Testes não paramétricos

A designação de "teste não paramétrico" deve-se ao facto de não ser necessário especificar a distribuição da população de onde provém a amostra (até agora, na maior parte dos casos, admitimos que tal população tinha distribuição normal ou, pelo menos, aproximadamente normal). Os métodos não paramétricos usam procedimentos que são aplicáveis independentemente da distribuição da população; quando muito, são por vezes exigidas algumas hipóteses como a de simetria ou a de continuidade da distribuição. Alguns destes métodos podem ser aplicados a dados qualitativos (relembremos que, com excepção dos intervalos de confiança e testes para proporções, todos os procedimentos estudados até agora são aplicáveis apenas a dados quantitativos). Outra situação em que os testes paramétricos são úteis, é aquela em que a dimensão da amostra é muito pequena e não se conhece a distribuição exacta da população.

1 Teste de aleatoriedade de uma amostra (teste dos *runs*)

Objectivo e pressupostos: A primeira hipótese em que nos baseamos para fazer inferências para uma população a partir de uma sua amostra é a de que esta é aleatória. O teste dos *runs* permite testar a veracidade desta hipótese. Pode ser aplicado a qualquer tipo de dados.

Hipóteses em teste:

H_0: a amostra é aleatória;
H_1: a amostra não é aleatória.

Como funciona o teste: o teste dos *runs* baseia-se na análise de uma sequência de dois tipos de símbolos, digamos, A e B. Um *run* é uma subsequência de símbolos iguais. Por exemplo, a sequência

$$ABAABBBBAAAAABBABAAABABBB$$

tem 12 *runs*:

$$A \quad B \quad AA \quad BBBB \quad AAAAA \quad BB \quad A \quad B \quad AAA \quad B \quad A \quad BBB.$$

Se os dois símbolos se apresentarem de forma aleatória não deverão ocorrer sequências do tipo

$$BBBBBBBBBBBBBBBBBBBBBBBAAAAAAAAAAAAAAAAAA$$

ou

$$ABABABABABABABABABABABABABABABABABABABABAB.$$

Valores muito pequenos ou muito grandes[1] do número de *runs* levam à rejeição de H_0.

O SPSS processa o teste apenas para variáveis do tipo $Numeric$. Assim, se pretendermos verificar se uma amostra é constituída por indivíduos seleccionados aleatoriamente, devemos escolher uma variável deste tipo para verificar a aleatoriedade (sempre no ficheiro original). Por exemplo, suponhamos que o ficheiro de dados é constituído por registos de 100 indivíduos relativos às variáveis "sexo"(modalidades: M e F), "clube de futebol preferido"(modalidades: $Benfica$,

[1] As "barreiras" correspondentes são estabelecidas pelo nível de significância que se pretende utilizar na realização do teste.

Porto, *Sporting*, ...) e "idade". As variáveis "sexo"e "clube de futebol preferido"são de tipo "String"pelo que o SPSS não as considera quando pretendemos efectuar o teste dos *runs*. Neste caso, fazemos o teste com a variável "idade". Se o ficheiro de dados tem apenas variáveis do tipo *String*, temos que codificar uma delas e usar essa para realizar o teste.

O procedimento consiste em comparar cada valor da amostra com um valor previamente fixado (média, mediana, moda ou outro fixado pelo utilizador), que é designado por *Cut Point* no SPSS. Neste caso, podemos pensar, por exemplo, que o símbolo A corresponde a uma diferença positiva ou nula e o símbolo B a uma diferença negativa. Assim, utilizando a média como *Cut Point*, à amostra

$$1 \quad 2 \quad 6 \quad 2 \quad 0 \quad 0 \quad 1 \quad 3 \quad 0 \quad 3,$$

cuja média é 1.8, corresponde a sequência

$$B \; A \; A \; A \; B \; B \; B \; A \; B \; A.$$

Para realizar o teste no SPSS utilizamos o trajecto *Analyze* $\rightarrow$ *Nonparametric Tests* $\rightarrow$ *Runs*.

