



Universidade Federal de Santa Catarina - UFSC
Departamento de Informática e Estatística - INE

ESTIMAÇÃO BAYESIANA PARA MODELOS DE PROPORÇÕES E CONTAGENS COM EXCESSO DE ZEROS

ADRIANO FERRETI BORGATTO

EXPERIMENTO 1

- A broca-da-cana é considerada letal aos canaviais;
- A espécie de parasita *Trichogramma galloi* foi usada para parasitar 128 ovos de *Anagasta kuehniella* (hospedeiro alternativo);
- O número de fêmeas do *T. galloi* usado para parasitar os ovos variou (2, 4, 8, 16, 32, 64, 128);
- O conjunto de dados é proveniente de um experimento inteiramente casualizado com 10 repetições;
- A variável de interesse está no número de ovos parasitados de *A. kuehniella*, dentre os 128 ovos.

BROCA-D A - C A N A



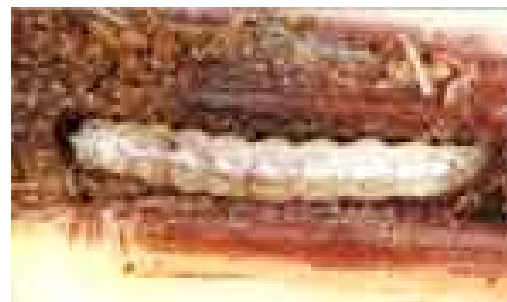
Adulto da broca da cana



Ovo da broca da cana

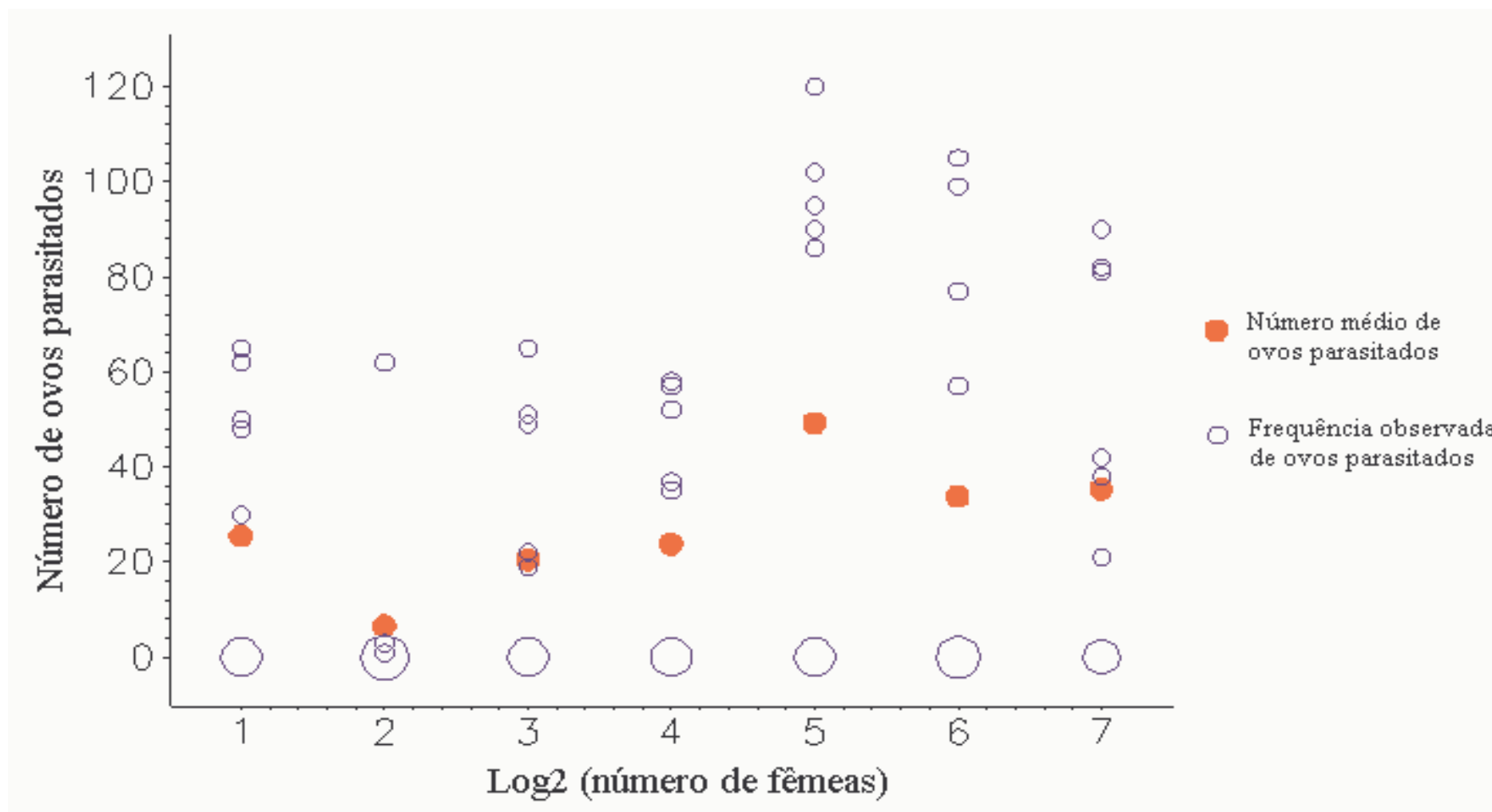


Espécie *Trichogramma*



Danos causados pela broca

Em 52,8% das observações, não foi parasitado nenhum ovo e a maior quantidade média de ovos parasitados ocorreu com 32 fêmeas.



CONSIDERAÇÕES

- i) a natureza discreta da variável resposta (proporções de ovos parasitados);
- ii) a variabilidade entre repetições para cada número de fêmeas do *T. galloi* (2, 4, 8, 16, 32, 64, 128);
- iii) o número de repetições em que nenhum ovo foi parasitado (excesso de zeros).

OBJETIVO

- Sugerir o número ideal de fêmeas de *T. galloi* que maximize o número de ovos parasitados de *A. kuehniella*;
- Usar a estimação Bayesiana na modelagem dos dados;
- Comparar o ajuste dos modelos usados na análise dos dados.

Modelos Estatísticos

Modelo Binomial

$$P(Y_i = y_i) = \binom{m_i}{y_i} p_i^{y_i} (1 - p_i)^{m_i - y_i}, \quad 0 \leq p_i \leq 1$$

$$E(Y_i) = m_i p_i \quad \text{e} \quad \text{Var}(Y_i) = m_i p_i (1 - p_i)$$

- Função de ligação: Logística.

Preditor Linear:

$$g(p_i) = \text{logit}(p_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_q x_i^q$$

em que $x_i = \frac{(z_i - \bar{z})}{\sigma_z}$ com $z_i = \log_2(\text{número de fêmeas})$

Modelo Beta-Binomial

$$Y_i | P \sim \text{Bin}(m_i, P) \quad \text{e} \quad P \sim \text{Beta}(a, b)$$

$$P(Y_i = y_i) = \binom{m_i}{y_i} \frac{\Gamma(a + y_i)\Gamma(m_i + b - y_i)}{\Gamma(m_i + a + b)} \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)}$$

$$E(Y_i) = m_i \left(\frac{a}{a + b} \right)$$

$$\text{Var}(Y_i) = m_i \left(\frac{a}{a + b} \right) \left(\frac{b}{a + b} \right) \left(\frac{m_i + a + b}{1 + a + b} \right)$$

Reparametrização:

$$p = \frac{a}{a+b} \quad \text{e} \quad \delta = \frac{1}{a+b}$$

$$P(Y_i = y_i) = \binom{m_i}{y_i} \frac{\Gamma\left(\frac{p_i}{\delta} + y_i\right) \Gamma\left(\frac{1-p_i}{\delta} + m_i - y_i\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right)} \frac{\Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{p_i}{\delta}\right) \Gamma\left(\frac{1-p_i}{\delta}\right)}$$

$$E(Y_i) = m_i p_i \quad \text{e} \quad \text{Var}(Y_i) = m_i p_i (1 - p_i) \left(\frac{m_i \delta + 1}{\delta + 1} \right)$$

- Função de ligação: Logística.

