ANÁLISE DE DADOS

«A terra é redonda (p<0,05)»

Professor Luís Loureiro

SUMÁRIO:

Formulário:

Intervalo confiança para a média (μ)	Intervalo confiança para proporção (π)
$\mu \in \pm \left]\bar{x} \pm t\frac{\hat{s}}{\sqrt{n}}\right[$ onde t é tal que P (-t<t<t) = 1-α $\hat{s} = s\sqrt{\frac{n}{n-1}}$	$\pi \in \pm \left]p \pm z\sqrt{\frac{pq}{N}}\right[$ onde z é tal que P (-z<z<zt) = 1-α
Teste *t* de uma média	Teste *t* (amostras independentes)
$t_o = \frac{\bar{x} - \mu_0}{\frac{\hat{s}}{\sqrt{n}}}$ *Graus de Liberdade* = n - 1	$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{(n_1-1)\hat{S}_1^2 + (n_2-1)\hat{S}_2^2}{n_1+n_2-2}\right]\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$ ou $t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{n_1S_1^2 + n_2S_2^2}{n_1+n_2-2}\right]\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$ *Graus de Liberdade* = n_1+n_2 - 2
Teste *t* (amostras emparelhadas)	Teste Z (diferença Proporções)
$t = \frac{\sum d}{\sqrt{\frac{n\sum d^2 - (\sum d)^2}{n-1}}}$ ou $t = \frac{\bar{d}}{\frac{\hat{S}_d}{\sqrt{n}}}$ $\hat{s} = s\sqrt{\frac{n}{n-1}}$ *Graus de Liberdade* = n - 1	$z = \frac{p_1 - p_2}{\sqrt{pq\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$ $p = \frac{x_1 + x_2}{n_1 + n_2}$
Teste *t* de significância de r	Teste U de Mann-Whitney
$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$ *ou* $t = r\sqrt{\frac{n-2}{1-r^2}}$ *Graus de Liberdade* = n-2	$U_1 = n_1n_2 + \left[\frac{n_1(n_1+1)}{2}\right] - R_1$ $U_2 = n_2n_1 + \left[\frac{n_2(n_2+1)}{2}\right] - R_2$ $z = \frac{U_o - \frac{n_An_B}{2}}{\sqrt{\frac{(n_A)(n_B)(n_A+n_B+1)}{12}}}$
Teste χ2 (tabelas 2x2)	Teste χ^2 com correcção de continuidade
$\chi^2 = \Sigma\frac{(F_o - F_e)^2}{F_e}$ Graus de Liberdade = (C-1)(L-1)	$\chi^2_c = \Sigma\frac{(\lvert F_o - F_e\rvert - 0{,}5)^2}{F_e}$ Graus de Liberdade = (C-1)(L-1)

«Uma informação completa sobre o meio que nos rodeia seria o
mesmo que ter um mapa à escala de um para um».
Borrow (2005)

1. Introdução ao conceito de inferência estatística

O que queremos dizer quando falamos de **inferência**? A palavra significa *levar a*, *concluir*. Por exemplo, com o objectivo de saber quanto tempo gastam em média os enfermeiros dos HUC a dar banho a um doente acamado, solicitámos a 150 enfermeiros dessa população (seleccionados aleatoriamente) que registassem o tempo que gastavam a dar banho a um doente acamado. Com bases nos dados colhidos nessa amostra, calculámos a média e o desvio padrão, tendo obtido $\overline{x}$ = 24,35 minutos e s = 4,02 minutos.

Estas estatísticas designam-se de estatísticas amostrais pois referem-se à amostra (n = 150), no entanto, o nosso objectivo era saber a média do tempo gasto pelos enfermeiros dos HUC, ou seja a população desse hospital. Podemos então assumir ou *concluir* esta média amostral ($\overline{x}$ = 24,35 minutos) como sendo uma estimativa do parâmetro populacional (μ - símbolo da média populacional), ou seja a média da população?

Um outro exemplo. Imagine que lhe oferecem uma moeda de 1€ dizendo que é *equilibrada*, mas se a lançar 10 vezes consecutivas obtém sempre 10 caras. Pensará que das duas uma, ou a moeda é *equilibrada*, ou não é *equilibrada*. Então deita mãos ao conhecimento que tem das probabilidades e verifica que a probabilidade de obter em 10 lançamentos de uma moeda honesta, 10 caras consecutivas, é 0,001, ou seja, 0,01%. O que decide? Decerto que se obtiver efectivamente 10 caras, o mais provável é que não seja honesta.

Estes dois exemplos ajudam-nos a introduzir nas duas formas mais usuais de inferência, a **estimação** e os **testes de hipóteses**. Na primeira procuramos estimar os valores dos parâmetros populacionais com a base nas estatísticas amostrais. Na segunda forma de inferência, lançamos uma hipótese e face aos dados tomamos uma decisão, isto é, com base nos dados amostrais avaliamos a probabilidade de determinadas diferenças ou relações (observados nas distribuições amostrais) serem devidas ao acaso ou à existência real na população, sempre tendo por base observações feitas em amostras.

1.1. Distribuições amostrais

É do seu conhecimento que quando realizamos uma investigação, seja no âmbito da Enfermagem ou noutro domínio científico, trabalhamos com amostras (subconjuntos da população). É com base nos dados colhidos junto dessas amostras, muitas vezes recorrendo apenas a uma amostra, que podemos chegar a conclusões e tomar decisões relativamente à população. Pode então perguntar: - *porque não necessitamos de estudar toda a população, bastando tomar uma amostra para fazer um julgamento sobre a população?*

As sondagens políticas realizadas, por exemplo, no dia das eleições (designadas *à boca da urna*) são um bom exemplo. O que verificamos é que a empresa que faz a sondagem, seleccionando **aleatoriamente** uma ou várias amostras **representativas** dos eleitores que foram votar e depois com base nos dados, produz uma afirmação, que em estatística se designa de **estimativa**, para a % de votantes nesse partido.

Muitas vezes verificamos que a estimativa não é só centrada num único valor, mas assume a forma de um intervalo; melhor ainda, quando a estimativa é centrada num único valor, muito raramente os valores são coincidentes com o que depois vem a acontecer no fim de se contarem todos os votos, a verdadeira proporção de votos nesse partido, o confronto da estimativa com a realidade.

Como é possível assumir que a proporção amostral de votantes (designada de p) num partido seja uma estimativa da proporção (π) de votantes nesse partido? Neste caderno vamos apenas apresentar a distribuição amostral para a média, ou seja, esqueçamos as proporções.

Relativamente ao exemplo inicial, com que base podemos assumir que a média do tempo gasto pelos enfermeiros a dar banho a doentes acamados, numa amostra de 150 enfermeiros, seja a estimativa do parâmetro populacional, ou melhor, o tempo gasto pela população de enfermeiros dos HUC? Vamos "solucionar" a questão apresentando alguns conceitos basilares ao funcionamento da inferência.

Um conceito que já encontrou anteriormente referenciado foi o de **amostras aleatórias**. O que significa? Uma amostra aleatória, significa que cada elemento da população tem uma probabilidade igual e independente de ser seleccionado.

A **aleatoriedade** está relacionada com a utilização das técnicas de inferência estatística, mas basta estar em presença de uma amostra aleatória para que se possa fazer uma inferência relativamente à população? De facto não, a aleatoriedade ombreia com um outro conceito, o de **representatividade**, ou seja, em investigação devemos trabalhar não só com amostras aleatórias, porque as técnicas de inferência o exigem, mas também com amostras representativas para que se possam fazer **generalizações** para a população.

Voltando ao exemplo do tempo gasto pelos enfermeiros a dar banho a um doente acamado, cujas estatísticas são apresentados no quadro 1.

Quadro 1 – Estatísticas resumo relativas à variável: tempo gasto a dar banho a um doente acamado

Variável:	n	$\overline{x}$	s
Tempo gasto a dar banho a doente acamado	150	24,35	4,02

Podemos afirmar que a média[1] de tempo gasto a dar banho a um doente acamado pelos enfermeiros dos H.U.C. é em média de 24,35 minutos, com um desvio padrão[2] (s) de 4,02 minutos? O que nos permite fazer tal afirmação? Ou seja, uma generalização quando não sabemos, na maioria das vezes a forma da distribuição da variável na população, neste caso do tempo gasto a dar banho aos doentes acamados? Sabemos para já que a finalidade da amostragem (trabalhar com amostras) é obter os valores que utilizamos posteriormente como estimativas dos parâmetros populacionais. Assim, podemos utilizar a média desta amostra ($\overline{x}$ = 24,35) como um estimador da média da população (μ).

Relativamente a estatísticas e parâmetros, no quadro seguinte são apresentados os símbolos que se utilizam nos casos em que se trata de uma estatística ou de um parâmetro populacional. É de notar que os parâmetros são letras do alfabeto grego.

[1] Recorde que as medidas de tendência central (media, moda e mediana) resumem uma distribuição num único valor que tendem a tipificar ou representar melhor um conjunto de dados. Ainda assim não caracterizam suficientemente todos os aspectos de um conjunto de dados e a forma como os seus valores se distribuem pela população ou amostra.
A média aritmética ($\overline{x}$) ou simplesmente média de n elementos de uma série de valores X_1, X_2, $X_3 \ldots X_n$ da variável, é o quociente entre o somatório desses n valores e o total destes, simbolicamente: $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$.

[2] O desvio padrão (s) e a variância (s^2) são medidas de dispersão ou variabilidade, e permitem-nos conhecer a forma como os valores da variável estatística se distribuem («dispersam») *em torno dos valores centrais*. O valor do desvio padrão é dado por: $s = \sqrt{\frac{\Sigma(x_i - \overline{x})^2}{n-1}}$.

Por exemplo os valores: 70, 72, 73, 78, 82, 82, 83, 90, 97 dizem respeito aos valores da glicémia em jejum (mg/dl) de uma amostra de n=9 utentes. A média será então $\overline{x} = \frac{727}{9} = 80{,}78\text{mg/dl}.$ e o desvio padrão:

$$s = \sqrt{\frac{(70-80{,}78)^2 + (72-80{,}78)^2 + (73-80{,}78)^2 (78-80{,}78)^2 + (82-80{,}78)^2 + (82-80{,}78)^2 + (83-80{,}78)^2 + (90-80{,}78)^2 + (97-80{,}78)^2}{9-1}} = 8{,}79\text{mg/dl}$$

Quadro 2 – Símbolos utilizados para representar algumas estatísticas amostrais e parâmetros populacionais

Designação:	**Estatísticas amostrais** (símbolos)	**Parâmetros populacionais** (símbolos)
Média	$\overline{X}$	μ
Desvio Padrão	s	σ
Variância	s^2	σ^2
Proporção	p	π
Correlação	r	ρ

Imagine que escolhíamos outras amostras aleatórias do mesmo tamanho de enfermeiros (n = 150) da população. É possível que encontrássemos valores ligeiramente diferentes para a média do tempo gasto a dar banho a um doente acamado. Poderíamos encontrar médias com valores ligeiramente inferiores ou superiores. Tal *flutuação* nas médias designa-se por **variabilidade amostral**. Ou seja, quando escolhemos outra amostra aleatória de 150 enfermeiros é possível que o valor seja diferente. Pode então perguntar se podemos tomar alguma decisão com base nesta *variabilidade amostral*? É claro que sim, obviamente levando em linha de conta essa mesma variabilidade.

Na realidade e como vamos verificar, a decisão que tomamos em aceitar que a média da amostra seja utilizada como média da população é fundamentada no conceito de **distribuição amostral**, e mais precisamente no **teorema do limite central**.

O conceito de **distribuição amostral** diz-nos que uma distribuição amostral de uma determinada estatística, neste caso a média, é a distribuição de valores que essa estatística assume em todas as amostras possíveis do mesmo tamanho, seleccionadas dessa população. Trata-se de uma distribuição teórica de probabilidades, representando todos os valores das amostras possíveis que poderiam ocorrer se extraíssem amostras infinitas do mesmo tamanho dessa população.

A questão que se nos levanta relativamente ao exemplo anterior do tempo médio gasto pelos enfermeiros a dar banho a um doente acamado, é saber em que medida é esta média ($\overline{X}$ = 24,35) representativa da média populacional, ou melhor, em que medida está a estatística amostral ($\bar{x}$) próxima do verdadeiro parâmetro populacional (μ).

Vamos exemplificar a ideia subjacente ao conceito de distribuição amostral das médias, a partir de 4 valores[3] respeitantes ao valor de estrógenio no plasma (gramas/100ml), no pós-parto imediato para uma população de N=4 e em que os valores são, 2,6; 3,3; 3,7 e 4,4.

[3] Trata-se apenas de um exemplo didáctico.

Como pode verificar os parâmetros populacionais são: μ=3,5 gramas/100ml e σ=0,65 gramas/100ml. Vamos seleccionar aleatoriamente todas as amostras diferentes de n=2 e calcular a média e o desvio padrão de cada uma.

$$\bar{x}_1 = \frac{2,60 + 3,3}{2} = 2,95 \qquad \bar{x}_2 = \frac{2,6 + 3,7}{2} = 3,15 \qquad \bar{x}_3 = \frac{2,6 + 4,4}{2} = 3,5$$

$$\bar{x}_4 = \frac{3,3 + 3,7}{2} = 3,5 \qquad \bar{x}_5 = \frac{3,3 + 4,4}{2} = 3,85 \qquad \bar{x}_6 = \frac{3,7 + 4,4}{2} = 4,05$$

Calculando a média das médias amostrais ($\mu_{\bar{x}}$) e o desvio padrão dos desvios padrões amostrais ($\sigma_{\bar{x}}$) teremos:

$$\mu_{\bar{x}} = \frac{2,95+3,15+3,5+3,5+3,85+4,05}{6} = 3,5 \text{ gramas/100ml.}$$

$$\sigma_{\bar{x}} = \sqrt{\frac{(2,95-3,5)^2 + (3,15-3,5)^2 + (3,5-3,5)^2 + (3,5-3,5)^2 + (3,85-3,5)^2 + (4,05-3,5)^2}{6}} = 0,376 \text{gramas/100ml.}$$

Repare na simbologia. A média das médias amostrais é representada por $\mu_{\bar{x}}$ e o desvio padrão por $\sigma_{\bar{x}}$.Vamos realizar o mesmo procedimento, mas para amostras de tamanho n=3. Como são possíveis quatro amostras, as médias serão:

$$\bar{x}_1 = \frac{2,60 + 3,3 + 3,7}{3} = 3,2 \qquad \bar{x}_2 = \frac{3,3 + 3,7 + 4,4}{3} = 3,8$$

$$\bar{x}_3 = \frac{3,7 + 4,4 + 2,6}{3} = 3,57 \qquad \bar{x}_4 = \frac{4,4 + 2,6 + 3,3}{3} = 3,43$$

Neste exemplo de amostra de n = 3, a média das médias amostrais (simbolicamente $\mu_{\bar{x}}$) e desvio padrão ($\sigma_{\bar{x}}$) são $\mu_{\bar{x}}$ = 3,5 e $\sigma_{\bar{x}}$ = 0.218. Conforme podemos verificar pelo quadro 3, comparativo do parâmetro populacional e a distribuição das médias amostrais para n = 2 e n = 3, a média amostral independente do tamanho dos n*, é igual à média população*, além disso à medida que se aumenta o tamanho da amostra o *desvio padrão é menor, diminuindo a dispersão*.

Quadro 3 – valores comparativos do parâmetro populacional
e distribuição de médias amostrais para amostras de n=2 e n=3

Parâmetro populacional	μ = 3,5	σ = 0.65
Distribuições amostrais n = 2 n = 3	$\mu_{\bar{x}}$ =3,5 $\mu_{\bar{x}}$ =3,5	$\sigma_{\bar{x}}$ =0,376 $\sigma_{\bar{x}}$ =0,218

Fazemos aqui uma breve introdução para apresentar o SPSS na versão 11.5, no entanto, se estiver à vontade com o software, poderá avançar de imediato para o ponto 1.3. (**Gerar amostras aleatórias no SPSS – continuação das distribuições amostrais**) deste caderno.

1.2. Breve introdução ao SPSS

O **SPSS** (Statistical Package for the Social Sciences), actualmente na versão 17, é um dos mais populares programas de análise de dados estatísticos. É de fácil manuseamento mesmo para quem não tiver qualquer formação específica sobre o programa, dado o seu interface amistoso. Apresentamos os comandos na versão 11.5, pois as mudanças ocorridas são pouco significativas.

É salientar que, se por um lado, o programa deixou de parte um conjunto de formas complexas e cálculos exaustivos, ele tem revelado o uso indevido de algumas ferramentas, como mais *gravoso*, a ausência de uma análise crítica dos resultados. O *software* gera resultados com base nas opções e decisões do investigador/utilizador, mas não censura as «asneiras» que lhe são solicitadas, tão pouco produz a interpretação dos resultados, isso é, felizmente, da responsabilidade do investigador/utilizador.

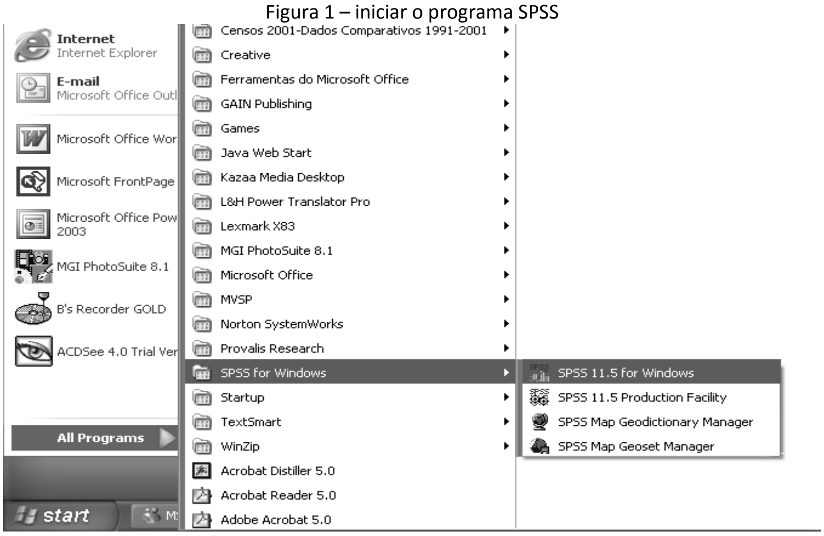
O exemplo apresentado de seguida diz respeito a um estudo em que foi utilizado o SPSS como programa de análise de dados. Como pode verificar, em nenhum momento o SPSS produz resultados como os apresentados, tão pouco pode servir de pretexto para o equívoco.

	FFT	Barthel	IMC
Nº doenças crónicas	-0.32 (p=1,45)	-0.38(p=0,03*)	0.75(p=0,04*)
Idade	-0.93 (p=1,05)	-0.065(p=0,04*)	0.012 (p=2,11)
Barthel	0.69 (p=0,04*)	X	-0.343(p=0,008**)
IMC	-0.65 (p=4,02)	-0.343(p=0,008**)	X

* Correlação significante p=0.05 ** Correlação significante p=0.01

Inicie o programa SPSS for Windows (neste caso é a versão 11.5) através **botão iniciar** da **barra de tarefas** que está no seu ambiente de trabalho, conforme a figura 1.

Figura 1 – iniciar o programa SPSS



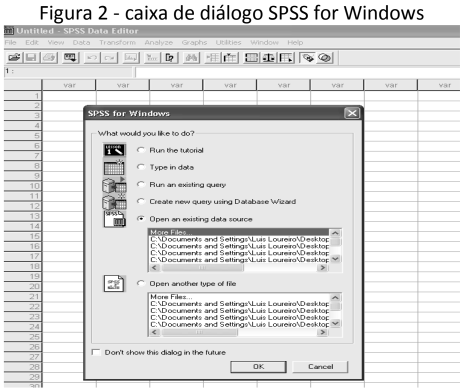
O programa contém várias janelas e caixas de diálogo que nos possibilitam visualizar todas as opções, bem como todo um conjunto de comandos facilmente acessíveis. Apresentaremos aqui os mais importantes e considerados essenciais no desenvolvimento das estatísticas apresentadas ao longo do caderno.

1.2.1. – O Editor de Dados (Data Editor)

Ao iniciar o programa aparece-lhe a caixa de diálogo **SPSS for Windows** (Figura 2). Independentemente do que estiver para realizar, mesmo que seja para construir uma base de dados de raiz, pode dar um clique sobre o botão **Cancel**.

Genericamente esta caixa de diálogo permite iniciar o manual do programa (**Run the tutorial**), construir uma base de dados nova (**Type in data**), importar dados através de uma *query* existente (**Run an existing query** ou **Create new query using Data Base Wizard**), abrir uma base de dados já existente (**Open na data source**) ou então abrir outro tipo de ficheiro (**Open another type of file**). A selecção de qualquer uma das tarefas é feita através da escolha do círculo adequado dando depois um clique em **OK** com o cursor.

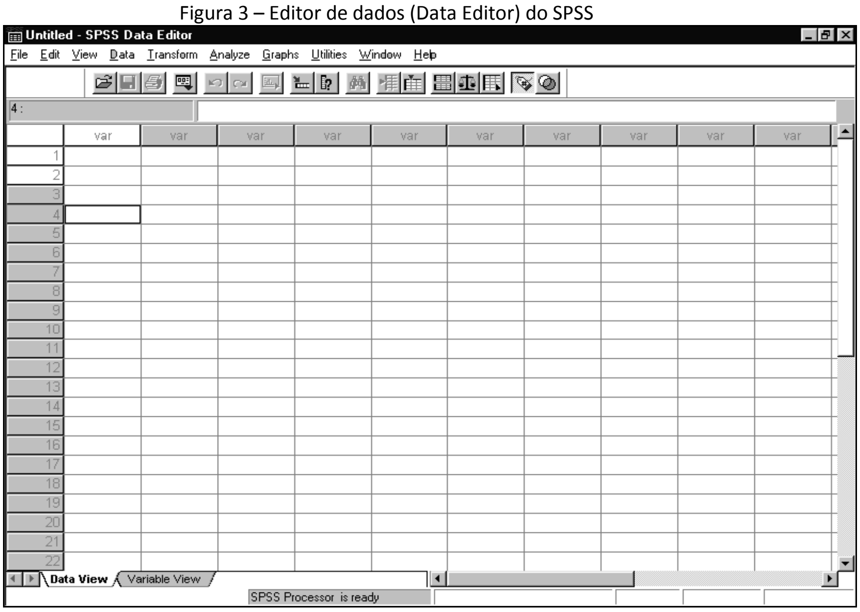
Figura 2 - caixa de diálogo SPSS for Windows



Se optar por dar um clique sobre o botão **Cancel** (opção mais adequada se não for um utilizador treinado) na caixa de diálogo **SPSS for Windows**, fica automaticamente em presença de uma janela designada de **SPSS Data Editor** ou **editor de dados** (figura 3)**.**

É a partir desta janela que são codificadas as variáveis e inseridos os dados colhidos nos nossos estudos. Desde a versão 15 que o SPSS já permite ter várias janelas do editor de dados abertas simultaneamente. O aspecto da janela e o conjunto de comandos disponíveis tornam a interface deste programa muito amistoso, como pode observar pela figura 3.

Figura 3 – Editor de dados (Data Editor) do SPSS

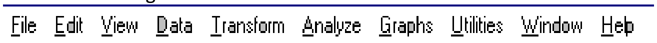


O editor de dados é semelhante a uma folha de cálculo do programa EXCEL, contendo nas colunas as variáveis (**var**) e nas linhas os casos, mas apenas isso, pois o EXCEL tem subjacente a folha de cálculo, enquanto o SPSS não deixa criar fórmulas nas células.

1.2.1.1. – Barra de menus do SPSS (menus e funções) e barra ferramentas

O editor de dados do SPSS não tem qualquer nome porque ainda não foi guardado, por isso a designação de **Untitled SPSS Data Editor**, no canto superior esquerdo da janela. No topo aparece a **barra de menus** (figura 4), cada um com as funções diferentes.

Figura 4 – barra de menus do editor de dados

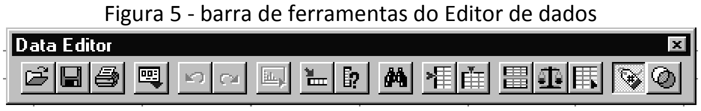


Cada um dos comandos tem as funções descritas seguidamente.

Nomes dos menus:	Função:
File	Permite criar, ler, abrir, imprimir e guardar ficheiros bem como abandonar o programa.
Edit	Permite apagar, copiar, colar e procurar dados, bem como aceder a um conjunto de opções diversas
View	Permite activar e desactivar a barra de botões bem como mostra os níveis codificados de cada variável.
Data	Permite realizar algumas funções tais como definir variáveis e casos, criar grupos de casos, etc.…
Transform	Permite produzir alterações nas variáveis bem como criar novas variáveis a partir de outras existentes.
Analyze	Permite realizar todas as estatísticas descritivas e inferenciais
Graphs	Permite criar e manipular gráficos
Utilities	Permite aceder à informação relativa a cada variável, controlo de lista de variáveis, etc...
Windows	Permite minimizar as janelas do SPSS
Help	Permite procurar ajuda do SPSS quanto ao funcionamento e à realização de estatísticas

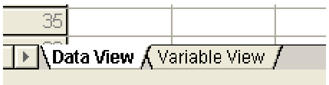
O programa apresenta ainda uma barra de ferramentas (figura 5). Clicando sobre os comandos da barra de ferramentas facilmente se tem acesso à sua informação. O que é possível realizar com a barra de ferramentas, também está disponível nos menus.

Figura 5 - barra de ferramentas do Editor de dados



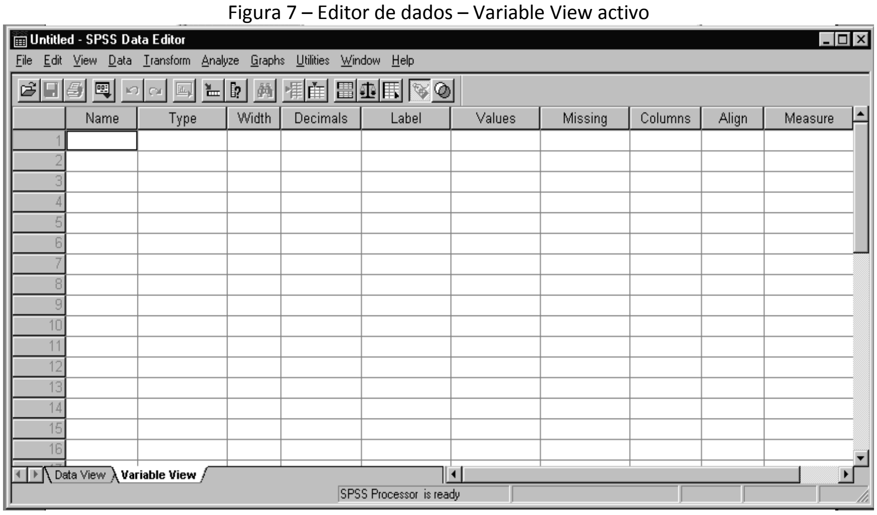
Ainda no editor de dados (figura 3) tem os modos de visionamento desta janela (figura 6). No **Data View** visionamos a base de dados, se tivermos inserido dados, caso contrário está em branco. Ao ar um clique no botão **Variable View** (figura 6), acede-se ao visionamento das variáveis para criar e codificar.

Figura 6 – modos de visionamento do Editor de dados



Por defeito a janela aparece com o **Data View** activo, mas poderá, com um simples clique sobre o botão **Variable View**, aceder ao visionamento das variáveis a codificar, conforme figura 7.

Figura 7 – Editor de dados – Variable View activo



Ainda relativamente ao Data Editor do SPSS, veja a figura 8 de uma base com dados inseridos. Nesta janela:

a) as linhas são os casos. Cada fila representa um caso ou uma observação. Por exemplo, cada resposta individual para um questionário é um caso;

b) as colunas são variáveis. Cada coluna representa um ser variável ou característica medida. Por exemplo, cada questão num questionário é uma variável;

c) as células contêm os valores. Cada célula contém um único valor de uma variável para cada caso. Como pode verificar, cada célula corresponde a um caso e uma variável. As células do SPSS não permitem apenas registar o valor;

d) o ficheiro de dados é rectangular. As suas dimensões são determinadas pelo número de casos e variáveis.

Repare que já tem o nome de «regímen de emagrecimento». A célula seleccionada[4] a cor cinza, corresponde ao índice de massa corporal do 6.º caso, na variável antes do regímen.

[4] Como pode verificar, a base de dados do exemplo é constituída por 4 variáveis e 18 casos. A célula acinzentada representa o caso n.º 6, da variável antes, e o seu valor é de 31,12.

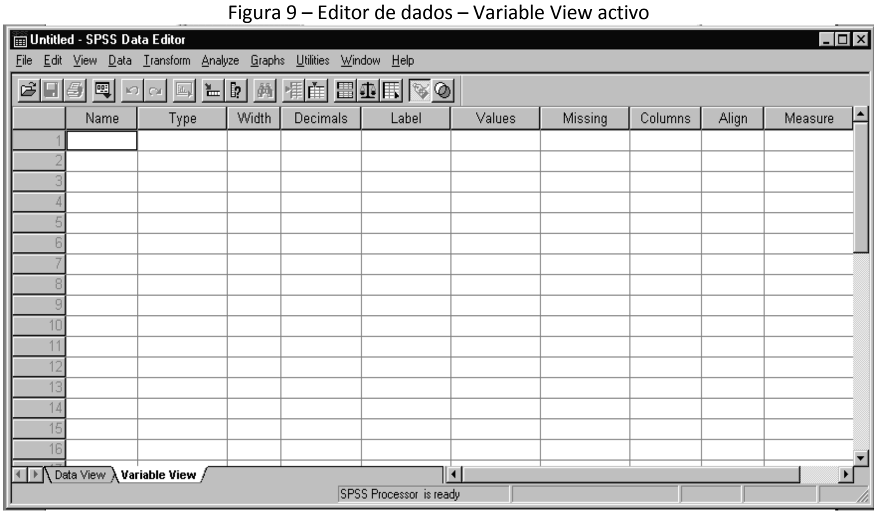
Figura 8 – Editor de dados (base existente)

	n_quest	sexo	antes	depois	var
1	1,00	masc	30,08	30,09	
2	2,00	femi	30,09	26,03	
3	3,00	masc	30,12	25,08	
4	4,00	femi	30,32	32,04	
5	5,00	masc	31,00	30,06	
6	6,00	femi	31,12	31,14	
7	7,00	masc	32,18	28,83	
8	8,00	femi	32,56	33,45	
9	9,00	masc	33,45	34,67	
10	10,00	femi	33,89	25,68	
11	11,00	masc	34,10	27,90	
12	12,00	femi	34,61	34,65	
13	13,00	masc	35,80	36,71	
14	14,00	femi	35,92	35,97	
15	15,00	masc	36,78	30,65	
16	16,00	femi	37,00	36,34	
17	17,00	masc	37,42	38,10	
18	18,00	femi	38,90	35,44	
19					

1.2.2. – Criação de uma base de dados no SPSS

O primeiro passo na construção de uma base de dados consiste na atribuição de nomes (designações) às variáveis para que estas não sejam uma massa de informação desordenada. Este processo é designado de codificação da base. O instrumento que deve utilizar para treinar deve ser codificado no Editor, opção **Variable View**. Dando um clique sobre o comando **Variable View** tem acesso à janela apresentada na figura 9 (a mesma da figura 7).

Figura 9 – Editor de dados – Variable View activo



As variáveis podem assim ser codificadas quanto ao nome (**Name**), ao tipo (**Type**), o n.º de casas decimais, **Label, Values, Missing values**, **Measure**, etc… Imagine que quer construir uma base de dados para um questionário que irá administrar a 100 enfermeiros do Hospital Pediátrico. Mesmo antes de inserir os dados no editor do SPSS, deve codificar a base. Imagine que o questionário era o seguinte:

N.º de questionário: _____

1. Sexo: Masculino ❑
   Feminino ❑

2. Idade:____ anos

3. Estado civil:
- Casada(o) ❑
- Solteira(o) ❑
- Divorciada(o) ❑
- Viúva(o) ❑
- União facto ❑

4. Anos de serviço: _____ anos

5. Formação em intervenção familiar
- Sim ❑
- Não ❑

Escala de opiniões dos enfermeiros acerca do envolvimento/participação dos pais nos cuidados

①	②	③	④	⑤
Discordo completamente	Discordo	Não concordo nem discordo	Concordo	Concordo completamente

1. Os pais devem ser considerados como um elemento pertencente à equipa prestadora de cuidados.
2. A participação dos pais nos cuidados deve ser negociada de acordo com os seus desejos e interesses.
3. A tomada de decisão relativamente aos cuidados a prestar à criança deve ter em conta a opinião dos pais.
4. A participação dos pais nos cuidados fortalece a ligação afectiva entre pais e filhos.
5. Os pais devem prestar os cuidados necessários para darem continuidade em casa.
6. O envolvimento dos pais nos cuidados não favorece o desenvolvimento da criança*.
7. O envolvimento dos pais nos cuidados dá maior segurança à criança.
8. O envolvimento dos pais nos cuidados permite uma melhor e mais rápida recuperação da criança.
9. Para envolver os pais nos cuidados é essencial que estes estejam receptivos.
10. O envolvimento dos pais nos cuidados não lhes dá poder de decisão e maior segurança*.

Todas as variáveis de qualquer questionário devem ter um nome que as identifique, isso é óbvio. Seja, o número de questionário, idade, estado civil, os itens das escalas, etc.

Reveja a figura 9. A 1.ª coluna, com o nome **Name** permite atribuir nomes às variáveis que pretende inserir. **Eis as regras que deve saber antes de lhe atribuir um nome** (na versão 11.5)**:**

a) Apenas permite utilizar um máximo de oito caracteres (**na versão 11.5)**, letras ou números, sendo que o primeiro terá que ser obrigatoriamente uma letra;

b) a variável nunca poderá acabar com um ponto (por exemplo sexo.);

c) a variável não poderá ter espaços em branco ou caracteres especiais, tais como &, ?, !, ´, *;

d) O nome da variável deverá ser único na base, assim se atribuir o nome sexo a uma variável, nunca o poderá utilizar para outra variável na mesma base;

e) por defeito, o SPSS atribui, se forem colocados dados na base sem ser atribuído nome às variáveis a designação **var00001**, **var00002**, etc…,como nome das variáveis;

f) as seguintes palavras, **ALL, NE, EQ, TO, LE, LT, BY, OR, GT, AND, NOT, GE, WITH**, nunca poderão aparecer como nomes pois fazem parte de um conjunto reservado de caracteres da sintaxe do SPSS.

1.2.2.1. – Codificação de variáveis no modo de visionamento Variable View

É obrigatório codificar as variáveis de forma a respeitar estas regras. Assim a variável estado civil, pode aparecer na coluna name como "**est_civ"**, a variável número de questionário como **n_quest**, etc… Posteriormente existe forma como veremos de colocar o nome da variável por extenso, ou seja, um local onde se poderá fazer a descrição detalhada da variável.

Atenção a criação da variável número de questionário é obrigatória. Nunca se esqueça, caso contrário poderá ter que inserir os dados novamente. É preferível aprender com os erros dos outros do que com os nossos. Numere sempre os questionários.

