

Análise de dados longitudinais: modelos, inferência e exemplos

Julio M Singer

Departamento de Estatística
Universidade de São Paulo
www.ime.usp.br/~jmsinger

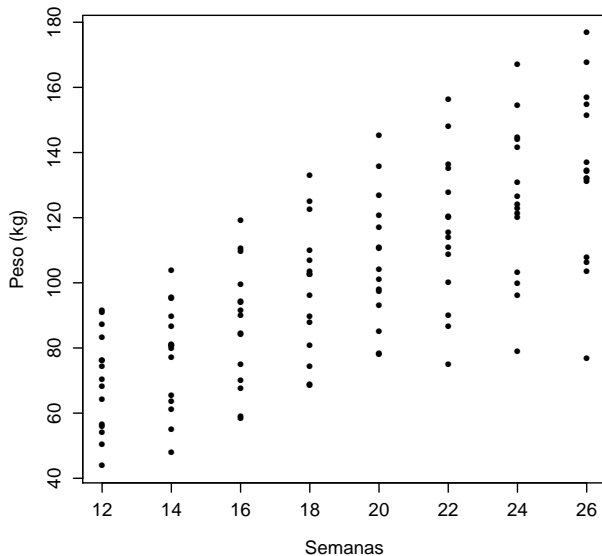
- Modelos de regressão para dados não-correlacionados e correlacionados
- Exemplos de dados longitudinais
- Modelos para dados gaussianos
- Modelos para dados não-gaussianos
- Exemplo comparativo
- O que ficou faltando

Exemplo EMBRAPA 1

Peso de bezerros (kg)

Semanas após nascimento							
12	14	16	18	20	22	24	26
54.1	65.4	75.1	87.9	98.0	108.7	124.2	131.3
91.7	104.0	119.2	133.1	145.4	156.5	167.2	176.8
64.2	81.0	91.5	106.9	117.1	127.7	144.2	154.9
70.3	80.0	90.0	102.6	101.2	120.4	130.9	137.1
68.3	77.2	84.2	96.2	104.1	114.0	123.0	132.0
43.9	48.1	58.3	68.6	78.5	86.8	99.9	106.2
87.4	95.4	110.5	122.5	127.0	136.3	144.8	151.5
74.5	86.8	94.4	103.6	110.7	120.0	126.7	132.2
50.5	55.0	59.1	68.9	78.2	75.1	79.0	77.0
91.0	95.5	109.8	124.9	135.9	148.0	154.5	167.6
83.3	89.7	99.7	110.0	120.8	135.1	141.5	157.0
76.3	80.8	94.2	102.6	111.0	115.6	121.4	134.5
55.9	61.1	67.7	80.9	93.0	100.1	103.2	108.0
76.1	81.1	84.6	89.8	97.4	111.0	120.2	134.2
56.6	63.7	70.1	74.4	85.1	90.2	96.1	103.6

Gráfico de dispersão para os dados da EMBRAPA

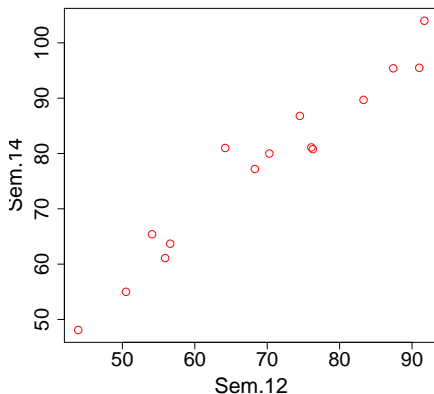


Exemplo EMBRAPA 3

- Regressão linear simples:

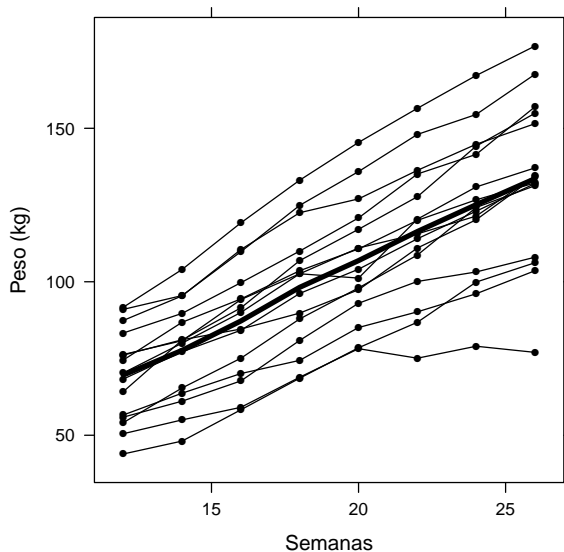
$$y_{ij} = \alpha + \beta x_j + e_{ij}$$

- $i = 1, \dots, 15, j = 1, \dots, 8$
- $e_{ij} \sim N(0, \sigma^2)$ não-correlacionados



Exemplo EMBRAPA 4

Gráfico de perfis para os dados da EMBRAPA



Exemplo EMBRAPA 5

- Regressão linear com interceptos aleatórios para dados longitudinais:

$$y_{ij} = \alpha + \beta x_j + a_i + d_{ij} = (\alpha + a_i) + \beta x_j + d_{ij}$$

- $i = 1, \dots, 15, j = 1, \dots, 8$
- $a_i \sim N(0, \sigma_a^2)$ não-correlacionados
- $d_{ij} \sim N(0, \sigma_e^2)$ não-correlacionados
- a_i e d_{ij} não-correlacionados
- Consequentemente
 - $\mathbb{V}(y_{ij}) = \sigma_a^2 + \sigma_e^2$
 - $\mathbb{Cov}(y_{ij}, y_{ik}) = \sigma_a^2$: covariância **intraunidades** amostrais
 - $\mathbb{Cov}(y_{lj}, y_{ik}) = 0$: covariância **interunidades** amostrais

