

# Inferência Bayesiana para Modelos Hierárquicos: parte 2

Prof. Caio Azevedo

## Exemplo dos dados de Potthof and Roy

$$Y_{jik} = \mu_{jik} + \xi_{ijk}, (\text{medida repetida, nível 1}),$$

$$\mu_{jik} = \mu_{jk} + \alpha_k + (\beta_1 + \gamma_k)(x_{jik} - 8), (\text{medida repetida, nível 1})$$

$$\mu_{jk} = \beta_0 + u_{jk} (\text{indivíduo, nível 2})$$

$$\xi_{jik} \stackrel{i.i.d.}{\sim} N(0, \sigma^2), u_{jk} \stackrel{i.i.d.}{\sim} N(0, \psi), \alpha_1 = \gamma_1 = 0$$

$$j = 1, 2, \dots, n_k, i = 1, 2, 3, 4 (\text{ano (condição de avaliação)}),$$

$$k = 1, 2 (\text{sexo - 1: feminino, 2: masculino}), n_{i1} = 11; n_{i2} = 16, \forall i$$

- As interpretações dos parâmetros são como dadas [aqui](#).
- Outros modelos podem ser ajustados (gama e log normal).

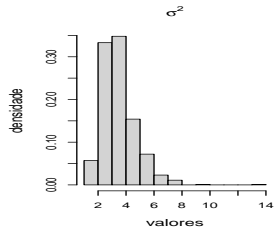
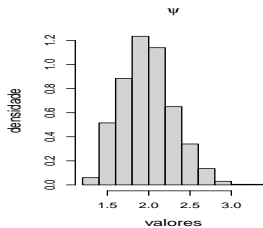
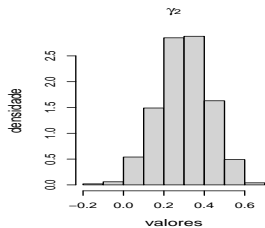
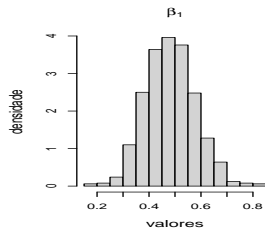
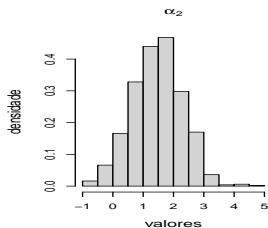
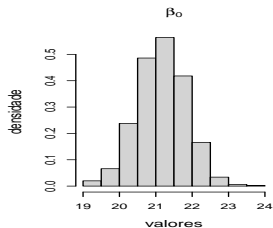
## Exemplo dos dados de Potthof and Roy

- Parâmetros MCMC: burn-in = 5000, thin(lag) = 50, número total de iterações = 55000, o que gera uma amostra válida das posteriores (após burn-in, com o que devido thin) de 1000 valores. Resultados [aqui](#).
- Prioris:  $\beta_0 \sim N(0, 1000)$ ,  $\alpha_2 \sim N(0, 1000)$ ,  $\beta_1 \sim N(0, 1000)$ ,  $\gamma_2 \sim N(0, 1000)$ ,  $\sigma^2 \sim \text{gama}(0.01, 0.01)$ ,  $\psi \sim \text{gama}(0.05, 0.05)$  ( $X \sim \text{gama}(a, b), \mathcal{E}(X) = a/b$ ).

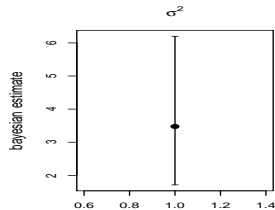
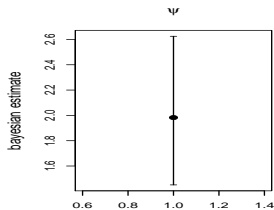
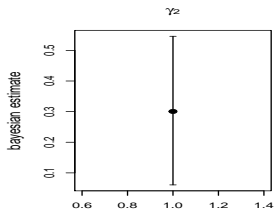
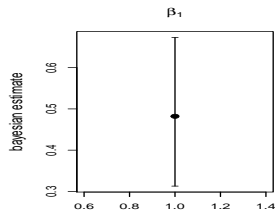
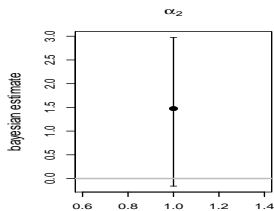
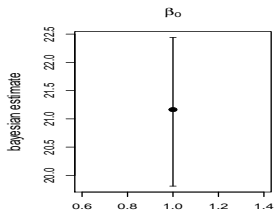
# Exemplo dos dados de Potthof and Roy

