Comparação de dois grupos (qualitativo)

Testes para proporções

Felipe Figueiredo

Instituto Nacional de Traumatologia e Ortopedia

Sumário

Discussão da leitura obrigatória da aula passada

Sumário

Considere a seguinte tabela de contingência:

	Lesão	Não tem lesão
Alongou-se	18	22
Não se alongou	211	189

Fonte: Larson & Farber 2013

- 4% dos que se alongaram tiveram lesão
- 48% dos que não se alongaram tiveram lesão

Pergunta

Como determinar se existe alguma relação entre as variáveis?

Isto é: **o desfecho é independente da exposição?**

Quais são as variáveis?

- Dependente: desfecho (categórica)
- Independente: exposição (categórica)

Esta relação pode ser expressa como

$$\text{desfecho} \sim \text{exposição}$$

Mas antes vamos ver o caso de uma única variável.

Sumário

Exemplo 1

Considere que 10% dos pacientes morrem após uma operação arriscada. Em uma amostra de 75 pacientes, observou-se que 16 pacientes morreram após a operação.

Como comparar o número de óbitos osbervado e o número esperado?

Fonte: Motulsky, 1995

Quais são as variáveis?

- Dependente: mortalidade (categórica)
- Independente: parâmetro fixo

Esta relação pode ser expressa como

$$\text{mortalidade} \sim 10\%$$

Exemplo 1

Exemplo 1

Considere que 10% dos pacientes morrem após uma operação arriscada.
Em uma amostra de 75 pacientes, observou-se que 16 pacientes morreram após a operação.

- O número observado de óbitos em 75 pacientes foi 16.

Exemplo 1

Exemplo 1

Considere que 10% dos pacientes morrem após uma operação arriscada.
Em uma amostra de 75 pacientes, observou-se que 16 pacientes morreram após a operação.

- O número observado de óbitos em 75 pacientes foi 16.
- O número esperado seria $75 \times 10\% = 7.5$

Exemplo 1

Exemplo 1

Considere que 10% dos pacientes morrem após uma operação arriscada.
Em uma amostra de 75 pacientes, observou-se que 16 pacientes morreram após a operação.

- O número observado de óbitos em 75 pacientes foi 16.
- O número esperado seria $75 \times 10\% = 7.5$
- A discrepância nos óbitos foi $16 - 7.5 = 8.5$

Observação

Se as frequências observadas forem iguais:

- diferença entre ambas (discrepância) $= 0$
-
-

- Esse aumento reflete uma mudança real na mortalidade?
- Em uma amostra qualquer com 75 pacientes esperaríamos observar 7.5 óbitos
- Em uma amostra específica poderíamos observar mais ou menos que isso
- Provavelmente algo próximo de 7.5

Pergunta

Se a mortalidade for 10%, qual é a probabilidade de se observar 16 ou mais óbitos em uma amostra de 75 pacientes?

- Podemos representar as contagens observadas e esperadas em uma tabela
- H_0: observamos uma amostra de uma população com 10% de mortalidade.
- As diferenças entre os dados observados e os esperados tem distribuição aproximadamente χ^2 (qui-quadrado)

Estatística de teste

$$\chi^2 = \frac{\sum(\text{observado} - \text{esperado})^2}{\text{esperado}}$$

A razão $\frac{\text{observado}}{\text{esperado}}$ é um princípio **central** na Estatística[1]

[1] Usada para comparação de valores quadráticos, incluindo variâncias

Observação

Se as frequências observadas forem iguais:

- diferença entre ambas (discrepância) $= 0$
- diferença $^2 = 0$
- $\chi^2 = 0$

Comparação de dois grupos (qualitativo)

Felipe Figueiredo

1 amostra
Objetivo da aula
Teste qui-quadrado para 1 proporção
2 amostras
Aprofundamento

Expectativa x Realidade

Observação

Se as frequências observadas forem iguais:

- diferença entre ambas (discrepância) $= 0$
- diferença ^{2} = 0
- $\chi^2 = 0$

Quanto maior o valor de de χ^2, maior a discrepância

Exemplo 1

	Observado	Esperado
Óbito	16	7.5
Vivo	59	67.5
Total	75	75

Estatística de teste:

$$\chi^2 = \frac{(16 - 7.5)^2}{7.5} + \frac{(59 - 67.5)^2}{67.5} =$$

$$= \frac{(8.5)^2}{7.5} + \frac{(-8.5)^2}{67.5} \approx 10.70$$

- H_0: não houve alteração da mortalidade do procedimento.
- Estatística de teste para a amostra: $\chi^2 = 10.7$.
- O teste χ^2 retorna $p = 0.0011$.

Resultado

(...) a mortalidade observada foi diferente de 10% ($p = 0.0011$).

Sumário

Definição

Uma tabela de contingência mostra as frequências observadas para duas variáveis categóricas.

