

# Modelos não lineares mistos

Prof. Caio Azevedo

# Relembrando

- MLM :  $\mu_{ij} = \mathbf{X}'_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{b}_j$ .
- MLGM :  $\mu_{ij} = g^{-1}(\mathbf{X}'_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{b}_j)$ .
- Nem sempre a relação entre as covariáveis ( $\mathbf{X}_j$ ) e os efeitos aleatórios é adequadamente descrita por um modelo linear ou por um modelo linear generalizado.
- Modelos não lineares (mistos) (MNLMM) podem ser mais apropriados em situações em que o comportamento das médias (em função de covariáveis como o tempo, por exemplo) não é satisfatoriamente modelado pelos MLM ou pelos MLGM.

# Limitações dos MLM e MLGM e vantagens dos MNLM

- O comportamento linear e não linear induzidos, respectivamente, pelo MLM e MLGM podem não ser adequados para representar o comportamento da média.
- Existência de assíntotas (inferior e/ou superior).
- As interpretações dos parâmetros de modelos não lineares podem ser mais adequadas do que aquelas associadas aos parâmetros do MLM e do MLGM (particularmente, polinômios de grau  $\geq 3$  possuem parâmetros de difícil interpretação), principalmente em termos do problema.

# Limitações dos MLM e MLGM e vantagens dos MNLM

- Os MNLM, em geral, apresentam poucos parâmetros com interpretações úteis em termos do problema.
- Garantia de que as médias preditas respeitarão o espaço paramétrico associado.
- Uma alternativa aos MNL(M) são os modelos que consideram polinômios fracionários.

## Exemplo 6: cinética de indometacina, Kwan et al (1976)

- Os dados correspondem à um experimento de farmacocinética (etapas pelas quais a droga passa desde a ministração, introdução do fármaco no organismo, como tomar um comprimido, até sua eliminação, processo pela qual o fármaco deixa o organismo definitivamente) da droga indometacina (um tipo de anti-inflamatório).
- Seis indivíduos receberam de modo intravenoso a mesma dose de indometacina, e tiveram sua concentração de droga no plasma (em mcg/ml) medidas 11 vezes entre 15 minutos e 8 horas após o medicamento ser ministrado.
- Estudar o comportamento da concentração ao longo do tempo.

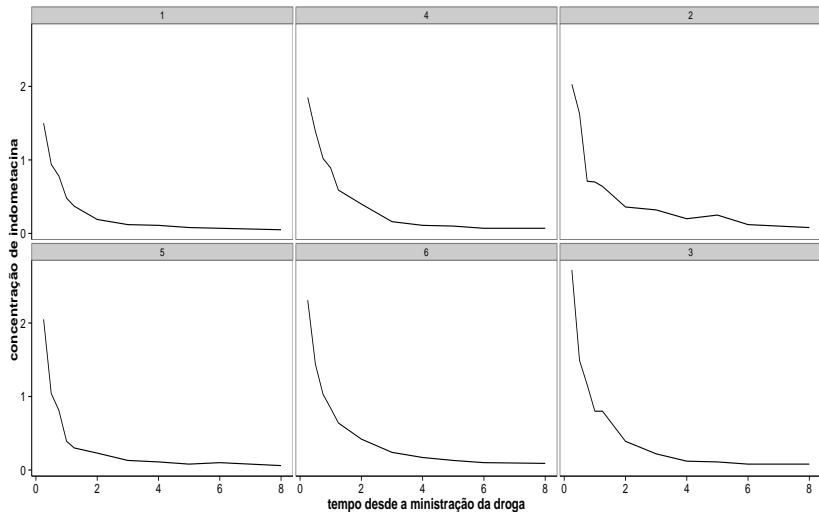
# Dados: indivíduo 1

Indivíduo	Tempo	Concetração
1	0,25	1,50
1	0,50	0,94
1	0,75	0,78
1	1,00	0,48
1	1,25	0,37
1	2,00	0,19
1	3,00	0,12
1	4,00	0,11
1	5,00	0,08
1	6,00	0,07
1	8,00	0,05

# Perfis médios e individuais

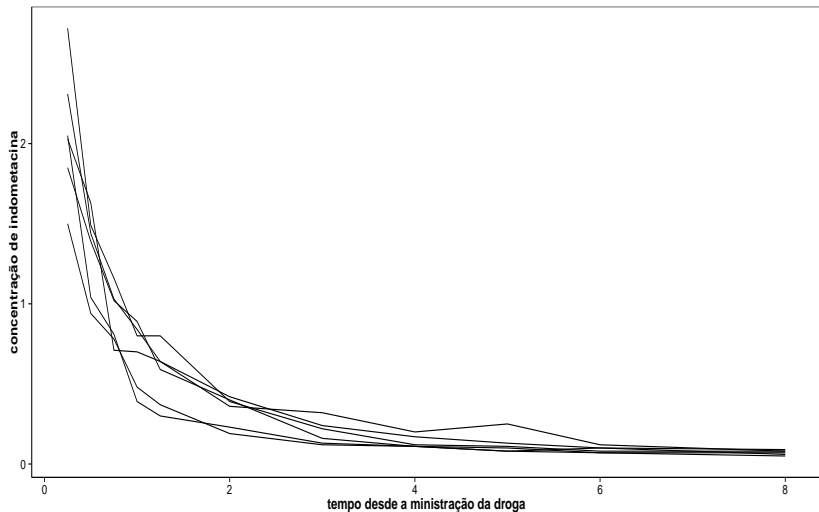
- Temos que  $n = 6$  e  $k_j = 11, j = 1, 2, \dots, 6$ , com um total de 66 observações.
- Estudo é completo (nenhuma observação foi perdida), balanceado (fo planejado para que todas as unidades experimentais fossem observadas em cada instante) e irregular (as distâncias entre os instantes não são as mesmas)

## Perfis individuais (separados)

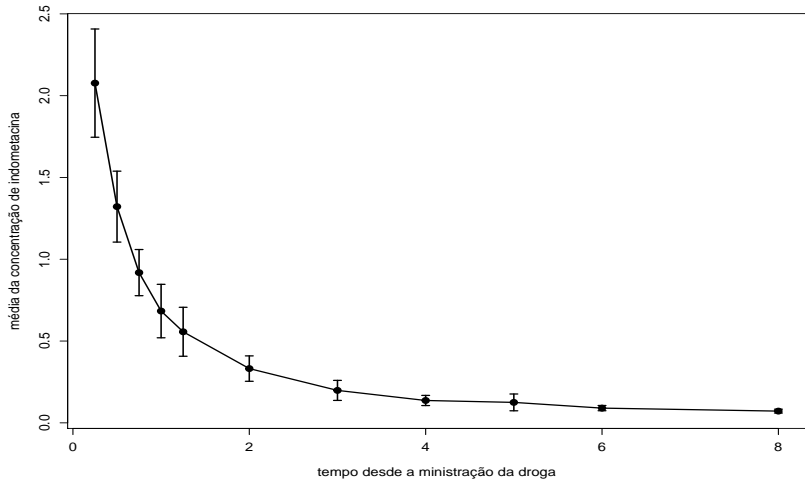




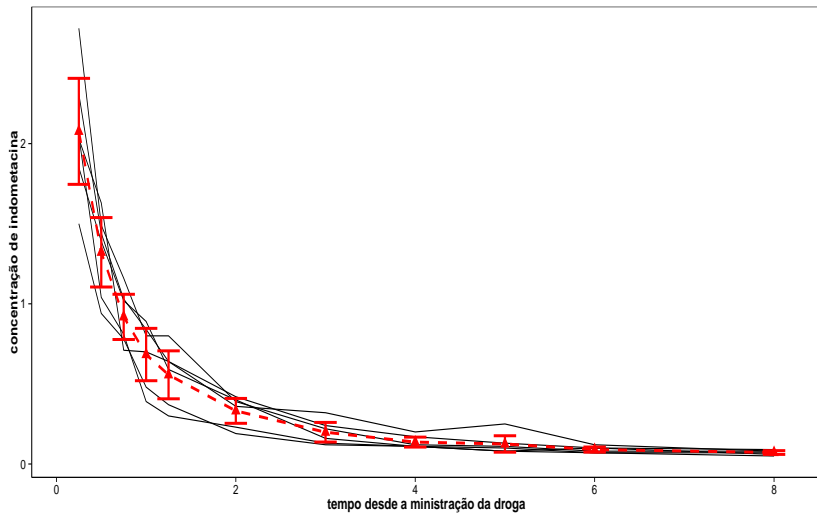
## Perfis individuais (juntos)



# Perfis médios



# Perfis médios e individuais



# Exemplo 7: crescimento de plantas de soja, Davidian and Giltinan (1995) (1976)

- Os dados correspondem à um experimento sobre crescimentos de dois tipos de soja: “Plant Introduction” # 416937 (P) um tipo experimental de cepa (parte da planta a que se cortou o caule e permanece viva no solo) e “Forrest” (F) (uma variedade comercial).
- O peso médio das folhas (em gramas) de seis plantas escolhidas aleatoriamente de cada parcela (área de cultivo) foi medida (aproximadamente) semanalmente entre duas e onze semanas depois de plantadas (embora os dias possam ser diferentes e o espaçamento não ser o mesmo).

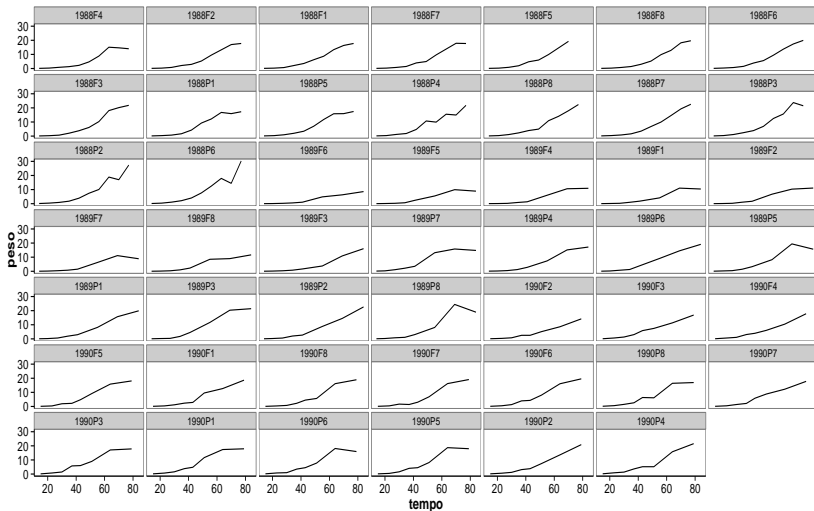
## Cont.

- O experimento foi conduzido em três anos diferentes: (1988, 1989 e 1990).
- Objetivo: comparar os tipos de plantas, quanto ao crescimento, ao longo do tempo, bem como avaliar seus comportamentos ao longo do tempo.