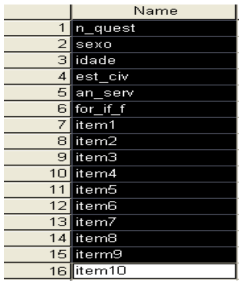
Muitas vezes utilizamos inventários os escalas que são constituídas por vários itens ou questões, neste exemplo, são 10 itens. Neste caso é comum atribuir como nome (**Name**) a designação de **item1**, **item2**…e sucessivamente até número **item10**. Posteriormente coloca-se o conteúdo do item por extenso no local adequado **(Label)**. Por exemplo, o 1.º item diz que «os pais devem ser considerados como um elemento pertencente à equipa prestadora de cuidados», na coluna **Name** ele deverá receber a designação **item1**.

Algumas variáveis têm diferentes categorias, por exemplo o estado civil tem cinco categorias, o sexo tem duas, os diferentes itens estão cotados em escala de Likert de 1 a 5 pontos, etc.…

Seguidamente mostramos a forma como vamos atribuir nomes às variáveis. As abreviaturas que usamos para inserir na coluna **Name** estão a **bold**. Os asteriscos (*) colocados nos itens 6 e 10 alertam para a necessidade de efectuar uma transformação de identidade à posteriori, isto é, os itens deverão ser recodificados. Só deverá fazer as recodificações depois de inseridos os dados na base, antes de efectuar qualquer tipo de análise estatística.

A figura 10 mostra a coluna **Name** com os nomes atribuídos a cada uma das variáveis do questionário.

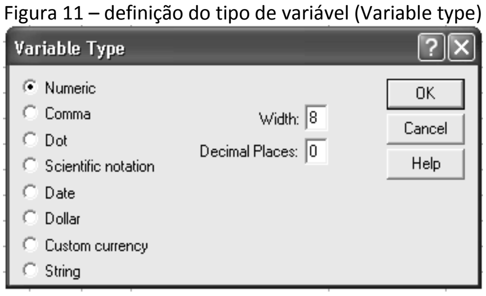
Figura 10 – definição de nomes na coluna
da janela Variable view



A seguir à coluna **Name** aparece no editor de dados a coluna **Type,** isto é, a definição do tipo de variável. Posicione o cursor sobre a célula e no botão acizentado [ ] tem acesso à definição do tipo de variável (figura 11). Por defeito aparece a variável como numérica (**numeric**). Convém salientar que na grande maioria das vezes é neste tipo que trabalhamos.

Pode ainda mexer no n.º de dígitos ou caracteres (**with**), por defeito aparece 8, e no n.º de casas decimais (**decimal place**), que é de 2. No exemplo das casas decimais, se tiver um enfermeiro que tiver 22 anos, ao assumir duas casas decimais aparece 22,00. Se colocar 0 casas decimais aparecerá 22 apenas, sem casas decimais. As duas colunas seguintes (with e decimal place têm a mesma função).

Figura 11 – definição do tipo de variável (Variable type)



Na coluna **Label** pode colocar por extenso o nome da variável, ou seja defini-la mais claramente. No fundo trata-se de definir aquilo que não é permitido no **Name** (veja a figura 14).

Na coluna **values** poderá atribuir números às variáveis categoriais (qualitativas). Só se aplica para os casos em que as variáveis assumem categorias de índole qualitativa. Por exemplo, a variável sexo pode assumir dois valores, em que 1 corresponde ao masculino e 2 ao feminino. Quando estiver a inserir os dados[5], basta colocar o n.º 1 e aparece-lhe a categoria masculino, ou o n.º 2 e aparece-lhe o feminino.

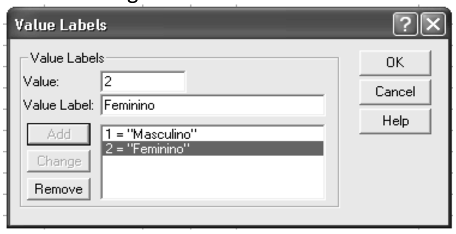
Por exemplo a variável idade se tiver em "bruto", isto é, idade real em anos, não faz sentido criar diferentes categorias, ou seja, como a variável n.º de questionário, não necessita

[5] Caso não lhe apareça apenas o número, na barra de ferramentas dê um clique sobre o botão que tem uma etiqueta.

de qualquer codificação prévia, já se a idade estiver em intervalos, por exemplo, 20-25 anos e 25-30 anos, deverá fazer uma codificação, atribuindo o n.º 1 à primeira categoria e por aí adiante. O mesmo para a variável estado civil.

Dando um clique na célula sobre o botão acizentado [] tem acesso à definição dos níveis ou categorias da variável. No **value** insere o valor e no **value label** o nome da categoria, conforme o exemplo da figura 12, seguido de clique sobre o botão **Add**. Se quiser retirar uma categoria, basta seleccioná-la e dar um clique sobre o botão **Remove**. Para alterar, se tivesse escrito com erro, bastaria um clique sobre o botão **Change**, depois de previamente seleccionada a categoria.

Figura 12 – Value Labels



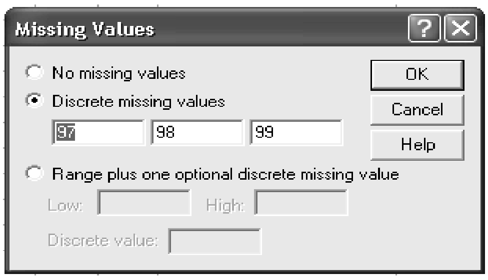
1.2.3. – As não respostas (*missing values*) no SPSS

Seguidamente apresenta-se a coluna das não respostas ou *Missing values*. Os missing´s são relativamente comuns nos nossos trabalhos. Os motivos para tal ocorrência são vários: a) pessoas esquecem-se de preencher alguns itens; b) não podem responder face à natureza contingente das respostas; c) dão respostas ilegíveis ou assinalam várias opções na mesma resposta. Existem varias formas de trabalhar as não respostas, no entanto elas podem ser codificar previamente.

Imagine que sempre que as pessoas não respondem por algum motivo a uma questão lhe atribuímos o valor 97 (nenhuma resposta), sempre que não se aplicar o valor 98 (se questionamos uma pessoa se tem religião e ele responde que não, parece óbvio que não irá responder na questão seguinte da frequência de actividades religiosas), e atribuímos o valor 99 para as circunstâncias em que as respostas são ilegíveis.

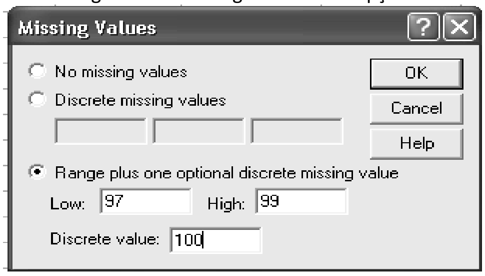
Um modo de codificar estas não respostas é colocando os três valores (97,98 e 99) nos espaços em branco do **Discrete missing values**, conforme a figura 13.

Figura 13 – Missing Values – 1.º opção



Outro modo é colocar um intervalo de valores entre 97 e 99 (figura 14), e além disso colocar um outro valor adicional, por exemplo 100.

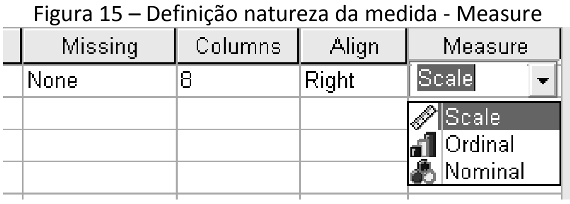
Figura 14 – Missing Values – 2.º opção



A coluna **Column** permite mudar a largura em termos de número de caracteres a incluir. Por defeito o SPSS atribui 8 caracteres (efeito puramente cosmético) mas isso não tem influência nos cálculos. A coluna **Align** permite alinhar o texto à esquerda, direita ou ao centro.

Por fim a coluna **Measure** (figura 15) permite codificar a variável quanto ao nível de medida. O SPSS agrupa os níveis intervalar e racional num único designado de Scale. Os níveis de medida nominal e ordinal são diferenciados. Aconselhamos vivamente, a não ser que a análise estatística o exija, mantenha sempre o nível scale independente do nível de medida das variáveis.

Figura 15 – Definição natureza da medida - Measure



A figura 16 apresenta o aspecto do editor de Dados – Variable View com as variáveis codificadas.

Figura 16 – Editor de Dados – Variable View

	Name	Type	Width	Deci	Label	Values	Missing	Colu	Align	Measure
1	n_quest	Numeric	8	2	número de questionário	None	None	8	Right	Scale
2	sexo	Numeric	8	2	sexo	{1,00, Masculino}...	97,00, 98,00, 99,	8	Right	Scale
3	idade	Numeric	8	2	idade	None	None	8	Right	Scale
4	est_civ	Numeric	8	2	Estado Civil	{1,00, Casada(o) }...	None	8	Right	Scale
5	an_serv	Numeric	8	2	tempo de exercício profissional	None	None	8	Right	Scale
6	for_if_f	Numeric	8	2	Formação em intervenção familiar	{,00, não}...	None	8	Right	Scale
7	item1	Numeric	8	2	1. Os pais devem ser considerados c	{1,00, discordo total	None	8	Right	Scale
8	item2	Numeric	9	2	2. A participação dos pais nos cuida	{1,00, discordo total	None	8	Right	Scale
9	item3	Numeric	8	2	3. A tomada de decisão relativament	{1,00, discordo total	None	8	Right	Scale
10	item4	Numeric	8	2	4. A participação dos pais nos cuida	{1,00, discordo total	None	8	Right	Scale
11	item5	Numeric	8	2	5. Os pais devem prestar os cuidado	{1,00, discordo total	None	8	Right	Scale
12	item6	Numeric	8	2	6. O envolvimento dos pais nos cuida	{1,00, discordo total	None	8	Right	Scale
13	item7	Numeric	8	2	7. O envolvimento dos pais nos cuida	{1,00, discordo total	None	8	Right	Scale
14	item8	Numeric	8	2	8. O envolvimento dos pais nos cuida	{1,00, discordo total	None	8	Right	Scale
15	iterm9	Numeric	8	2	9. Para envolver os pais nos cuidado	{1,00, discordo total	None	8	Right	Scale
16	item10	Numeric	8	2	10. O envolvimento dos pais nos cuid	{1,00, discordo total	None	8	Right	Scale

O aspecto do editor de Dados depois de inseridos os valores dos 100 questionários é o que apresentamos seguidamente na figura 17 (até ao caso 29).

Figura 17 – Editor de Dados – Data View

101 : n_quest

	n_quest	sexo	idade	est_civ	an_ser	for_if_f	item1	item2	item3	item4	item5	item6	item7	item8
1	1,00	Feminino	25,00	Solteira(o)	9	não	concordo t	concordo t	concordo t	concordo t	concordo t	concordo p	concordo t	concordo p
2	2,00	Feminino	46,00	Casada(o)	9	não	concordo t	concordo t	concordo t	concordo t	concordo p	concordo t	concordo t	concordo t
3	3,00	Feminino	28,00	Casada(o)	9	não	concordo t	concordo t	concordo t	concordo t	concordo p	concordo p	concordo t	concordo t
4	4,00	Feminino	42,00	Casada(o)	6	não	concordo t	discordo to	concordo p	concordo t	concordo t	concordo t	concordo t	concordo p
5	5,00	Feminino	33,00	Casada(o)	7	não	concordo t	concordo t	concordo p	concordo t	concordo p	não concor	concordo p	concordo p
6	6,00	Feminino	49,00	Casada(o)	6	não	concordo t	concordo t	concordo p	concordo p	concordo p	concordo p	concordo p	concordo t
7	7,00	Feminino	39,00	Casada(o)	6	sim	concordo t	concordo t	concordo t	concordo p	concordo p	concordo p	concordo p	concordo p
8	8,00	Feminino	40,00	Casada(o)	13	sim	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t
9	9,00	Masculino	53,00	Casada(o)	12	sim	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t
10	10,00	Feminino	38,00	Casada(o)	10	não	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo p	concordo t
11	11,00	Feminino	50,00	Casada(o)	6	sim	discordo pa	não concor	discordo to	concordo p	concordo p	concordo p	concordo t	concordo t
12	12,00	Feminino	25,00	Casada(o)	8	não	concordo t	concordo t	concordo p	concordo t	concordo t	concordo t	concordo p	concordo t
13	13,00	Feminino	23,00	Solteira(o)	9	não	concordo t	concordo t	concordo t	concordo t	concordo t	concordo p	concordo p	concordo p
14	14,00	Feminino	23,00	Solteira(o)	7	não	concordo p	concordo p	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t
15	15,00	Feminino	33,00	Casada(o)	9	não	concordo p	concordo t	concordo p	concordo t	concordo p	concordo t	concordo t	concordo t
16	16,00	Feminino	23,00	Solteira(o)	7	não	concordo t	concordo p	concordo p	concordo t	concordo t	concordo t	concordo t	concordo p
17	17,00	Feminino	35,00	Casada(o)	14	não	concordo p	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo p
18	18,00	Feminino	41,00	Solteira(o)	8	não	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t
19	19,00	Masculino	26,00	Solteira(o)	8	não	concordo p	concordo p	não concor	concordo t	concordo t	concordo p	concordo t	concordo p
20	20,00	Feminino	31,00	Solteira(o)	8	não	concordo p	concordo p	não concor	concordo t	concordo t	concordo p	concordo t	concordo p
21	21,00	Feminino	25,00	Casada(o)	9	não	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t
22	22,00	Feminino	51,00	Casada(o)	4	não	concordo p	concordo p	discordo pa	concordo p	concordo p	concordo p	concordo t	não concor
23	23,00	Feminino	23,00	Solteira(o)	8	não	concordo p	concordo t	concordo p	concordo t	concordo t	concordo t	concordo t	não concor
24	24,00	Feminino	53,00	Casada(o)	10	não	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t	concordo t
25	25,00	Feminino	53,00	Casada(o)	7	sim	concordo p	concordo p	concordo t	concordo p	concordo t	concordo p	concordo p	concordo p
26	26,00	Feminino	24,00	Solteira(o)	8	sim	concordo p	concordo t	concordo p	concordo p	concordo t	concordo p	concordo p	concordo p
27	27,00	Masculino	23,00	Solteira(o)	13	sim	concordo p	concordo t	concordo p	concordo p	concordo t	concordo p	concordo p	não concor
28	28,00	Feminino	61,00	Casada(o)	10	não	concordo t	concordo p	discordo to	concordo t	concordo t	concordo p	concordo t	não concor
29	29,00	Feminino	54,00	Casada(o)	6	não	concordo p	concordo p	concordo p	concordo t	concordo p	não concor	concordo p	concordo p

1.2.4. – Recodificação de variáveis e/ou criação de novas variáveis

Seguidamente vamos mostrar como deverá proceder às transformações de identidade dos itens 6 e 10. Essas transformações são feitas porque alguns itens das escalas e inventários são apresentados em sentido contrário (dizemos «pela negativa»). Para criar um score desses inventários ou escalas, os itens deverão ser recodificados para fazerem sentido no cálculo da pontuação total. Observe a lógica subjacente.

Os 3 itens seguintes (hipotéticos) foram retirados de uma escala de ansiedade perante a morte, com opção de resposta de 1 (discordo totalmente) a 5 (concordo totalmente).

① Discordo Totalmente	② Discordo	③ Não concordo Nem discordo	④ Concordo	⑤ Concordo Completamente

1. Tenho muito medo da morte	①	②	③	④	⑤
2. Os caixões deixam-se muito ansioso	①	②	③	④	⑤
3. Não tenho medo da morte	①	②	③	④	⑤

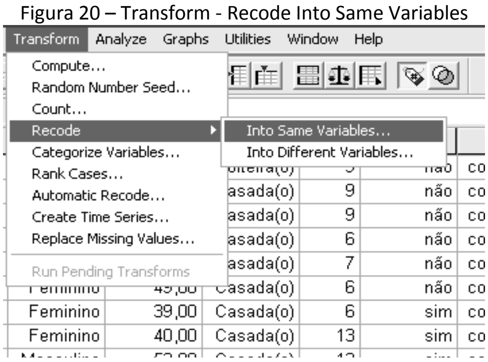
Repare que o item 3 é uma afirmação em sentido contrário dos outros dois itens. Se maiores pontuações correspondem a maior medo da morte, isso não acontece no item 3, para afirmar que se tem medo deverá assinalar-se com a cruz no n.º1 ou no n.º 2. Então antes de se calcular o score da escala, deverá proceder-se a uma transformação de identidade ou recodificação dos valores do item 3.

Assim o valor 1 passará a valer 5, o 2 a 4, o 3 mantém-se, o 4 a 2 e o 5 a 1. Assim os valores mais elevados neste item já se constituem como indicando maior medo da morte.

No SPSS esta operação só deverá ser feita depois de inseridos todos os dados. A forma é simples e rápida. Faz-se através do **Transform** da barra de menus, seguido da **opção Recode - Into Same Variables** (figura 20).

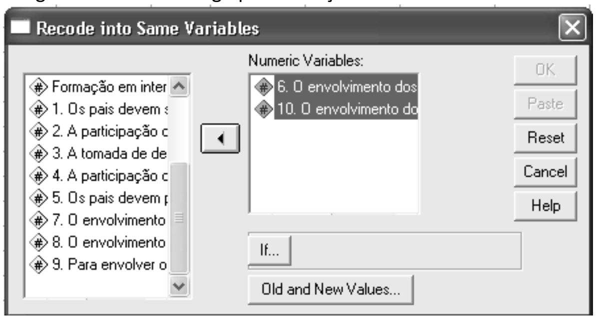
Poderia recodificar, criando duas novas variáveis na base (**Into Diferrent Variables**) mantendo os valores nas variáveis originais. Neste caso vamos efectuar a recodificação dentro das mesmas variáveis (**Into Same Variables**).

Figura 20 – Transform - Recode Into Same Variables



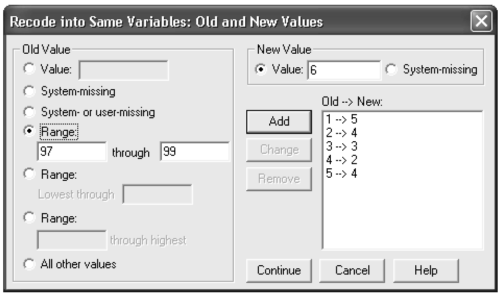
Ao solicitar a recodificação aparece-lhe a caixa de diálogo da figura 21. Nesta caixa tem dois campos distintos. Do lado esquerdo o campo (caixa) de variáveis passíveis de recodificação (todas as numéricas da base), e o campo (caixa) da direita é o campo que se destina às variáveis a recodificar (**Numeric Variables**). Deverá seleccionar as variáveis de interesse e move-las com o botão para o campo da direita. Poderá move-las de campo para campo com um clique duplo sobre a variável que ela passará automaticamente para o campo da direita e vice-versa.

Figura 21 – caixa diálogo para selecção de variáveis a recodificar



Depois de mover as variáveis de interesse (itens 6 e 10) para a caixa da direita, deverá proceder à sua recodificação, para isso deve premir com um clique o botão **Old and New values**, e aparece-lhe a caixa de diálogo da figura 22.

Figura 22 – Recodificar valores antigos/velhos e novos



Relembre que:

Valor originais **(Old Value)**		Novos valores **(New Value)**
1	→	5
2	→	4
3	→	3
4	→	2
5	→	1
Missing values	→	?

É simples proceder à recodificação. No rectângulo dos valores antigos (**Old Value**) deverá colocar o valor a recodificar, isto é o número 1 e simultaneamente no rectângulo **New**

Value colocará o novo valor, neste caso 5, seguido de um clique sobre o botão **Add**, e assim sucessivamente para todos os valores.

Relativamente aos missing values, se utilizou o sistema 97, 98 e 99, deverá colocá-los nos rectângulos do Range 97 to 99, e seleccionar seguidamente a opção do **system missing**. Posto isto a variável estará recodificada.

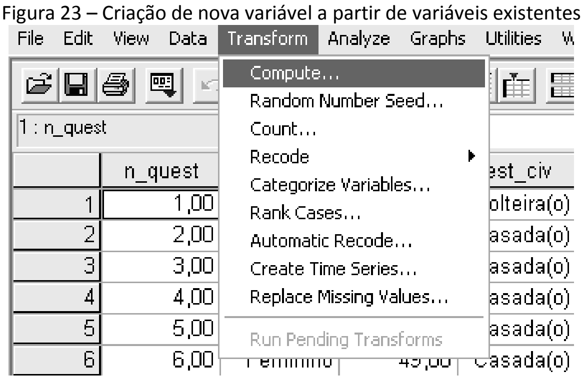
Como referimos, O SPSS permite-nos criar novas variáveis a partir das variáveis já existentes e inseridas na base, e isso é comum quando necessitamos de criar pontuações ou scores de escalas ou inventários, seja para as dimensões dos questionário ou mesmo para a pontuação global de um determinado instrumento.

Se designarmos por **«score das opiniões»** (do nosso exemplo) o nome da nova variável a criar, ela corresponderá ao somatório das pontuações de todos os itens, isto é:

sc_opini = item1+item2+item3+item4+item5+item6+item7+item8+item9+item10

Atribuímos a designação «**sc_opini»** pelo simples facto de apenas se poderem utilizar no SPSS 8 dígitos (versão 11.5). Os comandos do SPSS são apresentados seguidamente nas figuras 23 e 24.

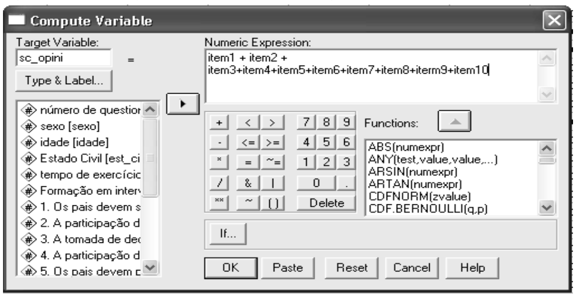
Figura 23 – Criação de nova variável a partir de variáveis existentes



A selecção do comando **Compute** faz aparecer uma caixa de diálogo com o nome de Compuite Variables. No rectângulo **Target Variable** coloca o nome da nova variável, essa variável irá corresponder (=) ao somatório dos itens de 1 a 10.

Os itens passam do campo das variáveis (campo das esquerda com todas as variáveis) para o novo campo da direita (**Numeric Expression:**) seleccionando os itens e deslocando-os com o botão () que se encontra na caixa. Neste modo terá que passar todos os itens, um a um. Repare que estão separados pelo sinal +.

Figura 24 – Criação de nova variável a partir de variáveis existentes



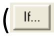
Repare que a caixa de diálogo da figura 24 encontra ainda um campo de funções (**Functions**), um botão () que lhe permitir aceder a outra caixa de diálogo (**Compute Variables: If Cases**) e um conjunto de teclas com símbolos e números (figura 25).

Figura 25 – teclas da calculator pad

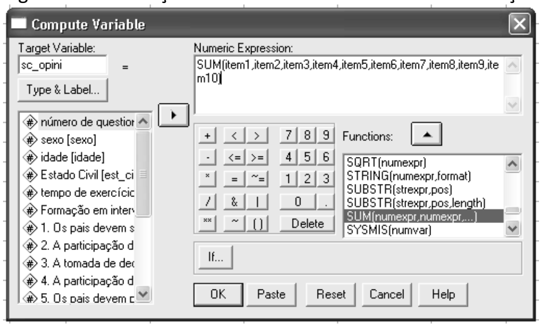
Conforme pode verificar pela figura 25, encontra a parte numérica do lado direito (e a tecla apagar), do lado esquerdo as operações básicas, e depois os símbolos, sendo que:

> maior do que
< menor do que
=> maior ou igual a
<= menor ou igual a
= igual
~= não igual a
| ou
& e
~ não
() parêntesis para separar variáveis e/ou operações

O SPSS permite, dentro desta caixa, a criação de somatórios mas fazendo recurso a outros dois modos. Um deles é recorrendo às funções (**Functions**) do campo de funções.

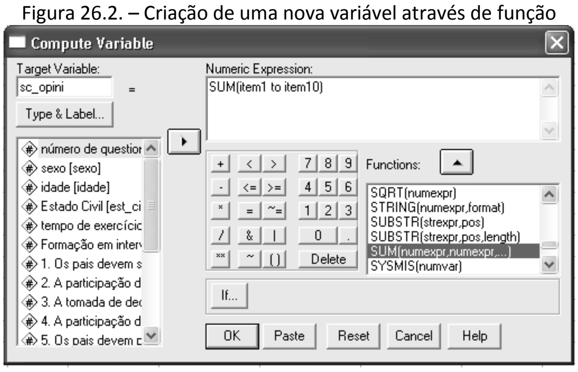
Se quiséssemos criar o mesmo somatório utilizando o mesmo nome da variável destino (**Target Variable**), no entanto, e conforme pode verificar pela figura 26, no campo da Numeric expression, deverá colocar a função **SUM(numexpr, numexpr,…)** da listagem de funções. Isso faz-se colocando o cursor sobre a função desejada, deslocando-a com o cursor sobre o botão (). Como pode verificar, os 10 itens que foram colocados dentro de parêntesis estão separados por vírgulas.

Figura 26.1. – Criação de uma nova variável através de função



Uma última forma de criar o somatório era recorrer à mesma função, mas colocando dentro de parêntesis a expressão **Sum(item1 to item10)**, conforme a figura 26.2.

Figura 26.2. – Criação de uma nova variável através de função



A realização destas operações de transformação dos dados, nomeadamente criando novas variáveis através das variáveis existentes, só deverá ser feito após a auditoria à base de dados, pois podem existir erros resultantes do processo de inserção de dados que se vão repercutir obviamente no valor das pontuações ou scores da escala.

1.2.5. – Guardar um ficheiro de dados do SPSS

Para guardar a base de dados ou seja o ficheiro que tem os seus dados, deverá faze-lo através do **File**, **Save as**… onde lhe atribuirá o nome (**base parceria**) e escolhera o sítio onde a vai guardar, neste caso Desktop.

Figura 27 – Guardar ficheiro

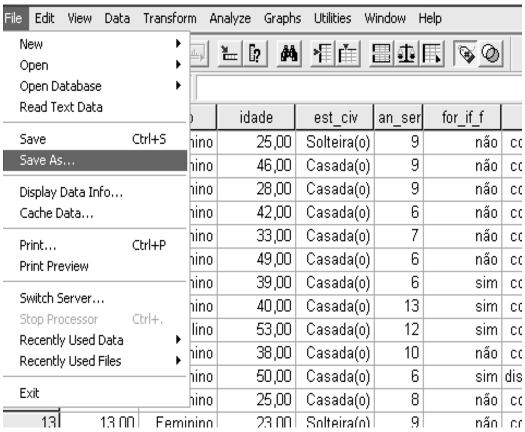
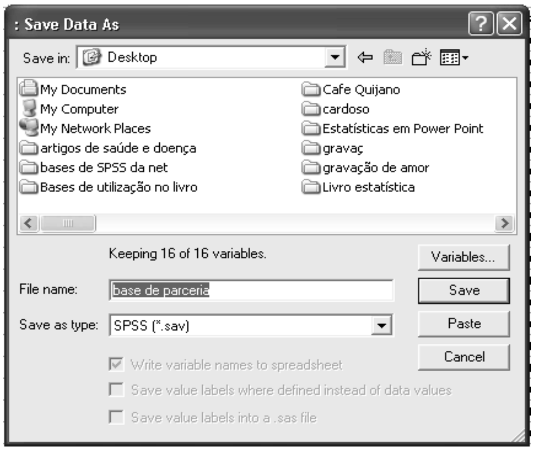


Figura 28 – atribuição de nome e localização



1.3 – Gerar amostras aleatórias no SPSS – continuação das distribuições amostrais

Retomando as distribuições amostrais. Se quiser criar amostras aleatórias a partir do SPSS, o processo é simples. O intuito é gerar repetidas amostras aleatórias e verificar a forma das distribuições amostrais das médias. Ou seja, com o programa pode por exemplo criar *k* amostras de igual tamanho (n) partindo por exemplo de uma população especifica de N valores. Utilize os valores seguintes para exercitar no SPSS.

62	71	27	37	40
47	53	63	31	32
74	64	34	35	28
22	20	25	44	50
33	26	55	58	19
41	33	42	63	25
52	53	50	39	45
48	19	75	49	50
76	68	27	52	40
71	54	57	27	30
50	53	52	18	20
26	65	27	22	35
70	59	61	73	60
52	25	37	20	45
66	55	53	31	50

Trata-se da idade de uma população de doentes com uma patologia específica que é abrangida por um centro de saúde com μ=44,81 anos. Construa a base de dados dando respectivamente a designação ***idade*** à variável, seguidamente insira os dados e guarde a base no seu disco rígido.

Se calcular a média e desvio padrão da população através do comando ***Descriptives*** do SPSS, a média obtida será (para a população) de 44,81 anos com um desvio padrão de 16,51 anos, isto sem seleccionar qualquer tipo de amostra aleatória.

O processo de criação de amostras aleatórias é simples. Seleccione o comando **Data** da barra de menus seguido de **Select Cases**. Depois deverá seleccionar **Random sample cases** - dê um clique no botão **Sample...** Dê um clique no circulo pequeno que tem a designação de **Exactly** e digite na caixa ao lado o n.º 30 (significa que o computador vai escolher aleatoriamente amostras de 30 utentes baseado na população) na caixa do lado *cases* ***from de first*** cases deverá colocar o tamanho da população, neste caso 75 seguido dos comandos **Continue** – **OK.**

A título de curiosidade, se quisesse calcular todas as amostras (sem reposição) de n= 30 de uma população N=75 utentes de um centro de saúde, a n.º de amostras seria de 7.818794306^{20} amostras possíveis. Não lhe recomendamos que esteja tanto tempo a solicitar ao SPSS que as faça. Basta para a sua compreensão que faça as 20 primeiras.

Atenção, que logo de acabar de executar estes procedimentos o programa cria na base (SPSS – data editor) um corte transversal nos números de alguns casos da 1.ª coluna (*cases*) criando uma nova variável paralela à variável idade com a designação **filter_$**, com os números 0 e 1. O programa usa este procedimento para seleccionar alguns casos e ignorar outros.

Para calcular a média da primeira amostra de n=30 basta que seleccione o comando **Statistics** – **Summarize** seguido de **Descriptive**s colocando a variável idade.

Para escolher uma nova amostra deverá seleccionar de novo o comando **DATA** – **Select Cases,** seguido de **OK,** ao aceder ele automaticamente cria uma nova amostra, não sendo necessário nenhum comando adicional senão o voltar simplesmente lá. Pode verificar na sua base que houve uma alteração nos casos seleccionados através da variável **filter_$**. Vamos apresentar um quadro com as médias das 10 primeiras amostras seleccionadas pelo SPSS.

Quadro 4 – médias resultantes de 10 amostra de n=30 seleccionadas aleatoriamente pelo SPSS de uma população N=75

Amostra n.º	*Média de cada amostra*
1	42,40
2	45,10
3	46,33
4	39,73
5	48,50
6	43,97
7	43,17
8	45,47
9	41,23
10	43,27

Como poderá verificar, a média da população (μ) é de 44,81 anos, e a média das médias amostrais para as 10 primeiras amostras (deveríamos ter utilizado todas) de n=30, pode verificar que é 43,91 anos, valor um pouco inferior.

Face ao que temos vindo a apresentar, constata desde logo que quando realizamos uma investigação, esta informação é muito útil, pois é impensável que um investigador utilize repetidas amostras de tamanho n; pelo contrário, recorre apenas a uma amostra, e a base para essa decisão é o que temos vindo a apresentar. O primeiro facto que observamos é que a média das médias amostrais é igual à média populacional ($\mu_{\bar{x}} = \mu$) e o desvio padrão passa a assumir uma outra designação, *erro padrão da média* (**SE**[6]), que é dado por $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. Este indicador - *erro padrão da média* é muito importante, pois permite-nos verificar a quantidade de variação esperada nas médias, em função do acaso. Se o seu valor for pequeno, há uma boa probabilidade da média da amostra estar próxima da média populacional, se for grande, é sinal que tal pode não acontecer, e a média ser muito diferente da verdadeira média populacional. $\sigma_{\bar{x}}$ depende, como podemos verificar, de dois factores, a dispersão da variável na população e da raiz quadrada de n.

Assim sendo, podemos falar de um importante teorema designado de teorema **do limite central**. Este teorema afirma que sempre que uma amostra aleatória de tamanho n é seleccionada de qualquer distribuição com média μ e desvio padrão σ, então a distribuição das médias amostrais ($\mu_{\bar{x}}$) será aproximadamente normal com média $\mu_{\bar{x}} = \mu$ e desvio padrão $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.

Resumindo:

1. a média da distribuição das médias amostrais será igual à média da população $\mu_{\bar{x}} = \mu$.

2. No caso da população base ser não normal, a distribuição das médias amostrais será aproximadamente normal para grandes amostras; à medida que se aumenta o tamanho amostral a distribuição das médias tende para a normalidade. A distribuição das médias amostrais torna-se normal para amostras ≥ 30.

3. O desvio padrão das distribuições amostrais é igual ao desvio padrão da população dividido pela raiz quadrada do tamanho da amostra $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.

[6] Como o programa SPSS utiliza a designação de *Standart Error* (*of mean*), para o erro padrão da média opta-se por utilizar a sigla *SE*, como se encontra na maioria dos artigos científicos publicados em língua Inglesa, já que a sigla SEM quer significar *Structural Equation Modeling*.

Lembra-se do conceito de **variabilidade amostral**, conduz-nos ao **erro amostral**, ou seja, diferentes amostras de tamanho igual e seleccionadas da mesma população produzem valores médios um pouco diferentes. Este erro amostral refere-se à flutuação dos valor da estatística de uma amostra para outra. Essa quantidade de erro amostral é a diferença entre o valor que se obtém da amostra e o parâmetro populacional, conforme se pode ver pelo quadro. A fórmula desse erro é dado pela diferença entre a média ($\overline{x}$) amostral e o parâmetro populacional (μ) (traduz-se por $\overline{x} - \mu$).

Nas nossas investigações esse erro não é conhecido porque na realidade não sabemos os parâmetros populacionais, o que se faz é estimar a quantidade desse erro. É de salientar que esse erro se deve à variabilidade amostral, por isso assume-se que é aleatório e está presente em todas as investigações que alguém possa fazer. Repare no exemplo seguinte para amostras de n=2:

$\mu_{\overline{x}}$	$\overline{x} - \mu$
2,95	- 0,55
3,15	- 0,35
3,5	0
3,85	0,35
4,05	0,55

Em termos práticos, o erro padrão ajuda-nos a verificar se um *estimador como a média é ou não um estimador preciso.* Se for grande o valor o estimador é **impreciso**, se for pequeno o estimador é **preciso**. Como apenas utilizamos habitualmente uma amostra, isto é, não fazemos todas as amostras aleatórias de tamanho n, utilizamos a média amostral como estimador da média da população. E como não conhecemos o desvio padrão da população, o seu valor é dado pelo erro padrão da média, ou seja, $\frac{s}{\sqrt{n}}$.

