



Bioestatística quantitativa aplicada

Edison Capp
Otto Henrique Nienov
Organizadores

Caroline Darski
Charles Francisco Ferreira
Cristiana Palma Kuhl
Fernanda Dapper Machado
Fernanda Vargas Ferreira
Hellen Meiry Grosskopf Werka
Johanna Ovalle Diaz
Marina Petter Rodrigues
Michele Strelow Moreira
Nadine de Souza Ziegler
Paula Barros Terraciano
Pedro Henrique Comerlato
Sinara Santos

Universidade Federal do Rio Grande do Sul
Faculdade de Medicina
Programa de Pós-Graduação em Ciências da Saúde:
Ginecologia e Obstetrícia

Bioestatística Quantitativa Aplicada

Porto Alegre 2020
UFRGS

U58b Universidade Federal do Rio Grande do Sul. Faculdade de Medicina. Programa de Pós-Graduação em Ciências da Saúde: Ginecologia e Obstetrícia Bioestatística quantitativa aplicada/ Universidade Federal do Rio Grande do Sul; organizadores: Edison Capp e Otto Henrique Nienov – Porto Alegre: UFRGS, 2020.

260p.

ISBN: 978-65-86232-43-1

E-Book: 978-65-86232-44-8

1. Epidemiologia e Bioestatística 2. Estatística 3. SPSS I. Capp, Edison, org. II. Nienov, Otto Henrique, org. III Título.

NLM: WA950

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
(Bibliotecária Shirlei Galarça Salort – CRB10/1929)

Endereço:

PPG em Ciências da Saúde: Ginecologia e Obstetrícia

FAMED – UFRGS

Rua Ramiro Barcellos, 2400/2º andar

CEP 900035-003 – Porto Alegre – RS

Telefone: +55 51 3308 5607

E-mail: ppggo@ufrgs.br

Editoração e diagramação: Edison Capp

Capa: Edison Capp, imagens: www.freepik.com/starline

Edison Capp
Otto Henrique Nienov
Organizadores

Caroline Darski
Charles Francisco Ferreira
Cristiana Palma Kuhl
Fernanda Dapper Machado
Fernanda Vargas Ferreira
Hellen Meiry Grosskopf Werka
Johanna Ovalle Diaz
Marina Petter Rodrigues
Michele Strelow Moreira
Nadine de Souza Ziegler
Paula Barros Terraciano
Pedro Henrique Comerlato
Sinara Santos

9 Correlações, risco, razão de chances e avaliação de testes diagnósticos

*Johanna Ovalle Diaz
Hellen Meiry Grosskopf Werka
Edison Capp
Otto Henrique Nienov*

Pensemos na seguinte situação hipotética: um professor de uma disciplina de bioestatística percebe que alguns dos seus alunos não estão apresentando um bom desempenho na realização das tarefas e, notando que estes não estavam dedicando tempo suficiente para o estudo do conteúdo, decide fazer um experimento com a turma. O professor solicita aos alunos que relatem o tempo que cada um dedicou em casa ao estudo do conteúdo e, em seguida, monta uma tabela com os dados de tempo de estudo (expresso em horas) e o desempenho obtido nas tarefas. A seguir, explica aos alunos que, se existe um relacionamento entre o tempo dedicado aos estudos e o desempenho nas tarefas, isso poderia facilmente ser observado em um gráfico (Figura 1).

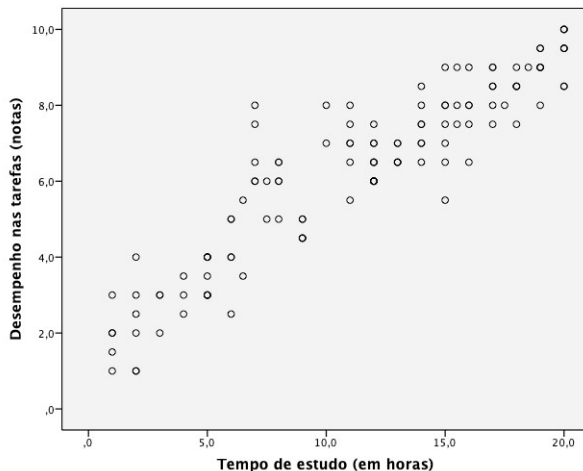


Figura 1. Relação entre o desempenho nas tarefas e o tempo de estudo dos alunos.

Utilizando o eixo X para o tempo de estudo (em horas) e o eixo Y para o desempenho nas tarefas (notas), o professor marca, no gráfico, a nota e o tempo de estudo de cada aluno. Visualmente parece haver alguma relação do tempo de estudo e o desempenho nas tarefas, pois quanto maior o tempo de estudo, melhor tende a ser o desempenho do aluno. Mas, como confirmar e quantificar essa relação?

Correlações bivariadas

Muitas vezes, o objetivo do estudo não é apenas observar se dois grupos possuem a mesma média, mas também se variam conjuntamente, se há alguma relação. Nestes casos, a estatística de interesse é o coeficiente de correlação. A correlação mede como as ordens de variáveis ou de ranqueamento estão relacionadas, ou seja, representa um relacionamento entre duas variáveis. Pode ser analisada por gráficos de dispersão e por coeficientes de correlação. No entanto, antes de calcularmos o coeficiente de correlação, é necessário verificar se os dados possuem valores discrepantes, que podem resultar em conclusões equivocadas.

Os gráficos de dispersão são outro método para avaliar a correlação entre duas variáveis de forma visual. Organiza-se um gráfico cartesiano de pontos e cada ponto corresponde a um sujeito do estudo, cujo posicionamento no gráfico representa os valores para as variáveis localizadas no eixo X e Y, respectivamente. O comportamento dos pontos como conjunto permite interpretar se existe ou não correlação entre as variáveis.

