



## Introdução à análise estatística com SPSS

### Guião nº4: Comparação de médias de duas amostras independentes

#### Investigação sobre a performance de um fármaco

Uma equipa de investigadores acredita que uma pequena dose de um determinado fármaco muda a capacidade das pessoas tomarem decisões correctas. A experiência consistiu em ter um grupo experimental que tomou o fármaco e um grupo controlo que tomou um placebo. Os scores obtidos das decisões correctas foram registados e os investigadores acreditam que a média dos scores do grupo experimental seja diferente da média do grupo de controlo. Os resultados estão na tabela 1.

**Tabela 1-** Resultados experimentais obtidos sobre a performance de um fármaco.

Grupo experimental				Grupo controlo			
Sujeito	Scores	Sujeito	Scores	Sujeito	Scores	Sujeito	Scores
1	390	8	425	15	446	22	440
2	494	9	421	16	749	23	471
3	386	10	407	17	599	24	501
4	323	11	386	18	460	25	492
5	660	12	550	19	390	26	392
6	406	13	470	20	477	27	578
7	345	14	393	21	556	28	398

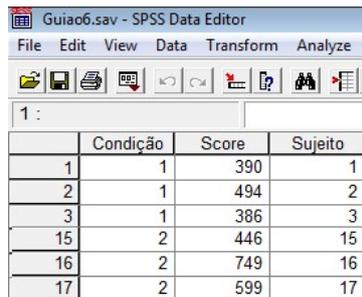
Identifique as variáveis e as hipóteses estatísticas em estudo.

#### Abrir o SPSS e construir o ficheiro com os dados

Abrir o SPSS e depois a opção **type in data** e clicar na janela do **variable view** para criar as variáveis. Como o problema é de amostras independentes é necessário criar três variáveis: a primeira variável designada por *condição*, na coluna do **label** indicará *condição da experiência* e na coluna **values** indicará que o 1=fármaco; 2=placebo; a segunda variável designada por *score* e na coluna do **label** indicará *score obtido na*



*experiência*. A terceira variável vai ser criada para indicar o elemento da amostra a que corresponde o tempo com o nome *sujeito* e na coluna do **label** indicará *elemento da amostra*. Na coluna **decimals**, colocar o valor 0 já que os dados não têm casas decimais. Depois da inserção dos dados na janela do **data view**, o resultado final deve ser como o apresentado na figura 1.



	Condição	Score	Sujeito
1	1	390	1
2	1	494	2
3	1	386	3
15	2	446	15
16	2	749	16
17	2	599	17

**Figura 1-** Representação dos 3 primeiros elementos do grupo experimental e do grupo de controlo, respectivamente.

## Explorar os dados

A primeira etapa é sempre examinar os dados e verificar se existe algum comportamento estranho.

O objectivo é criar uma tabela com médias e desvios-padrão, com a indicação sobre a distribuição dos dados através do histograma e da caixa de bigodes. Para o cálculo das médias e desvios-padrão, utiliza o comando **analyse/descriptive statistics /explore**. A variável **dependent** é o *score* e o **factor list** (variável independente que contém os grupos participantes) é a *condição*. No quadro **label cases by** colocar a variável *sujeito*. Na opção **statistics**, clicar na opção **descriptives**. Na opção **plots**, desactivar o **stem-and leaf** e activar a opção **normality plots with tests**. Depois clicar no **continue** e no **ok**.

Preencha a seguinte tabela:



**Tabela 2** - Resultados da estatística descritiva para variável *score* nos grupos experimental e de controle.

Grupos	n	Média	DP	Mediana	P <sub>25%</sub>	P <sub>75%</sub>
Experimental						
Controlo						

DP: Desvio-padrão

Responda às seguintes questões:

- Acha que os resultados descritivos apoiam a hipótese científica?
- Da análise da caixa de bigodes, existem outliers e extremos? No caso afirmativo, que valores têm os outliers/extremos.
- Da análise dos testes de ajustamento, que pode concluir para um  $\alpha=0.05$ ?