No output obtemos uma tabela onde entre outras informações surge o p-valor do teste efectuado: *Asymp. Sig.* $(2-tailed)$ (e *Exact Sig.* $(2-tailed)$, se seleccionarmos *Exact* antes do *OK* final).

2 Testes de ajustamento

2.1 Caso de uma amostra

Nesta secção estudamos testes que permitem verificar se a população de onde foi retirada a amostra tem determinada distribuição teórica (normal, exponencial, uniforme, ...).

2.1.1 Teste de ajustamento de Kolmogorov-Smirnov

Objectivo e pressupostos: O teste de Kolmogorov-Smirnov destina-se a averiguar se uma amostra pode ser considerada como proveniente de uma população com uma determinada distribuição. O teste é particularmente indicado para distribuições contínuas[2].

Hipóteses em teste:

H_0: a população tem uma determinada distribuição $\mathcal{D}$;
H_1: a população não tem a distribuição $\mathcal{D}$.

Como funciona o teste: comparam-se, para cada número real x, duas percentagens:

- a percentagem de valores da amostra inferiores ou iguais a x,
- a percentagem de valores da população inferiores ou iguais a x, admitindo que a população tem a distribuição $\mathcal{D}$ (o SPSS usa estimativas dos parâmetros desta distribuição para calcular tal percentagem).

[2] No entanto, no SPSS, o teste de Kolmogorov-Smirnov também está disponível para a distribuição discreta de Poisson.

Se o valor absoluto da maior das diferenças obtidas puder ser considerado suficientemente pequeno[3], então os dados levarão à aceitação da hipótese H_0.

No SPSS, o teste de Kolmogorov-Smirnov encontra-se em $Analyze \to Nonparametric\ Tests$, mas só está disponível para quatro distribuições: normal, uniforme, exponencial e Poisson. Como já foi referido acima, o SPSS usa estimativas dos parâmetros da distribuição em teste, não permitindo ao utilizador a especificação desses parâmetros. Assim, por exemplo, a hipótese "a classificação obtida pelos alunos na disciplina 1 tem distribuição $N(11.5, 3)$" **não** pode ser testada no SPSS com o teste de Kolmogorov-Smirnov. Por outro lado, a hipótese "a classificação obtida pelos alunos na disciplina 1 tem distribuição normal" pode ser testada no SPSS. Se aceitarmos esta hipótese, então a distribuição normal a considerar para a população é $N(\bar{x}, s_c)$, onde $\bar{x}$ e s_c são, respectivamente, a média e o desvio padrão corrigido da amostra.

Como consequência da utilização das estimativas acima referidas, o teste tende a aceitar a hipótese H_0 mais vezes do que deveria. Para resolver este problema foi proposta por Lilliefors uma correcção ao teste de Kolmogorov-Smirnov quando a distribuição em teste é normal. O teste de Kolmogorov-Smirnov com a correcção de de Lilliefors é então recomendado nesta situação. Está disponível no SPSS a partir de *Analyze* → *Descriptive Statistics* → *Explore*. Seleccionando *Plots* e optando por *Normality plots with tests*, o programa fornece o Q-Q plot Normal para a amostra e uma tabela de resultados de testes de ajustamento. Na figura seguinte apresentam-se outputs dos testes acima referidos.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Notas da disciplina 1	,069	70	,200*	,974	70	,145

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

One-Sample Kolmogorov-Smirnov Test

		Notas da disciplina 1
N		70
Normal Parameters[a,b]	Mean	11,531
	Std. Deviation	2,9112
Most Extreme Differences	Absolute	,069
	Positive	,054
	Negative	-,069
Kolmogorov-Smirnov Z		,574
Asymp. Sig. (2-tailed)		,897
Exact Sig. (2-tailed)		,874
Point Probability		,000

a. Test distribution is Normal.

b. Calculated from data.