A principal vantagem da correlação é que conseguimos demonstrar o comportamento do relacionamento: se positivo ou negativo (inverso). Além disso, podemos avaliar qualitativamente quanto a sua intensidade, conforme apresentado no quadro 1. O valor do intervalo de coeficientes de correlação varia de -1 (relacionamento negativo ou inverso perfeito) e +1 (relacionamento positivo ou direto perfeito). Um valor de zero ("0") indica nenhum relacionamento. No gráfico, quando existe uma correlação perfeita (-1 ou +1), os pontos formam uma linha reta inclinada. Quando positiva (ou direta), X e Y variam no mesmo sentido, ou seja, os

dois aumentam. Quando negativa (ou inversa), X e Y variam em sentidos opostos, ou seja, quando X aumenta, Y diminui.

Quadro 1. Avaliação qualitativa do grau de correlação entre duas variáveis.

Coeficiente de correlação	Intensidade da correlação
0,000	Nula ou inexistente
0,001 a 0,299	Fraca
0,300 a 0,599	Regular ou moderada
0,600 a 0,899	Forte
0,900 a 0,999	Muito forte
1,000	Plena ou perfeita

O que a correlação procura entender é como uma variável se comporta em um cenário onde outra está variando, visando identificar se existe alguma relação entre a variabilidade de ambas. Portanto, devemos ser cuidadosos ao interpretar coeficientes de correlação porque eles não dão indicação da direção da causalidade. As opções de testes estatísticos variam de acordo com a natureza das variáveis de estudo (Quadro 2).

Quadro 2. Tipos de correlação baseado na forma e comportamento das variáveis.

Correlação	Coeficiente	Tipo de variável
Pearson	r	Quantitativas normais
Spearman	ρ	Quantitativas não normais e qualitativas ordinais
Kendall	τ	Quantitativas não normais e qualitativas ordinais

Coeficiente de correlação de Pearson

O coeficiente de correlação de Pearson, também chamado de correlação linear ou r de Pearson, denominado pela letra r , descrito em 1896 por Karl Pearson, é uma medida de associação linear (simétrica), ou seja, permite avaliar a força de associação entre duas variáveis lineares. Isso quer dizer que as variáveis quantitativas devem estar normalmente distribuídas.

Duas variáveis podem estar perfeitamente relacionadas, mas se o relacionamento não for linear, o coeficiente de Pearson não é o mais apropriado para medir a associação entre as variáveis de estudo.

Coeficiente de correlação de Spearman

O coeficiente de correlação de Spearman, denominado pela letra grega ρ (ρ), descrito em 1904 por Charles Spearman, é uma estatística não paramétrica, que permite avaliar variáveis quantitativas se não estiverem normalmente distribuídas ou se possuírem categorias ordenadas, isto é, variáveis qualitativas ordinais. Uma vantagem do ρ de Spearman é que este não é afetado pelos valores atípicos ou extremos.

Coeficiente de correlação de Kendall

O coeficiente de correlação tau-b de Kendall, denominado pela letra grega tau (τ), descrito em 1938 por Maurice Kendall, é outra correlação não paramétrica, que permite associação para variáveis ordinais. Uma vantagem sobre o coeficiente de Spearman é a possibilidade de ser generalizado para um coeficiente de correlação parcial. Deve ser usada ao invés do coeficiente de Spearman quando temos um conjunto pequeno de dados com um grande número de postos empatados. Isso significa que, se você ordenar todos os escores e muitos deles apresentarem o mesmo posto, o tau de Kendall deve ser usado.

Quando uma das duas variáveis é do tipo nominal, não se pode fazer uma estatística de correlação. O que se costuma fazer nestas situações é aplicar um teste de Qui-quadrado de associação. No entanto, é possível calcular o coeficiente de correlação entre

duas variáveis, uma numérica e outra nominal, quando esta for dicotômica. O primeiro passo é codificar a variável nominal. Então, para se calcular a correlação, procede-se com o cálculo do coeficiente de correlação de Pearson da forma usual, que é chamado neste caso de coeficiente de correlação ponto-biserial.

Para exemplificar, vamos utilizar a situação hipotética mencionada no início deste capítulo, na qual um professor deseja verificar se existe uma relação entre o tempo de estudo (em horas) e desempenho nas tarefas (notas) da turma. Para isso, vamos utilizar o “Banco de dados 5.sav” (disponível em [https:// bit.ly/ bancosdedados](https://bit.ly/bancosdedados)). Lembre-se: é importante conhecer o banco de dados. Tome algum tempo para examiná-lo e conhecer as variáveis.

Primeiro, é necessário verificar a normalidade das variáveis “Tempo de estudo” e “Desempenho nas tarefas” através do teste de Shapiro-Wilk, visto no Capítulo 7. Observa-se que as variáveis apresentam distribuição não normal, logo, realizaremos uma correlação de Spearman. O procedimento para calcular os coeficientes de correlação de Pearson, r de Spearman e tau-b de Kendall são os mesmos, assim como a interpretação é feita da mesma forma.

