Testes de hipótese para tabelas de contingência: parte 1

Prof. Caio Azevedo

Exemplo 3: estudo sobre a inclinação (identificação) partidária estadunidense

- Considere o seguinte estudo, realizado para estudar a inclinação político-partidária estadunidense em relação aos gêneros masculino e feminino.
- Um total de 1557 mulheres e 1200 homens foram perguntados se possuem inclinação pelo partido democrata, republicano, ou se consideram ter uma inclinação "independente" (apartidária).
- Por simplicidade, por enquanto, vamos excluir a categoria "independente."

Exemplo 3 (cont.)

- Tabela de contingência (2×2) com os resultados da pesquisa.

		Inclinação partidária		
		Democrata	Republicano	Total
Gênero	Feminino	762	468	1230
	Masculino	484	477	961
Total	-	1246	945	2191

- Pergunta: as proporções de pessoas para cada inclinação partidária é a mesma entre os gêneros?

Modelagem probabilística: suposições

- Qual o mecanismo (probabilístico) gerador dos dados? Ou qual seria uma escolha razoável?

Modelagem probabilística: suposições

- Qual o mecanismo (probabilístico) gerador dos dados? Ou qual seria uma escolha razoável?
- Os totais de homens e mulheres foram fixados (totais marginais por linhas).

Modelagem probabilística: suposições

- Qual o mecanismo (probabilístico) gerador dos dados? Ou qual seria uma escolha razoável?
- Os totais de homens e mulheres foram fixados (totais marginais por linhas).
- Nenhuma outra informação é relevante (renda, escolaridade etc) e/ou de interesse.

Modelagem probabilística: suposições

- Qual o mecanismo (probabilístico) gerador dos dados? Ou qual seria uma escolha razoável?
- Os totais de homens e mulheres foram fixados (totais marginais por linhas).
- Nenhuma outra informação é relevante (renda, escolaridade etc) e/ou de interesse.
- Supondo que as pessoas respondem de modo independente, temos o produto de duas binomiais independentes.

Produto de binomiais (condicionalmente) independentes

- A tabela anterior é uma realização (amostra) possível, oriunda da seguinte estrutura:

		Inclinação partidária		
		Democrata	Republicano	Total
Gênero	Feminino	$N_{11}(\theta_{11})$	$N_{12}(\theta_{12})$	$n_{1.} = 1230$
	Masculino	$N_{21}(\theta_{21})$	$N_{22}(\theta_{22})$	$n_{2.} = 961$
Total	-	$N_{.1}$	$N_{.2}$	$n_{..} = 2191$

Produto de binomiais (condicionalmente) independentes

- Na tabela anterior, temos em que $N_{11}|F \sim \text{binomial}(n_{1.}, \theta_{11})$, $N_{21}|M \sim \text{binomial}(n_{2.}, \theta_{21})$ e $N_{11}|F \perp N_{12}|M$ (condicionalmente independentes).

Produto de binomiais (condicionalmente) independentes

- Na tabela anterior, temos em que $N_{11}|F \sim \text{binomial}(n_{1.}, \theta_{11})$, $N_{21}|M \sim \text{binomial}(n_{2.}, \theta_{21})$ e $N_{11}|F \perp N_{12}|M$ (condicionalmente independentes).
- Note que $N_{11}|F$ e $N_{21}|M$ são distribuições de probabilidade condicionais.

Produto de binomiais (condicionalmente) independentes

- Na tabela anterior, temos em que $N_{11}|F \sim \text{binomial}(n_{1.}, \theta_{11})$, $N_{21}|M \sim \text{binomial}(n_{2.}, \theta_{21})$ e $N_{11}|F \perp N_{12}|M$ (condicionalmente independentes).
- Note que $N_{11}|F$ e $N_{21}|M$ são distribuições de probabilidade condicionais.
- A notações $N_{11}|F$ e $N_{21}|M$ equivalem, respectivamente, à $N_{11}|n_{1.}$ e $N_{21}|n_{2.}$.

Produto de binomiais (condicionalmente) independentes

- Na tabela anterior, temos em que $N_{11}|F \sim \text{binomial}(n_{1.}, \theta_{11})$, $N_{21}|M \sim \text{binomial}(n_{2.}, \theta_{21})$ e $N_{11}|F \perp N_{12}|M$ (condicionalmente independentes).
- Note que $N_{11}|F$ e $N_{21}|M$ são distribuições de probabilidade condicionais.
- A notações $N_{11}|F$ e $N_{21}|M$ equivalem, respectivamente, à $N_{11}|n_{1.}$ e $N_{21}|n_{2.}$.
- Além disso $N_{.1} = N_{11} + N_{21}$, $N_{.2} = N_{12} + N_{22}$ e $n_{..} = N_{.1} + N_{.2} = n_{1.} + n_{2.}$.

Produto de binomiais (condicionalmente) independentes

- Portanto,

$$p(n_{i1}|n_{i.}) \equiv p(n_{i1}) = \frac{n_{i.}!}{n_{i1}!(n_{i.} - n_{i1})!}\theta_{i1}^{n_{i1}}(1-\theta_{i1})^{n_{i.}-n_{i1}}\mathbb{1}_{\{0,1,\ldots,n_{i.}\}}(n_{i1}),$$
$$i = 1,2$$

e

$$p(n_{11}, n_{21}|n_{1.}, n_{2.}) = p(n_{11}|n_{1.})p(n_{21}|n_{2.}) \equiv p(n_{11})p(n_{21})$$

Valores (frequências) observados

- Voltando ao exempo anterior:

		Inclinação partidária		
		Democrata	Republicano	Total
Gênero	Feminino	$n_{11} = 762$	$n_{12} = 468$	$n_{1.} = 1230$
	Masculino	$n_{21} = 484$	$n_{22} = 477$	$n_{2.} = 961$
Total	-	$n_{.1} = 1246$	$n_{.2} = 945$	$n_{..} = 2191$

Hipóteses de interesse

- $H_0 : \theta_{11} = \theta_{21}$ vs $H_1 : \theta_{11} \neq \theta_{21}$

Hipóteses de interesse

- $H_0 : \theta_{11} = \theta_{21}$ vs $H_1 : \theta_{11} \neq \theta_{21}$
- Note que se não rejeitarmos H_0, então $\theta_{12} = \theta_{22}$, caso contrário, $\theta_{12} \neq \theta_{22}$, pois $\theta_{i1} = 1 - \theta_{i2}, i = 1, 2$.

Hipóteses de interesse

- $H_0 : \theta_{11} = \theta_{21}$ vs $H_1 : \theta_{11} \neq \theta_{21}$
- Note que se não rejeitarmos H_0, então $\theta_{12} = \theta_{22}$, caso contrário, $\theta_{12} \neq \theta_{22}$, pois $\theta_{i1} = 1 - \theta_{i2}, i = 1, 2$.
- Não rejeitar H_0 implica em afirmar que as duas distribuições binomiais não são diferentes (homogeneidade entre as distribuições).