Exemplo EMBRAPA 6

- Resultado do ajuste do modelo de regressão linear simples
 - Mínimos quadrados ou Máxima verossimilhança
 - $\hat{\alpha} = 13.5$ com $EP = 7.8$
 - $\hat{\beta} = 4.7$ com $EP = 0.4$
- Resultado do ajuste do modelo de regressão linear com interceptos aleatórios
 - Máxima verossimilhança ou Mínimos quadrados/ANOVA para componentes de variância/Mínimos quadrados generalizados
 - $\hat{\alpha} = 13.5$ com $EP = 5.6$
 - $\hat{\beta} = 4.7$ com $EP = 0.1$

Efeito da correlação intraunidades amostrais

- Comparação de médias $X \sim N(\mu_X, \sigma^2)$ e $Y \sim N(\mu_Y, \sigma^2)$

- **Distribuições independentes:** $t = (\bar{X} - \bar{Y})/s\sqrt{2/n}$

$$s^2 = (2n)^{-1} \left[\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2 \right]$$

é uma estimativa de $\mathbb{V}(X_i) = \mathbb{V}(Y_i) = \sigma^2$

- **Distribuições pareadas:** $t_d = (\bar{X} - \bar{Y})/s_d\sqrt{1/n}$

$$s_d^2 = n^{-1} \left[\sum_{i=1}^n [(X_i - Y_i) - (\bar{X} - \bar{Y})]^2 \right]$$

é uma estimativa de $\mathbb{V}(X_i - Y_i) = 2\sigma^2 - 2\mathbb{C}\text{ov}(X_i, Y_i)$

- Se $\mathbb{C}\text{ov}(X_i, Y_i) > 0$, t_d tem mais poder para detectar diferenças do que t .

Exemplos não tão bem comportados 1

Tempo gasto para calçar meia (s)

Gênero	Treinamento		Gênero	Treinamento	
	Antes	Depois		Antes	Depois
m	6.0	4.9	f	6.7	4.9
m	5.3	5.3	f	4.8	3.8
m	8.5	5.8	f		4.4
m	5.1	5.6	f	6.8	8.5
m	4.9	6.2	f	9.8	7.9
m	11.0	10.3	f		3.9
m	9.2	6.7	f	7.5	3.7
m	10.0	5.5	f	6.3	5.0
m	9.9	8.7	f	4.7	3.8
m		5.6	f	6.0	5.1
m		8.4	f	8.1	6.6
m	6.9	6.6	f	10.3	7.2
f	4.5	8.3	f	8.0	7.7
f	5.8	5.4	f	6.3	4.9
f	5.2	4.4	f	8.8	8.0

Exemplos não tão bem comportados 2

Diâmetro sistólico da aorta por unidade de peso (mm/kg)

Peso	Semanas pós-concepção													
	27	28	29	30	31	32	33	34	35	36	37	38	39	40
AIG	9.4	10.3		8.7	8.2	6.7	6.1		5.6		5.1			4.9
AIG				6.1	6.1	6.2		5.4	5.2	4.9				
AIG				5.8	6.3	5.8	5.1	4.9	4.6					
AIG			9.7	9.2		9.5		7.3		6.1	5.4	4.8		4.5
AIG						6.4	5.8	5.2	4.7					
AIG					5.4	4.9	4.6	4.3						
AIG		8.3	8.5	8.6		7.9		6.2		5.5		4.2		
AIG		7.7	8.6	7.9			6.6		5.7					
AIG								5.9	6.1		5.4			4.1
AIG						7.0	6.5							
AIG				5.2		4.8	4.2	4.1			3.7			
AIG				6.2	6.1		6.2			6.0		5.3		
PIG						7.2	6.8		5.5			4.7		
PIG							7.1	8.0	7.7			6.5	5.6	
PIG					7.4	8.3	9.4	10.0	9.2			8.0		
PIG						7.7		6.6		5.5		4.6		
PIG								6.5					4.4	
PIG				7.6	8.6	9.3	8.0		6.6		5.0		4.7	
PIG				6.6	8.4		8.2	7.6		6.6				
PIG						7.1	6.3			6.1	5.9	5.7		4.8
PIG					8.5	8.4					4.9			
PIG			8.3	7.4		6.2			4.6		3.8			
PIG			9.8			9.1		7.3		5.3				
PIG	8.5													
PIG			10.9	10.7	9.4		8.0			5.8		4.9		

Exemplos não tão bem comportados 3

Nível de corrimento vaginal em três avaliações

Paciente	Avaliação			Paciente	Avaliação		
	inicial	14 d	21 d		inicial	14 d	21 d
1	1	0	0	26	2	0	0
2	2	0	.	27	2	3	.
3	1	0	0	28	3	0	1
4	2	0	0	29	2	2	1
5	2	1	.	30	2	0	0
6	2	.	.	31	3	.	0
7	2	.	.	32	0	.	0
8	1	1	1	33	1	1	0
9	3	0	0	34	1	0	0
10	2	1	2	35	1	0	0
11	2	1	3	36	1	1	0
12	1	1	0	37	0	0	1
13	2	0	0	38	0	0	1
14	2	0	0	39	1	0	0
15	2	1	1	40	2	0	0
16	2	1	1	41	1	1	0
17	2	1	0	42	1	0	0
18	1	0	0	43	2	.	.
19	1	0	0	44	2	2	.
20	2	0	0	45	2	0	1
21	1	1	0	46	2	.	0
22	3	1	0	47	3	1	0
23	3	0	0	48	3	0	0
24	2	1	1	49	2	1	1
25	2	0	0	50	3	0	0