No menu “Analisar”, “Correlacionar”, clique em “Bivariável...”. Na janela “Correlações de bivariável”, localize na tela da esquerda as variáveis “Tempo de estudo” e “Desempenho nas tarefas”, selecione-as e, com o auxílio do botão da seta, movas para a lista de “Variáveis”. O padrão do SPSS é a correlação de Pearson. Em “Coeficientes de correlação”, marque a opção “Spearman”. Em “Teste de Significância”, é possível selecionar se o teste é ou não de uma ou duas caudas (visto no Capítulo 5). Se a direção da associação for conhecida, ou seja, quando você tiver uma hipótese direcional, selecione “Unilateral”. Caso contrário, se você não souber a natureza do relacionamento, selecione “Bilateral”. Por fim, clique em “OK” ou “Colar”.

No arquivo de saída, é apresentado uma matriz dos coeficientes de correlação (*Correlations*) para as duas variáveis, contendo o valor do coeficiente de correlação r de Spearman, o p-valor e o tamanho da amostra (N). Para cada par de variáveis é apresentado um coeficiente de correlação. Cada variável é perfeitamente correlacionada com ela mesma, tendo um coeficiente

igual a 1,000. O coeficiente da correlação bivariada é encontrado na intersecção das duas variáveis de estudo. Os coeficientes de correlação significativos no nível de significância de 5% são identificados com um único asterisco e, os significativos no nível de significância de 1%, são identificados com dois asteriscos.

Desta forma, temos uma correlação positiva muito forte entre o tempo de estudo e o desempenho nas tarefas, com um coeficiente de $\rho = 0,915$ e $p\text{-valor} < 0,001$. O $p\text{-valor}$ significativo da correlação de Spearman indica a rejeição da hipótese de nulidade de que o coeficiente de correlação seja igual a zero, evidenciando uma relação significativa entre as variáveis testadas. Logo, o professor conclui que, quanto maior o tempo de estudo do aluno, melhor tende a ser o seu desempenho nas tarefas da disciplina.

Podemos apresentar o resultado da correlação através de um gráfico de dispersão. No menu "Gráficos", "Caixas de diálogo legadas", clique em "Dispersão/Ponto". Na janela "Dispersão/Ponto", marque a opção "Dispersão simples" e clique em "Definir". Na janela seguinte ("Diagrama de dispersão simples"), localize na tela da esquerda a variável "Desempenho nas tarefas" e a mova para o "Eixo Y" e, insira no "Eixo X", a variável "Tempo de estudo". Por fim, clique em "OK" ou em "Colar".

Novamente, no arquivo de saída, temos o gráfico de dispersão (Figura 1), onde percebemos que há uma relação significativa entre as variáveis testadas. Os gráficos de dispersão são úteis para visualizarmos a força de um relacionamento entre duas variáveis. O eixo X representa a variável independente e o eixo Y representa a variável dependente. Além disso, o gráfico nos permite verificar como as variáveis estão distribuídas e onde estão os valores atípicos.

Um gráfico de dispersão pode utilizar análise de regressão para estimar a força e direção do relacionamento entre as variáveis dependente e independente. Ao clicarmos duas vezes sobre o gráfico, abre-se uma janela de "Editor de gráfico", onde é possível editá-lo. Nesta janela, é possível "Adicionar uma linha de ajuste no total", clicando com o botão direito sobre o gráfico, que permite analisar o comportamento da dispersão dos dados. Também se obtém o coeficiente de determinação (R^2), que corresponde

ao quadrado do coeficiente de correlação e informa quanto da variação da variável dependente é explicada pela variável independente. No exemplo, a variabilidade do desempenho nas tarefas é explicada em 84,4% pelo tempo de estudo dos alunos. Além de tabelas e gráficos, podemos apresentar os resultados da correlação entre duas variáveis através de setas de oposto, equivalente e contrapeso (Figura 2).

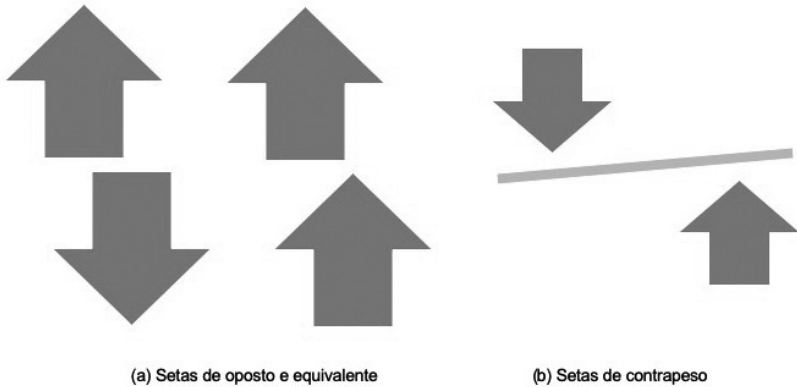


Figura 2. Apresentação de correlação entre duas variáveis.

Risco relativo e razão de chance

Os testes do Qui-quadrado indicam se duas variáveis dicotômicas estão significativamente relacionadas entre si. No entanto, não fornecem uma estatística para descrever a força do relacionamento. Para tabelas 2x2, também é possível calcular uma medida da intensidade da associação entre a presença de um fator e a ocorrência de um evento, ou seja, quanto da probabilidade de ocorrência da variável dependente (doença ou sem doença) deve-se à sua relação com a variável independente (exposição ou não exposição).

O risco relativo (RR) e a razão de chance (*Odds Ratio*, OR) representam as medidas desta associação. Ambas as estatísticas são estimativas de risco e, como tal, descrevem a probabilidade de que os indivíduos expostos a um determinado fator tenham uma doença em comparação com os indivíduos que não estão

expostos ao mesmo fator. A escolha do uso de RR ou de OR depende do desenho do estudo.