### **Aplicar o teste paramétrico t-student para diferença de médias, amostras independentes**

Vamos aplicar o teste t-student para diferença de médias, amostras independentes, sem eliminar os outliers e partir do princípio que os dados seguem uma distribuição normal.

Através do comando **analyse/compare means/independent samples t test**, o quadro referente ao **test variable** é utilizado para a variável dependente (*score*) e o quadro do **grouping variable** para a variável independente (*condição*). Ao clicar em **define groups**, iremos definir que o *grupo 1=1* (grupo experimental) e que *grupo 2=2* (grupo de controlo). Clicar no **continue** e depois no **OK**.

Responda às seguintes questões:

- Com base no teste de Levene, qual das linhas do teste t se deve utilizar?
- Escreva o valor do *t* (t), os graus de liberdade (df) e o valor do p-value (sig 2-tailed).



- Considerando um  $\alpha=0.05$ , será que os investigadores têm razão? Justifique analisando os valores amostrais obtidos na tabela 2.

### **Aplicar o teste não paramétrico U de Mann-Whitney para diferença de medianas, amostras independentes**

Como os dados não seguem distribuição normal, deve-se utilizar um teste não paramétrico.

Através do comando **analyse/non parametric tests/ 2-independent samples**, o quadro referente ao **test variable** é utilizado para a variável dependente (*scores*) e o quadro do **grouping variable** para a variável independente (*condição*). Ao clicar em **define groups**, iremos definir que o *grupo 1=1* (grupo experimental) e que *grupo 2=2* (grupo de controle). Clicar no **continue**. Na opção **exact** escolher o método *exact*. Clicar no **continue** e depois no **OK**. e depois no **OK**.

Responda às seguintes questões:

- Escreva o valor do U e o valor do p-value (sig 2-tailed). Considerando um  $\alpha=0.05$ , será que os investigadores têm razão?
- O resultado do teste U de Mann-Whitney é idêntico ao do teste t? Justifique em que circunstâncias irá ocorrer diferenças no p-values do teste de U de Mann-Whitney e no teste t.

### **Eliminação dos outliers/extremos**

A selecção de dados é vista como a criação de um filtro e é realizada através do comando: **data/selected cases**. A seguir temos que clicar em **if condition is satisfied** e depois no **if**. No novo quadro, a condição que vamos impor é: *score <600* no painel



direito. Depois é clicar no **Continue** e no **OK**. Sempre que se cria um filtro é necessário verificar se ficou bem feito. Verifique.

Repita as etapas da exploração dos dados e da aplicação dos testes e responda às seguintes questões:

- Da análise dos testes de ajustamento, que pode concluir?
- A conclusão relativa ao teste t continua válida?
- A conclusão relativa ao teste U de Mann-Whitney continua válida?
- Se os investigadores mudassem a questão de investigação para: “a média dos scores do grupo experimental é inferior á média do grupo de controlo”. Considerando um  $\alpha=0.05$ , será que os investigadores têm razão?

Num relatório ou artigo os resultados devem ser apresentados e escritos da seguinte forma:

**Tabela 3** - Resultados para a variável *score* nos grupos experimental e de controlo.

Grupos*	n	Média (M)	Desvio-padrão (SD)
Experimental			
Controlo			

\* teste t de amostras independentes;  $p=$  (2-tailed)

Complete os espaços:

O score médio da tomada de decisões do grupo de experimental ( $M=$  ;  $SD=$  ) foi significativamente diferente do score médio do grupo de controlo ( $M=$  ;  $SD=$  ):  $t(24) =$  ;  $p=$  (2-tailed);  $\alpha = 0.05$ , após remoção de dois sujeitos da análise. Este resultado confirma que o fármaco afecta a velocidade da tomada de decisões.