

# Tabelas de contingência: modelos de regressão e testes de hipóteses (parte 2)

Prof. Caio Azevedo

# Voltando ao exemplo 1: comparação de métodos de detecção de cárie

		Risco de cárie segundo o método convencional			
		Baixo	Médio	Alto	Total
Risco de cárie segundo o método simplificado	Baixo	11 ( $\theta_{11}$ )	5( $\theta_{12}$ )	0( $\theta_{13}$ )	16( $\theta_{1.}$ )
	Médio	14( $\theta_{21}$ )	34( $\theta_{22}$ )	7( $\theta_{23}$ )	55( $\theta_{2.}$ )
	Alto	2( $\theta_{31}$ )	13( $\theta_{32}$ )	11( $\theta_{33}$ )	26( $\theta_{3.}$ )
Total	-	27( $\theta_{.1}$ )	52( $\theta_{.2}$ )	18( $\theta_{.3}$ )	$n_{..} = 97$

# Hipóteses de interesse

- Homogeneidade (simetria) marginal

$$H_0 : \begin{cases} \theta_{1.} = \theta_{.1} \\ \theta_{2.} = \theta_{.2} \end{cases} \leftrightarrow H_0 : \begin{cases} \theta_{12} + \theta_{13} - \theta_{21} - \theta_{31} = 0 \\ \theta_{21} + \theta_{23} - \theta_{12} - \theta_{32} = 0 \end{cases}$$

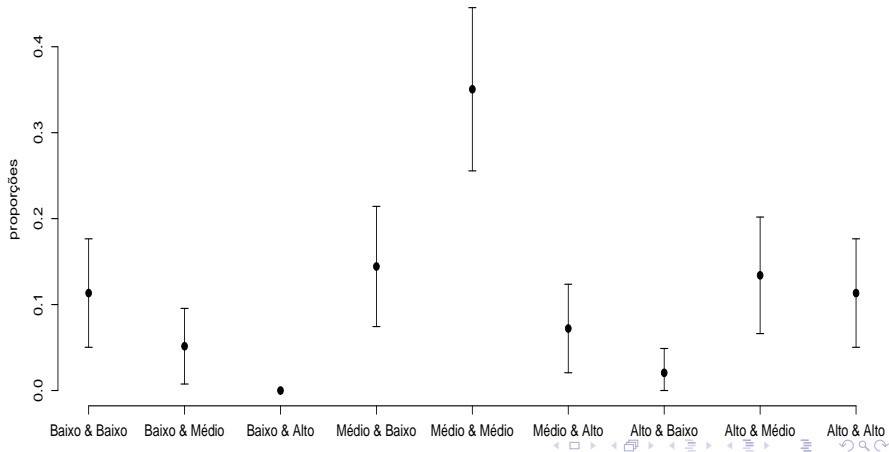
vs  $H_1$  : há pelo menos uma diferença

- Nesse caso  $\pi = (\theta_{11}, \theta_{12}, \theta_{13}, \theta_{21}, \theta_{22}, \theta_{23}, \theta_{31}, \theta_{32}, \theta_{33})'$  e

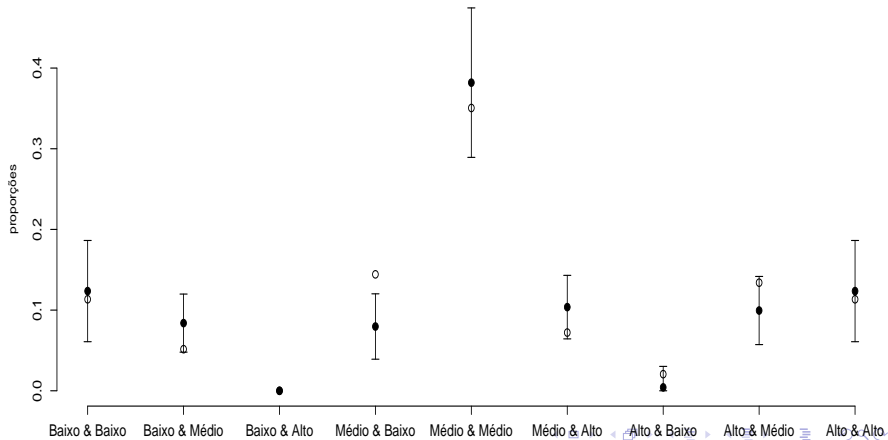
$$\theta = (\theta_{11}, \theta_{12}, \theta_{13}, \theta_{21}, \theta_{22}, \theta_{23}, \theta_{31}, \theta_{32})'$$

- Assim  $\mathbf{B} = \begin{bmatrix} 0 & 1 & 1 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}$ ;  $\mathbf{D} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

# Proporções (originais) estimadas



# Proporções estimadas sob $H_0$ (em preto) e observadas (em branco)



## Voltando ao Exemplo 7 ( fatores idade e duração da gestação)

idade	Dur. da gest.	N. de cigarros	Sobrevivência		Total
			Não	Sim	
<30	$\leq 260$	< 5	$50(\theta_{(11)11})$	$315(\theta_{(11)12})$	$n_{(11)1.} = 365$
		5+	$9(\theta_{(11)21})$	$40(\theta_{(11)22})$	$n_{(11)2.} = 49$
	>260	< 5	$24(\theta_{(12)11})$	$4012(\theta_{(12)12})$	$n_{(12)1.} = 4036$
		5+	$6(\theta_{(12)21})$	$459(\theta_{(12)22})$	$n_{(12)2.} = 465$
30+	$\leq 260$	< 5	$41(\theta_{(21)11})$	$147(\theta_{(21)12})$	$n_{(21)1.} = 188$
		5+	$4(\theta_{(21)21})$	$11(\theta_{(21)22})$	$n_{(21)2.} = 15$
	>260	<5	$14(\theta_{(22)11})$	$1594(\theta_{(22)12})$	$n_{(22)1.} = 1608$
		5 +	$1(\theta_{(22)21})$	$124(\theta_{(22)22})$	$n_{(22)2.} = 125$