Figura 1: Testes de normalidade.

[3] A "barreira" correspondente é estabelecida pelo nível de significância que se pretende utilizar na realização do teste.

O quadro *Tests of Normality* apresenta o teste de Kolmogorov-Smirnov com a correcção de Lilliefors e ainda outro teste de ajustamento desenvolvido especialmente para a lei normal, o teste de Shapiro-Wilk. Para amostras de dimensão superior ou igual a 30 aconselha-se o teste de Kolmogorov-Smirnov com a correcção de Lilliefors; para amostras de dimensão mais reduzida é mais indicado o teste de Shapiro-Wilk. Como aqui se trata de uma amostra de dimensão 70, a informação a reter é a de que o p-valor do teste é superior ou igual a 0.2. Este valor permite-nos aceitar a hipótese da normalidade da população para os níveis de significância habituais.

Observemos agora, no quadro *One-Sample Kolmogorov-Smirnov Test*, que o p-valor do teste de Kolmogorov-Smirnov sem a correcção de Lilliefors (*Exact Sig.*) é igual a 0.874. Este valor é muito superior a 0.2, o que vai de encontro ao foi referido anteriormente: o teste de Kolmogorov-Smirnov aceita mais facilmente a hipótese H_0 do que o teste de Kolmogorov-Smirnov com a correcção de Lilliefors.

Nota: No caso da lei exponencial, o teste só funciona bem para v.a.'s que tomem valores em $[0, +\infty[$. Assim, antes de efectuar o teste, devemos verificar se o extremo inferior da primeira classe está afastado de 0. Se assim for, criamos uma nova variável (em $Transform$) cujos valores são dados por $X - a$, onde X é a variável inicial e a é habitualmente o mínimo da amostra. Se o teste levar à aceitação de uma distribuição exponencial para a v.a. $X - a$, então podemos aceitar a hipótese de que X segue uma distribuição exponencial em $[a, +\infty[$.

2.1.2 Teste de ajustamento do qui-quadrado

Objectivo e pressupostos: O teste de ajustamento do qui-quadrado destina-se a averiguar se uma amostra pode ser considerada como proveniente de uma população com uma determinada distribuição sem restrições sobre esta.

Este teste também pode ser usado para verificar se as categorias de uma variável (mesmo do tipo *String*) estão equitativamente distribuídas.

Embora sendo de mais difícil implementação no SPSS, recorremos ao teste do qui-quadrado quando a distribuição teórica que pretendemos ajustar não é uma das quatro disponíveis no SPSS para o teste de Kolmogorov-Smirnov ou ainda quando, mesmo sendo uma destas, pretendemos especificar à partida os seus parâmetros.

Hipóteses em teste:

H_0: a população tem uma determinada distribuição $\mathcal{D}$;
H_1: a população não tem a distribuição $\mathcal{D}$.

Como funciona o teste: para a realização do teste, os dados têm que estar agrupados em k classes (intervalos ou categorias). No caso em que a distribuição $\mathcal{D}$ é contínua, tais classes podem ser baseadas nas classes do histograma. São comparadas duas quantidades:

- o número de valores observados em cada categoria (frequência observada, n_i);
- o número de valores que se teriam em cada categoria admitindo que a população tem a distribuição $\mathcal{D}$ (frequência esperada, e_i).

Se as diferenças entre $n_i - e_i,\ i = 1, \ldots, k$, forem "pequenas", então o teste levar-nos-á à aceitação de H_0.

Passos a dar para realizar o teste:

1. Especificar as classes e atribuir um valor a cada classe. Este valor pode ser, por exemplo, o seu ponto médio ou simplesmente $1, 2, 3, \ldots, k$. A este procedimento chamamos "categorização da variável".

