

Capítulo 1

Associação entre variáveis

Por vezes interessa estudar possíveis relações entre variáveis. Quantas vezes não nos perguntamos se duas variáveis são independentes ou pelo contrário dependentes? E se forem dependentes, qual o tipo e grau de dependência? A resposta a estas questões pode levar a diferentes métodos de análise estatística como já vimos, por exemplo, na comparação da localização entre dois grupos de observações (amostras emparelhadas versus amostras independentes).

Infelizmente não é fácil responder a estas questões porque não existe nenhuma forma completa de medir a dependência entre variáveis. Duas variáveis podem depender de tantas formas diferentes que se torna impossível encontrar uma forma única de avaliar a dependência. Medir o grau de dependência é mais ambicioso do que simplesmente testar a existência de alguma associação entre variáveis. É obviamente de interesse poder medir o grau de associação entre dois conjuntos de observações obtidos a partir de um dado conjunto de unidades experimentais (indivíduos por exemplo). Mas em geral, ficamos satisfeitos se conseguirmos, pelo menos, dizer se uma certa associação observada nos dados indica ou não uma associação na população de onde foram retirados.

Uma das formas mais habituais de associação entre variáveis é a linear. Uma relação puramente linear entre duas variáveis traduz-se num gráfico de dispersão onde os pontos se encontram dispostos sobre uma recta. Outras formas de associação incluem as relações exponenciais logarítmicas ou quadráticas.

Uma vez que as variáveis que observamos em Estatística estão sujeitas a erros ou outros factores aleatórios, não estamos à espera de observar relações matemáticas puras entre duas amostras de observações mas sim qualquer coisa do tipo apresentado nas figuras seguintes.

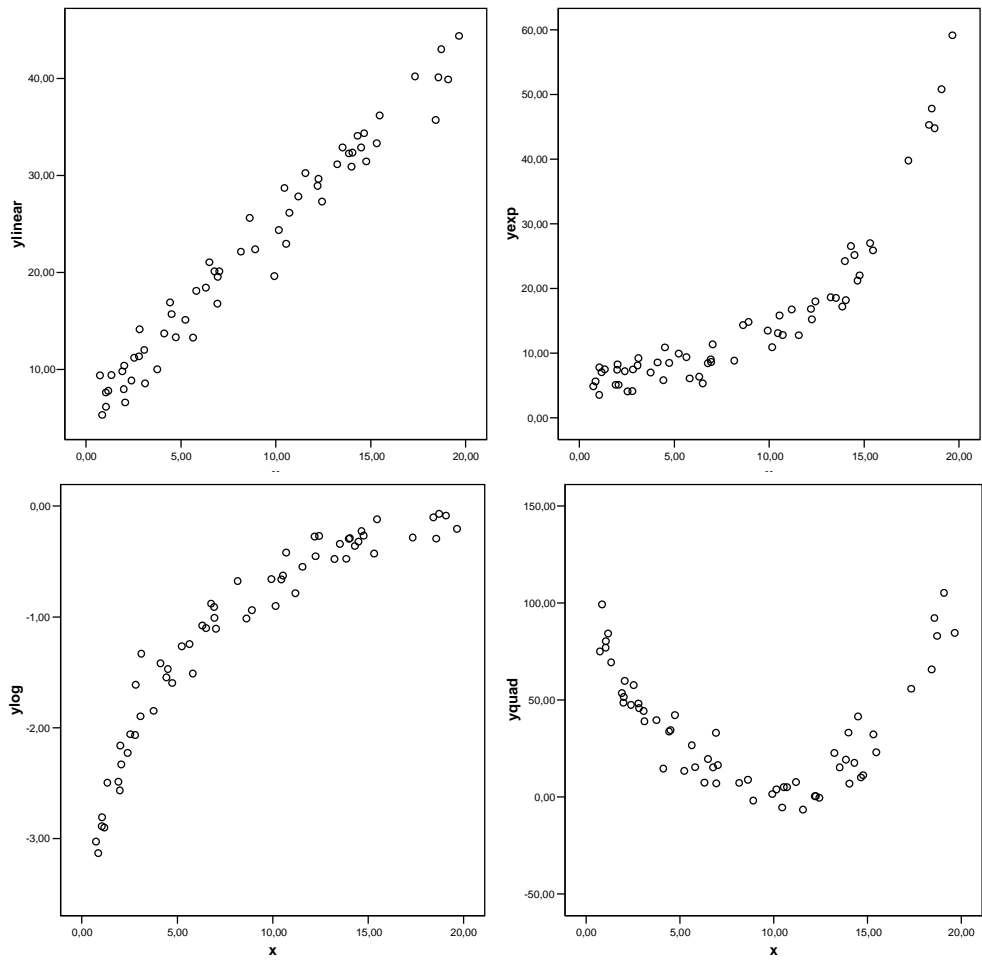


Figura 1.1: Relações entre variáveis aleatórias: linear, exponencial, logarítmica e quadrática

Como proceder então para medir ou avaliar a associação entre variáveis? Iremos abordar vários métodos que se aplicam em situações variadas.

