

Comparação de Modelos Bayesianos na Análise de Dados de Contagem de CD4



Emílio A. C. Barros
Jorge Alberto Achcar

Faculdade de Medicina de Ribeirão Preto-USP

Introdução

☞ Dados longitudinais de Contagem na presença de covariáveis.



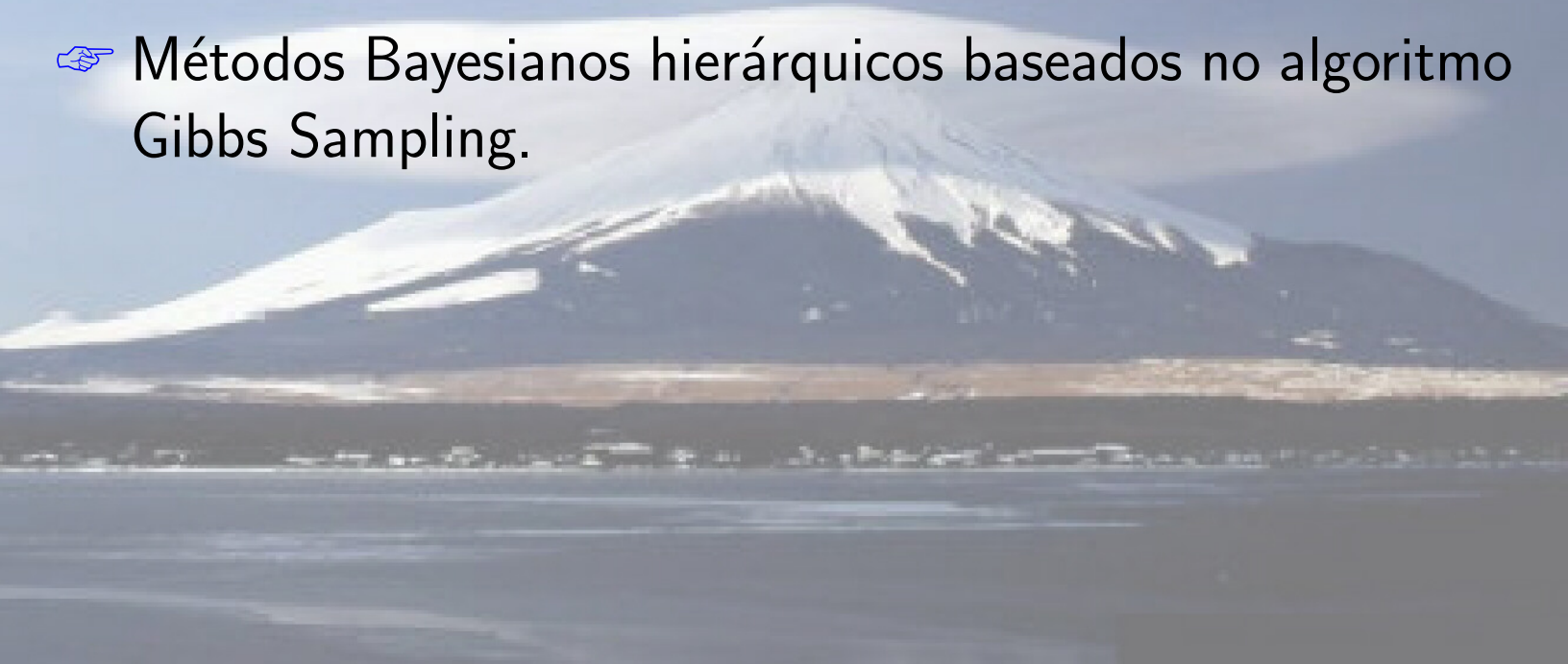
Introdução

- ➡ Dados longitudinais de Contagem na presença de covariáveis.
- ➡ Diferentes modelos, com diferentes estruturas de “fragilidade”.



