

Bioestatística F

Testes Diagnósticos

Enrico A. Colosimo

Departamento de Estatística
Universidade Federal de Minas Gerais
<http://www.est.ufmg.br/~enricoc>

2011

Testes Diagnósticos

- Uma das experiências mais rotineiras da prática clínica é a solicitação de um teste diagnóstico.
- Os objetivos incluem a triagem de pacientes, o diagnóstico de doenças e o acompanhamento ou prognóstico da evolução do paciente.
- Para chegar ao diagnóstico, existem várias possibilidades, com níveis de *certeza* que variam de acordo com as informações disponíveis.
- Como medir o nível de certeza de presença de uma doença após a observação de um teste positivo?
- Existe algum teste perfeito, que sempre acerta no diagnóstico?

Dados genéricos de um teste clínico

- Na etapa de pesquisa para a determinação do **grau de confiabilidade de um teste diagnóstico**, o pesquisador utiliza-o primeiramente em dois grupos muito específicos de pessoas: *doentes* e *não doentes*.
- Nesta etapa o diagnóstico (determinação do verdadeira *status* do paciente) é feito através de um teste *padrão ouro* (o melhor teste disponível sob certas condições).

Tabela: Esquema padrão de síntese dos dados para verificação da qualidade de um teste clínico.

| Doença | Teste | | Total |
|----------------------|----------------------|----------------------|-----------------|
| | Positivo ($T = +$) | Negativo ($T = -$) | |
| Presente ($D = 1$) | a | b | $n_1 = a + b$ |
| Ausente ($D = 0$) | c | d | $n_2 = c + d$ |
| Total | $a + c$ | $b + d$ | $n = n_1 + n_2$ |

Dados genéricos de um teste clínico

- Os eventos que estamos interessados são:
 - $\{T = +\}$ corresponde a teste positivo
 - $\{T = -\}$ corresponde a teste negativo
 - $\{D = 1\}$ corresponde a indivíduo portador da doença
 - $\{D = 0\}$ corresponde a indivíduo não portador da doença

Sensibilidade e especificidade

Na análise da qualidade de testes diagnósticos, interessa conhecer duas probabilidades condicionais que, pela sua importância, recebem nomes especiais:

- a **Sensibilidade (s)**: é a probabilidade de o teste ser positivo, dado que o indivíduo testado realmente tem a doença

$$s = P(T = + | D = 1)$$

- a **Especificidade (e)**: é a probabilidade de o teste ser negativo, dado que o indivíduo testado realmente não tem a doença

$$e = P(T = - | D = 0)$$

Calculando sensibilidade e especificidade

Recapitulando ...

Tabela: Esquema padrão de síntese dos dados para verificação da qualidade de um teste clínico.

| Doença | Teste | | Total |
|----------------------|----------------------|----------------------|-----------------|
| | Positivo ($T = +$) | Negativo ($T = -$) | |
| Presente ($D = 1$) | a | b | $n_1 = a + b$ |
| Ausente ($D = 0$) | c | d | $n_2 = c + d$ |
| Total | $a + c$ | $b + d$ | $n = n_1 + n_2$ |

Sensibilidade e especificidade, n_1 e n_2 fixos.

$$s = \frac{a}{n_1} = \frac{a}{a + b} ; e = \frac{d}{n_2} = \frac{d}{c + d}$$

Calculando sensibilidade e especificidade

Exemplo: Metástase de carcinoma hepático

Lind & Singer (1986) ¹ estudaram a qualidade da tomografia computadorizada para o diagnóstico de metástase de carcinoma de fígado, obtendo os resultados sintetizados na Tabela a seguir. Um total de 150 pacientes foram submetidos a dois exames: a tomografia computadorizada e a laparotomia. A laparotomia é o *padrão ouro* (classifica o paciente sem erro) para este caso.

¹LIND, S. E., SINGER, D. E. Diagnosing liver metastases: a Bayesian analysis. **Journal of Clinical Oncology**, v.3, p. 379-88, 1986.

Calculando sensibilidade e especificidade

Exemplo: Metástase de carcinoma hepático

Tabela: Resultados da tomografia computadorizada em 67 pacientes com metástase e 83 sem metástase do carcinoma hepático.

| Metástase de carcinoma hepático | Tomografia | | Total |
|------------------------------------|----------------------|----------------------|-------|
| | Positivo ($T = +$) | Negativo ($T = -$) | |
| Presente ($D = 1$) | 52 | 15 | 67 |
| Ausente ($D = 0$) | 9 | 74 | 83 |
| Total | 61 | 89 | 150 |

$$s = \frac{52}{67} = 0,776$$

$$e = \frac{74}{83} = 0,892$$

- A especificidade (89,2%) da tomografia computadorizada é maior que sua sensibilidade (77,6%).