Um primeiro passo fundamental é a construção de diagramas de dispersão. Quando duas variáveis são independentes, o diagrama de dispersão respectivo apresenta uma mancha de pontos aleatória (ou quando muito) um conjunto de pontos dispostos sobre uma recta horizontal.

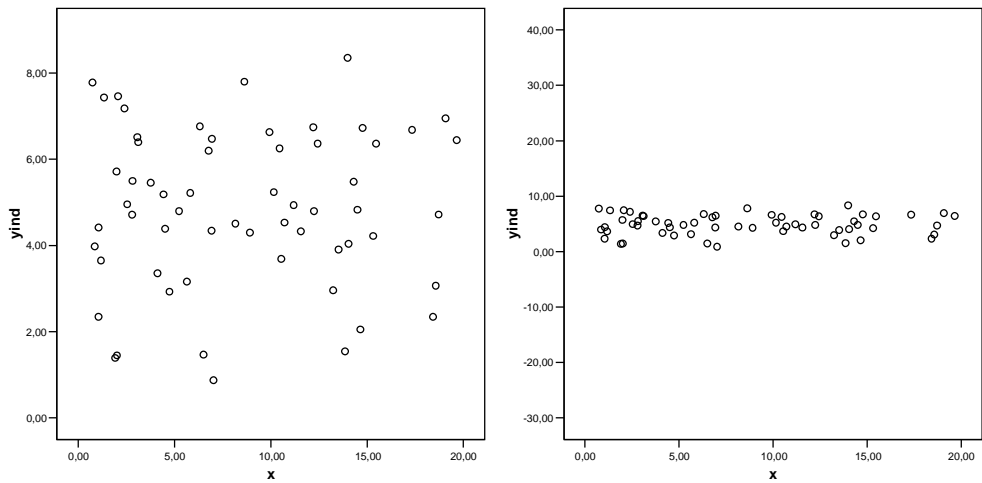


Figura 1.2: Diagramas de dispersão de variáveis independentes

Se a relação entre duas variáveis for linear, ao confrontarmos duas amostras num diagrama de dispersão devemos esperar observar um conjunto de pontos que se dispõem aproximadamente sobre uma recta. Por vezes os desvios em relação à recta são mínimos, mas noutras os pontos apresentam bastante dispersão tornando difícil a identificação da dita relação linear.

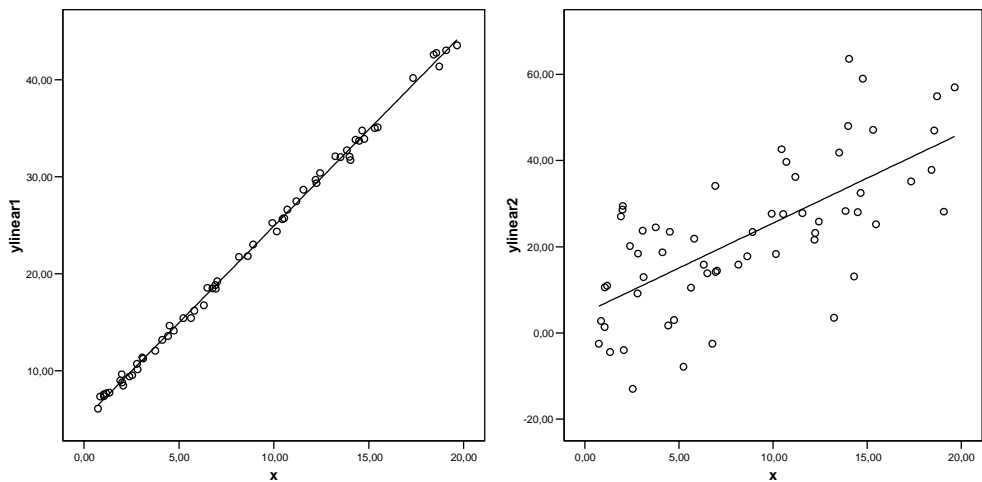


Figura 1.3: Relações entre variáveis aleatórias: fortemente linear e fracamente linear

O passo seguinte consiste em calcular medidas de associação que em geral requerem alguns pressupostos sobre o tipo de dados e o tipo de relação entre as variáveis. Quando as variáveis são ambas numéricas as medidas de associação são habitualmente designadas por **coeficientes de correlação**.

Um último passo habitual é o de realizar um teste de hipóteses para averiguar se os valores das medidas de associação observados nos dados são significativos, ou seja se podemos estatisticamente concluir a favor de uma associação na população.

1.1 Medidas de associação para dados numéricos ou ordinais

1.1.1 O coeficiente de correlação de Pearson

Dadas duas amostras de observações medidas numa escala de intervalos ou razões, podemos medir o grau de associação **linear** entre elas através do **coeficiente de correlação de Pearson** ou simplesmente **coeficiente de correlação amostral** (*Pearson product-moment correlation coefficient*).