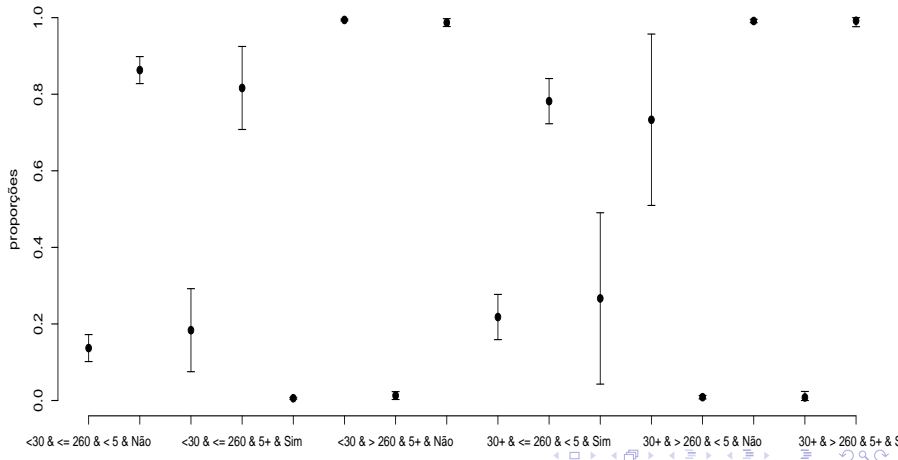
# Hipóteses de interesse

- Independência (sobrevivência e n. de cigarros) dentro de cada estrato (grupo: idade  $\times$  duração da gestação)

$$H_0 : \begin{cases} \theta_{(11)11} = \theta_{(11)21} \\ \theta_{(12)11} = \theta_{(12)21} \\ \theta_{(21)11} = \theta_{(21)21} \\ \theta_{(22)11} = \theta_{(22)21} \end{cases} \leftrightarrow \begin{cases} \theta_{(11)11} - \theta_{(11)21} = 0 \\ \theta_{(12)11} - \theta_{(12)21} = 0 \\ \theta_{(21)11} - \theta_{(21)21} = 0 \\ \theta_{(22)11} - \theta_{(22)21} = 0 \end{cases}$$

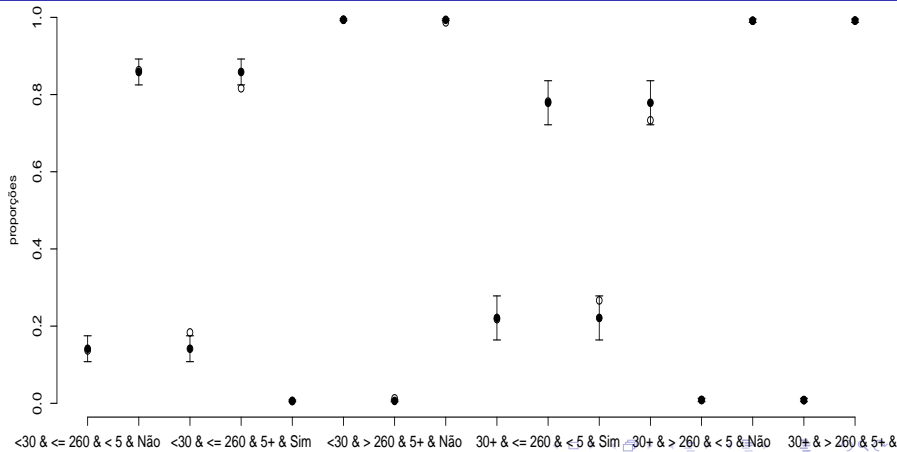
vs  $H_1$  : há pelo menos uma diferença

# Proporções (originais) estimadas





# Proporções estimadas sob $H_0$ (em preto) e observadas (em branco)



# Como estimar as proporções sob o modelo $\mathbf{A}\boldsymbol{\pi} = \mathbf{X}\boldsymbol{\beta}$ ?

- Como a matriz  $\mathbf{A}'\mathbf{A}$  quase sempre é não inversível, devemos encontrar uma matriz, digamos,  $\mathbf{H}$  e um vetor  $\mathbf{F}$ , de sorte que

$$\boldsymbol{\pi} = \mathbf{H}\boldsymbol{\beta} + \mathbf{F}.$$

- Assim, um estimador de  $\boldsymbol{\pi}$ , sob o modelo em questão, é dado por  $\hat{\boldsymbol{\pi}}_R = \mathbf{H}\hat{\boldsymbol{\beta}} + \mathbf{F}$ , em que  $\hat{\boldsymbol{\beta}}$  é o estimador MQG de  $\boldsymbol{\beta}$  (já visto anteriormente).

