Tabelas de contingência: modelos de regressão e testes de hipóteses (parte 1)

Prof. Caio Azevedo

Tabela de contingência $r \times s$: multinomial

		Variável 1 (resposta)					
		C_{11}	C_{12}	...	$C_{1(s-1)}$	C_{1s}	Total
Variável 2	C_{21}	$N_{11}(\theta_{11})$	$N_{12}(\theta_{12})$	...	$N_{1(s-1)}(\theta_{1(s-1)})$	$N_{1s}(\theta_{1s})$	$N_{1.}$
(resposta)	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	-	$N_{.1}$	$N_{.2}$	$\ldots$	$N_{.(s-1)}$	$N_{.s}$	$n_{..}$

Somente o total geral é fixado.

Tabela de contingência $r \times s$: multinomial

- $\boldsymbol{N} = (N_{11}, N_{12}, ..., N_{rs-1})' \sim \text{Multinomial}_{r \times s}(n_{..}, \boldsymbol{\theta})$.
- $\boldsymbol{\theta}^M = (\theta_{11}, \theta_{12}, ..., \theta_{rs-1})'$ $((r \times s) - 1)$ parâmetros), $\theta_{ij} \in (0,1)$ e $\left(\sum_{i=1}^{r-1}\sum_{j=1}^{s}\theta_{ij} + \sum_{j=1}^{s-1}\theta_{rj}\right) < 1$.
- $\boldsymbol{\pi} = \boldsymbol{\pi}^M = (\theta_{11}, \theta_{12}, ..., \theta_{rs-1}, \theta_{rs})'$ ($r \times s$ parâmetros), $\sum_{i=1}^{r}\sum_{j=1}^{s}\theta_{ij} = 1$.

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	-	$N_{.1}$	$N_{.2}$	...	$N_{.(s-1)}$	$N_{.s}$	$n_{..}$

Os totais marginais por linha ou coluna são fixados.

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

- $\boldsymbol{N}_i = (N_{i1}, N_{i2}, ..., N_{i(s-1)})' \sim \text{multinomial}_s(n_{i.}, \boldsymbol{\theta}_i), i = 1, ..., r.$
- $\boldsymbol{\theta}_i = (\theta_{i1}, \theta_{i2}, ..., \theta_{i(s-1)})'$ ($(s-1)$ parâmetros), $\theta_{ij} \in (0,1)$ e $\sum_{j=1}^{s-1} \theta_{ij} < 1, i = 1, ..., r.$
- $\boldsymbol{\pi}_i = (\theta_{i1}, \theta_{i2}, ..., \theta_{is})'$ (s parâmetros), $\sum_{j=1}^{s} \theta_{ij} = 1, i = 1, ..., r.$
- Defina $\boldsymbol{\theta}^{PM} = (\boldsymbol{\theta}_1', ..., \boldsymbol{\theta}_r')'$ ($r \times (s-1)$ parâmetros) e $\boldsymbol{\pi} = \boldsymbol{\pi}^{PM} = (\boldsymbol{\pi}_1', ..., \boldsymbol{\pi}_r')'$ ($r \times s$ parâmetros).

Hipóteses de interesse

- Para ambos os modelos probabilísticos (tabelas) muitas das (mas não todas as) hipóteses de interesse podem se escritas nas formas matriciais:
  - Multinomial
    - $H_0 : \boldsymbol{B}_{b\times(r\times s)}\boldsymbol{\pi}^{M}_{(r\times s)\times 1} = \boldsymbol{D}_{b\times 1}$ vs $H_1 : \boldsymbol{B}\boldsymbol{\pi}^{M} \neq \boldsymbol{D}$ (1).
    - $H_0 : \boldsymbol{C}_{c\times(r\times s-1)}\boldsymbol{\theta}^{M}_{(r\times s-1)\times 1} = \boldsymbol{M}_{c\times 1}$ vs $H_1 : \boldsymbol{C}\boldsymbol{\theta}^{M} \neq \boldsymbol{M}$ (2).
  - Produto de multinomiais
    - $H_0 : \boldsymbol{B}_{b\times(r\times s)}\boldsymbol{\pi}^{PM}_{(r\times s)\times 1} = \boldsymbol{D}_{b\times 1}$ vs $H_1 : \boldsymbol{B}\boldsymbol{\pi}^{PM} \neq \boldsymbol{D}$ (1).
    - $H_0 : \boldsymbol{C}_{c\times(r\times(s-1))}\boldsymbol{\theta}^{PM}_{(r\times(s-1))\times 1} = \boldsymbol{M}_{c\times 1}$ vs $H_1 : \boldsymbol{C}\boldsymbol{\theta}^{PM} \neq \boldsymbol{M}$ (2).
  - Em que $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ e $\boldsymbol{M}$ são conhecidos e não aleatórios, sendo que as duas primeiras matrizes tem posto linha completo em que $b \leq (r \times s)$ e $c \leq (r \times (s-1))$.