Valor preditivo positivo e negativo

- A sensibilidade e especificidade têm uma limitação séria: não ajudam a decisão da equipe médica que, recebendo um paciente com resultado positivo do teste, precisa avaliar se o paciente está ou não doente.
- Não se pode depender apenas da sensibilidade e da especificidade, pois estes índices são provenientes de uma situação em que há certeza total sobre o diagnóstico, o que não acontece no consultório médico.
- Dois outros índices refletem melhor a realidade prática:
 - **Valor Preditivo Positivo (VPP):** é a probabilidade do paciente estar realmente com a doença quando o resultado do teste é positivo

$$VPP = P(D = 1 | T = +)$$

- **Valor Preditivo Negativo:** é a probabilidade do paciente não estar doente quando o resultado do teste é negativo

$$VPN = P(D = 0 | T = -)$$

Exercício 1: História do Sr. Leonard Mlodnaw

Em 1989 o Sr. Leonard teve o Elisa + para HIV e o médico dele fez a seguinte afirmação:

"Você tem uma probabilidade de 999/1000 de morrer em uma década. Sinto muito."

Como o médico obteve este número?

- O valor de especificidade informado pelo laboratório:

$$e = P(T - |D = 0) = 999/1000.$$

e, portanto

$$P(T + |D = 0) = 1/1000!!$$

Ou seja, se com T+, a probabilidade de não ser HIV+, implica que ser HIV+ é 999/1000 !!!!!

- Ele confundiu com a verdadeira afirmação que deveria ser:

$$VPN = P(D = 1 | T+) = ?$$

Vamos entender a situação

- Caso: homem, heterossexual e não usuário de droga
-

$$\Omega = \{(D = 1, T+), (D = 0, T+), (D = 1, T-), (D = 0, T-)\}$$

- Suponha que p (prevalência) = $1/10000$.
- Para cada 10 000 pessoas do tipo caso espera-se que o teste detecte 11 testes + (um realmente doente e 10 não doentes)
- Conclusão: Para cada 11 Elisa + somente um seria realmente doente. Ou seja,

$$VPP = P(D = 1 | T = +) = 1/11$$

- O médico deveria ter dito:

"Não se preocupe, existe uma probabilidade de 10/11 de que você não está infectado pelo HIV."

Calculando VPP e VPN

- Para calcular o VPP e o VPN recorreremos ao **Teorema de Bayes**.
- Suponha que os eventos C_1 e C_2 formem uma partição de Ω e que suas probabilidades sejam conhecidas. Suponha ainda que para um evento A , se conheçam as probabilidades $P(A|C_i)$ para $i = 1, 2$. Então, para qualquer i :

Teorema de Bayes

$$P(C_i|A) = \frac{P(A|C_i) \times P(C_i)}{P(A|C_1) \times P(C_1) + P(A|C_2) \times P(C_2)}, i = 1, 2.$$

Exercício 2

Teorema de Bayes

Das pacientes de uma clínica de ginecologia, com idade acima de 40 anos, 60% são ou foram casadas e 40% são solteiras. Sendo solteira, a probabilidade de ter tido um distúrbio hormonal no último ano é de 10%, enquanto que para as demais essa probabilidade aumenta para 30%.

- Partição do espaço amostral: “casadas”, “solteiras”
- 1 Qual a probabilidade de uma paciente escolhida ao acaso ter tido um distúrbio hormonal?
- 2 Se a paciente sorteada tiver distúrbio hormonal, qual a probabilidade de ser solteira?

Calculando VPP e VPN

- $VPP = P(D = 1|T = +)$. Pelo Teorema de Bayes, temos:

$$\begin{aligned} P(D = 1|T = +) &= \frac{\overbrace{P(T = +|D = 1)}^{\text{sensibilidade}} \overbrace{P(D = 1)}^{\text{prevalência}}}{\underbrace{P(T = +|D = 1)}_{\text{sensibilidade}} \underbrace{P(D = 1)}_{\text{prevalência}} + \underbrace{P(T = +|D = 0)}_{1-\text{especificidade}} \underbrace{P(D = 0)}_{1-\text{prevalência}}} \\ &= \frac{ps}{ps + (1 - p)(1 - e)}, \end{aligned}$$

em que $p = P(D = 1)$ é a prevalência da doença na população de interesse (proporção de doentes, ou probabilidade de doença pré-teste).