- Questões de convergência. Provavelmente, devido à (elevada) correlação (à posteriori) entre os interceptos fixos e os aleatórios. Possíveis soluções: utilizar algoritmos auxiliares mais apropriados, reparametrizações entre outras.

# Distribuições a posteriori



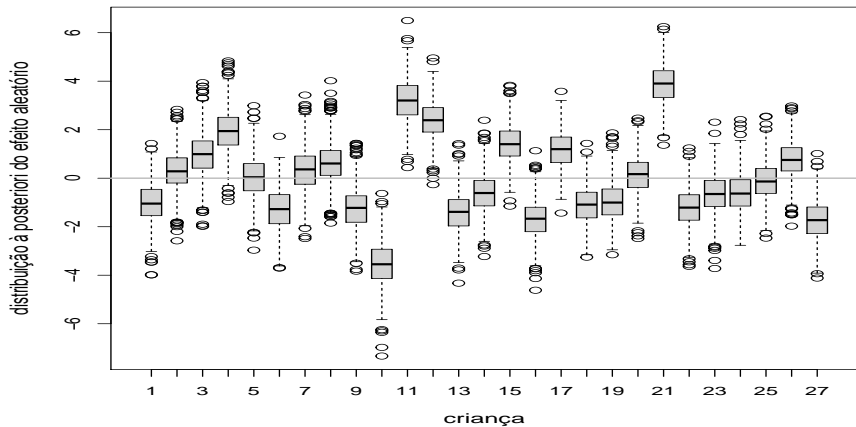
# Estimativas pontuais e intervalos de credibilidade



# Estimativas pontuais e intervalos de credibilidade

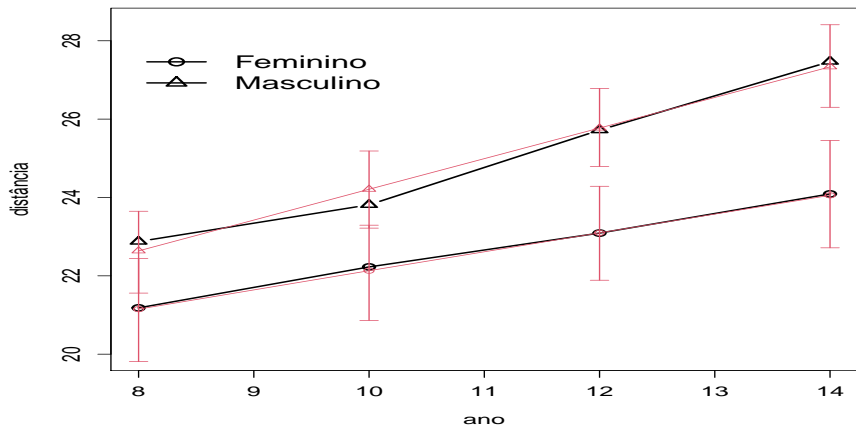
Parâmetro	EAP	MedAP	MoAP	DPAP	IC(%95)
$\beta_0$	21,19	21,21	21,31	0,64	[19,90 ; 22,42]
$\alpha_2$	1,43	1,43	1,11	0,86	[-0,29 ; 3,31]
$\beta_1$	0,48	0,48	0,46	0,09	[0,30 ; 0,66]
$\gamma_2$	0,30	0,29	0,25	0,12	[0,08 ; 0,54]
$\psi$	1,97	1,94	1,88	0,31	[1,46 ; 2,63]
$\sigma^2$	3,57	3,37	2,57	1,26	[1,72 ; 6,59]

