



Estatística Não Paramétrica

Prof. Adriana Andrade

Departamento de Matemática

andrade.ufrrj@gmail.com

<https://andradeadriana.github.io/praticaR/>

<https://cursos.ufrrj.br/posgraduacao/ppgeap/>

- › Introdução
- › Teste de Mann-Whitney
- › Teste de Wilcoxon
- › Teste de Kruskal-Wallis
- › Teste de Friedman
- › Teste de Qui-Quadrado
- › Correlação de Spearman

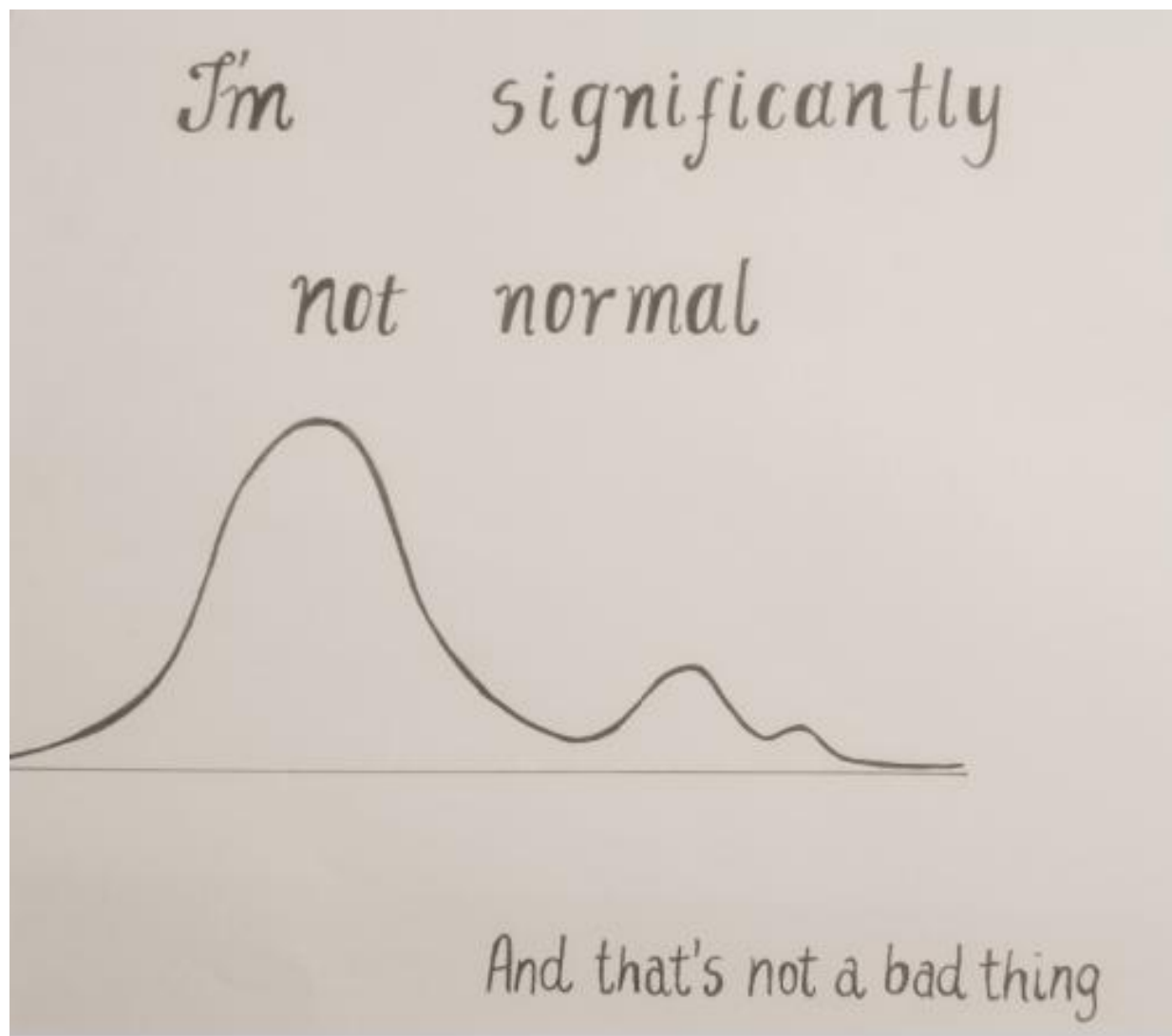
Introdução

- › Desenvolvida a partir dos estudos do inglês Ronald Fisher (meados do século XX).
- › Corresponde a um grande número de técnicas que são utilizadas amplamente.
- › ANOVA, método da máxima verossimilhança, dentre outros.
- › Fundamento importante em algumas das técnicas desenvolvidas por Fisher: Normalidade.

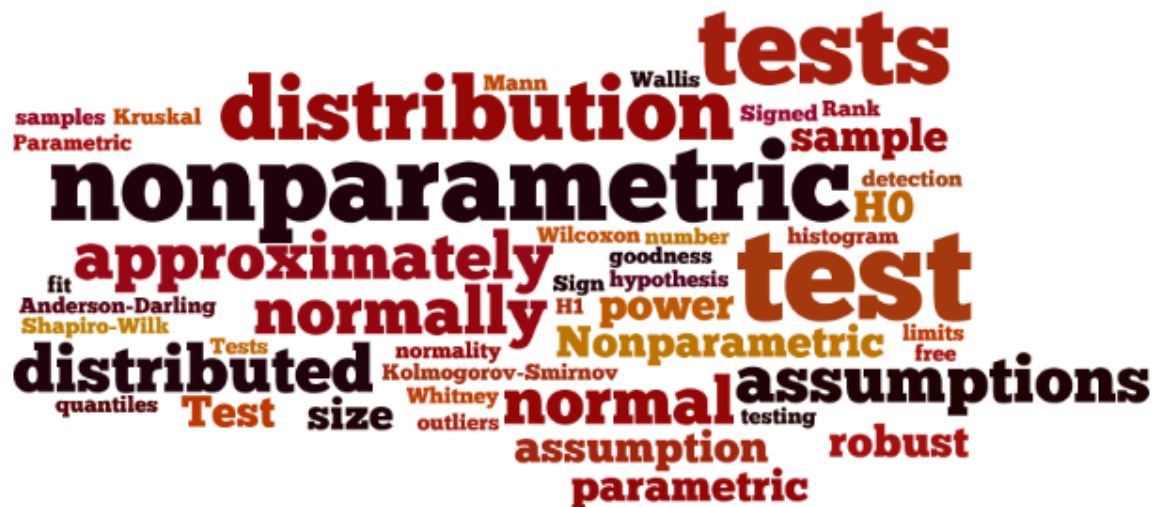
- › Apesar do Modelo Normal ter uso bastante difundido, em muitas circunstâncias, o fenômeno sob estudo não irá corresponder às características relacionadas com a Distribuição Normal;
- › O não reconhecimento dessas diferenças, seguido pelo uso de testes estatísticos que têm por requisito uma distribuição Normal, poderá ocasionar a obtenção de inferências inválidas e, portanto, a resultados incorretos.
- › Verificação da distribuição