   No SPSS, este procedimento corresponde à criação de uma nova variável em $Transform \to$ *Recode into Different Variables*. Na janela subsequente passamos para a direita a variável que se pretende categorizar. Na opção *Old and New Values* seleccionamos *Range* e escrevemos os extremos da primeira classe (o SPSS considera as classes abertas à esquerda e fechadas à direita). Em *Value* escrevemos a categoria correspondente (por exemplo, 1). A seguir clicamos em *Add*. Repetimos o processo até incluirmos todas as classes. Clicamos em *Continue* e escrevemos o nome da nova variável em *Name*. Um clique em *Change* seguido de *OK* cria a nova variável no ficheiro de dados.

2. Calcular os valores e_i (frequência esperada de cada classe). A frequência esperada de uma classe $]a, b]$ é dada por $n(F(b) - F(a))$, onde n é a dimensão da amostra e $F(q)$ representa a proporção de indivíduos da população com valores inferiores ou iguais a q, admitindo válida para tal população a distribuição $\mathcal{D}$ que estamos a testar. Para calcular estes valores podemos usar $Transform \to Compute\ Variable$, com *CDF and Noncentral CDF* em *Function Group* e $Cdf.*$ em *Functions and Special Variables* ($*$ corresponde à distribuição em teste).

3. $Analyze \to Nonparametric\ Tests \to Chi - Square$. Passamos a *nova variável* para *Test Variable List* e em *Expected Values* seleccionamos *Values*. Inscrevemos a frequência esperada da primeira classe, e_1, e *Add*. Repetimos o processo para todas as classes. Finalmente, *OK*.

Notas:

1. A soma das frequências esperadas tem que ser igual à soma das frequências observadas, isto é, igual a n. Para que isto se verifique, quando a distribuição $\mathcal{D}$ é contínua tomamos $e_1 = nF(a_1)$, sendo a_1 o extremo superior da primeira classe. Em todos os casos, a frequência esperada da última classe é igual a $n - soma\ das\ outras\ frequências\ esperadas$.

2. O p-valor do teste é calculado de forma aproximada. Considera-se que esta aproximação é boa desde que todas as frequências esperadas sejam superiores ou iguais a 5 e muito boa desde que todas as frequências esperadas sejam superiores ou iguais a 10. Assim, por vezes será conveniente reagrupar as classes primitivas; podemos, por exemplo, juntar numa só duas (ou mais) classes adjacentes.

3. O teste de ajustamento do quiquadrado também pode ser usado para verificar se as categorias de uma variável (mesmo do tipo *String*) estão equitativamente distribuídas (i.e., uniformemente distribuídas). Neste caso, basta seleccionar na janela do teste a opção *All categories equal* seguida de *OK*.

2.2 Caso de duas amostras independentes: teste de Kolmogorov-Smirnov

Objectivo e pressupostos: O teste de Kolmogorov-Smirnov para duas amostras independentes pretende averiguar se as duas amostras provêm de populações X e Y com a mesma distribuição. Estas amostras *não* têm que ter a mesma dimensão.

Hipóteses em teste:

H_0: X e Y têm a mesma distribuição;
H_1: X e Y não têm a mesma distribuição.

Como funciona o teste: para cada número real x, comparam-se as percentagem de valores de cada uma das amostras que são inferiores ou iguais a x.

Se o valor absoluto da maior das diferenças obtidas puder ser considerado suficientemente pequeno, então os dados levarão à aceitação da hipótese H_0.

Para realizar o teste de Kolmogorov-Smirnov para duas amostras independentes no SPSS, os dados correspondentes devem estar dispostos numa só coluna. Deve então criar-se outra coluna (com valores numéricos) que identifique a amostra de origem de cada uma das observações.

No SPSS, o teste de Kolmogorov-Smirnov para duas amostras independentes encontra-se em *Analyze* → *Nonparametric Tests* → *2 Independent Samples*. Selecciona-se a variável correspondente à coluna onde estão as duas amostras e passa-se para *Test Variable List*. Para *Grouping Variable* passa-se a variável que identifica a amostra (grupo) de origem de cada uma das observações. Finalmente identificam-se os dois grupos em *Define Groups*.

