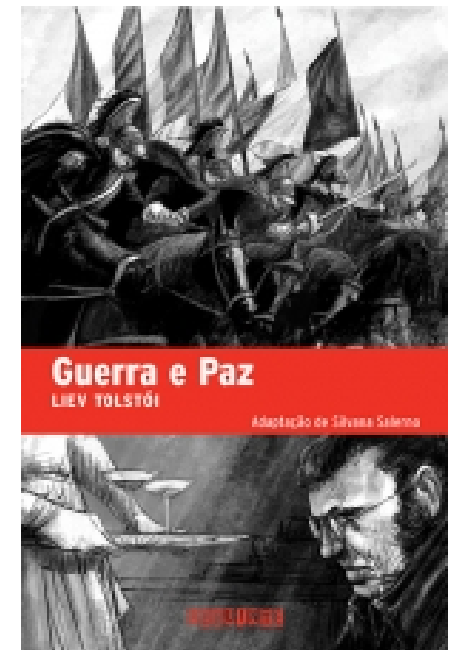
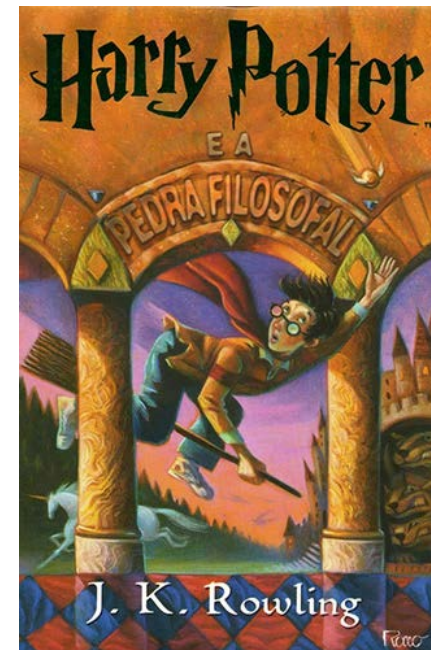
### Histograms



# CONCLUSÃO



- ❖ Valores críticos de  $Z= 1,96$  e  $-1,96$
- ❖  $Z= 3,67$
- ❖ Rejeita-se  $H_0$ .
  
- ❖ USAR O MENOR VALOR





# EXEMPLO 2 – IMC de Homens E mulheres

Os valores de IMC dos homens e das mulheres estão listados na tabela 2. Use o nível de significância de 0,05 para testar a afirmativa de que as medianas dos IMCs entre homens e mulheres são iguais.

TABELA 2 – MEDIDAS DE IMC		
	HOMENS	MULHERES
1	23,8	19,6
2	23,2	23,8
3	24,6	19,6
4	26,2	29,1
5	23,5	25,2
6	24,5	21,4
7	21,5	22,0
8	31,4	27,5
9	26,4	33,5
10	22,7	20,6
11	27,8	29,9
12	28,1	17,7
13	25,2	

- $H_0$ : Homens e mulheres tem valores de IMC com medianas iguais.
- $H_1$ : Homens e mulheres tem valores de IMC com medianas diferentes.



TABELA FINAL

	Valor	Posição	Posto
Mulher	17,7	1	1
Mulher	19,6	2	2,5
Mulher	19,6	3	2,5
Mulher	20,6	4	4
Mulher	21,4	5	5
Homem	21,5	6	6
Mulher	22,0	7	7
Homem	22,7	8	8
Homem	23,2	9	9
Homem	23,5	10	10
Homem	23,8	11	11,5
Mulher	23,8	12	11,5
Homem	24,5	13	13
Homem	24,6	14	14
Homem	25,2	15	15,5
Mulher	25,2	16	15,5
Homem	26,2	17	17
Homem	26,4	18	18
Mulher	27,5	19	19
Homem	27,8	20	20
Homem	28,1	21	21
Mulher	29,1	22	22
Mulher	29,9	23	23
Homem	31,4	24	24
Mulher	33,5	25	25