Se representarmos as amostras por X_1, \dots, X_n e Y_1, \dots, Y_n o coeficiente de correlação amostral é dado por

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}.$$

Este coeficiente pode também ser escrito como

$$\sum_{i=1}^n Z_{x_i} Z_{y_i}$$

onde Z_x e Z_y representam as observações padronizadas, ou seja, após subtrairmos a média e dividirmos pelo desvio padrão cada uma das amostras.

O coeficiente de correlação pode variar entre -1 e 1. R assume o valor 1 quando os pontos (x_i, y_i) estão exactamente sobre uma recta de declive positivo. Neste caso um aumento numa das variáveis corresponde necessariamente a um aumento na outra. R assume o valor -1 quando os pontos estão exactamente sobre uma recta de declive negativo. Nesta situação um aumento numa das variáveis corresponde a uma diminuição na outra. Estes dois casos correspondem ao máximo de associação linear que é possível observar entre duas amostras. Em geral observamos valores de R que em módulo são inferiores a 1. Quando as amostras são independentes o valor do coeficiente será próximo de zero ou mesmo zero. É difícil avaliar o significado de um valor observado para r . Por vezes $r = 0.6$ representa pouca associação e $r = 0.4$ pode representar muita associação.

Uma interpretação usual do coeficiente de correlação amostral passa por considerar o seu valor elevado ao quadrado, R^2 , a que se chama coeficiente de determinação. Uma vez que $-1 \leq R \leq 1$ o coeficiente de determinação está sempre entre 0 e 1. Este coeficiente tem uma interpretação directa: ele mede a percentagem de variabilidade de uma das variáveis explicada pela outra. Dito de outra forma, quando duas variáveis aleatórias estão associadas, a variação de uma delas implica algum tipo de variação na outra. Uma medida de variabilidade é a variância. Dizer que o coeficiente de determinação é por exemplo 0,6 significa que 60% da variância de uma das variáveis é provocada (explicada) pela outra variável e apenas 40% é de natureza independente.

Para termos uma leitura mais informativa deste coeficiente é usual proceder a um teste de hipóteses que testa as seguintes hipóteses:

$$H_0 : \rho = 0 \quad vs \quad H_1 : \rho \neq 0$$

onde ρ representa o coeficiente de correlação da população onde foram retirados os dados. Este teste tem como pressupostos que:

1. ambas as populações de onde foram retirados as amostras têm distribuição Normal,
2. a relação entre as variáveis é de forma linear, caso exista.

É importante saber que:

- no caso de variáveis Normais relacionadas linearmente, um coeficiente de correlação nulo é equivalente a independência. No entanto se as distribuições não forem Normais essa equivalência pode não se observar. Portanto, duma maneira geral, ao considerar uma relação de tipo linear, um coeficiente de correlação nulo não implica necessariamente independência.
- por outro lado, quando duas variáveis são independentes $\rho = 0$. Isto deverá traduzir-se num coeficiente de correlação amostral "próximo" de zero. No entanto, há situações em que $\rho = 0$ e as variáveis são dependentes, possivelmente com outro tipo de relação que não a linear. Portanto, o facto de observarmos um coeficiente de correlação nulo pode apenas significar que não existe uma relação linear ente as variáveis.

Na presença de diferentes padrões de associação podemos obter os mais variados valores para R , sempre no intervalo $[-1, 1]$. Pode acontecer que padrões completamente distintos conduzam ao mesmo valor de R . Portanto, R deve ser calculado sempre após a construção de um diagrama de dispersão para termos uma noção clara se estamos de facto a medir associação linear ou não. Em geral, para podermos concluir a favor de uma associação linear, o aspecto do diagrama de dispersão deve sugerir uma recta e não deve ficar muito alterado após a remoção aleatória de um conjunto reduzido de observações.

Resumindo, o coeficiente de correlação de Pearson mede o grau de associação linear entre duas variáveis medidas numa escala de intervalos ou razões. Se as variáveis tiverem distribuição Normal podemos efectuar um teste de hipóteses para averiguar se o coeficiente de correlação da população é significativamente diferente de zero, o que significará nesse contexto que as variáveis são independentes. Convém sempre construir um diagrama de dispersão para ter uma ideia sobre a linearidade da relação entre as variáveis.

No SPSS o coeficiente de Pearson pode ser obtido através do menu **Analyze / Correlate / Bivariate**.