3 Testes de localização

A média é o parâmetro de localização mais frequentemente utilizado em inferência estatística. No entanto, a mediana, que também é uma medida de tendência central das distribuições, pode constituir uma alternativa à média. De facto:

- a mediana não é influenciada por observações muito grandes ou muito pequenas;
- quando as distribuições são assimétricas, a mediana situa-se numa posição mais próxima do valor mais observado, podendo por isso ter mais sentido como medida de tendência central;
- quando as distribuições são simétricas, a mediana e a moda coincidem, possuindo assim o mesmo mérito como medida de tendência central.

Os testes aqui apresentados são especialmente indicados nas situações em que as amostras são de dimensão reduzida e a população não pode ser considerada normal.

3.1 Localização de uma população: teste dos sinais

Objectivo e pressupostos: O teste dos sinais para a localização de uma população é um teste para a sua mediana (μ). A hipótese H_0 é a de que μ é igual a um determinado valor especificado pelo utilizador (μ_0).

Pressupõe-se que a distribuição da população é contínua.

Hipóteses em teste:

H_0: $\mu = \mu_0$;
H_1: $\mu \neq \mu_0$

O teste também pode ser unilateral, i.e., a hipótese H_1 também pode ser

$$H_1 : \mu < \mu_0 \text{ ou } H_1 : \mu > \mu_0.$$

Como funciona o teste: o teste baseia-se no facto de que, se H_0 for verdadeira, então aproximadamente metade dos valores observados serão inferiores a μ_0. Assim, consideram-se as diferenças $x_i - \mu_0$ (ou $\mu_0 - x_i$), $i = 1, 2, \ldots, n$, aceitando-se H_0 se o número de diferenças com sinal negativo for aproximadamente[4] igual ao número de diferenças com sinal positivo.

No SPSS, o teste dos sinais aparece apenas na sua versão para duas amostras emparelhadas (2 *Related Samples*). Assim, para o utilizarmos com uma única amostra, começamos por criar uma nova variável com n valores todos iguais a μ_0. O teste processa-se da seguinte forma:

- *Analyze* → *Nonparametric Tests* → *2 Related Samples*.
- Na janela que aparece a seguir, seleccionar duas variáveis: aquela cuja mediana se pretende testar e a nova variável. Enviá-las para o quadro *Test Pair(s) List*.
- em *Test Type* seleccionar *Sign* (podemos ainda optar por *Exact* e por mais informações em *Options*) e *OK*.

Exemplo[5] Sabe-se que o rendimento familiar mediano numa determinada região é 600 euros/mês. Uma amostra aleatória constituída por 12 famílias de uma vila daquela região revelou os seguintes rendimentos:

$$440,\ 466\ 482,\ 518\ 603,\ 617,\ 636,\ 727,\ 774,\ 824,\ 961,\ 1056.$$

Esta amostra permite concluir que o rendimento mensal mediano na vila em causa é diferente do rendimento mensal mediano da região onde se insere?

Denotando por μ o rendimento mensal mediano naquela vila pretendemos testar

$$H_0 : \mu = 600 \text{ contra } H_1 : \mu \neq 600.$$

Na figura 2 podemos observar o correspondente output do SPSS. No quadro *Frequencies* podemos ver que há 8 diferenças ($600 - \mathit{rendimentos}$) negativas, 4 positivas e 0 nulas.

No quadro *Test Statistics* observamos que o p-valor do teste bilateral que estamos a efectuar é igual a 0.388 (Exact Sig. (2-tailed)). Perante este valor não devemos rejeitar H_0, pelo que não podemos concluir que o rendimento mediano das famílias da vila em causa é diferente do rendimento mensal mediano da região onde se insere.

