

Testes Não Paramétricos

DAI (Duas Amostras Independentes)

Prof. Lorí Viali, Dr.

<http://www.mat.ufrgs.br/viali/>

viali@mat.ufrgs.br

*Testes de
duas Amostras
Independentes*

Os testes

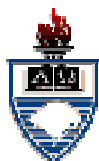
- ❖ *O teste Qui-Quadrado*
- ❖ *O teste exato de Fisher*
- ❖ *O teste de Kolmogorov-Smirnov*
- ❖ *O teste de U de Mann-Whitney*
- ❖ *O teste de Wilcoxon*
- ❖ *O teste de Siegel-Tukey*
- ❖ *O teste de Moses*

O teste Qui-Quadrado



O teste qui-quadrado

O teste χ^2 de duas ou mais amostras independentes pode ser utilizado para verificar a dependência ou independência entre as variáveis sendo consideradas. As variáveis devem estar tabuladas em tabelas de contingência. Para o caso de duas variáveis, tem-se uma tabela de dupla entrada.



Hipóteses e Cálculo

\mathcal{H}_0 : *As variáveis são independentes*

\mathcal{H}_1 : *As variáveis são dependentes*

A variável teste é:

$$\chi^2_v = \frac{\sum_{i=1}^k \sum_{j=1}^l (O_{ij} - E_{ij})^2}{E_{ij}}$$



Expressão alternativa

A variável teste é:

$$\chi_v^2 = \frac{\sum_{i=1}^k \sum_{j=1}^l (O_{ij} - E_{ij})^2}{E_{ij}} = \frac{\sum_{i=1}^k \sum_{j=1}^l O_{ij}^2}{E_{ij}} - n$$



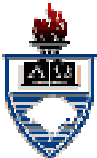
Onde:

r = número de linhas da tabela;

\mathcal{L} = número de colunas da tabela;

*O_{ij} = frequência observada na interseção da
linha i com a coluna j .*

*E_{ij} = número de casos esperados na interseção
da linha i com a coluna j .*



Onde:

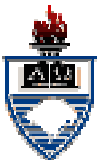
χ^2_v *é a estatística teste;*

$n = \sum_{i=1}^k \sum_{j=1}^l O_{ij}$ *= tamanho da amostra;*

$E_{ij} = n p_{ij}$ *são as frequências esperadas
de cada célula ij da tabela.*

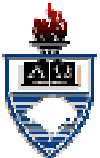


p_{ij} é a probabilidade de ocorrer uma observação na célula ij . Se as variáveis são supostamente independentes (\mathcal{H}_0 é Verdadeira), então $p_{ij} = p_i \cdot p_j$ onde p_i é a probabilidade marginal correspondente à linha “ i ” e p_j é a probabilidade marginal correspondente a coluna j .



Como não se conhecem as probabilidades marginais, elas devem ser estimadas através das correspondentes frequências relativas. Então:

$$E_{ij} = n p_{ij} = n p_{i.} \cdot p_{.j} =$$
$$n \cdot \frac{f_{i.}}{n} \cdot \frac{f_{.j}}{n} = \frac{f_{i.} f_{.j}}{n}$$



$$f_{i.} = \sum_{j=1}^{\ell} f_{ij} \quad e \quad f_{.j} = \sum_{i=1}^k f_{ij}$$



Exemplo



A tabela mostra os resultados de uma avaliação de satisfação com a compra de um novo modelo de automóvel de luxo. Teste a hipótese de que o novo modelo está agradando mais aos consumidores homens do que os consumidores mulheres.



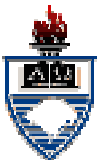
	<i>Avaliação</i>		
<i>Consumidores</i>	<i>Muito</i>	<i>Pouco</i>	<i>Não Satisfeito</i>
<i>Homens</i>	<i>30</i>	<i>20</i>	<i>15</i>
<i>Mulheres</i>	<i>25</i>	<i>5</i>	<i>5</i>



Hipóteses

H_0 : *Homens e mulheres estão igualmente satisfeitos.*

H_1 : *Homens e mulheres não estão igualmente satisfeitos.*



Totais marginais

<i>Consumidores</i>	<i>M</i>	<i>P</i>	<i>NS</i>	<i>Total</i>
<i>Homens</i>	<i>30</i>	<i>20</i>	<i>15</i>	<i>65</i>
<i>Mulheres</i>	<i>25</i>	<i>5</i>	<i>5</i>	<i>35</i>
<i>Total</i>	<i>55</i>	<i>25</i>	<i>20</i>	<i>100</i>



Frequências Esperadas

<i>Consumidores</i>	<i>M</i>	<i>P</i>	<i>NS</i>	<i>Total</i>
<i>Homens</i>	<i>35,75</i>	<i>16,25</i>	<i>13</i>	<i>65</i>
<i>Mulheres</i>	<i>19,25</i>	<i>8,75</i>	<i>7</i>	<i>35</i>
<i>Total</i>	<i>55</i>	<i>25</i>	<i>20</i>	<i>100</i>



Cálculo do Qui-Quadrado

<i>Consumidores</i>	<i>M</i>	<i>P</i>	<i>NS</i>	<i>Total</i>
<i>Homens</i>	<i>0,925</i>	<i>0,865</i>	<i>0,310</i>	<i>2,100</i>
<i>Mulheres</i>	<i>1,712</i>	<i>1,607</i>	<i>0,570</i>	<i>3,900</i>
<i>Total</i>	<i>2,642</i>	<i>2,473</i>	<i>0,880</i>	<i>5,990</i>



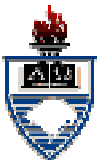
A estatística amostral

O grau de liberdade é:

$$v = (k - 1)(l - 1) = (2 - 1).(3 - 1) = 2$$

Então:

$$\chi^2_2 = \frac{\sum_{i=1}^2 \sum_{j=1}^3 (O_{ij} - E_{ij})^2}{E_{ij}} = 5,99$$



Qual a significância deste resultado?

Argumentos da função

DIST.QUIQUA

X	5,99	=	5,99
Graus_liberdade	2	=	2
Cumulativo	1	=	VERDADEIRO

= 0,949963373

Retorna a probabilidade de cauda esquerda da distribuição qui-quadrada.

Cumulativo é um valor lógico a ser retornado pela função: a função de distribuição cumulativa = VERDADEIRO, a função de densidade da probabilidade = FALSO.

Resultado da fórmula = 0,9500

[Ajuda sobre esta função](#)

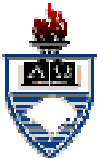
Este resultado deve ser subtraído de 1 para fornecer a cauda à direita. Assim a significância é 5%.



Tipos de Qui-Quadrado

O SPSS fornece ainda os seguintes valores do χ^2 :

- Qui-Quadrado de Pearson;
- Corrigido de Yates ou Correção de Continuidade;
- Razão de verossimilhança;
- Teste exato de Fisher;
- Qui-Quadrado de Mantel-Haenszel ou teste de associação linear ou ainda associação linear por linear.



Correção de Continuidade – Yates

Obs.: Só para tabelas 2x2

$$Q_c = \frac{\sum_{i=1}^k \sum_{j=1}^l [\max(0, |O_{ij} - E_{ij}| - 0,50)]^2}{E_{ij}}$$

Sob a hipótese nula de independência a estatística Q_c tem uma distribuição assintótica Qui-Quadrado com $(k-1) \cdot (l-1)$ g.l.



Razão de verossimilhança

$$G^2 = 2 \sum_{i=1}^k \sum_{j=1}^l O_{ij} \ln \left(\frac{O_{ij}}{E_{ij}} \right)$$

Quando as variáveis das linhas e colunas são independentes a estatística G^2 tem uma distribuição assintótica Qui-Quadrado com $(k-1) \cdot (l-1)$ g.l.