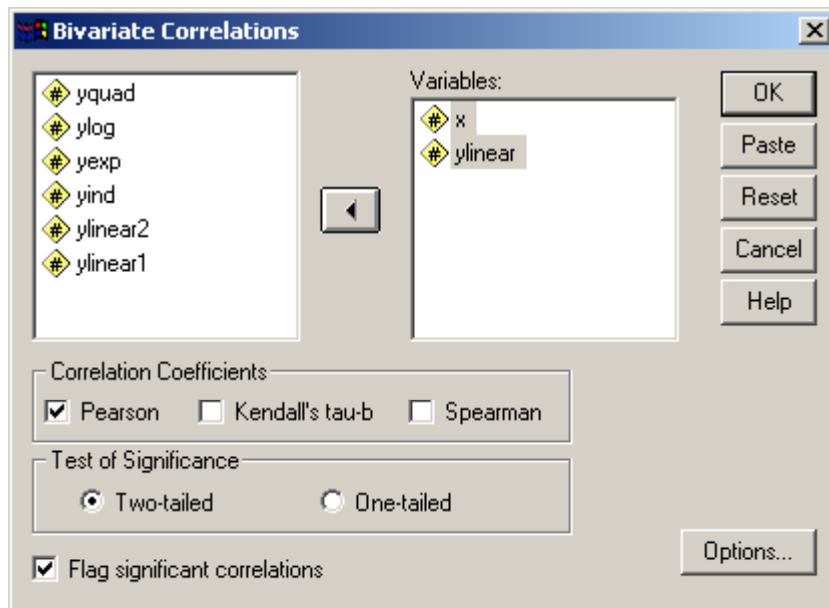


Figura 1.4: o menu da correlação entre duas amostras

Neste menu podem-se seleccionar mais do que duas variáveis, caso em que o SPSS fornece uma tabela de correlações para todas as combinações de pares de variáveis. O SPSS fornece também o p-value do teste acima referido, para cada par de variáveis.

1.1.2 O coeficiente de correlação de Spearman

Se estivermos perante duas variáveis medidas apenas numa escala ordinal, ou que apresentem uma relação não linear mas monótona (se uma aumenta a outra tem sempre tendência a aumentar (ou a diminuir)), o coeficiente de Pearson não pode ser aplicado. Mais ainda mesmo quando estamos em condições de aplicar o coeficiente de Pearson mas não podemos garantir os pressupostos da realização do teste de hipóteses a esse coeficiente, temos como alternativa o **coeficiente de correlação de Spearman** (*Spearman rank-order coefficient*). A ideia de construção deste coeficiente é bastante simples. Dadas duas amostras de observação ordenáveis, substitui-se cada um dos seus valores pela sua ordem de ordenação, em inglês *rank*. Por exemplo, se uma amostra de três valores for $x_1 = 2.1, x_2 = 1.7, x_3 = 4.8$, então os respectivos *ranks* serão $r_1 = 2, r_2 = 1, r_3 = 3$. Após substituir cada uma das amostras pelos seus *ranks* o coeficiente de Spearman não é mais do que o coeficiente de Pearson aplicado agora aos *ranks*.

De notar que se a relação entre as variáveis originais for monótona, a relação entre os *ranks* é necessariamente linear e portanto estamos em condições de calcular o coeficiente de Pearson.

Uma vez que as ordens variam sempre entre 1 e n (número de observações), pode-se re-

escrever o coeficiente da seguinte forma

$$R_s = 1 - \frac{6 \sum_{i=1}^n D_i^2}{n^3 - n}$$

onde D_i representa a diferença de *ranks* correspondentes a cada par de observações X_i, Y_i .

Tal como no caso do coeficiente de Pearson é possível testar as hipóteses

$$H_0 : \rho = 0 \quad vs \quad H_1 : \rho \neq 0.$$

Embora um coeficiente nulo não implique independência total (conforme já foi referido anteriormente) este teste é utilizado na prática para averiguar se a associação entre variáveis é significativa ou não, entendendo-se por associação uma correlação não nula.

Tal como para o coeficiente de Pearson, no SPSS o coeficiente de Spearman pode ser obtido através do menu **Analyse / Correalte / Bivariate**.

Mais uma vez, neste menu podem-se seleccionar mais do que duas variáveis, caso em que o SPSS fornece uma tabela de correlações para todas as combinações de pares de variáveis. O SPSS fornece também o p-value do teste acima referido, para cada par de variáveis.

1.1.3 O coeficiente de correlação τ de Kendall

Uma alternativa ao coeficiente de Spearman é o coeficiente τ de Kendall (*Kendall's tau coefficient*) que se aplica nas mesmas condições.

Podemos referir duas eventuais vantagens deste coeficiente sobre o anterior: se as amostras tiverem dimensão muito reduzida e valores repetidos, os resultados do teste ao coeficiente são mais precisos no caso presente; por outro lado, o coeficiente τ de Kendall pode ser generalizado para correlações parciais que são correlações medidas entre duas variáveis após remoção do efeito de uma possível terceira variável sobre ambas. Por exemplo se estudarmos a relação entre o domínio da linguagem ea altura de crianças em idade escolar iremos certamente encontrar uma associação. Mas essa associação não reflecte uma relação directa mas sim resulta do facto de ambas as variáveis estarem directamente relacionadas com a idade.

Uma diferença muito importante entre os dois coeficientes (Kendall e Spearman) reside na sua interpretação e na impossibilidade de comparar directamente valores provenientes de ambos. Embora o objectivo comum seja o de medir associação, a forma de o fazer é distinta.