Preditor Linear:

$$g(p_i) = \text{logit}(p_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_q x_i^q$$

Modelo Beta-Binomial Inflacionado de Zeros (ZIBB)

$$P(Y_i = y_i) = \begin{cases} w_i + (1 - w_i) \frac{\Gamma(\frac{1}{\delta})\Gamma(\frac{1-p_i}{\delta} + m_i)}{\Gamma(\frac{1}{\delta} + m_i)\Gamma(\frac{1-p_i}{\delta})} & y_i = 0 \\ (1 - w_i) \binom{m_i}{y_i} \frac{\Gamma(\frac{1}{\delta})\Gamma(\frac{p_i}{\delta} + y_i)\Gamma(\frac{1-p_i}{\delta} + m_i - y_i)}{\Gamma(\frac{1}{\delta} + m_i)\Gamma(\frac{p_i}{\delta})\Gamma(\frac{1-p_i}{\delta})} & y_i > 0 \end{cases}$$

$$E(Y_i) = (1 - w_i)m_i p_i$$

$$Var(Y_i) = \mu_i \left[(1 - p_i) \left(\frac{m_i \delta + 1}{\delta + 1} \right) + \left(\frac{w_i}{1 - w_i} \right) \mu_i \right]$$

- Função de ligação: Logística para p_i e w_i .

Preditores Lineares:

$$g(p_i) = \text{logit}(p_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_{q_1} x_i^{q_1}$$

$$g(w_i) = \text{logit}(w_i) = \gamma_0 + \gamma_1 g_i + \gamma_2 g_i^2 + \dots + \gamma_{q_2} g_i^{q_2}$$

ANÁLISE BAYESIANA

MODELO BINOMIAL

- Estimação de β_j

Distribuição a posteriori conjunta

$$\pi(\boldsymbol{\beta} | \mathbf{Y}) \propto L(\boldsymbol{\beta} | \mathbf{Y})\pi(\boldsymbol{\beta})$$

Distribuição a posteriori condicional

$$\pi(\beta_j | \boldsymbol{\beta}_{-j}, \mathbf{Y}) \propto L(\boldsymbol{\beta} | \mathbf{Y})\pi(\beta_j)$$

MODELO BETA-BINOMIAL

- Estimação de β_j e δ

Distribuição a posteriori conjunta

$$\pi(\boldsymbol{\beta}, \delta | \mathbf{Y}) \propto L(\boldsymbol{\beta}, \delta | \mathbf{Y})\pi(\boldsymbol{\beta})\pi(\delta)$$

Distribuições a posteriori condicionais

$$\pi(\beta_j | \boldsymbol{\beta}_{-j}, \delta, \mathbf{Y}) \propto L(\boldsymbol{\beta}, \delta | \mathbf{Y})\pi(\beta_j)$$

$$\pi(\delta | \boldsymbol{\beta}, \mathbf{Y}) \propto L(\boldsymbol{\beta}, \delta | \mathbf{Y})\pi(\delta)$$

MODELO ZIBB

- Estimaco de β_j , γ_k e δ

Distribuio a posteriori conjunta

$$\pi(\boldsymbol{\beta}, \boldsymbol{\gamma}, \delta \mid \mathbf{Y}) \propto L(\boldsymbol{\beta}, \boldsymbol{\gamma}, \delta \mid \mathbf{Y})\pi(\boldsymbol{\beta})\pi(\boldsymbol{\gamma})\pi(\delta)$$

Distribuies a posteriori condicionais

$$\pi(\beta_j \mid \boldsymbol{\beta}_{-j}, \boldsymbol{\gamma}, \delta, \mathbf{Y}) \propto L(\boldsymbol{\beta}, \boldsymbol{\gamma}, \delta \mid \mathbf{Y})\pi(\beta_j)$$

$$\pi(\gamma_k \mid \boldsymbol{\gamma}_{-k}, \boldsymbol{\beta}, \delta, \mathbf{Y}) \propto L(\boldsymbol{\beta}, \boldsymbol{\gamma}, \delta \mid \mathbf{Y})\pi(\gamma_k)$$

$$\pi(\delta \mid \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{Y}) \propto L(\boldsymbol{\beta}, \boldsymbol{\gamma}, \delta \mid \mathbf{Y})\pi(\delta)$$

P R I O R I : δ, β_j e γ_k

$\delta \sim \text{Gama}(c, d); c, d$ conhecidos

$\beta_j \sim N(\mu_j, \sigma_j^2); \mu_j, \sigma_j^2$ conhecidos , $j = 1, \dots, q_1$

$\gamma_k \sim N(\eta_k, \nu_k^2); \eta_k, \nu_k^2$ conhecidos , $k = 1, \dots, q_2$

Modelo Binomial

- Estimaco de β_j

Distribuio a posteriori conjunta

$$\pi(\beta_0, \beta_1, \dots, \beta_q \mid \mathbf{x}_i, \mathbf{y}, \mathbf{m}) \propto \exp \left\{ \frac{-1}{2\sigma_0^2} (\beta_0 - \mu_0)^2 \right\} \exp \left\{ \frac{-1}{2\sigma_1^2} (\beta_1 - \mu_1)^2 \right\} \\ \dots \exp \left\{ \frac{-1}{2\sigma_q^2} (\beta_q - \mu_q)^2 \right\} \prod_{i=1}^n \left[\frac{e^{y_i x_i \beta}}{(1 + e^{x_i \beta})^{m_i}} \right]$$

Distribuio a posteriori condicional

$$\pi(\beta_j \mid \boldsymbol{\beta}_{-j}, \mathbf{x}_i, \mathbf{y}, \mathbf{m}) \propto \exp \left\{ \frac{-1}{2\sigma_j^2} (\beta_j - \mu_j)^2 \right\} \prod_{i=1}^n \left[\frac{e^{y_i x_i \beta}}{(1 + e^{x_i \beta})^{m_i}} \right]$$

Modelo Beta-Binomial

- Estimaco de β_j e δ

Distribuio a posteriori conjunta

$$\pi(\boldsymbol{\beta}, \delta \mid \mathbf{x}_i, \mathbf{y}, \mathbf{m}) \propto \delta^{c-1} \exp\{-d\delta\} \exp\left\{\frac{-1}{2\sigma_0^2}(\beta_0 - \mu_0)^2\right\} \dots \exp\left\{\frac{-1}{2\sigma_q^2}(\beta_q - \mu_q)^2\right\}$$

$$\prod_{i=1}^n \left[\frac{\Gamma\left(\frac{e^{x_i\beta}}{(1+e^{x_i\beta})\delta} + y_i\right) \Gamma\left(\frac{1}{(1+e^{x_i\beta})\delta} + m_i - y_i\right) \Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right) \Gamma\left(\frac{e^{x_i\beta}}{(1+e^{x_i\beta})\delta}\right) \Gamma\left(\frac{1}{(1+e^{x_i\beta})\delta}\right)} \right]$$