Qui-Quadrado de Mantel-Haenszel

$$Q_{MH} = (n - 1) r^2$$

O Qui-Quadrado de Mantel-Haenszel testa a hipótese de que existe um relacionamento linear entre as duas variáveis. R^2 é a correlação de Pearson ($r\hat{o}$) entre as duas variáveis, que é calculado da seguinte forma:



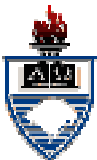
$$r = \frac{\text{cov}(X, Y)}{\sqrt{S_X S_Y}}$$

Onde:

$$\text{cov}(X, Y) = \sum_{i=1}^r \sum_{j=1}^c x_i y_j f_{ij} - \left(\sum_{i=1}^r x_i r_i \right) \left(\sum_{j=1}^c y_j c_j \right) / n$$

$$S_X = \sum_{i=1}^r x_i^2 r_i - \sum_{j=1}^r x_i r_i / n$$

$$S_Y = \sum_{j=1}^c y_j^2 c_j - \sum_{j=1}^c y_j c_j / n$$



Considerando uma tabela $r \times c$

	<i>1</i>	<i>2</i>	<i>...</i>	<i>c</i>	<i>Total</i>
<i>1</i>	f_{11}	f_{12}	<i>...</i>	f_{1c}	r_1
<i>2</i>	f_{21}	f_{22}	<i>...</i>	f_{2c}	r_2
<i>...</i>	<i>...</i>	<i>...</i>	<i>...</i>	<i>..</i>	<i>...</i>
<i>r</i>	f_{r1}	f_{r2}	<i>...</i>	f_{rc}	r_r
<i>Total</i>	c_1	c_2	<i>...</i>	c_{rc}	n



O teste exato de Fisher



O teste exato de Fisher

O teste de Fisher é útil para analisar dados discretos (nominais ou ordinais), quando os tamanhos das duas amostras são pequenos.

A cada indivíduo nos grupos é atribuído um dentre dois escores possíveis. Os escores são frequências em uma tabela 2×2 .



As amostras podem ser quaisquer dois grupos independentes tais como: homens e mulheres, empregados e desempregados, católicos e não-católicos, pais e mães, etc.



Disposição dos dados na prova de Fisher

	-	+	<i>Total</i>
<i>Grupo I</i>	\mathcal{A}	\mathcal{B}	$\mathcal{A} + \mathcal{B}$
<i>Grupo II</i>	\mathcal{C}	\mathcal{D}	$\mathcal{C} + \mathcal{D}$
<i>Total</i>	$\mathcal{A} + \mathcal{C}$	$\mathcal{B} + \mathcal{D}$	n



Os cabeçalhos são arbitrariamente indicados com sinais de "mais" e "menos", podem indicar duas classificações quaisquer: acima e abaixo da mediana, aprovado e reprovado, graduados em ciências e graduados em artes, a favor ou contra, etc.

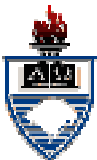


A prova determina se os dois grupos diferem na proporção em que se enquadram, nas duas classificações, ou seja, a prova determina se o Grupo I e o Grupo II diferem significativamente na proporção de sinais "mais" e "menos" atribuídos a cada um.



A estatística teste

A probabilidade de se observar determinado conjunto de frequências em uma tabela 2×2 , quando se consideram fixos os totais marginais, é dada pela distribuição hipergeométrica, isto é:



$$\begin{aligned}
 P(X = x) &= \frac{\binom{\mathcal{A} + \mathcal{C}}{\mathcal{A}} \binom{\mathcal{B} + \mathcal{D}}{\mathcal{B}}}{\binom{n}{\mathcal{A} + \mathcal{B}}} = \\
 &= \frac{(\mathcal{A} + \mathcal{B})! (\mathcal{C} + \mathcal{D})! (\mathcal{A} + \mathcal{C})! (\mathcal{B} + \mathcal{D})!}{n! \mathcal{A}! \mathcal{B}! \mathcal{C}! \mathcal{D}!}
 \end{aligned}$$



Exemplo



Suponha que os seguintes valores tenham sido observados: $\mathcal{A} = 10$, $\mathcal{B} = 0$, $\mathcal{C} = 4$ e $\mathcal{D} = 5$.

Então a tabela anterior seria:

	-	+	<i>Total</i>
<i>Grupo I</i>	10	0	10
<i>Grupo II</i>	4	5	9
<i>Total</i>	14	5	19



O valor da estatística, nesse caso, seria:

$$P = (10!9!14!5!)/(19!10!0!4!5!) = 1,08\%$$

Então sob H_0 , a probabilidade de dessa configuração ou uma mais extrema é de $p = 1,08\%$.



Esse exemplo foi simples em virtude da existência de uma célula com valor zero. Se nenhuma das frequências for zero, sob H_0 , podem ocorrer desvios "mais extremos" que devem ser levados em conta, pois o teste envolve a probabilidade daquela ocorrência ou de uma ocorrência ainda mais extrema?



Suponha, por exemplo, que os resultados de um teste fossem os da tabela:

	-	+	<i>Total</i>
<i>Grupo I</i>	1	6	7
<i>Grupo II</i>	4	1	5
<i>Total</i>	5	7	12



Com os mesmos totais marginais, uma situação mais extrema seria:

	-	+	Total
Grupo I	0	7	7
Grupo II	5	0	5
Total	5	7	12

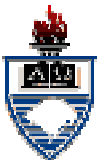


Se quisermos aplicar o teste a esses devemos somar as probabilidades das duas ocorrências.

Tem-se, então:

$$p_1 = (7!5!5!7!)/(12!1!6!4!1!) = 4,40\%.$$

$$p_2 = (7!5!5!7!)/(12!0!7!5!0!) = 0,13\%.$$



Logo:

$$p = p_1 + p_2 = 4,40\% + 0,13\% = 4,53\%.$$

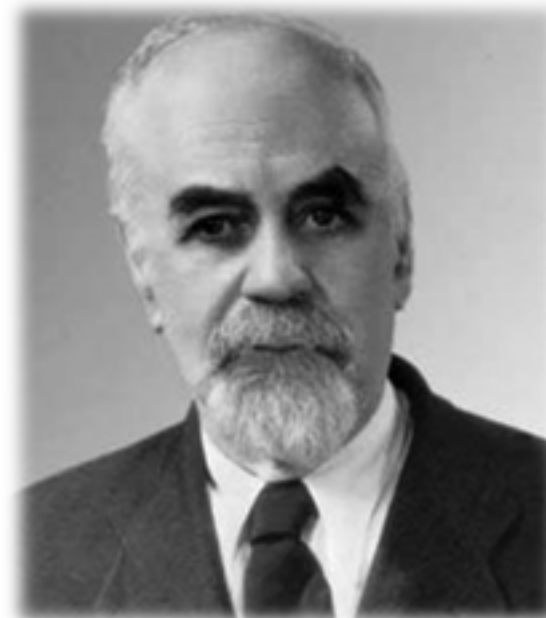
Isto é 4,53% é o valor-p que se deve utilizar para decidir se esses dados nos permitem rejeitar H_0 .



Pelo exemplo, pode-se verificar, que mesmo quando o menor valor não é muito grande, os cálculos do teste de Fisher se tornam longos. Por exemplo, se o menor valor for 2, deve-se determinar 3 probabilidades e somá-las. Se o menor valor de uma na célula é três, tem-se que determinar quatro probabilidades e somá-las e assim por diante.

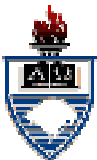


O teste de
Kolmogorov - Smirnov
ou $K-S$



Objetivos

A prova de Kolmogorov-Smirnov de duas amostras verifica se elas foram extraídas da mesma população (ou de populações com a mesma distribuição). A prova bilateral é sensível a qualquer diferença nas distribuições das quais se extraíram as amostras (posição central, dispersão ou assimetria).



A prova unilateral é utilizada para determinar se os valores da população da qual se extraiu uma das amostras são, ou não, estocasticamente maiores do que os valores da população que originou a outra amostra.



Metodologia

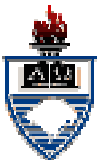
O teste utiliza as distribuições acumuladas. A prova de uma amostra verifica a concordância entre a distribuição de um conjunto de valores amostrais e uma distribuição teórica. A prova de duas amostras visa a concordância entre dois conjuntos de valores amostrais.



Se as duas amostras foram extraídas da mesma população, então se espera que as distribuições acumuladas das amostras estejam próximas. Se as distribuições estão “distantes” isto sugere que as amostras provenham de populações distintas e um desvio grande pode levar a rejeição da hipótese de nulidade.



O teste paramétrico equivalente é o t . Embora menos eficiente o $K-S$ é mais versátil pois trabalha apenas com as ordens das duas variáveis, sem se preocupar com o valor das mesmas. Ele envolve menos cálculos e apresenta menos restrições que o teste t .



Aplicação

Para aplicar a prova constrói-se a distribuição das frequências acumuladas relativas de cada uma das amostras, utilizando os mesmos intervalos (amplitude de classes) para cada uma delas. Em cada intervalo subtrai-se uma função da outra. A prova utiliza como estatística o maior destas diferenças.