O coeficiente de Kendall é muitas vezes descrito como uma medida de concordância entre dois conjuntos de classificações relativas a um conjunto de objectos ou experiências. Por exemplo, se pedirmos a dois professores para classificarem (por ordem crescente de qualidade) um conjunto de exames podemos tentar medir o grau de concordância entre os dois avaliadores através do coeficiente de Kendall. Basicamente este coeficiente mede a diferença entre a probabilidade de as classificações estarem na mesma ordem e a probabilidade de estarem em ordens diferentes. Do ponto de vista amostral estas probabilidades são dadas através das frequências relativas respectivas.

$$T = \frac{\# \text{concordâncias} - \# \text{discordâncias}}{\text{número total de pares possíveis}}$$

De notar que o conceito de concordância utilizado nesta expressão não equivale a classificações iguais em ambos os avaliadores, mas sim qualquer par que aponte num sentido comum de classificação. Assim, para determinar o número de concordâncias há que inicialmente ordenar as classificações de acordo com um dos avaliadores e a partir daí contar os pares que vão no mesmo sentido. Por exemplo, se as classificações forem

Objecto	a	b	c	d
Avaliador 1	3	4	2	1
Avaliador 2	3	1	4	2

após ordenarmos as classificações pelo avaliador 1 temos

Objecto	d	c	a	b
Avaliador 1	1	2	3	4
Avaliador 2	2	4	3	1

e o número de concordâncias é contado somando todas as classificações do avaliador 2 que vão no mesmo sentido (crescente) do avaliador 1. Assim, começando pelo *rank* 2, que é a primeira classificação do avaliador 2, temos duas classificações superiores (4 e 3) à sua direita (concordâncias), e apenas uma classificação inferior (1) (discordância). Para a classificação seguinte (4) temos 0 concordâncias (à direita) e 2 discordâncias. Finalmente para a classificação (3) temos 0 concordâncias e 1 discordância. No total temos 2 concordâncias e 4 discordâncias. O denominador do coeficiente de Kendall assume o valor $n(n-1)/2$ onde n é o número de objectos classificados, neste caso $n = 4$.

Devido à sua interpretação, o coeficiente de Kendall é utilizado para avaliar a fiabilidade de observações obtidas por diferentes avaliadores ou instrumentos de medida.

Tal como para os coeficientes de Pearson e Spearman é possível efectuar um teste de hipóteses para averiguar se a associação é significativa.

No SPSS o coeficiente de Kendall pode ser obtido através do menu **Analyse / Correlate / Bivariate**, o mesmo menu utilizado para obter os coeficientes de Pearson e Spearman. A versão aqui instalada (*Kendall's tau-b*) admite a ocorrência de empates nas observações de cada amostra.

Mais uma vez, neste menu podem-se seleccionar mais do que duas variáveis, caso em que o SPSS fornece uma tabela de correlações para todas as combinações de pares de variáveis. O SPSS fornece também o p-value do teste acima referido, para cada par de variáveis.

Quanto aos coeficientes de correlação parcial o SPSS apenas disponibiliza a versão mais clássica com base no coeficiente de Pearson. Estes coeficientes podem ser obtidos no menu **Analyse / Correlate / Partial**.

1.2 Medidas de associação para dados categóricos

Na presença de dados categóricos as metodologias para averiguar possíveis associações entre variáveis são distintas das apresentadas anteriormente e dependem do número de categorias das variáveis em questão.

Quando temos apenas duas categorias em cada uma de duas variáveis em estudo podemos organizar os dados numa tabela bidimensional com duas linhas e duas colunas, contendo as frequências observadas para cada combinação de categorias. Por exemplo, se estamos interessados em estudar se a ocorrência de uma certa patologia está relacionada com o sexo dos indivíduos podemos obter uma tabela como a que se segue baseada num conjunto de 100 pessoas:

Sexo	Patologia		Total
	Presente	Ausente	
Feminino	30	20	50
Masculino	15	35	50
Total	45	55	100

A este tipo de tabela chama-se **tabela de contingência** (*contingency table* ou *crosstab*). A tabela terá tantas linhas e colunas quantas as categorias em cada uma das variáveis.

Se existirem mais de duas variáveis em estudo podem-se construir tabelas com mais de duas dimensões. Neste caso a sua representação em papel ou numa folha de dados como a do SPSS torna-se um pouco mais carregada, havendo necessidade de decompor a tabela em várias tabelas bidimensionais.