Embora possam parecer muito semelhantes, o desvio padrão (s) e o erro padrão da média (SE) – Standart Error Of mean) utilizam-se na investigação com objectivos diferentes. Enquanto o desvio padrão se utiliza como medida de dispersão da distribuição (em relação à média), o erro padrão da média serve para caracterizar a média tomada enquanto estimador - a sua precisão.

Quando se pretende descrever apenas a variabilidade dos dados na amostra, o desvio padrão é dado por s = $\sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$. Quando se pretende usar as estatísticas amostrais como estimadores dos parâmetros populacionais, devemos proceder a algumas transformações nas fórmulas de forma a que o estimador não menospreze o parâmetro populacional, seja a variância ou o desvio padrão, devem ser multiplicadas por um factor de correcção ($\sqrt{\frac{n}{n-1}}$) ou então divididas por n-1 no denominador.

A fórmula com a correcção permite-nos verificar que o erro da estimativa é maior quando as amostras são pequenas, e diminui se aumentarmos a amostra. Por exemplo, se a amostra for de 20 com a aplicação do factor de correcção verificamos que o valor é grande ($\sqrt{\frac{15}{15-1}} = \sqrt{1,07}$), quando a amostra é de 150 o valor do factor de correcção é de 1,00. Se quisermos rescrever a fórmula do desvio padrão incorporando o factor de correcção teremos **s=** $\sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$.

Um exemplo pode-nos mostrar comparativamente um estimador tendencioso e não tendencioso da variância, colocando no denominador N e N-1. Imagine uma população hipotética de N=4, nomeadamente com os valores 8, 9, 10 e 11, a sua média μ=9,25 e σ^2=1.25.

Se escolhêssemos todas as amostras possíveis de n=2 e calculássemos a variância usando no denominador n e n-1, verificaríamos que quando o denominador é n-1 a média das variâncias amostrais é de 1,25 e quando o denominador é apenas n, o valor seria de 0,625.

Constatamos que nunca podemos predizer sem algum erro associado, bem como devemos considerar que o erro da estimação é maior em amostras pequenas e por consequência menor em amostras maiores, ou seja, diminui com o tamanho da amostra.

Vamos verificar como trabalha o SPSS quando apresenta medidas de tendência central e de dispersão. Para isso imagine que solicitávamos ao programa que calculasse a média, o desvio padrão e a variância, bem como o erro padrão da média para os valores de estrógenio no plasma (gramas/100ml), no pós-parto imediato para uma amostra que seleccionamos aleatoriamente de n=10.

Valores
2,6
3,7
3,3
3,1
3,5
3,1
3,7
3,5
3,0
3,9

As estatísticas calculadas através do SPSS são apresentadas no quadro 6, retirado directamente do output do SPSS.

Quadro 6 – Estatísticas resumo relativas aos valores de estrogénio no plasma (gramas/100ml)

Valores de estrógenio no plasma	Statistic	Std. Error
Mean (média)	3,3400	,125
Variance (variância)	,156	
Std. Deviation (desvio padrão)	,39497	

Conforme podemos verificar o SPSS, o valor da média da amostra é de 3,34 gramas/100ml, com um desvio padrão já estimado e aplicando o factor de correcção de 0,395 gramas/100ml, ou seja utilizando a fórmula $s=\sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}}$.

No que diz respeito ao erro padrão da média (SE) ele é de 0,125, ou seja, ($\frac{0{,}39497}{\sqrt{10}}$).

2. Estimação pontual e Intervalar para a média (μ)

Compreendido o conceito de distribuição amostral das médias, estamos aptos para começar a realizar inferências estatísticas. Um conceito importante em estatística e com o qual realizamos inferências é o de **estimação,** seja ela **pontual** ou **intervalar**. Trata-se simplesmente de, com base nos dados colhidos de uma amostra aleatória, estimar o valor populacional, quer a estimativa seja pontual (um único valor), quer a estimativa seja intervalar (um intervalo de valores).

Por exemplo, numa amostra aleatória de 30 recém-nascidos, a média ($\bar{X}$) do peso à nascença foi de 3,175 kg, com um desvio padrão (s) de 0,9 kg. Assim, se a amostra for aleatória, podemos fazer uma estimativa pontual da média populacional com base na media amostral, isto é, $\bar{x} \cong \mu$, no entanto como se observou na distribuição das médias amostrais, esta média está sujeita às flutuações devidas à variabilidade (as médias amostrais tendem a distribuir-se em torno da média populacional), ou seja, pode ser um pouco inferior ou superior em relação ao verdadeiro parâmetro populacional, o qual na realidade desconhecemos.

Então, é comum, para além da estimativa pontual de um parâmetro, que neste caso era de 3,175 kg, dispor de alguma forma de intervalo que indique a confiança que se pode depositar na *estimativa*, para isso constrói-se um **intervalo de confiança** (estimativa intervalar) para o verdadeiro valor populacional, isto é, um intervalo de valores no qual acreditamos que está contido o verdadeiro parâmetro populacional, com uma determinada probabilidade.

O cálculo da estimativa intervalar está condicionada por vários factores como o *grau de confiança quer desejamos*, o *tamanho da amostra* e o *conhecimento ou desconhecimento* do desvio padrão da população. Iremos centrarmo-nos apenas no caso em que não conhecemos o desvio padrão da variável na população (necessitamos de o estimar), e que representa a quase totalidade das nossas investigações, pelo que iremos fazer recurso a uma distribuição conhecida por ***distribuição t***[7], já que é adequada quando o desvio padrão é desconhecido e pode assim ser usada para qualquer tamanho de amostra (para grandes amostras os valores de *t* e de *z* são idênticos).

Como se mostrou, qualquer estimativa envolve erro. Este erro é inevitável na investigação, no entanto existe forma de o "controlar". Mesmo no caso dos intervalos de confiança, existe sempre a possibilidade do intervalo não conter o verdadeiro valor do parâmetro populacional, mas o mais provável é que contenha. Neste caso deveremos depositar o máximo de **confiança** no intervalo que vamos construir, ou seja, ter o máximo de confiança, sabendo que nunca se tem 100%.

Ao criar um intervalo de confiança, o investigador deposita o máximo de confiança possível (estatisticamente designa-se por **grau ou nível de confiança**), pois se não tem 100% de certeza, deve procurar ter o máximo possível (*quanto maior melhor*), por exemplo 95% ou 99% de confiança no intervalo que se constrói. Complementarmente, quando se aumenta a confiança diminui-se o erro, por exemplo, um grau de confiança de 95% acarreta um erro de 5%.

[7] A distribuição t é teoricamente adequada quando a distribuição é normal, opção particularmente importante quando a tamanho da amostra é ≤ 30.

Vejamos o exemplo do peso de uma amostra aleatória de 30 recém-nascidos à nascença em que a média do peso foi de 3,175 kg, e um desvio padrão de 900 gramas. Naturalmente que se poderia fazer uma estimativa pontual. No entanto e pelo que já sabemos relativamente à média, será mais seguro construir um intervalo com 95% de confiança.

Assim, teremos para o intervalo de confiança para a média:

$\mu \in \pm \left] \bar{x} \pm t \frac{\hat{s}}{\sqrt{n}} \right[$ onde t é tal que P(-t<**t**<t)= 1-α

Os sinais $\pm$ dizem-nos que o intervalo centrado na média implica um limite superior e um limite inferior. Como sabemos o erro padrão da média (*SE*) será $\frac{\hat{s}}{\sqrt{n}}$, logo o desvio padrão amostral (s) é igual a 0,9kg,

Calculando teremos: $\hat{s} = 0{,}9\sqrt{\frac{30}{30-1}} = 0{,}915$.

Logo, falta apenas procurar um valor *t* para prosseguirmos no cálculo do intervalo de confiança. Se a amostra fosse grande e/ou estivéssemos a trabalhar com o desvio padrão populacional (conhecido), seria fácil, pois esse valor é dado pela tabela de distribuição normal padronizada, mas como tal não acontece, vamos procurar o valor de *t* com recurso à tabela 1 (distribuição *t*) que nos falta, para tal devemos apenas recordar o grau/nível de confiança que desejamos e que é de 95% (ou seja, 1-α=95%) e o tamanho da nossa amostra, neste caso n=30.

Nas linhas da 1.ªa coluna encontram-se os **graus de liberdade**[8], neste caso o valor dos graus de liberdade será 29 (30–1). Para encontrar o valor basta simplesmente fazer a interpolação dos gl com o valor do nível de confiança. Se não encontrar o valor exacto dos gl, deve escolher o anterior imediato, por exemplo imagine *gl*=89, deveria utilizar o valor de *gl*=80. Como pode ver, o valor de *t* depende do nível de confiança desejado (ou erro assumido) e do n.º de graus de liberdade.

[8] Os graus de liberdade referem-se ao n.º de observações que podem variar depois de serem colocadas certas restrições aos dados. Um exemplo simples: um enfermeiro tem 5 doentes para serem acamados na sua enfermaria e resolve distribui-los aleatoriamente pela 5 camas que dispõe. Como pode verificar, todos são distribuídos aleatoriamente menos o último, que ficará na cama que restar não tendo escolha possível. Por analogia, se tivermos 10, 12, 14, 16, e 18, a sua média será de 14. Se calcular o desvio de cada um em relação à média e somar todos os desvios (positivos em negativos) verifica que o total é 0. Isto tem a ver com a própria definição de média. Os cinco números que escolhemos aleatoriamente são livres de variar (assumir qualquer valor) e são igualmente independentes pois a escolha de um não influenciou a escolha de outro, já os desvios não são independentes, pois a média restringe a liberdade que os potenciais valores podem assumir. No exemplo anterior apenas os valores de 4 deles (desvios) são livres de variar, o último não. Assim e porque os 4 primeiros desvios são conhecidos, o último é automaticamente encontrado, pois o somatório das diferenças deve dar 0. Mais fácil será pensar ainda num conjunto de 15 dados que dá $\bar{x}$ = 7,5. Quantos valores (dos dados) é que necessitamos saber para ter toda a informação relativamente a esta média? (n-1). Assim se conhecemos os primeiros 14 valores, o último é fixo. Mais à frente, em cada teste apresentaremos a forma de cálculo dos graus de liberdade que deve ser usada nesse teste.

Tabela 1 - Distribuição *t*

	Nível de confiança			
	.90	.95	.98	.99
	α para teste unilateral			
	0,05	0,025	0,01	0,005
	α para teste bilateral			
Graus de Liberdade	0,10	0,05	0.02	0,01
1	6.31	12.71	31.82	63.66
2	2.92	4.30	6.97	9.93
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5	2.02	2.57	3.37	4.03
6	1.94	2.45	3.14	3.71
7	1.90	2.37	3.00	3.45
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25
10	1.81	2.23	2.76	3.17
11	1.80	2.20	2.72	3.11
12	1.78	2.18	2.68	3.06
13	1.77	2.16	2.65	3.01
14	1.76	2.15	2.62	2.98
15	1.75	2.13	2.60	2.95
16	1.75	2.12	2.58	2.92
17	1.74	2.11	2.57	2.90
18	1.73	2.10	2.55	2.88
19	1.73	2.09	2.54	2.86
20	1.73	2.09	2.53	2.85
21	1.72	2.08	2.52	2.83
22	1.72	2.07	2.51	2.82
23	1.71	2.07	2.50	2.81
24	1.71	2.06	2.49	280
25	1.71	2.06	2.49	2.79
26	1.71	2.06	2.48	2.78
27	1.71	2:05	2.47	2.77
28	1.70	2.05	2.47	2.76
29	1.70	2.05	2.46	2.76
30	1.70	2.04	2.46	2.75
40	1.68	2.02	2.42	2.70
60	1.67	2.00	2.39	2.66
120	1.66	1.98	2.36	2.62
∞	1.65	1.96	2.33	2.58

No cabeçalho da tabela vemos o valor do nível de confiança ou o **alpha (α)**. Ou seja, podemos escolher o nível de confiança nas colunas através do nível de confiança ou do alpha que designa o erro, portanto ou um nível de confiança de 95% ou o erro máximo de 5%.

Neste caso, o valor do ***t*** (IC 95% e gl=29) é de 2.05. Resta-nos calcular o resto da fórmula, isto é, $\overline{x} \pm 2{,}05\frac{0{,}915}{\sqrt{30}}$. Bastará somar e subtrair à média o valor encontrado (0,342), ou seja, o intervalo de confiança de 95% para o valor populacional dos recém-nascidos oscilará entre 2,833 kg e 3.515 kg.

O que significa este valor? Significa tão simplesmente que possuímos 95% de confiança de que a média do peso dos recém-nascidos se situa entre 2,833 kg. e 3,515kg. Ou melhor, possui-se 95% de confiança de que a média real do peso dos recém-nascidos está contida no intervalo entre 2,833 kg **e** 3,515 kg**.**

Esta amplitude do intervalo[9] é estreita o que reflecte um estimador preciso. Isto implica se escolhêssemos diferentes amostras do tamanho desta da mesma população, e realizássemos o mesmo procedimento, 95% deles iriam conter o valor da média populacional.

[9] Na realidade a amplitude do intervalo de confiança depende de quatro coisas, a dispersão dos valores populacionais, do nível de confiança desejado, do erro e do tamanho da amostra.

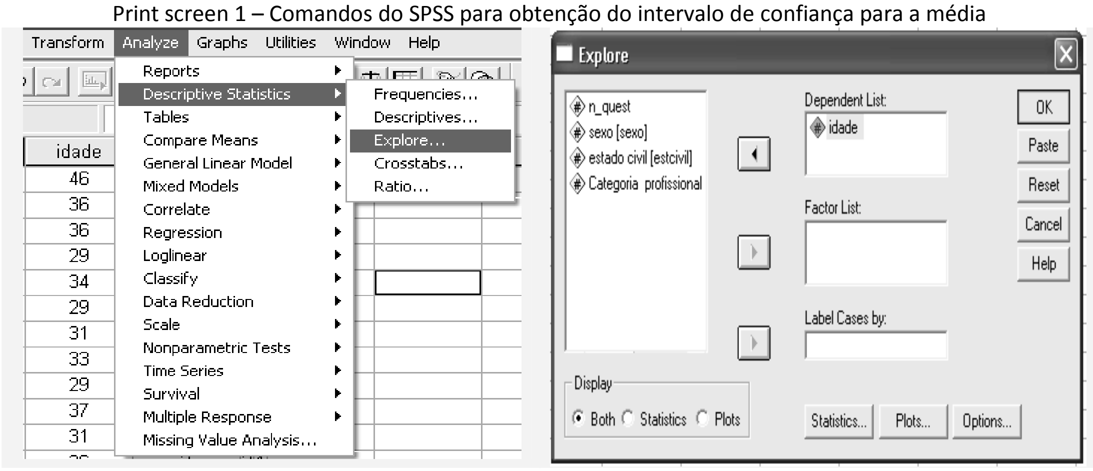
Convém ainda salientar que fomos escolher o alpha de 0,05 bicaudal (isto significa para as duas caudas), ou seja, para os dois lados da distribuição. Este conceito será explicado nos testes de hipóteses.

2.1. Estimação pontual e Intervalar para a média (μ) no SPSS

O programa SPSS também calcula intervalos de confiança para a média, opção que está presente sempre que se trabalhe com o comando **Explore**. Os comandos utilizados são os que aparecem seguidamente, e ao longo das próximas páginas.

Deixamos a designação de figuras para os comandos do SPSS. Nas próximas páginas ser-lhe-á atribuída a designação de *print screen* para as operações efectuadas com o SPSS[10].

Print screen 1 – Comandos do SPSS para obtenção do intervalo de confiança para a média



O quadro de resultados gerados pelo SPSS 11.5 é o que apresentamos seguidamente.

Quadro 6 – Estatísticas Resumo e IC 95% para a média da idade

		Estatística	Erro padrão
Média		34,35	,389
Intervalo de 95% de confiança para a média	Limite inferior	33,58	
	Limite superior	35,11	
Mediana		34,00	
Desvio padrão		6,312	

Como se pode observar, a média da idade é de 34,35 anos, com um desvio padrão de 6,312 anos. O erro padrão da média (SE) é de 0,389 anos e os valores do intervalo de 95% de confiança para a média oscilam entre os 33,58 e os 35,11 anos.

[10] Como referimos, os cálculos foram efectuados com a versão 11.5 do SPSS. Presentemente o software está na versão 17.0, ainda assim as diferenças e novidades de uma versão para a outra são reduzidas para aquilo que são os objectivos desta ferramenta.

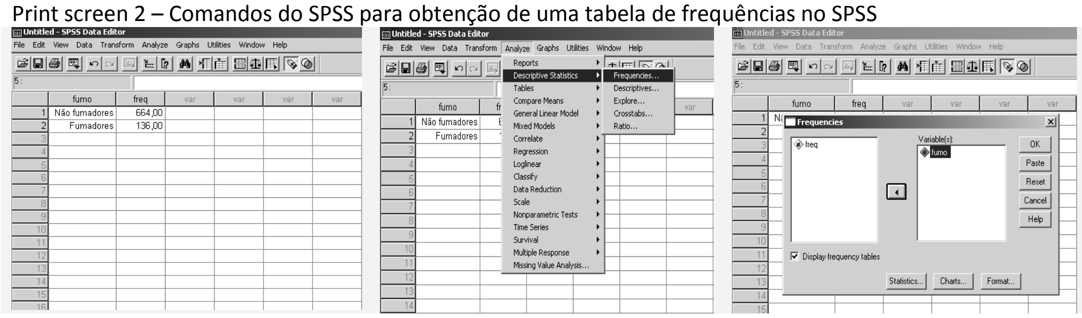
3. Estimação Intervalar para a proporção (π)

O caso dos intervalos de confiança para a proporção é análogo ao raciocínio desenvolvido no exemplo anterior[11].

Imagine que uma equipa de Professores da ESEnfC seleccionou aleatoriamente uma amostra de 800 alunos que frequentam a escola no ano de 2008, com intuito de avaliar a percentagem de estudantes de Enfermagem que são fumadores. Dos 800 alunos, verificou-se que 136 eram fumadores enquanto os restantes 664 afirmaram nunca ter fumado.

Como se referiu anteriormente para a média, poderemos também utilizar a proporção amostral como estimativa pontual da proporção populacional, sabendo do erro que isso acarreta. Para tal poderemos solicitar ao programa SPSS que faça os cálculos, criando uma tabela de frequências. Basta que execute os comandos seguintes (de acordo com os print screen´s):

Print screen 2 – Comandos do SPSS para obtenção de uma tabela de frequências no SPSS



Note que a base de dados (SPSS Data Editor) é ligeiramente diferente. A primeira variável designa-se de «fumo» e a 2.ª de «freq». A primeira variável tem duas categorias, nomeadamente *não fumadores* (ao qual foi atribuído o número 1) e *fumadores* (atribuído o número 2), a variável «freq» corresponde às frequências absolutas observadas. Para criar efectivamente a tabela de frequências deverá proceder a algumas transformações descritas no print screen 4.

O quadro 7, trata-se da *frequency table* gerada pelo SPSS, nome com que o programa designa o quadro. Os termos em português foram traduzidos por nós. Como pode observar, 136 estudantes (17,0%) afirmam serem fumadores e 664 (83,0%) não fumadores.

[11] Relembre que uma distribuição de proporções amostrais é a distribuição de probabilidades que indica quão prováveis são diversas proporções (p) calculadas a partir de amostras aleatórias tendo em conta o tamanho das amostras e a proporção populacional (π). Ainda que o Teorema do Limite Central se aplique em sentido estrito às médias, outras medidas, nomeadamente as proporções, têm idêntico comportamento em grandes amostras (n>30).

Quadro 7 – Distribuição absoluta e percentual de estudantes fumadores e não fumadores da ESEnfC

		Frequência	%	% válida	% acumulada
Validos	Não fumadores	664	83,0	83,0	83,0
	Fumadores	136	17,0	17,0	100,0
	Total	800	100,0	100,0	

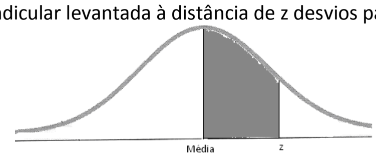
Se a partir desta estatística amostral (estimativa pontual) quisermos construir um intervalo com 95% de confiança para a proporção populacional de fumadores, a tarefa é acessível. Esse intervalo é dado por

$\pi \in \pm \left] p \pm z\sqrt{\frac{p.q}{n}} \right[$ onde z é tal que P(-z<**z**<z)= 1-α.

Neste caso, «p» é a proporção de «sucesso» (fumadores) e «q» a de «insucesso» (não fumadores). O valor de z=1.96 para IC com 95% é obtido a partir da tabela de distribuição normal padronizada z, conforme se pode verificar de seguida através da tabela 2.

Tabela 2 - Distribuição Normal Padronizada

Cada valor da tabela indica a proporção da área total sob a curva contida no segmento delimitado por uma perpendicular levantada na média e uma perpendicular levantada à distância de z desvios padrões unitários.



Ilustrando: 43,57% da área sob a curva normal estão entre a ordenada máxima e um ponto 1,52 desvios padrões adiante.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998

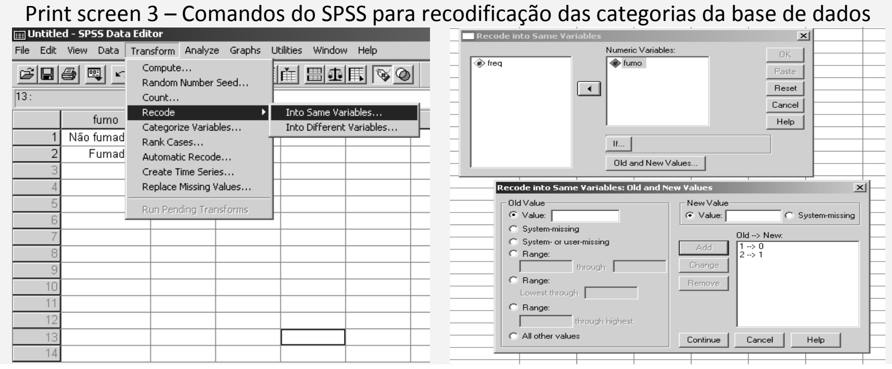
Assim substituindo os valores na formula temos: $0{,}17 \pm 1{,}96\sqrt{\dfrac{0{,}17x0{,}83}{800}}$. Deveremos pois somar e subtrair a 0,17 o valor de 0,026.

Podemos pois afirmar que a percentagem de estudantes fumadores na população representada por esta amostra, está contida no intervalo entre os 14,4% e os 19,6%, com um nível de confiança de 95%.

3.1. Estimação pontual e Intervalar para a proporção (π) no SPSS

Para efectuar estes cálculos no SPSS o procedimento é muito acessível, conforme se pode observar pelos *print screen* 3 e seguintes, no entanto necessita de procedimentos adicionais, nomeadamente recodificar os valores atribuídos à variável fumo. A alteração é feita através no comando **Transform** – **Recode** – **in to same variable: Old and New values**.

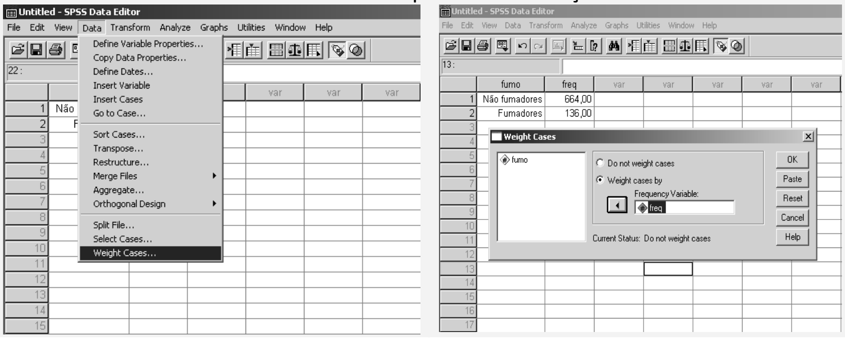
Print screen 3 – Comandos do SPSS para recodificação das categorias da base de dados



A transformação a efectuar é simples. Solicita ao programa que atribua o número 0 ao código 1, o número atribuído anteriormente aos não fumadores, passando os fumadores a ter o código 1 agora, deixando por isso o antigo código 2. De modo claro o valor 1 passou a valer 0 e o valor 2 passou a valer 1.

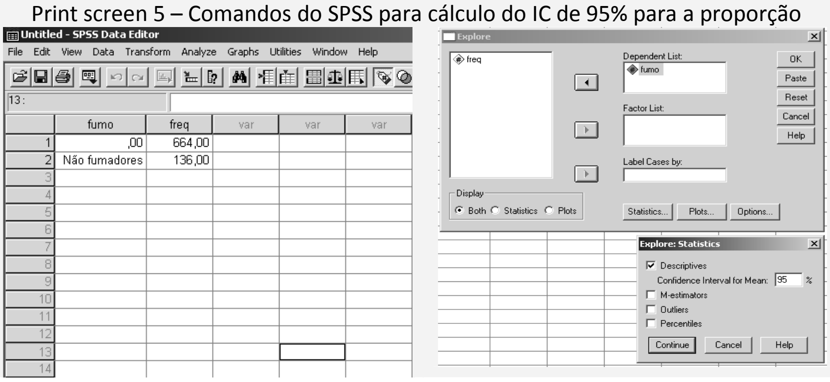
Seguidamente deverá aceder ao comando **Data**, seleccionar **Weight Cases** e depois seleccionar a variável «freq» e colocá-la com o cursor na opção **Frequency Variable**, seguido de **OK.**

Print screen 4 – Comandos do SPSS para transformação da base de dados



Como pode observar pelo *print screen* 5, a base de dados (SPSS Data Editor) já está modificada. Para aceder ao Explore, deverá fazer como fez para o cálculo do intervalo de confiança para a média através dos comandos **Analyze** – **Descriptive Statistics** – **Explore**.

Print screen 5 – Comandos do SPSS para cálculo do IC de 95% para a proporção



Se der um clique com o cursor no botão **Statistics** do **Explore** verifica que por defeito está seleccionada a 1.º opção e que nos interessa, nomeadamente **Confidence Interval for Mean: 95%**.

De seguida é apresentado o quadro com resultados dos cálculos efectuados pelo programa. Como pode verificar, os valores dos limites inferior (*Lower Bound*) e superior (Upper Bound) do intervalo de 95% de confiança para a proporção são respectivamente, 14,4% e 19,6%.

Quadro 8 – Intervalo de confiança para a proporção populacional de fumadores (N=800)

		Statistic
95% Confidence Interval	Lower Bound	,144
	Upper Bound	,196

Os intervalos de confiança, tanto para a média como para a proporção, têm sido preteridos na investigação ao longo das últimas décadas, favorecendo o uso dos testes de hipóteses[12]. Muitas razões poderão ser evocadas mas falaremos disso mais adiante. Ainda assim é possível encontrar artigos nas revistas científicas que utilizam os intervalos de confiança, mesmo que se observem algumas incorrecções ou imprecisões, como por exemplo a utilização indiscriminada dos símbolos $\pm$ a acompanhar a média ou a proporção. Na verdade ficamos sem saber, se se trata do desvio padrão, tomado enquanto medida de dispersão em relação à média, ou se o erro padrão da média, utilizado com intuitos de caracterizar a precisão do estimador. Estas críticas observam-se em 15% dos artigos científicos publicados e analisados por Altman e Bland (2005).

4. Cálculos dos intervalos para a média e proporção na maquina de calcular

É comum afirmar-se que as máquinas facilitam o trabalho de cálculo, mas não deve limitar-se a sua utilização ao simples exercício de cálculo, esquecendo o trabalho escrito e mental, imprescindíveis à realização de tarefas mais complexas, especificamente quando relacionadas com a investigação, dado que a estatística passou e deve ser assumida como um sistema lógico gerador de significados uma linguagem de investigação (Gameiro, 1998). Neste contexto apresentamos os comandos para o cálculo dos intervalos de confiança para ambos, média e proporção. São aplicados às máquinas TEXAS e CASIO.

O exemplo diz respeito a uma amostra aleatória de 83 grávidas seguidas no C. S. dos Lares, em que foi estudado o valor da excreção de pregnanodiol às 24 semanas de gestação. Os valores obtidos tinham uma média de 20,50 mg/24 horas e um desvio padrão de 2,8 mg/24 horas. Determinar o intervalo de 95% de confiança para média da população de referência.

Instruções para cálculo do intervalo de confiança para a média (μ)	
TEXAS	CASIO
1. Pressione o botão **STAT** e depois mude o cursor para a opção **TESTS**; 2. Seleccione a opção **8: TInterval…;** 3. Seleccione **Stats** pressionando a tecla **ENTER.** 4. Mude o cursor para $\overline{X}$: e depois introduza o valor da média da amostra ($\overline{X}$:20,50); 5. Mude o cursor para **Sx** e introduza o valor do desvio padrão amostral (**Sx:** 2,8); 6. Introduza o **n** (tamanho da amostra n:83); 7. Defina o nível de confiança desejado na opção **C-Level**:, neste caso **C-Level**:.95; 8. Seleccione a opção **Calculate**	1. Pressione o botão **MENU** e depois seleccione a opção INTR (tecla F4); 2. Seleccione a opção **t** (tecla F2); 3. Seleccione a opção **1-S** (tecla F1) 4. Na opção **Data**, seleccione a opção **var** (F2); 5. Ande com o cursor para a opção **C-Level** e defina a confiança desejada (**C-Level**:.95) e depois pressione a tecla EXE; 6. Introduza o valor de $\overline{X}$: ($\overline{X}$:20,50) – EXE; 7. Introduza o valor do desvio padrão amostral **xσn-1:** 2,8 seguido de EXE 8. Introduza o **n** (tamanho da amostra) – EXE; 9. Comando **EXECUTE**
Output (Resultados): (19.89, 21.11) $\overline{X}$:20,50 Sx: 2,8 n=83	

[12] Apesar das «virtudes» dos testes de hipóteses, muitas conclusões são abusivas, nomeadamente quando a interpretação se cinge apenas à significância estatística.

Como se pode observar, o intervalo de 95% de confiança para a média da população de referência situa-se entre os 19.89 mg/24 horas e os 21.11 mg/24 horas. Simbolicamente pode referir-se que 19.89 mg/24 horas $<\mu<$21.11 mg/24 horas com um nível de confiança de 95%. Tome em atenção as considerações efectuadas sobre o desvio padrão da amostra.

Agora vamos calcular o intervalo de confiança para a proporção. Imagine que numa amostra aleatória de 180 doentes jovens adultos (18-35 anos) com mais 15 dias de internamento em serviços de medicina, verificou-se uma prevalência de depressão de 22%. Determine o intervalo de 95% de confiança para a proporção de casos de depressão na respectiva população.

Instruções para cálculo do intervalo de confiança para a proporção (π)	
TEXAS	CASIO
1. Pressione o botão **STAT** e depois mude o cursor para a opção **TESTS**; 2. Desça com o cursor e seleccione a opção **A:1 – PropZInt…** - seguido de **ENTER** 3. Introduza em **X:** o número de sucessos (**X:40♣)** 4. Introduza em **n:** o tamanho da amostra (**n: 180**) 5. Defina na opção **C-Level**: o nível de confiança desejado (**C-Level:.95**); 8. Seleccione a opção **Calculate** e **ENTER.**	1. Pressione o botão **MENU** e depois seleccione a opção INTR (tecla F4); 2. Seleccione a opção **z** (tecla F1); 3. Seleccione a opção **1-P** (tecla F3) 4. Na opção **C-Level** defina a confiança desejada (**C-Level**:.95) e depois pressione a tecla **EXE**; 5. Introduza em **X:** o número de sucessos (**X:40♣)** **6.** Introduza o **n** (tamanho da amostra) – **EXE**; Comando **EXECUTE** seguido de **EXE**
Output (Resultados): (.16, .28) p:.222 n=180	

♣ Repare que não se coloca a percentagem de sucesso (p), mas sim o número de casos correspondentes, neste caso são 40 doentes (0,22x180)

Como se pode observar, a proporção da população de casos de depressão situa-se, com um intervalo de 95% de confiança, entre os 16,00% e os 28,00% Simbolicamente podem representar-se por **p($\pi\in$]0,16; 0,28[)= 0,95** ou então, **Int. Conf. 95%: 16,00% - 28,00%.**

4.1. Exercícios

1. Numa amostra aleatória de 93 mulheres entre as 8 e as 16 semanas pós-parto, seguidas na Maternidade da Boa Hora, realizou-se um estudo com a aplicação da "*Escala de Depressão Pós-Parto de Edimburgo*" (mínimo 0 pontos, máximo 30 pontos), tendo-se obtido uma média ($\bar{x}$) de 12.29 pontos com um desvio padrão (s) 6.32 pontos.

a) - Determine o *intervalo de 95% de confiança* para a média do nível de depressão medido através da escala indicada, na população representada pela referida amostra.

b) - Se pretendesse diminuir o erro da estimativa mantendo o grau de confiança, qual a estratégia a seguir? Justifique.

1.1. - Das 93 puérperas estudadas, 12 (12.90%) apresentavam uma pontuação na "*Escala de Depressão Pós-Parto de Edimburgo*" igual ou superior a 18, correspondendo a quadros depressivos graves.

a) Recorrendo à determinação do *intervalo de confiança*, verifique se este valor põe em causa a afirmação de que "15% das puérperas da comunidade da Boa Hora sofrem de depressão grave entre as 8 e as 16 semanas pós-parto" (utilize um IC de 95%).

2. O Centro de Saúde da Fonte Nova realizou um estudo com os idosos que recorreram à consulta de gerontologia durante o ano de 2006. Uma das variáveis estudadas foi a "*autonomia funcional nas actividades da vida diária*" medida através da escala de BARTHEL (0 a 100 pontos). Depois dos respectivos dados terem sido organizados, obteve-se a seguinte tabela:

K	F
00 - 20	9
20 - 40	13
40 - 60	38
60 - 80	35
80 - 100	14

a) Com base nos dados da tabela anterior, determine o intervalo de 95% de confiança para a média do nível *autonomia funcional nas actividades da vida diária* na população representada pela referida amostra.

3. O Centro de Saúde de Vila Branca em Janeiro de 2008 realizou um estudo por amostragem da situação de saúde das adolescentes entre os 13 e os 18 anos residentes no Bairro da Moita Velha. Na amostra de 56 adolescentes referida anteriormente verificou-se um valor médio de colesterol plasmático de 155.00 mg/100ml, com um desvio padrão de 12.50 mg/100ml. Tendo como base estes valores e considerando que aquela amostra é aleatória,

a) Estime o intervalo de 98% de confiança para a média dos valores médios de colesterol plasmático na população das adolescentes do Bairro da Moita Velha.

4. Na Unidade de Cuidados Intensivos aos Recém-Nascidos (UCIRN) da Maternidade Vida Nova realizou-se em 2007 um estudo com os recém-nascidos a termo com baixo peso (≤2500 g). Na amostra estudada de 76 mães de recém-nascidos de baixo peso, os valores de estriol sanguíneo na última semana da gravidez tinham uma média de 10.32 ng/dl, com desvio padrão de 2.14 ng/dl.

a) Com base nestes dados, estime um intervalo de 95% de confiança para a média de estriol no sangue da respectiva população.

b) Como poderia diminuir a amplitude do intervalo sem diminuir a confiança?