Enfoques para proposição de modelos lineares

- $\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\xi}_i$, $i = 1, \dots, N$ com
 - \mathbf{y}_i : $(n_i \times 1)$ vetor de respostas para a i -ésima unidade amostral
 - $\boldsymbol{\beta}$: $(p \times 1)$ vetor de parâmetros (**efeitos fixos**)
 - \mathbf{X}_i : $(n_i \times p)$ matriz de especificação do modelo (conhecida)
 - $\boldsymbol{\xi}_i$: $(n_i \times 1)$ vetor de erros aleatórios (não-correlacionados, $i = 1, \dots, N$)
- **Modelo marginal** (*population-averaged model* - modelo populacional ponderado): $\mathbb{E}(\boldsymbol{\xi}_i) = \mathbf{0}$, $\mathbb{V}(\boldsymbol{\xi}_i) = \mathbf{V}_i$
- **Modelo misto** (*subject-specific model* - modelo específico para indivíduos): $\boldsymbol{\xi}_i = \mathbf{Z}_i\mathbf{b}_i + \mathbf{e}_i$ com
 - \mathbf{Z}_i : $(n_i \times q)$ matriz de especificação do modelo (conhecida)
 - \mathbf{b}_i : $(q \times 1)$ vetor de **efeitos aleatórios** (não-correlacionados $i = 1, \dots, N$) com $\mathbb{E}(\mathbf{b}_i) = \mathbf{0}$, $\mathbb{V}(\mathbf{b}_i) = \mathbf{G}$
 - \mathbf{e}_i : $(n_i \times 1)$ vetor de erros aleatórios (não correlacionados $i = 1, \dots, N$) com $\mathbb{E}(\mathbf{e}_i) = \mathbf{0}$, $\mathbb{V}(\mathbf{e}_i) = \boldsymbol{\Sigma}_i$
 - \mathbf{e}_i e \mathbf{b}_i não correlacionados $i = 1, \dots, N$
 - Modelo marginal induzido: $\mathbb{V}(\boldsymbol{\xi}_i) = \mathbf{V}_i = \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^\top + \boldsymbol{\Sigma}_i$

Modelos lineares mistos

- Usualmente supõe-se
 - $\mathbf{b}_i \stackrel{\text{iid}}{\sim} \mathcal{N}_q(\mathbf{0}, \mathbf{G})$ e $\mathbf{e}_i \stackrel{\text{ind}}{\sim} \mathcal{N}_{n_i}(\mathbf{0}, \mathbf{\Sigma}_i)$
 - \mathbf{G} ($q \times q$) e $\mathbf{\Sigma}_i$ ($n_i \times n_i$) matrizes definidas positivas com elementos expressos como funções de **parâmetros de covariância** $\boldsymbol{\theta}$ ($m \times 1$) não relacionados funcionalmente com $\boldsymbol{\beta}$
 - Se $\mathbf{\Sigma}_i = \mathbf{I}_{n_i} \sigma^2$: **modelo homocedástico de independência condicional**
- Compactamente

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \mathbf{e}$$

com

$$\mathbf{y} = (\mathbf{y}_1^\top, \dots, \mathbf{y}_N^\top)^\top, \quad \mathbf{X} = (\mathbf{X}_1^\top, \dots, \mathbf{X}_N^\top)^\top, \quad \mathbf{Z} = \bigoplus_{i=1}^N \mathbf{Z}_i$$

$$\mathbf{b} = (\mathbf{b}_1^\top, \dots, \mathbf{b}_N^\top)^\top, \quad \mathbf{e} = (\mathbf{e}_1^\top, \dots, \mathbf{e}_N^\top)^\top$$

$$\mathbf{\Gamma} = \mathbf{I}_N \otimes \mathbf{G}, \quad \mathbf{\Sigma} = \bigoplus_{i=1}^N \mathbf{\Sigma}_i$$

- Consequentemente $\mathbb{V}(\mathbf{y}) = \mathbf{Z}\mathbf{\Gamma}\mathbf{Z}^\top + \mathbf{\Sigma}$

1 Estrutura Uniforme

$$\mathbf{V}_i(\boldsymbol{\theta}) = \begin{bmatrix} \sigma_a^2 + \sigma^2 & \sigma_a^2 & \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 + \sigma^2 & \sigma_a^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 & \sigma_a^2 + \sigma^2 & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 & \sigma_a^2 & \sigma_a^2 + \sigma^2 \end{bmatrix}$$

2 Não-estruturada (NE)

$$\mathbf{V}_i(\boldsymbol{\theta}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_4^2 \end{bmatrix}$$

3 Estrutura AR(1)

$$\mathbf{V}_i(\boldsymbol{\theta}) = \sigma^2 \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 \\ \phi & 1 & \phi & \phi^2 \\ \phi^2 & \phi & 1 & \phi \\ \phi^3 & \phi^2 & \phi & 1 \end{bmatrix}$$

- Dado θ [$\Gamma(\theta)$, $\Sigma(\theta)$ e $V(\theta)$]
 - BLUE de β : $\hat{\beta} = (\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{V}^{-1} \mathbf{y}$
 - BLUP de \mathbf{b} : $\hat{\mathbf{b}} = \Gamma \mathbf{Z}^\top \mathbf{V}^{-1} [\mathbf{I} - \mathbf{X} (\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{V}^{-1}] \mathbf{y}$
- Estimação de θ (Γ , Σ , V): Máxima verossimilhança restrita (REML)
 - Equações de estimação ($j = 1, \dots, m$):

$$-\frac{1}{2} \sum_{i=1}^N \text{tr} \{ \mathbf{V}_i^{-1}(\hat{\theta}) \dot{\mathbf{V}}_i(\hat{\theta}) \} - \frac{1}{2} \sum_{i=1}^N [\partial Q_i(\theta) / \partial \theta_j |_{\theta=\hat{\theta}}]$$
$$-\frac{1}{2} \sum_{i=1}^N \text{tr} \{ \mathbf{V}_i^{-1}(\hat{\theta}) \mathbf{X}_i^\top \mathbf{V}_i^{-1}(\hat{\theta}) \dot{\mathbf{V}}_i(\hat{\theta}) \mathbf{V}_i^{-1}(\hat{\theta}) \mathbf{X}_i \} = 0,$$

em que $\dot{\mathbf{V}}_i(\hat{\theta}) = [\partial \mathbf{V}_i(\theta) / \partial \theta_j]^\top_{\theta=\hat{\theta}}$ e

$$Q_i(\theta) = [\mathbf{y}_i - \mathbf{X}_i \hat{\beta}(\theta)]^\top \mathbf{V}_i^{-1}(\theta) [\mathbf{y}_i - \mathbf{X}_i \hat{\beta}(\theta)].$$