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\sigma_T = \sqrt{\frac{n_2 \mu_T}{6}}$$

$$Z = \frac{T - \mu_T}{\sigma_T}$$



## Dados

• HOMENS $n_1: 13$  $T_1: 187$ 

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\mu_T = 169$$

• MULHERES $n_2: 12$  $T_2: 138$ 

$$\sigma_T = \sqrt{\frac{n_2 \mu_T}{6}}$$

$$\sigma_T = 18,385$$



$$Z = \frac{T - \mu_T}{\sigma_T}$$

$$Z = 0,98$$



# Mann Whitney U-test Calculator

Group 1 values: (enter comma separated numbers)

23.8,23.2,24.6,26.2,23.5,24.5,21.5,31.4,26.4,22.7,27.8,28.1,25.2

Group 2 values: (enter comma separated numbers)

19.6,23.8,19.6,29.1,25.2,21.4,22.0,27.5,33.5,20.6,29.9,17.7

Clear Group 1

Clear Group 2

Escolher arquivo Nenhum arquivo selecionado

Perform Mann Whitney U-test

Sum of ranks for Group 1 is 187, for 13 samples

Sum of ranks for Group 2 is 138, for 12 samples

Value of U statistic is 60

For 5% two-tailed level, Critical value of U (from tables) is 41

As calculated  $U \geq U$  critical, we accept the Null Hypothesis

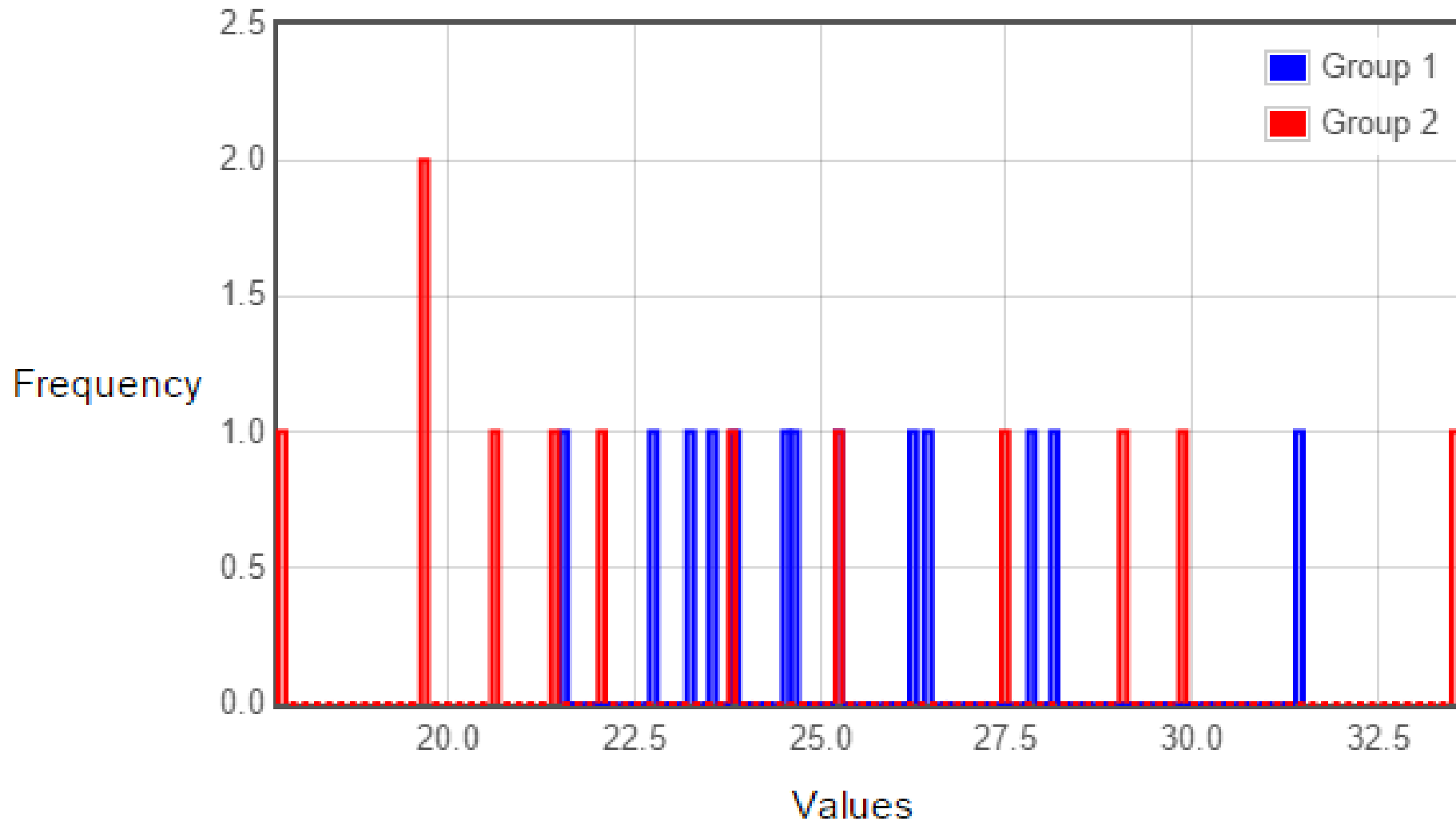
z score is 0.979071, will be quite inaccurate