Distribuio a posteriori condicional

$$\pi(\beta_j \mid \delta, \boldsymbol{\beta}_{-j}, \mathbf{x}_i, \mathbf{y}, \mathbf{m}) \propto \exp\left\{\frac{-1}{2\sigma_j^2}(\beta_j - \mu_j)^2\right\}$$

$$\prod_{i=1}^n \left[\frac{\Gamma\left(\frac{e^{x_i\beta}}{(1+e^{x_i\beta})\delta} + y_i\right) \Gamma\left(\frac{1}{(1+e^{x_i\beta})\delta} + m_i - y_i\right) \Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right) \Gamma\left(\frac{e^{x_i\beta}}{(1+e^{x_i\beta})\delta}\right) \Gamma\left(\frac{1}{(1+e^{x_i\beta})\delta}\right)} \right]$$

$$\pi(\delta \mid \boldsymbol{\beta}, \mathbf{x}_i, \mathbf{y}, \mathbf{m}) \propto \delta^{c-1} \exp\{-d\delta\} \prod_{i=1}^n \left[\frac{\Gamma\left(\frac{e^{x_i\beta}}{(1+e^{x_i\beta})\delta} + y_i\right) \Gamma\left(\frac{1}{(1+e^{x_i\beta})\delta} + m_i - y_i\right) \Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right) \Gamma\left(\frac{e^{x_i\beta}}{(1+e^{x_i\beta})\delta}\right) \Gamma\left(\frac{1}{(1+e^{x_i\beta})\delta}\right)} \right]$$

Modelo ZIBB

- Estimaco de β_j , γ_k e δ

Distribuio a posteriori conjunta

$$\begin{aligned}
 \pi(\boldsymbol{\beta}, \boldsymbol{\gamma}, \delta \mid \mathbf{x}_i, \mathbf{g}_i, \mathbf{y}, \mathbf{m}, \mathbf{z}) &\propto \delta^{c-1} \exp\{-d\delta\} \exp\left\{\frac{-1}{2\sigma_0^2}(\beta_0 - \mu_0)^2\right\} \dots \\
 &\exp\left\{\frac{-1}{2\sigma_{q_1}^2}(\beta_{q_1} - \mu_{q_1})^2\right\} \exp\left\{\frac{-1}{2\nu_0^2}(\gamma_0 - \eta_0)^2\right\} \dots \\
 &\exp\left\{\frac{-1}{2\nu_{q_2}^2}(\gamma_{q_2} - \eta_{q_2})^2\right\} \left\{ \prod_{i=1}^n z_i \left[\frac{1}{(1 + e^{g_i \gamma})} \frac{\Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta} + m_i\right) \Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta}\right)} \right] + \right. \\
 &\left. + \prod_{i=1}^n (1 - z_i) \left[\frac{e^{g_i \gamma}}{(1 + e^{g_i \gamma})} \frac{\Gamma\left(\frac{e^{x_i \beta}}{(1+e^{x_i \beta})\delta} + y_i\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta} + m_i - y_i\right) \Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right) \Gamma\left(\frac{e^{x_i \beta}}{(1+e^{x_i \beta})\delta}\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta}\right)} \right] \right\}
 \end{aligned}$$

Distribuição a posteriori condicional

$$\pi(\beta_j \mid \beta_{-j}, \gamma, \delta, \mathbf{x}_i, \mathbf{g}_i, \mathbf{y}, \mathbf{m}, \mathbf{z}) \propto \exp \left\{ \frac{-1}{2\sigma_j^2} (\beta_j - \mu_j)^2 \right\}$$

$$\left\{ \prod_{i=1}^n z_i \left[\frac{1}{(1 + e^{g_i \gamma})} \frac{\Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta} + m_i\right) \Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta}\right)} \right] + \right.$$

$$\left. + \prod_{i=1}^n (1 - z_i) \left[\frac{e^{g_i \gamma}}{(1 + e^{g_i \gamma})} \frac{\Gamma\left(\frac{e^{x_i \beta}}{(1+e^{x_i \beta})\delta} + y_i\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta} + m_i - y_i\right) \Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right) \Gamma\left(\frac{e^{x_i \beta}}{(1+e^{x_i \beta})\delta}\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta}\right)} \right] \right\}$$

$$\pi(\gamma_k | \gamma_{-k}, \beta, \delta, \mathbf{x}_i, \mathbf{g}_i, \mathbf{y}, \mathbf{m}, \mathbf{z}) \propto \exp \left\{ \frac{-1}{2\nu_k^2} (\gamma_k - \eta_k)^2 \right\}$$

$$\left\{ \prod_{i=1}^n z_i \left[\frac{1}{(1 + e^{g_i \gamma})} \frac{\Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta} + m_i\right) \Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta}\right)} \right] + \right.$$

$$\left. + \prod_{i=1}^n (1 - z_i) \left[\frac{e^{g_i \gamma}}{(1 + e^{g_i \gamma})} \frac{\Gamma\left(\frac{e^{x_i \beta}}{(1+e^{x_i \beta})\delta} + y_i\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta} + m_i - y_i\right) \Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right) \Gamma\left(\frac{e^{x_i \beta}}{(1+e^{x_i \beta})\delta}\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta}\right)} \right] \right\}$$

$$\pi(\delta | \gamma, \beta, \mathbf{x}_i, \mathbf{g}_i, \mathbf{y}, \mathbf{m}, \mathbf{z}) \propto \delta^{c-1} \exp\{-d\delta\}$$

$$\left\{ \prod_{i=1}^n z_i \left[\frac{1}{(1 + e^{g_i \gamma})} \frac{\Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta} + m_i\right) \Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta}\right)} \right] + \right.$$

$$\left. + \prod_{i=1}^n (1 - z_i) \left[\frac{e^{g_i \gamma}}{(1 + e^{g_i \gamma})} \frac{\Gamma\left(\frac{e^{x_i \beta}}{(1+e^{x_i \beta})\delta} + y_i\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta} + m_i - y_i\right) \Gamma\left(\frac{1}{\delta}\right)}{\Gamma\left(\frac{1}{\delta} + m_i\right) \Gamma\left(\frac{e^{x_i \beta}}{(1+e^{x_i \beta})\delta}\right) \Gamma\left(\frac{1}{(1+e^{x_i \beta})\delta}\right)} \right] \right\}$$

ALGORITMO

METROPOLIS - HASTINGS

A distribuição a posteriori condicional pode ser escrita através do produto de duas funções

$$\pi(\theta_j \mid \boldsymbol{\theta}_{-j}) \propto \Psi(\boldsymbol{\theta})q(\theta_j)$$

sendo que $q(\theta_j)$ tem a forma de uma distribuição conhecida para que através dela sejam gerados valores de θ_j .

Etapas do algoritmo:

- (i) simulação de amostras θ_j através de $q(\theta_j)$;
- (ii) cálculo da probabilidade de aceitação do novo valor

$$\alpha(\theta_j^t) = \min \left(1, \frac{\Psi(\theta^t)}{\Psi(\theta^{t-1})} \right);$$

- (iii) geração de um valor U a partir de uma distribuição Uniforme(0, 1);
- (iv) aceitação do ponto candidato θ_j^t se $U \leq \alpha(\theta_j^t)$ ou rejeição do mesmo em caso contrário;
- (v) retorno ao passo (i) até que a convergência seja alcançada, de acordo com algum critério estabelecido.

Valores Gerados de δ , β_j e γ_k

- Gerou-se uma cadeia com 100.000 valores para cada parâmetro, desprezando os primeiros 2.000 valores e selecionando um a cada 21 dos 98.000 valores restantes, formando uma amostra de 4.500 valores.
- Convergência:
 - análise gráfica dos valores gerados;
 - diagnóstico de Raftery & Lewis;
 - diagnóstico de Heidelberger & Welch.
- O fator de Bayes e o DIC foram usados para a escolha do modelo mais adequado.