- Podemos calcular as frequências esperadas, baseado
  - no tamanho das amostras
  - na H_0
- Comparação: frequência observada $\times$ frequência esperada

A tabela do exemplo 1 (óbitos) não é uma tabela de contingência! (Por que?)

Exemplo 8.1

Frequências observadas:

	doença progrediu	doença não progrediu
AZT	76	399
Placebo	129	332

- Existe relação entre o uso do AZT e a progressão da doença?
- Ou: nessa amostra o AZT foi mais eficiente que o placebo (rejeitar H_0)?

Quais são as variáveis?

- Dependente: desfecho (categórica)
- Independente: tratamento (categórica)

Esta relação pode ser expressa como

$$\text{progressão} \sim \text{grupo}$$

- H_0: o AZT não é mais eficaz que o placebo
- Pergunta: assumindo a H_0, qual seria a frequência esperada para a progressão da doença?
- Em outras palavras: quantos pacientes tiveram progressão na doença, em relação ao total?

Vamos começar pela primeira célula da tabela

Exemplo 8.1

Frequências observadas:

	progrediu	não progrediu	total
AZT	76	399	475
Placebo	129	332	461
total	205	731	936

- Proporção esperada $E = \frac{205}{936} \approx 0.2190 = 21.90\%$
- Frequência esperada (número): $475 \times 0.2190 = 104.025 \approx 104.0$

- Se a H_0 fosse verdadeira, esperaríamos que 104.0 pacientes tivessem a progressão da doença, usando o AZT.

- Mas observamos 76.
- Discrepância $|104.0 - 76| = 28$ pacientes

- Faltam os 3 outros valores esperados e discrepâncias

- Para simplificar, podemos usar a seguinte fórmula:

$$E = \frac{\text{total por linha} \times \text{total por coluna}}{\text{total da tabela}}$$

Exemplo 8.1

Frequências observadas:

	progrediu	não progrediu	total
AZT	76	399	475
Placebo	129	332	461
total	205	731	936

- AZT + Progressão = $\frac{205 \times 475}{936} = 104.0$
- AZT + Não progressão = $\frac{731 \times 475}{936} = 371.0$
- Placebo + Progressão = $\frac{205 \times 461}{936} = 101.0$
- Placebo + Não progressão = $\frac{731 \times 461}{936} = 360.0$

Colocando os valores em uma tabela semelhante:

Exemplo 8.1

Frequências esperadas:

	progrediu	não progrediu	total
AZT	104.0	371.0	475.0
Placebo	101.0	360.0	461.0
total	205.0	731.0	936.0

Observe que os totais esperados devem ser iguais aos observados!

- H_0: a progressão é independente do grupo de tratamento
- ou: não há relação entre o uso do AZT e a progressão da doença.

- Somamos as diferenças quadráticas entre o valor observado e o esperado

$$\chi^2 = \frac{\sum(\text{observado} - \text{esperado})^2}{\text{esperado}}$$

- Quanto maior o valor de de χ^2, maior a discrepância
- Fazemos o teste χ^2 e julgamos o p-valor

Exemplo 8.1

- $\text{AZT + P} = \dfrac{(76 - 104.0)^2}{104.0} = \dfrac{28^2}{104.0} \approx 7.54$
- $\text{AZT + NP} = \dfrac{(399 - 371.0)^2}{371.0} = \dfrac{28^2}{371.0} \approx 2.11$
- $\text{Placebo + P} = \dfrac{(129 - 101.0)^2}{101.0} = \dfrac{28^2}{101.0} \approx 7.76$
- $\text{Placebo + NP} = \dfrac{(332 - 360.0)^2}{360.0} = \dfrac{28^2}{360.0} \approx 2.18$

$$\chi^2 = 7.54 + 2.11 + 7.76 + 2.18 = 19.59$$

- Quanto maior for o valor da estatística de teste, menor será o p-valor.
- Calculamos a estatística de teste para a amostra e encontramos $\chi^2 = 19.59$
- O resultado deste teste é $p < 0.0001$.

- Se a H_0 for verdadeira, temos uma chance menor que 0.01% de observar ao acaso uma discrepância tão grande entre os valores observados e os esperados.
- Resultado: devemos rejeitar a H_0

Interpretação

Rejeitamos a hipótese de que o AZT não é mais eficiente que o placebo.

- O teste χ^2 é apenas uma aproximação da distribuição dos dados, que pode ser usado para amostras grandes.
- Vantagem: simples
- Desvantagem: a aproximação é ruim para amostras pequenas
- Nunca usar se alguma célula da tabela tiver valor < 5

O teste indicado para este cenário é o teste exato de Fisher

- Para as seguintes situações deve-se usar o teste exato de Fisher:
  1. Quando se tem amostras pequenas
  2. Quanto se tem amostras de tamanho moderado, e se tiver uma ferramenta computacional disponível

- Se sua amostra for enorme (milhares de dados), prefira o teste χ^2, pois:
  1. o cálculo do teste exato de Fisher pode ser lento
  2. a aproximação será boa

Sumário

Exemplo 8.1

Frequências observadas:

	doença progrediu	doença não progrediu
AZT	76	399
Placebo	129	332

- Existe relação entre o uso do AZT e a progressão da doença?
- Ou: nessa amostra o AZT foi mais eficiente que o placebo (rejeitar H_0)?