Hipóteses de interesse

- Em geral, é mais fácil escrever as hipóteses na forma (1), enquanto que a forma (2) implica numa simplicidade um pouco maior em termos de ajuste do modelo e testes de hipóteses.
- Trabalharemos com a forma (1).
- Exercícios: obter os resultados para a segunda forma.
- Quando utilizarmos as notações $\boldsymbol{\pi}$ e $\boldsymbol{\theta}$ sem os superíndices, estaremos nos referindo a ambos os modelos ou à um modelo específico, definido no (pelo) contexto.

Exemplo 3 (duas inclinações partidárias)

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

Hipóteses de interesse

- Homogeneidade entre as distribuições.
- $H_0 : \theta_{11} = \theta_{21}$ vs $H_1 : \theta_{11} \neq \theta_{21} \leftrightarrow H_0 : \theta_{11} - \theta_{21} = 0$ vs $H_1 : \theta_{11} - \theta_{21} \neq 0$.
- Nesse caso $\boldsymbol{\pi} = (\theta_{11}, \theta_{12}, \theta_{21}, \theta_{22})'$ e $\boldsymbol{\theta} = (\theta_{11}, \theta_{21})'$.
- Assim $\boldsymbol{B} = \left[\begin{array}{cccc} 1 & 0 & -1 & 0 \end{array} \right]$; $\boldsymbol{D} = [0]$.

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	$n_{..} = 2757$

Hipóteses de interesse

- Homegeneidade 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 $\boldsymbol{\theta} = (\theta_{11}, \theta_{12}, \theta_{21}, \theta_{22})'$.
- Assim $\boldsymbol{B} = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \end{bmatrix}; \boldsymbol{D} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Hipóteses de interesse

- Caso se rejeite H_0, as (sub)hipóteses de interesse são:
  - $H_0 : \theta_{11} = \theta_{21}$ vs $H_1 : \theta_{11} \neq \theta_{21} \leftrightarrow H_0 : \theta_{11} - \theta_{21} = 0$ vs $H_1 : \theta_{11} - \theta_{21} \neq 0$.
  - $H_0 : \theta_{12} = \theta_{22}$ vs $H_1 : \theta_{12} \neq \theta_{22} \leftrightarrow H_0 : \theta_{12} - \theta_{22} = 0$ vs $H_1 : \theta_{12} - \theta_{22} \neq 0$.

Hipóteses de interesse

- Para cada (sub) hipótese temos, respectivamente, que:

$$\boldsymbol{B} = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}; \boldsymbol{D} = \begin{bmatrix} 0 \end{bmatrix}.$$

e

$$\boldsymbol{B} = \begin{bmatrix} 0 & 1 & 0 & 0 & -1 & 0 \end{bmatrix}; \boldsymbol{D} = \begin{bmatrix} 0 \end{bmatrix}.$$

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

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

idade	N. de cigarros	Sobrevivência		
		Não	Sim	Total
<30	< 5	74 ($\theta_{(1)11}$)	4327($\theta_{(1)12}$)	$n_{(1)1.} = 4401$
	5+	15($\theta_{(1)21}$)	499($\theta_{(1)22}$)	$n_{(1)2.} = 514$
30+	< 5	55($\theta_{(2)11}$)	1741($\theta_{(2)12}$)	$n_{(2)1.} = 1796$
	5+	5($\theta_{(2)21}$)	135($\theta_{(2)22}$)	$n_{(2)2.} = 140$

Hipóteses de interesse

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

$$H_0 : \begin{cases} \theta_{(1)11} = \theta_{(1)21} \\ \theta_{(2)11} = \theta_{(2)21} \end{cases} \leftrightarrow \begin{cases} \theta_{(1)11} - \theta_{(1)21} = 0 \\ \theta_{(2)11} - \theta_{(2)21} = 0 \end{cases}$$

vs H_1 : há pelo menos uma diferença

- 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 $\boldsymbol{\theta} = (\theta_{(1)11}, \theta_{(1)21}, \theta_{(2)11}, \theta_{(2)21})'$.
- Assim $\boldsymbol{B} = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}$; $\boldsymbol{D} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

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

Hipóteses de interesse

- Nesse caso

$$\boldsymbol{\pi} = (\theta_{(11)11}, \theta_{(11)12}, \theta_{(11)21}, \theta_{(11)22}, \theta_{(12)11}, \theta_{(12)12}, \theta_{(12)21}, \theta_{(12)22},$$

$$\theta_{(21)11}, \theta_{(21)12}, \theta_{(21)21}, \theta_{(21)22}, \theta_{(22)11}, \theta_{(22)12}, \theta_{(22)21}, \theta_{(22)22})' \text{ e}$$

$$\boldsymbol{\theta} = (\theta_{(11)11}, \theta_{(11)21}, \theta_{(12)11}, \theta_{(12)21}, \theta_{(21)11}, \theta_{(21)21}, \theta_{(22)11}, \theta_{(22)21})'.$$