# Posteriori dos efeitos aleatórios

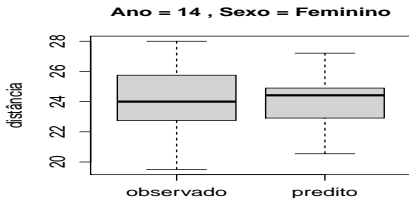
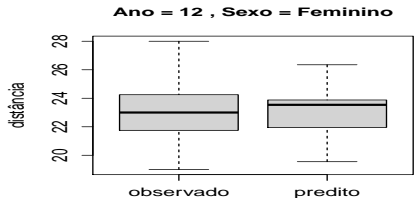
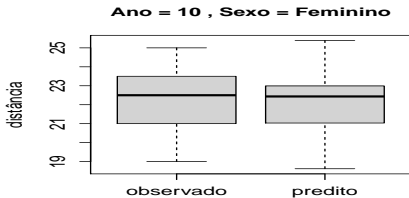
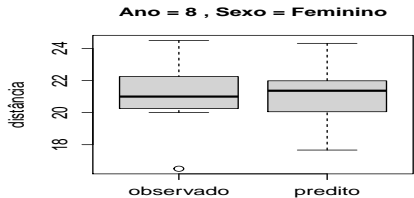




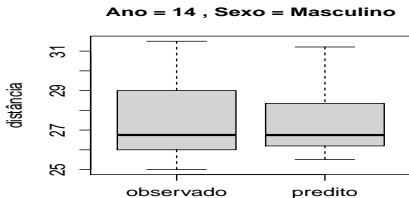
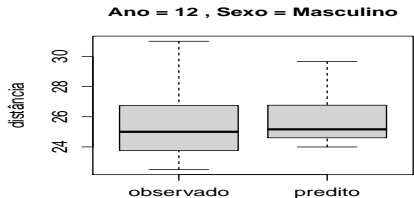
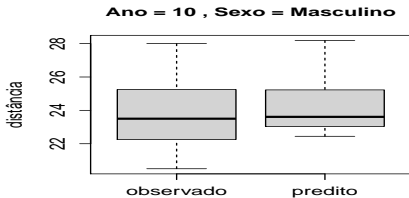
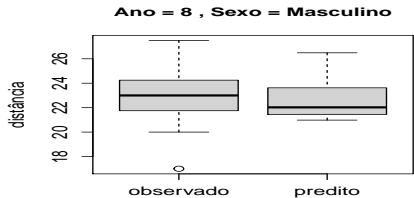
## Valores observados (preto) e preditos (vermelho)



