

ÍNDICE GERAL

I – INTRODUÇÃO.....	5
1. NOÇÕES GERAIS	6
2. POPULAÇÃO E AMOSTRA.....	9
3. MÉTODOS DE AMOSTRAGEM.....	11
3.1 AMOSTRAGENS PROBABILÍSTICAS E NÃO-PROBABILÍSTICAS	11
3.1.1 As Amostras Probabilísticas	11
3.1.1.1 TIPOS DE AMOSTRAGENS PROBABILÍSTICAS.....	12
3.1.1.2 TIPOS DE AMOSTRAGENS NÃO PROBABILÍSTICAS:.....	14
3.2 DETERMINAÇÃO DO TAMANHO DA AMOSTRA.....	16
3.2.1 Cálculo do Tamanho da Amostra para Populações Infinitas (>100.000 elementos).....	16
3.2.2 Cálculo do Tamanho da Amostra para Populações Finitas (<100.000 elementos).....	18
3.2.3 Determinação da Margem de Erro da Amostra.....	20
3.2.4 Determinação da Amostra sem conhecer os limites da população.....	20
3.3 INDIVÍDUO OU UNIDADE ESTATÍSTICA.....	21
3.4 VARIÁVEL ESTATÍSTICA.....	22
3.5 ESTATÍSTICA DESCRITIVA E ESTATÍSTICA INDUTIVA.....	24
3.6 PARÂMETRO E DADO ESTATÍSTICO.....	24
3.7 REPRESENTAÇÃO DE UMA VARIÁVEL ESTATÍSTICA	25
3.8 REDUÇÃO DE UMA VARIÁVEL ESTATÍSTICA.....	29
3.8.1 Média	30
3.8.2 Mediana	33
3.8.3 Moda	34
3.8.4 Quantis.....	38
3.8.5 Medidas de dispersão	40
3.8.6 Coeficiente de dispersão	49
3.9 AS CARACTERÍSTICAS DA DISTRIBUIÇÃO NORMAL.....	51
3.9.1 A curva normal e os desvios-padrão	53
4. ESTATÍSTICA PARAMÉTRICA E NÃO PARAMÉTRICA	59
5. TESTES NÃO PARAMÉTRICOS PASSO-A-PASSO	63
5.1 TESTE DO QUI-QUADRADO.....	63
5.2 PROCEDIMENTOS PARA ORDENAÇÃO DE RESULTADOS.....	67
5.3 TESTES PARA DUAS AMOSTRAS INDEPENDENTES.....	69
5.4 TESTE U DE MANN-WHITNEY.....	70
6. TESTES PARA DUAS AMOSTRAS RELACIONADAS	74
6.1 PROVA DE McNEMAR PARA A SIGNIFICÂNCIA DE MUDANÇAS.....	75

6.2	<i>CORRECÇÃO DE CONTINUIDADE</i>	78
6.3	<i>TESTE DOS SINAIS DE WILCOXON</i>	79
7.	TESTES PARA K AMOSTRAS INDEPENDENTES	82
7.1	<i>TESTE DE KRUSKAL – WALLIS</i>	82
8.	TESTES PARA K AMOSTRAS RELACIONADAS	87
8.1	<i>PROVA DE COCHRAN</i>	88
8.2	<i>TESTE DE FRIEDMAN</i>	92
9.	MEDIDAS DE CORRELAÇÃO E SUAS PROVAS DE SIGNIFICÂNCIA	97
9.1	<i>COEFICIENTE DE CORRELAÇÃO RHO DE SPEARMAN – RANK</i>	97
9.2	<i>O COEFICIENTE DE CONCORDÂNCIA DE KENDALL</i>	99
10.	TESTES PARAMÉTRICOS PASSO-A-PASSO	106
10.1	<i>CORRELAÇÃO</i>	106
10.2	<i>TESTE T DE STUDENT (NÃO RELACIONADO)</i>	111
10.3	<i>TESTE T DE STUDENT (RELACIONADO)</i>	115
10.4	<i>ANÁLISE DE VARIÂNCIA</i>	119
10.4.1	Análise da Variância com Igual Tamanho	119
10.4.1.1	TESTE DE TUKEY PARA COMPARAÇÃO ENTRE AS MÉDIAS	123
10.4.2	ANÁLISE DE VARIÂNCIA COM DIFERENTES TAMANHOS	125
10.4.2.1	TESTE DE TUKEY PARA COMPARAÇÃO ENTRE AS MÉDIAS	127
	ANEXOS	129
	<i>Anexo I</i>	130
	<i>Anexo II</i>	131
	<i>Anexo III</i>	132
	<i>Anexo IV</i>	133
	<i>Anexo V</i>	134
	<i>Anexo VI</i>	135
	<i>Anexo VII</i>	136
	<i>Anexo VIII</i>	137
	<i>Anexo IX</i>	138
	<i>Anexo X</i>	139

ÍNDICE DE ILUSTRAÇÕES

FIGURA 1: POPULAÇÃO E AMOSTRA.....	10
FIGURA 2: AMOSTRA ESTRATIFICADA.....	13
FIGURA 3: CONVERSÃO DOS NÍVEIS DE CONFIANÇA EM DESVIOS PADRÃO	17
QUADRO 1: QUADRO GERAL DE SERIAÇÃO	25
QUADRO 2: QUADRO DE FREQUÊNCIAS.....	26
QUADRO 3: QUADRO DE FREQUÊNCIAS.....	27
QUADRO 4: QUADRO DE FREQUÊNCIAS.....	28
QUADRO 5: QUADRO DO CÁLCULO DA MÉDIA	30
QUADRO 6: QUADRO DO CÁLCULO DA MÉDIA	31
QUADRO 7: CÁLCULO DA MÉDIA PONDERADA	32
FIGURA 4: CURVA SIMÉTRICA.....	35
FIGURA 5: CURVA ASSIMÉTRICA À DIREITA	36
FIGURA 6: CURVA ASSIMÉTRICA À ESQUERDA	36
FIGURA 7: CURVA ASSIMÉTRICA À DIREITA	39
FIGURA 8: CURVA SIMÉTRICA ACHATADA (PLATOCURTICA)	42
QUADRO 8: CÁLCULO DO DESVIO MÉDIO	44
QUADRO 9: CÁLCULO DO DESVIO MÉDIO PARA CLASSES	45
QUADRO 10: CÁLCULO DA VARIÂNCIA.....	45
QUADRO 11: CÁLCULO DA VARIÂNCIA.....	46
FIGURA 9: DISTRIBUIÇÃO NORMAL	51
FIGURA 10 - DISTRIBUIÇÃO LEPTOCÚRTICA E PLATICÚRTICA.....	51
FIGURA 11: % DE RESULTADOS 1 DP ACIMA E ABAIXO DA MÉDIA	54
FIGURA 12: POSIÇÃO DE UM RESULTADO.....	55
QUADRO 12: TABELA DE CONTINGÊNCIA (CROSSTAB).....	65
QUADRO 13: ORDENAÇÃO DE UM GRUPO EM FUNÇÃO DO N.º DE FACTORES DE RISCO	67
QUADRO 14: ORDENAÇÃO PARA AMOSTRAS RELACIONADAS OU EMPARELHADAS	69
QUADRO 15: RESULTADOS DO REGISTO DAS QUEIXAS	71
QUADRO 16: TESTE MCNEMAR.....	75
QUADRO 17: TABELA DE QUATRO CASAS PARA A PROVA DE SIGNIFICÂNCIA DE MUDANÇAS	77
QUADRO 18: RESULTADOS DO TESTE DO VOCABULÁRIO.....	79
QUADRO 19: NÚMERO DE IDEIAS RELEMBRADAS PARA TRÊS TIPOS DE TESTES	83
QUADRO 20: Q DE COCHRAN	89
QUADRO 21: Q DE COCHRAN	90
QUADRO 22: AVALIAÇÃO DE TRÊS TIPOS DE ILUSTRAÇÕES	93
QUADRO 23: COEFICIENTE DE CONTINGÊNCIA DE RHO DE SPEARMAN-RANK	99
QUADRO 24: POSTOS ATRIBUÍDOS A 6 CANDIDATOS A EMPREGO POR 3 CHEFES DE PESSOAL.....	100
FIGURA 13: DIAGRAMA DE DISPERSÃO DE PONTOS OU SCATTERPLOT OU SCATTERGRAM	107
QUADRO 25: CORRELAÇÃO DE PEARSON	110
QUADRO 26: TESTE T NÃO RELACIONADO	114
QUADRO 27: TESTE T RELACIONADO	117
QUADRO 28: DADOS DE 4 AMOSTRAS E RESPECTIVAS MÉDIAS	118

QUADRO 29: NOTAÇÃO PARA ANÁLISE DA VARIÂNCIA	119
QUADRO 30: RECIDIVAS POR TRATAMENTO.....	121
QUADRO 31: ANOVA	123
QUADRO 32: POST-HOC DAS RECIDIVAS	124
QUADRO 33: MÉDIAS POR TRATAMENTO	125
QUADRO 34: ANOVA	127
QUADRO 35: POST-HOC	128

I – INTRODUÇÃO

Desde séculos o homem tem, muitas vezes, tomado notas de coisas e de pessoas, não com o único fim de acumular números, mas com a esperança de utilizar os dados do passado para a resolução de problemas do presente assim como para a previsão de acontecimentos futuros. No entanto, o sucesso quanto a este objectivo só foi possível em data muito recente: só no final do século XIX e, sobretudo, no princípio do século XX é que, com a aplicação de probabilidades aos problemas sobre a interpretação dos dados recolhidos, foi possível resolver alguns deles.

A Estatística conquistou, hoje, o seu lugar entre as ciências. O poder do seu método é, sobretudo, afirmado nas últimas décadas e aplica-se, agora, nos domínios mais variados. Até aqui, só um pequeno número de pessoas se preocupou com estudos estatísticos, quer pela natureza das suas investigações, quer por causa da sua utilidade para as diferentes profissões. O valor e a importância do método estatístico residem no esforço para melhor compreender o nosso mundo, tão maravilhosamente complexo, tanto no ponto de vista físico como social, levam-nos a sonhar que ele se torne objecto de um conhecimento como as outras ciências. A vida corrente leva-nos a decisões para passar do conhecido ao desconhecido, da experiência à previsão.

Este curso tem por fim fornecer conhecimentos estatísticos (sem ter muitos conhecimentos matemáticos) e saber interpretar resultados que podem ser obtidos através de programas de computador.

1. NOÇÕES GERAIS

Para algumas pessoas, a Estatística não é senão um quadro de colunas mais ou menos longas de números que dizem respeito à população, à indústria ou ao comércio, como se vê frequentemente em revistas; para outras, ela dá gráficos mostrando a variação no tempo de um facto económico ou social, a produção ou os números relativos aos negócios de uma empresa, assim como se encontra nos escritórios de empresas privadas.

Tão diferenciados se apresentam os métodos estatísticos que não é possível estabelecer uma definição que os contenha a todos. Apesar disso, apresentamos a seguir uma definição que, embora necessariamente incompleta como qualquer outra, tem a vantagem de introduzir o aluno na matéria.

A Estatística tem como finalidade elaborar de uma síntese numérica que evidencie o que de mais generalizado e significativo exista num conjunto numeroso de observações.

O grande número de observações de que se parte reflecte uma diversidade tal que se torna ininteligível a sua interpretação. Para que, a partir dessa diversidade se possa começar a entender logo, torna-se necessário reduzir sucessivamente as observações, ganhando-se em generalidade o que se vai perdendo em individualidade.

A síntese implica, assim, que nos desprendamos do que é particular e individual para nos atermos ao que existe de mais geral no conjunto das observações; à medida que a síntese progride, vai-se perdendo o contacto com as particularidades imediatas.

Deste modo, a Estatística não se ocupa do que é excepcional, mas apenas do que é geral: não se interessa pelo indivíduo, mas por grupos de indivíduos; não se ocupa, em suma, de uma só medição, mas de um conjunto de medições.

Acrescente-se, ainda, que a síntese é numérica. Quer isto dizer que se prescinde inteiramente das palavras e dos recursos literários de mais ou menos efeito que elas possibilitam. Alcança-se a síntese pelo recurso exclusivo dos números.

Daí o afã com que frequentemente se escolhem os números de acordo com os argumentos. A Estatística é intrinsecamente uma disciplina não literária, manipula exclusivamente números e alcança a síntese ordenando-os e cooperando com eles.

“Estatística”, deriva de “status” que em latim significa Estado, e que só por si demonstra a ligação que sempre existiu entre ambos;

O primeiro levantamento estatístico remonta a 3050 a.C., no Egipto, tendo como objectivo informar o estado sobre recursos humanos e económicos.

No séc. XVII d.C., a disciplina de Estatística era já leccionada nas universidades alemãs, continuando com a finalidade de descrever as populações e as riquezas do Estado.

Ainda no séc. XVII, dá-se a expansão dos seus campos de investigação a áreas como:

- ❖ Saúde pública;
- ❖ Indústria;
- ❖ Comércio;
- ❖ Estudos Demográficos.

Os métodos de inferência estatística surgem com Jonh Graunt (1620-1674), um modesto comerciante, que tira conclusões válidas sobre uma população desconhecida por ele.

Fermat (1601-1665) e Pascal (1623-1662) permitem que o estudo do acaso tome uma expressão matemática, introduzindo o Cálculo das Probabilidades.

O Cálculo das Probabilidades e o aparecimento do Método dos mínimos quadrados, vêm credibilizar a Estatística conferindo-lhe a fundamentação matemática em que ela assenta hoje.

No séc. XVIII Lambert Quetelet (1796-1874) introduziu a Estatística em áreas como:

- ❖ Meteorologia;
- ❖ Antropometria;
- ❖ Ciências Sociais;
- ❖ Economia;
- ❖ Biologia.

Francis Galton (1822-1911), apresenta noções de regressão e correlação;

Karl Pearson (1857-1936) apresenta a mais bela e acabada teoria de Estatística, ficando também conhecido pelos seus coeficientes;

Fisher e os seus trabalhos sobre inferência Estatística deram também um grande contributo ao desenvolvimento da Estatística.

Em 1943, dá-se uma grande reviravolta, uma vez que o tratamento de dados deixa de ser feito manualmente e passa a ser computadorizado.

Método Estatístico:

- ❖ “Técnica usada para obter, analisar e apresentar dados numéricos”;
- ❖ “Método científico cuja finalidade é o estudo das propriedades numéricas dos conjuntos dos factos”.

O Método Estatístico, segundo a teoria de Cramer, pressupõe as seguintes fases:

- ❖ Recolha de dados estatísticos: obtenção da amostra a partir da população, devendo depurar e rectificar os dados estatísticos, que no seu conjunto são denominados série estatística.
- ❖ Descrição: conjunto de operações, numéricas ou gráficas, efectuadas sobre os dados estatísticos determinando a sua distribuição; procede-se à sua ordenação, codificação e representação por meio de quadros e tabelas.
- ❖ Análise: consiste em tirar conclusões sobre a distribuição da população, determinar o seu grau de confiança e ainda formular hipóteses, tentando verificá-las, quanto ao fenómeno em estudo.
- ❖ Predição: é uma previsão do comportamento do fenómeno em estudo, tendo em conta a definição da distribuição estatística.

2. POPULAÇÃO E AMOSTRA

População: somatório dos indivíduos ou elementos, com qualquer característica comum e que estão sujeitos a uma análise estatística, por terem interesse para o estudo.

Quanto à sua origem pode ser:

- Conjunto de pessoas;
- Conjunto de objectos;
- Conjunto de acontecimentos.

Quanto à sua natureza pode ser:

- Existente ou real;
- Hipotética ou não totalmente existente;

Pode ainda ser:

- Conjunto finito;
- Conjunto infinito.

Amostra: é um subconjunto retirado da população, que se supõe ser representativo de todas as características da mesma, sobre o qual será feito o estudo, com o objectivo de serem tiradas conclusões válidas sobre a população.

Amostragem: é o procedimento pelo qual um grupo de pessoas ou um subconjunto de uma população é escolhido com vista a obter informações relacionadas com um fenómeno, e de tal forma que a população inteira nos interessa esteja representada

O Plano de Amostragem serve para descrever a estratégia a utilizar para seleccionar a amostra.

Este plano fornece os detalhes sobre a forma de proceder relativamente à utilização de um método de amostragem para determinado estudo.

Logo que o investigador delimite a população potencial para o estudo, ele deve precisar os critérios de selecção dos seus elementos.

Certos critérios dizem respeito às características requeridas para que um elemento ou sujeito faça parte da amostra:

- A amplitude da idade dos participantes;
- Condições de saúde precisas.

Igualmente certos critérios definem os sujeitos que não farão parte do estudo:

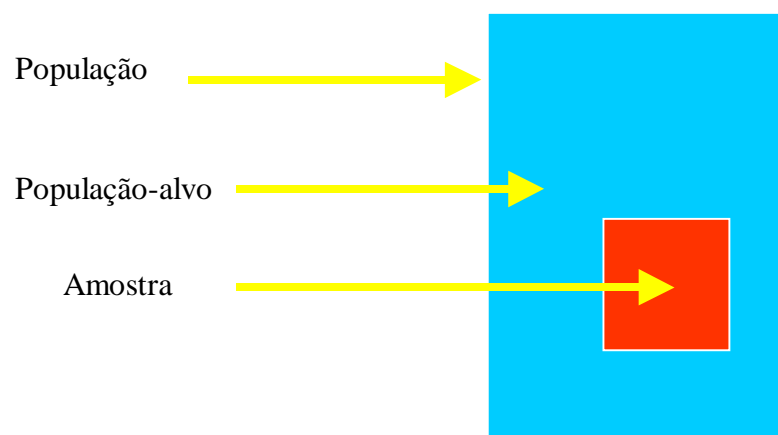


Figura 1: População e Amostra

Um investigador interessado pela readaptação após cirurgia de revascularização, pode concentrar-se somente nos sujeitos que tiveram uma única experiência deste tipo e excluírem os outros.

Uma amostra é dita representativa se as suas características se assemelham o mais possível às da população-alvo. É particularmente importante que a amostra represente não só as variáveis em estudo, mas também outros factores susceptíveis de exercer alguma influência sobre as variáveis estudadas, como a idade, o sexo, a escolaridade, o rendimento, etc.

A Representatividade avalia-se comparando as médias da amostra com as da população-alvo.

Como se ignora se todas as características da população estão presentes numa amostra dado que estas são muitas vezes desconhecidas, admite-se que existe sempre um grau de erro.

ERRO DE AMOSTRAGEM: é a diferença que existe entre os resultados obtidos numa amostra e os que teriam sido obtidos na população-alvo.

Duas soluções existem para reduzir ao mínimo o erro amostral:

1. Retirar de forma aleatória e um número suficiente de sujeitos que farão parte da amostra.
2. Procurar reproduzir o mais fielmente possível a população pela tomada em conta das características conhecidas desta.

3. MÉTODOS DE AMOSTRAGEM

3.1 AMOSTRAGENS PROBABILÍSTICAS E NÃO- PROBABILÍSTICAS

3.1.1 As Amostras Probabilísticas

Os métodos de amostragem probabilística servem para assegurar uma certa precisão na estimação dos parâmetros da população, reduzindo o erro amostral.

A principal característica dos métodos de amostragem probabilística reside no facto de que cada elemento da população tem uma probabilidade conhecida e diferente de zero, de ser escolhida, aquando da tiragem ao acaso para fazer parte da amostra.

O objectivo desta abordagem é obter a melhor representatividade possível.

3.1.1.1 TIPOS DE AMOSTRAGENS PROBABILÍSTICAS

Tipos de Amostragem:

- a) A Amostragem Aleatória Simples;
- b) A Amostragem Aleatória Estratificada;
- c) A Amostragem em Cachos;
- d) A Amostragem Sistemática.

AMOSTRAGEM ALEATÓRIA SIMPLES

A Amostragem aleatória simples é uma técnica segundo a qual cada um dos elementos (sujeitos) que compõe a população alvo tem igual probabilidade de ser escolhido para fazer parte de uma amostra. A amostragem aleatória simples consiste em elaborar uma lista numérica de elementos de onde se tira, com a ajuda de uma tabela de números aleatórios, uma série de números para constituir a amostra.

AMOSTRAGEM ALEATÓRIA ESTRATIFICADA

A Amostragem aleatória estratificada é uma variante da amostra aleatória simples. Esta técnica consiste em dividir a população alvo em subgrupos homogêneos chamados «estratos» e a seguir tirar de forma aleatória uma amostra de cada estrato. A Amostragem aleatória estratificada é utilizada quando a população inteira é reconhecida por certas características precisas, tais como a idade, o sexo, a incidência de uma condição de saúde, tudo isto para assegurar a melhor representatividade possível.

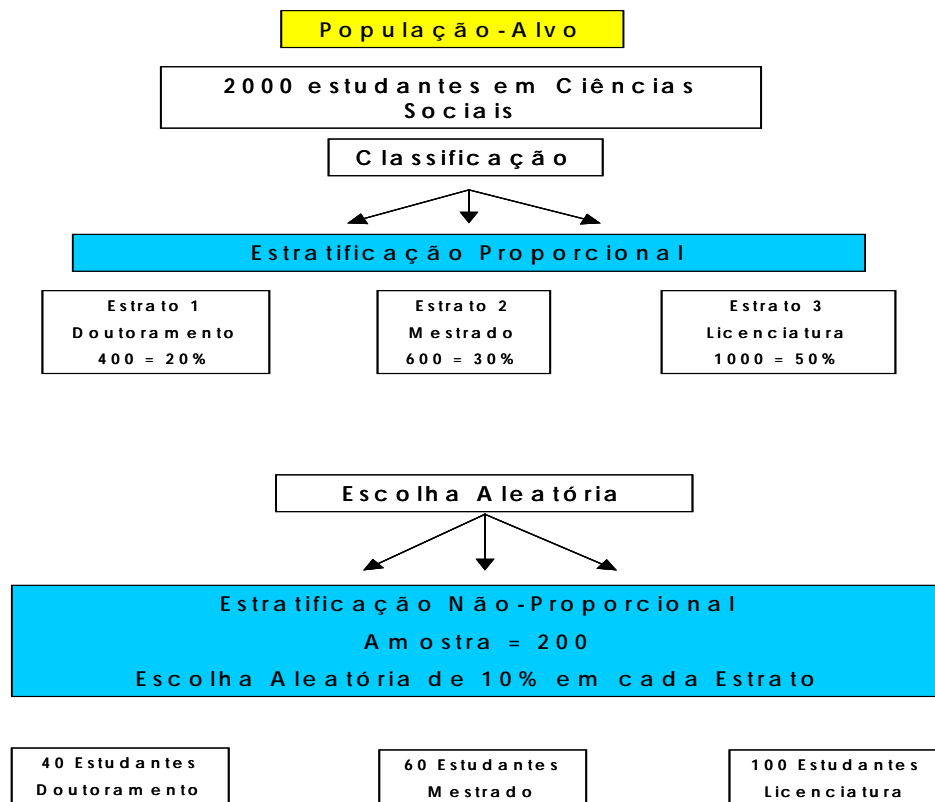


Figura 2: amostra estratificada

AMOSTRAGEM EM CACHOS

Consiste em retirar de forma aleatória os elementos por cachos em vez de unidades. É útil quando os elementos da população estão naturalmente por cachos e por isso devem ser tratados como grupos ou quando não é possível obter uma listagem de todos os elementos da população-alvo.

AMOSTRAGEM SISTEMÁTICA

Consiste quando existe uma lista ordenada de elementos da população. Esta técnica consiste K elementos dessa lista sendo o primeiro elemento da amostra retirado ao acaso.

O intervalo entre os elementos corresponde à razão entre o tamanho da população e da amostra.

Exemplo: Se pretender uma amostra de 100 indivíduos e a população for de 1000 o sistema será $1000:100=10$ (dez em dez é o sistema), isto é, será incluído um elemento da lista de 10 em 10 indivíduos a partir do 1.º n.º sorteado.

Importante

Se se utilizar uma amostragem por cachos ou outros tipos de agrupamentos, a amostra só é considerada probabilística se os grupos foram escolhidos ao acaso antes da repartição aleatória dos sujeitos nos grupos.

3.1.1.2 TIPOS DE AMOSTRAGENS NÃO PROBABILÍSTICAS:

É um procedimento de selecção segundo o qual cada elemento da população não tem a mesma probabilidade de ser escolhido para formar a amostra.

Este tipo de amostragem tem o risco de ser menos representativa que a probabilística no entanto é muitas vezes o único meio de construir amostras em certas disciplinas profissionais nomeadamente na área da saúde.

Tipos de Amostragens Não-Probabilísticas:

- a) A Amostragem Acidental ou de Conveniência;
- b) A Amostragem por Cotas;
- c) A Amostragem de Selecção Racional ou Tipicidade;
- d) A Amostragem por Redes ou Bola de Neve.

AMOSTRAGEM ACIDENTAL OU DE CONVENIÊNCIA

É formada por sujeitos facilmente acessíveis, que estão presentes num determinado local e momento preciso.

Exemplo: pessoas hospitalizadas. Um investigador pode ter acesso a uma unidade hospitalar para constituir uma amostra de pacientes hospitalizados.

Neste tipo de amostra tem a vantagem de ser simples em organizar e pouco onerosa, todavia este tipo de amostra provoca enviesamentos, pois nada indica que as primeiras 30 a 40 pessoas sejam

representativas da população-alvo. São utilizadas em estudos que não têm como finalidade a generalização dos resultados.

AMOSTRAGEM POR COTAS

Idêntica à amostragem aleatória estratificada diferindo desta apenas pelo facto dos sujeitos não serem escolhidos aleatoriamente no interior de cada estrato ou de cada grupo.