- Substituindo $\hat{\Gamma}$ e $\hat{\Sigma}$ nas expressões de $\hat{\beta}$ e $\hat{\mathbf{b}}$ obtemos BLUE e BLUP empíricos

Cálculo dos estimadores de MV e MV restrita

- Soluções explícitas das equações de estimação só em casos particulares
 - Análise de perfis com dados completos e matriz de covariâncias intraunidades amostrais não-estruturada (MANOVA) ou uniforme (ANOVA para medidas repetidas)
 - Modelos mistos com independência condicional e dados balanceados em relação ao tempo completos

- Algoritmo de **Newton-Raphson**:

$$\boldsymbol{\theta}^{(l)} = \boldsymbol{\theta}^{(l-1)} - \mathbf{H}^{-1}[\boldsymbol{\theta}^{(l-1)}] \mathbf{u}[\boldsymbol{\theta}^{(l-1)}], l = 1, 2, \dots$$

- Função escore: $\mathbf{u}(\boldsymbol{\theta}) = \partial l[\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}), \boldsymbol{\theta}] / \partial \boldsymbol{\theta}$
 - Matriz Hessiana: $\mathbf{H}(\boldsymbol{\theta}) = \partial^2 l[\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}), \boldsymbol{\theta}] / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top$
 - Regra de parada $\|\boldsymbol{\theta}^{(l)} - \boldsymbol{\theta}^{(l-1)}\| < \varepsilon, \varepsilon > 0$.
- Algoritmo de "**Scoring**" de Fisher: $\mathbf{H}(\boldsymbol{\theta})$ substituída por sua esperança
 - Algoritmo **EM**

Propriedades dos estimadores

- $\widehat{\beta}(\theta) \sim N\{\beta, \mathbf{V}_{\widehat{\beta}}(\theta)\}$ com $\mathbf{V}_{\widehat{\beta}}(\theta) = [\sum_{i=1}^N \mathbf{X}_i^\top \mathbf{V}_i^{-1}(\theta) \mathbf{X}_i]^{-1}$
- $\widehat{\theta} \approx N\{\theta, \mathbf{V}_{\widehat{\theta}}(\theta)\}$ com $\mathbf{V}_{\widehat{\theta}}(\theta)$ denotando uma matriz com dimensão $(m \times m)$ cujo elemento (r, s) , $r, s = 1, \dots, m$, é dado por $[\mathbf{V}_{\widehat{\theta}}(\theta)]_{rs} = \frac{1}{2} \sum_{i=1}^N \text{tr}\{\mathbf{V}_i^{-1}(\theta) \dot{\mathbf{V}}_{ir}(\theta) \mathbf{V}_i^{-1}(\theta) \dot{\mathbf{V}}_{is}(\theta)\}$
- $\widehat{\beta}(\widehat{\theta}) \approx N\{\beta, \mathbf{V}_{\widehat{\beta}}(\theta)\}$
- $\mathbf{V}_{\widehat{\beta}}(\theta)$ e $\mathbf{V}_{\widehat{\theta}}(\theta)$ podem ser estimadas por meio da inversa da matriz de informação de Fisher
- Resultados válidos mesmo sem a suposição de normalidade desde que N seja suficientemente grande

- Família exponencial

$$f(y_i|\theta_i, \phi) = c(y_i, \phi) \exp\{[y_i\theta_i - b(\theta_i)]/a(\phi)\}$$

- $\theta_i, i = 1, \dots, N$: parâmetros de interesse
 - $\phi > 0$: parâmetro de escala
 - a, b e c : funções de forma conhecida
 - $\mathbb{E}(y_i) = \mu_i = \mu_i(\theta_i) = b'(\theta_i)$
 - $\mathbb{V}(y_i) = b''(\theta_i)a(\phi) = a(\phi) \frac{\partial}{\partial \theta_i} \mu_i(\theta_i)$
 - $v_i[\mu_i(\theta_i)] = \frac{\partial}{\partial \theta_i} \mu_i(\theta_i)$: função de variância
- Modelo: $g[\mu(\theta_i)] = \mathbf{x}_i^\top \boldsymbol{\beta}$
 - $\mathbf{x}_i^\top \boldsymbol{\beta}$: **componente sistemática**
 - g : **função de ligação** (monótona e diferenciável)
 - $\theta_i = (g \circ \mu)^{-1}(\mathbf{x}_i^\top \boldsymbol{\beta})$
 - Se $g = \mu^{-1}$ então $\theta_i = \mathbf{x}_i^\top \boldsymbol{\beta}$ e g : **função de ligação canônica**

Modelos Lineares Generalizados 2

Em notação matricial: $\mathbf{G}(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta}$

- $\boldsymbol{\mu} = (\mu_1, \dots, \mu_N)^\top$
- $\mathbf{G}(\boldsymbol{\mu}) = [g(\mu_1), \dots, g(\mu_N)]^\top$
- $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)^\top$

Fazendo $\phi = 0$ sem perda de generalidade e $h = (g \circ \mu)^{-1}$:

$$\log L_N(\boldsymbol{\beta}) = \sum_{i=1}^N \left\{ y_i h(\mathbf{x}_i^\top \boldsymbol{\beta}) - b \left[h(\mathbf{x}_i^\top \boldsymbol{\beta}) \right] \right\} - K$$

Equações de estimação:

$$\sum_{i=1}^N [g' \{ \mu_i(\hat{\boldsymbol{\beta}}) \} v_i(\hat{\boldsymbol{\beta}})]^{-1} [y_i - \mu_i(\hat{\boldsymbol{\beta}})] \mathbf{x}_i = \mathbf{0}$$

Em forma matricial:

$$\mathbf{X}^\top [\mathbf{D}(\hat{\boldsymbol{\beta}})]^{-1} [\mathbf{y} - \boldsymbol{\mu}(\hat{\boldsymbol{\beta}})] = \mathbf{0}$$

com $\mathbf{D}(\boldsymbol{\beta}) = \bigoplus_{i=1}^N [g' \{ \mu_i(\boldsymbol{\beta}) \} v_i(\boldsymbol{\beta})]$