As medidas de associação e respectivos testes de hipóteses para dados organizados em tabelas de contingência estão disponíveis no SPSS através do menu **Analyze / Descriptive Statistics / Crosstabs**. Dentro deste menu, podem-se seleccionar várias medidas e testes através do botão **Statistics**. As medidas e testes abordados neste texto encontram-se assinaladas com um *tick* na figura seguinte:

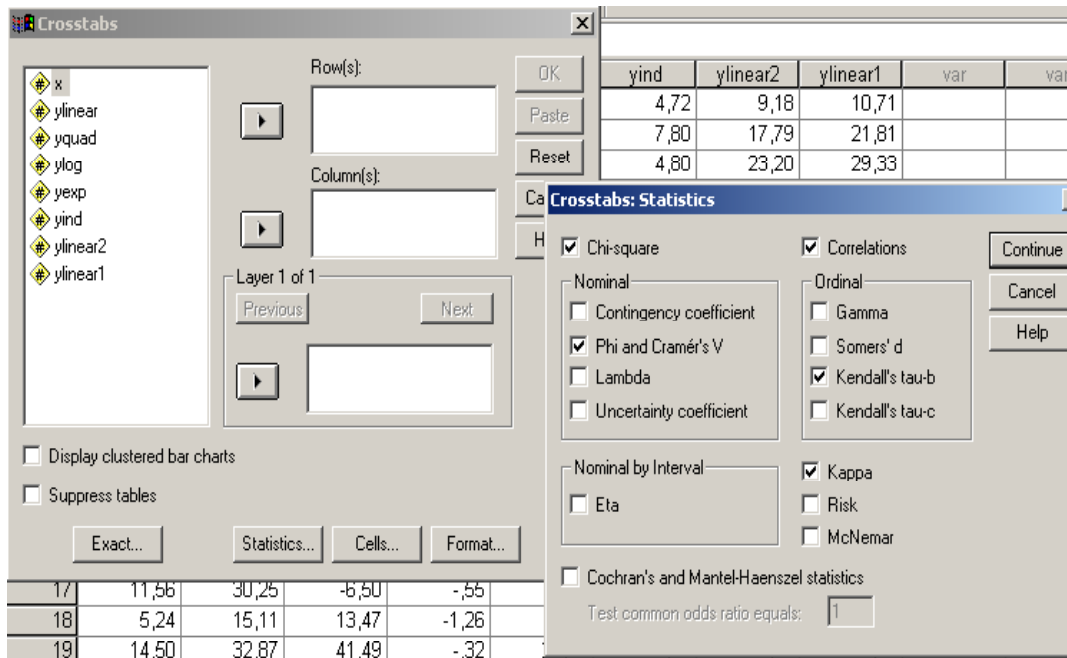


Figura 1.5: menu das medidas de associação para dados categóricos

1.2.1 O teste do χ^2

A partir de uma tabela de contingência pode-se calcular uma estatística a partir da qual é possível efectuar um teste de hipóteses designado qui-quadrado, χ^2 , para averiguar se as variáveis são independentes

H_0 : as variáveis são independentes *vs* H_1 : as variáveis são dependentes.

O mesmo teste também pode servir para averiguar se as frequências se distribuem de forma homogénea em todas as linhas (ou colunas) da tabela. Nesse caso diz-se que estamos perante um teste de homogeneidade.

Uma outra forma de aplicar o teste é em situações de ajustamento, ou seja, quando se quer averiguar se uma amostra provém de uma certa população com distribuição especificada. Nesse caso pretende-se comparar as frequências observadas (numa tabela de frequências ou num conjunto de classes) com as frequências esperadas de acordo com a dita distribuição.

Para determinar a estatística do teste de qui-quadrado começa-se por determinar quais as frequências que seriam esperadas caso as variáveis fossem completamente independentes. Estas frequências obtêm-se multiplicando as margens da tabela e dividindo pelo número total de observações (não esquecer que quando dois acontecimentos são independentes a sua probabilidade conjunta é o produto das probabilidades individuais). Depois calculam-se as diferenças entre as frequências observadas e as esperadas e somam-se através da expressão

$$X^2 = \sum_{\text{todas as células}} \frac{(O_i - E_i)^2}{E_i},$$

onde E_i representa a frequência esperada e O_i a observada.

Quando o número de observações é elevado a distribuição da estatística X^2 é aproximadamente a do χ^2 e daí o nome do teste.

Rejeita-se a hipótese de independência entre as variáveis quando o valor da estatística de teste é superior a um certo valor crítico (reflectindo grandes desvios entre as frequências observadas e esperadas).

No exemplo dado acima, o valor da estatística X^2 é de 9.091 e o *p-value* (assintótico) do teste de independência é 0.003, evidenciando a existência de associação entre o sexo e a presença da patologia aos níveis usuais de significância (1, 5 e 10%).

Resumindo:

O teste do χ^2 aplica-se sempre que quisermos averiguar a existência de dependência entre duas variáveis de tipo categórico.

Requisitos do teste: As frequências esperadas em cada classe não devem ser inferiores a 5 unidades sempre que o número total de observações é $n < 20$. Se $n > 20$ não deverá existir mais do que 20% das células com frequências esperadas inferiores a 5 nem deverá existir nenhuma célula com frequência esperada inferior a 1.

O SPSS fornece uma opção (*expected*) no botão *cells* do menu da figura 1.5 que produz uma tabela com os valores esperados permitindo ao utilizador verificar os pressupostos acima referidos.