5. Numa amostra aleatória de 180 doentes jovens adultos (18-35 anos) com mais 15 dias de internamento em serviços de medicina, verificou-se uma prevalência de depressão de 22%.

a) Determine o intervalo de 95% de confiança para a proporção de casos de depressão na respectiva população.

6. Numa amostra de 140 doentes idosos internados nos serviços de medicina do Hospital de S. Francisco 14 apresentavam úlcera de pressão.

a) Determinar o intervalo de 90% de confiança para a prevalência (em%) de úlceras de pressão na respectiva população.

7. Numa amostra aleatória de 56 jovens do 12ª Ano de Escolaridade da Escola Secundária do Pontal verificou-se um valor médio ($\overline{x}$) de colesterol plasmático de 155,00 mg/100 ml e um desvio padrão (s) de 12,50 mg/100 ml.

a) Determinar o intervalo para a média de colesterol plasmático da população de jovens que frequentam nessa escola o 12ª Ano de Escolaridade, com um nível de confiança de 95%

b) Assumindo a média amostral como uma boa estimativa pontual da média populacional, qual o erro máximo provável?

5. Introdução aos testes de hipóteses

Na investigação científica os testes estatísticos não são usados para não rejeitar ou rejeitar hipóteses, mas apenas para perceber o peso da evidência contra ou a favor delas.
(Motulsky, 1995).

Dividimos a estatística em dois grandes campos, a *estatística descritiva*, e a *estatística inferencial* que começámos a apresentar anteriormente.

A primeira traduz-se num conjunto de procedimentos que visam a organização, resumo e apresentação dos dados colhidos no decorrer de uma investigação (Reis, 1991; Hill e Hill, 2000; Pestana e Gageiro, 2000; Kachigan, 1991; Bluman, 1997). Depois de organizados e resumidos os dados, há que proceder por intermédio de um outro conjunto de técnicas, designadas de técnicas de inferência estatística, à *ampliação* do nosso conhecimento do conjunto de dados (amostra) para um conjunto mais amplo, a população. (Polit, 1996; Runyon *et al.*, 1996; Argyrous, 2000).

Agora vamos pois falar de inferência estatística através de **testes de hipóteses**. Os *testes de hipóteses* são um conjunto de técnicas de inferência estatística que avaliam a probabilidade de determinados desvios, diferenças, ou relações observados nas distribuições amostrais serem devidos ao acaso ou à existência real na população, obviamente tendo por base observações realizadas em amostras (subconjuntos representativos da população).

Imagine que um enfermeiro estava interessado em determinar se o *local de trabalho contribuía para uma diferença no stress profissional nos enfermeiros*. Para isso escolheu enfermeiros de dois serviços (urgências e cuidados de saúde primários), aos quais administrou uma escala de avaliação do stress. Elaborada a base de dados e inseridos os questionários, calculou os *scores* de stress de cada enfermeiro, tendo posteriormente calculado as medidas de tendência central e dispersão. Imagine que os dados eram os do quadro 9.

Quadro 9 – Estatísticas resumo dos scores do stress dos enfermeiros em função do local de trabalho

Local de trabalho:	n	$\overline{X}$	*s*
Serviço de urgências	35	72,60	8,16
Cuidados de Saúde Primários	33	68,26	7,98

O que o enfermeiro pretende saber é se a diferença observada entre as médias do stress (72,60 e 68,26) pode ser devida ao acaso ou à existência real na população. Servindo-se dos testes de hipóteses toma uma decisão.

Neste exemplo a *variável dependente* será o stress (medida através de uma escala de stress), os grupos de pertença (enfermeiros dos S.U. ou C.S.P.) são designados habitualmente por variável independente[13].

Como o nome indica, tratam-se de testes de hipóteses, assim referem-se a variáveis. Na generalidade uma hipótese é uma afirmação sobre uma relação entre variáveis, neste caso, stress e local de trabalho dos enfermeiros. Então, é com as variáveis que construímos hipóteses.

As hipóteses podem assumir diferentes designações. Uma distinção é entre *hipótese conceptual* e *hipótese estatística* [designada de H_0 – (agá zero)]. Esta distinção que nos parece pertinente é apresentada por Runyon *et al* (1996). Estes autores consagram primeiramente dois tipos de hipóteses, *conceptual* e a *estatística*[14] (nula). A primeira é uma afirmação que prevê uma relação entre variáveis (usualmente uma dependente e uma independente); a segunda descreve uma relação matemática entre dois ou mais parâmetros populacionais. Assim no referido estudo o enfermeiro investigador estabeleceu a seguinte hipótese: "*o stress profissional dos enfermeiros é diferente consoante o seu local de trabalho*".

De facto, nesta hipótese ele afirma que o stress (variável dependente) é diferente consoante o local de trabalho (variável independente). Já no que se refere à hipótese estatística, o investigador, ao seleccionar aleatoriamente os enfermeiros para os grupos, sabe que a amostra de enfermeiros do S.U. representa a população de enfermeiros que trabalham em serviços de urgência, bem como que a outra população é representada pela referida amostra (C.S.P.), ou seja, utiliza as médias amostrais como estimadores dos parâmetros populacionais. Ele pode traduzir essa linguagem por meio de símbolos, como a média da população dos enfermeiros de S.U. (μ_1) é igual à média dos enfermeiros de C.S.P. (μ_2), ou seja, $\mu_1=\mu_2$.

Uma das exigências da lógica inferencial é a apresentação de uma contra hipótese uma afirmação contrária à afirmação da hipótese estatística ou nula, e designada de hipótese, neste caso, $\mu_1 \neq \mu_2$. Atenção que muitas vezes a hipótese alternativa (ou também designada de hipótese de investigação) equivale à hipótese conceptual.

---

[13] Na maioria das vezes, a designação não é adequada. A sua utilização prende-se mais com o desenho de investigação utilizado.
[14] Na verdade deveremos pensar em duas hipóteses, uma nula (H_0) e outra alternativa (H_1) ou contra hipótese.

Vamos deter-nos por alguns momentos num exemplo simples que pode ajudar a compreender a formulação lógica das hipóteses e a sua ligação com o raciocínio estatístico. Lembra-se da proposição universal "todos os cisnes são brancos"? Pode provar essa afirmação? Naturalmente que não, pois não conhece todos os cisnes e é bem possível que possam existir deles de outra cor. A única certeza que tem prende-se com o facto de todos os cisnes que viu serem brancos, ou seja, a que lhe vem da experiência, mas se aparecesse um cisne azul, isso bastaria para provar que estava errado, ou seja, para falsificar ou provar como errada a sua proposição.

Isto permite-nos verificar uma regra da lógica que afirma que nenhuma afirmação ou declaração pode ser provada como verdadeira, apenas como falsa. Na investigação faz-se recurso a um procedimento lógico designado de *modus tollens,* também designado de procedimento de falsificação. Este procedimento assenta no facto de que uma observação singular poder conduzir-nos à conclusão de que a premissa é incorrecta. A utilização do *modus tollens* na estatística exige a existência de duas hipóteses, a hipótese nula ou estatística (H_0) e a hipótese alternativa ou contra hipótese (H_1). A hipótese nula é a nossa afirmação à priori, ou seja, aquela que se vai procurar mostrar como falsa, isto é, utilizamos os resultados para concluir se ela é uma afirmação falsa. Se rejeitamos a hipótese nula e a declaramos como falsa o que acontece é que utilizamos a hipótese alternativa como uma afirmação mais adequada. Assim, no exemplo anterior se considerarmos, face à evidência empírica que a afirmação de que o stress é igual nos dois grupos é incorrecta, rejeitamos a hipótese nula declarando-a falsa e aceitamos a hipótese alternativa como sendo uma afirmação (explicação) mais adequada. Como facilmente podemos ver, os procedimentos utilizados nos testes de hipóteses baseiam-se nas regras da inferência negativa, ou seja, parte-se sempre do pressuposto que a hipótese nula é verdadeira.

Uma outra analogia que nos pode permitir também ajudar a compreender os testes de hipóteses, prende-se com o conceito de *inocência* no sistema criminal. Como sabemos, todo o acusado é inocente até prova em contrário, por analogia com os testes de hipóteses a hipótese nula é verdadeira até que uma evidência suficiente mostre que essa afirmação é incorrecta.

Se quisermos traduzir de forma simbólica a construção das hipóteses será:

H_0: o acusado é inocente de ter cometido o crime X

H_1: o acusado é culpado de ter cometido o crime X

No sistema criminal, para que um indivíduo seja condenado é necessário apresentar provas, provas estas que devem ir para além de uma *dúvida razoável*. Face a essas provas, o juiz toma uma decisão, considera o acusado culpado ou não culpado. Se existem semelhanças entre o raciocínio jurídico e os testes de hipóteses, existem algumas diferenças, nomeadamente no conceito de prova. A rejeição da hipótese nula não constitui prova que a hipótese alternativa seja válida em definitivo, ou seja, é apenas uma evidência, muitas vezes provisória de que a hipótese nula é provavelmente incorrecta, ou seja, mesmo que ela tenha sido declarada como incorrecta há possibilidade dela ser verdadeira.

Com os testes de hipóteses fazemos o seguinte: formulamos uma H_0 e, face ao conjunto de dados amostrais, tomamos uma decisão, ou rejeitamos ou não rejeitamos a hipótese nula. No exemplo da hipótese que refere que o stress dos enfermeiros é diferente consoante o local de trabalho (H_1:$\mu_1 \neq \mu_2$) partimos do principio que nada há que justifique à priori que as médias sejam diferentes, portanto a nossa H_0 ou hipótese a testar será $\mu_1 = \mu_2$, e é sobre esta que tomamos uma decisão: ou se rejeita ou não se rejeita. Quando ela se rejeita é porque houve uma evidência suficiente para considerar que efectivamente os níveis de stress são diferentes, caso não se rejeite era porque as diferenças nas médias se deveriam ao acaso e por isso mesmo, consideramos $\mu_1 = \mu_2$.

5.1. Erros tipo I e tipo II

Se nos lembrarmos de algumas histórias, por exemplo de indivíduos que foram condenados à pena de morte e mais tarde, já depois de terem morrido, se veio a verificar que não tinham sido eles os autores dos crimes, isso alerta-nos para uma possibilidade importante, o erro que se pode cometer numa decisão do tribunal. De facto, a possibilidade de errar está sempre presente e acompanha qualquer decisão. Podemos libertar um indivíduo que é culpado ou então prender um indivíduo que é inocente. Normalmente, o que se faz na dúvida é decidir a favor do réu. Também na estatística podemos cometer dois tipos de erros, são eles, rejeitar a hipótese nula quando ela é verdadeira, e não rejeitar a hipótese nula quando ela é falsa, no entanto estes erros têm designações próprias, são eles erro tipo I e erro tipo II. Vejamos no que se segue (quadro 10) os dois tipos de erros aplicados tanto à decisão estatística como à decisão no tribunal.

Quadro 10 – Quadro comparativo de decisões relativas a testes de hipóteses e erros associados (Inclui, por analogia, as decisões no sistema criminal e eventuais erros)

	Decidimos rejeitar a H_0 Declarado culpado	Decidimos não rejeitar a H_0 Declarado não culpado
H_0 é verdadeira *(na realidade inocente)*	Erro tipo I (α)	Decisão correcta
H_0 é falsa *(realmente culpado)*	Decisão correcta	Erro tipo II (β)

Naturalmente que um investigador deve estar alerta para estes tipos de erros. Qual o mais importante? Os dois, apesar de se dar mais importância ao erro tipo I. No entanto, na estatística o investigador pode por exemplo controlar à partida a probabilidade de cometer um erro tipo I. É de salientar que o erro tipo I também recebe habitualmente a designação de *nível de significância* do teste e é representado pela letra grega alpha (α), e o erro tipo II é representado pela letra grega beta (β), conforme aparece no quadro anterior.

Para sistematizar ideias:

Erro tipo I[15]: designado por alfa (a), define-se como a probabilidade de rejeitar a hipótese nula quando ela é verdadeira.
Erro tipo II: designado por beta (β), define-se como a probabilidade de não rejeitar a hipótese nula quando ela é falsa.

Como podemos então controlar o erro na estatística? A probabilidade de cometer um erro tipo I, como iremos ver, é estabelecida à partida pelo investigador. Como corremos sempre um risco, podemos controlar a probabilidade de cometer um erro tipo I, estabelecendo um **nível de significância**, ou seja, o risco que se corre *à priori*. Esse erro é normalmente de 5% ou 1%, estes valores são estabelecidos pelo investigador antes de realizar o teste de hipótese.

A situação do controlo do erro tipo II é mais complexa, e está relacionada com o poder dos testes estatísticos. O risco do erro tipo II é afectado por variados factores, problemas relacionados com as amostras, desenhos experimentais utilizados, relacionamento entre variáveis, bem como o tipo de teste que estamos a utilizar. Normalmente quando se estabelece um critério muito rígido para o α, aumenta a probabilidade de se cometer um erro tipo II. Falaremos dele mais adiante.

[15] Tecnicamente o nível de significância, simbolizado por α indica o tamanho da área numa distribuição teórica de probabilidades e que corresponde à rejeição da hipótese nula.

5.2. As hipóteses nula e de investigação

Do ponto de vista da investigação, as hipóteses podem ser apresentadas quanto ao tipo, função etc..., mas como essa discussão vai para além dos objectivos deste caderno de apoio, iremos restringir a discussão ao mínimo. Veja a definição de Kerlinger (1980), apresentada seguidamente relativamente ao conceito de hipótese:

> "uma hipótese é um enunciado conjectural das relações entre duas ou mais variáveis. Hipóteses são sentenças declarativas e relacionam, de alguma forma, variáveis a variáveis. São enunciados de relações, e, como os problemas, devem implicar a testagem das relações enunciadas. Problemas e hipóteses são semelhantes. Ambos enunciam relações, só que os problemas são sentenças interrogativas, e as hipóteses sentenças afirmativas. Às vezes são idênticos em substância. Uma diferença importante, entretanto: as hipóteses geralmente estão mais próximas das operações de teste e pesquisa" (pág. 38).

Genericamente podemos dizer que a hipótese estabelece uma relação entre variáveis[16]. Como sabe, variáveis "são aqueles aspectos, propriedades ou factores, mensuráveis ou potencialmente mensuráveis, através dos valores que assumem, discerníveis num objecto de estudo" (Köche, 1979, pág 54). Assumindo que as variáveis são elementos constitutivos das hipóteses, facilmente verificamos que no caso da estatística trabalhamos com parâmetros, estimados a partir dos dados amostrais. Interessa-nos a hipótese de investigação, todavia a estatística não testa a hipótese de investigação, mas a hipótese nula. A hipótese nula é uma afirmação sobre um ou mais valores de parâmetros populacionais. Como verificou ela deve conter sempre a condição de igualdade. É sobre ela que recai o teste estatístico. Na realidade, a hipótese nula postula que as diferenças, as relações entre as variáveis se devem ao acaso amostral. Simbolicamente a hipótese nula pode aparecer como vemos pelo quadro 11.

Quadro 11 – Símbolos utilizados para representar algumas hipóteses nulas e de investigação

Hipótese nula (H_0):	Hipótese de investigação (H_1):
$\mu_1=\mu_2$	$\mu_1\neq\mu_2$
$\mu_1\leq\mu_2$	$\mu_1>\mu_2$
$\mu_1\geq\mu_2$	$\mu_1<\mu_2$
$\rho=0$	$\rho\neq0$
$\rho\leq0$	$\rho>0$
$\rho\geq0$	$\rho<0$

[16] Nem sempre isto acontece.

Como pode verificar, os exemplos estão relacionados com a diferença entre duas médias, ou então com a correlação, que simbolicamente, quando se refere ao parâmetro populacional representa-se por rho (ρ) em vez da letra *r*, o mesmo se passa com a média enquanto parâmetro populacional (μ).

Como pode notar, todas as hipóteses nulas contêm a condição de igualdade, no entanto algumas hipóteses de investigação (ou alternativas) estão direccionadas usando para isso o sinal > ou < (atribuindo-lhe direcção), como é o exemplo $\mu_1>\mu_2$ ou então $\rho>0$. Dois exemplos hipotéticos para essas hipóteses seriam por exemplo: "o stress profissional é superior nos enfermeiros que trabalham em serviço de urgências" ($\mu_1>\mu_2$), ou então, "existe relação positiva entre a idade e a incapacidade funcional dos doentes submetidos a prótese da anca..." ($\rho>0$).

Podemos então falar de outro conceito, o de direccionalidade das hipóteses. Vejamos com base no exemplo anterior, hipóteses direccionadas e não direccionadas no quadro 12.

Quadro 12 – Símbolos utilizados para representar algumas hipóteses de investigação e formulação da hipótese respectiva

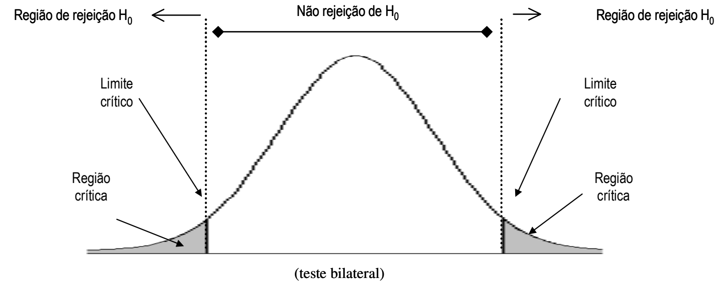
Simbolicamente	Hipótese formulada
$\mu_1\neq\mu_2$	"o stress profissional é diferente..."
$\mu_1>\mu_2$	"o stress profissional é superior ...."
$\rho\neq 0$	"existe relação entre..."
$\rho>0$	"existe relação positiva entre..."

De que forma a direcção de uma hipótese condiciona o teste? De uma forma simples. Os testes podem ser bilaterais e unilaterais (consoante a direccionalidade[17] das hipóteses) ou seja a região de rejeição da hipótese nula pode estar em apenas uma das caudas (direita ou esquerda) da distribuição ou então nas duas caudas.

Para facilitar a apreensão destes conceitos vamos apresentar os três esquemas de tomada de decisão num teste de hipóteses (figuras 1 e 2). O esquema envolve uma região de **não rejeição da H_0**, uma região de **rejeição da H_0** e por fim um limite crítico que corresponde a um valor teórico tabelado.

[17] A direccionalidade não é possível de testar em todos os testes de hipóteses.

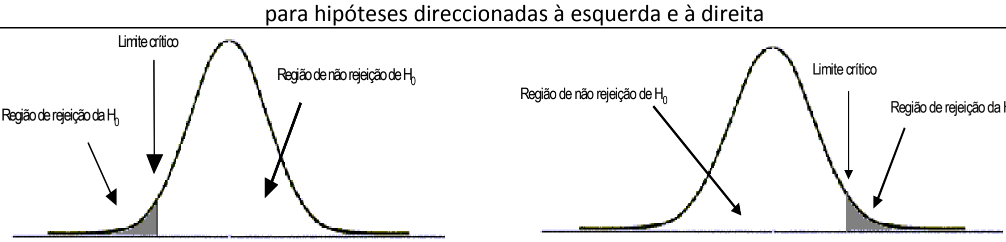
Figura 1 - Esquema de tomada de decisão nos testes de hipóteses
para hipótese não direccionada (bilateral)



O esquema mostra a região de rejeição de H_0 e de não rejeição da H_0. O limite crítico é um valor teórico de uma distribuição. É do confronto desse valor com o do nosso teste que tomamos a decisão. Se o valor do teste for superior ao valor teórico estabelecido e cair na região de rejeição da H_0 deveremos rejeitar a H_0, caso contrário não rejeitamos H_0.

Existem outras condicionantes, como veremos adiante. No caso do esquema a região de rejeição está de ambos os lados da distribuição (corresponde às áreas sob a curva que estão acizentadas), mas pode efectivamente estar apenas de um dos lados, basta para isso que as hipóteses sejam direccionadas, como se vê seguidamente. Se está só de um dos lados diz-se que o teste é unilateral à esquerda ou à direita.

Figura 2 - Esquema de tomada de decisão nos testes de hipóteses
para hipóteses direccionadas à esquerda e à direita



Recapitulando, com a realização do teste obtemos um resultado (estatística do teste) que posteriormente comparamos com os valores da distribuição teórica. O *limite crítico* é o limite que utilizamos como valor critério para rejeitar ou não rejeitar a hipótese nula (H_0). Assim e se o nosso valor cai na região de não rejeição da H_0, não rejeitamos a hipótese nula. Os valores do limite crítico estão tabelados para cada tipo de distribuição, só temos que os comparar com os valores obtidos no teste e tomar a decisão.

Para verificar se a região de rejeição da H_0 é em ambas as caudas ou só apenas numa delas, implica que saibamos ler as hipóteses e colocá-las de forma simbólica. Isso obriga a que o investigador verifique se são ou não direccionadas e em que sentido é afirmada a relação ou a diferença. A simbologia das hipóteses é simples de colocar.

Num relatório ou num projecto de investigação devem ser colocadas as hipóteses de investigação (também podem ter a designação de hipóteses alternativas ou contra hipótese) isto porque como a hipótese nula ou estatística é uma *afirmação matemática* por exemplo entre parâmetros, parece não fazer sentido colocar a hipótese que afirma a relação matemática, mas a hipótese que é uma afirmação da relação entre variáveis ou relativamente à previsão de uma relação esperada entre as variáveis.

5.3. Outros condicionalismos e especificidades presentes nos testes de hipóteses

Um outro condicionalismo dos testes de hipóteses prende-se com o número e tipo de grupos que estão a ser testados. Podemos trabalhar com uma só amostra ou com quantas tantas o desenho de investigação obrigar. Os grupos ou amostras podem ser considerados **independentes** ou **emparelhados**:

- **Independentes:** quando são sujeitos diferentes medidos na mesma variável dependente. Por exemplo no caso do stress dos enfermeiros são dois grupos independentes, ou seja, comparamos sujeitos diferentes.

- **Emparelhados:** quando temos os sujeitos medidos em diferentes situações ou momentos. Por exemplo quando temos uma hipótese de investigação que afirma que a frequência de um programa dirigido de educação para a saúde em fumadores diminui o consumo de tabaco, estamos a comparar o consumo de tabaco antes de iniciar o programa e depois da frequência do programa. Neste caso, são dois momentos (antes-depois) mas poderia ter três ou mais momentos. Para compreender melhor estes aspectos poderá encontrar referências essenciais em manuais de investigação, nomeadamente sobre desenhos ou planos de investigação.

Um outro condicionalismo dos testes de hipóteses prende-se com o nível de medida das variáveis. Isso obriga a que se escolham testes paramétricos (baseados em parâmetros) ou testes não paramétricos (algumas vezes aparecem designados também como *métodos de distribuição livre*, mas isso carece de uma explicação que incluiremos mais adiante). Um dos pressupostos que os testes paramétricos exigem é que a variável dependente seja no mínimo medida ao nível intervalar. Essa informação é útil para que se possa usar o quadro de testes de hipóteses que iremos apresentar mais adiante.

Não iremos fazer a distinção entre ao facto das hipóteses serem de relação[18] ou de diferença, isso prende-se com o propósito do teste, por isso o nosso quadro implica sempre que aconselhável, ao recurso e cálculo dos diversos índices que medem a força da relação entre as variáveis, para todos os testes apresentados, sem excepção.

Entretanto estamos pois aptos para realizar os primeiros passos, para isso deverá utilizar este esquema de passos.

OS PASSOS NUM TESTE DE HIPÓTESES:

1. PASSO: determinar o teste estatístico adequado.
Para se escolher o teste adequado devemos olhar para a natureza das hipóteses (se afirmam diferenças ou relações), o n.º e tipo de grupos que estão a ser testados, o nível de medida da variável dependente (e independente), no caso dos testes paramétricos devemos olhar para os pressupostos que eles exigem. **Consulte o guia de escolha dos testes que apresentamos seguidamente.**

2. PASSO: fixar o nível de significância do teste.
O nível de significância do teste deve ser estabelecido *à priori*. Normalmente utilizamos 0,05, mas pode ser utilizado 0,01. Não se esqueça que deve manter o critério de rejeição constante até ao fim da realização do teste.

3. PASSO: Determinar se o teste é unicaudal ou bicaudal.
Deve olhar atentamente a hipótese formulada. Se o teste apresentar direcção, deve ser unicaudal ou unilateral (para uma das caudas).

4. PASSO: calcular o valor do teste estatístico.
Agora resta-lhe aplicar o respectivo teste com a fórmula adequada. Efectivamente quando o teste é realizado no computador não existe esse problema, mas deve saber qual o resultado do teste que se deve ler, pois para o mesmo teste o computador pode utilizar fórmulas distintas, como é o caso do teste t para grupos independentes.

5. PASSO: determinar os graus de liberdade (GL).
Se o teste for realizado à mão, cada fórmula apresenta o modo adequado de determinar os graus de liberdade para consulta dos valores das distribuições teóricas. Este problema não se coloca igualmente quando recorremos ao SPSS, pois o output gerado pelo programa calcula os graus de liberdade em cada teste.

6. PASSO: comparar os valores teóricos com os valores do teste (tomar a decisão).
Se o teste for realizado manualmente, deve comparar os valores do teste com os valores teóricos da distribuição considerada e tomar uma decisão face à hipótese sob teste, a H0. Se utilizar o computador deverá ler os resultados na coluna da significância (*sig.*) do teste.

Relembre o exemplo da hipótese do stress nos enfermeiros. A hipótese de investigação era de que é "o nível de stress profissional é diferente consoante o local de trabalho dos enfermeiros". Foram comparados dois grupos de enfermeiros (S.U. e C.S.P.), trata-se neste caso de dois grupos independentes. O nível de medida da variável dependente pode considerar-se intervalar, relembre que se trata de um *score* de uma escala. Se utilizarmos o quadro seguinte (quadro 13) verificamos que com dois grupos independentes e o nível de

[18] Genericamente todas podem ser ou são de relação, no entanto é útil distingui-las para facilitar a compreensão.

medida da variável dependente medida no mínimo a um nível intervalar o teste adequado é o teste *t* para grupos independentes.

No entanto, a utilização deste teste está condicionada pelo cumprimento de um conjunto de pressupostos que iremos averiguar e apresentar sempre que falarmos de cada um dos testes.

Quadro 13 - Guia para a escolha do teste adequado

Hipóteses de diferença entre grupos (ou momentos)			
Nível de medição da variável dependente	N.º de grupos da variável independente	Tipo de amostras	Teste adequado
Mínimo intervalar	2	Independentes	Teste t (independentes)
Mínimo intervalar	2	Dependentes	Teste t (emparelhados)
Mínimo intervalar	≥ 2	Independentes	ANOVA
Mínimo intervalar	≥ 2	Dependentes	ANOVA medidas repetidas
Ordinal	2	Independentes	U de Mann-Witnhey
Ordinal	2	Dependentes	Postos por sinais de Wilcoxon
Ordinal	≥ 3	Independentes	H de Kruskal-Wallis
Ordinal	≥ 3	Dependentes	Friedman
Nominal (dicotómica)	2	Independentes	Qui-quadrado (Fisher)
Nominal (dicotómica)	2	Dependentes	McNemar
Nominal (dicotómica)	≥ 3	Independentes	Qui-quadrado
Nominal (3 ou mais)	≥ 3	Independentes	Q de Chocran
Hipóteses de relação entre variáveis			
Nível de medição da variável dependente	Medida estatística		Teste adequado
Intervalar ou racional	Coeficiente de correlação de Pearson (r)		t de significância do r
Ordinal	Coeficiente de correlação de Spearman (rs)		t de significância do r
Intervalar ou racional	Coeficiente de correlação de Kendall (τ)		z de significância do τ
Nominal	Coeficiente Phi (φ), V de Cramer		Qui quadrado ($\chi 2$)

Antes de iniciar a apresentação dos testes, vamos deter-nos por momentos num conceito muito utilizado em estatística e que é o de **significância estatística**. Muitas vezes esperamos impacientemente que exista relação entre as variáveis ou então que se encontrem diferenças estatisticamente significativas[19]. Em parte esta ânsia pode levar a interpretações erróneas dos resultados dos testes estatísticos.

É costume aparecer asteriscos (ou estrelas) nas tabelas junto a uma coluna que tem habitualmente duas designações possíveis (*Sig.* no programa SPSS ou então *p level* no STATISTICA). E então aparece * como sendo significativo, ** muito significativo e *** altamente significativo. Apesar de não constituir nenhum erro, muito pelo contrário, a significância estatística pode ser interpretada erroneamente, nomeadamente interpretando a significância estatística como *significância clínica*.

Isto porque quando realizamos o teste, comparamos o valor do teste com o valor teórico da distribuição considerada, se o valor do teste cai na região de rejeição da H_0 é costume dizer que essa diferença ou relação é estatisticamente significativa. Na estatística

[19] A significância estatística cria um problema interessante para não dizer grave, especialmente quando não se encontram resultados significativos pois obriga a atribuir um significado ao resultados, interpretá-los e discuti-los. É esse um dos objectivos da Estatística como "ferramenta supletiva" na Investigação em Enfermagem

inferencial, significativo quer dizer que essas diferenças não se devem eventualmente ao acaso para o nível de probabilidade apresentado, por exemplo 0,05.

No nosso exemplo do stress, imagine que encontramos diferenças significativas ao nível de 0,05 (rejeitamos H_0). O que significa este valor? Estamos a afirmar que essas diferenças provavelmente não se devem ao acaso mas a um efeito do local de trabalho no stress profissional dos enfermeiros. De facto os enfermeiros que trabalham em serviço de urgência apresentavam níveis de stress superiores, porque estão mais expostos a situações stressantes. Na realidade, os resultados parecem ter significado, por isso dizemos que as diferenças são estatisticamente significativas, o que não significa que sejam clinicamente relevantes. Se a hipótese nula é uma afirmação falsa, nós podemos afirmar que as diferenças são demasiado grandes para se deverem ao acaso. Temos então *evidência* para afirmar que os resultados não suportam a hipótese nula e que a hipótese alternativa pode ser aceite como uma descrição mais correcta para a diferença encontradas nas duas populações.

Por outro lado, verificamos que muitas vezes quando não se encontram diferenças significativas entre, por exemplo as médias de dois grupos, tal não significa que a hipótese esteja mal formulada ou então que o trabalho foi mal conduzido. Tal pode significar que o resultado obtido se dever ao acaso (flutuações aleatórias e erro amostral). Um dos problemas, que se levanta na discussão e interpretação de resultados que não apresentam diferenças significativas é a ambiguidade criada por este tipo de resultados. Estamos vocacionalmente situados para discutir o óbvio, a revisão da literatura vem nesse sentido muitas vezes, e quando é necessário interpretar um resultado deste tipo empeçamos e ficamos sem saber o que fazer. Mas os resultados são tão importantes encontrando como não encontrando diferenças, possivelmente mais no último caso, porque põem em causa o senso comum.

Ainda sobre os resultados dos testes e quando nos referimos aos asteriscos, convém salientar que a «sig.» ou «p level» é a abordagem feita pelos softwares estatísticos em que o valor encontrado corresponde "à probabilidade de obter um valor da estatística amostral de teste no mínimo tão extremo como o resultado dos dados amostrais, na suposição de a hipótese nula ser verdadeira" (Triola, 1999, pág. 180). Ou seja, nível de significância (p level ou sig.) corresponde à probabilidade da H_0 ser rejeitada quando na realidade é verdadeira. Quanto menor o nível de significância (p) associado à distribuição, menor o risco associado à rejeição de H_0, optando pela alternativa (H_1).

6. Testes de hipóteses paramétricos e não paramétricos

세 미만, 30-40세, 40-50세, 50세 이상으로 구분하여 분석에 활용하였고, 학력은 중졸 이하, 고교 재학 및 졸업, 대학 재학 및 졸업 이상으로 구분하였다. 경제력은 조사 대상자의 판단에 따라 상, 중, 하로 나누어 분석하였다.

본 설문도구의 Cronbach's alpha 값은 0.731로 내적 신뢰성이 있다고 볼 수 있다.

3. 자료 분석

연구 모형을 검증하기 위해 다음과 같이 자료 분석을 실시하였다. 먼저, 각 설문문

Variable	Frequency (%)	Familiarity		Knowledge		Prejudice		Social distance	
		Mean	p-value	Mean	p-value	Mean	p-value	Mean	p-value
Total	1,040 (100.0%)	0.40		0.000		0.000		6.53	
Gender									
Female	522 (50.2%)	0.37	0.088	-0.030	0.334	0.029	0.354	6.92	<0.001
Male	518 (49.8%)	0.43		0.030		-0.029		6.13	
Age(yrs)									
< 30	297 (28.6%)	0.32	0.047	0.177	<0.001	-0.526	<0.001	6.04	0.001
30 - 39	241 (23.2%)	0.41		0.164		-0.206		6.40	
40 - 49	231 (22.2%)	0.43		0.070		0.233		6.72	
≥ 50	271 (26.0%)	0.44		-0.400		0.560		7.00	
Education									
≤Middle school	182 (17.5%)	0.37	0.070	-0.363	<0.001	0.462	<0.001	6.60	0.897
High school	388 (37.3%)	0.36		0.086		-0.004		6.48	
≥College	470 (45.2%)	0.44		0.212		-0.176		6.54	
Economic Level									
Low	309 (29.7%)	0.39	0.625	-0.147	0.006	0.128	0.028	6.55	0.976
Middle	570 (54.8%)	0.39		0.044		-0.054		6.52	
High	161 (15.5%)	0.44		0.126		-0.053		6.50	

Nas últimas duas décadas assistimos a um desenvolvimento sem precedentes ao nível da investigação em Enfermagem e isso, reflectiu-se no aparecimento de um vasto conjunto de trabalhos empíricos relativos a temáticas como *sofrimento na doença*, *ansiedade perante a morte e o morrer*, *qualidade de vida*, *crenças e atitudes acerca dos doentes e doenças mentai*s, *satisfação no trabalho, conforto* entre muitos outros. Na sua maioria seguem abordagens quantitativas e recorrem às ferramentas estatísticas de tomada de decisão, especificamente **os testes de hipóteses**.

Repare no exemplo que aparece de seguida (Loureiro *et al*., 2008, p. 39). Trata-se de um estudo relacionado com as crenças e atitudes acerca das doenças e dos doentes mentais. Os autores procuravam através de medidas e testes estatísticos (Análise de Regressão Múltipla), verificar quais as variáveis (crenças por exemplo) que são preditoras de uma atitude de *separatismo hostil e estigmatizante* dos doentes (variável critério). A hipótese de investigação era de que «*as crenças acerca dos doentes e doenças mentais, as causas quanto à etiologia e as variáveis de domínio sócio-demográfico e contextual permitem prever a atitude de separatismo hostil e estigmatizante dos doentes*».