Modelos Lineares Generalizados 3

- Distribuição assintótica:

$$\sqrt{N}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \rightarrow N\{\mathbf{0}, [\mathbf{I}(\boldsymbol{\beta})]^{-1}\}$$

- $\mathbf{I}(\boldsymbol{\beta}) = \mathbb{E}_{\boldsymbol{\beta}} \left[-\frac{\partial^2}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \log L(\boldsymbol{\beta}) \right] = \phi \{ \mathbf{D}(\boldsymbol{\beta})^\top [\mathbf{V}(\boldsymbol{\beta})]^{-1} \mathbf{D}(\boldsymbol{\beta}) \}$

- $\mathbf{V}(\boldsymbol{\beta}) = \oplus_{i=1}^N v_i(\boldsymbol{\beta})$

- $\hat{\mathbf{I}}(\hat{\boldsymbol{\beta}}) = \hat{\phi} \{ \mathbf{D}(\hat{\boldsymbol{\beta}})^\top [\mathbf{V}(\hat{\boldsymbol{\beta}})]^{-1} \mathbf{D}(\hat{\boldsymbol{\beta}}) \}$

- $\hat{\phi} = (N - p)^{-1} \sum_{i=1}^N \{ [y_i - \mu_i(\hat{\boldsymbol{\beta}})]^2 / [\mathbf{V}_i(\hat{\boldsymbol{\beta}})] \}$

- Para $g = \mu^{-1}$ (ligação canônica), equações de estimação:

$$\sum_{i=1}^N [y_i - \mu_i(\hat{\boldsymbol{\beta}})] \mathbf{x}_i = \mathbf{0}$$

- Exemplo: distribuição gaussiana, função de ligação identidade

$$\sum_{i=1}^N [y_i - \mathbf{x}_i^\top \hat{\boldsymbol{\beta}}] \mathbf{x}_i = \mathbf{0}$$

Equações de estimação generalizadas 1

Consideremos y_i tal que

- $\mathbb{E}(y_i) = \mu_i$ e $\mathbb{V}(y_i) = \phi v(\mu_i)$
- sem especificação da forma da distribuição
- $g(\mu_i) = \mathbf{x}_i^\top \boldsymbol{\beta}$

Consideremos $U_i = [y_i - \mu_i(\boldsymbol{\beta})] / \{\phi v[\mu_i(\boldsymbol{\beta})]\}$

- $\mathbb{E}(U_i) = 0$ e $\mathbb{V}(U_i) = \{\phi v[(\mu_i(\boldsymbol{\beta}))]\}^{-1}$
- $Q[\mu_i(\boldsymbol{\beta}); y_i] = \int_{y_i}^{\mu_i(\boldsymbol{\beta})} (y_i - t) / \{\phi v[\mu_i(\boldsymbol{\beta})]\} dt$ funciona como uma log-verossimilhança
- Função de **quase-log-verossimilhança**:
 $Q[\boldsymbol{\mu}(\boldsymbol{\beta}); \mathbf{y}] = \sum_{i=1}^N Q[\mu_i(\boldsymbol{\beta}); y_i]$
- Equações de estimação generalizadas (**GEE**):

$$\frac{\partial}{\partial \boldsymbol{\beta}} Q[\mu_i(\boldsymbol{\beta}); y_i] = \sum_{i=1}^N \frac{\partial \mu_i(\hat{\boldsymbol{\beta}})}{\partial \boldsymbol{\beta}} \{v[\mu_i(\hat{\boldsymbol{\beta}})]\}^{-1} [y_i - \mu_i(\hat{\boldsymbol{\beta}})] = \mathbf{0}$$

Equações de estimação generalizadas 2

- Equações de estimação generalizadas (GEE):

$$\sum_{i=1}^N \mathbf{D}_i(\hat{\boldsymbol{\beta}})^\top [\mathbf{V}_i(\hat{\boldsymbol{\beta}})]^{-1} [\mathbf{y}_i - \boldsymbol{\mu}_i(\hat{\boldsymbol{\beta}})] = \mathbf{0}$$

- $\mathbf{D}_i(\boldsymbol{\beta}) = (\partial/\partial\boldsymbol{\beta})\boldsymbol{\mu}_i(\boldsymbol{\beta})$
- Distribuição assintótica:

$$\sqrt{N}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \rightarrow N\{\mathbf{0}, [\mathbf{I}(\boldsymbol{\beta})]^{-1}\}$$

- $\mathbf{I}(\boldsymbol{\beta}) = \phi \sum_{i=1}^N \mathbf{D}_i(\boldsymbol{\beta})^\top [\mathbf{V}_i(\boldsymbol{\beta})]^{-1} \mathbf{D}_i(\boldsymbol{\beta})$
- $\hat{\mathbf{I}}(\hat{\boldsymbol{\beta}}) = \hat{\phi} \{\mathbf{D}(\hat{\boldsymbol{\beta}})^\top [\mathbf{V}(\hat{\boldsymbol{\beta}})]^{-1} \mathbf{D}(\hat{\boldsymbol{\beta}})\}$
- $\hat{\phi} = (N - p)^{-1} \sum_{i=1}^N \{[y_i - \mu_i(\hat{\boldsymbol{\beta}})]^2 / [\mathbf{V}_i(\hat{\boldsymbol{\beta}})]\}$

Equações de estimação generalizadas 3

Para observações correlacionadas, consideremos y_{ij} tal que

- $\mathbb{E}(y_{ij}) = \mu_{ij}$ e $\mathbb{V}(y_{ij}) = \phi v(\mu_{ij})$
- $\mathbf{V}_i = \mathbf{A}_i^{1/2} \mathbf{R}_i \mathbf{A}_i^{1/2}$
 - $\mathbf{A}_i^{1/2} = \bigoplus_{j=1}^{n_i} \sqrt{\phi v(\mu_{ij})}$
 - \mathbf{R}_i é uma **matriz de correlações de trabalho**
- sem especificação da forma da distribuição
- $g(\mu_{ij}) = \mathbf{x}_i^\top \boldsymbol{\beta}$
- $\mathbf{R}_i = \mathbf{R}_i(\boldsymbol{\alpha})$
- Equações de estimação generalizadas (GEE):

$$\sum_{i=1}^N \mathbf{D}_i(\hat{\boldsymbol{\beta}})^\top [\mathbf{V}_i(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}})]^{-1} [\mathbf{y}_i - \boldsymbol{\mu}_i(\hat{\boldsymbol{\beta}})] = \mathbf{0}$$