AMOSTRAGEM POR SELECÇÃO RACIONAL OU POR TIPICIDADE

Tem por base o julgamento do investigador para constituir uma amostra de sujeitos em função do seu carácter típico.

Por exemplo: o estudo de casos extremos ou desviantes como uma patologia rara ou uma instituição.

AMOSTRAGEM POR REDES OU BOLA DE NEVE

Consiste em escolher sujeitos que seriam difíceis de encontrar de outra forma. Toma-se por base, redes sociais amizades e conhecimentos.

Por exemplo: Imigrantes de Leste.

Quando o investigador encontra sujeitos que satisfazem os critérios escolhidos pede-lhes que indiquem outras pessoas de características similares.

3.2 DETERMINAÇÃO DO TAMANHO DA AMOSTRA

Os tamanhos das amostras são relativos, isto é, depende do tamanho da população. Para determinar as amostras existem várias fórmulas, consoante o parâmetro em critério. As mais utilizadas na saúde são as que se baseiam na percentagem do fenómeno:

3.2.1 Cálculo do Tamanho da Amostra para Populações Infinitas (>100.000 elementos)

A amostra depende da:

1. Extensão do universo;
2. Do Nível de Confiança;
3. Do Erro Máximo permitido;
4. Da percentagem com que o fenómeno se verifica.

Fórmula:
$$n = \frac{\sigma^2 \cdot (p \cdot q)}{e^2}$$

n= Tamanho da amostra

σ = Nível de confiança escolhido expresso em n desvios padrão (s)

p = % com o qual o fenómeno se verifica

q = % complementar (100-p)

e = Erro máximo permitido

Se desejarmos um nível de confiança bastante alto – superior a 99% aplica-se a fórmula dos três desvios.

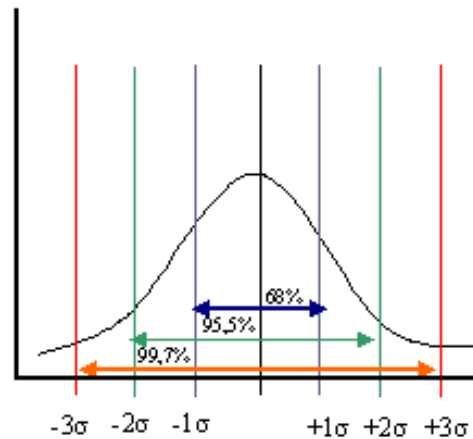


Figura 3: conversão dos níveis de confiança em desvios padrão

Logo, o desvio (s)² seria igual a $3^2 = 9$

Se o erro máximo for de 2% o e^2 será igual a $2^2 = 4$

Exemplo: Se for possível admitir que o número de captações de água em profundidade se situam por volta dos 50%, não ultrapassando esta %, então $p=50$ e, conseqüentemente, $q=100-50$ ou seja 50. Assim, tem-se a equação

$$n = 9 \cdot \frac{(50 \cdot 50)}{4} = 5625$$

Isto é, para atender às exigências estabelecidas, o n.º de captações a analisar seria 5625.

Se todavia, for aceite o nível de confiança de 95% (2 desvios) e um erro máximo de 5% o n.º de elementos será bem menor.

os cálculos.

$$n = \frac{\sigma^2 \cdot (p \cdot q)}{e^2}$$

$$n = 4 \cdot \frac{(50 \cdot 50)}{25} = 400$$

Convém lembrar que sempre que não seja possível estimar uma percentagem do fenómeno, deve utilizar-se sempre $p=50$

3.2.2 Cálculo do Tamanho da Amostra para Populações Finitas (<100.000 elementos)

A amostra depende da:

1. Extensão do universo;
2. Do Nível de Confiança;
3. Do Erro Máximo permitido;
4. Da percentagem com que o fenómeno se verifica.

Fórmula

$$n = \frac{\sigma^2 p q N}{e^2 (N-1) + \sigma^2 p q}$$

Onde:

n = Tamanho da amostra

σ = nível de confiança escolhido, expresso em números de desvios padrão

p = percentagem do fenómeno

q = percentagem complementar

N = tamanho da população

e^2 = erro máximo permitido

Exemplo: Verificar quantos dos 100 empregados de uma cantina cumprem correctamente as normas de higiene e segurança do trabalho.

Presume-se que esse n.º não seja superior a 30% do total; deseja-se um nível de confiança de 95% (2 desvios) e tolera-se um erro até 3%.

Então, $n=90,4$ Logo deverão ser pesquisados 90 empregados.

Confirme aplicando a fórmula

Mas, se a população fosse de 10.000 empregados, com os mesmos critérios anteriormente referidos, então:

$$n = \frac{4 \cdot (30.70) \cdot 10000}{(9.9999) + 4 \cdot (30.70)} = 853$$

DETERMINAÇÃO DO TAMANHO DA AMOSTRA

O tamanho "óptimo" de uma amostra, não depende tanto do tamanho da população mas sim de dois parâmetros estatísticos: a margem de erro e o nível de confiança

Margem de erro – Uma amostra representa aproximadamente (e nunca exactamente) uma população. A medida deste "aproximadamente" é a chamada margem de erro, e é lido assim: se uma pesquisa tem uma margem de erro de 2% e a Doença Cardíaca teve 25% de prevalência na amostra recolhida, podemos dizer que, naquele instante, na população, ela terá uma prevalência entre 23% e 27% (25% menos 2% e 25% mais 2%).

Nível de confiança – As pesquisas são feitas com um parâmetro chamado nível de confiança, geralmente de 95%. Estes 95% querem dizer o seguinte: se realizarmos uma outra pesquisa, com uma amostra do mesmo tamanho, nas mesmas datas e locais e com o mesmo instrumento de recolha de dados, há uma probabilidade de 95% de que os resultados sejam os mesmos (e uma probabilidade de 5%, é claro, de que tudo difira).

Quando já se efectuou uma pesquisa e se deseja conhecer a margem de erro utilizada aplica-se:

$$\text{Fórmula} \quad \sigma_p = \sqrt{\frac{(p \cdot q)}{n}}$$

Onde:

n = Tamanho da amostra

σ_p = Erro padrão ou desvio da percentagem com que se verifica determinado fenómeno

p = percentagem do fenómeno

q = percentagem complementar

3.2.3 Determinação da Margem de Erro da Amostra

Exemplo: Numa pesquisa efectuada com 1000 adultos, verificou-se que 30% bebem café pelo menos uma vez por dia. Qual a probabilidade de que tal resultado seja verdadeiro para todo o universo.

$$\sigma_p = \sqrt{\frac{(30 \cdot 70)}{1000}}$$

$$\sigma_p = 1,45$$

Como o valor encontrado (margem de erro) corresponde a um desvio, então para dois desvios (95,5%), temos $1,45 \cdot 2 = 2,90$.

Para 3 desvios é o triplo (4,35).

Isto significa que, por exemplo, para um nível de confiança de 95% (2 desvios) o resultado da pesquisa apresentará como margem de erro 2,90 para mais ou menos.

É provável, portanto, que o n.º de consumidores de café esteja entre 27,10% ($30\% - 2,90$) e 32,90% ($30\% + 2,90$).

3.2.4 Determinação da Amostra sem conhecer os limites da população

$$\text{Fórmula} \quad \frac{Z^2_{(\alpha/2)} \cdot p(1-p)}{d^2}$$

Em que

p = fenómeno]

α = erro

{Se IC (intervalo de confiança)=95%, α =5% (0,05)} Então

$$\alpha/2 = 0,05/2 = 0,025$$

$$Z(\alpha/2) = Z(0,025) = 1,96$$

d = número de desvios

Assim para um fenómeno que tenha uma prevalência de 25%, os resultados seriam:

$$n = \frac{1,96^2 * 0,25(1-0,25)}{(0,05)^2} = 288$$

Em termos estatísticos consideramos que uma amostra é: Pequena se $n < 30$

CUIDADOS A TER NA ESCOLHA DA AMOSTRA:

1. Imparcialidade: todos os elementos devem ter a mesma probabilidade e oportunidade de serem escolhidos;
2. Representatividade: deve conter em proporção todas as características que a população possui, qualitativa e quantitativamente, de modo a que não se torne tendenciosa;
3. Tamanho: suficientemente grande de modo a fornecer as principais características, por outro lado pequena para economizar tempo, dinheiro e pessoal.

3.3 INDIVÍDUO OU UNIDADE ESTATÍSTICA

O estudo Estatístico recai sobre a amostra, no entanto este é feito de modo pormenorizado a cada um dos elementos da amostra, que são designados por Indivíduo ou Unidade Estatística.

Unidade Estatística: é o factor elementar, o objecto de análise, que independentemente da sua natureza tem que possuir uma definição precisa.

As principais características de uma boa unidade Estatística são:

1. Propriedade ou adequação ao objectivo da investigação;
2. Clareza;
3. Mensurabilidade;
4. Comparabilidade.

No estudo de cada unidade Estatística, surgem resultados individuais com os quais são feitas as inferências sobre a população. Estes resultados têm o nome de Dado Estatístico.

Dado Estatístico: é o resultado do estudo efectuado a cada unidade Estatística tendo em conta a sua individualidade, sendo este depois tratado de modo a permitir inferir sobre a colectividade que a integra (população).

3.4 VARIÁVEL ESTATÍSTICA

Ao ser efectuada uma análise Estatística a uma população, os aspectos (características) que se têm em conta, um ou vários, são denominados por Variável Estatística.

Uma variável Estatística pode ser:

1. Qualitativa: se é a sua natureza que varia de elemento para elemento.

As variáveis qualitativas dividem-se em:

- a. Variáveis nominais: quando o seu significado só se entende em função do nome e o número ou código que se lhe atribua não nos dá nenhuma informação (sexo, cor de olhos, grau de parentesco, tipo de patologia, presença/ausência de factores de risco, etc.).
- b. Variáveis ordinais: quando existe uma ordenação possível (gravidade de uma lesão, classe social, grau de escolaridade, etc.).

2. Quantitativa: se é a sua intensidade que varia de elemento para elemento, tornando-a mensurável ou referenciável.

As variáveis quantitativas dividem-se em:

- a. Variáveis discretas: assume valores isolados, normalmente inteiros (n.º de filhos, n.º de factores de risco, n.º de dependentes, n.º de respostas, etc)
- b. Variáveis contínuas: em que é possível qualquer operação aritmética, podendo assumir qualquer valor real (altura, peso, IMC, distância, etc).

Tendo em conta o número de atributos (características) que estão a ser estudadas, as variáveis podem ser:

1. Unidimensionais: se apenas corresponde a um atributo
2. Bidimensionais: se corresponde a dois atributos;
3. Pluridimensionais: se corresponde a vários atributos.

Modalidade: é toda a manifestação possível de uma variável, isto é, as várias hipóteses de resposta, podendo elas ser duas ou mais.

As modalidades têm obrigatoriamente que ser:

1. Incompatíveis: cada unidade Estatística não pode pertencer simultaneamente a duas ou mais modalidades;
2. Exaustivas: todas as unidades Estatísticas têm que ser inseridas numa modalidade.

A escolha das modalidades deve ser feita de acordo com as informações possuídas. No entanto, surgem situações em que há necessidade de se aumentar uma modalidade suplementar.

3.5 ESTATÍSTICA DESCRITIVA E ESTATÍSTICA INDUTIVA

A Estatística Descritiva recolhe, organiza e analisa os dados de uma amostra, sem tirar qualquer conclusão sobre um grupo maior.

A Estatística Indutiva recolhe, organiza, analisa e estabelece relações entre os dados para fazer inferências sobre a população. Com base nos resultados obtidos sobre a amostra podemos inferir conclusões válidas sobre a população (este ramo da Estatística já exige a utilização de recursos matemáticos especiais, nomeadamente a Teoria das Probabilidades).

Assim, a Estatística Indutiva permite-nos fazer inferências sobre a população e chegar a leis e a teorias.

3.6 PARÂMETRO E DADO ESTATÍSTICO

O parâmetro é toda a função definida a partir dos dados numéricos de uma população.

Exemplo: consideremos as seguintes notas em Estatística - 10 11 10 15 9

$$\text{Média} = \sum x_i / n = 55 / 5 = 11$$

O valor 11 é o parâmetro (resultado da média aritmética).

O dado estatístico é toda a função definida a partir dos dados numéricos de uma amostra.

Exemplo: consideremos a amostra: 10 10

$$\text{Média} = \sum x_i / n = 20 / 2 = 10 \quad \text{O valor 10 é o dado estatístico}$$

3.7 REPRESENTAÇÃO DE UMA VARIÁVEL ESTATÍSTICA

Depois de termos definido algumas noções básicas de estatística, tratar-se-á, a seguir, da segunda fase de um estudo estatístico. Como já referimos, os dados numéricos recolhidos registam-se em séries estatísticas e, para serem analisados, devem ser ordenados e representados em quadros e em gráficos.

Quando trabalhamos com uma variável discreta ou descontínua falamos em seriação e quando trabalhamos com uma variável contínua falamos em classificação.

SERIAÇÃO DE UMA AMOSTRA

Como já referimos anteriormente, uma seriação implica que a variável seja discreta (exemplo: número de filhos de um casal, número de divisões de uma casa, etc.).

O quadro geral de uma seriação está expresso no quadro que se segue.

Quadro 1: Quadro Geral de Seriação

Variável	Frequências absolutas	Frequências acumuladas	Frequências relativas	Frequências relativas acumuladas
X1	n1	n1	$\frac{n1}{N}$	$\frac{n1}{N}$
X2	n2	n1+ n2	$\frac{n2}{N}$	$\frac{n1+n2}{N}$
X3	n3	n1+ n2+ n3	$\frac{n3}{N}$	$\frac{n1+ n2+n3}{N}$
...
Xi	ni	n1+ n2+ ...+ni	$\frac{ni}{N}$	$\frac{n1+ n2+...+ ni}{N}$
...
Xk	nk	n1+ n2+ ...+nk	$\frac{nk}{N}$	1
Total	N		1	

Legenda: Distribuição de frequências de uma variável discreta

Na primeira coluna temos os valores que a nossa variável X toma (x1, x2, ..., xk). Esta coluna pode ser também denominada por xi.

Na segunda coluna apresentamos as frequências absolutas simples ou efectivas, isto é, o número de vezes que cada valor foi observado ou o número de vezes que cada valor surge na amostra.

Esta coluna pode também ser denominada por n_i . Devemos lembrar aqui que o número total de indivíduos pertencentes à população/ou amostra chama-se efectivo total e designa-se por n , devendo sempre verificar-se que $\sum n_i = n$.

Na terceira coluna podemos visualizar as frequências acumuladas, isto é, a soma das frequências absolutas correspondentes a valores inferiores a um determinado valor. Esta coluna pode ser também denominada por $N(x_i)$. Assim,

$$N(x_i) = \sum_{x_j < x_i} n_j$$

Na quarta coluna apresentamos as frequências relativas simples. Esta coluna pode ser também denominada por f_i , em que $f_i = n_i/n$ e, de tal modo que, $\sum f_i = 1$; a frequência relativa é expressa, em geral, em percentagem.

Na quinta coluna temos as frequências relativas acumuladas, isto é, a soma das frequências relativas correspondentes a valores inferiores a um determinado valor. Esta coluna pode ser também denominada por $F(x_i)$; a frequência relativa acumulada é expressa, em geral, em percentagem.

Exemplo: Construa o quadro de frequências para a variável X que toma os seguintes valores:

x_i 14 14 13 12 11 12 10 13 12

Quadro 2: Quadro de Frequências

x_i	n_i	$f_i = n_i/n$	$N(x_i)$	$F(x_i)$
10	1	1/9	1	1/9
11	1	1/9	2	2/9
12	3	3/9	5	5/9
13	2	2/9	7	7/9
14	2	2/9	9	9/9
$\sum n_i = 9$		$\sum f_i = 1$	em que $n = 9$	

CLASSIFICAÇÃO DE UMA AMOSTRA

Como já referimos anteriormente, uma classificação implica que a variável seja contínua (exemplo: a temperatura de um corpo, a altura de uma pessoa, a duração de certo fenómeno, etc. - variáveis relacionadas com o espaço, o tempo ou a massa).

Quadro 3: Quadro de Frequências

Classes	Marcas da classe	Frequências absolutas	Frequências acumuladas	Frequências relativas	Frequências relativas acumuladas
c0-c1	X'1	n1	n1	$\frac{n1}{N}$	$\frac{n1}{N}$
c1-c2	X'2	n2	n1+ n2	$\frac{n2}{N}$	$\frac{n1+n2}{N}$
c2-c3	X'3	n3	n1+ n2+ n3	$\frac{n3}{N}$	$\frac{n1+ n2+n3}{N}$
...
ci-1-ci	X'i	ni	n1+ n2+ ...+ni	$\frac{ni}{N}$	$\frac{n1+ n2+...+ ni}{N}$
...
ci-1-ck	X'k	nk	n1+ n2+ ...+nk	$\frac{nk}{N}$	1
Total	—			1	

Legenda: Distribuição de frequências de uma variável contínua

Na primeira coluna temos as classes. Por convenção, as classes são abertas superiormente, com excepção da última classe, naturalmente.

Na segunda coluna temos as marcas da classe. Esta coluna pode ser também designada por x'i. A marca de uma classe é o ponto médio dessa classe, ou seja, é o ponto equidistante dos extremos de uma classe.

$$\text{Exemplo: classe } c0-c1 \quad x'1 = (c0 + c1) / 2$$

Na terceira coluna apresentamos as frequências absolutas simples ou efectivas, ou seja, o número de vezes que os valores de determinada classe foram observados. Esta coluna pode ser também denominada por ni. Mais uma vez, $\sum ni = n$ (número total de indivíduos pertencentes à população/ou amostra, a que se chama efectivo total).

Na quarta coluna apresentamos as frequências acumuladas, isto é, a soma das frequências absolutas correspondentes a valores inferiores a um determinado valor. Esta coluna pode ser também denominada por N(ai).

Na quinta coluna temos as frequências relativas simples. Esta coluna pode ser também denominada por f_i , em que $f_i = n_i/n$ e, de tal modo, que $\sum f_i = 1$.

Na sexta coluna apresentamos as frequências relativas acumuladas, isto é, a soma das frequências relativas correspondentes a valores inferiores a um determinado valor. Esta coluna pode ser também denominada por $F(\alpha_i)$.

Numa classificação é habitual representarmos por K o número de classes (em geral K varia entre 5 e 20, inclusive) e por A a amplitude (em que $A = x_{\text{máximo}} - x_{\text{mínimo}}$).

Então,

- Se K é dado $\alpha = A/K$
- Se α é dado $K = A/\alpha$, sendo α a amplitude do intervalo de classe e A a amplitude do intervalo da amostra

Exemplo: Construa o quadro de frequências com os seguintes dados:

Classes	[20-23[[23-26[[26-29[[29-32[[32-35[[35-38]
n_i	2	5	7	10	4	2

Quadro 4: Quadro de Frequências

Classes	n_i	f_i	$N(\alpha_i)$	$F(\alpha_i)$	$x' \cdot i$
[20-22]	2	2/30	2	2/30	21
[23-25]	5	5/30	7	7/30	24
[26-28]	7	7/30	14	14/30	27
[29-31]	10	10/30	24	24/30	30
[32-34]	4	4/30	28	28/30	33
[35-38]	2	2/30	30	30/30	36

$\sum n_i = 30$ $\sum f_i = 1$ em que $n = 30$ e $\alpha = 3$

3.8 REDUÇÃO DE UMA VARIÁVEL ESTATÍSTICA

CONCEITO DE REDUÇÃO E SUA CONVENIÊNCIA

Anteriormente definimos o conceito de variável estatística e construíram-se quadros e gráficos estatísticos com vista a uma descrição numérica e gráfica de uma variável estatística. Naturalmente, os gráficos permitem uma primeira síntese das informações registadas nos quadros.

Por outro lado, por simples aproximação das curvas de frequências absolutas ou relativas de duas ou mais variáveis, podemos fazer uma primeira comparação entre elas.

Há, porém, necessidade de sintetizarmos toda a informação respeitante a uma variável estatística, resumindo-se os dados a um pequeno número de elementos que bastam para caracterizá-la. Tal síntese consiste na redução de dados e os elementos numéricos obtidos designam-se por parâmetros da variável estatística.

Feita a representação dos dados estatísticos por meio de quadros e/ou de gráficos, importa fazer sobre os mesmos um estudo no sentido de se poder chegar a conclusões.

Para tal, impõe-se um trabalho de simplificação que consiste em proceder a sínteses, em reduzir grandes quantidades de dados a números simples que permitam uma análise rápida e uma fácil comparação com outras séries da mesma natureza ou de natureza diferente.

Tais números são designados, habitualmente, por características, parâmetros ou medidas e são agrupados em categorias conforme o tipo de informação que fornecem.

MEDIDAS DE TENDÊNCIA CENTRAL OU DE POSIÇÃO

Sob esta designação agrupam-se os parâmetros que, ou nos indicam algo de associável ao núcleo ou centro da distribuição, ou nos permitem compartimentá-la. Vamos considerar as seguintes medidas de tendência central ou de posição: média, mediana, moda e quantis.

3.8.1 Média

MÉDIA ARITMÉTICA

A média aritmética é o ponto de equilíbrio dos dados, isto é, tendo um conjunto de n valores x_1, x_2, \dots, x_n de uma variável X é o quociente entre a soma desses valores e o número deles.

A média aritmética pode ser representada pela seguinte fórmula matemática:

$$\bar{x} = \sum x_i / n$$

Numa amostra seriada os valores x_1, x_2, \dots, x_k ocorrem n_1, n_2, \dots, n_k vezes, respectivamente, a média aritmética ser

$$\bar{x} = \sum n_i x_i / n = \sum f_i x_i$$

Exemplo: Calcule a média aritmética, sabendo que

x_i	4	6	7	8	10
n_i	2	4	5	3	2

Quadro 5: Quadro do cálculo da média

x_i	n_i	$n_i x_i$
4	2	8
6	4	24
7	5	35
8	3	24
10	2	20

$\Sigma=16$ $\Sigma=111$

Então,

$$\bar{x} = \sum n_i x_i / n = 111/16 = 6,94$$

Numa amostra classificada a fórmula definidora da média não se pode aplicar directamente porque não conhecemos os valores exactos da variável estatística, mas apenas o número de observações dentro de cada classe, isto é, quando os dados estão agrupados em classes, para o cálculo da média simples, devemos considerar o ponto médio de cada classe (marca) como representativo de todos os valores nela incluídos, pelo que aquela pode ser representada pela fórmula matemática seguinte:

$$\bar{x}' = \Sigma nix'i / n = \Sigma fix'i$$

Exemplo: Calcule a média aritmética, sabendo que

Classes	[3-5[[5-7[[7-9[[9-11[[11-13]
ni	2	4	5	3	2

Quadro 6: Quadro do cálculo da média

Classes	ni	x'i	nix'i
[3-5[2	3,5	7
[5-7[4	5,5	22
[7-9[5	7,5	37,5
[9-11[3	9,5	28,5
[11-13]	2	12	24
$\Sigma=16$		$\Sigma=119$	

Então:

$$\bar{x}' = \Sigma nix'i / n = 119/16 = 7,43$$

Temos de salientar que quando usamos a marca da classe estamos a colocar um certo erro de agrupamento, pelo que devemos considerar a Correção de Sheppard, de tal modo que:

$$\bar{x} \text{ é semelhante a } \bar{x}'$$

Relativamente às propriedades da média aritmética podemos enunciar duas:

- Somando ou subtraindo uma constante a todos os valores observados, a média resultante ficará aumentada ou diminuída, respectivamente, dessa constante;
- Multiplicando ou dividindo os valores observados por uma constante diferente de zero, a média resultante ficará multiplicada ou dividida, respectivamente, por essa constante.

MÉDIA ARITMÉTICA PESADA OU PONDERADA:

Por sua vez associa-se a x_1, x_2, \dots, x_k certos factores de ponderação ou pesos p_1, p_2, \dots, p_k que dependem do significado ou importância atribuída às observações. Assim, a fórmula matemática da média será:

$$\bar{x}_p = \sum \text{pixi} / p_i$$

Exemplo 9: Um professor de matemática quer saber a média ponderada das suas avaliações nas quatro turmas em que lecciona, sabendo que o teste tinha uma ponderação de 30% e o trabalho uma ponderação de 70%:

Turma A - Média da nota do teste =65%	Média da nota do trabalho =78%
Turma B - Média da nota do teste =60%	Média da nota do trabalho =70%
Turma C - Média da nota do teste =40%	Média da nota do trabalho =28%
Turma D - Média da nota do teste =80%	Média da nota do trabalho =75%

Determine a média ponderada das quatro turmas em conjunto.

$$\bar{x}_p = \sum \text{pixi} / p_i = [((65 \cdot 30) / 100) + ((78 \cdot 70) / 100) + ((60 \cdot 30) / 100) + ((70 \cdot 70) / 100) + ((40 \cdot 30) / 100) + ((28 \cdot 70) / 100) + ((80 \cdot 30) / 100) + ((75 \cdot 70) / 100)] / 4 = 62,3$$

Se preferirmos em quadro a resolução será

Quadro 7: Cálculo da média ponderada

x_i	p_i	pixi
65	30	1950
60	30	1800
40	30	1200
80	30	2400
78	70	5460
70	70	4900
28	70	1960
75	70	5250
Σ	400	24920

$$\bar{x}_p = \sum \text{pixi} / p_i = 24920 / 400 = 62,3$$

Para além da média aritmética e da média aritmética ponderada, temos também a média geométrica, a média harmónica e a média quadrática (a estes três tipos de médias não iremos fazer qualquer referência).