## Cont.

- Logo, para tamanhos amostrais suficientemente grandes,  
 $\hat{\pi}_R \approx N_q \left( \mathbf{H}\beta + \mathbf{F}, \mathbf{H} (\mathbf{X}'\boldsymbol{\Sigma}_F^{-1}\mathbf{X})^{-1} \mathbf{H}' \right)$ , em que  $\boldsymbol{\Sigma}_F$  é como vista anteriormente.
- Portanto, intervalos de confiança podem ser construídos com base na distribuição assintótica de  $\hat{\pi}_R$ , utilizando um estimador consistente para  $\boldsymbol{\Sigma}_F$  (estimador esse já visto anteriormente).

## Cont. (Exemplo 1)

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 1 \\ -1 & -1 & -2 & -1 & -1 & -2 \end{bmatrix}, \mathbf{F} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

## Cont. (exemplo 7)

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \mathbf{F} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

# Método delta univariado

- Considere uma tabela gerada por uma multinomial (os resultados valem, com as devidas adaptações, para o modelo produto de multinomiais.)
- Temos que o emv de cada  $\theta_{ij}$  é dado por  $\hat{\theta}_{ij} = \frac{N_{ij}}{n_{..}}$ .
- Para  $n_{..}$  suficientemente grande

$$\hat{\theta}_{ij} \approx N\left(\theta_{ij}, \frac{\theta_{ij}(1 - \theta_{ij})}{n_{..}}\right)$$

ou, de modo equivalente,

$$\sqrt{n_{..}} \left( \hat{\theta}_{ij} - \theta_{ij} \right) \xrightarrow[n_{..} \rightarrow \infty]{D} N(0, \theta_{ij}(1 - \theta_{ij}))$$

# Método delta univariado

- Defina  $\tau_{ij} = h(\theta_{ij}) = \theta_{ij}(1 - \theta_{ij})$ . Pelo princípio da invariância, temos que o emv de  $\tau_{ij}$  é  $\hat{\tau}_{ij} = \hat{\theta}_{ij}(1 - \hat{\theta}_{ij})$ .
- Qual a distribuição assintótica de  $\hat{\tau}_{ij}$ ?
- A obtenção pode ser feita através do método Delta (univariado).
- Seja  $\hat{\theta}$  uma variável aleatória de sorte que, para  $n$  suficientemente grande,

$$\hat{\theta} \approx N(\theta, \sigma^2).$$

# Aplicação

- Defina  $\hat{\tau} = g(\hat{\theta})$ . Então, para  $n$  suficientemente grande,

$$\hat{\tau} \approx N(g(\theta), \sigma^2 [\psi(\theta)]^2),$$

em que  $\psi(\theta) = \frac{d}{d\theta}g(\theta)$ .

- Nesse caso,  $g(\theta_{ij}) = \theta_{ij}(1 - \theta_{ij})$  e  $\psi(\theta_{ij}) = 1 - 2\theta_{ij}$ .
- Assim, para  $n_{..}$  suficientemente grande,

$$\hat{\theta}_{ij} \approx N\left(\theta_{ij}(1 - \theta_{ij}), \frac{\theta_{ij}(1 - \theta_{ij})(1 - 2\theta_{ij})^2}{n_{..}}\right)$$



# Método delta multivariado

- Seja  $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_p)'$  de sorte que, para  $n$  suficientemente grande

$$\hat{\boldsymbol{\theta}} \approx N_p(\boldsymbol{\theta}, \boldsymbol{\Sigma}).$$

- Defina  $\hat{\boldsymbol{\tau}} = (g_1(\hat{\boldsymbol{\theta}}), g_2(\hat{\boldsymbol{\theta}}), \dots, g_r(\hat{\boldsymbol{\theta}}))'$ , então, para  $n$  suficientemente grande,

$$\hat{\boldsymbol{\tau}} \approx N_r(\boldsymbol{\tau}, \boldsymbol{\Psi}\boldsymbol{\Sigma}\boldsymbol{\Psi}'),$$

em que  $\boldsymbol{\tau} = (g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}), \dots, g_r(\boldsymbol{\theta}))'$ .

# Método delta multivariado

- E

$$\Psi = \begin{bmatrix} \frac{\partial}{\partial \theta_1} g_1(\boldsymbol{\theta}) & \frac{\partial}{\partial \theta_2} g_1(\boldsymbol{\theta}) & \frac{\partial}{\partial \theta_3} g_1(\boldsymbol{\theta}) & \dots & \frac{\partial}{\partial \theta_p} g_1(\boldsymbol{\theta}) \\ \frac{\partial}{\partial \theta_1} g_2(\boldsymbol{\theta}) & \frac{\partial}{\partial \theta_2} g_2(\boldsymbol{\theta}) & \frac{\partial}{\partial \theta_3} g_2(\boldsymbol{\theta}) & \dots & \frac{\partial}{\partial \theta_p} g_2(\boldsymbol{\theta}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial \theta_1} g_r(\boldsymbol{\theta}) & \frac{\partial}{\partial \theta_2} g_r(\boldsymbol{\theta}) & \frac{\partial}{\partial \theta_3} g_r(\boldsymbol{\theta}) & \dots & \frac{\partial}{\partial \theta_p} g_r(\boldsymbol{\theta}) \end{bmatrix}.$$

# Aplicação

- Estimador do log da razão de chances (produto de multinomiais):

$$\pi = \frac{\lambda_1}{\lambda_2}, \lambda_i = \frac{\theta_{i1}}{1 - \theta_{i1}}, i = 1, 2, \eta = \ln \pi.$$