- $\mathbf{V}_i(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \phi [\mathbf{A}_i(\boldsymbol{\beta})]^{1/2} \mathbf{R}_i(\boldsymbol{\alpha}) [\mathbf{A}_i(\boldsymbol{\beta})]^{1/2}$
- Na solução das GEE, $\mathbf{R}_i(\boldsymbol{\alpha})$ é substituída por um estimador consistente

Equações de estimação generalizadas 4

- Distribuição assintótica:

$$\sqrt{N}(\hat{\beta} - \beta) \rightarrow N\{\mathbf{0}, \mathbf{W}^{-1}\mathbf{M}\mathbf{W}^{-1}\}$$

- $\mathbf{W} = \sum_{i=1}^N \mathbf{D}_i(\beta)^\top [\mathbf{V}_i(\alpha, \beta)]^{-1} \mathbf{D}_i(\beta)$
- $\mathbf{M} = \sum_{i=1}^N \mathbf{D}_i(\beta)^\top [\mathbf{V}_i(\alpha, \beta)]^{-1} \mathbb{V}(\mathbf{y}_i) [\mathbf{V}_i(\alpha, \beta)]^{-1} \mathbf{D}_i(\beta)$
- Se a matriz de covariâncias de trabalho, $\mathbf{V}_i(\alpha, \beta)$ estiver **bem especificada**, i.e. $\mathbf{V}_i(\alpha, \beta) = \mathbb{V}(\mathbf{y}_i)$,

$$\sqrt{N}(\hat{\beta} - \beta) \rightarrow N\{\mathbf{0}, [\mathbf{W}(\alpha, \beta)]^{-1}\}$$

- Mesmo que a matriz de covariâncias de trabalho não esteja bem especificada, $\hat{\beta}$ é um **estimador consistente** de β

- Estimadores de ϕ , α :

- $\hat{e}_{ij} = [y_{ij} - \mu_{ij}(\hat{\beta})] / \sqrt{v_i[\mu_{ij}(\hat{\beta})]}$

- $\hat{\phi} = \sum_{i=1}^N \sum_{j=1}^{n_i} \hat{e}_{ij}^2 / \sum_{i=1}^N n_i$

- $\hat{\alpha}_{jk} = (N\hat{\phi})^{-1} \sum_{i=1}^N \hat{e}_{ij}\hat{e}_{jk}$ [$\mathbb{V}(\mathbf{y}_i)$ não-estruturada]

- $\hat{\alpha} = (N\hat{\phi})^{-1} \sum_{i=1}^N \frac{1}{n_i-1} \sum_{i < n_i-1} \hat{e}_{ij}\hat{e}_{i,j-1}$ [$\mathbb{Cov}(\mathbf{y}_i)$ autorregressiva]

- Estimador de $\mathbf{W}^{-1}\mathbf{M}\mathbf{W}^{-1}$: substituir $\mathbb{V}(\mathbf{y}_i)$ por

$$[\mathbf{y}_i - \boldsymbol{\mu}_i(\hat{\beta})][\mathbf{y}_i - \boldsymbol{\mu}_i(\hat{\beta})]^\top$$

- Condicionalmente a um vetor de efeitos aleatórios \mathbf{b}_i ,

$$f(y_{ij}|\mathbf{b}_i, \theta_{ij}, \phi) = c(y_{ij}, \phi) \exp\{[y_{ij}\theta_{ij} - b(\theta_{ij})]/a(\phi)\}$$

- $\mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G})$
- $g(\mu_{ij}) = g[\mathbb{E}(y_{ij}|\mathbf{b}_i)] = \mathbf{x}_{ij}^\top \boldsymbol{\beta} + \mathbf{z}_{ij}^\top \mathbf{b}_i$
- g : função de ligação
- Estimação envolve algoritmos mais complicados
- Interpretação dos parâmetros é mais delicada

Exemplo comparativo

Comparar eficácia de dois tipos de escova de dentes quanto à capacidade de remoção de placa bacteriana ao longo de 60 dias

Escova monobloco



Índice de placa bacteriana 1

Método de cálculo do índice de placa bacteriana

Study conducted at the School of Dentistry, USP to evaluate the efficiency of

- a) types of toothbrush (conventional and monoblock)
- b) use of toothpaste (yes or no)

with respect to removal of dental plaque

Dental plaque index = $\frac{\# \text{ of stained regions}}{\text{total } \# \text{ of teeth}}$

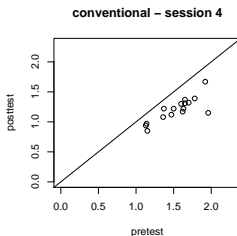
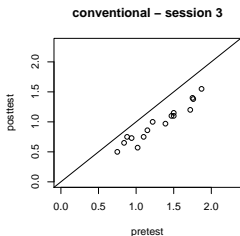
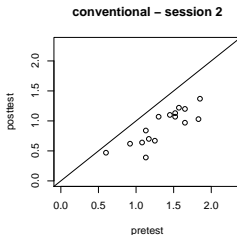
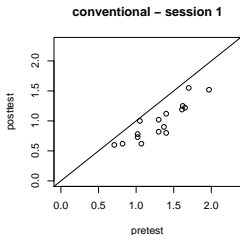