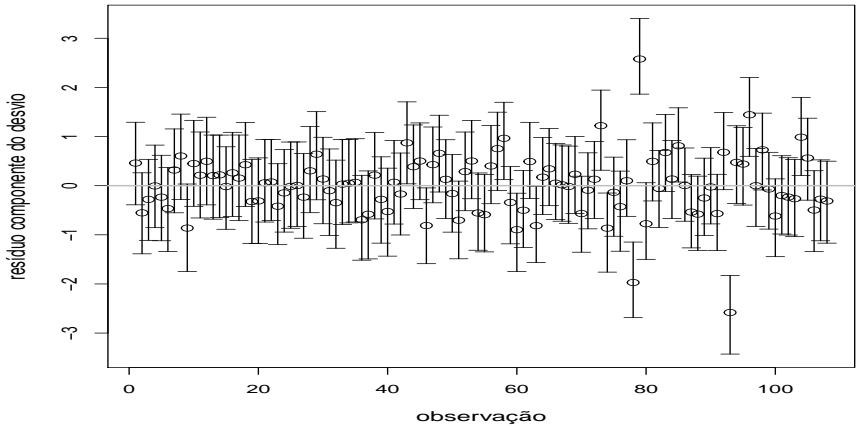
# Valores observados e preditos



# Valores observados e preditos



# Medianas a posteriori e IC(95%) do RCD



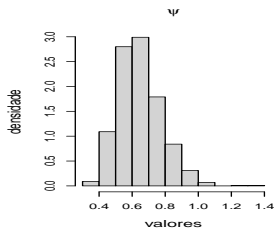
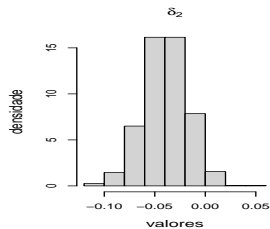
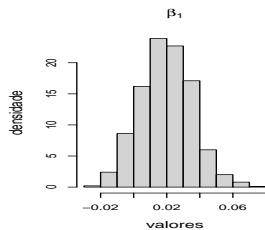
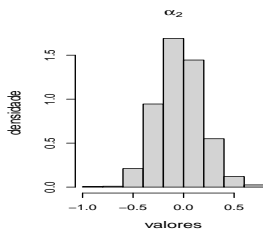
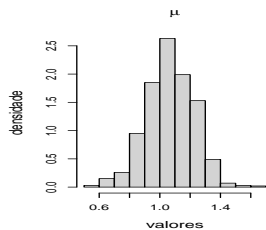
# Exemplo dos ataques epiléticos

- Utilizaremos o mesmo modelo considerado em [aqui](#).
- Prioris:  $\mu, \alpha, \beta_1, \delta_2 \stackrel{ind.}{\sim} N(0, 1000)$ ,  $\psi \sim \text{gama}(0, 05; 0, 05)$ ,  
 $\mathcal{E}(\psi) = 1$ ,  $\mathcal{V}(\psi) = 20$  (prioris vagas).

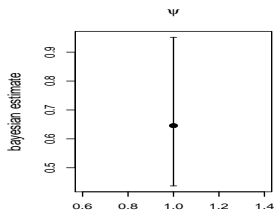
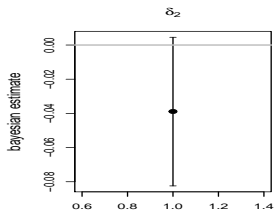
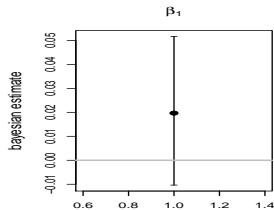
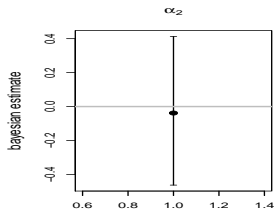
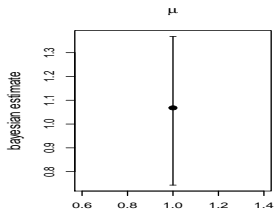
# Exemplo dos ataques epiléticos

- Parâmetros MCMC: burn-in = 50000, thin(lag) = 50 , número total de iterações = 100000, o que gera uma amostra válida das posteriores (após burn-in, com o que devido thin) de 1000 valores. Resultados [aqui](#).
- Questões de convergência. Provavelmente, devido à (elevada) correlação (à posteriori) entre os interceptos fixos e os aleatórios. Possíveis soluções: utilizar algoritmos auxiliares mais apropriados, reparametrizações entre outras.