Hipóteses

\mathcal{H}_0 : *As amostras são da mesma pop.*

\mathcal{H}_1 : *As amostras não são da mesma pop.*

Inicialmente ordenam-se as $t = m + n$ observações de forma crescente. Considera-se os estimadores S_1 e S_2 de F_1 e F_2 , isto é:

$$S_1(x) = k_1/m \text{ e } S_2(x) = k_2/n$$



Onde $k_1 =$ número de valores $X_i \leq x$;

$k_2 =$ número de valores $Y_j \leq x$;

Define-se:

$$D = \max |S_1(x) - S_2(x)|$$



Rejeitamos \mathcal{H}_0 ao nível α de significância se:

$$\mathcal{D} = \max |S_1(x) - S_2(x)| \geq \mathcal{D}_\alpha \text{ onde}$$

$$P(\mathcal{D} \geq \mathcal{D}_\alpha) = \alpha$$



Exemplo:

Os resultados de duas amostras \mathcal{A} e \mathcal{B} são:

\mathcal{A}		\mathcal{B}	
7,49	7,37	7,28	7,48
7,35	7,51	7,35	7,31
7,54	7,50	7,52	7,22
7,48	7,52	7,50	7,41
7,48	7,46	7,52	7,45



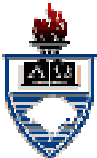
Verifique se existe uma diferença significativa entre as duas amostras.

Tem-se:

$$H_0: F_1(x) = F_2(x)$$

$$H_1: F_1(x) \neq F_2(x)$$

Fazer no Excel e depois no SPSS!



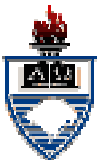
A tabela

Se o $n > 40$ e o teste é bilateral, então o

valor crítico é dado por: $d = 1,36 \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$

Se o teste é unilateral, então o valor

crítico é dado por: $\chi_2^2 = 4 \mathcal{D}^2 \frac{n_1 n_2}{n_1 + n_2}$



Exemplo



Amostras de $n_1 = n_2 = 50$ valores das opiniões de diretores financeiros de grandes e pequenas empresas mostraram os resultados da tabela seguinte, medidos em uma escala Likert de 5 pontos:



Amostras de $n_1 = n_2 = 50$ valores das opiniões de diretores de empresas Grande e Pequenas.

<i>Escala</i>	<i>Grandes</i>	<i>Pequenas</i>
<i>1</i>	<i>5</i>	<i>15</i>
<i>2</i>	<i>8</i>	<i>13</i>
<i>3</i>	<i>10</i>	<i>10</i>
<i>4</i>	<i>15</i>	<i>8</i>
<i>5</i>	<i>12</i>	<i>4</i>
<i>Total</i>	<i>50</i>	<i>50</i>



Teste a hipótese de que opiniões dos diretores dos dois tipos de empresa são divergentes.



Determinação das Diferenças

<i>Escala</i>	<i>Grandes</i>	<i>Pequenas</i>	$Fr_1(x)$	$Fr_2(x)$	$ D $
1	5	15	0,10	0,30	0,20
2	8	13	0,26	0,56	0,30
3	10	10	0,46	0,76	0,30
4	15	8	0,76	0,92	0,16
5	12	4	1,00	1,00	0,00
<i>Total</i>	50	50			0,30



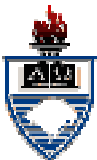
Como as amostras são grandes $n > 40$, o qui-quadrado deve ser utilizado. Assim:

$$d = 1,36 \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = 0,27$$



Conclusão

A menos de um erro de 5% (significância), posso afirmar que as opiniões dos diretores financeiros de empresas grandes e pequenas são divergentes.



Observação:

O SPSS utiliza a estatística Z (Smirnov, 1948), obtida seguinte forma:

$$Z = \text{Max}_j |D_j| \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

onde o nível de significância é calculado utilizando a aproximação de Smirnov, fornecida para o teste de uma amostra.



O teste U de Mann-Whitney



Objetivos

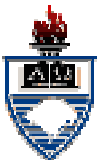
Comprovar se dois grupos independentes foram ou não extraídos da mesma população.

Requisitos

Grau de mensuração seja pelo menos ordinal.

Substitui

O teste t para amostras independentes.



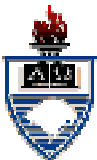
H_0 : A e B apresentam a mesma distribuição.

H_1 : A é maior do que B (teste unilateral).



Metodologia

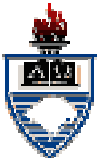
Sejam n_1 = número de casos no menor dos dois grupos independentes e n_2 = número de casos no maior grupo. Primeiramente combinam-se as observações ou escores de ambos os grupos, relacionando-os por ordem ascendente.



Nessa ordenação ascendente, consideram-se os valores algébricos do grupo $n = n_1 + n_2$, isto é, os postos mais baixos são atribuídos aos maiores valores (negativos se houver).



Focaliza-se agora um dos grupos, por exemplo, o grupo que apresenta n_1 casos. O valor de \mathcal{U} (a estatística teste) é o número de vezes que um escore no grupo com n_2 casos precede um escore no grupo com n_1 casos no grupo ordenado formado por $n = n_1 + n_2$ casos.



Exemplo



Suponha um grupo experimental com $n_1 = 3$ casos e um grupo de controle n_2 com 4 casos. Admita-se que os escores sejam os seguintes:

<i>Experimental</i>	9	11	15	
<i>Controle</i>	6	8	10	13

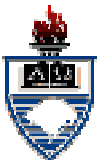


Para determinar \mathcal{U} , ordenam-se primeiro os escores de forma crescente, tendo o cuidado de identificar a qual grupo cada um pertence (\mathcal{E} ou \mathcal{C}):

6	8	9	10	11	13	15
\mathcal{C}	\mathcal{C}	\mathcal{E}	\mathcal{C}	\mathcal{E}	\mathcal{C}	\mathcal{E}



Considera-se agora o grupo de controle e conta-se o número de escores E que precedem cada escore do grupo de controle.



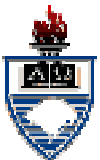
Nenhum score E precede o score C igual a 6. Isto também é verdade para o score $C = 8$. O próximo score C é 10 e é precedido por um score E . O último score C , o 13, é antecedido por dois scores E .



Assim, $\mathcal{U} = 0 + 0 + 1 + 2 = 3$. O número de vezes que um escore E vem antes de um escore C é igual a 3, isto é, $\mathcal{U} = 3$.



A distribuição amostral de \mathcal{U} , sob \mathcal{H}_0 , é conhecida e pode-se então determinar-se a probabilidade associada à ocorrência, sob \mathcal{H}_0 , de qualquer valor de \mathcal{U} tão extremo quanto o valor observado.

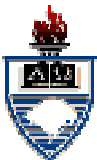


Amostras bem pequenas

Quando nem n_1 e nem n_2 são superiores a 8, pode-se utilizar o conjunto J (Siegel) para determinar a probabilidade exata associada à ocorrência, sob H_0 , de qualquer \mathcal{U} tão extremo quanto o valor observado.



O conjunto J é formado por seis tabelas separadas, uma para cada valor de n_2 , com $3 \leq n_2 \leq 8$. Para determinar a probabilidade, sob H_0 , associada aos dados é necessário entrar com os valores de n_1 , n_2 e \mathcal{U} .



No exemplo dado, tem-se: $n_1 = 3$, $n_2 = 4$ e $\mathcal{U} = 3$. A tabela de $n_2 = 4$ do conjunto \mathcal{J} mostra que $\mathcal{U} \leq 3$ tem probabilidade de ocorrência, sob \mathcal{H}_0 , de $p = 0,20 = 20\%$.



Observação 1:

As probabilidades fornecidas são unilaterais. Para um teste bilateral, deve-se duplicar o valor da probabilidade apresentado em cada tabela.



Observação 2:

Caso o valor observado de \mathcal{U} seja grande e não conste da tabela, existe a possibilidade de ter-se tomado o grupo “errado” no cálculo de \mathcal{U} .

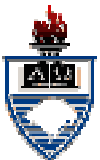
Neste caso, pode-se utilizar a transformação:

$\mathcal{U} = n_1 \cdot n_2 - \mathcal{U}'$, onde \mathcal{U}' é o valor que não foi encontrado na tabela.



Amostras médias

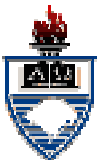
Se n_2 representar o tamanho da maior das duas amostras e for maior do que 8, o conjunto de tabelas J não poderá mais ser utilizado. Quando $9 \leq n_2 \leq 20$, pode-se utilizar tabela \mathcal{K} (Siegel).



Essa tabela fornece valores críticos de U para os níveis de significância de 0,001, 0,01, 0,025 e 0,05 para um teste unilateral. Para um teste bilateral, os níveis de significância são dados por: 0,002, 0,02, 0,05 e 0,10.