Quando o numerador corresponde a eventos que ocorrem de maneira prospectiva, como em estudos experimentais e de coorte prospectiva, a medida de associação denomina-se RR. O RR corresponde a medida da força da associação entre a exposição e o evento (desfecho) e mede a incidência ou risco da doença (variável dependente) acontecer entre os expostos e os não expostos (à variável independente) num período de tempo.

Outras análises podem ser realizadas em estudos que avaliam o efeito de um tratamento. A Redução do Risco Absoluto (RRA) representa a diferença do risco entre os expostos e não expostos e equivale a quantos eventos a menos se produzem ao expor ou não os sujeitos. A Redução do Risco Relativo (RRR) é a diferença do risco entre os expostos e não expostos, mas com denominador constituído apenas pelos não expostos. É possível calcular também o Número Necessário a Tratar (NNT), que equivale ao número de indivíduos que se deve tratar para evitar o evento indesejado.

Quando o numerador corresponde a eventos que ocorreram, como em estudos de caso-controle e de coorte retrospectiva, a medida de associação denomina-se OR e é definido como a probabilidade de que um evento ocorra dividido pela probabilidade de que ele não ocorra. O OR pode ser utilizado como uma estimativa.

Como mencionado, para a análise dos RR e OR é necessário realizar a análise da distribuição de frequência em tabelas de contingência 2x2. Portanto, é necessário que as variáveis sejam qualitativas dicotômicas. A tabela 1 indica a estrutura de organização da tabela de contingência 2x2 para as análises de RR e OR. Como no SPSS a tabela de contingência é organizada de acordo com a ordem numérica dos códigos atribuídos para as categorias de resposta, é necessário verificar a apresentação da tabela antes de realizar a análise. Caso necessário, recodifica-se as variáveis, como visto no Capítulo 4, para obtermos a tabela de contingência corretamente.

Tabela 1. Tabela de contingência 2x2 para as análises de RR e OR.

	Doentes	Não-doentes	Total
Expostos	A	B	A+B
Não-expostos	C	D	C+D
Total	A+C	B+D	A+B+C+D

Onde:

$$RR = A*(A+B)/C*(C+D)$$

$$RRA = [A/(A+B)] - [C/(C+D)]$$

$$RRR = (1 - RR)*100$$

$$NNT = 1/RRA$$

$$OR = (A/B)/(C/D) = (A*D)/(B*C)$$

Como os resultados são decorrentes de uma razão, o valor de nulidade para esses é de 1, ou seja, se o intervalo de confiança incluir um valor inteiro igual a 1, não será possível supor que o fator está associado ao evento. Logo, significa ausência de associação e o risco de acontecer a doença é idêntico, independe da presença ou não da exposição. Se o resultado é > 1 , há uma associação possivelmente causal, ou seja, o risco aumenta com a presença da exposição. Se < 1 , há uma associação possivelmente de proteção, ou seja, o risco diminui na presença da exposição. A interpretação é a mesma, no entanto, no RR avaliamos o risco, enquanto que no OR avaliamos a probabilidade ou chance de o evento ocorrer.

Para exemplificar, vamos utilizar o “Banco de dados 6.sav” (disponível em [https:// bit.ly/bancosdedados](https://bit.ly/bancosdedados)), que contém variáveis qualitativas dicotômicas como tabagismo, câncer de pulmão, obesidade, hipertensão arterial e diabetes mellitus; variáveis quantitativas como massa corporal, IMC, glicemia de jejum, HDL-Colesterol; além de dois escores de neuropatia periférica, calculados através dos instrumentos MNSI e NDS (*Neuropathy Disability Score*). Lembre-se: é importante conhecer o banco de dados. Tome algum tempo para examiná-lo e conhecer as variáveis.

Vamos supor que um estudo de coorte acompanhou 100 indivíduos com hábito de fumar e 100 indivíduos sem o hábito

de fumar para observar o desenvolvimento de câncer no pulmão. Pensando na estrutura de organização da tabela de contingência 2x2, a doença é o câncer no pulmão, a exposição é o hábito de fumar, o total de indivíduos de cada grupo é 100 e a amostra total do estudo é 200. Ao final do estudo, observou-se que 33 tabagistas e 7 não-tabagistas desenvolveram câncer no pulmão (Tabela 2). Portanto, iremos verificar o risco de ocorrência de câncer no pulmão entre tabagistas e não tabagistas.

Tabela 2. Tabela de contingência 2x2 para avaliação do risco de ocorrência de câncer no pulmão em indivíduos com hábito de fumar.

	Com câncer de pulmão	Sem câncer de pulmão	Total
Tabagista	33	67	100
Não-tabagista	7	93	100
Total	40	160	200

Primeiro, devemos lembrar que no SPSS a tabela de contingência é organizada de acordo com a ordem numérica dos códigos atribuídos pelo pesquisador. Por isso, é necessário verificar se a combinação de códigos obedece a estrutura de organização da tabela de contingência 2x2. Para isso, no menu "Analisar", "Estatísticas descritivas", clique em "Tabela de referência cruzada...". Na janela de "Tabulações cruzadas", localize na tela da esquerda a variável "Tabagismo" (exposição) e a insira na "Linha(s)". Em seguida, na mesma tela, selecione a variável "Câncer de pulmão" (desfecho) e a mova para "Coluna(s)". Clique em "Ok" ou "Colar".

No arquivo de saída, é possível perceber que a tabela de contingência 2x2 obtida não corresponde ao formato correto. Portanto, para analisar o RR e o OR precisamos recodificar as variáveis tabagismo ("0" = "Tabagista" e "1" = "Não tabagista") e câncer no pulmão ("0" = "Com câncer" e "1" = "Sem câncer"), para obtermos a estrutura de organização correta da tabela de contingência 2x2.