Calculando VPP e VPN

- $VPN = P(D = 0 | T = -)$, e da mesma forma (via Teorema de Bayes), temos:

$$\begin{aligned} P(D = 0 | T = -) &= \frac{\overbrace{P(T = - | D = 0)}^{\text{especificidade}} \overbrace{P(D = 0)}^{1-\text{prevalência}}}{\underbrace{P(T = - | D = 0)}_{\text{especificidade}} \underbrace{P(D = 0)}_{1-\text{prevalência}} + \underbrace{P(T = - | D = 1)}_{1-\text{sensibilidade}} \underbrace{P(D = 1)}_{\text{prevalência}}} \\ &= \frac{(1-p)e}{p(1-s) + (1-p)e} \end{aligned}$$

Calculando VPP e VPN

Exemplo: Metástase de carcinoma hepático (continuação)

Para uma população cuja **prevalência de metástase de carcinoma de fígado é de 2%**, o VPP da tomografia computadorizada é

$$VPP = \frac{0,02 \times 0,78}{0,02 \times 0,78 + (1 - 0,02) \times (1 - 0,89)} = 0,13$$

e o VPN é

$$VPN = \frac{(1 - 0,02) \times 0,89}{(1 - 0,02) \times 0,89 + 0,02 \times (1 - 0,78)} = 0,99.$$

Calculando VPP e VPN

Exemplo: Metástase de carcinoma hepático (continuação)

Observações:

- O VPP é baixo enquanto o VPN é bastante alto.
- Se o resultado da tomografia computadorizada é negativo, a probabilidade de não haver metástase é de 99%.
- Os valores de VPP e VPN são diretamente ligados ao de prevalência.

- **Probabilidade de Falso-Positivo (PFP):**
 $PFP = P(D = 0 | T = +) = 1 - P(D = 1 | T = +) = 1 - VPP$
- **Probabilidade de Falso-Negativo (PFN):**
 $PFN = P(D = 1 | T = -) = 1 - P(D = 0 | T = -) = 1 - VPN$
- **Atenção:** não há na literatura padronização relativa a nomes dos índices de um teste diagnóstico. PFP e PFN são frequentemente usadas para as quantidades $1 - s$ e $1 - e$.

Decisões incorretas

Exemplo: Teste ELISA para detecção do HIV

- Entre as várias técnicas para detectar a presença do HIV, a primeira a se difundir no Brasil foi o teste de ELISA (*Enzymelinked immunosorbent assay*).
- Em 1985 esta técnica foi comercializada por vários laboratórios americanos.
- O Laboratório ABBOTT relatou, em seus testes preliminares, sensibilidade de 95% e especificidade de 99,8%.
- Com estes valores (sensibilidade e especificidade) mais a prevalência podemos calcular VPP, VPN e conseqüentemente PFP e PFN.

Decisões incorretas

Exemplo: Teste ELISA para detecção do HIV

Relembrando

$$VPP = \frac{ps}{ps + (1 - p)(1 - e)} \text{ e } PFP = 1 - VPP$$

- Para uma prevalência de 2% ($p = 1/50 = 0,02$) temos:

$$VPP = \frac{0,02 \times 0,95}{0,02 \times 0,95 + (1 - 0,02) \times (1 - 0,998)} = 0,9065$$
$$PFP = 1 - 0,9065 = 0,0935$$

- Valor preditivo positivo de 90,65% e probabilidade de falso positivo de 9,35%.
- Da mesma forma (sabendo e , s e p) obtemos $VPN = 99,89\%$ e $PFN = 0,11\%$.

Decisões incorretas

Exemplo: Teste ELISA para detecção do HIV

- Variando os valores da prevalência temos os seguintes resultados para VPP, VPN, PFP e PFN:

Tabela: Valores de predição (VPP e VPN) e proporção de falsos resultados (PFP e PFN) do teste ELISA para detecção do HIV, versão ABBOTT.

| Prevalência | VPP (%) | VPN (%) | PFP (%) | PFN (%) |
|-------------|---------|---------|---------|---------|
| 1/100.000 | 0,47 | 100,00 | 99,53 | 0,00 |
| 1/10.000 | 4,54 | 100,00 | 95,46 | 0,00 |
| 1/1.000 | 32,21 | 99,99 | 67,79 | 0,01 |
| 1/500 | 48,77 | 99,99 | 51,23 | 0,01 |
| 1/200 | 70,47 | 99,99 | 29,53 | 0,01 |
| 1/100 | 82,75 | 99,99 | 17,25 | 0,01 |
| 1/100 | 90,65 | 99,89 | 9,35 | 0,11 |