Índice de placa bacteriana 2

		Sessão 1		Sessão 2		Sessão 3		Sessão 4	
		Antes	Depois	Antes	Depois	Antes	Depois	Antes	Depois
1	Convencional	1.05	1	1.13	0.84	1.15	0.86	1.13	0.94
2	Convencional	1.07	0.62	0.92	0.62	1.02	0.57	1.15	0.85
3	Convencional	0.82	0.62	1.52	1.07	1.39	0.97	1.78	1.39
4	Convencional	1.37	0.9	1.65	1.2	1.75	1.4	1.92	1.67
5	Convencional	1.97	1.52	1.3	1.07	1.5	1.15	1.65	1.37
6	Convencional	1.3	0.82	1.17	0.7	0.75	0.5	1.47	1.12
7	Convencional	1.61	1.19	1.52	1.13	1.22	1	1.63	1.22
8	Convencional	1.02	0.73	1.08	0.64	0.94	0.73	1.14	0.97
9	Convencional	1.62	1.25	1.45	1.1	1.1	0.75	1.7	1.32
10	Convencional	1.65	1.22	1.57	1.22	1.47	1.1	1.62	1.17
11	Convencional	1.02	0.78	0.6	0.47	0.88	0.75	1.36	1.08
12	Convencional	0.71	0.6	1.13	0.39	0.84	0.65	1.65	1.31
13	Convencional	1.7	1.55	1.85	1.37	1.87	1.55	1.6	1.3
14	Convencional	1.3	1.02	1.65	0.97	1.72	1.2	1.37	1.22
15	Convencional	1.4	0.8	1.83	1.03	1.76	1.38	1.96	1.15
16	Convencional	1.4	1.12	1.25	0.67	1.5	1.1	1.5	1.22
17	Monobloco	1.66	1.63	1.36	1.16	1.52	0.88	1.41	1.2
18	Monobloco	1.02	0.8	0.92	0.82	1.1	0.76	1.28	1.15
19	Monobloco	0.75	0.67	1	0.92	1	0.87	1.15	1.1
20	Monobloco	1.29	1.23	0.91	0.76	1.14	0.94	1.35	0.97
21	Monobloco	1.27	1.2	1.2	0.95	1.1	1	1.37	1.17
22	Monobloco	1.07	0.85	1.39	1.25	1.39	1.25	1.28	1.21
23	Monobloco	1.35	1.21	1.42	1.17	1.42	1.19	1.42	1.23
24	Monobloco	1.32	1.02	1.6	1.4	1.35	1.02	1.5	1.25
25	Monobloco	1.66	1.61	1.5	1.36	1.72	1.41	1.69	1.44
26	Monobloco	1.3	1.07	0.84	0.61	0.88	0.61	0.96	0.57
27	Monobloco	1.57	1.2	1.5	1.07	1.15	1	1.25	1.05
28	Monobloco	1.67	1.5	1.47	1.32	1.07	0.97	1.5	1.37
29	Monobloco	0.91	0.67	0.96	0.62	1.09	0.53	1.12	0.37
30	Monobloco	1.06	0.7	1	0.85	1.15	0.93	1.12	1
31	Monobloco	2.3	2	1.37	1.25	1.4	1.32	2.15	1.9
32	Monobloco	1.15	1	1.23	1.11	1.15	1.07	1.26	1

Índice de placa bacteriana 6

Índice de placa bacteriana pré- e pós-escovação (escova convencional)



Índice de placa bacteriana 4

- Singer and Andrade (1997, Biometrics)
- Singer, Nobre and Sef (2002, Statistical Modelling)
- Nobre and Singer (2007, Biometrical Journal)
- Nobre and Singer (2011, Journal of Applied Statistics)
- Alencar, Singer and Rocha (2011, submetido)
 - Índice de placa pré-teste (x) nulo implica índice de placa pós-teste (y) nulo
 - Índices de placa pré- e pós-teste não-negativos
 - Dados heterocedásticos [a resposta é não-negativa e $E(y) \leq E(x)$]
 - A relação entre os índices pré- e pós-teste pode ser não-linear
 - As observações da mesma criança podem ser correlacionadas
- O modelo deveria incorporar todas essas características

Índice de placa bacteriana 5

Modelo proposto

$$y_{ijd} = \beta_{jd} x_{ijd}^{\gamma_{jd}} \varepsilon_{ijd},$$

- $i = 1, \dots, 16$
- $j = 0$ (convencional), 1 (monobloco)
- $d = 1, 2, 3, 4$
- ε_{ijd} erro aleatório não negativo

Modelo linearizado

$$\ln y_{ijd} = \ln(\beta_{jd}) + \gamma_{jd} \ln(x_{ijd}) + \ln(\varepsilon_{ijd})$$

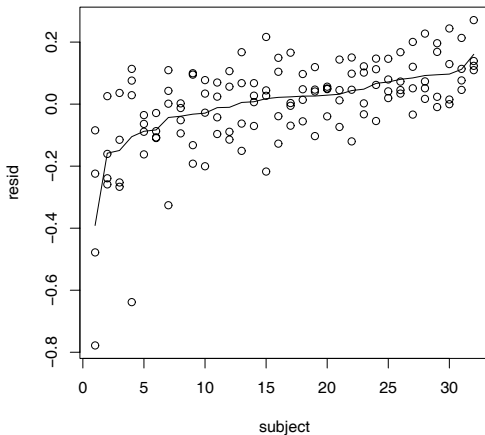
$$\ln y_{ijd} = \lambda_{jd} + \gamma_{jd} \ln(x_{ijd}) + \delta_{ijd}$$

$$\delta_{ijd} \sim N(0, \sigma^2)$$

- $\lambda_{jd} = \ln(\beta_{jd})$
- δ_{ijd} independentes (análise preliminar)
- $\varepsilon_{ijd} \sim \text{lognormal}$