z critical (5%, two-tailed) is 1.959964

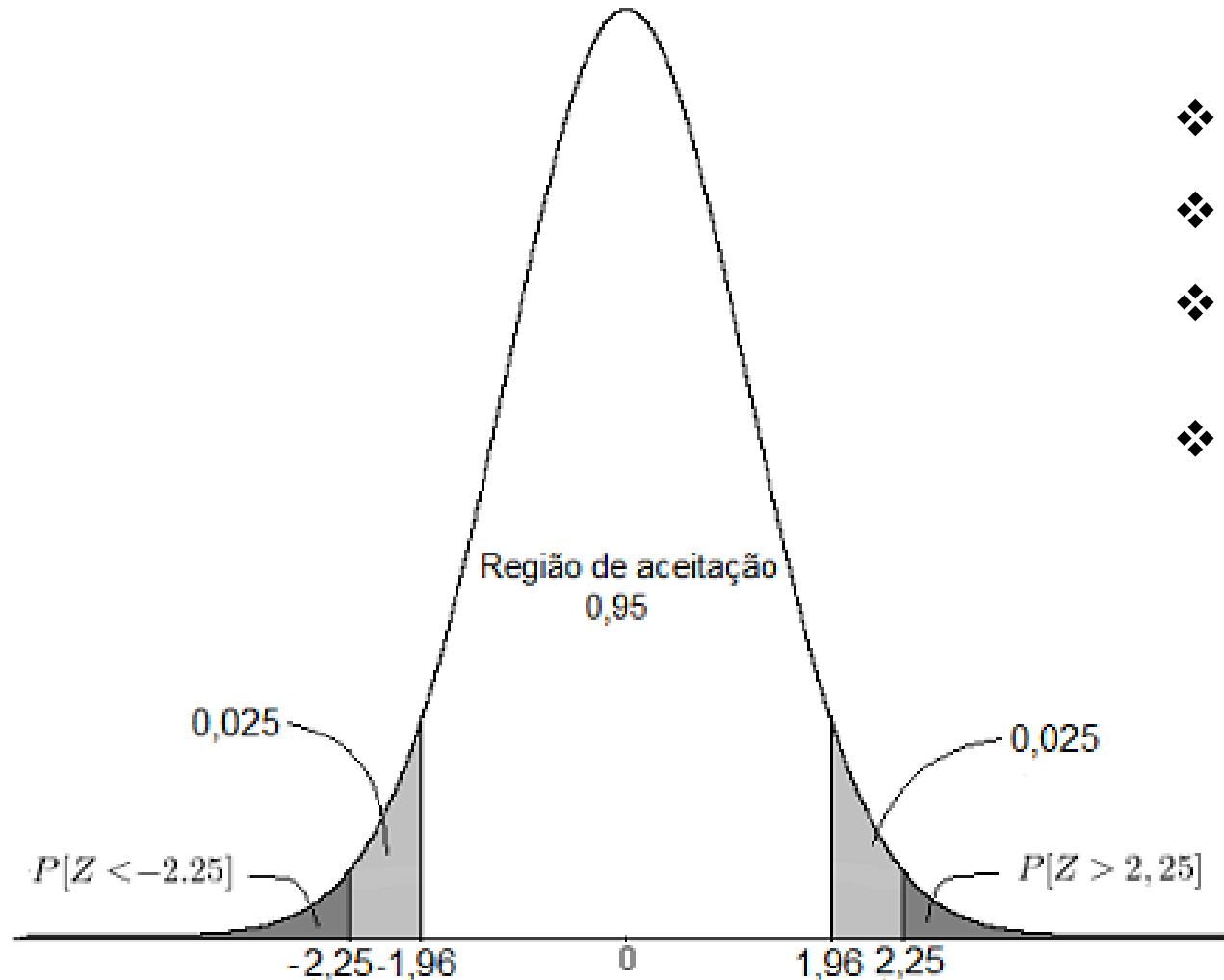
pvalue is 0.327545



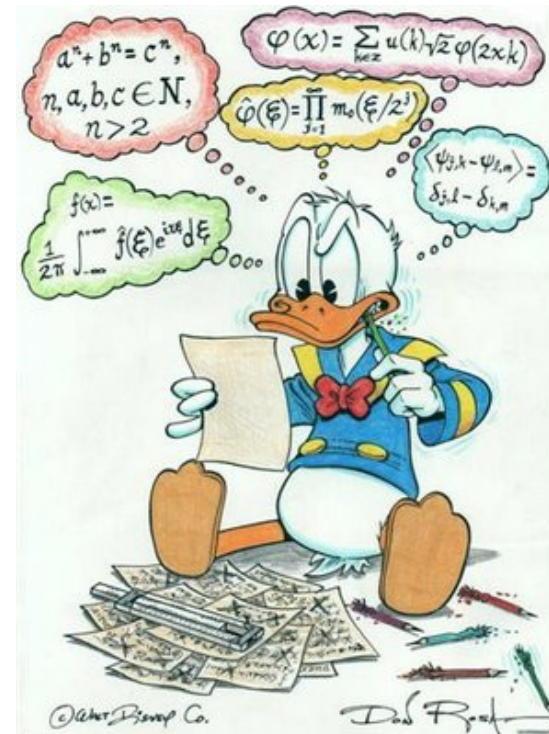
# Histograms



# CONCLUSÃO



- ❖ Valores críticos de  $Z= 1,96$  e  $-1,96$
- ❖  $Z= 0,98$
- ❖  $H_0$  Não pode ser rejeitado
- ❖ USAR O MENOR VALOR



# REFERÊNCIAS

- **Critical Values of the Mann-Whitney U.** Disponível em: <[http://www.lesn.appstate.edu/olson/stat\\_directory/Statistical%20procedures/Mann\\_Whitney%20U%20Test/Mann-Whitney%20Table.pdf](http://www.lesn.appstate.edu/olson/stat_directory/Statistical%20procedures/Mann_Whitney%20U%20Test/Mann-Whitney%20Table.pdf)>. acesso em: 27/05/2015.
- **Teste de Man-Whitney .** Disponível em: <<http://leg.ufpr.br/~silvia/CE055/node95.html>>. acesso em : 26/05/2015.
- **Teste de Man-Whitney, UFSC.** Disponível em: <[http://www.inf.ufsc.br/~verav/Testes\\_de\\_Hipoteses/Testes\\_nao\\_parametricos\\_Mann-Whitney.pdf](http://www.inf.ufsc.br/~verav/Testes_de_Hipoteses/Testes_nao_parametricos_Mann-Whitney.pdf)>. Acesso em: 26/05/2015.
- **Testes Não Paramétricos, FGV.** Disponível em: <<http://epge.fgv.br/we/Graduacao/Estatistica1/2009/2?action=AttachFile&do=get&target=teste-dos-sinais-wilcoxon-e-mann-whitney.pdf>>. acesso em: 26/05/2015.
- **Scistatcalc: Mann Whitney U-test Calculator.** Disponível em: <<http://scistatcalc.blogspot.co.uk/2013/10/mann-whitney-u-test-calculator.html>>. acesso em : 24/05/2015.
- **TRIOLA, Mario F. Introdução a estatística. 9ª Edição.** Rio de Janeiro : LTC, 2005.
- **TRIOLA, Mario F. Introdução a estatística. 10ª Edição.** Rio de Janeiro : LTC, 2008.

# OBRIGADA