Decisões incorretas

Exemplo: Teste ELISA para detecção do HIV

- Considerando-se a população total de um país, a AIDS é uma doença de prevalência baixa.
- Os resultados da tabela mostram que em um programa de uso do teste em larga escala, grande parte dos pacientes com **resultado positivo** consiste na realidade de **falsos-positivo** (VPP baixo e PFP alto).
- Por outro lado VPN é alto e PFN é baixo para prevalência menores que 1/100. Ou seja há pouquíssimos falsos-negativos.
- Um resultado positivo deve ser reconfirmado através de teste baseado em tecnologia diferente do ELISA.

Escolha entre testes diagnósticos

- Idealmente, os testes utilizados devem ter alta sensibilidade e especificidade.
- Na prática nem sempre existem testes disponíveis com características ótimas.
- É comum a situação em que há mais de um teste ou os valores da sensibilidade e especificidade podem ser calibrados. Desta forma, surge a necessidade da comparação entre eles para uma escolha mais adequada de acordo com o seu objetivo. O objetivo muitas vezes pode ser visto como triagem (prevalência baixa) ou consultório médico (prevalência alta).

Escolha entre testes diagnósticos

Tabela: VPP e VPN para alguns valores de sensibilidade e especificidade para prevalências de 1% e 90%.

| <i>s</i> | <i>e</i> | Pravalência de 1% | | Pravalência de 90% | |
|----------|----------|-------------------|------------|--------------------|------------|
| | | <i>VPP</i> | <i>VPN</i> | <i>VPP</i> | <i>VPN</i> |
| 0,99 | 0,99 | 0,5000 | 0,9999 | 0,9989 | 0,9167 |
| 0,99 | 0,90 | 0,0909 | 0,9999 | 0,9889 | 0,9091 |
| 0,99 | 0,80 | 0,0476 | 0,9999 | 0,9780 | 0,8989 |
| 0,90 | 0,99 | 0,4762 | 0,9990 | 0,9988 | 0,5238 |
| 0,90 | 0,90 | 0,0833 | 0,9989 | 0,9878 | 0,5000 |
| 0,90 | 0,80 | 0,0435 | 0,9987 | 0,9759 | 0,4706 |
| 0,80 | 0,99 | 0,4469 | 0,9980 | 0,9986 | 0,3548 |
| 0,80 | 0,90 | 0,0748 | 0,9978 | 0,9863 | 0,3333 |
| 0,80 | 0,80 | 0,0388 | 0,9975 | 0,9730 | 0,3077 |

Escolha entre testes diagnósticos

- Quando **prevalência é baixa**, VPP é mais influenciado pela especificidade.
 - VPN é pouco influenciado tanto pela sensibilidade quanto pela especificidade.
- Para **prevalência alta**, o VPP é próximo de 1, independente dos valores da sensibilidade e da especificidade.
 - VPN é influenciado mais pela sensibilidade do que pela especificidade.

Escolha entre testes diagnósticos

Conclusões

- 1 Um teste com alta especificidade deve ser usado quando a prevalência da doença é relativamente baixa (*doença rara*), mesmo que o teste tenha relativamente baixa sensibilidade.
- 2 Um teste com alta sensibilidade deve ser usado quando a prevalência da doença é alta (*doença comum*), mesmo que o teste tenha relativamente baixa especificidade.

Combinação de testes diagnósticos

- Muitas vezes, para o diagnóstico de certa doença, dispomos apenas de testes com VPP e VPN baixo.
- Se existe um bom teste, este pode ser muito caro ou oferece grande risco ao paciente.
- Nestas circunstâncias, uma opção frequentemente usada é o uso de uma **combinação de testes** mais simples.
- A associação entre os testes eleva a qualidade do diagnóstico.

Combinação de testes diagnósticos

- Quando dois testes são usados para se chegar a um diagnóstico é preciso saber como são obtidos os índices de qualidade do teste múltiplo.
- As maneiras mais simples de se formar um teste múltiplo a partir de resultados de dois testes são os esquemas **em paralelo** e **em série**.