QUADRO 4 – Resultados da ARM (método stepwise): Preditores: crenças, causas quanto à etiologia e restantes variáveis sócio-demográficas e contextual. Critério: Separatismo hostil e estigmatizante (N = 834)

Modelos	r^2 ajustado	r^2_{Change}
1.º - a	.241	.242
2.º - a+b	.270	.029
3.º - a+b+c	.290	.021
4.º - a+b+c+d	.309	.020
5.º - a+b+c+d+e	.317	.008
6.º - a+b+c+d+e+f	.323	.007
Preditores retidos no modelo 6		
a. Perigosidade		.304***
b. Incurabilidade		.192***
c. Idade		.158***
d. Responsabilidade individual		.128***
e. Doença como condição médica		-.096**
f. Causas associadas ao envelhecimento		.091**

**p<.01; ***p<.001

> Relativamente à análise de regressão efectuada (quadro 4), tendo como critério a atitude de *separatismo hostil e estigmatizante*, verifica-se que o modelo 6 explica 34.2% da variância, no entanto, até ao modelo 4, inclusive, é explicada 30.0% da variância. Os preditores com maior "peso" na atitude de *separatismo hostil e estigmatizante*, são a *crença na perigosidade* (β=.337), a crença na *incurabilidade* (β=.184), *idade* (β=.114), e a *crença na responsabilidade individual* (β=.144). Quer isto dizer que uma atitude que remete para a segregação dos doentes, confinando-os ao espaço psiquiátrico, é mais provável naqueles indivíduos que acreditam que os doentes são perigosos, incuráveis, com idade mais avançada e que tendem a ver o doente como o único responsável pelo seu estado de saúde/doença (Loureiro *et al*., 2008, p. 38).

Trata-se apenas de um dos muitos exemplos estudados e publicados nas mais variadas revistas científicas existentes. A referência à Análise de Regressão Múltipla, utilizada para testar a hipótese enunciada, serve de pretexto para introduzir o primeiro conjunto de testes designados de **testes paramétricos**.

Os testes **paramétricos**, sem dúvida os mais divulgados e utilizados, referem-se à utilização de parâmetros, como, por exemplo, a média (μ) ou o desvio padrão (σ), e partem do princípio que conhecemos a forma da distribuição da qual provêm as observações que fazemos. Por isso, muitas vezes assumimos que os dados provêm de uma distribuição normal e realizamos os testes a partir das estatísticas amostrais, que utilizamos como estimadores dos parâmetros populacionais.

Estes testes, considerados mais potentes e robustos têm no entanto alguns pressupostos[20] que devem estar cumpridos previamente à sua utilização. Exigem por exemplo que a variável dependente seja medida a um nível intervalar ou racional (costuma dizer-se nível intervalar ou superior), mas não se ficam por aqui. Existem consoante a tipologia de alguns testes, pressupostos que necessitam de ser avaliados (verificados) para que seja possível a sua utilização, isto é, para que possamos assumir como válidos dos seus resultados e assim podermos retirar conclusões.

No entanto existe outro conjunto de testes designados de **não paramétricos** ou como são referidos frequentemente, **testes de distribuição livre**. Este uso indistinto na maioria dos manuais esquece muitas vezes que os termos não paramétrico e de distribuição livre não são sinónimos, tal como refere Bradley (1968):

> The terms nonparametric and distribution-free are not synonymous (…) Popular usage, however, has equated the terms (…) Roughly speaking, a nonparametric test is test one which makes no hypothesis about the value of a parameter in a statistical density function,

[20] Verifica-se em alguns manuais a utilização indiscriminada, como tradução de *assumptions*, quer os termos assunções, suposições ou mesmo pressupostos. De facto, se assunções e pressupostos forem tomados por suposição, uma proposição admitida como verdadeira para efeitos dedutivos, podemos usar uma ou outra, dado que pressuposto se pode entender como um supor antecipadamente, e assunção como o acto de assumir. Assim a nossa opção vai pela utilização do termo pressupostos.

> whereas a distribution-free test is one which makes no assumptions about the precise form of the sampled population. (p. 15)

Também a designação «não paramétrico» não é *precisa* (Mosteller & Rourke, 1993), pois muitas estatísticas paramétricas lidam como parâmetros, como, por exemplo, a mediana da distribuição. Apesar destes testes serem com regularidade colocados em 2.º plano, excepção para o teste do qui-quadrado (χ^2), são um conjunto de ferramentas adequadas a muitas situações e constrangimentos de quem desenvolve uma investigação. A este respeito, Syegel (1979, p. 34-35) refere que:

> Em muitas das provas estatísticas não paramétricas (...) os dados são transformados de números para postos ou mesmo para sinais. Tais métodos podem suscitar a crítica de não utilizarem toda a informação contida na amostra, ou de desperdiçarem informações. A resposta a tais objecções encontra-se nas perguntas seguintes: (a) dos métodos disponíveis – paramétricos e não paramétricos – qual utiliza mais adequadamente a informação contida na amostra? (b) é importante que as conclusões da pesquisa tenham aplicação geral, ao invés de se aplicarem apenas a populações com distribuição normal?

É de salientar que este tipo de testes, apesar de não estarem condicionados por um conjunto tão rígido de pressupostos como os paramétricos, admitem, por exemplo que as distribuições tenham forma idêntica e a distribuição seja simétrica, bem como que a distribuição básica seja contínua.

Os testes não paramétricos apresentam algumas *vantagens* e *desvantagens* relativamente aos seus equivalentes paramétricos. Relativamente às *vantagens*, elas passam por serem testes mais fáceis de entender, permitindo mesmo testar hipóteses em que não estão envolvidos parâmetros e quando as variáveis não estão normalmente distribuídas. Relativamente às *desvantagens*, elas situam-se no facto dos testes não paramétricos tenderem a usar menos informação e não serem tão potentes como os paramétricos, por isso também são menos sensíveis.

6.1. Testes paramétricos e seus pressupostos

Existem, consoante a tipologia de alguns testes, pressupostos que necessitam de ser verificados para que seja possível a sua utilização, isto é, para que possamos assumir como válidos dos seus resultados e assim podermos retirar conclusões. Referenciamos aqui os que são comuns nos testes apresentados seguidamente.

a) Normalidade das distribuições

O pressuposto mais comummente referenciado para a utilização dos testes paramétricos é o da **distribuição normal da variável na população**[21]**.** (H_0: A variável x segue distribuição normal).

A verificação da normalidade da distribuição pode ser feita no programa SPSS através de dois testes, o teste de Kolmogorov-Smirnov com correcção de Lilliefors quando o n>50 e o teste de Shapiro-Wilk quando n≤50.

Quando este pressuposto não se encontra cumprido, alguns caminhos são possíveis:

i. Tentar uma transformação da variável em questão para a normalidade (consultar Petrie & Sabin, 2000);

ii. Optar pelo teste não paramétrico equivalente;

iii. No caso de amostras n≥ 30, é comum alguns autores afirmarem que podemos assumir a normalidade. Esta decisão é em nosso entender discutível. Convém salientar que na investigação em Enfermagem, que envolve muitas vezes ao investigação em contexto saúde/doença, parece à partida complexo seleccionar amostras que não sejam representativas, mas também aleatórias da população que se pretende estudar.

Como regra básica devem ser usadas amostras colhidas de forma independente, e aleatória sempre que tal for possível, caso isso não aconteça, deve-se clarificar muito bem qual a população para a qual se pode generalizar os resultados (Polit, 1996).

Assim sendo, aconselhamos que não se olhe para o *n* de 30 como *tamanho sagrado*, olhe sim para a natureza da temática sob estudo e para o desenho de investigação.

b) Homogeneidade da variância

Exige-se também que exista **homogeneidade da variância** ($H_0 : \sigma_1^2 = \sigma_2^2$). A sua verificação é realizada através do teste de Levene[22]. Este pressuposto não é restritivo à utilização dos testes paramétricos, ou porque se pode realizar o teste usando uma fórmula alternativa, ou porque pode ser assumido quando se cumprem certas condições, por exemplo, amostras grandes em dimensão, o que será apresentado adiante.

b) O nível de medida

Um último pressuposto prende-se com a natureza da variável «dependente». Esta deve ser medida a um nível intervalar ou racional (dependendo dos testes). Como compreende, este pressuposto exige o conhecimento dos níveis de medida.

[21] Este pressuposto é discutível no caso das grandes amostras.
[22] Tanto o teste *F* de homogeneidade da variância como o teste de Bartlett são sensíveis aos casos em que a distribuição não segue uma curva de Gauss (Markowski & Markowski, 1990). Levene (1960) propôs um teste de homogeneidade da variância que é robusto mesmo quando as distribuições não seguem distribuição normal.

6.2. Teste *t* para a média

Dimensão Burnout:	n	$\overline{x}$	s	*t*	p
Despersonalização	571	5.81	5.34	13.07	.000

Fonte: Patrick & Larry (2007)
Hipótese: H_1: o nível médio de burnout dos enfermeiros na dimensão *despersonalização* é diferente dos valores normativos da população.

O primeiro teste que é apresentado trata-se do teste *t* para uma amostra, isto é, inferência para a média de uma população. Tal como se referiu anteriormente, ao nível dos pressupostos, a variável deve seguir distribuição normal na população, o nível de medida da variável deve ser intervalar ou racional. A amostra também deverá aleatória.

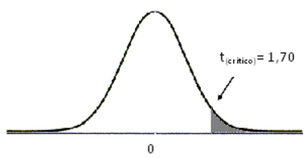
A estatística *t de* uma amostra é calculada pela fórmula: $t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$ e tem distribuição *t* com n-1 graus de liberdade.

Imagine que se queria testar a hipótese (utilizando um erro tipo I de 0,05) de que «o teor médio de alumínio no sangue dos doentes hemodializados há mais de 1 ano no centro de hemodiálise da Rocha é superior a 50 µg/dl". O estudo foi realizado numa amostra aleatória de 35 desses doentes e o valor médio observado foi de 52,19 µg/dl e o desvio padrão de 8,23 µg/dl. Simbolicamente teremos:

H_0:µ=50µg/dl H_1:µ>50µg/dl

Como pode verificar, a hipótese é direccionada à direita, conforme a figura 3. Antes mesmo de calcular a estatística do teste, vamos determinar o valor do *t* teórico ou *t* da tabela que será o valor do limite crítico. Como o alpha é de 0,05, unilateral (unicaudal) com 34 graus de liberdade (gl= n-1), o valor do *t* teórico será 1,70, conforme a tabela de distribuição *t*

Figura 3 - Esquema de tomada de decisão para teste com região de rejeição à direita



Não esqueça que na tabela deverá sempre utilizar o valor dos graus de liberdade imediatamente anteriores ao valor dado pela expressão n-1. Neste caso deverá utilizar gl=30.

Tabela Distribuição *t* (excerto)

	α para teste unilateral			
	0,05	0,025	0,01	0,005
	α para teste bilateral			
Graus de Liberdade	0,10	0,05	0.02	0,01
...	....	....	....	....
29	1.70	2.05	2.46	2.76
30	**1.70**	2.04	2.46	2.75
40	1.68	2.02	2.42	2.70

O passo seguinte consiste no cálculo do valor do *t* do teste. Se o valor da estatística do teste *t* for superior ao valor tabelado (1,70), poderemos rejeitar a hipótese nula.

$$t = \frac{52{,}19 - 50{,}00}{8{,}23/\sqrt{35}} = 1{,}574$$

Como pode verificar o valor do *t* do teste ou *t* observado é inferior ao *t* crítico ou da tabela, caindo por isso na região de não rejeição da H_0, ou seja não se rejeita a hipótese nula. Embora a média de alumínio no sangue na amostra seja de 52,19 µg/dl, este valor não constitui prova empírica suficiente, admitindo um erro tipo I máximo de 5%, de que na população de doentes hemodializados há mais de 1 ano no centro de hemodiálise da Rocha é superior a 50 µg/dl. Simbolicamente teremos ($t_{(34)}$=1,574; p>0,05).

Repare que neste exemplo assumimos o pressuposto da distribuição normal do teor de alumínio no sangue da população de referência recorrendo ao teste de Shapiro-Wilk, pois o n≤50 e obteve-se uma valor p>0,05 (será mostrado adiante). É de salientar que se utilizou a valor do desvio padrão partindo do pressupondo que ele foi estimado ($\hat{s}$), no entanto se aquando do seu cálculo utilizar n-1 no denominador não necessita de o estimar posteriormente.

6.2.1. Cálculo do teste *t* na máquina de calcular

Instruções para cálculo do teste *t* para uma média	
TEXAS	CASIO
1. Pressione o botão **STAT** e depois mude o cursor para a opção **TESTS**; 2. Seleccione a opção **2: T- TEST…** seguido de **ENTER;** 3. Seleccione **Stats** pressionando a tecla **ENTER.** 4. Introduza os valores: **µ0**:**50** $\overline{X}$:**52.19** **Sx -1:8,23** **n: 35** 5. Introduza a hipótese de investigação simbolicamente [**µ: > µ0**] 6. **Calculate** e **ENTER**.	1. Pressione o botão **MENU** e depois seleccione a opção **STAT**, **EXE**; 2. Seleccione a opção **TEST** (tecla F2); 3. Seleccione a opção **t** (tecla F2); 4. Seleccione a opção **1-S** (tecla F1) 5. Na opção **Data**, seleccione a opção **var** (F2); 6. Introduza a hipótese de investigação simbolicamente [**µ: > µ0**] 7. Introduza os valores: **µ0:50** $\overline{X}$:**52.19** **xσn-1:8,23** **n: 35** 8. **EXECUTE** seguido de **EXE**.

Output (Resultados): µ > 50 t = 1,574 p =0,07 $\overline{X}$ =52.19 s = 8,23 n = 35

6.2.2. O cálculo do teste no SPSS

Utilize a base de dados (**t_uma_amostra_opinião.sav**) respeitante aos valores relativos às opiniões das pessoas relativas à perigosidade dos doentes e doenças mentais. Trata-se de uma amostra de n=216 indivíduos com idade ≥ 18 anos, utentes de um centro de Saúde do concelho de Cantanhede.

Começamos por proceder à verificação do pressuposto da **distribuição normal da variável na população** recorrendo ao SPSS. Este software disponibiliza dois testes consoante o tamanho da amostra.

Se $n \leq 50 \rightarrow$ Teste de Shapiro-Wilk (S-W);

Se $n>50 \rightarrow$ Teste de Kolmogorov-Smirnov com correcção de significância de Lilliefors;

Dado que o tamanho da amostra são 216 adultos utentes de um centro de saúde, o teste a utilizar será o teste de Kolmogorov-Smirnov (K-S). Também deveremos escolher previamente o erro tipo ou nível de significância como critério para rejeitar ou não rejeitar a H_0 (H_0: a variável x segue distribuição normal na população). Neste caso vamos utilizar um erro tipo I ou alpha de 0,05.

Os comandos para a realização do teste são apresentados no quadro 14.

Quadro 14 - Comandos para realização dos testes de normalidade (K-S e S-W)

Comando do SPSS (acção):	Comentário:
1. Seleccione do menu as opções: **Analyze** → **Descriptive Statistics** → **Explore**	Acede à caixa de diálogo do Explore que permite realizar os testes de normalidade.
2. Na caixa de diálogo Explore, seleccione a variável *Perigosidade* (com um clique sobre a variável) e depois com o botão ▶ coloque-a na **Dependent List:**	O **Dependent List:** é o espaço onde se colocam as variáveis que procuramos testar.
3. Dê um clique sobre o botão **Plots…**	Acede a nova caixa de diálogo, **Explore Plots**
4. Na caixa **Explore Plots** seleccione com um clique sobre o quadrado a opção **Normality plots with tests** [☒]	Selecciona automaticamente os testes de normalidade, assim como os gráficos «Normal Q-Q Plot» e «Detrended Normal Q-Q Plot».
5. Clique em **Continue** e posteriormente no botão **OK**	Vai gerar um output de resultados dos testes

No quadro seguinte estão apresentados os resultados dos testes, nomeadamente Kolmogorov-Smirnov (K-S) e Shapiro-Wilk (S-W). Dado que o n>50, deveremos ler os resultados do K-S. Para a conclusão basta olhar o valor da significância do teste. Como o valor do alpha definido *à priori* pelo investigador foi de α=0,05 e o valor da significância associada ao teste é de p=,200, valor superior ao α, significa que a variável em questão segue distribuição normal na população, pelo que podemos prosseguir para a realização do teste *t* (One-Sample Test).

Quadro 15 – Resultados dos testes de normalidade

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	**Sig.**	Statistic	df	Sig.
Perigosidade	,055	216	**,200**[*]	,993	216	,407

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

A hipótese colocada a teste refere que «as pessoas manifestam, em média, uma opinião de que os doentes mentais são perigosos». Em termos simbólicos, as hipóteses serão:

H_0:μ= 3,50 pontos H_1:μ> 3,50 pontos

A utilização do valor 3,50 pontos corresponde ao *score* considerado «limite» para que a opinião seja favorável à discriminação, pelo facto dos doentes serem considerados perigosos e imprevisiveis. Os comandos para a realização do teste são apresentados no quadro 16.

Quadro 16 – Comandos para realização do teste *t* para a média de uma amostra

Comando do SPSS (acção):	**Comentário:**
1. Seleccione do menu as opções: **Analyze** $\rightarrow$ **Compare Means** $\rightarrow$ **One-Sample T Test…**	Permite-lhe aceder à caixa de diálogo do teste
2. Na caixa de diálogo seleccione a variável *Perigosidade* (com um clique sobre a variável) e depois com o botão ▶ coloque-a na lista do **Test variable(s):**	O **Test variable(s)** é o espaço onde se colocam a variáveis a testar.
3. Na opção **Test Value**: coloque o valor de 3,50	Esta opção serve para colocar o valor da média alegada (μ0).
6. Clique no botão **<u>OK</u>**	Vai gerar um output de resultados do teste

Os resultados são apresentados nos quadros seguintes (17 e 18). O primeiro quadro diz respeito às estatísticas resumo e como pode verificar, a $\overline{x}$= 3,69 pontos, s=0,99 pontos e o erro padrão da média (se) é de 0,07.

Quadro 17 – Estatísticas resumo relativas à opinião acerca das perigosidade dos doentes e doenças

	n	Mean	s	se
Perigosidade	216	3,69	,99	,07

Relativamente aos resultados do teste *t*, podemos observar pela leitura do quadro 18 que o valor do $t_{(teste)}$= 2,813, com 215 graus de liberdade e uma significância associada de 0,005. No entanto o teste foi calculado como se a hipótese não fosse direccionada (2-tailed), então deveremos dividir o valor da significância por 2, ou seja, 0,0025. A conclusão do teste leva à rejeição da hipótese nula.

Poderemos então concluir que o valor ($\overline{x}$ = 3,69) constitui prova empírica suficiente, admitindo um erro tipo I máximo de 5%, de que na população existe a opinião de que os doentes e as doenças mentais acarretam perigosidade ($t_{(215)}$=2,813; p<0,05).

Quadro 18 – resultados da aplicação do teste *t* para uma amostra (One-Sample T Test…)

	Test Value = 3.5					
					95% Confidence Interval of the Difference[23]	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Perigosidade	2,813	215	,005	,19	,06	,32

6.3. Teste *t* para amostras independentes

	Grupos:	n	$\overline{x}$	s	*t*	p
EAPM (total)	Enfermeiros Técnicos de diagnóstico e terapêutica	482 113	51,58 32,98	15,76 11,13	14.655	.000

Fonte: Loureiro (2004)
Hipótese em estudo (H_1) «o nível médio de ansiedade perante a morte é diferente dos Enfermeiros para os Técnicos de diagnóstico e terapêutica.

Em muitas situações da investigação, construímos hipóteses em que testamos se dois grupos (cada um composto por sujeitos diferentes) diferem ou não relativamente a uma dada média[24] (a variável deve ser medida a um nível intervalar ou racional).

Tomemos como exemplo um estudo em que se pretendeu comparar o nível médio de triglicerídios no plasma (em mg/dl) da população adulta rural e da população adulta urbana. A hipótese de investigação (H_1) refere que «o valor de triglicerídios no plasma é diferente da população adulta em meio rural, comparativamente com a população adulta do meio urbano». Observe o quadro seguinte.

Quadro 19 – Estatísticas resumo relativas ao nível de triglicerídios no plasma (mg/dl) das populações adultas residentes em meio rural e urbano

Local de residência:	n	$\overline{x}$	*s*
Rural	56	135.34	21.34
Urbano	47	146.21	16.28

Como pode verificar a variável «independente» é composta por dois grupos

[23] Relativamente aos intervalos de confiança (95% Confidence Interval of the Difference), o seu cálculo é dado por $\overline{x} \pm t\frac{s}{\sqrt{n}}$ com gl=n-1. Como sabe o valor de *t* crítico depende do nível de confiança desejado e dos graus de liberdade.

[24] O objectivo do teste *t* é avaliar a diferença entre as médias de dois grupos independentes

(categorias ou níveis da variável), são eles a população adulta residente em meio rural e a população adulta residente em meio urbano. Já a variável dependente (v*alor de triglicerídios no plasma em mg/dl*) cumpre o pressuposto do nível de medida da variável «dependente» (intervalar ou superior).

Se relembrar a simbologia das hipóteses, a proposição que está aqui apresentada é a *hipótese de investigação*, ou seja H_1: $\mu_1 \neq \mu_2$, logo a H_0:$\mu_1 = \mu_2$. Como a hipótese não é direccionada, a região de rejeição da H_0 envolve as duas caudas, ou seja, o teste é bicaudal. Neste caso, deveremos utilizar o teste *t* de comparação de duas médias, designado de teste *t* para grupos independentes (*t test for independent groups*).

Como nível de significância vamos utilizar 0,05 para todos os testes.

Este teste comporta **três pressupostos**, nomeadamente o nível de medida (da variável dependente) que deve ser intervalar ou superior, deve existir homogeneidade da variância[25] e a variável dependente deve seguir distribuição normal na população. Estes dois últimos pressupostos irão ser verificados mais adiante aquando da realização do teste *t* com o SPSS.

Face aos resultados do teste, duas decisões podem ser tomadas em relação à hipótese sob teste:

a) Não rejeitamos H_0, ou seja, assumimos face aos resultados do teste que a diferença do valor de triglicerídios no plasma se deve ao acaso amostral; o facto de se viver em meio rural ou urbano não evidencia uma diferença (estatisticamente significativa) no valor de triglicerídios no plasma;

b) Rejeitamos H_0, ou seja, existe de facto uma evidência (para α=0,05), que a diferença (das médias) no nível de triglicerídios no plasma é estatisticamente significativa. A diferença vai para além daquilo que seria provável esperar pelo acaso. Neste caso costumamos dizer que *há evidência empírica para afirmar*[26] que o valor de triglicerídios no plasma é diferente dos adultos residentes em meio rural para os residentes em meio urbano, ao nível de 0,05. Isto significa, como já tínhamos dito, que a hipótese de investigação se considera uma explicação mais adequada, razão pela qual optamos por ela.

O teste é realizado através dos cálculos indicados pela fórmula:

[25] Este pressuposto não é restritivo à utilização do teste, pois, mesmo não existindo homogeneidade da variância pode utilizar-se o teste *t.*

[26] Esta leitura formal não é comum na descrição dos resultados do teste quando estamos a escrever um relatório ou artigo científico, mas disso falaremos adiante.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left[\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\right]\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$$

Logo, substituindo os valores

$$t = \frac{135{,}34 - 146{,}21}{\sqrt{\left[\frac{(56-1)21{,}34^2 + (47-1)16{,}28^2}{56+47-2}\right]\left[\frac{1}{56}+\frac{1}{47}\right]}} = -2{,}8616$$

A fórmula apresentada anteriormente pressupõe a homogeneidade da variância como parte do princípio que os valores das variâncias são estimados.

O valor do teste obtido é designado de $t_{(\text{observado/teste})}$. Face à hipótese apresentada, verificamos que a região de rejeição é bilateral (ambas as caudas da distribuição). Devemos então comparar este valor com o valor do *t* teórico ($t_{(\text{teórico})}$).

Para obter o valor do $t_{(\text{observado})}$, valor que procuramos na tabela seguinte, necessitamos do nível de significância do teste (0,05/2)[27] e do número de graus de liberdade (gl). Assim $gl=n_1+n_2-2$, logo gl=101.

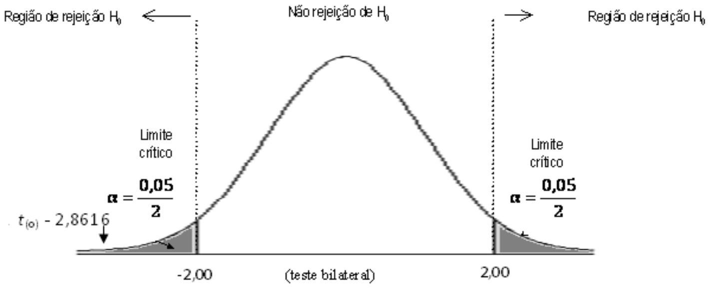
Tabela Distribuição *t* (excerto)

	***α* para teste unilateral**			
	0,05	**0,025**	**0,01**	**0,005**
	***α* para teste bilateral**			
Graus de Liberdade	**0,10**	**0,05**	**0.02**	**0,01**
1	6.31	12.71	31.82	63.66
2	2.92	4.30	6.97	9.93
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
60	1.67	2.00	2.39	2.66
120	1.66	1.98	2.36	2.62
∞	1.65	1.96	2.33	2.58

O valor do $t_{(\text{teórico})}$= 2,00, é o limite crítico para ambas as caudas. Na figura 3 mostra-se a decisão a tomar face aos valores do $t_{(\text{teórico})}$ e do $t_{(\text{observado})}$.

[27] Repare que na tabela já está apresentado o alpha para o teste unilateral e bilateral. Neste caso deverá ler directamente na coluna α de 0,05 para teste bilateral.

Figura 4 – Esquema de tomada de decisão
para teste de hipótese com região de rejeição bilateral e alpha = 0,05



Como podemos ver pela figura o t_{teste} cai na região de rejeição da H_0 ($t_{teste} > t_{observado}$), ou seja, rejeita-se H_0, o que significa que há evidência empírica para afirmar que o valor o valor de triglicerídios no plasma é diferente da população adulta residente em meio rural, comparativamente com a população adulta a residir em meio urbano **($t_{(101)}$= -2,8616; p<0,05).**

A fórmula de cálculo do teste *t* usada anteriormente assume que existe homogeneidade da variância ($\sigma_1^2 = \sigma_2^2$), quando tal facto não acontece o cálculo do teste é dado por outra fórmula, designado *t corrigido** (Sachs, 1984). Esta fórmula tem a vantagem de poder comparar variâncias que não são homogéneas bem como amostras de diferentes tamanhos. O teste é realizado através dos cálculos realizados pela fórmula:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \text{ e em que } gl = \frac{1}{\left[\frac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right] / (n_1 - 1) + \left[\frac{\frac{s_2^2}{n_1}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right]}$$

6.3.1. O cálculo do teste *t* para amostras independentes na máquina

Deverá utilizar os valores apresentados no quadro 19.

Instruções para cálculo do teste *t* as médias de dois grupos independentes	
TEXAS	CASIO
1. Pressione o botão **STAT** e depois mude o cursor para a opção **TESTS**; 2. Seleccione a opção **4: 2-SampTest…** seguido de **ENTER;** 3. Seleccione **Stats** pressionando a tecla **ENTER.** 4. Introduza os valores de $\overline{X}$, **Sx e n** para cada um dos grupos. 5. Introduza a hipótese de investigação simbolicamente [**$\mu_1 \neq \mu_2$**] 6. Na opção **Pooled** seleccione **Yes**; 6. **Calculate** e **ENTER**.	1. Pressione o botão **MENU** e depois seleccione a opção **STAT**, **EXE**; 2. Seleccione a opção **TEST** (tecla F2); 3. Seleccione a opção **t** (tecla F2); 4. Seleccione a opção **2-S** (tecla F2) 5. Na opção **Data**, seleccione a opção **var** (F2); 6. Introduza a hipótese de investigação simbolicamente [**$\mu_1 \neq \mu_2$**] e os valores da $\overline{X}$, **desvio padrão** (x1σn-1) e **n** para cada grupo; 7. Na opção **Pooled** seleccione **On** (tecla F2); 8. **EXECUTE** seguido de **EXE**.
Output (Resultados): $\mu_1 \neq \mu_2$ t =-2.862 p =,005 df =101	

6.3.2. O cálculo do teste *t* para amostras independentes no SPSS

Vamos utilizar o programa SPSS para realizar o teste t. Para todos os testes daqui em diante vamos utilizar como erro tipo I (alpha) admissível o valor de 0,05. O exemplo que utilizamos, diz respeito a um estudo relacionado com o «efeito» da morte e do processo de morrer dos doentes, na ansiedade dos enfermeiros perante a morte.

Foram comparados dois grupos de enfermeiros quanto ao seu nível de ansiedade perante a morte, uns que trabalham em Cuidados de Saúde Primários (C.S.P.) e outros em Serviço de Urgências (S.U.). A hipótese de investigação refere que «o nível de ansiedade perante a morte é diferente consoante o local de trabalho dos enfermeiros». Simbolicamente temos H_0: $\mu_1 = \mu_2$ e H_1: $\mu_1 \neq \mu_2$.

Para testar a hipótese deve utilizar-se a base «das_testet.sav». O 1º passo consiste na verificação do pressuposto da normalidade (variável dependente: «das»), que se faz através dos comandos que aparecem no quadro 20.

Quadro 20 - Comandos para realização dos testes de normalidade (K-S e S-W)

Comando do SPSS (acção):	Comentário:
1. Seleccione do menu as opções: **Analyze** → **Descriptive Statistics** → **Explore**	Acede à caixa de diálogo do Explore que permite realizar os testes de normalidade.
2. Na caixa de diálogo Explore, seleccione a variável *das* (com um clique sobre a variável) e depois com o botão ▶ coloque-a na **Dependent List:**	O **Dependent List:** é o espaço onde se colocam as variáveis que procuramos testar.
3. Dê um clique sobre o botão **Plots…**	Acede a nova caixa de diálogo, **Explore Plots**
4. Na caixa **Explore Plots** seleccione com um clique sobre o quadrado a opção **Normality plots with tests** [☒]	Selecciona automaticamente os testes de normalidade, assim como os gráficos «Normal Q-Q Plot» e «Detrended Normal Q-Q Plot».
5. Clique em **Continue** e posteriormente no botão **OK**	Vai gerar um output de resultados dos testes

O SPSS apresenta os resultados dos dois testes[28], neste caso deveremos utilizar o teste de Shapiro-Wilk dado n=20. Do extenso conjunto de quadros e gráficos resultantes da opção Explore do SPSS, seleccionou-se apenas o quadro 21, relativo aos resultados de ambos os testes.

Quadro 21 – Resultados dos testes K-S e S-W

	Kolmogorov-Smirnov[(a)]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	**Sig.**
das	,179	20	,091	,934	20	**,185**

[a] Lilliefors Significance Correction

[28] Não esqueça: - Se n≤ 50 deve usar-se o teste de Shapiro-Wilk.
- Se n>50 deve usar-se o teste de Kolmogorov-Smirnov

Como o valor da significância associada ao teste de Shapiro-Wilk foi de p=,185, valor superior ao erro tipo I que estabelecemos *à priori*, tal significa que a variável em questão segue distribuição normal na população.

Como regra para decisão relativamente ao pressuposto da normalidade em ambos os testes (K-S e S-W), devemos tomar a seguinte decisão: se a significância do teste (sig) for ≥ ao nível de significância fixado pelo investigador, assumimos que a variável em questão segue distribuição normal na população.

O passo seguinte é verificar se existe homogeneidade da variância, isto é, H_0: $\sigma_1^2 = \sigma_2^2$. O SPSS disponibiliza o teste de Levene, no entanto, como este pressuposto não é impeditivo à realização do teste *t*, encontrando-se nas várias opções dos testes paramétricos. Assim tanto se pode ir ao comando **Explore** onde se testa a normalidade, como o próprio software o calcula em simultâneo ao teste *t*, por defeito. Os comandos para a realização do teste de Levene e teste *t* são apresentados no quadro seguinte.

Quadro 22 – Comandos para realização do teste *t* para grupos independentes e teste de Levene

Comando do SPSS (acção):	Comentário:
1. Seleccione do menu as opções: **Analyze** → **Compare Means** → **Independents Samples T Test.**	Permite-lhe aceder à caixa de diálogo do teste
2. Na caixa de diálogo seleccione a variável *das* (com um clique sobre a variável) e depois com o botão ▶ coloque-a na lista do **Test variable(s):**	O **Test variable(s)** é o espaço onde se colocam as variáveis dependentes.
3. Seleccione a variável *local* com um clique e depois com o botão ▶ coloque-a no **Grouping Variable:**	O **Grouping Variable** é o espaço onde se coloca a variáveis «independente», isto é os grupos.
4. Dê um clique sobre **Define Groups**.	Surge a caixa de definição de grupos onde se inserem os códigos de cada grupo. Repare que aparece como: **local ? ?**
5. Na área do **Group 1:** introduza o código dos enfermeiros de CP (1), e no **Group 2:** o código dos enfermeiros de CSP (2).	Esta operação identifica os grupos que estão a ser comparados e aos quais atribuímos diferentes números.
6. Clique em **Continue** e posteriormente no botão **OK**	Vai gerar um output de resultados do teste

Os resultados do teste (retirados do Output – SPSS Viewer) são-nos apresentados nos quadros 23 e 24. O quadro 23 apresenta, além do tamanho de cada grupo, as medidas resumo da ansiedade perante a morte (das) para cada grupo, nomeadamente a média (Mean), o desvio padrão (Std. Deviation) e o erro padrão da média (Std. Error Mean). O arredondamento às centésimas[29] foi realizado a partir do próprio programa.

[29] No Output dê um clique duplo sobre o quadro (table) em questão. Ao ficar activo os limites do quadro passam a tracejado. Seleccione todas as células de interesse com o cursor do rato e depois dê um clique com o botão direito do rato e aceda à opção **Cell Properties**. Na opção **Value** coloque na célula **Decimals** o número 2 que significa duas casas decimais. Depois clique no botão **OK.**

Quadro 23 – Estatísticas resumo da ansiedade perante a morte (das)
em função do local de trabalho de dos enfermeiros

local	N	Mean	Std. Deviation	Std. Error Mean
CSP	11	9,00	2,37	,71
SU	9	10,00	2,40	,80

Como podemos observar a média da ansiedade medida pelas «*das*» é de 9,00 pontos para os enfermeiros de CSP, com um desvio padrão de 2,37 pontos. Os enfermeiros que trabalham em Serviço de Urgências, apresentam uma média de 10,0 pontos, superior à dos enfermeiros de CSP, com um desvio padrão de 2,40 pontos.

Como pode observar pelo quadro 24, o SPSS produz dois testes *t* de modo automático, uma para o caso em que podemos assumir a homogeneidade da variância, e neste caso deve ler-se na linha **Equal variances assumed,** e outro, para os casos em que tal não acontece (**Equal variances not assumed**). Como referimos anteriormente, quer exista, quer não exista homogeneidade da variância, isso não é restritivo à utilização do teste *t*, mas a decisão adequada relativamente aos testes *t*, está dependente do resultado do teste de Levene (Levene's Test for Equality of Variances).