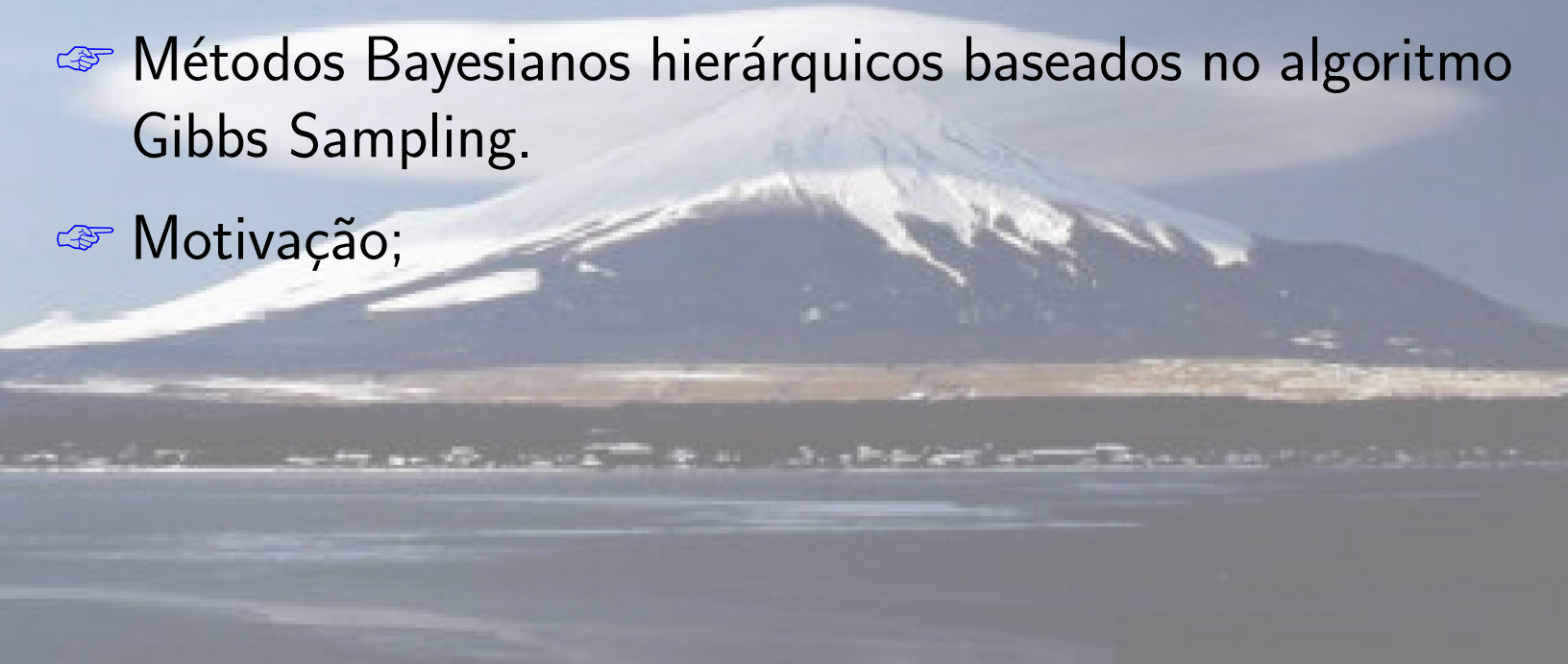
Introdução

- ➡ Dados longitudinais de Contagem na presença de covariáveis.
- ➡ Diferentes modelos, com diferentes estruturas de “fragilidade”.
- ➡ Métodos Bayesianos hierárquicos baseados no algoritmo Gibbs Sampling.



Introdução

- ➡ Dados longitudinais de Contagem na presença de covariáveis.
- ➡ Diferentes modelos, com diferentes estruturas de “fragilidade”.
- ➡ Métodos Bayesianos hierárquicos baseados no algoritmo Gibbs Sampling.
- ➡ Motivação;



Introdução

- ➡ Dados longitudinais de Contagem na presença de covariáveis.
- ➡ Diferentes modelos, com diferentes estruturas de “fragilidade”.
- ➡ Métodos Bayesianos hierárquicos baseados no algoritmo Gibbs Sampling.
- ➡ Motivação;
 - ➡ Estudo prospectivo, aberto e randomizado.