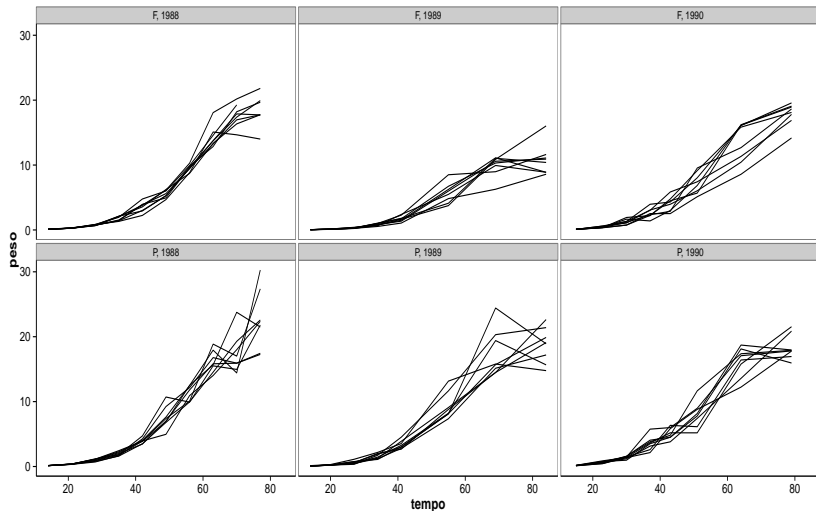
# Dados da parcela 1988F1

parcela	variedade	ano	tempo (semanas)	peso
1988F1	F	1988	14	0,11
1988F1	F	1988	21	0,26
1988F1	F	1988	28	0,67
1988F1	F	1988	35	2,11
1988F1	F	1988	42	3,56
1988F1	F	1988	49	6,23
1988F1	F	1988	56	8,71
1988F1	F	1988	63	13,35
1988F1	F	1988	70	16,34
1988F1	F	1988	77	17,75

# Perfis individuais por variedade x anos (separados)

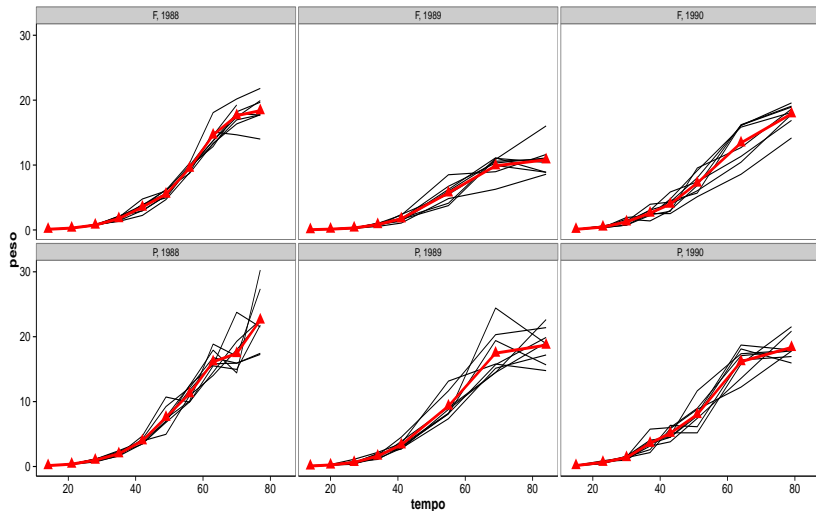


# Perfis individuais por variedade x anos (juntos)





# Perfis individuais e médios por variedade x anos



# Modelo não linear (sem efeito aleatório)

$$Y_j = f(\phi_j, \mathbf{X}_j, \mathbf{W}_j) + \xi_j; j = 1, \dots, n(\text{indivíduo})$$

- Em que  $\mathbf{X}_j$  representa covariáveis (matriz de planejamento) de interesse associadas aos efeitos fixos.
- $\phi_j = \mathbf{X}_j \beta$ .
- $\mathbf{W}_j$ : outras covariáveis como o tempo, por exemplo.
- $f(., ., .)$  é uma função geral, real e diferenciável e não linear em pelo menos uma componente do vetor  $\phi_j$ .
- $\xi_j \stackrel{i.i.d}{\sim} N(0, \sigma^2)$ .

## Modelo não lineares: exemplos

- $(M_1) : Y_j = \phi_1 + \phi_2 \exp(\phi_3/w_j) + \xi_j.$
- $(M_2) : Y_j = \phi_1 - \phi_2 (w_j + \phi_3)^{-1} + \xi_j.$
- $(M_3) : Y_j = \phi_1 w_j^{\phi_2} + \xi_j.$
- $(M_4) : Y_j = \phi_1 / (1 + \exp(-(w_j - \phi_2)/\phi_3)) + \xi_j.$
- $(M_5) : Y_j = \phi_1 + \phi_2 w_j - e^{\phi_3 + \phi_4 w_j} + \xi_j.$

Para os modelos  $M_1, M_2, M_4$ ,  $\phi_j = (\phi_1, \phi_2, \phi_3)' = (\beta_1, \beta_2, \beta_3)'$ , ou seja  $\mathbf{X}_j = \mathbf{I}_3$ . Para o modelo  $M_3$   $\phi_j = (\phi_1, \phi_2)' = (\beta_1, \beta_2)'$ , ou seja  $\mathbf{X}_j = \mathbf{I}_2$ , enquanto que para o modelo  $M_5$   $\phi_j = (\phi_1, \phi_2, \phi_3, \phi_4)' = (\beta_1, \beta_2, \beta_3, \beta_4)'$ , ou seja  $\mathbf{X}_j = \mathbf{I}_4$ . Em todos os casos, a matriz  $\mathbf{W}_j$  corresponde à variável  $w_j$ , que pode ser o tempo.

## Cont.

- Os modelos  $M_1$  e  $M_5$  podem ser apresentados de uma outra forma:

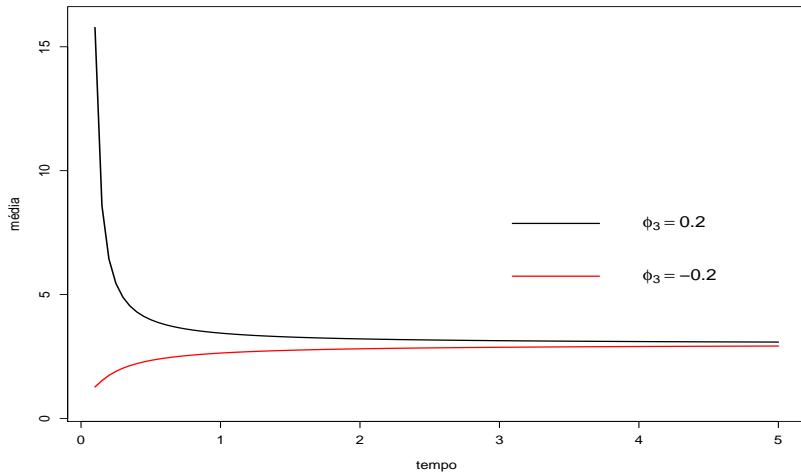
$M_1$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}; \mathbf{X}_j = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & w_j^{-1} \end{bmatrix}$$

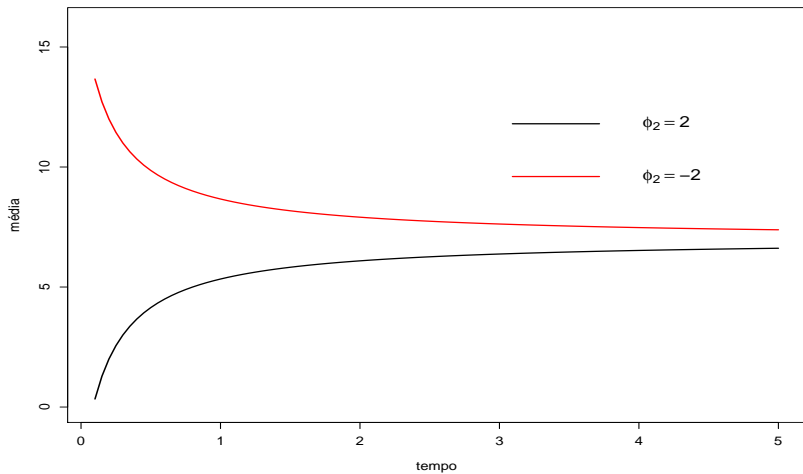
$M_5$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}; \mathbf{X}_j = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & w_j \end{bmatrix}$$