Hipóteses de interesse

- $H_0 : \theta_{11} = \theta_{21}$ vs $H_1 : \theta_{11} \neq \theta_{21}$
- Note que se não rejeitarmos H_0, então $\theta_{12} = \theta_{22}$, caso contrário, $\theta_{12} \neq \theta_{22}$, pois $\theta_{i1} = 1 - \theta_{i2}, i = 1, 2$.
- Não rejeitar H_0 implica em afirmar que as duas distribuições binomiais não são diferentes (homogeneidade entre as distribuições).
- Ou seja, sob H_0 ($\theta_{11} = \theta_{21} = \theta$), $N_{11}|F \sim$ binomial$(n_{1.}, \theta)$ e $N_{21}|M \sim$ binomial$(n_{2.}, \theta)$ (possuem a mesma distribuição).

Hipóteses de interesse

- Lembremos que, duas v.a.'s X e Y são independentes $\leftrightarrow$ $p(x,y) = p(x)p(y), \forall x, y \in \Omega$. Ou, de modo equivalente, $\leftrightarrow p(x|y) = p(x), \forall x, y \in \Omega$ ou $p(y|x) = p(y), \forall x, y \in \Omega$.
- Assim, como temos distribuições condicionais, não rejeitar H_0 equivale a concluir que a inclinação partidária independe do gênero.

Hipóteses de interesse

- Como testar as hipóteses acima?

Hipóteses de interesse

- Como testar as hipóteses acima?
- Sabemos que $\widehat{\theta}_{i1} = \dfrac{N_{i1}}{n_{i.}}$ é o estimador de MV de $\theta_{i1}, i = 1, 2$.

Hipóteses de interesse

- Como testar as hipóteses acima?
- Sabemos que $\widehat{\theta}_{i1} = \dfrac{N_{i1}}{n_{i.}}$ é o estimador de MV de $\theta_{i1}, i = 1, 2$.
- Portanto, sob as condições de regularidade, $\widehat{\theta}_{i1} \approx N\left(\theta_{i1}, \frac{\theta_{i1}(1-\theta_{i1})}{n_{i.}}\right)$, para $n_{i.}$ suficientemente grande, i=1,2.

Estatísticas do Teste

- Além disso, $\widehat{\theta}_{11} \perp \widehat{\theta}_{21}$.

Estatísticas do Teste

- Além disso, $\widehat{\theta}_{11} \perp \widehat{\theta}_{21}$.
- Note que, sob H_0, $\theta_{11} = \theta_{21} = \theta$ e $\widehat{\theta}_{i1} \approx N\left(\theta, \frac{\theta(1-\theta)}{n_{i.}}\right)$, para $n_{i.}$ suficientemente grande, $i = 1, 2$.

Estatísticas do Teste

- Além disso, $\widehat{\theta}_{11} \perp \widehat{\theta}_{21}$.
- Note que, sob H_0, $\theta_{11} = \theta_{21} = \theta$ e $\widehat{\theta}_{i1} \approx N\left(\theta, \frac{\theta(1-\theta)}{n_{i.}}\right)$, para $n_{i.}$ suficientemente grande, $i = 1, 2$.
- Assim, sob H_0, $\widehat{\theta}_{11} - \widehat{\theta}_{21} \approx N\left(0, \theta(1-\theta)\left(\frac{1}{n_{1.}} + \frac{1}{n_{2.}}\right)\right)$, para $n_{i.}$ suficientemente grande, $i = 1, 2$.

Estatísticas do Teste

- Além disso, $\widehat{\theta}_{11} \perp \widehat{\theta}_{21}$.
- Note que, sob H_0, $\theta_{11} = \theta_{21} = \theta$ e $\widehat{\theta}_{i1} \approx N\left(\theta, \frac{\theta(1-\theta)}{n_{i.}}\right)$, para $n_{i.}$ suficientemente grande, $i = 1, 2$.
- Assim, sob H_0, $\widehat{\theta}_{11} - \widehat{\theta}_{21} \approx N\left(0, \theta(1-\theta)\left(\frac{1}{n_{1.}} + \frac{1}{n_{2.}}\right)\right)$, para $n_{i.}$ suficientemente grande, $i = 1, 2$.
- Considere, ainda, $\widehat{\theta} = \dfrac{n_{1.}\widehat{\theta}_{11} + n_{2.}\widehat{\theta}_{21}}{n_{1.} + n_{2.}}$.

Estatísticas do Teste

- Logo, sob H_0, $Z_t = \frac{\widehat{\theta}_{11} - \widehat{\theta}_{21}}{\sqrt{\widehat{\theta}(1-\widehat{\theta})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \approx N(0,1)$, para n_i suficientemente grande, $i = 1, 2$, pois $\widehat{\theta}$ é um estimador consistente de θ.

Estatísticas do Teste

- Logo, sob H_0, $Z_t = \frac{\widehat{\theta}_{11} - \widehat{\theta}_{21}}{\sqrt{\widehat{\theta}(1-\widehat{\theta})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \approx N(0,1)$, para n_i suficientemente grande, $i = 1, 2$, pois $\widehat{\theta}$ é um estimador consistente de θ.

- Portanto, um teste para testar as hipóteses acima, consiste em rejeitar H_0 se $z_t \geq z_c$ ou $z_t \leq -z_c$, em que

$$P(Z \geq z_c | H_0) = \alpha/2, \alpha \in (0,1), Z \sim N(0,1)$$

Estatísticas do Teste

- e $z_t = \frac{\widetilde{\theta}_{11} - \widetilde{\theta}_{21}}{\sqrt{\widetilde{\theta}(1-\widetilde{\theta})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ e $\widetilde{\theta} = \frac{n_{1.}\widetilde{\theta}_{11} + n_{2.}\widetilde{\theta}_{21}}{n_{1.} + n_{2.}}$, lembrando que $\widetilde{(.)}$, denota a estimativa (valor numérico).
- Alternativamente, rejeita-se H_0 se $p - valor < \alpha$, em que $p = p - valor = 2P(Z > |z_t||H_0)$.

Voltando ao exemplo

- Temos que $\widetilde{\theta}_{11} = \dfrac{n_{11}}{n_{1.}} = 0,6195122$, $\widetilde{\theta}_{21} = \frac{n_{21}}{n_{2.}} = 0,503642$ e $\widetilde{\theta} = \frac{n_{1.}\widetilde{\theta}_{11} + n_{2.}\widetilde{\theta}_{21}}{n_{1.} + n_{2.}} = \frac{n_{11} + n_{21}}{n_{..}} = 0,5686901$.

Voltando ao exemplo

- Temos que $\widetilde{\theta}_{11} = \dfrac{n_{11}}{n_{1.}} = 0,6195122$, $\widetilde{\theta}_{21} = \frac{n_{21}}{n_{2.}} = 0,503642$ e $\widetilde{\theta} = \frac{n_{1.}\widetilde{\theta}_{11} + n_{2.}\widetilde{\theta}_{21}}{n_{1.} + n_{2.}} = \frac{n_{11} + n_{21}}{n_{..}} = 0,5686901$.
- Logo, $z_t = 5,434158$ e $p < 0,0001$ (p-valor). Assim, rejeita-se H_0.