- Temos que o estimador de máxima verossimilhança de  $\eta$  é dado por

$$\hat{\eta} = \ln \hat{\pi} = \ln \left( \frac{\frac{\hat{\theta}_{11}}{1 - \hat{\theta}_{11}}}{\frac{\hat{\theta}_{21}}{1 - \hat{\theta}_{21}}} \right) = \ln \left( \frac{N_{11} N_{22}}{N_{12} N_{21}} \right) = \ln N_{11} + \ln N_{22} - \ln N_{12} - \ln N_{21},$$

em que  $\hat{\theta}_{i1} = \frac{N_{i1}}{n_i}$ ,  $i = 1, 2$ , devido à propriedade da invariância dos estimadores de MV.

# Aplicação

- Idéia: queremos obter a distribuição assintótica de  $\ln \hat{\pi} = \ln \hat{\lambda}_1 - \ln \hat{\lambda}_2$ , em que  $\ln \hat{\lambda}_i = \ln \frac{\hat{\theta}_{i1}}{1-\hat{\theta}_{i1}}, i = 1, 2$ .
- Como  $\hat{\lambda}_i, i = 1, 2$  são mutuamente independentes e são funções não-lineares de  $\hat{\theta}_i$ , aplica-se o método em cada um (separadamente) e depois utiliza-se a linearidade e independência de  $\ln \hat{\pi}$  em termos de  $\ln \hat{\lambda}_i, i = 1, 2$ .
- Nesse caso, para  $n_{i.}, i = 1, 2$  suficientemente grandes,  $\ln \hat{\lambda}_i \approx N\left(\ln \lambda_i, \frac{\theta_{i1}(1-\theta_{i1})}{n_{i.}} (\psi(\theta_{i.}))^2\right)$ , em que  $\psi(\theta_{i.}) = \frac{\partial}{\partial \theta_{i1}} \lambda_i$ .

## Aplicação

- Mas,  $\frac{\partial}{\partial \theta_{i1}} \lambda_i = \frac{1}{\theta_{i1}(1-\theta_{i1})}$ ,  $i = 1, 2$ .
- Logo,  $\sigma^2(\psi(\theta))^2 = \frac{\theta_{i1}(1-\theta_{i1})}{n_i} \frac{1}{\theta_{i1}^2(1-\theta_{i1})^2} = \frac{n_i}{\theta_{i1}(1-\theta_{i1})}$ .
- Assim, temos, para  $n_i$ ,  $i = 1, 2$  suficientemente grandes, que

$$\ln \hat{\lambda}_i = \ln \left( \frac{\hat{\theta}_{i1}}{1 - \hat{\theta}_{i1}} \right) \approx N_1 \left( \ln \lambda_i, \frac{1}{n_i \cdot \theta_{i1}(1 - \theta_{i1})} \right), i = 1, 2$$

- Portanto, como  $\hat{\lambda}_1$  e  $\hat{\lambda}_2$ , são mutuamente independentes,

$$\begin{aligned} \mathcal{V}(\hat{\eta}) &= \mathcal{V}(\ln \hat{\lambda}_1 - \ln \hat{\lambda}_2) = \frac{1}{n_{1.} \theta_{11}(1 - \theta_{11})} + \frac{1}{n_{2.} \theta_{21}(1 - \theta_{21})} \\ &= \frac{1}{n_{1.} \theta_{11}} + \frac{1}{n_{1.}(1 - \theta_{11})} + \frac{1}{n_{2.} \theta_{21}} + \frac{1}{n_{2.}(1 - \theta_{21})} \end{aligned}$$

# Aplicação

- Logo

$$\begin{aligned}\widetilde{\mathcal{V}}(\widehat{\eta}) = \sigma_{\eta}^2 &= \frac{1}{n_{1.}\widetilde{\theta}_{11}} + \frac{1}{n_{1.}(1-\widetilde{\theta}_{11})} + \frac{1}{n_{2.}\widetilde{\theta}_{21}} + \frac{1}{n_{2.}(1-\widetilde{\theta}_{21})} \\ &= \frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}\end{aligned}$$

- Portanto, como já tinha sido visto anteriormente, para  $n_{i.}, i = 1, 2$  suficientemente grandes,  $\widehat{\eta} \approx N(\eta, \sigma_{\eta}^2)$ .

# Mais sobre modelos de regressão

- Em geral, as hipóteses de interesse, relativas as tabelas de contingência geradas pelos modelos multinomial e produto de multinomiais, podem ser escritas como:

$$\mathbf{F}(\boldsymbol{\pi}) = \mathbf{X}\boldsymbol{\beta},$$

em que  $\mathbf{F}(\cdot)$  é uma matriz cujas componentes são funções de interesse.

- Exemplo

$$\ln(\mathbf{A}\boldsymbol{\pi}) = \mathbf{X}\boldsymbol{\beta},$$

nesse caso,  $\ln(\cdot)$  aplica o logaritmo natural em cada elemento do vetor  $\mathbf{A}\boldsymbol{\pi}$ .

# Modelos funcionais lineares

- Estudaremos modelos da forma

$$\mathbf{A}_{a \times g} \ln(\mathbf{G}_{g \times q} \boldsymbol{\pi}_{q \times 1}) = \mathbf{X}_{a \times p} \boldsymbol{\beta}_{p \times 1}$$

em que  $\mathbf{A}$  e  $\mathbf{G}$  são matrizes conhecidas.