Inconvenientes do teste:

1. Uma vez que a distribuição da estatística de teste é apenas aproximada (assintótica), para amostras pequenas o valor do *p-value* poderá conter um erro apreciável. No caso de tabelas 2×2 e sempre que $n \leq 20$ deve-se recorrer ao teste de Fisher que fornece valores exactos para os *p-values* do teste.
2. Devido à natureza discreta da contagem das frequências o valor da estatística do χ^2 vem acrescida de um erro. No caso de tabelas 2×2 deve-se utilizar uma correcção à continuidade (fornecida pelo SPSS).

Inconvenientes da estatística do χ^2 enquanto medida de associação

A estatística X^2 utilizada no teste do χ^2 é uma medida de associação entre duas variáveis já que assume valores próximos de zero quando as variáveis são independentes e valores elevados (positivos) quando existe dependência. No entanto, ao contrário do que acontecia com os coeficientes de assimetria, esta medida não está limitada ao intervalo $[0, 1]$ e o seu valor máximo depende do número total de observações. Surgem então algumas propostas de coeficientes de associação para dados categóricos que passamos a expor de seguida.

1.2.2 O coeficiente de Cramér

O coeficiente de Cramér é uma medida de associação entre duas variáveis medidas numa escala categórica. Portanto pode ser aplicado em situações onde a informação se encontra distribuída

por categorias nominais não ordenáveis.

Este coeficiente obtém-se directamente a partir da estatística X^2 descrita na secção anterior através da expressão

$$C = \sqrt{\frac{X^2}{n(l-1)}}$$

onde n representa o número total de observações e l representa o mínimo entre o número de linhas e colunas da tabela de contingência.

Vantagens do coeficiente de Cramér:

- o seu valor está limitado ao intervalo $[0, 1]$.
- quando as variáveis são totalmente independentes $C = 0$.
- quanto maior a associação maior o valor do coeficiente.
- o coeficiente pode ser determinado em situações onde mais nenhum coeficiente (dos já expostos) pode ser aplicado.
- ao contrário da estatística X^2 , o coeficiente pode ser aplicado para comparar tabelas de contingência de dimensão diferente ou baseadas em amostras de dimensão diferente.

Desvantagens do coeficiente:

- quando $C = 1$ pode não haver associação perfeita entre as duas variáveis. A associação só é perfeita se o número de linhas for igual ao número de colunas.
- o coeficiente de Cramér está sujeito aos mesmos pressupostos do teste do qui-quadrado se pretendermos testar o seu significado.
- este coeficiente não deve ser comparado directamente com outros. Se os dados forem ordinais podemos calcular o coeficiente de Cramér mas não devemos comparar directamente o seu valor com o valor do coeficiente de Pearson. Embora o coeficiente aumente com o grau de associação as diferenças na magnitude não têm uma interpretação directa.

A partir do valor do coeficiente de Cramér é possível efectuar um teste às hipóteses

H_0 : as variáveis são independentes *vs* H_1 : as variáveis são dependentes.

1.2.3 O coeficiente Φ

O coeficiente Φ é muito semelhante ao coeficiente de Cramér e foi proposto inicialmente apenas para tabelas de contingência 2×2 . Neste caso o teste de independência que se pode efectuar pode ser baseado no teste exacto de Fisher fornecendo valores mais exactos que os do coeficiente

de Cramér. Para tabelas 2×2 com conteúdo representado pelas letras $\begin{matrix} A & B \\ C & D \end{matrix}$ o coeficiente é dado por

$$R_\phi = \frac{|AD - BC|}{\sqrt{(A + B)(C + D)(A + C)(B + D)}}$$

Se retirarmos o módulo do numerador obtemos um coeficiente que pode assumir valores negativos detectando assim um sentido na associação entre as duas variáveis.

No que respeita a vantagens e desvantagens do coeficiente, elas são idênticas às do coeficiente de Cramér.

1.3 Medidas de associação no SPSS

Se as variáveis estiverem representadas numa escala pelo menos ordinal, os coeficientes de correlação (e respectivos testes de associação) encontram-se no menu **Analyze / Correlate / Bivariate** (ver figura 1.4).

Neste menu podem-se seleccionar mais do que duas variáveis, caso em que o SPSS fornece uma tabela de correlações para todas as combinações de pares de variáveis. O SPSS fornece também o p-value do teste à hipótese nula de correlação igual a zero, para cada par de variáveis.

As medidas de associação e respectivos testes de hipóteses para dados organizados em tabelas de contingência estão disponíveis no SPSS através do menu **Analyze / Descriptive Statistics / Crosstabs** (ver figura 1.5). Neste caso há que seleccionar as medidas de interesse no botão **Statistics**. Conforme poderá ver este menu também disponibiliza alguns coeficientes de correlação para dados ordinais, incluindo o coeficiente τ de Kendall descrito anteriormente.

Primeiramente há que introduzir os dados da tabela de contingência e caso uma das colunas contenha as frequências de cada célula é necessário seleccionar o menu **Data / Weight cases** por forma a atribuir os pesos correspondentes.