RESULTADOS

Tabela 1: Valores do DIC para os modelos: binomial, beta-binomial e ZIBB

Modelos	Preditores lineares	DIC
binomial	$\text{logit}(p) = \beta_0$	4.747,50
	$\text{logit}(p) = \beta_0 + \beta_1 x_i$	4.538,60
	$\text{logit}(p) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$	4.538,60
	$\text{logit}(p) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3$	4.375,82
	$\text{logit}(p) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \beta_4 x_i^4$	4.295,41
beta binomial	$\text{logit}(p) = \beta_0$	446,42
	$\text{logit}(p) = \beta_0 + \beta_1 x_i$	447,24
	$\text{logit}(p) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$	449,15
	$\text{logit}(p) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3$	440,55
	$\text{logit}(p) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \beta_4 x_i^4$	442,53
ZIBB	Fixando-se: $\text{logit}(w) = \gamma_0$	
	$\text{logit}(p) = \beta_0$	417,95
	$\text{logit}(p) = \beta_0 + \beta_1 x_i$	413,47
	$\text{logit}(p) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$	414,54
	$\text{logit}(p) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3$	398,00
	$\text{logit}(p) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \beta_4 x_i^4$	398,03

Tabela 3: Estimativas dos parâmetros para o modelo ZIBB com preditor linear para as proporções o polinômio de terceiro grau e para os zeros uma constante

Parâmetros	Médias	Desvio Padrão	Intervalos de credibilidade	
			2,5%	97,5%
β_0	-0,1844	0,2149	-0,6185	0,2290
β_1	2,2370	0,4236	1,4264	3,1029
β_2	-0,0885	0,1545	-0,4034	0,2215
β_3	-0,9714	0,2154	-1,4091	-0,5436
γ_0	0,0972	0,2376	-0,3699	0,5651
δ	0,1698	0,0491	0,0967	0,2870

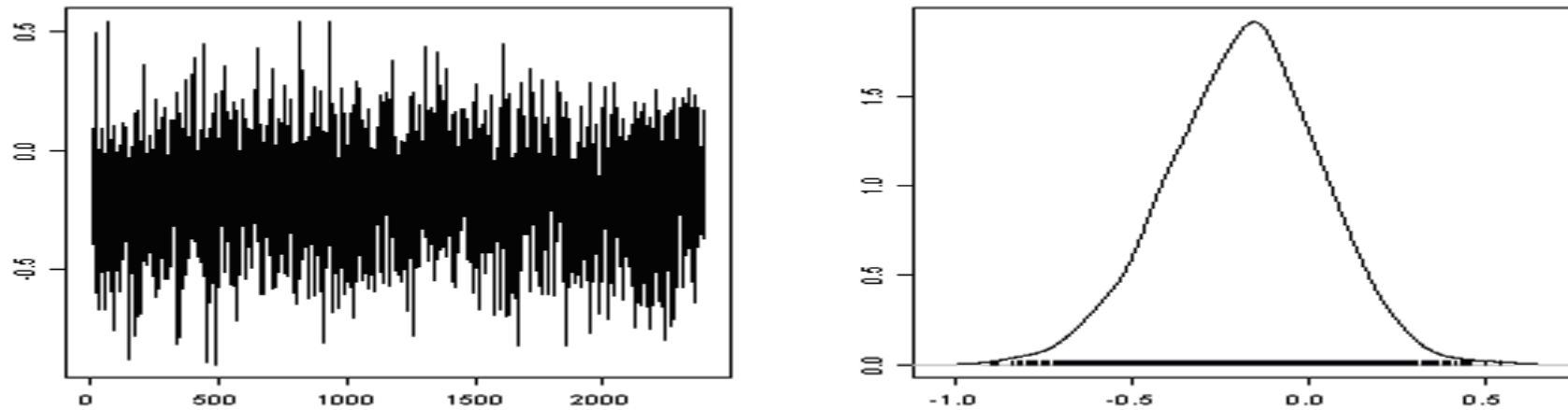


Figura 1: Trajetória da cadeia ao longo das iterações e a aproximação da densidade a posteriori de β_0 para o modelo ZIBB.

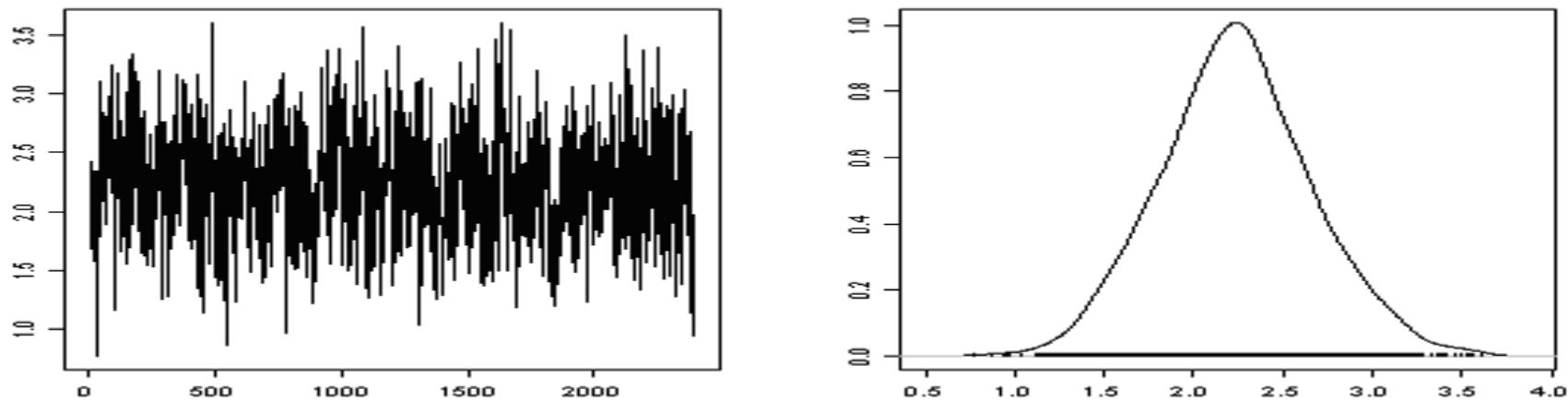


Figura 2: Trajetória da cadeia ao longo das iterações e a aproximação da densidade a posteriori de β_1 para o modelo ZIBB.

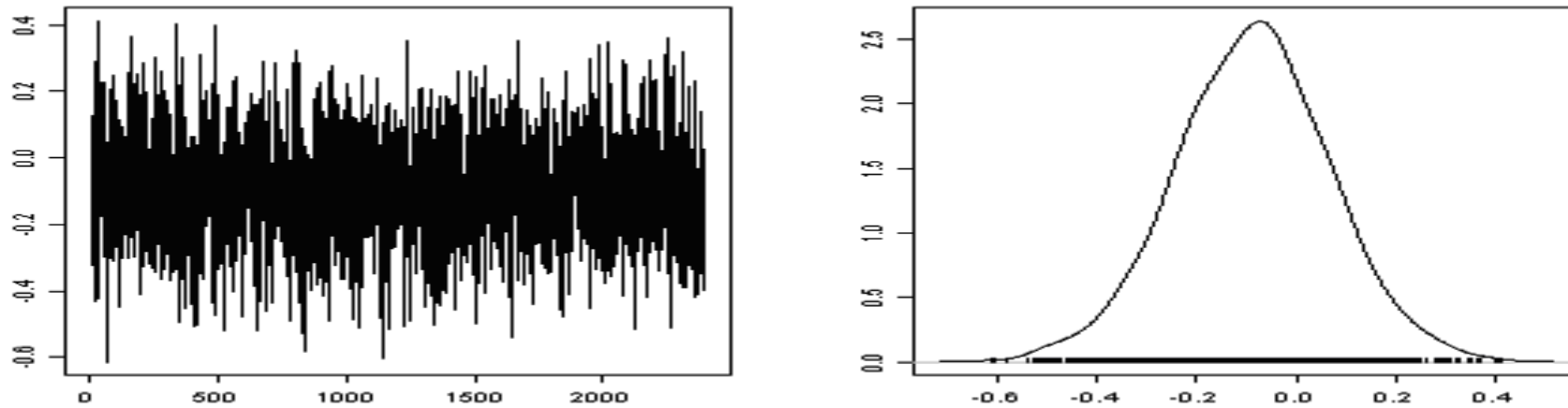


Figura 3: Trajetória da cadeia ao longo das iterações e a aproximação da densidade a posteriori de β_2 para o modelo ZIBB.