- Para outros modelos, consultar referências e pesquisar.



## Exemplo 3 (duas inclinações partidárias)

- Homogeneidade entre as distribuições. Lembremos que

$$\eta = \ln \frac{\theta_{11}/(1 - \theta_{11})}{\theta_{21}/(1 - \theta_{21})} = \ln \frac{\theta_{11}/\theta_{12}}{\theta_{21}/\theta_{22}} \text{ e } \theta_{i1} + \theta_{i2} + \theta_{i3} = 1, i = 1, 2.$$

- $H_0 : \eta = 0$  vs  $H_1 : \eta \neq 0 \leftrightarrow H_0 : \ln \theta_{11} - \ln \theta_{12} = \ln \theta_{21} - \ln \theta_{22}$  vs  $H_1 : \ln \theta_{11} - \ln \theta_{12} \neq \ln \theta_{21} - \ln \theta_{22}$ .
- Nesse caso  $\boldsymbol{\pi} = (\theta_{11}, \theta_{12}, \theta_{21}, \theta_{22})'$  e  $\boldsymbol{\beta} = (\alpha)$ .
- Assim  $\mathbf{G} = \mathbf{I}_4$ , e  $\mathbf{G}\boldsymbol{\pi} = \boldsymbol{\pi} = (\theta_{11}, \theta_{12}, \theta_{21}, \theta_{22})'$ .

## Exemplo 3 (duas inclinações partidárias)

- $\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}; \mathbf{X} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$

- Modelo

$$\mathbf{A} \ln(\mathbf{G}\boldsymbol{\pi}) = \mathbf{X}\boldsymbol{\beta} \Rightarrow \begin{bmatrix} \ln \theta_{11} - \ln \theta_{12} \\ \ln \theta_{21} - \ln \theta_{22} \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha \end{bmatrix}$$

## Exemplo 3 (três inclinações partidárias)

- Homogeneidade entre as distribuições. Sejam

$$\eta_1 = \ln \frac{\theta_{11}/(1 - \theta_{11})}{\theta_{21}/(1 - \theta_{21})}, \eta_2 = \ln \frac{\theta_{12}/(1 - \theta_{12})}{\theta_{22}/(1 - \theta_{22})}$$

- $H_0 : \begin{cases} \eta_1 = 0 \\ \eta_2 = 0 \end{cases} \leftrightarrow H_0 :$

$$\begin{cases} \ln \theta_{11} - \ln(\theta_{12} + \theta_{13}) = \ln \theta_{21} - \ln(\theta_{22} + \theta_{23}) \\ \ln \theta_{12} - \ln(\theta_{11} + \theta_{13}) = \ln \theta_{22} - \ln(\theta_{21} + \theta_{23}) \end{cases} \text{ vs}$$

$H_1$  : há pelo menos uma diferença

- Nesse caso  $\pi = (\theta_{11}, \theta_{12}, \theta_{13}, \theta_{21}, \theta_{22}, \theta_{23})'$  e  $\beta = (\alpha_1, \alpha_2)'$ .

## Exemplo 3 (três inclinações partidárias)

■ Assim  $\mathbf{G} =$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}, \mathbf{G}\boldsymbol{\pi} = \begin{bmatrix} \theta_{11} \\ \theta_{12} + \theta_{13} \\ \theta_{21} \\ \theta_{22} + \theta_{23} \\ \theta_{12} \\ \theta_{11} + \theta_{13} \\ \theta_{22} \\ \theta_{21} + \theta_{23} \end{bmatrix}.$$

## Exemplo 3 (três inclinações partidárias)

■ Assim  $\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$ .

■  $\mathbf{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$ .

## Exemplo 3 (três inclinações partidárias)

- Modelo

$$\mathbf{A} \ln(\mathbf{G}\boldsymbol{\pi}) = \mathbf{X}\boldsymbol{\beta} \Rightarrow \begin{bmatrix} \ln \theta_{11} - \ln(\theta_{12} + \theta_{13}) \\ \ln \theta_{21} - \ln(\theta_{22} + \theta_{23}) \\ \ln \theta_{12} - \ln(\theta_{11} + \theta_{13}) \\ \ln \theta_{22} - \ln(\theta_{21} + \theta_{23}) \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_1 \\ \alpha_2 \\ \alpha_2 \end{bmatrix}$$

# Exemplo 1: comparação de métodos de detecção de cárie

- Independência entre os métodos de detecção de cárie

$H_0 : \theta_{ij} = \theta_i \cdot \theta_j, \forall i, j$  vs  $H_1 :$  há pelo menos uma diferença

- Nesse caso  $\pi = (\theta_{11}, \theta_{12}, \theta_{13}, \theta_{21}, \theta_{22}, \theta_{23}, \theta_{31}, \theta_{32}, \theta_{33})'$  e

$\beta = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8, \alpha_9)'$ .