# Distribuições a posteriori



# Estimativas pontuais e intervalos de credibilidade

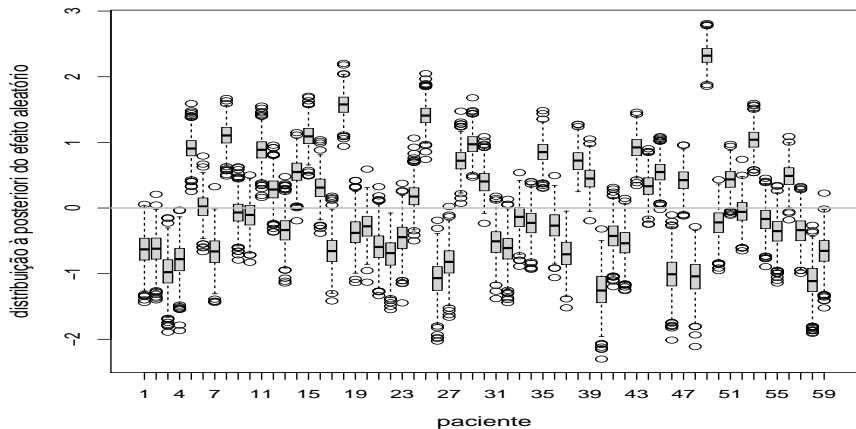




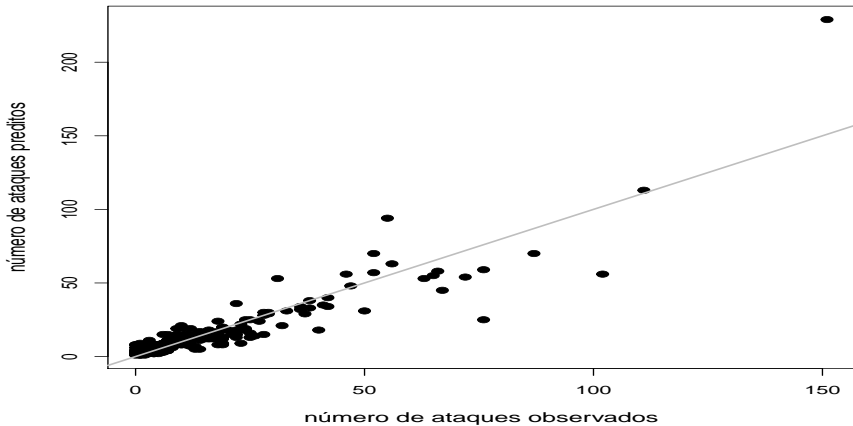
## Estimativas pontuais e intervalos de credibilidade

Parâmetro	EAP	MedAP	MoAP	DPAP	IC(%95)
$\mu$	1,06	1,06	1,01	0,16	[ 0,75 ; 1,35]
$\alpha_2$	-0,02	-0,01	-0,03	0,22	[-0,48 ; 0,40 ]
$\beta_1$	0,02	0,02	0,02	0,02	[ -0,01 ; 0,05]
$\delta_2$	-0,04	-0,04	-0,05	0,02	[-0,08 ; 0,00]
$\psi$	0,65	0,64	0,74	0,13	[0,44 ; 0,96 ]