- › **Análise Gráfica**
 - › Histograma
 - › Boxplot
 - › QQ-plot
- › **Testes de Distribuição**
 - › Shapiro-Wilk
 - › Kolgomorov-Smirnov

Métodos Não Paramétricos



- ✓ Ausência de Normalidade
- ✓ Variâncias Heterogêneas
- ✓ **Tamanho amostral pequeno**



- › São aplicáveis até mesmo às variáveis nominais e ordinais. Diferentemente da estatística paramétrica, cujas técnicas, em grande medida, são destinadas às variáveis intervalares.
- › Possuem simplicidade de cálculos e são aplicáveis para análise de pequenas amostras ou ainda quando se tem algum questionamento relacionado à adoção de algum delineamento experimental.

- › Os métodos não-paramétricos independem da distribuição de probabilidade da população envolvida;
- › Os testes de não-paramétricos também são denominados de testes livres de distribuições.
- › A terminologia não-paramétrica por vezes é associada com a inexistência de um parâmetro de referência, entretanto, alguns deles se baseiam em algum parâmetro, como a mediana. Mas, tais testes não têm por pressuposto uma distribuição específica;

Vantagens	Desvantagens
Aplicável em diversas situações, por não ter por requisito a necessidade de uma distribuição específica.	Menor poder que testes paramétricos.
Cálculos mais simplificados.	Menos eficientes.
Aplicáveis a variáveis qualitativas e quantitativas ordinais	Perda de informação no processo de cálculo da estatística.

Tal como é inadequado a utilização de testes paramétricos quando não se cumprem os pressupostos necessários, também deverá ser evitada a utilização dos testes não-paramétricos em situações em que prevalecem as condições de utilização dos testes paramétricos, pois estes (paramétricos) são mais eficientes.

Métodos Não-Paramétricos

Medidas Descritivas

- › Se a distribuição não é Normal, a média geralmente não é uma medida de tendência central adequada para descrever a variável sob a análise.
- › Nos casos de distribuições assimétricas, a mediana é uma melhor opção.

Medida	Distribuições Normais	Distribuições Assimétricas
Tendência	Média	Mediana
Dispersão	Desvio Padrão	Amplitude interquartilica

- › Em boa parte das estatísticas não-paramétricas as medidas que serão submetidas aos testes não utiliza a magnitude dos valores observados.
- › As estatísticas não paramétricas, em geral, são calculadas com os **postos dos valores observados**.
- › O **posto (*rank*)** de um número é atribuído às observações originais com base na **posição que o valor assume no conjunto ordenado em ordem crescente**.

- Exemplo: Dois grupos de indivíduos, A e B, participaram da resolução de um exercício. O grupo A recebeu treinamento, enquanto o grupo B não foi treinado. Os tempos de resolução foram anotados para os indivíduos de cada grupo.

A	B
3	14
6	6
2	2
5	13
5	12
1	8

› Rankeamento dos tempos observados.

Valores observados <i>ordenados</i>	Grupo	Posição no <i>Rol</i> <i>ordenado</i>	Posto
Valor	Grupo	Posição	Posto
1	A	1	1
2	A	2	2,5
2	B	3	2,5
3	A	4	4
5	A	5	5,5
5	A	6	5,5
6	A	7	7,5
6	B	8	7,5
8	B	9	9
12	B	10	10
13	B	11	11
14	B	12	12

Quando ocorrer empates,
o posto será a média da soma das posições nas
quais se deu o empate.

- › Em geral, os testes não-paramétricos se baseiam na **soma dos postos**;
- › *Se H_0 é verdadeira, os postos estarão distribuídos de forma equilibrada entre as amostras e as somas dos postos serão semelhantes;*
- › Se H_0 é falsa, uma amostra tenderá a ter mais postos baixos, com uma menor soma de postos, enquanto a outra apresentará a maior soma de postos, já que possui os maiores postos. Diz-se que uma distribuição possui uma dominância estocástica sobre a outra.

Grupo	Soma dos Postos
A	26
B	52

1 – Teste Mann-Whitney

- › Teste equivalente ao teste *t-Student* ;
- › Aplicado para comparar duas amostras independentes;
- › Não exige nenhuma hipótese sobre a distribuição populacional e suas variâncias;
- › Testa se a localização dos dados é a mesma para duas populações independentes.

› Hipóteses

H_0 : Distribuições são iguais.

H_1 : Distribuição não são iguais.

- › O teste de Mann-Whitney (*estatística U*) também é uma opção interessante no caso da variável sob estudo ter uma distribuição qualitativa ordinal, a exemplo das variáveis com que utilizam a escala Likert.
- › Ao invés de utilizar os valores arbitrariamente atribuídos aos níveis da variável, serão utilizados os postos na estatística de teste.

Etapas - Para $n_1 + n_2 \geq 20$ (Siegel, 1975)

1 – Obter os postos para o conjunto, sendo R_1 e R_2 a soma dos postos atribuídos aos valores do grupo 1 e 2 respectivamente;

2 – Calcular para cada grupo:

$$U_1 = n_1 n_2 + \frac{n_1 (n_1 + 1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2 (n_2 + 1)}{2} - R_2 = n_1 n_2 - U_1$$