Como visto no Capítulo 4, no menu "Transformar", clique em "Recodificar em variáveis diferentes...". Na caixa de diálogo "Recodificar em variáveis diferentes", localize na tela da esquerda a

variável "Tabagismo" e mova, com o auxílio do botão da seta, para a lista de "Variável numérica -> Variável de saída". Em seguida, informe um novo nome e rótulo para a variável em "Nome" e "Rótulo", respectivamente. Clique em "Alterar". Clique em "Valores antigo e novo". Na janela seguinte, em "Valor", informe o valor antigo "0" e insira o novo valor "1". Clique em "Adicionar". Em seguida, informe o valor antigo "1" e insira o novo valor "0". Clique em "Adicionar". Por fim, clique em "Continuar" e, em "OK" ou "Colar". A nova variável aparecerá automaticamente no final da planilha "Visualização de dados". Informe o valor dos novos rótulos na planilha "Visualização da variável". Agora, repita o processo para recodificar a variável "Câncer de pulmão".

Com as variáveis recodificadas, vamos realizar a análise de RR e OR. No menu "Analisar", "Estatísticas descritivas", clique em "Tabela de referência cruzada...". Na janela de "Tabulações cruzadas", localize na tela da esquerda a nova variável "Tabagismo" (exposição) e a insira na "Linha(s)". Em seguida, na mesma tela, selecione a nova variável "Câncer de pulmão" (desfecho) e a mova para "Coluna(s)". Em "Exato", marque a opção de "Monte Carlo". Clique "Continuar". Para realizar o cálculo do RR e do OR no SPSS, clique em "Estatísticas" e marque as opções "Qui Quadrado" e "Risco". Clique "Continuar". Em "Células", marque as opções "Observado" em "Contagens", "Linha" e "Total" em "Porcentagens" e, para "Residuais", marque a opção "Padronizado ajustado". Clique em "Continuar". Por fim, clique em "Ok" ou "Colar".

No arquivo de saída, temos a tabela de contingência 2x2 (*Crosstabulation*), que fornece os dados necessários para o cálculo do RR e OR. O teste de correção de continuidade de Yates indica associação significativa entre tabagismo e câncer de pulmão ($p < 0,001$), evidenciada pelos resíduos. Independentemente do tipo de estudo, o SPSS produz a análise de ambos RR e OR, e compete ao pesquisador selecionar o tipo de resultado que está adequado ao seu estudo.

O SPSS, além do cálculo do RR e OR, também produz seus respectivos intervalos de confiança de 95% (IC95%). Para um p -valor $\leq 0,05$, o IC95% representa o intervalo onde deve estar o parâmetro, ou seja, o RR e OR verdadeiros. Se o valor 1, que se refere à nulidade

da associação, não estiver contido no IC95%, temos uma confiança que na população de onde nossa amostra foi extraída, o achado da amostra é significativo. Desta forma, ser tabagista aumenta o risco de câncer no pulmão em 4,71 vezes (IC95%: 2,19-10,15; $p < 0,001$) e aumenta a chance de câncer no pulmão em 6,54 vezes (IC95%: 2,73-15,68; $p < 0,001$) em comparação a não ser tabagista.

Avaliação de testes diagnósticos

O teste diagnóstico tem como objetivo determinar se um indivíduo tem ou não a doença. Para avaliar a acurácia deste teste, o pesquisador pode fazer uso dos parâmetros de sensibilidade e especificidade. Uma forma eficiente de demonstrar a relação antagônica entre a sensibilidade e a especificidade são as Curvas de Características de Operação do Receptor (Curvas ROC). A Curva ROC é uma ferramenta poderosa para medir e especificar problemas no desempenho dos testes diagnósticos, pois permite avaliar a variação da sensibilidade e especificidade para diferentes valores de corte.

Para isso, é necessário um teste de referência, também chamado de padrão ouro, que é capaz de diferenciar um indivíduo doente de um não doente, pois acredita-se que este tenha a medida "verdadeira". É utilizado como comparação na avaliação da qualidade (performance) de testes índices, que são medidos numa escala numérica (variável quantitativa) e, assim, a curva ROC permite estabelecer um ponto de corte que define os resultados deste teste como sendo positivo e negativo (variável dicotômica).

A reprodutibilidade, repetibilidade ou precisão de um teste diagnóstico consiste na consistência dos resultados quando o exame se repete pelo mesmo examinador ou por outros examinadores. A comparação dos resultados pode ser apresentada através do índice de concordância Kappa, visto no Capítulo 6.

A acurácia ou validade do teste refere-se à probabilidade de um teste determinar o verdadeiro valor do que está sendo medido (ser positivo nos doentes e negativo nos não doentes). Como mencionado, para avaliar a acurácia de um teste diagnóstico o pesquisador pode fazer uso dos índices de sensibilidade e especificidade.

A sensibilidade é a probabilidade de um teste detectar os sujeitos verdadeiramente positivos (doentes). Testes sensíveis são utilizados para reduzir possibilidades diagnósticas, como por exemplo, em triagens sorológicas em bancos de sangue para prevenção de transmissão de infecções nas quais a não detecção de casos acarretará risco para a população. Nestes casos, o ponto de corte deverá ser estabelecido tendo como objetivo alcançar maior sensibilidade do teste para que não ocorram resultados falsos negativos.

