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

Prof. Caio Azevedo

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

- Como a matriz $\boldsymbol{A}'\boldsymbol{A}$ quase sempre é não inversível (a matriz $\boldsymbol{A}$, e, geral, terá tem pelo menos uma coluna igual a zero), devemos encontrar uma matriz, digamos, $\boldsymbol{H}$ e um vetor $\boldsymbol{F}$, de sorte que (a matriz $\boldsymbol{H}$ guarda relação com a matriz $\boldsymbol{X}$ e o vetor $\boldsymbol{\beta}$, enquanto que o vetor $\boldsymbol{F}$ serve, em geral, para recuperar os parâmetros ausentes no produto $\boldsymbol{A\pi}$):

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

- Assim, um estimador de $\boldsymbol{\pi}$, sob o modelo em questão, é dado por $\widehat{\boldsymbol{\pi}}_R = \boldsymbol{H}\widehat{\boldsymbol{\beta}} + \boldsymbol{F}$, em que $\widehat{\boldsymbol{\beta}}$ é o estimador MQG de $\boldsymbol{\beta}$ (já visto anteriormente, slide 60 de
http://www.ime.unicamp.br/~cnaber/aula_tabelas_de_

Cont.

- Logo, para tamanhos amostrais suficientemente grandes (depende do modelo probabilístico adotado), $\widehat{\boldsymbol{\pi}}_R \approx N_q\left(\boldsymbol{H}\boldsymbol{\beta} + \boldsymbol{F}, \boldsymbol{H}\left(\boldsymbol{X}'\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{H}'\right)$, em que $\boldsymbol{\Sigma}_{\boldsymbol{F}}$ é como vista anteriormente.
- Portanto, intervalos de confiança podem ser construídos com base na distribuição assintótica de $\widehat{\boldsymbol{\pi}}_R$, utilizando um estimador consistente para $\boldsymbol{\Sigma}_{\boldsymbol{F}}$ (estimador esse já visto anteriormente, slide 59 de `http://www.ime.unicamp.br/~cnaber/aula_tabelas_de_contingencia_MDR_TH_1S_2017.pdf`).

Voltando ao Exemplo 3 (três inclinações partidárias)

		Inclinação partidária			
		Democrata	Independente	Republicano	Total
Gênero	Feminino	762 (θ_{11})	327(θ_{12})	468(θ_{13})	$n_{1.} = 1557$
	Masculino	484(θ_{21})	239(θ_{22})	477(θ_{23})	$n_{2.} = 1200$
Total	-	1246	566	945	$n_{..} = 2757$

Exemplo 3 (três inclinações partidárias)

- Homogeneidade entre as distribuições.
- $H_0 : \begin{cases} \theta_{11} = \theta_{21} \\ \theta_{12} = \theta_{22} \end{cases} \leftrightarrow H_0 : \begin{cases} \theta_{11} - \theta_{21} = 0 \\ \theta_{12} - \theta_{22} = 0 \end{cases}$ vs
  H_1 : há pelo menos uma diferença
- Nesse caso $\boldsymbol{\pi} = (\theta_{11}, \theta_{12}, \theta_{13}, \theta_{21}, \theta_{22}, \theta_{23})'$ e $\boldsymbol{\beta} = (\alpha_1, \alpha_2)'$.

Exemplo 3 (três inclinações partidárias)

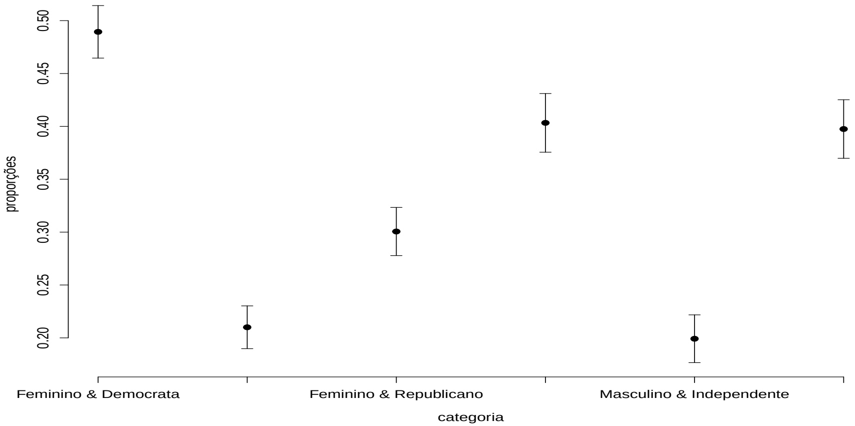
- Assim $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}; \boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$.
- Modelo