Quadro 24 – resultados da aplicação do teste *t*.
Incluindo os resultados do teste de Levene

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
das	Equal variances assumed	**,000**	**1,000**	**-,935**	**18**	**,362**
	Equal variances not assumed			-,933	17,13	,364

Tal como nos outros testes, também no teste de Levene, o investigador deve definir *à priori* o nível de significância do teste (α=0,05), de modo a rejeitar ou não rejeitar a hipótese da homogeneidade da variância, cuja representação esquemática das hipóteses são respectivamente, H_0: $\sigma_1^2 = \sigma_2^2$ e H_1: $\sigma_1^2 \neq \sigma_2^2$. Neste caso o resultado do teste de Levene é de p=1,000, isto é superior ao alpha (0,05) estabelecido pelo investigador, pelo que se assume que as variâncias são homogéneas[30]. Assim os valores do teste *t* devem ser lidas na 1.ª linha

[30] A regra é simples.
► Se o valor da significância (sig) do teste de Levene for > 0,05 (erro assumido pelo investigador), devemos assumir a homogeneidade das variâncias, assim os resultados do teste *t* (t-test for Equality of Means) para uma decisão relativamente à hipótese são lidos na linha **Equal variances assumed**.
► Se o valor da significância (sig) do teste de Levene for < 0,05, devemos assumir a heterogeneidade das variâncias, assim os resultados do teste *t* para uma decisão relativamente à hipótese de investigação são lidos na linha **Equal variances not assumed**.

(**Equal variances assumed**). Caso o valor da sig. (p) fosse <0,05, deveríamos ler os resultados do teste na 2.º linha.

Os resultados do teste *t*, mostram que $t_{(observado)}$=**-,935, com gl=18 e p=,362**[31]. Neste caso não se rejeita a H_0, ou seja, as diferenças não são estatisticamente significativas. Como pode observar o valor da significância ou *p* do teste é > 0,05 (erro máximo tipo I estabelecido pelo investigador).

Como se apresentam na forma tabelar os resultados e qual a forma de os descrever? A forma que se apresenta é a mais comummente referenciada nas revistas e jornais científicos e adequada. Primeiramente apresentamos o quadro (quadro 25) e posteriormente o texto que acompanha a decisão.

Quadro 25 - Resultado da aplicação do teste *t* de student para grupos independentes.
Variável independente: local de trabalho dos enfermeiros; variável dependente: scores da das

Local de trabalho:	n	$\overline{X}$	s	*t*	p
Cuidados Saúde Primários	11	9,00	2,37	-,935	,362
Serviço Urgências	9	10,00	2,40		

No sentido de testar a hipótese que *"o nível de ansiedade perante a morte é diferente consoante o local de trabalho dos enfermeiros"*, procedeu-se à realização do *t* para grupos independentes, tendo-se previamente avaliado os pressupostos subjacentes à realização deste teste. Ambos os resultados dos testes (teste de Shapiro-Wilk com p=0,185; teste de Levene com p=1,000) permitiram prosseguir com a realização do teste *t*.

Os resultados mostram que o nível médio de ansiedade perante a morte é nos enfermeiros de cuidados de saúde primários de $\overline{X}$=9,00 (s=2,47) e nos enfermeiros que trabalham em Serviço de Urgências de $\overline{X}$ = 10,00 (s=2,40), **contudo as diferenças encontradas não se mostram estatisticamente significativas ($t_{(18)}$=-0,935; p=0,362)**, pelo que o local de trabalho dos enfermeiros não se constitui como um factor de diferenciação nos níveis de ansiedade perante a morte.

Repare na simbologia utilizada **($t_{(18)}$=-0,935; p=0,362)**. Contempla os graus de liberdade, o valor do *t* observado e nível de significância associado. As medidas resumo como a média e o desvio padrão estão arredondados às centésimas e os valores do *t* do teste e o valor do *p* estão arredondados às milésimas.

[31] Muitas vezes tem-se este pensamento: - «Não se rejeita a H_0, porque a probabilidade dela ser verdadeira é de 36,2%». Na realidade esta ideia é incorrecta, o que o valor de p=,362 indica é que caso a H_0 seja verdadeira, a probabilidade da sua ocorrência é >0,05. Isto leva-nos a não rejeitar a H_0, e a declarar as diferenças como não significativas estatisticamente.

6.3.3. Considerações suplementares

Uma das características do teste *t,* é que ele é robusto a violações dos seus pressupostos, especialmente a violações de normalidade e de homogeneidade da variância quando as amostras são suficientemente grandes (>25) e de igual tamanho (Runyon *et al,* 1996; Polit,1996). Outros autores como Freund & Simon (1997) sugerem que cumpridos os pressupostos, se $n_1 \leq 30$ e $n_2 \leq 30$, podemos utilizar a 1.ª formula apresentada (variâncias combinadas), posição partilhada por Pestana & Gajeiro (2000), no entanto sugerem que quando não se cumpre o pressuposto da homogeneidade da variância e $n_1 > 30$ e $n_2 > 30$, se deve utilizar a formula *t corrigida.* D` Hainaut (1997) alerta para todas as situações referenciadas anteriormente acrescentando que a dimensão das amostras não deverá ser *excessivamente diferente.* A nossa posição vai pela avaliação e cumprimento dos pressupostos para a utilização deste teste.

Para treinar o teste *t* para grupos independentes no SPSS, utilize a base de dados «**dap_medo.sav**». Trata-se de um estudo realizado em Enfermeiros e outros profissionais de saúde cuja hipótese refere que «o medo da morte é inferior nos enfermeiros (grupo 1), comparativamente aos restantes profissionais de saúde (grupo 2). Utilize um erro máximo tipo I de 0,05. Teste os pressupostos.

6.3.4. Exercícios

1. De modo a testar a hipótese de que “*o nível de depressão pós-parto é mais elevado nas puérperas primíparas do que nas puérperas multíparas*”, os dados foram analisados em separado, tendo-se obtido as seguintes estatísticas:

Paridade	n	$\bar{x}$	s
Primíparas	54	13.03	6.15
Multíparas	39	10.95	7.86

a) - Admitindo que “*o nível de depressão pós-parto*” é uma variável normal e existe homogeneidade das variâncias das duas amostras, efectue os cálculos relativos ao teste de hipótese adequado e indique a conclusão com um nível de significância de 5%.

2. Para testar a hipótese de que "*o teor de magnésio sérico é superior nos lactentes com suplemento de vitamina D*", seleccionou-se uma amostra aleatória e a partir dos respectivos dados calcularam-se as seguintes estatísticas:

Suplemento vit. D	n	$\overline{x}$	s
Sim	34	1.74 mEq/l	0.22 mEq/l
Não	21	1.63 mEq/	0.18 mEq/l

a) - Qual o teste *t* para diferenças de médias que se aplica neste caso?

b) - O que conclui em relação à hipótese formulada? Utilize um erro tipo I máximo de 0.05.

3. Num estudo que tinha como objectivo avaliar o conhecimento dos adolescentes sobre práticas sexuais de risco, foram comparados dois grupos de adolescentes (sexo masculino e sexo feminino). A hipótese do estudo refere que "*o nível de conhecimento sobre práticas sexuais de risco é superior nas adolescentes*". A amostra foi seleccionada aleatoriamente e a partir dos respectivos dados calcularam-se as seguintes estatísticas:

Sexo:	n	$\overline{x}$	s
Masculino	157	31.56	10.45
Feminino	206	38.81	11.08

a) - O que conclui em relação à hipótese formulada? Utilize um erro tipo I máximo de 0,01.

6.4. Teste *t* para amostras emparelhadas

Existe ainda um outro teste *t,* também de diferença de médias, mas aplicado às médias de duas amostras emparelhadas. É um teste paramétrico que permite verificar a diferença de duas médias, mas quando é avaliado o mesmo[32] grupo de indivíduos em dois momentos distintos, ou seja, permite verificar se há alguma alteração e de que forma de um 1.º momento (antes) para um 2.º momento (depois).

Um exemplo simples. 18 adolescentes obesas (com *Índice de Massa Corporal* > 30) sujeitaram-se a um determinado *regímen* de regulação do peso corporal com objectivo de reduzir o IMC, tendo-se registado os seguintes valores do *Índice de Massa Corporal* antes e 6 meses após de iniciar o referido regímen.

Antes	6 meses após
30.08	30.09
30.09	26.03
30.12	25.08
30.32	32.04
31.00	30.06
31.12	31.14
32.18	28.83
32.56	33.45
33.45	34.67
33.89	25.68
34.10	27.90
34.61	34.65
35.80	36.71
35.92	35.97
36.78	30.65
37.00	36.34
37.42	38.10
38.90	35.44

A hipótese nula (H_0) do teste é que as diferenças (D) do 1.º para o 2.º momentos serão D=0, simbolicamente teremos H_0: $\mu_D=0$. A hipótese de investigação será H_1: $\mu_D\neq0$. O α a utilizar será de 5%.

O teste é realizado através dos cálculos indicados por qualquer uma das duas fórmulas:

[32] Existem outras formas de emparelhamento, dependendo da natureza do estudo. A título de exemplo, um estudo realizado na década de 90 em que emparelhando pai e mãe, se avaliava o nível de stress de ambos, numa situação que envolvia o facto de ter um filho menor a ser submetido a um tratamento de quimioterapia.

$$t=\frac{\sum D}{\sqrt{\frac{n\sum D^2-(\sum D)^2}{n-1}}}$$

$gl = n-1$

$$t=\frac{\bar{d}}{\frac{\hat{s}_d}{\sqrt{n}}}$$

$gl = n-1$

Calculando as diferenças individuais do peso, antes e depois (D) do regímen, para utilizar a 1.ª fórmula teremos:

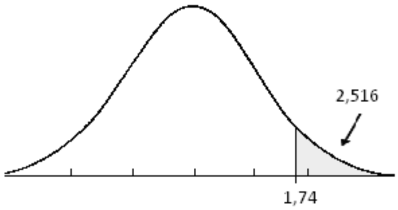
Antes	6 meses Após	*D*	*D²*
30,08	30,09	-,01	,00
30,09	26,03	4,06	16,48
30,12	25,08	5,04	25,40
30,32	32,04	-1,72	2,96
31,00	30,06	,94	,88
31,12	31,14	-,02	,00
32,18	28,83	3,35	11,22
32,56	33,45	-,89	,79
33,45	34,67	-1,22	1,49
33,89	25,68	8,21	67,40
34,10	27,90	6,20	38,44
34,61	34,65	-,04	,00
35,80	36,71	-,91	,83
35,92	35,97	-,05	,00
36,78	30,65	6,13	37,58
37,00	36,34	,66	,44
37,42	38,10	-,68	,46
38,90	35,44	3,46	11,97
		$\Sigma D = 32,51$	$\Sigma D^2 = 216,3535$

Substituindo na fórmula teremos:

$$t=\frac{32,51}{\sqrt{\frac{18x(216,3535)-(32,51)^2}{18-1}}}=2,516$$

O valor do nosso $t_{(observado)}$= 2,52, resta procurar na tabela o valor do t teórico. Os graus de liberdade neste teste são de (n–1), logo 17. Consultando os valores do excerto da tabela *t* apresentada abaixo (região de rejeição de H_0 unilateral à direita), verifica-se que o valor do $t_{(teórico)}$ = 1,74.

Figura 5 – Esquema de tomada de decisão
para teste de hipótese com região de rejeição unilateral à direita alpha = 0,05 (gl=17)



Como se pode verificar, há evidência empírica para afirmar que o regímen contribui para a diminuição do Índice de Massa Corporal das adolescentes ($t_{(17)}$= 2,516; p<0,05).

Tabela 1 - Distribuição *t* (excerto)

	α para teste unilateral			
	0,05	**0,025**	**0,01**	**0,005**
	α para teste bilateral			
Graus de Liberdade	**0,10**	**0,05**	**0.02**	**0,01**
1	6.31	12.71	31.82	63.66
2	2.92	4.30	6.97	9.93
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
17	**1.74**	2.11	2.57	2.90

Relativamente à direccionalidade das hipóteses para os casos em que são ou não direccionadas.

Bicaudal	Unicaudal à esquerda	Unicaudal à direita
H_0:μ_D = 0	H_0:μ_D $\geq$ 0	H_0:μ_D $\leq$ 0
H_1:μ_D $\neq$ 0	H_1:μ_D < 0	H_1:μ_D > 0

Também poderá utilizar a fórmula alternativa de modo simples. Suponha que em 49 mulheres foi detectada uma depressão moderada ou grave (acima de 12 pontos na escala DNPF) no pós-parto. Nesse sentido foi planeado um plano de intervenção de enfermagem. De modo a testar a hipótese de que a «o plano de intervenção de enfermagem aplicado diminui significativamente o nível de depressão pós-parto», apresentam-se a seguir os resultados da aplicação da escala antes e depois dessa intervenção:

	Antes (a)	Depois (b)	Diferença (d) [b-a)
$\bar{x}$	17.79	13.54	-4.25
s	4.36	2.68	1.63

a) Qual a conclusão do teste da hipótese com um nível de significância de 0.05?

$$t = \frac{\bar{d}}{\frac{\hat{s}_d}{\sqrt{n}}} \qquad gl=n-1$$

Na fórmula substitua os símbolos pelos valores correspondentes.

6.4.1. O teste *t* para amostras emparelhadas na máquina de calcular

O cálculo do teste *t* para amostras emparelhadas é fácil e simples com recurso à máquina de calcular. A hipótese é de que «o número de cigarros fumados diariamente diminui depois da frequência de um curso breve de educação para a saúde». Foi avaliado o número de cigarros consumidos diariamente antes e depois do curso em 10 mulheres. Simbolicamente, teremos H_0:$\mu_D \leq 0$ e H_1:$\mu D > 0$. Os dados são apresentados de seguida.

Sujeito	Antes	Depois
a	31	32
b	30	28
c	31	27
d	35	29
e	34	35
f	32	28
g	31	24
h	34	26
i	36	35
j	38	33

O primeiro passo consiste no cálculo manual das diferenças entre os valores antes e depois. É com as diferenças que vamos realizar o teste e tomar uma decisão quanto à hipótese. Deverá trabalhar com as listas, nomeadamente introduzindo as diferenças na lista L1. De seguida deverá calcular os valores da média e o desvio padrão da amostra (já corrigido).

Sujeito	Antes	Depois	d
a	31	32	-1
b	30	28	2
c	31	27	4
d	35	29	6
e	34	35	-1
f	32	28	4
g	31	24	7
h	34	26	8
i	36	35	-1
j	38	33	5
			$\overline{X}$ = 3,50
			$\hat{S}_d$ = 3,17
			n=10

De seguida deverá seguir as seguintes instruções. Dado que apresentamos os valores de $\overline{d}$ ($\overline{\text{x}}$= 3,50) e de $\hat{\text{s}}_{\text{d}}$ = 3,17 e n=10. Não esqueça que $\hat{\text{s}}_{\text{d}}$ equivale ao desvio padrão da amostra.

Instruções para cálculo do teste *t* para grupos emparelhados	
TEXAS	CASIO
1. Pressione o botão **STAT**. 2. Mova o cursor para a opção **2: T-Test**; 3. ENTER 4. Seleccione **Stats** pressionando a tecla **ENTER.** 4. Introduza os valores: μ0 : 0 $\overline{\text{X}}$:3,5 **Sx**:3,17 **n:** 10 μ: >μ0 5. **Calculate** e **ENTER**.	1. Pressione o botão **MENU** e depois seleccione a opção **STAT**, **EXE**; 2. Seleccione a opção **TEST** (tecla F3); 3. Seleccione a opção **t** (tecla F2); 4. Seleccione a opção **1-S** (tecla F1) 5. Na opção **Data**, seleccione a opção **var** (F2); 6. Introduza: μ : >μ0 μ0 : 0 $\overline{\text{X}}$:3,5 **xσn-1**:3,17 **n:** 10 7. **EXE**.
Output (Resultados): μ>0 t =3.49 p =0,0034 $\overline{\text{X}}$ =3,5 **Sx**=3,17 **n: 10**	

6.4.2. O cálculo do teste *t* para amostras emparelhadas no SPSS

Neste exemplo pretendemos testar o efeito de um regímen de controlo do peso em adolescentes obesas, nomeadamente na redução do IMC. Trata-se do exemplo já apresentado anteriormente. A hipótese de investigação refere que o «regímen de controlo de peso diminui o IMC das adolescentes obesas»**.** Simbolicamente teremos: H_0:$\mu_d \leq 0$ e H_1:$\mu_d > 0$.

A base deste exemplo designa-se de «**t_emparelhado.sav**». Vão analisar-se os resultados obtidos numa amostra aleatória de 18 adolescentes obesas. Como já se referiu, torna-se necessário verificar o pressuposto da normalidade subjacente ao teste *t*, cujos resultados estão apresentados no quadro 26. Os comandos para verificar este pressuposto foram apresentados aquando da realização do teste *t* para grupos independentes (quadro 22).

Quadro 26 – Resultados dos testes K-S e S-W

	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
IMC_antes	,143	18	,200(*)	,929	18	,189
IMC_6mesesdepois	,145	18	,200(*)	,949	18	,406

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

Como se pode observar do quadro anterior, em ambos os casos (IMC antes e depois) o valor da significância obtida pelo teste de Levene é > 0,05, pelo que estamos dentro das condições impostas à realização do teste paramétrico *t*, logo podemos prosseguir com os comandos necessários à sua obtenção com o objectivo de comparar as médias das duas amostras emparelhadas.

Os comandos para o cálculo do teste *t* para amostras emparelhadas aparecem no quadro seguinte.

Quadro 27 – comandos para realização do teste *t* para grupos independentes

Comando do SPSS (acção):	Comentário:
1. Seleccione do menu as opções: **Analyze** → **Compare Means** → **Paired Samples T Test.**	Permite-lhe aceder à caixa de diálogo do teste
2. Na caixa de diálogo seleccione as variáveis *IMC_antes* e *IMC_6mesesdepois* e depois com o botão ▶ coloque-a na lista do **Paired variables:**	O **Paired variables** é o espaço onde se colocam as variáveis que vamos comparar
3. Clique em **Continue** e posteriormente no botão **OK**	Vai gerar um output de resultados do teste

De seguida o SPSS gera um Output com os seguintes quadros. O quadro seguinte (28), estatísticas resumo, mostra que a média do IMC antes da frequência do regímen foi de IMC= 33,63 e depois veio a situar-se nos 31,82.

Quadro 28 – Estatísticas resumo do IMC antes e depois da frequência do regímen

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	IMC_antes	33,63	18	2,86	,68
	IMC_depois	31,82	18	4,07	,96

O quadro 29 mostra os resultados do cálculo do coeficiente de correlação de Pearson e respectivo teste de significância. Deve haver sempre correlação entre os dois grupos (momentos) para se utilizar este teste. Se não existir correlação entre os dois grupos ou o seu valor for pequeno, significa que o emparelhamento não foi útil. A correlação obtida (r=,665) é forte, positiva, linear e estatisticamente significativa entre as variáveis (p<0,05).

Quadro 29 – Resultados do calculo do r de Pearson e respectivo teste de significância

		N	Correlation	Sig.
Pair 1	IMC_antes & IMC_depois	18	,665	,003

Relativamente ao quadro 30, encontram-se a média das diferenças emparelhadas relativas aos dois momentos (1,81), o que denota uma diminuição nos valores do IMC após a frequência do regímen de controlo do peso, o desvio padrão e o erro padrão da média. Para se saber se esta diferença é significativa observam-se os valores do teste *t*.

Quadro 30 – Resultados da aplicação do teste *t* para grupos emparelhados aos valores do IMC

		Paired Differences					
		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	IMC_antes - IMC_depois	1,81	3,05	,72	2,516	17	,022

Pode então concluir-se que o valor da diferença das médias (1,81) é estatisticamente significativa **($t_{(17)}$=2,516;p=,011[33])**, o que denota que o regímen teve um efeito positivo, reduzindo o Índice de Massa Corporal das adolescentes obesas.

A apresentação dos resultados (quadro 31) é conclusão relativamente à hipótese são apresentados de seguida.

Quadro 31 – Resultados da aplicação do teste *t* para grupos emparelhados aos valores do IMC

IMC	$\overline{x}$	s	*t*	p
Antes Depois	33,63 31,82	2,86 4,07	2,516	,011

No sentido de testar a hipótese de que «o regímen de controlo de peso altera o IMC das adolescentes obesas», os resultados do teste *t* para grupos emparelhados revelou uma diferença estatisticamente significativa ($t_{(17)}$=2,516;p=,011) nos valores do IMC das adolescentes. Mais se observa que a média do IMC era de 33,61 (s=2,86) antes da frequência do regímen vindo a obter-se uma $\overline{x}$ = 31,82 (s=4,07) 6 meses após o regímen.

6.4.3. Exercícios

1. De modo a testar a hipótese de que «a aplicação do creme *Kerniprev* reduz os valores de bilirrubina indirecta no sangue dos recém-nascidos", foi feita uma experiência aplicando o referido produto em 24 recém-nascidos. Apresentam-se no quadro seguinte os valores médios e os desvios padrão (em mg/dl) da bilirrubinémia <u>antes</u> e 6 horas <u>depois</u> da aplicação.

	Antes	Depois	Diferença (d)
$\overline{x}$	9,6	9.3	-0.3
s	1.2	1.6	0.9

a) Qual a conclusão do teste da hipótese com um nível de significância de 0.05?

[33] Não esqueça que a hipótese é direccionada, logo o valor deverá ser dividido por 2.

6.5. ANOVA one-way

Uma das limitações do *teste t* (para amostras independentes) assenta no facto de apenas nos permitir comparar dois grupos, e muitas vezes nos estudos que realizamos, temos necessidade de comparar três ou mais grupos.

Podemos questionar-nos se não será possível utilizar o teste *t*, por exemplo, para comparar a satisfação profissional dos enfermeiros (variável dependente) em termos de três categorias/grupos (variável independente), nomeadamente enfermeiros de nível I, especialistas e chefes. Se realizássemos o teste *t,* comparando os grupos dois a dois, seriam necessários três testes. Mais, se admitirmos um erro tipo I máximo de 0,05, verificamos um aumento considerável deste erro. Ao fazermos *n* comparações de valores médios dois a dois, a probabilidade de erro é de $(1 – (1- \alpha)^n)$x100. Assim se tivermos que comparar três grupos, cada um com a sua média, teremos $(1 – (1\text{-}0,05)^3)$x100, o que equivale a um erro de **14,3%** (Runyon *et al*., 1996).

Nestes casos teremos que recorrer a um outro teste para as situações que envolvem três ou mais grupos, a **ANOVA** (acrónimo de **análise de variância**). A análise da variância pode ainda designar-se de *análise da variância de um critério* (ou ANOVA de uma via – *one-way* ANOVA) nos casos que envolve apenas uma variável independente[34], caso tenhamos duas variáveis independentes a tem a designação de ANOVA com dois factores ou *two-way* ANOVA.

Comecemos com um exemplo simples para compreender a lógica deste teste. Um enfermeiro pretende comparar a *preocupação com o peso,* que é uma das dimensões da *imagem corporal* em três tipos de doentes crónicos, *insuficientes renai*s, *ostomizados* e *hipertensos*. A hipótese do estudo refere que «a preocupação com o peso é diferente consoante o tipo de doença crónica» e vamos utilizar um alpha de 0,05.

Simbolicamente a hipótese nula (H_0) é definida da seguinte forma: $H_0 : \mu_1 = \mu_2 =.... \mu_k$. A H_1 afirma que pelo menos uma das médias é diferente. A ANOVA considera a existência de duas fontes de variação que contribuem para a *variação total* nos resultados, e procede a uma partição da variabilidade total ou variação total em duas componentes (costuma dizer-se que se efectua uma partição das variâncias), *a variância sistemática* (pode assumir outras

[34] A variável independente pode assumir outras designações, como por exemplo **factor**. Um factor (ou V.I.) é constituído por grupos, que podem designar-se de tratamentos, níveis, ou condições da V.I., estando muitas vezes dependentes do desenho de investigação.

designações como *variação factorial*, ou *variação entre tratamentos*) e a *variância de erro* (também pode assumir a designação de *variação residual*).

A variância sistemática refere-se à variância *entre os grupos* (**B**etween groups) e a *variância de erro* (***W**ithin groups)* deve-se a *flutuações aleatórias* dentro dos próprios grupos. O que o teste nos permite concluir é se *a variância é maior entre os grupos ou dentro dos grupos*, ou seja se a variância sistemática é diferente da variância de erro.

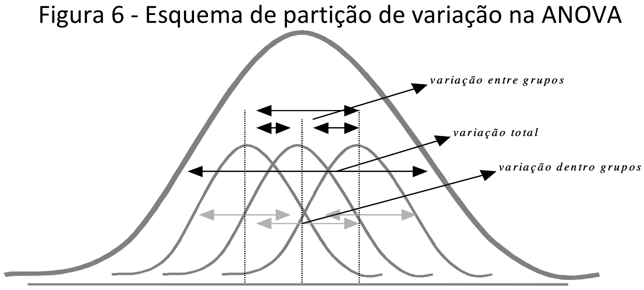
Simbolicamente as hipóteses podem ser também traduzidas da seguinte forma:

H_0: $\sigma^2_B = \sigma^2_W$.

H_0: $\sigma^2_B \neq \sigma^2_W$.

Podem pois ocorrer duas situações, a variância entre os grupos é maior que a variância dentro dos grupos ($\sigma^2_B \neq \sigma^2$) e isso pode levar-nos à rejeição da H_0, ou então pode acontecer outra situação, $\sigma^2_B = \sigma^2_W$ e então não rejeitamos H_0. Para observar a partição da variação observe a figura seguinte.

Figura 6 - Esquema de partição de variação na ANOVA



Vamos verificar como funciona a ANOVA para a hipótese que referimos anteriormente. Repare no quadro seguinte 32. Cada coluna apresenta os *scores* resultantes do somatório dos itens (pontuações brutas) da dimensão preocupação com o peso. Estão também calculadas as médias para cada um dos grupos, e a média global ($\overline{X}_G$).

Quadro 32 – *scores* obtidos pelos doentes crónicos relativos à dimensão *preocupação com o peso* - Inclui as médias

IRC	Ostomizados	Hipertensos	
30	40	20	↑
32	42	22	*Variação*
34	40	26	*dentro*
38	48	**28**	*dos grupos*
40	52	30	↓
$\overline{X}_1 = 34,8$	$\overline{X}_2 = 44,4$	$\overline{X}_3 = 25,2$	$\overline{X}_G = 34,8$
←*Variação entre os grupos* →			

Como pode observar no grupo de doentes com IRC os *scores* variam entre 30 e 40 pontos, no grupo de Ostomizados variam entre 40 e 58 pontos, e no grupo de Hipertensos variam entre 20 e 30 pontos. Entre os grupos as médias variam de 44,40 e 25,2 pontos.

Observe a lógica da partição da variação. Tomemos o exemplo do doente com IRC que obteve um *score* de 30 pontos. O seu *score* desvia-se da média dos scores do seu grupo em 4,8 pontos (*variação dentro do grupo*), o seu grupo (IRC) desvia-se da média global ($\overline{X}_G$) em 0 pontos (*variação entre grupos*). Este doente que obteve 30 pontos desvia-se da média global em 4,8 pontos (*variação total*).

Por outras palavras, para este doente teremos:

Variação dentro dos grupos = 4,8

Variação entre grupos = 0,0

Variação total = 4,8

Quando os desvios como este são obtidos para todos os sujeitos, podemos juntar toda a informação e construir uma razão que agrega a variação entre os grupos e a variação dentro dos grupos, designada de F ratio, que constitui a distribuição estatística que é utilizada na ANOVA:

$$F = \frac{\text{Variabilidade entre os grupos}}{\text{Variabilidade dentro dos grupos}}$$

Assim se as médias forem iguais a variação entre os grupos será zero, ou seja, não rejeitamos a hipótese nula. Quando as médias são diferentes, e se rejeita a hipótese nula, isso pode ser atribuído a dois factores, erro amostral e ao efeito da variável independente, logo a razão também será:

$$F = \frac{\text{Efeito da variável independente} + \text{Erro amostral}}{\text{Erro amostral}}$$

Vamos calcular manualmente a ANOVA para observar como se processa o teste.

Quadro 33 – Cálculos necessários para a obtenção do SS dentro dos grupos

IRC	$(X - \overline{X}_1)$	$(X - \overline{X}_1)^2$	OST.	$(X - \overline{X}_2)$	$(X - \overline{X}_2)^2$	HIP	$(X - \overline{X}_3)$	$(X - \overline{X}_3)^2$
30	- 4,8	23,04	40	- 4,4	19,36	20	- 5,2	27,04
32	- 2,8	7,84	42	- 2,4	5,76	22	- 3,2	10,24
34	- 0,8	0,64	40	-4,4	19,36	26	0,8	0,64
38	3,2	10,24	48	3,6	12,96	28	2,8	7,84
40	5,2	27,04	52	7,6	57,76	30	4,8	23,04
$\overline{X}_1 = 4,8$		$\Sigma x^x = 69,16$	$\overline{X}_2 = 44,40$		$\Sigma x^x = 115,2$	$\overline{X}_3 = 25,2$		$\Sigma x^x = 68,80$

Os doentes IRC apresentam um nível médio de preocupação com o peso de 34.80 pontos, os ostomizados uma média de 44,40 pontos e os hipertensos de 34,80 pontos. A média dos três grupos (média global) será:

$$\overline{x} = \frac{34{,}8 + 44{,}4 + 25{,}2}{3} = 34{,}8 \text{ pontos.}$$

Conforme pode verificar no quadro 33, estão calculados para cada grupo os quadrados dos desvios de cada valor em relação à média do seu grupo, e a respectiva soma dos quadrados. Como pode verificar o somatório dos quadrados é no grupo de IRC de 69,16, no grupo de ostomizados de 115,2 e nos hipertensos de 68,80.

O cálculo da ANOVA implica o conceito de somatório dos quadrados (SQ) (Sum of Squares - *SS*) que não é mais do que o quadrado dos desvios em relação à media (relembre o cálculo da variância). O somatório que apresentámos anteriormente é o somatório dos quadrados dentro dos grupos (SS_W), para cada um dos grupos, por exemplo nos IRC o Σx_1^2 = 69,16. Este somatório permite-nos verificar a variação da pontuação de cada doente relativamente ao seu grupo (IRC).

O somatório dos quadrados dentro dos grupos será:

$SS_W = (\Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2)$

$SS_W = 69{,}16 + 115{,}2 + 68{,}80 = 253{,}16$

Se o processo anterior nos permite encontrar a variação dos scores de cada indivíduo em relação à média do seu grupo, falta-nos verificar a variação da média de cada grupo em relação a média global. A variação entre os grupos, vai designa-se por soma do quadrado entre os grupos (SS_B). Relembre as médias encontradas anteriormente.

Para obter o SS_B usamos um processo idêntico ao anterior, conforme o quadro 34.

Quadro 34 – Cálculos necessários para a obtenção do SS entre os grupos

	$(x - \overline{x}_g)$	x_g^2	n	$X_g^2 . n$
$\overline{x}$ 1 = 34,8	(34,8- 34,8) = 0	0	5	0
$\overline{x}$ 2 = 44,4	(44,4 – 34,8) = 9,6	92,16	5	460,8
$\overline{x}$ 3 = 25,.2	(25,2 – 34,8) = -9,6	92,16	5	460,8
$\overline{x}_g$ = 34.8				SS_B = 921,6

Neste caso será: $SS_B = 0 + 460{,}8 + 460{,}8 = 921{,}6$

Então para encontrar a variabilidade total, que se designa soma dos quadrados total, ela não é mais que a soma das duas fontes de variância, entre e dentro dos grupos:

$SS_T = SS_w + SS_B$

$SS_T = 253{,}16 + 921{,}6 = 1174{,}76$

Então o cálculo do F será: $F = \dfrac{SS_B/df_B}{SS_W/df_W} = \dfrac{MS_B}{MS_W}$

Em que SS_B/df_B corresponde à *média dos quadrados entre os grupos* (MS_B) e SS_W/df_W corresponde à média dos quadrados dentro dos grupos (MS_W).

Os graus de liberdade são respectivamente:
$df_B = k–1$.
$df_w = N–k$.
$df_T = N–1$.
O F do teste é dado por: $\dfrac{MS_B}{MS_W}$

Teremos então o valor do $F_{(observado/teste)}$= 460,8/21,10 = 21,84. Sintetizando, podemos verificar através do quadro 35, todos os valores calculados até aqui, respeitantes à ANOVA.

Quadro 35 - resultados do cálculo da ANOVA

Fontes de variância	Soma dos Quadrados (SS)	Graus de liberdade (df)	Média dos Quadrados (MS)	F
Entre os Grupos	921,6	2	613,200	24,381
Dentro dos grupos	253,16	12	21,10	
Total	1174,76	14		

O passo seguinte consiste na comparação do $F_{(observado/teste)}$ com o $F_{(teórico)}$. Se o valor do teste for superior ao valor teórico estamos aptos a rejeitar a hipótese nula, caso contrário não rejeitamos H_0.

A utilização da tabela é análoga ao que temos vindo a fazer, no entanto considera os graus de liberdade do numerador (df_B) com os graus de liberdade do denominador (df_W). Neste caso $df_B = 2$ e $df_W = 12$. A tabela da distribuição F (versão reduzida) é para $\alpha = 0{,}05$

Como pode observar, o valor crítico será de 3,8853. Como O $F_{(observado)}$= 21,84 e $F_{(crítico)}$ = 3,885, rejeitamos a hipótese nula admitindo um erro máximo de 0,05. Quer isto dizer que $F_{(2;12)}$ =24,381, p<0,05, pelo que a dimensão preocupação com o peso é diferente consoante a patologia dos doentes. Repare que se coloca na simbologia os valores dos graus de liberdade (df) que utilizamos para fazer a leitura na tabela ($F_{(2;12)}$).

Tabela 3 - Distribuição F-Snedecor (α=0,05)

df_W	df_B								
	1	2	3	4	5	6	7	8	9
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38
3	10,13	9,552	9,277	9,117	9,014	8,941	8,887	8,845	8,812
4	7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041	5,999
5	6,608	5,786	5,410	5,192	5,050	4,950	4,876	4,818	4,773
6	5,987	5,143	4,757	4,534	4,387	4,284	4,207	4,147	4,099
7	5,591	4,737	4,347	4,120	3,972	3,866	3,787	3,726	3,677
8	5,318	4,459	4,066	3,838	3,688	3,581	3,501	3,438	3,388
9	5,117	4,257	3,863	3,633	3,482	3,374	3,293	3,230	3,179
10	4,965	4,103	3,708	3,478	3,326	3,217	3,136	3,072	3,020
11	4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948	2,896
12	4,747	**3,885**	3,490	3,259	3,106	2,996	2,913	2,849	2,796
13	4,667	3,806	3,411	3,179	3,025	2,915	2,832	2,767	2,714
14	4,600	3,739	3,344	3,112	2,958	2,848	2,764	2,699	2,646
15	4,543	3,682	3,287	3,056	2,901	2,791	2,707	2,641	2,588
16	4,494	3,634	3,239	3,007	2,852	2,741	2,657	2,591	2,538
17	4,451	3,592	3,197	2,965	2,810	2,699	2,614	2,548	2,494
18	4,414	3,555	3,160	2,928	2,773	2,661	2,577	2,510	2,456
19	4,381	3,522	3,127	2,895	2,740	2,628	2,544	2,477	2,423
20	4,351	3,493	3,098	2,866	2,711	2,599	2,514	2,447	2,393
21	4,325	3,467	3,073	2,840	2,685	2,573	2,488	2,421	2,366
22	4,301	3,443	3,049	2,817	2,661	2,549	2,464	2,397	2,342
23	4,279	3,422	3,028	2,796	2,640	2,528	2,442	2,375	2,320
24	4,260	3,403	3,009	2,776	2,621	2,508	2,423	2,355	2,300
25	4,242	3,385	2,991	2,759	2,603	2,490	2,405	2,337	2,282
26	4,225	3,369	2,975	2,743	2,587	2,474	2,388	2,321	2,266
27	4,210	3,354	2,960	2,728	2,572	2,459	2,373	2,305	2,250
28	4,196	3,340	2,947	2,714	2,558	2,445	2,359	2,291	2,236
29	4,183	3,328	2,934	2,701	2,545	2,432	2,346	2,278	2,223
30	4,171	3,316	2,922	2,690	2,534	2,421	2,334	2,266	2,211
40	4,085	3,232	2,839	2,606	2,450	2,336	2,249	2,180	2,124
60	4,001	3,150	2,758	2,525	2,368	2,254	2,167	2,097	2,040
120	3,920	3,072	2,680	2,447	2,290	2,175	2,087	2,016	1,959
∞	3,842	2,996	2,605	2,372	2,214	2,099	2,010	1,938	1,880

A ANOVA por si só, e no caso em que rejeitamos a H_0 não nos permite concluir que $\mu_1 \neq \mu_2 \neq \mu_3$. Trata-se de um teste designado de *omnibus*, isto é, testa a hipótese na sua generalidade, como um todo. Caso se rejeite a hipótese nula (H_0) necessitamos de realizar um conjunto de procedimentos *à posterior* para identificar quais as médias que diferem entre si de modo significativo, ou pela análise da sua tendência, por exemplo tendência linear, se for significativa a diferença e o factor for ordinal, ou por comparações múltiplas e que incluem um conjunto de testes designados de *post hoc*.