Este conjunto de tabelas fornece valores críticos de \mathcal{U} e não probabilidades exatas (como as J). Isto é, se um valor observado de \mathcal{U} , para $n_1 \leq 20$ e $9 \leq n_2 \leq 20$, não superar o valor da tabela, pode-se rejeitar \mathcal{H}_0 , a um dos níveis de significância indicados.



Amostras médias – Determinação de \mathcal{U}

Para valores grandes de n_1 e n_2 , o método para determinar \mathcal{U} é trabalhoso. Um processo alternativo com resultados idênticos, consiste em atribuir posto 1 ao valor mais baixo do grupo combinado ($n_1 + n_2$) valores, o posto 2 ao valor seguinte e assim por diante.



Então:

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - \mathcal{R}_1$$

ou

$$U = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - \mathcal{R}_2$$

onde \mathcal{R}_1 = soma dos postos atribuídos ao grupo n_1 e \mathcal{R}_2 = soma dos postos atribuídos ao grupo n_2 .



Por exemplo, se $n_1 = 6$ e $n_2 = 13$, um valor de $U = 12$ permite rejeitar H_0 ao nível $\alpha = 0,01$ em uma prova unilateral e rejeitar H_0 ao nível $\alpha = 0,02$ em uma prova bilateral.



Exemplo



Para ilustrar o processo vamos utilizar amostras pequenas. Assim:

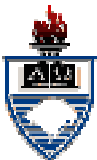
<i>Escore E</i>	<i>Posto</i>	<i>Escore C</i>	<i>Posto</i>
78	7	110	9
64	4	70	5
75	6	53	3
46	1	51	2
82	8		
<i>Soma</i>	$R_2 = 26$	<i>Soma</i>	$R_1 = 19$



Aplicando a fórmula anterior segue:

$$U = 4.5 + 5.(5 + 1) / 2 - 26 = 9$$

O menor dos dois valores de U é aquele cuja distribuição amostral constitui a base da tabela \mathcal{K} (Siegel).



Amostras grandes

Nem a tabela J e nem a K podem ser utilizadas quando $n_2 > 20$.

Mann e Whitney mostraram (1947), que à medida que n_1 e n_2 aumentam, a distribuição amostral de U tende rapidamente para a distribuição normal, com:



Média

$$\mu_{\mathcal{V}} = E(\mathcal{V}) = \frac{n_1 n_2}{2}$$

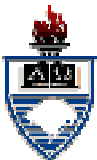
e

$$\sigma_{\mathcal{V}} = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

Então:

$$z = \frac{\mathcal{V} - \mu_{\mathcal{V}}}{\sigma_{\mathcal{V}}} = \frac{\mathcal{V} - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

É assintoticamente $\mathcal{N}(0; 1)$.



Empates

A prova de Mann-Whitney supõe que os escores representem uma distribuição basicamente contínua. Numa distribuição contínua a probabilidade de um empate é zero. Todavia, como a mensuração tem uma precisão limitada, os empates podem ocorrer.



Admite-se que as observações que estejam empatadas, tenham, na realidade, escores diferentes, e que esta diferença é muita pequena para ser detectada pelo instrumento de medida.



Assim quando ocorrem empates atribui-se a cada um dos valores empatados a média dos postos que lhes seriam atribuídas se não houvesse empate.



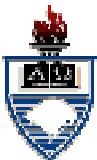
Se os empates ocorrem entre dois ou mais valores do mesmo grupo, o valor de \mathcal{U} não é afetado. Mas se os empates ocorrem entre duas ou mais observações envolvendo os dois grupos, então o valor de \mathcal{U} é afetado.



Embora, os efeitos práticos dos empates sejam desprezíveis existe uma correção para empates que deve ser utilizada com a aproximação normal para grandes amostras.



O efeito dos postos empatados modifica a variabilidade do conjunto de postos. Assim, a correção deve ser aplicada ao desvio padrão da distribuição amostral de U . Com esta correção o desvio padrão é dado por:



$$\sigma_v = \sqrt{\left(\frac{n_1 n_2}{n(n-1)} \right) \left(\frac{n^3 - n}{12} - \sum T \right)}$$

Onde $n = n_1 + n_2$

$$T = (t^3 - t) / 12$$

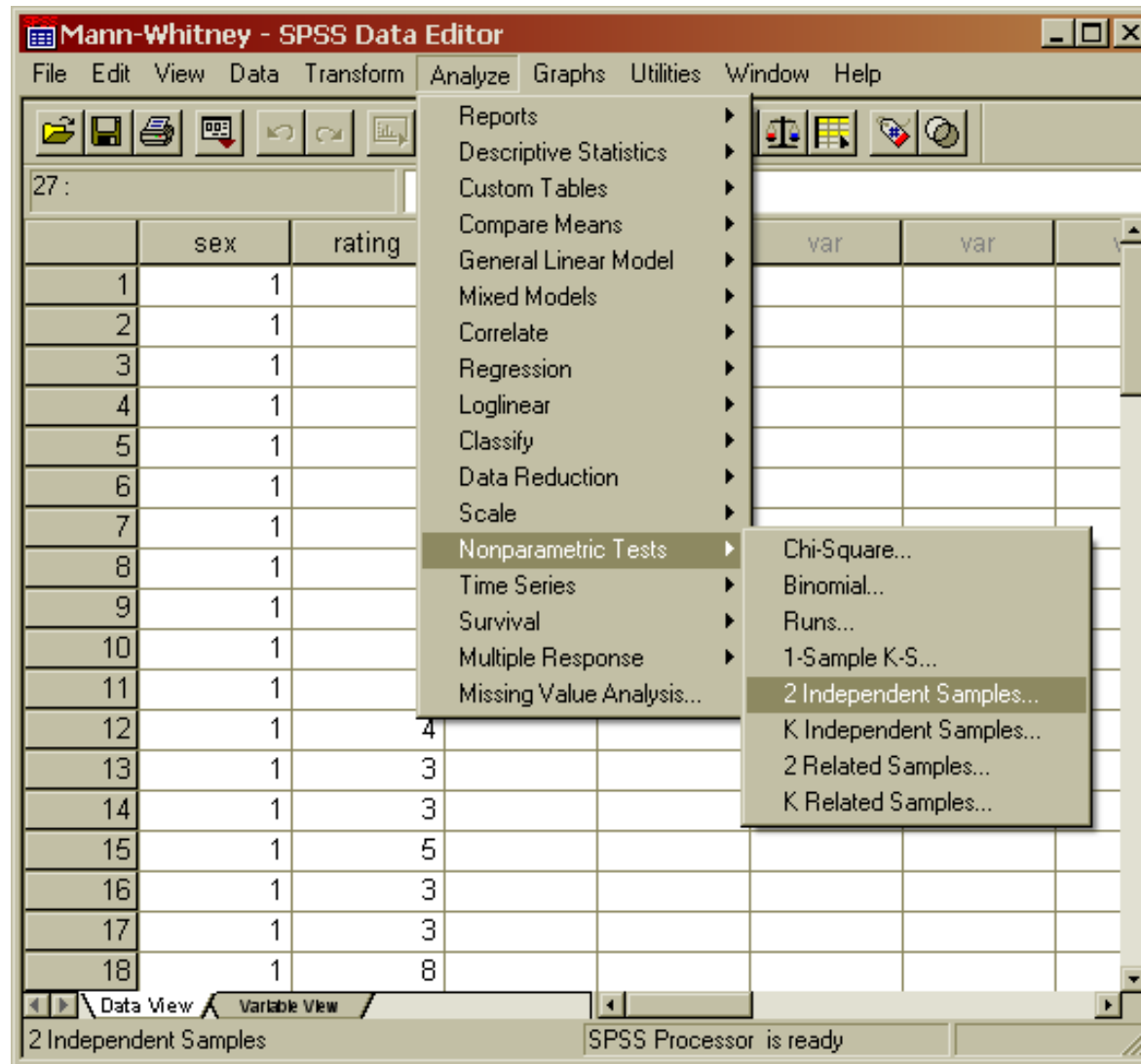
$t =$ número de escores empatados para um determinado posto.