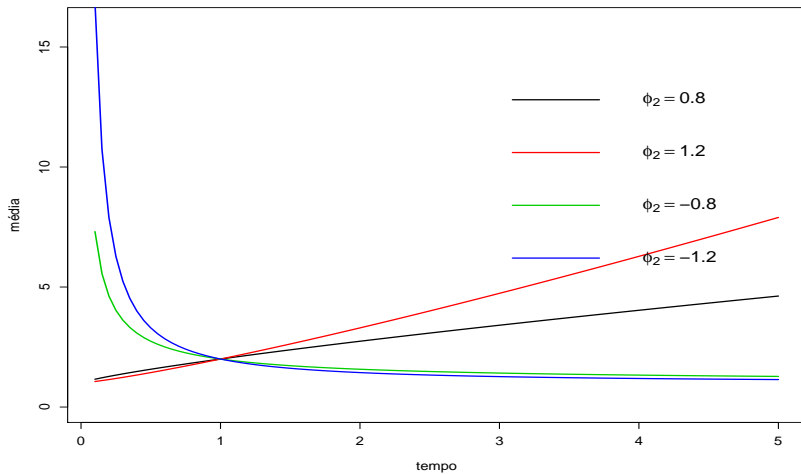
# Modelo 1: $\phi_1 = 1, \phi_2 = 2$



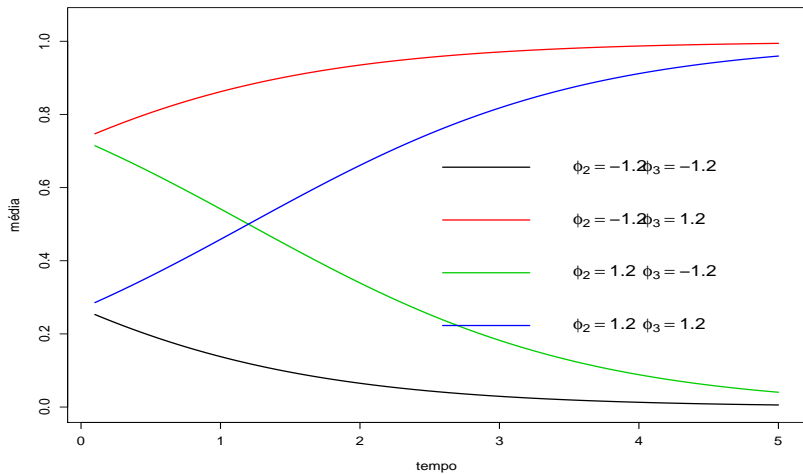
## Modelo 2: $\phi_1 = 7, \phi_3 = 0,2$



# Modelo 3: $\phi_1 = 1$

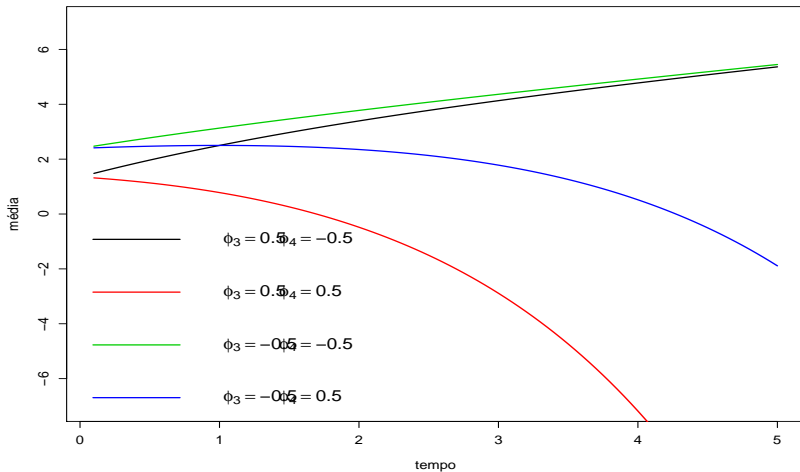


# Modelo 4: $\phi_1 = 1$





# Modelo 5: $\phi_1 = 3, \phi_2 = 0,5$



# Modelo não linear misto

$$Y_{ij} = f(\phi_{ij}, \mathbf{X}_{ij}, \mathbf{W}_{ij}) + \xi_{ij}; j = 1, \dots, n(\text{indivíduo}); i = 1, \dots, k_j(\text{medida repetida})$$

- Em que  $\mathbf{X}_{ij}$  representa covariáveis (matrizes de planejamento) de interesse associadas aos efeitos fixos ( $\mathbf{A}_{ij}$ ) e aleatórios ( $\mathbf{B}_{ij}$ ).
- $\phi_{ij} = \mathbf{A}_{ij}\beta + \mathbf{B}_{ij}\mathbf{b}_j$ .
- $\mathbf{W}_{ij}$  : outras covariáveis como o tempo, por exemplo.
- $f(.,.,.)$  é uma função geral, real e diferenciável e não linear em pelo menos uma componente do vetor  $\phi_{ij}$ .
- $\xi_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$ ,  $\mathbf{b}_j \stackrel{i.i.d.}{\sim} N_q(\mathbf{0}, \Psi)$ ,  $\xi_{ij} \perp \mathbf{b}_j \forall i, j$ .

## Modelo não linear misto (forma vetorial)

$$\mathbf{Y}_j = \mathbf{f}(\phi_j, \mathbf{X}_j, \mathbf{W}_j) + \boldsymbol{\xi}_j; j = 1, \dots, n$$

$$\mathbf{Y}_j = \begin{bmatrix} Y_{1j} \\ Y_{2j} \\ \vdots \\ Y_{kj} \end{bmatrix}; \mathbf{X}_j = \begin{bmatrix} X_{1j} \\ X_{2j} \\ \vdots \\ X_{kj} \end{bmatrix}; \mathbf{A}_j = \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{kj} \end{bmatrix}; \mathbf{B}_j = \begin{bmatrix} B_{1j} \\ B_{2j} \\ \vdots \\ B_{kj} \end{bmatrix};$$

## Modelo não linear misto (forma vetorial) cont.

$$\phi_j = \begin{bmatrix} \phi_{1j} \\ \phi_{2j} \\ \vdots \\ \phi_{k_jj} \end{bmatrix}; \xi_j = \begin{bmatrix} \xi_{1j} \\ \xi_{2j} \\ \vdots \\ \xi_{k_jj} \end{bmatrix}; \mathbf{f}(\phi_j, \mathbf{X}_j, \mathbf{W}_j) = \begin{bmatrix} f(\phi_{1j}, \mathbf{X}_{1j}, \mathbf{W}_{1j}) \\ f(\phi_{2j}, \mathbf{X}_{2j}, \mathbf{W}_{2j}) \\ \vdots \\ f(\phi_{k_jj}, \mathbf{X}_{k_jj}, \mathbf{W}_{k_jj}) \end{bmatrix};$$

$$\mathbf{W}_j = \begin{bmatrix} \mathbf{W}_{1j} \\ \mathbf{W}_{2j} \\ \vdots \\ \mathbf{W}_{k_jj} \end{bmatrix}$$

# Modelo não lineares mistos: exemplos

- $(M_1) : Y_{ij} = \phi_{1j} + \phi_{2j} \exp(\phi_{3j}/w_{ij}) + \xi_{ij}.$
- $(M_2) : Y_{ij} = \phi_{1j} - \phi_{2j} (w_{ij} + \phi_{3j})^{-1} + \xi_{ij}.$
- $(M_3) : Y_{ij} = \phi_{1j} w_{ij}^{\phi_{2j}} + \xi_{ij}.$
- $(M_4) : Y_{ij} = \phi_{1j} / (1 + \exp(-(w_{ij} - \phi_{2j})/\phi_{3j})) + \xi_{ij}.$
- $(M_5) : Y_{ij} = \phi_{1j} + \phi_{2j} w_{1j} - e^{\phi_{3j} + \phi_{4j} w_{ij}} + \xi_{ij}.$

## Modelo 1 misto

■  $Y_{ij} = \phi_{1j} + \phi_{2j} \exp(\phi_{3j}/w_{ij}) + \xi_{ij}$ .

$\phi_{1j} = \beta_1 + b_{1j}$ ,  $\phi_{2j} = \beta_2$  e  $\phi_{3j} = \beta_3$ , nesse caso,  $\beta = (\beta_1, \beta_2, \beta_3)'$ ,

$b_j = b_{1j} \stackrel{i.i.d.}{\sim} N(0, \psi)$  e

$$\phi_j = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} b_{1j} \\ 0 \\ 0 \end{bmatrix}$$

$$A_{ij} = I_3, B_{ij} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

# Modelo 1 misto

- $Y_{ij} = \phi_{1j} + \phi_{2j} \exp(\phi_{3j}/w_{ij}) + \xi_{ij}$ .

$\phi_{1j} = \beta_1 + b_{1j}$ ,  $\phi_{2j} = \beta_2 + b_{2j}$  e  $\phi_{3j} = \beta_3$ , nesse caso,

$\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)'$ ,  $\mathbf{b}_j = (b_{1j}, b_{2j})'$   $i.i.d.$   $N_2(\mathbf{0}, \boldsymbol{\Psi})$  e

$$\boldsymbol{\phi}_j = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} b_{1j} \\ b_{2j} \\ 0 \end{bmatrix}$$

$$\mathbf{A}_{ij} = \mathbf{I}_3, \mathbf{B}_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

# Modelo 1 misto

- $Y_{ij} = \phi_{1j} + \phi_{2j} \exp(\phi_{3j}/w_{ij}) + \xi_{ij}$ .

$\phi_{1j} = \beta_1 + \mathbf{b}_{1j}$ ,  $\phi_{2j} = \beta_2 + \mathbf{b}_{2j}$  e  $\phi_{3j} = \beta_3 + \mathbf{b}_{3j}$ , nesse caso,

$\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)'$ ,  $\mathbf{b}_j = (b_{1j}, b_{2j}, b_{3j})' \stackrel{i.i.d.}{\sim} N_3(\mathbf{0}, \boldsymbol{\Psi})$  e

$$\boldsymbol{\phi}_j = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} b_{1j} \\ b_{2j} \\ b_{3j} \end{bmatrix}$$

$$\mathbf{A}_{ij} = \mathbf{I}_3, \mathbf{B}_{ij} = \mathbf{I}_3$$

Em todos os casos, a matriz  $\mathbf{W}_{ij}$  corresponde à variável  $w_{ij}$ , que pode ser o tempo.



## Modelo 3 misto

- $Y_{ij} = \phi_{1j} w_{ij}^{\phi_{2j}} + \xi_{ij}$ .

$\phi_{1j} = \beta_1 + \mathbf{b}_{1j}$  e  $\phi_{2j} = \beta_2$ , nesse caso,  $\boldsymbol{\beta} = (\beta_1, \beta_2)'$ ,

$\mathbf{b}_j = \mathbf{b}_{1j} \stackrel{i.i.d}{\sim} N(0, \psi)$  e

$$\boldsymbol{\phi}_j = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \mathbf{b}_{1j} \\ 0 \end{bmatrix}$$

$$\mathbf{A}_{ij} = \mathbf{I}_2, \mathbf{B}_{ij} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

## Modelo 3 misto

- $Y_{ij} = \phi_{1j} w_{ij}^{\phi_{2j}} + \xi_{ij}$ .

$\phi_{1j} = \beta_1 + b_{1j}$  e  $\phi_{2j} = \beta_2 + b_{2j}$ , nesse caso,  $\boldsymbol{\beta} = (\beta_1, \beta_2)'$ ,

$\mathbf{b}_j = (b_{1j}, b_{2j}) \stackrel{i.i.d}{\sim} N_2(\mathbf{0}, \boldsymbol{\Psi})$  e

$$\boldsymbol{\phi}_j = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} b_{1j} \\ b_{2j} \end{bmatrix}$$

$$\mathbf{A}_{ij} = \mathbf{I}_2, \mathbf{B}_{ij} = \mathbf{I}_2$$

# Estimação

- Parâmetros para estimar  $(\boldsymbol{\beta}^T, \mathbf{b}^T, \boldsymbol{\theta}^T, \sigma^2)^T$ , em que  $\boldsymbol{\Psi} = \boldsymbol{\Psi}(\boldsymbol{\theta})$ .
- Verossimilhança completa:

$$L(\boldsymbol{\beta}, \mathbf{b}, \boldsymbol{\theta}, \sigma^2) \propto \exp \left\{ - \sum_{j=1}^n \sum_{i=1}^{k_j} \frac{(y_{ij} - f(\boldsymbol{\phi}_{ij}, \mathbf{X}_{ij}, \mathbf{W}_{ij}))^2}{2\sigma^2} \right\} \sigma^{-\sum_{j=1}^n k_j/2}$$
$$\times \exp \left\{ -0,5 \sum_{j=1}^n \mathbf{b}'_j \boldsymbol{\Psi}^{-1} \mathbf{b}_j \right\} |\boldsymbol{\Psi}|^{-n/2}$$

## Verossimilhança marginal

$$L(\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2) \propto \sigma^{-\sum_{j=1}^n k_j/2} \prod_{j=1}^n \int_{\mathbb{R}^q} \exp \left\{ - \sum_{i=1}^{k_j} \frac{(y_{ij} - f(\phi_{ij}, \mathbf{X}_{ij}, \mathbf{W}_{ij}))^2}{2\sigma^2} \right\} \\ \times \exp \{ -0,5 \mathbf{b}_j' \boldsymbol{\Psi}^{-1} \mathbf{b}_j \} |\boldsymbol{\Psi}|^{-1/2} d\mathbf{b}$$

A integral acima não tem solução explícita (lembrando que  $\phi_{ij}$  é função de  $\mathbf{b}_j$ ).