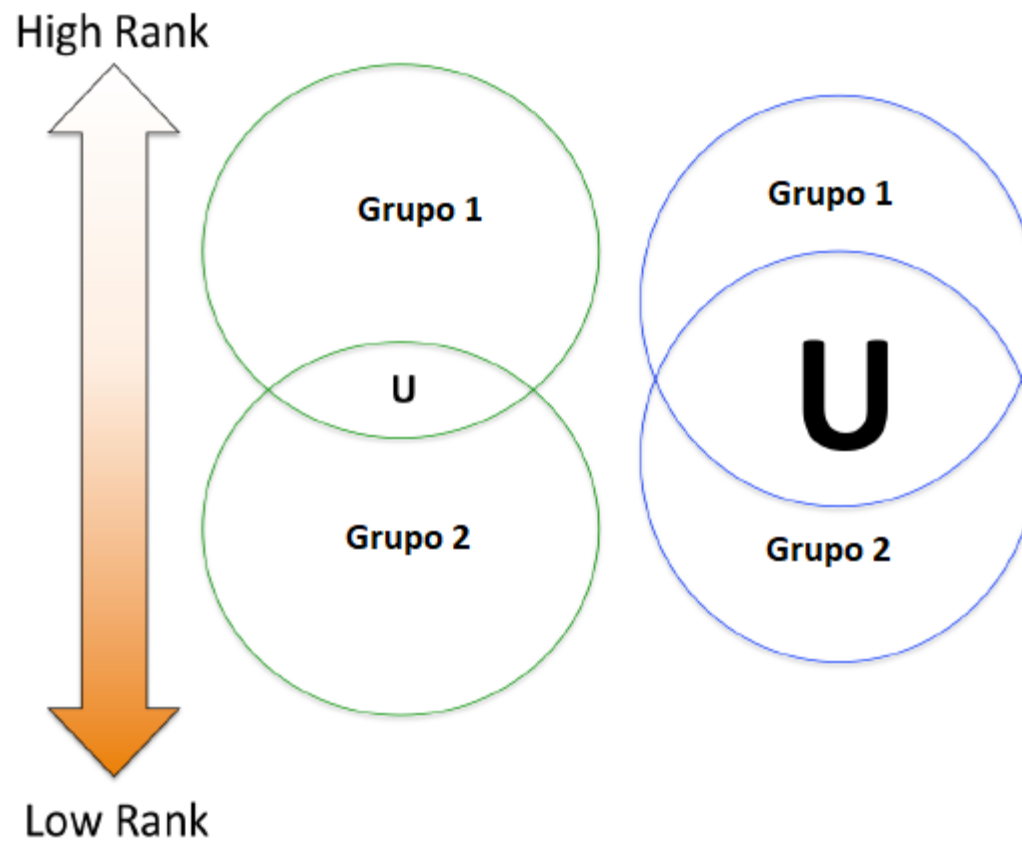
3 - Determina-se U como o menor entre U_1 e U_2 ;

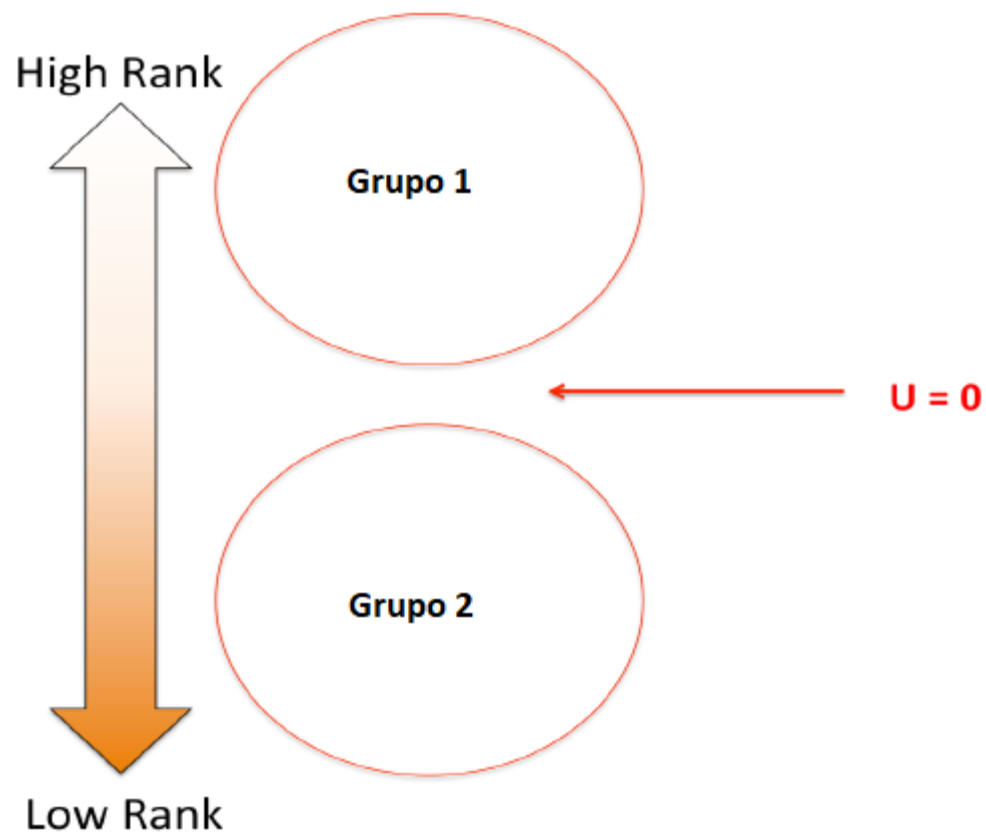
- › A estatística U fornece o grau de sobreposição nos postos entre os dois grupos;

U pequeno = Grande diferença entre os grupos;

U grande = Pequena diferença entre os grupos;

- › *Quanto menor o valor de U, maior a evidência de que as populações são diferentes.*





4i - Consultar a tabela com os valores críticos do Teste de Mann-Whitney;

- H_0 é rejeitada se o menor valor de U entre as amostras for menor ou igual ao valor crítico tabelado.

4.ii-Quando há muitos empates, e ambas as amostras tem tamanhos iguais ou superiores a 10, pode fazer-se a aproximação à função de distribuição normal. Obter a estatística U padronizada pela Normal:

$$z_U = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

› Caso $p\text{-valor} < \alpha$, deveremos rejeitar a hipótese nula.

- › Exemplo: Foi realizado um delineamento inteiramente ao acaso para avaliar a diferença entre dois tratamentos A e B no peso (gramas) das mudas aos 30 dias após a semeadura.

A	22	23	28	29	30	33			
B	36	19	30	34	33	42	40	35	37

Tratamento	A	B
Mediana	28,5	35

1º) Obter os postos para os valores observados

Tratamento	Peso	Posto
B	19	1
A	22	2
A	23	3
A	28	4
A	29	5
A	30	6,5
B	30	6,5
A	33	8,5
B	33	8,5
B	34	10
B	35	11
B	36	12
B	37	13
B	40	14
B	42	15

Soma dos Postos

$$R_A = 29$$

$$R_B = 91$$

2º) Cálculo da Estatística U

$$U_A = 46$$

$$U_B = 8 \rightarrow U = 8$$

> Software R

```
wilcox.test(Peso~Tratamento, data=data)
```

Wilcoxon rank sum test with continuity correction

data: Peso by Tratamento

$W = 8$, p-value = 0.02895

alternative hypothesis: true location shift is not equal to 0

No R, o teste de Mann-Whitney é fornecido pelo default da função `wilcox.test`

Rejeita-se a hipótese de que a população de pesos submetidos ao tratamento A seja igual à população de pesos que receberam o tratamento B.

2 – Teste Wilcoxon

- › O teste de Wilcoxon compara os postos (ranks) de dois grupos ou amostras pareadas ou repetidas;
- › Também chamado de teste de Mann-Whitney-Wilcoxon para dados pareados;

Amostras Pareadas (Relacionadas, Combinadas e/ou Repetidas)

- › Medidas realizadas sobre o mesmo indivíduo em momentos distintos de tempo.
- › O pareamento objetiva reduzir o efeito de fontes de variação que poderiam influenciar no resultado esperado.