3.8.2 Mediana

A mediana de uma série de n observações x_1, x_2, \dots, x_n de uma variável X é o valor que ocupa a posição central quando as observações estão ordenadas por ordem crescente ou decrescente, isto é, a mediana de uma variável estatística é o valor dessa variável tal que a frequência dos valores que lhe são inferiores é a mesma que a frequência dos valores que lhe são superiores. Representa-se, habitualmente, por Md .

A mediana é usada quando na amostra há valores excêntricos em relação a outros valores.

Para o cálculo da mediana, temos de considerar duas situações: o caso em que N é ímpar e o caso em que N é par.

N É ÍMPAR:

A mediana é um valor observado, de tal modo que o lugar que ocupa é dado pela fórmula

$$Md = (N + 1) / 2$$

Exemplo: Determine a mediana para a seguinte série de dados:

5 9 8 7 6

Ordenando por ordem crescente, vem 5 6 7 8 9

Como N é ímpar, então $Md = (N+1) / 2 = (5+1) / 2 = 6/2 = 3$

Se a mediana ocupa a terceira posição ou terceiro termo, o seu valor é 7

N É PAR:

A mediana não coincide com nenhum valor observado ficando compreendida entre dois valores centrais - classe mediana; convencionou-se tomar para mediana a média destes dois valores. A posição que a mediana ocupa é dada pela fórmula

$$\text{Md} = \text{média dos valores que se encontram na posição } N / 2 \text{ e } (N / 2) + 1$$

Exemplo: Determine a mediana para a seguinte série de dados

5 6 12 9 8 7

Ordenando por ordem crescente, vem 5 6 7 8 9 12

Como N é par, então $N/2 = 6/2 = 3$; $(N/2 + 1) = 3+1 = 4$

Assim, a classe mediana é ocupada pelas posições 3 e 4, ou seja, pelos valores 7 e 8, pelo que a

$$\text{Md} = (7+8) / 2 = 15/2 = 7,5$$

3.8.3 Moda

A moda (ou valor modal) de uma série de n valores x_1, x_2, \dots, x_n de uma variável X é o valor onde a frequência atinge o máximo (relativo). Representa-se, habitualmente, por Mo.

AMOSTRA SERIADA:

A moda é o valor da variável com maior efectivo, isto é, se uma variável é discreta, a(s) moda(s) é(são) o(s) valor(es) da variável estatística que se observa(m) com maior frequência.

Exemplo: Determine a moda para a seguinte série de valores

x_i	4	6	8	10	2
n_i	1	3	5	4	2

A moda é 8.

AMOSTRA CLASSIFICADA:

Neste caso falamos em classe modal, tomando-se para moda a marca da classe. Assim, a moda de uma variável contínua X é todo o valor de X para o qual a função de frequência (absoluta ou relativa) atinge um máximo local.

Exemplo: Determine a moda para a seguinte série de valores

classes	[2-4[[4-6[[6-8[[8-10[[10-12]	
ni		1	7	4	10	2

A classe modal é [8-10[, pelo que a moda = $(8+10) / 2 = 18/2=9$

RELAÇÃO EMPÍRICA ENTRE A MÉDIA, A MEDIANA E A MODA

Pearson chegou à conclusão de que para as distribuições unimodais e moderadamente assimétricas, vigora a seguinte relação empírica:

$$\text{Média} - \text{Moda} = 3 (\text{Média} - \text{Mediana})$$

Vejamos, agora, a relação existente entre estas três medidas no caso das figuras que se seguem

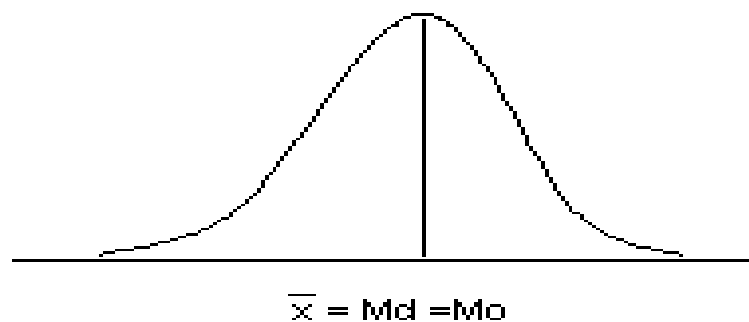


Figura 4: Curva simétrica

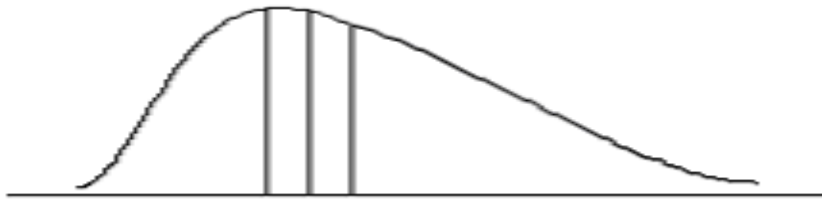


Figura 5: Curva assimétrica à direita

Média > Moda
Média > Mediana
Mediana > Moda

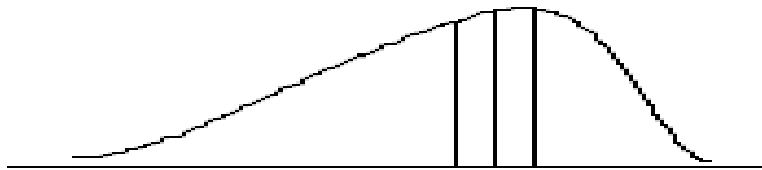


Figura 6: Curva assimétrica à esquerda

Moda > Mediana
Moda > Média
Mediana > Média

A UTILIZAÇÃO DA MÉDIA ARITMÉTICA, DA MEDIANA E DA MODA

MÉDIA ARITMÉTICA:

A medida de tendência central mais usada é a média aritmética, que apresenta em relação à mediana e à moda vantagens apreciáveis, tais como:

- É facilmente calculável;
- É a que melhor se presta a ulteriores análises estatísticas;
- Depende de todos os valores da série;
- É uma medida de tendência central particularmente estável, variando o menos possível de amostra para amostra extraídas da mesma população;
- Pode ser tratada algebricamente.

Existem, todavia, casos em que a informação fornecida pela mediana ou pela moda, parece ser mais completa do que a fornecida pela média aritmética, como passamos a indicar.

MEDIANA:

Esta medida de tendência central deverá utilizar-se quando:

- Se deseja conhecer o ponto médio exacto da distribuição;
- Existem resultados extremos que afectariam grandemente a média. Importa salientar que a mediana não é influenciada pelos resultados extremos - exemplo: na série 10,13,15,16,18,19,21, tanto a média como a mediana são 16; se substituirmos 21 por 50 e os restantes resultados permanecerem os mesmos, a mediana é 16 e a média será de 20,1;
- A distribuição é truncada, isto é, incompleta nas extremidades - exemplo: desejamos medir os tempos de reacção em Psicologia e as várias respostas ultrapassam as capacidades de medida do aparelho; vemo-nos, assim, impossibilitados de calcular a média, já que não dispomos de todos os valores da variável, sendo, no entanto, possível calcular o valor da mediana, já que conhecemos o número de medidas efectuadas.

MODA:

Esta medida de tendência central deverá utilizar-se quando:

- Uma medida rápida e aproximada da tendência central for suficiente;
- Se deseja conhecer o valor mais típico de uma distribuição - exemplos: descrever o estilo de vestido usado pela mulher média, salário preponderante em determinada empresa.

É de salientar que, apesar das vantagens apresentadas para a moda, esta tem a desvantagem de ser de determinação imprecisa, no caso das variáveis contínuas.

3.8.4 Quantis

Chama-se quantil de ordem p com $0 \leq p \leq 1$ e representa-se, habitualmente, por C_p ao valor de x tal que $F(x)=p$.

Alguns quantis têm denominações especiais:

QUARTIS:

Os quartis dividem a série ordenada em 4 partes iguais, contendo cada uma delas $1/4$ ou 25% das observações.

$Q_1=1^\circ$ quartil (corresponde ao quantil de ordem $p=1/4$)

$Q_2=2^\circ$ quartil (corresponde ao quantil de ordem $p=1/2$)

$Q_3=3^\circ$ quartil (corresponde ao quantil de ordem $p=3/4$)

Assim, Q_1 é o valor da variável estatística que deixa atrás de si 25% das observações; Q_2 é o valor da variável estatística que deixa atrás de si 50% das observações e Q_3 é o valor da variável estatística que deixa atrás de si 75% das observações. A $(Q_1 - Q_3)$ chama-se intervalo interquartil e é o intervalo ao qual pertencem 50% das observações, deixando 25% para a direita e 25% para a esquerda.

É de notar que dizer que os quartis dividem a série em 4 partes iguais não significa que, por exemplo, os intervalos (Q_1, Q_2) e (Q_2, Q_3) têm a mesma amplitude, mas sim que contêm o mesmo número de observações.

DECIS:

Os decis dividem a série ordenada em 10 partes iguais, contendo cada uma delas $1/10$ ou 10% das observações.

$D_1=1^\circ$ decil (corresponde ao quantil de ordem $p=1/10$)

$D_2=2^\circ$ decil (corresponde ao quantil de ordem $p=2/10$)

CENTIS:

Os centis dividem a série ordenada em 100 partes iguais, contendo cada uma delas $1/100$ ou 1% das observações.

$C1=1^{\circ}$ centil (corresponde ao quantil de ordem $p=1/100$)

$C2=2^{\circ}$ centil (corresponde ao quantil de ordem $p=2/100$)

RELAÇÕES ENTRE QUARTIS, DECIS, CENTIS E MEDIANA:

Como podemos observar na figura 18

$Q1=C25$
 $Q2=Md=D5=C50$
 $Q3=C75$
 $D1=C10$
 $D2=C20$

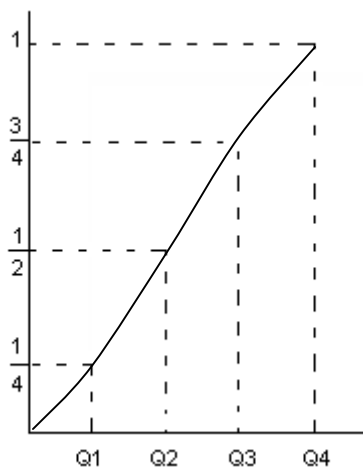


Figura 7: Curva assimétrica à direita

3.8.5 Medidas de dispersão

Uma medida de tendência central não nos dá, só por si, uma informação exaustiva da distribuição considerada; pelo contrário, a capacidade que se lhe atribui de representar os elementos de uma distribuição depende do modo como estes se concentram ou dispersam em torno dela. Assim, podemos dizer que os parâmetros de tendência central não são suficientes para caracterizar uma série estatística, apesar de a mediana e os quantis darem já uma ideia sumária do modo como estão distribuídas as observações.

Consideremos o seguinte exemplo: Dois grupos de alunos com as seguintes classificações

A	2	3	10	16	19
B	8	9	10	11	12

A média e a mediana é 10 e, contudo, estas distribuições são muito diferentes. Com efeito, enquanto no grupo A as notas apresentam desvios muito grandes, na distribuição B todos os valores se aproximam de 10. A dispersão ou variabilidade da primeira série é mais acentuada do que na segunda.

Quer dizer: distribuições com a mesma tendência central podem apresentar aspectos bastante diferentes no que concerne à dispersão ou variabilidade, e à medida que esta dispersão aumenta, menos significativas da distribuição vão sendo as medidas de tendência central.

Assim, para melhor caracterizarmos uma distribuição, temos de considerar, além das medidas de tendência central, uma outra medida que exprima o grau de dispersão ou variabilidade dos dados.

Vamos considerar as seguintes medidas de dispersão: amplitude total, amplitude interquartis, desvio médio, variância, desvio padrão e coeficiente de dispersão ou de variação.

AMPLITUDE TOTAL

A amplitude total é a diferença entre o maior valor e o menor valor, isto é, a amplitude total de uma variável estatística é a diferença entre o valor máximo e o valor mínimo dos valores

observados. É a forma mais simples de avaliar a dispersão dos dados, de tal modo que quanto maior for a amplitude total maior é a dispersão dos dados.

A amplitude total pode ser também denominada de intervalo total ou campo de variação; representa-se, habitualmente, por A e apenas usa valores extremos.

Numa amostra seriada:

$$A = x_{\text{máximo}} - x_{\text{mínimo}}$$

Numa amostra classificada:

$$A = \text{extremo superior da última classe} - \text{extremo inferior da primeira classe}$$

Se alguma destas classes for de amplitude indeterminada não é possível definir o intervalo de variação.

A amplitude total apresenta as seguintes desvantagens:

Embora seja fácil de calcular, a amplitude total depende somente dos valores extremos, que são, geralmente, os menos frequentes e os menos significativos de uma distribuição, desprezando-se os valores intermédios que são os mais frequentes. Além disso, os valores extremos são vulgarmente anómalos e muito variáveis, consoante a amostra que se retire de uma população, de tal modo que duas distribuições podem ter a mesma amplitude total, mas dispersões muito diferentes.

Outro inconveniente da amplitude total é consequência de não tomar em consideração as frequências das observações.

Exemplo: Calcule a amplitude total do grupo G, sabendo que:

G= 2 3 10 16 19

$$A = x_{\text{máximo}} - x_{\text{mínimo}} = 19 - 2 = 17$$

AMPLITUDE INTERQUARTIS

Os quartis fornecem indicação quanto à forma como as observações se distribuem em torno da mediana.

Como o 1º e o 3º quartis representam valores abaixo dos quais estão, grosso modo, respectivamente, 25% e 75% das observações, entre eles existirão, assim, 50% das observações centrais. Consequentemente, quanto mais aproximados estiverem estes quartis, maior será a concentração das observações em torno da mediana.

A amplitude interquartis pode ser definida como a diferença entre o Quartil 3 e o Quartil 1. Esta medida de dispersão pode ser também denominada de intervalo interquartis ou intervalo quartílico.

Como podemos observar na figura 8, quanto mais achatada é a curva, maior é a amplitude e quanto maior é a amplitude interquartilica mais dispersa é a distribuição.

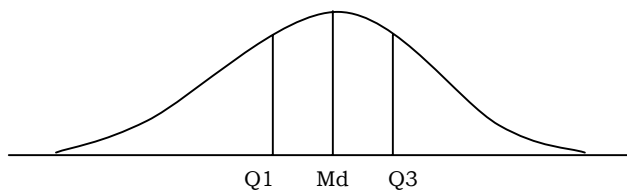


Figura 8: Curva simétrica achatada (platocúrtica)

Exemplo: Calcule a amplitude interquartis, sabendo que $Q3=177,46$ e $Q1=166,88$.

$$Q = Q3 - Q1 = 177,46 - 166,88 = 10,58$$

Podemos também calcular a amplitude semi-interquartis ou intervalo inter-quartilico ou intervalo semi-quartil ou desvio quartílico ou amplitude semi-interquartilico, que se representa, habitualmente, por Q e pode ser definida do seguinte modo:

$$Q = (Q3 - Q1) / 2$$

A distribuição é tanto mais dispersa quanto maior for a amplitude interquartis.

Podemos enumerar as vantagens e as desvantagens do uso da amplitude interquartis e da amplitude semi-interquartis:

Estas medidas são mais completas do que a amplitude total, porque usam dois valores menos extremos (Q1 e Q3). No entanto, têm ainda a limitação de não entrarem em linha de conta com a disposição das frequências nos intervalos definidos pelos valores separados - exemplo: a amplitude interquartis será a mesma, quer as 50% das observações se acumulem num só ponto, quer estejam uniformemente distribuídas por esse intervalo interquartis.

As medidas de dispersão que passaremos a descrever não têm esta limitação, porquanto o seu cálculo depende de todos os valores da série.

DESVIO

Dados n valores x_1, x_2, \dots, x_n de uma variável X , chama-se desvio de cada valor x_i em relação à constante c , a diferença de x_i para c , isto é,

$$x_i - c$$

Note-se que os desvios da variável X em relação a c , isto é, $(x_1-c), (x_2-c), \dots, (x_n-c)$ constituem os n valores da variável $X-c$.

DESVIO MÉDIO

Falamos em desvio médio quando consideramos os desvios de cada valor x_i em relação à média aritmética, isto é:

$$x_i - \bar{x}$$

O simples total destes desvios não pode ser utilizado como medida de dispersão, por ser identicamente nulo. De facto, para n valores singulares, ter-se-á:

$$\sum (x_i - \bar{x}) = 0$$

No entanto, o quociente entre a soma dos módulos destes desvios e o número deles, já pode ser considerado como medida de dispersão

$$D.M. = \sum |x_i - \bar{x}| / n$$

Exemplo 20: Calcule o desvio médio para A = 4 5 3

Quadro 8: Cálculo do Desvio Médio		
x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $
4	$4 - 4 = 0$	0
5	$5 - 4 = 1$	1
3	$3 - 4 = -1$	1
$\Sigma = 12$	$3 - 4 = -1$	$\Sigma = 2$

$$\bar{x} = \Sigma x_i / n = 12/3 = 4$$

$$D.M. = \Sigma |x_i - \bar{x}| / n = 2/3 = 0,67$$

Observação: também se utiliza o desvio médio em relação a qualquer outra medida de posição central.

Numa amostra seriada temos:

$$D.M. = \Sigma n_i |x_i - \bar{x}| / n = \Sigma f_i |x_i - \bar{x}|$$

Se os valores da variável estiverem tabelados de modo que cada valor x_i corresponda a frequência absoluta n_i , o desvio médio é igual à soma dos produtos das frequências pelos valores absolutos dos respectivos desvios em relação à média, dividida pelo efectivo da distribuição.

Numa amostra classificada, os desvios em relação à média aritmética são calculados a partir dos pontos médios de cada classe, ou seja:

$$D.M. = \Sigma n_i |x'_i - \bar{x}| / n = \Sigma f_i |x'_i - \bar{x}|$$

Exemplo: Calcule o desvio médio para

classes	[4-6[[6-8]
ni	1	2

Quadro 9: Cálculo do Desvio Médio para classes

Classes	ni	x' i	x' i- x'	x' i- x'	ni x' i- x'
[4-6[1	4,5	4,5-6,2	1,7	1,7
[6-8]	2	7	7-6,2	0,8	1,6

$\Sigma=2,4$

$\Sigma=3,3$

$$\bar{x}' = \Sigma nix'i / n = [(1x4,5) + (2x7)] / 3 = 6,2$$

Então, o desvio médio é

$$D.M.=\Sigma ni |x'i-x'| / n = 3,3 / 3 = 1,1$$

O desvio médio apresenta a seguinte desvantagem:

Embora dependa de todos os valores observados, o desvio médio tem a desvantagem de considerar os valores absolutos dos desvios, o que impede o seu tratamento algébrico.

VARIÂNCIA

Outra maneira de eliminarmos os sinais dos desvios, consiste em elevá-los ao quadrado. Por isso, em vez da média dos valores absolutos dos desvios considera-se a média dos quadrados dos desvios. Obtém-se, assim, uma outra medida de dispersão bastante usada - a variância.

Dados n valores x_1, x_2, \dots, x_n de uma variável X, chama-se variância e representa-se, habitualmente, por s^2 ou s^2_x a média aritmética dos quadrados dos desvios em relação à média dessas valores, isto é:

$$S^2 = \Sigma (xi- \bar{x})^2 / n$$

Exemplo 22: Calcule a variância para X=17,18,19,20,21

Resolução 22:

$$\bar{x} = \Sigma xi / n = (17+18+19+20+21) / 5 = 19$$

Então, a variância é:

Quadro 10: Cálculo da Variância

xi	(xi- x)	(xi- x) ²
17	17-19=-2	4
18	18-19=-1	1
19	19-19=0	0
20	20-19=1	1
21	21-19=2	4
N=5		$\Sigma=10$

$$S^2 = \Sigma (xi- \bar{x})^2 / n = 10/5 = 2$$

Se x_1, x_2, \dots, x_n ocorrem n_1, n_2, \dots, n_k vezes, respectivamente, temos

Para uma amostra seriada:

$$S^2 = \sum n_i (x_i - \bar{x})^2 / n = \sum f_i (x_i - \bar{x})^2$$

Para uma amostra classificada:

$$S'^2 = \sum n_i (x'_i - \bar{x}')^2 / n = \sum f_i (x'_i - \bar{x}')^2$$

Exemplo 23: Calcule a variância para a distribuição cuja média é 103

Classes	[85-90[[90-95[[95-100[[100-105[[105-110[[110-115[[115-120]
ni	12	25	38	85	93	16	9

Quadro 11: Cálculo da Variância

Classes	ni	x'_i	$(n_i x'_i)$	$(x'_i - \bar{x}')$	$(x'_i - \bar{x}')^2$	$n_i (x'_i - \bar{x}')^2$
85-90	12	87	1044	-15,5	240,25	2883
90-95	25	92	2300	-10,5	110,25	2756,25
95-100	38	97	3686	-5,5	30,25	1149,5
100-105	85	102	8670	-0,5	0,25	21,25
105-110	93	107	9951	4,5	20,25	1883,25
110-115	16	112	1792	9,5	90,25	1444
115-120	9	117,5	1057,5	15	225	2025
	278		28500,5			12162,25

$$\begin{aligned} \bar{x}' &= \sum (n_i x'_i) / n \\ &= 28500,5 / 278 \\ &= 102,5 \end{aligned}$$

Então, a variância é

$$S'^2 = \sum n_i (x'_i - \bar{x}')^2 / n = 12162,25 / 278 = 43,7491$$

Podemos, agora, enumerar as propriedades da variância:

- Somando ou subtraindo uma constante a todos os valores observados, a variância resultante permanecerá inalterada;

- Multiplicando ou dividindo todos os valores observados por uma constante diferente de zero, a variância resultante virá multiplicada ou dividida, respectivamente, pelo quadrado dessa constante.

CORRECÇÃO DE SHEPPARD:

Ao calcular-se a média e a variância da amostra classificada através da distribuição empírica das marcas, comete-se um certo erro (erro de agrupamento), pois supomos que as observações agrupadas em cada classe têm todas o valor da respectiva marca. Todavia, existem fórmulas correctivas devidas a Sheppard, isto é, na variância, ao valor calculado deve subtrair-se $1/12$ ao quadrado da amplitude das classes (a)

$s^2x = s^2x' - a^2/12$ assim, no nosso exemplo anterior, a variância corrigida era

$$s^2x = 43,7491 - \frac{5^2}{12}$$

$$s^2x = 41.666$$

DESVIO PADRÃO

O desvio padrão pode ser definido como a raiz quadrada da variância, representando-se, habitualmente, por s_x , isto é:

$$s_x = \sqrt{s^2x}$$

Ainda que a variância nos dê uma boa informação sobre a distribuição ou variabilidade dos valores observados em relação à sua média, apresenta, no entanto, a desvantagem de não se exprimir na mesma unidade a que estão referidos os dados iniciais. Contudo, esta desvantagem poderá ser eliminada se extrairmos a raiz quadrada da variância. A nova medida chama-se desvio padrão ou desvio quadrático.

Numa amostra seriada, temos:

$$s_x = \sqrt{s^2x}$$

Exemplo: Calcule o desvio padrão, sabendo que a variância de uma amostra seriada é 2.

$$s^2x = 2 = 1,414$$

Numa amostra classificada, temos:

$$s'x = \sqrt{s^2x'}$$

Exemplo: Calcule o desvio padrão, sabendo que a variância corrigida de uma amostra classificada é 4327,16.

$$s'x = \sqrt{s^2x'} \\ = \sqrt{4327,16} = 65,78$$

A UTILIZAÇÃO DAS MEDIDAS ABSOLUTAS DE DISPERSÃO

Amplitude total

Utiliza-se quando:

- Os dados forem muito raros ou demasiado dispersos para se justificar o cálculo de uma medida mais precisa de dispersão;
- For apenas necessário o conhecimento dos resultados extremos;
- Desejamos um índice muito rápido de dispersão.

Amplitude interquartílica

Utiliza-se quando:

- A mediana é a medida de tendência central usada;
- Existirem resultados extremos que poderiam afectar o desvio padrão de uma maneira desproporcionada;
- A distribuição é truncada;
- A distribuição apresenta uma forte assimetria.

Desvio médio

Utiliza-se quando:

- Desejamos ponderar todos os desvios em relação à média de acordo com a sua grandeza;
- Os desvios extremos influenciarem indeterminadamente o desvio padrão.

Desvio padrão e Variância

Utilizam-se quando:

- Se procura uma medida de dispersão em relação com a curva normal;
- Tiverem de ser calculados posteriormente coeficientes de correlação e outras estatísticas;
- Se desejar obter uma medida que se revista de um máximo de estabilidade;
- Se se trata somente de descrever uma distribuição prefere-se o desvio padrão à variância.

A variância intervém sobretudo na análise estatística.

3.8.6 Coeficiente de dispersão

As medidas de dispersão a que anteriormente nos referimos são medidas que se exprimem na mesma unidade dos dados e, sendo assim, torna-se impossível comparar entre si as dispersões de duas distribuições cujos valores não se refiram à mesma unidade.