Introdução

- ➡ Dados longitudinais de Contagem na presença de covariáveis.
- ➡ Diferentes modelos, com diferentes estruturas de “fragilidade”.
- ➡ Métodos Bayesianos hierárquicos baseados no algoritmo Gibbs Sampling.
- ➡ Motivação;
 - ➡ Estudo prospectivo, aberto e randomizado.
 - ➡ Pacientes infectados pelo HIV (Unidade Especial de Terapia de Doenças Infeciosas - UETDI/HCFMRP-USP).

⇒ Virgens de tratamento.



- ⇒ Virgens de tratamento.
- ⇒ Idade superior a 18 anos.



- ⇒ Virgens de tratamento.
- ⇒ Idade superior a 18 anos.
- ⇒ Contagem de linfócitos T CD4+ inferior a 350 cél/mm^3 .

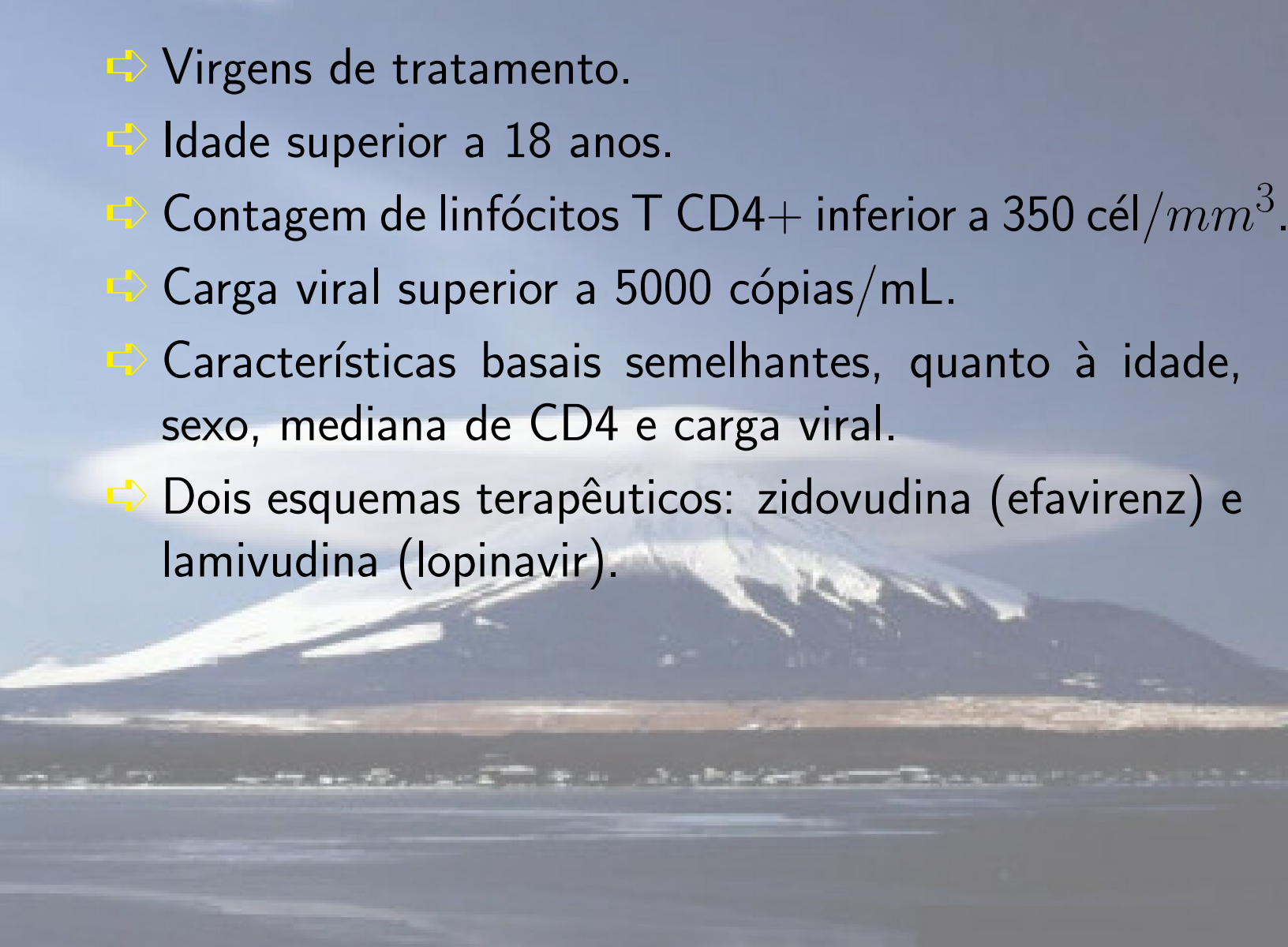


- ⇒ Virgens de tratamento.
- ⇒ Idade superior a 18 anos.
- ⇒ Contagem de linfócitos T CD4+ inferior a 350 cél/mm^3 .
- ⇒ Carga viral superior a 5000 cópias/mL.



- ⇒ Virgens de tratamento.
- ⇒ Idade superior a 18 anos.
- ⇒ Contagem de linfócitos T CD4+ inferior a 350 cél/mm^3 .
- ⇒ Carga viral superior a 5000 cópias/mL.
- ⇒ Características basais semelhantes, quanto à idade, sexo, mediana de CD4 e carga viral.



- 
- ⇒ Virgens de tratamento.
 - ⇒ Idade superior a 18 anos.
 - ⇒ Contagem de linfócitos T CD4+ inferior a 350 cél/mm^3 .
 - ⇒ Carga viral superior a 5000 cópias/mL.
 - ⇒ Características basais semelhantes, quanto à idade, sexo, mediana de CD4 e carga viral.