Voltando ao exemplo

- Temos que $\widetilde{\theta}_{11} = \dfrac{n_{11}}{n_{1.}} = 0,6195122$, $\widetilde{\theta}_{21} = \frac{n_{21}}{n_{2.}} = 0,503642$ e $\widetilde{\theta} = \frac{n_{1.}\widetilde{\theta}_{11} + n_{2.}\widetilde{\theta}_{21}}{n_{1.} + n_{2.}} = \frac{n_{11} + n_{21}}{n_{..}} = 0,5686901$.
- Logo, $z_t = 5,434158$ e $p < 0,0001$ (p-valor). Assim, rejeita-se H_0.
- Portanto, as distribuições relativas à inclinação partidária não são as mesmas entre os gêneros e, assim, há dependência entre inclinação partidária e gênero.

Abordagens alternativas

- Uma outra opção é utilizar o TRV (teste da razão de verossimilhanças) na sua forma assintótica (exercício).

Abordagens alternativas

- Uma outra opção é utilizar o TRV (teste da razão de verossimilhanças) na sua forma assintótica (exercício).
- Outra opção: testes qui-quadrado, os quais são baseados em estatísticas que possuem, sob H_0, distribuição assintótica qui-quadrado com um certo número de graus de liberdade.

Abordagens alternativas

- Uma outra opção é utilizar o TRV (teste da razão de verossimilhanças) na sua forma assintótica (exercício).
- Outra opção: testes qui-quadrado, os quais são baseados em estatísticas que possuem, sob H_0, distribuição assintótica qui-quadrado com um certo número de graus de liberdade.
- Modelos de regressão.

Considerando o conjunto de dados em sua íntegra

- Considere agora os dados originais (tabela de contingência (2×3), ou seja:

		Inclinação partidária			
		Democrata	Independente	Republicano	Total
Gênero	Feminino	762	327	468	1557
	Masculino	484	239	477	1200
Total	-	1246	566	945	2757

- Pergunta: as proporções de pessoas para cada inclinação partidária é a mesma entre os gêneros?

Considerando o conjunto de dados em sua íntegra

- Considere agora os dados originais (tabela de contingência (2×3), ou seja:

		Inclinação partidária			
		Democrata	Independente	Republicano	Total
Gênero	Feminino	762	327	468	1557
	Masculino	484	239	477	1200
Total	-	1246	566	945	2757

- Pergunta: as proporções de pessoas para cada inclinação partidária é a mesma entre os gêneros?
- Como responder à pergunta de interesse?

Modelo probabilístico

- Representação populacional

		Inclinação partidária			
		Democrata	Independente	Republicano	Total
Gênero	Feminino	$N_{11}(\theta_{11})$	$N_{12}(\theta_{12})$	$N_{13}(\theta_{13})$	$n_{1.}$
	Masculino	$N_{21}(\theta_{21})$	$N_{22}(\theta_{22})$	$N_{23}(\theta_{23})$	$n_{2.}$
Total	-	$N_{.1}$	$N_{.2}$	$N_{.3}$	$n_{..}$

- Temos que:

$\mathbf{N}_1 = (N_{11}, N_{12})'|F \sim \text{Trinomial}(n_{1.}, \boldsymbol{\theta}_1), \boldsymbol{\theta}_1 = (\theta_{11}, \theta_{12})'$e

$\mathbf{N}_2 = (N_{21}, N_{22})'|M \sim \text{Trinomial}(n_{2.}, \boldsymbol{\theta}_2), \boldsymbol{\theta}_2 = (\theta_{21}, \theta_{22})'$ e

$\mathbf{N}_1|F \perp \mathbf{N}_2|M$, ou seja, o produto de duas distribuições trinomiais independentes.

Produto de trinomiais (condicionalmente) independentes

- Portanto,

$$
\begin{aligned}
p(n_{i1}, n_{i2} | n_{i.}) &\equiv p(n_{i1}, n_{i2}) = p(\mathbf{n}_i) \\
&= \frac{n_{i.}!}{n_{i1}! n_{i2}! (n_{i.} - n_{i1} - n_{i2})!} \theta_{i1}^{n_{i1}} \theta_{i2}^{n_{i2}} (1 - \theta_{i1} - \theta_{i2})^{n_{i.} - n_{i1} - n_{i2}} \\
&\times \mathbb{1}_{\{0,1,\ldots,n_{i.}\}}(n_{i1}) \mathbb{1}_{\{0,1,\ldots,n_{i.} - n_{i1}\}}(n_{i2}), i = 1, 2
\end{aligned}
$$

e

$$
p(\mathbf{n}_1, \mathbf{n}_2 | n_{1.}, n_{2.}) = p(\mathbf{n}_1 | n_{1.}) p(\mathbf{n}_2 | n_{2.}) \equiv p(\mathbf{n}_1) p(\mathbf{n}_2)
$$

Hipóteses de interesse

- Neste caso, temos que as hipóteses de interesse são

$$H_0 : \begin{cases} \theta_{11} = \theta_{21} = \theta_1 \\ \theta_{12} = \theta_{22} = \theta_2 \end{cases} \quad \text{vs } H_1 : \text{há pelo menos uma diferença}$$

- Note que, neste caso, se H_0 for verdade, então $\theta_{13} = \theta_{23}$ e, além disso, teremos que as distribuições trinomiais são idênticas (homogeneidade) e, portanto, independência.
- Como testar as hipóteses acima?

Testes de hipótese

- Uma opção: teste da razão de verossimilhanças (versão assintótica). Exercício.
- Outra opção: modelos de regressão (veremos um pouco à respeito, adiante).
- Outra opção: comparar o que os dados nos informa com o que esperamos sob H_0.

Teste de hipóteses (exercício)

- Os estimadores de MV de $\theta_{ij}, \forall, i, j$, irrestritos e sob H_0 são, respectivamente,

$$\widehat{\theta}_{ij} = \frac{N_{ij}}{n_{i.}}, \text{ e}$$

$$\widehat{\theta}^0_{ij} = \widehat{\theta}_j = \frac{\sum_{i=1}^{2} n_{i.}\widehat{\theta}_{ij}}{n_{..}} = \frac{\sum_{i=1}^{2} N_{ij}}{n_{..}} = \frac{N_{.j}}{n_{..}}, N_{.j} = \sum_{i=1}^{2} N_{ij}.$$

- As respectivas estimativas, são dadas por:

$$\widetilde{\theta}_{ij} = \frac{n_{ij}}{n_{i.}} \text{ e}$$

$$\widetilde{\theta}^0_{ij} = \widetilde{\theta}_j = \frac{\sum_{i=1}^{2} n_{i.}\widetilde{\theta}_{ij}}{n_{..}} = \frac{\sum_{i=1}^{2} n_{ij}}{n_{..}} = \frac{n_{.j}}{n_{..}}, n_{.j} = \sum_{i=1}^{2} n_{ij}.$$

Teste de hipóteses

- Frequências observáveis: N_{ij} (variável aleatória).
- Frequências esperadas sob o modelo probabilístico (produto de trinomiais independentes): $E(N_{ij}) = n_{i.}\theta_{ij}, \forall i, j$.
- Frequências esperadas sob o modelo probabilístico e H_0: $E(N_{ij})^0 = n_{i.}\theta_j, \forall i, j$.
- Por simplicidade, retiramos o condicionamento das contas acima e de algumas outras que serão apresentadas.