$$\boldsymbol{A}\boldsymbol{\pi} = \boldsymbol{X}\boldsymbol{\beta} \Rightarrow \begin{bmatrix} \theta_{11} \\ \theta_{21} \\ \theta_{12} \\ \theta_{22} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_1 \\ \alpha_2 \\ \alpha_2 \end{bmatrix}, \text{ assim } \theta_{i3} = 1 - \theta_{i1} - \theta_{i2}, i = 1, 2$$

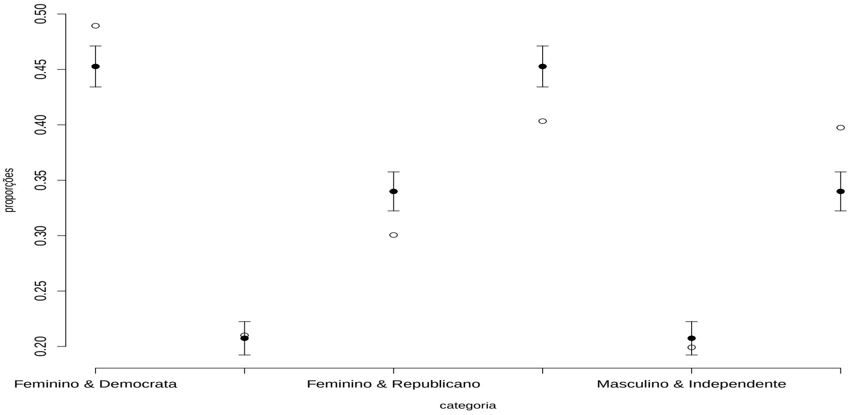
Cont. (Exemplo 1)

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

Proporções (originais) estimadas



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



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 (θ_{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 $\boldsymbol{\pi} = (\theta_{11}, \theta_{12}, \theta_{13}, \theta_{21}, \theta_{22}, \theta_{23}, \theta_{31}, \theta_{32}, \theta_{33})'$ e $\boldsymbol{\theta} = (\theta_{11}, \theta_{12}, \theta_{13}, \theta_{21}, \theta_{22}, \theta_{23}, \theta_{31}, \theta_{32})'$.
- Assim $\boldsymbol{B} = \begin{bmatrix} 0 & 1 & 1 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}$; $\boldsymbol{D} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Cont. (Exemplo 1)

Relembrando

$$\boldsymbol{A}\pi = \boldsymbol{X}\beta \Rightarrow \begin{bmatrix} \theta_{11} \\ \theta_{12} \\ \theta_{13} \\ \theta_{21} \\ \theta_{22} \\ \theta_{23} \\ \theta_{31} \\ \theta_{32} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \\ \alpha_2 + \alpha_3 - \alpha_4 \\ -\alpha_2 + \alpha_4 + \alpha_6 \end{bmatrix}.$$

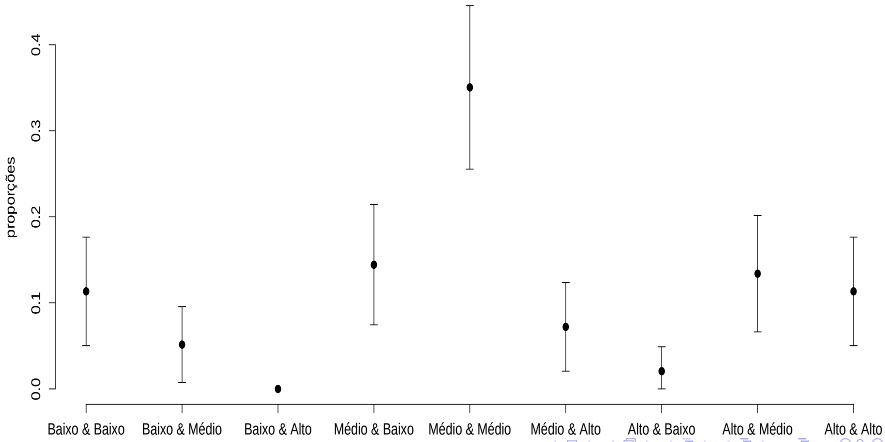
Assim, temos que $\theta_{33} = 1 - \theta_{11} - \theta_{12} - \theta_{13} - \theta_{21} - \theta_{22} - \theta_{23} - \theta_{31} - \theta_{32}$