 - ⇒ Dois esquemas terapêuticos: zidovudina (efavirenz) e lamivudina (lopinavir).

- ⇒ Virgens de tratamento.
- ⇒ Idade superior a 18 anos.
- ⇒ Contagem de linfócitos T CD4+ inferior a 350 cél/mm^3 .
- ⇒ Carga viral superior a 5000 cópias/mL.
- ⇒ Características basais semelhantes, quanto à idade, sexo, mediana de CD4 e carga viral.
- ⇒ Dois esquemas terapêuticos: zidovudina (efavirenz) e lamivudina (lopinavir).

👉 Objetivo: Comparar a contagem de células CD4 considerando os dois esquemas terapêuticos (efavirenz e lopinavir).

Tabela 1: Contagem de CD4 em pacientes infectados pelo HIV (Colares, 2007).

Ind	Droga	Tempo	CD4	Ind	Droga	Tempo	CD4	Ind	Droga	Tempo	CD4
1	LPV	1	202	12	EFV	1	2	23	LPV	1	301
1	LPV	2	284	12	EFV	2	41	23	LPV	2	NA
1	LPV	3	300	12	EFV	3	57	23	LPV	3	399
1	LPV	4	443	12	EFV	4	64	23	LPV	4	390
1	LPV	5	557	12	EFV	5	107	23	LPV	5	380
1	LPV	6	532	12	EFV	6	87	23	LPV	6	523
2	LPV	1	296	13	EFV	1	159	24	LPV	1	26
2	LPV	2	550	13	EFV	2	195	24	LPV	2	194
2	LPV	3	499	13	EFV	3	195	24	LPV	3	117
2	LPV	4	584	13	EFV	4	212	24	LPV	4	149
2	LPV	5	723	13	EFV	5	NA	24	LPV	5	170
2	LPV	6	617	13	EFV	6	NA	24	LPV	6	222
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
11	EFV	1	44	22	EFV	1	42	33	EFV	1	75
11	EFV	2	NA	22	EFV	2	278	33	EFV	2	117
11	EFV	3	NA	22	EFV	3	232	33	EFV	3	285
11	EFV	4	NA	22	EFV	4	286	33	EFV	4	221
11	EFV	5	NA	22	EFV	5	353	33	EFV	5	341
11	EFV	6	NA	22	EFV	6	380	33	EFV	6	474