# Resolução da integral

- Quadratura (gaussiana).
- Quadratura adaptativa.
- Aproximação da verossimilhança por um modelo linear misto.
- Aproximação de Laplace.

# Resolução da integral

- Aproximação do integrando usando o primeiro termo da expansão em séries de Taylor em torno do valor esperado condicional dos efeitos aleatórios.
- Aproximação do integrando usando o primeiro termo da expansão em séries de Taylor em torno da moda condicional dos efeitos aleatórios.
- Monte carlo.
- Algoritmo EM, SEM, SAEM, MCEM.

# Discutiremos brevemente

- Quadratura adaptativa.
- Aproximação da verossimilhança por um modelo linear misto (método de Lindstrom and Bates).
- Aproximação de Laplace.

# Verossimilhança

- Vamos utilizar a seguinte decomposição da matriz  $\Psi^{-1} = \sigma^{-2} \mathbf{\Delta}' \mathbf{\Delta}$ , em que  $\mathbf{\Delta}$  (associado à decomposição de Cholesky) é uma matriz de precisão ( $N = \sum_{j=1}^n k_j$ ). Assim:

$$\begin{aligned} L(\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2) &\propto (\sigma^2)^{-(N+nq)/2} \\ &\times \prod_{j=1}^n \int_{\mathbb{R}^q} \exp \left\{ - \sum_{i=1}^{k_j} \frac{(y_{ij} - f(\boldsymbol{\phi}_{ij}, \mathbf{X}_{ij}, \mathbf{W}_{ij}))^2 + \|\mathbf{\Delta} \mathbf{b}_j\|^2}{2\sigma^2} \right\} \\ &\times |\mathbf{\Delta}| d\mathbf{b}_j \end{aligned} \quad (1)$$

em que  $\|\cdot\|^2$  é a norma Euclidiana elevada ao quadrado.



# Método de Lindstrom and Bates

- Tal método consiste em iterar entre dois passos: um passo de mínimos quadrados não lineares penalizado (MQNLP) e um passo de modelos lineares mistos (MLM).
- No passo MQNLP, consideramos uma estimativa provisória de  $\Delta$  e obtemos estimativas provisórias de  $\mathbf{b}_j$  e  $\beta$  minimizando

$$\sum_{j=1}^n \left[ \sum_{i=1}^{k_j} (y_{ij} - f(\phi_{ij}, \mathbf{X}_{ij}, \mathbf{W}_{ij}))^2 + \|\Delta \mathbf{b}_j\|^2 \right]$$

## Cont.

- No passo MLM a matriz  $\mathbf{\Delta}$  é atualizada baseada na expansão de primeira ordem em séries de Taylor da função  $f(., ., .)$  em torno de estimativas provisórias de  $\beta$  e  $\mathbf{b}_j$ , as quais serão denotadas por  $\tilde{\beta}^{(w)}$  e  $\tilde{\mathbf{b}}_j^{(w)}$ . Defina ainda

$$\tilde{\mathbf{w}}_j^{(w)} = \mathbf{y}_j - \mathbf{f}_j(\tilde{\beta}^{(w)}, \tilde{\mathbf{b}}_j^{(w)}) + \tilde{\mathbf{X}}_j^{(w)} \tilde{\beta}^{(w)} + \tilde{\mathbf{Z}}_j^{(w)} \tilde{\mathbf{b}}_j^{(w)}$$

em que  $\tilde{\mathbf{X}}_j^{(w)} = \left. \frac{\partial \mathbf{f}_j}{\partial \beta'} \right|_{\tilde{\beta}^{(w)}, \tilde{\mathbf{b}}_j^{(w)}}$  e  $\tilde{\mathbf{Z}}_j^{(w)} = \left. \frac{\partial \mathbf{f}_j}{\partial \mathbf{b}_j'} \right|_{\tilde{\beta}^{(w)}, \tilde{\mathbf{b}}_j^{(w)}}$

## Cont.

- Assim a log-verossimilhança aproximada (para estimar  $\mathbf{\Delta}$ ) é dada por:

$$\begin{aligned} l_{MLM}(\boldsymbol{\beta}, \mathbf{\Delta}, \sigma^2) &= -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2} \sum_{j=1}^n \left\{ \ln |\boldsymbol{\Sigma}_j(\mathbf{\Delta})| \right. \\ &+ \sigma^{-2} \left[ \mathbf{w}_j^{(w)} - \tilde{\mathbf{X}}_j^{(w)} \tilde{\boldsymbol{\beta}}^{(w)} \right]' \boldsymbol{\Sigma}_j^{-1}(\mathbf{\Delta}) \\ &\times \left. \left[ \mathbf{w}_j^{(w)} - \tilde{\mathbf{X}}_j^{(w)} \tilde{\boldsymbol{\beta}}^{(w)} \right] \right\} \end{aligned} \quad (2)$$

em que  $\boldsymbol{\Sigma}_j(\mathbf{\Delta}) = \mathbf{I} + \tilde{\mathbf{Z}}_j^{(w)} \mathbf{\Delta}^{-1} (\mathbf{\Delta}^{-1})' \tilde{\mathbf{Z}}_j^{(w)'} (veja slide)$

## Cont.

- De modo semelhante ao que é feito nos MLM, podemos obter os valores que maximizam a função acima, em relação à  $\beta$  e  $\sigma^2$ , explicitamente em função de  $\Delta$  e trabalhar com a verossimilhança perfilada de  $\Delta$  para estimá-la.

## Cont.

- Pode-se também trabalhar com a logverossimilhança restrita (veja slides sobre MLM), ou seja:

$$l_R(\sigma^2, \mathbf{\Delta}) = l(\tilde{\beta}(\mathbf{\Delta}), \mathbf{\Delta}, \sigma^2) - \frac{1}{2} \sum_{j=1}^n \ln |\sigma^{-2} \tilde{\mathbf{X}}_j^{(w)'} \mathbf{\Sigma}_j(\mathbf{\Delta}) \tilde{\mathbf{X}}_j^{(w)}|$$

em que  $l(\tilde{\beta}(\mathbf{\Delta}), \mathbf{\Delta}, \sigma^2)$  é dada por (2), com  $\beta$  substituído por  $\tilde{\beta}$ .

- O algoritmo alterna entre os passos MQNLP e MLM, até que algum critério de convergência seja alcançado.
- Para teoria assintótica, IC e Testes de hipótese, veja

[http://www.ime.unicamp.br/~cnaber/aula\\_Mod\\_Lin\\_Misto\\_](http://www.ime.unicamp.br/~cnaber/aula_Mod_Lin_Misto_)

ADI\_2S\_2018.pdf

# Aproximação de Laplace

- A integral que queremos aproximar é a constante na expressão abaixo:

$$L(\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2) \propto \prod_{j=1}^n \int_{\mathbb{R}^q} \exp \left\{ - \sum_{i=1}^{k_j} \frac{(y_{ij} - f(\boldsymbol{\phi}_{ij}, \mathbf{X}_{ij}, \mathbf{W}_{ij}))^2 + \|\boldsymbol{\Delta} \mathbf{b}_j\|^2}{2\sigma^2} \right\} \\ \times (2\pi\sigma^2)^{-(k_j+1)/2} |\boldsymbol{\Delta}| d\mathbf{b}$$

- Seja  $g(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \mathbf{b}_j) = \sum_{i=1}^{k_j} (y_{ij} - f(\boldsymbol{\phi}_{ij}, \mathbf{X}_{ij}, \mathbf{W}_{ij}))^2 + \|\boldsymbol{\Delta} \mathbf{b}_j\|^2$ .

# Aproximação de Laplace

- Defina:

$$\begin{aligned}\tilde{\mathbf{b}}_j &= \operatorname{argmin}_{\mathbf{b}_j} g(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \mathbf{b}_j) \\ g'(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \mathbf{b}_j) &= \frac{\partial g(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \mathbf{b}_j)}{\partial \mathbf{b}_j} \\ g''(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \mathbf{b}_j) &= \frac{\partial^2 g(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \mathbf{b}_j)}{\partial \mathbf{b}_j \partial \mathbf{b}_j'}\end{aligned}$$

- Considere a expansão em séries de Taylor de segunda ordem de  $g$  em torno de  $\tilde{\mathbf{b}}_j$ :

$$g(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \mathbf{b}_j) \approx g(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \tilde{\mathbf{b}}_j) + \frac{1}{2} (\mathbf{b}_j - \tilde{\mathbf{b}}_j)' g''(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \tilde{\mathbf{b}}_j) (\mathbf{b}_j - \tilde{\mathbf{b}}_j)$$

(note que  $g'(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \tilde{\mathbf{b}}_j) = \mathbf{0}$ )

## Cont.

- A aproximação de Laplace da verossimilhança é dada por:

$$L(\boldsymbol{\beta}, \boldsymbol{\theta}, \sigma^2) \approx (2\pi\sigma^2)^{-N/2} |\boldsymbol{\Delta}|^n \\ \times \exp \left[ -\frac{1}{2\sigma^2} \sum_{j=1}^n g(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \tilde{\mathbf{b}}_j) \right] \prod_{i=1}^n |g''(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \tilde{\mathbf{b}}_j)|^{-1/2}$$

- Temos que (veja Pinheiro and Bates (2009), pag. 316-317)

$$g''(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \tilde{\mathbf{b}}_j) \approx \mathbf{G}(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j) = \frac{\partial \mathbf{f}_j(\boldsymbol{\beta}, \mathbf{b}_j)}{\partial \mathbf{b}_j} \bigg|_{\tilde{\mathbf{b}}_j} \frac{\partial \mathbf{f}_j(\boldsymbol{\beta}, \mathbf{b}_j)}{\partial \mathbf{b}'_j} \bigg|_{\tilde{\mathbf{b}}_j} + \boldsymbol{\Delta}' \boldsymbol{\Delta}$$



- A aproximação de Laplace da logverossimilhança modificada é dada por:

$$l_{LA}(\beta, \sigma^2, \mathbf{\Delta}) = -\frac{N}{2} + n \ln |\mathbf{\Delta}| - \frac{1}{2} \left\{ \sum_{j=1}^n \ln |G(\beta, \mathbf{\Delta}, \mathbf{y}_j)| + \sigma^{-2} \sum_{j=1}^n g(\beta, \mathbf{\Delta}, \mathbf{y}_j, \tilde{\mathbf{b}}_j) \right\}$$

- O procedimento itera entre a maximização de  $l_{LA}$  em relação à  $\beta, \sigma^2, \mathbf{\Delta}$  e da minimização de  $g(\beta, \mathbf{\Delta}, \mathbf{y}_j)$  com relação à  $\mathbf{b}_j$ .