- Assim

$$\boldsymbol{B} = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix};$$

$$\boldsymbol{D} = \begin{bmatrix} 0, 0, 0, 0 \end{bmatrix}'.$$

Testes de hipótese

- Como testar as hipóteses de interesse?
- Para cada vetor de parâmetros $\boldsymbol{\pi}$ e $\boldsymbol{\theta}$ e cada modelo probabilístico (multinomial e produto de multinomiais), os estimadores usuais são as proporções amostrais (frequência relativas):

$$
\begin{aligned}
\widehat{\boldsymbol{\pi}}^{M} &= [\widehat{\theta}_{ij}] = \left[\frac{N_{ij}}{n_{..}}\right] \text{(multinomial)} \\
\widehat{\boldsymbol{\pi}}^{PM} &= [\widehat{\theta}_{ij}] = \left[\frac{N_{ij}}{n_{i.}}\right] \text{(produto de multinomiais)}
\end{aligned}
$$

- Os estimadores de máxima verossmilhança coincidem com tais estimadores.

Testes de hipótese

- O vetor $\widehat{\boldsymbol{\theta}}$ corresponde ao vetor $\widehat{\boldsymbol{\pi}}$ sem algum(ns) elemento(s) (veja as definições anteriores).
- Trabalharemos com o vetor $\widehat{\boldsymbol{\pi}}$.
- Problema: (a matriz de) informação de Fisher para qualquer um dos dois modelos probabilísticos é singular (determinante igual à 0) se a cacularmos em função de $\boldsymbol{\pi}$ (mesmo se todos os θ_{ij} forem diferentes de 0).

Testes de hipótese

- Se a calcularmos em função de $\boldsymbol{\theta}$, ela será não singular (se todos os θ_{ij} forem diferentes de 0).
- O problema acima compromete a obtenção da distribuição assintótica dos estimadores de MV de $\boldsymbol{\pi}$.
- Portanto, não utilizaremos os resultados assintóticos associados aos estimadores de MV.
- Contudo, podemos usar outro resultado.

Exemplo

- Seja $N_1, ..., N_n \overset{i.i.d.}{\sim}$ Bernoulli(θ), então $\widehat{\theta} = \frac{\sum_{i=1}^{n} N_i}{n}$, logo para n suficientemente grande, $\widehat{\theta} \approx N\left(\theta, \frac{\theta(1-\theta)}{n}\right)$.
- Para $\widehat{\pi} = (\widehat{\theta}_1, \widehat{\theta}_2) = (\widehat{\theta}, 1 - \widehat{\theta})$, temos que ($\theta_1 = \theta$, $\theta_2 = 1 - \theta$)

$$\boldsymbol{I}(\boldsymbol{\pi}) = n \begin{bmatrix} \theta_1 & \theta_1\theta_2 \\ \theta_1\theta_2 & \theta_2 \end{bmatrix}$$

em que $det(\boldsymbol{I}(\boldsymbol{\pi})) = 0$.

TCL para vetores aleatórios

- Sejam $\boldsymbol{X}_1, \boldsymbol{X}_2, ... \overset{i.i.d.}{\sim} \boldsymbol{X}$, $\boldsymbol{X} = (X_1, ..., X_p)' \sim D_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, em que $D_p(.,.)$ significa qualquer distribuição p-variada tal que $\mathcal{E}(\boldsymbol{X}) = \boldsymbol{\mu}$, $Cov(\boldsymbol{X}) = \boldsymbol{\Sigma}$, ambos com todas as componentes finitas, então

$$\sqrt{n}\left(\overline{\boldsymbol{X}} - \boldsymbol{\mu}\right) \quad \underset{n\to\infty}{\xrightarrow{D}} \quad N_p(\mathbf{0}, \boldsymbol{\Sigma}) \tag{1}$$

em que

$\overline{\boldsymbol{X}} = (\overline{X}_1, ..., \overline{X}_p)' = \left(\frac{1}{n}\sum_{j=1}^{n} X_{j1}, ..., \frac{1}{n}\sum_{j=1}^{n} X_{jp}\right)'$

- Ou seja, para n suficientemente grande, $\overline{\boldsymbol{X}} \approx N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$.

Matriz de dados

Indivíduo	Variável 1	Variável 2	...	Variável p
1	X_{11}	X_{12}	$\dots$	X_{1p}
2	X_{21}	X_{22}	$\dots$	X_{2p}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	X_{n1}	X_{n2}	$\dots$	X_{np}

Multinomial

$$\widehat{\boldsymbol{\pi}}^{M} = \frac{1}{n_{..}} \begin{bmatrix} \sum_{k=1}^{n_{..}} N_{11k} \\ \sum_{k=1}^{n_{..}} N_{12k} \\ \vdots \\ \sum_{k=1}^{n_{..}} N_{rsk} \end{bmatrix} \tag{2}$$

em que $N_{ijk} \stackrel{i.i.d.}{\sim} \text{Bernoulli}(\theta_{ij}), k = 1, 2, ..., n_{..}$.