Ex1: Pontuações pré-teste e pós-teste dos mesmos indivíduos;

Ex2: Status antes e depois de intervenção;

Observações não são *independentes!*

- › Este teste é equivalente ao teste t pareado;
- › Leva em consideração a magnitude da diferença para cada par;

Hipóteses:

H_0 : Não há diferença entre os grupos

H_1 : Há diferença entre os grupos

- › O teste de Wilcoxon se baseia nos **postos** das **diferenças intrapares**;
- › No caso de uma comparação entre um tratamento A e outro B, tendo A produzido maiores valores que B, as diferenças (A-B) com sinais “+” serão em maior quantidade;
- › Se os tratamentos apresentarem o mesmo efeito, as diferenças “+” e “-” tendem a ser anular.

› Etapas:

- 1) Atribuir postos (colocar em ordem crescente) às **diferenças de cada par** (d_i), desconsiderando os sinais. No caso de empate, atribuir a média dos postos empatados;
- 2) Identificar cada posto pelo sinal “+” ou “-”, seguindo a diferença (d_i) que representa.
- 3) Abater do n o número de zeros, isto é, os pares onde $d_i = 0$;
- 4) Determinar W : a menor das **soma** de postos de mesmo sinal (*);

$$\mu_T = \frac{n(n+1)}{4}; \sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}; z_{calc} = \frac{W - \mu_T}{\sigma_T}$$

› Exemplo:

Um pesquisador mediu a colinesterase sérica em agricultores que aplicaram inseticida em plantas de interesse comercial. Foram feitas duas coletas de sangue em cada pessoa: uma antes da aplicação do inseticida e outra 24h após. O que pode ser afirmar quanto ao efeito da exposição ao inseticida sobre o nível de colinesterase no sangue desses agricultores?

A hipótese nula que se deseja testar é:

› H_0 : o nível de colinesterase é o mesmo antes e após a aplicação do inseticida.

Teste de Wilcoxon

Colinesterase total (micromol/mL de plasma) em 17 agricultores do sexo masculino: dosagens antes e após uma sessão de aplicação de inseticidas em plantas.

Indivíduo	Antes	Depois	$d_i=A-D$
1	8,3	6,84	1,46
2	6,7	5,98	0,72
3	7,8	7,1	0,7
4	9,3	8,38	0,92
5	6,5	6,07	0,43
6	10,5	10,22	0,28
7	6,9	5,87	1,03
8	7,5	7,28	0,22
9	6,6	6,15	0,45
10	6,7	6,26	0,44
11	7,5	7,46	0,04
12	7,5	7,69	-0,19
13	8,1	7,95	0,15
14	8,8	9,15	-0,35
15	7,6	7,56	0,04
16	9,4	9,07	0,33
17	7,2	6,78	0,42

d_i	Rank	Signed Rank
0,04	1,5	1,5
0,04	1,5	1,5
0,15	3	3
-0,19	4	-4
0,22	5	5
0,28	6	6
0,33	7	7
-0,35	8	-8
0,42	9	9
0,43	10	10
0,44	11	11
0,45	12	12
0,7	13	13
0,72	14	14
0,92	15	15
1,03	16	16
1,46	17	17

Soma postos (+) = 141

Soma postos (-) = - 12

```
> wilcox.test(colinesterase ~ Tempo, paired=T, data=data)
```

```
Wilcoxon signed rank test with continuity correction
```

```
data: colinesterase by Tempo
```

```
V = 141, p-value = 0.002445
```

```
alternative hypothesis: true location shift is not equal to 0
```

- › Conclui-se que houve uma redução significativa nos níveis de colinesterase sérica após a exposição ao inseticida.

No R, o teste de Wilcoxon é fornecido pela função `wilcox.test` com o argumento `paired = TRUE`

3 – Teste Kruskal-Wallis

- › Este teste é equivalente a uma ANOVA de um fator (experimento inteiramente ao acaso);
- › Enquanto a análise de variância possui o requisito de normalidade e dos resíduos e homocedasticidade das variâncias, o teste KW não impõem essas restrições.

Extensão do teste de Mann-Whitney, aplicado na comparação de três ou mais grupos independentes;
Baseia-se na atribuição de **postos** aos valores observados.

Hipóteses:

H_0 : Não existe diferença entre os k grupos

H_1 : Existe diferença entre os k grupos

ETAPAS:

- 1) Dispor em ordem crescente as observações de todos os k grupos, atribuindo-lhes postos de 1 até n . Caso haja empates, atribuir o posto médio;
- 2) Determinar o valor da **soma dos postos** para cada um dos k grupos: $R_i, i = 1, 2, \dots, k$;
- 3) Calcular a estatística de teste:

$$\chi^2_{cal} = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{(R_i)^2}{n_i} - 3(n+1)$$

- › Essa estatística tem, aproximadamente, distribuição Qui-quadrado com graus de liberdade $k-1$;

- › Os dados a seguir são de uma experiência clássica agrícola para avaliar o rendimento de culturas divididas em quatro grupos diferentes. O objetivo da análise consiste em avaliar se os grupos são distintos.

G1	G2	G3	G4
83	81	81	77
84	83	91	78
89	83	93	79
89	84	94	80
90	88	95	81
91	89	96	81
91	90	100	82
92	91	101	
94	91		
96			

- › Obter postos dos dados e soma dos postos em cada grupo.

G1	G2	G3	G4
11	6,5	6,5	1
13,5	11	23	2
17	11	27	3
17	13,5	28,5	4
19,5	15	30	6,5
23	17	31,5	6,5
23	19,5	33	9
26	23	34	
28,5	23		
31,5			
210	139,5	213,5	32