Teste de hipóteses

- Frequências observadas: n_{ij} (número).
- Estimadores das frequências esperadas sob o modelo probabilístico e H_0: $E_{ij} = n_{i.}\widehat{\theta}_{.j} = \frac{n_{i.}N_{.j}}{n_{..}}$.
- Estimativas das frequências esperadas sob o modelo probabilístico e H_0: $e_{ij} = n_{i.}\widetilde{\theta}_{.j} = \frac{n_{i.}n_{.j}}{n_{..}}$.
- Estatística do teste $Q_H = \sum_{i=1}^{2}\sum_{j=1}^{3}\frac{(N_{ij}-E_{ij})^2}{E_{ij}}$.
- Estatística (numérica) do teste $q_H = \sum_{i=1}^{2}\sum_{j=1}^{3}\frac{(n_{ij}-e_{ij})^2}{e_{ij}}$

Teste de hipóteses

- Quanto maior o valor de q_H, maior será a evidência contra H_0.
- Sob H_0, pode-se provar que $Q_H \approx \chi^2_2$, para n suficientemente grande.
- Dessa forma, rejeita-se H_0 se $q_H \geq q_c$, em que $P(X \geq q_c | H_0)$, ou se $p = p - valor = P(X \geq q_H | H_0) < \alpha$, com $X \sim \chi^2_1$.

Voltando ao exemplo

		Inclinação partidária			
		Democrata	Independente	Republicano	Total
Gênero	Feminino	762	327	468	1557
	Masculino	484	239	477	1200
Total	-	1246	566	945	2757

Voltando ao exemplo

- Estimativas das frequências esperadas. Por exemplo, $e_{11} = \frac{1557 \times 1246}{2757}$ e assim por diante.
- A seguir apresentamos as frequências observadas (e entre parênteses as esperadas).

		Inclinação partidária		
		Democrata	Independente	Republicano
Gênero	Feminino	762 (703,67)	327 (319,65)	368 (533,68)
	Masculino	484 (542,33)	239 (246,35)	477 (411,32)

Resultados

- Neste caso, $q_H = 30,07$ e $p - valor < 0,0001$. Assim, concluí-se que não há homegeneidade entre as distribuições. Ou seja, existe dependência entre gênero e inclinação partidária.
- Pergunta: como identificar, em detalhes, como é essa relação?

Tabela de contingência $r \times s$: produto de multinomiais independentes

		Variável 1 (resposta)					
		C_{11}	C_{12}	...	$C_{1(s-1)}$	C_{1s}	Total
Variável 2	C_{21}	$N_{11}(\theta_{11})$	$N_{12}(\theta_{12})$	...	$N_{1(s-1)}(\theta_{1(s-1)})$	$N_{1s}(\theta_{1s})$	$n_{1.}$
(explicativa)	C_{22}	$N_{21}(\theta_{21})$	$N_{22}(\theta_{22})$	...	$N_{1(s-1)}(\theta_{2(s-1)})$	$N_{2s}(\theta_{2s})$	$n_{2.}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	
	C_{2r}	$N_{r1}(\theta_{r1})$	$N_{r2}(\theta_{r2})$	...	$N_{r(s-1)}(\theta_{r(s-1)})$	$N_{rs}(\theta_{rs})$	$n_{r.}$
Total	-	$N_{.1}$	$N_{.2}$	...	$N_{.(s-1)}$	$N_{.s}$	$n_{..}$

Modelo probabilístico

- Seja $\mathbf{N}_i = (N_{i1}, N_{i2}, ..., N_{i(s-1)}), i = 1, ..., r$, assim, temos que

$$\mathbf{N}_i | C_{2i} \sim \text{Multinomial}(n_{i.}, \boldsymbol{\theta}_i), \boldsymbol{\theta}_i = (\theta_{i1}, \theta_{i2}, .., \theta_{i(s-1)})'.$$

- Além disso, $\mathbf{N}_i | C_{2i} \perp \mathbf{N}_j | C_{2j}, \forall i, j, i \neq j, i, j \in \{1, 2..., r\}$.
- Defina também $\mathbf{N} = (\mathbf{N}_1, ..., \mathbf{N}_r)'$ e $\mathbf{C} = (C_{21}, ..., C_{2r})'$.

Modelo probabilístico

- Assim, sendo A_i o conjunto que define o suporte da distribuição multinomial e lembrando que $n_{is} = n_i - \sum_{j=1}^{(s-1)} n_{ij}$ e $\theta_{is} = 1 - \sum_{j=1}^{(s-1)} \theta_{ij}, i = 1, 2, .., r$, temos que

$$p(\mathbf{n}_i | C_{2i}) \equiv p(\mathbf{n}_i) = \frac{n_{i.}!}{\prod_{j=1}^{s} n_{ij}} \prod_{j=1}^{s} \theta_{ij}^{n_{ij}} \mathbb{1}_{A_i}(\mathbf{n}_i)$$

e

$$p(\mathbf{n}|\mathbf{C}) = \prod_{i=1}^{r} p(\mathbf{n}_i | C_{2i}) = \prod_{i=1}^{r} \frac{n_{i.}!}{\prod_{j=1}^{s} n_{ij}} \prod_{i=1}^{s} \theta_{ij}^{n_{ij}} \mathbb{1}_{A_i}(\mathbf{n}_i)$$

Hipóteses de interesse

- Neste caso, temos que as hipóteses de interesse são

$$H_0 : \begin{cases} \theta_{11} = \theta_{21} = ... = \theta_{r1} = \theta_1 \\ \theta_{21} = \theta_{22} = ... = \theta_{r2} = \theta_2 \\ \vdots \\ \theta_{1(s-1)} = \theta_{2(s-1)} = ... = \theta_{r(s-1)} = \theta_{(s-1)} \end{cases}$$

vs

H_1 : há pelo menos uma diferença

Obteanção dos estimadores de MV

- Irrestritos: considerando $\boldsymbol{\theta}_i = (\theta_{i1}, ..., \theta_{is})'$, maximizar

$$l^*(\boldsymbol{\theta}_i, \lambda) = \sum_{j=1}^{s} n_{ij} \ln(\theta_{ij}) + \lambda \left(\sum_{j=1}^{s} \theta_{ij} - 1 \right)$$

para $i = 1, ..., r$

- Restritos (sob H_0 apresentada anteriormente): maximizar

$$l^{**}(\boldsymbol{\theta}, \lambda) = \sum_{j=1}^{s} n_{.j} \ln(\theta_j) + \lambda \left(\sum_{j=1}^{s} \theta_j - 1 \right)$$

em que $\boldsymbol{\theta} = (\theta_1, ..., \theta_s)$ e $n_{.j} = \sum_{i=1}^{s} n_{ij}$,

Teste de hipóteses

- Frequências observáveis: N_{ij} (variável aleatória).
- Frequências esperadas sob o modelo probabilístico (produto de multinomiais independentes): $E(N_{ij}) = n_{i.}\theta_{ij}, \forall i, j$.
- Frequências esperadas sob o modelo probabilístico e H_0: $E(N_{ij})^0 = n_{i.}\theta_j, \forall i, j$.