Produto de Multinomiais

$$\widehat{\boldsymbol{\pi}}^{PM} = \begin{bmatrix} \widehat{\boldsymbol{\pi}}_1^M \\ \widehat{\boldsymbol{\pi}}_2^M \\ \vdots \\ \widehat{\boldsymbol{\pi}}_r^M \end{bmatrix} \tag{3}$$

em que $\widehat{\boldsymbol{\pi}}_i^M = \frac{1}{n_{i.}}\left(\sum_{k=1}^{n_{i.}} N_{i1k}, \sum_{k=1}^{n_{i.}} N_{i2k}, ..., \sum_{k=1}^{n_{i.}} N_{isk}\right)'$, $i = 1, ..., r$.

Lembrando que $\widehat{\boldsymbol{\pi}}_i^M \perp \widehat{\boldsymbol{\pi}}_{i'}^M, \forall i \neq i'$.

Testes de hipótese

- Como cada componente do estimador $\widehat{\boldsymbol{\pi}}$ corresponde à uma média de variáveis aleatórias iid Bernoulli(θ_{ij}), utilizaremos os resultados apresentados em (1).
- Entretanto, mesmo assim, a matriz de variâncias e covariâncias de $\widehat{\boldsymbol{\pi}}$ é singular.
- Portanto, o estimador $\widehat{\boldsymbol{\pi}}$ converge(ria) em distribuição para uma distribuição normal multivariada singular.

Testes de hipótese

- Contudo, voltando ao conjunto de hipóteses $\boldsymbol{B\pi} = \boldsymbol{D}$, temos que, se a matriz $\boldsymbol{B}$ for tal que $Cov(\boldsymbol{B}\widehat{\boldsymbol{\pi}}) = \boldsymbol{B\Sigma_{\pi}B}'$ ($Cov(\widehat{\boldsymbol{\pi}}) = \boldsymbol{\Sigma_{\pi}}$) seja não singular, o resultado sobre convergência em distribuição, de $\boldsymbol{B}\widehat{\boldsymbol{\pi}}$, para uma normal multivariada não singular, é válido.
- Exemplos de matriz $\boldsymbol{B}$ que levam à matriz $Cov(\boldsymbol{B\pi})$ ser singular: $\boldsymbol{B} = \boldsymbol{I}$ ou se houver alguma dependência linear entre suas linhas.

Matriz de variâncias e covariâncias exata de π (modelo multinomial)

$$
\begin{aligned}
\boldsymbol{\Sigma}_{\boldsymbol{\pi}} &= Cov(\widehat{\boldsymbol{\pi}}) = \\
& \frac{1}{n_{..}} \begin{bmatrix}
\theta_{11}(1-\theta_{11}) & -\theta_{11}\theta_{12} & -\theta_{11}\theta_{13} & \ldots & -\theta_{11}\theta_{rs} \\
-\theta_{11}\theta_{12} & \theta_{12}(1-\theta_{12}) & -\theta_{12}\theta_{13} & \ldots & -\theta_{12}\theta_{rs} \\
-\theta_{11}\theta_{13} & -\theta_{12}\theta_{13} & \theta_{13}(1-\theta_{13}) & \ldots & -\theta_{13}\theta_{rs} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-\theta_{11}\theta_{rs} & -\theta_{12}\theta_{rs} & -\theta_{13}\theta_{rs} & \ldots & \theta_{rs}(1-\theta_{rs})
\end{bmatrix} \\
&= \frac{1}{n_{..}}\left[\text{diag}(\boldsymbol{\pi}) - \boldsymbol{\pi}\boldsymbol{\pi}'\right] = \frac{1}{n_{..}}\boldsymbol{\Sigma}
\end{aligned} \tag{4}
$$

Matriz de variâncias e covariâncias exata de π (modelo produto de multinomiais)

$$\boldsymbol{\Sigma}_{\boldsymbol{\pi}} = Cov(\widehat{\boldsymbol{\pi}}) = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\pi}_1} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\boldsymbol{\pi}_2} & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{\Sigma}_{\boldsymbol{\pi}_r} \end{bmatrix} = \begin{bmatrix} \frac{1}{n_{1.}}\boldsymbol{\Sigma}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \frac{1}{n_{2.}}\boldsymbol{\Sigma}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \frac{1}{n_{r.}}\boldsymbol{\Sigma}_r \end{bmatrix}$$

em que $\boldsymbol{\Sigma}_{\boldsymbol{\pi}_i} = \dfrac{1}{n_{i.}}\left[\mathrm{diag}(\boldsymbol{\pi}_i) - \boldsymbol{\pi}_i\boldsymbol{\pi}_i'\right] = \dfrac{1}{n_{i.}}\boldsymbol{\Sigma}_i, i = 1, 2, ..., r$ ou seja, é uma matriz semelhante à (4), de dimensão $s \times s$ dividida por $n_{i.}$.