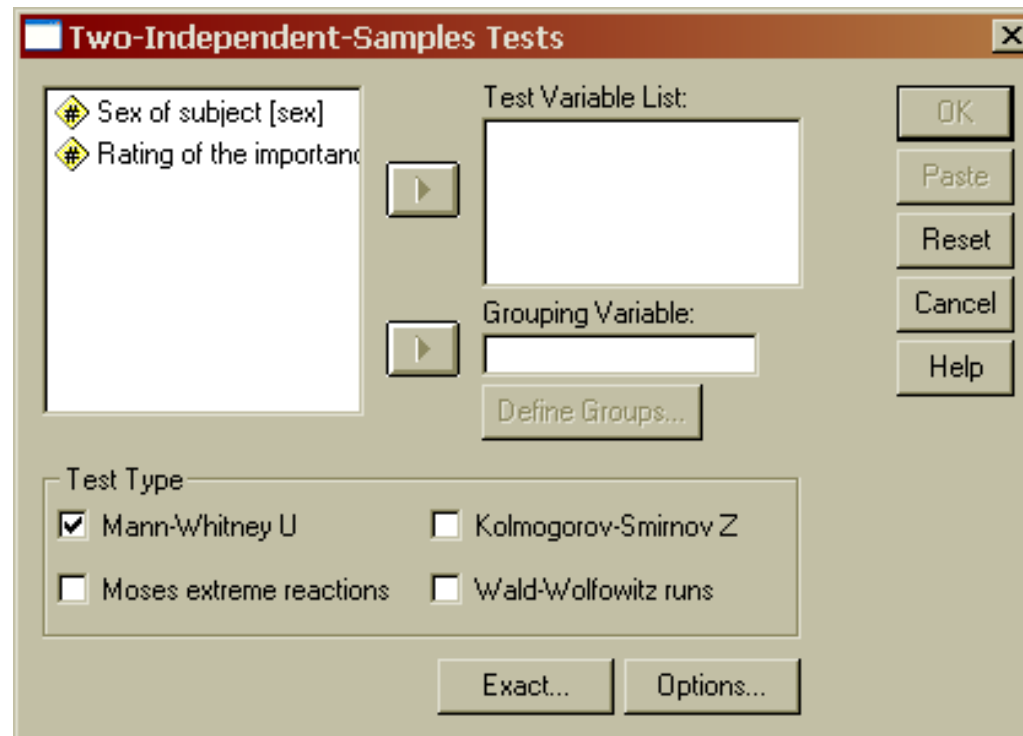
Exemplo



Clique conforme figura

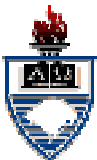


Isso abrirá a seguinte caixa de diálogos:

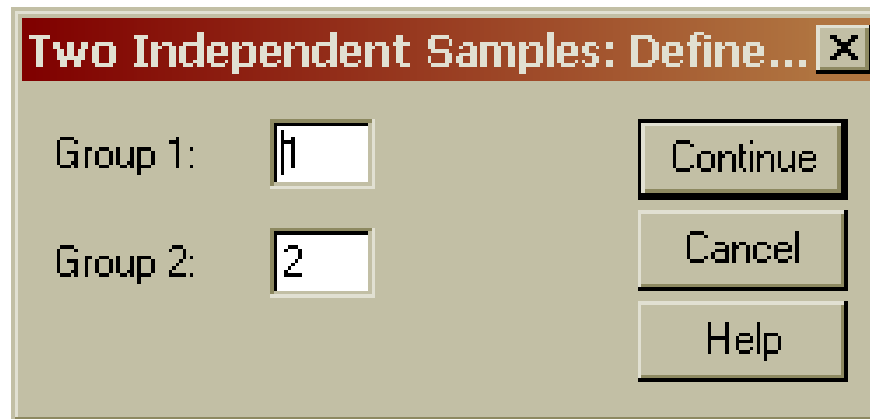


Coloque Rating ... Como Test Variable

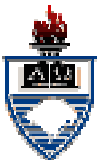
List e Sex of subject como Grouping Variable.



Clique em Define Groups



Entre os códigos, conforme planilha.



Test Statistics

	<i>RATING Rating of the importance of body as characteristic in a partner</i>
<i>Mann-Whitney U</i>	<i>147,500</i>
<i>Wilcoxon W</i>	<i>357,500</i>
<i>Z</i>	<i>-1,441</i>
<i>Asymp. Sig. (2-tailed)</i>	<i>,150</i>
<i>Exact Sig. [2*(1-tailed Sig.)]</i>	<i>,0157</i>

a Not corrected for ties.

b Grouping Variable: SEX Sex of subject



Conclusão:

Não é possível afirmar que existe diferença entre homens e mulheres quanto a importância que eles atribuem a forma do corpo do companheiro.

$$U = 147,50, n_1 = 20, n_2 = 20, p = 15,70\%$$

bilateral.

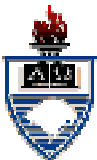


O teste de Wilcoxon



Objetivos

O teste de Wilcoxon investiga se existe diferença na posição de duas populações. Introduzido em 1945 com o nome de Teste da Soma das ordens (Rank Sum Test) destacou-se na área não paramétrica pelo seu poder.



Requisitos

As duas amostras são aleatórias e independentes.

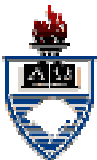
Substitui

O teste t para amostras independentes.



H_0 : Os grupos A e B são da mesma população.

H_1 : Os grupos A e B não são da mesma população.



Metodologia

Sejam X_1, X_2, \dots, X_m e Y_1, Y_2, \dots, Y_n
($m \geq n$). Forma-se um único grupo de
 $k = m + n$ observações ordenadas de forma
crescente. Define-se: $W = \sum_{j=1}^n O_j$

Onde O_j representa a ordem de Y_j na
classificação conjunta dos $k = m + n$ valores.



As hipóteses são:

$$\mathcal{H}_0: \Delta = 0$$

$$\mathcal{H}_1: \Delta > 0$$

$$\Delta < 0$$

$$\Delta \neq 0$$

Rejeitamos \mathcal{H}_0 se $\mathcal{W} \geq \mathcal{W}_\alpha$ onde $P(\mathcal{W} \geq \mathcal{W}_\alpha) = \alpha$ nas hipóteses unilaterais e metade desse valor na bilateral.



A hipótese unilateral é mais recomendável pois a ideia é de que uma população é em média maior do que a outra.



Observações:

- (i) Os valores máximo e mínimo de W ocorrem quando \mathcal{Y}_j ocupa respectivamente as n últimas ou as n primeiras observações na classificação conjunta $k = m + n$. Tais valores correspondem as seguintes situações:*



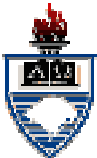
$$W_{\text{máx}} \rightarrow X X \dots X Y Y \dots Y$$

$$W_{\text{mín}} \rightarrow Y Y \dots Y X X \dots X$$

E assim, tem-se:

$$W_{\text{mín}} = \sum_{j=1}^n j = \frac{n(n+1)}{2} e$$

$$W_{\text{máx}} = \sum_{j=m+1}^k j = \frac{n(2m+n+1)}{2}$$

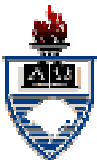


(ii) A média (mediana) dos possíveis valores de

$$\mathcal{W}, \text{ sob } H_0 \text{ é: } \mathcal{W}_{med} = \frac{n(m+n+1)}{2}$$

(iii) A amplitude do intervalo de variação de \mathcal{W}

$$\text{é: } A_{\mathcal{W}} = \mathcal{W}_{m\acute{a}x} - \mathcal{W}_{m\acute{i}n} = mn$$



(iv) \mathcal{W} é uma variável discreta.

(v) n é o tamanho da menor amostra.

(vi) A distribuição de \mathcal{W} , sob \mathcal{H}_0 , é simétrica em relação a sua média. Como consequência: \mathcal{W}_α

$$= n(m + n + 1) - \mathcal{W}_{1-\alpha}$$

Ou seja:

$$P(\mathcal{W} \leq \mathcal{W}_\alpha) = P[\mathcal{W} \leq n(m+n+1) - \mathcal{W}_{1-\alpha}]$$



E x e m p l o



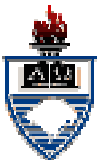
Suponha que se tenha dois grupos, um denominado de experimental e outro de controle, conforme valores da tabela.



<i>Valores</i>	<i>Experimental</i>		<i>Controle</i>	
	<i>Escore</i>	<i>Posto</i>	<i>Escore</i>	<i>Posto</i>
<i>1</i>	<i>25</i>	<i>18</i>	<i>12</i>	<i>10</i>
<i>2</i>	<i>5</i>	<i>3</i>	<i>16</i>	<i>15</i>
<i>3</i>	<i>14</i>	<i>13</i>	<i>6</i>	<i>4</i>
<i>4</i>	<i>19</i>	<i>17</i>	<i>13</i>	<i>12</i>
<i>5</i>	<i>0</i>	<i>1</i>	<i>13</i>	<i>11</i>
<i>6</i>	<i>17</i>	<i>16</i>	<i>3</i>	<i>2</i>
<i>7</i>	<i>15</i>	<i>14</i>	<i>10</i>	<i>7</i>
<i>8</i>	<i>8</i>	<i>6</i>	<i>10</i>	<i>8</i>
<i>9</i>	<i>8</i>	<i>5</i>	<i>11</i>	<i>9</i>
<i>Total</i>		<i>93</i>		<i>W = 78</i>

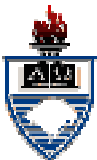


Como as duas amostras são iguais e não apresentam empates entre os grupos o valor da estatística de Wilcoxon é a menor das duas somas de postos obtida. Nesse caso, $W = 78$.