A especificidade é a probabilidade de um teste detectar os sujeitos verdadeiramente negativos (não doentes). Testes específicos são utilizados para confirmar a presença de doença, como por exemplo, em indicações cirúrgicas, onde deve-se evitar resultados falsos positivos. Nestes casos, o ponto de corte deve ser definido de tal forma que aumente a especificidade do teste.

Para se realizar as análises de sensibilidade e especificidade é necessário que utilizemos dois métodos ou instrumentos diferentes, mas que tenham o mesmo objetivo, sendo um o padrão ouro. O teste ideal, com 100% de sensibilidade e especificidade, raramente existe na prática. Isso porque, na maioria das vezes, a tentativa de melhorar a sensibilidade diminui a especificidade e vice-versa.

O valor preditivo positivo (VPP) mede a proporção de doentes entre os positivos pelo teste, ou seja, a probabilidade de doença quando o teste é positivo. Quanto mais específico, melhor o VPP (redução de falsos positivos). O valor preditivo negativo (VPN) mede a proporção de saudáveis entre os negativos pelo teste, ou seja, a probabilidade de ausência de doença quando o teste é negativo. Quanto mais sensível, melhor o VPN (redução de falsos negativos).

A razão de verossimilhança positiva (RVP) é a probabilidade de um resultado positivo em alguém com a doença dividida pela probabilidade do mesmo resultado em alguém sem a doença, ou seja, sensibilidade/1-especificidade. A razão de verossimilhança negativa (RVN) é a probabilidade de um resultado negativo em alguém com a doença dividida pela probabilidade do mesmo resultado em alguém sem a doença, ou seja, 1-sensibilidade/especificidade. Desta forma, os resultados dos testes padrão e índice podem ser organizados numa tabela de contingência 2x2 (Tabela 3).

Tabela 3. Tabela de contingência 2x2 para avaliar a relação de teste-doença.

		Padrão-ouro		Total
		Doente	Não doente	
Teste índice	Positivo	A (Verdadeiro positivo)	B (Falso positivo)	A+B
	Negativo	C (Falso negativo)	D (Verdadeiro negativo)	C+D
Total		A+C	B+D	A+B+C+D

Onde:

$$\text{ACURÁCIA} = A+D/A+B+C+D$$

$$\text{SENSIBILIDADE} = A/A+C$$

$$\text{ESPECIFICIDADE} = D/B+D$$

$$\text{VPP} = A/A+B$$

$$\text{VPN} = D/C+D$$

$$\text{RVP} = [A/(A+C)]/[B/(B+D)]$$

$$\text{RVN} = [C/(A+C)]/[D/(B+D)]$$

Curva ROC

Quando o resultado do teste é expresso como uma variável quantitativa, deve-se procurar um ponto de corte ("*cutoff point*"), cujo valor determinará a interpretação do teste como sendo positivo ou negativo. Para isso, é utilizada a curva ROC, que avalia o desempenho dos testes e a variação da sensibilidade e especificidade para diferentes valores de corte num gráfico que apresenta a proporção dos verdadeiros positivos (sensibilidade), no eixo Y, versus a proporção dos falsos positivos (1-especificidade), no eixo X. Desta forma, a Curva ROC é um gráfico de sensibilidade.

Neste tipo de representação gráfica, a área abaixo da curva (AUC) define a probabilidade do teste em detectar os verdadeiros positivos e os verdadeiros negativos, ou seja, expressa a acurácia do teste. A curva ROC é uma forma de representar a relação antagônica entre a sensibilidade e a especificidade. Para construir uma curva ROC traça-se um diagrama que represente a sensibilidade em

função da proporção de falsos positivos (1-especificidade) para um conjunto de valores de corte e , a linha diagonal, corresponde a um teste que é positivo ou negativo, aleatoriamente (Figura 3).

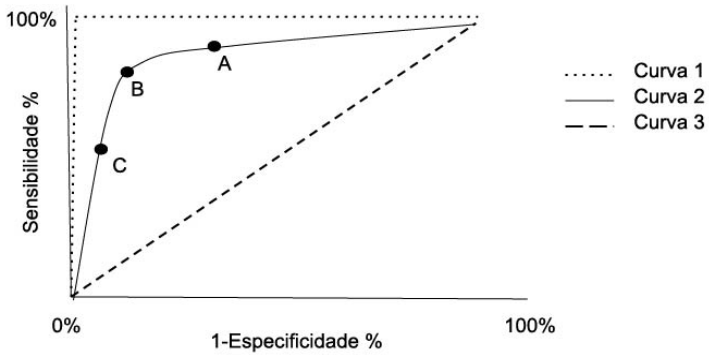


Figura 3. Curva ROC. 1, AUC 100%; 2, AUC 85%; 3, AUC 50%. Pontos de corte A, B e C com sensibilidade e especificidade respectivas.

Na curva 1, é representado o ponto de maior acurácia do teste (100%). No entanto, como mencionado, as características de sensibilidade e especificidade geralmente apresentam um comportamento antagônico. A curva 2 representa uma acurácia de 85%, que significa que em 85% das vezes em que for utilizado, o teste discriminará os verdadeiros positivos e negativos, mas também significa que em 15% das vezes fornecerá um falso resultado. A curva 3, expressa a acurácia de 50% e significa que o teste é tão bom quanto qualquer escolha ao acaso. Em geral, uma acurácia maior ou igual a 70% é considerada satisfatória.

O valor escolhido como ponto de corte influenciará as características do teste. Na curva 2, quanto menor o ponto de corte (ponto A), maior é a sensibilidade, mas menor é a especificidade. Neste caso, se pretendemos um teste muito sensível e menos específico, escolhe-se este ponto para obter uma menor proporção de falsos negativos e uma maior proporção de falsos positivos. O ponto de corte localizado no ponto B permite pôr em evidência os valores para os quais existe maior otimização da sensibilidade em função da especificidade. Já quanto maior o ponto de corte (ponto C), maior especificidade, porém, menor sensibilidade.