Teste de hipóteses

- Os estimadores de MV de $\theta_{ij}, i = 1, ..., r, j = 1, ..., s$, irrestritos e sob H_0 são, respectivamente,

$$\widehat{\theta}_{ij} = \frac{N_{ij}}{n_{i.}}, \text{ e}$$

$$\widehat{\theta}_{ij}^0 = \widehat{\theta}_j = \frac{\sum_{i=1}^r n_{i.}\widehat{\theta}_{ij}}{n_{..}} = \frac{\sum_{i=1}^r N_{ij}}{n_{..}} = \frac{N_{.j}}{n_{..}}, N_{.j} = \sum_{i=1}^r N_{ij}.$$

- As respectivas estimativas, são dadas por:

$$\widetilde{\theta}_{ij} = \frac{n_{ij}}{n_{i.}} \text{ e}$$

$$\widetilde{\theta}_{ij}^0 = \widetilde{\theta}_j = \frac{\sum_{i=1}^r n_{i.}\widetilde{\theta}_{ij}}{n_{..}} = \frac{\sum_{i=1}^r n_{ij}}{n_{..}} = \frac{n_{.j}}{n_{..}}, n_{.j} = \sum_{i=1}^r n_{ij}$$

.

Teste de hipóteses

- Frequências observadas: n_{ij} (número).
- Estimadores das frequências esperadas sob o modelo probabilístico e H_0: $E_{ij} = n_{i.}\widehat{\theta}_{.j} = \frac{n_{i.}N_{.j}}{n_{..}}$.
- Estimativas das frequências esperadas sob o modelo probabilístico e H_0: $e_{ij} = n_{i.}\widetilde{\theta}_{.j} = \frac{n_{i.}n_{.j}}{n_{..}}$.
- Estatística do teste $Q_H = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(N_{ij}-E_{ij})^2}{E_{ij}}$.
- Estatística (numérica) do teste $q_H = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{ij}-e_{ij})^2}{e_{ij}}$

Teste de hipóteses

- Quanto maior o valor de q_H, maior será a evidência contra H_0.
- Sob H_0, pode-se provar que $Q_H \approx \chi^2_{(r-1)(s-1)}$, para n suficientemente grande.
- Dessa forma, rejeita-se H_0 se $q_H \geq q_c$, em que $P(X \geq q_c | H_0)$, ou se $p = p - valor = P(X \geq q_H | H_0) < \alpha$, com $X \sim \chi^2_{(r-1)(s-1)}$.

Exemplo 1: comparação de métodos de detecção de cárie

- Suponha que um pesquisador lhe apresente a seguinte tabela de contingência, resumindo os dados coletados por ele, oriundos de um determinado experimento:

		Risco de cárie segundo o método convencional			
		Baixo	Médio	Alto	Total
Risco de cárie segundo o método simplificado	Baixo	11	5	0	16
	Médio	14	34	7	55
	Alto	2	13	11	26
Total	-	27	52	18	97

Relembrando

- Suponha que: selecionou-se, ao acaso, através de um processo de amostragem aleatória simples sem reposição, 97 indivídios (de um possível grupo de interesse?). Em cada um deles, as duas técnicas foram aplicadas.
- Podemos considerar que a tabela obtida é uma dentre várias possíveis, obteníveis ao se replicar o experimento. Ou seja, ela é uma amostra de uma população de interesse.
- Possível modelo probabilístico apropriado: multinominal com 9 classes (exaustivas).

Modelo probabilístico (Multinomial)

		Risco de cárie segundo o método convencional			
		Baixo	Médio	Alto	Total
Risco de cárie segundo o método simplificado	Baixo	$N_{11}(\theta_{11})$	$N_{12}(\theta_{12})$	$N_{13}(\theta_{13})$	$N_{1.}(\theta_{1.})$
	Médio	$N_{21}(\theta_{21})$	$N_{22}(\theta_{22})$	$N_{23}(\theta_{23})$	$N_{2.}(\theta_{2.})$
	Alto	$N_{31}(\theta_{31})$	$N_{32}(\theta_{32})$	$N_{33}(\theta_{33})$	$N_{3.}(\theta_{3.})$
Total	-	$N_{.1}(\theta_{.1})$	$N_{.2}(\theta_{.2})$	$N_{.3}(\theta_{.3})$	$n_{..}$

Estrutura probabilística

- Sejam $\mathbf{N} = (N_{11}, N_{12}, N_{13}, N_{21}, N_{22}, N_{23}, N_{31}, N_{32})'$, $\mathbf{N}_c = (N_{.1}, N_{.2})'$ e $\mathbf{N}_s = (N_{1.}, N_{2.})'$. Então
- $\mathbf{N} \sim \text{Multinomial}(n_{..}, \boldsymbol{\theta}), \boldsymbol{\theta} = (\theta_{11}, \theta_{12}, \theta_{13}, \theta_{21}, \theta_{22}, \theta_{23}, \theta_{31}, \theta_{32})'$,

  $\mathbf{N}_c \sim \text{Multinomial}(n_{..}, \boldsymbol{\theta}_c), \boldsymbol{\theta}_c = (\theta_{.1}, \theta_{.2})'$ e

  $\mathbf{N}_s \sim \text{Multinomial}(n_{..}, \boldsymbol{\theta}_s), \boldsymbol{\theta}_s = (\theta_{1.}, \theta_{2.})'$.

Perguntas

- Os métodos de detecção de cáries classificam os indivíduos de forma independente?
- Hipótese de independência: $H_0 : \theta_{ij} = \theta_{i.}\theta_{.j}, \forall i, j$ vs H_1 : há pelo menos uma diferença.
- Lembrando que o interesse principal é medir o grau de concordância entre os métodos quanto a classificação, a independência é uma característica favorável ou desfavorável ?
- Já discutimos que, duas situações favoráveis à existência de concordância entre os métodos são: concordância absoluta e concordância marginal.

Concordância absoluta

		Risco de cárie segundo o método convencional			
		Baixo	Médio	Alto	Total
Risco de cárie segundo o método simplificado	Baixo	θ_{11}	0	0	$\theta_{1.}$
	Médio	0	θ_{22}	0	$\theta_{2.}$
	Alto	0	0	θ_{33}	$\theta_{3.}$
Total	-	$\theta_{.1}$	$\theta_{.2}$	$\theta_{.3}$	$\theta_{..} = 1$

- Ou seja $\theta_{i.} = \theta_{.j} = \theta_{ij}, \forall i = j$ (concordância absoluta).
- Claramente, a independência é incompatível com essa situação.

Concordância marginal

		Risco de cárie segundo o método convencional			
		Baixo	Médio	Alto	Total
Risco de cárie segundo	Baixo	θ_{11}	θ_{12}	θ_{13}	θ_1
o método simplificado	Médio	θ_{21}	θ_{22}	θ_{23}	θ_2
	Alto	θ_{31}	θ_{32}	θ_{33}	θ_3
Total	-	θ_1	θ_2	θ_3	$\theta_{..} = 1$

- $\theta_{i.} = \theta_{.j} = \theta_i, \forall i = j$ (concordância marginal).
- Neste caso, não, necessariamente, tal situação é incompatível com a existência de independência.