LPV: Lopinavir; EFV: Efavirenz; NA: Dados não observados

Formulação do Modelo

☞ Distribuição de Poisson,

$$P(Y_{ij} = y_{ij}) = \frac{e^{-\lambda_{ij}} \lambda_{ij}^{y_{ij}}}{y_{ij}!},$$



Formulação do Modelo

☞ Distribuição de Poisson,

$$P(Y_{ij} = y_{ij}) = \frac{e^{-\lambda_{ij}} \lambda_{ij}^{y_{ij}}}{y_{ij}!},$$

☞ $y_{ij} = 0, 1, 2, \dots$

☞ $i = 1, \dots, n$ (tamanho amostral).

☞ $j = 1, \dots, k$ (número de tempos).

👉 Modelo 1,

$$\lambda_{ij} = \alpha_j \exp(\beta_j x_{ij} + w_i)$$



👉 Modelo 1,

$$\lambda_{ij} = \alpha_j \exp(\beta_j x_{ij} + w_i)$$

- ⇒ $x_{ij} = 0$ indica que o indivíduo recebeu o tratamento a base de efavirenz
- ⇒ $x_{ij} = 1$ indica que o indivíduo recebeu o tratamento a base de lopinavir.
- ⇒ $w_i \stackrel{iid}{\sim} N(0, \tau^2)$

👉 Modelo 1,

$$\lambda_{ij} = \alpha_j \exp(\beta_j x_{ij} + w_i)$$

⇒ $x_{ij} = 0$ indica que o indivíduo recebeu o tratamento a base de efavirenz

⇒ $x_{ij} = 1$ indica que o indivíduo recebeu o tratamento a base de lopinavir.

⇒ $w_i \stackrel{iid}{\sim} N(0, \tau^2)$

👉 Modelo 2,

$$\lambda_{ij} = w_i \alpha_j \exp(\beta_j x_{ij}),$$

👉 Modelo 1,

$$\lambda_{ij} = \alpha_j \exp(\beta_j x_{ij} + w_i)$$

⇒ $x_{ij} = 0$ indica que o indivíduo recebeu o tratamento a base de efavirenz

⇒ $x_{ij} = 1$ indica que o indivíduo recebeu o tratamento a base de lopinavir.

⇒ $w_i \stackrel{iid}{\sim} N(0, \tau^2)$

👉 Modelo 2,

$$\lambda_{ij} = w_i \alpha_j \exp(\beta_j x_{ij}),$$

⇒ $w_i \stackrel{iid}{\sim} \text{Gamma}(\phi^{-1}; \phi^{-1})$

👉 Modelo 3,

$$\lambda_{ij} = \left(\sum_{l=1}^r \alpha_{lj} w_{li} \right) \exp(\beta_j x_{ij}),$$



👉 Modelo 3,

$$\lambda_{ij} = \left(\sum_{l=1}^r \alpha_{lj} w_{li} \right) \exp(\beta_j x_{ij}),$$

⇒ $w_{li} \stackrel{iid}{\sim} \text{Gamma}(\phi_l^{-1}; \phi_l^{-1})$



Análise Bayesiana Para o Modelo 1

☞ $\alpha_j \sim \text{Gamma}(a; b)$; a, b conhecidos;

☞ $\beta_j \sim N(c; d^2)$; c, d conhecidos;



Análise Bayesiana Para o Modelo 1

➡ $\alpha_j \sim \text{Gamma}(a; b)$; a, b conhecidos;

➡ $\beta_j \sim N(c; d^2)$; c, d conhecidos;

➡ $w_i \sim N(0, \tau^2)$.



Análise Bayesiana Para o Modelo 1

☞ $\alpha_j \sim \text{Gamma}(a; b)$; a, b conhecidos;

☞ $\beta_j \sim N(c; d^2)$; c, d conhecidos;

☞ $w_i \sim N(0, \tau^2)$.