# Quadratura adaptativa

- A idéia é utilizar parte dos resultados da AL e substituir a integral de uma normal multivariada por sucessivas integrais de normais padrão independentes.
- Sejam  $z_j, w_j, j = 1, 2, \dots, N_Q$ , respectivamente, as abscissas e os pesos para a integração por quadratura Gaussiana (unidimensional) baseado na  $N(0,1)$ .
- Do resultado da AL temos que o integrando  $\exp(-g(\beta, \Delta, \mathbf{y}_j, \mathbf{b}_j)/(2\sigma^2))$  pode ser aproximado por uma distribuição  $N_q(\tilde{\mathbf{b}}_j, \sigma^2 G^{-1}(\beta, \Delta, \mathbf{y}_j))$

# Quadratura adaptativa

■ Assim

$$\begin{aligned} & \int \exp(-g(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \mathbf{b}_j)/(2\sigma^2)) d\mathbf{b}_j = \int \sigma^q |\mathbf{G}(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j)|^{-1/2} \\ & \times \exp\left\{-g\left[\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \tilde{\mathbf{b}}_j + \sigma \mathbf{G}^{-1/2}(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j) \mathbf{z}\right] / (2\sigma^2) + \|\mathbf{z}\|^2/2\right\} \\ & \times \exp(-\|\mathbf{z}\|^2/2) d\mathbf{z} \\ & \approx \sigma^q |\mathbf{G}(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j)|^{-1/2} \\ & \times \sum_{j_1=1}^{N_Q} \dots \sum_{j_q=1}^{N_Q} \exp\left\{-g\left[\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \tilde{\mathbf{b}}_j + \sigma \mathbf{G}^{-1/2}(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j) \mathbf{z}_j\right] / (2\sigma^2) \right. \\ & \left. + \|\mathbf{z}_j\|^2/2\right\} \prod_{k=1}^q w_{jk} \end{aligned}$$

em que  $\mathbf{z}_j = (z_{j1}, \dots, z_{jq})$ .

## Cont.

- Aproximação da logverossimilhança por quadratura adaptativa:

$$\begin{aligned}l_{AQA}(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\Delta}) &= -\frac{N}{2} \ln(2\pi\sigma^2) + n \ln |\boldsymbol{\Delta}| - \frac{1}{2} \sum_{j=1}^n \ln |\mathbf{G}(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j)| \\ &= \sum_{j=1}^n \left( \sum_{j_1=1}^{N_Q} \dots \sum_{j_q=1}^{N_Q} \exp \left\{ -g \left[ \boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j, \tilde{\mathbf{b}}_j \right. \right. \right. \\ &\quad \left. \left. \left. + \sigma \mathbf{G}^{-1/2}(\boldsymbol{\beta}, \boldsymbol{\Delta}, \mathbf{y}_j) \mathbf{z}_j \right] / (2\sigma^2) \right. \right. \\ &\quad \left. \left. \left. + \|\mathbf{z}_j\|^2 / 2 \right\} \prod_{k=1}^q w_{jk} \right)\end{aligned}$$

- A logverossimilhança acima pode ser maximizada através de algoritmos de otimização adequados.

# Resíduos normalizados

- $R_j = D_j \frac{Y_j - f(\hat{\phi}_j, X_j, W_j)}{\sqrt{\hat{\sigma}^2}}$ ,  $j = 1, 2, \dots, n$ , em que  $D_j$  corresponde ao inverso da matriz da decomposição de Cholesky da matriz de correlação induzida pelo modelo (depende da estrutura dos efeitos aleatórios).

# Recursos computacionais

- As funções “nls” (default no R) e “nlme” (pacote nlme) ajustam, respectivamente, modelos não lineares e modelos não lineares mistos.

- A primeira estima os parâmetros via mínimos quadrados

(ponderados) não lineares, minimização de

$$\sum_{j=1}^n (Y_j - f(\phi_j, \mathbf{X}_j, \mathbf{W}_j))^2, \text{ em relação à } \beta.$$

- Estimador para o  $\sigma^2$  (função nls) :

$$\hat{\sigma}^2 = \frac{1}{n-p} \sum_{j=1}^n \left( Y_j - f(\hat{\phi}_j, \mathbf{X}_j, \mathbf{W}_j) \right)^2.$$

# Recursos computacionais

- A segunda o faz usando máxima verossimilhança ou máxima verossimilhança restrita (método de Lindstrom and Bates, veja [https://www.jstor.org/stable/2532087?seq=1#metadata\\_info\\_tab\\_contents](https://www.jstor.org/stable/2532087?seq=1#metadata_info_tab_contents)).
- Um ponto interessante é que ambas solicitam que o usuário forneçam estimativas iniciais para os parâmetros do preditor não linear.

## Cont.

- Há uma certa flexibilidade na escolha do preditor não linear.
- Ele pode ser escrito diretamente na função ou ser inserido como uma função.
- Em geral é melhor entrar com o preditor não linear como uma função na qual conste a respectiva derivada (auxilia na convergência dos algoritmos).
- Há a possibilidade de criar uma outra função através da função “selfStart” na qual é possível inserir o cálculo automático das estimativas iniciais.



## Cont.

- Há vários preditores não lineares implementados com esse pacote (derivada e estimativa iniciais) já estão disponíveis na função (*input* : covariável, *dose* : valor inicial de interesse).

Nome	Modelo
SSasymp	$Asym + (R0 - Asym) * \exp(-\exp(lrc) * input)$
SSasympOff	$Asym * (1 - \exp(-\exp(lrc) * (input - c0)))$
SSasympOrig	$Asym * (1 - \exp(-\exp(lrc) * input))$
SSbiexp	$A1 * \exp(-\exp(lrc1) * input) + A2 * \exp(-\exp(lrc2) * input)$
SSfol	$Dose * \exp(IKe + IKa - ICl) * (\exp(-\exp(IKe) * input) - \exp(-\exp(IKa) * input)) * (\exp(IKa) - \exp(IKe))$

## Cont.

- ( $x$  : covariável)

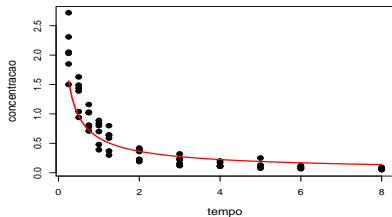
Nome	Modelo
SSfpl	$A + (B - A)/(1 + \exp((xmid - input)/scal))$
SSgompertz	$Asym * \exp(-b2 * b3^x)$
SSlogis	$Asym/(1 + \exp((xmid - input)/scal))$
SSmicmen	$Vm * input/(K + input)$
SSweibull	$Asym - Drop * \exp(-\exp(lrc) * x^{pwr})$

## Voltando ao Exemplo 6

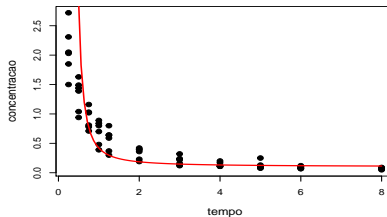
- Ajustou-se os modelos  $M_1$ ,  $M_2$  e  $M_3$  (não houve convergência para os modelo  $M_4$  e  $M_5$ )
- $(M_1) : Y_{ij} = \phi_{1j} + \phi_2 \exp(\phi_3/w_{ij}) + \xi_{ij}$ .
- $(M_2) : Y_{ij} = \phi_{1j} - \phi_2 (w_{ij} + \phi_3)^{-1} + \xi_{ij}$ .
- $(M_3) : Y_{ij} = \phi_{1j} w_{ij}^{\phi_2} + \xi_{ij}$ .
- $\phi_{1j} = \phi_1 + b_j$ ,  $b_j \stackrel{i.i.d.}{\sim} N(0, \psi)$ .
- $\xi_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$ ,  $\xi_{ij} \perp b_j$ .
- Em todos os modelos,  $i = 1, 2, \dots, 11$  (condição de avaliação),  
 $j = 1, 2, \dots, 6$  (indivíduo).

# Sobre os valores iniciais do processo iterativo

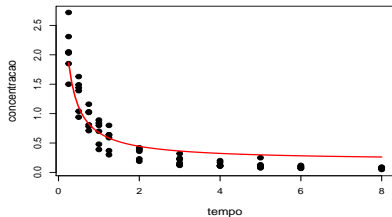
M3,  $\phi_1 = 0.59, \phi_2 = -0.7$



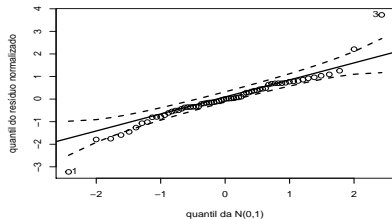
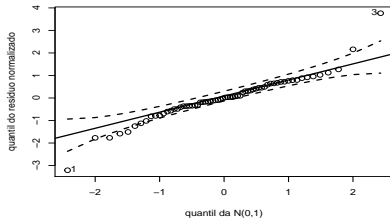
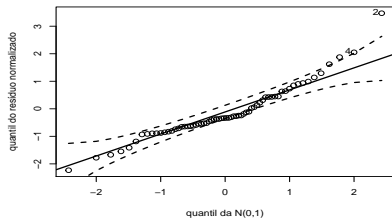
M1,  $\phi_1 = 0.05, \phi_2 = 0.05, \phi_3 = 2$



M2,  $\phi_1 = 0.2, \phi_2 = -0.5, \phi_3 = 0.05$



# Resíduos normalizados com envelopes baseados na $N(0,1)$



# Resíduos normalizados com envelopes baseados no modelo

Gráfico de quantil-quantil (modelo M3)

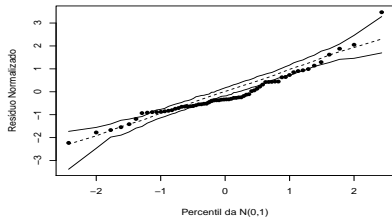


Gráfico de quantil-quantil (modelo M1)