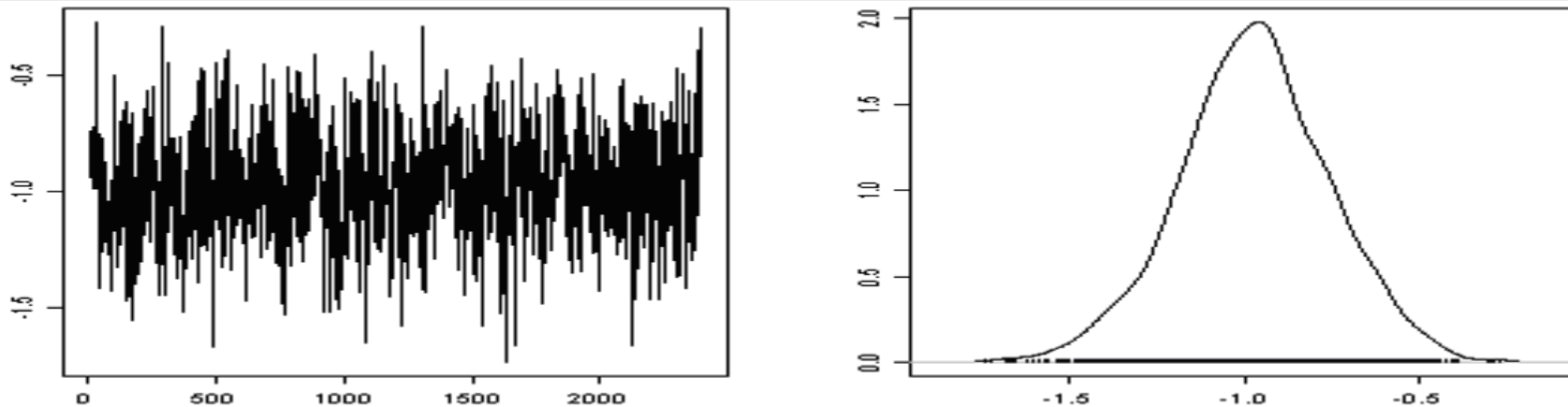


Figura 4: Trajetória da cadeia ao longo das iterações e a aproximação da densidade a posteriori de β_3 para o modelo ZIBB.

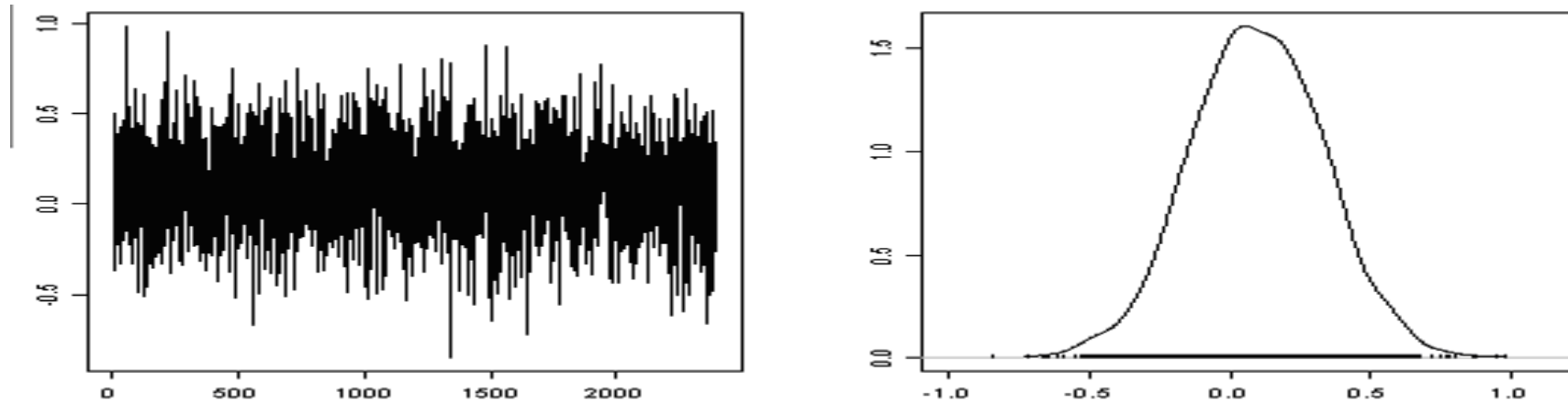


Figura 5: Trajetória da cadeia ao longo das iterações e a aproximação da densidade a posteriori de γ_0 para o modelo ZIBB.

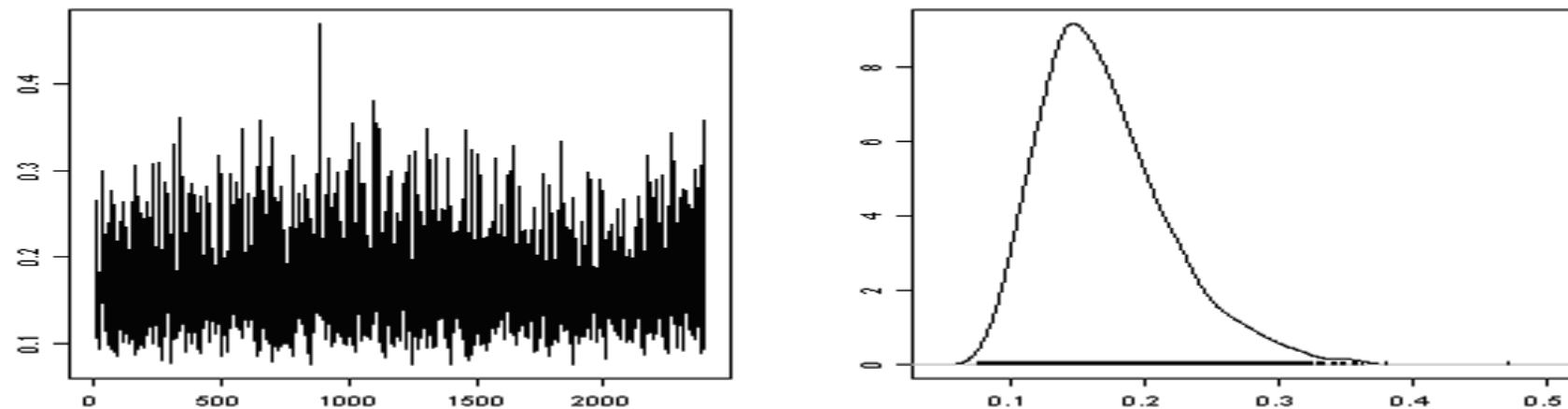


Figura 6: Trajetória da cadeia ao longo das iterações e a aproximação da densidade a posteriori de δ para o modelo ZIBB.