Saída típica de um programa

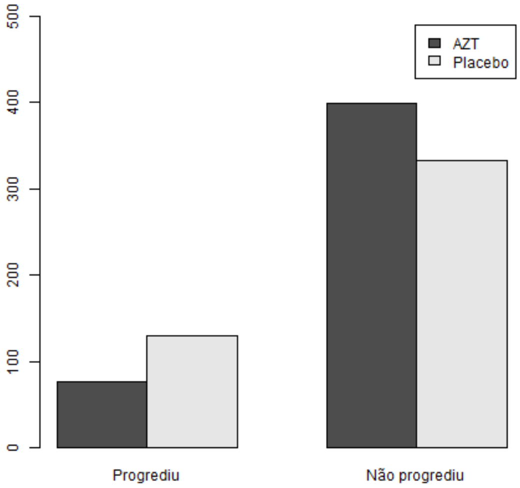
Teste Qui-quadrado

```
Pearson's Chi-squared test with Yates'
 continuity correction

data:  exemplo8.1
X-squared = 18.944, df = 1,
 p-value = 1.346e-05
```

Teste exato de Fisher

```
Fisher's Exact Test for Count Data

data:  exemplo8.1
p-value = 9.24e-06
alternative hypothesis: true odds ratio
 is not equal to 1
95 percent confidence interval:
 0.3512693 0.6818650
sample estimates:
odds ratio
 0.4905877
```

Visualização - gráfico de barra



Atenção

NÃO use gráfico de pizza!

- É uma visualização ineficiente
- Nosso olho é "bom" para julgar distâncias/comprimentos
- Nosso olho é ruim para julgar áreas
- Indicado **apenas** quando as categorias são muito discrepantes

Cleveland (1985)

"Data that can be shown by pie charts always can be shown by a dot chart.

This means that judgements of position along a common scale can be made instead of the less accurate angle judgements."

Sumário

E quando temos mais do que duas categorias?

- E quando temos mais do que duas categorias?
- Resposta: procedemos como no caso anterior, mas precisamos considerar os graus de liberdade do teste χ^2

$$gl = (l - 1)(c - 1) = (\text{linhas} - 1) \times (\text{colunas} - 1)$$

- Obs: no caso 2×2 temos $gl = (2 - 1) \times (2 - 1) = 1 \times 1 = 1$

Exemplo 3

Em dois hospitais, os resultados de 575 autópsias foram comparados com as causas de morte listadas nos atestados. Um dos hospitais que participou do estudo era comunitário (A); o outro era universitário (B).

Hospital	Precisão confirmada	Falta de informações	Recodificação incorreta
A	157	18	54
B	268	44	34

Os resultados sugerem práticas diferentes no preenchimento de atestados de óbito nos dois hospitais?

Fonte: Aula Hacker & Simões (2008 - Fiocruz)

- H_0: Dentro de cada categoria do status do atestado, as proporções de atestados de óbitos no hospital A são idênticas ao hospital B.
- H_1: As proporções não são idênticas
- Graus de liberdade:

$$(l-1) \times (c-1) = (2-1) \times (3-1) = 1 \times 2 = 2$$

- Dependente: qualidade do preenchimento (categórica)
- Independente: hospital (categórica)

Esta relação pode ser expressa como

$$\text{preenchimento} \sim \text{hospital}$$

Teste exato de Fisher

```
Fisher's Exact Test for Count Data

data:  exemplo3
p-value = 2.575e-05
alternative hypothesis: two.sided
```

Teste Qui-quadrado

```
Pearson's Chi-squared test

data:  exemplo3
X-squared = 21.523, df = 2, p-value = 2.12e-05
```

- Estatística de teste $\chi^2 = 21.52$
- p-valor: $p < 0.001$
- Rejeitamos H_0 ao nível de significância de $\alpha = 0.05$.

Resultado

Há associação entre o hospital e o status do atestado.

Conclusão

Parece que o hospital A tem maior proporção de atestados incorretos.

Sumário

- O teste de Fisher é um teste de independência entre os grupos
- O teste Qui-quadrado é uma boa aproximação, para N grande

Sumário

Leitura obrigatória

- Capítulo 26.
- Capítulo 27, pular a seção: Calculando o poder

Leitura recomendada

- Capítulo 29: Outros testes de tabelas de contingência
- Capítulo 27, seção: Calculando o poder