$= 1 - \alpha_1 - \alpha_2 - 2\alpha_3 - \alpha_4 - \alpha_5 - 2\alpha_6$

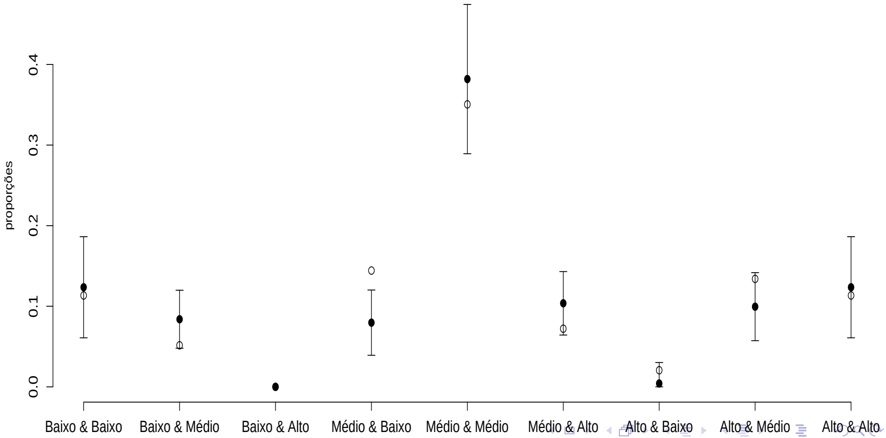
Cont. (Exemplo 1)

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

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	≤ 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	$24(\theta_{(12)11})$	$4012(\theta_{(12)12})$	$n_{(12)1.} = 4036$
		$5+$	$6(\theta_{(12)21})$	$459(\theta_{(12)22})$	$n_{(12)2.} = 465$
$30+$	≤ 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

Cont. (exemplo 7)

- Relembrando

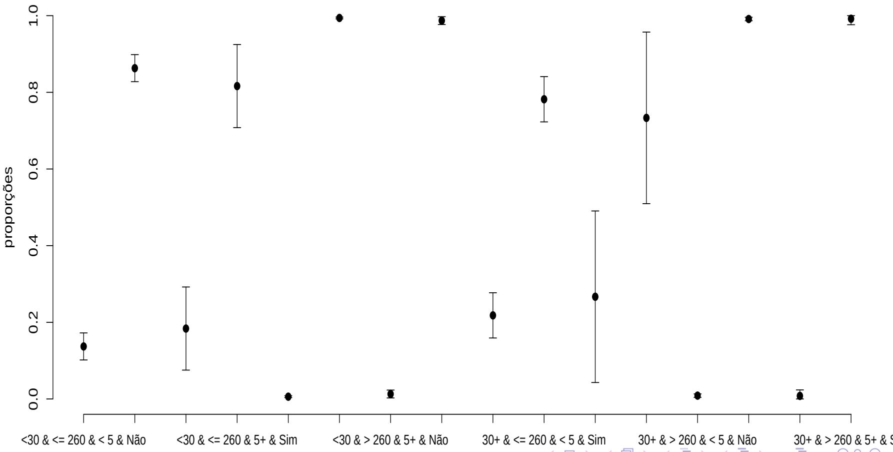
$$\boldsymbol{A\pi} = \boldsymbol{X\beta} \Rightarrow \begin{bmatrix} \theta_{(11)11} \\ \theta_{(11)21} \\ \theta_{(12)11} \\ \theta_{(12)21} \\ \theta_{(21)11} \\ \theta_{(21)21} \\ \theta_{(22)11} \\ \theta_{(22)21} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_1 \\ \alpha_2 \\ \alpha_2 \\ \alpha_3 \\ \alpha_3 \\ \alpha_4 \\ \alpha_4 \end{bmatrix}, \text{assim}, \theta_{(ij)k2} =$$

$1 - \theta_{(ij)k1}, \forall i, j, k \in \{1, 2\}$.