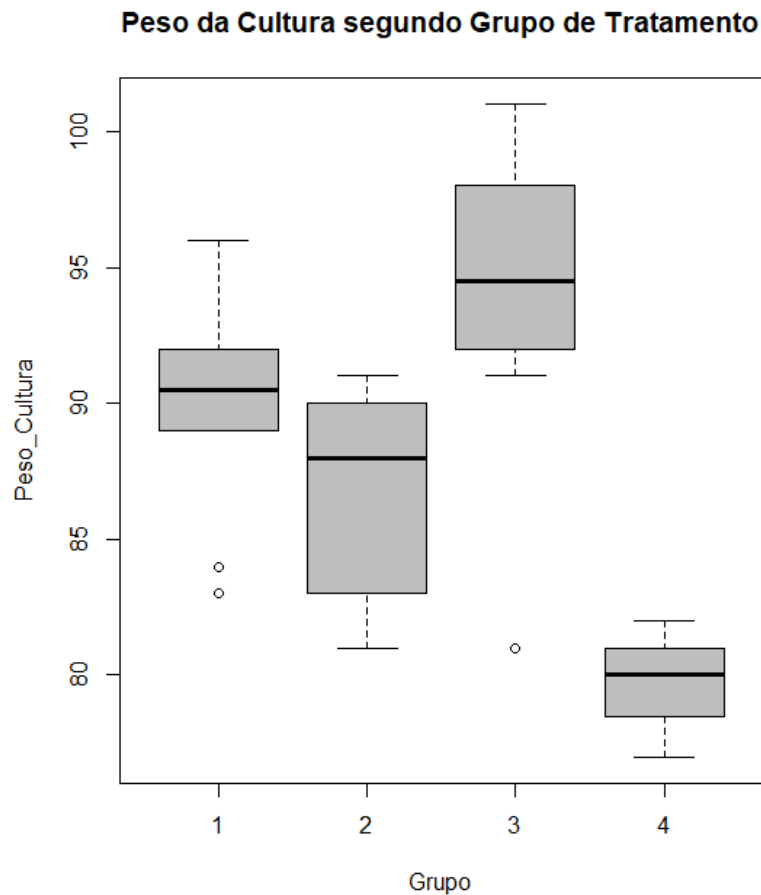
› Software R

› `kruskal.test(Cultura_Kruskal$cultura,Cultura_Kruskal$grupo)`

› Kruskal-Wallis rank sum test

› data: `Cultura_Kruskal$cultura` and `Cultura_Kruskal$grupo`

› Kruskal-Wallis chi-squared = 20.337, df = 3, p-value = 0.0001445



Teste Kruskal-Wallis
Rejeitou a hipótese de que as
todos os grupos são iguais.

Entretanto, o Teste KW
não indica quais grupos
são diferentes entre si.

- › **Análise Post Hoc – Teste de Dunn**
- › O teste de Dunn realiza comparações múltiplas entre os pares de grupos, tal como o teste Tukey. É computado a partir das médias dos postos.
- › O Teste de Dunn possibilita o ajuste do p-valor nas comparações múltiplas com a correção de Bonferroni.

Testes para diferença entre grupos – Teste de Dunn

Pares	P-values	
	Sem ajuste	Com ajuste
1 - 2 :	-0,1139	-0,6835
1 - 3 :	-0,1135	-0,6812
2 - 3 :	(0,0102)*	-0,0611
1 - 4 :	(0,0004)*	(0,0024)*
2 - 4 :	(0,0145)*	-0,0867
3 - 4 :	(0,0000)*	(0,0001)*

4 – Teste Friedman

- › Conceitualmente, pode ser considerado uma extensão do teste de Wilcoxon para mais de duas Amostras ou Grupos Pareados (Medidas Repetidas);
- › Também pode ser utilizado como teste equivalente a ANOVA 2 fatores – Delineamentos em Blocos ao Acaso;

Hipóteses:

H_0 : Não existe diferença entre os k grupos

H_1 : Existe diferença entre os k grupos

- › Seja um delineamento em blocos casualizados, com k tratamentos e b blocos ou medidas repetidas. Dentro de cada um dos b blocos, as observações são ordenadas em ordem crescente, e atribuídos números de ordem. Em seguida, somam-se os números de ordem (atribuídos por bloco) dentro de cada um dos tratamentos (R_i);
- › Determinar o valor da soma dos postos para cada um dos k grupos: R_i , $i = 1, 2, \dots, k$;

- › Calcular a estatística de teste:

$$\chi^2_{cal} Q' = \frac{12}{bk(k+1)} \sum_{i=1}^k \frac{(R_i)^2}{n_i} - 3b(k+1)$$

- › A estatística $\chi^2_{cal} Q'$: mensura se existe ou não pelo menos uma diferença significativa entre dois ou mais tratamentos levando em consideração os blocos ou medidas repetidas.

Exemplo:

Foi realizado um experimento em Blocos casualizados com a finalidade de estudar os efeitos na produção média diária de leite (litros) da administração de raízes e tubérculos, como suplementação de inverno na alimentação de vacas em lactação. O experimento contou com 4 tipos de suplementos (tratamentos) e 5 raças (Blocos). Verifique se há diferença entre os tipos de suplementação.

- › Existem suplementos mais eficientes na produção de leite?

Vacas	Suplementação			
	S	M	A	BD
Gir	6.4	12.0	11.2	10.9
Holandesa	11.6	10.9	6.2	11.6
Jersey	11.5	11.4	10.9	6.2
Nelore	10.4	7.1	12.1	11.1
Guzerá	12.4	6.6	10.1	11.8

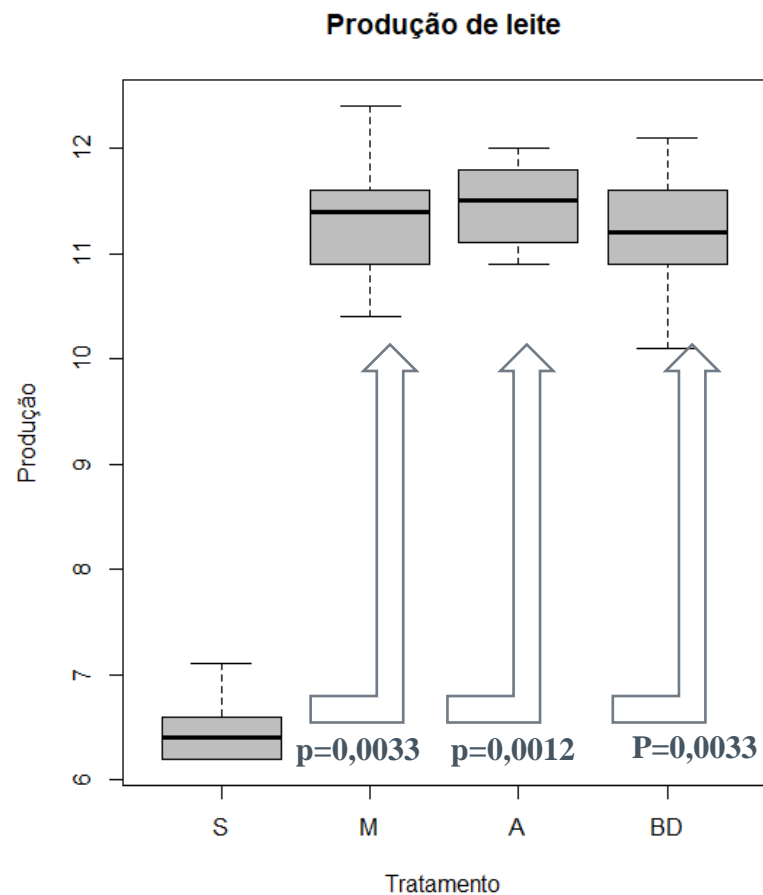
> Software R

- › Friedman rank sum test

- › data: `VacasFriedman$Producao`, `VacasFriedman$Trat` and `VacasFriedman$Blocos`