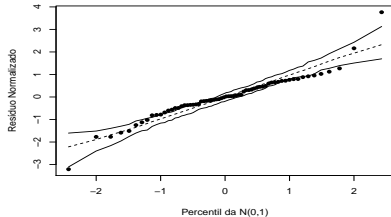
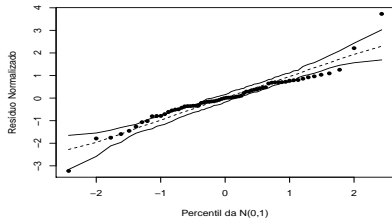
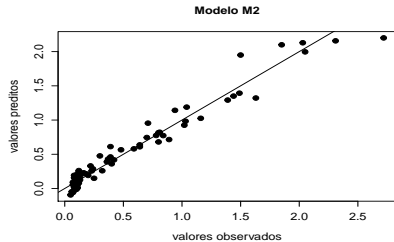
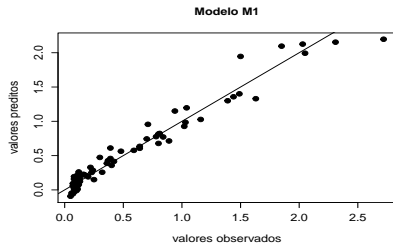
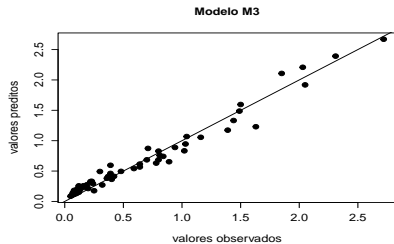


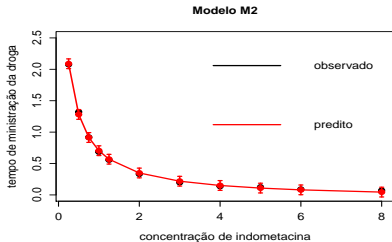
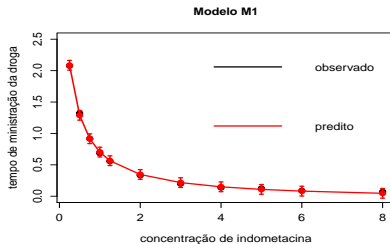
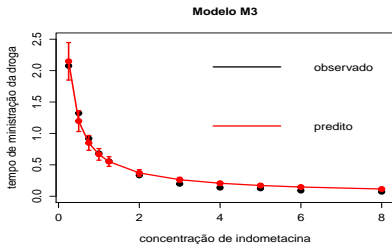
Gráfico de quantil-quantil (modelo M2)



# Valores individuais observados e preditos



# Médias observadas e previstas (via valores previstos indiv.)



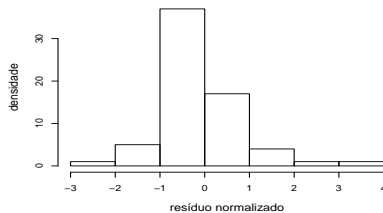
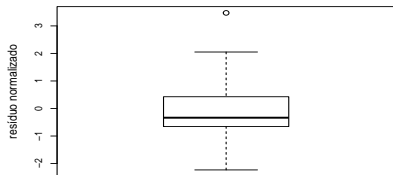
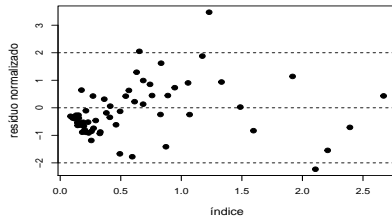
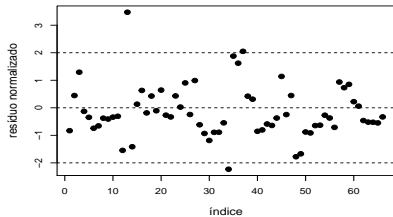


# Estatísticas de comparação de modelos e somas de quadrados de resíduos

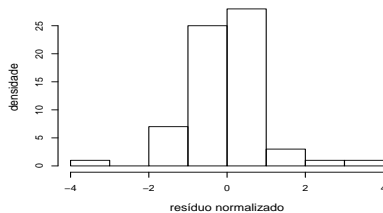
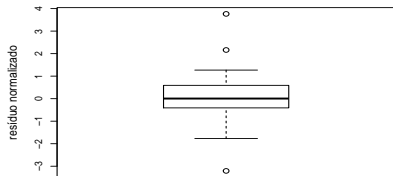
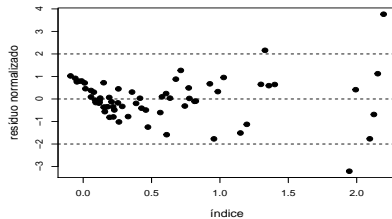
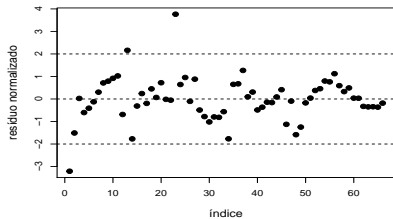
	AIC	BIC
$M_3$	-72,95	-64,20
$M_2$	-51,12	-41,18
$M_1$	-51,80	-40,85

$M_3$	$M_2$	$M_1$
0,915	0,924	0,924

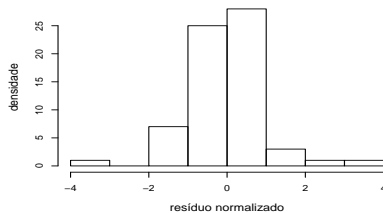
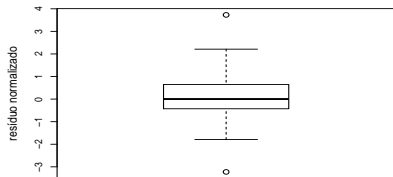
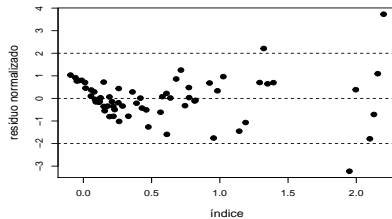
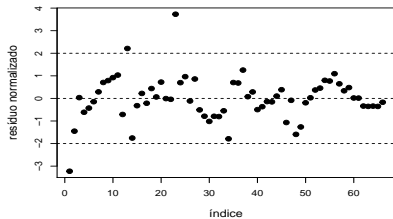
# Gráficos de resíduos normalizados M3



# Gráficos de resíduos normalizados M1



# Gráficos de resíduos normalizados M2



# Estimativa dos parâmetros do modelo M1

Parâmetro	Est.	EP	Estat. t	IC (95%)	p-valor
$\phi_1$	3,23	0,41	2,42	[4,04 ; 7,84]	< 0,0001
$\phi_2$	-3,29	0,39	-4,06	[-2,53 ; -8,42]	< 0,0001
$\phi_3$	-0,26	0,05	-0,37	[-0,16 ; -5,09]	< 0,0001

Interpretação: valores esperados quando  $w \rightarrow 0$ ,  $\mathcal{E}(Y) = \widetilde{\phi}_1 = 3,23$  e  $w \rightarrow \infty$ ,  $\mathcal{E}(Y) = \widetilde{\phi}_1 + \widetilde{\phi}_2 = -0,06$ .

# Estrutura de variância e covariância (M1)

- $\mathcal{E}(Y_{ij}) = \phi_1 + b_j + \phi_2 \exp(\phi_3/w_{ij}) + \xi_{ij} =$   
 $\phi_1\phi_2 \exp(\phi_3/w_{ij}) + \mathcal{E}(b_j) + EM(\xi_{ij}) = \phi_1\phi_2 \exp(\phi_3/w_{ij}).$
- $\mathcal{V}(Y_{ij}) = \mathcal{V}(\phi_1 + b_j + \phi_2 \exp(\phi_3/w_{ij}) + \xi_{ij}) = \mathcal{V}(b_j) + \mathcal{V}(\xi_{ij}) = \psi + \sigma^2.$
- $Cov(Y_{ij}, Y_{i'j}) =$   
 $Cov(\phi_1 + b_j + \phi_2 \exp(\phi_3/w_{ij}) + \xi_{ij}, \phi_1 + b_j + \phi_2 \exp(\phi_3/w_{ij}) + \xi_{ij})$   
 $= \mathcal{V}(b_j) = \sigma^2.$
- Exercício: fazer o mesmo para os modelos M1 e M3.

# Inferência Bayesiana

- Similar à abordagem apresentada para os os modelos lineares mistos e generalizados mistos.
- Verossimilhança (completa):

$$L(\beta, \phi, \sigma^2) = \phi^{\sum_{j=1}^n k_j/2} \exp \left\{ - \sum_{j=1}^n \sum_{i=1}^{k_j} \frac{(y_{ij} - f(\phi_{ij}, \mathbf{X}_{ij}, \mathbf{W}_{ij}))^2}{2} \right\} \\ \times \exp \left\{ - \sum_{j=1}^n \frac{b_j^2}{2\sigma^2} \right\} \sigma^{-n/2} \quad (3)$$

# Inferência Bayesiana

- Prioris usuais:  $\beta \sim N_p(\mu_\beta, \Psi_\beta)$  (em que, usualmente,  $\Psi_\beta = \text{diag}(\psi_{\beta_1}, \dots, \psi_{\beta_p})$ ,  $\sigma^2 \sim IG(\alpha_{\sigma^2}, \beta_{\sigma^2})$  e  $\psi \sim IG(\alpha_\phi, \beta_\phi)$  (ou gama, para esses dois últimos)).
- Priori conjunta :  $p(\beta, \sigma^2, \phi) = p(\beta)p(\sigma^2)p(\phi)$
- Posteriori:  $p(\beta, \phi, \sigma^2 | \mathbf{y}, \mathbf{x}) \propto L(\beta, \phi, \sigma^2)p(\beta)p(\sigma^2)p(\phi)$ .
- A obtenção das posterioris marginais não é analiticamente viável.  
Uma solução: utilizar algoritmos MCMC.