[4] Estabelecido pelo nível de significância do teste.
[5] Guimarães, R.C. e Sarsfield Cabral, J.A. (2007) Estatística (2^a edição) McGraw-Hill.

Descriptive Statistics

	N	Percentiles		
		25th	50th (Median)	75th
Rendimento das famílias	12	491,0000	626,5000	811,5000
Mediana sob H0	12	600,0000	600,0000	600,0000

Sign Test

Frequencies

		N
Mediana sob H0 - Rendimento das famílias	Negative Differences[a]	8
	Positive Differences[b]	4
	Ties[c]	0
	Total	12

a. Mediana sob H0 < Rendimento das famílias

b. Mediana sob H0 > Rendimento das famílias

c. Mediana sob H0 = Rendimento das famílias

Test Statistics[b]

	Mediana sob H0 - Rendimento das famílias
Exact Sig. (2-tailed)	,388[a]
Exact Sig. (1-tailed)	,194
Point Probability	,121

a. Binomial distribution used.

b. Sign Test

Figura 2: Teste dos sinais.

O valor 0.194 (Exact Sig. (1-tailed)) é o p-valor do teste unilateral em que a hipótese alternativa está de acordo com a tendência da amostra: há mais diferenças negativas do que positivas. Assim, a hipótese alternativa natural é a que vai no sentido $600 - \textit{rendimentos} < 0$, i.e., H_1: $\mu > 600$. Outra forma de concluir que este é o teste unilateral que naturalmente deve ser considerado, consiste em observar, na tabela *Descriptive Statistics*, que a mediana da amostra (626.5) é superior a 600.

Recorde-se que o p-valor do teste unilateral "sugerido" pela amostra é metade do p-valor do correpondente teste bilateral. Assim, sendo 1 o valor máximo possível para o p-valor do teste bilateral, tal valor máximo é 0.5 para o teste unilateral acima mencionado.

A referência Ties (ligações) que aparece no quadro $Frequencies$ indica o número de diferenças nulas, as quais não são a favor nem contra H_0. Quando há ligações, o procedimento habitual consiste em eliminar da amostra as observações que as provocam. Repetimos então o teste com a amostra resultante (necessariamente de menor dimensão).

NOTA: Se a distribuição da população for <u>simétrica</u> devemos usar o **teste de Wilcoxon**.

Relativamente ao teste do sinal, o teste de Wilcoxon tem a vantagem de ser mais potente, i.e., é menor a probabilidade de se cometer o erro de aceitar H_0 sendo H_0 falsa.

No SPSS, este teste processa-se como o teste dos sinais seleccionando *Wilcoxon* (em vez de *Sign*).

3.2 Duas amostras emparelhadas: teste dos sinais

Objectivo e pressupostos: Usa-se este teste quando se pretende analisar o efeito de determinado factor sobre a localização de uma distribuição contínua. Para efectuar o teste usam-se duas amostras emparelhadas (relativas ao "antes"e ao "depois").

É um teste para a diferença de medianas, que denotamos por μ_D.

Hipóteses em teste:

H_0: $\mu_D = 0$;
H_1: $\mu_D \neq 0$.

O teste também pode ser unilateral, i.e., a hipótese H_1 também pode ser

$$H_1 : \mu_D < 0 \quad \text{ou} \quad H_1 : \mu_D > 0.$$

Como funciona o teste: designando por $(x_1, x_2, \ldots, x_n)$ e $(y_1, y_2, \ldots, y_n)$ as duas amostras emparelhadas, consideram-se as diferenças $y_i - x_i$, $i = 1, 2, \ldots, n$. O teste processa-se depois como no caso do teste dos sinais para uma amostra.

NOTA: Se a distribuição da variável "diferenças" puder ser considerada simétrica, devemos usar o **teste de Wilcoxon**.

3.3 Duas amostras independentes: teste U de Mann-Whitney

Objectivo e pressupostos: O teste de Mann-Whitney é apropriado para averiguar se são iguais as medianas μ_X e μ_Y de duas populações contínuas e independentes, X e Y, resp.. As duas amostras envolvidas *não* têm que ter a mesma dimensão.