Testes de hipótese

- Com efeito, sob a validade da não singularidade de $\boldsymbol{B\Sigma_{\pi}B}'$ e sob H_0, temos que, para o caso multinomial

$$\sqrt{n}\left(\boldsymbol{B}\widehat{\boldsymbol{\pi}}^{M} - \boldsymbol{D}\right) \xrightarrow[n\to\infty]{D} N_b(\boldsymbol{0}, \boldsymbol{B\Sigma B}') \quad (5)$$

em que b é o número de linhas da matriz $\boldsymbol{B}$. Já para o caso produto de multinomiais, temos que:

$$\left(\boldsymbol{B}\boldsymbol{D}_{\sqrt{n}}\widehat{\boldsymbol{\pi}}^{PM} - \boldsymbol{D}\right) \xrightarrow[n\to\infty]{D} N_b(\boldsymbol{0}, \boldsymbol{B\Sigma B}') \quad (6)$$

em que $\boldsymbol{D}_{\sqrt{n}} = \text{diag}(n_{1.}^{-1/2}\boldsymbol{I}_s, .., n_{r.}^{-1/2}\boldsymbol{I}_s)$, em que $\boldsymbol{I}_s$ é uma matriz identidade de ordem s.

Testes de hipótese

- Além disso, temos que

$$\widehat{\mathbf{\Sigma}} \quad \xrightarrow[n \to \infty]{P} \quad \mathbf{\Sigma} \tag{7}$$

em que $\widehat{\mathbf{\Sigma}}$ corresponde à matriz $\mathbf{\Sigma}$ com os parâmetros substituídos pelos respectivos estimadores (para os dois modelos).

- Portanto, para n (ou $n_i, i = 1, 2, ..., r$) suficientemente grande e sob H_0, $\boldsymbol{B}\widehat{\boldsymbol{\pi}}^{M} \approx N_b\left(\boldsymbol{D}, \boldsymbol{B}\boldsymbol{\Sigma}_{\boldsymbol{\pi}}\boldsymbol{B}'\right)$ e $\boldsymbol{B}\widehat{\boldsymbol{\pi}}^{PM} \approx N_b\left(\boldsymbol{D}, \boldsymbol{B}\boldsymbol{\Sigma}_{\boldsymbol{\pi}}\boldsymbol{B}'\right)$.

$\boldsymbol{\Sigma}$: modelo multinomial

$$
\begin{aligned}
\boldsymbol{\Sigma} &= \begin{bmatrix}
\theta_{11}\left(1-\theta_{11}\right) & -\theta_{11}\theta_{12} & -\theta_{11}\theta_{13} & \ldots & -\theta_{11}\theta_{rs} \\
-\theta_{11}\theta_{12} & \theta_{12}\left(1-\theta_{12}\right) & -\theta_{12}\theta_{13} & \ldots & -\theta_{12}\theta_{rs} \\
-\theta_{11}\theta_{13} & -\theta_{12}\theta_{13} & \theta_{13}\left(1-\theta_{13}\right) & \ldots & -\theta_{13}\theta_{rs} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-\theta_{11}\theta_{rs} & -\theta_{12}\theta_{rs} & -\theta_{13}\theta_{rs} & \ldots & \theta_{rs}\left(1-\theta_{rs}\right)
\end{bmatrix} \\
&= \text{diag}(\boldsymbol{\pi}) - \boldsymbol{\pi}\boldsymbol{\pi}' = n_{..}\boldsymbol{\Sigma}_{\boldsymbol{\pi}} \qquad (8)
\end{aligned}
$$

Assim, $\boldsymbol{\Sigma}_{\boldsymbol{\pi}} = \dfrac{1}{n_{..}}\boldsymbol{\Sigma}$

$\boldsymbol{\Sigma}$ produto de multinominais

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Sigma}_3 & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{\Sigma}_r \end{bmatrix}$$

em que $\boldsymbol{\Sigma}_i = n_{i.}\boldsymbol{\Sigma}_{\boldsymbol{\pi}_i}, i = 1, .., r$ é uma matriz como em (8) de dimensão $s \times s$. Assim, $\boldsymbol{\Sigma}_{\boldsymbol{\pi}_i} = \frac{1}{n_{i.}}\boldsymbol{\Sigma}_i$ e $\boldsymbol{\Sigma}_{\boldsymbol{\pi}} = \text{diag}\left(\boldsymbol{\Sigma}_{\boldsymbol{\pi}_1}, \boldsymbol{\Sigma}_{\boldsymbol{\pi}_2}, ..., \boldsymbol{\Sigma}_{\boldsymbol{\pi}_r}\right)$