- Note que  $H_0$  equivale à

$H_0 : \ln \theta_{ij} = \ln \theta_i + \ln \theta_j, \forall i, j$  vs  $H_1 :$  há pelo menos uma diferença

- Para cada casela, temos

$$H_0 : \left\{ \begin{array}{l} \ln \theta_{11} = \ln \theta_{1.} + \ln \theta_{.1} \\ \ln \theta_{12} = \ln \theta_{1.} + \ln \theta_{.2} \\ \ln \theta_{13} = \ln \theta_{1.} + \ln \theta_{.3} \\ \ln \theta_{21} = \ln \theta_{2.} + \ln \theta_{.1} \\ \ln \theta_{22} = \ln \theta_{2.} + \ln \theta_{.2} \\ \ln \theta_{23} = \ln \theta_{2.} + \ln \theta_{.3} \\ \ln \theta_{31} = \ln \theta_{3.} + \ln \theta_{.1} \\ \ln \theta_{32} = \ln \theta_{3.} + \ln \theta_{.2} \\ \ln \theta_{33} = \ln \theta_{3.} + \ln \theta_{.3} \end{array} \right.$$







■ Finalmente,

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

## ■ Modelo

$$\mathbf{A} \ln(\mathbf{G}\boldsymbol{\pi}) = \mathbf{X}\boldsymbol{\beta} \Rightarrow \begin{bmatrix} \ln \theta_{11} \\ \ln \theta_{12} \\ \ln \theta_{13} \\ \ln \theta_{21} \\ \ln \theta_{22} \\ \ln \theta_{23} \\ \ln \theta_{31} \\ \ln \theta_{32} \\ \ln \theta_{33} \\ \ln \theta_{1.} + \ln \theta_{.1} \\ \ln \theta_{1.} + \ln \theta_{.2} \\ \ln \theta_{1.} + \ln \theta_{.3} \\ \ln \theta_{2.} + \ln \theta_{.1} \\ \ln \theta_{2.} + \ln \theta_{.2} \\ \ln \theta_{2.} + \ln \theta_{.3} \\ \ln \theta_{3.} + \ln \theta_{.1} \\ \ln \theta_{3.} + \ln \theta_{.2} \\ \ln \theta_{3.} + \ln \theta_{.3} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \\ \alpha_7 \\ \alpha_8 \\ \alpha_9 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \\ \alpha_7 \\ \alpha_8 \\ \alpha_9 \end{bmatrix}$$

## Exemplo 7: efeitos de certos fatores na sobrevivência de recém nascidos (considerando apenas o fator idade)

- Independência (sobrevivência e n. de cigarros) dentro de cada estrato (grupo: idade)

$$H_0 : \begin{cases} \theta_{(1)11} = \theta_{(1)21} \\ \theta_{(2)11} = \theta_{(2)21} \end{cases} \Leftrightarrow \begin{cases} \eta^{(1)} = 0 \\ \eta^{(2)} = 0 \end{cases} \Leftrightarrow$$

$$\begin{cases} \ln \theta_{(1)11} - \ln \theta_{(1)12} = \ln \theta_{(1)21} - \ln \theta_{(1)22} \\ \ln \theta_{(2)11} - \ln \theta_{(2)12} = \ln \theta_{(2)21} - \ln \theta_{(2)22} \end{cases}$$

vs  $H_1$  : há pelo menos uma diferença,

## Exemplo 7: efeitos de certos fatores na sobrevivência de recém nascidos (considerando apenas o fator idade)

- em que  $\eta_{(i)} = \ln \frac{\theta_{(i)11}/(1 - \theta_{(i)11})}{\theta_{(i)21}/(1 - \theta_{(i)21})} = \ln \frac{\theta_{(i)11}/\theta_{(i)12}}{\theta_{(i)21}/\theta_{(i)22}}, i = 1, 2.$
- Nesse caso  $\boldsymbol{\pi} = (\theta_{(1)11}, \theta_{(1)12}, \theta_{(1)21}, \theta_{(1)22}, \theta_{(2)11}, \theta_{(2)12}, \theta_{(2)21}, \theta_{(2)22})'$  e  $\boldsymbol{\beta} = (\alpha_1, \alpha_2)'$ .
- Assim  $\mathbf{G} = \mathbf{I}_8, \mathbf{G}\boldsymbol{\pi} = \boldsymbol{\pi}.$

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

# Exemplo 7: efeitos de certos fatores na sobrevivência de recém nascidos (considerando apenas o fator idade)

## ■ Modelo

$$\mathbf{A} \ln(\mathbf{G}\boldsymbol{\pi}) = \mathbf{X}\boldsymbol{\beta} \Rightarrow \begin{bmatrix} \ln \theta_{(1)11} - \ln \theta_{(1)12} \\ \ln \theta_{(1)21} - \ln \theta_{(1)22} \\ \ln \theta_{(2)11} - \ln \theta_{(2)12} \\ \ln \theta_{(2)21} - \ln \theta_{(2)22} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_1 \\ \alpha_2 \\ \alpha_2 \end{bmatrix}$$

# Ajuste do modelo

- Mínimos quadrados generalizados (MQG): minimizar a seguinte forma quadrática em função de  $\beta$

$$\left(\widehat{\mathbf{F}} - \mathbf{X}\beta\right)' \boldsymbol{\Sigma}_{\mathbf{F}}^{-1} \left(\widehat{\mathbf{F}} - \mathbf{X}\beta\right),$$

em que  $\widehat{\mathbf{F}} = \mathbf{A} \ln(\mathbf{G}\widehat{\boldsymbol{\pi}})$ ,  $\boldsymbol{\Sigma}_{\mathbf{F}} = \text{Cov}(\widehat{\mathbf{F}})$ , a qual dependerá da forma funcional ( $\mathbf{F}(\boldsymbol{\pi})$ ) adotada e  $\widehat{\boldsymbol{\pi}}$  são as proporções amostrais.