Exemplo: Distribuição A: $\bar{x}_A=30$ $s_A=10$
 Distribuição B: $\bar{x}_B=600$ $s_B=20$

Qual é a distribuição mais dispersa? Se compararmos os desvios padrões é a B, porque tem maior desvio padrão. Mas a variação de 20 para 600 é muito maior do que 10 para 30. Assim, em vez de compararmos os desvios padrões, aplicamos outra medida de dispersão relativa que é o coeficiente de variação ou de dispersão, que pode ser definido pela fórmula

$$V = sx / \bar{x}$$

Exemplo 26: A distribuição dos pesos e das alturas de um grupo de estudantes de determinada Universidade conduziu aos seguintes resultados:

X: Pesos	Média=57,5Kg	Desvio Padrão=7,5Kg
Y: Alturas	Média=170cm	Desvio Padrão=7,1cm

Determine o coeficiente de dispersão para cada uma das distribuições e, depois, indique em qual delas a dispersão relativa é maior.

$$V_x = sx / \bar{x} = 7,5\text{Kg} / 57,5\text{Kg} = 0,130 = 13\%$$

$$V_y = sy / \bar{Y} = 7,1\text{cm} / 170\text{cm} = 0,042 = 4,2\%$$

Assim, podemos dizer que a dispersão relativa é mais acentuada na distribuição dos pesos (X).

A dispersão é maior na distribuição que tiver maior coeficiente de dispersão.

Se pretendermos estabelecer comparações entre dispersões absolutas, devemos usar o desvio padrão, de tal modo que quanto maior for o desvio padrão maior será a dispersão.

Se pretendermos estabelecer comparações entre dispersões relativas, devemos usar um coeficiente de dispersão, de tal modo que quanto maior for o coeficiente de dispersão V maior será a dispersão.

3.9 AS CARACTERÍSTICAS DA DISTRIBUIÇÃO NORMAL

A distribuição normal apresenta uma curva em forma de sino, como mostra a figura 9. A sua principal característica é a de as três medidas de tendência central - média, mediana e moda - Se encontrarem todas no mesmo ponto da curva, ou seja, todas terem o mesmo valor ou, pelo menos, valores muito próximos. Se Os elementos que constituem uma distribuição estão muito próximos ou muito dispersos, encontraremos formas semelhantes à da figura 10.

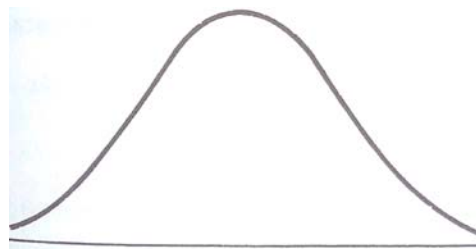


Figura 9: Distribuição normal

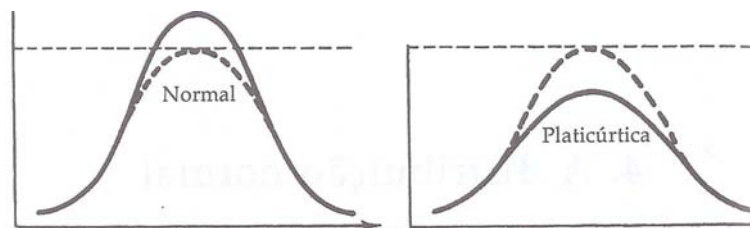


Figura 10 - Distribuição leptocúrtica e platicúrtica

Estas não são distribuições normais, apesar de a média, a mediana e a moda se encontrarem todas no mesmo ponto (é isto que confere simetria à distribuição); a distribuição normal tem sempre a forma de um sino. Como foi «descoberta» pelo matemático Gauss, também lhe chamamos distribuição gaussiana.

A maior parte dos dados recolhidos com organismos vivos têm este padrão. Podemos observar que, devido à forma da curva, há poucos resultados muito baixos e poucos resultados muito elevados (a curva «cai» nos extremos esquerdo e direito, o que se deve às baixas frequências encontradas), enquanto a maioria dos resultados se encontram junto à média. Vamos debruçar-

nos sobre o padrão de resultados muito em breve, mas nesta fase referiremos outra característica muito importante da distribuição normal. Teóricamente, a curva nunca toca o eixo horizontal, mas aproxima-se dele infinitamente. Esta é uma propriedade matemática da distribuição que não se reflecte na recolha de dados «real». Não nos cruzamos com seres humanos com dimensões gigantescas ou micros' cóspicas!

Então as propriedades da distribuição normal são as seguintes:

- 1) É simétrica;
- 2) Tem forma de sino;
- 3) A média, a mediana e a moda encontram-se no mesmo ponto da curva;
- 4) Tem duas pontas que nunca tocam o eixo horizontal.

Podemos perguntar-nos quão rígida é a distribuição normal relativamente aos elementos. Por outras palavras, quanto pode uma curva desviar-se da forma de sino e continuar a ser considerada normal? Geralmente, usamos duas abordagens para tomarmos este tipo de decisão; na verdade, este problema é mais importante do que o leitor pode pensar, pois existem testes estatísticos, descritos mais à frente, que só podem realizar-se se os elementos forem normalmente distribuídos. Uma das abordagens baseia-se na observação dos dados «por averiguação», para lhe darmos um nome mais respeitável. Se o conjunto de números for extenso, tornar-se-á mais fácil desenhar uma distribuição de frequências. A outra abordagem reside em seguir um dos procedimentos matemáticos para determinar se um conjunto de resultados é normalmente distribuído. A versão do teste do quiquadrado que incluímos neste manual é um desses procedimentos. Na verdade, é improvável que nesta fase da sua carreira estatística necessite de saber com grande precisão se uma distribuição é considerada normal ou não, pelo que o teste gráfico deve bastar. No entanto, deve ser sensível ao problema.

3.9.1 A curva normal e os desvios-padrão

Suponhamos que temos um conjunto de números cuja média é 50 e cujo desvio padrão é 5. Chamamos a este valor (centímetros, segundos, pontos numa escala, ou outra coisa qualquer) um desvio padrão. Dez centímetros, segundos, etc., seriam dois desvios padrão e quinze centímetros, segundos, etc., três desvios padrão ... sempre com referência ao conjunto através do qual obtivemos o valor 5. É como se pudéssemos tirar o desvio padrão e transformá-lo numa unidade de medida de uma escala; é como se disséssemos que uma polegada são 2,54 cm. Nunca misturaríamos polegadas e centímetros nos mesmos cálculos, mas poderíamos converter uma unidade na outra. Do mesmo modo, não misturamos resultados de desvios padrão com resultados reais, mas convertemos uma escala na outra.

Voltemos às proporções de números em diferentes partes da distribuição. Se retirarmos uma parte da curva entre a média, que esta marcada no eixo horizontal da figura que se segue como 50, e um desvio padrão, marcado no eixo horizontal como 55, sabemos que devemos ter cerca de um terço de todos os resultados neste conjunto, porque é o que acontece sempre com a distribuição normal. De uma forma mais precisa, a proporção exacta do número total de resultados que se encontram entre a média e um desvio-padrão acima da média (50 e 55 neste caso) é 34,13 %. Como a distribuição normal é simétrica, deve verificar-se a mesma coisa abaixo da média, isto é, devemos ter outros 34,13 % dos resultados entre os valores 50 e 45 - sendo 45 o valor da média menos um desvio-padrão de 5 pontos. Observemos as duas partes a sombreado. A aritmética elementar diz-nos que 68,26 % do total dos resultados se encontram entre os valores 45 e 55, pertencendo 31,74 % aos valores extremos que se encontram nos outros dois lados. Mais uma vez, a simetria da curva significa que, para esta proporção, metade de 31,74 %, ou seja, 15,87 %, encontra-se de cada um dos lados da distribuição. Por outras palavras, cerca de 16 % de todos os números neste conjunto serão menores do que 45, sendo a mesma quantidade maior do que 55.

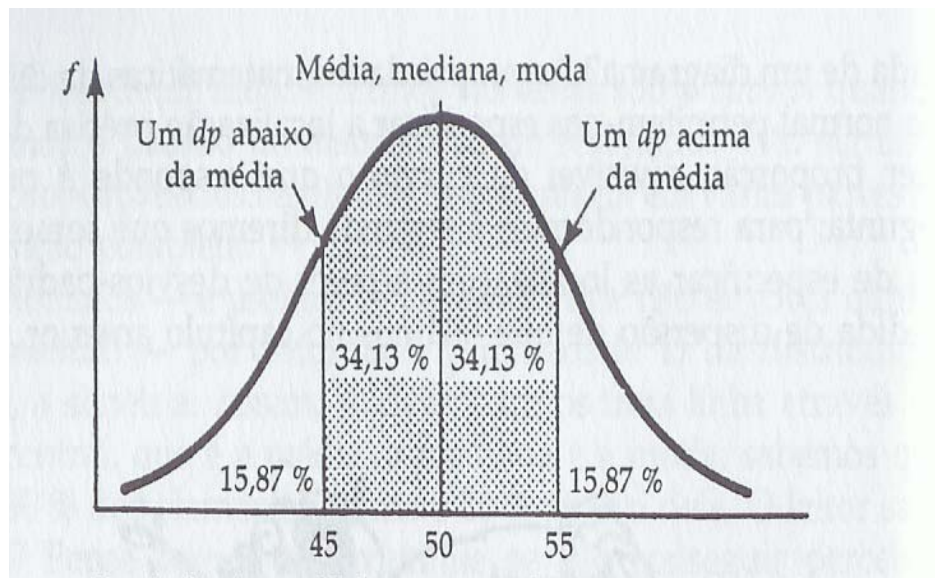


Figura 11: % de resultados 1 dp acima e abaixo da média

Exemplo: Suponhamos que o professor obtém os resultados de um teste de leitura feito a 200 crianças. Os resultados são normalmente distribuídos com uma média de 60 e um desvio-padrão de 8. A partir das propriedades da distribuição normal, sabemos que cerca de dois terços dos resultados, isto é, aqueles que foram obtidos com cerca de 136 crianças, encontrar-se-ão entre os 52 e os 68 pontos. Cerca de 32 crianças (16 %) terão resultados abaixo de 52 e cerca de 32 terão resultados acima de 68. Já os referimos todos. Suponhamos então que os pais de uma criança que obteve 68 gostariam de saber algo acerca do progresso do seu filho. Quando souberam que o resultado da criança estava «acima da média», ficaram contentes, mas gostariam de saber, posteriormente, quão acima da média se encontra, relativamente aos outros 50 % de crianças que também obtiveram resultados «acima da média». Por outras palavras, os pais pretendem saber qual a posição relativa do desempenho do filho. Se os resultados estivessem todos muito perto da média, sendo a nota máxima 68, os pais continuariam encantados. Ficariam, porém, menos satisfeitos se soubessem que a nota máxima tinha sido 90, com um grande conjunto de notas altas, acima de 70. No entanto, o professor sabe que o desvio-padrão das notas foi 8 e, por isso, um terço de todos os resultados estava entre 60 e 68. Sabendo que 50 % dos resultados obtidos estavam «abaixo da média», podemos perceber que a posição desta criança está ao nível de 84 % dos resultados, na parte superior de todos os resultados. Afinal, os pais têm razões para estarem contentes! Se a criança tivesse obtido 76, os pais teriam muito mais razões para estarem orgulhosos, pois saberiam que o seu filho estava acima de 98 % das outras crianças (nota 76 e dois desvios-padrão acima da média); uma nota 84 colocaria o menino na posição invejável de estar acima de 99,87 % das outras crianças - por outras palavras, num grupo de 200 crianças,

estaria, muito provavelmente, no topo. Os desvios-padrão cortam proporções fixas da distribuição normal, a partir da média e até ao infinito (pelo menos teoricamente), nas duas direcções. Deve certificar-se de que percebeu como se obtém a posição relativa da nota 76 (isto é, $50\% + 33\% + 15\%$) e como se calcula que neste grupo de crianças existem outras quatro com notas acima de 76. Veja se consegue calcular a nota que colocaria a criança na posição, menos invejável, de estar apenas a quatro lugares do fim.

A resposta é 44. Para obtermos esta nota necessitamos de saber que nota representa dois desvios-padrão abaixo da média ou que nota corresponde a 2%. Partindo de 60, a média, se lhe subtrairmos o valor de dois desvios-padrão - 16, duas vezes o valor de 8, que é um desvio-padrão - obteremos 44. Devemos ter cuidado e não misturar os valores dos desvios-padrão com os resultados reais. Neste exemplo não subtraímos o valor 2 da média de 60, apesar de querermos o resultado que estava dois desvios-padrão abaixo dele. Subtraímos 16 pontos, pois este é o número que corresponde a dois desvios-padrão para este conjunto de resultados.

Resultados z Nos exemplos considerados os resultados encontravam-se sempre na média, ou exactamente um, dois ou três desvios-padrão acima ou abaixo dela. Temos, porém, de examinar resultados que não sejam tão facilmente convertíveis para desvios-padrão. Suponhamos, por exemplo, que uma criança com pais ansiosos obteve uma nota 64 num teste de leitura. A posição da criança na curva seria a metade da distância, no eixo horizontal, entre o resultado da média (60) e um desvio-padrão acima (68).

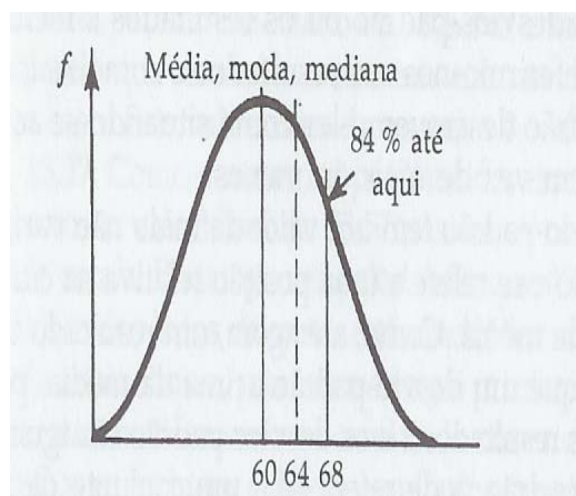


Figura 12: posição de um resultado

A posição da criança é exactamente a meio entre os pontos 60 e 68. Significará isto que a sua posição no grupo é o ponto central entre a média de 50 % e 84 % da nota 68? Isto é, encontrar-se-á a criança acima de 67 % dos colegas? Olhemos cuidadosamente para as duas porções da curva que está dividida pela linha ao nível da nota 64. Serão simétricas? Não - e aqui temos um problema que torna o cálculo de uma posição relativa muito mais complicado e cansativo do que gostaríamos. Quanto mais nos afastamos da média, menos resultados correspondem às diferentes proporções. Assim, se tivermos duas porções entre 60 e 64 e entre 64 e 68, haverá menos resultados neste último intervalo. Haverá ainda menos no intervalo seguinte, entre as notas 68 e 72, e assim sucessivamente. Isto também é verdadeiro para os resultados abaixo da média, mas, neste caso, são os resultados mais elevados, e não os mais baixos, que se encontram mais perto da média. Há muito menos resultados entre 44 e 48 do que entre 48 e 52, apesar de, em ambos os casos, a variação de notas ser de 4 pontos, ou seja, meio desvio-padrão. Quando olhamos para a forma de uma distribuição normal, o tamanho diferente das proporções que cada desvio-padrão compreende parece óbvio. No entanto, o problema de decidir a posição relativa de uma nota 64, quando comparada com os resultados, não desapareceu. Como podemos determiná-la? A resposta é dada através de resultados z . Os resultados z correspondem a desvios-padrão e, na verdade, são virtualmente a mesma coisa, excepto no facto de um resultado z se referir sempre à posição de um ponto em relação a média. Isto vai tornar-se claro em breve. Para já, pensemos que um resultado z de 1 é a mesma coisa que um desvio-padrão de 1, que um resultado z de 2 é um dp 2, e assim por diante. Como não há, virtualmente, nada numa distribuição normal depois do terceiro desvio-padrão ou resultado z - em qualquer das direcções -, é raro que os desvios-padrão ou os resultados z incluam o valor 4. É comum referirmo-nos aos resultados z como mais ou menos; aos desvios-padrão descrevemo-los como situando-se acima ou abaixo da média, em vez de mais ou menos. Um desvio-padrão tem um valor definido não variável, enquanto um resultado z se refere a uma posição relativa na curva e é referido em função da média. Como, até agora, um resultado z tem o mesmo significado que um desvio-padrão acima da média, podemos considerar que os resultados z e os desvios-padrão são iguais. No entanto, um desvio-padrão pode referir-se a um conjunto de resultados que distem um desvio-padrão de qualquer ponto da curva, enquanto os resultados z têm posições fixas. Um resultado z de + 1 corresponde exactamente a um desvio-padrão acima da média, e não a qualquer conjunto de resultados que constituam um desvio-padrão. Voltemos ao problema do resultado de 64 e à sua posição relativa. Sabemos que a sua posição é exactamente metade de um desvio-padrão acima da média, pelo que lhe damos um resultado z de + 0,5.

Há tabelas que nos permitem ver muito facilmente onde os resultados z se situam na curva normal (Anexo VIII). Vejamos como utilizá-las. Utilizaremos o nosso exemplo de 64, cujo valor z é $+0,5$.

Lemos o valor na primeira coluna da esquerda, encabeçada por z , até chegarmos ao valor 0,5. Olhamos para a coluna à direita e vemos o número 19,15. Temos de somar 50%, de modo a obtermos o valor 69,15. Sabemos então que há 69,15% dos resultados abaixo de 64 e 30,85% acima. Devemos arredondar os valores para 69% e 31%, respectivamente. Consideremos outro exemplo, desta vez com o valor 65. Este valor está 5 pontos acima da média e o desvio-padrão para o conjunto é de 8. Um resultado de 5 pontos acima da média é $5/8$ de desvio-padrão acima da média. Se fizermos as contas, sabemos que z é $+0,63$. Como se encontra acima da média, o seu valor é positivo. Voltemos à tabela 51. Como z tem, desta vez, duas casas decimais, os procedimentos vão ser um pouco diferentes. O valor imediatamente à direita (22,7) é a percentagem correcta para um resultado z de 0,6. No entanto, o nosso resultado é 0,63, pelo que temos de andar três colunas da tabela até ao valor 0,03, no topo. Este valor, somado ao valor 0,6, dá-nos o z de 0,63 - ou seja, 23,57. Como o nosso z é positivo, devemos somar-lhe 50% para obtermos o valor final de 73,57. Assim, a nota 65 está à frente de 74% da escala. Podemos ver pela tabela que 49% de todas as notas em cada um dos lados da curva estão incluídas num z de 2,33 ou um bocadinho mais abaixo, para sermos mais precisos). Notemos que, matematicamente, as caudas da curva nunca tocam o eixo horizontal, nem incluem todos os resultados possíveis.

Reparemos agora na posição relativa de uma pessoa que obtenha um resultado abaixo da média, digamos uma nota 41 na amostra original. Esta nota está 19 pontos abaixo da média, apenas um pouco menos do que dois desvios-padrão. Para sermos precisos, está $19/8$ ou 2,375 abaixo. O seu z será $-2,375$. Na tabela SI iem anexo vemos que um z de $+2,3$ inclui 48,93% dos resultados, mas o nosso resultado z é o valor um pouco superior de 2,375. A nossa tabela só pode ser usada com duas casas decimais, pelo que vamos arredondar este valor para 2,38. Paramos, desta vez, junto da coluna de 0,08 e obtemos o valor 49,13. Assim, um z de $+2,38$ inclui $50\% + 49,13\% = 99,13\%$ de todos os resultados. Até agora tudo bem, mas o problema é que o nosso valor era negativo. Basta virarmos a nossa curva ao contrário e trabalharmos com a sua imagem ao espelho. Assim, com o nosso valor $-2,38$ sabemos que 99,13% de todas as notas da distribuição estão acima dele e apenas 0,87% abaixo. Se considerarmos esta pequena proporção de 1%, devemos esperar que, na nossa amostra de 200 indivíduos, 1%, ou seja, dois indivíduos tenham notas inferiores a 41. No

outro extremo das notas, devemos esperar que apenas dois alunos tenham notas de 19 ou mais pontos acima da média, ou seja, notas que excedam os 79%.

O modo de obter o valor z é dado pela expressão formal

$$Z = \frac{\text{desvio da nota em relação à média}}{\text{desvio-padrão}}$$

Se o desvio em relação à média tiver um sinal positivo ou negativo, se estiver acima ou abaixo da média, respectivamente, z ficará com o sinal correcto.

Nota: : tenha cuidado quando trabalhar com z e dp , de modo a usá-los sempre que os dados através dos quais foram obtidos sigam uma distribuição normal Ou aproximadamente normal. De outro modo, arranjará confusões

4. ESTATÍSTICA PARAMÉTRICA E NÃO PARAMÉTRICA

Paramétricos: calcula as diferenças numéricas exactas entre os resultados.

Não paramétricos: apenas consideram se certos resultados são superiores ou inferiores a outros resultados.

REQUISITOS PARA UTILIZAÇÃO DE TESTES PARAMÉTRICOS

Quando se pretende empregar um teste t de Student ou uma análise da variância para fazer comparações entre amostras (testes paramétricos), existe uma lista de requisitos que inclui, entre outros:

1. que a variável tenha sido mensurada num nível mínimo intervalar;
2. que a distribuição seja simétrica e mesocurtica;
3. que a característica estudada (variável) tenha distribuição normal numa dada população.

Sempre que não se pode, honestamente, admitir a simetria e a normalidade de distribuição, ou os dados foram recolhidos num nível de mensuração inferior ao intervalar, devemos recorrer a testes que não incluem a normalidade da distribuição ou nível intervalar de mensuração. Esses testes chamam-se não paramétricos

VANTAGENS DOS TESTES NÃO-PARAMÉTRICOS

Podem ser utilizados, mesmo quando os seus dados só podem ser medidos num nível ordinal, isto é, quando for apenas possível ordená-los por ordem de grandeza) podem ser utilizados mesmo quando os seus dados são apenas nominais, isto é, quando os sujeitos podem apenas ser classificados em categorias.

PODER DE UM TESTE

O poder de um teste é a probabilidade de rejeitarmos a H_0 quando ela é realmente nula

Os testes mais poderosos (os que têm maior probabilidade) de rejeição de H_0 , são testes que possuem pré-requisitos mais difíceis de satisfazer (testes paramétricos como t e F).

As alternativas não paramétricas exigem muito menos pré-requisitos mas produzem testes de significância com menos poder que os correspondentes paramétricos.

EM CONSEQUÊNCIA

Ao rejeitar-se a H_0 sem preencher as exigências mínimas dos testes paramétricos, é mais provável que essa rejeição seja falsa (se rejeitar a H_0 quando ela é verdadeira comete um erro de tipo I; se aceitar a H_0 quando ela é falsa comete um erro de tipo II). Quando os requisitos de um teste paramétrico são violados, torna-se impossível conhecer o seu poder e a sua dimensão (α)

É obvio que os investigadores querem, a todo o custo, rejeitar a H_0 quando ela é mesmo falsa, evitando um erro de tipo I.

O teste ideal seria aquele que $\alpha=0$ e $\beta=1$, o que implicaria que o teste conduziria sempre à decisão correcta, contudo este teste ideal raramente existe.

A probabilidade do erro de 1ª espécie deve ser reduzida, fixando α teórico em 0,1; 0,05 ou 0,01. o valor fixado para α depende da importância que se dá ao facto de rejeitar a H_0 quando esta é verdadeira.

Uma ilustração deste ponto de vista pode ser feita com o seguinte exemplo:

Uma pessoa é inocente até prova do contrário

- H_0 : A pessoa é inocente
- H_1 : A pessoa é culpada

Erro I: A pessoa é condenada mas está inocente

Erro II: A pessoa é absolvida mas é culpada

Naturalmente a justiça procura reduzir a possibilidade de ocorrer o erro de 1ª espécie, pois entende-se que é mais grave condenar inocentes que absolver criminosos.

Para certos sistemas judiciais um $\alpha = 0,1$ é demasiado elevado, optando por $\alpha=0,01$; noutros sistemas judiciais pode admitir que $\alpha=0,05$ é um valor razoável.

ASSIM ...

Fixada a probabilidade do erro de tipo I (dimensão do teste), o teste mais potente é aquele em que a escolha da região crítica minimiza a probabilidade do erro de 2ª espécie. Diz-se também que esta região crítica é a mais potente.

Facilmente se conclui que o teste mais potente é aquele que, uma vez fixada a probabilidade de rejeitar a H_0 , quando ela é verdadeira, maximiza a potência ou a capacidade para rejeitar a mesma hipótese quando esta é falsa.

PRESSUPOSTOS

Para saber se uma variável é simétrica dividimos o coeficiente assimetria (Skewness) pelo erro padrão e se o resultado estiver entre 2 e -2 a distribuição é simétrica.

Para saber se uma variável é mesocurtica dividimos o coeficiente de achatamento (Kurtosis) pelo erro padrão e se o resultado estiver entre 2 e -2 a distribuição é mesocurtica.

Mas se os resultados de um teste paramétrico, não cumpriram com os requisitos (no mínimo dados intervalares; distribuição simétrica, mesocurtica e normal), então não têm interpretação significativa.

Quando acontecem estes factos, a maioria dos investigadores opta por testes de significância não-paramétricos.

PARA ESCOLHER QUALQUER TIPO DE TESTE ESTATÍSTICO

Distinguir se a nossa amostra é constituída pelos mesmos sujeitos em todas as situações ou se é formada por diferentes sujeitos para cada situação

INTER-SUJEITOS OU DESIGN NÃO-RELACIONADO

este tipo de design é utilizado quando um indivíduo ou objecto é avaliado apenas uma vez. A comparação é efectuada entre os grupos de sujeitos/ objectos cujos resultados são não-relacionados.

Desvantagem: conjunto das diferenças individuais na forma como os sujeitos reagem ou respondem à tarefa.

INTRA-SUJEITOS OU DESIGN RELACIONADO

A comparação é feita entre os mesmos sujeitos (sujeitos do mesmo grupo).

A importância destes designs é a eliminação de quaisquer particularidades individuais, uma vez que ficam igualizadas em todas as situações.

Desvantagem: Efeito de memória e aprendizagem.