$\widehat{\boldsymbol{\Sigma}}$ e $\widehat{\boldsymbol{\Sigma}}_{\pi}$ modelo multinomial

$$
\begin{aligned}
\widehat{\boldsymbol{\Sigma}} &= \begin{bmatrix}
\widehat{\theta}_{11}\left(1-\widehat{\theta}_{11}\right) & -\widehat{\theta}_{11}\widehat{\theta}_{12} & -\widehat{\theta}_{11}\widehat{\theta}_{13} & \ldots & -\widehat{\theta}_{11}\widehat{\theta}_{rs} \\
-\widehat{\theta}_{11}\widehat{\theta}_{12} & \widehat{\theta}_{12}\left(1-\widehat{\theta}_{12}\right) & -\widehat{\theta}_{12}\widehat{\theta}_{13} & \ldots & -\widehat{\theta}_{12}\widehat{\theta}_{rs} \\
-\widehat{\theta}_{11}\widehat{\theta}_{13} & -\widehat{\theta}_{12}\widehat{\theta}_{13} & \widehat{\theta}_{13}\left(1-\widehat{\theta}_{13}\right) & \ldots & -\widehat{\theta}_{13}\widehat{\theta}_{rs} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-\widehat{\theta}_{11}\widehat{\theta}_{rs} & -\widehat{\theta}_{12}\widehat{\theta}_{rs} & -\widehat{\theta}_{13}\widehat{\theta}_{rs} & \ldots & \widehat{\theta}_{rs}\left(1-\widehat{\theta}_{rs}\right)
\end{bmatrix} \\
&= \operatorname{diag}(\widehat{\boldsymbol{\pi}}) - \widehat{\boldsymbol{\pi}}\widehat{\boldsymbol{\pi}}' \qquad (9)
\end{aligned}
$$

e $\widehat{\boldsymbol{\Sigma}}_{\pi} = \dfrac{1}{n}\widehat{\boldsymbol{\Sigma}}$

$\widehat{\mathbf{\Sigma}}$ e $\widehat{\mathbf{\Sigma}}_{\pi}$ produto de multinominais

$$\widehat{\mathbf{\Sigma}} = \begin{bmatrix} \widehat{\mathbf{\Sigma}}_1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{\Sigma}}_2 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{\Sigma}}_3 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \vdots & \widehat{\mathbf{\Sigma}}_r \end{bmatrix}$$

em que $\widehat{\mathbf{\Sigma}}_i, i = 1, .., r$ é uma matriz como em (9) de dimensão $s \times s$.
Além disso, $\widehat{\mathbf{\Sigma}}_{\boldsymbol{\pi}_i} = \frac{1}{n_{i.}}\widehat{\mathbf{\Sigma}}_i$. Logo, $\widehat{\mathbf{\Sigma}}_{\boldsymbol{\pi}} = diag(\widehat{\mathbf{\Sigma}}_{\boldsymbol{\pi}_1}, \widehat{\mathbf{\Sigma}}_{\boldsymbol{\pi}_2}, ..., \widehat{\mathbf{\Sigma}}_{\boldsymbol{\pi}_r})$

Teste para Hipóteses de interesse

- Assim, para testar as hipóteses $H_0 : \boldsymbol{B\pi} = \boldsymbol{D}$ vs $\boldsymbol{B\pi} \neq \boldsymbol{D}$, qualquer que seja o modelo probabilístico, podemos utilizar a estatística

$$Q = (\boldsymbol{B}\widehat{\boldsymbol{\pi}} - \boldsymbol{D})' \left(\boldsymbol{B}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\pi}}\boldsymbol{B}'\right)^{-1} (\boldsymbol{B}\widehat{\boldsymbol{\pi}} - \boldsymbol{D})$$

  a qual, sob H_0 é tal que $Q \xrightarrow[n \to \infty]{D} \chi^2_b$, em que b é o número de linhas da matriz $\boldsymbol{B}$.

- Demostração: De (5), (6) e (7) e teorema de Slutsky (exercício).

Teste para Hipóteses de interesse

- Naturalmente, as linhas da matriz $\boldsymbol{B}$ (assim como da matriz $\boldsymbol{C}$) têm de ser linearmente independentes e definir hipóteses não equivalentes.
- Assim, rejeitamos H_0 se p-valor $\leq \alpha$, em que p-valor $\approx P(X \geq q|H_0)$, em que $X \sim \chi^2_b$ e q é o valor calculado da estatística Q.

Resumo

- Multinomial
  - $\widehat{\boldsymbol{\pi}} = [\widehat{\theta}_{ij}] = \left[\dfrac{N_{ij}}{n_{..}}\right]$
  - $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\pi}} = \dfrac{1}{n_{..}}\left[\text{diag}(\widehat{\boldsymbol{\pi}}) - \widehat{\boldsymbol{\pi}}\widehat{\boldsymbol{\pi}}'\right]$.
- Produto de Multinomiais
  - $\widehat{\boldsymbol{\pi}} = [\widehat{\theta}_{ij}] = \left[\dfrac{N_{ij}}{n_{i.}}\right]$
  - $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\pi}} = diag(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\pi}_1}, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\pi}_2}, ..., \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\pi}_r})$, em que $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\pi}_i} = \dfrac{1}{n_{i.}}\left[\text{diag}(\widehat{\boldsymbol{\pi}}_i) - \widehat{\boldsymbol{\pi}}_i\widehat{\boldsymbol{\pi}}_i'\right], i = 1, ..., r$.

