Bioestatística

INFERÊNCIA ESTATÍSTICA

Silvia Shimakura

Testes de hipóteses

- Intervalos de confiança são a maneira mais informativa de apresentar achados de um estudo.
- Mas pode haver interesse em investigar se o efeito em estudo é igual a um valor específico, ou se temos evidência razoável para excluir este valor.
- Testes de hipóteses nos dá uma estrutura para fazermos isso.

Teste de hipótese

- **Hipótese**
  - É uma afirmativa sobre uma característica da população
- **Teste de hipótese**
  - É um protocolo para testar uma afirmativa sobre uma característica da população
  - Auxilia na tomada de decisões sobre a população com base em informações amostrais

Exemplo

- Estudo com 560 pacientes com câncer, classificados segundo sexo e tipo de tumor cerebral (glioma e meningioma).

Sexo	Tipo de tumor				Total
	Glioma		Meningioma		
	N	%	N	%	
Feminino	129	55,6	103	44,4	232
Masculino	280	85,4	48	14,6	328
Total	409	73,0	151	27,0	560

- Existe diferença nas proporções de homens e mulheres com glioma, ou a diferença pode ter ocorrido ao acaso na amostra?

Procedimento geral

1) Estabeleça as hipóteses (H_0 e H_a)
2) Estabeleça o nível de significância desejado.
3) Decida o teste a ser usado e cheque suas suposições
4) Calcule a estatística de teste, supondo que H_0 é verdadeira
5) Calcule a probabilidade de observar a estatística de teste (valor-p), supondo que H_0 é verdadeira.
6) Decida se existe evidência suficiente para rejeitar a hipótese nula (↓valor-p ↑evidência contra H_0)
7) Formule suas conclusões e interpretação dos resultados

Procedimento geral

1) Estabeleça as hipóteses (H_0 e H_a)

H_0: Não há diferença nas proporções de homens e mulheres com glioma

H_a: Há diferença nas proporções de homens e mulheres com glioma

Ao fazer um teste chegamos a dois resultados possíveis:

- **Rejeitar H_0** em favor de H_a
- **Não rejeitar H_0** e concluímos que não temos evidência contra H_0

Procedimento geral

2) Estabeleça o nível de significância desejado

Desejamos que a probabilidade de errar ao rejeitar H_0, quando H_0 é verdadeira seja de no máximo 5%.

- **Possíveis erros de decisão:**
  - **Erro Tipo I:** rejeitar H_0, quando H_0 é verdadeira (falso positivo)
  - **Erro Tipo II:** não rejeitar H_0, quando H_0 é falsa (falso negativo)

Procedimento geral

2) Estabeleça o nível de significância desejado

Desejamos que a probabilidade de errar ao rejeitar H_0, quando H_0 é verdadeira seja de no máximo 5%.

- **Possíveis erros de decisão:**
  - **Erro Tipo I:** rejeitar H_0, quando H_0 é verdadeira (falso positivo) → P(Erro Tipo I)=α
  - **Erro Tipo II:** não rejeitar H_0, quando H_0 é falsa (falso negativo) →P(Erro Tipo II)=β

Procedimento geral

2) Estabeleça o nível de significância desejado

Desejamos que a probabilidade de errar ao rejeitar H_0, quando H_0 é verdadeira seja de no máximo 5%.

- **Possíveis erros de decisão:**
  - **Erro Tipo I:** rejeitar H_0, quando H_0 é verdadeira (falso positivo) → P(Erro Tipo I)=α → nível de significância
  - **Erro Tipo II:** não rejeitar H_0, quando H_0 é falsa (falso negativo) →P(Erro Tipo II)=β

Procedimento geral

3) Decida o teste a ser usado e cheque suas suposições

Deseja-se comparar proporções de indivíduos com certa característica em grupos independentes → Teste χ^2

Teste Qui-Quadrado χ^2

- Usado pra comparar dois ou mais grupos independentes de indivíduos (ex. Homem/Mulher) e uma variável categórica (duas ou mais categorias) de cada indivíduo (ex. melhora/sem mudança/piora)
- Vamos tomar o caso mais simples: 2 grupos e uma variável dicotômica (presente/ausente)
- Temos os dados apresentados na seguinte tabela de contigência:

	Característica		
Grupo	Presente	Ausente	Total
1	a	b	n_1
2	c	d	n_2
Total	r_1	r_2	N

Exemplo

- Estudo com 560 indivíduos. Tabela de sexo de pacientes versus tipo de tumor cerebral (glioma e meningioma).

Sexo	Tipo de tumor				Total
	Glioma		Meningioma		
	N	%	N	%	
Feminino	129	55,6	103	44,4	232
Masculino	280	85,4	48	14,6	328
Total	409	73,0	151	27,0	560

Procedimento geral

4) Calcule a estatística de teste, supondo que H_0 é verdadeira

Se H_0 fosse verdadeira, qual seria a tabela esperada?

Procedimento geral

Se H_0 fosse verdadeira, qual seria a tabela esperada?