AMOSTRAS EMPARELHADAS

Igualizam-se sujeitos diferentes mas emparelhados, em termos de idade, sexo, profissão e outras características gerais que parecem importantes para cada pesquisa em particular.

estes tipos de designs podem ser considerados de designs relacionados, uma vez que é controlado nas suas características relevantes.

Desvantagem: Dificuldade em encontrar sujeitos que permitam o emparelhamento de todas as características relevantes.

Dificuldades arranjar grandes amostras.

5. TESTES NÃO PARAMÉTRICOS PASSO-A-PASSO

5.1 TESTE DO QUI-QUADRADO

O Qui-Quadrado (χ^2) de aderência consiste em comparar os dados obtidos experimentalmente com os dados esperados de acordo com a lei. Das comparações surgem diferenças que podem ser grandes ou pequenas: se forem grandes, a hipótese nula (H_0) que pressupõe um bom ajustamento deverá ser rejeitada em favor da hipótese alternativa (H_1); se forem pequenas, a hipótese nula não será rejeitada e as diferenças são atribuíveis ao acaso.

O χ^2 de independência serve para ajudar a decidir se as duas variáveis estão ou não "amarradas" uma à outra por uma relação de dependência.

QUANDO UTILIZAR

Utiliza-se quando os dados são nominais, pelo que em vez de se medirem resultados dos sujeitos apenas se podem distribuir os sujeitos por uma ou mais categorias.

O Qui-Quadrado (χ^2) testa a hipótese experimental que prevê quantos sujeitos de cada grupo são distribuídos por uma determinada categoria.

O χ^2 é um teste estatístico não paramétrico, sendo um dos mais utilizados e bastante aplicado em diferentes planeamentos experimentais. O χ^2 é muito usado mesmo ao nível da estatística multivariada (no sentido de obter o grau de aderência entre o modelo obtido e o teórico).

OBJECTIVO

O objectivo é comparar frequências observadas com frequências teóricas ou esperadas, ou seja, verificar o seu grau de aproximação, que pode ser grande ($=0$) ou pequeno (>0).

Assim, o objectivo é comparar as frequências observadas em cada uma das células de uma tabela de contingência com as diferenças esperadas. O teste compara o número de sujeitos que se distribuem por uma determinada categoria com o número de sujeitos que se esperaria se distribuíssem por essa mesma categoria, caso não existissem diferenças.

O teste do X^2 reflecte o tamanho das diferenças entre as frequências observadas e esperadas. Para ser significativo, o valor de X^2 deverá ser igual ou superior aos valores críticos da tabela (Anexo I).

INSTRUÇÕES PASSO-A-PASSO

1. Calcular as frequências esperadas (E) para cada célula, multiplicando os dois totais parciais relevantes para cada uma e dividindo pelo número total de sujeitos.

2. Calcular X^2 :

$$X^2 = \frac{\sum (O-E)^2}{E}$$

em que:

O = frequências observadas para cada célula;

E = frequências esperadas para cada célula.

3. Calcular os graus de liberdade:

$$g.l. = (r-1) (c-1)$$

em que

r = número de linhas da tabela de contingência

c = número de colunas da tabela de contingência

Se X^2 observado $>$ X^2 crítico rejeita-se H_0

Se X^2 observado $<$ X^2 crítico não se rejeita H_0

Especificando para uma tabela de dupla entrada 2X2:

$$g.l. = (\text{número de colunas} - 1) (\text{número de linhas} - 1) = 1 \times 1 = 1$$

A	B	A+B
C	D	C+D
A+C	B+D	N

$$X^2 = \frac{N (AD - CB)^2}{(A+C) (C+D) (A+C) (B+D)}$$

Quando uma tabela tem mais de duas entradas:

g.l. = (número de colunas - 1) (número de linhas - 1)

$$X^2 = \frac{(\text{total de linha}) (\text{total de coluna})}{(\text{total geral})}$$

Exemplo: Suponha que quer estudar se os estudantes de ciências sociais utilizam um método de estudo significativamente diferente daquele que é utilizado pelos estudantes de tecnologia. Escolheram-se dois grupos, um composto por 50 estudantes de ciências sociais e o outro por 50 estudantes de tecnologia. Enviou-se um questionário aos 100 estudantes pedindo-lhes que indicassem em qual das seguintes categorias melhor se enquadra o seu método de estudo: estudo diário regular, estudo irregular em períodos concentrados de trabalho intensivo ou uma mistura dos dois. Foram recebidas 44 respostas dos estudantes de ciências sociais e 42 dos estudantes de tecnologia. Colocou-se a hipótese experimental (H1) de que o tipo de estudo dependia do curso frequentado

Os resultados são apresentados na forma de uma tabela 2X3, designada por tabela de contingência (crosstab). As "células" que representam cada uma das categorias são numeradas de 1 a 6.

Quadro 12: Tabela de Contingência (Crosstab)

	Tipo de estudo			Total parcial do número de estudantes
	Regular	Irregular	Misto	
Grupo 1 Estudantes de Ciências Sociais	1 6 E=8,19	2 15 E=11,77	3 23 E=24,05	44
Grupo 2 Estudantes de Tecnologia	4 10 E=7,81	5 8 E=11,23	6 24 E=22,95	42
Totais dos padrões de estudo	16	23	47	N=86

Resolução 49:

Instruções Passo-a-Passo:

1. calcular as frequencia esperadas (E)

$$\text{Célula 1: } E = 16 \times 44 / 86 = 8,19$$

$$\text{Célula 2: } E = 23 \times 44 / 86 = 11,77$$

$$\text{Célula 3: } E = 47 \times 44 / 86 = 24,05$$

$$\text{Célula 4: } E = 16 \times 42 / 86 = 7,81$$

$$\text{Célula 5: } E = 23 \times 42 / 86 = 11,23$$

$$\text{Célula 6: } E = 47 \times 42 / 86 = 22,95$$

2. Aplicar a fórmula:

$$X^2 = \frac{\sum (O-E)^2}{E}$$

$$X^2 = \frac{(6-8,19)^2}{8,19} + \frac{(15-11,77)^2}{11,77} + \frac{(23-24,05)^2}{24,05} + \frac{(10-7,81)^2}{7,81} + \frac{(8-11,23)^2}{11,23} + \frac{(24-22,95)^2}{22,95}$$

$$X^2 = 0,59 + 0,89 + 0,05 + 0,61 + 0,93 + 0,05 = 3,12$$

3. Calcular os graus de liberdade (gl)

$$g.l. = (r - 1) (c - 1) = (2 - 1) (3 - 1) = 2$$

4. Consultar a tabela (Anexo I)

Para

p=0,05 e gl=2

x² critico=5,99

Conclusões: Dado que o valor observado de X² é apenas de 3,12, ou seja, inferior ao valor crítico de 5,99 para p < 0,05, o resultado da experiência não é significativo. Aceita-se hipótese nula de que os padrões de estudo dos estudantes de ciências sociais e de tecnologia não diferem, rejeitando-se desta forma a nossa hipótese experimental (H1).

5.2 PROCEDIMENTOS PARA ORDENAÇÃO DE RESULTADOS

Os testes U de Mann-Whitney; Wilcoxon; H de Kruskal-Wallis; rho de Spearman; tau de Kendall e Friedman, exigem o recurso a ordenações de resultados para efectuar os seus cálculos. Neste sentido começaremos por explicar os procedimentos de ordenação de resultados

Ordenamento global de resultados (designs não-relacionados para sujeitos diferentes):

1. Atribua a ordem mais baixa 1 ao resultado mais baixo, a ordem seguinte 2 ao resultado seguinte mais baixo, e por aí fora.
2. Os resultados nulos de 0 são contados como o resultado mais baixo possível, sendo-lhe atribuída a ordem mais baixa.
3. Quando existem resultados idênticos são-lhe atribuídas ordens médias, calculadas com base na globalidade das ordens que deviam ter sido atribuídas a estes resultados.

Exemplo: ordenar um grupo em função do n.º de factores de risco:

Quadro 13: Ordenação de um grupo em função do n.º de factores de Risco

N.º de Factores de risco	Ordem
6	4
3	1
12	7
4	2
7	5
5	3
8	6

Se existir um zero este deve ser considerado o valor mais baixo.

Nos casos em que existem resultados iguais utiliza-se a \bar{x} dos lugares que devia ocupar.

Exemplo:

N.º de FR.		Ordem
1		2
2		4
1		2
4		6,5
1		2
3		5
4		6,5
6		9
5		8

Assim os sujeitos com um Factor de risco são 3 (1+1+1) que ocupariam o 1.º - 2.º - 3.º lugar

Então $3+2+1=6:(1+1+1)=2$

Com 4 factores de risco temos 2 sujeitos que ocupariam o 6.º e 7.º lugar

Então $6+7=13:(1+1)=6,5$

Ordenamento de diferenças entre resultados (designs relacionados para os mesmos sujeitos ou emparelhados):

1. Em geral, a atribuição de ordens às diferenças entre resultados efectua-se tal como fizemos para os resultados, sendo atribuída a ordem mais baixa à menor diferença e por aí adiante;
2. Diferenças idênticas entre resultados são ordenadas da mesma forma que resultados idênticos, atribuindo-se uma ordem média resultante da globalidade de ordens que essas diferenças deveriam ocupar;
3. Resultados nulos de 0 são contados como o resultado mais baixo possível quando se calculam diferenças entre resultados;
4. Contudo, quando existe igualdade entre resultados que originem uma diferença nula entre as situações experimentais, estes não são ordenados, sendo retirados da análise;
5. Diferenças positivas e negativas são ordenadas em conjunto como se se tratasse de um ordenamento simples de resultados, ignorando os sinais positivos e negativos.

Exemplo: Suponha que quer ordenar as diferenças entre o número de frases correctas que um grupo de crianças com suspeita de perda auditiva produziu antes da colocação de um aparelho auditivo e após a colocação daquele.

Quadro 14: Ordenação para Amostras Relacionadas ou Emparelhadas

Sujeitos	Nº de frases correctas antes do aparelho auditivo	Nº de frases correctas depois do aparelho auditivo	diferenças	ordem
1	5	6	1	2
2	5	7	2	4
3	2	3	1	2
4	1	5	4	6,5
5	4	5	1	2
6	2	5	3	5
7	1	5	4	6,5
8	4	4	0	-
9	1	7	6	9
10	1	6	5	8

Ao contrário do que acontece nos casos das amostras relacionadas quando a diferença entre 2 situações é nula nas amostras relacionadas a este tipo de resultado não é atribuída nenhuma ordem, sendo que o resultado nem sequer é considerado na análise. No ordenamento de resultados negativos em amostras relacionadas ignoram-se os sinais quando se ordenam os resultados.

5.3 TESTES PARA DUAS AMOSTRAS INDEPENDENTES

Ao estudarmos as diferenças entre dois grupos podemos utilizar grupos relacionados/emparelhados ou grupos independentes. No caso de duas amostras independentes determinamos se as diferenças nas amostras constituem uma evidência convincente de uma diferença nos processos de tratamento a elas aplicados.

Conquanto o uso de duas amostras relacionadas em projectos de pesquisa tenha méritos indiscutíveis, a sua aplicação, em geral, não é prática. Frequentemente, a natureza da variável dependente impede a utilização dos indivíduos como seus próprios controlos, tal como ocorre quando a variável dependente é o suicídio tentado; um problema que pode acontecer uma única vez. Pode ser também impossível delinear um projecto que utilize pares de dados, talvez por

desconhecimento, por parte do investigador, de variáveis úteis que possam formar pares, ou pela impossibilidade de obter mensurações adequadas de alguma variável de reconhecida importância ou, enfim, porque simplesmente não se dispõe de “pares” adequados.

Quando a utilização de duas amostras relacionadas não é prática ou adequada, podemos utilizar duas amostras independentes. Em tais projectos, as duas amostras podem ser obtidas por um de dois métodos:

- a) Podem ser extraídas aleatoriamente de duas populações;
- b) Podem decorrer da atribuição aleatória de dois tratamentos aos membros de uma amostra de origem arbitrária.

Nota: Em nenhum destes casos se exige que as amostras tenham o mesmo tamanho.

5.4 TESTE U DE MANN-WHITNEY

Quando Utilizar:

Dadas duas amostras, de tamanhos n_1 e n_2 , é possível, mediante a prova U de Mann-Whitney, saber se ambas as amostras podem ser consideradas provenientes da mesma população.

Como já se sabe, a estatística paramétrica só pode ser usada desde que os dados tenham sido mensurados, no mínimo, no nível intervalar. Além disso, as amostras devem ser aleatórias, independentes e a variável observacional precisa de ter distribuição normal na população.

O teste U de Mann-Whitney deve ser utilizado em *designs* com **duas situações**, não-relacionados, quando são utilizados sujeitos **diferentes** em cada uma das situações experimentais.

Exemplo: Suponha que quer investigar o n.º de queixas dolorosas durante um tratamento a um mesmo problema terapêutico (controlada a gravidade e a extensão da lesão) em que se utilizaram duas técnicas diferentes. A hipótese experimental supõe que é durante a utilização da técnica B que o doente se apresenta mais queixoso.

Quadro 15: Resultados do registo das queixas

Paciente	A	Ordem (1)	B	Ordem (2)
1	3	3	9	11
2	4	4	7	9
3	2	1,5	5	5,5
4	6	7,5	10	12
5	2	1,5	6	7,5
6	5	5,5	8	10
TOTAL	22	T1=23	45	T2=55
MÉDIA	3,67		7,5	

RACIONAL

O racional que está por trás do teste U de Mann-Whitney é bastante semelhante ao do teste de Wilcoxon. A diferença fundamental entre as duas reside no facto do segundo se aplicar a designs relacionados e o U se aplicar a designs não-relacionados, utilizando, portanto, sujeitos diferentes. O teste de Wilcoxon analisa as diferenças entre a performance dos mesmos sujeitos (ou pares de sujeitos emparelhados) submetidos a duas situações experimentais. Com um design não-relacionado não temos uma base que nos permita comparar diferenças entre pares de resultados. Assim, o teste U de Mann-Whitney ordena os resultados de todos os sujeitos em ambas as situações como se fossem apenas um conjunto simples de resultados.

Se as diferenças entre as situações forem aleatórias, como é postulado pela hipótese nula, então os resultados devem ser aproximadamente os mesmos e, consequentemente, as ordens devem ser também aproximadamente as mesmas para as duas situações. Se houver uma preponderância de ordens altas ou baixas numa situação ou na outra, então é porque a diferença no total dos resultados ordenados para cada situação é devida aos efeitos previstos da variável independente e não ao acaso. Se a soma total das ordens for muito baixa para uma das situações, então terá de haver uma preponderância de ordens elevadas na outra situação. *Quanto menor for U mais significativas serão as diferenças entre as ordens das duas situações.*

INSTRUÇÕES PASSO-A-PASSO PARA CALCULAR W

Ordene todos os resultados para ambos os grupos como se se tratasse de um conjunto único de resultados, atribuindo a ordem 1 ao resultado inferior e assim sucessivamente.

O ordenamento global de todos os resultados é apresentado em Ordens(1) e Ordens (2).

Adicione as ordens totais para o grupo 1 e grupo 2 em separado.

$$T1=23 \text{ e } T2=55$$

Selecione o maior total das ordens.

$$T2=55$$

Calcule o valor de U através da fórmula

$$U = n_1.n_2 + \frac{N_x (n_x + 1)}{2} - T_x$$

em que:

n_1 =número de sujeitos no grupo 1

$$n_1=6$$

n_2 =número de sujeitos no grupo 2

$$n_1=6$$

T_x =maior total de ordens

$$T_x=T_2=55$$

N_x =número de sujeitos do grupo com o maior total de ordens

$$N_x=6$$

Cálculo de U

$$U = 6 \times 6 + \frac{6 \times 7}{2} - 55 = 36+21-55=2$$

Quando existe o mesmo número de sujeitos em cada situação rapidamente se torna mais fácil verificar qual o maior total de ordens. Normalmente, é preferível ter um número idêntico de sujeitos em cada grupo; no entanto, se tiver de utilizar um número desigual de sujeitos em cada grupo, poderá fazê-lo. Em caso de dúvida, calcule U para ambas as ordens totais, seleccionando o n apropriado a cada caso e, depois, considere o U mais pequeno.

CONSULTA DA SIGNIFICÂNCIA NA TABELA

As Tabelas (Anexo II) apresentam-lhe os valores críticos de U nos diferentes níveis de significância de testes unicaudais e bicaudais, para as diferentes combinações de n_1 e n_2 dos dois grupos.

O procedimento mais usual é começar por verificar se o valor de U é significativo ao nível de significância de $p < 0,05$ para testes unicaudais) ou para testes bicaudais. Uma vez que previmos que a técnica B produzia mais queixas, poderemos consultar a Tabela para testes unicaudais. Localizando $n_1=6$ na linha superior e $n_2=6$ na coluna do lado esquerdo, encontraremos o valor crítico de U na intercepção dos dois.

Conclusão: Uma vez que se convencionou utilizar o nível inferior de U, o nosso valor de $U=2$ deve ser igual ou inferior ao valor crítico 7, o que acontece, podemos, desta forma, rejeitar a hipótese nula e aceitar que existem diferenças significativas ao nível da dor ($p < 0,05$).

6. TESTES PARA DUAS AMOSTRAS RELACIONADAS

Empregam-se as provas estatísticas de duas amostras quando o investigador deseja determinar se dois tratamentos são diferentes ou se um tratamento é “melhor” do que o outro. Em cada caso, compara-se o grupo em que se aplicou o tratamento com outro que não sofreu nenhum tratamento ou que sofreu tratamento diferente.

Em tais comparações de dois grupos observam-se, por vezes, diferenças significativas que não são resultantes do tratamento aplicado.

Uma das maneiras de superar a dificuldade decorrente da introdução de diferenças “extrínsecas” entre dois grupos consiste em utilizar na pesquisa duas amostras relacionadas, isto é, relacionar de alguma forma as duas amostras estudadas. Tal relacionamento pode ser conseguido utilizando-se cada indivíduo como seu próprio controlo ou então formando pares de indivíduos e, em seguida, associando os dois membros de cada par às duas condições. Quando um indivíduo “serve como o seu próprio controlo”, ele é submetido a ambos os tratamentos em ocasiões diferentes. Quando se utiliza o método do emparelhamento devem procurar seleccionar-se, para cada par, indivíduos que sejam tão semelhantes quanto possível em relação a quaisquer variáveis extrínsecas que possam influenciar os resultados da pesquisa.

Sempre que possível, o método de utilização do indivíduo como o seu próprio controlo (contrabalançando a ordem em que se aplicam os tratamentos ou métodos) é preferível ao método de emparelhamento. E a razão disso é que é limitada a nossa capacidade para formar os pares adequadamente, em consequência do nosso desconhecimento das variáveis relevantes que determinam o comportamento. A validade por emparelhamento está na razão directa do investigador para determinar como formar os pares, e essa capacidade é quase sempre muito limitada. Essa dificuldade é contornada quando se utiliza cada indivíduo como seu próprio controlo; não se pode pretender relacionamentos mais precisos do que a própria identidade.

6.1 PROVA DE McNEMAR PARA A SIGNIFICÂNCIA DE MUDANÇAS

Quando utilizar:

A prova de McNemar para a significância de mudanças é particularmente aplicável aos **planeamentos do tipo “antes e depois”**, em que **cada indivíduo é utilizado como o seu próprio controlo** e a **mensuração se faz ao nível de uma escala nominal ou ordinal**. Pode, assim, ser usada para testar a eficiência de determinada técnica (reuniões, folhetos, visita, etc.) sobre as preferências eleitorais a respeito de vários candidatos.

Nestes casos, cada pessoa pode servir como o seu próprio controlo, utilizando-se a mensuração em escala nominal para avaliar as alterações da situação “após” em relação à situação “antes”.

Exemplo: Suponha que um profissional de saúde está interessado em estudar os comportamentos resultantes da iniciação de obesos à prática do exercício físico. Este profissional observou ao longo dos anos que os obesos utilizam preferencialmente o elevador para se dirigirem à sua consulta cujo consultório era no 1.º andar. Coloca a hipótese de que os obesos que tiveram como terapêutica exercício físico começariam a usar preferencialmente as escadas. A fim de testar a hipótese o técnico observa 25 doentes em que ministrou como exercício físico caminhar uma hora por dia cinco vezes por semana. Decorrido um mês de exercício ele observa os mesmos 25 doentes e faz a classificação comportamentos. Os dados são os seguintes:

Quadro 16: Teste McNemar

Preferencia após 30º dias de exercício

Preferência antes da terapêutica	Preferencia após 30º dias de exercício	
	Escadas	Elevador
Elevador	4 (A)	14 (B)
Escadas	4 (C)	3 (D)

HIPÓTESE DE NULA:

H0: Para os obesos que modificaram a sua atitude, a probabilidade de mudar o percurso do elevador para as escadas (PA) é igual à probabilidade de mudar de mudar das escadas para o elevador (PD) e ambas são iguais a $\frac{1}{2}$. Isto é,

$$H0: PA=PD=1/2$$

$$H1: PA>PD$$

PROVA ESTATÍSTICA:

Utiliza-se a prova de McNemar para significância de mudanças, porque o estudo utiliza duas amostras relacionadas, é do tipo antes-e-depois e utiliza a escala de medida nominal (classificativa).

Nível de significância:

$p=0,05$ $N=25$

Distribuição amostral:

A distribuição Qui-quadrado com 1 grau de liberdade dá uma boa aproximação da distribuição amostral de Qui-quadrado, tal como calculada pela fórmula.

Região de Rejeição:

Como H_1 especifica o sentido da diferença prevista, a região de rejeição é unilateral. Consiste de todos os valores de Qui-quadrado que são tão grandes que acusem uma probabilidade unilateral, associada à sua ocorrência sob H_0 não superior a 0,05.

Decisão:

Estamos interessados nos obesos cujo comportamento acusa alteração: representados nas células A e D.

Para os dados, temos:

$$X^2 = \frac{(|A-D| - 1)^2}{A+D} = \frac{(|4-3| - 1)^2}{4+3} = \frac{(|1| - 1)^2}{7} = \frac{0}{7} = 0$$

Sendo 0 valor observado de $x^2=0$, devemos consultar a tabela (Anexo I) para obter o valor crítico, não esquecendo que temos uma amostra unicaudal a um nível de 0,05. consultando a tabela observamos um x^2 crítico de 5,41.

Racional

Para comparar a significância de qualquer mudança observada, por este método, constrói-se uma tabela de frequências de 4 casas para representar o 1º e o 2º conjunto de reacções dos mesmos

indivíduos. As características gerais desta tabela são as que se apresentam a seguir, em que se utilizam os sinais “+” e “-” para indicar diferentes reacções.

Quadro 17: Tabela de quatro casas para a prova de significância de mudanças

		ANTES	
		-	+
DEPOIS	+	A	B
	-	C	D

Note-se que os casos que acusam modificações entre a 1ª e a 2ª reacção aparecem nas células A e D. Um indivíduo é localizado na célula A se passou de “+” para “-” e na célula D se passou de “-” para “+”. Na ausência de modificação, o indivíduo é classificado na célula B (reacção “+” antes e depois) ou na célula C (reacção “-” antes e depois).

Como A e D representa o número total de indivíduos que acusam modificação, a perspectiva, sob a hipótese de nulidade, seria que $\frac{1}{2} (A+D)$ acusassem modificações num sentido e $\frac{1}{2} (A+D)$ acusassem modificações noutro sentido. Ou seja, $\frac{1}{2} (A+D)$ é a frequência esperada, sob H_0 , tanto na célula A como na célula D.

Na prova de McNemar de significância de mudança, estamos interessados apenas nas células A e D. Portanto, A=número de casos observados na célula A, D=número de casos observados na célula D e $\frac{1}{2} (A+D)$ =número esperado de casos tanto nas células A como D, então

$$X^2 = \frac{(A-D)^2}{A+D} \quad \text{com graus de liberdade}=1$$

6.2 CORRECÇÃO DE CONTINUIDADE

A aproximação, pela distribuição do Qui-quadrado, da distribuição amostral da fórmula torna-se excelente se se introduzir uma correcção de continuidade. Tal correcção é necessária, porque se utilizou uma distribuição contínua (Qui-quadrado) para aproximar uma distribuição discreta. Quando todas as frequências esperadas são pequenas, tal aproximação pode ser fraca. A correcção de continuidade (Yates, 1934) constitui uma tentativa de remoção dessa fonte de erro.

Com a correcção de continuidade, tem-se:

$$X^2 = \frac{(|A-D| - 1)^2}{A+D} \quad \text{com graus de liberdade}=1$$

Esta expressão indica que se deve subtrair 1 do valor absoluto da diferença entre A e D antes de elevar ao quadrado. O grau de significância de qualquer valor observado de Qui-quadrado, tal como calculado através da fórmula, é determinado mediante referência a uma Tabela (Anexo I). Se o valor observado de Qui-quadrado é igual a, ou maior do que, o valor exibido na Tabela para determinado nível de significância com gl=1, a implicação é que existe efeito “significativo” nas reacções “antes” e “depois”.

INSTRUÇÕES PASSO-A-PASSO PARA CALCULAR X^2

Enquadrar as frequências observadas numa tabela de 4 casas.

Determinar as frequências esperadas nas células A e D

$$E = 1/2 (A+D)$$

Se as frequências esperadas são inferiores a 5, empregar a prova binomial em substituição à prova de McNemar.

Se as frequências esperadas não são inferiores a 5, calcular o valor de X^2 através da fórmula

$$X^2 = \frac{(|A-D| - 1)^2}{A+D}$$

Mediante referência à Tabela (Anexo I), determinar a probabilidade, sob H_0 , associado a um valor tão grande quanto o valor observado de X^2 . Se se trata de uma prova unilateral, dividir por

2 o valor da probabilidade exibido na Tabela. Se o valor de p , dado pela Tabela para o valor observado de X^2 com $gl=1$, não supera p , rejeita-se H_0 em favor de H_1 .