Índice de placa bacteriana 6

Resíduos ordenados por criança



Modelo log-normal misto

$$y_{ijd}^* = \lambda_{jd} + \gamma_{jd} \ln(x_{ijd}) + b_{ij} + e_{ijd}$$

$$\mathbf{y}_{ij}^* = [\ln(y_{ij1}), \ln(y_{ij2}), \ln(y_{ij3}), \ln(y_{ij4})]^\top$$

$$\mathbf{e}_{ij} = [e_{ij1}, e_{ij2}, e_{ij3}, e_{ij4}]^\top$$

$$b_{ij} \sim N(0, \sigma_b^2) \text{ independentes}$$

$$\mathbf{e}_{ij} \sim N(\mathbf{0}, \mathbf{V}_j) \text{ independentes}$$

$$b_{ij} \text{ e } \mathbf{e}_{ij} \text{ independentes}$$

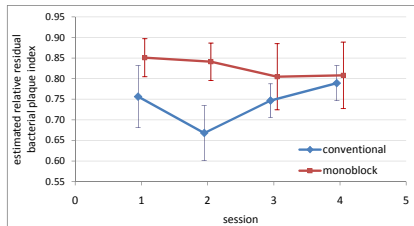
$$\mathbf{V}_j = \begin{cases} \text{diag}\{\tau_1^2, \tau_1^2, \tau_2^2, \tau_2^2\}, & \text{se } j = 0 \\ \text{diag}\{\tau_2^2, \tau_2^2, \tau_1^2, \tau_1^2\}, & \text{se } j = 1 \end{cases} .$$

Índice de placa bacteriana 8

- Resultados indicam que a hipótese $\gamma_{jd} = 1$ é plausível ($p = 0.31$)
- Parametros de interesse:

$$\beta_{jd}^* = \mathbb{E}(y_{ijd}|x_{ijd} = 1) = \beta_{jd} \exp[(\sigma_b^2 + v_{jd}^2)/2]$$

Índices de placa bacteriana residual estimados e intervalos de confiança (95%)



Modelo linear generalizado misto

$$\begin{aligned}y_{ijd}|b_{ij} &\sim \text{Gamma}(\mu_{ijd}, \phi) \\ \boldsymbol{\mu}_{ij} &= \mathbb{E}(\mathbf{y}_{ij}|b_{ij}) \\ \ln(\boldsymbol{\mu}_{ij}) &= \mathbf{X}_{ij}\boldsymbol{\beta} + \mathbf{1}_4 b_{ij} \\ b_{ij} &\sim N(0, \sigma_b^2) \text{ independentes}\end{aligned}$$

- $\boldsymbol{\mu}_{ij} = (\mu_{ij1}, \dots, \mu_{ij4})^\top$
- $\mathbb{V}(\mathbf{y}_{ij}) = \phi \mathbf{A} \boldsymbol{\mu}_{ij}$
- ϕ parâmetro de dispersão
- $\mathbf{A} \boldsymbol{\mu}_{ij} = \text{diag}(\mu_{ij1}^2, \mu_{ij2}^2, \mu_{ij3}^2, \mu_{ij4}^2)$.

Modelo linear generalizado com matriz de covariância de trabalho

$$\begin{aligned}y_{ijd} &\sim \text{Gamma}(\mu_{ijd}, \phi_{jd}) \\ \boldsymbol{\mu}_{ij} &= \mathbb{E}(\mathbf{y}_{ij}) \\ \ln(\boldsymbol{\mu}_{ij}) &= \mathbf{X}_{ij}\boldsymbol{\beta} \\ \text{Corr}(y_{ijd}, y_{ijd'}) &= \alpha, \quad d \neq d',\end{aligned}$$

$$\phi_{jd} = \begin{cases} \phi, & \text{if } (j=0 \text{ e } d=3 \text{ or } 4) \text{ ou } (j=1 \text{ e } d=1 \text{ ou } 2) \\ \phi + \phi_1, & \text{em caso contrario.} \end{cases}$$

Estima $\beta_{jd}^* = \mathbb{E}(y_{ijd})$ diretamente.

Resultados

Par.	LNLM				GLMM				GLMWCM			
	Completos		Sem outliers		Completos		Sem outliers		Completos		Sem outliers	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE
β_{01}^*	0.756	0.039	0.749	0.025	0.747	0.032	0.746	0.024	0.749	0.025	0.749	0.025
β_{02}^*	0.668	0.034	0.691	0.024	0.667	0.028	0.690	0.022	0.669	0.030	0.691	0.023
β_{03}^*	0.747	0.021	0.747	0.020	0.744	0.020	0.745	0.019	0.746	0.018	0.746	0.018
β_{04}^*	0.789	0.022	0.789	0.021	0.786	0.021	0.786	0.021	0.788	0.017	0.788	0.017
β_{11}^*	0.851	0.024	0.851	0.023	0.849	0.023	0.849	0.022	0.851	0.024	0.851	0.024
β_{12}^*	0.841	0.023	0.841	0.023	0.838	0.023	0.838	0.022	0.840	0.021	0.840	0.021
β_{13}^*	0.805	0.041	0.818	0.028	0.797	0.034	0.816	0.027	0.801	0.033	0.819	0.027
β_{14}^*	0.808	0.041	0.843	0.029	0.808	0.034	0.839	0.027	0.814	0.040	0.842	0.025
σ_{ϵ}^2	0.005	0.002	0.004	0.002	0.005	0.002	0.003	0.002				
τ_{12}^2	0.033	0.007	0.014	0.003	0.024	0.005	0.013	0.003				
τ_{22}^2	0.007	0.002	0.008	0.002	0.007	0.001	0.008	0.002				
α			1						0.331	0.139	0.247	0.087
ϕ									0.009	0.002	0.009	0.002
ϕ_1									0.018	0.009	0.006	0.003

Valores-P para testes de inexistência de interação e efeito de sessão (tempo)

Modelo	Interação	Efeito de sessão	
	Efeito	Convencional	Monobloco
LNLMM	0.0103	0.0062	0.3871
GLMM	0.0109	0.0037	0.4320
GLWC	0.0088	0.0038	0.5443

- Estimativas e valores-P similares sob os três modelos
- Interação entre Tipo de escova e Sessão significativa
- A eficácia da escova varia ao longo do tempo apenas para a escova convencional

Valores-P para testes de efeito de escova

Modelo	Efeito de escova			
	Sessão 1	Sessão 2	Sessão 3	Sessão 4
LNLMM	0.0207	<0.0001	0.2656	0.8398
GLMM	0.0125	<0.0001	0.1727	0.5965
GLWC	0.0040	<0.0001	0.1268	0.5446

- Há diferença significativa entre escovas apenas nas sessões 1 e 2.
- Depois de um mês as eficácias são equivalentes