- › Friedman chi-squared = 9.3673, df = 3, p-value = 0.02479

Pós-teste Friedman



5 – Teste Qui-Quadrado

- › Possibilita duas versões de teste:
 - 1) Verificação da adequação ou o ajuste entre uma distribuição de frequências observadas e uma distribuição teórica;
 - 2) Teste de associação/independência entre duas variáveis qualitativas;
- › O objetivo é verificar se a distribuição das frequências observadas se diferencia significativamente das frequências esperadas;

› A estatística de qui-quadrado é dado por:

$$\chi^2 = \frac{\sum_{i=1}^k (f_{\text{observado}} - f_{\text{esperado}})^2}{f_{\text{esperado}}}$$

Hipóteses:

H_0 : As variáveis são independentes

H_1 : As variáveis não são independentes

Exemplo: Verificar a frequência de ocorrência de 4 tipos de sangue em uma dada raça

Classes	A	B	AB	O
Freq. Observada	230	470	170	130
Freq. esperada	180	480	200	140

$$\chi^2 = (230-180)^2/180 + (470-480)^2/480 + (170-200)^2/200 + (130-140)^2/140$$

$$\chi^2_{\text{calc}} = 16.04$$

$$\chi^2_{\text{tab}} = 9,25$$

Logo rejeita-se H_0 com 2,5% de probabilidade de erro.

- › Exemplo 1: Em duas amostras tiradas de solos diferentes foram encontradas duas espécies de tatuzinho: Oniscus e Armadilidium. Pretende-se testar a hipótese de que a predominância das espécies de tatuzinho independe do tipo de solo.

Tipo de Solo	Espécie Tatuzinho	
	Oniscus	Armadilidium
Solo Argiloso	14	6
Solo Calcáreo	22	46

› Software R

› `chisq.test(x)`

› Pearson's Chi-squared test

› data: x

› X-squared = 9.061, df = 1, p-value = 0.002611

Considerações:

- › No caso de análises de amostras com menos de 20 indivíduos e nas tabelas 2x2, a frequência observada mínima não pode ser menor que 5 (neste caso, usar a correção de continuidade de Yates);

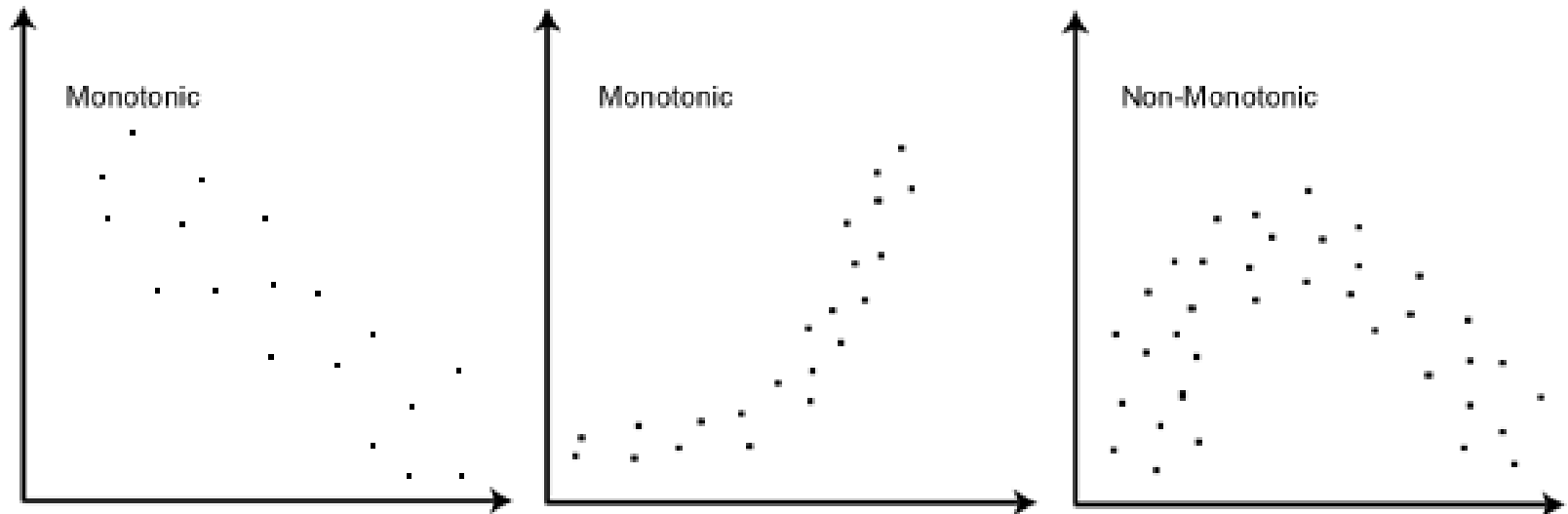
6 – Correlação de Spearman

Coeficiente de Correlação de Spearman

- › Mensura a correlação entre duas variáveis;
- › Este teste não supõe que as variáveis envolvidas tenham uma distribuição em particular, sendo portanto um coeficiente de correlação não-paramétrico;
- › Particularmente útil quando uma (ou ambas) variável(eis) é(são) quantitativas discretas ou qualitativas ordinais;
- › Calculado sobre a ordenação (postos) dos dados obtidos, dentro de cada variável, daí calcula-se o Coeficiente de Correlação de Pearson entre posto-x e posto-y.

Coeficiente de Correlação de Spearman

- › Mensura relações lineares e não-Lineares monotônicas;



Coeficiente de Correlação de Spearman

- › Hipóteses:
- › $H_0 : \rho_S = 0$ (não existe correlação entre as variáveis)
- › $H_1 : \rho_S \neq 0$ (existe correlação entre as variáveis)

Coeficiente de Correlação de Spearman

- › Coeficiente de Correlação de Spearman é dado por:

$$\rho_s = 1 - \frac{6 \sum d_i^2}{n^3 - n} \quad -1 \leq \rho_s \leq +1$$

Sendo d_i a diferença entre os postos de ordem i .

- › A estatística do teste em função do Coeficiente de Correlação de Spearman é dada por:

$$t_{calc} = \rho_s \sqrt{\frac{n-2}{1-\rho_s^2}}$$

Fixar α , $t_{tab} = t_{(n_1+n_2-2; \alpha\%)}$;

Coeficiente de Correlação de Spearman

- › A correlação positiva indica que os postos das duas variáveis seguem aproximadamente o mesmo padrão de distribuição.
- › A correlação negativa indica que há uma inversão dos valores dos postos da variável Y em relação à variável X.
- › Um coeficiente cujo valor seja próximo de zero sugere a não existência de correlação entre as duas variáveis.