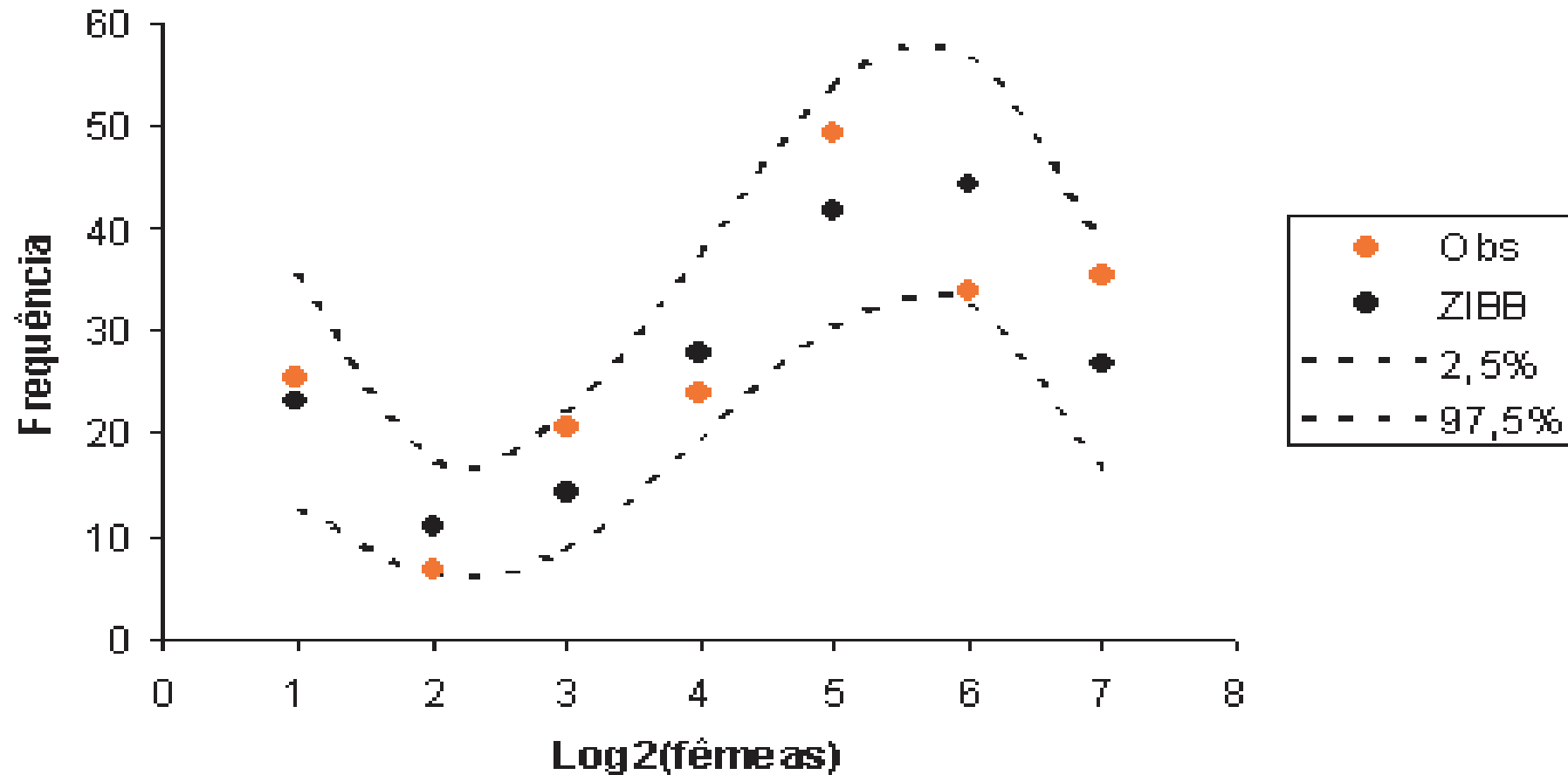


Figura 7: Valores observados, preditos e intervalos de credibilidade a 95% de probabilidade para o modelo ZIBB com preditor linear para as proporções o polinômio de terceiro grau e para os zeros uma constante

$$g(\hat{p}_i) = \log \left(\frac{\hat{p}_i}{1 - \hat{p}_i} \right) = -0,1844 + 2,2370x_i - 0,0885x_i^2 - 0,9714x_i^3$$

$$g(\hat{w}_i) = \log \left(\frac{\hat{w}_i}{1 - \hat{w}_i} \right) = 0,0972$$

- **Número ideal de fêmeas**

Maximizando-se o número ideal de fêmeas para parasitar a maior quantidade de ovos encontrou-se 52 fêmeas.

- **Proporções de ovos parasitados e não parasitados**

$$\hat{p} = \frac{e^{1,057}}{1 + e^{1,057}} = 0,742 \quad \text{e} \quad \hat{w} = \frac{e^{0,0972}}{1 + e^{0,0972}} = 0,5243$$

- **Número esperado de ovos parasitados**

$$E(X) = (1 - w)mp = 45,2$$

CONCLUSÕES

- O modelo ZIBB com preditor linear para as proporções o polinômio de terceiro grau e para os zeros uma constante ajustou-se melhor aos dados do ensaio de controle biológico, de acordo com o fator de Bayes e o DIC;
- Para o critério de comparação da adequabilidade dos modelos, os resultados encontrados no fator de Bayes e no DIC são semelhantes;
- Ajustando o modelo escolhido pelo fator de Bayes e pelo DIC, obteve-se que 52 fêmeas do parasita *T. galloi* é o número ideal para maximizar o número de ovos parasitados de *A. kuehniella*.

EXPERIMENTO 2

Delineamento inteiramente casualizado:

- ▶ 2 tratamentos, 4 semanas e 8 repetições;

$$\left\{ \begin{array}{l} MGM : \text{ milho geneticamente modificado MON810;} \\ MC : \text{ milho convencional com aplicação de inseticida.} \end{array} \right.$$

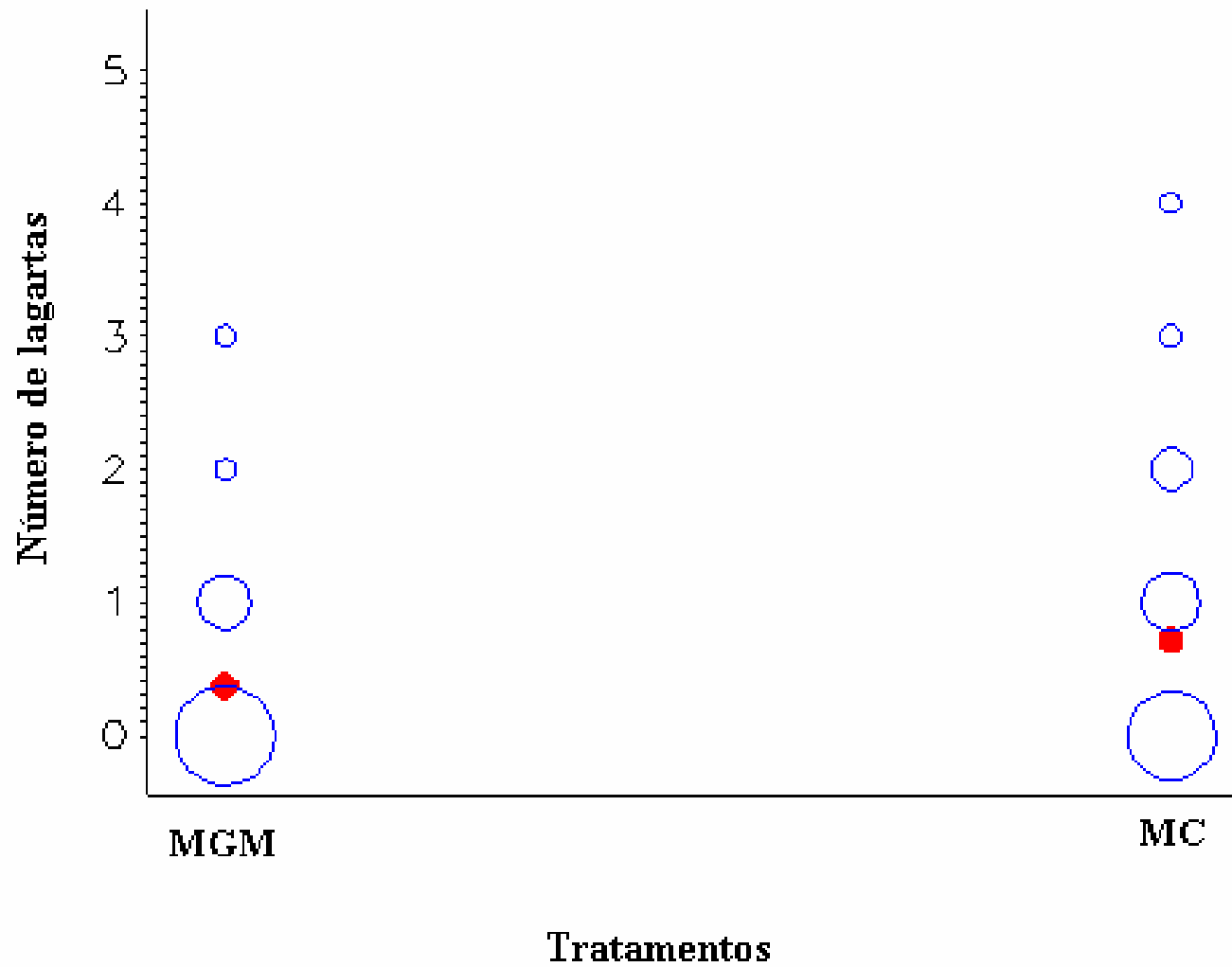
- ▶ parcelas com $1250m^2$;