6.5.1. Comparações *à posteriori* (não planeadas)

É comum designar este processo como *comparações à posteriori*. *Posteriori* no sentido de "depois dos factos". Como não estabelecemos à partida hipóteses relativamente às diferenças nas médias dos grupos, isto é não dizemos quais vão ser diferentes ou quais vão ser iguais, este procedimento mostra-se muito útil e é o mais comum na investigação. Estas comparações, pelo motivo referido anteriormente designam-se também por comparações não planeadas.

Um dado importante, os procedimentos à posteriori procuram conciliar dois objectivos, por um lado proteger-nos quanto à inflação do α criada pelo número de comparações, por outro necessitamos de comparações com poder suficiente para detectar as diferenças entre as médias e assim encontrar as diferenças.

Existe uma grande panóplia de testes *post hoc* disponíveis no SPSS e que nos indicam de modo imediato quais as médias que apresentam diferenças significativas entre si. Um dos testes *post hoc* que procura conciliar os objectivos anteriores é o **Tukey's honestly significant difference (HSD).**

6.5.2. A ANOVA realizada no SPSS (inclui testes *post hoc*)

Antes mesmo de iniciar o cálculo da ANOVA no SPSS relembramos que ela está sujeitas aos mesmos pressupostos do teste *t* para grupos independentes.

Imagine que queríamos testar (α = 0,05) a hipótese de que «o tempo gasto (em minutos) semanalmente na prática de exercício físico é diferente consoante a profissão dos técnicos de saúde». A base deste exemplo designa-se de **«anova.sav».**

Para testar o pressuposto da normalidade deverá proceder conforme as instruções do quadro 36.

Quadro 36 - comando para realização dos testes de normalidade (K-S e S-W)

Comando do SPSS (acção):	Comentário:
1. Seleccione do menu as opções: **Analyze** → **Descriptive Statistics** → **Explore**	Acede à caixa de diálogo do Explore que permite realizar os testes de normalidade.
2. Na caixa de diálogo Explore, seleccione a variável *tempo* (com um clique sobre a variável) e depois com o botão ► coloque-a na **Dependent List:**	O **Dependent List:** é o espaço onde se colocam as variáveis que procuramos testar.
3. Dê um clique sobre o botão **Plots…**	Acede a nova caixa de diálogo, **Explore Plots**
4. Na caixa **Explore Plots** seleccione com um clique sobre o quadrado a opção **Normality plots with tests** [☒]	Selecciona automaticamente os testes de normalidade, assim como os gráficos «Normal Q-Q Plot» e «Detrended Normal Q-Q Plot».
5. Clique em **Continue** e posteriormente no botão **OK**	Vai gerar um output de resultados dos testes

Relativamente ao pressuposto da normalidade, o quadro 37 apresenta os resultados do teste de Kolmogorov-Smirnov com correcção de significância de Lilliefors (n> 50). Como pode observar a variável tempo segue distribuição normal na população dado que o valor da significância do teste (p=,200) é superior a 0,05 (erro tipo I definido pelo investigador).

Quadro 37 – Resultados dos testes de Normalidade

	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
tempo	,072	80	,200(*)	,980	80	,260

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

Relativamente ao pressuposto da homogeneidade da variância, ele pode ser realizado em simultâneo com a ANOVA, no entanto esse pressuposto, sob determinadas circunstâncias, pode não ser restritivo à realização da ANOVA. Quando não existe homogeneidade da variância, o SPSS calcula os testes alternativos, nomeadamente o Brown-Forsythe ou então o teste de Welch. Para tal, deverá, no cálculo da ANOVA, solicitar esses testes na opção Options. No quadro 38 estão descritos os comandos para a realização da ANOVA.

Quadro 38 – comandos para realização do teste *t* para grupos independentes e teste de Levene

Comando do SPSS (acção):	Comentário:
1. Seleccione do menu as opções: **Analyze** → **Compare Means** → **One-Way ANOVA...**	Permite-lhe aceder à caixa de diálogo do teste
2. Na caixa de diálogo (**One-Way ANOVA**) seleccione a variável *tempo* (com um clique sobre a variável) e depois com o botão ▶ coloque-a na lista do **Dependent List:**	O **Dependent List** é o espaço onde se colocam as variáveis dependentes.
3. Seleccione a variável *profissão* com um clique e depois com o botão ▶ coloque-a no **Factor:**	O **Factor** é o espaço onde se coloca a variável «independente», isto é os grupos.
4. Dê um clique sobre o botão **Options** e seleccione as opções **Descriptive** e **Homogeneity of variance test**	Este comando permite calcular as medidas resumo para cada grupo, assim como realiza o teste de Levene.
5. Dê um clique sobre o botão Continue seguido de OK.	O computador realiza o teste e apresenta o Output de resultados.

O Output resultante apresenta em 1.º lugar um quadro de resultados com as estatísticas apresentadas no quadro 39.

Como pode observar, encontram-se o tamanho de cada grupo, a média ($\bar{x}$), o desvio padrão (s), o erro padrão da média, o intervalo de 95% de confiança para a média de cada grupo, e os valores mínimo e máximo observado em cada grupo.

Quadro 39 - Resultados do cálculo da opção Descriptives do SPSS (inclui medidas resumo e IC 95%)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Médicos	20	203,15	26,46	5,92	190,77	215,53	156	250
Enfermeiros	20	165,85	29,29	6,55	152,14	179,56	102	210
Assistentes Sociais	20	186,75	29,56	6,61	172,91	200,59	124	240
Psicologos	20	127,30	26,78	5,99	114,77	139,83	89	194
Total	80	170,76	39,65	4,43	161,94	179,59	89	250

No quadro 40 está apresentado o resultado do teste de Levene de homogeneidade da variância. Como o valor da Sig. =0,992, superior ao α = 0,05, podemos assumir que existe homogeneidade da variância e prosseguir com o cálculo da ANOVA.

Quadro 40 – Resultados dos testes de Levene

Levene Statistic	df1	df2	Sig.
,033	3	76	,992

O quadro 41 refere-se à ANOVA propriamente dita. Os resultados mostram que a variação entre grupos (Between Groups) é maior que a variação dentro dos grupos (Within Groups). O valor da estatística do teste 27,246, sendo a significância associada de 0,000. Estes resultados permitem concluir que existe evidência empírica para considerar as diferenças significativas (**$F_{(3;76)}$= 27,246;p=,000).**

Quadro 41 – Resultados da aplicação da ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	64353,438	3	21451,146	27,246	,000
Within Groups	59835,050	76	787,303		
Total	124188,488	79			

Dado que a ANOVA é um teste *omnibus*, há que verificar onde se situam as diferenças, ou seja, entre que grupos as diferenças são estatisticamente significativas (ao nível de 0,05). Para tal vamos recorrer aos procedimentos *post hoc*. Os comandos situam-se dentro dos comandos da ANOVA, bastando seleccionar o botão «Post Hoc…» (quadro38). Da panóplia de testes existentes, vamos optar pelo Tukey's-b. Os resultados são-nos apresentados no quadro 42

Quadro 42 – Resultados do teste *post hoc* Tukey's-b aplicado às médias do tempo gasto em minutos em função da profissão

Profissão	N	Subset for alpha = .05 1	2	3
Psicólogos	20	127,30	---	---
Enfermeiros	20	---	165,85	---
Assistentes Sociais	20	---	---	186,75
Médicos	20	---	---	203,15

Como pode observar pelo quadro, são os Psicólogos que gastam em média menos tempo semanal em exercício físico ($\overline{x}$=127,30), que diferente significativamente ao nível de 0,05 de todos os grupos profissionais. De seguida quem gasta em média mais tempo são os Enfermeiros ($\overline{x}$=165,85), que também se distinguem de modo significativo dos restantes grupos profissionais. Apesar dos Assistentes Sociais ($\overline{x}$=186,75) e os Médicos ($\overline{x}$=203,15) serem os que apresentam valores médios superiores, não se distinguindo entre si (p>0,05), eles distinguem-se de modo significativo dos restantes grupos.

Os resultados devem ser apresentados conforme o quadro 43.

Quadro 43 - Resultados da aplicação do ANOVA de um critério.
Variável independente (Factor): profissão; variável dependente: tempo gasto (em minutos)

Grupos:	N	$\overline{x}$	s	F	p
Médicos	20	203,15	26,46	27,246	,000
Enfermeiros	20	165,85	29,29		
Assistentes Sociais	20	186,75	29,56		
Psicólogos	20	127,30	26,78		

6.5.3. Exercício

1. No sentido de testar a hipótese de que "a distância social está relacionada com o contacto/familiaridade com as doenças mentais", compararam-se três grupos de indivíduos em função do seu nível de contacto (baixo contacto; moderado contacto; elevado contacto). As medidas estatísticas e os resultados dos testes calculados com auxílio do SPSS são apresentados seguidamente:

Estatísticas resumo da distância social

Nível de contacto	n	Mean	Std. Deviation
Baixo contacto	142	2,994	,905
Elevado contacto	31	3,503	,953
Moderado contacto	42	3,224	1,020
Total	215	3,113	,948

Tests of Normality

	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Distância	,062	215	,051	,989	215	,099

a Lilliefors Significance Correction

Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
,411	2	212	,664

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	7,235	2	3,617	4,139	,017
Within Groups	185,281	212	,874		
Total	192,516	214			

Multiple Comparisons -LSD

(I) Contacto	(J) Contacto	Mean Difference (I-J)	Sig.
Baixo contacto	Elevado contacto	-,5089*	,007
	Moderado contacto	-,2294	,164
Elevado	Baixo contacto	,5089	,007
	Moderado contacto	,2794	,208
Moderado contacto	Baixo contacto	,2294	,164
	Elevado contacto	-,2794	,208

a) – Face aos testes utilizados, tome uma decisão (α=0,05) face à hipótese formulada apresentando correctamente a decisão e utilizando a simbologia adequada.

6.6. Coeficiente de correlação r de Pearson e teste de significância da correlação

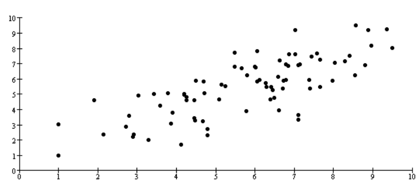
Muitas vezes recorremos a uma medida estatística designada de coeficiente de correlação r de Pearson. Esta medida varia entre $\pm$ 1,00, passando por 0 (ausência de relação). A correlação mede pois a presença, direcção e magnitude da relação entre duas variáveis (x; y) e aplica-se às situações em que ambas as variáveis são medidas a um nível intervalar ou racional e a relação entre as variáveis é linear.

Como podemos interpretar o valor do coeficiente de correlação? Quando é que o seu valor é considerado elevado (forte), moderado ou baixo? O sinal dá-nos a direcção. Se for positivo, as variáveis covariam no mesmo sentido, se for negativo covariam em sentidos opostos. Como guia pode utilizar os valores apresentados seguidamente (Pestana & Gajeiro, 2002), no entanto são apenas uma referência[35], pois há que ter em conta que a interpretação do r é ampla e deve ter sempre em conta, a natureza das variáveis em estudo; a significância do coeficiente e a finalidade do cálculo do r.

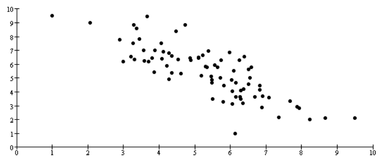
Valores do r		Interpretação do valor encontrado
r de 0,00 a ± 0,20	→	Indica relação muito baixa
r de ± 0,20 a ± 0,40	→	Indica relação baixa, presente mas ligeira
r de ± 0,40 a ± 0,70	→	Indica relação moderada ou acentuada
r de ± 0,70 a ± 0,89	→	Indica relação forte
r de ± 0,90 a ± 1,00	→	Indica relação muito forte ou perfeita ± (1,00)

[35] Trata-se mesmo de uma referência, pois alguns autores consideram em termos absolutos que um valor de r=0,25 é fraco, 0,5 é moderado e 0,8 é forte, ou seja, depende da natureza do estudo.

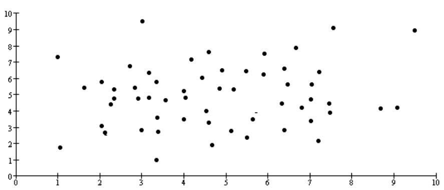
Seguidamente estão apresentados os diagramas de dispersão que implicam diferentes tipos de relação e inclusive a ausência de relação. Como deverá recordar-se, os diagramas de dispersão (Scatterplots no SPSS) permitem dar resposta, entre outras, às seguintes questões: a) as variáveis estão de algum modo relacionadas? b) a eventual relação é linear ou de tipo não linear? c) as mudanças de variação de y dependem de x?; d) existem outliers?



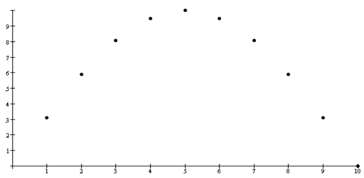
Relação linear positiva



Relação linear negativa



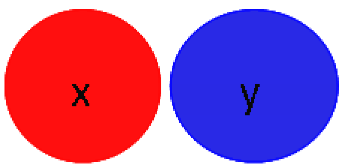
Ausência de relação



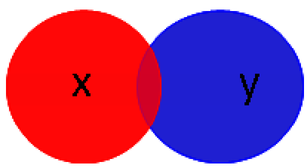
Relação não linear

O coeficiente de correlação tem associado a si um outro coeficiente, o *coeficiente de determinação* (r^2) e que nos dá a percentagem de variação explicada (r^2x100) de uma variável que é devida à outra. A percentagem de variação não explicada é dada por $1-r^2$

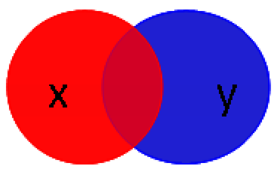
Se um valor do r for 0,8 e outro 0,4, tal não significa que a primeira associação seja o dobro da segunda. A comparação dos dois coeficientes deve ser feita em termos do coeficiente de determinação, que indica a percentagem de variação de uma variável explicada pela outra. Assim, na primeira relação 64% da variação de uma variável é explicada por outra e vice-versa, valor quatro vezes superior ao da segunda (16%). É de salientar que r como o r^2 não exprimem relações de causalidade, porque a associação de x com y é igual à associação de y com x. Observe os diferentes diagramas de Venn representam diferentes valores do r^2.



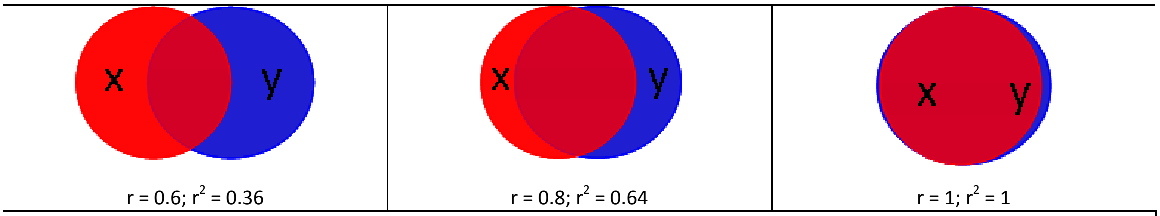
$r = 0$; $r^2 = 0$

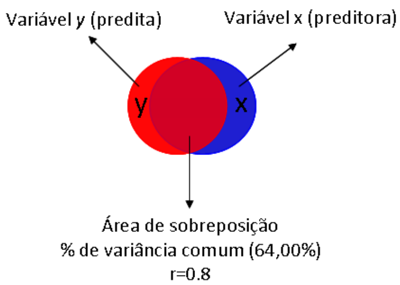


$r = 0.2$; $r^2 = 0.04$



$r = 0.4$; $r^2 = 0.16$





A medida (r) é calculada através da seguinte fórmula:

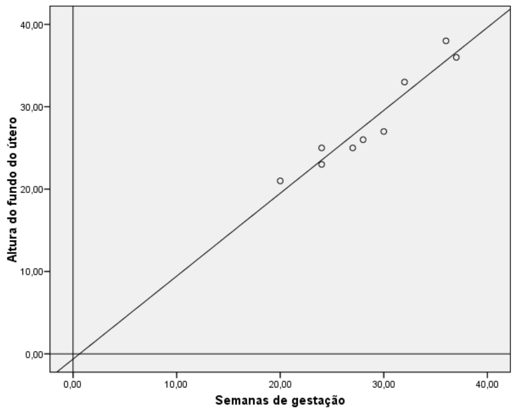
$$r = \frac{S_{xy}}{S_x S_y} = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\Sigma x)^2][n\sum y^2 - (\Sigma y)^2]}}$$

Numa amostra de mulheres grávidas (n=9) foi realizado um estudo com objectivo de estudar a relação positiva entre as *semanas de gestação* (S.G.) e a *altura do fundo do útero* (A.F.U.). Os dados colhidos são apresentados seguidamente:

S. G. (x)	A.F.U (y)
20	21
24	23
24	25
27	25
28	26
30	27
32	33
36	38
37	36

O valor da medida calculada através da máquina revelou um valor de r=0,96; r^2=0,92 (r^2%=91,52%). Com o auxílio do SPSS realizou-se o diagrama de dispersão onde foi traçada a recta de regressão (y=a+b_x) que apresentamos de seguida.

Gráfico 1 – Diagrama de dispersão da altura do fundo do útero
em função das semanas de gestação



Face aos valores encontrados, podemos verificar que estamos perante uma correlação positiva, muitíssimo forte e linear entre as variáveis, isto é, à medida que aumentam as semanas de gestão, aumenta o fundo do colo do útero das mulheres, no entanto esta relação é ou não estatisticamente significativa? Para o saber deveremos realizar o teste de significância da correlação de Pearson que é calculado pela fórmula:

$$t \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

gl=n-2

O teste vai dizer-nos se a relação se deve ao acaso amostral ou se existe de facto na população. Trata-se de uma amostra da população (n=9 mulheres), e deveremos pensar na possibilidade da correlação no universo ser zero, daí que a hipótese nula seja H_0:$\rho \leq 0$, enquanto a hipótese de investigação é H_1:$\rho > 0$. O teste, para α=0,05, vai dizer-nos se a relação é estatisticamente significativa e neste caso rejeita-se H_0, ou se não o é, e a declaramos como não significativa, isto é, não rejeito H_0.

Temos outros casos em que as hipóteses são direccionadas, por exemplo, «existe uma relação negativa entre o número de dias em coma e a capacidade motora dos doentes». Neste caso a região de rejeição situa-se apenas numa das caudas (da esquerda) e simbolicamente teremos H_0:$\rho \geq 0$ e H_1:$\rho < 0$. Mas também existem casos em que as hipóteses não são direccionadas, por exemplo quando pretendemos testar a hipótese de que «a crença acerca da

Incurabilidade das doenças mentais está relacionada com a crença na perigosidade dos doentes». As hipóteses seriam H_0:ρ=0 e H_1:ρ≠0.

Voltando ao exercício apresentado, o teste de significância da correlação vai ser realizado com recurso à distribuição *t*.

$$t \frac{0,96}{\sqrt{\frac{1-0,92}{7}}} = 8,98$$

gl=9-2

Fazendo recurso à tabela da distribuição *t* (excerto), dado que o teste é unicaudal e o α=0,05, com 7 graus de liberdade, o valor do *t* teórico é de *t*=1,90. Assim comparando os valores, podemos verificar que o *t* observado cai na região de rejeição da H_0. Então poderemos afirmar que estamos perante uma correlação positiva, muitíssimo forte, linear e estatisticamente significativa (r=0,96 (r^2=0,92); p<0,05) entre as variáveis.

Tabela Distribuição *t* (excerto)

	α para teste unilateral			
	0,05	**0,025**	**0,01**	**0,005**
	α para teste bilateral			
Graus de Liberdade	**0,10**	**0,05**	**0.02**	**0,01**
1	6.31	12.71	31.82	63.66
.	.	.	.	.
6	1.94	2.45	3.14	3.71
7	**1.90**	2.37	3.00	3.45
8	1.86	2.31	2.90	3.36

O cálculo do coeficiente assume que os sujeitos foram independente e aleatoriamente seleccionados da população de interesse e que ambas as variáveis medidas seguem distribuição normal bivariada (ambas as variáveis). Finalmente um último pressuposto a *homoscedasticidade*, isto é, para cada valor de x a variação dos valores de y devem ser sensivelmente os mesmos, e vice-versa. Normalmente considera-se que quando as amostras têm um tamanho razoável, o não cumprimento destes últimos pressupostos, tem «pouco» efeito no valor do teste estatístico.

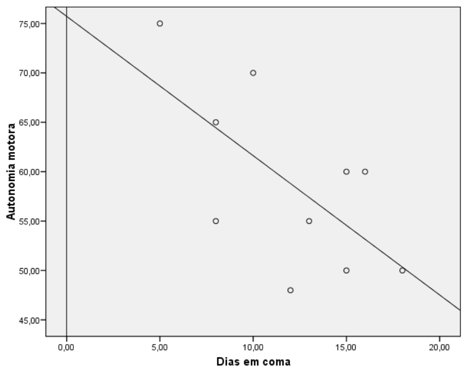
6.6.1. R de Pearson e teste de significância da correlação na máquina de calcular

Vejamos outro exemplo que permite utilizar a máquina de calcular. Num estudo pretendeu-se testar a seguinte hipótese: «há uma relação negativa entre o número de dias em coma e a capacidade motora dos doentes». Os dados colhidos junto de uma amostra de 10 doentes são apresentados seguidamente.

Dias de coma (x)	Autonomia motora (y)
5	75
8	65
8	55
10	70
16	60
12	48
15	50
13	55
18	50
15	60

Apresentamos primeiramente o diagrama de dispersão das variáveis.

Gráfico 2 – Diagrama de dispersão da autonomia funcional em função dos dias de coma dos doentes



A hipótese vai ser testada com recurso à máquina. O nível de significância (α) será de 0,05.

Instruções para cálculo do teste *t* as médias de dois grupos independentes	
TEXAS	CASIO
1. Pressione o botão **STAT** e depois mude o cursor para a opção **EDIT,** seleccionado a opção **1: Edit…** e **ENTER**; 2. Introduza os valores das variáveis nas listas L1 e L2; 3. Pressione novamente o botão **STAT** e mude o cursor para **TESTS,** seleccionando a opção **E:LinRegTTest…** seguido de **ENTER.** 4. Substitua os valores nas opções do **LinRegTTest** **Xlist: L1** (assegure que é a variável x) **Ylist: L2** (assegure que é a variável y) **Freq:1** **β & ρ:** ≠0 **<u><0</u>** >0 (seleccione a simbologia da hipótese, neste caso, **<u><0</u>**) RegEQ: Calculate e ENTER	1. Pressione o botão **MENU** e depois seleccione a opção **STAT**, **EXE**; 2. Introduza os valores de x e y na list1 e List2; 3. Seleccione a opção **TEST** (tecla F3); 4. Seleccione a opção **t** (tecla F2); 5. Seleccione a opção **REG** (tecla F3); 6. Substitua os valores nas opções: **β & ρ:** colocar a simbologia da hipótese (**β & ρ:<0)** **Xlist: List1** **Ylist: List 2** **Freq:1** **Execute:**

O output de resultados mostra que estamos perante uma correlação negativa, forte, linear e estatisticamente significativa (r=-0,65; r^2=0,43; p=0,021).

Output (Resultados) das máquinas: LinearReg tTest t = -2.4328 p =0.02 df = 8 a = 75.723 a = -1.41 s = 7.24 r = -0.65 r^2 = 0.43

É de salientar que quando não existem os valores para colocar nas listas, apenas o valor do r e o tamanho da amostra, deverá fazer uso da fórmula para cálculo da significância da correlação. Pode experimentar a partir do exercício seguinte.

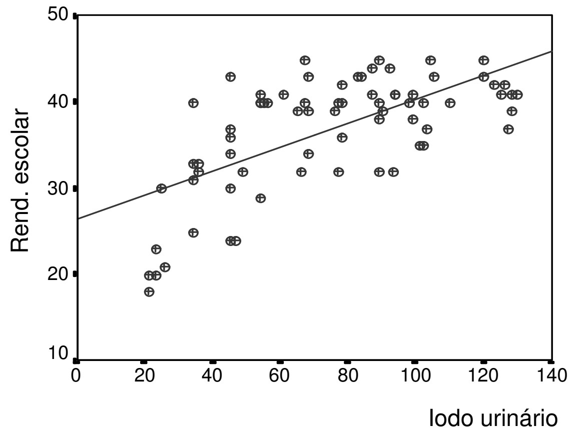
1. Numa amostra de 78 crianças que frequentam a escola do Bairro dos Plasmas, realizou-se ainda um estudo da correlação entre a <u>*excreção urinária de iodo*</u> (em µgr/gr de creatinina) e o <u>*rendimento escolar*</u> medido através de uma escala própria (mínimo 15 pontos e máximo 45 pontos). Obtiveram-se os seguintes resultados, incluindo o diagrama de dispersão efectuado:

r = 0.63

b = 0.14

a = 26.40

GRÁFICO 3 - Diagrama de dispersão do rendimento escolar das crianças da Escola C+S do Souto (1997) com hipertrofia da tiróide em função dos seus valores de excreção urinária de iodo (μgr/gr de creatinina)



a) Tendo em conta as estatísticas o diagrama de dispersão, como classifica a relação entre a *excreção urinária de iodo* e a *aprendizagem escolar* das crianças na amostra estudada?

b) Considerando a hipótese: "*existe correlação entre excreção urinária de iodo e a aprendizagem escolar das crianças*", pode afirmar-se que a correlação observada é significativa para $\alpha = 0.05$? O que é que a sua conclusão significa em termos de inferência estatística?

6.6.2. R de Pearson e teste de significância da correlação no SPSS

Para efectuar o cálculo do coeficiente de correlação de Pearson no SPSS deverá utilizar a base «correlação_Pearson». Trata de um estudo realizado junto de uma amostra de 217 adultos (idade≥ 18 anos) e em que a hipótese de investigação refere que «a crença na incurabilidade das doenças mentais está relacionada com a crença na perigosidade dos doentes mentais». Utilize um erro tipo I máximo de 0,05.

O 1.º passo consiste na verificação do pressuposto da normalidade para ambas as variáveis (bistribuição normal bivariada). Os comandos são apresentados no quadro 44.

Quadro 44 - comandos para realização dos testes de normalidade (K-S e S-W)

Comando do SPSS (acção):	**Comentário:**
1. Seleccione do menu as opções: **Analyze** → **Descriptive Statistics** → **Explore**	Acede à caixa de diálogo do Explore que permite realizar os testes de normalidade.
2. Na caixa de diálogo Explore, seleccione as variáveis incurabilidade e perigosidade (com um clique sobre as variáveis) e depois com o botão ▶ coloque-as na **Dependent List:**	O **Dependent List:** é o espaço onde se colocam as variáveis que procuramos testar.
3. Dê um clique sobre o botão **Plots…**	Acede a nova caixa de diálogo, **Explore Plots**
4. Na caixa **Explore Plots** seleccione com um clique sobre o quadrado a opção **Normality plots with tests** [☒]	Selecciona automaticamente os testes de normalidade, assim como os gráficos «Normal Q-Q Plot» e «Detrended Normal Q-Q Plot».
5. Clique em **Continue** e posteriormente no botão **OK**	Vai gerar um output de resultados dos testes

Os resultados do teste de Kolmogorov-Smirnov são apresentados no quadro 45. Lembre que se optou por este teste dado o tamanho da amostra ser >50. Para ambas as variáveis o valor da significância do teste é >0,05, o que permite concluir que as duas variáveis seguem distribuição normal na população.

Quadro 45 – Resultados dos testes de Normalidade

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
incurabilidade	,055	217	,200[*]	,989	217	,111
perigosidade	,053	217	,200[*]	,993	217	,420

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

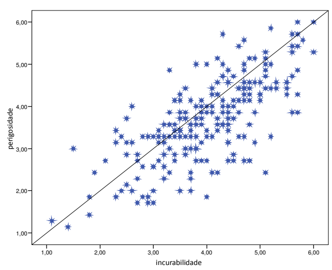
Os comandos descritos no quadro 46 permitem-lhe criar o diagrama de dispersão.

Quadro 46 - comandos para realização o diagrama de dispersão

Comando do SPSS (acção):	Comentário:
1. Seleccione do menu as opções: **Graphs** → **Legacy Dialogs** → **Scatter/Dot…** → Simple Scatter → **Define**	Acede aos gráficos, especificamente os Scatterplots.
2. Na caixa de diálogo **Simple Scatter Plot** seleccione a variável *incurabilidade* e depois com o botão ▶ coloque-as X-axis. Faça o mesmo procedimento para a variável perigosidade, mas colocando-a no Y-axis	Esta caixa de diálogo permite-lhe colocar as variáveis que vão ser utilizadas na construção do diagrama.
3. Dê um clique sobre o botão **OK**	Constrói o gráfico que aparece no output.
4. Na output aparece-lhe o gráfico. Se quiser que o programa trace a recta, de um clique duplo sobre o gráfico e acede ao **Chart Editor.** Na barra de ferramentas dê um clique sobre o ícone . Automaticamente vai criar uma nova caixa de diálogo designada de **Properties.** Dê um clique sobre o botão **Close. A** recta encontra-se no gráfico	Selecciona automaticamente os testes de Permite-lhe efectuar alterações no gráfico.
5. Regresse ao Output para visionar o gráfico.	

O gráfico, designado de diagrama de dispersão é apresentado seguidamente. Como pode observar, a distribuição da nuvem de pontos indicia uma relação linear entre as variáveis, como pode observar.

Gráfico 4 – Diagrama de dispersão da crença na incurabilidade em função da crença na perigosidade



O passo seguinte consiste no cálculo do coeficiente de correlação e respectivo teste de significância. Utilize os comandos a seguir:

Quadro 47 – Comandos para o calculo do coeficiente de correlação de Pearson e teste de significância

Comando do SPSS (acção):	Comentário:
1. Seleccione do menu as opções: **Analyze** → **Correlate** → **Bivariate…**	Permite-lhe aceder à caixa de diálogo do cálculo do coeficiente de correlação e respectivo teste de significância
2. Acede à caixa de diálogo **Bivariate Correlations**. Na caixa de diálogo seleccione ambas as variáveis com o cursor e coloque-as com o botão ▶ na lista **Variable(s):**	**Variables:** é o espaço onde se colocam a variáveis que pretendemos correlacionar e sobre as quais é também realizado o teste de significância. Por defeito está seleccionado o coeficiente de **Pearson**, mas tem outros como o **Kendall's tau-b** e o coeficiente de **Spearman**. Por defeito o teste é bicaudal (Two-Tailed), mas se quiser pode seleccionar One-Tailed.
3. Clique sobre o botão OK	Vai gerar um output de resultados do teste

Os resultados do valor da medida estatística e do teste de significância são apresentados no quadro 48. Como pode verificar o valor do coeficiente de correlação é r=0.710 e o nível de significância associado ao teste é de sig=,000, ou seja, muito inferior ao erro máximo assumido no início do estudo (α=0,05). Estamos pois perante uma correlação positiva, linear, forte e estatisticamente significativa. Se tomar atenção, pode observar que junto ao valor do coeficiente aparecem dois asteriscos (,710**) que remetem para uma nota no rodapé da tabela [*** Correlation is significant at the 0.01 level (2-tailed)*]. Isto significa que o próprio programa considera que a correlação é estatisticamente significativa, ao nível de 0,01.

Quadro 48 - Resultados do cálculo do r de Pearson e respectivo teste de significância aos scores das crenças na perigosidade e incurabilidade (n=217)

		incurabilidade	perigosidade
incurabilidade	Pearson Correlation	1	,710(**)
	Sig. (2-tailed)	---	,000
	N	217	217
perigosidade	Pearson Correlation	,710(**)	1
	Sig. (2-tailed)	,000	---
	N	217	217

** Correlation is significant at the 0.01 level (2-tailed).

A forma de apresentar os resultados relativos a esta hipótese, quer em quadro, quer do ponto de vista da linguagem considerada mais adequada e com validade científica, são apresentados de seguida.

Quadro 49 - Matriz de correlação bivariada de Pearson entre as crenças na incurabilidade e perigosidade dos doentes e doenças mentais (n = 217)

	Crença na incurabilidade
Crença na perigosidade	r = ,710 (r^2 = 0,5041) p = ,000

Procedeu-se ao cálculo do coeficiente de correlação de Pearson e respectivo teste significância para testar a hipótese de que «a crença na incurabilidade das doenças mentais está relacionada com a crença na perigosidade dos doentes mentais». Encontramos uma correlação positiva, forte, linear e estatisticamente significativa (r=0,710; p<0,05) entre as variáveis. A percentagem de variação explicada (r^2) é de 51,41%, valor que se considera bom face à natureza do estudo permite observar que estas crenças ou estereótipos, relacionadas entre si, podem ser um factor importante em termos de discriminação social dos doentes mentais. Em média, são os indivíduos que crêem que as doenças psiquiátricas têm um cunho de incurabilidade, que apresentam valores médios mais elevados na crença da perigosidade e imprevisibilidade desses doentes.

6.7. Teste z de diferença de proporções

Apresentamos anteriormente (ponto 6.1.1.) um teste de diferença de médias para dois grupos independentes, no entanto também podemos comparar duas proporções populacionais através do teste Z de diferença de proporções.

O teste é realizado através dos cálculos indicados pela fórmula:

$$z = \frac{p_1 - p_2}{\sqrt{pq\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

Sendo que $p = \frac{x_1 + x_2}{n_1 + n_2}$

Vejamos um exemplo. No Serviço de Medicina do Hospital dos Anjos dos 214 doentes internados no ano transacto, 27 desenvolveram úlceras de pressão, enquanto nos 361 doentes internados no mesmo período no Serviço de Medicina da Clínica do Beato, desenvolveram úlceras de pressão 29. Teste a hipótese de que "a prevalência de úlceras de pressão no Hospital dos Anjos é diferente da prevalência de úlceras na Clínica do Beato (a=0,05).