Empates

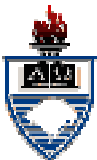
Quando ocorrem empates entre valores dos dois grupos, ou seja, entre X e Y , a média das ordens dos valores empatados é utilizada no cálculo de W e o cálculo é realizado da mesma forma que anteriormente.



Considere os seguintes valores de duas amostras X e Y :

	X	Y
1	2,3	1,8
2	3,2	2,3
3	3,8	2,3
4	4,5	3,2

Esses valores em um única amostra ordenada seriam:



<i>Valores</i>	1,8	2,3	2,3	2,3	3,2	3,2	3,8	4,5
<i>Grupo</i>	γ	χ	γ	γ	χ	γ	χ	χ
<i>Postos</i>	1	2	3	4	5	6	7	8
<i>Empates</i>	1	3	3	3	5,5	5,5	7	8

Então:

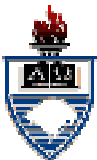
$$W = 1 + 3 + 3 + 5,5 = 12,5.$$

$$W = 3 + 5,5 + 7 + 8 = 23,5.$$



Observação:

Empates entre os valores de X e entre os valores de Y apenas não afetam o valor da estatística W , mas afetam a sua distribuição sob \mathcal{H}_0 .



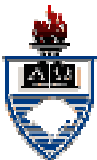
Aproximação pela normal

Quando n e m crescem os valores de \mathcal{W} podem ser aproximados por uma distribuição normal de média:

$$\mu_{\mathcal{W}} = E(\mathcal{W}) = \frac{n(m+n+1)}{2}$$

e desvio padrão:

$$\sigma_{\mathcal{W}} = \sqrt{\frac{mn(m+n+1)}{12}}$$

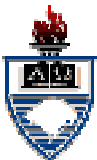


Aproximação pela normal

Quando n e m crescem os valores de \mathcal{W} podem ser aproximados por uma distribuição

normal de média:
$$\mu_{\mathcal{W}} = E(\mathcal{W}) = \frac{n(m+n+1)}{2}$$

e desvio padrão:
$$\sigma_{\mathcal{W}} = \sqrt{\frac{mn(m+n+1)}{12}}$$



Correção de continuidade

Em geral é recomendável aplicar-se uma correção de continuidade na aproximação pela normal. Essa correção consiste em somar ou subtrair o valor 0,5 ao valor de W conforme se esteja calculando valores na parte inferior ou superior da curva.

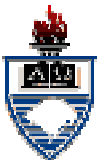


Por exemplo:

Se $m = 8$, $n = 4$ e $\mathcal{W} = 35$. O limite superior exato é 7,7%.

Aproximando pela normal, sem correção, temos valor-p = 6,32%

Utilizando a correção o valor passa para valor-p = 7,44%.



Distribuição sob \mathcal{H}_0

Para ilustrar a distribuição sob \mathcal{H}_0 de \mathcal{W} .

Considere-se $m = 4$ e $n = 2$. Com essa configuração o número de combinações (agrupamentos)

possíveis é:
$$\binom{6}{4} = \binom{6}{2} = 15$$



<i>Agrupamento</i>	W_0	<i>Agrupamento</i>	W_0
$Y Y X X X X$	3	$X Y X X X Y$	8
$Y X Y X X X$	4	$X X Y Y X X$	7
$Y X X Y X X$	5	$X X Y X Y X$	8
$Y X X X Y X$	6	$X X Y X X Y$	9
$Y X X X X Y$	7	$X X X Y Y X$	9
$X Y Y X X X$	5	$X X X Y X Y$	10
$X Y X Y X X$	6	$X X X X Y Y$	11
$X Y X X Y X$	7		



De onde obtém-se a distribuição:

W_0	$P(W = W_0)$	$P(W \geq W_0)$	$P(W \leq W_0)$
3	0,0667	0,0667	1,0000
4	0,0667	0,1333	0,9333
5	0,1333	0,2667	0,8667
6	0,1333	0,4000	0,7333
7	0,2000	0,6000	0,6000
8	0,1333	0,7333	0,4000
9	0,1333	0,8667	0,2667
10	0,0667	0,9333	0,1333
11	0,0667	1,0000	0,0667



Observações:

Considerando os resultados anteriores, tem-se:

$$(i) \mathcal{P}(\mathcal{W} = \mathcal{W}_0) = \mathcal{P}[\mathcal{W} = n(m + n + 1) - \mathcal{W}_0]$$

$$(ii) \mathcal{P}(\mathcal{W} \geq \mathcal{W}_0) = \mathcal{P}[\mathcal{W} \leq n(m + n + 1) - \mathcal{W}_0]$$

(iii) A distribuição é simétrica em torno da média

$$E(\mathcal{W}) = n(m + n + 1)/2$$



Empates:

No caso de observações empatadas a distribuição de W se altera e como consequência os níveis de significância das tabelas que são feitas sem empates se tornam apenas aproximações.



Exemplo:

Para ilustrar considere-se duas amostras de tamanhos $m = 3$ e $n = 2$, onde os valores dos postos 3 e 4 são iguais. Os possíveis arranjos bem como a distribuição da estatística W , para essa situação, são as seguintes:



<i>Agrupamento</i>	W_0	<i>Agrupamento</i>	W_0
$\gamma\gamma x x x$	3	$x\gamma x\gamma x$	5,5
$\gamma x\gamma x x$	4,5	$x\gamma x x\gamma$	7
$\gamma x x\gamma x$	4,5	$x x\gamma\gamma x$	7
$\gamma x x\gamma x$	6	$x x\gamma x\gamma$	8,5
$x\gamma\gamma x x$	5,5	$x x x\gamma\gamma$	8,5

A distribuição de W_0 , para essa situação, será:



Distribuição de \mathcal{W} sob \mathcal{H}_0

\mathcal{W}_0	$P(\mathcal{W} = \mathcal{W}_0)$	$P(\mathcal{W} \geq \mathcal{W}_0)$
3	0,10	1,00
4,5	0,20	0,90
5,5	0,20	0,70
6	0,10	0,50
7	0,20	0,40
8,5	0,20	0,20

Assim, por exemplo, se $\mathcal{W} = 8,5$,

$P(\mathcal{W} \geq 8,5) = 0,20$, mas pela tabela tem-se: 10%.



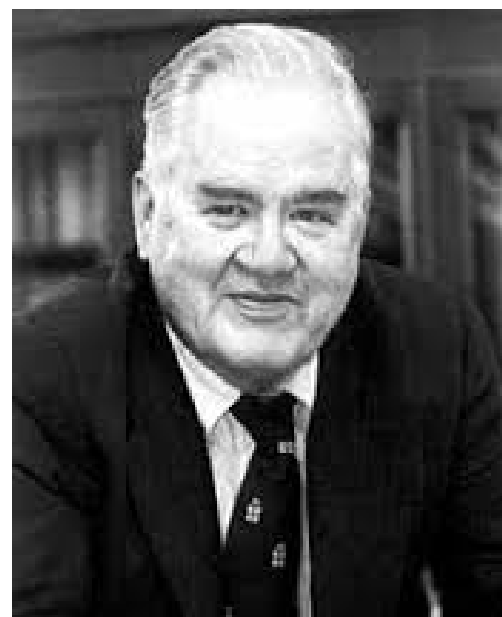
Exercício



Cinco mulheres e dez homens foram submetidos a um teste de aptidão para exercer determinada função. Eles foram avaliados por meio de uma escala de 0 a 10. Os resultados estão na tabela. Se você fosse o diretor com qual grupo trabalharia? Resolva utilizando o Excel e o SPSS.



O teste de Siegel-Tukey

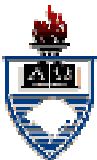


Objetivos

*Desenvolvido por Sidney Siegel (1916
1961) e John Wilder Tukey (1915 – 2000)
em 1960 o teste é utilizado para verificar se
existe uma diferença significativa entre as
variâncias de duas populações.*



O teste de Siegel-Tukey é um dos muitos testes de dispersão que são, também conhecidos como testes de escala ou espalhamento, utilizados para verificar se as variâncias de duas populações independentes são homogêneas.