Variável resposta: número de lagartas grandes. As contagens foram feitas semanalmente, durante 4 semanas;

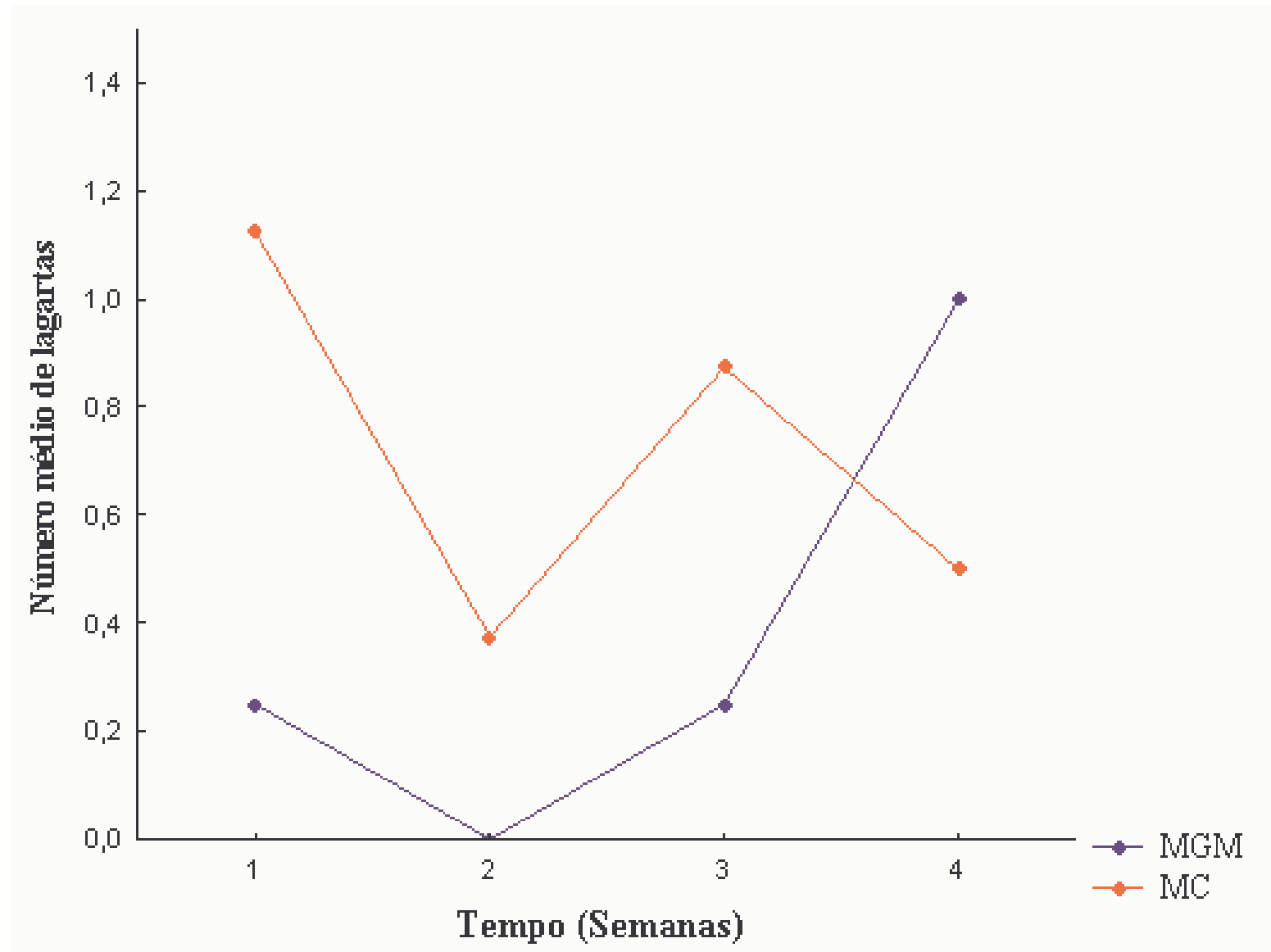
Frequências observadas do número de lagartas por tratamento

Número de lagartas	MGM		MC	
	Freq.	%	Freq	%
0	23	71,9	18	56,3
1	7	21,9	8	25,0
2	1	3,1	4	12,5
3	1	3,1	1	3,1
4	0	0,0	1	3,1
Média	0,375		0,719	

Em 64,1% das observações, não foi encontrada nenhuma lagarta
(os círculos são proporcionais às frequências e ● representa a média)



Frequências observadas do número de lagartas por tratamento



CONSIDERAÇÕES

- i) a natureza discreta da variável resposta (contagem do número de lagartas);
- ii) o número de repetições em que nenhuma lagarta foi encontrada (excesso de zeros);
- iii) análise longitudinal.

OBJETIVO

- Avaliar a eficiência do MGM em relação ao MC no controle da lagarta-do-cartucho;
- Usar a estimação Bayesiana na modelagem dos dados;
- Comparar o ajuste dos modelos usados na análise dos dados.

Modelos Estatísticos

Modelo Poisson

$$Y_{ij} \sim \text{Poisson}(\lambda_{ij})$$

- Função de ligação: Logarítmica.

Preditor Linear:

$$\log(\lambda_{ij}) = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij}$$

Priori:

$$\mu \sim N(0, 10^3) \quad \alpha_i \sim N(0, 10^3) \quad \tau_j \sim N(0, 10^3) \quad (\alpha\tau)_{ij} \sim N(0, 10^3)$$

Modelo ZIP

$$P(Y_{ij} = y_{ij}) = \begin{cases} w + (1 - w)e^{-\lambda_{ij}} & y_{ij} = 0 \\ (1 - w) \frac{e^{-\lambda_{ij}} \lambda_{ij}^{y_{ij}}}{y_{ij}!} & y_{ij} > 0 \end{cases}$$

Preditor Linear:

$$\log(\lambda_{ij}) = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij}$$

$$\text{logit}(w) = \mu$$

Priori:

$$\mu \sim N(0, 10^3) \quad \alpha_i \sim N(0, 10^3) \quad \tau_j \sim N(0, 10^3) \quad (\alpha\tau)_{ij} \sim N(0, 10^3)$$

Totais de zeros observados, por tratamento, ao longo do tempo

Tratamentos	Semanas				Totais de zeros
	1	2	3	4	
MGM	6	8	6	3	23
MC	4	5	3	6	18
Totais de zeros	10	13	9	9	

Modelo ZIP com efeito aleatório

Devido à natureza longitudinal das observações, acredita-se que haja correlação entre as observações no tempo.