- A obtenção de  $\boldsymbol{\Sigma}_{\mathbf{F}}$  dar-se-á pelo método Delta.



# Ajuste do modelo

- Estimador MQG (como visto anteriormente):

$$\hat{\beta}^* = (\mathbf{X}'\Sigma_{\mathbf{F}}^{-1}\mathbf{X})^{-1} \mathbf{X}'\Sigma_{\mathbf{F}}^{-1}\hat{\mathbf{F}}$$

- Como desconhecemos  $\Sigma_{\mathbf{F}}$ , a substituímos, na fórmula de  $\hat{\beta}^*$ , por um estimador consistente, denotado por  $\hat{\Sigma}_{\mathbf{F}}$ . Ou seja, trabalhamos com o estimador

$$\hat{\beta} = \left(\mathbf{X}'\hat{\Sigma}_{\mathbf{F}}^{-1}\mathbf{X}\right)^{-1} \mathbf{X}'\hat{\Sigma}_{\mathbf{F}}^{-1}\hat{\mathbf{F}},$$

em que, novamente,  $\hat{\Sigma}_{\mathbf{F}}$  dependerá da forma funcional  $(\mathbf{F}(\boldsymbol{\pi}))$  adotada.

# Ajuste do modelo

- Para tamanhos amostrais  $n_{..}$  (multinomial) ou  $n_i$  (produto de multinomiais),  $i = 1, \dots, r$  suficientemente grandes e sob  $H_0$  temos que

$$\hat{\beta} \approx N_p(\beta, (\mathbf{X}'\Sigma_{\mathbf{F}}^{-1}\mathbf{X})^{-1}).$$

- Esse resultado pode ser utilizado para construir intervalos de confiança para cada componente do vetor  $\beta$ , utilizando a estimativa de  $\Sigma_{\mathbf{F}}$ , obtida através de um estimador consistente.

# Ajuste do modelo

- O ajuste do modelo pode ser testado através da estatística

$$Q_F = \left( \widehat{\mathbf{F}} - \mathbf{X}\widehat{\boldsymbol{\beta}} \right)' \widehat{\boldsymbol{\Sigma}}_F^{-1} \left( \widehat{\mathbf{F}} - \mathbf{X}\widehat{\boldsymbol{\beta}} \right).$$

- Sob  $H_0$  e para tamanhos amostrais  $n_{..}$  (multinomial) ou  $n_i$  (produto de multinomiais),  $i = 1, \dots, r$  suficientemente grandes,  $Q_F \approx \chi^2_{(a-p)}$ , em que  $a$  é o número de linhas da matriz  $\mathbf{X}$  e  $p$  é o número de parâmetros do vetor  $\boldsymbol{\beta}$ .
- Assim, rejeitamos  $H_0$  se p-valor  $\leq \alpha$ , em que p-valor  $\approx P(X \geq q_F | H_0)$ ,  $X \sim \chi^2_{(a-p)}$  e  $q_F$  é o valor calculado da estatística  $Q_F$ .

# Ajuste do modelo

- Dado que o modelo está bem ajustado, podemos ainda testar hipóteses acerca do vetor  $\beta$ , da forma:

$$H_0 : \mathbf{C}_{(c \times p)} \beta_{(p \times 1)} = \mathbf{M}_{(c \times 1)} \text{ vs } H_1 : \mathbf{C}_{(c \times p)} \beta_{(p \times 1)} \neq \mathbf{M}_{(c \times 1)}$$

- Para tal, podemos considerar a estatística usual:

$$Q_C = (\mathbf{C}\hat{\beta} - \mathbf{M})' \left[ \mathbf{C}(\mathbf{X}'\hat{\Sigma}_F^{-1}\mathbf{X})^{-1}\mathbf{C}' \right]^{-1} (\mathbf{C}\hat{\beta} - \mathbf{M})$$

# Ajuste do modelo

- Sob  $H_0$  e para tamanhos amostrais  $n_{..}$  (multinomial) ou  $n_i$  (produto de multinomiais),  $i = 1, \dots, r$  suficientemente grandes,  $Q_C \approx \chi_C^2$ .
- Assim, rejeitamos  $H_0$  se p-valor  $\leq \alpha$ ,  $\approx P(X \geq q_C | H_0)$ ,  $X \sim \chi_C^2$  e  $q_C$  é o valor calculado da estatística  $Q_C$ .

# Ajuste do modelo

- Em nosso caso,  $\mathbf{F}(\boldsymbol{\pi}) = \mathbf{A} \ln(\mathbf{G}\boldsymbol{\pi})$ . Assim,  $\boldsymbol{\Sigma}_F = \boldsymbol{\Psi}\boldsymbol{\Sigma}_\pi\boldsymbol{\Psi}'$ , em que

$$\boldsymbol{\Psi} = \frac{\partial}{\partial \boldsymbol{\pi}} \mathbf{A} \ln(\mathbf{G}\boldsymbol{\pi}) = \mathbf{A}\mathbf{B}\mathbf{G},$$

e  $\mathbf{B}$  é uma matriz diagonal composta pelo inverso de cada um dos elementos do vetor  $\mathbf{G}\boldsymbol{\pi}$ .