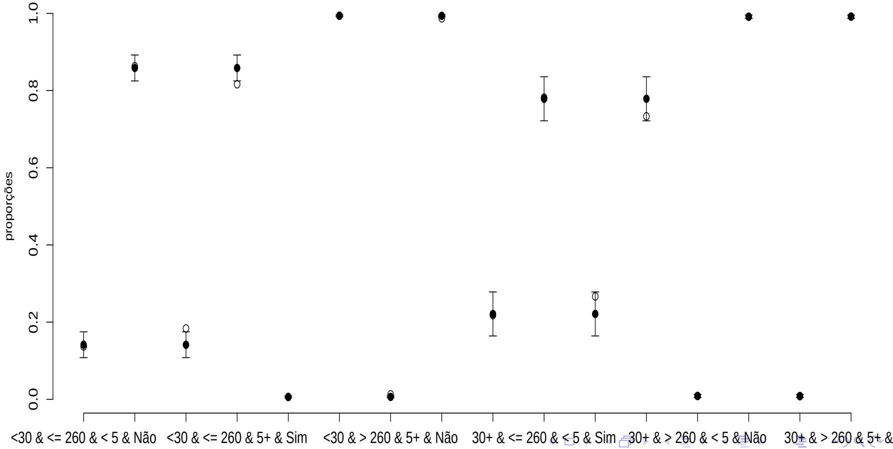
Cont. (exemplo 7)

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

Proporções (originais) estimadas



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



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:

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

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

- Exemplo

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

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

Modelos funcionais lineares

- Estudaremos modelos da forma

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

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

- Para outros modelos, consultar referências e pesquisar.

Exemplo 3 (duas inclinações partidárias)

- Homogeneidade entre as distribuições. Lembremos que $\eta = \ln \dfrac{\theta_{11}/(1-\theta_{11})}{\theta_{21}/(1-\theta_{21})} = \ln \dfrac{\theta_{11}/\theta_{12}}{\theta_{21}/\theta_{22}}$ 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 $\boldsymbol{G} = \boldsymbol{I}_4$, e $\boldsymbol{G}\boldsymbol{\pi} = \boldsymbol{\pi} = (\theta_{11}, \theta_{12}, \theta_{21}, \theta_{22})'$.

Exemplo 3 (duas inclinações partidárias)

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

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

Exemplo 3 (três inclinações partidárias)

- Homogeneidade entre as distribuições. Sejam

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

- $H_0 : \begin{cases} \eta_1 = 0 \\ \eta_2 = 0 \end{cases} \leftrightarrow H_0 :$
$$\begin{cases} \ln\theta_{11} - \ln(\theta_{12}+\theta_{13}) = \ln\theta_{21} - \ln(\theta_{22}+\theta_{23}) \\ \ln\theta_{12} - \ln(\theta_{11}+\theta_{13}) = \ln\theta_{22} - \ln(\theta_{21}+\theta_{23}) \end{cases} \quad \text{vs}$$
H_1 : há pelo menos uma diferença

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

Exemplo 3 (três inclinações partidárias)

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

Exemplo 3 (três inclinações partidárias)

- Assim $\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$.
- $\boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$.

Exemplo 3 (três inclinações partidárias)

- Modelo

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

Exemplo 1: comparação de métodos de detecção de cárie

- Independência entre os métodos de detecção de cárie

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

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

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

- 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

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

Exemplo 1: comparação de métodos de detecção de cárie

- Portanto,

$$\boldsymbol{G} = \begin{bmatrix} & & & & \boldsymbol{I}_9 & & & & \\ \mathbf{1}_3' & & & & \mathbf{0}_3' & & & & \mathbf{0}_3' \\ \mathbf{0}_3' & & & & \mathbf{1}_3' & & & & \mathbf{0}_3' \\ \mathbf{0}_3' & & & & \mathbf{0}_3' & & & & \mathbf{1}_3' \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{G}\boldsymbol{\pi} = \begin{bmatrix} \boldsymbol{\pi} \\ \theta_{1.} \\ \theta_{2.} \\ \theta_{3.} \\ \theta_{.1} \\ \theta_{.2} \\ \theta_{.3} \end{bmatrix}.$$

Exemplo 1: comparação de métodos de detecção de cárie

- Também,

$$
\boldsymbol{A} = \left[\begin{array}{ccccccccccccccc}
\multicolumn{9}{c}{\boldsymbol{I}_9} & \multicolumn{6}{c}{\mathbf{0}_{9\times 6}} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1
\end{array}\right]
$$