⇒ Segundo estágio da análise Bayesiana hierárquica,

$\tau^2 \sim \text{IG}(f, g)$; f, g conhecidos.

Análise Bayesiana Para o Modelo 2

☞ Mesmas distribuições a priori para α_j e β_j .



Análise Bayesiana Para o Modelo 2

➡ Mesmas distribuições a priori para α_j e β_j .

➡ $w_i \sim \text{Gamma}(\phi^{-1}; \phi^{-1})$.



Análise Bayesiana Para o Modelo 2

➡ Mesmas distribuições a priori para α_j e β_j .

➡ $w_i \sim \text{Gamma}(\phi^{-1}; \phi^{-1})$.

➡ Segundo estágio da análise Bayesiana hierárquica,

$\phi \sim \text{Gamma}(f; g)$; f, g conhecidos.



Análise Bayesiana Para o Modelo 3

☞ $\beta_j \sim N\left(\beta_j^*, d^2\right)$; d^2 conhecido.



Análise Bayesiana Para o Modelo 3

☞ $\beta_j \sim N(\beta_j^*, d^2)$; d^2 conhecido.

☞ $\hat{\beta}_j$: médias a posteriori estimadas para β_j do modelo 2.

☞ $\beta_j^* = \hat{\beta}_j$.



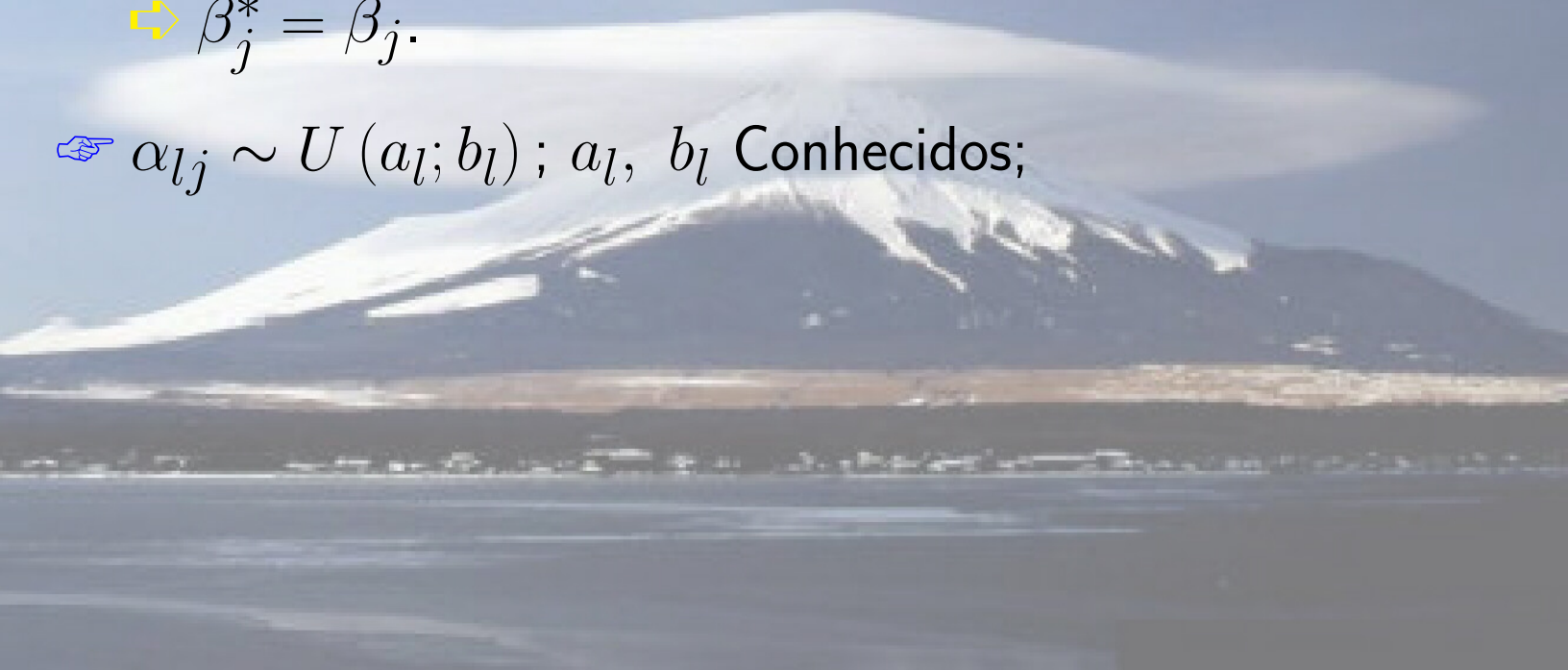
Análise Bayesiana Para o Modelo 3

☞ $\beta_j \sim N(\beta_j^*, d^2)$; d^2 conhecido.

☞ $\hat{\beta}_j$: médias a posteriori estimadas para β_j do modelo 2.

☞ $\beta_j^* = \hat{\beta}_j$.

☞ $\alpha_{lj} \sim U(a_l; b_l)$; a_l, b_l Conhecidos;



Análise Bayesiana Para o Modelo 3

➡ $\beta_j \sim N\left(\beta_j^*, d^2\right)$; d^2 conhecido.

➡ $\hat{\beta}_j$: médias a posteriori estimadas para β_j do modelo 2.