Grupo	Resposta		Total
	Presente	Ausente	
1	$n_1 \times \frac{r_1}{N}$	$n_1 \times \frac{r_2}{N}$	n_1
2	$n_2 \times \frac{r_1}{N}$	$n_2 \times \frac{r_2}{N}$	n_2
Total	r_1	r_2	N

Procedimento geral

Se H_0 fosse verdadeira, qual seria a tabela esperada?

Grupo	Resposta		Total
	Presente	Ausente	
1	$n_1 \times \frac{r_1}{N}$	$n_1 \times \frac{r_2}{N}$	$n_1 = 232$
2	$n_2 \times \frac{r_1}{N}$	$n_2 \times \frac{r_2}{N}$	$n_2 = 328$
Total	$r_1 = 409$	$r_2 = 151$	$N = 560$

Procedimento geral

Se H_0 fosse verdadeira, qual seria a tabela esperada?

Grupo	Tipo de Tumor		Total
	Glioma	Meningioma	
1	$232 \times \frac{409}{560}$	$232 \times \frac{151}{560}$	$n_1 = 232$
2	$328 \times \frac{409}{560}$	$328 \times \frac{151}{560}$	$n_2 = 328$
Total	$r_1 = 409$	$r_2 = 151$	$N = 560$

Procedimento geral

Observados

Grupo	Tipo de Tumor		Total
	Glioma	Meningioma	
1	129	103	232
2	280	48	328
Total	409	151	560

Esperados

Grupo	Tipo de Tumor		Total
	Glioma	Meningioma	
1	169,4	62,6	232
2	239,6	88,4	328
Total	409	151	560

Procedimento geral

Observados

Grupo	Tipo de Tumor		Total
	Glioma	Meningioma	
1	129	103	232
2	280	48	328
Total	409	151	560

Esperados

Grupo	Tipo de Tumor		Total
	Glioma	Meningioma	
1	169,4	62,6	232
2	239,6	88,4	328
Total	409	151	560

$$X^2 = \frac{\sum_{caselas} (Obs - Esp)^2}{Esp} \sim distribuição\ \chi^2\ com\ (linhas - 1) \times (colunas - 1)\ graus\ de\ liberdade$$

Procedimento geral

Observados

Grupo	Tipo de Tumor		Total
	Glioma	Meningioma	
1	129	103	232
2	280	48	328
Total	409	151	560

Esperados

Grupo	Tipo de Tumor		Total
	Glioma	Meningioma	
1	169,4	62,6	232
2	239,6	88,4	328
Total	409	151	560

$$X^2 = \frac{\sum_{caselas} (Obs - Esp)^2}{Esp} \sim distribuição\ \chi^2\ com\ 1\ grau\ de\ liberdade$$

$$X^2 = \frac{(129-169,4)^2}{169,4} + \frac{(103-62,6)^2}{62,6} + \frac{(280-239,6)^2}{239,6} + \frac{(48-88,4)^2}{88,4} = 61,12$$

Procedimento geral

Observados

Grupo	Tipo de Tumor		Total
	Glioma	Meningioma	
1	129	103	232
2	280	48	328
Total	409	151	560

Esperados

Grupo	Tipo de Tumor		Total
	Glioma	Meningioma	
1	169,4	62,6	232
2	239,6	88,4	328
Total	409	151	560

$$X^2 = \frac{\sum_{caselas} (Obs - Esp)^2}{Esp} \sim distribuição\ \chi^2\ com\ 1\ grau\ de\ liberdade$$

$$X^2 = \frac{(129-169{,}4)^2}{169{,}4} + \frac{(103-62{,}6)^2}{62{,}6} + \frac{(280-239{,}6)^2}{239{,}6} + \frac{(48-88{,}4)^2}{88{,}4} = 61{,}12$$

E agora?

Procedimento geral

5) Supondo H_0 verdadeira, calcule a probabilidade de observar a estatística de teste (valor-p).

Se a hipótese nula for verdadeira, qual é a probab. do resultado observado 61,12 ter ocorrido por mero acaso?

$$X^2 = 61{,}12 \rightarrow P\left(X^2 \geq 61{,}12\right)$$

Procedimento geral

5) Suponha que H_0 é verdadeira e calcule a probabilidade de observar a estatística de teste (valor-p).

Se a hipótese nula for verdadeira, qual é a probab. do resultado observado 61,12 ter ocorrido por mero acaso?

valor-p

$$X^2 = 61{,}12 \;\rightarrow\; P\left(X^2 \geq 61{,}12\right) \ll 0{,}001$$

Procedimento geral

6) Decida se existe evidência suficiente para rejeitar a hipótese nula (↓valor-p ↑evidência contra H_0)

$$X^2 = 61{,}12 \rightarrow P\left(X^2 \geq 61{,}12\right) \ll 0{,}001$$

Como valor-p=0,1%<nível de significância 5% então rejeitamos H_0.

Procedimento geral

7) Formule suas conclusões e interpretação dos resultados

Há fortes evidências (valor-p<0,1%) de que pacientes com tumor cerebral do sexo feminino são menos propensas a ter gliomas do que pacientes com tumor cerebral do sexo masculino.