Hipóteses em teste:

H_0: $\mu_X = \mu_Y$;
H_1: $\mu_X \neq \mu_Y$.

O teste também pode ser unilateral, i.e., a hipótese H_1 também pode ser

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

Como funciona o teste: juntam-se as duas amostras numa só (amostra combinada) identificando a origem (X ou Y) de cada elemento desta nova amostra. Ordena-se a amostra combinada e observam-se as posições (*ranks*) ocupadas pelos elementos X e pelos elementos Y. Se X e Y ficarem aleatoriamente distribuídos, então a decisão final será favorável a H_0 (pois, sob H_0, a mediana da população subjacente à amostra combinada será igual às medianas μ_X e μ_Y). Caso contrário, rejeita-se H_0. Neste caso, se, por exemplo, as maiores observações estiverem mais frequentemente associadas à amostra de X, é possível inferir que $\mu_X > \mu_Y$.

X	Y	**X**	Y	Y	**X**	Y	**X**	**X**
1	2	3	4	5	6	7	8	9

Situação que favorece H_0

Y	Y	Y	**X**	Y	**X**	**X**	**X**	**X**
1	2	3	4	5	6	7	8	9

Situação que favorece H_1

O teste é desenvolvido com base na soma das posições (*Sum of Ranks*) ocupadas, na amostra ordenada, pelos elementos da amostra inicial de menor dimensão. Por exemplo, na primeira situação acima indicada, este valor é $2 + 4 + 5 + 7 = 18$.

No caso de haver uma ou mais observações iguais nas duas amostras iniciais (empates=*Ties*) é feita uma "correcção" usando as posições médias das observações empatadas.

No SPSS, o teste U de Mann-Whitney processa-se de modo análogo ao teste de Kolmogorov-Smirnov para duas amostras independentes.

Exemplo[6]: Num estudo sobre os efeitos de dois calmantes a administrar a reclusos violentos em situações de motim, obteve-se a colaboração de 15 voluntários. Em situações provocadas foram disparados dardos que injectavam as substâncias narcóticas (do tipo A em 8 reclusos, do tipo B nos outros 7) medindo-se o tempo, em segundos, que demoraram a fazer efeito. Os dados obtidos foram os seguintes:

Substância A	143.0	134.0	130.5	172.8	151.7	137.4	139.4	158.6
Substância B	155.6	149.7	217.1	153.2	136.4	154.0	138.6	

Com base nas duas amostras observadas, podemos concluir que há diferença significativa no tempo que as duas substâncias demoram a actuar?

O output obtido para o teste de Mann Whitney é o apresentado na figura 3.

Ranks

	Substância utilizada	N	Mean Rank	Sum of Ranks
Tempo (seg) até fazer efeito	A	8	7,00	56,00
	B	7	9,14	64,00
	Total	15		

Test Statistics[b]

	Tempo (seg) até fazer efeito
Mann-Whitney U	20,000
Wilcoxon W	56,000
Z	-,926
Asymp. Sig. (2-tailed)	,355
Exact Sig. [2*(1-tailed Sig.)]	,397[a]
Exact Sig. (2-tailed)	,397
Exact Sig. (1-tailed)	,198
Point Probability	,031

a. Not corrected for ties.

b. Grouping Variable: Substância utilizada

Figura 3: Teste U de Mann Whitney.

[6]Pestana, D.D. e Velosa, S.F. (2006) *Introd. à Probabilidade e à Estatística*, Fund. Calouste Gulbenkian, 2^a ed.

O p-valor do teste bilateral é 0.397 (Exact Sig. (2-tailed)) o qual, para os níveis de significância usuais, nos leva a aceitar a hipótese da igualdade das medianas. Concluímos assim que não há diferença significativa no tempo que os dois narcóticos levam a actuar.