➡ $\beta_j^* = \hat{\beta}_j$.

➡ $\alpha_{lj} \sim U(a_l; b_l)$; a_l, b_l Conhecidos;

➡ $w_{li} \sim \text{Gamma}\left(\phi_l^{-1}; \phi_l^{-1}\right)$.

Análise dos Dados de Contagem de CD4

👉 Software *Winbugs*.



Análise dos Dados de Contagem de CD4

👉 Software *Winbugs*.

👉 Foram simuladas 1005000 amostras.



Análise dos Dados de Contagem de CD4

- ➡ Software *Winbugs*.
- ➡ Foram simuladas 1005000 amostras.
- ➡ 5000 amostras foram descartadas (“burn-in-samples”).



Análise dos Dados de Contagem de CD4

- ➡ Software *Winbugs*.
- ➡ Foram simuladas 1005000 amostras.
- ➡ 5000 amostras foram descartadas (“burn-in-samples”).
- ➡ Foram consideradas as amostras $100^a, 200^a, 300^a, \dots$



Análise dos Dados de Contagem de CD4

- ➡ Software *Winbugs*.
- ➡ Foram simuladas 1005000 amostras.
- ➡ 5000 amostras foram descartadas (“burn-in-samples”).
- ➡ Foram consideradas as amostras $100^a, 200^a, 300^a, \dots$
- ➡ Amostra final de tamanho 10000 para cada parâmetro.

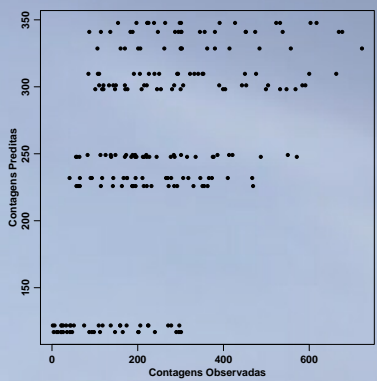


Análise dos Dados de Contagem de CD4

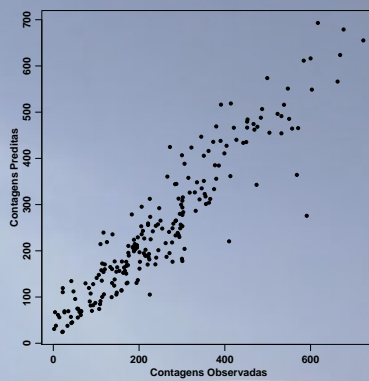
- ➡ Software *Winbugs*.
- ➡ Foram simuladas 1005000 amostras.
- ➡ 5000 amostras foram descartadas (“burn-in-samples”).
- ➡ Foram consideradas as amostras 100^a , 200^a , 300^a , ...
- ➡ Amostra final de tamanho 10000 para cada parâmetro.