Simbolicamente H_0: $p_1=p_2$ e H_1: $p_1 \neq p_2$. A hipótese não é direccionada logo a região de rejeição é bilateral.

Assim para no Serviço de Medicina do Hospital dos Anjos, o valor do p_1 é dado por 27/214, ou seja, p=0,13, enquanto no Serviço de Medicina da Clínica do Beato o valor do p_2 corresponde a 29/361, isto é, p_2=0,08.

Substituindo os valores na formula teremos:

$$z = \frac{0{,}13 - 0{,}08}{\sqrt{(0{,}10x0{,}90)\left[\frac{1}{214} + \frac{1}{361}\right]}} = 1{,}787$$

Sendo que $p = \frac{27+29}{214+361} = 0{,}097$

Dado que o valor do z teórico ou z da tabela é de 1,96, não existe evidência empírica que suporte que a prevalência de úlceras de pressão seja diferente consoante o hospital (z=1,7919; p>0,05).

6.7. Teste z de diferença de proporções na máquina de calcular

Vamos apresentar o cálculo do teste na máquina a partir do exercício anterior.

Instruções para cálculo do teste *z* de diferença de proporções	
TEXAS	CASIO
1. Pressione o botão **STAT** e depois mude o cursor para a opção **TESTS**; 2. Seleccione a opção **6: 2- PropZTest…** seguido de **ENTER;** 3. Introduza os valores: X1: **27 (casos)** **N1: 214** X2: **29 (casos)** **N2: 214** **$p_1 \neq p_2$** 4. **Calculate** e **ENTER**.	1. Pressione o botão **MENU** e depois seleccione a opção **STAT**, **EXE**; 2. Seleccione a opção **TEST** (tecla F3); 3. Seleccione a opção **z** (tecla F2); 4. Seleccione a opção **2-P** (tecla F4) 5. Introduza a hipótese de investigação [**$p_1 \neq p_2$**] 6. Introduza os valores: X1: **27 (casos)** **N1: 214** X2: **29 (casos)** **N2: 214** 8. **EXECUTE** seguido de **EXE**.
Output (Resultados): **$p_1 \neq p_2$** z=1.7919 p=0.073143 $\hat{p}_1 = .126$ $\hat{p}_2 = .080$ $\hat{p} = .097$	

7. Teste U de Mann-Witnhey

Um outro tipo de testes a que podemos recorrer no âmbito da estatística inferencial são os testes não paramétricos, como referimos no ponto 6.1. Muitas vezes assistimos a alguma resistência dado o facto deste tipo de testes apresentarem menor poder, no entanto eles começam progressivamente a tornar-se numa ferramenta preciosa de trabalho na investigação no âmbito da Enfermagem e são muitas vezes os mais adequados.

O primeiro dos testes que apresentamos designa-se por teste U de Mann-Witnhey. É o correspondente não paramétrico do teste *t* de student para amostras independentes (Siegel & Castellan, 1988). Aplica-se nos casos em que temos dois grupos independentes e o nível de medida da variável «dependente» e ordinal ou superior (intervalar e racional). É a alternativa adequada quando não se cumprem determinados pressupostos como a normalidade das distribuições e ainda quando as amostras são de reduzida dimensão.

Este, como a maioria dos testes não paramétricos trabalha com o que podemos considerar postos, posições ou mesmo ordens, na tradução de *rank* do inglês. Os valores da variável, sejam scores ou de outro tipo, são ordenados segundo um critério, por exemplo do menor para o maior (ordem crescente ou ascendente).

Vejamos um exemplo simples e didáctico. Num estudo recente procurou-se avaliar o nível de distância social de adolescentes (rapazes e raparigas) relativamente a uma vinheta que descrevia um caso de um rapaz com 17 anos, a quem tinha sido diagnosticada uma esquizofrenia. Especificamente, o investigador procurava verificar se existia maior ou menor distância social relativamente *à capacidade de uma amizade estável* com alguém como a pessoa descrita na vinheta. A opção de resposta variava entre 1 (sem qualquer problema) a 10 (de modo nenhum).

A hipótese de investigação referia que «a distância social relativamente às pessoas a quem foi diagnosticada esquizofrenia, difere de rapazes para raparigas». Os dados relativos ao estudo são apresentados de seguida.

Rapazes (♂) - A	Raparigas (♀) - B
10	1
7	3
2	4
5	9
8	6

O 1.º passo consiste em ordenar as pontuações atribuídas pelos 10 adolescentes do menor para o maior valor.

Pontuações:	1	2	3	4	5	6	7	8	9	10
Postos/posições	1	2	3	4	5	6	7	8	9	10
Grupo:	♀	♂	♀	♀	♂	♀	♂	♂	♀	♂

O 2.º passo consiste em encontrar o somatório dos postos para cada um dos grupos. No exemplo o somatório dos postos do grupo $A_♂$ (R_1) =2+5+7+8+10=32 e o Somatório dos postos do grupo $B_♀$ (R_2) = 1+3+4+6+9 =23

Se dividirmos os somatórios dos postos pelo número de casos teremos a média dos postos ou posições (em inglês *mean rank*). No grupo A (♂) a média dos postos é **6,40** o que equivale a 32/5, enquanto no grupo B (♀) será de **4,60** e que corresponde a 23/5.

Como pode verificar $R_1+R_2 = N(N+1)/2$, ou seja, 55. Caso a hipótese nula seja verdadeira esperaríamos obter sensivelmente uma média de postos igual em ambos os grupos, caso que não encontramos aqui. Agora resta saber se a diferença vai para lá do acaso amostral e nesse caso rejeitamos a H_0, ou pelo contrário, declaramos que não existe evidência empírica para rejeitar H_0.

O 3.º passo consiste no cálculo do valor U a utilizar como U observado para comparar com a distribuição teórica. O seu valor é dado pelo cálculo das fórmulas:

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

$$U_2 = n_2 n_1 + \left[\frac{n_2(n_2+1)}{2}\right] - R_2$$

Substituindo os símbolos nas fórmulas e efectuados os cálculos temos:

$$U_1 = (5)(5) + \left[\frac{5(6)}{2}\right] - 32$$

$$U_1 = 8{,}00$$

$$U_2 = (5)(5) + \left[\frac{5(6)}{2}\right] - 23$$

$$U_2 = 17{,}00$$

Para que possamos declarar a diferença como estatisticamente significativa, isto é, rejeitar H_0, deveremos utilizar como $U_{(observado)}$ o menor dos U´s observados, neste caso U_1=8,00. Na comparação com o $U_{(teórico)}$), para declarar a diferença como estatisticamente significativa, o valor do $U_{(observado)}$ deverá ser menor ou igual ao $U_{(teórico)}$. Neste caso e pela leitura da tabela 4, verificamos que o valor do $U_{(teórico)}$ é de 2 (para n_1=5 e n_2=5), ou seja, as diferenças encontradas não são estatisticamente significativas.

Tabela 4 - Valores críticos da Estatística U para α = 0,05 (Teste bicaudal)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
2	---	---	---	---	---	---	---	0	0	0	0	1	1	1	1	1	2	2	2	2
3	---	---	---	---	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	---	---	---	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5	---	---	0	1	**2**	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	---	---	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	---	---	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	---	0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	---	0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10	---	0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	---	0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	---	1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	---	1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	---	1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
15	---	1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	---	1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	---	2	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18	---	2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	---	2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	---	2	8	13	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

(Fonte: Siegel & Castellan, 1988

No caso em que uma das amostras ou ambas sejam > 20 poderemos fazer uma aproximação à distribuição normal padronizada. O teste é realizado através dos cálculos indicados pela fórmula:

$$z = \frac{U_o - \frac{n_A n_B}{2}}{\sqrt{\frac{(n_A)(n_B)(n_A + n_B + 1)}{12}}}$$

7.1. Teste U de Mann-Witnhey no SPSS

Para utilizar o SPSS na realização do teste U deverá utilizar a base «prestígio.sav». A hipótese de investigação refere que «o prestígio profissional atribuído aos enfermeiros é diferente consoante a comunidade tenha ou não cuidados de Enfermagem no domicílio». A comunidade A tem cuidados de Enfermagem ao domicílio e a comunidade B tem.

H_0: As duas populações são iguais em tendência central.

H_1: As duas populações não são iguais em tendência central.

Quadro 50 – comandos para realização do teste *t* para grupos independentes e teste de Levene

Comando do SPSS (acção):	Comentário:
1. Seleccione do menu as opções: **Analyze** → **Nonparametric tests** → **2 Independents Samples…**	Permite-lhe aceder à caixa de diálogo do teste
2. Na caixa de diálogo (**Two-Independent-Samples Tests**) seleccione a variável *presti* (com um clique sobre a variável) e depois com o botão ▶ coloque-a na lista do **Test Variable List:**	O **Test Variable List:** é o espaço onde se colocam as variáveis dependentes.
3. Seleccione a variável *comunidade* com um clique e depois com o botão ▶ coloque-a no **Grouping Variable:**	O **Grouping Variable** é o espaço onde se coloca a variáveis «independentes», isto é os grupos.
4. Dê um clique sobre **Define Groups**.	Surge a caixa de definição de grupos onde se inserem os códigos de cada grupo. Repare que aparece como: **comunidade ? ?**
5. Na área do **Group 1:** introduza o código da comunidade A (1), e no **Group 2:** o código da comunidade B (2).	Esta operação identifica os grupos que estão a ser comparados e aos quais atribuímos diferentes números. Por defeito o SPSS tem o teste U seleccionado.
6. Clique em **Continue** e posteriormente no botão **OK**	Vai gerar um output de resultados do teste

Os resultados são apresentados nos quadros 51 e 52. Relativamente ao quadro 52 pode observar-se que as médias dos postos ou das posições são respectivamente 43,23 (1297,00/30) na comunidade A e de 20,50 (656/32) na comunidade B. Logo daqui poderá depreender que os scores mais elevados na escala de prestígio são atribuídos pela comunidade A, ainda que tal não signifique que esta diferença seja estatisticamente significativa.

Quadro 51 - Média dos postos/posições e somatório dos postos para cada grupo

	comunidade	N	Mean Rank	Sum of Ranks
presti	comunidade A	30	43,23	1297,00
	Comunidade B	32	20,50	656,00
	Total	62		

O quadro 52 mostra os valores do U (128,00), W de Wilcoxon (656,00) e o z observado (da aproximação à normal), assim como o valor da significância [Asymp. Sig. (2-tailed)] que é respectivamente 0,000**.** Como se pode observar, os resultados da significância do teste mostram para um erro tipo I de 0,05, que as duas distribuições diferem em tendência central (sig=0,000). Poderemos concluir que as diferenças encontradas no prestígio profissional atribuído aos enfermeiros é diferente consoante a comunidade.

Quadro 52 – Resultados do teste U Mann-Whitney

	presti
Mann-Whitney U	128,000
Wilcoxon W	656,000
Z	-4,995
Asymp. Sig. (2-tailed)	,000

a Grouping Variable: comunidade

A apresentação dos resultados do teste U de Mann-Whitney em quadros é muitas vezes feita de modo indistinto, apresentando as médias dos postos, os somatórios dos postos, o valor do U e depois a significância obtida. Uma forma simples e correcta é a que se apresenta de seguida (quadro 53), mas para isso deverá calcular as medianas para cada grupo.

Quadro 53 - Resultado da aplicação do teste U Mann-Whitney
Variável independente: comunidade; variável dependente: scores da escala prestigio

Comunidade:	n	Mediana	Média Postos	*z*	p
Comunidade A	30	7,00	43,23	-4,995	,000
Comunidade B	32	3,00	20,50		

No sentido de testar a hipótese de que "o prestígio profissional atribuído aos enfermeiros é diferente consoante a comunidade tenha ou não cuidados de Enfermagem no domicílio", procedeu-se à realização do U de Mann-Whitney.

Os resultados mostram que o nível mediano do prestígio atribuído aos profissionais de enfermagem é na comunidade (A) de 7,00 pontos (média dos postos = 43,23), enquanto na comunidade B a mediana observada foi de 3,00 pontos (média dos postos = 20,50), o que permite observar que o nível mediano de prestigio atribuído aos profissionais de enfermagem é maior na comunidade A. Estas diferenças encontradas entre os grupos revelaram-se estatisticamente significativas (z=-4,995;p=,000).

7. 2. Exercício

1. Considerando os valores apresentados de seguida relativos à glicemia em jejum (mg/dl) de 2 grupos de sujeitos, teste a hipótese (α=0,05) de que "a população da qual provém o Grupo A tem valores de glicemia em jejum diferentes do grupo B".

Grupo A	B
77	81
125	87
86	98
96	72
110	115
100	85
94	76
88	
117	

8. Teste do quiquadrado (χ^2) – tabelas 2x2 e LxC

Em muitas circunstâncias na investigação em Enfermagem pretendemos testar hipóteses que colocam em relação variáveis qualitativas, sejam de nível nominal e/ou ordinal. Por exemplo poderemos querer verificar se as crenças religiosas estão relacionadas com a atitude perante as dádivas de sangue dos dadores, se determinado tipo de doença está relacionado com determinado estilo de vida, se por exemplo a frequência de um curso de preparação para o parto está relacionado com o tipo de parto, etc…

Estes testes também podem ser aplicados para amostras emparelhadas (ex. teste McNemar para avaliar a mudança de opinião/tendência), isto é, podemos investigar se a opinião relativamente ao consumo de tabaco (prejudicial/não prejudicial) muda depois da frequência de um curso.

O teste do quiquadrado (χ^2) indica a concordância ou pelo contrário, a discrepância das frequências observadas (F_o) com as frequências esperadas (F_e) obtidas pelo cálculo de probabilidades. Uma aplicação útil do (χ^2) nas tabelas de contingência reside na possibilidade que nos oferece de comparar frequências observadas (F_o) com as frequências esperadas (F_e), com o intuito de analisar a relação (associação) entre duas variáveis, isto é, se as variáveis estão associadas ou se pelo contrário, elas são independentes.

Se χ^2=0, verifica-se coincidência entre as frequências observadas e esperadas (teóricas), se o χ^2>0, verifica-se uma discrepância entre as frequências, mas devemos comparar os valores do $\chi^2_{(observado)}$ com o valor do $\chi^2_{(teórico)}$, de modo a ver se rejeitamos ou não a H_0.

Simbolicamente

H_o: $p_1=p_2=p_3=p_4$, isto é as variáveis são independentes:

H_1: $p_1 \neq p_2 \neq \ldots p_4$, as variáveis estão relacionadas.

Os testes que vamos apresentar aplicam-se a duas situações, para tabelas de 2x2, ou seja, uma variável com duas categorias em linha e uma variável com duas categorias em coluna (equivale a 4 células), e também os casos das tabelas LxC (em inglês RxC) em que pelo menos uma das variáveis (em linha ou coluna) tem três categorias, por exemplo 3x2 (cinco células).

Vejamos como se realiza o teste do χ^2 a partir de um exemplo aplicado. No quadro estão os dados referentes a um estudo em que se procurava testar a hipótese de que «existe relação entre a suplementação alimentar precoce com vitamina D e a erupção dos incisivos inferiores (até aos 6 meses ou após os 6 meses de idade)». A amostra é constituída por 110 indivíduos.

Os dados do quadro 54 referem-se a um estudo cuja hipótese de investigação referia que «existe relação entre a suplementação alimentar precoce com vitamina D e a erupção dos incisivos inferiores (até aos 6 meses ou após os 6 meses de idade)». A amostra é constituída por 110 crianças.

Quadro 54 – Dados relativos à erupção dos incisivos em função da suplementação alimentar precoce com vitamina D

Suplemento Vit. D: / Erupção dos incisivos:	Sim	Não	Total
Até aos 6 meses	36	14	50
Depois dos 6 meses	32	28	60
Total	68	42	110

As frequências observadas (F_o) estão nas células, agora resta-nos calcular as frequências esperadas (F_e), isto é aquelas que seriam de esperar caso a hipótese nula seja verdadeira, ou seja, para que não exista relação entre a variável em coluna e em linha (as variáveis são independentes).

As frequências esperadas por célula (F_e) são dadas por:

$$Fe = \frac{(\text{total em linha})\,x\,(\text{total em coluna})}{\text{total de observações}}$$

Isto é, o cálculo das frequências esperadas por célula é resultado da razão do produto dos totais marginais da categoria em linha e em coluna pelo total das observações. Assim relativamente à 1.ª célula do quadro 54, cuja frequência observada é 36, a frequência esperada respectiva será: $Fe = \frac{(50)\,x\,(68)}{110} = 30{,}9$

No quadro 55 estão calculadas as frequências esperadas por célula.

Quadro 55 – Cálculo das frequências esperadas para os dados relativos à erupção dos incisivos em função da suplementação alimentar precoce com vitamina D

Suplemento Vit. D: / Erupção dos incisivos:	**Sim**	**Não**	**Total**
Até aos 6 meses	E_1 = (50x68)/110 = **30,9**	E_2 = (60x68)/110 = **19,1**	50
Depois dos 6 meses	E_3 = (60x68)/110 = **37,1**	E_4= (60x42)/110 = **22,9**	60
Total	68	42	110

O grau de discrepância entre as frequências, observadas e esperadas (cálculo do $\chi^2_{(observado)}$), é-nos dado pela fórmula:

$$\chi^2 = \Sigma \frac{(F_o - F_e)^2}{F_e}$$

Assim teremos:

$$\chi^2 = \Sigma \frac{(36-30,9)^2}{30,9} + \frac{(14-19,1)^2}{19,1} + \frac{(32-31,7)^2}{31,7} + \frac{(28-22,9)^2}{22,9} = 4,026$$

O valor 4,026 corresponde ao $\chi_{(observado)}$, que deve ser comparado com um ($\chi^2_{teórico}$). Para isso deveremos ter presente o erro tipo I máximo, neste caso será α=0,05, e os graus de liberdade que correspondem a (C-1)(L-1), isto é n.º de categorias em linha subtraído 1 e número de categorias em coluna, também subtraído 1. Logo teremos um grau de liberdade [(2-1)(2-1)]. Resta pois comparar observar o valor do ($\chi^2_{teórico}$) que aparece na tabela 4 (distribuição do χ^2) com o $\chi_{(observado)}$.

Como pode verificar, o valor do ($\chi^2_{teórico}$) = 3.84146, e o valor do o $\chi_{(observado)}$=4,026. Este valor é superior significa que $\chi_{(observado)}$ «cai» na região de rejeição da H_0, ou seja, a evidência empírica permite-nos afirmar que há uma relação significativa entre a suplementação alimentar precoce com vitamina D e a erupção dos incisivos inferiores (até aos 6 meses ou após os 6 meses de idade), isto é $\chi^2_{(1)}$ = 4,026; p<0,05).

Tabela 4 – Distribuição do χ^2

df	α			
	.050	.025	.010	.005
1	3.84146	5.02389	6.63490	7.87944
2	5.99146	7.37776	9.21034	10.59663
3	7.81473	9.34840	11.34487	12.83816
4	9.48773	11.14329	13.27670	14.86026
5	11.07050	12.83250	15.08627	16.74960
6	12.59159	14.44938	16.81189	18.54758
7	14.06714	16.01276	18.47531	20.27774
8	15.50731	17.53455	20.09024	21.95495
9	16.91898	19.02277	21.66599	23.58935
10	18.30704	20.48318	23.20925	25.18818
11	19.67514	21.92005	24.72497	26.75685
12	21.02607	23.33666	26.21697	28.29952
13	22.36203	24.73560	27.68825	29.81947
14	23.68479	26.11895	29.14124	31.31935
15	24.99579	27.48839	30.57791	32.80132
16	26.29623	28.84535	31.99993	34.26719
17	27.58711	30.19101	33.40866	35.71847
18	28.86930	31.52638	34.80531	37.15645
19	30.14353	32.85233	36.19087	38.58226
20	31.41043	34.16961	37.56623	39.99685
21	32.67057	35.47888	38.93217	41.40106
22	33.92444	36.78071	40.28936	42.79565
23	35.17246	38.07563	41.63840	44.18128
24	36.41503	39.36408	42.97982	45.55851
25	37.65248	40.64647	44.31410	46.92789
26	38.88514	41.92317	45.64168	48.28988
27	40.11327	43.19451	46.96294	49.64492
28	41.33714	44.46079	48.27824	50.99338
29	42.55697	45.72229	49.58788	52.33562
30	43.77297	46.97924	50.89218	53.67196

Caso a relação seja declarada como estatisticamente significativa, deveremos, apenas nesse caso, indicar o sentido da relação. Uma forma simples é comparar as frequências observadas e esperadas, isto é, através dos resíduos (R), que correspondem a F_o-F_e, em cada célula. No exemplo anterior poderemos optar pelos resíduos cuja diferença tiver sinal positivo para encontrar as células que indicam o sentido da relação. No programa SPSS o método utilizado é diferente, mas será explicado adiante.

O teste do qui-quadrado tem pressupostos para ser utilizado. Deverá ter sempre presente que apenas pode realizar o teste quando os pressupostos estão cumpridos, embora existam formas de «fazer cumprir» os pressupostos, nomeadamente agregando categorias. A agregação de categorias que não tem em conta a lógica e natureza das variáveis em estudo, assim como a temática, pode revelar-se uma asneira. Devemos aprender com os erros dos outros como referem Hill & Hill (2000).

8.1. Critérios

Os critérios são conhecidos como critérios de Cochran.

Para tabelas de 2x2, temos:

❶ Utilizar o teste exacto de Fisher se n<20;

❷ Se n>40 utilizar χ^2_c[36] ou n>>40 utilizar χ^2;

❸ Se 20 ≤ n ≤40 utilizar χ^2_c (se nenhuma frequência esperada inferior a 5)

Para tabelas LxC devemos utilizar o χ^2 (se no **máximo 20%** das células com frequências esperadas inferiores a cinco e **nenhuma** inferior a 1).

Vamos agora efectuar o cálculo do quiquadrado na máquina. Não sabemos se é do seu conhecimento, mas vários estudos tem evidenciado a elevada prevalência de lombalgias nas crianças em idade escolar e tem-se colocado a hipótese de que esse fenómeno se relaciona com o modo habitual de transportar o material escolar. Para verificar essa hipótese, realizou-se um estudo amostral, cujos dados se apresentam no quadro 56.

[36] Correcção de continuidade. O valor do $\chi^2_{corrigido}$ ou correcção de continuidade de YATES é utilizado algumas vezes no cálculo do χ^2 nas tabelas 2x2. Algumas vezes é utilizado quando as frequências esperadas por célula são inferiores a 10 (Jaccard & Becker, 1990).
A fórmula de cálculo com correcção de Yates:

$$\chi^2_c = \Sigma \frac{(|F_o - F_e| - 0{,}5)^2}{F_e}.$$

A utilização deste teste é «controversa» pelo facto de ser mais conservador que o teste do χ^2 sem correcção.

Quadro 56 – Dados relativos à prevalência de lombalgia em função
Do modo habitual de transporte do material escolar

	Modo habitual de transporte			
	Mochila com rodas	Mochila nos dois ombros	Outro	Total
C/ lombalgias	12 (18,46%)	19 (20,21%)	34 (29,31%)	65 (23,64%)
S/ lombalgias	53 (81,54%)	75 (79,79%)	82 (70,69%)	210 (76,33%)
Total	65 (100,00%)	94 (100,00%)	116 (100,00%)	275 (100,00%)

Nota: entre parêntesis as percentagens em coluna

8.2. Teste do quiquadrado (χ^2) – tabelas 2x2 e LxC na máquina calcular

Vamos testar a hipótese recorrendo então à máquina de calcular cujas instruções aparecem de seguida.

Instruções para cálculo do χ^2 (tabelas 2x2 e LxC)	
TEXAS	CASIO
1. Pressione primeiro a tecla **2nd** e depois o tecla **x^{-1} (**opção **MATRX).** 2. Avance para o comando **EDIT** e seleccione a opção **1:[A]** seguindo de **ENTER.** 3. Coloque o número de linhas e colunas, neste caso será: **MATRIX [A] 2 x 3 (duas linhas e três colunas)** **Insira os valores.** **[12 19 34]** **[53 75 82]** 4. Pressione a tecla **STAT** e depois seleccione a opção **TESTS.** 5. Com o cursor seleccione a opção **C: χ^2-Test**… seguido de **ENTER.** Vai aparecer a seguinte linha: **χ^2-Test** **Observed: [A]** **Expected: [B]** **Calculate Draw** **6.** Seleccione a opção **Calculate** e depois tecla **ENTER** * Para saber o valor das frequências esperadas volte ao comando MATRIX, seleccione a opção [B] seguido de ENTER.	1. Seleccione com o cursor a opção MAT seguido da tecla EXE: Aparece a seguinte linha de comandos: **Matrix** **Mat A : None** Deverá na opção **none** introduzir **2 x 3** (duas linhas e três colunas). Depois pressione a tecla EXE e introduza os valores na matriz **A** **[12 19 34]** **[53 75 82]** 2. Depois de introduzidos os dados seleccione a tecla **MENU** e seleccione a opção **STATS**. 3. Seleccione a opção **TEST** (tecla F3); 4. Seleccione a opção **CHI** (tecla F3) Vai aparecer a seguinte linha: **χ^2-Test** **Observed: Mat A** **Execute** 5. Seleccione a opção Execute e depois a tecla EXE. * Na CASIO Para saber o valor das frequências esperadas volte ao comando MATRIX, e depois seleccione pressionando a tecla SHIFT seguida da tecla **(-)** que tem a opção **ANS** Vai encontrar a linha Mat Ans : 2x 3 Pressione a tecla EXE.
Output (Resultados): χ^2=3,6438 p=0.16171 df=2	

Como pode verificar pela aplicação do teste χ^2, não existe evidência suficiente para sustentar a hipótese de relação entre as duas variáveis ($\chi^2_{(2)}$= 3,6438; p=0,16171).

8.3. Teste do quiquadrado (χ^2) – tabelas 2x2 e LxC no SPSS

O calculo do χ^2 no SPSS é também ele simples. Utilize a base «**quiquadrado.sav**». A base de dados deve ser utilizada para testar a hipótese de que «há relação entre a opinião relativamente a co-incineração (favor e contra) e o facto das pessoas viverem em proximidade ou distanciadas geograficamente de uma co-incineradora». O erro tipo máximo será de 0,05.

Os comandos para a realização do teste são apresentados no quadro 57.

Quadro 57 – comandos para realização do teste *t* para grupos independentes e teste de Levene

Comando do SPSS (acção):	Comentário:
1. Seleccione do menu as opções: **Analyze** → **Descriptive Statistics** → **Crosstabs…**	Permite-lhe aceder à caixa de diálogo Crosstabs onde se realiza o teste
2. Na caixa de diálogo (**Crosstabs**) seleccione a variável opinião (com um clique sobre a variável) e depois com o botão ► coloque-a no campo da **Row(s):**. Deverá proceder da mesma forma para a variável residen, mas colocando-a no campo **Column(s).**	O **Crosstabs** é o espaço onde se colocam as variáveis que submetemos a teste
3. Dê um clique sobre o botão **Statistics** e depois seleccione a opção **Chi-square.** Clique sobre o botão **Continue.**	Permite seleccionar o teste do quiquadrado.
4. Acede à caixa de diálogo: **Crosstabs Cell Display.** Dê um clique sobre o botão **Cells…** e depois no campo **Residuals** seleccione a opção **Adjusted Standardized**	Permite seleccionar diferentes campos, como tipo de percentagens (linha e/ou coluna), resíduos, frequência esperadas, etc… Os resíduos ajustados estandardizados permitem uma leitura eficiente do sentido da relação, caso esta seja estatisticamente significativa como é o caso
6. Clique em **Continue** e posteriormente no botão **OK**	Vai gerar um output de resultados do teste

Os resultados do output aparecem nos quadros seguintes. O primeiro quadro (58) apresenta distribuição das frequências observadas, esperadas e inclui os resíduos ajustados estandardizados.

Quadro 58 – Distribuição das opiniões dos participantes em função do seu local de residência

Opinião		Residência		Total
		Próxima	Distante	
Favor	Count	167	205	372
	Expected Count	205,1	166,9	372,0
	Adjusted Residual	-5,7	5,7	
Contra	Count	230	118	348
	Expected Count	191,9	156,1	348,0
	Adjusted Residual	5,7	-5,7	
Total	Count	397	323	720
	Expected Count	397,0	323,0	720,0

O 2.º quadro (59) apresenta os resultados do teste do qui-quadrado. Repare na informação que o programa dá em rodapé do quadro: «b: 0 cells (,0%) have expected count less than 5. The minimum expected count is 156,12». Esta informação torna-se útil e essencial para verificar os critérios para a realização do teste.

Quadro 59 – Resultados do teste do qui-quadrado

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	32,667(b)	1	,000

b 0 cells (,0%) have expected count less than 5. The minimum expected count is 156,12.

Dado que os resultados do teste do qui-quadrado mostram uma relação estatisticamente significativa **($\chi^2_{(1)}$= 32,667;p=,000)** entre as variáveis, resta situar essas diferenças ou seja, analisar o sentido da relação. Para isso poderemos recorrer aos resíduos ajustados estandardizados. Neste caso as células cujo valor seja > 1,96, são aquelas que indicam o sentido da relação. Neste caso situam-se na 2.ª e 3.ª células. Caso não tenha os resíduos ajustados estandardizados[37], calcule os resíduos «simples», isto é, a diferença entre observadas e esperadas por células, mas atenção só faz sentido se a relação entre as variáveis for significativa, isto é, no caso em que se rejeita a hipótese nula, caso contrário não fará nenhum sentido.

Em muitas circunstâncias encontramos nas revistas e jornais científicos aplicações do teste do quiquadrado, apresentando relações significativas, mas não se encontra qualquer informação sobre o sentido da relação (Haberman, 1973).

Caso não queira ser «inocente» nas suas conclusões, especificamente quando encontra relações estatisticamente significativas, pode servir-se de dois coeficientes que são medidas de associação entre as variáveis, nomeadamente o phi (ϕ) aplica-se a dado nominais nas tabelas de 2x2 e é dado pela formula: $\phi = \sqrt{\frac{\chi^2}{N}}$. Ou então o V de Cramer que se aplica a tabelas LxC e é-nos dado pela fórmula $V = \sqrt{\frac{\chi^2}{N(k-1)}}$. Se o valor da medida for <0,4, estamos perante associações baixas entre as variáveis (Liebetrau (1983).

[37] Trata-se da diferença entre as frequências observadas e esperadas por células, divididas por uma estimativa do seu erro padrão. Os resultados dos resíduos estandardizados são expressos em unidades de desvio padrão, abaixo ou acima da média. É um meio precioso para a leitura das tabelas.

8.4. Exercícios

1. De modo a testar a hipótese de que «existe relação entre a frequência ou não de um programa de preparação psicoprofilática para o parto e a ocorrência de depressão pós-parto», realizou-se um estudo do qual se colheram os seguintes dados:

P. Psicoprofilática: / Depressão pós parto:	Sim	Não
Não	35	45
Sim	2	11

a) Teste a hipótese (alpha=0,05).

2. De modo a testar a hipótese de que «existe relação entre a frequência ou não de um programa de preparação psicoprofilática para o parto e a avaliação da experiência de parto», realizou-se um estudo do qual se colheram os seguintes dados:

P. Psicoprofilática: / Avaliação do parto:	Sim	Não
Experiência positiva	47	59
Experiência negativa	21	53

a) Realize o respectivo teste e indique a sua conclusão face aos resultados (com um nível de significância de 5%).

3. No sentido de verificar se «as atitudes perante os doentes mentais estão relacionadas com as representações que as pessoas têm das doenças», colheram-se dados junto de uma amostra de 834 residentes em Penacova, tendo-se obtido os seguintes resultados:

Representações: / Atitudes:	Representação Criminalizada	Representação Médica	Total
Atitude tolerante	---	534	534
Atitude de indiferença	145	---	145
Atitude defensiva	155	---	155
Total	300	534	834

a) Tome uma decisão face à hipótese formulada (α=0,05) apresentando correctamente a decisão e utilizando a simbologia adequada. Apresente o sentido da relação se for caso disso.

9. Síntese….

Leia com atenção:

1.º - Quando se refere que algo é *estatisticamente significativo* tal não quer dizer que seja *clinicamente significativo*. Provavelmente esse efeito ou diferença não é "nulo" (na estatística), mas pode ser clinicamente irrelevante, como o contrário também é verdadeiro. A rejeição da hipótese nula, recorrendo por exemplo ao teste *t,* pode evidenciar que existe um efeito significativo e esse efeito pode ser também clinicamente significativo, no entanto amostras grandes (com pequena variação) podem conduzir a um problema, declarando a diferença estatisticamente significativa quando ela não tem qualquer significado clínico. Por exemplo, um valor de correlação r = .15 (n=1100) com p = < .05 é de facto estatisticamente significativo*,* mas no contexto de uma investigação pode tornar-se ilusório o seu poder explicativo;

2.º - Dois estudos sobre a mesma temática, mas em amostras diferentes produzem valores de *p value significativos* (*p=.05 e p=.001*). Podemos supor que as diferenças são maiores no segundo estudo? Se as amostras forem iguais é bem possível que sim, mas um valor *p* = .05 num estudo com amostras de reduzida dimensão, pode reflectir um efeito maior que um p = .001 num estudo que recorreu a uma amostra de grande dimensão;

3.º - Por exemplo, o *burnout* é estudado em duas amostras de enfermeiros com a mesma dimensão e características semelhantes, numa o valor encontrado foi de p=.04 e noutra p = .06. Isso não significa que apenas no primeiro estudo é que existem diferenças e efeito significativos. O efeito varia de amostra para amostra, assim como o *p value* encontrado, logo a utilização de um nível de significância de 0,05 não deve ser assumido como um axioma;

4.º - Quando num estudo se refere que uma diferença não é significativa, isso não significa que as médias sejam iguais ou que não exista um *efeito substantivo*, significa sim que não houve evidência suficiente forte para *provar* que essas diferenças eram significativas, se se falhou poderá ser pelo facto do tamanho da amostra ser inadequado;

5.º - Se um investigador encontra resultados estatisticamente significativos com uma amostra de 30 indivíduos, isso não significa que uma amostra de 30 sujeitos seja suficiente para planear um estudo. Aleatoriedade e representatividade são essenciais. O valor de 30 indivíduos como amostra, considerado uma «amostra grande», ou ainda para fugir à validação dos pressupostos dos testes paramétricos, são dois mitos. Uma amostra grande com 30 indivíduos provavelmente não será uma grande amostra. Depende da natureza da temática sob estudo.

6.º - Assumir erradamente que o valor de *p* estabelece a validade da hipótese de investigação. O valor de *p* (se <.05) não estabelece a probabilidade da hipótese de investigação ser verdadeira, a sua função é determinar e existe uma "evidência suficiente" para rejeitar H_0.

7.º. Se assumimos inicialmente um valor para o alpha de .05, ele deve manter-se, e nesse contexto o *p value* e os respectivos asteriscos gerados pelo computador não são um *substituto*. Se assumimos como critério para a rejeição de H_0 um α =.05, não devemos andar para tras e substituir o α para .01, só porque o *p* obtido pelo computador foi de .002.