- Finalmente,

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

- Modelo

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

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 $\eta_{(i)} = \ln \dfrac{\theta_{(i)11}/(1-\theta_{(i)11})}{\theta_{(i)21}/(1-\theta_{(i)21})} = \ln \dfrac{\theta_{(i)11}/\theta_{(i)12}}{\theta_{(i)21}/\theta_{(i)22}}, i = 1, 2.$
- 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 $\boldsymbol{G} = \boldsymbol{I}_8$, $\boldsymbol{G}\boldsymbol{\pi} = \boldsymbol{\pi}$.
- $\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$

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

- Modelo

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

Ajuste do modelo

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

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

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

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

Ajuste do modelo

- Estimador MQG (como visto anteriormente):
  $\widehat{\boldsymbol{\beta}}^{*} = \left(\boldsymbol{X}'\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}\widehat{\boldsymbol{F}}$

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

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

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

Ajuste do modelo

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

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

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

Ajuste do modelo

- O ajuste do modelo pode ser testado através da estatística

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

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

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

- Para tal, podemos considerar a estatística usual:

$$Q_C = \left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right)' \left[\boldsymbol{C}(\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{F}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{C}'\right]^{-1} \left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right)$$

Ajuste do modelo

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

Ajuste do modelo

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

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

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

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

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-valor $< 0,0001$. Assim, rejeita-se a hipótese de independência entre os métodos de classificação.
- Exemplo 7: (grupo= idade): $q_F = 3,97$,p-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:
  - $\boldsymbol{N} \sim \text{multinomial}_r(n_{..}, \boldsymbol{\theta})$, $\boldsymbol{\theta} = (\theta_1, ..., \theta_{(r-1)})'$ e $\boldsymbol{\pi} = (\theta_1, ..., \theta_r)'$.
  - $\boldsymbol{N}_i \overset{ind.}{\sim} \text{multinomial}_s(n_{i.}, \boldsymbol{\theta}_i), i = 1, 2, ..., r$, $\boldsymbol{\theta}_i = (\theta_{i1}, ..., \theta_{i(s-1)})'$,

    $\boldsymbol{\pi}_i = (\theta_{i1}, ..., \theta_{is})'$, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', ..., \boldsymbol{\theta}_r')'$ e $\boldsymbol{\pi} = (\boldsymbol{\pi}_1', .., \boldsymbol{\pi}_r')'$.
- Estruturas de regressão:
  - Linear: $\boldsymbol{A\pi} = \boldsymbol{X\beta}$.
  - Log-linear: $\boldsymbol{A}\ln(\boldsymbol{G\pi}) = \boldsymbol{X\beta}$.

Resumos dos dois modelos de regressão vistos

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

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

  $Q_C = \left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right)'\left[\boldsymbol{C}(\boldsymbol{X}'\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{F}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{C}'\right]^{-1}\left(\boldsymbol{C}\widehat{\boldsymbol{\beta}} - \boldsymbol{M}\right)$

Resumos dos dois modelos de regressão vistos

- Em que:
  - Modelo linear: $\widehat{\boldsymbol{F}} = \boldsymbol{A}\widehat{\boldsymbol{\pi}}$ e $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{F}} = \boldsymbol{A}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\pi}}\boldsymbol{A}'$.
  - Modelo log-linear: $\widehat{\boldsymbol{F}} = \boldsymbol{A}\ln\left(\boldsymbol{G}\widehat{\boldsymbol{\pi}}\right)$ e $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{F}} = \widehat{\boldsymbol{\Psi}}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\pi}}\widehat{\boldsymbol{\Psi}}'$, em que $\widehat{\boldsymbol{\Psi}} = \boldsymbol{A}\widehat{\boldsymbol{B}}\boldsymbol{G}$ e $\widehat{\boldsymbol{B}}$ é como definida anteriormente.
- Para $n_{..}$ ou $n_{i.}, i = 1, ..., r$ suficientemente grandes, $\widehat{\beta} \approx N_p(\beta, \left(\boldsymbol{X}'\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}\boldsymbol{X}\right)^{-1})$, e, além disso, sob H_0, $Q_F \approx \chi^2_{(a-p)}$ e $Q_C \approx \chi^2_c$.