Aplicação da metodologia aos exemplos anteriores

- Exemplo 3 (duas inclinações partidárias): $q = 29,72$, p-valor $< 0,0001$.
- Exemplo 3 (três inclinações partidárias): $q = 30,02$, p-valor $< 0,0001$.
- Exemplo 1 : $q = 8,69$, p-valor $= 0,0130$.
- Exemplo 7: (grupo= idade): $q = 2,70$,p-valor$= 0,2596$.
- Exemplo 7: (grupo= idade$\times$ duração da gestação): $q = 2,50$,p-valor$= 0,6450$.

Modelos de regressão lineares

- Todas as hipóteses (nulas) discutidas anteriormente (e outras) podem ser representadas através da seguinte estrutura de regressão:

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

em que $\boldsymbol{A}$ e $\boldsymbol{X}$ são matrizes de planejamento conhecidas e $\boldsymbol{\beta}$ é um vetor de parâmetros de interesse (de modo a representar as hipóteses a serem testedas e/ou o modelo a ser ajustado). O índice q depende do modelo probabilístico adotado (multinomial ou produto de multinomiais).

Exemplo 3 (duas inclinações partidárias)

- Homogeneidade entre as distribuições.
- $H_0 : \theta_{11} = \theta_{21}$ vs $H_1 : \theta_{11} \neq \theta_{21} \leftrightarrow H_0 : \theta_{11} - \theta_{21} = 0$ vs $H_1 : \theta_{11} - \theta_{21} \neq 0$.
- Nesse caso $\boldsymbol{\pi} = (\theta_{11}, \theta_{12}, \theta_{21}, \theta_{22})'$ e $\boldsymbol{\beta} = (\alpha)$.
- Assim $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \boldsymbol{X} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
- Modelo

$$\boldsymbol{A}\pi = \boldsymbol{X}\beta \Rightarrow \begin{bmatrix} \theta_{11} \\ \theta_{21} \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha \end{bmatrix}$$

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

$$\boldsymbol{A}\pi = \boldsymbol{X}\beta \Rightarrow \begin{bmatrix} \theta_{11} \\ \theta_{21} \\ \theta_{12} \\ \theta_{22} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_1 \\ \alpha_2 \\ \alpha_2 \end{bmatrix}$$

Exemplo 1: comparação de métodos de detecção de cárie

- 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 $\boldsymbol{\beta} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6)'$.
- Note que H_0 equivale à

$$H_0 : \begin{cases} \theta_{31} = \theta_{12} + \theta_{13} - \theta_{21} \\ \theta_{32} = \theta_{21} + \theta_{23} - \theta_{12} \end{cases}$$

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

- Além disso,

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

- Modelo

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

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)

$$H_0 : \begin{cases} \theta_{(1)11} = \theta_{(1)21} \\ \theta_{(2)11} = \theta_{(2)21} \end{cases} \leftrightarrow \begin{cases} \theta_{(1)11} - \theta_{(1)21} = 0 \\ \theta_{(2)11} - \theta_{(2)21} = 0 \end{cases}$$

vs H_1 : há pelo menos uma diferença

- 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 $\boldsymbol{\beta} = (\alpha_1, \alpha_2)'$.

Exemplo 7

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

  e

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

Exemplo 7

- Modelo

$$\boldsymbol{A}\boldsymbol{\pi} = \boldsymbol{X}\boldsymbol{\beta} \Rightarrow \begin{bmatrix} \theta_{(1)11} \\ \theta_{(1)21} \\ \theta_{(2)11} \\ \theta_{(1)21} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_1 \\ \alpha_2 \\ \alpha_2 \end{bmatrix}.$$

Exemplo 7: efeitos de certos fatores na sobrevivência de recém nascidos (os fatores idade e dur. da gest.)

- 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

Exemplo 7

- Nesse caso

$$\boldsymbol{\pi} = (\theta_{(11)11}, \theta_{(11)12}, \theta_{(11)21}, \theta_{(11)22}, \theta_{(12)11}, \theta_{(12)12}, \theta_{(12)21}, \theta_{(12)22},$$

$$\theta_{(21)11}, \theta_{(21)12}, \theta_{(21)21}, \theta_{(21)22}, \theta_{(22)11}, \theta_{(22)12}, \theta_{(22)21}, \theta_{(22)22})'$$

e $\boldsymbol{\beta} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)'$.