Coeficiente de Correlação de Spearman

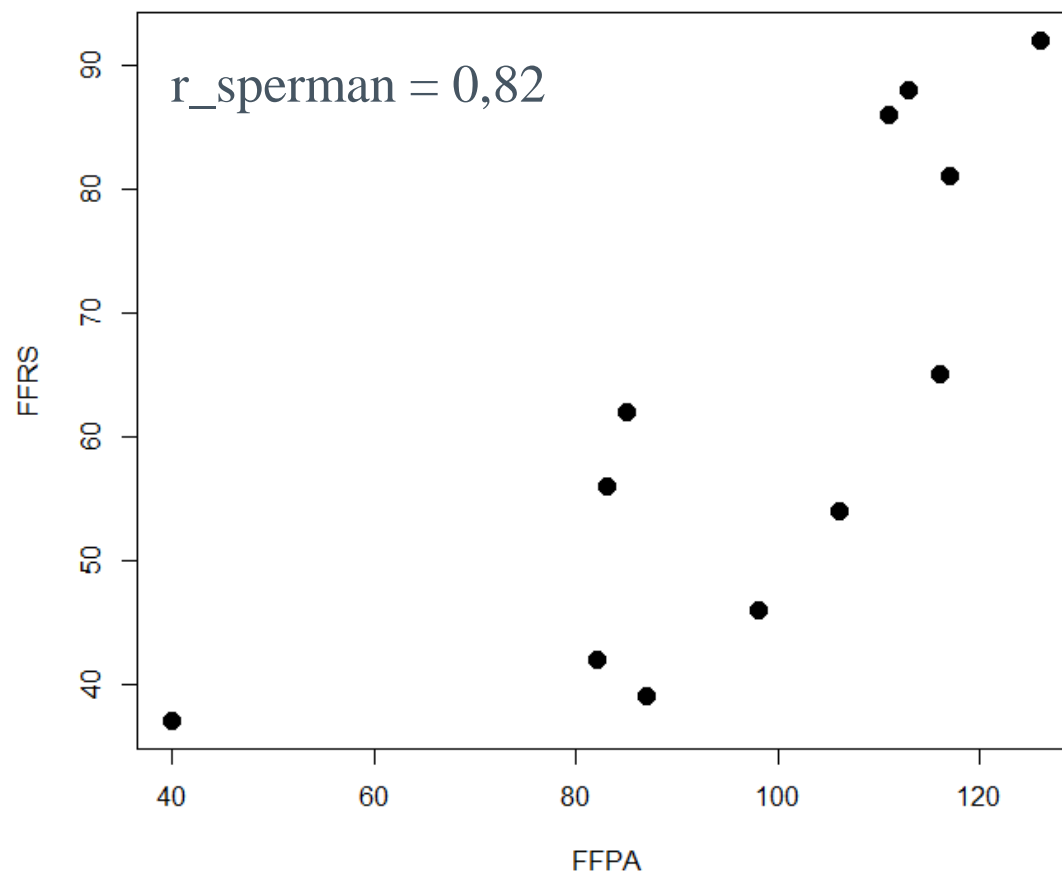
Exemplo: Em um experimento com 12 vasos contendo uma planta de feijão, cultivar carioca. Foram avaliadas a fitomassa fresca da parte aérea (FFPA) e do sistema radicular (FFSR) aos 45 dias após a semeadura. Verificar se as variáveis observadas estão significativamente correlacionadas.

Parcela	1	2	3	4	5	6	7	8	9	10	11	12
FFPA	82	98	87	40	116	113	111	83	85	126	106	117
FFSR	42	46	39	37	65	88	86	56	62	92	54	81

> Software R

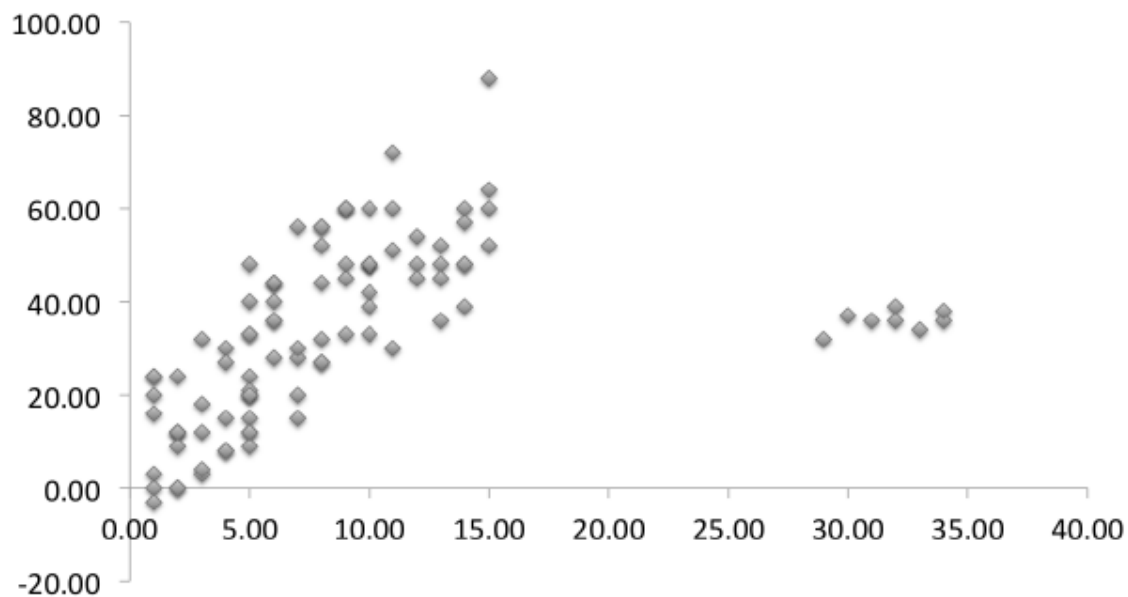
- › `plot(feijao$FFPA,feijao$FFSR, main="Diagrama de Dispersão - FFPA x FFSR",cex=1.5,pch=16,xlab="FFPA",ylab="FFRS")`
- › `cor.test(feijao$FFPA,feijao$FFSR,method = c("spearman"))`