Neste caso, se pretendemos um teste muito específico e menos sensível, escolhe-se o ponto C, obtendo-se uma menor proporção de falsos positivos e uma maior proporção de falsos negativos.

O valor do ponto de corte é definido como um valor que pode ser selecionado, arbitrariamente, pelo pesquisador entre os valores possíveis para a variável de decisão. Acima do ponto de corte, o paciente é classificado como positivo (teste positivo, presença de doença) e, abaixo, é classificado como negativo (teste negativo, ausência de doença). Para cada ponto de corte são calculados valores de sensibilidade e especificidade, dispostos no gráfico.

Para exemplificar, no "Banco de dados 6.sav" (disponível em [https:// bit.ly/bancosdedados](https://bit.ly/bancosdedados)), vamos construir a curva ROC do NDS, utilizando como padrão ouro o MNSI, ambos instrumentos de rastreamento para neuropatia periférica. O SPSS não realiza os cálculos de sensibilidade, especificidade, VPP e VPN, porém fornece as informações na tabela de contingência 2x2 que permitem realizar os cálculos.

Para calcular esses valores, a priori, foi definido um ponto de corte de 3,0 para classificar os sujeitos como sendo neuropatas ("0") e não neuropatas ("1") pelo teste índice (NDS) e, para o padrão ouro (MNSI), um ponto de corte de 2,5 é utilizado para classificar os sujeitos da mesma forma. Os valores foram atribuídos desta forma para obtermos a correta estrutura de organização da tabela de contingência 2x2.

No menu "Analisar", "Estatísticas descritivas", clique em "Tabela de referência cruzada...". Insira na "Linha(s)" a variável que classifica os sujeitos pelo NDS e, em "Coluna(s)", a variável que classifica os sujeitos pelo MNSI. Em "Exato...", selecione "Monte Carlo". Clique em "Continuar". Em "Estatísticas...", selecione as opções "Qui-quadrado" e "Capa". Clique em "Continuar". Em "Células...", marque "Observado" em "Contagens", "Total" em "Porcentagens" e, em "Residuais", selecione "Padronizado ajustado". Clique em "Continuar".

No arquivo de saída, os resultados do teste Exato de Fisher indicam haver associação significativa entre os dois testes ($p < 0,001$). Porém, o teste de Kappa indica uma concordância regular entre os métodos (0,330), apesar de significativa ($p < 0,001$). Estes resultados nos indicam uma baixa sensibilidade do método índice (NDS) para rastreamento de neuropatia periférica com o ponto de corte 3,0.

Na tabela de contingência estão disponíveis as informações necessárias para os cálculos da sensibilidade, especificidade, VPP e VPN:

$$\text{ACURÁCIA} = 15+143/200 = 0,79 = 79,0\%$$

$$\text{SENSIBILIDADE} = 15/(15+40) = 0,272 = 27,2\%$$

$$\text{ESPECIFICIDADE} = 143/(143+2) = 0,986 = 98,6\%$$

$$\text{VPP} = 15/(15+2) = 0,882 = 88,2\%$$

$$\text{VPN} = 143/(143+40) = 0,781 = 78,1\%$$

Como esperado, o NDS apresenta uma baixa sensibilidade. Como se trata de um teste de triagem, seria interessante utilizarmos um ponto de corte que apresente uma maior sensibilidade. Lembre-se que, no caso do rastreamento da neuropatia periférica, os valores do NDS produzem o *screening* positivo a partir de pontos de corte. Desta forma, as modificações dos pontos de corte podem produzir desempenhos diferentes quanto aos valores de sensibilidade, especificidade, VPP e VPN.

Uma forma de se analisar o desempenho dos pontos de corte é através da Curva ROC. Para que seja possível realizar esta análise no SPSS, é necessário que uma variável associada ao método de referência (padrão ouro) para o diagnóstico seja qualitativa em escala dicotômica (positivo ou negativo) e, a variável que representa o teste índice que terá os pontos de corte analisados, seja do tipo quantitativa.

No menu "Analisar", clique na opção "Curva ROC". Na janela "Curva ROC", localize na tela da esquerda as variáveis e insira, com o auxílio do botão da seta, a variável quantitativa ("NDS") em "Variável de teste", e mova a variável qualitativa relacionada ao critério de referência ("NEUROMNSI") em "Variável de estado". Na

opção "Valor da variável de estado", insira o valor "0" que se refere ao valor positivo para o teste de rastreamento ("0" = "Neuropata"). Em "Exibir", selecione as opções "Curva ROC" para gerar o gráfico; a opção "Com linha de referência diagonal" para gerar a linha que representa a hipótese de nulidade; "Erro padrão e intervalo de confiança" para gerar o intervalo de confiança associado ao erro padrão e; "Coordenar pontos da curva ROC" para gerar a tabela com valores de sensibilidade e especificidade nos diferentes pontos de corte. Por fim, clique em "Ok" ou "Colar".

No arquivo de saída, no primeiro quadro (*Case Processing Summary*), é indicado o número de casos positivos e negativos conforme a variável de referência. No quadro seguinte (*Area Under the Curve*), temos o valor de AUC (Area), o erro padrão, o p-valor e o IC95%. Lembre-se que a AUC é uma medida de desempenho que serve para determinar a acurácia do teste. Um teste totalmente incapaz de discriminar indivíduos doentes e não doentes, teria uma área sob a curva de 0,500 (50%), correspondendo a H_0 . Acima de 0,700 (70%) é considerado desempenho satisfatório. No exemplo, o valor de AUC de 0,831 indica que os valores do NDS têm uma probabilidade de 83% de detectar os resultados positivos (IC95%: 0,769-0,893; $p < 0,001$).