Exemplo 7

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

Exemplo 7

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

Exemplo 7

- Modelo

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

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 $\widehat{\boldsymbol{F}} = \boldsymbol{A}\widehat{\pi}$, $\boldsymbol{\Sigma}_{\boldsymbol{F}} = \text{Cov}\left(\widehat{\boldsymbol{F}}\right) = \boldsymbol{A}\boldsymbol{\Sigma}_{\pi}\boldsymbol{A}'$, $\widehat{\pi}$ são as proporções amostrais e $\boldsymbol{\Sigma}_{\boldsymbol{\pi}}$ a matriz de covariâncias de $\boldsymbol{\pi}$ calculadas conforme o modelo probabilístico (multinomial ou produto de multinomiais).

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

Ajuste do modelo

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

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

em que $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\pi}}$ depende do modelo probabilistico (como visto anteriormente).

Ajuste do modelo

- Para tamanhos amostrais $n_{..}$ (multinomial) ou $n_{i.}$ (prroduto de multinomiais), $i = 1, ..., r$ suficientemente grandes e sob H_0 temos que $\widehat{\boldsymbol{\beta}} \approx N_p(\beta, \left(\boldsymbol{X}'\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}\boldsymbol{X}\right)^{-1})$.
- Esse resultado pode ser utilizado para constuir intervalos de confiança para cada componente do vetor $\boldsymbol{\beta}$, utilizando a estimativa de $\boldsymbol{\Sigma}_{\boldsymbol{F}}$.

Ajuste do modelo

- O ajuste do modelo pode ser testado através da estatística

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

- 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 $\approx P(X \geq q_F | H_0)$, $X \sim \chi^2_{(a-p)}$ 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 $\leq \alpha$, $p - valor \approx P(X \geq q_C | H_0)$, $X \sim \chi^2_c$ e q_C é o valor calculado da estatística Q_C.

Aplicação da metodologia aos exemplos anteriores

- Exemplo 3 (duas inclinações partidárias): $q_F = 29,72$, p-valor $< 0,0001$. Estimativa dos parâmetros (erro-padrão) : $0,57(0,01)$. Assim, rejeita-se a hipótese de independência entre gênero e inclinação partidária.
- Exemplo 3 (três inclinações partidárias): $q_F = 30,02$, p-valor $< 0,0001$. Estimativa dos parâmetros (erro-padrão) : $0,45(0,01); 0,21(0,01)$. Assim, rejeita-se a hipótese de independência entre gênero e inclinação partidária.

Aplicação da metodologia aos exemplos anteriores

- Exemplo 1 : $q_F = 8,69$, p-valor $= 0,0130$.

Parâmetro	Estimativa	EP
α_1	0,123	0,032
α_2	0,084	0,019
α_3	0,001	0,003
α_4	0,081	0,021
α_5	0,381	0,047
α_6	0,103	0,020

Aplicação da metodologia aos exemplos anteriores

- Exemplo 7: (grupo= idade): $q_F = 2,70$,p-valor$= 0,2596$. Portanto, não se rejeita a hipótese de independência entre n. de cigarros e sobrevivência do recém nascido para todos os estratos. Estimativa dos parâmetros (erro-padrão) : $0,017(0,002); 0,03(0,004)$.
- Exemplo 7: (grupo= idade$\times$ duração da gestação): $q_F = 2,50$,p-valor$= 0,6450$. Portanto, não se rejeita a hipótese de independência entre n. de cigarros e sobrevivência do recém nascido para todos os estratos. Estimativa dos parâmetros (erro-padrão) : $0,142(0,017); 0,006(0,001); 0,221(0,029); 0,009(0,002)$.

Redução de modelos (teste $\boldsymbol{C}\boldsymbol{\beta} = \boldsymbol{M}$)

- Para fins de ilustração.
- Exemplo 3, com as três inclinações partidárias: $H_0 : \alpha_1 = \alpha_2$ vs $H_1 : \alpha_1 \neq \alpha_2 \Leftrightarrow H_0 : \alpha_1 - \alpha_2 = 0$ vs $H_1 : \alpha_1 - \alpha_2 \neq 0$.
- Ou seja, equivale a testar se as quatro probabilidades: de ser democrata e republicanos são iguais entre si e entre os sexos.
- $\boldsymbol{C} = [1 - 1]$ e $\boldsymbol{M} = 0$, qc $= 279{,}54$, $p < 0,0001$.

Redução de modelos (teste $C\beta = M$)

- Exemplo 1.
- $H_0 : \begin{cases} \alpha_1 = \alpha_2 \\ \alpha_1 = \alpha_4 \\ \alpha_1 = \alpha_6 \end{cases} \leftrightarrow H_0 : \begin{cases} \alpha_1 - \alpha_2 = 0 \\ \alpha_1 - \alpha_4 = 0 \\ \alpha_1 - \alpha_6 = 0 \end{cases}$ vs

  H_1 : há pelo menos uma diferença
- Equivale a testar se as probabilidades de classificação baixo-baixo, baixo-médio, médio-baixo, e médio-alto são iguais entre si.
- Nesse caso $C = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$ e $M = [0 \;\; 0 \;\; 0]'$, qc

  $= 1{,}20$ e p $= 0{,}7522$.