Nota: O teste unilateral a considerar seria o teste de hipótese alternativa $\mu_A < \mu_B$ (observar os valores *Mean Rank* no quadro *Ranks* ou comparar as medianas das duas amostras iniciais).

4 Teste de independência do qui-quadrado

Objectivo e pressupostos: O teste de independência do qui-quadrado permite verificar a independência entre duas variáveis de qualquer tipo que se apresentem agrupadas numa tabela de contingência.

Este teste não deve ser utilizado se mais do que 20% das frequências esperadas sob a hipótese da independência forem inferiores a 5 ou se alguma delas for igual a 0.

Hipóteses em teste:

H_0: As variáveis são independentes;
H_1: As variáveis não são independentes.

Note-se que a hipótese alternativa não tem nenhuma indicação sobre o tipo de associação entre as variáveis.

Como funciona o teste: comparam-se as frequências observadas de cada uma das $p \times q$ células, n_{ij}, com as correspondentes frequências esperadas sob a hipótese da independência, e_{ij}, através do valor

$$\chi^2 = \sum_{i=1}^{p} \sum_{j=1}^{q} \frac{(n_{ij} - e_{ij})^2}{e_{ij}},$$

que é usado para o cálculo do coeficiente de contingência de Pearson. Se este valor é suficientemente pequeno[7], o que significa que as diferenças $n_{ij} - e_{ij}$ são pequenas, então somos conduzidos à aceitação de H_0.

No output do SPSS, o valor χ^2 é designado por *Pearson Chi-Square*.

Como foi referido acima, este teste não deve ser utilizado se mais do que 20% das frequências e_{ij} forem inferiores a 5 ou se alguma delas for igual a 0. Se fizer sentido, podemos tentar ultrapassar este problema agregando classes adjacentes.

O caso especial das tabelas 2×2: o valor χ^2 deve ser "corrigido"por

$$\chi^2(corrigido) = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(|n_{ij} - e_{ij}| - 0.5)^2}{e_{ij}}.$$

No output do SPSS, este valor é designado por *Continuity Correction*. Além disso deve ter-se em consideração o seguinte:

[7] A "barreira" correspondente é estabelecida pelo nível de significância do teste

- Quando $n > 40$, devemos usar o p-valor correspondente a *Continuity Correction*.
- Quando $20 \leq n \leq 40$, podemos usar o p-valor correspondente a *Continuity Correction*, desde que nenhuma das frequências esperadas seja inferior a 5. Se isto acontecer, devemos usar um teste alternativo: o **teste exacto de Fisher**.
- Quando $n < 20$, devemos usar o teste exacto de Fisher em qualquer caso.

O teste de independência do qui-quadrado está disponível no SPSS em *Analyse* → *Descriptive Statistics* → *Crosstabs* → *Statistics* → *Chi-square*.

Este trajecto fornece, no output, todos os valores acima indicados, incluindo os referentes ao teste exacto de Fisher, quando tal se justifica.

Na figura 4 apresenta-se um exemplo de output para uma tabela 2×2.

grupo * resposta Crosstabulation

Count

		resposta		
		nao	sim	Total
grupo	A	6	3	9
	B	16	0	16
Total		22	3	25

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	6,061[b]	1	,014	,037	,037
Continuity Correction[a]	3,315	1	,069		
Likelihood Ratio	6,889	1	,009	,037	,037
Fisher's Exact Test				,037	,037
N of Valid Cases	25				

a. Computed only for a 2x2 table

b. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 1,08.

Figura 4: Teste dos sinais.

Como $n = 25$ está entre 20 e 40 e há frequências esperadas inferiores a 5, devemos usar o teste exacto de Fisher. O p-valor a considerar é então 0.037, perante o qual devemos rejeitar a hipótese da independência para o nível de significância usual de 0.05. Assim, podemos dizer que a resposta dos indivíduos é influenciada pelo grupo a que pertencem.