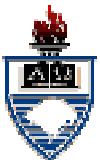
Hipóteses

$$\mathcal{H}_0: \sigma_1^2 = \sigma_2^2$$

$$\mathcal{H}_1: \sigma_1^2 \neq \sigma_2^2$$

$$\sigma_1^2 > \sigma_2^2$$

$$\sigma_1^2 < \sigma_2^2$$

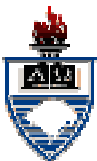


Suposições

Este teste tem por base as seguintes

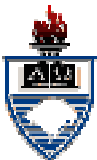
suposições:

- (i) Cada amostra foi selecionada aleatoriamente;*
- (ii) As duas amostras são independentes;*
- (iii) O nível de mensuração é pelo menos ordinal.*
- (iv) As duas populações tem a mesma mediana.*



Metodologia

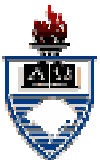
O procedimento de cálculo do teste de Siegel-Tukey para a equivalência da variabilidade é idêntico ao empregado no teste U de Mann-Whitney, exceto pelo fato de que o teste emprega um protocolo diferente para o cálculo dos postos.



Enquanto o procedimento do teste de Mann-Whitney é utilizado para identificar diferenças na tendência central (especificamente do valor mediano) o teste de Siegel-Tukey é projetado para identificar diferenças na variabilidade.



A suposição básica do teste toma por base a premissa de que na distribuição global dos $N = n + m$ escores, a distribuição dos escores no grupo com maior variância irá conter valores mais extremos (isto é, escores que serão muito altos ou muito baixos).



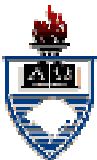
A suposição básica do teste toma por base a premissa de que na distribuição global dos $N = n + m$ escores, a distribuição dos escores no grupo com maior variância irá conter valores mais extremos (isto é, escores que serão muito altos ou muito baixos).



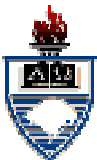
O procedimento para o cálculo dos postos para o teste de Siegel-Tukey é similar ao de Mann-Whitney, mas emprega uma forma diferente de atribuição dos escores.



Enquanto no teste de Mann-Whitney os escores são atribuídos de forma crescente ao leque dos valores de mínimo para máximo o de Siegel-Tukey a atribuição é feita de forma alternada.



O posto 1 é atribuído ao menor escore enquanto que o posto 2 ao maior. O posto 3 é atribuído ao segundo maior escore e o 4 ao segundo menor escore. O posto 5 é atribuído ao terceiro menor escore e 6 ao terceiro maior escore e assim sucessivamente.



E x e m p l o



Para testar novas drogas um grupo de 12 pacientes depressivos foram aleatoriamente colocados em um de dois grupos. Seis pacientes ficaram no Grupo 1 onde receberam o antidepressivo Elatrix por um período de seis meses. Os demais pacientes foram colocados no Grupo 2 e receberam o antidepressivo Euphyria durante o mesmo período.



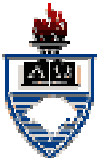
Suponha que o pesquisador testou o nível de depressão nos dois grupos antes de iniciar o tratamento e não verificou diferença. Após os seis meses os dois grupos foram avaliados por um Psiquiatra (que não conhecia quem era de qual grupo) sobre o nível de depressão.



Os escores que foram atribuídos as pacientes dos dois grupos foram.

<i>Grupo 1</i>	10	10	9	1	0	0
<i>Grupo 2</i>	6	6	5	5	4	4

O fato de que a média e a mediana dos dois grupos são equivalentes sugere que não existe diferença na eficácia das duas drogas.



Uma inspeção nos dois grupos, contudo revela que existe uma diferença aparente de variabilidade entre os dois grupos. Especificamente a droga Elatrix diminui a depressão em algumas pessoas enquanto aumenta em outras. O pesquisador quer comparar a variabilidade dos dois grupos.



O número total de sujeitos é $N = 12$, onde existe $n = 6$ sujeitos no Grupo 1 e $m = 6$ sujeitos no Grupo 2. Os dados para os dois grupos estão resumidos na Tabela.

<i>Sujeito</i>	5,1	6,1	4,1	5,2	6,2	3,2	4,2	1,2	2,2	3,1	1,1	2,1
<i>Escores</i>	0	0	1	4	4	5	5	6	6	9	10	10
<i>Postos</i>	1	4	5	8	9	12	11	10	7	6	3	2
<i>Postos</i>	2,5	2,5	5	8,5	8,5	11,5	11,5	8,5	8,5	6	2,5	2,5



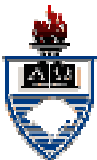
Assim as somas dos postos dos grupos 1 e 2 são iguais a:

$$\mathcal{R}_1 = 21 \text{ e } \mathcal{R}_2 = 57.$$

Os valores de \mathcal{U}_1 e \mathcal{U}_2 são dados por:

$$\mathcal{U}_1 = nm + \frac{n(n+1)}{2} - \mathcal{R}_1 = 6 \cdot 6 + \frac{6 \cdot (6+1)}{2} - 21 = 36$$

$$\mathcal{U}_2 = nm + \frac{m(m+1)}{2} - \mathcal{R}_2 = 6 \cdot 6 + \frac{6 \cdot (6+1)}{2} - 57 = 0$$

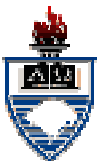


Observações:

(a) Note que os valores de \mathcal{U}_1 e \mathcal{U}_2 não podem ser negativos

(b) Se Mann-Whitney for utilizado a seguinte relação pode ser usada para checagem: $n.m = \mathcal{U}_1 + \mathcal{U}_2$.

(c) Para pequenos valores de n e m a significância de \mathcal{U} é obtida das tabelas.



(d) O valor de \mathcal{U} pode ser aproximado por uma normal de média $(nm)/2$ e variância $[nm(n + m - 1)]/12$.

(e) A correção de continuidade é obtida por $z = (|\mathcal{U} - (nm)/2| - 0,5)/DP$.



Para o exemplo dado, se a normal fosse adequada, o que não é o caso, o valor de z obtido seria igual a:

$$z = \frac{\left| \mathcal{U} - \frac{mn}{2} \right| - 0,5}{\sqrt{\frac{nm(n+m+1)}{12}}} = \frac{\left| 0 - \frac{6 \cdot 6}{2} \right| - 0,5}{\sqrt{\frac{6 \cdot 6(6+6+1)}{12}}} = -2,80$$

Se não fosse utilizado a CC o valor de z obtido seria igual a -2,88.



(f) A aproximação normal pode ser corrigida para empates quando eles forem excessivos.

Neste caso a variância é dada por:

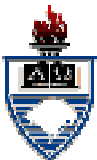
$$\mathcal{V}(\mathcal{V}) = \frac{nm(n+m+1)}{12} - \frac{nm \sum_{i=1}^s (t_i^3 - t_i)}{12(n+m)(n+m-1)}$$

Onde s = número de conjuntos de empates.



Para o exemplo dado, se a normal fosse adequada, o que não é o caso, e a correção para empates fosse utilizada o valor de z seria igual a:

$$z = \frac{U - \frac{mn}{2}}{\sqrt{\frac{nm(n+m+1)}{12} - \frac{nm \sum_{i=1}^s (t_i^3 - t_i)}{12(n+m)(n+m-1)}}} = \frac{0 - \frac{6.6}{2}}{\sqrt{\frac{6.6(6+6+1)}{12} - \frac{6.6 \cdot 30}{12(6+6)(6+6-1)}}} = -2,91$$



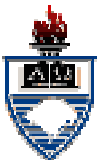
$$\begin{aligned}
 z &= \frac{\left| \mathcal{V} - \frac{mn}{2} \right| - 0,5}{\sqrt{\frac{nm(n+m+1)}{12} - \frac{nm \sum_{i=1}^s (t_i^3 - t_i)}{12(n+m)(n+m-1)}}} = \\
 &= \frac{\left| 0 - \frac{6 \cdot 6}{2} \right| - 0,5}{\sqrt{\frac{6 \cdot 6(6+6+1)}{12} - \frac{6 \cdot 6 \cdot 30}{12(6+6)(6+6-1)}}} = -2,83
 \end{aligned}$$

Onde $\sum_{i=1}^s (t_i^3 - t_i) = 5(2^3 - 2) = 30$ uma vez que

todos os 5 empates foram entre 2 valores.



(g) O teste de Siegel-Tukey é baseado na hipótese de que as medianas dos dois grupos são iguais. Se isto não acontecer é necessário ajustar os escores. Isto pode ser feito subtraindo a diferença entre as duas medianas do grupo com a mais alta mediana ou somando esta diferença ao grupo com a menor mediana.