Preditor Linear:

$$\log(\lambda_{ikj}) = \mu + \alpha_i + b_{k(i)} + \tau_j + (\alpha\tau)_{ij}$$

$$\text{logit}(w) = \mu$$

Priori:

$$\begin{aligned} \mu &\sim N(0, 10^3) & \alpha_i &\sim N(0, 10^3) & \tau_j &\sim N(0, 10^3) & (\alpha\tau)_{ij} &\sim N(0, 10^3) \\ b_{k(i)} &\sim N(0, \sigma) & \sigma &\sim \text{Gamma}(0.001, 0.001) \end{aligned}$$

Valores Gerados

- Gerou-se uma cadeia com 100.000 valores para cada parâmetro, desprezando os primeiros 1.000 valores e selecionando um a cada 25 dos 99.000 valores restantes, formando uma amostra de 3.960 valores.
- Convergência:
 - análise gráfica dos valores gerados;
 - diagnóstico de Raftery & Lewis;
 - diagnóstico de Heidelberger & Welch.
- O DIC foi usado para a escolha do modelo mais adequado.

R E S U L T A D O S

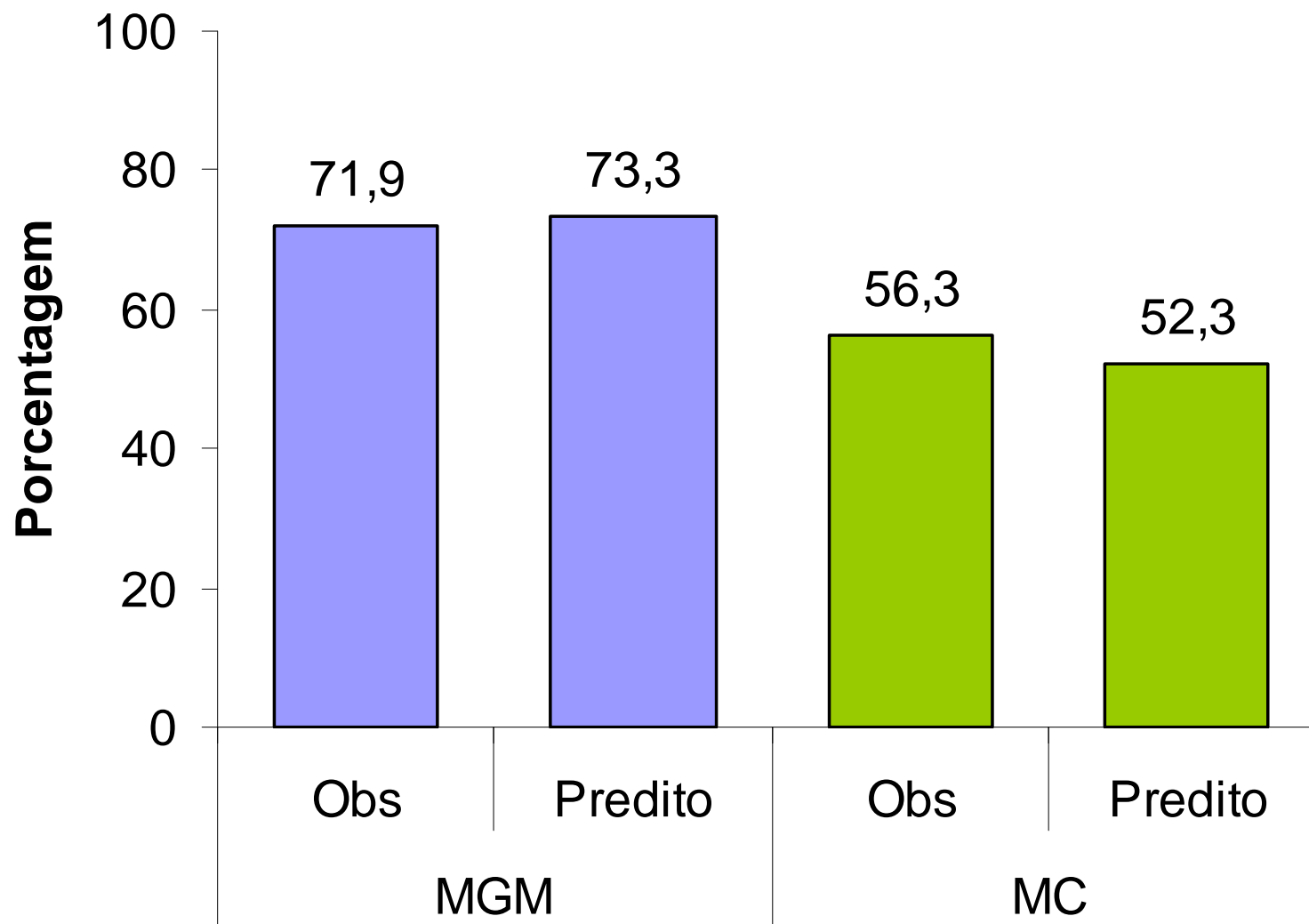
Valores do DIC para os modelos: Poisson, ZIP e ZIP com efeito aleatório

Modelos	Preditores lineares	DIC
Poisson	$\log(\lambda) = \mu$	134,7
	$\log(\lambda) = \mu + \alpha_i$	133,2
	$\log(\lambda) = \mu + \alpha_i + \tau_j$	132,2
	$\log(\lambda) = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij}$	127,1
Fixando-se: $\text{logit}(w) = \mu$		
ZIP	$\log(\lambda) = \mu$	131,9
	$\log(\lambda) = \mu + \alpha_i$	131,3
	$\log(\lambda) = \mu + \alpha_i + \tau_j$	130,5
	$\log(\lambda) = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij}$	127,1
Fixando-se: $\text{logit}(w) = \mu$		
ZIP com ef. aleat.	$\log(\lambda) = \mu + \phi$	130,3
	$\log(\lambda) = \mu + \alpha_i + \phi$	130,6
	$\log(\lambda) = \mu + \alpha_i + \tau_j + \phi$	129,4
	$\log(\lambda) = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij} + \phi$	126,8

Estimativa do número médio de lagartas usando o modelo
 Poisson: $\log(\lambda) = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij}$

Semanas	Tratamentos	Observados	Preditos	Intervalos de credib.	
				2,5%	97,5%
1	MGM	0,25	0,27	0,04	0,71
	MC	1,125	1,10	0,47	1,94
2	MGM	0,0	0,04	0,00	0,22
	MC	0,375	0,34	0,07	0,82
3	MGM	0,25	0,27	0,05	0,70
	MC	0,875	0,85	0,37	1,56
4	MGM	1,0	0,91	0,38	1,60
	MC	0,5	0,61	0,20	1,20

Porcentagem observada e esperada de zeros, por tratamento
para o modelo Poisson: $\log(\lambda) = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij}$



CONCLUSÕES

- ▶ Não existe evidências de que o modelo ZIP tenha melhorado a análise.
- ▶ Não há indícios de que a inclusão do efeito aleatório faz com que o modelo ajustado seja melhor.
- ▶ Não foi possível observar diferença estatística entre os tratamentos (MGM e MC), em relação ao número de lagartas.