Tabela 2: Critério DIC.

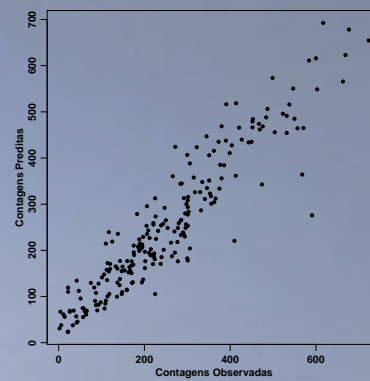
Modelo	<i>Deviance Information Criterion (DIC)</i>
Modelo 1 (ausência de w_i)	18010.800
Modelo 1 (presença w_i)	4914.000
Modelo 2	4861.450
Modelo 3 ($r = 2$)	3659.560
Modelo 3 ($r = 3$)	2783.210
Modelo 3 ($r = 4$)	2430.550



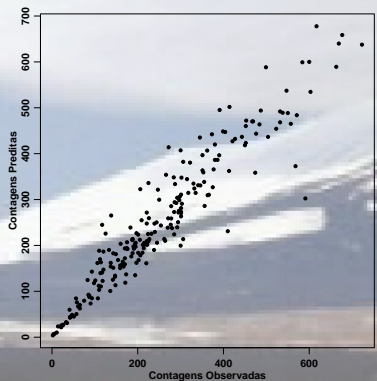
(a) Modelo 1 (ausência de w_i)



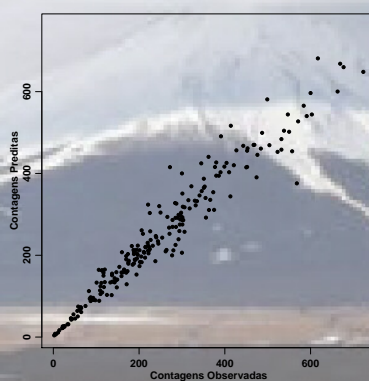
(b) Modelo 1 (presença de w_i)



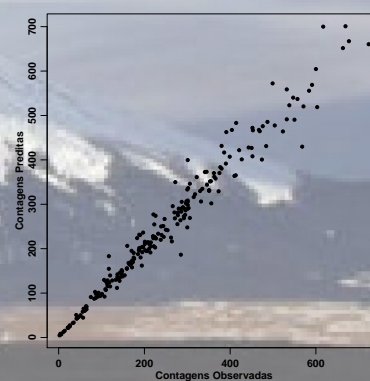
(c) Modelo 2



(d) Modelo 3 ($r=2$)



(e) Modelo 3 ($r=3$)



(f) Modelo 3 ($r=4$)

Figura 1: Gráficos dos valores observados *versus* os valores preditos.

Tabela 3: Soma do quadrado das diferenças entre os valores observados e os valores preditos.

Modelo	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$
Modelo 1 (ausência w_i)	4093784.0
Modelo 1 (presença w_i)	722850.9
Modelo 2	722642.4
Modelo 3 ($r = 2$)	552654.9
Modelo 3 ($r = 3$)	312354.2
Modelo 3 ($r = 4$)	176012.7

Tabela 4: Média a posteriori e intervalos de credibilidade para os parâmetros do modelo 3, $r = 4$.

Parâmetro	Média a Posteriori	Desvio Padrão	Intervalo de Credibilidade
β_1	0.394	0.233	(-0.040; 0.877)
β_2	0.157	0.140	(-0.121; 0.433)
β_3	0.446	0.145	(0.164; 0.738)
β_4	0.489	0.156	(0.193; 0.810)
β_5	0.338	0.140	(0.071; 0.624)
β_6	0.299	0.135	(0.038; 0.572)

OBRIGADO
PELA
ATENÇÃO

The background of the image is a sunset over a desert landscape. The sky transitions from a deep purple at the top to a bright orange near the horizon where the sun is setting. In the foreground, the silhouettes of three camels with riders are visible, moving from left to right. In the background, the silhouettes of several pyramids are visible against the horizon.

eacbarros@gmail.com