6.3 TESTE DOS SINAIS DE WILCOXON

Quando utilizar

O teste de Wilcoxon deve ser utilizado num *design* experimental relacionado, com **duas situações** experimentais quando são utilizados os **mesmos sujeitos** ou sujeitos **emparelhados** em ambas as situações.

Exemplo: Suponha que quer investigar se existe alguma diferença na quantidade de vocabulário utilizado por crianças que usam um aparelho auditivo ou por crianças que não usam. Este é um bom exemplo dum caso em que é essencial a utilização de sujeitos emparelhados. Como é óbvio, não é possível utilizar os mesmos sujeitos, uma vez que nenhuma criança que não precisa de usar aparelho auditivo usa um mesmo tempo. Por outro lado, não podemos escolher aleatoriamente os sujeitos para cada grupo. Pode dar-se o caso, por exemplo, de os sujeitos que usam aparelho auditivo serem mais velhos. Qualquer efeito encontrado neste grupo pode ficar a dever-se unicamente a esta diferença. Os dois grupos “com aparelho” e “sem aparelho” necessitam de ser emparelhados em termos de idade, sexo, inteligência e todas as outras variáveis que achemos necessário serem controladas. Apresentamos depois às crianças um teste que meça o seu vocabulário, traduzindo-o em resultados, tal como é mostrado na tabela seguinte.

Quadro 18: Resultados do teste do vocabulário

Par de sujeitos	Situação A (com aparelho)	Situação B (sem aparelho)	d (A-B)	Ordem de d	Ordem das diferenças positivas	Ordem das diferenças negativas
1	3	5	-2	5 (-)		5
2	4	5	-1	2 (-)		2
3	3	2	+1	2 (+)	2	
4	1	5	-4	8,5 (-)		8,5
5	5	4	+1	2 (+)	2	
6	2	5	-3	7 (-)		7
7	3	5	-2	5 (-)		5
8	4	4	0	-		
9	1	5	-4	8,5 (-)		8,5
10	3	5	-2	5 (-)		5
TOTAL	29	45			4	41
MÉDIA	2,9	4,5				

RACIONAL

O objectivo do teste dos sinais de Wilcoxon é comparar as *performances* de cada sujeito (ou pares de sujeitos) no sentido de verificar se existem diferenças significativas entre os seus resultados nas duas situações. Os resultados da Situação B são subtraídos dos da Situação A e à diferença resultante (d) é atribuído o sinal mais (+) ou, caso seja negativa, o sinal menos (-). Estas diferenças são ordenadas em função da sua grandeza (independentemente do sinal positivo ou negativo). O ordenamento assim obtido é depois apresentado separadamente para os resultados positivos e negativos. O menor dos valores deste segundo, dá-lhe o valor de uma “estatística” designada por W , que pode ser consultada na Tabela de significância apropriada.

A ideia é que se existirem apenas diferenças aleatórias, tal como é postulado pela hipótese nula, então haverá aproximadamente o mesmo número de ordens elevadas e de ordens inferiores tanto para as diferenças positivas como negativas. Se se verificar uma preponderância de baixos resultados para um dos lados, isso significa a existência de muitos resultados elevados para o outro lado, indicando uma diferença em favor de uma das situações, superior àquilo que seria de esperar se os resultados se devessem ao acaso. Dado que a estatística W reflecte o menor total de ordens, quanto menor for o W mais significativas serão as diferenças nas ordenações entre as duas situações.

INSTRUÇÕES PASSO-A-PASSO PARA CALCULAR W

1. Calcule a diferença d entre cada par de resultados, atribuindo o sinal mais ou menos.

Veja a coluna $d(A-B)$

2. Ordene as diferenças por ordem de grandeza desde a ordem inferior até à superior, ignorando os sinais positivos e negativos.

Veja a coluna ordenamento de d

3. Em separado, junte também a ordenação correspondente aos sinais diferentes (+ ou -).

Veja os totais para ordenamentos de diferenças positivas e de diferenças negativas nas respectivas colunas

4. Considere o menor dos totais das ordens como W.

Valor observado de $W=4$, uma vez que o total das ordens para as diferenças positivas é o menor

5. Conte o número de pares de sujeitos N (não considere as igualdades).

$$N=10-1=9$$

CONSULTA DA SIGNIFICÂNCIA NA TABELA

A tabela anexada (Anexo III) apresenta-lhe o nível de significância de w tanto para os testes unicaudais como bicaudais. Na coluna da esquerda encontra os valores de N. Uma vez que não efectuámos uma previsão da direcção (como por exemplo, que obteríamos melhores resultados no vocabulário de criança em jardim de infância) teremos de utilizar os níveis de significância para uma hipótese bicaudal. Seleccione o valor adequado $N=9$ e verifique ao longo dessa linha se o valor de W é significativo. Uma vez que se convencionou utilizar o menor valor das ordens, então o valor obtido de W terá de ser igual ou inferior ao valor crítico da Tabela. Como o valor obtido $W=4$ é inferior ao valor crítico de 6 para $p<0,05$ (bicaudal), pode rejeitar a hipótese nula e concluir que existe uma diferença significativa entre os resultados no vocabulário dos dois grupos de sujeitos emparelhados.

Suponha que tinha efectuado uma previsão numa dada direcção, por exemplo, que as crianças que usam aparelho auditivo (Situação B) obtêm resultados mais elevados no teste de vocabulário. O valor obtido de $W=4$, é inferior a 6, que é o valor crítico de W para $p<0,025$ (hipótese unicaudal), uma probabilidade inferior e, consequentemente, mais significativa do que o nível de significância para uma hipótese bicaudal $p<0,05$.

7. TESTES PARA K AMOSTRAS INDEPENDENTES

O investigador pode precisar de decidir se diversas variáveis independentes devem ser consideradas como procedentes da mesma população. Os valores amostrais quase sempre são um tanto diferentes e o problema é determinar se as diferentes amostras observadas sugerem realmente diferenças entre as populações ou se são apenas variações casuais que podem ser esperadas entre amostras aleatórias da mesma população.

Apresentamos técnicas para **comparar a significância de diferenças entre três ou mais grupos de amostras independentes**, ou seja, para comprovar a hipótese de nulidade de que K amostras independentes tenham sido extraídas da mesma população ou de populações idênticas.

As provas não-paramétricas têm a vantagem de permitir estudar, quanto à significância, dados que são inerentemente classificados (escala nominal) ou se apresentam em postos (escala ordinal).

7.1 TESTE DE KRUSKAL-WALLIS

REQUISITOS PARA O USO DO TESTE DE KRUSKAL-WALLIS

O teste de Kruskal-Wallis pressupõe as seguintes condições para o seu adequado uso:

- a) Comparação de três ou mais amostras independentes;
- b) O teste de Kruskal-Wallis não pode ser usado para testar diferenças numa única amostra de respondentes mensurados mais de uma vez;
- c) Dados cujo nível de mensuração seja no mínimo ordinal;
- d) Esta prova exige dados que possam ser ordenados e aos quais, por isso mesmo, seja possível atribuir postos ou ordens;
- e) O tamanho mínimo de cada amostra deve ser de 6 para se poder recorrer ao χ^2 .

Quando $n > 5$ por grupo de respondentes, a significância de H pode ser determinada por recorrência à Tabela do Qui-quadrado (Anexo I). Para testar diferenças entre amostras de tamanho inferior a 6, deve recorrer a tabelas especiais (Anexo IV).

QUANDO UTILIZAR

Este teste pode ser considerado uma extensão do teste U de Mann-Whitney quando necessitamos de utilizar **três** ou **mais** situações. Deve ser utilizado em *designs não-relacionados* quando sujeitos **diferentes** são distribuídos por **três** ou **mais** situações.

Exemplo: Suponha que estamos interessados em descobrir se existem diferenças no acesso a uma página da internet em função da característica: muito ilustrada, com algumas ilustrações e, sem ilustrações. Alocámos três páginas na internet com o mesmo assunto e título durante 4 meses. A seguir verificámos o número de vezes que foram acedidas durante quatro sábados seguidos. Os resultados foram.

Quadro 19: Número de ideias relembradas para três tipos de testes

	Sujeitos do grupo 1 (página muito ilustrada)		Sujeitos do grupo 2 (página com algumas ilustrações)		Sujeitos do grupo 3 (Página sem ilustrações)	
	Resul	Ordem	Resul	Ordem	Result.	Ordem
Sabado 1	19	10	14	6	12	3,5
Sabado 2	21	11	15	7	12	3,5
Sabado 3	17	9	9	1	13	5
Sabado 4	16	8			10	2
TOTAL	73	38	38	14	47	14
MÉDIA	18,25		12,67		11,75	

RACIONAL

Este teste pretende determinar se os resultados são significativamente diferentes para três ou mais grupos. Uma vez que todos os resultados foram, em princípio, obtidos por sujeitos diferentes a única forma de verificarmos as diferenças entre as situações é ordená-las em conjunto, como se se tratassem apenas de um conjunto de resultados, tal como havíamos efectuado no teste U de Mann-Whitney. Isto acontece porque, não temos uma base para comparar resultados dos mesmos sujeitos ou de sujeitos emparelhados em diferentes situações, como com o teste U de Mann-Whitney para *designs* relacionados.

Este ordenamento global, quando posteriormente adicionamos as ordens de cada coluna em separado, permite-nos obter o total das ordens para cada situação. Se existirem apenas diferenças aleatórias entre as situações, como é postulado na hipótese nula, é de esperar que ordens altas e

baixas se distribuam de forma aproximadamente equivalente pelas diferentes situações. Mas, se pelo contrário, houver uma preponderância de altos ou baixos resultados em qualquer uma das situações, é provável que tal facto reflecta diferenças significativas devidas à variável independente.

O valor das diferenças entre os totais das ordens é dado pela estatística designada por H. Desde que a hipótese experimental preveja a existência de diferenças significativas entre as situações, o valor que obtivermos de H deverá ser igual ou superior ao valor crítico da Tabela, para que possa ser considerado significativo.

INSTRUÇÕES PASSO-A-PASSO PARA CALCULAR H

Ordene todos os grupos do design como se se tratasse apenas de um conjunto de resultados, atribuindo a Ordem 1 ao menor resultado e assim sucessivamente.

Para um ordenamento global dos resultados, veja as colunas ordem para os grupos 1, 2 e 3, em que todos os resultados são considerados em conjunto

Adicione os totais das ordens para cada situação.

Calcule o valor de H a partir da fórmula

$$H = \frac{12}{N(N+1)} \sum \frac{T_c^2}{n_c} - 3(N+1)$$

em que:

N=número total de sujeitos

N=11

n_c =número de sujeitos em cada grupo

$n_1=4; n_2=3; n_3=4$

T_c =total de ordens para cada situação, ou seja,

os totais das ordens para cada coluna

$T_1=38; T_2=14; T_3=14$

T_c^2 =total das ordens para cada situação, cada um elevado ao quadrado

$T_1^2=38^2; T_2^2=14^2; T_3^2=14^2$

$\sum T_c^2$ = soma dos quadrados dos totais das ordens para cada situação dividido
 n_c pelo número de sujeitos dessa situação:

$$38^2/4 + 14^2/3 + 14^2/4$$

Cálculo de H

$$H = \frac{12}{N(N+1)} \sum \frac{T_c^2}{n_c} - 3(N+1)$$

$$H = \frac{12}{11(11+1)} (38^2/4 + 14^2/3 + 14^2/4) - 3(11+1)$$

$$H = \frac{12}{132} (1444/4 + 196/3 + 196/4) - 3 \times 12$$

$$H = 0,091 (361 + 65,33 + 49) - 36$$

$$H = 43,255 - 36 = 7,26$$

Calcule os graus de liberdade, ou seja, o número de situações (C) menos uma.

$$gl = C - 1 = 3 - 1 = 2$$

CONSULTA DA SIGNIFICÂNCIA NA TABELA

A Tabela (Anexo IV) utiliza-se em experiências com três grupos de sujeitos, e com um máximo de cinco sujeitos em cada grupo. Para um maior número de sujeitos, deve ser utilizada a Tabela do Qui-quadrado (Anexo I). Quando não são utilizados mais de três grupos, poderá localizar na coluna da esquerda da Tabela o número de sujeitos de cada grupo. Localize então a combinação que procura (no nosso caso: 4, 4, 3). Note que a ordem do número de sujeitos não é importante, mas a combinação apropriada na Tabela é 4, 4 e 3. Para essa combinação encontrará os valores críticos de H para várias probabilidades. Se o valor de H que obteve for igual ou superior ao valor crítico de um determinado nível de significância pode rejeitar a hipótese nula. No nosso exemplo, o valor obtido de $H=7,26$ é superior ao valor crítico de 7,1439 para $p<0,01$, pelo que podemos aceitar a hipótese experimental a este nível de significância.

Se possuir mais de três situações, e/ou sujeitos em cada situação, deverá procurar o valor crítico na Tabela do Qui-quadrado. Repare que para isso terá que calcular os graus de liberdade. Localize os valores dos graus de liberdade (no nosso exemplo, $gl=2$) ao longo da coluna do lado esquerdo e verifique ao longo da linha os valores críticos para as diferentes probabilidades. O

valor que obtivemos $H=7,26$ é superior ao valor crítico de 5,99 para $p \leq 0,05$, pelo que podemos aceitar que o resultado é significativo a este nível. Dará conta que esta probabilidade é menos significativa de que quando utilizamos a Tabela anterior. Isso acontece porque essa Tabela é especialmente concebida para nos dar as probabilidades com um pequeno número de sujeitos e de situações.

Notará também que o teste de Kruskal-Wallis apenas lhe pode dizer que existem diferenças globais nos resultados entre as situações experimentais. Na tabela apresentada parece existir uma tendência para consultar páginas com mais ilustrações do que sem ilustrações. Mas para poder testar se essa tendência realmente existe, terá de utilizar um teste de tendência.

8. TESTES PARA K AMOSTRAS RELACIONADAS

São estatísticas para **comprovar a significância de diferenças entre três ou mais grupos**, ou seja, para comparar a hipótese de nulidade de que K (3 ou mais) amostras tenham sido extraídas da mesma população ou de populações idênticas.

As circunstâncias exigem, por vezes, o recurso a um experimento que nos permita estudar simultaneamente mais de duas amostras ou condições. Quando se trata de comparar três ou mais amostras ou condições de um experimento, é necessário aplicar uma prova estatística que indique se há alguma diferença global entre as K amostras ou condições, antes que possamos cogitar de comprovar a significância da diferença entre duas amostras quaisquer.

Só quando uma prova global (prova de K amostras) nos autoriza a rejeitar a hipótese nula é que podemos empregar um processo para determinar diferenças significativas entre duas quaisquer das K amostras.

Estas provas não-paramétricas têm a vantagem de permitir o estudo da significância de dados que, inerentemente, se apresentam apenas sob a forma classificativa ou em postos.

Há dois planos básicos para comparar K grupos:

1. No primeiro deles, as K amostras de igual tamanho são postas em correspondência de acordo com determinado(s) critério(s) que podem afectar os valores das observações. Nalguns casos, essa correspondência obtém-se comparando os mesmos indivíduos ou casos sob todas as K condições ou então cada um dos N grupos pode ser mensurado sob todas as K condições. Em tais planos, devem usar-se provas estatísticas para K amostras relacionadas;
2. O segundo plano envolve K amostras aleatórias independentes (não necessariamente do mesmo tamanho), uma de cada população. Em tais casos, devemos usar as provas estatísticas para K amostras independentes.

8.1 PROVA DE COCHRAN

Quando utilizar:

A prova de McNemar para duas amostras pode ser estendida para aplicação a pesquisas que envolvem mais de duas amostras. Essa extensão, que constitui a prova Q de Cochran para K amostras relacionadas, proporciona um método para **comparar se três ou mais conjuntos correspondentes de frequências ou proporções diferem entre si significativamente**. A correspondência pode basear-se em características relevantes dos diferentes indivíduos ou no facto de os mesmos indivíduos serem observados sob condições diferentes. A prova Q de Cochran adapta-se especialmente ao caso em que os dados se apresentam numa **escala nominal ou sob a forma de informação ordinal dicotomizada**.

Exemplo: Suponha que estamos interessados em saber se a atitude de um entrevistador influencia a aceitação de participação num estudo por inquérito. Poderemos treinar o entrevistador para efectuar as suas entrevistas de três maneiras diferentes:

1. Demonstrando interesse, cordialidade, entusiasmo;
2. Demonstrando formalismo, reserva e cortesia;
3. Demonstrando modo abrupto, formalismo e aspereza.

Exemplo: O entrevistador visitaria 3 grupos de 18 casas, aplicando o tipo 1 de entrevista a um grupo, o tipo 2 a outro grupo e o 3 ao terceiro grupo. Teríamos, assim, 18 conjuntos de potenciais inquiridos com três deles correspondendo em cada conjunto. Em cada conjunto atribuir-se-iam aleatoriamente aos três membros as três condições (tipos de entrevista). Teríamos, então, 3 amostras relacionadas (correspondentes) com 18 elementos cada uma (N=18). Poderíamos, pois, comprovar se as diferenças fundamentais nos tipos de entrevista influenciariam o número de respostas afirmativas “sim” dadas para aceitação de participação pelos 3 grupos de correspondentes.

Quadro 20: Q de Cochran

Conjunto	Resposta à entrevista 1	Resposta à entrevista 2	Resposta à entrevista 3
1	1	1	1
2	2	2	1
3	1	2	1
4	1	1	1
5	2	1	1
6	2	2	1
7	2	2	1
8	1	2	1
9	2	1	1
10	1	1	1
11	2	2	2
12	2	2	2
13	2	2	1
14	2	2	1
15	2	2	1
16	2	2	2
17	2	2	1
18	2	2	1

Respostas “Sim” (1) e “Não” (2) dadas por donas de casa a três tipos de entrevistas

HIPÓTESE NULA:

A probabilidade de um “Sim” é a mesma para os três grupos de entrevistas.

H1: As probabilidades de um “Sim” diferem conforme o tipo de entrevista.

PROVA ESTATÍSTICA:

Escolhe-se a prova Q de Cochran, porque os dados se referem a mais de dois grupos relacionados (K=3) e apresentam-se dicotomizados sob a forma “Sim” e “Não”.

NÍVEL DE SIGNIFICÂNCIA:

p=0,01 N=18

DISTRIBUIÇÃO AMOSTRAL:

Sob a hipótese de nulidade, Q tem uma distribuição aproximadamente Qui-quadrado com $gl=1$. Isto é, a probabilidade associada à ocorrência, sob H_0 , de qualquer valor tão grande quanto um valor observado de Q pode ser determinada mediante referência à Tabela.

REGIÃO DE REJEIÇÃO:

Consiste em todos os valores de Q tão grandes quanto a probabilidade associada à sua ocorrência, sob H_0 , não seja superior a $p=0,01$.

DECISÃO:

Recodificámos “Sim” por 1 e “Não” por 0.

Quadro 21: Q de Cochran

Conjunto	Resposta à entrevista 1	Resposta à entrevista 2	Resposta à entrevista 3	Li	Li ²
1	0	0	0	0	0
2	1	1	0	2	4
3	0	1	0	1	1
4	0	0	0	0	0
5	1	0	0	1	1
6	1	1	0	2	4
7	1	1	0	2	4
8	0	1	0	1	1
9	1	0	0	1	1
10	0	0	0	0	0
11	1	1	1	3	9
12	1	1	1	3	9
13	1	1	0	2	4
14	1	1	0	2	4
15	1	1	0	2	4
16	1	1	1	3	9
17	1	1	0	2	4
18	1	1	0	2	4
Total	G1=13	G2=13	G3=3	ΣLi=29	Σ Li²63

Li=número total de respostas “Sim” para cada linha

K=número de colunas

N=número de linhas

Substituindo estes valores na fórmula, vem:

$$Q = \frac{(K-1) \left[K \sum G_j^2 - (\sum G_j)^2 \right]}{K \sum L_i - \sum L_i^2} = \frac{(3-1) \left[3 \times (13^2 + 13^2 + 3^2) - 29^2 \right]}{3 \times 29 - 63} = 16,7$$

Em que $K = n.^{\circ}$ de grupos
 $G_j = n.^{\circ}$ total de sucessos

A Tabela (Anexo I) indica que $Q \geq 16,7$ tem uma probabilidade de ocorrência, sob H_0 , $p < 0,001$, quando $gl = K-1 = 3-1 = 2$. Essa probabilidade é inferior ao nível de significância de $p = 0,01$. O valor de Q está na região de rejeição e, conseqüentemente, a nossa decisão é rejeitar H_0 em favor de H_1 , concluindo que o número de respostas “Sim” difere significativamente em relação aos tipos 1, 2 e 3 de entrevista.

RACIONAL

Se os dados de uma pesquisa se dispõem numa tabela de dupla entrada com N linhas e K colunas, é possível testar a hipótese de nulidade de que a proporção ou frequência de respostas de determinado tipo seja a mesma em cada coluna, exceptuando as diferenças devidas ao acaso. Cochran mostrou que se a hipótese de nulidade é verdadeira, isto é, se não há diferença na probabilidade, digamos de “Sucesso” sob cada condição (o que equivale a dizer que os “Sucessos” ou “Fracassos” se distribuem aleatoriamente pelas linhas e colunas da tabela de dupla entrada), então, se o número de linhas é muito pequeno

$$Q = \frac{K(K-1) \sum (G_j - \bar{G})^2}{K \sum L_i - \sum L_i^2}$$

tem distribuição aproximadamente Qui-quadrado com $gl = K-1$, em que:

G_j = número total de “Sucessos” na coluna j

\bar{G} = média dos G_j

L_i = número total de “sucessos” na linha i

Uma expressão equivalente à fórmula anterior, e facilmente dedutível dela, mas que simplifica os cálculos é:

$$Q = \frac{(K-1) \left[K \sum G_j^2 - (\sum G_j)^2 \right]}{K \sum Li - \sum Li^2}$$

INSTRUÇÕES PASSO-A-PASSO PARA CALCULAR Q

Para dados dicotomizados, atribuir o score “1” a cada “Sucesso” e o score “0” a cada “Falha”.

Dispor os dados numa tabela K.N, com K colunas e N linhas (N=número de condições em cada um dos grupos).

Determinar o valor de Q, aplicando a fórmula.

A significância do valor observado de Q pode ser determinado mediante a observação do Anexo I, pois Q recorre à distribuição do Qui-quadrado com $gl=K-1$. Se a probabilidade associada à ocorrência, sob H_0 , de um valor tão grande quanto um valor observado de Q não supere p , rejeita-se H_0 .

8.2 TESTE DE FRIEDMAN

Quando utilizar:

Este teste pode ser considerado uma extensão do teste de Wilcoxon, quando é necessário utilizar **três ou mais** situações experimentais. Deve ser utilizado para um *design* **relacionado** quando os **mesmos** sujeitos (ou sujeitos **emparelhados**) são distribuídos por **três ou mais** situações experimentais.

Exemplo: Suponha que um editor de livros de estatística produziu uma série de livros e quer escolher de entre três tipos de ilustrações aquele que é mais eficaz para os estudantes. É pedido a oito universitários que classifiquem as obras numa escala de cinco pontos, desde “nada boa” até “muito boa”. Obtiveram-se os resultados apresentados na tabela seguinte.

Quadro 22: Avaliação de três tipos de ilustrações

Sujeitos	Situação 1 (Ilustração A)		Situação 2 (Ilustração B)		Situação 3 (Ilustração C)	
	Result.	Ordem	Result.	Ordem	Result.	Ordem
1	2	1	5	3	4	2
2	1	1	5	3	3	2
3	3	1	5	2,5	5	2,5
4	3	2	5	3	2	1
5	2	1	3	2	5	3
6	1	1	4	2,5	4	2,5
7	5	3	3	2	2	1
8	1	1	4	3	3	2
TOTAL	18	11	34	21	28	16
MÉDIA	2,25		4,25		3,50	

RACIONAL

Uma vez que se trata de um *design* relacionado no qual o mesmo sujeito obtém resultados em todas as situações, é permitido comparar os resultados de cada sujeito através de todas as situações, no sentido de verificarmos em que situação obtém maiores e menores resultados.

Uma vez que existem mais do que duas situações, não é possível calcular as diferenças nos resultados de duas situações, como era o caso do teste de Wilcoxon. Pelo contrário, o ordenamento dos resultados de cada sujeito para as três condições será feita horizontalmente ao longo das linhas, tal como mostra a tabela. Por exemplo, aos resultados do sujeito 1, respectivamente 2 na Situação 1, 5 na Situação 2 e 4 na Situação 3, são atribuídas três ordens, do menor resultado para o maior: Ordem 1 para a Situação 1, Ordem 2 para a Situação 3 e Ordem 3 para a Situação 2; este procedimento é semelhante para todos os sujeitos. Claro que se existissem quatro situações experimentais, os resultados de cada sujeito seriam ordenados de 1 a 4.

O próximo passo é calcular os totais de ordens para cada situação. Se existirem apenas diferenças aleatórias entre os resultados de todas as situações, como é postulado pela hipótese nula, é de esperar que estes totais sejam aproximadamente iguais partindo do princípio de que surgiriam algumas ordens baixas (baixos resultados) e algumas ordens altas (altos resultados). Contudo, se as situações forem significativamente diferentes, é de esperar que se obtenham totais das ordens

significativamente diferentes, com algumas situações a terem uma preponderância de ordens baixas e outras uma preponderância de ordens altas. O tamanho das diferenças entre os totais das ordens é-nos dado por uma estatística designada por X_R^2 . Se o valor de X_R^2 for igual ou superior aos valores críticos das Tabelas C e D (Anexo V), isso implica que as diferenças nos totais das ordens são suficientemente grandes para que se possam considerar significativas.