Concordância marginal e independência

		Risco de cárie segundo o método convencional			
		Baixo	Médio	Alto	Total
Risco de cárie segundo o método simplificado	Baixo	θ_1^2	$\theta_1\theta_2$	$\theta_1\theta_3$	θ_1
	Médio	$\theta_1\theta_2$	θ_2^2	$\theta_2\theta_3$	θ_2
	Alto	$\theta_1\theta_3$	$\theta_2\theta_3$	θ_3^2	θ_3
Total	-	θ_1	θ_2	θ_3	$\theta_{..} = 1$

- $\theta_{i.} = \theta_{.j}, \forall i = j$ (concordância marginal) e $\theta_{ij} = \theta_{i.}\theta_{.j}, \forall i, j$ (independência).
- Note, entretanto, que a validade simultânea dos dois conjuntos de hipóteses restrige, por demais, a configuração da tabela.

Voltando à hipótese de independência

- Novamente, vamos comparar as frequências observadas com as esperadas sob a validade de H_0.
- Frequências observáveis: N_{ij} (variável aleatória).
- Frequências esperadas sob o modelo probabilístico (multinomial): $E(N_{ij}) = n_{..}\theta_{ij}, \forall i, j$.
- Frequências esperadas sob o modelo probabilístico e H_0: $E(N_{ij})^0 = n_{..}\theta_{i.}\theta_{.j}, \forall i, j$.

Teste de hipóteses

- Os estimadores de MV de $\theta_{ij}, i = 1, ..., r, j = 1, ..., s$, irrestritos e sob H_0 são, respectivamente,

$$\widehat{\theta}_{ij} = \frac{N_{ij}}{n_{..}}, \text{ e}$$

$$\widehat{\theta}_{ij}^{0} = \widehat{\theta}_{i.}\widehat{\theta}_{.j} = \frac{N_{i.}}{n_{.J}}\frac{N_{.j}}{n_{..}} = \frac{N_{i.}N_{.j}}{n_{..}^2}.$$

- As respectivas estimtivas, são dadas por:

$$\widetilde{\theta}_{ij} = \frac{n_{ij}}{n_{..}}, \text{ e}$$

$$\widetilde{\theta}_{ij}^{0} = \widetilde{\theta}_{i.}\widetilde{\theta}_{.j} = \frac{n_{i.}}{n_{..}}\frac{n_{.j}}{n_{..}} = \frac{n_{i.}n_{.j}}{n_{..}^2}.$$

Teste de hipóteses

- Frequências observadas: n_{ij} (número).
- Estimadores das frequências esperadas sob o modelo probabilístico e H_0: $E_{ij} = n_{..}\widehat{\theta}_{i.}\widehat{\theta}_{.j} = \frac{N_{i.}N_{.j}}{n_{..}}$.
- Estimativas das frequências esperadas sob o modelo probabilístico e H_0: $e_{ij} = n_{i.}\widetilde{\theta}_{i.}\widetilde{\theta}_{.j} = \frac{n_{i.}n_{.j}}{n_{..}}$.
- Estatística do teste $Q_H = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(N_{ij}-E_{ij})^2}{E_{ij}}$.
- Estatística (numérica) do teste $q_H = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{ij}-e_{ij})^2}{e_{ij}}$

Teste de hipóteses

- Quanto maior o valor de q_H, maior será a evidência contra H_0.
- Sob H_0, pode-se provar que $Q_H \approx \chi^2_{(r-1)(s-1)}$, para n suficientemente grande.
- Dessa forma, rejeita-se H_0 se $q_H \geq q_c$, em que $P(X \geq q_c|H_0)$, ou se $p = p - valor = P(X \geq q_H|H_0) < \alpha$, com $X \sim \chi^2_{(r-1)(s-1)}$.

Voltando ao Exemplo 1: frequências observadas e (esperadas)

		Risco de cárie segundo o método convencional		
		Baixo	Médio	Alto
Risco de cárie segundo o método simplificado	Baixo	11(4,45)	5 (8,58)	0(2,97)
	Médio	14 (15,31)	34 (29,48)	7 (10,21)
	Alto	2(7,24)	13 (13,94)	11(4,82)

Resultados

- Neste caso, $q_H = 27,65$ e $p - valor < 0,0001$. Assim, conclui-se que os métodos não classificam o risco de cárie de forma independente.
- Em geral, o testes aqui apresentados (homogeneidade e independência) devem ser utilizados quando $e_{ij} \geq 5$. O que não é caso.
- Alternativas: TRV, inferência bayesiana, calcular o p-valor através de métodos de reamostragem (bootstrap), modelos de regressão, teste exato de Fisher (no caso de tabelas 2×2).

Resíduos ajustados (por casela)

- Podemos ter uma idéia sobre a natureza da estrutura de dependência utilizandos os chamados resíduos ajustados (ou padronizados) sugeridos por Haberman (1973). Tal expediente também pode ser utilizado no exemplo anterior.
- Haberman (1973) sugere ao invés de utilizar-se $\frac{N_{ij}-E_{ij}}{\sqrt{E_{ij}}}$ (que seria uma escolha natural), utilizar-se

$$R_{ij} = \frac{N_{ij} - E_{ij}}{\sqrt{E_{ij}(1-\widehat{\theta}_{i.})(1-\widehat{\theta}_{.j})}},$$

devido à este último ter uma aproximação mais acurada para a distribuição normal (para $n_{..}$ suficientemente grande, sob H_0).

Resíduos ajustados (por casela)

- Assim, para cada célula, podemos calcular $r_{ij} = \frac{n_{ij} - e_{ij}}{\sqrt{e_{ij}(1-\widetilde{\theta}_{i.})(1-\widetilde{\theta}_{.j})}}$. Espera-se que se tivermos um pequeno número de caselas, sob H_0, quase todas apresentem $|r_{ij}| < 2$ e, se tivermos um número elevado, quase todas apresentem $|r_{ij}| < 3$.

Resíduos ajustados

		Risco de cárie segundo o método convencional		
		Baixo	Médio	Alto
Risco de cárie segundo o método simplificado	Baixo	3,99	-1,96	-2,08
	Médio	-0,59	1,86	-1,69
	Alto	-2,67	-0,43	3,64

Os resultados indicam classificação dependente para quase todas as caselas, em particular para as classificações concordantes.

Exemplo 4: a provadora de chá

- Em 1935, Sir Ronald A. Fisher, em seu livro *The Design of experiments*, descreveu o seguinte experimento.
- Os britânicos, em geral, gostam de tomar chá com leite.
- Ele convidou uma colega para distinguir se, na prepração do chá, leite ou o próprio chá foi primeiramente adicionado à xícara.
- Sua colega experimentou oito xícaras de chá, sendo que em quatro primeiramente o chá foi adicionado e nas outras quatro, primeiramente o leite.

Exemplo 4: a provadora de chá

- Para cada uma das xícaras, foi perguntado à ela o que primeiramente foi adicionado.
- Ela sabia que existiam quatro xícaras de cada tipo sem, no entanto, saber quais eram.
- Pergunta: existe uma concordância positiva entre a opnião dela e a verdadeira ordem?