Por exemplo, a tabela dada anteriormente relacionando os variáveis *sexo* e *patologia* pode ser introduzida no SPSS da seguinte forma: primeiro criam-se duas variáveis numéricas com os nomes *sexo* e *patologia*. Para cada uma delas atribuem-se etiquetas (*labels*) tais como 0 - feminino e 1 - masculino para a variável *sexo*, e 0 - ausente e 1 - presente para a variável *patologia*. Finalmente cria-se uma variável *freq* com as respectivas frequências.

	sexo	patologia	freq
1	0	0	20
2	0	1	30
3	1	0	35
4	1	1	15
5			
6			
7			

1.4 Medidas de concordância entre conjuntos de classificações

Uma questão de bastante interesse em diversas áreas de aplicação da Estatística é a de comparar um conjunto de classificações atribuídas por diferentes avaliadores (ou observadores) para o mesmo conjunto de objectos, indivíduos ou questões. Pretende-se assim saber se os métodos de classificação utilizados pelos dois (ou mais) avaliadores são consistentes e consequentemente fiáveis ou fidedignos (em inglês este procedimento é usualmente designado por *interjudge reliability* ou *interobserver reliability*).

Para atingir este objectivo recorre-se usualmente a medidas de concordância entre os conjuntos de classificações as quais incluem algumas das medidas de associação referidas nas secções anteriores.

Quando as classificações são medidas numa escala de intervalos ou razões e apenas temos dois avaliadores o **coeficiente de Pearson** pode ser aplicado e interpretado como uma medida de concordância entre as classificações. Quando as classificações apenas são registadas numa escala ordinal e continuamos com apenas dois avaliadores então os coeficientes de Spearman ou τ de Kendall fornecem uma medida de concordância. Em especial é muito utilizado o **coeficiente τ de Kendall** para esta situação.

Seguidamente iremos referir outras medidas de concordância vulgarmente aplicadas.

1.4.1 Medida de fiabilidade simples

Numa abordagem mais simples, a concordância entre dois conjuntos de classificações (rankings) sobre os mesmos objectos pode ser obtida dividindo o número de classificações concordantes (neste caso cada objecto terá exactamente a mesma classificação nos dois avaliadores) sobre o número total de objectos classificados. Por exemplo, se as classificações forem as da tabela seguinte esta medida assume o valor $\frac{2}{5} = 0,4$, ou seja 20%.

objecto	avaliador 1	avaliador 2
1	2	3
2	4	4
3	3	5
4	1	1
5	5	2

1.4.2 Dados categóricos e a estatística K de Cohen

Por vezes a classificação pedida aos avaliadores não é a de atribuição de ranks (ordenar os objectos) mas sim a de atribuir uma categoria a cada objecto. Por exemplo pode-se pedir a um conjunto de médicos que cada qual atribua um de k tratamentos a cada um de n doentes. Neste caso as classificações registam-se numa escala nominal e podemos utilizar a estatística K de Cohen para medir o grau de concordância entre os avaliadores (médicos).

Esta medida tem a vantagem de tomar em linha de conta o grau de concordância que se pode esperar observar mesmo quando os avaliadores atribuem as categorias ao acaso. Assim a estatística tem a seguinte expressão:

$$K = \frac{\text{número concordâncias observadas} - \text{número de concordâncias esperadas por mero acaso}}{\text{número total de observações} - \text{número de concordâncias esperadas por mero acaso}}$$

No SPSS o coeficiente de concordância K de Cohen encontra-se disponível no menu **Analyze / Descriptive Statistics / Crosstabs**. No entanto ele só é calculado em tabelas 2×2 , ou seja para comparar dois avaliadores, e que apenas atribuam uma classificação dicotómica aos objectos em causa.

1.4.3 O coeficiente de concordância τ de Kendall para dois avaliadores

Na presença de classificações atribuídas por dois avaliadores e medidas numa escala ordinal o coeficiente mais utilizado é o τ de Kendall descrito anteriormente.

1.4.4 O coeficiente de concordância de Kendall, W , para vários avaliadores

Kendall também propôs um coeficiente para mais do que dois avaliadores que aparece na literatura pela designação *Kendall coefficient of concordance* W . Com base nesse coeficiente é possível efectuar um teste às hipóteses

$$H_0 : \text{as classificações são independentes} \quad \text{vs} \quad \text{as classificações estão associadas}$$

No SPSS o coeficiente W de Cohen está disponível no menu **Analyze / Nonparametric Testes / K related samples**.

Bibliografia

- Siegel, S. & Castellan, N.J., *Nonparametric statistics for the behavioral sciences*, McGraw Hill, 1988.
- Schweigert, W., *Research methods and statistics for psychology*, Brooks/Cole Publishing Company, Pacific Grove California, 1994.
- Kinncar, P. & Gray, C., *SPSS for Windows made simple*, Psychology Press Ltd, 2000.

Notas:

Com excepção do coeficiente de correlação de Pearson e da medida simples de fiabilidade entre avaliadores, pode-se encontrar uma descrição pormenorizada de todos os coeficientes abordados neste texto no livro de Siegel e Castellan.

O coeficiente de Pearson é tratado em praticamente todos os livros introdutórios de Estatística.