Combinação de testes diagnósticos

Combinação em paralelo (sensibilidade)

- O resultado do teste será considerado positivo **se pelo menos um dos testes apresentar resultado positivo**. É de maior utilidade em casos de urgência.

| Teste A | Teste B | Teste em paralelo |
|---------|---------|-------------------|
| - | - | - |
| - | + | + |
| + | - | + |
| + | + | + |

Sensibilidade

$$\begin{aligned}P(T = + | D = 1) &= P(T_{A+} \cup T_{B+} | D = 1) \\ &= P(T_{A+} | D = 1) + P(T_{B+} | D = 1) - P(T_{A+} \cap T_{B+} | D = 1) \\ s &= s_A + s_B - s_A \times s_B\end{aligned}$$

Combinação de testes diagnósticos

Combinação em paralelo (especificidade)

- Pode-se calcular a especificidade de um teste em paralelo admitindo-se que os resultados dos dois testes **são independentes**.

| Teste A | Teste B | Teste em paralelo |
|---------|---------|-------------------|
| - | - | - |
| - | + | + |
| + | - | + |
| + | + | + |

Especificidade

$$\begin{aligned}P(T = - | D = 0) &= P(T_{A-} \cap T_{B-} | D = 0) \\ &= P(T_{A-} | D = 0) \times P(T_{B-} | D = 0) \\ e &= e_A \times e_B\end{aligned}$$

Combinação de testes diagnósticos

Combinação em série

- Os testes são aplicados consecutivamente, sendo o segundo aplicado apenas se o primeiro apresentar resultado positivo. Esse procedimento é indicado em situações onde não há necessidade de rápido atendimento e quando o paciente poder ser acompanhado ao longo do tempo.

| Teste A | Teste B | Teste em série |
|---------|---------------|----------------|
| - | desnecessário | - |
| + | - | - |
| + | + | + |

Combinação de testes diagnósticos

Combinação em série (sensibilidade)

| Teste A | Teste B | Teste em série |
|---------|---------------|----------------|
| - | desnecessário | - |
| + | - | - |
| + | + | + |

Sensibilidade

$$\begin{aligned}P(T = +|D = 1) &= P(T_{A+} \cap T_{B+}|D = 1) \\ &= P(T_{A+}|D = 1) \times P(T_{B+}|D = 1) \\ s &= s_A \times s_B\end{aligned}$$

Combinação de testes diagnósticos

Combinação em série (especificidade)

| Teste A | Teste B | Teste em série |
|---------|---------------|----------------|
| - | desnecessário | - |
| + | - | - |
| + | + | + |

Especificidade

$$\begin{aligned}P(T = - | D = 0) &= P(T_{A-} \cup T_{B-} | D = 0) \\ &= P(T_{A-} | D = 0) + P(T_{B-} | D = 0) \\ &\quad - P(T_{A-} | D = 0) \times P(T_{B-} | D = 0) \\ e &= e_A + e_B - e_A \times e_B\end{aligned}$$

Combinação de testes diagnósticos

Exemplo: Diagnóstico de câncer pancreático

Imagine um paciente idoso com dores persistentes nas costas e no abdômen e perda de peso. Na ausência de uma explicação para estes sintomas, a possibilidade de câncer do pâncreas é frequentemente levantada. É comum para se verificar esta possibilidade diagnóstica, que ambos os testes de **ultra-som** e **tomografia computadorizada** do pâncreas sejam solicitados.

Tabela: Sensibilidade e especificidade dos testes de ultra-som e tomografia computadorizada no diagnóstico do câncer de pâncreas, individualmente e em conjunto (dados hipotéticos).

| Teste | Sensibilidade (%) | Especificidade (%) |
|--------------------|-------------------|--------------------|
| A: Ultrassom | 80 | 60 |
| B: Tomografia | 90 | 90 |
| Paralelo: A ou B + | 98 | 54 |
| Série: A e B + | 72 | 96 |

Necessidade da combinação de testes

Há pelo menos duas situações em que a necessidade de combinação de testes surge naturalmente

- **Triagem:** é um tipo de procedimento que visa classificar pessoas assintomáticas quanto à probabilidade de terem ou não a doença.
 - É aplicado em grande número de pessoas de uma população.
 - A triagem não faz um diagnóstico, mas aponta as pessoas com maior probabilidade de estarem doentes. Essas são submetidas a um teste diagnóstico para comprovar ou não a presença da doença.
 - É indicada em caso de doença séria, se o tratamento na fase assintomática é mais benéfico do que na fase sintomática e em casos de alta prevalência.
 - O teste de triagem é considerado o primeiro dos testes de uma combinação em série

Necessidade da combinação de testes

- **Diagnóstico individual:** aparecem os dois tipos de combinação: em série e paralelo.
 - A combinação em paralelo é usada em casos de urgência ou para pacientes residentes em lugares distantes.
 - A combinação em série é usada em consultórios e clínicas hospitalares e em casos de testes caros e arriscados.