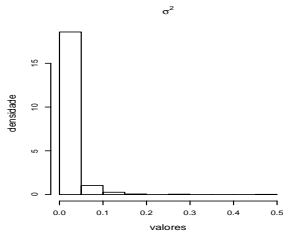
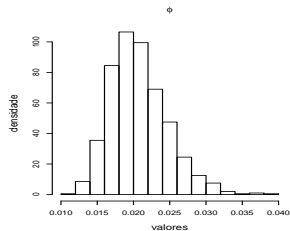
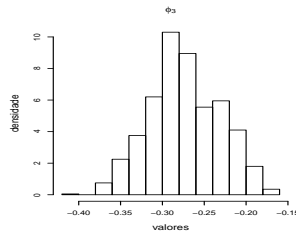
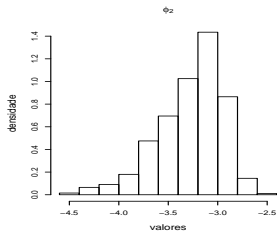
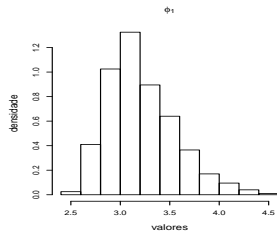
# Análise dos dados de indometacina

- Modelo 1 (M1), com as seguintes priors:  $\phi_1 \sim N(1, 100)$ ,  
 $\phi_2 \sim N(1, 100)$ ,  $\phi_1 \sim N(0, 2; 100)$ ,  $\sigma^2 \sim \text{gama}(0, 1; 0, 1)$ ,  
 $\phi \sim \text{gama}(0, 1; 0, 1)$ (escolha também pautada por questões de convergência).
- Análise de convergência: três conjuntos de cadeias, de tamanho 50000, foram geradas, de forma independente, a partir de pontos (diferentes) gerados de forma aleatória. Resultados <http://www.ime.unicamp.br/~cnaber/MNL1Normal.zip>.

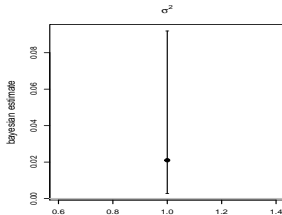
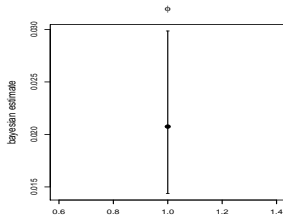
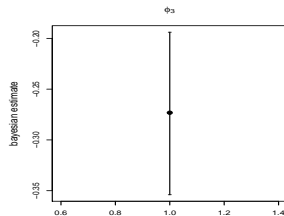
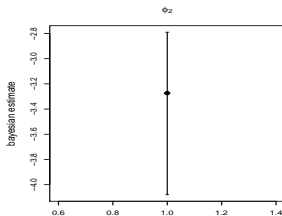
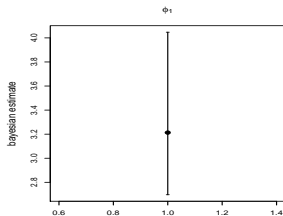
# Análise dos dados de indometacina

- Parâmetros MCMC: burn-in = 50000, thin(lag) = 50, número total de valor = 100000, o que gera uma amostra válida das posteriores (após burn-in, com o que devido thin) de 1000 valores.

# Distribuições a posteriori



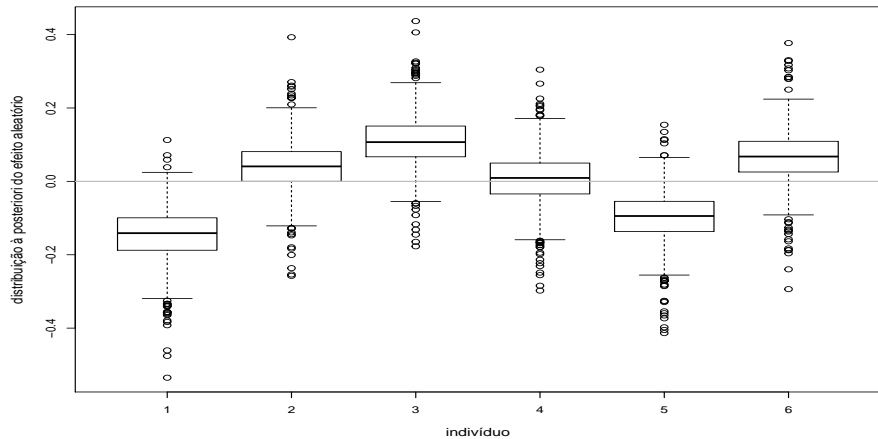
# Estimativas pontuais e intervalos de credibilidade



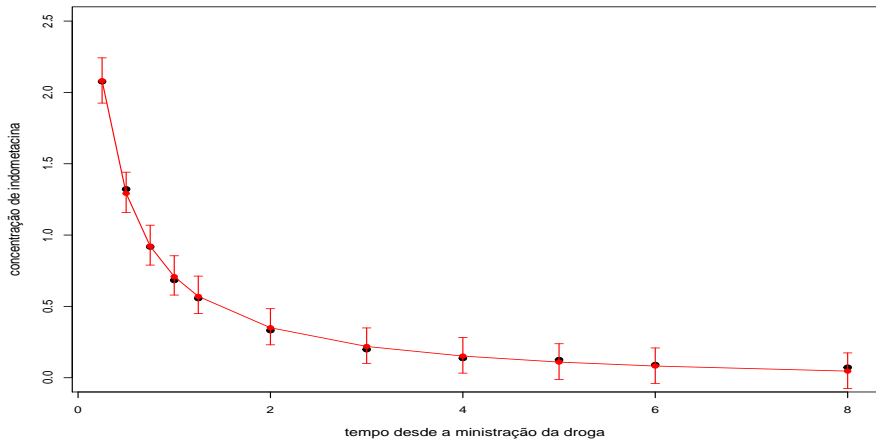
# Estimativas pontuais e intervalos de credibilidade

Parâmetro	EAP	MedAP	MoAP	DPAP	IC(%95)
$\phi_1$	3,21	3,15	3,13	0,35	[2,70 ; 4,05]
$\phi_2$	-3,27	-3,21	-3,18	0,33	[-4,08 ; -2,79]
$\phi_3$	-0,27	-0,28	-0,27	0,04	[-0,35 ; -0,19]
$\phi$	0,02	0,02	0,02	0,00	[ 0,01 ; 0,03]
$\sigma^2$	0,02	0,01	0,01	0,03	[< 0,01 ; 0,09]

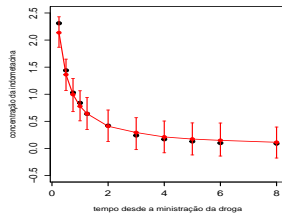
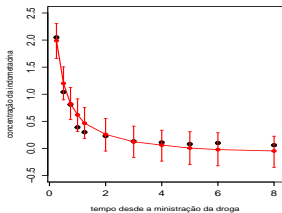
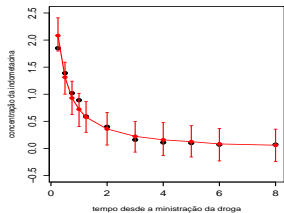
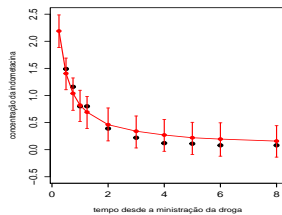
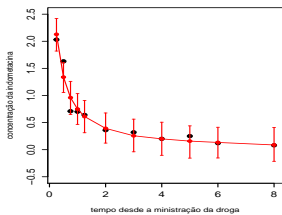
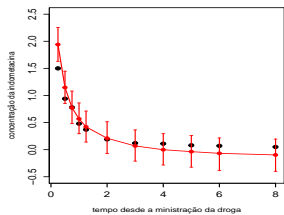
# Posteriori dos efeitos aleatórios



# Médias observadas (preto) e preditos (vermelho)

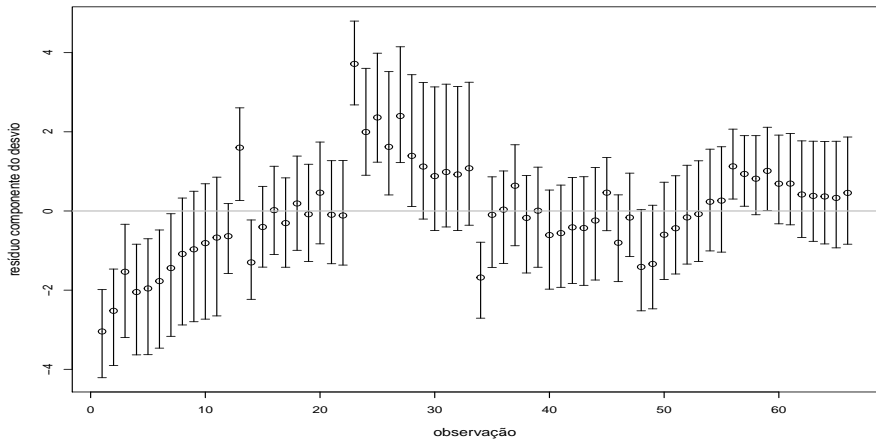


# Valores observados (preto) e preditos (vermelho)





# Mediana a posteriori e IC (95%) do RN



# Resíduos normalizados com envelopes baseados no modelo

Gráfico de quantil-quantil (modelo M3)

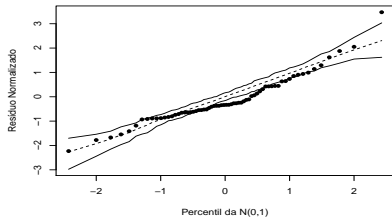


Gráfico de quantil-quantil (modelo M1)

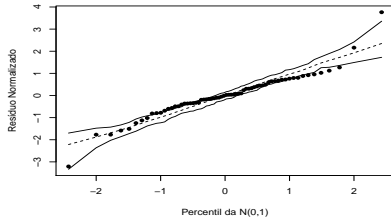
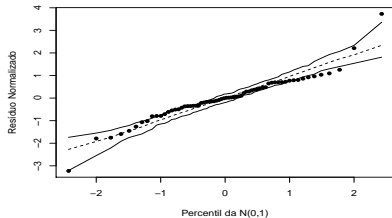
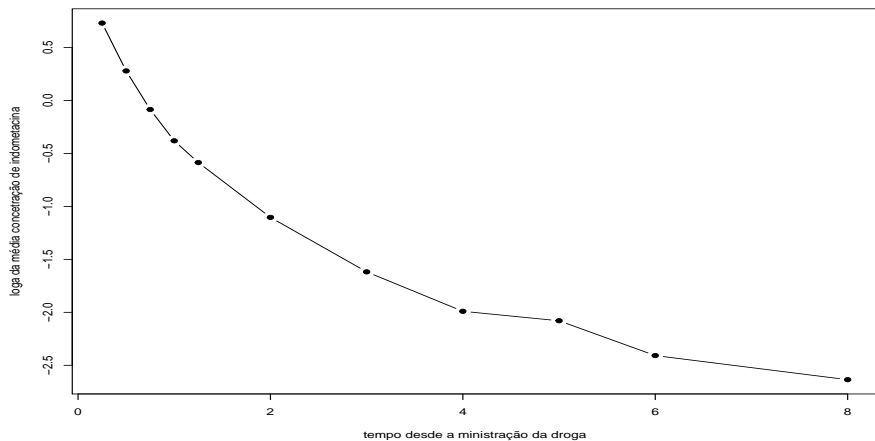


Gráfico de quantil-quantil (modelo M2)