Exemplo 4 (cont.)

- Tabela de contingência (2×2) com os resultados do experimento.

		Resposta da colega		
		leite	chá	Total
Primeiro ingrediente	leite	3	1	4
	chá	1	3	4
Total	-	4	4	8

Produto de binomiais (condicionalmente) independentes

- A tabela anterior é uma realização (amostra) possível, oriunda da seguinte estrutura:

		Resposta da colega		
		leite	chá	Total
Primeiro ingrediente	leite	$N_{11}(\theta_{11})$	$N_{12}(\theta_{12})$	$n_{1.} = 4$
	chá	$N_{21}(\theta_{21})$	$N_{22}(\theta_{22})$	$n_{2.} = 4$
Total	-	$n_{.1} = 4$	$n_{.2} = 4$	$n_{..} = 8$

Hipóteses de interesse e estatística do teste

- Hipóteses de interesse $H_0 : \theta_{11} = \theta_{21} = \theta$ vs $H_1 : \theta_{11} > \theta_{21}$.
- Os tamanhos amostrais são pequenos para se usar os testes assintóticos.
- A idéia de Fisher: calcular a probabilidade de ocorrência de todas as configurações possíveis da tabela de contingência original, com os totais (linha e coluna) originais preservados, que sejam pelo menos tão favoráveis à H_1, em relação à tabela observada, utilizando um mecanismo que não dependa de parâmetros (desconhecidos).

Estatística do teste

- Nesse caso, dado que os totais em cada linha e coluna são fixados, o conhecimento de n_{11} determina as outras três frequências.
- Temos, sob H_0, que: $N_{11} \sim \text{binomial}(n_{1.}, \theta)$ e $N_{21} \sim \text{binomial}(n_{2.}, \theta)$, em que $N_{11} \perp N_{21}$.
- Além disso, sob H_0, $Z = N_{11} + N_{21} \sim \text{binomial}(n_{1.} + n_{2.}, \theta)$.
- Assim, sob H_0, $N_{11}|Z = z \sim \text{hipergeométrica}(n_{..}, n_{1.}, z = n_{.1})$, como será demonstrado a seguir.

Demonstração

$$P(N_{11}=n_{11}|Z=z)=\frac{P(N_{11}=n_{11},Z=z)}{P(Z=z)}$$

$$=\frac{P(N_{11}=n_{11},N_{21}=z-n_{11})}{P(Z=z)}\overset{ind.}{=}\frac{P(N_{11}=n_{11})P(N_{21}=z-n_{11})}{P(Z=z)}$$

$$=\frac{\binom{n_{1.}}{n_{11}}\theta^{n_{11}}(1-\theta)^{n_{1.}-n_{11}}\binom{n_{2.}}{z-n_{11}}\theta^{z-n_{11}}(1-\theta)^{n_{2.}-(z-n_{11})}}{\binom{n_{1.}+n_{2.}}{z}\theta^{z}(1-\theta)^{n_{1.}+n_{2.}-z}}$$

$$\overset{n_{1.}+n_{2.}=n_{..}}{=}\frac{\binom{n_{1.}}{n_{11}}\binom{n_{..}-n_{1.}}{z-n_{11}}}{\binom{n_{..}}{z}}\mathbb{1}_A(n_{11})$$

em que A é o respectivo indicar da distribuição hipergeométrica.

Teste exato de Fisher

- Calcular o p-valor, com base na distribuição hipergeométrica, considerando a probabilidade de ocorrência de todas as possíveis tabelas que atendam às restrições anteriores.
- $H_0 : \theta_{11} = \theta_{21}$ vs $H_1 : \begin{cases} \theta_{11} > \theta_{21} & \text{ou} \\ \theta_{11} < \theta_{21} & \text{ou} \\ \theta_{11} \neq \theta_{21} \end{cases}$
- No nosso problema, $N_{11}|Z = z \sim$ hipergeométrica$(8, 4, 4)$.

Teste exato de Fisher: cálculo do p-valor

- Seja $n_{11} = a$
- Se $H_1 : \theta_{11} > \theta_{21}$, então $p - valor = P(N_{11} \geq a | N_{11} + N_{21} = z) = \sum_{k \geq a} \frac{\begin{pmatrix} n_{1.} \\ k \end{pmatrix} \begin{pmatrix} n_{..} - n_{1.} \\ z - k \end{pmatrix}}{\begin{pmatrix} n_{..} \\ z \end{pmatrix}}$

Teste exato de Fisher: cálculo do p-valor

- Se $H_1 : \theta_{11} < \theta_{21}$, então $p - valor = P(N_{11} \leq a | N_{11} + N_{21} = z) =$

$$\sum_{k \leq a} \frac{\begin{pmatrix} n_{1.} \\ k \end{pmatrix} \begin{pmatrix} n_{..} - n_{1.} \\ z - k \end{pmatrix}}{\begin{pmatrix} n_{..} \\ z \end{pmatrix}}$$

- Se $H_1 : \theta_{11} \neq \theta_{21}$, então p-valor $=$ soma das probabilidades menores ou iguais a probabilidade da tabela observada.

Voltando ao Exemplo 4 (cont.)

- A única tabela mais favorável à hipótese alternativa, sob as restições, é

		Resposta da colega		
		leite	chá	Total
Primeiro ingrediente	leite	4	0	4
	chá	0	4	4
Total	-	4	4	8

- Então temos que calcular $p - valor = P(N_{11} = 3|Z = 4) + P(N_{11} = 4|Z = 4) = 0,22864 + 0,0143 = 0,2428$, (respectivamente a tabela original e a tabela acima). Portanto, não temos evidências para rejeitar a independência.

Teste de McNemar

- Voltemos ao exemplo 3 (risco de cárie), desconsiderando a categoria "risco médio".

		Risco de cárie segundo o método convencional	
		Baixo	Alto
Risco de cárie segundo	Baixo	11	0
método simplificado	Alto	2	11

Modelo probabilístico (Multinomial)

		Risco de cárie segundo o método convencional		
		Baixo	Alto	Total
Risco de cárie segundo	Baixo	$N_{11}(\theta_{11})$	$N_{12}(\theta_{12})$	$N_{1.}(\theta_{1.})$
o método simplificado	Alto	$N_{21}(\theta_{21})$	$N_{22}(\theta_{22})$	$N_{2.}(\theta_{2.})$
Total	-	$N_{.1}(\theta_{.1})$	$N_{.2}(\theta_{.2})$	$n_{..}$

- Hipóteses de interesse: $H_0 : \theta_{i.} = \theta_{.j}, \forall i = j$ (concordância marginal) vs H_1 pelo menos uma diferença
- Nesse caso, basta testar $H_0 : \theta_{1.} = \theta_{.1}$ vs $H_1 : \theta_{1.} \neq \theta_{.1}$, pois $\theta_{1.} = 1 - \theta_{2.}$ e $\theta_{.1} = 1 - \theta_{.2}$.