- Logo,  $\widehat{\boldsymbol{\Sigma}}_F = \mathbf{A}\widehat{\mathbf{B}}\mathbf{G}\widehat{\boldsymbol{\Sigma}}_\pi\mathbf{G}'\widehat{\mathbf{B}}'\mathbf{A}'$ ,  $\widehat{\boldsymbol{\Sigma}}_\pi$  é como vista anteriormente e  $\widehat{\mathbf{B}}$  é uma matriz diagonal composta pelo inverso de cada um dos elementos do vetor  $\mathbf{G}\widehat{\boldsymbol{\pi}}$ .

## Aplicação da metodologia aos exemplos anteriores

- Exemplo 3 (duas inclinações partidárias):  $q_F = 29,38$ , p-valor  $< 0,0001$ . Assim, rejeita-se a hipótese de independência entre gênero e inclinação partidária.
- Exemplo 3 (três inclinações partidárias):  $q_F = 191,85$ , p-valor  $< 0,0001$ . Portanto, rejeita-se a hipótese de independência entre gênero e inclinação partidária.

## Aplicação da metodologia aos exemplos anteriores

- Exemplo 1 :  $q_F = 153,03$ ,  $p\text{-valor} < 0,0001$ . Assim, rejeita-se a hipótese de independência entre os métodos de classificação.
- Exemplo 7: (grupo= idade):  $q_F = 3,97$ ,  $p\text{-valor} = 0,1373$ . Portanto, não se rejeita a hipótese de independência entre n. de cigarros e sobrevivência do recém nascido para todos os estratos.



# Resumos dos dois modelos de regressão vistos

- Consideramos um, dos dois seguintes modelos:
  - $\mathbf{N} \sim \text{multinomial}_r(n_{\cdot}, \boldsymbol{\theta})$ ,  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_{(r-1)})'$  e  $\boldsymbol{\pi} = (\theta_1, \dots, \theta_r)'$ .
  - $\mathbf{N}_i \stackrel{\text{ind.}}{\sim} \text{multinomial}_s(n_{i\cdot}, \boldsymbol{\theta}_i)$ ,  $i = 1, 2, \dots, r$ ,  $\boldsymbol{\theta}_i = (\theta_{i1}, \dots, \theta_{i(r-1)})'$ ,  
 $\boldsymbol{\pi}_i = (\theta_{i1}, \dots, \theta_{ir})'$ ,  $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_r)'$  e  $\boldsymbol{\pi} = (\boldsymbol{\pi}_1, \dots, \boldsymbol{\pi}_r)'$ .
- Estruturas de regressão:
  - Linear:  $\mathbf{A}\boldsymbol{\pi} = \mathbf{X}\boldsymbol{\beta}$ .
  - Log-linear:  $\mathbf{A} \ln(\mathbf{G}\boldsymbol{\pi}) = \mathbf{X}\boldsymbol{\beta}$ .

## Resumos dos dois modelos de regressão vistos

- Estimador (MQG):;  $\hat{\beta} = \left( \mathbf{X}' \hat{\Sigma}_F^{-1} \mathbf{X} \right)^{-1} \mathbf{X}' \hat{\Sigma}_F^{-1} \hat{\mathbf{F}}$ .
- Estatística de ajuste do modelo:  $Q_F = \left( \hat{\mathbf{F}} - \mathbf{X} \hat{\beta} \right)' \hat{\Sigma}_F^{-1} \left( \hat{\mathbf{F}} - \mathbf{X} \hat{\beta} \right)$ .
- Estatística para testes de hipóteses do tipo

$H_0 : \mathbf{C}\beta = \mathbf{M}$  vs  $H_1 : \mathbf{C}\beta \neq \mathbf{M}$ :

$$Q_C = \left( \mathbf{C} \hat{\beta} - \mathbf{M} \right)' \left[ \mathbf{C} \left( \mathbf{X}' \hat{\Sigma}_F^{-1} \mathbf{X} \right)^{-1} \mathbf{C}' \right]^{-1} \left( \mathbf{C} \hat{\beta} - \mathbf{M} \right)$$

# Resumos dos dois modelos de regressão vistos

- Em que:

- Modelo linear:  $\hat{\mathbf{F}} = \mathbf{A}\hat{\boldsymbol{\pi}}$  e  $\hat{\boldsymbol{\Sigma}}_F = \mathbf{A}\hat{\boldsymbol{\Sigma}}_\pi\mathbf{A}'$ .

- Modelo log-linear:  $\hat{\mathbf{F}} = \mathbf{A} \ln(\mathbf{G}\hat{\boldsymbol{\pi}})$  e  $\hat{\boldsymbol{\Sigma}}_F = \hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Sigma}}_\pi\hat{\boldsymbol{\Psi}}'$ , em que  $\hat{\boldsymbol{\Psi}} = \mathbf{A}\hat{\mathbf{B}}\mathbf{G}$  e  $\hat{\mathbf{B}}$  é como definida anteriormente.

- Para  $n_{..}$  ou  $n_{i.}$ ,  $i = 1, \dots, r$  suficientemente grandes,

$$\hat{\boldsymbol{\beta}} \approx N_p(\boldsymbol{\beta}, (\mathbf{X}'\boldsymbol{\Sigma}_F^{-1}\mathbf{X})^{-1}), \text{ e, além disso, sob } H_0, Q_F \approx \chi_{(a-p)}^2 \text{ e}$$

$$Q_C \approx \chi_c^2.$$