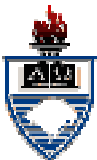


O teste de Moses



Objetivos

O teste de Moses serve para comparar as dispersões de duas populações. Ele não exige que as populações tenham a mesma mediana como o teste de Siegel-Tukey, isto é, ele pode ser aplicado quando as duas populações diferem na “posição”.

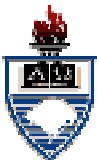


Requisitos

Grau de mensuração seja pelo menos ordinal.

Substitui

Ele é uma aplicação do teste de Wilcoxon às variâncias dos dados originais quando agrupadas de forma conveniente.



$$\mathcal{H}_0: \sigma_1^2 = \sigma_2^2$$

$$\mathcal{H}_1: \sigma_1^2 \neq \sigma_2^2$$

$$\sigma_1^2 > \sigma_2^2$$

$$\sigma_1^2 < \sigma_2^2$$



Metodologia

Sejam X_1, X_2, \dots, X_m e Y_1, Y_2, \dots, Y_n .

Escolhe-se um valor arbitrário $k \geq 2$ e dividimos os “ m ” valores X em m_1 grupos aleatórios com “ k ” elementos em cada um, desprezando sobras se existirem. Repetir o mesmo procedimento para os valores de Y , isto é, obtendo n_1 grupos.



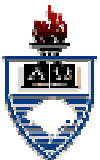
Para cada grupo calcula-se a soma dos quadrados dos desvios, ou seja:

$$C_i = \sum_{r=1}^k (X_{ir} - \bar{X}_i)^2 = \sum_{r=1}^k X_{ir}^2 - \frac{\left(\sum_{r=1}^k X_{ir}\right)^2}{k}$$

para $i = 1, 2, \dots, m_1$

$$D_j = \sum_{r=1}^k (Y_{jr} - \bar{Y}_j)^2 = \sum_{r=1}^k Y_{jr}^2 - \frac{\left(\sum_{r=1}^k Y_{jr}\right)^2}{k}$$

para $j = 1, 2, \dots, n_1$



Os valores C_i e D_j são as variações dos grupos i e j , ou ainda, as variâncias desses grupos multiplicadas por $(k - 1)$.

Assim para testar as variâncias aplica-se o teste de Man-Whitney aos valores C_i e D_j .



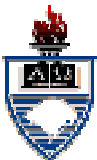
Observações:

(i) *Fixando um mesmo nível de significância as conclusões podem variar em função de:*

(a) *Número de grupos formados (k é qualquer);*

(b) *A estrutura dos subgrupos. Para um mesmo*

k , são possíveis $\binom{m}{k}$ e $\binom{n}{k}$ combinações.



(ii) *Não é recomendável testar vários valores de k até que se obtenha a conclusão mais adequada.*



Procedimento

O número de subamostras derivadas de cada Grupo não precisa ser o mesmo. O número de escores em cada subgrupo deve ser tal que os produtos n_1k e m_1k incluam o maior número possível de escores. A situação ótima seria $n_1k = n$ e $m_1k = m$ o que nem sempre será possível.



Após a determinação das amostras, determina-se a média de cada uma e em seguida a variação. Tendo os valores das variações o procedimento é o de Mann-Witney sobre os escores das variações.



E x e m p l o



Para testar novas drogas um grupo de 12 pacientes depressivos foram aleatoriamente colocados em um de dois grupos. Seis pacientes ficaram no Grupo 1 onde receberam o antidepressivo Elatrix por um período de seis meses. Os demais pacientes foram colocados no Grupo 2 e receberam o antidepressivo Euphyria durante o mesmo período.



Suponha que o pesquisador testou o nível de depressão nos dois grupos antes de iniciar o tratamento e não verificou diferença. Após os seis meses os dois grupos foram avaliados por um Psiquiatra (que não conhecia quem era de qual grupo) sobre o nível de depressão.



Os escores que foram atribuídos as pacientes dos dois grupos foram.

<i>Grupo 1</i>	10	10	9	1	0	0
<i>Grupo 2</i>	6	6	5	5	4	4

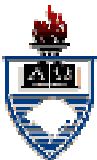
O fato de que a média e a mediana dos dois grupos são equivalentes sugere que não existe diferença na eficácia das duas drogas.



Contudo o exame dos dois grupos revela que pode haver uma diferença significativa na variabilidade dos dois grupos, isto é, o Grupo 1 apresenta uma variabilidade maior do que o Grupo 2. Assim, os dados sugerem que a droga Elatrix pode, de fato, diminuir a depressão em alguns sujeitos, mas aumentar em outros.



O número total de sujeitos envolvidos é $N = 12$ com $n = 6$ sujeitos no Grupo 1 e $m = 6$ sujeitos no Grupo 2. Estes conjuntos devem ser divididos em subamostras menores do que o tamanho dos grupos. Assim a melhor opção é escolher $k = 2$, tendo-se, desta forma, 3 amostras de tamanho 2 em cada grupo e nenhuma sobra.



As amostras devem ser selecionadas aleatoriamente de cada um dos dois grupos. Neste caso, teremos 3 amostras de tamanho 2 de cada um dos grupos. As amostras, as médias e as variações são apresentadas na tabela da próxima lâmina.



<i>Grupo 1</i>	<i>Média</i>	<i>Variação</i>	<i>Escores</i>
1, 10	5,5	40,5	4,5
10, 0	5	50	6
9, 0	4,5	40,5	4,5
		R_1	15
<i>Grupo 2</i>	<i>Média</i>	<i>Variação</i>	<i>Escores</i>
4, 4	4,5	0	1
5, 6	6	0,5	2,5
5, 6	4,5	0,5	2,5
		R_2	6



Neste caso os escores \mathcal{U}_1 e \mathcal{U}_2 serão:

$$\mathcal{U}_1 = nm + \frac{n(n+1)}{2} - \mathcal{R}_1 = 3.3 + \frac{3.(3+1)}{2} - 15 = 0$$

$$\mathcal{U}_2 = nm + \frac{m(m+1)}{2} - \mathcal{R}_2 = 3.3 + \frac{3.(3+1)}{2} - 6 = 9$$

O menor valor é a estatística \mathcal{U} , que neste caso, é igual a zero. Conforme tabela J, a significância deste resultado é de 0,05.



Assim é possível concluir que existe uma variabilidade maior, nos escores de depressão, no Grupo 1 que recebeu a droga Elatrix, do que no Grupo que recebeu a droga Euphyria (Grupo 2).



Aproximação pela Normal

Embora os grupos sejam pequenos demais para uma aproximação pela normal, isto será feito apenas para ilustração do procedimento.

Neste caso a estatística teste será:



$$z = \frac{v - \frac{mn}{2}}{\sqrt{\frac{nm(n+m+1)}{12}}} = \frac{0 - \frac{3.3}{2}}{\sqrt{\frac{3.3(3+3+1)}{12}}} = -1,96$$

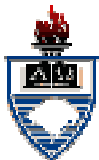
Este resultado concorda com o anterior e fornece uma significância de 5% bilateral.

Se a correção de continuidade for utilizada o valor de z passará para:



$$z = \frac{\left| \mathcal{U} - \frac{mn}{2} \right| - 0,5}{\sqrt{\frac{nm(n+m+1)}{12}}} = \frac{\left| 0 - \frac{3.3}{2} \right| - 0,5}{\sqrt{\frac{3.3(3+3+1)}{12}}} = -1,75$$

Neste caso este resultado não será mais significativo a 5% bilateral, mas apenas a 5% unilateral.



Referências:

MANN, Henry B., WHITNEY, Donald R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. Annals of Mathematical Statistics, v. 18, n. 1, p. 50-60, 1947.

MOSES, L. E. A Two-Sample Test. Psychometrika, v. 17, n. 3, 1952, p. 239-47. doi:10.1007/BF02288755.

SHESKIN, David J. Handbook of Parametric and Nonparametric Statistical Procedures. 4th ed. Boca Raton (FL): Chapman & Hall/CRC, 2007.



Referências:

SIEGEL, Sidney, TUKEY, John W. A Nonparametric Sum of Ranks Procedure for Relative Spread in Unpaired Samples. Journal of the American Statistical Association, v. 55, n. 291, Sep., 1960, p. 429-45.

WILCOXON, Frank. Individual comparisons by ranking methods (PDF). Biometrics Bulletin, v. 1, n. 6, p. 80-83, Dec 1945.

<http://www.real-statistics.com/non-parametric-tests/wilcoxon-rank-sum-test/>