INSTRUÇÕES PASSO-A-PASSO PARA CALCULAR W

Ordene os resultados de cada sujeito em separado, ao longo de cada linha, atribuindo a Ordem 1 ao menor resultado e por aí adiante.

(Veja as colunas das Ordens na tabela. Note que a ordem para cada linha de resultados corresponde às ordens 1, 2 e 3, dado existirem três situações).

Calcule o total das ordens para cada situação.

Calcule o valor de X_R^2 a partir da fórmula

$$X_R^2 = \frac{12}{NC(C+1)} \sum T_c^2 - 3N(C+1)$$

em que

C=número de situações

C=3

N=número de sujeitos

N=8

T_c =total de ordens para cada situação

$T_1=11; T_2=21; T_3=16$

T_c^2 =quadrado do total de ordens para cada situação

$T_1^2=11^2; T_2^2=21^2; T_3^2=16^2$

$\sum T_c^2$ =soma dos quadrados dos totais das ordens para cada situação: $11^2+21^2+16^2$

Cálculo de X_R^2

$$X_R^2 = \frac{12}{(8 \times 3)(3+1)} (11^2 + 21^2 + 16^2) - 3 \times 8(3+1)$$

$$X_R^2 = \frac{12}{24 \times 4} (121 + 441 + 256) - (24 \times 4)$$

$$X_R^2 = \frac{12}{96} \times 818 - 96$$

$$X_R^2 = 6,25$$

Calcule os graus de liberdade, ou seja, o número de situações menos uma.

$$(gl = C - 1 = 3 - 1 = 2)$$

CONSULTA DA SIGNIFICÂNCIA NA TABELA

Existem duas tabelas para consultar os valores críticos de Xr^2 . Uma delas, a Tabela C (Anexo V), é utilizada quando o número de situações e de sujeitos é pequeno. A Tabela C (1) apresenta os valores de Xr^2 para três situações quando N (número de sujeitos) se situa entre 2 e 9. A Tabela C (2) apresenta os valores de Xr^2 para quatro situações quando N (número de sujeitos) é de 2, 3 ou 4. A Tabela D (Anexo V) é a tabela de distribuição do Qui-quadrado; pode utilizá-la quando a amostra de sujeitos for superior às das Tabelas C (1) e C (2), uma vez que o Xr^2 tem uma distribuição semelhante à do Qui-quadrado.

A Tabela que deve utilizar para consultar o valor de Xr^2 , no caso desta experiência, é a Tabela C (1), uma vez que se trata de 8 sujeitos expostos a 3 situações experimentais. Aquilo que temos de fazer é encontrar a coluna apropriada para N (número de sujeitos ou pares de sujeitos emparelhados) e descobrir na coluna p a probabilidade mais próxima que seja inferior aos níveis de significância convencionais. Consultando as probabilidades para N=8, o valor obtido de $Xr^2=6,25$ é equivalente a uma probabilidade de $p<0,047$, que é inferior aos níveis de significância convencionais ($p<0,05$). Para considerarmos o nível de significância de $p<0,01$ o nosso valor de Xr^2 teria de ser 9,00, dado que $p<0,009$ é inferior a $p<0,01$. Se o valor de Xr^2 não for apresentado na Tabela, deverá considerar o valor seguinte mais próximo quando consulta as probabilidades. Por exemplo, se o valor de Xr^2 for 5,95 terá de considerar a probabilidade apresentada para 5,25, ou seja, $p<0,079$, que é superior a $p<0,05$ e, conseqüentemente, não significativa. Para consultar os valores da Tabela C (2) deverá proceder tal como para a Tabela C (1).

Se tiver mais situações e/ou sujeitos e tiver de consultar a Tabela D, aquilo que tem a fazer é localizar os valores dos graus de liberdade ao longo da coluna da esquerda (no nosso exemplo, $gl=2$, ou seja, número de situações-1). Depois siga ao longo da linha de probabilidades até que encontre um dos níveis de significância convencionais. O valor que obtivemos de $Xr^2=6,25$ é superior ao valor crítico de 5,99 apresentado na Tabela do Qui-quadrado, pelo que podemos aceitar que os nossos resultados são significativos ao nível de significância de $p<0,05$. Apesar disso, e dado que o nosso valor de Xr^2 é inferior ao valor crítico de 9,21 para $p<0,01$, não podemos rejeitar a hipótese nula a este nível de significância.

A partir da análise estatística da experiência pode concluir que as crianças mostram preferências significativamente diferentes pelos três tipos de ilustrações. Em função das médias apresentadas na tabela, sabemos que preferem a Ilustração B, que recolheu as ordens mais elevadas, seguindo-se a Ilustração C e, por último, a Ilustração A. Contudo, o teste de Friedman pode apenas indicar que existem diferenças globais entre as situações. Para verificar se existe uma tendência para uma determinada ordem de preferência das Ilustrações, necessita de utilizar um teste de tendência.

9. MEDIDAS DE CORRELAÇÃO E SUAS PROVAS DE SIGNIFICÂNCIA

Frequentemente, o investigador quer saber se dois conjuntos de scores estão relacionados e qual o grau desse relacionamento.

Apresentam-se medidas não-paramétricas de correlação e de provas estatísticas para determinar a probabilidade associada à ocorrência de uma correlação tão grande quanto a observada na amostra, sob a hipótese de nulidade de que as variáveis sejam não-relacionadas na população.

Mas é de muito maior interesse podermos afirmar se determinada associação observada numa amostra de scores indica, ou não, a probabilidade de associação entre as variáveis na população da qual se extraiu a amostra. O coeficiente de correlação, por si só, representa o grau de associação. As provas de significância sobre aquele coeficiente determinam, a um certo nível de probabilidade, se existe a associação na população da qual se extraiu a amostra que serviu de base para o cálculo do coeficiente.

9.1 COEFICIENTE DE CORRELAÇÃO RHO DE SPEARMAN-RANK

CONDIÇÕES DE UTILIZAÇÃO

Este tipo de coeficiente de correlação utiliza-se quando temos:

- Teste não paramétrico (semelhante a uma distribuição livre), isto é, não coloca restrições quanto à forma da distribuição;
- Escala de medida no mínimo ordinal.

Pode acontecer que os caracteres estudados não sejam mensuráveis, mas podem ser ordenados ou classificados. Por exemplo, se se considera um grupo de candidatos a um certo lugar, eles podem ser examinados segundo dois pontos de vista: conhecimentos e personalidade. Estas duas qualidades não podem ser medidas, mas é possível para cada uma delas efectuar uma classificação dos candidatos. Podemos, assim, examinar se existe correlação entre estes dois caracteres; cada par de dados (x_i, y_i) é formado pelas ordens ocupadas por um candidato nas duas classificações.

Formulário:

$$\rho = 1 - \frac{6(\sum d_i^2)}{N(N^2 - 1)}$$

Em que:

d_i = diferença entre as posições nas duas variáveis, isto é, $d_i = x_i - y_i$

Para tal, temos que dar valores às posições: à pontuação mais baixa damos o valor 1 e assim sucessivamente.

Se as duas classificações são iguais, d_i é sempre zero e então $r=1$ e a correlação é perfeita. Se as ordens mais altas de uma classe estão associadas às mais baixas da outra r torna-se negativo e se as duas classificações são inversas $\rho=-1$.

Então,

$$-1 \leq \rho \leq 1$$

INTERPRETAÇÃO

O coeficiente de correlação obtido pode ser interpretado, tal como o coeficiente de correlação momento-produto de Brawais-Pearson (ver este coeficiente em testes paramétricos).

Exemplo: Calcule ρ , sabendo que:

xi	18	17	14	13	13	12	11	9	7	5
yi	24	27	17	22	19	20	14	11	3	6

Quadro 23: Coeficiente de Contingência de RHO de Spearman-Rank

xi	posição	yi	posição	d	d ²
18	1	24	2	-1	1
17	2	27	1	1	1
14	3	17	6	-3	9
13	4,5	22	3	1,5	2,25
13	4,5	19	5	-0,5	0,25
12	6	20	4	2	4
11	7	14	7	0	0
9	8	11	8	0	0
7	9	3	10	-1	1
5	10	6	9	1	1

$\Sigma=19,5$

$$\rho = 1 - \frac{6 \times 19,5}{10(10^2-1)} = 0,88$$

9.2 O COEFICIENTE DE CONCORDÂNCIA DE KENDALL

Quando utilizar:

O coeficiente de concordância de Kendall é uma **medida da relação entre vários conjuntos de postos de N objectos ou indivíduos**.

Quando temos K conjuntos de postos, podemos determinar a associação entre eles utilizando o coeficiente de concordância de Kendall (W).

RACIONAL E EXEMPLO

Como solução do problema da determinação da concordância global entre K conjuntos de postos, poderia parecer razoável determinar os r 's entre todos os pares possíveis de postos e então calcular a média desses coeficientes para determinar a associação global. Adoptando este procedimento, o processo torna-se impraticável, uma vez que temos de calcular inúmeros coeficientes de correlação de postos.

Assim, o cálculo de W é muito mais simples e W tem uma relação linear com o valor médio de r_s relativo a todos os grupos. Denotando r_{sav} o valor médio dos coeficientes de correlação por postos de Spearman, Kendal mostrou que

$$r_{sav} = \frac{KW - 1}{K - 1}$$

Outro processo consiste em imaginar como se apresentariam dados se não houvesse concordância alguma entre os conjuntos de postos e, em seguida, como se apresentariam se houvesse concordância perfeita. O coeficiente de concordância seria, então, um índice de divergência entre a concordância efectiva acusada pelos dados e a concordância máxima possível (perfeita). De modo aproximado, W é um coeficiente desta natureza.

Suponhamos que 3 chefes de pessoal estejam encarregados de entrevistar 6 candidatos a emprego e classificá-los em postos, separadamente, segundo a capacidade de cada um para preencher a vaga. A tabela seguinte dá os 3 conjuntos independentes de postos atribuídos pelos chefes de pessoal X , Y e Z aos candidatos a a f ; a última linha da tabela R_j dá as somas dos postos atribuídos a cada candidato.

Quadro 24: Postos atribuídos a 6 candidatos a emprego por 3 chefes de pessoal

		Candidato					
		a	b	c	d	e	f
Chefe	X	1	6	3	2	5	4
	Y	1	5	6	4	2	3
	Z	6	3	2	5	4	1
	R_j	8	14	11	11	11	8

Se os chefes de pessoal apresentassem perfeita concordância nos seus julgamentos sobre os candidatos, isto é, se tivessem atribuído postos aos candidatos da mesma ordem, então um candidato teria recebido três postos 1 e assim a sua soma de postos, R_j , seria $1+1+1=3=K$. O candidato que os 3 chefes tivessem considerado em segundo lugar receberia $R_j=2+2+2=6=2K$ e o menos promissor dos candidatos seria $R_j=6+6+6=18=NK$.

De facto, no caso da concordância perfeita entre os 3 chefes, as várias somas de postos, R_i , seriam 3,6,9,12,15,18, muito embora não necessariamente nesta ordem. Em geral, quando há concordância perfeita entre K conjuntos de postos, obtemos, para valores significativos de R_i , a sequência K, 2K, 3K, ..., NK.

Por outro lado, se não tivesse havido concordância entre os 3 chefes, então os diversos R_j seriam aproximadamente iguais.

Deste exemplo, vemos que o grau de concordância entre os K julgamentos é reflectido pelo grau de variância entre as N somas de postos. W, coeficiente de concordância, é função desse grau de variância.

Assim:

$$W = \frac{S}{1/12 K^2 (N^3 - N)}$$

em que:

S=soma dos quadrados dos desvios observados a contar da média dos R_j , isto é, $S = \sum (R_j - \sum R_j / N)^2$;

K=número de conjuntos de postos;

N=número de entidades (objectos ou indivíduos a que se atribuem postos);

$1/12 K^2 (N^3 - N)$ =valor máximo possível da soma dos quadrados dos desvios, isto é, o valor de S que ocorreria no caso de concordância perfeita entre os K conjuntos de postos.

Então,

$$\bar{\chi} = (8 + 14 + 11 + 11 + 11 + 8) / 6 = 10,5$$

$$S = (8-10,5)^2 + (14-10,5)^2 + (11-10,5)^2 + (11-10,5)^2 + (11-10,5)^2 + (8-10,5)^2 = \\ 6,25 + 12,25 + 0,25 + 0,25 + 0,25 + 6,25 = 25,5$$

Pelo que

$$W = \frac{S}{1/12 K^2 (N^3 - N)} = \frac{25,5}{1/12 \times 3^2 (6^3 - 6)} = \frac{25,5}{1/12 \times 9 \times 210} = \\ = \frac{25,5}{157,5} = 0,16$$

$W=0,16$ exprime o grau de concordância entre os 3 colegas fictícios ao atribuírem postos aos 6 candidatos a emprego.

É de salientar que a razão porque W não pode tomar valores negativos é que, quando estão em jogo mais de dois conjuntos de postos, esses conjuntos não podem ser completamente discordantes. Por exemplo, se X e Y discordam, e se X também discorda de Z , então Y e Z certamente concordarão. Isto é, quando estão em jogo mais de 3 juízes, concordância e discordância não são opostos simétricos. K juízes podem todos concordar, mas é impossível todos eles discordarem completamente. Portanto, W deve ser zero ou um número positivo.

NOTA: Quando ocorrem empates, atribuem-se às observações empatadas a média dos postos que lhes caberiam se não houvesse empates.

O efeito dos empates é reduzir o valor de W . Se a proporção de empates é pequena, o efeito é desprezível, podendo continuar a usar-se a fórmula anterior. Se a proporção de empates é grande, pode introduzir-se uma correcção que aumentará ligeiramente o valor de W em relação ao valor que se apresentaria sem correcção. O elemento correctivo é:

$$T = \frac{\sum (t^3 - t)}{12}$$

Em que t =número de observações num grupo empatados em relação a um dado posto

Então, a fórmula será:

$$W = \frac{S}{1/12 K^2 (N^3 - N) - \sum T}$$

INSTRUÇÕES PASSO-A-PASSO PARA CALCULAR W

Seja N=número de objectos ou indivíduos a serem classificados em postos e K=número de juízes classificadores. Dispor os postos observados numa tabela do tipo K X N.

Para cada indivíduo ou objecto determinar R_i , soma dos postos atribuídos áquele indivíduo pelos K juízes.

Determinar a média dos R_j . Expressar cada R_j como desvio a contar dessa média. Elevar ao quadrado esses desvios e somá-los, obtendo S.

Se a proporção de empates nos K conjuntos é grande, para calcular W utilizar a fórmula

$$W = \frac{S}{1/12 K^2 (N^3 - N) - \sum T}$$

Em caso contrário, usar

$$W = \frac{S}{1/12 K^2 (N^3 - N)}$$

O método para determinar se o valor observado é significativamente diferente de zero depende do tamanho de N:

Se $N \leq 7$ a Tabela (Anexo VI) dá os valores críticos de S associados com os W's significativos aos níveis de 0,05 e de 0,01;

Se $N > 7$ podemos usar $X^2 = K (N - 1)W$ para calcular um valor de X^2 cuja significância, para g.l.=N-1, possa ser comparada recorrendo à Tabela do X^2 (Anexo I).

CONSULTA DA SIGNIFICÂNCIA NA TABELA

PROVA DE SIGNIFICÂNCIA

PEQUENAS AMOSTRAS:

Utilizando-se a Tabela esta dá o valor de S para W's significativos aos níveis de 0,05 e de 0,01. Esta Tabela é aplicável para K de 3 a 20 e N de 3 a 7. Se um valor observado de S é igual ou superior ao valor exibido na Tabela para um dado nível de significância, então H₀ pode ser rejeitada áquele nível.

Por exemplo, vimos que quando K=3 chefes de pessoal fictícios classificaram N=6 candidatos a emprego, a concordância dos julgamentos foi W=0,16. A Tabela indica que o S associado àquele valor de W (25,5) não é significativo, uma vez que S crítico (103,9) é maior do que S observado (25,5).

Quando N>7, a distribuição é aproximadamente uma distribuição de Qui-quadrado com g.l.=N-1, vindo:

$$X^2 = \frac{S}{1/12 KN (N + 1)}$$

Note-se que:

$$\frac{S}{1/12 KN (N + 1)} = K (N - 1) W$$

E, portanto, $X^2 = K (N - 1) W$, pelo que se pode usar esta fórmula que é mais fácil para determinar a probabilidade associada à ocorrência, sob H₀, de qualquer valor tão grande quanto o W observado.

Se o valor de X² observado iguala ou supera o valor da Tabela (Anexo I) para um dado nível de significância e um particular valor de g.l.=N-1, então H₀ (não há relação entre os K conjuntos de postos) pode ser rejeitada áquele nível.

INTERPRETAÇÃO DE W

Um valor elevado ou significativo de W pode ser interpretado como indicando que os observadores ou juizes estão aplicando essencialmente os mesmos padrões ao atribuírem postos aos N indivíduos em estudo. Frequentemente a sua ordenação conjunta pode servir como “padrão”, especialmente quando não há critérios externos relevantes para ordenar os indivíduos.

Kendall sugere que a melhor estimativa da “verdadeira” classificação de N objectos, quando W é significativo, seja dada pela ordem das várias somas de postos, R_j . Se se aceita o critério segundo o qual diversos juizes concordarem na atribuição de postos a N indivíduos, então a melhor estimativa da “verdadeira” ordenação desses indivíduos segundo aquele critério é dada pela ordem da soma dos postos. Assim, a nossa melhor estimativa seria dada ao candidato a ou f , pois em ambos os casos $R_j=0$, o menor valor observado.

10. TESTES PARAMÉTRICOS PASSO-A-PASSO

10.1 CORRELAÇÃO

Quando estudamos um grupo relativamente a dois caracteres vemos, como já dissemos, que pode existir uma relação entre eles.

Se medirmos os raios de várias circunferências e também os seus perímetros verificamos que existe uma relação entre eles que é constante; neste caso temos "dependência funcional". Isto quer dizer que existe uma fórmula exprimindo a medida do segundo em função da do primeiro: $P=2\pi r$.

Suponhamos agora que registamos, durante todos os dias de um certo período de tempo, a altura máxima da maré no porto da Figueira da Foz e o número de passageiros da linha Porto-Lisboa. Vê-se bem que entre as duas estatísticas assim obtidas não se encontra nenhuma relação. Diremos que os dois caracteres são "independentes".

Mas entre estes dois casos extremos, a dependência funcional e a independência, a vida corrente oferece um grande número de exemplos onde existe uma relação entre os caracteres, mas em que existe uma impossibilidade de a formular por uma lei: é o caso da "dependência estatística". Existe uma dependência estatística entre as pessoas que morrem com febre tifóide e as pessoas vacinadas contra esta doença.

Desde que os dois caracteres sejam tais que as suas variações sejam sempre no mesmo sentido, ou em sentidos contrários, pressentimos que os caracteres estejam ligados entre si: dizemos, então, que eles estão correlacionados ou que existe uma correlação entre eles.

Estes métodos de correlação foram criados por Sir Francis Galton, que trabalhou juntamente com Pearson, nos fins do século XIX.

A correlação e a regressão são dois aspectos que andam sempre muito ligados, pertencendo à Estatística Descritiva. Assim, importa fazermos a distinção entre eles:

- A correlação pode ser definida como o grau de semelhança no sentido das variações entre os valores correspondentes dos dois caracteres, isto é, a correlação preocupa-se com a descrição da relação entre variáveis.
- A regressão é usada quando tentamos prever uma variável quando conhecemos a outra.

TIPOS DE COEFICIENTE DE CORRELAÇÃO

Basicamente, podemos considerar dois tipos de coeficientes de correlação:

- Coeficiente de correlação momento-produto de Brawais-Pearson, cujo símbolo é " r ", e que é uma técnica de estatística paramétrica;
- Coeficiente de correlação Rho de Spearman-Rank, cujo símbolo é " ρ ", e que é uma técnica de estatística não paramétrica.

Devemos salientar que, para o cálculo das correlações, é necessário termos sempre duas medidas para cada sujeito.

REPRESENTAÇÃO GRÁFICA

À representação gráfica da correlação chamamos diagrama de dispersão de pontos ou scatterplot ou scattergram e, genericamente, toma a seguinte forma:

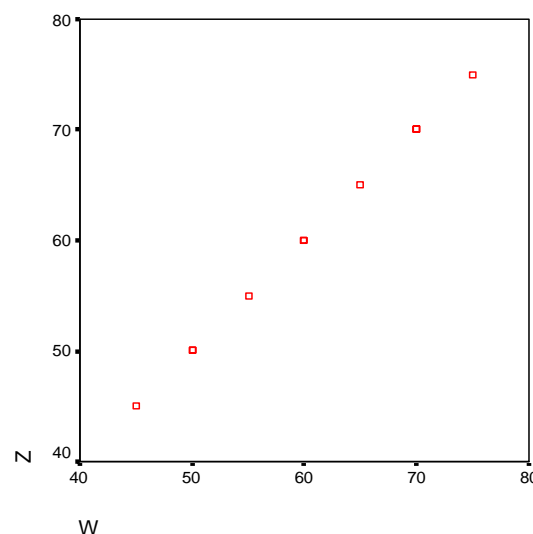


Figura 13: Diagrama de dispersão de pontos ou scatterplot ou scattergram

Suponhamos que temos duas séries estatísticas formadas pelos valores x_i e y_i de dois caracteres. Podemos fazer uma representação gráfica dos dados representando todos os pontos (x_i, y_i) e obtendo a nuvem de pontos.

Segundo os dados, a nuvem de pontos pode apresentar diversos aspectos.

Por exemplo os pontos podem distribuir-se na semelhança de uma linha recta ou de uma curva: isto sugere a existência de uma relação funcional entre X e Y .

Os pontos podem ser dispersos e colocados ao acaso no plano: caso das marés no porto da Figueira da Foz e o número de passageiros na linha Porto-Lisboa (independência). Também pode acontecer que os pontos cubram uma porção do plano da qual se pode definir o contorno; esta forma sugere que as duas variáveis estão ligadas. Limitemo-nos ao caso mais simples em que a nuvem tem uma forma alongada lembrando uma elipse e suponhamos que a sua orientação é tal que desde que X cresça, a variável Y também cresce. A forma desta nuvem sugere a possibilidade da existência de uma recta tal que os valores estimados por esta recta, a partir dos valores de x_i , sejam boas aproximações dos valores de y_i . Nós podemos determinar pelo método dos mínimos quadrados uma recta tal que a soma dos quadrados dos desvios seja mínima. Esta recta é chamada recta de regressão de Y em X ou recta de estimação de Y em X .

Mas, poderíamos, de um modo semelhante, procurar uma recta tal que os valores de x estimados ao longo desta recta, a partir de y_i , constituam igualmente boas aproximações de x_i . Esta recta é chamada recta de regressão de X em Y ou recta de estimação de X em Y .

Normalmente, estas rectas são distintas uma da outra. Elas serão confundidas quando existe ligação funcional linear e são perpendiculares quando há independência. Compreendemos, assim, que a correlação entre os caracteres é tanto maior quanto maior as rectas de regressão estejam mais próximas uma da outra.

COEFICIENTE DE CORRELAÇÃO MOMENTO-PRODUTO DE BRAVAIS-PEARSON

Condições de utilização:

Este tipo de coeficiente de correlação utiliza-se quando:

1. As duas variáveis são contínuas;
2. A distribuição se aproxima da distribuição normal;
3. É preferível para distribuições unimodais;
4. Escala intervalar de medida.

Formulário

$$r = \frac{(\sum XY) - \bar{X} \bar{Y}}{N \cdot s_x s_y}$$

Então:

$$-1 \leq r \leq 1$$

Interpretação:

O coeficiente de correlação obtido pode se interpretado com base em:

Para Cardoso:

$r \leq 0,2$	Correlação muito baixa (valores desprezíveis)
$0,2 < r \leq 0,5$	Correlação baixa
$0,5 < r \leq 0,7$	Valores significativos
$0,7 < r \leq 0,9$	Alta correlação
$0,9 < r \leq 1$	Muito alta correlação

Para Borg:

$0,20 < r \leq 0,35$	Ligeira relação entre as variáveis, embora já possam ser estatisticamente significativas
$0,35 < r \leq 0,65$	Correlação estatisticamente significativa para além do nível de 1%
$0,65 < r \leq 0,85$	Correlações que tornam possíveis predições do grupo de que são dignas
$r > 0,85$	Íntima relação entre as variáveis correlacionadas

Para Byrman e Cramer,
se Eta, r, Rho, phi:

- $\leq 0,2$ Correlação muito fraca e sem significância
- $0,2 < r \leq 0,39$ Correlação fraca
- $0,4 < r \leq 0,69$ Correlação moderada
- $0,7 < r \leq 0,89$ Correlação forte
- $0,9 < r \leq 1$ Correlação muito elevada

Coefficiente de correlação dá-nos:

- A direcção que é indicada pelo sinal + ou -
- A intensidade ou força que é dada pelo valor que varia entre -1 e 1. Se a correlação for zero não existe correlação entre as variáveis (exemplo: cor dos olhos e inteligência).