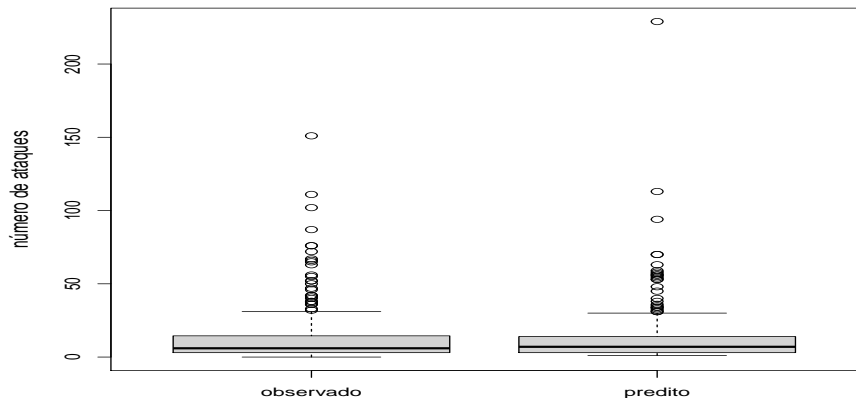
# Posteriori dos efeitos aleatórios



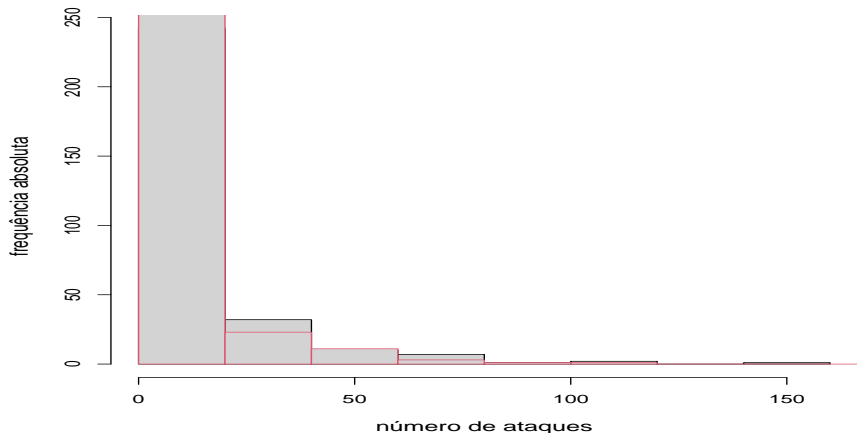
# Valores observados e preditos



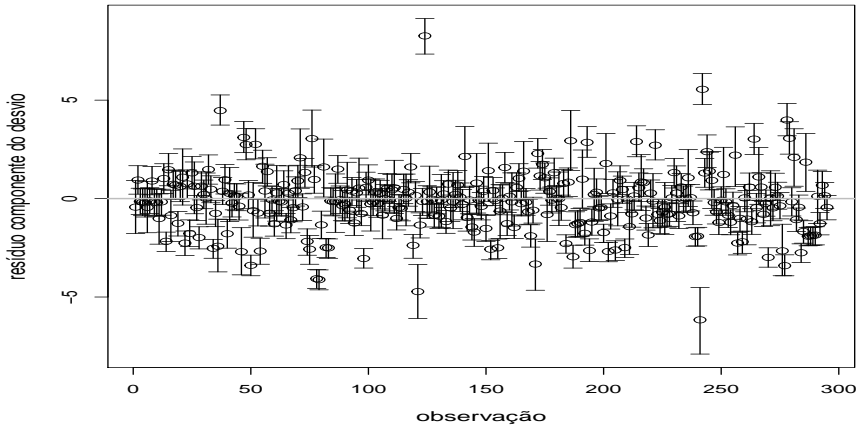
# Distribuições observadas e previstas



# Distribuições observadas (preto) e previstas (vermelho)



# Medianas a posteriori e IC(95%) do RCD



# M.R. para os dados do exemplo (regressão logística)

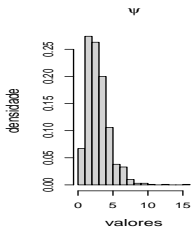
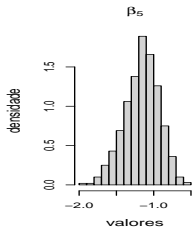
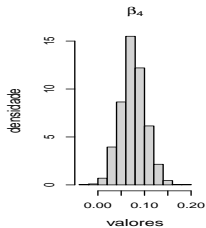
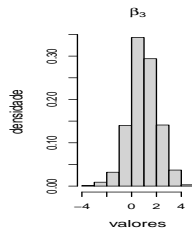
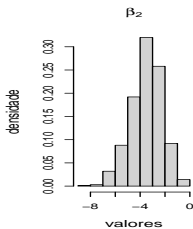
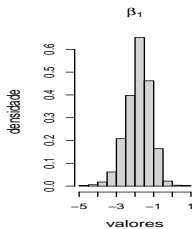
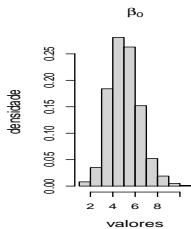
- Mesmo modelo usado [aqui](#).
- Prioris:  $\beta_i \stackrel{iid}{\sim} N(0, 1000)$ ,  $\psi \sim \text{gama}(0.05, 0.05)$ ,  $\mathcal{E}(\psi) = 1$ .

## Exemplo dos dados da condicação respiratória

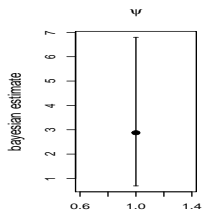
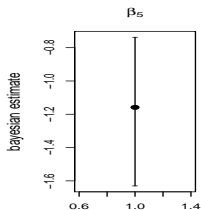
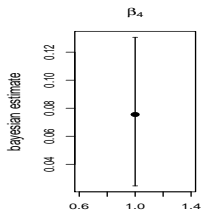
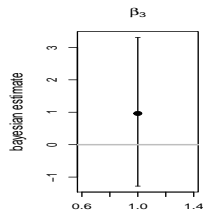
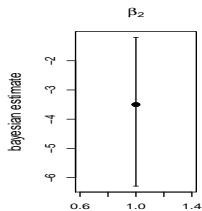
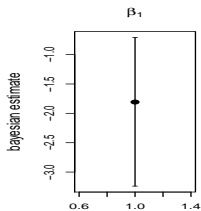
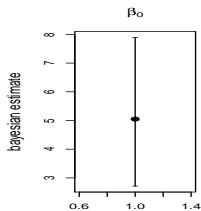
- Parâmetros MCMC: burn-in = 5000, thin(lag) = 50, número total de iterações = 55000, o que gera uma amostra válida das posteriores (após burn-in, com o que devido thin) de 1000 valores. Resultados [aqui](#).
- Questões de convergência. Provavelmente, devido à (elevada) correlação (à posteriori) entre os interceptos fixos e os aleatórios. Possíveis soluções: utilizar algoritmos auxiliares mais apropriados, reparametrizações entre outras.