No terceiro quadro (*Coordinates of the Curve*), é possível verificar a relação dos diferentes pontos de corte de sensibilidade e 1-especificidade. A sensibilidade é a proporção de neuropatas com os resultados do teste maior que o ponto de corte e 1-especificidade é a proporção de não neuropatas com os resultados do teste maior que o de ponto de corte. O ponto de corte de equilíbrio será escolhido de acordo com o objetivo do teste. O gráfico gerado permite complementar a análise para a tomada de decisão (Figura 4).

O menor valor de corte é o valor mínimo de teste observado menos 1 e, o maior valor de corte, é o valor máximo de teste observado mais 1. Todos os outros valores de corte são as médias de dois valores de teste observados ordenados consecutivos. No ponto de corte, "0" é equivalente a assumir que todos são neuropatas. No ponto de corte "6" é equivalente a assumir que todos são não neuropatas. Ambos os extremos são insatisfatórios. É preciso escolher um ponto de corte que devidamente equilibre as necessidades de sensibilidade e especificidade.

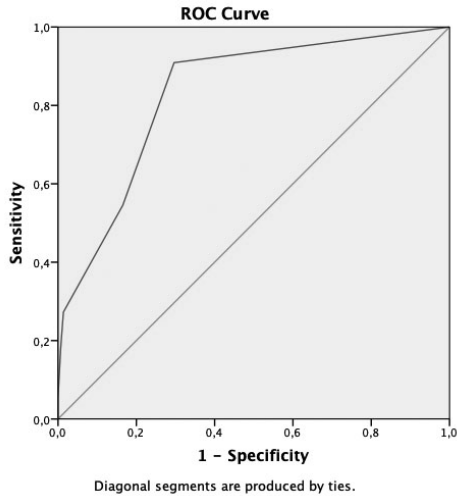


Figura 4. Curva ROC para avaliação da sensibilidade e 1-especificidade do teste índice (NDS) em comparação ao padrão ouro (MNSI).

Se considerarmos o ponto de corte de 2,5, temos o valor de sensibilidade de 0,273 (27,3%) e 1-especificidade de 0,014 (1,4%). Assim, aproximadamente 27,3% dos sujeitos com neuropatia periférica seriam corretamente identificadas e 1,4% dos sujeitos sem neuropatia periférica poderiam ser incorretamente identificadas como neuropatas (98,6% de especificidade). Por outro lado, se considerarmos o ponto de corte de 1,5, teríamos valores de sensibilidade de 0,545 e 1-especificidade de 0,166, ou seja, 54,5% dos neuropatas estariam classificados corretamente e 16,6% dos não neuropatas estariam classificados incorretamente como neuropatas (83,4% de especificidade). Note que os valores deste quadro representam as melhores orientações para os quais devemos considerar os pontos de cortes, mas não inclui as estimativas de erro. Portanto, não há garantia da exatidão da sensibilidade ou especificidade para um dado ponto de corte no quadro.

A escolha do ponto de corte será decidida pela necessidade de aumentar a sensibilidade ou a especificidade: uma maior sensibilidade para detectar os doentes aumentará as chances de falsos positivos, enquanto que uma maior especificidade para

determinar quem não esteja doente, que pode aumentar as chances de falsos negativos. Logo, dependerá do pesquisador a seleção do ponto de corte de equilíbrio entre sensibilidade e especificidade.

Referências

Barton, B, Peat, J. Medical Statistics: A guide to SPSS. Data analysis and Critical Appraisal. 2. ed. John Wiley & Sons, Ltda. 2014. 411 p.

Callegari-Jacques SM. Bioestatística: princípios e aplicações. Porto Alegre: ArtMed, 2011. 255p.

Campbell MJ, Machin D & Walters SJ. Medical Statistics, a Textbook for the Health Sciences. 4. ed. 2007. 344 p.

Chagas, E.F.B. Módulo 3: Estatística Analítica II - Análise de dados categóricos e teste diagnóstico no SPSS. 2016. 44 p.

Field A. Descobrendo a estatística usando o SPSS. Tradução: Lorí Viali. 2. ed. Porto Alegre: Artmed, 2009. 684 p.

Schoenbach VJ & Rosamond W. Understanding the Fundamentals of Epidemiology: An Evolving Text. Chapel Hill, North Carolina. 2000. 584 p.

Exercícios sugeridos

1. Utilizando o "Banco de dados 6.sav":

- a) Há correlação entre massa corporal e glicemia de jejum? Realize o teste estatístico, interprete os resultados e construa o gráfico.
- b) Há correlação entre massa corporal e HDL-Colesterol? Realize o teste estatístico, interprete os resultados e construa o gráfico.

2. Utilizando o "Banco de dados 6.sav":

- a) Há correlação entre IMC e glicemia de jejum? Realize o teste estatístico, interprete os resultados e construa o gráfico.
- b) Há correlação entre IMC e HDL-Colesterol? Realize o teste estatístico, interprete os resultados e construa o gráfico.

3. Utilizando o "Banco de dados 6.sav":

- a) Qual a chance de ser obeso e ter hipertensão? Realize o teste estatístico e interprete os resultados.
- b) Qual a chance de ser obeso e ter diabetes? Realize o teste estatístico e interprete os resultados.