Exemplo: Considere as classificações obtidas por 10 alunos nas disciplinas X, Y, Z e W:

Quadro 25: Correlação de Pearson

Estudante	X	Y	Z	W	XY	XZ	XW	YZ	YW	ZW
1	75	75	45	45	5625	3375	3375	3375	3375	2025
2	70	70	50	50	4900	3500	3500	3500	3500	2500
3	70	70	50	50	4900	3500	3500	3500	3500	2500
4	65	65	55	55	4225	3575	3575	3575	3575	3025
5	60	60	60	60	3600	3600	3600	3600	3600	3600
6	60	60	60	60	3600	3600	3600	3600	3600	3600
7	55	55	65	65	3025	3575	3575	3575	3575	4225
8	50	50	70	70	2500	3500	3500	3500	3500	4900
9	50	50	70	70	2500	3500	3500	3500	3500	4900
10	45	45	75	75	2025	3375	3375	3375	3375	5625
Σ	600	600	600	600	36900	35100	35100	35100	35100	36900

Sabe-se que:

$$\bar{X}=60$$

$$\bar{Y}=60$$

$$\bar{Z}=60$$

$$\bar{W}=60$$

$$\Sigma XY=36900$$

$$\Sigma XZ=35100$$

$$s^2x= 90$$

$$s^2y= 90$$

$$s^2z= 90$$

$$s^2w= 90$$

- a) Calcule o coeficiente de correlação entre X e Y e represente graficamente.
- b) Calcule o coeficiente de correlação entre X e Z e represente graficamente.

Resolução 47:

a)

$$r = \frac{(\sum XY) - \bar{X} \bar{Y}}{N s_x s_y}$$

Então,

$$r = \frac{\frac{36900}{10} - 60 \times 60}{\sqrt{90} \times \sqrt{90}}$$

$$r = 1$$

Existe uma relação positiva muito forte entre X e Y

b)

$$r = \frac{(\sum XZ) - \bar{X} \bar{Z}}{N s_x s_Z}$$

Então,

$$r = \frac{\frac{35100}{10} - 60 \times 60}{\sqrt{90} \times \sqrt{90}}$$

$$r = -1$$

Existe uma relação negativa muito forte entre X e W, isto é, quanto maior é X menor é Y e vice versa.

10.2 TESTE T DE STUDENT (NÃO RELACIONADO)

CARACTERÍSTICAS DO TESTE T:

1. Teste para a comparação de médias;
2. Distribuição com forma leptocúrtica, isto é, as caudas da distribuição são mais grossas do que na distribuição normal;
3. Contínua;
4. Simétrica;
5. De forma campanular;
6. Varia de mais infinito a menos infinito;
7. Média = 0 e desvio padrão da variável de acordo com n;

8. O uso de T pressupõe que a variável observacional tenha na população (de onde foi colhida a amostra) distribuição normal;
9. Não existe uma distribuição t, mas sim uma família. Para cada n - tamanho da amostra - existe uma distribuição (uma curva) específica;
10. À medida que n cresce, t tende a z (a aproximação entre t e z começa a ficar "boa" a partir de $n \geq 30$).

Requisitos para utilizar o teste t:

1. Quando as duas amostras têm a mesma homogeneidade de variância;
2. Os dois grupos tenham a mesma distribuição aproximada à normal - leptocúrtica;
3. Escala de medida intervalar.

QUANDO UTILIZAR

Utiliza-se para designs experimentais com duas situações testando uma variável independente, quando nessas situações se encontram sujeitos diferentes - designs não relacionados.

O teste t não relacionado é equivalente ao teste não paramétrico U de Mann-Whitney; ambos comparam diferenças entre dois grupos.

OBJECTIVO

O objectivo é comparar a quantidade da variabilidade devida às diferenças previstas nos resultados entre dois grupos com a variabilidade total nos resultados dos sujeitos. As diferenças previstas são calculadas como uma diferença entre os resultados médios entre os dois grupos.

A estatística t representa o tamanho da diferença entre as médias para os dois grupos, tomando em consideração a variância total. Para que o valor observado de t seja significativo terá de ser igual ou superior aos valores críticos de t apresentados na tabela.

INSTRUÇÕES PASSO-A-PASSO

Elevar ao quadrado cada resultado individual para ambos os grupos em separado:

1. Adicionar os totais dos resultados ao quadrado para cada grupo;
2. Elevar ao quadrado todos os resultados individuais para cada grupo;
3. Calcular a média para cada grupo;
4. Calcular t:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\frac{\sum x_1^2 - (\sum x_1)^2}{n_1} + \frac{\sum x_2^2 - (\sum x_2)^2}{n_2}}{(n_1 - 1) + (n_2 - 1)}} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

em que:

\bar{x}_1 = média do grupo 1

\bar{x}_2 = média do grupo 2

$\sum x_1^2$ = soma dos quadrados para o grupo 1

$\sum x_2^2$ = soma dos quadrados para o grupo 2

$(\sum x_1)^2$ = resultados totais do grupo 1 ao quadrado

$(\sum x_2)^2$ = resultados totais do grupo 2 ao quadrado

n_1 = número de sujeitos do grupo 1

n_2 = número de sujeitos do grupo 2

Calcular g.l.:

$$g.l. = (n_1 - 1) + (n_2 - 1)$$

Se t observado > t crítico rejeita-se H_0

Se t observado < t crítico não se rejeita H_0

Exemplo: para verificar se duas dietas para emagrecer são igualmente eficazes, um médico separou ao acaso um conjunto de pacientes em dois grupos. Cada paciente seguiu a dieta designada para o seu grupo durante 4 meses. O médico registou a perda de peso em kg de cada paciente por grupo. Os dados estão apresentados no quadro que se segue:

Quadro 26: Teste t não relacionado

Grupo 1 (dieta 1)		Grupo 2 (dieta 2)	
Resultados	Resultados ao quadrado	Resultados	Resultados ao quadrado
10	100	2	4
5	25	1	1
6	36	7	49
3	9	4	16
9	81	4	16
8	64	5	25
7	49	2	4
5	25	5	25
6	36	3	9
5	25	4	16
Total	$\Sigma x_1 = 64$	$\Sigma x_2 = 37$	$\Sigma x_2^2 = 165$
Média	$\bar{\chi}_1 = 6,4$	$\bar{\chi}^2 = 3,7$	

Instruções Passo-a-Passo:

1. construir a tabela

2. $\Sigma x_1 = 64$

$\Sigma x_2 = 37$

$\Sigma x_1^2 = 450$

$\Sigma x_2^2 = 165$

3. $(\Sigma x_1)^2 = 64 \times 64 = 4096$

$(\Sigma x_2)^2 = 37 \times 37 = 1369$

4. $\bar{\chi}_1 = 6,4$

$\bar{\chi}_2 = 3,7$

$$t = \frac{6,4 - 3,7}{\sqrt{\frac{\frac{450 - 4096}{10} + \frac{165 - 1369}{10}}{9 + 9} \left(\frac{1}{10} + \frac{1}{10} \right)}} = \frac{2,7}{3,806 \cdot 0,2} = 3,547$$

$$6. \text{ g.l.} = (n1 - 1) + (n2 - 1) = (10 - 1) + (10 - 1) = 18$$

O valor de t observado (3,547) é superior ao valor crítico de 2,878 (Anexo VII), indicando um nível de significância de $p < 0,05$ para uma hipótese unicaudal ou $p < 0,01$ para uma hipótese bicaudal.

10.3 TESTE T DE STUDENT (RELACIONADO)

Quando utilizar:

Utiliza-se para designs experimentais com duas situações testando uma variável independente, quando os mesmos sujeitos (ou emparelhados) se encontram em ambas as situações - design relacionado. O teste t relacionado é equivalente ao teste não paramétrico de Wilcoxon.

OBJECTIVO

O objectivo é comparar as diferenças entre as duas situações experimentais com a variabilidade total nos resultados. Quando os mesmos sujeitos são usados em ambas as situações podem comparar-se pares de resultados obtidos por cada indivíduo quando sujeito a ambas as situações.

A estatística t apresenta o tamanho das diferenças entre os resultados dos sujeitos para as duas situações. Para que seja significativo o valor de t terá de ser igual ou superior aos valores críticos da tabela (Anexo VII).

INSTRUÇÕES PASSO-A-PASSO

1. Calcular as diferenças entre os resultados dos sujeitos subtraindo os resultados da situação B para a situação A
2. Elevar essas diferenças ao quadrado
3. Calcular o somatório das diferenças obtidas (Σd)
4. Calcular o somatório do quadrado das diferenças (Σd^2)
5. Elevar ao quadrado as diferenças totais $(\Sigma d)^2$
6. Calcular t:

$$t = \frac{\Sigma d}{\sqrt{\frac{N \Sigma d^2 - (\Sigma d)^2}{N - 1}}}$$

em que

Σd = soma das diferenças dos resultados A e B

Σd^2 = soma dos quadrados das diferenças

$(\Sigma d)^2$ = soma das diferenças elevadas ao quadrado

N = número de sujeitos

7. Calcular g.l.:

$$g.l. = N - 1$$

Se t observado > t crítico rejeita-se H0

Se t observado < t crítico não se rejeita H0

Exemplo: Para verificar se a eficácia de uma dieta era influenciada pelo ministrar de um fármaco, um médico decidiu administrar, a um grupo de indivíduos que o tinham procurado para perder peso, um placebo em conjunto com uma dieta que já havia administrado um mês antes aos mesmos sujeitos. Referiu aos seus casos que aquele medicamento servia para perder apetite e ajudava a queimar gorduras.

Registou a perda de peso que tinha ocorrido nos 30 dias antecedentes à tomada de placebo e trinta dias após o placebo. Os resultados estão no quadro que se segue

Quadro 27: Teste t relacionado

Sujeito	Situação A (com placebo)	Situação B (só com dieta)	d (A-B)	d ²
1	10	2	8	64
2	5	1	4	16
3	6	7	-1	1
4	3	4	-1	1
5	9	4	5	25
6	8	5	3	9
7	7	2	5	25
8	5	5	0	0
9	6	3	3	9
10	5	4	1	1
Total	64	37	$\Sigma d = 27$	$\Sigma d^2 = 151$
Média	6,4	3,7		

Resolução 50:

Instruções Passo-a-Passo:

1. construir tabela

2. calcular as médias

3. $\Sigma d = 27$

4. $\Sigma d^2 = 151$

5. $(\Sigma d)^2 = 27 \times 27 = 729$

6. proceder aos calculos

$$t = \frac{27}{\sqrt{\frac{10 \times 151 - (27)^2}{10 - 1}}} = \frac{27}{\sqrt{86,78}} = 27 / 9,316 = 2,90$$

7. g.l. = N - 1 = 10 - 1 = 9

Como o valor observado de t de 2,90 é superior ao valor crítico de 2,821 ao nível de significância de $p < 0,01$ (unicaudal) ou $p < 0,02$ (bicaudal), e a média de perda de peso é superior com o placebo, o médico concluiu que o efeito de placebo é positivo na perda de peso

O ponto 10.3 explica como comparar médias de duas populações, com base em amostras dessas populações. Mas às vezes é preciso comparar médias de mais de duas populações. Por exemplo, para verificar se pessoas com diferentes níveis socioeconómicos, isto é, alto, médio e baixo têm, em média, o mesmo peso corporal, é preciso comparar médias de três populações.

Para comparar médias de mais de duas populações aplica-se o teste F, na forma que a seguir se descreve, desde que a variável em estudo tenha distribuição normal ou aproximadamente normal. Mas antes de mostrar como se faz esse teste, convém apresentar um exemplo.

Imagine que quatro grupos, todos com cinco elementos, conduziram aos dados apresentados no quadro 28. As médias dessas amostras estão na última linha desse quadro. Será que as diferenças das médias das amostras são suficientemente grandes para que se possa afirmar que as médias das populações são diferentes? Para responder a esta pergunta, é preciso um teste estatístico.

Quadro 28: Dados de 4 amostras e respectivas médias

	A	B	C	D
	11	8	5	4
	8	5	7	4
	5	2	3	2
	8	5	3	0
	8	5	7	0
Σ	40	25	25	10
\bar{x}	8	5	5	2

10.4 ANÁLISE DE VARIÂNCIA

10.4.1 Análise da Variância com Igual Tamanho

Se a variável em estudo tem distribuição normal ou aproximadamente normal, para comparar mais de duas médias aplica-se o teste F .

Primeiro, é preciso estudar as *causas de variação*. Por que é os dados variam? Uma explicação é o facto de as amostras provirem de populações diferentes. Outra explicação é o acaso, porque até mesmo os dados provenientes de uma mesma população variam.

O teste F é feito através de uma *análise de variância*, que separa a variabilidade devido aos "tratamentos" (no exemplo, devido às amostras terem provindo de populações diferentes) da variabilidade residual, isto é, devido ao acaso. Para aplicar o teste F é preciso fazer uma série de cálculos, que exigem conhecimento da notação.

O quadro 29 apresenta os dados de k tratamentos, cada um com r repetições (no exemplo, denominam-se repetições os elementos da mesma amostra). A soma das repetições (r) de um mesmo tratamento constitui o *total* desse tratamento (T). O *total geral* (ΣT) é dado pela soma dos k totais de tratamentos.

Quadro 29: Notação para análise da variância

	1	2	3	...	K	Total
	x_{11}	x_{21}	x_{31}	...	x_{k1}	
	x_{12}	x_{22}	x_{32}	...	x_{k2}	
	
	
	
	x_{1r}	x_{2r}	x_{3r}		x_{kr}	
Total	T_1	T_1	T_1		T_1	$\Sigma T = \Sigma x$
N.º de repetições	r	r	r		r	$N = kr$
Média	\bar{x}_1	\bar{x}_2	\bar{x}_3		\bar{x}_k	

Para fazer a análise de variância é preciso calcular as seguintes quantidades:

1.º passo:

- a) os graus de liberdade (gl) dos grupos: $k - 1$
- b) gl do total: $n-1$
- c) gl dos resíduos: $(n-1)-(k-1) = n-k$

2.º passo:

- a) calcular o valor de Correção (C) que é dado pelo total geral ao quadrado e dividido pelo número de dados.

FORMULA:
$$C = \frac{(\sum x)^2}{n}$$

- b) calcular a Soma dos Quadrados Total (SQT)

FORMULA:
$$SQT = \sum x^2 - C$$

- c) calcular a Soma do Quadrado do Total de cada repetição (SQTr)

FORMULA:
$$SQTr = \frac{\sum T^2}{r} - C$$

- d) calcular a Soma dos Quadrados dos Resíduos (SQR)

FORMULA:
$$SQR = SQT - SQTr$$

- e) calcular o Quadrado médio do Total de cada repetição (QMTr)

FORMULA:
$$QMTr = \frac{SQTr}{k-1}$$

- f) calcular o Quadrado médio do Total do Resíduo (QMR)

FORMULA:
$$QMR = \frac{SQR}{n-k}$$

g) calcular o valor de F

FORMULA:
$$F = \frac{QMT_r}{QMR}$$

3.º Passo: comparar o F calculado com o valor dado na tabela de F, ao nível de significância estabelecido, observando os k-1 graus de liberdade no numerador e os n-k graus de liberdade no denominador (coluna da esquerda).

Exemplo: Um profissional de saúde recém contratado para acompanhar um conjunto de atletas de alta competição, verificou, pelos registos clínicos deixados pelo seu antecessor, que alguns atletas com o mesmo tipo de lesão (em grau e extensão) tinham mais recidivas que outros, apesar das condições de treino e o tempo de recuperação ser o mesmo. Colocou a hipótese de que tal acontecimento se podia dever às diferentes terapêuticas que eram utilizadas para tratar as mesmas lesões. Os resultados podem ser observados no quadro que se segue:

Quadro 30: recidivas por tratamento

Tratamento	Tratamento	Tratamento	Tratamento
A	B	C	D
11	8	5	4
8	5	7	4
5	2	3	2
8	5	3	0
8	5	7	0
Σ 40	25	25	10
\bar{x} 8	5	5	2

1.º passo:

- a) os graus de liberdade (gl) dos grupos: $k - 1 = 4 - 1 = 3$
- b) gl do total: $n - 1 = 20 - 1 = 19$
- c) gl dos resíduos: $n - k = 20 - 4 = 16$

2.º passo:

- a) calcular o valor de Correção (C) que é dado pelo total geral ao quadrado e dividido pelo número de dados.

$$C = \frac{(\sum x)^2}{N} = \frac{(11+8+5+4+8+5+7+4+5+2+3+2+8+5+3+8+5+7)^2}{20} = \frac{100^2}{20} = 500$$

- b) calcular a Soma dos Quadrados Total (SQT)

$$\begin{aligned} SQT &= \sum x^2 - C = \\ 11^2 + 8^2 + 5^2 + 4^2 + 8^2 + 5^2 + 7^2 + 4^2 + 5^2 + 2^2 + 3^2 + 2^2 + 8^2 + 5^2 + 3^2 + 8^2 + 5^2 + 7^2 &- 500 \\ &= 658 - 500 \\ &= 158 \end{aligned}$$

- c) calcular a Soma do Quadrado do Total de cada repetição (SQTr)

$$\begin{aligned} SQTr &= \frac{\sum T^2}{R} - C = \frac{40^2 + 25^2 + 25^2 + 10^2}{5} - 500 \\ &= 590 - 500 \\ &= 90 \end{aligned}$$

- d) calcular a Soma dos Quadrados dos Resíduos (SQR)

$$\begin{aligned} SQR &= SQT - SQTr = 158 - 90 \\ &= 68 \end{aligned}$$

- e) calcular o Quadrado médio do Total de cada repetição (QMTr)

$$QMTr = \frac{SQTr}{k-1} = \frac{90}{3} = 30$$

- f) calcular o Quadrado médio do Resíduo (QMR)

$$QMR = \frac{SQR}{n-k} = \frac{68}{16} = 4,25$$

- g) calcular o valor de F

$$F = \frac{QMTr}{QMR} = \frac{30}{4,25} = 7,06$$

3.º passo: ir à tabela F para um nível de significância (p) de 5% (0,05) e observar qual o F teórico para 3 e 16 graus de liberdade (anexo). Como o valor calculado (7,06) é maior que o da tabela (3,24), concluímos que as médias das recidivas diferem em função do tratamento, para um nível de significância de 0,05.

A acompanhar este comentário, os valores calculados devem ser apresentados num quadro, da seguinte forma:

Quadro 31: ANOVA

Causas de variação	gl	SQ	QM	F	p
Tratamentos	3	90	30	7,06	<0,05
Resíduo	16	68	4,25		
Total	19	158			

Mas, como se pode observar, apesar da tabela mostrar que existem diferenças significativas, não nos informa, que tratamentos é que produzem diferenças e quais são semelhantes. Sempre que as diferenças são significativas, e só nesse caso, temos que proceder às comparações à posteriori (Post-Hoc). Podemo-nos socorrer de diversos testes (LSD; Bonferroni; Sidak; Scheffe; SNK; Tukey; etc.), a grande diferença entre eles reside no tipo de distribuição em que assentam e no tipo de ajustamento). Apresentaremos de seguida apenas o teste de Tukey, por ser dos mais utilizados e o mais simples de calcular, quando recorremos ao cálculo manual.

10.4.1.1 TESTE DE TUKEY PARA COMPARAÇÃO ENTRE AS MÉDIAS

O teste Tukey permite estabelecer a diferença mínima significante, ou seja, a menor diferença entre as médias que deve ser tomada como significativa em determinado nível de significância. Essa diferença (dms) é dada por:

$$dms = q \sqrt{\frac{QMR}{r}}$$

Onde q é um valor dado em tabela (Anexo____)

QMR é o quadrado médio do residuo da ANOVA

r é o número de repetições

assim, se consultarmos a tabela verificamos que o q para comparar quatro tratamentos com 16 gl no residuo é de 4,05. como QMR=4,25 e r=5, temos:

$$dms = 4,05 \sqrt{\frac{4,25}{5}} = 3,73$$

De acordo com o teste de Tukey, duas médias são estatisticamente diferentes sempre que o valor absoluto da diferença entre elas for igual ou superior ao valor da dms.

Passemos então à observação dos valores:

Quadro 32: post-hoc das recidivas			
Pares de médias	Valor absoluto da diferença	dms	p
A-B	(8-5) 3	3,73	ns
A-C	(8-5) 3		ns
A-D	(8-2) 6		<0,05
B-C	(5-5) 0		ns
B-D	(5-2) 3		ns
C-D	(5-2) 3		ns

É fácil de observar que só existem diferenças entre a média dos tratamentos A e a média dos tratamentos D, em que o tratamento D é aquele com que se obtém, significativamente, menos recidivas

10.4.2 ANÁLISE DE VARIÂNCIA COM DIFERENTES TAMANHOS

O pesquisador, nem sempre tem amostras do mesmo tamanho, mesmo assim é possível conduzir uma análise da variância (ANOVA). Aliás todos os cálculos, com excepção SQTr, são feitos da mesma forma em ambas as situações.

Assim em vez de fazer a soma dos quadrados pela fórmula

$$SQTr = \frac{\sum T^2}{r} - C$$

Utiliza:

FORMULA:
$$SQTr = \frac{T_1^2}{r_1} + \frac{T_2^2}{r_2} + \dots + \frac{T_k^2}{r_k} - C$$

Onde C é a correcção definida anteriormente

Para mais fácil de entender observe o seguinte exemplo:

Quadro 33: médias por tratamento

	Tratamento A	Tratamento B	Tratamento C
	15	23	19
	10	16	15
	13	19	21
	18	18	14
	15		16
\bar{x}	84	76	85

1.º passo:

- os graus de liberdade (gl) dos grupos: $k - 1 = 3 - 1 = 2$
- gl do total: $n - 1 = 15 - 1 = 14$
- gl dos resíduos: $n - k = 14 - 2 = 12$

2.º passo:

- a) calcular o valor de Correção (C) que é dado pelo total geral ao quadrado e dividido pelo número de dados.

$$C = \frac{(\sum x)^2}{n} = \frac{245^2}{15} = 4001,67$$

- b) calcular a Soma dos Quadrados Total (SQT)

$$SQT = \sum x^2 - C = 15^2 + 10^2 + 13^2 + \dots + 16^2 - 4001,67 = 159,33$$

- c) calcular a Soma do Quadrado do Total de cada repetição (SQTrat)

$$SQTrat = \frac{T_1^2}{r_1} + \frac{T_2^2}{r_2} + \dots + \frac{T_k^2}{r_k} - C$$

$$SQTr = \frac{84^2}{6} + \frac{76^2}{4} + \frac{85^2}{5} - 4001,67 = 63,33$$

- d) calcular a Soma dos Quadrados dos Resíduos (SQR)

$$SQR = SQT - SQTr = 159,33 - 63,33 = 96$$

- e) calcular o Quadrado médio do Total de cada repetição (QMTr)

$$QMTr = \frac{SQTr}{k-1} = \frac{63,33}{2} = 31,67$$

- f) calcular o Quadrado médio do Resíduo (QMR)

$$QMR = \frac{SQR}{n-k} = \frac{96}{12} = 8$$

- g) calcular o valor de F

$$F = \frac{QMTr}{QMR} = \frac{31,67}{8} = 3,96$$

3.º passo: ir à tabela F para um nível de significância (p) de 5% (0,05) e observar qual o F teórico para 2 e 12 graus de liberdade (anexo). Como o valor calculado (3,96) é maior que o da tabela (3,89), concluímos que as médias dos sujeitos diferem em função do tratamento, para um nível de significância de 0,05.

Os valores calculados devem ser apresentados da seguinte forma:

Quadro 34: ANOVA

Causas de variação	gl	SQ	QM	F	p
Tratamentos	2	63,33	31,67	3,96	<0,05
Resíduo	12	96	8		
Total	14	159,33			

Como as diferenças são significativas, temos que proceder às comparações à posteriori (Post-Hoc). O resultado do teste de Tukey é o seguinte:

10.4.2.1 TESTE DE TUKEY PARA COMPARAÇÃO ENTRE AS MÉDIAS

O teste Tukey para amostras com tamanhos diferentes é dada pela seguinte fórmula:

$$dms = q \sqrt{\left(\frac{1}{r_i} + \frac{1}{r_j} \right) \frac{QMR}{2}}$$

No caso do exemplo, para comparar a média de A com a média de B tem-se:

$$dms (A;B) = 3,77 \quad \sqrt{\left(\frac{1}{6} + \frac{1}{4} \right) \frac{8,00}{2}}$$

dms (A;B) = 4,87

No caso do exemplo, para comparar a média de A com a média de C tem-se:

$$dms (A;C) = 3,77 \quad \sqrt{\left(\frac{1}{6} + \frac{1}{5} \right) \frac{8,00}{2}}$$

dms (A;C) = 4,57

No caso do exemplo, para comparar a média de B com a média de C tem-se:

$$\text{dms (B;C)} = 3,77 \quad \sqrt{\left(\frac{1}{4} + \frac{1}{5}\right) \frac{8,00}{2}}$$

$$\text{dms (B;C)} = 5,06$$

De acordo com o teste de Tukey, duas médias são estatisticamente diferentes sempre que o valor absoluto da diferença entre elas for igual ou superior ao valor da dms.

Passemos então à observação dos valores:

Quadro 35: post-hoc			
Pares de médias	Valor absoluto da diferença	dms	p
A-B	14-19 = 5	4,87	<0,05
A-C	14-17 = 3	4,57	ns
B-C	19-17 = 2	5,06	ns

Conclui-se que em média A é significativamente diferente de B, ao nível de significância de 0,05.

ANEXOS

Anexo I

TABELA DO QUIQUADRADO (χ^2)

Anexo II

TABELA DE U

Anexo III

TABELA DE WILCOXON

Anexo IV

TABELA H

Anexo V

TABELAS DE FRIEDMAN

Anexo VI

TABELA S (TABELA DE KENDAL)

Anexo VII

TABELA T

Anexo VIII

TABELA Z

Anexo IX

TABELA F

Anexo X

TABELA DOS VALORES DE q