Diagrama de Dispersão - FFPA x FFRS



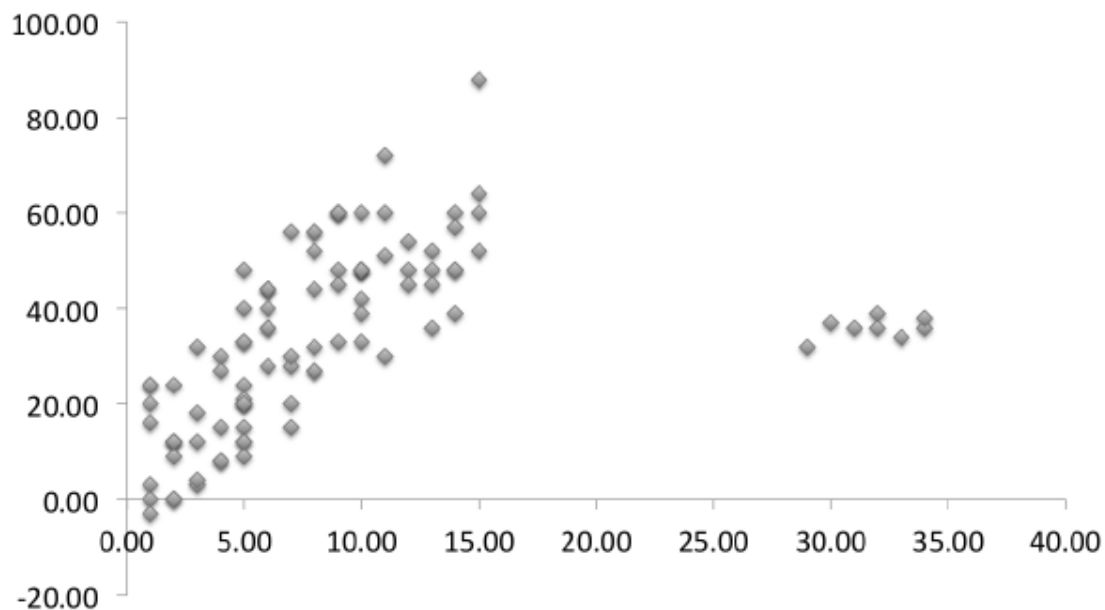
Coeficiente de Correlação de Spearman

Outliers



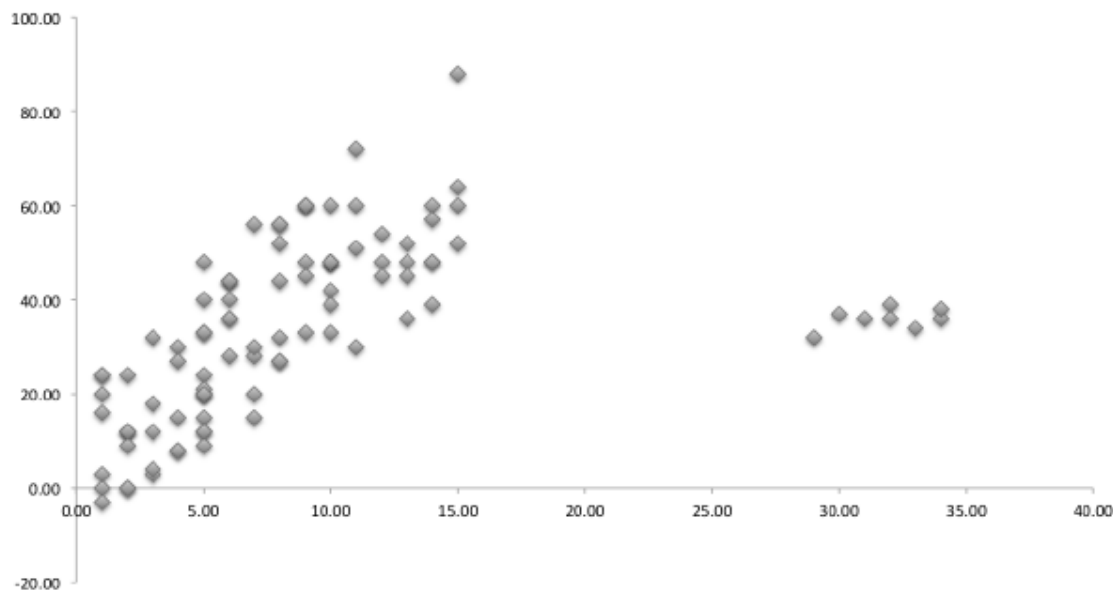
Coeficiente de Correlação de Spearman

Pearson's $r = .42$

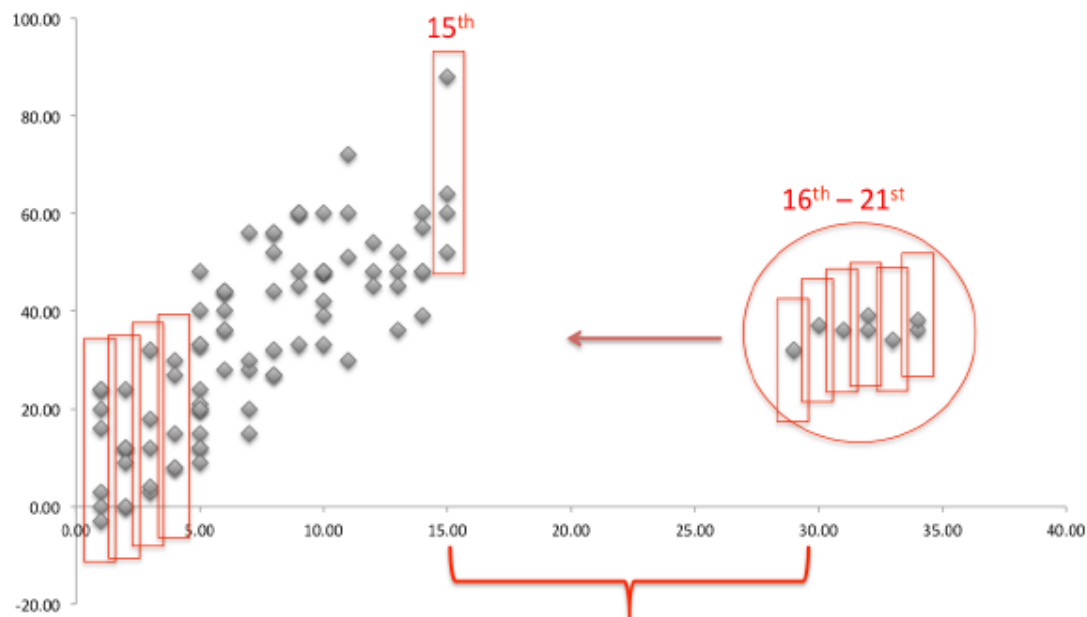


Coeficiente de Correlação de Spearman

Spearman's $r = .73$



Coeficiente de Correlação de Spearman



Data Setup	Parametric test	Non-Parametric test
1 Variable 2 Categories Between Subjects	independent t-test	Mann-Whitney U test
1 Variable 2 Categories Within-Subjects	paired t-test	Wilcoxon Signed Rank Test
1 Variable >2 Categories Between Subjects	One-way ANOVA	Kruskal Wallis Test
1 Variable >2 Categories Within Subjects	repeated measures ANOVA	Friedman test
2 variables	Pearson's r	Spearman's ρ (rho)

- › Bibliografia
- › CONOVER, W. J. Practical Nonparametric Statistics
- › LEHAMN, E. L. Nonparametrics: Statistical Methods Based on Ranks
- › SIEGEL, S. Estatística Não-Paramétrica para as Ciências do Comportamento
- › MARTINS, G.A. Estatística Geral e Aplicada

- › <https://cran.r-project.org/>
- › <https://www.rstudio.com/>