# Distribuições a posteriori



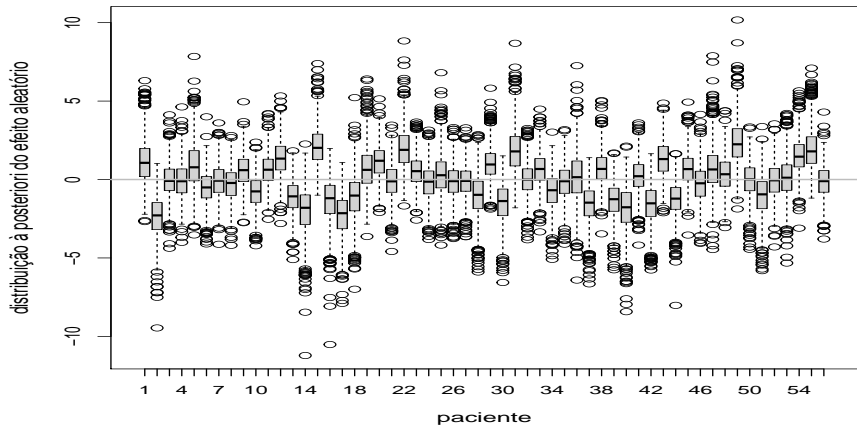
# Estimativas pontuais e intervalos de credibilidade



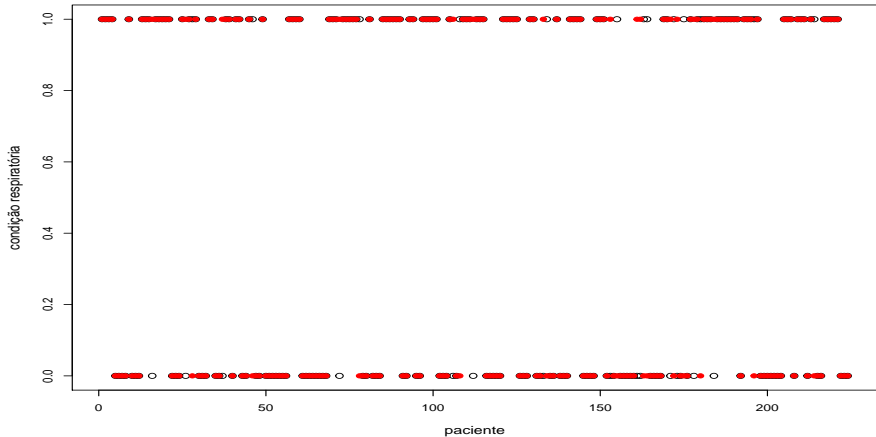
# Estimativas pontuais e intervalos de credibilidade

Parâmetro	EAP	MedAP	MoAP	DPAP	IC(%95)
$\beta_0$	5,16	5,06	4,54	1,43	[2,57 ; 8,09]
$\beta_1$	-1,87	-1,83	-1,91	0,67	[-3,24 ; -0,63]
$\beta_2$	-3,54	-3,49	-2,92	1,38	[-6,38 ; -1,08]
$\beta_3$	0,96	0,90	1,32	1,15	[-1,31 ; 3,23]
$\beta_4$	0,08	0,08	0,04	0,03	[0,02 ; 0,13]
$\beta_5$	-1,17	-1,15	-1,33	0,23	[-1,63 ; -0,76]
$\psi$	2,98	2,69	1,62	1,55	[0,77 ; 6,87]

# Posteriori dos efeitos aleatórios



# Valores observados (preto) e preditos (vermelho)



# Medianas a posteriori e IC(95%) do RCD

# Exercício

- Analise todos os conjuntos de dados vistos até agora, de forma apropriada, com algum modelo que você entender ser o mais apropriado, do ponto de vista bayesiano.
- Comparação de modelos também são oportunas.