# Log das médias das concentrações $\times$ tempo



# Dados de indometacina: modelo gama

- Vamos ajustar um modelo gama log-linear misto.

$$Y_{ij}|b_j \stackrel{ind.}{\sim} \text{gama}(\mu_{ij}, \phi)$$

$$\log(\mu_{ij}) = \beta_0 + \beta_1 w_{ij}, j = 1, 2, \dots, 6 (\text{indiv\u00edduo})$$

$, i = 1, \dots, 11 (\text{tempo de medi\u00e7\u00e3o (condi\u00e7\u00e3o de avalia\u00e7\u00e3o)})$

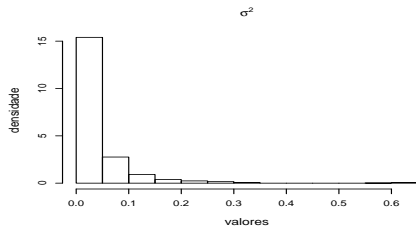
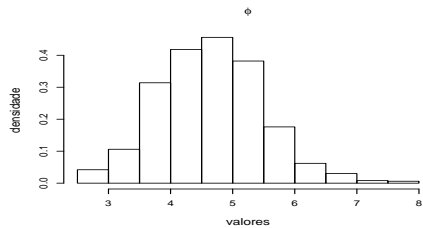
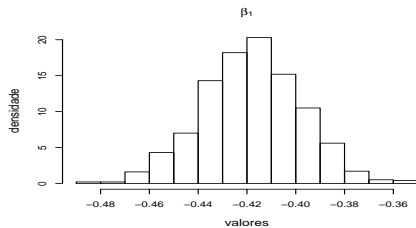
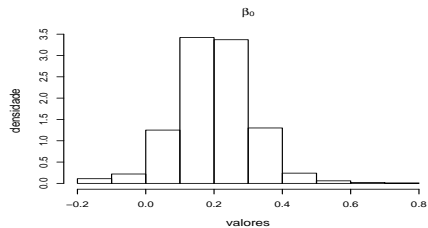
$$b_j \stackrel{i.i.d}{\sim} N(0, \sigma^2)$$

em que  $\mathcal{E}(Y_{ij}|b_j) = \mu_{ij}$ ,  $\mathcal{V}(Y_{ij}|b_j) = \frac{\mu_{ij}^2}{\phi}$ . Esperan\u00e7as e vari\u00e2ncias marginais s\u00e3o an\u00e1logas \u00e0quelas apresentadas nos slides

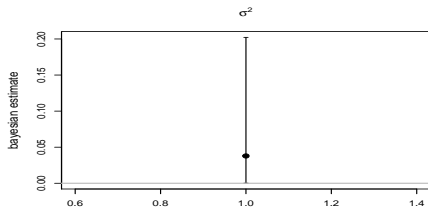
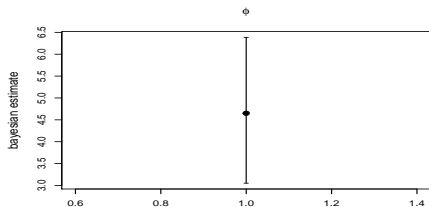
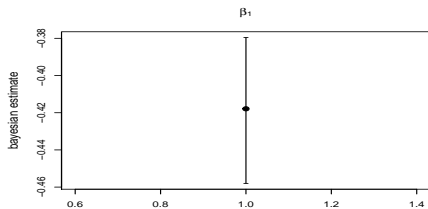
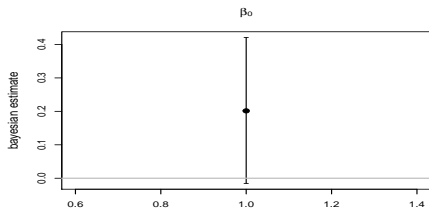
# Parâmetros MCMC

- Priors:  $\beta_0 \sim N(0, 100)$ ,  $\beta_1 \sim N(0, 100)$ ,  $\sigma^2 \sim \text{gama}(0, 1; 0, 1)$ ,  $\phi \sim \text{gama}(0, 1; 0, 1)$ (escolha também pautada por questões de convergência).
- Análise de convergência: três conjuntos de cadeias, de tamanho 50000, foram geradas, de forma independente, a partir de pontos (diferentes) gerados de forma aleatória. Resultados <http://www.ime.unicamp.br/~cnaber/MNL1Gama.zip>.
- Parâmetros MCMC: burn-in = 10000, thin(lag) = 30, número total de valor = 40000, o que gera uma amostra válida das posteriores (após burn-in, com o que devido thin) de 1000 valores.

# Distribuições a posteriori



# Estimativas pontuais e intervalos de credibilidade

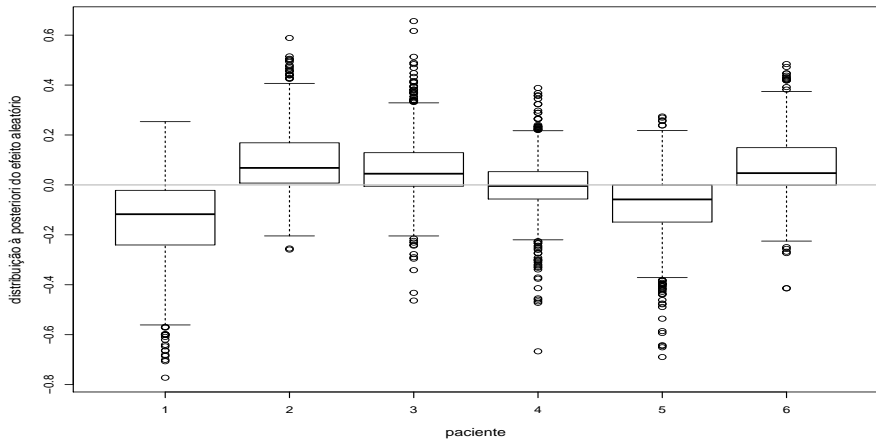


# Estimativas pontuais e intervalos de credibilidade

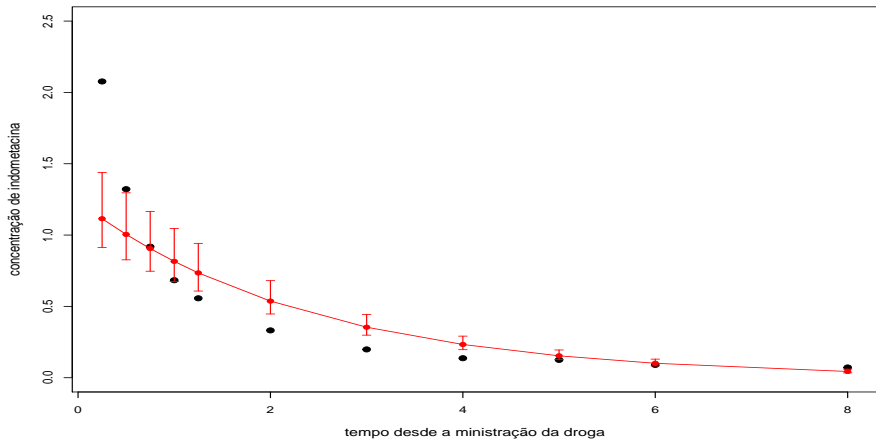
Parâmetro	EAP	MedAP	MoAP	DPAP	IC(%95)
$\phi_1$	3,21	3,15	3,13	0,35	[2,70 ; 4,05]
$\phi_2$	-3,27	-3,21	-3,18	0,33	[-4,08 ; -2,79]
$\phi_3$	-0,27	-0,28	-0,27	0,04	[-0,35 ; -0,19]
$\phi$	0,02	0,02	0,02	0,00	[ 0,01 ; 0,03]
$\sigma^2$	0,02	0,01	0,01	0,03	[< 0,01 ; 0,09]



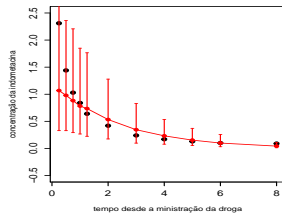
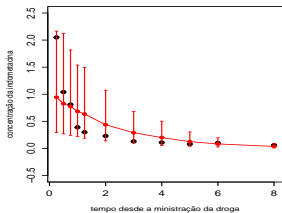
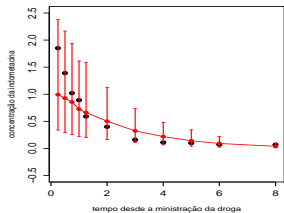
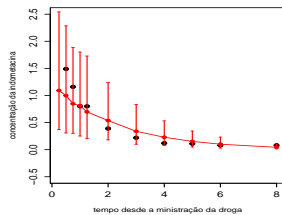
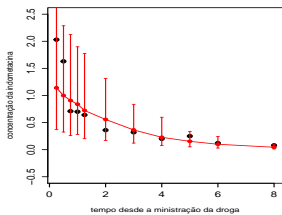
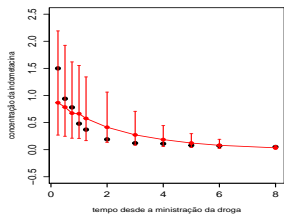
# Posteriori dos efeitos aleatórios



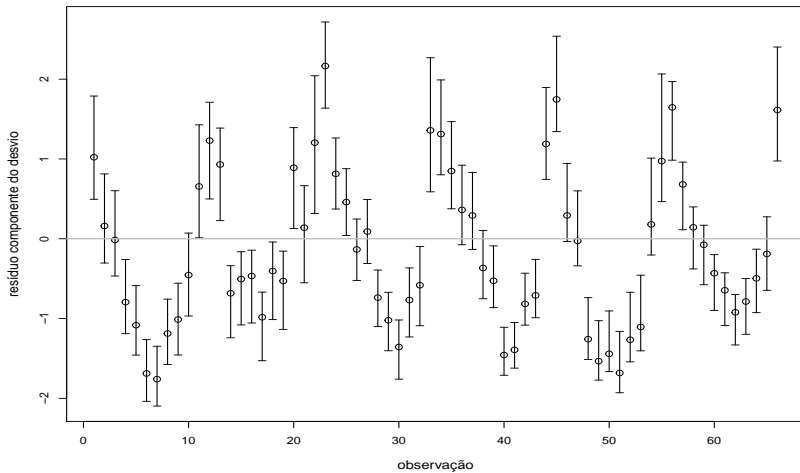
# Médias observadas (preto) e preditos (vermelho)



# Valores observados (preto) e preditos (vermelho)



# Mediana a posteriori e IC (95%) para o rcd



# Exercício

- Proponha um modelo linear generalizado misto não-linear para dos dados de indometacina, interpretando os parâmetros, ajustando o modelo e verificando a qualidade do ajuste, do ponto de vista bayesiano.