Aspectos relevantes

- Par concordante: $N_{11} + N_{22}$.
- Par discordante: $N_{12} + N_{21}$.
- Par discordante do tipo 1: N_{12}.
- Par discordante do tipo 2: N_{21}.
- Note que

$$
\begin{aligned}
H_0 : \theta_{1.} = \theta_{.1} \quad & \Leftrightarrow \quad \theta_{11} + \theta_{12} = \theta_{11} + \theta_{21} \Leftrightarrow \theta_{12} + \theta_{12} = \theta_{21} + \theta_{12} \\
& \Leftrightarrow \quad \frac{\theta_{12}}{\theta_{12} + \theta_{21}} = \frac{1}{2} \Leftrightarrow \frac{n_{..}\theta_{12}}{n_{..}\theta_{12} + n_{..}\theta_{21}} = \frac{1}{2} \\
& \Leftrightarrow \quad \frac{\mu_{12}}{\mu_{12} + \mu_{21}} = \frac{1}{2} \Leftrightarrow \theta = \frac{1}{2}
\end{aligned}
$$

em que $\theta = \dfrac{\theta_{12}}{\theta_{12} + \theta_{21}} = \dfrac{\mu_{12}}{\mu_{12} + \mu_{21}}$ e $\mu_{ij} = n_{..}\theta_{ij}$ (valor esperado).

Aspectos relevantes

- Assim, as hipóteses anteriores podem ser reecritas como: $H_0 : \theta = 1/2$ vs $H_1 : \theta \neq 1/2$.
- Podemos considerar também $H_1 : \theta > 1/2$, ou seja, a proporção de classificados, simultaneamente, como risco baixo pelo método simplifcado e como risco alto pelo método convencional é maior do que que a proporção de classificados, simultaneamente, como risco alto pelo método simplificado e risco baixo pelo método convencional.
- Se $H_1 : \theta < 1/2$, teremos uma interpretação contrária ao item anterior.

Aspectos relevantes

- Suponha que $n_{..}$ não tenha sido fixado e, portanto $N_{ij} \sim \text{Poisson}(\mu_{ij})$ pode ser uma suposição razoável.
- Lembremos que N_{12} é o número de pares discordantes do tipo 1 e N_{21} é o número de pares discordantes do tipo 2.
- Além disso, temos que a distribuição de $N_{12} = n_{12} | N_{12} + N_{21} = s \sim \text{binomial}(s, \theta)$.
- Sob $H_0 : \theta = 1/2$, $P_{H_0}(N_{12} = n_{12} | N_{12} + N_{21} = s) = \left(\begin{array}{c} s \\ n_{12} \end{array} \right) \left(\frac{1}{2}\right)^s$

Demonstração do resultado

- Temos que

$$P(N_{12} = n_{12} | N_{12} + N_{21} = s) = \frac{P(N_{12} = n_{12}, N_{12} + N_{21} = s)}{P(N_{12} + N_{21} = s)}$$

$$= \frac{P(N_{12} = n_{12}, N_{21} = s - n_{12})}{P(N_{12} + N_{21} = s)} = \frac{P(N_{12} = n_{12})P(N_{21} = s - n_{12})}{P(N_{12} + N_{21} = s)}$$

$$= \frac{\frac{e^{-\mu_{12}}\mu_{12}^{n_{12}}}{n_{12}!}\frac{e^{-\mu_{21}}\mu_{21}^{s-n_{12}}}{(s-n_{12})!}}{\frac{e^{-(\mu_{12}+\mu_{21})}(\mu_{12}+\mu_{21})^{s}}{s!}}$$

$$= \frac{s!}{n_{12}!(s-n_{12})!}\left(\frac{\mu_{12}}{\mu_{12}+\mu_{21}}\right)^{n_{12}}\left(\frac{\mu_{21}}{\mu_{12}+\mu_{21}}\right)^{2-n_{12}}$$

$$= \frac{s!}{n_{12}!(s-n_{12})!}\theta^{n_{12}}(1-\theta)^{2-n_{12}}\mathbb{1}_{\{0,1,\ldots,s\}}(n_{12})$$

Teste de McNemar exato: p-valor

- Seja $n_{12} = a$ o número de pares discordantes do tipo 1.
- Se $H_1 : \theta < 1/2$, então

$$p - valor = P(N_{12} \leq a | N_{12} + N_{21} = s) = \left(\frac{1}{2}\right)^s \sum_{k=0}^{a} \left(\begin{array}{c} s \\ k \end{array} \right)$$

- Se $H_1 : \theta > 1/2$, então

$$p - valor = P(N_{12} \geq a | N_{12} + N_{21} = s) = \left(\frac{1}{2}\right)^s \sum_{k=a}^{s} \left(\begin{array}{c} s \\ k \end{array} \right)$$

- Se $H_1 : \theta \neq 1/2$, então

$$p - valor = 2\left(\frac{1}{2}\right)^s \min \left\{ \sum_{k=a}^{s} \left(\begin{array}{c} s \\ k \end{array} \right), \sum_{k=0}^{a} \left(\begin{array}{c} s \\ k \end{array} \right) \right\}$$

Teste de McNemar exato aplicado ao exemplo

- No nosso problema, s $= 4$, $n_{12} = 0$, $H_1 : \theta \neq 1/2$. Assim,

$$p - valor = 2\left(\frac{1}{2}\right)^2 \min\left\{\sum_{k=0}^{2}\left(\begin{array}{c} s \\ k \end{array}\right), \sum_{k=0}^{0}\left(\begin{array}{c} s \\ k \end{array}\right)\right\} = 0,5$$

- Logo, temos indícios de que existe concordância marginal entre os métodos de detecação de cáries.
- Exercício: repetir o procedimentos considerando as outras duas combinações de níveis de classificação (baixo, médio) e (médio,alto).

Teste de McNemar assintótico

- Seja $Y = N_{12}|N_{12} + N_{21} = s$. Já vimos que, sob H_0, $Y \sim \text{binomial}(s, 1/2)$.
- Defina $Z_t = \frac{Y - s/2}{\sqrt{s/4}}$. Para n suficientemente grande, $Z_t \approx N(0,1)$ e, consequentemente, $Q_t = Z_t^2 \approx \chi_1^2$.
- Podemos usar Q_t como estatística do teste. Neste caso, a estatística calculada é

  $q_t = \frac{(y - s/2)^2}{s/4} = \frac{(2y - s)^2/4}{s/4} = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}}$.

Teste de McNemar assintótico

- Assim, rejeitamos H_0 se $q_t \geq q_c$, em que $P(Q \geq q_c|H_0)$, ou, alternativamente, se $p - valor = P(Q \geq q_c|H_0) \leq \alpha$, $Q \sim \chi^2_1$, para $H_1 : \theta \neq 1/2$.
- Para as outras hipóteses alternativas, é melhor usar Z_t, da forma usual, ou seja, rejetar H_0 se $p - valor = P(Z \geq z_t|H_0) \leq \alpha$ ou $p - valor = P(Z \leq z_t|H_0) \leq \alpha$ para $H_1 : \theta > 1/2$ ou $H_1 : \theta < 1/2$, respectivamente, em que $Z \sim N(0,1)$ e z_t é o valor calculado da estatística Z_t.